

INVERSE PROBLEMS IN
PROPAGATION-BASED X-RAY PHASE
CONTRAST IMAGING AND TOMOGRAPHY:
STABILITY ANALYSIS AND
RECONSTRUCTION METHODS

Dissertation
zur Erlangung des
mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"

der Georg-August-Universität Göttingen
im Promotionsprogramm "Mathematical Sciences"
der Georg-August University School of Science (GAUSS)

vorgelegt von
Simon Maretzke
aus Gifhorn

Göttingen, 2019

Betreuungsausschuss

Prof. Dr. Thorsten Hohage,
Institut für Numerische und Angewandte Mathematik,
Georg-August-Universität Göttingen

Prof. Dr. Tim Salditt,
Institut für Röntgenphysik,
Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission

Referent: Prof. Dr. Thorsten Hohage

Koreferent: Prof. Dr. Tim Salditt

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Russell Luke,
Institut für Numerische und Angewandte Mathematik,
Georg-August-Universität Göttingen

Prof. Dr. Gerlind Plonka-Hoch,
Institut für Numerische und Angewandte Mathematik,
Georg-August-Universität Göttingen

Prof. Dr. Claus Ropers,
IV. Physikalisches Institut,
Georg-August-Universität Göttingen

Dr. Frank Werner,
Institut für Mathematische Stochastik,
Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 04.03.2019

Contents

I	Introduction and Background	1
1	Introduction	2
1.1	From Röntgen's first radiograph to X-ray phase contrast imaging	2
1.2	The GINIX: an exemplary XPCI-setup	5
1.3	Image reconstruction: XPCI and XPCT as inverse problems	6
1.4	Mathematical challenges in XPCI and XPCT	7
1.5	Outline and structure of the thesis	8
2	Background and Preliminaries	10
2.1	Physical model of XPCI	10
2.1.1	Basic wave-optical model	10
2.1.2	The refractive index in the hard X-ray regime	11
2.1.3	Step-by-step modeling of the image-formation process	11
2.1.4	Image reconstruction: the inverse problem of XPCI	14
2.1.5	Model adjustments for real-world XPCI setups	14
2.2	Focus on Fresnel propagation	18
2.2.1	On Fresnel numbers and imaging-regimes	18
2.2.2	Mathematical properties of the Fresnel propagator	21
2.3	Tomographic imaging model	22
2.3.1	Relation to the Radon transform	23
2.3.2	Analytical properties	24
2.3.3	Full angular sampling	25
2.4	Forward operators	26
2.4.1	Basic nonlinear forward model	26
2.4.2	Linearized CTF-model	26
2.4.3	Homogeneity constraints and pure phase objects	27
2.4.4	Multiple holograms	28
2.4.5	Tomographic models for XPCT	29
2.4.6	Domains and constraints	29
2.4.7	Well-definedness, continuity and differentiability	30

2.5	Inverse problems and ill-posedness	36
2.5.1	Uniqueness	36
2.5.2	Ill-posedness despite uniqueness	40
2.6	Discretization	42
2.6.1	General remarks	42
2.6.2	Detailed strategy	42
2.7	Existing reconstruction methods	45
2.7.1	Generic methods	45
2.7.2	Phase reconstruction in XPCI	46
2.7.3	Two-step reconstruction methods for XPCT	49
2.7.4	All-at-once reconstruction methods for XPCT	50

II Summary and Discussion of the Publications 52

3	Stability estimates for linearized near-field phase retrieval in X-ray phase contrast imaging	53
3.1	Motivation	53
3.1.1	Teaser: the impact of support constraints in practice	54
3.1.2	Lipschitz-stability estimates and their meaning	55
3.2	Summary of the results	56
3.2.1	Setting and assumptions	56
3.2.2	Principal results	57
3.2.3	Ideas of the proofs	60
3.3	Supplement: stability of phase contrast tomography	63
3.3.1	Stability of tomographic reconstruction	64
3.3.2	Projection-based stability estimates	64
3.3.3	3D-based stability estimates	65
3.3.4	Interpretation of the results	66
3.4	Supplement: numerical stability computations	67
3.4.1	Contrast-minimizing modes	67
3.4.2	Numerical method	68
3.4.3	Numerical stability results	71
3.5	Supplement: stability by non-negativity constraints	73
3.6	Supplement: the completely holographic regime – a doubly sweet spot of XPCI .	75
3.6.1	Asymptotic linearity of XPCI for low Fresnel numbers	75
3.6.2	Practical realization of completely holographic XPCI	77
3.7	Conclusions	78

4	Locality Estimates for Fresnel-wave-propagation and stability of X-ray phase contrast imaging with finite detectors	81
4.1	Motivation	81
4.2	Summary of the results	83
4.2.1	Basic setting	83
4.2.2	Preliminary results and goal of the article	84
4.2.3	Assessment by Gaussian wave-packets:	85
4.2.4	General locality theory	87
4.2.5	Stability estimates for spline-objects	88
4.2.6	Improved bounds for real-valued objects	90
4.2.7	Extension to linearized XPCI	91
4.3	Conclusions	91
5	Regularized Newton methods for X-ray phase contrast and general imaging problems	95
5.1	Motivation	95
5.2	Summary of the results	97
5.2.1	Iteratively regularized Gauss-Newton Method (IRGNM)	97
5.2.2	IRGNM applied to XPCI	99
5.2.3	Regularizing Newton-Kaczmarz for XPCT	100
5.3	Supplement: comparison to theoretical predictions	102
5.3.1	Comparison to stability theory from Article 1	103
5.3.2	Comparison to resolution estimates from Article 2	104
5.4	Conclusions	104
6	Generalized SART-Methods for Tomographic Imaging	107
6.1	Motivation	107
6.1.1	The problem of complexity in tomographic reconstruction	107
6.1.2	Classical Kaczmarz-iterations and SART	108
6.1.3	The idea of generalizing SART	109
6.2	Summary of the results	110
6.2.1	General setting	110
6.2.2	Properties of tomographic projectors and notation	111
6.2.3	The Generalized SART-principle	112
6.2.4	Admissible penalty-functionals	114
6.2.5	Applications: XPCT and beyond	116
6.2.6	Remarks on discretization	119
6.2.7	Numerical examples	120
6.3	Supplement: High-performance implementation	124

6.3.1	General considerations	124
6.3.2	Parallelization	126
6.3.3	Efficient implementation by distributed design	127
6.4	Supplement: Application to a large-scale data set	130
6.4.1	Description of the data set	130
6.4.2	Reconstruction methods	130
6.4.3	Results	131
6.5	Conclusions	134
7	Discussion and Outlook	137
III	Complete Articles	139
8	Complete Article 1: Stability estimates for linearized near-field phase retrieval in X-ray phase contrast imaging	140
8.1	Introduction	140
8.2	Imaging problems and main results	142
8.2.1	Physical model	142
8.2.2	Weak object limit and principal inverse problem	144
8.2.3	Homogeneous and non-absorbing objects	145
8.2.4	Stability estimates	145
8.3	Stability analysis of inverse problem 8.1	146
8.3.1	Principal approach	146
8.3.2	Reduction to Fourier completion problem	147
8.3.3	Stability result for stripe-shaped supports	149
8.3.4	Characterization of the least stable modes	150
8.3.5	Numerical Validation	152
8.4	Stability analysis of inverse problem 8.2	153
8.4.1	Preparations and Fourier domain splitting	153
8.4.2	Estimate for the central CTF-minimum	155
8.4.3	Estimate for the first order CTF-zeros	157
8.4.4	Global stability results	159
8.4.5	Optimality	161
8.5	Image reconstruction from two measurements	162
8.6	Discussion and conclusions	164
9	Complete Article 2: Locality estimates for Fresnel-wave-propagation and stability of X-ray phase contrast imaging with finite detectors	166
9.1	Introduction	167

9.2	Background	168
9.2.1	Basic setting	168
9.2.2	Properties of the Fresnel propagator	171
9.2.3	Preliminary results	173
9.3	Assessment by Gaussian wave-packets	174
9.3.1	The Gaussian-beam solution	174
9.3.2	Gaussian wave-packets	175
9.3.3	Resolution estimates via Gaussian wave-packets	175
9.4	Locality estimates for complex-valued objects	180
9.4.1	Basic idea and preliminaries	180
9.4.2	Principal leakage estimates	181
9.4.3	Explicit leakage bounds for rectangular domains	182
9.4.4	Stability estimates	185
9.5	Stability estimates for spline objects	185
9.5.1	Multi-variate B-splines	186
9.5.2	Quasi-band-limitation of B-splines	186
9.5.3	Stability estimates	188
9.5.4	Application: resolution estimates	189
9.6	Improved estimates for real-valued objects	190
9.6.1	Quasi-symmetric propagation principle	190
9.6.2	Construction of improved leakage bounds	192
9.6.3	Stability estimates for spline objects	195
9.6.4	Application: resolution estimates	197
9.7	Extension to the phaseless case: application to linearized XPCI	198
9.7.1	Leakage estimates	198
9.7.2	Stability estimates	198
9.7.3	Improved estimates for real-valued objects	200
9.8	Conclusions	200
	Appendices	201
9.A	Fresnel-propagation and frequency shifts	201
9.B	Quasi-band-limitation of B-splines	202
10	Complete Article 3: Regularized Newton methods for X-ray phase contrast and general imaging problems	206
10.1	Introduction	207
10.2	Regularized Newton methods for imaging	208
10.2.1	Basic approach	208
10.2.2	Parameter choice and constraints	210
10.2.3	Newton-Kaczmarz methods	211

10.3	Application to propagation-based phase contrast	212
10.4	Phase contrast tomography of a colloidal crystal	216
10.5	Conclusions	221
11	Complete Article 4: Generalized SART-Methods for Tomographic Imaging	223
11.1	Introduction	223
11.2	Background	225
11.2.1	Tomographic imaging model	225
11.2.2	Inverse problem and a priori constraints	226
11.2.3	Reconstruction methods	227
11.2.4	Convergence of Kaczmarz-iterations and relation to Tikhonov regularization	227
11.2.5	Contribution	228
11.3	The Generalized SART-Principle	229
11.3.1	Preparations: notation and analysis of the projectors	229
11.3.2	L^2 -SART: a promising example	230
11.3.3	Generalized SART framework	232
11.4	Admissible penalty functionals	234
11.4.1	Preliminary insights	234
11.4.2	(Weighted) L^2 -penalties	234
11.4.3	Gradient-penalties	237
11.4.4	L^q -penalties	241
11.5	Applications	242
11.5.1	Noise-model-adapted GenSART	242
11.5.2	Regularized Newton-Kaczmarz-GenSART	246
11.5.3	Extensions	248
11.6	Numerical examples	249
11.6.1	Implementation	249
11.6.2	Robust tomography test case	251
11.6.3	Newton-Kaczmarz-GenSART for experimental XPCT-data	253
11.7	Conclusions	256
	Appendices	257
11.A	Geometry of the Projectors	257
11.B	Projectors and Gradients	257
11.C	Admissibility of L^q -Penalties	259
11.D	Poisson-noise-adapted data fidelity	260
A	Supplementary proofs for numerical stability computations	261
A.1	Characterization of contrast-minimizing modes	261
A.2	Symmetries of contrast-minimizing modes	263

Nomenclature: recurrent notation

Abbreviations

Short form	Long form
e.g.	<i>exempli gratia</i> (latin for “for example”)
i.e.	<i>id est</i> (latin for “that is” or “in other words”)
w.r.t.	with respect to
w.l.o.g.	without loss of generality
l.h.s. / r.h.s.	left-hand side / right-hand side
1D / 2D / 3D	one-/two-/three-dimensional
CT	computed tomography
XPCI	(propagation-based) phase contrast imaging
XPCT	(propagation-based) phase contrast tomography
GINIX	Göttingen Instrument for Nano-Imaging with X-rays (exemplary high-resolution XPCI-setup, see §1.2)
CTF	contrast-transfer-function
FFT	fast Fourier transform
FBP	filtered back-projection
CG	conjugate gradient method
(Gen)SART	(generalized) simultaneous algebraic reconstruction technique

Physical parameters

Notation	Description	Introduced in
δ	refractive decrement of the imaged sample (refractive index: $n = 1 - \delta + i\beta$)	§2.1.2
β	absorption decrement of the imaged sample	§2.1.2
ϕ	phase image	§2.1.3.2
μ	absorption image	§2.1.3.2
I	intensity of the measured hologram(s)	§2.1.3.4
λ, k	X-ray wavelength λ and -wavenumber $k = 2\pi/\lambda$	
$d_{(\text{eff})}$	(effective) propagation-distance between sample and detector	fig. 2.1, §2.1.5.1
$\Delta x_{\text{pix}}^{(\text{eff})}$	(effective) pixel size of the detector	§2.1.5.1
$\mathfrak{f}_{(b)}$	Fresnel number (of a lengthscale b)	§2.1.3.1, §2.2.1
$\mathfrak{f}_{\text{pix}}$	Fresnel number associated with the effective pixel size, $\mathfrak{f}_{\text{pix}} = \mathfrak{f}_b$ for $b = \Delta x_{\text{pix}}^{(\text{eff})}$	§2.2.1.4
$\mathfrak{f}_{(b)}$	modified Fresnel number (of a lengthscale b), $\mathfrak{f}_{(b)} = 2\pi\mathfrak{f}_{(b)}$	§2.2.1.6
$c_{\beta/\delta}$	constant ratio of absorption and refraction in the case of homogeneous objects, $c_{\beta/\delta} = \beta/\delta = \mu/\phi$	§2.4.3
ν	homogeneity-parameter, $\nu = \arctan(c_{\beta/\delta})$	§2.4.3

Sets and spaces

Notation	Description
\overline{A}	closure of a set A
A^c	complement of a set $A \subset \mathbb{R}^m$, $A^c = \mathbb{R}^m \setminus A$
$\mathbf{1}_A$	indicator function of $A \subset \mathbb{R}^m$, $\mathbf{1}_A(\mathbf{x}) = 1$ for $\mathbf{x} \in A$ and $\mathbf{1}_A(\mathbf{x}) = 0$ for $\mathbf{x} \in A^c$
$B[\mathbf{a}; r]$	closed ball of radius r around $\mathbf{a} \in \mathbb{R}^m$, $B[\mathbf{a}; r] = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{x} - \mathbf{a} \leq r\}$
\mathbb{S}^{m-1}	unit-sphere in \mathbb{R}^m , $\mathbb{S}^{m-1} := \{\mathbf{n} \in \mathbb{R}^m : \mathbf{n} = 1\}$
$\text{supp}(f)$	support of a function $f : A \rightarrow \mathbb{C}$, $\text{supp}(f) = \overline{\{\mathbf{x} \in \mathbb{R}^m : f(\mathbf{x}) \neq 0\}}$
$L^2(\mathbb{R}^m)$	space of square-integrable functions, $L^2(\mathbb{R}^m) = \{f : \mathbb{R}^m \rightarrow \mathbb{C} : \int_{\mathbb{R}^m} f ^2 d\mathbf{x} < \infty\}$
$L^2(\Omega)$	space L^2 -functions supported in $\Omega \subset \mathbb{R}^m$, $L^2(\Omega) = \{f \in L^2(\mathbb{R}^m) : \text{supp}(f) \subset \Omega\}$
$L^2(\Omega, \mathbb{R})$	space of real-valued L^2 -functions, $L^2(\Omega, \mathbb{R}) = \{f : \mathbb{R}^m \rightarrow \mathbb{R} : f \in L^2(\Omega)\}$

Basic operations

Notation	Description
$ \mathbf{v} $	Euclidean norm of a vector $\mathbf{v} \in \mathbb{R}^n$
$\mathbf{v} \cdot \mathbf{w}$	Euclidean inner product of vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$
\bar{z}	conjugate of a complex number $z \in \mathbb{C}$
$ z $	modulus of a complex number $z \in \mathbb{C}$
$\text{Re}(z), \text{Im}(z)$	real- and imaginary part of a complex number $z \in \mathbb{C}$
$\ \cdot\ _V$	norm of a normed vector space V
$\langle \cdot, \cdot \rangle_X$	inner product of a Hilbert space X
$\ \cdot\ _{(L^2)}, \langle \cdot, \cdot \rangle_{(L^2)}$	L^2 -norm and inner product, $\langle f, g \rangle_{(L^2)} = \int f(\mathbf{x}) \overline{g(\mathbf{x})} d\mathbf{x}$, $\ f\ _{(L^2)}^2 = \langle f, f \rangle_{(L^2)}$
$f _K$	restriction of a function $f : \mathbb{R}^m \rightarrow \mathbb{C}$ to a subset $K \subset \mathbb{R}^m$, $f _K = \mathbf{1}_K \cdot f$

Special operators

Notation	Description	Introduced in
id_A	identity on a set A , $\text{id}_A : A \rightarrow A; a \mapsto a$	
\mathcal{F}	Fourier transform, $\mathcal{F}(f)(\boldsymbol{\xi}) := (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \exp(-i\boldsymbol{\xi} \cdot \mathbf{x}) f(\mathbf{x}) d\mathbf{x}$	
$\mathcal{D}^{(f)}$	Fresnel propagator (to the modified Fresnel number f)	§2.1.3.1 , §2.2.2
$m_{\mathfrak{f}}, n_{\mathfrak{f}}, k_{\mathfrak{f}}$	kernel functions associated with $\mathcal{D}^{(f)}$	§2.2.2
$\mathcal{N}^{(f)}$	general nonlinear forward map of XPCI	§2.4.1
$\mathcal{N}_\nu^{(f)}$	nonlinear forward map of XPCI for homogeneous objects	§2.4.3
$\mathcal{T}^{(f)}$	linearized forward map of XPCI for general objects	§2.4.2
$\mathcal{S}_\nu^{(f)}$	linearized forward map of XPCI for homogeneous objects	§2.4.3
$s_\nu^{(f)}$	contrast-transfer-function (CTF) associated with $\mathcal{S}_\nu^{(f)}$	§2.4.3
$F^{(f_1, \dots, f_\ell)}$	forward map to multiple modified Fresnel numbers f_1, \dots, f_ℓ ($F \in \{\mathcal{N}, \mathcal{N}_\nu, \mathcal{T}, \mathcal{S}_\nu\}$)	§2.4.4
$\mathcal{P}_\theta, \mathcal{P}_\theta$	parallel-beam projector for a direction $\boldsymbol{\theta} \in \mathbb{S}^2$ or angle $\theta \in \mathbb{R}$	§2.3
$\mathcal{P}_{\text{full}}$	parallel-beam projector for a full set of angles $\theta \in [0; \pi)$	§2.3.3
$T^* : Y \rightarrow X$	adjoint of a bounded linear map $T : X \rightarrow Y$ (X, Y : Hilbert)	
$F'[f]$	Fréchet-derivative of a nonlinear operator F evaluated at f	§2.4.7.3

Part I

Introduction and Background

Chapter 1

Introduction

1.1 From Röntgen’s first radiograph to X-ray phase contrast imaging

Ever since the CONRAD RÖNTGEN published the first radiograph of his wife’s hand in 1895 [173], imaging with X-rays has enabled unprecedented non-destructive views into the internal workings of formerly unknown realms. The principle of classical X-ray radiography is simple: X-rays emitted by some source are partially absorbed as they traverse matter, such that the induced profile in the X-ray intensity behind an imaged sample yields a two-dimensional (2D) projection-image of its three-dimensional (3D) structure. Besides its immediate use in medical diagnosis of bone fractures for example, the approach permits *3D*-imaging via the technique of *computed tomography* (CT), as first implemented by CORMACK [44, 45] and HOUNSFIELD [103] in the 1960s: Radiographs acquired from different perspectives (by either rotating the imaged object or the X-ray-source and -detector) allow to resolve 3D-variations of the local X-ray absorption. Via mathematical tools developed by JOHANN RADON already in 1917, such a 3D-image can be *reconstructed* from the measured series of conventional 2D X-ray images.

The remarkable ability of classical radiography and CT to “look into things” is physically based on the *high penetration depth* of (hard) X-rays in matter compared to other types of radiation like visible light or electrons: provided a sufficiently high photon-energy, residual radiation is transmitted even through objects of thickness in the order of centimeters or larger. While this principal selling-point of X-rays enables imaging of macroscale objects such as entire human organs, their small wavelengths in the range between 0.01 and 10 nanometers in principle also allow to resolve structures down to the *nanoscale*. Yet, when trying to apply X-ray imaging to small objects of size in the order of micrometers or below, the high penetration depth becomes problematic: microscale samples that are composed mainly of light chemical elements, such as single biological cells, are quasi completely transparent to (hard) X-rays so that the induced absorption contrast in a radiograph may be imperceptibly low. Importantly, however, this does not mean that such specimens are *invisible* to X-rays since non-*absorbing* does not imply non-*interacting*: in analogy to glass or water for visible light, even transparent materials may *refract* X-rays to a significant degree and thereby become perceptible, see fig. 1.1(a). Although also refraction of X-rays is much weaker than for visible light, it may still exceed X-ray absorption by several orders of magnitude especially for biological soft tissue and other light-element materials, thus promising massively improved contrast.

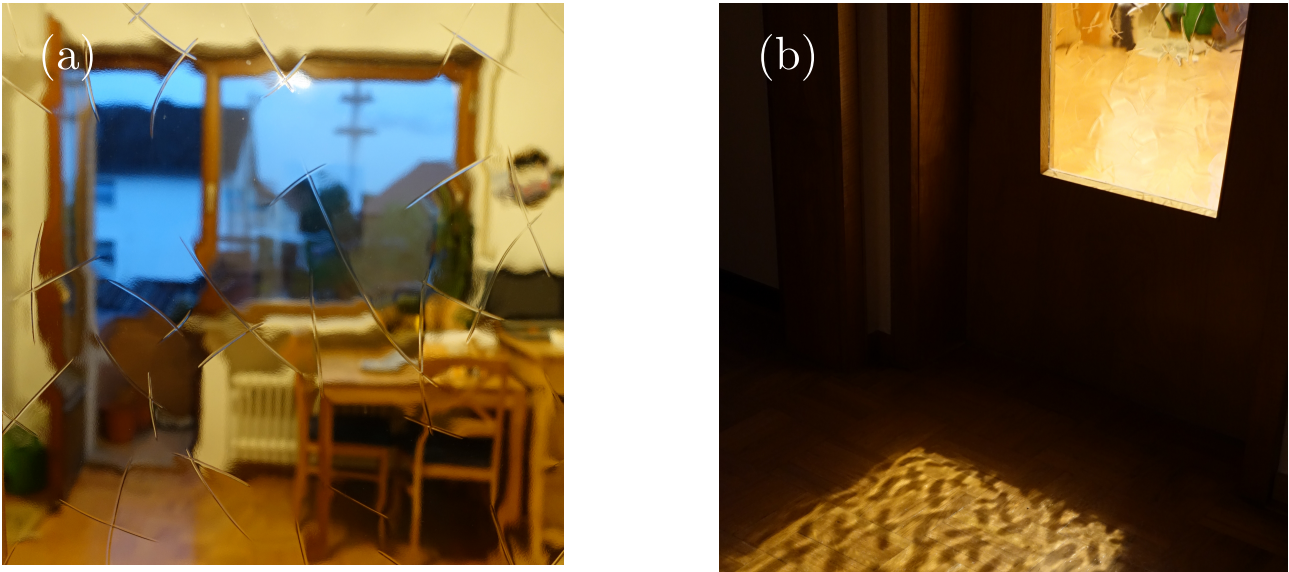


Figure 1.1. Phase contrast in everyday life (photos by the author). (a) Refraction of light renders the pattern in the glass-window visible although it is practically non-absorbing. (b) The *hologram* formed on the floor behind the illuminated window encodes the refracting structure as intensity-variations.

While the potential benefits of refraction-based X-ray imaging have been known for a long time, its practical implementation is hampered by physical details: in the standard model of X-rays as classical electromagnetic waves, refraction manifests as *phase-shifts* of the transmitted X-ray wave-fronts, which is why refraction-sensitive imaging methods are denoted as *phase contrast* techniques. Measuring phases, however, brings about fundamental difficulties:

- (1) *The phase problem:* X-ray detectors are sensitive only to the *intensity* of the incident radiation, i.e. to the (squared) amplitude of the wave-field, whereas the phase-information is lost in the measurement process. Hence, phase contrast requires an experimental setup that establishes phase-sensitivity by encoding phases into measurable intensities.
- (2) *Coherence requirements:* As phases may only be observed indirectly via wave-interference according to point (1), phase contrast effects typically require sufficiently coherent radiation to be observable, i.e. in particular ideally monochromatic X-rays.

For a long time, the low coherent photon-flux produced by available X-ray sources (after sufficient monochromatization and collimation with optical filters and slits) has indeed ruled out the possibility of exploiting the high-resolution potential of X-ray imaging via phase contrast techniques. A prominent exception is given by *X-ray crystallography*. Already shortly after the Nobel-Prize-awarded works of VON LAUE [204] and BRAGG [27] in the early 20th century, the method revealed the positions of atoms in a wide range of crystals and thereby enabled revolutionary insights in material science and structural biology. The key property of crystals lies in their periodic structure, providing a large number of copies of the object of interest, the crystal's unit cell, which amplifies the induced signal in the measured diffraction patterns by a huge factor. This effect may compensate a low intensity of the illuminating X-ray beam.

Extending the success of X-ray crystallography to general, non-crystalline specimens required massive increases in coherent flux. This has only recently been achieved with the development of novel X-ray sources during the past few decades. Indeed, as remarked in [158],

coherence had been lacking for so long that the first experimental demonstrations of X-ray phase contrast [187, 207, 40, 123, 162, 146] in the 1990s initially came as a surprise: the discovery was a consequence of the advent of *third-generation synchrotrons*, providing sufficiently coherent X-ray sources for the first time – about 100 years after Röntgen’s first radiograph. Today, free-electron laser sources produce so much coherent flux that even single macro-molecules yield enough contrast to be imaged without being embedded in a crystal [15].

Different phase-sensitive imaging methods have been proposed, see [160, 158] for reviews. This thesis, however, is exclusively concerned with *propagation-based X-ray Phase-Contrast Imaging (XPCI*)*, also known as *in-line holography* or *near-field propagation imaging*. Interestingly, the method uses essentially the same setup as classical X-ray radiography (see §1.2) – only coherence and physical lengthscales make the difference. The required phase-sensitivity is solely established by free-space propagation of the X-rays between sample and detector: after interaction with the imaged object, *diffraction*, i.e. self-interference of the X-rays partially encodes the imprinted phase-shifts in the X-ray wave-field into measurable intensities at the detector. The analogue of this effect for visible light and glass can be seen in fig. 1.1(b), where the floor takes the role of the detector visualizing the resulting near-field diffraction pattern, called *hologram*. Similarly as CT, XPCI thus involves a non-trivial *computational problem*: in order to obtain an interpretable image, the sample-characterizing phase-shifts have to be numerically reconstructed from the measured hologram via suitable algorithms.

Promoted by ongoing progress in optical control of X-rays [178, 54, 11, 95] and ever brighter and more coherent sources [4, 189, 55], XPCI has developed into a promising technique over the past 20 years. Imaging microscale specimens at resolutions of a few tens of nanometers is now possible [13]. Meanwhile, the required X-ray dose in XPCI is low compared to other X-ray imaging methods [85], which translates into relatively short acquisition times and low radiation-damage of the imaged sample. Analogously as classical CT-scanners compute a 3D-image from a series of 2D-radiographs, XPCI can furthermore be extended to a 3D-imaging technique by acquiring holograms for different incident directions of the X-rays, as first demonstrated in [42]. The approach, denoted as **X-ray Phase-Contrast Tomography (XPCT)**, has been applied for example to investigate brain-tissue down to the cellular level in 3D without invasive sample-preparation requirements [193]. Recently, even time-resolved XPCT, i.e. phase contrast imaging in *four* dimensions has been successfully demonstrated [174, 177]. In addition to large-scale X-ray sources, also state-of-the-art laboratory sources enable phase contrast imaging [120] and first clinical applications of such setups are in preparation [28]. The present work, however, focuses on XPCI and XPCT at resolutions below 100 nanometers, which still requires the brilliance of modern synchrotron sources.

As progress on the experimental side breaches ever new frontiers, deep *mathematical* understanding and tailored algorithms for the involved image reconstruction problem become more and more important in order not to bottleneck the capabilities of XPCI and XPCT. Promoting the success of these emerging imaging techniques by contributions on the mathematical side is the principal goal of this thesis. The subsequent sections 1.2 to 1.5 give an outline of the specific problems to be addressed as well as of the applied methodology.

*Both in the abbreviation XPCI and in referring to “phase contrast” we usually omit the specification “propagation-based” in the following since no other phase contrast techniques are considered.

1.2 The GINIX: an exemplary XPCI-setup

Figure 1.2 shows the Göttingen Instrument for Nano-Imaging with X-rays (GINIX) – an exemplary setup for high-resolution XPCI installed at the third-generation synchrotron PETRA III, DESY, Hamburg. While the theory and algorithms of this work are not limited to a specific experimental setup, we will frequently make reference to the GINIX when it comes to identifying practically relevant ranges of physical parameters. Moreover, all of the experimental XPCI- and XPCT-data considered in this thesis has been acquired at the GINIX.

In the synchrotron-source, the X-rays are generated by an undulator that forces accelerated electrons onto oscillatory trajectories via alternating magnetic fields, causing photons to be emitted at a small bandwidth of energies. The GINIX-setup in fig. 1.2 furthermore contains a cascade of different X-ray optics, that serve to form a highly coherent beam from the generated X-rays, emanating from the nanometer-sized source spot at the wave-guide exit. As indicated in fig. 1.2, we will treat this part of the setup as a black-box and simply summarize it as the “X-ray source” in the following, referring to [111, 178] for experimental details. In this abstract picture, the XPCI-setup is appealingly simple, merely consisting of an unknown object placed between source and detector – without requiring any further optical elements.

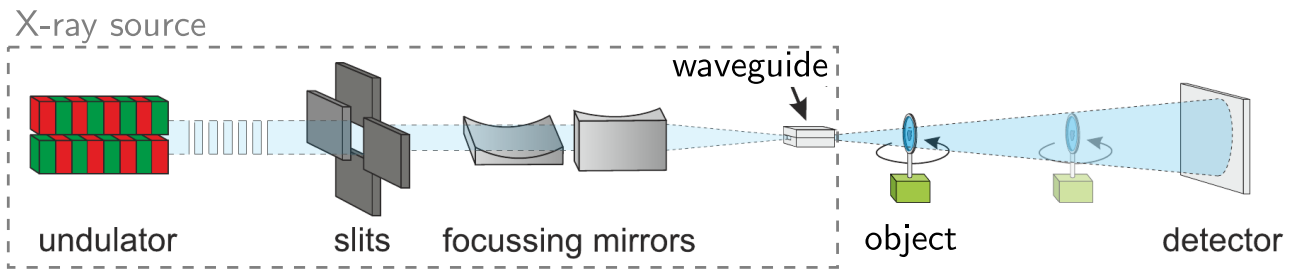


Figure 1.2. Sketch of an exemplary XPCI-setup at a third-generation synchrotron: the Göttingen Instrument for Nano-Imaging with X-rays (GINIX) [111, 178]. For details, see text.

For a correct intuition of the GINIX-setup and the involved physical parameters, it is important to note that the sketch in fig. 1.2 is *not to scale* but the following relations hold for the experimental data considered in this work:

- Source(waveguide)-to-detector-distance (fixed): $d_{0,2} \approx 5$ meters.
- Source-to-object-distance (variable): in the order of *centimeters*.
- Object diameter: in the order of 10 to 100 *micrometers*.
- Detector size: ≈ 1.33 *centimeters* (2048×2048 square pixels of aspect length $6.54 \mu\text{m}$)
- Opening angle of the X-ray beam: in the order of 0.1 degrees.

The divergent-beam geometry of the GINIX-setup gives rise to a *geometrical magnification* of the imaged object on the detector by a factor of $d_{0,2}/d_{0,1} \sim 100$. Accordingly, the setup constitutes a *lensless X-ray microscope*. As indicated in fig. 1.2, the sample may be rotated to acquire a tomographic series of holograms for XPCT.

For a qualitative impression, fig. 1.3 shows examples of holograms of different samples recorded at the GINIX (corrected for lateral intensity-variations of the X-ray beam). A general

feature of the near-field diffraction patterns is that they typically contain a mixture of direct contrast, i.e. undistorted features of the imaged object, and additional wavy fringe-patterns.

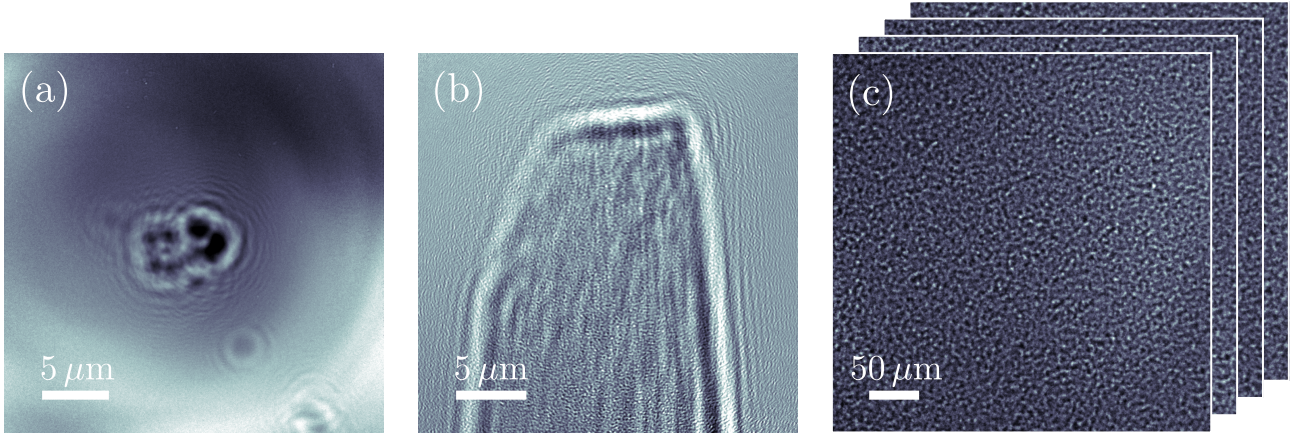


Figure 1.3. Examples of XPCI data (holograms) for different specimens, measured at the GINIX: (a) an ensemble of freeze-dried *Deinococcus radiodurans* bacteria, (b) a tip-shaped test-structure made of nanoporous glass, (c) human brain tissue (cerebellum). The images plot the detected intensity relative to an empty-beam image, i.e. to the intensity-distribution without an object in the beam. In (c), three holograms have been acquired of the same sample at different sample-to-detector-distances (see fig. 1.2) to increase the information-content of the data.

1.3 Image reconstruction: XPCI and XPCT as inverse problems

In order to obtain interpretable images in XPCI, the sample-induced phase-shifts (and possible absorption) have to be *numerically reconstructed* from measured holograms of the kind shown in fig. 1.3. In the case of XPCT, tomographic reconstruction as in classical CT is furthermore required to compute a 3D-image from the tilt-series of 2D-images. Mathematically, XPCI and XPCT thus fit into the framework of *inverse problems*. The problems amount to inverting a mathematical mapping, denoted as the *forward map* F , that describes the relation between the sample-parametrization f (the sought 2D- or 3D-image) and the observable hologram-data g :

$$F : \{\text{admissible images}\} \rightarrow \{\text{possible hologram-data}\} \quad (1.3.1)$$

While the *forward problem* of computing data $F(f) = g$ from a given image f is described by fundamental physics, well-understood and stably solvable, the associated inverse problem is that of greater practical interest: after all, f is unknown and only g can be measured directly. However, reconstructing f from the data g in XPCI is known to be ambiguous and unstable to perturbations of g in general [108, 157, 32] (mainly because of the missing phase-information for the measured holograms, compare §1.1), i.e. *ill-posed* in the sense of HADAMARD [82]:

Definition 1.1 (Well-posedness and Ill-posedness [82]). A problem is called *well-posed* if

- (a) a solution exists,

- (b) the solution is unique,
- (c) the solution depends continuously on the data.

Otherwise, it is called *ill-posed*.

Non-uniqueness in XPCI means that structurally different specimens give rise to identical hologram-data and are thus indistinguishable, which is clearly undesirable. Although not as obvious, also condition (c) in definition 1.1 is vital: if the reconstruction does not depend continuously on the data, then small measurement-inaccuracies, as arising in any real-world imaging experiment, may lead to arbitrary errors in the recovered images. To enable faithful image reconstruction in real-world settings, well-posedness thus needs to be (re-)established either by a careful choice of the imaging-setting, i.e. by acquiring “sufficient” data, or by suitable *regularization*, exploiting *a priori knowledge* on the imaged sample. Analyzing the ill- or well-posedness of the inverse problems in XPCI and XPCT allows to identify general potentials and limitations of the imaging-techniques beyond specific example-images.

1.4 Mathematical challenges in XPCI and XPCT

While some mathematical aspects of XPCI and XPCT are relatively well-understood by now, such as the question of uniqueness [108, 157, 32, 140], several open problems remain, some of which can be identified in the example-holograms in fig. 1.3. The following (incomplete) list of mathematical challenges in X-ray phase contrast imaging will be addressed in this thesis:

- *Stability*: The hologram in fig. 1.3(a) shows pronounced variations of the background-intensity (caused by imperfect flat-field correction, as will be detailed in §2.1.5.2). In order to provide images of practical value, reconstruction needs to be robust against such and other data-errors. To guarantee this, the ill- or well-posedness of the involved inverse problems needs to be quantified by *stability estimates* beyond mere uniqueness.
- *Required number of holograms*: The data in fig. 1.3(c) consists of three holograms acquired at different sample-to-detector-distances, compared to only a single hologram in the other examples. To ensure efficient imaging in terms of acquisition times and radiation dose, it is vital to understand how much data is really needed for faithful image reconstruction.
- *Exploitation of a priori constraints*: Typically, some basic properties of the imaged sample are known *a priori*. The sample in fig. 1.3(a), for example, can be seen to occupy only a small region in the center of the imaged field-of-view. Similarly, the sample contours in (b) may be readily identified in the hologram. Imposing such prior knowledge as constraints reduces the number of admissible images and may thereby facilitate reconstruction.
- *Finite field-of-view (FoV)*: The hologram in fig. 1.3(b) is not fully captured by the square detector: the fringe-pattern can be seen to reach out further than the imaged FoV. The situation is even more severe in fig. 1.3(c) where the sample exceeds the FoV in diameter, so that there are not only fringes leaving the FoV but also some that enter from the outside. Theory on the impact of these effects on the recovered images is scarce.

- *Nonlinearity*: The forward map F from §1.3, that models the relation between the sought images and the measured hologram-data, is nonlinear in general. This poses challenges both in theoretical analysis and in designing practical image reconstruction algorithms.
- *Mixed phase- and absorption contrast*: Some specimens are not only phase-shifting but also absorb X-rays to a non-negligible degree. This raises the question whether and how phase- and absorption contrast in the holograms may be disentangled.
- *Large-scale data*: Especially in XPCT, the numerical size of the processed 3D-data sets (typically arrays with $\sim 10^9$ floating-point numbers) is so large that computational costs of image reconstruction remains a significant issue. In particular, the required time to recover the image should not substantially exceed the experimental acquisition time.

The above aspects will be analyzed mathematically with the goal of improving the quality of the reconstructed images. To date, the most widely used methods for XPCI (and XPCT) at synchrotrons are based on a simplified image-formation model in terms of so-called *contrast-transfer-functions* (CTF) [42, 194, 124, 96, 122, 121]. While the CTF-model allows computationally fast image reconstruction and thereby meets the efficiency requirements arising from large-scale data, it relies on a *linearization* of contrast that is valid only for weakly interacting samples. In addition to the resulting inability to account for nonlinearity, it will be seen in the course of this thesis that standard CTF-based schemes also bear shortcomings in terms of exploiting a priori knowledge, for example. On the contrary, more flexible classes of algorithms tend to increase the required computation times by an unacceptably large factor. Tailoring image reconstruction methods that provide a reasonable compromise between computational requirements and versatility is a major motivation for this work.

1.5 Outline and structure of the thesis

The overall goal of this thesis is to develop a theoretical understanding of- and algorithmic solutions for the challenges outlined in §1.4. Its principal structure is that of a *cumulative dissertation*, also known as *compilation thesis* or *thesis by publication*: the present work is a compilation of four journal articles [144, 142, 143, 141], that touch on different topics in XPCI and XPCT but partly also address more general aspects of inverse problems in imaging. This thesis brings the findings of the articles together in a unified notation, outlining connections between the different contributions and also drawing joint conclusions beyond the scope of the individual works. The topics of the considered articles are as follows (cf. §1.4):

- **Article 1** [144] analyzes *stability* of image reconstruction in XPCI and XPCT and highlights the significance of *a priori constraints* for this. The derived quantitative stability estimates shed a light on the required *number of holograms* in different settings as well as on possibilities of *disentangling phase- and absorption contrast*.
- **Article 2** [142] studies the effects of incomplete hologram-data due to a *finite field-of-view* on image reconstruction, extending the stability analysis of Article 1 to this setting and drawing conclusions on the achievable *resolution* in XPCI.
- **Article 3** [143] proposes regularized Newton-type methods as image reconstruction algorithms that account for *nonlinearity* and flexibly incorporate *constraints*.

- **Article 4** [141] derives efficient computational schemes for a certain class of reconstruction algorithms (Kaczmarz methods) for tomographic inverse problems. This helps to reduce the computational costs of the proposed method for XPCT from Article 3, rendering it applicable also for *large-scale data*.

Supplements: In addition to reviewing and discussing already published results, this thesis also contains original, unpublished contributions both on mathematical theory and on algorithmic aspects of XPCI and XPCT, extending the findings of the presented articles. For clarity, the titles of the sections with significant novel results begin with the word “**Supplement**”.

Structure: This thesis is divided into three parts:

- **Part I (chapters 1 and 2)** introduces the basic physical and mathematical background of XPCI and XPCT, outlining general concepts and notation used throughout this work.
- **Part II** contains the core chapters of the thesis. **Chapters 3 to 6** contains summaries (or rather reviews[†]) of the articles 1 to 4, respectively, explaining key findings but omitting technical details and lengthy mathematical proofs where possible. Each summary is accompanied by a thorough motivation and discussion in the light of the other articles and optional *supplements* of the kind explained above. **Chapter 7** concludes the monographic part of this thesis by discussing possible directions of future research.
- **Part III (chapters 8 to 11)** contains the complete articles 1 to 4 in their published forms, up to minor corrections and notational adjustments. In particular, it serves as an appendix for mathematical details that have been omitted in the summaries of part II.

Despite the principal cumulative form, the unpublished, monographic parts I and II are intended to be self-contained, i.e. readable without reference to the complete articles, hopefully leaving the reader with new insights and view-points on the fascinating technique of X-ray phase contrast imaging.

[†]Indeed, while the presentation of the results is generally briefer than in the complete articles, some aspects of particular importance to this thesis may even be treated to greater detail.

Chapter 2

Background and Preliminaries

2.1 Physical model of XPCI

The following section introduces the underlying physical model of *image-formation* in XPCI, describing the dependence of the measured holograms from the sought images. Firstly, a basic model is derived for an idealized XPCI-setting in §2.1.1 to §2.1.3. In §2.1.5, it is then discussed in how far this model is applicable to real-world imaging-setups as sketched in fig. 1.2 and what adjustments are required. More in-depth treatments of the physical background of XPCI can be found in the books [160, 3] as well as in several PhD-theses [39, 12, 120, 171].

2.1.1 Basic wave-optical model

Although the underlying quantum-mechanical processes are complicated in principle, it is a well-known fact that the propagation of X-rays and their interaction with matter is well-described by classical electrodynamics [3]. On lengthscales larger than the size of atoms, the predominant microscopic interactions given by Compton- and Thomson-scattering as well as photo-electric absorption, can be summarized to surprising accuracy by parametrizing materials via a spatially varying refractive index n . Accordingly, X-rays behave like electromagnetic waves in an optically inhomogeneous medium, analogous to visible light passing through a (slightly opaque) lens. Moreover, anisotropies in the scattering-interaction are often negligible so that the polarization of the waves is irrelevant.

Within the scope of this work, X-rays may thus be described by a wave-equation governing the evolution of the scalar *wave-field* $\psi(x, y, z, t)$: $(n^2/c_0^2)\partial_t^2\psi - \Delta\psi = 0$ (c_0 : vacuum speed of light, ∂_t^2 : second derivative in time t , Δ : Laplace-operator w.r.t. the spatial coordinates x, y, z). Furthermore, we restrict to sufficiently coherent, *monochromatic** X-rays, that can be described by a time-harmonic wave-field $\psi(\mathbf{x}, t) = \exp(i\omega t)\Psi(\mathbf{x})$ of a specific frequency ω . The spatial *complex-valued* amplitude Ψ of the wave-field is then described by the *Helmholtz equation*:

$$\Delta\Psi + n^2k^2\Psi = 0. \quad (k = \omega/c_0: \text{wave-number}) \quad (2.1.1)$$

*Although this is easily assumed here by a theoretician, the experimental difficulties in generating sufficiently coherent X-rays have been a major obstacle to imaging methods like XPCI for a long time, see chapter 1.

	δ	β	β/δ
Glucose (density 1.56 g/cm ³)	$5.48 \cdot 10^{-6}$	$1.56 \cdot 10^{-8}$	0.00285
Gold (density 19.32 g/cm ³)	$4.90 \cdot 10^{-5}$	$5.16 \cdot 10^{-6}$	0.105

Table 2.1. Exemplary values of the refractive- and absorption-decrements δ, β at photon-energy $E = 7.9$ keV according to tabulated values in [92].

2.1.2 The refractive index in the hard X-ray regime

The frequency ω of the X-ray wave-field is related to the energy E of the X-ray photons via Planck's formula $E = \hbar\omega$. The XPCI experiments considered in this work operate within the *hard X-ray regime* with beam energies in the order of $E \sim 10$ keV (keV: kilo electron volts). In this part of the electromagnetic spectrum, the refractive index n of matter deviates only slightly from unity. Therefore, it is typically written in the form

$$n = 1 - \delta + i\beta \quad (\text{i: imaginary unit}). \quad (2.1.2)$$

As will be detailed in §2.1.3.2, the real-valued spatially varying parameters δ and β govern *refraction* and *absorption* of X-rays by the material. Therefore, we call δ the *refractive decrement* and β the *absorption-decrement* (or simply absorption). We emphasize the following:

For physical reasons, δ and β are always non-negative in the hard X-ray regime.

Table 2.1 gives exemplary values of δ and β for glucose C₆H₁₂O₆ (biological material) and gold (heavy metal) at photon-energy $E = 7.9$ keV. We note that

- δ and β are both very small compared to one, typically $\lesssim 10^{-4}$. In particular, this means that the following standard approximation is *practically exact*:

$$n^2 = (1 - \delta + i\beta)^2 \approx 1 - 2\delta + 2i\beta. \quad (2.1.3)$$

- $\beta \ll \delta$: for glucose, β is about 350 times smaller and even for gold one still has $\beta/\delta \approx 0.1$. *Refraction is thus much stronger than absorption for hard X-rays.*

More details on the quantities δ , β and the underlying physics are e.g. given in [3, 39, 12]. Let us just note that probing δ , as done in XPCI, yields images in *natural contrast*:

For hard X-rays, the refractive decrement δ is proportional to local electron density.

2.1.3 Step-by-step modeling of the image-formation process

Based upon the Helmholtz-model (2.1.1) for the propagation of X-rays, an XPCI-setup as in fig. 1.2 may be viewed in a simplified and abstract manner as sketched in fig. 2.1: plane waves are incident along the optical axis, scatter on the sample parametrized by its refractive index n and then propagate over a distance d to the detector. Accordingly, image-formation may be divided into three steps: scattering-interaction, free-space wave-propagation and detection. In the following, we will derive a specific mathematical description for each of these steps and

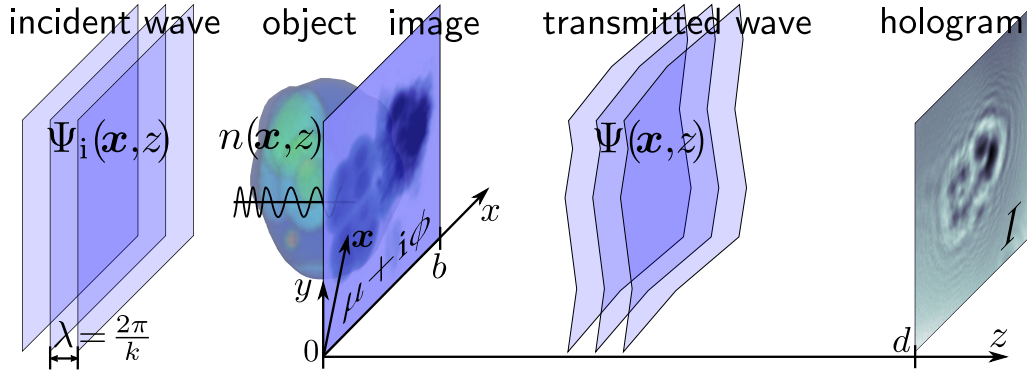


Figure 2.1. Sketch of the basic physical model of XPCI.

finally combine them to obtain a complete *image-formation model*. We choose coordinates such that the optical axis is the z -axis and denote the remaining lateral dimensions by $\mathbf{x} = (x, y)$.

The principal problem is as follows: given the sample's refractive index n , what is the wave-field Ψ in the detector-plane $z = d$? Describing this relation in terms of the full Helmholtz-equation (2.1.1) is possible, but would result in a rather abstract description in terms of the solution to a complicated partial differential equation. Therefore, some additional approximations to the basic model will be introduced, that are standard and highly accurate for the considered hard X-ray setting and enable a simpler description.

2.1.3.1 (Free-space) wave-propagation in paraxial approximation

The incident X-rays are modeled by a plane wave $\Psi_i(\mathbf{x}, z) \propto \exp(ikz)$. For the considered energies $E \sim 10$ keV, the wavelength $\lambda = 2\pi/k$ is in the order 0.1 nanometers, which is much smaller than the lengthscales of variations of n for typical samples in XPCI. Consequently, the wave-field perturbations induced by the object in fig. 2.1 will be coarse compared to the oscillations of the incident plane waves. This motivates to write the total, perturbed wave field in the form $\Psi(\mathbf{x}, z) = \exp(ikz)\tilde{\Psi}(\mathbf{x}, z)$ with slowly varying envelope $\tilde{\Psi}$. Upon inserting this ansatz into the Helmholtz-equation (2.1.1), the second derivative $\partial_z^2\tilde{\Psi}$ may be neglected compared to terms $k\partial_z\tilde{\Psi}$ and $k^2\tilde{\Psi}$. This is the so-called *paraxial* (or Schrödinger- or one-way-) *approximation*, yielding the paraxial Helmholtz equation for the envelope $\tilde{\Psi}$:

$$2ki\partial_z\tilde{\Psi} + \Delta_{\mathbf{x}}\tilde{\Psi} + (n^2 - 1)k^2\tilde{\Psi} = 0. \quad (\Delta_{\mathbf{x}}: \text{Laplacian w.r.t. } \mathbf{x} = (x, y)) \quad (2.1.4)$$

Free-space (Fresnel-)propagation: Based on (2.1.4), we may describe the evolution of the envelope $\tilde{\Psi}$ as the X-rays propagate from the sample-plane $z = 0$ to the detector at $z = d$ in fig. 2.1. As the gap in between is *free space* (air or vacuum in practice), $n = 1$ holds identically in this domain so that (2.1.4) reduces to $2ki\partial_z\tilde{\Psi} + \Delta_{\mathbf{x}}\tilde{\Psi} = 0$. Assuming that the space is infinite in the lateral x - and y -dimensions, the solution to this PDE may be expressed via the *Fresnel propagator* \mathcal{D} : writing $\Psi_{z_0} : \mathbf{x} \mapsto \Psi(\mathbf{x}, z_0)$ for the wave-field at $z = z_0$, we have

$$\tilde{\Psi}_d(\mathbf{x}) = \mathcal{D}\left(\tilde{\Psi}_0((\cdot) \cdot b)\right)(\mathbf{x}/b), \quad \mathcal{D}(f) := \mathcal{F}^{-1}\left(\exp\left(-\frac{i\xi^2}{4\pi\mathfrak{f}}\right) \cdot \mathcal{F}(f)\right) \quad (2.1.5)$$

where $\mathcal{F}(f)(\xi) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} f(\mathbf{x}) \exp(-i\xi \cdot \mathbf{x}) d\mathbf{x}$ denotes the Fourier transform (here: $m = 2$) and $\mathfrak{f}_{(b)} = kb^2/(2\pi d)$ is the dimensionless Fresnel number associated with the lengthscale b .

Dimensionless formulation: (2.1.5) implies that free-space propagation in XPCI-setups is described by only a single parameter \mathfrak{f} – provided that the wave-field Ψ is expressed in terms of *dimensionless* coordinates $\mathbf{x}' = \mathbf{x}/b$ by identifying an arbitrary physical length b with unity. From hereon, all quantities of dimension “length” will therefore be assumed to be dimensionless, implicitly measuring the corresponding physical length in multiples of some reference-scale b . Details on Fresnel propagation, length-scales and their relation are discussed in §2.2.

2.1.3.2 Scattering interaction in projection-approximation

In fig. 2.1, the interaction of the incident X-rays with the sample takes place between the entrance-plane at $z = -L$ and exit-plane at $z = 0$ – everywhere else the refractive index n equals one. Typically, the sample-thickness L is sufficiently small such that *diffraction* of the X-rays may be neglected while they traverse the object, which means that the scattering-interaction is well-described by *ray-optics*[†]. Mathematically, this corresponds to neglecting the lateral Laplacian in the paraxial Helmholtz equation (2.1.4), which leads to

$$\begin{aligned} 2ki\partial_z\tilde{\Psi}(\mathbf{x}, z) + (n^2 - 1)k^2\tilde{\Psi}(\mathbf{x}, z) &= 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^2, z \in [-L; 0] \\ \Rightarrow \tilde{\Psi}_0 &= \exp\left(\frac{ik}{2} \int_{-L}^0 (n^2 - 1) dz\right) \tilde{\Psi}_{-L} \stackrel{\delta, \beta \ll 1}{\approx} \exp\left(-k \int_{-L}^0 (\beta + i\delta) dz\right) \tilde{\Psi}_{-L}. \end{aligned} \quad (2.1.6)$$

with $\tilde{\Psi}_z(\mathbf{x}) = \tilde{\Psi}(\mathbf{x}, z)$ as in §2.1.3.1. The model (2.1.6) is known as the *projection-approximation* as the sample-induced perturbation of the wave-field $\tilde{\Psi}$ is described in terms of line-integrals over δ and β along the incident z -direction, which can be seen as a 2D-*projection* of the 3D-sample – the *image* in fig. 2.1. δ and β determine the phase- and magnitude, respectively, of the exponential factor in (2.1.6) by which the wave-field $\tilde{\Psi}$ is modulated. Therefore, we call

$$\phi : \mathbf{x} \mapsto k \int_{-L}^0 \delta(\mathbf{x}, z) dz \quad \text{and} \quad \mu : \mathbf{x} \mapsto k \int_{-L}^0 \beta(\mathbf{x}, z) dz \quad (2.1.7)$$

the *phase-* and *absorption-images*.

The projection-approximation is valid if β and δ do not vary on lengthscales finer than $r_{\text{proj}} \sim (2\lambda L)^{1/2}$ [166, 39, 46]. This condition can be interpreted as a *resolution limit*: when aiming to resolve finer scales than r_{proj} , the approximation is no longer sufficiently accurate. On the contrary, if resolution is limited to scales $\gtrsim r_{\text{proj}}$ by other effects, which is typically the case in XPCI, it may be applied without a significant loss in accuracy. Rigorous estimates of the modeling error due to the projection-approximation are given in [108].

2.1.3.3 X-ray detection and the phase-problem

The final stage of image-formation is detection, which is subject to the *phase-problem*: real-world X-ray detectors may not measure the full complex-valued wave-field Ψ (or $\tilde{\Psi}$) but only its intensity that is (up to a scaling constant) given by the squared modulus $I = |\Psi|^2 = |\tilde{\Psi}|^2$. The physical reason becomes evident upon recalling that the actual X-ray waves $\psi(\mathbf{x}, t) = \Psi(\mathbf{x}) \exp(i\omega t)$ oscillate in time at a frequency $\omega = c_0 k$. For hard X-rays, phase-differences

[†]On the contrary, the model of free-space propagation between object and detector from §2.1.3.1 is *wave-optical* and thus incorporates diffraction. This is consistent as typically $d \gg L$ by several orders of magnitude.

in $\Psi(\mathbf{x})$ thus correspond to time-lapses between the oscillations of ψ at different positions in the order of 10^{-18} seconds or less, which is not resolvable by any available technology. On the contrary, the intensity corresponds to transferred energy per time and area and is thus accessible via *time-averaging* measurements, i.e. the time-scale of oscillations need not be resolved.

2.1.3.4 Complete model

Combining the different image-formation steps from §2.1.3.1 to §2.1.3.3 yields a description how the object-parametrizing quantities δ, β are encoded in the intensity of the resulting hologram:

$$I = \left| \tilde{\Psi}_d \right|^2 \stackrel{(2.1.5)}{=} \left| \mathcal{D}(\tilde{\Psi}_0) \right|^2 \stackrel{(2.1.6)}{=} \left| \mathcal{D} \left(\tilde{\Psi}_{-L} \cdot \exp(-\mu - i\phi) \right) \right|^2 = |\mathcal{D}(\exp(-\mu - i\phi))|^2 \quad (2.1.8)$$

with phase- and absorption-images ϕ, μ as defined by (2.1.7). The final equality in (2.1.8) follows from the assumption of incident plane waves $\Psi(\mathbf{x}, z) = \Psi_i(\mathbf{x}, z) = \exp(ikz)$ for all $z \leq -L$, which is equivalent to $\tilde{\Psi}(\mathbf{x}, z) = 1$ by definition of the wave-field-envelope $\tilde{\Psi}$.

2.1.4 Image reconstruction: the inverse problem of XPCI

Throughout this work, we are mainly concerned with *image-reconstruction* in XPCI, which corresponds to the *inverse problem* associated with the derived (forward-)model (2.1.8):

Inverse Problem 2.1 (Image reconstruction in XPCI). For some set A , reconstruct the phase- and absorption image $h = \mu + i\phi \in A$ from one or several hologram(s) of the form

$$I^{\text{obs}} \approx |\mathcal{D}(\exp(-\mu - i\phi))|^2. \quad (2.1.9)$$

The set of admissible images A depends on available *a priori* knowledge on the images ϕ, μ . Moreover, the usage of “ \approx ” in the formulation of inverse problem 2.1 emphasizes that real-world XPCI experiments never provide exact data in terms of the model (2.1.8), due to the underlying idealizations, noise and other effects discussed in §2.1.5.

As solving inverse problem 2.1 implicitly requires to recover the lost phase-information at the detector, see §2.1.3.3, it is also referred to as *phase retrieval* or *phase reconstruction*.

2.1.5 Model adjustments for real-world XPCI setups

The model of the preceding section §2.1.3 has been derived for the idealized, schematical XPCI-setup in fig. 2.1. In the following, we discuss different adjustments of this model in order to make it more applicable to real-world XPCI setups. Many of the discussed problems and conclusions drawn in this thesis may only be understood within a more realistic picture of XPCI. Therefore, a brief presentation seems necessary here.

2.1.5.1 Illumination by a point-source and the Fresnel-Scaling-Theorem

Possibly the most obvious difference of the schematical model in fig. 2.1 to the real-world experimental setup in fig. 1.2 lies in the geometry of the illuminating beam: in experimental

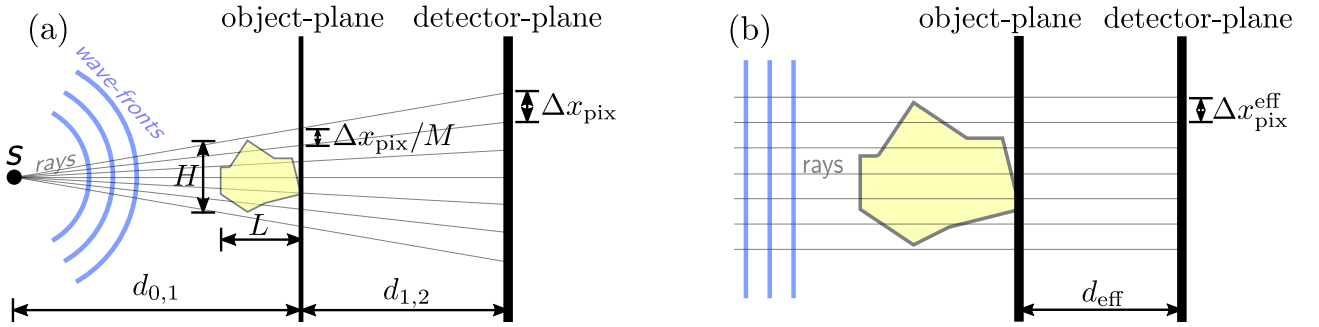


Figure 2.2. (a) Geometrical sketch of a realistic XPCI setup, characterized by a divergent X-ray beam as produced by a point-source s . Under certain conditions on the geometrical parameters $H, L, \Delta x$, the setting may be mathematically modeled to good approximation by an *effective* parallel-beam geometry as sketched in (b). For details, see text.

realizations, the object is illuminated by a *divergent beam* (or cone beam), emanating from a very small, ideally point-shaped source-spot. Even if the point-source was ideal, it would evidently not produce an X-ray wave-field given by a plane wave, but concentric *spherical* waves. In a geometrical optics picture, this means that the corresponding rays are not parallel but emanate radially from the source spot. The real-world, divergent beam geometry is sketched in fig. 2.2(a). The modified geometry affects different steps of the XPCI-model:

- *Interaction* (§2.1.3.2): The projection-approximation remains valid but the rays traversing the object are no longer parallel. However, for samples of small thickness L and lateral size H (see fig. 2.2(a)) the beam-divergence *within* the object is negligible so that the rays may be assumed to be approximately parallel. Quantitatively, assuming an approximate parallel-beam geometry as in fig. 2.2(b) is accurate down to object-features of size $\gtrsim LH/(2d_{0,1})$. As long as the resolution of the imaging-setup is limited to larger scales due to other effects, the parallel-beam model from §2.1.3.2 thus need not be modified.
- *Free-space propagation* (§2.1.3.1): By the *Fresnel-Scaling-Theorem* (see e.g. [160]), propagation from object- to detector-plane in the divergent-beam setting of fig. 2.2(a) is equivalent to propagation in the *effective* parallel-beam geometry in fig. 2.2(b) if
 - (1) Object-features are enlarged according to the geometrical magnification $M = (d_{0,1} + d_{1,2})/d_{0,1}$ of the divergent-beam setup, i.e. if $f \in \{\mu, \phi\}$ is a phase- or absorption image in the object-plane fig. 2.2(a), its effective analogue in (b) is $f_{\text{eff}}(\mathbf{x}) := f(\mathbf{x}/M)$.
 - (2) An effective propagation-distance $d_{\text{eff}} = d_{1,2}/M$ is assumed.

All in all, we see that the ideal parallel-beam model from §2.1.3 may be adjusted to real-world divergent-beam geometries at relative ease. Furthermore, we note that the magnifying effect of a divergent beam is vital for high-resolution imaging: in a true parallel-beam setup, the lateral resolution would be limited by the pixel-size Δx_{pix} of the detector, which is typically in the order of micrometers, compare §1.2. In contrast, object-features down to the *effective pixel size* $\Delta x_{\text{pix}}^{\text{eff}} = \Delta x_{\text{pix}}/M$ may be resolved in a divergent beam geometry, as seen from fig. 2.2(a).

2.1.5.2 Non-plane wave illumination and flat-field correction

Real-world X-ray sources never produce wave-fields given by ideal plane waves (or ideal spherical waves in divergent-beam geometry). Yet, the assumption may be readily relaxed within the given model by assigning an initial wave-field envelope $P := \tilde{\Psi}_{-L} \neq 1$ in the interaction-model (2.1.6). The *probe-function* P describes the lateral profile of the illuminating paraxial X-ray beam. According to (2.1.8), the image-formation model with non-trivial probe reads

$$I = |\mathcal{D}(P \cdot \exp(-\mu - i\phi))|^2. \quad (2.1.10)$$

Flat-field correction: The underlying complex-valued probe-field P in an XPCI experiment is typically unknown. While some methods have been demonstrated to enable probe reconstructions [83, 84, 170, 171], the problem is that P is *not static* in practice so that it cannot be measured once and for all. It is therefore standard to apply a heuristic correction, known as *flat-field correction* or *empty-beam division*, in order to obtain holograms that resemble those under a hypothetical illumination by plane waves. The idea is to use the approximation(!)

$$\frac{|\mathcal{D}(P \cdot \exp(-\mu - i\phi))|^2}{|\mathcal{D}(P)|^2} \approx |\mathcal{D}(\exp(-\mu - i\phi))|^2 \quad (2.1.11)$$

Both the numerator and the denominator on the l.h.s. of (2.1.11) are accessible by measurements, where the latter corresponds to $\phi = \mu = 0$ in (2.1.10), i.e. to an *empty-beam image*, acquired without a sample in the beam.

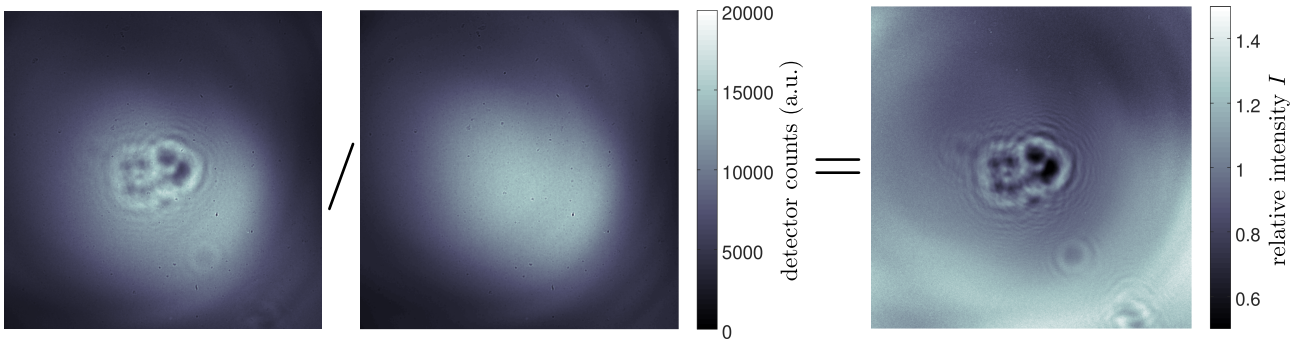


Figure 2.3. Demonstration of flat-field-correction for data acquired at the GINIX. Due to minor changes in the X-ray beam-profile between the acquisition of the hologram (left panel) and empty-beam image (middle), significant background-variations remain in corrected hologram (right panel).

There are both theoretical and practical issues associated with the flat-field correction:

- *Theoretical:* For non-plane wave probes P , (2.1.11) is only approximately valid. Analytical and numerical studies [101, 83] show however that empty-beam division is fairly accurate for the smooth probing beams produced by the waveguide in the GINIX-setup fig. 1.2, up to a *resolution limit* in the order of the source spot size.
- *Practical:* In experiments, it is difficult to produce a constant probe P over longer time-scales than minutes due to slight drifts in the setup and/or time-varying processes of the X-ray source. Hence, the “ P ” in numerator and denominator in (2.1.11) differ in practice,

especially in the case of phase contrast *tomography*, where total acquisition times are in the order of hours. Importantly, this typically leads to *imperfect* flat-field correction beyond the theoretical limitations, as seen in the real-data example shown in fig. 2.3: because of minor differences of the in the hologram and empty-beam image, spatial variations of the beam-profile are partly over- and under-compensated by the empty-beam division, so that significant background-variations remain in the (supposedly) corrected hologram.

In this thesis, we will exclusively consider holograms that have been divided by an empty-beam image. As seen from fig. 2.3, this gives rise to an important issue:

Flat-field corrected holograms (especially in XPCT) are often corrupted by slowly varying background variations that may exceed the actual diffraction-signal in magnitude.

2.1.5.3 Partial coherence

The derived XPCI-model, even with the adjustments from §2.1.5.1 and §2.1.5.2, is strictly valid only for fully coherent X-rays. In practice, one has the following situation:

- *Partial spatial coherence*: the source-spot (the waveguide-exit for the setup in fig. 1.2) has a finite extent and photons emitted from different positions do not interact fully coherently at the detector.
- *Partial temporal coherence*: the source does not emit photons of a single frequency ω (or energy $E = \hbar\omega$) but stochastically according to some probability-density $p(\omega)$ of some characteristic width. Photons of different frequency do not interact coherently.

Effects of partial coherence have been considered already in the first practical demonstrations of XPCI, see e.g. [166, 41]. Essentially, both spatial- and temporal incoherence manifest by damping the contrast for object-structures finer than some length r_{coh} depending on the degree of coherence, whereas features of size $\gtrsim r_{\text{coh}}$ are represented in the holograms quasi as if the illumination was fully coherent. Accordingly, partial coherence limits the resolution of XPCI. For the GINIX, the estimated limiting resolution is $r_{\text{coh}} = 20$ nm according to [12].

2.1.5.4 Real-world detectors and Poisson-noise

So far, it has been assumed that the full intensity $I = |\tilde{\Psi}_d|^2$ in the detector-plane is measured. In practice, different modifications to this ideal settings occur, as outlined in the following.

Finite field-of-view: Real-world detectors may cover only a bounded sub-domain $K \subset \mathbb{R}^2$ of the infinite detector-plane. The effects of this limitation are studied in detail in chapter 4.

Pixelation: Detecting devices are composed of a finite-number of *pixels*, that measure scalar intensity-values I_i according to their sensitivities $\omega_i : \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0}$ (ideally $\omega_i(\mathbf{x}) = 1$ for $\mathbf{x} \in \mathbb{D}_i$ and $\omega_i(\mathbf{x}) = 0$ otherwise, where $\mathbb{D}_i \subset \mathbb{R}^2$ is the area covered by the pixel):

$$I_i = \int_{\mathbb{R}^2} I(\mathbf{x})\omega_i(\mathbf{x}) \, d\mathbf{x} \quad \text{for } i = 1, 2, \dots, M. \quad (2.1.12)$$

Measured holograms thus provide *discrete* data $\mathbf{I} = (I_i)_{i=1}^M \in \mathbb{R}^M$. In particular, this limits the resolution by which the underlying continuous intensity I is sampled to the pixel-size Δx_{pix} . This resolution-limit carries over to reconstructed images ϕ, μ .

Poisson statistics: X-rays are quantized into discrete photons. As a consequence, ideal detector-pixels (without dead-times, electronic noise, etc.) measure *integer* numbers $I_i^{\text{obs}} \in \mathbb{N}_0$ of incident photons, that are *Poisson-distributed random variables* $I_i^{\text{obs}} \sim \text{Poi}(tI_i)$. The expected value of the I_i^{obs} is given by the product of the true intensity I_i and the exposure-time $t > 0$ (in suitable units), see e.g. [99]. The stochastic nature of the observations I_i^{obs} constitutes a natural source of *data-errors*: even for perfectly calibrated detectors without electronic noise, the Poisson-statistics of the counted photons gives rise to noisy holograms. These *Poisson-errors* are typically the predominant noise-component in XPCI.

2.2 Focus on Fresnel propagation

Mathematically, the basic image-formation model in (2.1.8) is surprisingly simple: up to the map \mathcal{D} , the relation between object-quantities ϕ, μ and the data I only involves *pointwise* operations: $(\phi, \mu) \mapsto \exp(-\mu - i\phi)$ and $\mathcal{D}(\dots) \mapsto |\mathcal{D}(\dots)|^2$. As a consequence, quasi all non-trivial properties of the map $(\phi, \mu) \mapsto I$ relate to the Fresnel propagator \mathcal{D} . The following section is therefore dedicated to gaining a physical and mathematical understanding of Fresnel propagation.

2.2.1 On Fresnel numbers and imaging-regimes

2.2.1.1 Why propagation?

For a first insight on the significance of Fresnel propagation, it is illustrative to investigate what would happen if the detector in an XPCI-setup was placed directly behind the sample, i.e. $d = 0$ in fig. 2.1, such that the X-rays do not undergo free-space propagation after interacting with the sample. This eliminates \mathcal{D} from the image-formation model (2.1.8), so that

$$I = |\mathcal{D}(\exp(-\mu - i\phi))|^2 = |\exp(-\mu - i\phi)|^2 = \exp(-2\mu). \quad (2.2.1)$$

Accordingly, the intensities I would only encode the absorption-image μ , but are completely insensitive to the phase-image ϕ . In other words, *Fresnel propagation enables phase contrast*. Details of contrast-formation are determined by the *Fresnel number*(s).

2.2.1.2 Meaning of the Fresnel number

The Fresnel number $\mathfrak{f} = b^2/(\lambda d)$ has been introduced as a dimensionless condensation of the setup parameters in XPCI. Its physical meaning can be seen from the definition of the Fresnel propagator in (2.1.5): the action of \mathcal{D} modulates the Fourier-components $\mathcal{F}(f)(\boldsymbol{\xi})$ of an image f by the phase-factor $\exp(-i\boldsymbol{\xi}^2/(4\pi\mathfrak{f}))$, which implies the following:

- For *low spatial frequencies* $|\boldsymbol{\xi}| \lesssim \mathfrak{f}^{1/2}$, it holds that $\exp(-i\boldsymbol{\xi}^2/(4\pi\mathfrak{f})) \approx 1$ so that the modulation is quasi without effect. Accordingly, image-features of size coarser than $\sim \mathfrak{f}^{-1/2}$ are practically retained upon Fresnel propagation.

- For *high spatial frequencies* $|\boldsymbol{\xi}| \gtrsim \mathfrak{f}^{1/2}$, the assigned phases $\exp(-i\boldsymbol{\xi}^2/(4\pi\mathfrak{f}))$ differ significantly from unity and oscillate with the frequency $|\boldsymbol{\xi}|$. Hence, image-features finer than $\sim \mathfrak{f}^{-1/2}$ are significantly affected by Fresnel propagation.

The physical effect modeled by \mathcal{D} is coherent (near-field) *diffraction*, i.e. self-interference of the propagating X-ray wave-field. The above observations show that wave-field structures smaller than some specific scale are strongly distorted by diffraction (*holographic lengthscales*), whereas coarser, *non-holographic scales* only exhibit minor diffractive effects. The Fresnel number \mathfrak{f} defines the threshold between holographic and non-holographic scales.

2.2.1.3 Scale-dependence of the Fresnel number

It is important to note that *the* Fresnel number of an imaging-setup does not exist. Instead, a Fresnel number is always assigned with (implicit) reference to a lateral lengthscale b that defines unit-length in the underlying dimensionless coordinates, see §2.1.3.1. The Fresnel number is thus *scale-dependent* as can be seen from the behavior of the Fresnel propagator under rescaling of the coordinates: if $\mathcal{D}^{(\mathfrak{f}_*)}$ denotes the propagator for some specific $\mathfrak{f} = \mathfrak{f}_*$, then it holds for any function $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ and $f_\sigma(\mathbf{x}) := f(\mathbf{x}/\sigma)$ with $\sigma > 0$ that

$$\mathcal{D}^{(\mathfrak{f})}(f_\sigma)(\mathbf{x}) = \mathcal{D}^{(\sigma^2\mathfrak{f})}(f)(\mathbf{x}/\sigma) \quad (2.2.2)$$

The interpretation of (2.2.2) is as follows: $\mathcal{D} = \mathcal{D}^{(\mathfrak{f})}$ acts on features of size σ in the same manner as $\mathcal{D}^{(\mathfrak{f}_\sigma)}$ with $\mathfrak{f}_\sigma := \sigma^2\mathfrak{f}$ on features of size 1. Accordingly, each length scale σ has its own Fresnel number \mathfrak{f}_σ , that governs near-field diffraction on this particular scale.

Assigning a Fresnel number to different scales enables a compact description of the observations in §2.2.1.2: the identified threshold-scale $\sigma = \mathfrak{f}^{-1/2}$ for diffractive effects corresponds to a Fresnel number $\mathfrak{f}_\sigma = \sigma^2\mathfrak{f} = 1$. Thus, a lengthscale σ is holographic if and only if $\mathfrak{f}_\sigma \lesssim 1$.

2.2.1.4 Propagation-regimes

The phase- and absorption-images ϕ and μ for real-world samples typically involve features on a wide range of lengthscales: from the coarsest scale, defined by the diameter of the total object, to the smallest resolvable lengthscale, which may for example be the effective pixel-size[‡] $\Delta x_{\text{pix}}^{\text{eff}}$, compare 2.1.5.1. The coarsest and finest scales define the largest- and smallest relevant Fresnel numbers $\mathfrak{f}_{\text{max}}, \mathfrak{f}_{\text{min}}$, respectively. Based on these, *imaging-regimes* of XPCI may be defined. For orientation, fig. 2.4 shows exemplary (simulated) holograms for some of the settings:

- (1) *Contact regime* $\mathfrak{f}_{\text{min}} \gg 1$: There are no holographic scales and thus no detectable effects of Fresnel propagation. Hence, the model (2.2.1) applies and there is *no phase contrast*.
- (2) *Direct contrast regime* $\mathfrak{f}_{\text{min}} \sim 1, \mathfrak{f}_{\text{max}} \gg 1$: Only the smallest scales are holographic, so that phase contrast is visible only for the sharpest image-features (edges) (fig. 2.4(c),(d)).
- (3) *Holographic regime* $\mathfrak{f}_{\text{min}} \ll 1 \ll \mathfrak{f}_{\text{max}}$: There is a wide range of both holographic- and non-holographic scales (fig. 2.4(e),(f)).

[‡]Indeed, it is very common to express the Fresnel number of an XPCI-setup w.r.t. the lengthscale $b = \Delta x_{\text{pix}}^{\text{eff}}$. We denote this pixel-size-based Fresnel number as $\mathfrak{f}_{\text{pix}}$.

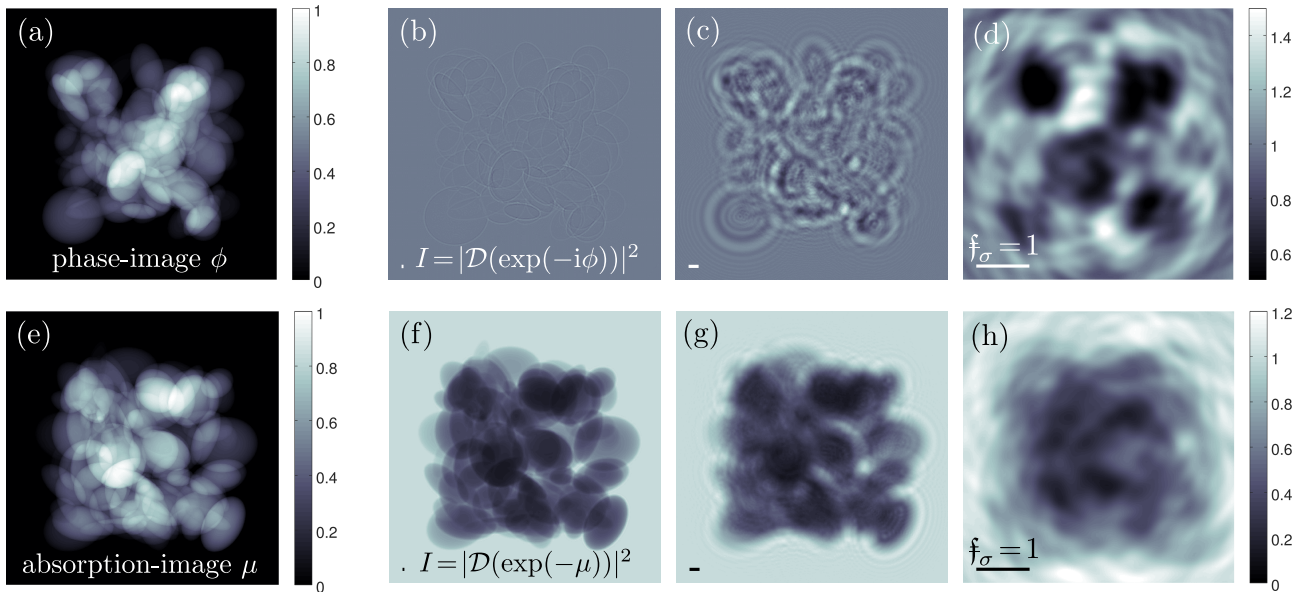


Figure 2.4. Simulated holograms at different Fresnel numbers \mathfrak{f} for pure phase contrast ($\mu = 0$, top row) and pure absorption contrast ($\phi = 0$, bottom row). (a),(e) assumed phase- or absorption image ϕ, μ . (b)–(d), (f)–(h) resulting hologram-intensities. The scale-bars indicate the length σ that corresponds to a Fresnel number $\mathfrak{f}_\sigma = 1$. The holograms are characteristic of the propagation-regimes **(2)** (images (b),(f)), **(3)** (images (c),(g)) and **(4)** (images (d),(h)), respectively, defined in §2.2.1.4.

- (4) *Deeply holographic regime*: Only the coarsest scales are non-holographic (fig. 2.4(g),(h)).
- (5) *Completely holographic regime* $\mathfrak{f}_{\max} \sim 1$: All relevant scales are holographic.
- (6) *Far-field regime* $\mathfrak{f}_{\max} \ll 1$: \mathcal{D} can be approximated by the far-field limit, see §2.2.2.2.

In this thesis, we will *exclusively* consider XPCI-data from the settings **(3)** and **(4)**, which are the standard regimes for high-resolution XPCI at synchrotrons. Setting **(1)** is the regime of classical X-ray radiography of macroscale objects such as entire human organs, whereas the regime **(2)** applies to X-ray imaging down to micrometer resolutions (micro-CT, small-animal-imaging, ...) as routinely performed with laboratory X-ray sources, see e.g. [120] for details. Possibilities and potential benefits of imaging in regime **(5)** will be discussed in §3.6.

From fig. 2.4(e)–(h), it is seen that *absorption contrast* is strong in holograms for any Fresnel number \mathfrak{f} – only the structural similarity between hologram and absorption-image μ decreases as the setting becomes more holographic. For *phase contrast* (fig. 2.4(a)–(d)), the overall contrast in the hologram seems to increase as \mathfrak{f} becomes smaller. However, the true picture is more subtle, in accordance with the scale-dependence of Fresnel propagation discussed in §2.2.1.3:

Non-holographic scales induce low phase contrast (low-frequency instability of XPCI).

Thus, the perceived contrast-increase is thus due to more and more scales becoming holographic as \mathfrak{f} decreases – phase contrast on holographic scales in general does *not* increase with $1/\mathfrak{f}$.

2.2.1.5 The holographic nature of XPCI and delineation from CDI

Regime **(6)** in §2.2.1.4 is the setting of *coherent diffractive imaging* (CDI), see e.g. [146, 147, 182], which can be seen as the far-field limit of the *near-field* technique XPCI. Yet, a fundamental change occurs in the transition from XPCI to CDI as explained in the following.

As seen in the raw holograms in fig. 2.3, both the unscattered X-ray beam and the sample-induced scattering-signal manifest in XPCI-data. In the notation of §2.1.5.2, we have

$$I = |\mathcal{D}(P \cdot \exp(-\mu - i\phi))|^2 = |\Psi_t|^2 + 2\operatorname{Re}(\overline{\Psi_t} \cdot \Psi_s) + |\Psi_s|^2, \quad (2.2.3)$$

where $\Psi_t = \mathcal{D}(P)$ and $\Psi_s = \mathcal{D}(P \cdot (\exp(-\mu - i\phi) - 1))$ are the transmitted (primary) and scattered (secondary) wave-fields, respectively, at the detector. Image reconstruction in XPCI mostly uses the term $2\operatorname{Re}(\overline{\Psi_t} \cdot \Psi_s)$ which models *interference* between primary- and secondary wave, exploiting the *holographic reference-signal* provided by the transmitted beam Ψ_t to determine the unknown scattered part Ψ_s .

In CDI, on the contrary, the contributions $|\Psi_t|^2$ and $2\operatorname{Re}(\overline{\Psi_t} \cdot \Psi_s)$ are typically negligible in the recorded far-field diffraction data for reasons detailed in e.g. [139, §2.5.3], so that (2.2.3) reduces to $I \approx |\Psi_s|^2$, i.e. holographic signal-components do not manifest. As a consequence, the mathematical structure of image reconstruction in CDI strongly differs from XPCI. Thus:

Theory and algorithms derived in this thesis in general do not generalize to CDI.

2.2.1.6 Modified Fresnel numbers – for notational convenience

It is convention to work with Fresnel numbers $\mathfrak{f}_{(b)} = b^2(\lambda d)$, using the wavelength λ as reference-scale. Yet, it often turns out to be more convenient to instead employ the reciprocal wavenumber $1/k$ as a lengthscale, which leads to the *modified Fresnel number*

$$\mathfrak{f}_{(b)} := kb^2/d = 2\pi \bar{\mathfrak{f}}_{(b)}. \quad (2.2.4)$$

As indicated by the extra “bar”, the difference in definition between $\mathfrak{f}_{(b)}$ and $\bar{\mathfrak{f}}_{(b)}$ is analogous to Planck’s constants h and $\hbar = h/(2\pi)$. Defining the Fresnel propagator in terms of the modified Fresnel number tends to eliminate 2π -factors from the arising formulas, which simplifies lengthy expressions. We will mostly use $\mathfrak{f}_{(b)}$, but sometimes also $\bar{\mathfrak{f}}_{(b)}$ when it seems suitable. For brevity, we will often omit the prefix “modified” and simply refer to $\mathfrak{f}_{(b)}$ as “Fresnel number”.

2.2.2 Mathematical properties of the Fresnel propagator

In the second part of this section, we present some fundamental mathematical properties of the Fresnel propagator \mathcal{D} . Derivations of the stated properties can be found for example in [160].

2.2.2.1 Definition and basic properties on \mathbb{R}^m

Although \mathcal{D} arises naturally in a two-dimensional setting, it will turn out to be useful to generalize the definition to arbitrary dimensions \mathbb{R}^m :

$$\mathcal{D} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m); f \mapsto \mathcal{F}(m_{\mathfrak{f}} \cdot \mathcal{F}^{-1}(f)) \quad \text{with} \quad m_{\mathfrak{f}}(\boldsymbol{\xi}) := \exp(-i\boldsymbol{\xi}^2/(2\mathfrak{f})) \quad (2.2.5)$$

defines a bounded linear operator. Moreover, \mathcal{D} has the following properties:

- *Unitarity*: \mathcal{D} preserves the L^2 -norm $\|f\|_{(L^2)} := (\int_{\mathbb{R}^m} |f(\mathbf{x})|^2 d\mathbf{x})^{1/2}$ of functions,

$$\|\mathcal{D}(f)\|_{L^2} = \|f\|_{L^2} \quad \text{for all } f \in L^2(\mathbb{R}^m). \quad (2.2.6)$$

Its inverse/adjoint is given by back-propagation: $\mathcal{D}^{-1}(f) = \mathcal{D}^*(f) = \mathcal{F}^{-1}(m_{-\mathfrak{f}} \cdot \mathcal{F}(f))$.

- *Translation invariance*: as a *Fourier-multiplier*, i.e. an operator given by a multiplication in Fourier space, \mathcal{D} is translation-invariant, i.e. commutes with coordinate-shifts:

$$\mathcal{D}T_{\mathbf{a}} = T_{\mathbf{a}}\mathcal{D} \quad \text{for all } \mathbf{a} \in \mathbb{R}^m, T_{\mathbf{a}} : f \mapsto f(\cdot + \mathbf{a}). \quad (2.2.7)$$

- *Rotational invariance*: as $m_{\mathfrak{f}}(\boldsymbol{\xi})$ only depends on the modulus $|\boldsymbol{\xi}|$, \mathcal{D} is furthermore invariant under orthogonal coordinate-transforms (rotations and reflections):

$$\mathcal{D}\mathcal{A} = \mathcal{A}\mathcal{D} \quad \text{for all } \mathcal{A} : f \mapsto f(\mathcal{A}(\cdot)), \mathcal{A} \in \mathbb{R}^{m \times m} : \mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A} = \text{id}_{\mathbb{R}^m}. \quad (2.2.8)$$

2.2.2.2 Alternate forms

Fourier-multipliers may be expressed as convolution-operators (with a possibly singular kernel). The *convolution-form* of the Fresnel propagator \mathcal{D} reads ($f \in L^2(\mathbb{R}^m) \cap L^1(\mathbb{R}^m)$)

$$\begin{aligned} \mathcal{D}(f)(\mathbf{x}) &= (n_{\mathfrak{f}} * f)(\mathbf{x}) = \int_{\mathbb{R}^m} k_{\mathfrak{f}}(\mathbf{x} - \mathbf{y})f(\mathbf{y})d\mathbf{y} \quad \text{for all } \mathbf{x} \in \mathbb{R}^m \\ k_{\mathfrak{f}} &= u_0\mathfrak{f}^{\frac{m}{2}} \cdot n_{\mathfrak{f}}, \quad n_{\mathfrak{f}}(\mathbf{x}) = \exp(i\mathfrak{f}\mathbf{x}^2/2), \quad u_0 = \exp(-im\pi/4). \end{aligned} \quad (2.2.9)$$

By manipulating the convolution-integral in (2.2.9), the following *alternate form* of the Fresnel propagator may be obtained:

$$\mathcal{D}(f)(\mathbf{x}) = u_0\mathfrak{f}^{\frac{m}{2}} n_{\mathfrak{f}}(\mathbf{x}) \cdot \mathcal{F}(n_{\mathfrak{f}} \cdot f)(\mathfrak{f}\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^m. \quad (2.2.10)$$

Far-field limit: An advantage of the alternate form (2.2.10) is that one may readily read-off the behavior of \mathcal{D} in the far-field limit $\mathfrak{f} \rightarrow 0$: if f has compact support, w.l.o.g. $f(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^m$ with $|\mathbf{x}| > \frac{1}{2}$, and $\mathfrak{f} \ll 1$, then $n_{\mathfrak{f}} \cdot f \approx f$ holds true so that

$$\mathcal{D}(f)(\mathbf{x}) \approx u_0\mathfrak{f}^{\frac{m}{2}} n_{\mathfrak{f}}(\mathbf{x}) \cdot \mathcal{F}(f)(\mathfrak{f}\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^m. \quad (2.2.11)$$

In the far-field limit, \mathcal{D} is thus essentially given by a Fourier transform, up to a rescaling of the coordinates and multiplication with the function $u_0\mathfrak{f}^{\frac{m}{2}} n_{\mathfrak{f}}$, that has constant modulus.

2.3 Tomographic imaging model

According to the physical model derived in §2.1, holograms encode *projection-images* $\phi = k \int \delta dz$ and $\mu = k \int \beta dz$ of the sample's refraction- and absorption-decrements δ, β . Notably, this implies that the imaging-setup is completely insensitive to variations of δ, β along the incident z -direction. In order to resolve the full 3D object-structure, XPCI has to be combined with *tomographic* techniques: by rotating the sample in the setup as suggested in fig. 1.2, holograms may be acquired for different *incident directions* $\boldsymbol{\theta} \in \mathbb{S}^2$ ($\mathbb{S}^{m-1} = \{\mathbf{x} \in \mathbb{R}^m : |\mathbf{x}| = 1\}$)

of the X-rays relative to the object (in §2.1, $\boldsymbol{\theta}$ was assumed to coincide with the z -axis). The encoded images are then given by projections of δ, β along $\boldsymbol{\theta}$, i.e. $\phi_{\boldsymbol{\theta}} = k\mathcal{P}_{\boldsymbol{\theta}}(\delta)$ and $\mu_{\boldsymbol{\theta}} = k\mathcal{P}_{\boldsymbol{\theta}}(\beta)$ with *parallel-beam projector* $\mathcal{P}_{\boldsymbol{\theta}}$ defined by

$$\mathcal{P}_{\boldsymbol{\theta}}(f)(x, y) := \int_{\mathbb{R}} f(x\mathbf{n}_x + y\mathbf{n}_y + z\boldsymbol{\theta}) \, dz \quad \text{for all } x, y \in \mathbb{R}, \quad (2.3.1)$$

where $\mathbf{n}_x, \mathbf{n}_y \in \mathbb{S}^2$ are chosen such that $\{\boldsymbol{\theta}, \mathbf{n}_x, \mathbf{n}_y\}$ forms an orthonormal basis of \mathbb{R}^3 . The considered tomographic geometry and the action of the projector is illustrated in fig. 2.5.

Inverse problem: Projections $\{\mathcal{P}_{\boldsymbol{\theta}}(f)\}_{\boldsymbol{\theta} \in \Theta}$ for a suitable set of incident directions $\Theta \subset \mathbb{S}^2$ allow to reconstruct the underlying 3D-density f . This is the principle of computed tomography (CT), which is routinely exploited in medical CT-scanners as well as in several other imaging modalities. We are concerned with the inverse problem of *X-ray phase contrast tomography* (XPCT) which, by the basic XPCI-model (2.1.8) and the argument above, reads as follows:

Inverse Problem 2.2 (Image reconstruction in XPCT). For some set A , reconstruct the 3D-density $f := k\beta + ik\delta \in A$ from a tomographic series of holograms $\{I_{\boldsymbol{\theta}}^{\text{obs}}\}_{\boldsymbol{\theta} \in \Theta}$, $\Theta \subset \mathbb{S}^2$ with

$$I_{\boldsymbol{\theta}}^{\text{obs}} \approx |\mathcal{D}(\exp(-\mathcal{P}_{\boldsymbol{\theta}}(f)))|^2 \quad \text{for all } \boldsymbol{\theta} \in \Theta. \quad (2.3.2)$$

Inverse problem 2.2 amounts to a combination of the image reconstruction in XPCI (inverse problem 2.1) and classical tomographic reconstruction as performed by CT-scanners.

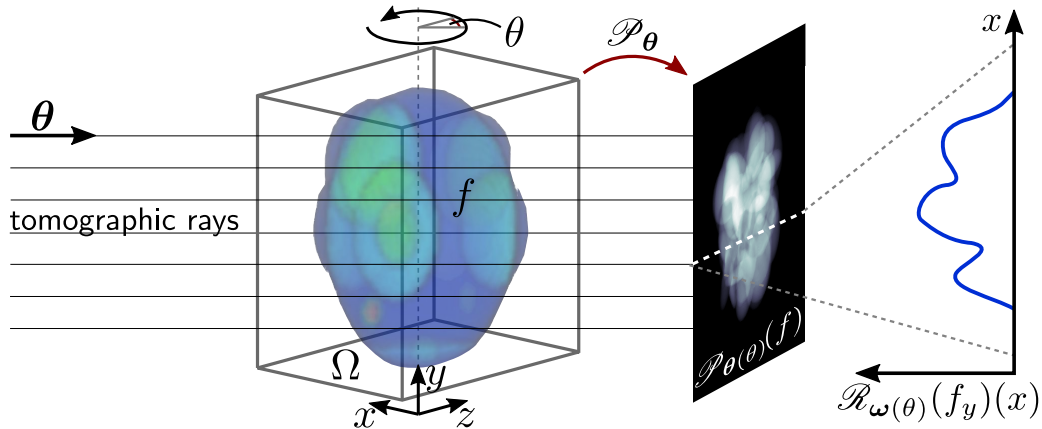


Figure 2.5. Geometrical sketch of the tomographic imaging model used to describe XPCT: the parallel-beam projectors $\mathcal{P}_{\boldsymbol{\theta}}$ integrate a 3D-object density $f (= k\beta + ik\delta)$ along the incident direction $\boldsymbol{\theta} \in \mathbb{S}^2$ of the X-rays, yielding 2D-projection images $p_{\boldsymbol{\theta}} = \mathcal{P}_{\boldsymbol{\theta}}(f)$ (which then induce holograms $I_{\boldsymbol{\theta}} = |\mathcal{D}(\exp(-p_{\boldsymbol{\theta}}))|^2$). Lines in the image $p_{\boldsymbol{\theta}}$ provide values of the Radon-transform \mathcal{R} of 2D-slices $f_y : (x, z) \mapsto f(x, y, z)$ of the 3D-object f . For details, see text.

2.3.1 Relation to the Radon transform

The most common approach to acquire tomographic data sets is to rotate the sample around a single fixed axis, the so-called *tomographic axis* or *tilt axis* (w.l.o.g. the y -axis throughout this

thesis). The directions $\boldsymbol{\theta}$ are then parametrized by the *incident-angle* $\theta \in [0; 2\pi)$ enclosed with the z -axis, as indicated in fig. 2.5, and we write

$$\mathcal{P}_\theta := \mathcal{P}_{\boldsymbol{\theta}(\theta)} \quad \text{with} \quad \boldsymbol{\theta}(\theta) := (\sin(\theta), 0, \cos(\theta)) \quad \text{for} \quad \theta \in \mathbb{R} \quad (2.3.3)$$

We define by $f_y : (x, z) \mapsto f(x, y, z)$ two-dimensional *slices* of a three-dimensional function $f : \mathbb{R}^3 \rightarrow \mathbb{C}$ in planes perpendicular to the tomographic axis. Moreover, we identify θ with a direction in \mathbb{R}^2 , setting $\boldsymbol{\omega}(\theta) := (-\cos(\theta), \sin(\theta)) \in \mathbb{S}^1$. Then it holds that

$$\mathcal{P}_{\boldsymbol{\theta}(\theta)}(f)(x, y) = \mathcal{R}_{\boldsymbol{\omega}(\theta)}(f_y)(x) = \mathcal{R}(f_y)(\boldsymbol{\omega}(\theta), x) \quad \text{for all} \quad x, y, \theta \in \mathbb{R}, \quad (2.3.4)$$

where $\mathcal{R}_{\boldsymbol{\omega}}$ denotes the two-dimensional *Radon transform* (for fixed $\boldsymbol{\omega} \in \mathbb{S}^1$), as defined in standard literature on computed tomography [155, 110]:

$$\mathcal{R}_{\boldsymbol{\omega}}(f^{(2d)})(x) := \mathcal{R}(f^{(2d)})(\boldsymbol{\omega}, x) := \int_{\mathbb{R}} f^{(2d)}(x\boldsymbol{\omega} + z\boldsymbol{\omega}_\perp) dz. \quad (2.3.5)$$

for integrable functions $f^{(2d)} : \mathbb{R}^2 \rightarrow \mathbb{C}$, mutually orthogonal directions $\boldsymbol{\omega}, \boldsymbol{\omega}_\perp \in \mathbb{S}^1$ and $x \in \mathbb{R}$. According to (2.3.4), lines $\mathcal{P}_{\boldsymbol{\theta}(\theta)}(f)(\cdot, y)$ of projection images provide Radon-data $\mathcal{R}_{\boldsymbol{\omega}(\theta)}(f_y)$ of the corresponding 2D-slice f_y , as visualized in fig. 2.5.

2.3.2 Analytical properties

Based on the identification with a (slice-wise) 2D-Radon transform, a vast amount of theory is available on parallel-beam tomography. Here, we recall the most important aspects to which will be referred throughout this thesis. Details and proofs of the given statements can be found e.g. in the books [155, 110] as well as in the author's master's thesis [139].

2.3.2.1 Boundedness and adjoints (backprojections)

For any bounded domains $\Omega \subset \mathbb{R}^3$ and $\Omega^{(2d)} \subset \mathbb{R}^2$ and $\boldsymbol{\theta} \in \mathbb{S}^2, \boldsymbol{\omega} \in \mathbb{S}^1$, the maps

$$\mathcal{P}_\boldsymbol{\theta} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^2) \quad \text{and} \quad \mathcal{R} : L^2(\Omega^{(2d)}) \rightarrow L^2(\mathbb{S}^1 \times \mathbb{R}) \quad (2.3.6)$$

define bounded linear operators. The corresponding *adjoints* are given by

$$\mathcal{P}_\boldsymbol{\theta}^*(p)(\boldsymbol{x}) = \begin{cases} p(\boldsymbol{n}_x \cdot \boldsymbol{x}, \boldsymbol{n}_y \cdot \boldsymbol{x}) & \text{if } \boldsymbol{x} \in \Omega \\ 0 & \text{else} \end{cases} \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^3, \quad (2.3.7a)$$

$$\mathcal{R}^*(g)(\boldsymbol{x}) = \begin{cases} \int_{\mathbb{S}^1} g(\boldsymbol{\omega}, \boldsymbol{\omega} \cdot \boldsymbol{x}) d\boldsymbol{\omega} & \text{if } \boldsymbol{x} \in \Omega^{(2d)} \\ 0 & \text{else} \end{cases} \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^2. \quad (2.3.7b)$$

The operations in (2.3.7) are *back-projections*, that uniformly smear out projections along the incident directions associated with $\boldsymbol{\theta}, \boldsymbol{\omega}$ (integrated over all $\boldsymbol{\omega}$ in the case of \mathcal{R}^*).

2.3.2.2 Fourier-Slice-Theorem

For the theory of tomographic reconstruction, it is essential to quantify what information a 2D-projection provides on the underlying 3D-density. This question was essentially already answered by JOHANN RADON himself [168] in terms of the following theorem:

Theorem 2.1 (Fourier-Slice-Theorem). *For $\Omega \subset \mathbb{R}^3, \Omega^{(2d)} \subset \mathbb{R}^2$ bounded, let $f \in L^2(\Omega)$, $f^{(2d)} \in L^2(\Omega^{(2d)})$ and $\boldsymbol{\theta} \in \mathbb{S}^2, \boldsymbol{\omega} \in \mathbb{S}^1$. Then it holds that*

$$\mathcal{F}(\mathcal{R}_{\boldsymbol{\omega}}(f^{(2d)}))(\xi) = (2\pi)^{1/2} \mathcal{F}(f^{(2d)})(\xi\boldsymbol{\omega}) \quad \text{for all } \xi \in \mathbb{R} \quad (2.3.8)$$

$$\mathcal{F}(\mathcal{P}_{\boldsymbol{\theta}}(f))(\xi_x, \xi_y) = (2\pi)^{1/2} \mathcal{F}(f)(\xi_x \mathbf{n}_x + \xi_y \mathbf{n}_y) \quad \text{for all } (\xi_x, \xi_y) \in \mathbb{R}^2. \quad (2.3.9)$$

The classical result (2.3.8) states that Radon-data $\mathcal{R}_{\boldsymbol{\omega}}(f^{(2d)})$ provides samples of the Fourier transform $\mathcal{F}(f^{(2d)})$ on lines through the origin $\{\xi\boldsymbol{\omega} : \xi \in \mathbb{R}\} \subset \mathbb{R}^2$. (2.3.9) generalizes this, stating that projections of 3D-densities provide Fourier-data on planes through the origin.

2.3.2.3 Range conditions and tomographic consistency

Projections $\mathcal{P}_{\boldsymbol{\theta}_1}(f), \mathcal{P}_{\boldsymbol{\theta}_2}(f)$ for $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2 \in \mathbb{S}^2$ correspond to views of one and the same object under different perspectives. As a consequence, such images are not arbitrary but correlated in a characteristic manner. These correlations are exhaustively described by the HELGASON-LUDWIG consistency conditions for the Radon transform [91, 134]:

Theorem 2.2 (Tomographic consistency conditions [91, 134]). *For $\Omega^{(2d)} \subset \mathbb{R}^2$ bounded, let $f^{(2d)} \in L^2(\Omega^{(2d)})$. Then the following holds in a Lebesgue “almost-everywhere” sense:*

- (1) $\mathcal{R}(f^{(2d)})$ is an even function: $\mathcal{R}(f^{(2d)})(-\boldsymbol{\omega}, -x) = \mathcal{R}(f^{(2d)})(\boldsymbol{\omega}, x)$ for all $\boldsymbol{\omega} \in \mathbb{S}^1, x \in \mathbb{R}$.
- (2) For all $\ell \in \mathbb{N}_0$, the ℓ -th order moments of $\boldsymbol{\omega} \mapsto \int_{\mathbb{R}} x^\ell \mathcal{R}(f^{(2d)})(\boldsymbol{\omega}, x) dx$ are homogeneous polynomials of order ℓ in $\boldsymbol{\omega}$.

Theorem 2.2 shows that tomographic data is highly structured and contains *redundancies*: certain image-components of individual projections may be inferred from data at other tomographic angles via (2). On the other hand, a well-known consequence of (1) is that Radon-data need not be measured for all $\boldsymbol{\omega} \in \mathbb{S}^1$, corresponding to a *full rotation* of the object in a tomographic setup, but data for *half* a rotation is sufficient.

The HELGASON-LUDWIG conditions can be shown to be sufficient in a suitable setting (see e.g. [155, section II.4]): data $g(\boldsymbol{\omega}, x)$ that satisfies (1) and (2) is the Radon-transform of a function. Moreover, theorem 2.2 may be readily adapted to 3D-projectors $\mathcal{P}_{\boldsymbol{\theta}}$ via (2.3.4).

2.3.3 Full angular sampling

The relation to the Radon-transform and the exhaustive amount of theory available for the latter motivates to consider the special case of tomographic data with *full angular sampling*: projections $\{\mathcal{P}_{\boldsymbol{\theta}}(f)\}_{\boldsymbol{\theta} \in \Theta}$ acquired over a *densely sampled* range of incident-angles $\Theta \subset \mathbb{R}$ covering 180 degrees or more, i.e. with Θ modulo $\pi = [0; \pi)$, provide complete Radon-data $\mathcal{R}(f_y)$

for any slice f_y of f . By theorem 2.1, this is furthermore equivalent to obtaining complete representation of the Fourier transform of the object-density $\mathcal{F}(f)$, which renders the analysis of the tomographic reconstruction problem $\{\mathcal{P}_\theta(f)\}_{\theta \in \Theta} \mapsto f$ particularly simple. Therefore, we model this important setting by an individual forward operator:

$$\mathcal{P}_{\text{full}} : L^2(\Omega) \rightarrow L^2([0; \pi) \times \mathbb{R}^2); \quad \mathcal{P}_{\text{full}}(f)(\theta, \mathbf{x}) = \mathcal{P}_\theta(\mathbf{x}). \quad (2.3.10)$$

Well-definedness, continuity and other properties of $\mathcal{P}_{\text{full}}$ may be readily adapted from §2.3.2.

In practice, projections may clearly only be acquired for a *finite* set of incident angles $\Theta = (\theta_1, \theta_2, \dots, \theta_{N_{\text{proj}}})$. Provided that the θ_j (modulo π) sample the interval $[0; \pi)$ in a sufficiently uniform and dense manner, however, the true tomographic operator $\mathcal{P}_\Theta : f \mapsto (\mathcal{P}_\theta(f))_{\theta \in \Theta}$ is typically well-approximated by the idealization $\mathcal{P}_{\text{full}}$. Notably, this approximation, which is studied in detail for example in [155, §V.1], provides the basis for tomographic reconstruction via filtered back-projection methods, see §2.7.3.1.

In this thesis, we will often use $\mathcal{P}_{\text{full}}$ in the theoretical analysis of XPCT, taking advantage of its favorable analytical properties induced by the close relation to $\mathcal{R}(f_y)$ and $\mathcal{F}(f)$. On the contrary, the proposed practical reconstruction algorithms do *not* rely on any specific angular sampling scheme but can be applied for *arbitrary* tomographic projectors $\{\mathcal{P}_\theta\}_{\theta \in \Theta}$, $\Theta \subset \mathbb{S}^2$.

2.4 Forward operators

The physical model from §2.1 describes the basic relation between the sought images and the measurable data in XPCI. For analysis by the tools of inverse problems, it is convenient to phrase this model in terms of a *forward operator* $F : X \rightarrow Y$ mapping between the space of admissible objects X and the space of possible data Y . This approach also provides a natural language to distinguish different settings encountered in XPCI, characterized by different types of available data and additional a priori knowledge on the imaged sample. Similarly, the approach may be adapted to describe *XPCT* by combining the forward maps of XPCI with the tomographic imaging model from §2.3.

2.4.1 Basic nonlinear forward model

The basic forward operator of XPCI is defined by (2.1.8): if $h = \mu + i\phi$ denotes the phase- and absorption-image, the induced *contrast* in the hologram I is modeled by the map

$$\mathcal{N} : h \mapsto |\mathcal{D}(\exp(-h))|^2 - 1 = I - 1. \quad (2.4.1)$$

Notably, \mathcal{N} is *nonlinear* due to the pointwise exponential and the squared modulus operation. Moreover, the model is non-injective for certain choices of the object-space $X \ni h$, i.e. μ and ϕ may not be unambiguously recovered from the data I , as will be seen in §2.5.1. For these reasons, different modifications of the basic forward map \mathcal{N} are introduced in the following.

2.4.2 Linearized CTF-model

If the interaction of the X-rays with imaged object is sufficiently weak, the resulting image h will be small in a suitable sense. In such a setting, the induced contrast in the hologram I may

be linearized to good approximation, leading to a *linear forward map* \mathcal{T} :

$$I - 1 = \mathcal{N}(h) = \mathcal{T}(h) + \mathcal{O}(h^2) \quad \text{with} \quad \mathcal{T}(h) := -2\text{Re}(\mathcal{D}(h)). \quad (2.4.2)$$

The linearization in (2.4.2) is commonly known as *CTF-model* and widely used in practical XPCI [42, 194, 124, 96, 122, 121]. The acronym ‘‘CTF’’ refers to the *contrast-transfer-functions* s_0 and c_0 , plotted in fig. 2.6, that describe the phase- and absorption contrast in Fourier space:

$$\mathcal{T}(\mu + i\phi) = -2\mathcal{F}^{-1} \left(\underbrace{\sin\left(\frac{|\boldsymbol{\xi}|^2}{2f}\right)}_{=:s_0(\boldsymbol{\xi})} \mathcal{F}(\phi) + \underbrace{\cos\left(\frac{|\boldsymbol{\xi}|^2}{2f}\right)}_{=:c_0(\boldsymbol{\xi})} \mathcal{F}(\mu) \right) =: \mathcal{S}(\phi) + \mathcal{C}(\mu). \quad (2.4.3)$$

As discussed in [194], the linear CTF-model is valid for images $h = \mu + i\phi$ with weak absorption $\mu \ll 1$ and *slowly varying* phase, characterized by sufficiently small phase-gradients $|\nabla\phi|$.

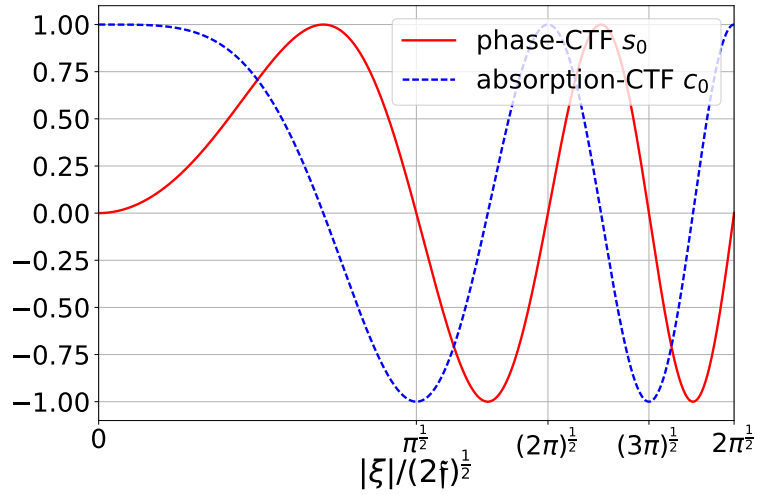


Figure 2.6. Plot of the (radially symmetric) contrast-transfer-functions (CTFs) s_0 , c_0 from (2.4.3).

According to (2.4.2) and (2.4.3), the Fourier transform of the linearized contrast is given by a superposition of the Fourier transforms of phase- and absorption-image ϕ , μ , modulated by the CTFs s_0 and c_0 , respectively:

$$\mathcal{F}(I - 1) \approx \mathcal{F}(\mathcal{T}(h)) = -2s_0 \cdot \mathcal{F}(\phi) - 2c_0 \cdot \mathcal{F}(\mu). \quad (2.4.4)$$

For once, this implies that it is generally hard to clearly attribute a feature in the data to either ϕ or μ , which raises the question of *uniqueness* of image reconstruction. Moreover, the *zeros* of the oscillatory functions s_0 and c_0 , which are plotted in fig. 2.6, correspond to Fourier-components $\mathcal{F}(\phi)(\boldsymbol{\xi})$ or $\mathcal{F}(\mu)(\boldsymbol{\xi})$ that are underrepresented in the hologram-data $I - 1$ and are thus hard to recover. In particular, note that the phase-CTF s_0 has a second-order zero at $\boldsymbol{\xi} = 0$ (whereas $c_0(0) = 1$), which means that coarse (non-holographic) features in the phase-image ϕ induce low contrast, in agreement with the observations in section 2.2.1.4.

2.4.3 Homogeneity constraints and pure phase objects

For pure substances with a fixed chemical formula (e.g. $\text{C}_6\text{H}_{12}\text{O}_6$: glucose), the refractive- and absorption decrements δ and β at a given photon-energy E are mutually proportional $\beta/\delta =$

$c_{\beta/\delta} = \text{const.}$ For samples with a sufficiently homogeneous composition (ideally made of a single material only), the resulting proportionality of phase- and absorption-image may be imposed as a constraint by fixing $\phi - i\mu = e^{-i\nu}\varphi$ for a real-valued function φ and $\nu = \arctan(c_{\beta/\delta})$. This is known as a *homogeneity-* or *single-material constraint* [161] and widely used in practice.

Indeed, note that the definition of homogeneity here includes in particular the special case $\nu = 0$, i.e. $\varphi = \phi$ and $\mu = 0$ identically, corresponding to the assumption of a *pure phase-* or *non-absorbing object*. This model is an excellent approximation for specimens composed of only light chemical elements according to the discussion in §2.1.2. In general, note that practically relevant values of ν will always be close to zero since $\beta \ll \delta$ and thus $\nu \approx c_{\beta/\delta} \ll 1$.

The practical benefit of imposing homogeneity is that only a single real-valued image φ has to be recovered from the data instead of both ϕ and μ , i.e. the number of unknowns is reduced. The constraint may be incorporated into the more general forward-models. The nonlinear forward operator becomes

$$\mathcal{N}_\nu(\varphi) := \mathcal{N}(ie^{-i\nu}\varphi) = |\mathcal{D}(\exp(-ie^{-i\nu}\varphi))|^2 - 1 \quad (2.4.5)$$

Importantly, the linearized forward model may be expressed via a single modified CTF:

$$\begin{aligned} \mathcal{S}_\nu(\varphi) &:= \mathcal{T}(ie^{-i\nu}\varphi) = -2\mathcal{F}^{-1} \left(\cos(\nu) \sin\left(\frac{|\boldsymbol{\xi}|^2}{2f}\right) \mathcal{F}(\varphi) + \sin(\nu) \cos\left(\frac{|\boldsymbol{\xi}|^2}{2f}\right) \mathcal{F}(\varphi) \right) \\ &= -2\mathcal{F}^{-1} \left(s_\nu \cdot \mathcal{F}(\varphi) \right) \quad \text{with} \quad s_\nu(\boldsymbol{\xi}) := \sin\left(\frac{|\boldsymbol{\xi}|^2}{2f} + \nu\right). \end{aligned} \quad (2.4.6)$$

The CTF s_ν exhibits an analogous oscillatory behavior as the functions s_0, c_0 in fig. 2.6 and its zeros, i.e. frequencies $\boldsymbol{\xi} \in \mathbb{R}^m$ with $s_\nu(\boldsymbol{\xi}) \approx 0$, likewise correspond to Fourier-components $\mathcal{F}(\varphi)(\boldsymbol{\xi})$ that induce low contrast in the hologram-data $\mathcal{S}_\nu(\varphi)$, compare §2.4.2.

2.4.4 Multiple holograms

In order to obtain more complete information on the unknown images ϕ, μ , it is standard to acquire multiple holograms at different distances between sample and detector [42, 212, 122, 121]. Settings with more than a single hologram may be modeled by stacking any of the above forward maps to a *vector-valued* operator in some sense: if I_1, I_2, \dots, I_ℓ denote the holograms at the considered Fresnel numbers f_1, f_2, \dots, f_ℓ and $F^{(f_i)} : h \mapsto I_i - 1$ with $F \in \{\mathcal{N}, \mathcal{N}_\nu, \mathcal{T}, \mathcal{S}_\nu\}$ the associated forward maps (linear or nonlinear, incorporating homogeneity or not), the relation between the unknown image h and the total data is described by the forward map

$$F^{(f_1, \dots, f_\ell)} : h \mapsto \begin{pmatrix} F^{(f_1)}(h) \\ \vdots \\ F^{(f_\ell)}(h) \end{pmatrix} = \begin{pmatrix} I_1 - 1 \\ \vdots \\ I_\ell - 1 \end{pmatrix}. \quad (2.4.7)$$

The benefit of acquiring multiple holograms may be most easily seen for the case of a CTF-model with homogeneity constraint $F = \mathcal{S}_\nu$ (cf. §2.4.3): Missing Fourier-components $\mathcal{F}(\varphi)(\boldsymbol{\xi})$ in the first hologram $\mathcal{S}_\nu^{(f_1)}(\varphi) = I_1 - 1$ due to zeros of the corresponding CTF, $s_\nu^{(f_1)}(\boldsymbol{\xi}) \approx 0$ may be completed via a second one $\mathcal{S}_\nu^{(f_2)}(\varphi) = I_2 - 1$ since $s_\nu^{(f_2)}$ has different zeros from $s_\nu^{(f_1)}$ if $f_1 \neq f_2$ so that $|s_\nu^{(f_2)}(\boldsymbol{\xi})|$ may be sufficiently large. Analogously, frequencies $\boldsymbol{\xi} \in \mathbb{R}^m$ for which

both CTFs are small, $s_\nu^{(f_1)}(\boldsymbol{\xi}) \approx 0$ and $s_\nu^{(f_2)}(\boldsymbol{\xi}) \approx 0$, may be compensated by a third hologram $\mathcal{S}_\nu^{(f_3)}(\varphi) = I_3 - 1$ with $|s_\nu^{(f_3)}(\boldsymbol{\xi})|$ large and so on. This idea can be formalized by systematically choosing f_1, f_2, \dots such that the maxima and minima of the $|s_\nu^{(f_i)}|$ optimally compensate each other [212]. Notably, however, the problem of low phase contrast at low frequencies may *not* be resolved via such a compensation since $s_\nu^{(f_i)}(0) = \sin(\nu) \ll 1$ holds independently of f_i (recall from §2.4.3 that $\nu \ll 1$ in practice), so that Fourier-components $\mathcal{F}(\varphi)(\boldsymbol{\xi})$ for low spatial frequencies $\boldsymbol{\xi} \approx 0$ are equally underrepresented in *all* holograms.

2.4.5 Tomographic models for XPCT

As detailed in §2.3, the phase- and absorption images $\phi_{(\boldsymbol{\theta})} = k\mathcal{P}_{\boldsymbol{\theta}}(\delta)$ and $\mu_{(\boldsymbol{\theta})} = k\mathcal{P}_{\boldsymbol{\theta}}(\beta)$ yield projections of the 3D-decrements δ and β (scaled by the wavenumber k), respectively, under some incident direction $\boldsymbol{\theta} \in \mathbb{S}^2$. By composing these relations with any of the forward maps $F \in \{\mathcal{N}^{(f_1, \dots, f_\ell)}, \mathcal{N}_\nu^{(f_1, \dots, f_\ell)}, \mathcal{T}^{(f_1, \dots, f_\ell)}, \mathcal{S}_\nu^{(f_1, \dots, f_\ell)}\}$ from sections 2.4.1 to 2.4.4, the associated forward model for XPCT is obtained:

$$F_{\text{PCT}} : f \mapsto (F(\mathcal{P}_{\boldsymbol{\theta}}(f)))_{\boldsymbol{\theta} \in \Theta} = (I_{\boldsymbol{\theta}} - 1)_{\boldsymbol{\theta} \in \Theta} \quad \text{with} \quad f = k\beta + ik\delta, \quad \Theta \subset \mathbb{S}^2. \quad (2.4.8)$$

$I_{\boldsymbol{\theta}}$ denotes the predicted hologram(s) (one or several) for the tomographic incident direction $\boldsymbol{\theta}$. The special case of full angular sampling (compare §2.3.3) is modeled by the operator

$$F_{\text{PCT}}^{\text{full}} : f \mapsto F^{(\text{lat})}(\mathcal{P}_{\text{full}}(f)) \quad \text{with} \quad F^{(\text{lat})}(p)(\boldsymbol{\theta}, \cdot) := F(p(\boldsymbol{\theta}, \cdot)). \quad (2.4.9)$$

2.4.6 Domains and constraints

For a complete mathematical description of XPCI and XPCT, it is necessary to specify suitable object- and data-spaces X and Y for the forward maps $F : X \supset A \rightarrow Y$ of §2.4.1 to §2.4.5.

2.4.6.1 General object- and data-spaces

Object space: In general, we will consider objects from the Hilbert-space of square-integrable functions on \mathbb{R}^m with $m \in \{2, 3\}$, i.e. $\mu + i\phi \in A$ or $\delta + i\beta \in A$ with

$$A \subset L^2(\mathbb{R}^m) := \{f : \mathbb{R}^m \rightarrow \mathbb{C} : \|f\|_{L^2}^2 = \int_{\mathbb{R}^m} |f(\mathbf{x})|^2 d\mathbf{x} < \infty\}. \quad (2.4.10)$$

Data space: All forward maps $F \in \{\mathcal{N}, \mathcal{N}_\nu, \mathcal{T}, \mathcal{S}_\nu\}$ have been defined such that they map onto the *contrast* $I - 1$ with respect to the constant background-intensity 1. Other than the hologram I itself, the contrast $I - 1$ induced by L^2 -images is also in L^2 , as will be seen in §2.4.7. Hence, the natural data-space for the case of a single hologram is $Y := L^2(\mathbb{R}^m)$.

For forward maps with more than one hologram I_1, \dots, I_ℓ as considered in §2.4.4 and §2.4.5, the natural data-space is given by the Cartesian product $Y := L^2(\mathbb{R}^m)^\ell \ni (I_1 - 1, \dots, I_\ell - 1)$.

2.4.6.2 Homogeneity constraints

Restrictions to homogeneous objects (including the special case of pure phase objects) may be imposed by considering any of the forward maps $F \in \{\mathcal{N}_\nu, \mathcal{S}_\nu\}$ from §2.4.3 on a set of

real-valued functions:

$$A \subset L^2(\mathbb{R}^m, \mathbb{R}) = \{f : \mathbb{R}^m \rightarrow \mathbb{R} : f \in L^2(\mathbb{R}^m)\}. \quad (2.4.11)$$

2.4.6.3 Support constraints

In practice, the approximate size of the imaged sample will typically be known. Mathematically, this type of a priori knowledge translates to a bound on the *support* of the sample-induced decrements δ and β of the refractive index (overbar denotes set-closure):

$$\text{supp}(f) := \overline{\{\mathbf{x} \in \mathbb{R}^3 : f(\mathbf{x}) \neq 0\}} \subset \Omega \quad (2.4.12)$$

for $f \in \{\delta, \beta\}$ and some bounded *support-domain* $\Omega \subset \mathbb{R}^3$. This also restricts the support of the resulting projection-images, $\text{supp}(\phi), \text{supp}(\mu) \subset \Omega^{(2d)}$ for some bounded $\Omega^{(2d)} \subset \mathbb{R}^2$. Such *support constraints* may be imposed by restricting the set of admissible objects to

$$A \subset L^2(\Omega) := \{f \in L^2(\mathbb{R}^m) : \text{supp}(f) \subset \Omega\}. \quad (2.4.13)$$

2.4.6.4 Non-negativity constraints

As noted in §2.1.2, β and δ are always non-negative for physical reasons. By definition, the same also holds true for the phase- and absorption-image ϕ and μ . Hence, it makes sense to restrict to object-functions with non-negative real- and imaginary parts:

$$A \subset \mathcal{K}_+ := \{f : \mathbb{R}^m \rightarrow \mathbb{C} : \text{Re}(f(\mathbf{x})) \geq 0 \text{ and } \text{Im}(f(\mathbf{x})) \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^m\}. \quad (2.4.14)$$

Linear and nonlinear constraints: Non-negativity constraints differ fundamentally from homogeneity- and support constraints. The latter restrict the object-space to closed (\mathbb{R} -)linear subspaces $L^2(\mathbb{R}^m, \mathbb{R}) \subset L^2(\mathbb{R}^m)$ and $L^2(\Omega) \subset L^2(\mathbb{R}^m)$, respectively, retaining the principal Hilbert-space geometry of $L^2(\mathbb{R}^m)$ (when regarded as a *real* Hilbert-space). On the contrary, non-negativity is a *nonlinear* constraint by nature – the associated sets $\mathcal{K}_+ \cap L^2(\Omega, \mathbb{R})$ are no vector-spaces but have the geometry of a closed convex cone.

As a consequence, non-negativity constraints are more complicated to analyze theoretically and also more difficult to enforce in practice. In particular, imposing non-negativity always requires a *nonlinear* reconstruction method – even when a linear forward model is assumed.

2.4.7 Well-definedness, continuity and differentiability

In the following, we show that the different forward maps are well-defined and continuous on the object- and data-spaces from §2.4.6. Moreover, it will be shown that the nonlinear maps are *differentiable* in the sense that they admit local linear approximations. These properties are important to devise and analyze algorithms for the ultimate goal of *inverting* these maps.

2.4.7.1 Boundedness of the basic linearized forward maps

For the linearized forward maps \mathcal{I} and \mathcal{S}_ν from §2.4.2 and §2.4.3, continuity as operators on $L^2(\mathbb{R}^m)$ is easy to show, owing to their close relation to the Fresnel propagator \mathcal{D} :

Theorem 2.3 (Boundedness of the linearized forward-operators).

- $\mathcal{T} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is a continuous \mathbb{R} -linear operator with $\|\mathcal{T}\| = 2$
- $\mathcal{S}_\nu : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is a continuous linear operator with $\|\mathcal{S}_\nu\| = 2$

Proof. $\mathcal{T} : h \mapsto -2\operatorname{Re}(\mathcal{D}(h))$ is \mathbb{R} -linear (but not \mathbb{C} -linear!) by linearity of \mathcal{D} and \mathbb{R} -linearity of the pointwise real part $\operatorname{Re}(\cdot)$. For continuity, we note that $|\operatorname{Re}(z)| \leq |z|$ for all $z \in \mathbb{C}$. Together with unitarity of $\mathcal{D} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ (see §2.2.2), this implies that

$$\|\mathcal{T}(h)\| = \|2\operatorname{Re}(\mathcal{D}(h))\| \leq 2\|\mathcal{D}(h)\| = 2\|h\| \quad \text{for all } h \in L^2(\mathbb{R}^m), \quad (2.4.15)$$

where the bound is attained for any h of the form $h = \mathcal{D}^{-1}(\tilde{h})$ with $\tilde{h} \in L^2(\mathbb{R}^m, \mathbb{R})$ real-valued. Thus, $\mathcal{T} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is bounded, i.e. continuous, with $\|\mathcal{T}\| = 2$.

Linearity of \mathcal{S}_ν follows from its Fourier-multiplier form, $\mathcal{S}_\nu : h \mapsto \mathcal{F}^{-1}(\mathbf{m} \cdot \mathcal{F}(h))$ with factor $\mathbf{m} = -2s_\nu$. Moreover, boundedness and continuity of $\mathbf{m} : \mathbb{R}^m \rightarrow \mathbb{C}$ implies boundedness of $\mathcal{S}_\nu : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ with $\|\mathcal{S}_\nu\| = \max_{\boldsymbol{\xi} \in \mathbb{R}^m} |\mathbf{m}(\boldsymbol{\xi})| = 2 \max_{\boldsymbol{\xi} \in \mathbb{R}^m} |s_\nu(\boldsymbol{\xi})| = 2$. \square

2.4.7.2 Well-definedness and continuity of the nonlinear forward map

By theorem 2.3, it is very natural to study the linearized forward maps \mathcal{T} and \mathcal{S}_ν as operators in $L^2(\mathbb{R}^m)$. As these maps arise as limiting cases of the general *nonlinear* model $\mathcal{N} : h \mapsto |\mathcal{D}(\exp(-h))|^2 - 1$, it is logical to aim for proving that also the latter is well-defined and continuous as a map on suitable L^2 -spaces. This will be achieved in the present section, followed by a proof of *Fréchet-differentiability* of \mathcal{N} in §2.4.7.3. Yet, the analysis turns out to be much more involved than for the linearized case. Indeed, it seems that a rigorous treatment of nonlinear XPCI on L^2 -spaces is to date not available in the literature: in [140], well-definedness and differentiability of \mathcal{N} is only shown on L^∞ and the analysis of \mathcal{N} in L^2 in [49] is for once limited to settings without absorption and moreover contains a wrong proof of differentiability, as detailed further below. Therefore, a thorough analysis is given in the following.

The nonlinear operations \exp and $|\cdot|^2$ require some (intermediate) work on general L^p -spaces:

$$\begin{aligned} L^p(\mathbb{R}^m) &= \{f : \mathbb{R}^m \rightarrow \mathbb{C} : \|f\|_{L^p}^p := \int_{\mathbb{R}^m} |f|^p dx < \infty\} \quad \text{for } 1 \leq p < \infty, \\ L^\infty(\mathbb{R}^m) &= \{f : \mathbb{R}^m \rightarrow \mathbb{C} : \|f\|_{L^\infty} := \operatorname{esssup}_{\mathbf{x} \in \mathbb{R}^m} |f(\mathbf{x})| < \infty\} \end{aligned} \quad (2.4.16)$$

where esssup denotes the supremum of a function up to modifications on sets of zero Lebesgue-measure. We define $L^p(\Omega, \mathbb{R})$ for $\Omega \subset \mathbb{R}^m$ analogously to the L^2 -case, compare §2.4.6.

For further analysis, we cast the nonlinear forward operator to the form

$$\begin{aligned} \mathcal{N}(h) &= |\mathcal{D}(1) + \mathcal{D}(\exp(-h) - 1)|^2 - 1 = 2\operatorname{Re}(\mathcal{D}(\exp(-h) - 1)) + |\mathcal{D}(\exp(-h) - 1)|^2 \\ &= -\mathcal{T}(o(h)) + |\mathcal{D}(o(h))|^2 \quad \text{with } o(h) := \exp(-h) - 1. \end{aligned} \quad (2.4.17)$$

Studying o instead of $\exp(\cdot)$ has the advantage that $o(h)(\mathbf{x}) = 0$ whenever $h(\mathbf{x}) = 0$ (whereas $\exp(0) = 1$), which implies that $\operatorname{supp}(o(h)) \subset \operatorname{supp}(h)$ for any h . Yet, the problem remains that the involved nonlinear operations o and $|\cdot|^2$ in (2.4.17) – as opposed to the linear maps \mathcal{T} and \mathcal{D} – are not well-defined as maps $L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ but only on very specific L^p -spaces:

- $o : L^p(\Omega) \rightarrow L^q(\Omega)$ for $\Omega \subset \mathbb{R}^m$ may only be well-defined (and continuous) if $p = \infty$.
- $|\cdot|^2 : L^p(\mathbb{R}^m) \rightarrow L^q(\mathbb{R}^m, \mathbb{R})$ is well-defined (and continuous) if and only if $p = 2q$.

In the following, we outline how these subtleties may be circumvented.

The pointwise exponential: The fact that o is well-defined only as a map $L^\infty \rightarrow L^q$ is due to the super-algebraic growth of the pointwise exponential $h \mapsto \exp(-h)$ for h with negative real part, which infinitely amplifies singularities of the kind $\operatorname{Re}(h(\mathbf{x})) \rightarrow -\infty$. In the context of XPCI however, function-values $\operatorname{Re}(h(\mathbf{x})) < 0$ correspond to the *hypothetical* case of *negative absorption* $\mu = \operatorname{Re}(h) < 0$. In other words, the undesirable exponential growth is an artifact of the mathematical model without physical relevance! This justifies to fix the problem by adjusting the forward model in either of the following ways:

- (1) Restrict to the physically relevant part of the domain of definition:

$$L_{\geq 0}^p(\Omega) := \{f \in L^p(\Omega) : \operatorname{Re}(f(\mathbf{x})) \geq 0 \text{ for almost all } \mathbf{x} \in \mathbb{R}^m\}. \quad (2.4.18)$$

- (2) Modify the forward operator \mathcal{N} , replacing o by a map \tilde{o} with more favorable analytical properties that satisfies $\tilde{o}(h) = o(h)$ for all $h \in L_{\geq 0}^p(\Omega)$.

Both the analytical difficulties induced by pointwise exponentials and the above solutions have been already noted for related inverse problems in transmission imaging [53].

For now, we pursue idea (1), coming back to approach (2) in §2.4.7.3. On the set $\mathbb{C}_{\geq 0} := \{z \in \mathbb{C} : \operatorname{Re}(z) \geq 0\}$, we have the properties $|1 - \exp(-z)| \leq |z|$ and $|\exp(-z_1) - \exp(-z_2)| \leq |z_1 - z_2|$ for all $z, z_1, z_2 \in \mathbb{C}_{\geq 0}$. By applying these relations pointwise, we obtain

$$\|o(h)\|_{L^p} = \|1 - \exp(-h)\|_{L^p} \leq \|h\|_{L^p} \quad (2.4.19a)$$

$$\|o(h_1) - o(h_2)\|_{L^p} = \|\exp(-h_2) - \exp(-h_1)\|_{L^p} \leq \|h_1 - h_2\|_{L^p} \quad (2.4.19b)$$

for all $h, h_1, h_2 \in L_{\geq 0}^p(\mathbb{R}^m)$ and $1 \leq p \leq \infty$. Since $\operatorname{supp}(o(h)) \subset \operatorname{supp}(h)$ holds for all h , the derived bounds show well-definedness (2.4.19a) and (Lipschitz-)continuity (2.4.19b) of

$$o : L_{\geq 0}^p(\mathbb{R}^m) \rightarrow L^p(\mathbb{R}^m) \quad \text{for any } \Omega \subset \mathbb{R}^m, 1 \leq p \leq \infty. \quad (2.4.20)$$

The squared modulus: The problem concerning the admissible domains for $f \mapsto |f|^2$ can be resolved by restricting to compactly supported functions, as explained in the following.

In order for a term of the form $|\mathcal{D}(f)|^2$ as in (2.4.17) to be in $L^2(\mathbb{R}^m)$, it is necessary and sufficient that $\mathcal{D}(f) \in L^4(\mathbb{R}^m)$. While the Fresnel propagator \mathcal{D} has so far only been considered as a unitary map $L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$, it may also be shown that $\mathcal{D} : L^1(\mathbb{R}^m) \rightarrow L^\infty(\mathbb{R}^m)$ is well-defined and bounded, using its alternate form (2.2.10) and continuity of $\mathcal{F} : L^1(\mathbb{R}^m) \rightarrow L^\infty(\mathbb{R}^m)$. By combining the boundedness on L^2 and L^1 via the Riesz-Thorin interpolation theorem, we find more generally that

$$\mathcal{D} : L^p(\mathbb{R}^m) \rightarrow L^q(\mathbb{R}^m) \quad \text{for all } 1 \leq p \leq 2, q = p/(p-1) \quad (2.4.21)$$

is well-defined and bounded. In particular, we have $\mathcal{D}(f) \in L^4(\mathbb{R}^m)$ and thus $|\mathcal{D}(f)|^2 \in L^2(\mathbb{R}^m)$ for all $f \in L^{4/3}(\mathbb{R}^m)$. Now, for $f \in L^2(\Omega)$ supported within a bounded domain $\Omega \subset \mathbb{R}^m$, it

may be used that $L^q(\Omega) \subset L^p(\Omega)$ with continuous embedding for all $1 \leq p \leq q \leq \infty$, which implies that $f \in L^{4/3}(\mathbb{R}^m)$ and hence $|\mathcal{D}(f)|^2 \in L^2(\mathbb{R}^m)$. More generally, the argument shows well-definedness and continuity of the *Fresnel-intensity-operator*

$$\mathcal{I}_{\mathcal{D}} : L^p(\Omega) \rightarrow L^q(\mathbb{R}^m); f \mapsto |\mathcal{D}(f)|^2 \quad (2.4.22)$$

for any bounded $\Omega \subset \mathbb{R}^m$ and all $p \leq 2$, $q \geq \frac{1}{2}p/(p-1)$, i.e. in particular for $p = q = 2$.

By combining the results for the different sub-operators in (2.4.17), we may finally prove well-definedness and continuity of \mathcal{N} :

Theorem 2.4 (Well-definedness and continuity of the nonlinear forward map). *For $\Omega \subset \mathbb{R}^m$ bounded, $\mathcal{N} : L^2_{\geq 0}(\Omega) \rightarrow L^2(\mathbb{R}^m)$; $h \mapsto |\mathcal{D}(\exp(-h))|^2 - 1$ is well-defined and continuous.*

Proof. According to (2.4.17), we have $\mathcal{N} = -\mathcal{S} \circ o + \mathcal{I}_{\mathcal{D}} \circ o$, i.e. \mathcal{N} is a composition of the continuous maps $o : L^2_{\geq 0}(\Omega) \rightarrow L^2(\Omega)$, $\mathcal{S} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$ and $\mathcal{I}_{\mathcal{D}} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$. \square

2.4.7.3 Fréchet-differentiability

We aim to prove (Fréchet-)differentiability of \mathcal{N} : a map $F : X \supset U \rightarrow Y$ (X, Y : real Banach spaces, $U \subset X$ open) is called (Fréchet-)differentiable (at $f_0 \in U$) if there exists a bounded linear operator $F'[f_0] : X \rightarrow Y$ – the (Fréchet-)derivative – such that

$$\lim_{f_0+h \in U, h \rightarrow 0} \frac{\|F(f_0+h) - F(f_0) - F'[f_0]h\|_Y}{\|h\|_X} = 0, \quad (2.4.23)$$

Intuitively, this means that F is locally well-described by its first-order Taylor approximation $F(f) \approx F(f_0) + F'[f_0](f - f_0)$ for $f \approx f_0$, i.e. by a *local linearization*. This property is crucial as it provides the mathematical basis for nonlinear image reconstruction via gradient-descent- or Newton-type-methods, see §2.7.1. Details on mathematical properties of Fréchet-derivatives can be found in standard monographs on functional analysis, see e.g. [213, §4]. Here, we will only need that the sum- and chain-rule hold as expected: if $F, G : X \supset U \rightarrow Y$ and $H : V \rightarrow Z$ with $F(U) \subset V$ are Fréchet-differentiable, then so are $F + G$ and $H \circ F$ with

$$(F + G)[f]' = F'[f] + G'[f] \quad \text{and} \quad (H \circ F)'[f] = H'[F(f)] \circ F'[f], \quad f \in U. \quad (2.4.24)$$

The forward operator \mathcal{N} is a composition of the linear map \mathcal{D} and the smooth pointwise nonlinear functions \exp and $|\cdot|^2$, so that some form of differentiability is expected to hold. Indeed, it has been shown in [139] that $\mathcal{N} : L^\infty(\Omega) \rightarrow L^2(\mathbb{R}^m)$ is Fréchet-differentiable for bounded $\Omega \subset \mathbb{R}^m$. Here, we aim extend this result to a favorable pure L^2 -setting $\mathcal{N} : L^2_{\geq 0}(\Omega) \rightarrow L^2(\mathbb{R}^m)$, for which continuity has been proven in theorem 2.4. Yet, in addition to the subtleties dealt with in §2.4.7.2, new technical problems arise:

- Fréchet-differentiability is a *local* property that is only well-defined for points within open sets, whereas $L^2_{\geq 0}(\Omega) \subset L^2(\Omega)$ has *empty interior*, i.e. the only open set $U \subset L^2_{\geq 0}(\Omega)$ (w.r.t. the topology of $L^2(\Omega)$) is $U = \emptyset$. A statement like “ $\mathcal{N} : L^2_{\geq 0}(\Omega) \rightarrow L^2(\mathbb{R}^m)$ is differentiable” thus would not make any sense. A work-around is to prove differentiability for *extensions* $\tilde{\mathcal{N}} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$ such that $\tilde{\mathcal{N}}(h) = \mathcal{N}(h)$ for all $h \in L^2_{\geq 0}(\Omega)$. This work-around follows idea (2) from §2.4.7.2 to fix the pointwise exponential.

- $o : h \mapsto \exp(-h) - 1$ may *not* be extended to a Fréchet-differentiable operator on $L^2(\Omega)$: by [53, Example A.5], there exists *no* function $\text{e}\ddot{\text{x}}\text{p} : \mathbb{C} \rightarrow \mathbb{C}$ with $\text{e}\ddot{\text{x}}\text{p}(-z) = \exp(-z)$ for all $z \in \mathbb{C}_{\geq 0}$ such that the induced extension $\tilde{o} : L^2(\Omega) \rightarrow L^2(\Omega); h \mapsto \text{e}\ddot{\text{x}}\text{p}(-h) - 1$ of o is differentiable. Yet, there *do* exist such extensions that are continuous from $L^p(\Omega)$ to $L^p(\Omega)$ and differentiable as maps $\tilde{o} : L^p(\Omega) \rightarrow L^q(\Omega)$ for all $1 \leq q < p \leq \infty$ and $\Omega \subset \mathbb{R}^m$, according to [53, Proposition 3.9].[§]

Notably, difficulties in proving differentiability of \mathcal{N} only arise for o , the other operators entering in (2.4.17) are well-behaved: $|z+h|^2 - 2\text{Re}(\bar{z} \cdot h) - |z|^2 = |h|^2$ holds true for all $z, h \in \mathbb{C}$, which implies that $|\cdot|^2 : L^{2p}(\mathbb{R}^m) \rightarrow L^p(\mathbb{R}^m)$ is differentiable for all $1 \leq p \leq \infty$. Likewise, the *linear* maps \mathcal{D} and \mathcal{F} are differentiable when bounded. By analogous argument as in proving continuity (see §2.4.7.2), this implies differentiability of $\mathcal{I}_{\mathcal{D}} : L^p(\Omega) \rightarrow L^q(\mathbb{R}^m); f \mapsto |\mathcal{D}(f)|^2$ for any bounded $\Omega \subset \mathbb{R}^m$ and all $p \leq 2, q \geq \frac{1}{2}p/(p-1)$. The derivative is given by

$$\mathcal{I}'_{\mathcal{D}}[f]h = 2\text{Re} \left(\overline{\mathcal{D}(f)} \cdot \mathcal{D}(h) \right) \quad \text{for all } f, h \in L^p(\Omega). \quad (2.4.25)$$

Combining the differentiability-results for the operators $o, \mathcal{I}_{\mathcal{D}}, \mathcal{F}$ yields the following theorem:

Theorem 2.5 (Fréchet-differentiability of the nonlinear forward map). *Let $\Omega \subset \mathbb{R}^m$ be bounded. Then there exists a continuous operator $\tilde{\mathcal{N}} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$ such that $\tilde{\mathcal{N}}(h) = \mathcal{N}(h)$ for all $h \in L^2_{\geq 0}(\Omega)$. Furthermore, $\tilde{\mathcal{N}}$ is Fréchet-differentiable both as a map $L^2(\Omega) \rightarrow L^q(\mathbb{R}^m)$ and $L^q(\Omega) \rightarrow L^2(\mathbb{R}^m)$ for any $q > 2$. For $f \in L^2_{\geq 0}(\Omega)$, the derivative is given by*

$$\tilde{\mathcal{N}}'[f]h = -2\text{Re} \left(\overline{\mathcal{D}(\exp(-f))} \cdot \mathcal{D}(\exp(-f) \cdot h) \right) \quad \text{for all } h \in L^2(\Omega). \quad (2.4.26)$$

Proof. We take an extension $\tilde{o} : h \mapsto \text{e}\ddot{\text{x}}\text{p}(-h) - 1$ of the kind introduced above and set $\tilde{\mathcal{N}} := -\mathcal{F} \circ \tilde{o} + \mathcal{I}_{\mathcal{D}} \circ \tilde{o}$. Since $\tilde{o}(h) = o(h)$ if $\text{Re}(h) \geq 0$ almost everywhere, $\tilde{\mathcal{N}}$ is an extension of \mathcal{N} , i.e. $\tilde{\mathcal{N}}(h) = \mathcal{N}(h)$ for all $h \in L^2_{\geq 0}(\Omega)$. Moreover, continuity of $\tilde{o} : L^2(\Omega) \rightarrow L^2(\Omega)$ implies that $\tilde{\mathcal{N}} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$ is continuous by an analogous argument as in theorem 2.4.

Now let $q > 2$. Then $\tilde{o} : L^q(\Omega) \rightarrow L^2(\Omega)$ is differentiable and so are $\mathcal{F} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$ and $\mathcal{I}_{\mathcal{D}} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$, which shows differentiability of $\tilde{\mathcal{N}} : L^q(\Omega) \rightarrow L^2(\mathbb{R}^m)$. Likewise, \tilde{o} is differentiable as a map $L^2(\Omega) \rightarrow L^p(\Omega)$ for $p := q/(q-1) < 2$ and $\mathcal{I}_{\mathcal{D}}, \mathcal{F}$ are differentiable from $L^p(\Omega)$ to $L^q(\mathbb{R}^m)$ (for \mathcal{F} , this follows from boundedness of $\mathcal{D} : L^p(\mathbb{R}^m) \rightarrow L^q(\mathbb{R}^m)$), which implies differentiability of $\tilde{\mathcal{N}} : L^2(\Omega) \rightarrow L^q(\mathbb{R}^m)$.

The explicit formula (2.4.26) for $\tilde{\mathcal{N}}'$ follows by combining the derivatives for the different sub-operators via the chain-rule (2.4.24), using (2.4.25) and that $\tilde{o}'[f]h = \exp(-f) \cdot h$ must hold true for $f \in L^2_{\geq 0}(\Omega)$, since \tilde{o} extends $o : h \mapsto \exp(-h) - 1$ in a differentiable manner. \square

Sadly, theorem 2.5 only shows \mathcal{N} to be *almost* Fréchet-differentiable from L^2 to L^2 . This remaining flaw of the result may – once again – be fixed by reducing the nonlinear forward model to the physically relevant setting: if the operator \mathcal{N} is modified to incorporate the fact that experimental hologram-data may not be acquired in the whole space \mathbb{R}^m , but only within a *bounded* domain $K \subset \mathbb{R}^m$ covered by the detector, differentiability is established on L^2 :

[§]This is the point where the differentiability-proof in [49] is wrong: the theorems in [53] are incorrectly cited to conclude that $E : L^2(\Omega, \mathbb{R}) \rightarrow L^\infty(\Omega); u \mapsto \exp(-iu)$ is differentiable. Like the extensions \tilde{o} of $h \mapsto \exp(-h) - 1$ considered here, however, E is only differentiable as a map $L^p(\Omega, \mathbb{R}) \rightarrow L^q(\Omega)$ for $1 \leq q < p \leq \infty$.

Corollary 2.6 (Fréchet-differentiability on L^2). *Within the setting of theorem 2.5, let $K \subset \mathbb{R}^m$ be bounded. Then $\tilde{\mathcal{N}}_K : L^2(\Omega) \rightarrow L^2(K)$; $h \mapsto \tilde{\mathcal{N}}(h)|_K$ is Fréchet-differentiable.*

Proof. By theorem 2.5, $\tilde{\mathcal{N}}_K$ is differentiable as a map $L^2(\Omega) \rightarrow L^q(K)$ for $q > 2$. Hence, the claim follows by continuity of the embedding $\iota : L^q(K) \hookrightarrow L^2(K)$ for bounded K . \square

The bottom-line: While quite a few mathematical technicalities had to be addressed in order to describe the nonlinear forward model of XPCI as a differentiable operator on L^2 -spaces, we emphasize that our adjustments of the setting (domain-restrictions, operator-extensions, etc.) do not affect the modeled physics. In the following, we will therefore omit the involved technical details and refer to the results in a simplified form:

For all practical purposes, the nonlinear forward map of XPCI $\mathcal{N} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$ is well-defined, continuous and Fréchet-differentiable.

2.4.7.4 Extension to homogeneity constraints, multiple holograms and XPCT

The basic results on well-definedness and regularity of the basic forward operators may be readily extended to the settings considered in §2.4.3 to §2.4.5. Details are given below.

Homogeneity constraints: According to §2.4.3, the transition to homogeneous objects corresponds to composing the forward map \mathcal{N} with the linear-bounded operator $\iota_\nu : L^2(\Omega, \mathbb{R}) \rightarrow L^2(\Omega)$; $\varphi \mapsto ie^{-i\nu}\varphi$. Thus, well-definedness, continuity and Fréchet-differentiability carry over to $\mathcal{N}_\nu = \mathcal{N} \circ \iota_\nu : L^2(\Omega, \mathbb{R}) \rightarrow L^2(\mathbb{R}^m)$ with derivative $\mathcal{N}'_\nu[f]h = \mathcal{N}'[\iota_\nu(f)]\iota_\nu(h)$.

Multiple holograms: As seen in §2.4.4, settings with multiple holograms are modeled by forward maps $F^{(f_1, \dots, f_\ell)} : f \mapsto (F^{(f_1)}(f), \dots, F^{(f_\ell)}(f))$ with $F \in \{\mathcal{N}, \mathcal{N}_\nu, \mathcal{I}, \mathcal{I}_\nu\} : L^2(\Omega, (\mathbb{R})) \rightarrow L^2(\mathbb{R}^m)$. If F is linear-bounded or nonlinear, continuous and Fréchet-differentiable, then the same holds true for $F^{(f_1, \dots, f_\ell)} : L^2(\Omega, (\mathbb{R})) \rightarrow L^2(\mathbb{R}^m)^\ell$. The derivative is given by

$$(F^{(f_1, \dots, f_\ell)})' [f]h = \left((F^{(f_1)})' [f]h, \dots, (F^{(f_\ell)})' [f]h \right). \quad (2.4.27)$$

XPCT: In the XPCT-model from §2.4.5, $F \in \{\mathcal{N}^{(f_1, \dots, f_\ell)}, \mathcal{N}_\nu^{(f_1, \dots, f_\ell)}, \mathcal{I}^{(f_1, \dots, f_\ell)}, \mathcal{I}_\nu^{(f_1, \dots, f_\ell)}\}$ is composed with parallel-beam projectors \mathcal{P}_θ . For any $\theta \in \mathbb{S}^2$, $\Omega \subset \mathbb{R}^3$ bounded, $\Omega_\theta := \text{supp}(\mathcal{P}_\theta(\mathbf{1}_\Omega)) \subset \mathbb{R}^2$ is also bounded and $\mathcal{P}_\theta : L^2(\Omega, \mathbb{K}) \rightarrow L^2(\Omega_\theta, \mathbb{K})$ is well-defined and continuous for $\mathbb{K} \in \{\mathbb{C}, \mathbb{R}\}$. Thus, if $F : L^2(\Omega_\theta, \mathbb{K}) \rightarrow L^2(\mathbb{R}^2)^\ell$ is linear-bounded or continuous and Fréchet-differentiable, so is $F \circ \mathcal{P}_\theta : L^2(\Omega, \mathbb{K}) \rightarrow L^2(\mathbb{R}^2)^\ell$ and, by the chain-rule (2.4.24),

$$(F \circ \mathcal{P}_\theta)' [f]h = F'[\mathcal{P}_\theta(f)]\mathcal{P}_\theta(h). \quad (2.4.28)$$

Analogously as for multiple holograms, this result extends to $F_{\text{PCT}} : L^2(\Omega, \mathbb{K}) \rightarrow L^2(\mathbb{R}^2)^{\ell \cdot N_{\text{proj}}}$; $f \mapsto (F(\mathcal{P}_\theta(f)))_{\theta \in \Theta}$, where $N_{\text{proj}} \in \mathbb{N}$ the number of elements of $\Theta \subset \mathbb{S}^2$.

Similarly, continuity and differentiability may also be shown for XPCT-models assuming full angular sampling, i.e. for $F_{\text{PCT}}^{\text{full}} : L^2(\Omega) \rightarrow L^2([0; \pi) \times \mathbb{R}^2)$ as defined by (2.4.9).

2.5 Inverse problems and ill-posedness

In the preceding section, we have shown that image-formation in XPCI and XPCT may be described mathematically by forward maps $F : X \supset A \rightarrow Y$ (between L^2 -spaces X, Y) and that these maps are analytically well-behaved in the sense that F is continuous (even differentiable). Essentially, it has thus been shown that the *forward problems* of computing holograms from given images μ, ϕ or δ, β are *well-posed* in the sense of HADAMARD [82], see definition 1.1:

- (a) A solution exists
- (b) The solution is unique
- (c) The solution depends continuously on the data

Yet, of greater practical interest are the associated *inverse problems* of image-reconstruction, i.e. the problem of inverting F . In the present section, we will see that the inverse problems 2.1 and 2.2 of XPCI and XPCT are indeed *ill-posed* in general: depending on the specific setting, the unknown images need not be uniquely determined by the hologram-data so that condition (b) may be violated, as detailed in §2.5.1. Furthermore, even in settings where uniqueness holds, criterion (c) may fail to be satisfied, which implies that image reconstruction may be arbitrarily susceptible to data-errors in principle, as will be discussed in §2.5.2.

In general, uniqueness and ill-posedness of an inverse problem depend strongly on details of the associated forward map $F : X \supset A \rightarrow Y$. In particular, the following aspects matter:

- (1) *Data-richness*: the more information is provided by the measurements, the more likely they permit unique and stable reconstruction. For XPCI, it is thus highly relevant how many holograms are acquired at different Fresnel numbers f_1, \dots, f_ℓ (see §2.4.4) and for what set of tomographic incident directions Θ in the XPCT-case.
- (2) *A priori knowledge*: the smaller the *constraint-set* A , the easier it is to distinguish any two images $h_1, h_2 \in A$ with $h_1 \neq h_2$ based on the measured data.

Furthermore, it is relevant whether the fully nonlinear XPCI-model $F \in \{\mathcal{N}, \mathcal{N}_\nu\}$ is considered or whether the linearization $F \in \{\mathcal{T}, \mathcal{S}_\nu\}$ is assumed: for example, it may happen that uniqueness holds for the linearized but not for the nonlinear model.

All subsequent results on uniqueness and ill-posedness of XPCI and XPCT that are given without a reference are considered common knowledge – no claim of originality is made.

2.5.1 Uniqueness

We review the uniqueness theory of XPCI. Extensions to XPCT will be discussed in §2.5.1.4. Uniqueness of an inverse problem is equivalent to *injectivity* of the associated forward map F .

2.5.1.1 Sources of non-uniqueness

The phase-wrapping ambiguity: Upon considering the nonlinear forward map $\mathcal{N}(h) = |\mathcal{D}(\exp(-h))|^2 - 1$, an obvious and rather trivial source of non-uniqueness can be seen to be

the periodicity of the exponential: if $h_1 = \mu_1 + i\phi_1$ and $h_2 = \mu_2 + i\phi_2$ differ at each point only by multiples of 2π in the phase-images ϕ_1, ϕ_2 , their exponentials are equal,

$$\exp(-h_1) = \exp(-h_2) \iff h_1(\mathbf{x}) - h_2(\mathbf{x}) \in 2\pi i\mathbb{Z} \quad \text{for almost all } \mathbf{x} \in \mathbb{R}^m. \quad (2.5.1)$$

The ambiguity in (2.5.1) describes the well-known effect of *phase-wrapping*. In XPCI, it implies that the phase-image ϕ may in general only be recovered up to pointwise increments by multiples of 2π . While one might argue that this is no problem because any *phase* is only well-defined modulo 2π anyway, phase-wrapped (i.e. wrongly phased) projections $\phi_{\boldsymbol{\theta}} = k\mathcal{P}_{\boldsymbol{\theta}}(\delta)$ translate to severe systematic distortions in the refractive decrement δ in XPCT – while the latter quantity actually *does* have a sound physical meaning beyond any 2π -periodicity!

In practice, however, the phase-wrapping ambiguity typically has a minor effect on XPCI. Firstly, the imaged samples are often sufficiently small and weakly interacting such that the induced phase-images satisfy $0 \leq \phi_{(\boldsymbol{\theta})} < 2\pi$ (for all incident directions $\boldsymbol{\theta}$). Secondly, real-world images ϕ are typically smooth to a certain degree such that jumps of $\phi(\mathbf{x}_i)$ by more than $\pm\pi$ between adjacent sampling points \mathbf{x}_i (pixels) can be precluded. If such a smoothness-condition is known to hold, the ambiguity is resolved up to a global additive constant in $2\pi\mathbb{Z}$.

The phase-problem: A second and more severe source of non-uniqueness in XPCI is the phase-problem: the phase of the propagated wave-field $\propto \mathcal{D}(\exp(-h))$ is lost in the detection-process as only the squared modulus $|\mathcal{D}(\exp(-h))|^2$ is measured. In the linearized forward map $\mathcal{T} : h \mapsto -2\text{Re}(\mathcal{D}(h))$, the phase-problem manifests as the restriction to the real part of the Fresnel-data $\mathcal{D}(h)$, which likewise destroys half the information in some sense. Image reconstruction in XPCI implicitly requires to recover the missing phase (or imaginary part). Uniqueness of such *phase retrieval problems* has been extensively studied in the literature, especially for the related problem where \mathcal{D} is replaced by a Fourier transform \mathcal{F} (the far-field limit, §2.2.2.2), see [149, 118, 66, 182, 136] for reviews. The uniqueness results presented in the following are based on the *holographic* nature of phase retrieval in XPCI, compare §2.2.1.5.

2.5.1.2 Examples of uniqueness and non-uniqueness in XPCI

General non-uniqueness for a single hologram: The most challenging XPCI-setting is arguably image reconstruction from a single hologram without further constraints, i.e. retrieval of both phase- and absorption-image ϕ and μ without assuming homogeneity. Inverse problem 2.1 is non-unique in this setting as can be easily seen for the linearized model:

Theorem 2.7 (General non-uniqueness of linearized XPCI). $\mathcal{T} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is not injective, its null-space being $\text{kern}(\mathcal{T}) := \{h \in L^2(\mathbb{R}^m) : \mathcal{T}(h) = 0\} = i\mathcal{D}^{-1}(L^2(\mathbb{R}^m, \mathbb{R}))$.

Note that the non-uniqueness is severe: $\text{kern}(\mathcal{T})$ contains half of the whole space $L^2(\mathbb{R}^m)$ in the sense of the orthogonal decomposition $L^2(\mathbb{R}^m) = \mathcal{D}^{-1}(L^2(\mathbb{R}^m, \mathbb{R})) \oplus i\mathcal{D}^{-1}(L^2(\mathbb{R}^m, \mathbb{R}))$.

Non-uniqueness for the nonlinear model is technically more challenging to show. Examples of provably non-unique structures are given by so-called *phase-vortices* [157].

Theorem 2.8 (General non-uniqueness of nonlinear XPCI (phase-vortex example [157])). Let $h_{\pm} : \mathbb{R}^2 \rightarrow \mathbb{C}; \mathbf{x} \mapsto a(|\mathbf{x}|) \pm i\nu \arctan_2(\mathbf{x})$, where $\nu \in \mathbb{N}$, $a : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ is some smooth function

and $\arctan_2 : \mathbb{R}^2 \setminus \{0\} \rightarrow [0; 2\pi)$; $(r \cos(\varphi), r \sin(\varphi)) \mapsto \varphi$ assigns the angle enclosed by a vector in \mathbb{R}^2 with the x_1 -axis. Then

$$|\mathcal{D}(\exp(-h_+))|^2 = |\mathcal{D}(\exp(-h_-))|^2 \quad \text{but} \quad h_+ \neq h_- \quad \text{almost everywhere.} \quad (2.5.2)$$

Note that the example does *not* show non-uniqueness of $\mathcal{N} : L^2_{\geq 0}(\Omega) \rightarrow \mathbb{R}^m$, as the phase-images $\phi_{\pm} = \text{Im}(h_{\pm})$ from theorem 2.8 are not compactly supported, indeed $h_{\pm} \notin L^2(\mathbb{R}^m)$.

Uniqueness under homogeneity constraints and for two holograms: Based on the insights from theorems 2.7 and 2.8, it is commonly argued that at least two holograms and/or a homogeneity constraint are necessary for faithful image reconstruction in XPCI, compare §2.4.3 and §2.4.4. Indeed, uniqueness of the linearized model is established in these settings:

Theorem 2.9 (Uniqueness of linearized XPCI for homogeneity constraints or two holograms).

- $\mathcal{S}_{\nu} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is injective for any $\nu \in \mathbb{R}$ (and any Fresnel number \mathfrak{f})
- $\mathcal{T}^{(\mathfrak{f}_1, \mathfrak{f}_2)} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)^2$ is injective for any Fresnel numbers $\mathfrak{f}_1 \neq \mathfrak{f}_2$

Proof. Let $\varphi \in L^2(\mathbb{R}^m)$ such that $\mathcal{S}_{\nu}(\varphi) = 0$. Then $\mathcal{F}(\mathcal{S}_{\nu}(\varphi)) = s_{\nu} \cdot \mathcal{F}(\varphi) = 0$. Thus, $\mathcal{F}(\varphi)(\boldsymbol{\xi}) = 0$ for almost all $\boldsymbol{\xi} \in \mathbb{R}^m$ for which $s_{\nu}(\boldsymbol{\xi}) \neq 0$. As the CTF s_{ν} is non-zero almost everywhere in \mathbb{R}^m , this implies that $\mathcal{F}(\varphi) = 0$ and hence $\varphi = 0$ in an L^2 -sense. This proves that $\text{kern}(\mathcal{S}_{\nu}) = \{0\}$, i.e. that \mathcal{S}_{ν} is injective.

To prove injectivity of $\mathcal{T}^{(\mathfrak{f}_1, \mathfrak{f}_2)}$, we write the contrast in Fourier space via a CTF-matrix:

$$\begin{aligned} \mathcal{F}(\mathcal{T}^{(\mathfrak{f}_1, \mathfrak{f}_2)}(h))(\boldsymbol{\xi}) &= \begin{pmatrix} \mathcal{F}(\mathcal{T}^{(\mathfrak{f}_1)}(h))(\boldsymbol{\xi}) \\ \mathcal{F}(\mathcal{T}^{(\mathfrak{f}_2)}(h))(\boldsymbol{\xi}) \end{pmatrix} \stackrel{(2.4.3)}{=} -2\mathbf{S}^{(\mathfrak{f}_1, \mathfrak{f}_2)}(\boldsymbol{\xi}) \cdot \begin{pmatrix} \mathcal{F}(\text{Im}(h))(\boldsymbol{\xi}) \\ \mathcal{F}(\text{Re}(h))(\boldsymbol{\xi}) \end{pmatrix} \\ \mathbf{S}^{(\mathfrak{f}_1, \mathfrak{f}_2)}(\boldsymbol{\xi}) &:= \begin{pmatrix} \sin(\boldsymbol{\xi}^2/(2\mathfrak{f}_1)) & \cos(\boldsymbol{\xi}^2/(2\mathfrak{f}_1)) \\ \sin(\boldsymbol{\xi}^2/(2\mathfrak{f}_2)) & \cos(\boldsymbol{\xi}^2/(2\mathfrak{f}_2)) \end{pmatrix} \end{aligned} \quad (2.5.3)$$

According to (2.5.3), having $\mathcal{F}(\mathcal{T}^{(\mathfrak{f}_1, \mathfrak{f}_2)}(h))(\boldsymbol{\xi}) = 0$ but $\mathcal{F}(h)(\boldsymbol{\xi}) \neq 0$ requires the determinant of $\mathbf{S}^{(\mathfrak{f}_1, \mathfrak{f}_2)}(\boldsymbol{\xi}) \in \mathbb{R}^{2 \times 2}$ to vanish. As $\det(\mathbf{S}^{(\mathfrak{f}_1, \mathfrak{f}_2)}(\boldsymbol{\xi})) = \sin(\boldsymbol{\xi}^2/(2\mathfrak{f}_1) - \boldsymbol{\xi}^2/(2\mathfrak{f}_2))$ is non-zero for almost all $\boldsymbol{\xi} \in \mathbb{R}^m$, injectivity of $\mathcal{T}^{(\mathfrak{f}_1, \mathfrak{f}_2)}$ follows by an analogous argument as for \mathcal{S}_{ν} . \square

The results in theorem 2.9 have been extended to the *nonlinear* XPCI-model in [108], yet only under the additional assumption of *compactly supported* images. As will be seen in §2.5.1.3, support constraints indeed guarantee uniqueness already under much weaker assumptions.

2.5.1.3 General uniqueness of XPCI under support constraints

It is illustrative to write the general XPCI forward maps in a form such that the *holographic*- and *non-holographic* components of the contrast may be identified, compare §2.2.1.5:

$$\mathcal{T}(h) = -2\text{Re}(h) = -\mathcal{D}(o) - \overline{\mathcal{D}(o)} \quad \text{with} \quad o := h \quad (2.5.4a)$$

$$\mathcal{N}(h) = -\mathcal{T}(o) + |\mathcal{D}(o)|^2 = -\mathcal{D}(o) - \overline{\mathcal{D}(o)} + |\mathcal{D}(o)|^2 \quad \text{with} \quad o := 1 - \exp(-h), \quad (2.5.4b)$$

As \mathcal{D} is invertible, o can thus be uniquely recovered from XPCI-data if the superposition of $\mathcal{D}(o)$, $\overline{\mathcal{D}(o)}$ and $|\mathcal{D}(o)|^2$ (in the nonlinear case) on the r.h.s. of (2.5.4) may be disentangled.

Now assume that the image $h : \mathbb{R}^m \rightarrow \mathbb{C}$ has compact support. Then the same holds true for $o \in \{h, 1 - \exp(-h)\}$. From the alternate form of the Fresnel propagator (2.2.10) and the Paley-Wiener theorem, it follows that $\mathcal{D}(o)$ is an *entire (analytic) function*, i.e. $g_1 : \boldsymbol{\xi} \mapsto \mathcal{D}(o)(\boldsymbol{\xi})$ has a unique extension to \mathbb{C}^m with an everywhere convergent Taylor-series. The same holds true for $g_2 := \overline{\mathcal{D}(o)}$ and $g_3 := |\mathcal{D}(o)|^2$. The principal result of [140] is that the entire functions g_1, g_2, g_3 indeed behave so differently in \mathbb{C}^m that they may be distinguished – solely based on the assumption of a compact support of o (or h). This implies that $\mathcal{D}(o)$ and thus o may be uniquely reconstructed from $g \in \{\mathcal{T}(h), \mathcal{N}(h)\}$. In the linear setting, this means that $h = o$ is uniquely determined. For the nonlinear model $g = \mathcal{N}(h)$, $o = 1 - \exp(-h)$ determines h up to the phase-wrapping ambiguity discussed in §2.5.1.1.

Analyticity of $\mathcal{D}(o)$ and hence of the data g for compactly supported images h has another surprising and counter-intuitive consequence: if the XPCI-data g is known on an arbitrary open set $U \subset \mathbb{R}^m$, then it may be *completed* $g|_U \mapsto g$ in a unique manner by analytic continuation. Together with the insights above, this enables a surprisingly strong uniqueness-result [140]:

Theorem 2.10 (Uniqueness of XPCI under support constraints [140]). *Let $\Omega \subset \mathbb{R}^m$ be bounded and let $K \subset \mathbb{R}^m$ contain an open set. Then the following holds true:*

- Uniqueness of linearized XPCI: $\mathcal{T}_K : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m); h \mapsto \mathcal{T}(h)|_K$ is injective
- Essential uniqueness of nonlinear XPCI: $\mathcal{N}_K : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m); h \mapsto \mathcal{N}(h)|_K$ is injective up to phase-wrapping: if $\mathcal{N}(h_1)|_K = \mathcal{N}(h_2)|_K$ for $h_1, h_2 \in L^2(\Omega)$, then

$$h_1(\mathbf{x}) - h_2(\mathbf{x}) \in 2\pi i\mathbb{Z} \quad \text{for almost all } \mathbf{x} \in \mathbb{R}^m \quad (2.5.5)$$

The general statement in theorem 2.10 trivially carries over to settings with multiple holograms and/or additional constraints. The possible restriction of the data to $K \subset \mathbb{R}^m$ implies that – theoretically – holograms recorded by an arbitrarily small detector are sufficient for unique image reconstruction. In chapter 4, this will be seen to be merely a mathematical curiosity.

2.5.1.4 Extension to XPCT

Image reconstruction in XPCT (inverse problem 2.2) amounts to a combination of XPCI and tomographic inversion. Accordingly, uniqueness holds if both steps are uniquely solvable, i.e. if the following conditions hold:

- (1) For each incident direction $\boldsymbol{\theta} \in \Theta$, the hologram-data $I_{\boldsymbol{\theta}}^{\text{obs}}$ uniquely determines the projection-image $\mu_{\boldsymbol{\theta}} + i\phi_{\boldsymbol{\theta}} = \mathcal{P}_{\boldsymbol{\theta}}(f)$.
- (2) The projections $\{\mathcal{P}_{\boldsymbol{\theta}}(f)\}_{\boldsymbol{\theta} \in \Theta}$ uniquely determine the 3D-density f .

Uniqueness results for XPCT may thus be readily obtained by combining the well-established uniqueness-theory for (parallel-beam-) computed tomography (see e.g. [155, §II.3, §VI.2]) with the results from §2.5.1.2 and §2.5.1.3. In particular, theorem 2.10 implies that the inverse problem of XPCT is uniquely solvable for compactly supported objects in a very general setting (hologram-data $\{I_{\boldsymbol{\theta}}^{\text{obs}}|_K\}_{\boldsymbol{\theta} \in \Theta}$ for any $K \subset \mathbb{R}^2$ with non-empty interior and any set $\Theta \subset \mathbb{S}^2$ that contains an open interval of tomographic incident angles), see [140] for details.

2.5.2 Ill-posedness despite uniqueness

2.5.2.1 Data-errors and their origin

According to the theorems in §2.5.1, in particular theorem 2.10, image reconstruction in XPCI and XPCT is unique under extremely mild assumptions. However, uniqueness only guarantees that image reconstruction is feasible in hypothetical settings with *exact data*, whereas a large number of effects give rise to *data-errors* in real-world imaging:

- *Poisson- and other noise* ϵ_{noise} : As discussed in §2.1.5.4, the discrete number of detected photons gives rise to inevitable noise in the measured holograms. Additionally, noise may be induced by the detecting devices or by (cosmic) background radiation, for example.
- *Model errors* ϵ_{model} : The image-formation model XPCI is based on several approximations, so that that the true relation F^{true} between the sought images $f \in \{\phi, \mu, \delta, \beta\}$ and the data g may differ from the assumed model $F : f \mapsto g$. This gives rise to errors $\epsilon_{\text{model}} = F^{\text{true}}(f) - F(f)$ in the data g . Assuming a linear forward map (see §2.4.2) instead of the true nonlinear one induces such model errors. The same is true for all effects from §2.1.5, where particularly large errors may arise from flat-field correction (§2.1.5.2).
- *Discretization errors* ϵ_{discr} : Even if exact hologram-data g and a perfect forward model F was available, the true image f could still not be reconstructed because it lives (according to our model) in an infinite-dimensional function space. As detailed in §2.6, only a discretely sampled representation $\mathbf{f} \in \mathbb{R}^N$ may be computed by numerical reconstruction algorithms, based on a discretized model $F_{\text{discr}} : \mathbf{f} \mapsto g$. The transition to a discretized model induces additional errors $\epsilon_{\text{discr}} = F(f) - F_{\text{discr}}(\mathbf{f})$.

The total data-error is a superposition of the above contributions (and possibly others):

$$g^{\text{obs}} = g^{\text{ideal}} + \epsilon \quad \text{with} \quad \epsilon = \epsilon_{\text{noise}} + \epsilon_{\text{model}} + \epsilon_{\text{discr}}. \quad (2.5.6)$$

In particular, we conclude that data-errors are inevitable in real-world settings.

2.5.2.2 Stability and implications of ill-posedness

By the discussion in §2.5.2.1, *stability* to data-errors is required in order to yields satisfactory results in practice: if the data-errors ϵ in (2.5.6) are sufficiently small, the reconstructed images should ideally be close to the ground truth. Yet, even if a forward map $F : X \supset A \rightarrow Y$ is injective (uniqueness) and thus invertible, its inverse $F^{-1} : F(A) \rightarrow A$ may be *discontinuous*. In this case, there is no guarantee that reconstructions from perturbed data $I^{\text{obs}} = F(f) + \epsilon$ come anywhere close to the sought image f : due to the lack of continuity, the reconstruction-error $\|f - F^{-1}(F(f) + \epsilon)\|_X$ may be large even for arbitrarily small errors ϵ *in the data*. Despite uniqueness, the problem of inverting F is then ill-posed in the sense of definition 1.1.

On the contrary, if *well-posedness* holds, i.e. existence *and* continuity of the inverse F^{-1} , then at least convergence $F^{-1}(F(f) + \epsilon) \rightarrow f$ for vanishing data-errors $\epsilon \rightarrow 0$ is guaranteed. Hence, the impact of data-errors on the reconstruction is limited. Yet, it should be emphasized that the distinction between well- and ill-posed problems is purely *qualitative*: well-posedness alone does not bound the reconstruction-error in terms of ϵ . Obtaining such error-estimates necessitates a *quantitative* stability analysis as will be carried out in chapter 3.

2.5.2.3 Ill-posedness of XPCI beyond non-uniqueness

A thorough analysis of ill-posedness of XPCI, in particular for nonlinear forward maps, is beyond the scope here. Rather, we restrict to demonstrating that ill-posedness (beyond non-uniqueness) plays a role in general. We show this for simplest unique setting, given by a linear reconstruction of a homogeneous object from a single hologram (compare theorem 2.9):

Theorem 2.11 (General ill-posedness of linearized XPCI under homogeneity constraints). *For any Fresnel number $\mathfrak{f} \neq 0$ and $\nu \in \mathbb{R}$, the inverse of $\mathcal{S}_\nu : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is discontinuous, i.e. the associated inverse problem is ill-posed.*

Proof. By theorem 2.9, $\mathcal{S}_\nu : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is injective and hence invertible. As $\mathcal{S}_\nu(\varphi) = \mathcal{F}^{-1}(s_\nu \cdot \mathcal{F}(\varphi))$ for all $\varphi \in L^2(\mathbb{R}^m)$, the inverse is necessarily given by

$$\mathcal{S}_\nu^{-1} : \mathcal{S}_\nu(L^2(\mathbb{R}^m)) \rightarrow L^2(\mathbb{R}^m); g \mapsto \mathcal{F}^{-1}((1/s_\nu) \cdot \mathcal{F}(g)). \quad (2.5.7)$$

Due to zeros of the CTF s_ν , the inverse factor $1/s_\nu$ (that is well-defined almost everywhere) is unbounded: for any $\varepsilon > 0$, the domains $\Omega_\varepsilon := \{\boldsymbol{\xi} \in \mathbb{R}^m : |s_\nu(\boldsymbol{\xi})| \leq \varepsilon\}$ have non-zero (indeed infinite) Lebesgue-measure and it holds that $|1/s_\nu(\boldsymbol{\xi})| \geq \varepsilon^{-1}$ for all $\boldsymbol{\xi} \in \Omega_\varepsilon$.

Hence, there exist functions $0 \neq \varphi_\varepsilon \in L^2(\mathbb{R}^m)$ whose Fourier transforms are ε -concentrated around the CTF-zeros, i.e. $\text{supp}(\mathcal{F}(\varphi_\varepsilon)) \subset \Omega_\varepsilon$, for any $\varepsilon > 0$. If we set $g_\varepsilon := \mathcal{S}_\nu(\varphi_\varepsilon)$, then also $\text{supp}(\mathcal{F}(g_\varepsilon)) = \text{supp}(s_\nu \cdot \mathcal{F}(\varphi_\varepsilon)) \subset \Omega_\varepsilon$ and $g_\varepsilon \in \mathcal{S}_\nu(L^2(\mathbb{R}^m)) \setminus \{0\}$ since \mathcal{S}_ν is injective and $\varphi_\varepsilon \neq 0$. Moreover, the constructed functions satisfy

$$\begin{aligned} \|\mathcal{S}_\nu^{-1}(g_\varepsilon)\| &= \|\mathcal{F}(\mathcal{S}_\nu^{-1}(g_\varepsilon))\| = \|(1/s_\nu) \cdot \mathcal{F}(g_\varepsilon)\| \stackrel{\text{supp}(\mathcal{F}(g_\varepsilon)) \subset \Omega_\varepsilon}{=} \|(1/s_\nu) \cdot \mathcal{F}(g_\varepsilon)|_{\Omega_\varepsilon}\| \\ &\geq \left(\min_{\boldsymbol{\xi} \in \Omega_\varepsilon} |1/s_\nu(\boldsymbol{\xi})| \right) \cdot \|\mathcal{F}(g_\varepsilon)|_{\Omega_\varepsilon}\| = \varepsilon^{-1} \cdot \|g_\varepsilon\| \quad \text{for all } \varepsilon > 0. \end{aligned} \quad (2.5.8)$$

The estimate (2.5.8) shows that there does *not* exist a constant $M > 0$ such that $\|\mathcal{S}_\nu^{-1}(g)\| \leq M\|g\|$ for all $g \in \mathcal{S}_\nu(L^2(\mathbb{R}^m))$. Hence, \mathcal{S}_ν^{-1} is an unbounded linear map, i.e. discontinuous. \square

Ill-posedness despite uniqueness may likewise be shown for XPCI of general, inhomogeneous objects from two holograms in the setting of theorem 2.9. According to the proof of theorem 2.11, the ill-posedness is caused by the zeros of the CTF s_ν . These define Fourier-components of the image φ that induce low contrast in the hologram-data (see §2.4.2) and thus have to be amplified strongly upon reconstruction, thereby also amplifying the data-errors in the associated Fourier-frequencies. This is the reason why practitioners often seek to eliminate the CTF-zeros by acquiring holograms at several Fresnel numbers, compare §2.4.4.

2.5.2.4 Ill-posedness of tomographic reconstruction

In addition to image reconstruction in XPCI, also the tomographic reconstruction step involved in XPCT, $\{\mathcal{P}_\theta(f)\}_{\theta \in \Theta} \mapsto f$, is generally ill-posed on L^2 -spaces, even when uniqueness holds. By §2.3.2, the ill-posedness may be analyzed via known results for the inversion of the Radon-transform. Accordingly, it strongly depends on the tomographic incident directions $\Theta \subset \mathbb{S}^2$: while tomographic reconstruction in the case of full angular sampling (see §2.3.3), which corresponds to inverting the operator $\mathcal{P}_{\text{full}} : L^2(\Omega) \rightarrow L^2([0; \pi] \times \mathbb{R}^2)$, is only mildly ill-posed [155,

§IV.3], severe ill-posedness holds in the case of a *missing wedge*, i.e. if the tomographic angles cover a range of less than 180 degrees [155, §VI.2].

In general, tomographic reconstruction is a fairly standard, well-studied inverse problem, which is why the analysis in this thesis focuses on the image reconstruction problem of XPCI.

2.6 Discretization

2.6.1 General remarks

So far, we have always considered models of XPCI and XPCT where both the sought images $f \in \{\delta, \beta, \phi, \mu\}$ and hologram-intensities I are described as functions in continuous space \mathbb{R}^m . As seen in §2.1.5.4, real-world experiments however provide *discrete* intensity-measurements $\mathbf{I} \in \mathbb{R}^M$. Likewise, numerical reconstruction algorithms may only compute discrete, i.e. *pixelated*, representations $\mathbf{f} \in \mathbb{R}^N$ of the sought images because of finite computer-memory. Yet, there are still good reasons to consider descriptions in continuous space:

- (1) Continuous models are often closer to physical reality. For example, the spatial structure of real-world specimens is not expected to match any “natural” discretization and hence may be described more accurately in terms of continuous functions.
- (2) Although they require to work in infinite-dimensional function-spaces, continuous models often have favorable analytical properties that are not necessarily retained upon discretization. A prime example in XPCI is given by the distinct forms of the Fresnel propagator \mathcal{D} , compare §2.2.2: though equivalent in a continuous setting, standard discretizations of (2.2.5) and (2.2.10) do *not* yield identical discretized propagators [203].

In order to avoid complicated notation due to continuous descriptions in the theoretical sections and discrete ones in the algorithmic parts, discretization is regarded as *implicit* throughout this thesis: all inverse problems, results and also the principal reconstruction algorithms are phrased in terms of continuous quantities. Discretization is only explicitly discussed when it is not straightforward, i.e. when it deviates from the principal strategy outlined in §2.6.2.

Notably, discretization of inverse problems in L^p -spaces, as studied in this work, is indeed such a standard problem that there even exist software-solutions that do this job in a widely automated manner: the open-source project OPERATOR DISCRETIZATION LIBRARY (ODL) [1] for the programming language PYTHON allows to assign algorithms in terms of continuous-space models, automatically generating a runnable, discretized version of it via a strategy similar to the one explained below.

2.6.2 Detailed strategy

An image reconstruction algorithm defined in a continuous setting consists of a chain of *operations* applied to L^2 -functions:

$$g^{\text{obs}} \xrightarrow{O_1} (f_{1,1}, \dots, f_{1,n_1}) \xrightarrow{O_2} \dots \xrightarrow{O_{k-1}} (f_{k-1,1}, \dots, f_{k-1,n_{k-1}}) \xrightarrow{O_k} f^{\text{recon}} \quad (2.6.1)$$

where g^{obs} is the data, $f^{\text{recon}} \in L^2(\Omega)$ is the final reconstruction and $f_{i,j} \in L^2(\Omega_{i,j})$ denote intermediate results of the different operations O_i . The principal idea of a discretization lies in identifying (2.6.1) with an algorithm on *vectors* $\mathbf{g}^{\text{obs}} \in \mathbb{C}^M$, $\mathbf{f}^{\text{recon}} \in \mathbb{C}^N$, $\mathbf{f}_{i,j} \in \mathbb{C}^{N_{i,j}}$:

$$\mathbf{g}^{\text{obs}} \xrightarrow{O_1^{\text{discr}}} (\mathbf{f}_{1,1}, \dots, \mathbf{f}_{1,n_1}) \xrightarrow{O_2^{\text{discr}}} \dots \xrightarrow{O_{k-1}^{\text{discr}}} (\mathbf{f}_{k-1,1}, \dots, \mathbf{f}_{k-1,n_{k-1}}) \xrightarrow{O_k^{\text{discr}}} \mathbf{f}^{\text{recon}}, \quad (2.6.2)$$

which – other than the continuous-space version (2.6.1) – can be implemented on a computer.

Clearly, the identification between (2.6.1) and (2.6.2) should be such that the discretization-error is low: if \mathbf{g}^{obs} (approximately) parametrizes g^{obs} , then the result $\mathbf{f}^{\text{recon}}$ from (2.6.2) should approximate its correspondent f^{recon} from (2.6.1). To this end, functions g^{obs} , f^{recon} , f_i and operations O_i need to be discretized in a *mutually consistent* way, as detailed in the following.

2.6.2.1 Discretization of functions

We identify continuous and discretized functions $f \in L^2(\Omega)$ and $\mathbf{f} \in \mathbb{C}^N$ based on a disjoint partitioning of $\Omega = \bigcup_{i=1}^N \mathbb{P}_i$ into subdomains \mathbb{P}_i (of equal volume $\int_{\mathbb{P}_i} d\mathbf{x} = v > 0$, for simplicity), which can be thought of as *pixels* of an image (or *voxels* in a 3D-setting). Discrete and continuous quantities are mutually related by linear *sampling*- and *interpolation* operators $S : L^2(\Omega) \rightarrow \mathbb{R}^N$ and $E : \mathbb{R}^N \rightarrow L^2(\Omega)$. Though there are other possibilities, we consider sampling via averaging[¶] over the domains \mathbb{P}_i , combined with piecewise constant interpolation:

$$S(f)_i := \frac{1}{v} \int_{\mathbb{P}_i} f(\mathbf{x}) d\mathbf{x} \quad \text{for all } i = 1, \dots, N \quad (2.6.3)$$

$$E(\mathbf{f})(\mathbf{x}) := \begin{cases} f_i & \text{if } \mathbf{x} \in \mathbb{P}_i \\ 0 & \text{if } \mathbf{x} \notin \Omega \end{cases} \quad \text{for all } \mathbf{x} \in \mathbb{R}^m. \quad (2.6.4)$$

for $f \in L^2(\Omega)$ and $\mathbf{f} = (f_i)_{i=1}^N \in \mathbb{C}^N$. Note that $S(E(\mathbf{f})) = \mathbf{f}$ for all $\mathbf{f} \in \mathbb{C}^N$.

The operators S and E identify vectors $\mathbf{f} \in \mathbb{C}^N$ with a finite-dimensional subspace $E(\mathbb{C}^N) \subset L^2(\Omega)$ formed by piecewise constant functions $f = E(\mathbf{f})$, i.e. pixelated images in some sense. This identification provides the basic relation between the involved *objects* in the continuous and discretized algorithms (2.6.1), (2.6.2) – what remains is to relate the *operations* O_i, O_i^{discr} .

2.6.2.2 Discretization of operations

We consider general n -ary operations of the form $O : L^2(\Omega_1) \times \dots \times L^2(\Omega_n) \rightarrow L^2(\Omega_{n+1})$. Given discretizations of the spaces $L^2(\Omega_j)$ in terms of sampling- and interpolation operators $S_j : L^2(\Omega_j) \rightarrow \mathbb{C}^{N_j}$ and $E_j : \mathbb{C}^{N_j} \rightarrow L^2(\Omega_j)$, a canonical discretization of O is obtained by identifying input- and output-functions with vectors in \mathbb{C}^{N_j} :

$$O_{\text{discr}} : \mathbb{C}^{N_1} \times \dots \times \mathbb{C}^{N_n} \rightarrow \mathbb{C}^{N_{n+1}}; (\mathbf{f}_1, \dots, \mathbf{f}_n) \mapsto S_{n+1}(O(E_1(\mathbf{f}_1), \dots, E_n(\mathbf{f}_n))). \quad (2.6.5)$$

From studying the forward operators in §2.4, it can be seen that the most relevant operations arising in XPCI and XPCT comprise pointwise operations, L^2 -norms, parallel-beam projectors \mathcal{P}_θ , Fourier transforms \mathcal{F} and the Fresnel propagator \mathcal{D} . Upon discretization $O \rightsquigarrow O_{\text{discr}}$, all of these operations transform in the expected manner:

[¶]Note that the sampling operator may *not* be defined in the possibly expected manner, $S : f \mapsto (f(\mathbf{x}_i))_{i=1}^N$ for some $(\mathbf{x}_i)_{i=1}^N \subset \Omega$, since point-evaluations $f \mapsto f(\mathbf{x}_i)$ are not well-defined for L^2 -functions.

- *Pointwise* operations on functions, $O(f_1, \dots, f_n)(\mathbf{x}) = o(f_1(\mathbf{x}), \dots, f_n(\mathbf{x}))$ for all $\mathbf{x} \in \Omega$ and some scalar map $o : \mathbb{C} \times \dots \times \mathbb{C} \rightarrow \mathbb{C}$ ($o \in \{+, -, \cdot, /, (\cdot)^\gamma, \exp, \sin, \cos, \operatorname{Re}, \operatorname{Im}, |\cdot|, \dots\}$), exactly correspond to the associated *elementwise* vector-operations, $O_{\text{discr}}(\mathbf{f}_1, \dots, \mathbf{f}_n)_j = o(f_{1j}, \dots, f_{nj})$ for $\mathbf{f}_i = (f_{ij})_{j=1}^N \in \mathbb{C}^N$, since

$$O(E(\mathbf{f}_1), \dots, E(\mathbf{f}_n)) = E(O_{\text{discr}}(\mathbf{f}_1, \dots, \mathbf{f}_n)) \quad \text{for all } \mathbf{f}_1, \dots, \mathbf{f}_n \in \mathbb{C}^N. \quad (2.6.6)$$

- *Integrals* of functions $f = E(\mathbf{f}) \in E(\mathbb{C}^N)$ are represented by (*weighted*) *sums*:

$$\int_A E(\mathbf{f})(\mathbf{x}) \, d\mathbf{x} \stackrel{(2.6.4)}{=} \sum_{i=1}^N v_i(A) f_i, \quad v_i(A) := \int_{A \cap \mathbb{P}_i} d\mathbf{x} \quad \text{for } \mathbf{f} = (f_i)_{i=1}^N \in \mathbb{C}^N. \quad (2.6.7)$$

In particular, L^p -norms of functions are thus discretized by p -norms of vectors:

$$\|E(\mathbf{f})\|_{L^p}^p = \int_{\Omega} |E(\mathbf{f})(\mathbf{x})|^p \, d\mathbf{x} = \sum_{i=1}^N \underbrace{v_i(\Omega)}_{=v} |f_i|^2 = v \cdot \|\mathbf{f}\|_p^p, \quad (2.6.8)$$

- *Tomographic projections* $\mathcal{P}_\theta(f)$ amount to evaluating line-integrals over f and thus transform to weighted sums over vectors $\mathbf{f} \in \mathbb{C}^N$ for discretized functions $f = E(\mathbf{f})$, analogously as seen above. Details on- and specific variants of discretizations of tomographic projectors can be found for example in [209].
- *Fourier transforms* of functions $f = E(\mathbf{f}) \in E(\mathbb{C}^N)$ transform to *discrete Fourier transforms* (DFTs) of the underlying vectors \mathbf{f} :

$$\begin{aligned} \mathcal{F}(E(\mathbf{f}))(\boldsymbol{\xi}) &= (2\pi)^{-m/2} \int_{\Omega} \exp(-i\boldsymbol{\xi} \cdot \mathbf{x}) E(\mathbf{f})(\mathbf{x}) \, d\mathbf{x} \\ &\approx v(2\pi)^{-m/2} \sum_{i=1}^N \exp(-i\boldsymbol{\xi} \cdot \mathbf{x}_i) f_i \propto \text{DFT}(\mathbf{f})(\boldsymbol{\xi}), \quad \boldsymbol{\xi} \in \mathbb{R}^m \end{aligned} \quad (2.6.9)$$

for suitably chosen points $\mathbf{x}_i \in \mathbb{P}_i$ (for example the center-of-mass of \mathbb{P}_i). Quantitative estimates on the accuracy of the approximation in (2.6.9) are given for instance in [61].

In the standard case where the \mathbb{P}_i are identical rectangular pixels or voxels uniformly covering a finite Cartesian grid of size $N = N_1 \cdot \dots \cdot N_m$ and for frequencies $(\boldsymbol{\xi}_i)_{i=1}^N \subset \mathbb{R}^m$ arranged on an analogous grid, the map $\mathbf{DFT} : \mathbb{C}^N \rightarrow \mathbb{C}^N; \mathbf{f} \mapsto (\text{DFT}(\mathbf{f})(\boldsymbol{\xi}_i))_{i=1}^N$ may be evaluated at a computational complexity of $\mathcal{O}(N \log N)$ via *fast Fourier transforms* (FFTs). The same holds for true its inverse \mathbf{DFT}^{-1} .

- The *Fresnel propagator* \mathcal{D} may be formulated in different manners, compare §2.2.2:

$$\mathcal{D}(f) = \mathcal{F}^{-1}(m_{\mathfrak{f}} \cdot \mathcal{F}(f)) = k_{\mathfrak{f}} * f = u_0 n_{\mathfrak{f}} \cdot \mathcal{F}(n_{\mathfrak{f}} \cdot f)(\mathfrak{f}(\cdot)) \quad (2.6.10)$$

Each expression on the r.h.s. may be discretized via the replacement-rules for functions and operations derived above, leading to *different* discrete versions, all with their individual pros and cons. Indeed, designing accurate and efficient discretizations of \mathcal{D} remains a topic of current research [174, 203]. In this thesis, we exclusively use the most common form, based on the first variant on the r.h.s. of (2.6.10) (\odot : element-wise product):

$$\mathcal{D}_{\text{discr}}(\mathbf{f}) := \mathbf{DFT}^{-1}(\mathbf{m}_{\mathfrak{f}} \odot \mathbf{DFT}(\mathbf{f})) \approx \mathcal{D}(E(\mathbf{f})), \quad \mathbf{m}_{\mathfrak{f}} := (m_{\mathfrak{f}}(\boldsymbol{\xi}_i))_{i=1}^N. \quad (2.6.11)$$

Analogously to the continuous Fresnel propagator \mathcal{D} , this discretization is unitary as a map $\mathcal{D}_{\text{discr}} : \mathbb{C}^N \rightarrow \mathbb{C}^N$ with inverse $\mathcal{D}_{\text{discr}}^{-1} : \mathbf{f} \mapsto \mathbf{DFT}^{-1}(\overline{\mathbf{m}_{\mathfrak{f}}} \odot \mathbf{DFT}(\mathbf{f}))$.

2.6.2.3 Periodicity artifacts and padding

The Fourier transform \mathcal{F} is defined for functions in *infinite* space \mathbb{R}^m , whereas approximations by DFTs inherently assume *periodic* functions: finite sampling in Fourier space on a Cartesian grid with spacing $\Delta\xi$ induces $2\pi/\Delta\xi$ -periodicity in real-space (in each dimension). As a consequence, FFT-based numerical Fresnel propagation of images via the discretization $\mathcal{D}_{\text{discr}}$ in (2.6.11) implicitly imposes *periodic boundary conditions* at the edges of the image, which may lead to artifacts (*aliasing*). To suppress these, images may be enlarged by adding pixels of value zero at its boundaries (*zero-padding*) prior to applying $\mathcal{D}_{\text{discr}}$, thereby increasing the size of the computational domain and thus reducing the impact of periodicity. In Fourier space, this has the effect of refining the sampling-rate $\Delta\xi$. For details on numerical Fresnel propagation, padding and sampling-requirements, we refer to [75, §3.1.7], [139, §4.4.2] and [203].

2.7 Existing reconstruction methods

After the theoretical introduction and preliminary analysis of the inverse problems 2.1 and 2.2 of XPCI and XPCT, the present, final section of this chapter outlines some existing algorithms that may be applied to reconstruct images f from measured hologram-data $g^{\text{obs}} \approx F(f)$ *in practice*. By the general ill-posedness of image reconstruction (see §2.5), it is typically not desirable (and neither possible) to achieve this by applying the exact inverse F^{-1} of the forward map F to the data g^{obs} . Instead, practical reconstruction algorithms must necessarily involve some *regularization*, i.e. they apply some *stabilized* inverse F_α^{-1} which is obtained by supplementing the inverse problem with additional a priori knowledge on the unknown image, typically expressed as smallness in some norm. In practice, this may be achieved via various different algorithmic strategies, as discussed in standard textbooks on inverse problems, see e.g. [60, 112, 154]. The necessity of regularization thus introduces some arbitrariness to image reconstruction in XPCI and XPCT, which is why a wide range of algorithms exist for these tasks, all with their pros and cons. The following overview on reconstruction methods is not meant to be exhaustive but merely to cover the most relevant approaches for this thesis.

2.7.1 Generic methods

As shown in §2.4, all considered inverse problems in XPCI or XPCT can be phrased as an equation $F(f) \approx g^{\text{obs}}$ with a linear-bounded or nonlinear but Fréchet-differentiable operator $F : X \rightarrow Y$ on Hilbert-spaces X, Y . For such a setting, a wide range of reconstruction methods have been proposed in the literature (see e.g. [60, 112, 180, 179]) that are *generic* in the sense that they may be applied independently of specific properties of F . Examples include:

- *Quadratic Tikhonov regularization*: A reconstruction f^{recon} is computed as a compromise between closeness to a prior $f_0 \in X$ ($\|f - f_0\|_X$ small) and a small residual $\|F(f^{\text{recon}}) - g^{\text{obs}}\|_Y$, where the balance is determined by the regularization parameter $\alpha > 0$:

$$f^{\text{recon}} \in \underset{f \in X}{\operatorname{argmin}} \left\| \|F(f) - g^{\text{obs}}\|_Y^2 + \alpha \|f - f_0\|_X^2 \right. \quad (2.7.1)$$

- *Generalized Tikhonov regularization* (also called *variational reconstruction methods*): The squared norms in (2.7.1) are replaced by more general *data-fidelity-* and *penalty-functionals*

\mathcal{S}, \mathcal{R} , which increases flexibility in terms of exploiting a priori knowledge:

$$f^{\text{recon}} \in \underset{f \in X}{\operatorname{argmin}} \mathcal{S}(g^{\text{obs}}; F(f)) + \alpha \mathcal{R}(f) \quad (2.7.2)$$

- *Landweber's method*: Perform a finite number of gradient-descent steps on the quadratic residual $\|F(f) - g^{\text{obs}}\|_Y^2$ ($F'[f]^*$: adjoint of the Fréchet-derivative, compare §2.4.7.3)

$$\begin{aligned} f_{k+1} &:= f_k - \omega_k F'[f]^*(F(f_k) - g^{\text{obs}}) \quad \text{for } k = 0, 1, \dots, k_{\text{stop}} - 1, \quad \omega_k > 0 \\ f^{\text{recon}} &:= f_{k_{\text{stop}}} \end{aligned} \quad (2.7.3)$$

- *Regularized Newton-type methods*: Tikhonov-type reconstructions are iteratively computed on the basis of local *linearizations* $F(f) \approx F(f_k) + F'[f_k](f - f_k)$ to avoid algorithmic problems for nonlinear F . Such methods will be discussed in detail in chapter 5.

Owing to the abstract operator-formulation of XPCI and XPCT, all of the above methods (and many more) may be used for image reconstruction. For example, Tikhonov regularization has been applied to the *nonlinear* inverse problem of XPCI in [47, 49]. However, the generic nature of these basic algorithms has a significant drawback: as no specific mathematical properties of the inverse problems are exploited, computational costs are typically high. In particular, note that the minimizers in (2.7.1) and (2.7.2) generally have to be computed by generic iterative optimization algorithms that may require many steps to converge.

2.7.2 Phase reconstruction in XPCI

In the following, we present some specific algorithms for phase reconstruction, i.e. for solving the inverse problem 2.1 of XPCI. As the presented methods take advantage of the specific problem-structure, they may perform better than the generic approaches from §2.7.1. Note that we will only consider methods that are applicable in the *holographic regime* as encountered in high-resolution XPCI at synchrotrons (see §2.2.1.4) – for an overview on algorithms for the *direct contrast regime*, which is of minor interest to this work, we refer to [120].

2.7.2.1 Direct CTF-inversion

Probably the most widely used image reconstruction methods for holographic XPCI are still the direct CTF-inversion schemes that have been proposed already in the first experimental demonstrations of holographic XPCI and XPCT [40, 39, 42].

Formally, the approach boils down to quadratic Tikhonov regularization applied to the *linearized* inverse problem of XPCI for *homogeneous* objects, as modeled by the CTF-operator $F = \mathcal{S}_\nu^{(f_1, \dots, f_\ell)}$ (see §2.4.2 and §2.4.3): given hologram-data $g^{\text{obs}} = (g_1^{\text{obs}}, \dots, g_\ell^{\text{obs}}) \approx \mathcal{S}_\nu^{(f_1, \dots, f_\ell)}(\varphi)$ with $g_\ell^{\text{obs}} = I_\ell^{\text{obs}} - 1$ acquired at $\ell \in \mathbb{N}$ different Fresnel numbers f_1, \dots, f_ℓ , the phase- and absorption-image $\mu + i\phi = ie^{-i\nu}\varphi$ is reconstructed by computing

$$\varphi^{\text{recon}} \in \underset{\varphi \in L^2(\mathbb{R}^m, \mathbb{R})}{\operatorname{argmin}} \left(\sum_{i=1}^{\ell} \|\mathcal{S}_\nu^{(f_i)}(\varphi) - g_i^{\text{obs}}\|_{L^2}^2 \right) + \|w^{1/2} \cdot \mathcal{F}(\varphi)\|_{L^2}^2, \quad (2.7.4)$$

where the weighting-function $w > 0$ allows to set the degree of regularization in each Fourier-frequency individually. Typically, w is chosen such that a different regularization is applied in holographic and non-holographic spatial frequencies. For the CTF reconstructions considered in this thesis, we use the weighting proposed in [39] (erfc: complementary error function):

$$w(\boldsymbol{\xi}) := \alpha_1 r(|\boldsymbol{\xi}|) + \alpha_2 (1 - r(|\boldsymbol{\xi}|)), \quad r(\xi) := \frac{\operatorname{erfc}(\xi - \xi_{\text{cut}}/\sigma_{\text{cut}})}{\operatorname{erfc}(-\xi_{\text{cut}}/\sigma_{\text{cut}})}$$

with $\alpha_1, \alpha_2 > 0$, $\xi_{\text{cut}} := (\pi f)^{1/2}$, $\sigma_{\text{cut}} = \frac{\pi}{50}$. (2.7.5)

Thus far, the method is still *generic* in the sense of §2.7.1. Its XPCI-specific part consists in computing the minimizer φ^{recon} from (2.7.4) in a *direct* manner via an analytical solution: owing to the Fourier-multiplier form of the CTF-operators $\mathcal{S}_\nu^{(f_i)} : \varphi \mapsto \mathcal{F}^{-1}(s_\nu^{(f_i)} \cdot \mathcal{F}(\varphi))$, the first-order optimality-condition to (2.7.4) is linear and diagonalizes in Fourier space:

$$(2.7.4) \Leftrightarrow \left(\sum_{i=1}^{\ell} (\mathcal{S}_\nu^{(f_i)})^* (\mathcal{S}_\nu^{(f_i)}(\varphi^{\text{recon}}) - g_i^{\text{obs}}) \right) + \mathcal{F}^{-1}(w \cdot \mathcal{F}(\varphi^{\text{recon}})) = 0 \quad (2.7.6a)$$

$$\Leftrightarrow \left(\sum_{i=1}^{\ell} s_\nu^{(f_i)} \cdot (s_\nu^{(f_i)} \cdot \mathcal{F}(\varphi^{\text{recon}}) - \mathcal{F}(g_i^{\text{obs}})) \right) + w \cdot \mathcal{F}(\varphi^{\text{recon}}) = 0 \quad (2.7.6b)$$

$$\Leftrightarrow \varphi^{\text{recon}} = \mathcal{F}^{-1} \left(\frac{\sum_{i=1}^{\ell} s_\alpha^{(f_i)} \cdot \mathcal{F}(g_i^{\text{obs}})}{w + \sum_{i=1}^{\ell} |s_\alpha^{(f_i)}|^2} \right). \quad (2.7.6c)$$

Importantly, discretized forms of solution-formula (2.7.6c) may be implemented at computational costs of ℓ FFTs and a single inverse FFT, which renders the approach highly efficient. This is the main selling point of direct CTF-inversion schemes of the kind (2.7.6c). Similar inversion-formulas may also be derived in the absence of homogeneity constraints, i.e. for the general linear XPCI-model $F = \mathcal{F}^{(f_1, \dots, f_\ell)}$. Yet, this seems to be rarely done in practice.

Limitations: The computationally efficient direct CTF-inversion formula (2.7.6c) is fundamentally based on linearity and diagonality of the optimality-condition (2.7.6a) in Fourier space. Importantly, this rules out extensions of the approach to certain relevant settings:

- *Nonlinear reconstructions?* To account for nonlinearities in XPCI, the CTF-operator \mathcal{S}_ν in (2.7.4) would have to be exchanged by its nonlinear correspondent \mathcal{N}_ν . However, the resulting optimality-condition, i.e. the analogue of (2.7.6a), is then no longer linear nor diagonal in Fourier space and no closed-form solution is known to hold.
- *Non-negativity constraints?* If the minimization in (2.7.4) is restricted to non-negative functions $\varphi \geq 0$, then the optimality-condition (2.7.6a) turns into a *nonlinear* equation for φ^{recon} , that does not have a (known) closed-form solution.
- *Support constraints?* The diagonality of the formula (2.7.6c) in Fourier space is related to *translation-invariance in real-space*: if we define shift-operators $T_{\mathbf{a}} : f \mapsto f(\cdot - \mathbf{a})$, then any reconstruction-method of the form $\varphi^{\text{recon}} = R(g_1^{\text{obs}}, \dots, g_\ell^{\text{obs}}) := \sum_{i=1}^{\ell} \mathcal{F}^{-1}(\lambda_i \cdot \mathcal{F}(g_i^{\text{obs}}))$ satisfies $R(T_{\mathbf{a}}g_1^{\text{obs}}, \dots, T_{\mathbf{a}}g_\ell^{\text{obs}}) = T_{\mathbf{a}}R(g_1^{\text{obs}}, \dots, g_\ell^{\text{obs}})$ for all $\mathbf{a} \in \mathbb{R}^m$ and $g_i^{\text{obs}} \in L^2(\mathbb{R}^m)$. Imposing a support constraint, $\text{supp}(\varphi^{\text{recon}}) \subset \Omega$ for some $\emptyset \neq \Omega \subsetneq \mathbb{R}^m$, necessarily breaks

this translation-invariance as it requires that $R(T_{\mathbf{a}}g_1^{\text{obs}}, \dots, T_{\mathbf{a}}g_\ell^{\text{obs}})|_{\mathbb{R}^m \setminus \Omega} = 0$ for arbitrary data $g_1^{\text{obs}}, \dots, g_\ell^{\text{obs}}$ under *all* possible shifts $\mathbf{a} \in \mathbb{R}^m$. The associated loss of diagonality in Fourier space renders an exact solution of the optimality-condition (a *non-sparse* linear system) numerically infeasible for discretized problems of practically relevant size.

We conclude that all of the above settings are fundamentally incompatible with direct CTF-inversion. This constitutes a major motivation to seek alternate reconstruction methods.

2.7.2.2 Alternating projection-type methods

Constrained and/or nonlinear reconstructions in XPCI are to date mostly computed via a class of algorithms which we refer to as *alternating-projection* (AP) methods [76, 48, 14, 165, 87]. AP-type algorithms, such as the error-reduction (Gerchberg-Saxton) algorithm [74], the hybrid-input-algorithm [65] and relaxed averaged alternating reflections (RAAR) [135] are commonly applied to the classical phase retrieval problem of recovering an image f from the magnitude of its Fourier transform $|\mathcal{F}(f)|^2$ [182, 66]. Such problems arise in X-ray coherent diffractive imaging (CDI), the far-field-imaging analogue of XPCI (compare §2.2.1.5). The algorithmic approach may be readily adapted to the *near-field* XPCI-setting.

As detailed in [17], the general mathematical idea behind AP-type methods is to phrase image reconstruction as a (non-convex) feasibility problem: both measured data and available a priori on the unknown image f are expressed via *constraint-sets* $C_i := \{f : f \text{ fits the } i\text{th data}\}$ for $i = 1, \dots, \ell$ and $C_{\ell+i} := \{f : f \text{ satisfies the } i\text{th constraint}\}$ for $i = 1, \dots, k$, such that the reconstruction is cast to finding an element in the intersection, $f^{\text{recon}} \in \bigcap_{i=1}^{\ell+k} C_i$. In order to find the solution, AP-algorithms iteratively and alternately *project* onto the sets C_i , i.e. compute an element in C_i with minimal (L^2 -)distance to the current image:

$$P_{C_i}(f) \in \underset{g \in C_i}{\operatorname{argmin}} \|f - g\|^2 \quad (2.7.7)$$

The exact usage of projectors P_{C_i} is what makes the difference between the specific AP-type methods. The RAAR-algorithm for example also makes use of *reflectors* $R_{C_i} : f \mapsto 2P_{C_i}(f) - f$. However, the simplest variant is to compute image-iterates f_k by cyclically applying the P_{C_i} in some fixed order $i_1, \dots, i_{\ell+k}$, which corresponds to the Gerchberg-Saxton algorithm:

$$f_{k+1} = P_{C_{i_{\ell+k}}} \dots P_{C_{i_1}}(f_k) \quad \text{for } k = 0, 1, \dots, k_{\text{stop}} - 1. \quad (2.7.8)$$

AP-methods take advantage of the specific problem-structure of XPCI in that the projectors P_{C_i} may be evaluated via efficient analytical formulas for all of the constraint-sets C_i that typically arise in reconstruction problems. For example, recovering a pure phase image ϕ from $\ell \in \mathbb{N}$ holograms $I_\ell^{\text{obs}} = |\mathcal{D}^{(f_\ell)}(\exp(-i\phi))|^2$ (exact data for simplicity) subject to a support constraint $\operatorname{supp}(\phi) \subset \Omega$ corresponds to finding $f := \exp(-i\phi) \in \bigcap_{i=1}^{\ell+2} C_i$ with

$$C_i := \{f : |\mathcal{D}(f)|^2 = I_\ell^{\text{obs}}\} \quad \Rightarrow \quad P_{C_i}(f) = \mathcal{D}^{-1} \left((I_\ell^{\text{obs}})^{\frac{1}{2}} \cdot \operatorname{sign}(\mathcal{D}(f)) \right), \quad i \leq \ell \quad (2.7.9a)$$

$$C_{\ell+1} := \{f : |f| = 1\} \quad \Rightarrow \quad P_{C_{\ell+1}}(f) = \operatorname{sign}(f) \quad (2.7.9b)$$

$$C_{\ell+2} := \{f : \operatorname{supp}(f - 1) \in \Omega\} \quad \Rightarrow \quad P_{C_{\ell+2}}(f) = 1 + (f - 1)|_\Omega \quad (2.7.9c)$$

with $\operatorname{sign}(z) = z/|z|$ for complex-numbers $z \in \mathbb{C} \setminus \{0\}$ and $\operatorname{sign}(0) \in \{z \in \mathbb{C} : |z| = 1\}$ arbitrary. Although the sets $C_1, \dots, C_{\ell+1}$ are *non-convex*, all P_{C_i} in (2.7.9) are efficiently computable.

2.7.3 Two-step reconstruction methods for XPCT

Image reconstruction in XPCT does not only require phase retrieval as in XPCI but also tomographic reconstruction. The standard approach is to solve these two sub-problems independently and subsequently: given XPCT-data of the form $I_{\boldsymbol{\theta}}^{\text{obs}} \approx 1 + F(\mathcal{P}_{\boldsymbol{\theta}}(f))$ for incident directions $\boldsymbol{\theta} \in \Theta$ and any XPCI-forward map F , the following two steps are performed:

- (1) (*phase reconstruction*) For each incident direction $\boldsymbol{\theta} \in \Theta$, reconstruct the tomographic projection $p_{\boldsymbol{\theta}} = \mathcal{P}_{\boldsymbol{\theta}}(f)$ from the hologram-data $I_{\boldsymbol{\theta}}^{\text{obs}} \approx 1 + F(p_{\boldsymbol{\theta}})$.
- (2) (*tomographic inversion*) Recover the 3D-object f from the projection-data $(p_{\boldsymbol{\theta}}^{\text{recon}})_{\boldsymbol{\theta} \in \Theta} \approx (\mathcal{P}_{\boldsymbol{\theta}}(f))_{\boldsymbol{\theta} \in \Theta}$ obtained in step (1).

Step (1) simply corresponds to image reconstruction in XPCI and thus can be implemented using any of the methods from §2.7.2. The tomographic inversion step (2), on the other hand, is a standard problem encountered in a wide range of other imaging modalities beyond X-ray phase contrast, such as medical CT and electron-tomography. As a consequence, a vast number of algorithms are available to solve this sub-problem, too large to be reviewed here. In particular, variational schemes of the kind (2.7.2) are an area of active research for tomographic problems. Similarly as in medical CT [163], however, by far the most commonly used algorithm in XPCT is still *filtered back-projection* (FBP) because it is for once well-understood and also sufficiently computationally efficient to remain applicable for large-scale tomographic data sets composed of $\mathcal{O}(10^3)$ images with $\mathcal{O}(10^6)$ pixels each. Therefore, FBP will be used as a reference for the methods proposed in this thesis and its principle is outlined in the following.

2.7.3.1 Filtered back-projection (FBP)

In its most basic form, FBP is an implementation of an exact inversion-formula for the 2D-Radon-transform \mathcal{R} (see §2.3.1) that goes back to the pioneering work [168] of JOHANN RADON himself. As such it guarantees exact reconstruction from fully angularly sampled tomographic data of the form $p_{\text{full}} = \mathcal{P}_{\text{full}}(f)$, compare §2.3.3. In real-world settings with projections $p_{\boldsymbol{\theta}}^{\text{obs}} \approx \mathcal{P}_{\boldsymbol{\theta}}(f)$ acquired for *finitely* many incident directions $\boldsymbol{\theta} \in \Theta$, the approach is no longer theoretically exact but may still yield highly accurate results for appropriate, sufficiently densely sampled sets Θ . FBP amounts to computing a reconstruction via the formula

$$f^{\text{recon}} := \sum_{\boldsymbol{\theta} \in \Theta} \mathcal{P}_{\boldsymbol{\theta}}^* (\mathcal{F}^{-1} (w_{\boldsymbol{\theta}} \cdot \mathcal{F}(p_{\boldsymbol{\theta}}^{\text{obs}}))), \quad (2.7.10)$$

i.e. by back-projecting *filtered* versions of the measured projections $p_{\boldsymbol{\theta}}^{\text{obs}}$. The weighting-factors $w_{\boldsymbol{\theta}}$ depend on Θ and may also be used to impose regularization. Mathematical details on the choice of $w_{\boldsymbol{\theta}}$ and on FBP in general can be found in standard textbooks on CT [155, 110, 33].

As the filtering-step in (2.7.10) may be efficiently implemented based on FFTs, the computational costs of an FBP-reconstruction are relatively low, essentially amounting to that of one back-projection along all incident directions, $(p_{\boldsymbol{\theta}})_{\boldsymbol{\theta} \in \Theta} \mapsto \sum_{\boldsymbol{\theta} \in \Theta} \mathcal{P}_{\boldsymbol{\theta}}^*(p_{\boldsymbol{\theta}})$. Moreover, FBP-algorithms are easy to *parallelize*, which is crucial for high-performance computations that are required in the case of large-scale data sets. A drawback of FBP is that it is incompatible with imposing support constraints or non-negativity on the recovered object f^{recon} due to the inherent linearity and translation-invariance of the filtering-based approach, analogously as discussed for direct CTF-inversion in §2.7.2.1.

2.7.4 All-at-once reconstruction methods for XPCT

Alternatively to the *sequential* two-step approach presented in §2.7.3, image reconstruction in XPCT may also be achieved by so-called *joint-* or *all-at-once* methods that perform phase reconstruction of the holograms and tomographic inversion in a combined manner. Different algorithms implementing this principal concept have been proposed recently [139, 175, 176, 119], which will be reviewed in §2.7.4.1 to §2.7.4.3.

The general motivation for joint methods is to exploit *tomographic consistency* to stabilize the phase reconstruction step: by recovering a 3D-object density f^{recon} directly from the complete tomographic hologram-series $(I_{\theta}^{\text{obs}})_{\theta \in \Theta}$, it is implicitly imposed that the corresponding projection-data $(p_{\theta}^{\text{recon}})_{\theta \in \Theta} = (\mathcal{P}_{\theta}(f^{\text{recon}}))_{\theta \in \Theta}$ is consistent in terms of the Helgason-Ludwig conditions from theorem 2.2. On the contrary, this is in general *not* the case if the projections p_{θ} are reconstructed from I_{θ}^{obs} for each $\theta \in \Theta$ independently without reference to the underlying 3D-object. Effectively, joint reconstruction thus imposes an additional *consistency constraint* in the phase reconstruction step: the set of possible solutions A in the (implicitly) solved inverse problem “reconstruct $(p_{\theta})_{\theta \in \Theta} \in A$ from $(I_{\theta}^{\text{obs}})_{\theta \in \Theta}$ ” is restricted to

$$A \subset C_{\text{tomog}} := \{(p_{\theta})_{\theta \in \Theta} : p_{\theta} = \mathcal{P}_{\theta}(f) \text{ for all } \theta \in \Theta \text{ for an admissible } f\}, \quad (2.7.11)$$

which is expected to reduce ill-posedness by §2.5 and thereby improve reconstruction-quality.

A further benefit of joint reconstruction methods is that additional constraints, such as support- and/or non-negativity, may be imposed directly on the 3D-object f rather than on its projections $p_{\theta} = \mathcal{P}_{\theta}(f)$. As will be further discussed in chapter 3, this often permits a tighter confinement of the admissible solutions.

2.7.4.1 Joint reconstruction by generic methods

Conceptually the simplest way to construct all-at-once methods is to apply any of the generic algorithms from §2.7.1 to the full inverse problem of XPCT $F_{\text{PCT}}(f) \approx (I_{\theta}^{\text{obs}} - 1)_{\theta \in \Theta}$. Tomographic consistency is then fully encoded in the structure of the employed forward operator $F_{\text{PCT}} : f \mapsto (F(\mathcal{P}_{\theta}(f)))_{\theta \in \Theta}$ and thus automatically imposed by the reconstruction algorithm. Such a generic all-at-once approach has been implemented in [139].

In general, while genericity yields flexibility in terms of the choice of the underlying XPCI-model F (homogeneous, nonlinear and/or multiple holograms?), the main drawback of the approach is computational performance: In high-resolution XPCT, the numerical size of the inverse problem of XPCT ($\gtrsim 10^9$ dimensions of discrete object- and data-spaces) is typically too large to compute joint reconstructions by generic methods in tolerable runtimes.

2.7.4.2 Iterative reprojection phase retrieval

In [175], a joint reconstruction algorithm termed “iterative reprojection phase retrieval” (IRP) has been proposed that is based on an adaptation of the alternating-projection (AP) approach from §2.7.2.2 to the 3D-setting of XPCT. The principal idea is to reconstruct tomographic projections $(p_{\theta})_{\theta \in \Theta}$ via standard AP-iterations from the holograms $(I_{\theta}^{\text{obs}})_{\theta \in \Theta}$ for all θ and to enforce tomographic consistency of the projection-iterates $(p_{\theta,k})_{\theta \in \Theta}$ by interlacing *reprojection* steps: intermediate 3D-objects are reconstructed by (approximate) least-square-minimization

$f_k \in \operatorname{argmin}_{f \in A} \|(\mathcal{P}_\theta(f))_{\theta \in \Theta} - (p_{\theta,k})_{\theta \in \Theta}\|^2$ and then $(p_{\theta,k})_{\theta \in \Theta}$ is replaced by the reprojected object $(\mathcal{P}_\theta(f_k))_{\theta \in \Theta}$ before proceeding to the next sweep of AP-iterations. In the light of the general AP-scheme from §2.7.2.2, this can be viewed as (approximate) projection-steps of the iterates $(p_{\theta,k})_{\theta \in \Theta}$ onto the tomographic consistency constraint-set C_{tomo} defined in (2.7.11).

The pros and cons of the IRP-algorithm are similar to the generic approach from §2.7.4.1: while easily adaptable to different XPCT-models and constraints by the flexible AP-structure, the recurrent tomographic reconstructions to be computed give rise to high computational costs.

2.7.4.3 3D-phase reconstruction

Another approach to joint reconstruction has been proposed independently in [119, 176]. It is based on the observation that the CTF-based forward operators of linearized XPCI *commute* with the parallel-beam projectors \mathcal{P}_θ , i.e. for any well-behaved function f and any $\theta \in \mathbb{S}^2$

$$T(\mathcal{P}_\theta(f)) = \mathcal{P}_\theta(T^{(3d)}(f)) \quad \text{for } T \in \{\mathcal{T}, \mathcal{S}_\nu\}. \quad (2.7.12)$$

$T^{(3d)} : L^2(\mathbb{R}^3) \rightarrow L^2(\mathbb{R}^3)$ denotes an analogue of the image-formation operator T acting on *3D-objects* instead of 2D-images. The relation (2.7.12) follows from the fact that $\mathcal{T}, \mathcal{S}_\nu$ are given by isotropic Fourier-multipliers, as will be detailed at a later point in §3.3.3.

For linear XPCT-data of the form $(T(\mathcal{P}_\theta(f)))_{\theta \in \Theta} \approx (I_\theta^{\text{obs}} - 1)_{\theta \in \Theta}$, (2.7.12) implies that the order of phase- and tomographic reconstruction may be *reversed* compared to the standard two-step procedure from §2.7.3: if $\text{tomoRec}(\cdot)$ denotes any tomographic-inversion method such that $\text{tomoRec}((p_\theta)_{\theta \in \Theta}) \approx f$ if $p_\theta \approx \mathcal{P}_\theta(f)$ for all $\theta \in \Theta$ (for example FBP, compare §2.7.3.1), then an application of this algorithm to the tomographic *hologram-series* yields

$$\text{tomoRec}((I_\theta - 1)_{\theta \in \Theta}) \stackrel{(2.7.12)}{\approx} \text{tomoRec}((\mathcal{P}_\theta(T^{(3d)}(f)))_{\theta \in \Theta}) \approx T^{(3d)}(f). \quad (2.7.13)$$

Accordingly, *3D-hologram data* $T^{(3d)}(f)$ is obtained, from which the object f may be recovered in a second step by applying a *3D-variant* of any of the phase reconstruction algorithms from §2.7.2. As detailed in [176, 174], tomographic consistency is implicitly encoded in the constructed 3D-phase retrieval problem and thus automatically exploited by the approach. In this sense, the method achieves the principal goal of all-at-once algorithms for XPCT even though, strictly speaking, phase- and tomographic reconstruction are still performed sequentially.

3D-phase retrieval $T^{(3d)}(f) \mapsto f$ may be implemented in a highly efficient manner via direct CTF-inversion, compare §2.7.2.1. Moreover, even if the step is performed by an iterative method (to impose additional constraints), the total algorithm may amount to much less computational effort than those from §2.7.4.1 and §2.7.4.2. The reason is that the parallel-beam projectors \mathcal{P}_θ are eliminated from the problem via the initial “tomoRec”-step (2.7.13) so that no costly tomographic (back-)projections have to be computed in the iterations. The downside of the presented 3D-phase retrieval approach is that it seems to be fundamentally limited to *linear* XPCT-models, as the exploited relation (2.7.12) is no longer valid in nonlinear settings.

Designing joint reconstruction algorithms for XPCT that are both efficiently implementable and based on the *nonlinear* XPCI-model is a main concern of the algorithmic chapters 5 and 6.

Part II

Summary and Discussion of the Publications

Chapter 3

Stability estimates for linearized near-field phase retrieval in X-ray phase contrast imaging

Article 1 [144] deals with stability of XPCI in the linearized (CTF-)regime. Accordingly, the theoretical study investigates how robust the required image reconstruction step is to errors in the measured data. Thereby, it sheds a light on the controversial question how many holograms have to be acquired to enable accurate imaging in real-world XPCI.

3.1 Motivation

Ever since the pioneering works of Peter Cloetens and co-authors [40, 42, 39] it has been widely accepted that at least two holograms acquired at different object-to-detector-distances are necessary for faithful image reconstruction in holographic XPCI [108, 157, 158, 32]. The typical reasoning is that the zeros of the oscillatory contrast-transfer-functions (CTFs), see §2.4.2, correspond to Fourier-components of the object that are practically invisible in a single hologram. As seen in theorem 2.11, this gives rise to ill-posedness. The situation is considered to be even more severe when phase-shifts and absorption are not coupled by a homogeneity constraint (compare §2.4.3), but have to be reconstructed as independent images. Indeed, we have seen in §2.5.1.2 that image reconstruction may be non-unique in this setting. Physically, this non-uniqueness relates to the well-known *twin-image problem* of holography [72, 73].

On the other hand, recent works have demonstrated that accurate reconstructions from a single hologram can be achieved, given a careful treatment of the CTF-zeros [150] or by imposing *support constraints* (see §2.4.6) on the unknown image [14, 13], i.e. exploiting a priori knowledge on the size of the imaged sample. The success of the latter approach is in line with theorem 2.10, stating that a compact support guarantees uniqueness of image reconstruction in XPCI. While this uniqueness result seems to strongly contradict some observations by previous authors, a few aspects should be emphasized: for once, it can be shown that support constraints rule out the existence of phase-vortices [140], which serve as a standard counter-example [157] to prove non-uniqueness (see theorem 2.8). Likewise, support-knowledge allows to circumvent the twin-image problem, as will be seen in §3.2.3.2. Moreover, while it is perfectly true that CTF-zeros correspond to missing information when images of theoretically infinite extent are

to be reconstructed, support constraints allow to infer the missing Fourier frequencies as will be detailed in §3.2.3.2. Finally, it is a well-known fact in inverse problems that mere uniqueness does *not* guarantee reconstructability in any practical sense, as the reconstruction might still be arbitrarily *ill-posed*, i.e. *unstable* to errors in the data, as discussed in §2.5.2.

To assess in which settings accurate image reconstruction is feasible for real-world, imperfect data, the uniqueness theorem 2.10 has to be supplemented with *stability estimates*. This is the principal goal of the presented article. Due to the analytical difficulties arising from nonlinearity, the analysis is restricted to the linearized (CTF-) forward maps from §2.4.2. This is justified by the large regime-of-validity of the CTF-model found in several studies, see e.g. [194, 96].

3.1.1 Teaser: the impact of support constraints in practice

Analogously as uniqueness in theorem 2.10, we will establish stability of image reconstruction under the assumption of support constraints. In order to get an intuition for such constraints and how much can be gained from them *in practice*, we consider numerical reconstructions from exemplary real XPCI-data given by a single hologram I^{obs} shown in fig. 3.1(a). The underlying sample is a pure phase object to very good approximation (see §2.4.3) and moderately weak so that the sought phase-image ϕ can be recovered by linear CTF-inversion schemes (§2.7.2.1) from the data $I^{\text{obs}} - 1 \approx \mathcal{S}_0(\phi)$. Moreover, it can be inferred from the plotted hologram in fig. 3.1(a) that the sample only fills a small subdomain in the center of the imaged field-of-view, so that a support constraint $\text{supp}(\phi) \subset \Omega$ ($\Leftrightarrow \phi|_{\Omega^c} = 0$) holds for a suitable domain $\Omega \subset \mathbb{R}^2$.

To assess the effects of support-knowledge, we compute numerical reconstructions ϕ^{recon} via CTF-inversion schemes with and without support constraints, i.e. (up to discretization):

- (1) Standard CTF-inversion without support constraint ($\alpha_1 = 10^{-3}$, $\alpha_2 = 10^{-2}$):

$$\phi_1^{\text{recon}} \in \underset{\phi \in L^2(\mathbb{R}^2, \mathbb{R})}{\text{argmin}} J_{\text{CTF}, \alpha_1, \alpha_2}(\phi), \quad (3.1.1)$$

where $J_{\text{CTF}, \alpha_1, \alpha_2}$ denotes the Tikhonov-functional defined by (2.7.4) and (2.7.5).

- (2) CTF-inversion supplemented with a support constraint (α_1 and α_2 as in (1)):

$$\phi_2^{\text{recon}} \in \underset{\phi \in L^2(\Omega, \mathbb{R})}{\text{argmin}} J_{\text{CTF}, \alpha_1, \alpha_2}(\phi) \quad (3.1.2)$$

with $\Omega \subset \mathbb{R}^2$ given by the circular-region marked in fig. 3.1(a)

- (3) CTF-inversion with support and *without regularization* ($\alpha_1 = \alpha_2 = 0$ and Ω as in (2)):

$$\phi_3^{\text{recon}} \in \underset{\phi \in L^2(\Omega, \mathbb{R})}{\text{argmin}} J_{\text{CTF}, 0, 0}(\phi). \quad (3.1.3)$$

Note: While the constrained methods (2) and (3) are easy to write down, solving the associated discretized optimization problems requires an *iterative* method and is thus much harder than for the standard method (1), which may implemented efficiently as a *direct* inversion in Fourier space, compare §2.7.2.1. We will get back to this *algorithmic* problem in §5.1.

The reconstructed phase-images ϕ^{recon} from (1) to (3) are plotted in fig. 3.1(b)–(d), respectively. Additionally, subfigures (e)–(g), i.e. the bottom-row of fig. 3.1, show zooms of the red-dashed square for the corresponding images in the top-row. We observe the following:

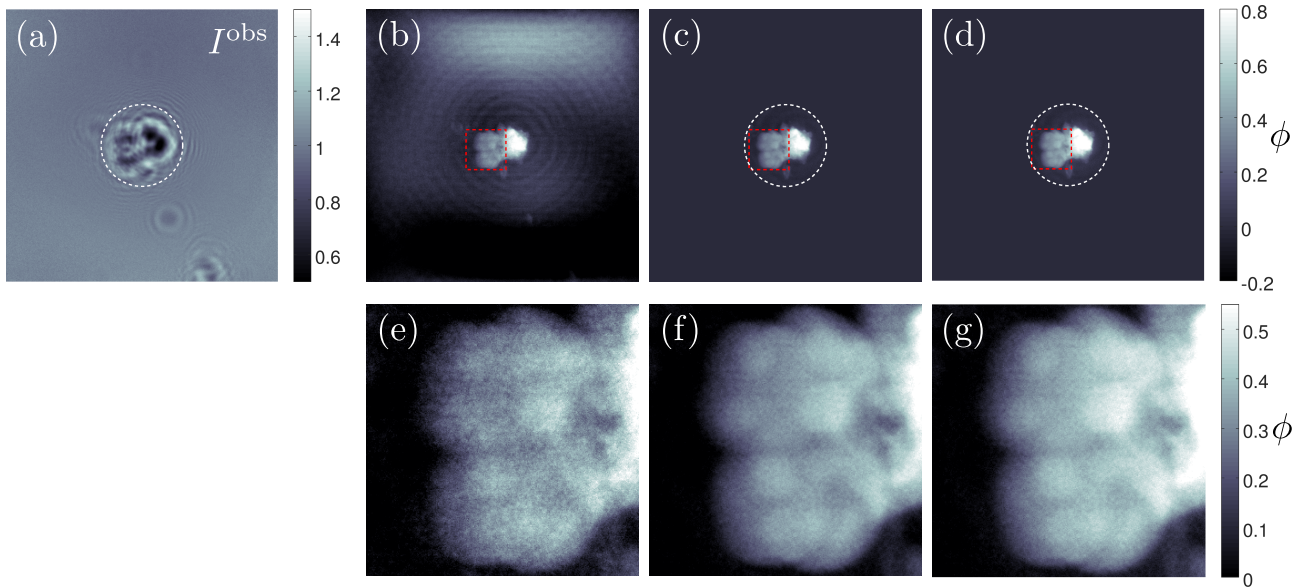


Figure 3.1. Effect of a support constraint on XPCI-reconstructions of a pure phase object.

(a) Measured hologram I^{obs} (real data for freeze-dried *d. radiodurans* bacteria imaged at the GINIX (see §1.2); 2048×2048 pixels of effective size $\Delta x_{\text{pix}}^{\text{eff}} \approx 13.4$ nm at Fresnel number $f_{\text{pix}} \approx 1.16 \cdot 10^{-4}$). The white-dashed circle in marks the assumed support in the reconstructions in (c) and (d).

(b)–(d) Linear CTF-reconstructions (see §2.7.2.1) of the phase-image ϕ : (b) Standard method without support constraint, regularization parameters $\alpha_1 = 10^{-3}$, $\alpha_2 = 10^{-2}$. (c) α_1, α_2 as in (b) but imposing the white-dashed circle as a support. (d) support as in (c) but *without* regularization, $\alpha_1 = \alpha_2 = 0$.

(e)–(g) Zooms of the red-dashed square for the respective images in the top-row.

- Reconstruction **(1)** (no support) shows pronounced bright- and dark artifacts in the background. These are due to slight background variations in the hologram in fig. 3.1(a) arising from imperfect flat-field correction (see §2.1.5.2), which are amplified in the reconstruction due to low-frequency instability of the CTF, $s_0(\boldsymbol{\xi}) \approx 0$ for $\boldsymbol{\xi} \approx 0$ (see §2.4.2).
- In reconstruction **(2)** (with support), background-variations may not arise outside the circular domain Ω because $\phi^{\text{recon}}|_{\Omega^c} = 0$ is imposed. Yet, notably, the low-frequency artifacts are also suppressed *within* Ω , revealing a stabilizing effect of the support constraint.
- The zooms in fig. 3.1(e),(f) furthermore reveal that also fine structural details come out much more clearly in **(2)** than in **(1)**, as noise is suppressed considerably.
- The non-regularized reconstruction **(3)** is of similar quality as **(2)**. On the contrary, such a reconstruction *completely fails* without a support constraint (results not shown). Accordingly, the support-knowledge seems to make regularization redundant in the example, indicating that the ill-posedness arising from the CTF-zeros (see §2.5.2) is eliminated.

The stability analysis of the presented article provides a theoretical explanation for the observed strong stabilizing effect of support constraints in practical image reconstruction.

3.1.2 Lipschitz-stability estimates and their meaning

Although weaker notions of stability are more commonly used in the theory of inverse problems (known as conditional stability estimates, see e.g. [105]), we find that *Lipschitz-stability* is a

suitable concept for the considered linearized XPCI-problems. An inverse problem described by a forward map $F : A \subset X \rightarrow Y$ between normed spaces with a set of admissible solutions A is said to be *Lipschitz-stable* if a stability-estimate of the form

$$\|F(f_1) - F(f_2)\|_Y \geq C_{\text{stab}} \|f_1 - f_2\|_X \quad \text{for all } f_1, f_2 \in A \quad (3.1.4)$$

holds for some *stability constant* $C_{\text{stab}} > 0$. Intuitively, the bound (3.1.4) ensures that any substantial modification of the object from f_1 to f_2 results in sufficient *contrast* $\|F(f_1) - F(f_2)\|_Y$ in the observable data, i.e. that any such change is *visible* in practice.

If (3.1.4) is satisfied, F is injective with Lipschitz-continuous inverse $F^{-1} : F(A) \rightarrow A$: $\|F^{-1}(g_1) - F^{-1}(g_2)\|_X \leq C_{\text{stab}}^{-1} \|g_1 - g_2\|_Y$ for all $g_1, g_2 \in F(A)$. Most importantly, this implies *robustness to data-errors*: if an object $f^\dagger \in A$ is to be recovered from perturbed measurements $g^{\text{obs}} = F(f^\dagger) + \epsilon \in F(A)$, then the resulting deviation of the reconstructed object $f^{\text{recon}} := F^{-1}(g^{\text{obs}})$ from f^\dagger is bounded by

$$\|f^\dagger - f^{\text{recon}}\|_X = \|F^{-1}(F(f^\dagger)) - F^{-1}(F(g^{\text{obs}}))\|_X \leq C_{\text{stab}}^{-1} \|\epsilon\|_Y. \quad (3.1.5)$$

According to (3.1.5), the data-errors are at most amplified by a *finite* factor C_{stab}^{-1} upon reconstruction. In particular, this implies *well-posedness* of the inverse problem by definition 1.1: if $\epsilon \rightarrow 0$, then $f^{\text{recon}} \rightarrow f^\dagger$. Quantitatively, (3.1.5) furthermore reveals the following:

To guarantee truly small object-deviations due to data-errors, i.e. stability in a practical sense, the stability constant C_{stab} should be as large as possible.

On the contrary, if C_{stab} is very small, the error-amplification predicted by (3.1.5) may be too large to allow for accurate reconstruction at realistic noise-levels $\|\epsilon\|_Y$.

In the presented article, we exclusively consider *linear*-bounded forward maps F and $A \subset X$ given by linear subspaces. In this case, Lipschitz-stability (3.1.4) is equivalent to

$$C_{\text{stab}} = \inf_{f \in A, \|f\|_X=1} \|F(f)\|_Y > 0. \quad (3.1.6)$$

Moreover, a linear inverse problem is well-posed *if and only if* it is Lipschitz-stable.

3.2 Summary of the results

3.2.1 Setting and assumptions

We consider the setting of linearized XPCI, as described by the CTF-based forward models from sections 2.4.2 to 2.4.4. Note that we study the forward maps in arbitrary dimensions $m \in \{1, 2, 3, \dots\}$ although the physical XPCI-setting corresponds to $m = 2$ dimensions. The benefit of this will be seen in §3.3. Continuous and complete measurements I^{obs} of the hologram-intensities are assumed, which may be corrupted by noise. It is furthermore assumed that a bound on the maximal support $\Omega \subset \mathbb{R}^m$ of the unknown image $h \in L^2(\Omega)$ is available. The choices of Ω will typically be domains of diameter 1 in some sense. Importantly, this implies that the (modified) Fresnel numbers \mathfrak{f} are implicitly defined w.r.t. the support size, see §2.2.1. Typical experimental values of \mathfrak{f} in synchrotron-based XPCI are thus very large, $10^2 \lesssim \mathfrak{f} \lesssim 10^5$.

The principal inverse problem, describing image reconstruction from a single hologram under a support constraint without additional a priori knowledge, is of the following form:

Inverse Problem 3.1 (Linearized XPCI of weak general objects). For a given support $\Omega \subset \mathbb{R}^m$, recover a complex-valued image $h = \mu + i\phi \in L^2(\Omega)$ from a noisy hologram

$$I^{\text{obs}} - 1 \approx \mathcal{I}(h).$$

Solving inverse problem 3.1 is commonly regarded as practically infeasible, in spite of being uniquely solvable for bounded domains Ω according to theorem 2.10. To improve stability, practitioners typically assume *homogeneity* of the imaged object (including in particular the case of *pure phase objects*, see §2.4.3), which corresponds to the following setting:

Inverse Problem 3.2 (Linearized XPCI of weak homogeneous objects). For a given support $\Omega \subset \mathbb{R}^m$, recover a real-valued image $\varphi \in L^2(\Omega, \mathbb{R})$ from a noisy hologram

$$I^{\text{obs}} - 1 \approx \mathcal{I}_\nu(\varphi).$$

Alternatively to restricting the images to be recovered, it is also commonly argued that the acquisition of additional holograms may stabilize image reconstruction. Here, the case of *two* holograms acquired at different Fresnel numbers $f_1 \neq f_2$ is considered, as described by the following forward operator (compare §2.4.2 and §2.4.4 for notation):

$$\mathcal{I}^{(f_1, f_2)}(h) = \begin{pmatrix} \mathcal{I}^{(f_1)}(h) \\ \mathcal{I}^{(f_2)}(h) \end{pmatrix} = \begin{pmatrix} \mathcal{I}^{(f_1)}(\phi) + \mathcal{C}^{(f_1)}(\mu) \\ \mathcal{I}^{(f_2)}(\phi) + \mathcal{C}^{(f_2)}(\mu) \end{pmatrix} \quad \text{for } h = \mu + i\phi \in L^2(\mathbb{R}^m). \quad (3.2.1)$$

Inverse Problem 3.3 (Linearized XPCI from two holograms). For a given support $\Omega \subset \mathbb{R}^m$, recover a complex-valued image $h = \mu + i\phi \in L^2(\Omega)$ from *two* noisy holograms

$$\begin{pmatrix} I_1^{\text{obs}} - 1 \\ I_2^{\text{obs}} - 1 \end{pmatrix} \approx \mathcal{I}^{(f_1, f_2)}(h).$$

3.2.2 Principal results

3.2.2.1 General objects (inverse problem 3.1)

The principal result of the article is that inverse problems 3.1 to 3.3 are *well-posed* if the support-domain Ω is bounded, i.e. Lipschitz-stability holds true in the sense of §3.1.2. It is most surprising this is even true for the most challenging setting, given by inverse problem 3.1:

Theorem 3.1 (Well-posedness and stability estimate for inverse problem 3.1). *Let the support-domain Ω be given by a stripe of width 1, w.l.o.g. $\Omega := [-\frac{1}{2}; \frac{1}{2}] \times \mathbb{R}^{m-1}$. Then it holds that*

$$C_{\text{IP1}}(\Omega, f) := \inf_{h \in L^2(\Omega), \|h\|=1} \|\mathcal{I}(h)\| > 0 \quad (3.2.2)$$

i.e. inverse problem 3.1 is well-posed. The stability constant satisfies the asymptotic estimate

$$C_{\text{IP1}}(\Omega, f) \geq (2\pi f)^{\frac{1}{4}} \left(1 - \frac{3}{8f} + \mathcal{O}(f^{-2}) \right) \exp(-f/8) \quad \text{for } f \rightarrow \infty. \quad (3.2.3)$$

If $\Omega = [1/2; 1/2]^m$ is a cube, the r.h.s. of (3.2.3) may be improved by a factor of $m^{1/2}$.

Notably, the asymptotic bound (3.2.3) on the stability constant C_{IP1} decays exponentially in the modified Fresnel number \mathfrak{f} . Indeed, numerical eigenvalue-computations in the article (cf. §8.3.5) indicate that the bound is *not* overly pessimistic but asymptotically sharp for large \mathfrak{f} . Accordingly, C_{IP1} decays dramatically fast in \mathfrak{f} : while $C_{\text{IP1}}(\Omega, 10) \geq 0.77$ (neglecting the $\mathcal{O}(\mathfrak{f}^{-2})$ -terms in (3.2.3)), one only has $C_{\text{IP1}}(\Omega, 10^2) \geq 1.8 \cdot 10^{-5}$ and $C_{\text{IP1}}(\Omega, 10^3) \geq 4.5 \cdot 10^{-54}$. As the factor C_{IP1}^{-1} governs data-error-amplification in image reconstruction according to §3.1.2, the estimate thus allows for dramatic instability at Fresnel numbers $\mathfrak{f} \gtrsim 10^2$, i.e. inverse problem 3.1 may be excessively ill-*conditioned* despite being well-*posed* under support constraints.

Beyond the worst-case stability estimate (3.2.2), the article also provides specific bounds for arbitrary image-modes, based on the singular value decomposition of a related operator, see §8.3.4: for a cube-shaped support-domain $\Omega = [1/2; 1/2]^m$, it is found that the least stable modes are given by prolate spheroidal wave-functions of low order, modulated by the Fresnel-factor $\exp(i\mathfrak{f}\mathbf{x}^2/4)$. Accordingly, the image-components that are most difficult to reconstruct in inverse problem 3.1 are *low-frequency modes*, whereas high-frequency components induce strong contrast in the hologram-data and may thus be stably recovered.

3.2.2.2 Homogeneous objects (inverse problem 3.2)

The stability results from theorem 3.1 clearly remain valid for inverse problem 3.2 as it differs from inverse problem 3.1 merely by a restriction of the admissible objects. Notably, the stability bound $C_{\text{IP1}} \gtrsim \mathfrak{f}^{1/4} \exp(-\mathfrak{f}/8)$ may be furthermore improved to an algebraic decay in \mathfrak{f} :

Theorem 3.2 (Well-posedness and stability estimate for inverse problem 3.2). *Let the support-domain be a ball of diameter 1, $\Omega := B[\mathbf{a}; \frac{1}{2}]$ for any $\mathbf{a} \in \mathbb{R}^m$. Then the stability constant of inverse problem 3.2 is bounded by*

$$C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) := \inf_{\varphi \in L^2(\Omega, \mathbb{R}), \|\varphi\|=1} \|\mathcal{S}_\nu(\varphi)\| \geq \max \left\{ \min \{c_1, c_2 \mathfrak{f}^{-1}\}, \min \{c_3 \nu, c_4 \mathfrak{f}^{-\frac{1}{2}}\} \right\} \quad (3.2.4)$$

for some constants $c_j > 0$ that depend only on the dimension m . In particular, $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) \geq \mathcal{O}(\mathfrak{f}^{-1})$ for $\nu = 0$ and $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) \geq \mathcal{O}(\mathfrak{f}^{-1/2})$ for $\nu > 0$ as $\mathfrak{f} \rightarrow \infty$.

The stability estimate in theorem 3.2 is less explicit than in theorem 3.1 as the values of the constants c_j are not given. Indeed, the proof does provide numerical values for these. However, they are expected to be highly non-optimal and therefore omitted here. On the contrary, the *order* of the asymptotic decay of the stability constant, $C_{\text{IP2}}(\Omega, \mathfrak{f}, 0) \gtrsim \mathfrak{f}^{-1}$ and $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) \gtrsim \mathfrak{f}^{-1/2}$ for $\nu > 0$ as $\mathfrak{f} \rightarrow \infty$, is shown to be optimal in the article, see §8.4.5.

Remark. Although the constant C_{IP2} has been defined as an infimum over *real-valued* $\varphi \in L^2(\Omega, \mathbb{R})$ in theorem 3.2, the value is the same for general, complex-valued arguments of \mathcal{S}_ν , i.e. $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) = \inf_{\varphi \in L^2(\Omega), \|\varphi\|=1} \|\mathcal{S}_\nu(h)\|$. This property will be needed in theorem 3.3 and can be proven using the fact that \mathcal{S}_ν commutes with the pointwise real-part $\text{Re}(\cdot)$.

3.2.2.3 Reconstruction from two holograms (inverse problem 3.3)

A similar improvement of the stability constant as in the homogeneous case is obtained in the two-holograms-settings of inverse problem 3.3. Indeed, it is found that the stability of the

latter problem is related to that of inverse problem 3.2 for $\nu = 0$. This insight gives rise to the following theorem:

Theorem 3.3 (Stability estimate for inverse problem 3.3). *Let $\mathcal{S}_0^{(f_-)}$ denote the forward operator of inverse problem 3.2 for $\nu = 0$ and $\mathbf{f} = \mathbf{f}_- := |\mathbf{f}_1^{-1} - \mathbf{f}_2^{-1}|^{-1}$. Then*

$$\|\mathcal{I}^{(f_1, f_2)}(h)\| \geq 2^{-\frac{1}{2}} \|\mathcal{S}_0^{(f_-)}(h)\| \quad \text{for all } h \in L^2(\mathbb{R}^m). \quad (3.2.5)$$

In particular, for any support-domain $\Omega \subset \mathbb{R}^m$, the following stability estimate holds true:

$$C_{\text{IP}_3}(\Omega, \mathbf{f}_-) := \inf_{h \in L^2(\Omega): \|h\|=1} \|\mathcal{I}^{(f_1, f_2)}(h)\| \geq 2^{-\frac{1}{2}} C_{\text{IP}_2}(\Omega, \mathbf{f}_-, 0) \quad (3.2.6)$$

Notably, the stability constant in theorem 3.3 depends on the difference between the (reciprocal) Fresnel numbers $\mathbf{f}_1, \mathbf{f}_2$. According to the bounds (3.2.4) and (3.2.6), a favorably large stability constant $C_{\text{IP}_3}(\Omega, \mathbf{f}_-)$ is thus obtained for pairs $\mathbf{f}_1, \mathbf{f}_2$ that differ strongly in value, i.e. if the two holograms are acquired in significantly distinct XPCI-setups.

3.2.2.4 Algebraic properties of C_{IP^*} and adaptation to other geometries

Theorems 3.1 to 3.3 assume very specific geometries of the support-domain Ω , which might seem to restrict their applicability at the first glance. However, basic algebraic properties of the stability constants provide a simple handle for generalization of the results:

Theorem 3.4 (Algebraic properties of the stability constants). *Let $C_{\text{IP}^*}(\Omega, \mathbf{f})$ be either of the stability constants $C_{\text{IP}_1}(\Omega, \mathbf{f})$ or $C_{\text{IP}_2}(\Omega, \mathbf{f}, \nu)$ for fixed ν . Then the following properties hold:*

- (a) Translational- and rotational invariance: $C_{\text{IP}^*}(\mathbf{a} + A \cdot \Omega, \mathbf{f}) \geq C_{\text{IP}^*}(\Omega, \mathbf{f})$ for all $\Omega \subset \mathbb{R}^m$, $\mathbf{f} > 0$, $\mathbf{a} \in \mathbb{R}^m$ and orthogonal matrices $A \in \mathbb{R}^{m \times m}$
- (b) Monotonicity in Ω : $C_{\text{IP}^*}(\Omega_1, \mathbf{f}) \geq C_{\text{IP}^*}(\Omega_2, \mathbf{f})$ for all $\Omega_1 \subset \Omega_2 \subset \mathbb{R}^m$, $\mathbf{f} > 0$.
- (c) Scale invariance: $C_{\text{IP}^*}(r\Omega, \mathbf{f}) = C_{\text{IP}^*}(\Omega, r^2\mathbf{f})$ for all $r > 0$, $\Omega \subset \mathbb{R}^m$, $\mathbf{f} > 0$.
- (d) Monotonicity in \mathbf{f} : $C_{\text{IP}^*}(\Omega, \mathbf{f}_1) \geq C_{\text{IP}^*}(\Omega, \mathbf{f}_2)$ for all $0 < \mathbf{f}_1 \leq \mathbf{f}_2$, $\Omega \subset \mathbb{R}^m$.

Proof. Let $T \in \{\mathcal{I}, \mathcal{S}_\nu\}$. Then property (a) follows from the fact that T commutes with shifts and orthogonal transformations of the coordinates: $T(h(\mathbf{a} + A(\cdot))(\mathbf{x})) = T(h)(\mathbf{a} + A\mathbf{x})$ for any $h \in L^2(\mathbb{R}^m)$ and $\mathbf{x} \in \mathbb{R}^m$. For (b) note that $L^2(\Omega_1) \subset L^2(\Omega_2)$. Hence, it holds that

$$C_{\text{IP}^*}(\Omega_1, \mathbf{f}) = \inf_{h \in L^2(\Omega_1), \|h\|=1} \|T(h)\| \geq \inf_{h \in L^2(\Omega_2), \|h\|=1} \|T(h)\| = C_{\text{IP}^*}(\Omega_2, \mathbf{f}).$$

Property (c) follows by a rescaling of the coordinate system $\mathbf{x}' = \mathbf{x}/r$, which implicitly transforms the Fresnel number to $\mathbf{f}' = r^2\mathbf{f}$ by its scale-dependence explained in §2.2.1.3. Finally, (d) is a consequence of (b) and (c): upon setting $\Omega_1 := \Omega$ and $\Omega_2 := (\mathbf{f}_2/\mathbf{f}_1)^{1/2} \cdot \Omega \supset \Omega_1$, one has

$$C_{\text{IP}^*}(\Omega, \mathbf{f}_1) = C_{\text{IP}^*}(\Omega_2, \mathbf{f}_1) \stackrel{(c)}{=} C_{\text{IP}^*}(\Omega_1, \mathbf{f}_2) \stackrel{(b)}{\geq} C_{\text{IP}^*}(\Omega_2, \mathbf{f}_2) = C_{\text{IP}^*}(\Omega, \mathbf{f}_2). \quad \square$$

Despite its simplicity, theorem 3.4 bears two important consequences:

- (1) If the exact support-domain Ω is geometrically complicated, properties (a)–(c) ensure that the resulting stability constant may be bounded from below by that for one of the simple domains treated in theorems 3.1 to 3.3. For example, whenever Ω has a finite diameter $d := \sup_{\mathbf{x}, \mathbf{y} \in \Omega} |\mathbf{x} - \mathbf{y}| < \infty$, then $\Omega \subset B[\mathbf{a}; d/2]$ for some ball and thus

$$C_{\text{IP}^*}(\Omega, \mathbf{f}) \stackrel{(b)}{\geq} C_{\text{IP}^*}(B[\mathbf{a}; d/2], \mathbf{f}) \stackrel{(a)}{=} C_{\text{IP}^*}(B[0; d/2], \mathbf{f}) \stackrel{(c)}{=} C_{\text{IP}^*}(B[0; 1/2], d^2 \mathbf{f}) \quad (3.2.7)$$

- (2) The monotonic decrease in \mathbf{f} implies that the function $\mathbf{f} \mapsto C_{\text{IP}^*}(\Omega, \mathbf{f})$ is well-approximated by a finite number of samples $(\mathbf{f}_j, C_{\text{IP}^*}(\Omega, \mathbf{f}_j))$ as it limits the variation between subsequent samples $\mathbf{f}_j < \mathbf{f}_{j+1}$: $C_{\text{IP}^*}(\Omega, \mathbf{f}_j) \geq C_{\text{IP}^*}(\Omega, \mathbf{f}) \geq C_{\text{IP}^*}(\Omega, \mathbf{f}_{j+1})$ for all $\mathbf{f} \in [\mathbf{f}_j; \mathbf{f}_{j+1}]$. Importantly, this means that the functional dependence $\mathbf{f} \mapsto C_{\text{IP}^*}(\Omega, \mathbf{f})$ is accessible by numerical computations, as will be exploited in §3.4.

3.2.3 Ideas of the proofs

3.2.3.1 General objects (inverse problem 3.1)

The stability analysis for inverse problem 3.1 is based on the principle of holographic image reconstruction, that earned DENNIS GABOR the 1971 Nobel Prize in Physics [72, 71]. The main observation is that the linearized contrast in a hologram may be written as a superposition of a propagated image and a *back-propagated twin-image*, the pointwise complex conjugate \bar{h} :

$$-\mathcal{I}(h) = 2\text{Re}(\mathcal{D}(h)) = \mathcal{D}(h) + \mathcal{D}^{-1}(\bar{h}) \quad (3.2.8)$$

Gabor concluded that an application of the back-propagator \mathcal{D}^{-1} to a hologram would thus recover the focused image h , perturbed by some fringes originating from the twin-image contribution $\mathcal{D}^{-2}(\bar{h})$, and thereby permit *qualitative* image reconstruction.

In the present work, Gabor's principle is exploited in a converse manner: if a support-bound $\text{supp}(h) \subset \Omega$ is available, then the perturbing twin-image-contribution can be eliminated by propagating it into focus and then restricting the data to the complement of the support:

$$-\mathcal{D}(\mathcal{I}(h))|_{\Omega^c} = \mathcal{D}^2(h)|_{\Omega^c} + \underbrace{\bar{h}|_{\Omega^c}}_{=0} = \mathcal{D}^2(h)|_{\Omega^c} \quad \text{for all } h \in L^2(\Omega). \quad (3.2.9)$$

According to (3.2.9), the transformation yields incomplete Fresnel-data $\mathcal{D}^2(h)|_{\Omega^c}$. To obtain a stability estimate from this, the L^2 -norm is taken on both sides of (3.2.9) and it is exploited that $\|\mathcal{D}(f)|_A\| \leq \|f\|$ for all $f \in L^2(\mathbb{R}^m)$ and any measurable set $A \subset \mathbb{R}^m$:

$$\|\mathcal{I}(h)\| \geq \|\mathcal{D}(\mathcal{I}(h))|_{\Omega^c}\| = \|\mathcal{D}^2(h)|_{\Omega^c}\| \quad \text{for all } h \in L^2(\Omega). \quad (3.2.10)$$

By noting that $\mathcal{D}^2 = \mathcal{D}^{(f/2)}$ and expressing the propagator $\mathcal{D}^{(f/2)}$ via the alternate form (2.2.10), the right-hand side of (3.2.10) may furthermore be rewritten as the Fourier transform of a Fresnel-modulated image $\tilde{h} := n_{f/2} \cdot h$ with $n_{f/2}(\mathbf{x}) = \exp(i\mathbf{f}\mathbf{x}^2/4)$: $\|\mathcal{D}^2(h)|_{\Omega^c}\| = \|\mathcal{F}(\tilde{h})|_{(f/2\cdot\Omega)^c}\|$. Inserting this into expression into (3.2.10) finally gives

$$\|\mathcal{I}(h)\| \geq \|\mathcal{F}(\tilde{h})|_{(f/2\cdot\Omega)^c}\| \quad \text{for all } h \in L^2(\Omega), \tilde{h} = n_{f/2} \cdot h. \quad (3.2.11)$$

Since $\|\tilde{h}\| = \|h\|$, the bound (3.2.11) can be interpreted as a *relative* stability estimate: the reconstruction of an image h from XPCI-data $\mathcal{S}(h)$ is at least as stable as the reconstruction of \tilde{h} from its Fourier-data outside the domain $\mathfrak{f}/2 \cdot \Omega$. Reconstructing functions $\tilde{h} \in L^2(\Omega)$ from partial Fourier-data in turn constitutes a well-studied class of problems, known under the keyword of *analytic continuation* in the literature because of the analyticity of the Fourier transforms of compactly supported functions. In particular, the considered problem is *well-posed*, $\|\mathcal{F}(\tilde{h})|_{(\mathfrak{f}/2 \cdot \Omega)^c}\| \geq C_{\text{IP1}} \|\tilde{h}\|$ for some $C_{\text{IP1}} > 0$, provided that Ω is bounded along at least one dimension. Quantitative bounds on the Lipschitz-constant C_{IP1} can be obtained for the special case of rectangular domains Ω by reducing the setting to an analytic-continuation-problem on 1D-intervals, as analyzed in [185, 186]. By (3.2.11), these results for reconstructions from incomplete Fourier data directly translate into stability estimates for linearized XPCI. In particular, the asymptotic expansions for the eigenvalues of some related compact self-adjoint operator provided in [186] give rise to the asymptotic formula (3.2.3). Furthermore, the full singular value decomposition of this operator induces a characterization of stability for individual image-modes in terms of prolate spheroidal wave functions.

3.2.3.2 Homogeneous objects (inverse problem 3.2)

The stability proof for inverse problem 3.2 works completely differently from the above approach for the case of general, non-homogeneous objects. It exploits that contrast is non-negligible in all Fourier-frequencies except for the zeros of the CTF s_ν , as seen from

$$|\mathcal{F}(\mathcal{S}_\nu(\varphi))(\boldsymbol{\xi})| = |s_\nu(\boldsymbol{\xi})| \cdot |\mathcal{F}(\varphi)(\boldsymbol{\xi})| = \left| \sin\left(\frac{\boldsymbol{\xi}^2}{2\mathfrak{f}}\right) \right| \cdot |\mathcal{F}(\varphi)(\boldsymbol{\xi})|. \quad (3.2.12)$$

The reconstruction of L^2 -images $\varphi \in L^2(\mathbb{R}^m, \mathbb{R})$ from $\mathcal{S}_\nu(\varphi)$ is *unstable* because $\mathcal{F}(\varphi)$ may be arbitrarily concentrated around the zeros of s_ν , in which case the relative contrast $\|\mathcal{S}_\nu(\varphi)\|/\|\varphi\|$ is small according to (3.2.12). This is illustrated in fig. 3.2(a).

In order to understand how support constraints introduce stability in the CTF-model, it is crucial to note that such an arbitrary concentration of $\mathcal{F}(\varphi)$ around the CTF-zeros is *not* possible if φ is compactly supported. Indeed, there is a well-known class of theorems in Fourier-analysis, known as *uncertainty principles*, stating that a spatial confinement of a signal in real-space – as imposed by a support constraint – induces a *minimum* width of structures in Fourier space. Hence, $\mathcal{F}(\varphi)$ may only be finitely peaked at the CTF-zeros so that neighboring Fourier-frequencies will always give rise to some finitely small total data-contrast, see fig. 3.2(b).

To obtain quantitative estimates from this insight, the Fourier-domain is decomposed into stable regions and such around the CTF-zeros (\sqcup : disjoint union):

$$\mathbb{R}^m = D_\varepsilon \sqcup \bigsqcup_{j=0}^{\infty} B_j \quad \text{with} \quad D_\varepsilon := \{\boldsymbol{\xi} \in \mathbb{R}^m : |s_\nu(\boldsymbol{\xi})| \geq \sin(\varepsilon)\}, \quad 0 < \varepsilon < \pi/6, \quad (3.2.13)$$

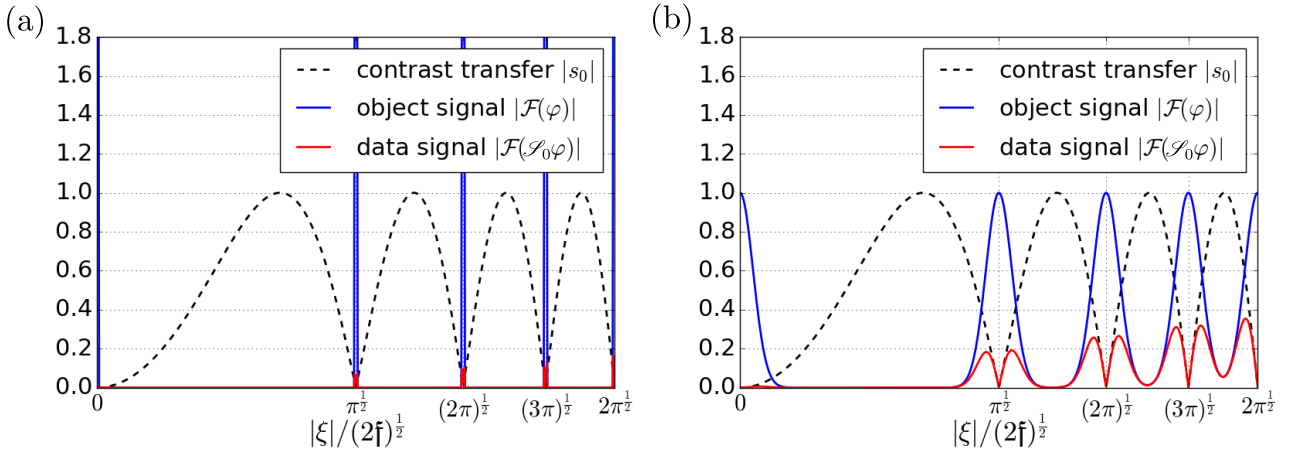


Figure 3.2. Illustration of the principal argument for stability of inverse problem 3.3: the black-dashed line plots the radially symmetric CTF s_ν for $\nu = 0$. Blue and red solid lines show examples of Fourier space signals of objects φ and their images under the forward operator \mathcal{S}_ν , respectively. (a) General images $\varphi \in L^2(\mathbb{R}^m, \mathbb{R})$: $\mathcal{F}(\varphi)$ may be arbitrarily peaked at the zero-manifolds of s_ν . (b) Support constraint $\varphi \in L^2(\Omega, \mathbb{R})$, Ω bounded: $\mathcal{F}(\varphi)$ is smooth and has finitely sharp peaks.

where B_0 is some small disc around the origin and B_j are annular neighborhoods around the j th zero-manifold of s_ν at $|\boldsymbol{\xi}| = (2f)^{\frac{1}{2}}(j\pi - \nu)^{\frac{1}{2}}$, $j \in \mathbb{N}$. With (3.2.12), this decomposition yields

$$\begin{aligned} \|\mathcal{S}_\nu(\varphi)\|^2 &= \|\mathcal{F}(\mathcal{S}_\nu(\varphi))\|^2 = \|(s_\nu \cdot \mathcal{F}(\varphi))|_{D_\varepsilon}\|^2 + \sum_{j=0}^{\infty} \|(s_\nu \cdot \mathcal{F}(\varphi))|_{B_j}\|^2 \\ &\geq \sin(\varepsilon)^2 \|\mathcal{F}(\varphi)s|_{D_\varepsilon}\|^2 + \sum_{j=0}^{\infty} \|(s_\nu \cdot \mathcal{F}(\varphi))|_{B_j}\|^2, \end{aligned} \quad (3.2.14)$$

where the bound in the second line follows by definition of the stable sub-domain D_ε .

Within the domains B_j , $|s_\nu|$ is not bounded from below but exhibits first- or second order zero-crossings so that the remaining terms on the r.h.s. of (3.2.14) have to be estimated via a different approach. It is shown in the article that estimates of the form

$$\|(s_\nu \cdot \mathcal{F}(\varphi))|_{B_j}\|^2 \geq \lambda_j(\varepsilon) \|\mathcal{F}(\varphi)|_{B_j}\|^2 + f\eta_j(\varepsilon) \int_{B_j^-} \Delta |\mathcal{F}(\varphi)(\boldsymbol{\xi})|^2 d\boldsymbol{\xi} \quad (3.2.15)$$

with some $\lambda_j(\varepsilon) > 0$, $\eta_j(\varepsilon) \geq 0$ and $B_j^- \subset B_j$ hold true for all $j \in \mathbb{N}_0$, where Δ is the Laplace-operator. The integral with the Laplacian term is then bounded by exploiting the smoothness of $\mathcal{F}(\varphi)$ induced by the support constraint, as characterized by the following lemma:

Lemma 3.5. *Let $g \in L^2(\Omega)$ with support in $\Omega = B[0; R]$ and Fourier transform $\hat{g} := \mathcal{F}(g)$. Let Δ be the Laplacian on \mathbb{R}^m . Then it holds for any measurable set $B \subset \mathbb{R}^m$ that*

$$\int_B -\Delta |\hat{g}|^2 d\boldsymbol{\xi} \leq 2R^2 \|g\|^2. \quad (3.2.16)$$

By substituting the bounds (3.2.15) into (3.2.14), applying lemma 3.5 to the arising Laplacian integral over $\bigcup_j B_j^-$ and carefully choosing the parameter ε , theorem 3.2 is obtained.

3.2.3.3 Reconstruction from two holograms (inverse problem 3.3)

To prove stability of the reconstruction from two holograms, it is again useful to consider the contrast in Fourier space: for $h = \mu + i\phi$ with $\mu, \phi \in L^2(\mathbb{R}^m, \mathbb{R})$, it holds that

$$\mathcal{F}(\mathcal{I}^{(f_1, f_2)}(h))(\boldsymbol{\xi}) = -2 \underbrace{\begin{pmatrix} \sin(\boldsymbol{\xi}^2/(2f_1)) & \cos(\boldsymbol{\xi}^2/(2f_1)) \\ \sin(\boldsymbol{\xi}^2/(2f_2)) & \cos(\boldsymbol{\xi}^2/(2f_2)) \end{pmatrix}}_{\mathbf{S}_{f_1, f_2}(\boldsymbol{\xi})} \cdot \begin{pmatrix} \mathcal{F}(\phi)(\boldsymbol{\xi}) \\ \mathcal{F}(\mu)(\boldsymbol{\xi}) \end{pmatrix} \quad (3.2.17)$$

for (almost) all $\boldsymbol{\xi} \in \mathbb{R}^m$, where the Fourier transform is meant component-wise. The minimal singular value of the matrix $\mathbf{S}_{f_1, f_2}(\boldsymbol{\xi})$ is $\sigma_{\min}(\boldsymbol{\xi})^2 = 1 - |\cos(\boldsymbol{\xi}^2/(2f_-))| \geq \frac{1}{2} \sin(\boldsymbol{\xi}^2/(2f_-))^2 = \frac{1}{2} s_{0, f_-}(\boldsymbol{\xi})^2$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$, where the phase-CTF s_{0, f_-} to the Fresnel number $\mathbf{f} = \mathbf{f}_- := |\mathbf{f}_1^{-1} - \mathbf{f}_2^{-1}|^{-1}$ has been identified. By elementary linear algebra, this implies that

$$\begin{aligned} |\mathcal{F}(\mathcal{I}^{(f_1, f_2)}(h))(\boldsymbol{\xi})|^2 &\geq 4\sigma_{\min}(\boldsymbol{\xi})^2 (|\mathcal{F}(\phi)(\boldsymbol{\xi})|^2 + |\mathcal{F}(\mu)(\boldsymbol{\xi})|^2) \\ &\geq 2s_{0, f_-}(\boldsymbol{\xi})^2 (|\mathcal{F}(\phi)(\boldsymbol{\xi})|^2 + |\mathcal{F}(\mu)(\boldsymbol{\xi})|^2) \end{aligned} \quad (3.2.18)$$

for almost all $\boldsymbol{\xi} \in \mathbb{R}^m$. Integrating (3.2.18) over \mathbb{R}^m yields the sought L^2 -estimate

$$\|\mathcal{I}^{(f_1, f_2)}(h)\|^2 \geq 2 (\|s_{0, f_-}(\boldsymbol{\xi})^2 \cdot \mathcal{F}(\phi)\|^2 + \|s_{0, f_-}(\boldsymbol{\xi})^2 \cdot \mathcal{F}(\mu)\|^2) = \frac{1}{2} \|\mathcal{S}_0^{(f_-)}(h)\|^2, \quad (3.2.19)$$

which relates stability of inverse problem 3.3 to inverse problem 3.2, leading to theorem 3.3.

3.3 Supplement: stability of phase contrast tomography

A benefit in having proved stability for linearized XPCI independently of the dimension m lies in enabling an extension of the results from 2D-imaging to X-ray phase contrast *tomography*, i.e. to inverse problem 2.2 in the linear regime. This is established in the present supplement.

We consider linearized XPCT data for objects $f \in L^2(\Omega)$ of compact support $\Omega \subset \mathbb{R}^m$:

$$I_{\boldsymbol{\theta}}^{\text{obs}} \approx 1 + T(\mathcal{P}_{\boldsymbol{\theta}}(f)) \quad \text{for } \boldsymbol{\theta} \in \Theta \subset \mathbb{S}^2. \quad (3.3.1)$$

where the image-formation operator $T \in \{\mathcal{I}, \mathcal{S}_\nu\}$ is any of the forward maps from inverse problems 3.1 and 3.2. The stability analysis may be readily extended to multiple holograms, in particular to the setting of inverse problem 3.3, but we omit this case for notational simplicity.

We furthermore widely restrict to the case of full angular sampling where the set of incident directions Θ covers a dense range of tomographic angles of 180 degrees, compare §2.3.3. The reason for this restriction is that stability theory of pure tomographic reconstruction (i.e. no phase contrast, $T = \text{id}$ in (3.3.1)) is simple in this setting, as will be explained in §3.3.1. In principle, however, the presented stability approach is applicable for arbitrary sets of incident directions Θ . According to §2.4.5, fully sampled linear XPCT-data is modeled by

$$I^{\text{obs}}(\boldsymbol{\theta}, \mathbf{x}) \approx 1 + T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f))(\boldsymbol{\theta}, \mathbf{x}) \quad \text{for } \boldsymbol{\theta} \in [0; \pi), \mathbf{x} \in \mathbb{R}^m \quad (3.3.2)$$

where $T^{(\text{lat})}(p)(\boldsymbol{\theta}, \cdot) := T(p(\boldsymbol{\theta}, \cdot))$ applies T along the lateral dimensions and $\mathcal{P}_{\text{full}}(f)(\boldsymbol{\theta}, \mathbf{x}) = \mathcal{P}_{\boldsymbol{\theta}}(f)(\mathbf{x})$ as defined in §2.3.3.

For functions $p \in L^2([0; \pi) \times \mathbb{R}^2)$, let $\|p\| := (\int_0^\pi \int_{\mathbb{R}^2} |p(\boldsymbol{\theta}, \mathbf{x})|^2 d\mathbf{x} d\boldsymbol{\theta})^{1/2}$ denote the standard L^2 -norm. According to §2.3.2, $\mathcal{P}_{\text{full}} : L^2(\Omega) \rightarrow L^2([0; \pi) \times \mathbb{R}^2)$ is well-defined and continuous for any bounded domain $\Omega \subset \mathbb{R}^3$. Moreover, since $T : L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2)$ is continuous, so is $T^{(\text{lat})} : L^2([0; \pi) \times \mathbb{R}^2) \rightarrow L^2([0; \pi) \times \mathbb{R}^2)$ with the same operator-norm $\|T^{(\text{lat})}\| = \|T\|$.

3.3.1 Stability of tomographic reconstruction

By boundedness of T , it follows from (3.3.2) that linearized XPCT cannot be more stable than a reconstruction from pure projection-data $\mathcal{P}_{\text{full}}(f)$ (up to a constant) since $\|T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f))\| \leq \|T\| \|\mathcal{P}_{\text{full}}(f)\|$ for all f . As noted in §2.5.2.4 however, already the inversion of $\mathcal{P}_{\text{full}}$ alone is *ill-posed* in L^2 -norm: for any object-domain $\Omega \subset \mathbb{R}^3$ that contains an open set, there exists *no* constant $C_{\mathcal{P}}(\Omega) > 0$ such that $\|\mathcal{P}_{\text{full}}(f)\| \geq C_{\mathcal{P}}(\Omega) \|f\|$ for all $f \in L^2(\Omega)$. Hence, Lipschitz-stability estimates may neither hold for XPCT.

Yet, by the close relation of $\mathcal{P}_{\text{full}}$ to the 2D Radon transform \mathcal{R} (see §2.3.1), known results for \mathcal{R} (see e.g. [155, § II.5]) imply that stability holds w.r.t. weighted norms in Fourier space:

$$\|\mathcal{P}_{\text{full}}(f)\|_{\mathcal{R}^{-1}} = \|f\|, \quad \|p\|_{\mathcal{R}^{-1}} := \|w_{\mathcal{R}^{-1}} \cdot \mathcal{F}^{(\text{lat})}(p)\|, \quad w_{\mathcal{R}^{-1}}(\theta, \xi_1, \xi_2) := \frac{|\xi_1|^{1/2}}{(2\pi)^{1/2}}. \quad (3.3.3a)$$

$$\|\mathcal{P}_{\text{full}}(f)\| = \|f\|_{\mathcal{R}}, \quad \|f\|_{\mathcal{R}} := \|w_{\mathcal{R}} \cdot \mathcal{F}(f)\|, \quad w_{\mathcal{R}}(\xi_1, \xi_2, \xi_3) := \frac{(2\pi)^{1/2}}{(\xi_1^2 + \xi_3^2)^{1/4}}. \quad (3.3.3b)$$

The identities can be proven via the Fourier-Slice-Theorem 2.1. As the stability estimates (3.3.3a) and (3.3.3b) are *equalities*, they can be thought of as to provide full and lossless characterizations of the information content of tomographic data. However, their interpretations slightly differ: the estimate (3.3.3a) states that data-errors that are bounded not only in L^2 -norm, but even in the *stronger* norm $\|\cdot\|_{\mathcal{R}^{-1}}$ that amplifies high Fourier-frequencies, lead to bounded object-errors in L^2 -norm. On the other hand, (3.3.3b) states that data-errors that are only bounded in L^2 -norm give rise to bounded object errors in the *weaker* norm $\|\cdot\|_{\mathcal{R}}$, which underrepresents high- compared to low frequencies. Accordingly, both estimates express the well-known property that tomographic reconstruction is susceptible in high spatial frequencies but very stable in the low-frequency regime, yet from slightly different perspectives.

Remark 3.6. The identities in (3.3.3) are typically cast to *Sobolev-space estimates* [93, 155]: if $\|f\|_{H^s}^2 := \int_{\mathbb{R}^2} (1 + |\boldsymbol{\xi}|^2)^s |\mathcal{F}(f)(\boldsymbol{\xi})|^2 d\boldsymbol{\xi}$ and $\|f\|_{H_{\text{lat}}^s}^2 := \int_{\mathbb{S}^1} \int_{\mathbb{R}} (1 + \xi^2)^s |\mathcal{F}^{(\text{lat})}(p)(\boldsymbol{\omega}, \xi)|^2 d\xi d\boldsymbol{\omega}$, then for any bounded $\Omega \subset \mathbb{R}^2$ there exist constants $0 < c(\Omega) \leq C(\Omega) < \infty$ such that $c(\Omega) \|f\|_{H^s} \leq \|\mathcal{R}(f)\|_{H_{\text{lat}}^{s+1/2}} \leq C(\Omega) \|f\|_{H^s}$ for all $f \in L^2(\Omega)$ and $s \geq 0$. Analogous bounds can be obtained for $\mathcal{P}_{\text{full}}$. While such estimates are more commonly used to express stability of tomographic reconstruction, the equalities in (3.3.3) are more useful for the argument to be made.

3.3.2 Projection-based stability estimates

By the argument in §3.3.1, the best possible stability results for linearized XPCT are given by analogues of the bounds (3.3.3a) and (3.3.3b), which hold true for pure projection-data. Indeed, an estimate of the second kind is straightforward to obtain by applying the Lipschitz-stability bounds for XPCI for each tomographic incident angle individually:

Theorem 3.7 (Stability estimates for linearized XPCT (2D-based)). *Let $T \in \{\mathcal{T}, \mathcal{S}_\nu\} : L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2)$ and let $C_{\text{IP}^*}^{(2d)}(\Omega^{(2d)})$ be the corresponding stability constant for $\Omega^{(2d)} \subset \mathbb{R}^2$ from theorem 3.1 or 3.2. Furthermore, let $\Omega \subset \mathbb{R}^3$ be bounded and let $\Omega_\theta = \text{supp}(\mathcal{P}_\theta(\mathbf{1}_\Omega))$ be the maximal support of $\mathcal{P}_\theta(f)$ for $f \in L^2(\Omega)$ and $\theta \in [0; \pi)$. Then it holds that*

$$\|T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f))\| \geq \left(\min_{\theta \in [0; \pi)} C_{\text{IP}^*}^{(2d)}(\Omega_\theta) \right) \|f\|_{\mathcal{R}}. \quad (3.3.4)$$

Proof. We have

$$\|T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f))\|^2 = \int_0^\pi \int_{\mathbb{R}^2} |T(\mathcal{P}_\theta(f))(\mathbf{x})|^2 \, d\mathbf{x}d\theta = \int_0^\pi \|T(\mathcal{P}_\theta(f))\|^2 \, d\mathbf{x}d\theta$$

Since $\mathcal{P}_\theta(f) \in L^2(\Omega_\theta)$ for all $\theta \in [0; \pi)$, each of the L^2 -norms in the integral on the r.h.s. can be estimated via the stability-estimate for T from theorem 3.1 or 3.2, $\|T(\mathcal{P}_\theta(f))\| \geq C_{\text{IP}^*}^{(2d)}(\Omega_\theta) \|\mathcal{P}_\theta(f)\|$. By combining this estimate with (3.3.3b), the assertion follows:

$$\begin{aligned} \|T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f))\|^2 &\geq \int_{\mathbb{S}^1} C_{\text{IP}^*}^{(2d)}(\Omega_\theta)^2 \|\mathcal{P}_\theta(f)\|^2 \, d\mathbf{x}d\theta \geq \left(\min_{\theta \in [0; \pi)} C_{\text{IP}^*}^{(2d)}(\Omega_\theta) \right) \|\mathcal{P}_{\text{full}}(f)\|^2 \\ &= \left(\min_{\theta \in [0; \pi)} C_{\text{IP}^*}^{(2d)}(\Omega_\theta) \right)^2 \|f\|_{\mathcal{R}}^2. \quad \square \end{aligned}$$

3.3.3 3D-based stability estimates

In order to obtain an analogue of (3.3.3a), a different approach is pursued. In [119, 176], it is pointed out that the linearized XPCT problem can be related to an XPCI-setting in $m = 3$ dimensions, which has been used to derive *all-at-once* reconstruction methods for XPCT, compare §2.7.4.3. As will be shown in the following, the surprising relation to 3D-XPCI may also be used as a tool to analyze stability of XPCT. The basic mathematical principle is the intertwining property of the parallel-beam projector \mathcal{P}_θ for isotropic Fourier-multipliers:

Lemma 3.8. *Let $\mathbf{m} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{C}$ be a bounded function and let $T : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ with $\mathcal{F}(T(f))(\boldsymbol{\xi}) := \mathbf{m}(|\boldsymbol{\xi}|) \cdot \mathcal{F}(f)(\boldsymbol{\xi})$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$ be the associated isotropic Fourier-multiplier. Then it holds for all $\boldsymbol{\theta} \in \mathbb{S}^2$ and $f \in L^2(\mathbb{R}^3)$ with compact support that*

$$T(\mathcal{P}_\theta(f)) = \mathcal{P}_\theta(T(f)). \quad (3.3.5)$$

Proof. Since $f \in L^2(\mathbb{R}^3)$ is compactly supported, it holds that $\mathcal{P}_\theta(f) \in L^2(\mathbb{R}^2)$ so that the l.h.s. of (3.3.5) is well-defined. By the Fourier-Slice-Theorem 2.1, it furthermore follows that

$$\begin{aligned} \mathcal{F}(T(\mathcal{P}_\theta(f))) (\xi_x, \xi_y) &= \mathbf{m}(|(\xi_x, \xi_y)|) \cdot \mathcal{F}(\mathcal{P}_\theta(f))(\xi_x, \xi_y) \\ &\stackrel{(2.3.9)}{=} (2\pi)^{\frac{1}{2}} \mathbf{m}(|(\xi_1, \xi_2)|) \cdot \mathcal{F}(f)(\xi_x \mathbf{n}_x + \xi_y \mathbf{n}_y) \\ &= (2\pi)^{\frac{1}{2}} \mathbf{m}(|\xi_x \mathbf{n}_x + \xi_y \mathbf{n}_y|) \cdot \mathcal{F}(f)(\xi_x \mathbf{n}_x + \xi_y \mathbf{n}_y) \\ &= (2\pi)^{\frac{1}{2}} \mathcal{F}(T(f))(\xi_x \mathbf{n}_x + \xi_y \mathbf{n}_y) \\ &\stackrel{(2.3.9)}{=} \mathcal{F}(\mathcal{P}_\theta(T(f))) (\xi_x, \xi_y) \quad \text{for all } \xi_x, \xi_y \in \mathbb{R}. \quad (3.3.6) \end{aligned}$$

The claim now follows by inverse Fourier transform. \square

As the linearized forward maps of XPCI $T \in \{\mathcal{T}, \mathcal{S}_\nu\}$ relate to isotropic Fourier-multipliers (CTF-operators), lemma 3.8 may be applied to the XPCT-problem to interchange T and \mathcal{P}_θ in (3.3.1). For the full angular sampling case in (3.3.2), this yields the identity $T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f)) = \mathcal{P}_{\text{full}}(T(f))$. By combining this approach with the stability results from theorems 3.1 and 3.2 for $m = 3$ dimensions, alternative bounds to that in theorem 3.7 are obtained:

Theorem 3.9 (Stability estimates for linearized XPCT (3D-based)). *Let $T \in \{\mathcal{T}, \mathcal{S}_\nu\} : L^2(\mathbb{R}^3) \rightarrow L^2(\mathbb{R}^3)$ and let $C_{\text{IP}^*}^{(3\text{d})}(\Omega)$ be the corresponding stability constant from theorem 3.1 or 3.2 for a bounded domain $\Omega \subset \mathbb{R}^3$. Then it holds for all $f \in L^2(\Omega)$ and $\boldsymbol{\theta} \in \mathbb{S}^2$ that*

$$T(\mathcal{P}_\boldsymbol{\theta}(f)) = \mathcal{P}_\boldsymbol{\theta}(T(f)) \quad (3.3.7)$$

$$T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f)) = \mathcal{P}_{\text{full}}(T(f)) \quad (3.3.8)$$

$$\|T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f))\|_{\mathcal{R}^{-1}} = \|T(f)\| \geq C_{\text{IP}^*}^{(3\text{d})}(\Omega)\|f\|. \quad (3.3.9)$$

Proof. If $T = \mathcal{S}_\nu$, T is an isotropic Fourier-multiplier and so (3.3.7) holds by lemma 3.8. In the case $T = \mathcal{T}$, we may employ the CTF-formulation $\mathcal{T}(f) = \mathcal{C}_0(\text{Re}(f)) + \mathcal{S}_0(\text{Im}(f))$ from (2.4.3) to verify that

$$\begin{aligned} \mathcal{T}(\mathcal{P}_\boldsymbol{\theta}(f)) &= \mathcal{C}_0(\text{Re}(\mathcal{P}_\boldsymbol{\theta}(f))) + \mathcal{S}_0(\text{Im}(\mathcal{P}_\boldsymbol{\theta}(f))) = \mathcal{C}_0(\mathcal{P}_\boldsymbol{\theta}(\text{Re}(f))) + \mathcal{S}_0(\mathcal{P}_\boldsymbol{\theta}(\text{Im}(f))) \\ &= \mathcal{P}_\boldsymbol{\theta}(\mathcal{C}_0(\text{Re}(f))) + \mathcal{P}_\boldsymbol{\theta}(\mathcal{S}_0(\text{Im}(f))) = \mathcal{P}_\boldsymbol{\theta}(\mathcal{T}(f)). \end{aligned} \quad (3.3.10)$$

Here, it has been exploited that $\mathcal{P}_\boldsymbol{\theta}$ acts independently on real- and imaginary parts and lemma 3.8 has been applied for the isotropic Fourier-multipliers $\mathcal{C}_0, \mathcal{S}_0$ from the first to the second line. Accordingly, (3.3.7) is also true for $T = \mathcal{T}$. Thus, $T(\mathcal{P}_\boldsymbol{\theta}(f)) = \mathcal{P}_\boldsymbol{\theta}(T(f))$ holds in general. Moreover, (3.3.7) yields (3.3.8) since $T^{(\text{lat})}(\mathcal{P}_{\text{full}}(f))(\boldsymbol{\theta}, \cdot) = T(\mathcal{P}_{\boldsymbol{\theta}(\boldsymbol{\theta})}(f)) = \mathcal{P}_{\boldsymbol{\theta}(\boldsymbol{\theta})}(T(f)) = \mathcal{P}_{\text{full}}(T(f))(\boldsymbol{\theta}, \cdot)$ for all $\boldsymbol{\theta} \in [0; \pi)$ by definition of $\mathcal{P}_{\text{full}}$, compare §2.3.3.

By combining (3.3.8) with (3.3.3a) and the Lipschitz stability estimates from theorems 3.1 and 3.2, the assertion follows:

$$\|T(\mathcal{P}_{\text{full}}(f))\|_{\mathcal{R}^{-1}} = \|\mathcal{P}_{\text{full}}(T(f))\|_{\mathcal{R}^{-1}} \stackrel{(3.3.3a)}{=} \|T(f)\| \geq C_{\text{IP}^*}^{(3\text{d})}(\Omega)\|f\|. \quad \square$$

3.3.4 Interpretation of the results

It is interesting to note that the obtained stability results for XPCT in theorems 3.7 and 3.9, unlike their pure-projection-data-analogues (3.3.3a) and (3.3.3b), show significant differences:

- The first relation in (3.3.9) is an *equality*, thus yielding an *exact* correspondence between stability of the XPCT-problem and 3D-XPCI. No such relation holds in theorem 3.7.
- The value of the stability constant in theorem 3.9 depends on the size of the 3D-support domain Ω , whereas, in theorem 3.7, the projected supports Ω_θ enter. A 3D-support allows for a tighter description of an object geometry, as illustrated by the following example:

Example 3.10 (3D- vs projected support for a spherical shell). For some $R > r \geq 0$, let $\Omega := \{\boldsymbol{x} \in \mathbb{R}^3 : r \leq |\boldsymbol{x}| \leq R\}$ and let Ω_θ be defined as in theorem 3.7. Then the projected supports $\Omega_\theta = B[0; R] \subset \mathbb{R}^2$ are given by balls of radius R (independently of r) for all $\boldsymbol{\theta} \in [0; \pi)$, whereas the 3D-support Ω becomes more and more sparse as the inner shell-radius r tends to R .

In general, it can be seen that *holes* in the interior of a 3D-support Ω never manifest in the associated projections $\Omega_\theta = \text{supp}(\mathcal{P}_\boldsymbol{\theta}(\mathbf{1}_\Omega))$ for any incident direction $\boldsymbol{\theta} \in \mathbb{S}^2$.

- The XPCI-stability constants C_{IP^*} enter for different dimensions $m \in \{2, 3\}$. Even when 3D- and projected supports are equally tight, as is the case for ball-domains $\Omega = B[0; R]$ ($\Rightarrow \Omega_\theta = B[0; R]$), a slightly larger, i.e. better, value is often obtained for the 3D-constant $C_{\text{IP}^*}^{(3\text{d})}(\Omega)$ than for $C_{\text{IP}^*}^{(2\text{d})}(\Omega_\theta)$. This effect can be identified in the bounds from theorem 3.1 for square- and cube-shaped Ω and will also show in the numerical results of §3.4.3.

According to these observations, we conclude that the 3D-based estimate from theorem 3.9 gives the stronger stability statement. Comparing to the purely tomographic estimate (3.3.3a), the practical meaning of the derived stability bound (3.3.9) can be summarized as follows:

Reconstructions in linearized XPCT are less stable than pure tomographic inversion by no more than a constant factor $C_{\text{IP}^}^{(3\text{d})}(\Omega)$.*

3.4 Supplement: numerical stability computations

As mentioned before, the obtained bound for the stability constant $C_{\text{IP}1}$ of inverse problem 3.1 is found to be quite sharp. On the contrary, the stability result for homogeneous objects (theorem 3.2) only predicts the asymptotic behavior of the stability constant $C_{\text{IP}2}$ with the Fresnel number, $C_{\text{IP}2} \sim \mathfrak{f}^{-\gamma}$ with $\gamma \in \{1, \frac{1}{2}\}$, up to numerical constants for which only rough bounds are obtained. In order to validate stability of image reconstruction in a specific XPCI-setup, this result is not sufficient but quantitative values of $C_{\text{IP}2}$ are needed that are as sharp as possible. The aim of the following supplement is therefore to derive (numerical) strategies that enables accurate computation of the sought stability constant $C_{\text{IP}2}(\Omega, \mathfrak{f}, \nu)$ for a given imaging setting parametrized by $\Omega, \mathfrak{f}, \nu$. Although the presentation here will be limited to inverse problem 3.2 for simplicity, note that the approach also enables stability-computations for inverse problem 3.3 by virtue of theorem 3.3. Moreover, the results presented below may also be adapted to inverse problem 3.1, enabling stability computations beyond the analytical bound from theorem 3.1.

As the required theory underlying to the numerical stability computations is technical in parts, the proofs of the theorems in §3.4.1 and §3.4.2 are postponed to appendix A.

3.4.1 Contrast-minimizing modes

For the computation of the sought stability constant it is a highly relevant and non-trivial question, whether the stability estimate is actually attained by any image, i.e. whether there exists a *contrast-minimizing mode* $\varphi_0 \in L^2(\Omega, \mathbb{R})$ with $\|\varphi_0\| = 1$, satisfying

$$\|\mathcal{S}_\nu(\varphi_0)\| = \inf_{\varphi \in L^2(\Omega, \mathbb{R}), \|\varphi\|=1} \|\mathcal{S}_\nu(\varphi)\| = C_{\text{IP}2}(\Omega, \mathfrak{f}, \nu). \quad (3.4.1)$$

To prove existence of contrast-minimizing modes, we need to rewrite the contrast $\|\mathcal{S}_\nu(\varphi)\|$:

Lemma 3.11. *Let $\Omega \subset \mathbb{R}^m$ be bounded, $\mathfrak{f} > 0$ and $\nu \in \mathbb{R}$. Then it holds that*

$$\|\mathcal{S}_\nu(\varphi)\|^2 = \langle \varphi, 2\varphi - 2\mathcal{K}_{\mathfrak{f}, \nu}(\varphi) \rangle \quad \text{for all } \varphi \in L^2(\Omega), \quad (3.4.2)$$

where $\mathcal{K}_{\mathfrak{f},\nu} : L^2(\Omega) \rightarrow L^2(\Omega)$ is a selfadjoint and compact linear operator defined by

$$\mathcal{K}_{\mathfrak{f},\nu}(\varphi) = \mathcal{F}^{-1}(\hat{k}_{\mathfrak{f},\nu} \cdot \mathcal{F}(\varphi))|_{\Omega} = (k_{\mathfrak{f},\nu} * \varphi)|_{\Omega} = \int_{\Omega} k_{\mathfrak{f},\nu}(\cdot - \mathbf{y})\varphi(\mathbf{y}) \, d\mathbf{y} \quad (3.4.3)$$

$$\hat{k}_{\mathfrak{f},\nu}(\boldsymbol{\xi}) = \cos\left(\frac{\boldsymbol{\xi}^2}{\mathfrak{f}} + 2\nu\right), \quad k_{\mathfrak{f},\nu}(\mathbf{x}) = (\mathfrak{f}/2)^{m/2} \cos\left(\frac{\mathfrak{f}\mathbf{x}^2}{4} - (2\nu + m\pi/4)\right) \quad (3.4.4)$$

Lemma 3.11 implies that the stability constant can be obtained by computing the maximal eigenvalue of the selfadjoint compact operator $\mathcal{K}_{\mathfrak{f},\nu}$ and that contrast-minimizing modes are given by the associated eigenfunctions:

Theorem 3.12 (Existence and characterization of contrast-minimizing modes). *Let $\Omega \subset \mathbb{R}^m$ be bounded, $\mathfrak{f} > 0$ and $\nu \in \mathbb{R}$ and let $\mathcal{K}_{\mathfrak{f},\nu}$ be the operator from lemma 3.11. Then either of the following cases hold true:*

- (1) $\mathcal{K}_{\mathfrak{f},\nu}$ has a positive eigenvalue. Then there exist contrast-minimizing modes $\varphi_0 \in L^2(\Omega, \mathbb{R})$ and these are exactly the eigenfunctions of $\mathcal{K}_{\mathfrak{f},\nu}$ to its maximal eigenvalue $\lambda_{\mathfrak{f},\nu} > 0$, i.e. $\mathcal{K}_{\mathfrak{f},\nu}(\varphi_0) = \lambda_{\mathfrak{f},\nu}\varphi_0$, and

$$C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) = (2 - 2\lambda_{\mathfrak{f},\nu})^{1/2}. \quad (3.4.5)$$

Moreover, there are finitely many linearly independent contrast-minimizing modes and any such mode is infinitely smooth in the interior of Ω .

- (2) $\mathcal{K}_{\mathfrak{f},\nu}$ does not have positive eigenvalues. Then $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) = 2^{1/2}$.

3.4.2 Numerical method

3.4.2.1 General numerical strategy and its limits

Theorem 3.12 implies that stability may be assessed via numerical eigenvalue computations of the operator $\mathcal{K}_{\mathfrak{f},\nu}$. Since $\mathcal{K}_{\mathfrak{f},\nu}$ is compact, it may be uniformly approximated by a suitable sequence of discretizations $\mathbf{K}_{\mathfrak{f},\nu,N} \in \mathbb{R}^{N \times N}$, $\lim_{N \rightarrow \infty} \|E_N \mathbf{K}_{\mathfrak{f},\nu,N} E_N^* - \mathcal{K}_{\mathfrak{f},\nu}\| = 0$ for some isometric embeddings $E_N : \mathbb{R}^N \rightarrow L^2(\Omega)$. Hence, $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) = (2 - 2\lambda_{\mathfrak{f},\nu})^{1/2}$ may be approximated by computing the maximal eigenvalue $\lambda_{\mathfrak{f},\nu,N}$ of the matrix $\mathbf{K}_{\mathfrak{f},\nu,N}$ for sufficiently large N , in which case the associate eigenvectors provide approximations $\varphi_{0,N}$ of a contrast-minimizing mode φ_0 . Since any such φ_0 is smooth according to theorem 3.12, the convergence $\varphi_{0,N} \rightarrow \varphi_0$ and $\lambda_{\mathfrak{f},\nu,N} \rightarrow \lambda_{\mathfrak{f},\nu}$ for $N \rightarrow \infty$ can be expected to be reasonable fast for any standard discretization schemes that approximate functions from coarse to finer and finer scales as N increases.

Discretizations may be obtained by sampling functions φ on points $\{\mathbf{x}_i\}_{i=1}^N \subset \Omega$, setting $\boldsymbol{\varphi} := (\varphi(\mathbf{x}_j))_{j=1}^N$ and $\mathbf{K}_{\mathfrak{f},\nu,N} := (k_{\mathfrak{f},\nu}(\mathbf{x}_i - \mathbf{x}_j)\omega_j)_{i,j=1}^N$ with quadrature-weights ω_j such that

$$(\mathbf{K}_{\mathfrak{f},\nu,N} \boldsymbol{\varphi})_i = \sum_{j=1}^N k_{\mathfrak{f},\nu}(\mathbf{x}_i - \mathbf{x}_j)\omega_j\varphi(\mathbf{x}_j) \approx \int_{\Omega} k_{\mathfrak{f},\nu}(\mathbf{x}_i - \mathbf{y})\varphi(\mathbf{y}) \, d\mathbf{y} = \mathcal{K}_{\mathfrak{f},\nu}(\varphi)(\mathbf{x}_i) \quad (3.4.6)$$

for all sufficiently smooth $\varphi \in L^2(\Omega, \mathbb{R})$ and $i = 1, \dots, N$. However, the convolution-kernel $k_{\mathfrak{f}, \nu}$ is highly oscillatory: for the unit-ball $\Omega = B[0; 1/2]$, the fastest oscillations of $k_{\mathfrak{f}, \nu}(\mathbf{x} - \mathbf{y})$ for $\mathbf{x}, \mathbf{y} \in \Omega$ correspond to a half-period of $2\pi/\mathfrak{f} = 1/\mathfrak{f}$. From sampling theory, we thus expect that the integral operator $\mathcal{K}_{\mathfrak{f}, \nu}$ needs to be sampled at this resolution at least, which translates to a dimension $N \gtrsim \mathfrak{f}^m$ of the discretizing space \mathbb{R}^N . Recalling that the regime of practical relevance for XPCI (and XPCT, see §3.3) is $10^2 \lesssim \mathfrak{f} \lesssim 10^5$ and $m \in \{2, 3\}$, we find that N is thus typically so large that precomputing the matrix $\mathbf{K}_{\mathfrak{f}, \nu, N} \in \mathbb{R}^{N \times N}$ is *computationally prohibitive*. Even *matrix-free* eigenvalue-algorithms, that require only matrix-vector-products $\varphi \mapsto \mathbf{K}_{\mathfrak{f}, \nu, N} \varphi$, which can be implemented in $\mathcal{O}(2N \log(2N))$ operations via FFTs owing to the convolutional structure of $\mathcal{K}_{\mathfrak{f}, \nu, N}$, become infeasible at Fresnel numbers $\mathfrak{f} \gtrsim 10^{9/m}$.

3.4.2.2 Exploiting symmetries

We seek to eliminate this fatal dependence on the dimensionality m by exploiting spatial *symmetries* of the problem. As the CTF-operator \mathcal{S}_ν is *isotropic*, it is invariant under orthogonal transformations of the coordinate system such as rotations or reflections $\mathbf{x} \mapsto -\mathbf{x}$. When also the domain Ω is invariant under one of these operations the operator $\mathcal{K}_{\mathfrak{f}, \nu}$ will therefore commute with the corresponding symmetry-transformation. In particular, this implies that the symmetries will manifest in the contrast-minimizing modes φ_0 in some manner.

Consequently, one may have the following *conjecture* (or hope): for a ball-domain $\Omega = B[0; 1/2]$, where the problem is rotationally invariant, there exists a contrast-minimizing mode that is *radially-symmetric*. As a benefit, finding such a mode – and thus the stability constant C_{IP2} – would reduce to determining its dependence on the radial coordinate, i.e. to finding a *one-dimensional* function. Interestingly, this conjecture turns out to be only *almost* true and proving it is by far more involved than one might expect (see appendix A.2):

Theorem 3.13 (Symmetries of contrast-minimizing modes). *Let $\mathfrak{f} > 0$, $\nu \in \mathbb{R}$ and $\Omega \subset \mathbb{R}^m$ bounded. Then the following holds true:*

- (a) (point-symmetry) *If $\Omega = -\Omega$, then there exists either an even or an odd contrast-minimizing mode, i.e. φ_0 from theorem 3.12 may be chosen such that*

$$\varphi_0(-\cdot) = \varphi_0 \quad \text{or} \quad \varphi_0(-\cdot) = -\varphi_0. \quad (3.4.7)$$

- (b) (radial symmetry) *If $\Omega = B[0; \frac{1}{2}]$, then there exists either a radially symmetric or a radially anti-symmetric contrast-minimizing mode, i.e. φ_0 from theorem 3.12 may be chosen such that, for some function $\varphi_0^{\text{rad}} \in L^2([0; \frac{1}{2}])$, either of the following applies:*

- (1) $\varphi_0(\mathbf{x}) = |\mathbf{x}|^{-\frac{m-1}{2}} \varphi_0^{\text{rad}}(|\mathbf{x}|)$ for all $\mathbf{x} \in \mathbb{R}^m \setminus \{0\}$. Then φ_0^{rad} is an eigenfunction of a selfadjoint compact integral operator $\mathcal{K}_{\mathfrak{f}, \nu}^{\text{rad}, s} : L^2([0; \frac{1}{2}]) \rightarrow L^2([0; \frac{1}{2}])$ defined by

$$\begin{aligned} \mathcal{K}_{\mathfrak{f}, \nu}^{\text{rad}, s}(\varphi_0^{\text{rad}})(r) &= \int_0^{\frac{1}{2}} K_{\mathfrak{f}, \nu}^{\text{rad}, s}(r, \rho) \varphi_0^{\text{rad}}(\rho) \, d\rho \\ K_{\mathfrak{f}, \nu}^{\text{rad}, s}(r, \rho) &= \left(\frac{\mathfrak{f}}{2}\right)^{\frac{m}{2}} (r\rho)^{\frac{m-1}{2}} \cos\left(\frac{\mathfrak{f}(r^2 + \rho^2)}{4} - \gamma\right) \begin{cases} 2 \cos(\frac{1}{2}r\rho) & (m=1) \\ 2\pi J_0(\frac{1}{2}r\rho) & (m=2) \\ 4\pi \text{sinc}(\frac{1}{2}r\rho) & (m=3) \end{cases} \end{aligned} \quad (3.4.8)$$

to its maximal eigenvalue, which is $\lambda_{\mathfrak{f},\nu}$. Then $\tilde{\varphi}_0 : \mathbf{x} \mapsto |\mathbf{x}|^{-\frac{m-1}{2}} \tilde{\varphi}_0^{\text{rad}}(|\mathbf{x}|)$ for any $\tilde{\varphi}_0^{\text{rad}} \in L^2([0; \frac{1}{2}])$ with $\mathcal{K}_{\mathfrak{f},\nu}^{\text{rad},s}(\tilde{\varphi}_0^{\text{rad}}) = \lambda_{\mathfrak{f},\nu} \tilde{\varphi}_0^{\text{rad}}$ is also a contrast-minimizing mode.

(2) $\varphi_0(\mathbf{x}) = (\boldsymbol{\theta}_0 \cdot \frac{\mathbf{x}}{|\mathbf{x}|}) |\mathbf{x}|^{-\frac{m-1}{2}} \varphi_0^{\text{rad}}(|\mathbf{x}|)$ for all $\mathbf{x} \in \mathbb{R}^m \setminus \{0\}$ and some $\boldsymbol{\theta}_0 \in \mathbb{S}^{m-1}$. Then analogous statements to (1) hold for a modified integral operator $\mathcal{K}_{\mathfrak{f},\nu}^{\text{rad},a}$ with kernel

$$K_{\mathfrak{f},\nu}^{\text{rad},a}(r, \rho) = \left(\frac{\mathfrak{f}}{2}\right)^{\frac{m}{2}} (r\rho)^{\frac{m-1}{2}} \sin\left(\frac{\mathfrak{f}(r^2 + \rho^2)}{4} - \gamma\right) \begin{cases} 2 \sin(\frac{\mathfrak{f}}{2} r \rho) & (m=1) \\ 2\pi J_1(\frac{\mathfrak{f}}{2} r \rho) & (m=2) \\ 4\pi \text{sinc}_2(\frac{\mathfrak{f}}{2} r \rho) & (m=3) \end{cases} \quad (3.4.9)$$

Here, the abbreviation $\gamma := 2\nu - m\pi/4$ is used, $J_n : x \mapsto \frac{1}{\pi} \int_0^\pi \cos(nt - x \sin(t)) dt$ denote n th order Bessel functions of the first kind and $\text{sinc}_2 : x \mapsto (\sin(x) - x \cos(x))/x^2$.

Note that the complicated part in the proof of theorem 3.13 is the existence of radially (anti)-symmetric modes. The result is shown in appendix A for general isotropic Fourier-multipliers, beyond the particular case of the CTF-operator \mathcal{S}_ν .

3.4.2.3 Improved numerical method for ball-domains

The gist of theorem 3.13 is that the computation of $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu)$ for the special case of ball-shaped domains $\Omega = B[0; 1/2] \subset \mathbb{R}^m$ can be reduced to eigenvalue problems for the integral operators $\mathcal{K}_{\mathfrak{f},\nu}^{\text{rad},s}, \mathcal{K}_{\mathfrak{f},\nu}^{\text{rad},a}$ on the 1D-interval $[0; 1/2]$ – independently of the true spatial dimension m . Accordingly, we sample the kernels $K_{\mathfrak{f},\nu}^{\text{rad},(\cdot)}$ on equally spaced radii $\{r_i\}_{i=1}^N = \{\rho_j\}_{j=1}^N = \{(1+2j)/(4N) : j = 0, 1, \dots, N-1\} \subset [0; 1/2]$ and compute the maximal eigenvalues of the resulting matrices $\mathbf{K}_{\mathfrak{f},\nu,N}^{(\cdot)} := (K_{\mathfrak{f},\nu}^{\text{rad},(\cdot)}(r_i, \rho_j))_{i,j=1}^N \in \mathbb{R}^{N \times N}$. This corresponds to approximating the integrals in (3.4.8) via the midpoint-rule. By the sampling requirements arising from the oscillatory behavior of the integral kernels, that have been discussed further above, a dimension $N \gtrsim \mathfrak{f}/2$ is required for a good approximation of the continuous problem. We therefore consider choices $N(\sigma) = \lceil \sigma \max(100, \mathfrak{f}/2) \rceil$ with sampling factors $\sigma \gtrsim 1$.

Convergence of the derived numerical algorithm is exemplarily studied for $m \in \{2, 3\}$, $\nu \in \{0, \arctan(0.1)\}$ and $\mathfrak{f} \in \{10^3, 10^4\}$: approximations $C_{\text{IP}_2}^{(N)}(\Omega, \mathfrak{f}, \nu) = (2 - 2\lambda_{\mathfrak{f},\nu,N})^{1/2}$ are computed with $N(\sigma)$ according to sampling factors $\sigma = 2^{-1}, 2^{-0.5}, \dots, 2^4$. As the true stability constant $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu)$ is unknown, we compute the relative deviations from the (supposedly highly accurate) result for $\sigma = 32$, $\epsilon(\sigma) := |C_{\text{IP}_2}^{(N(\sigma))}(\Omega, \mathfrak{f}, \nu) - C_{\text{IP}_2}^{(N(32))}(\Omega, \mathfrak{f}, \nu)| / |C_{\text{IP}_2}^{(N(32))}(\Omega, \mathfrak{f}, \nu)|$. The results are plotted in fig. 3.3. The plots indicate highly uniform, algebraic convergence, $\epsilon(\sigma) \sim 1/\sigma^2$, to the true stability constant $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu)$. Notably, the approximation is typically already as good as $\epsilon(\sigma) \lesssim 1\%$ for σ between 1 and 2, i.e. practically without oversampling. Recalling that this translates to matrices $\mathbf{K}_{\mathfrak{f},\nu,N}^{(\cdot)}$ of size $N(\sigma) \times N(\sigma)$ with $N(\sigma) = \lceil \sigma \mathfrak{f} / (4\pi) \rceil$ for large \mathfrak{f} , the convergence results thus imply that accurate stability computations are feasible for modified Fresnel numbers $\mathfrak{f} \leq 10^6$ on current workstations.

Contrary to the general setting, see §3.4.2.1, this limit may however *not* be easily overcome by matrix-free algorithms as the convolutional-structure of the original operator $\mathcal{K}_{\mathfrak{f},\nu}$ does not carry over to the radial ones $\mathcal{K}_{\mathfrak{f},\nu}^{\text{rad},s}$ and $\mathcal{K}_{\mathfrak{f},\nu}^{\text{rad},a}$, as can be seen from (3.4.8) and (3.4.9).

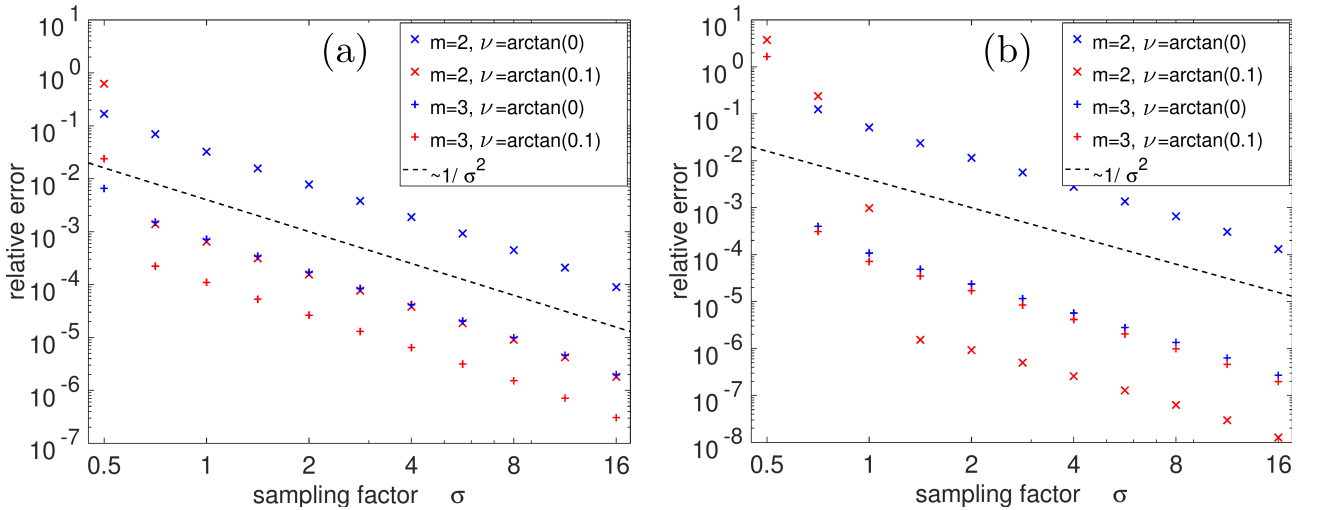


Figure 3.3. Convergence experiments for numerical stability computations for ball-domains $\Omega = B[0; 1/2]$. Plotted are the relative deviations $\epsilon(\sigma)$ of the approximated stability constants $C_{\text{IP}_2}^{(N(\sigma))}(\Omega, \mathbf{f}, \nu)$ at various sampling factors σ from the value obtained at high oversampling, $\sigma = 32$, for different values of m and ν at Fresnel numbers (a) $\mathbf{f} = 10^3$ and (b) $\mathbf{f} = 10^4$, respectively.

3.4.3 Numerical stability results

Stability computations for ball-domains $\Omega = B[0; 1/2]$ are carried out for the whole parameter-space of interest: for dimensions $m = 1, 2, 3$, Fresnel numbers $\mathbf{f} \in [1; 5 \cdot 10^5]$ and $\nu \in \arctan(c_{\beta/\delta})$ corresponding to fixed ratios $\mu/\phi = \beta/\delta = c_{\beta/\delta} \in \{0, 1\%, 2\%, 5\%, 10\%\}$ between absorption and phase contrast, compare §2.4.3. The results are obtained for a sampling factor $\sigma = 2$. Additionally, the computations are repeated for $\sigma = 1$, which allows to estimate the error as $|C_{\text{IP}_2}^{(N(2))}(\Omega, \mathbf{f}, \nu) - C_{\text{IP}_2}(\Omega, \mathbf{f}, \nu)| \approx \frac{1}{3}|C_{\text{IP}_2}^{(N(2))}(\Omega, \mathbf{f}, \nu) - C_{\text{IP}_2}^{(N(1))}(\Omega, \mathbf{f}, \nu)|$ according to the observed convergence rates in §3.4.2.3. The maximum estimated error is $\approx 1.4\%$, where the deviation is much smaller for most of the investigated parameter-triples (\mathbf{f}, ν, m) .

Let us first discuss the contrast-minimizing modes, obtained as a by-product of the computations. It is found that these come in only two different classes, which are exemplarily represented by the results for $m = 2$, $\mathbf{f} = 10^4$ and $\nu \in \{0, \arctan(0.1)\}$ shown in fig. 3.4(e),(f): for $\nu = 0$ (pure phase contrast) the obtained modes are always unimodal low-frequency “blobs” as depicted in fig. 3.4(e) and the same holds true for $\nu > 0$ (non-vanishing absorption) as long as \mathbf{f} is sufficiently small. For $\nu > 0$ and large \mathbf{f} on the other hand, the computed modes become oscillatory as seen in fig. 3.4(f). This is to be expected as the contrast-minimizing modes should concentrate at the least stable parts in Fourier space, which, for $\nu > 0$, is not necessarily around zero-frequency $\boldsymbol{\xi} = 0$ but near the first-order CTF-zero, as seen from the analysis in §3.2.3.2.

The obtained approximations for $C_{\text{IP}_2}(\Omega, \mathbf{f}, \nu)$ are plotted in fig. 3.4(a)–(d) as solid curves. For comparison, also the analytical bounds from theorem 3.2 are shown (dashed curves). *Qualitatively*, the numerically determined dependence $\mathbf{f} \mapsto C_{\text{IP}_2}(\Omega, \mathbf{f}, \nu)$, in particular the asymptotic decay rates $C_{\text{IP}_2}(\Omega, \mathbf{f}, \nu) \sim \mathbf{f}^{-\gamma}$ for $\mathbf{f} \rightarrow \infty$, turns out as expected from the analytical bound in theorem 3.1. Quantitatively, however, some interesting observations can be made in addition:

- The numerical values for $C_{\text{IP}_2}(\Omega, \mathbf{f}, \nu)$ always exceed the analytical bound by factors ~ 10 . This shows that numerical (re-)computations are actually worth the effort.

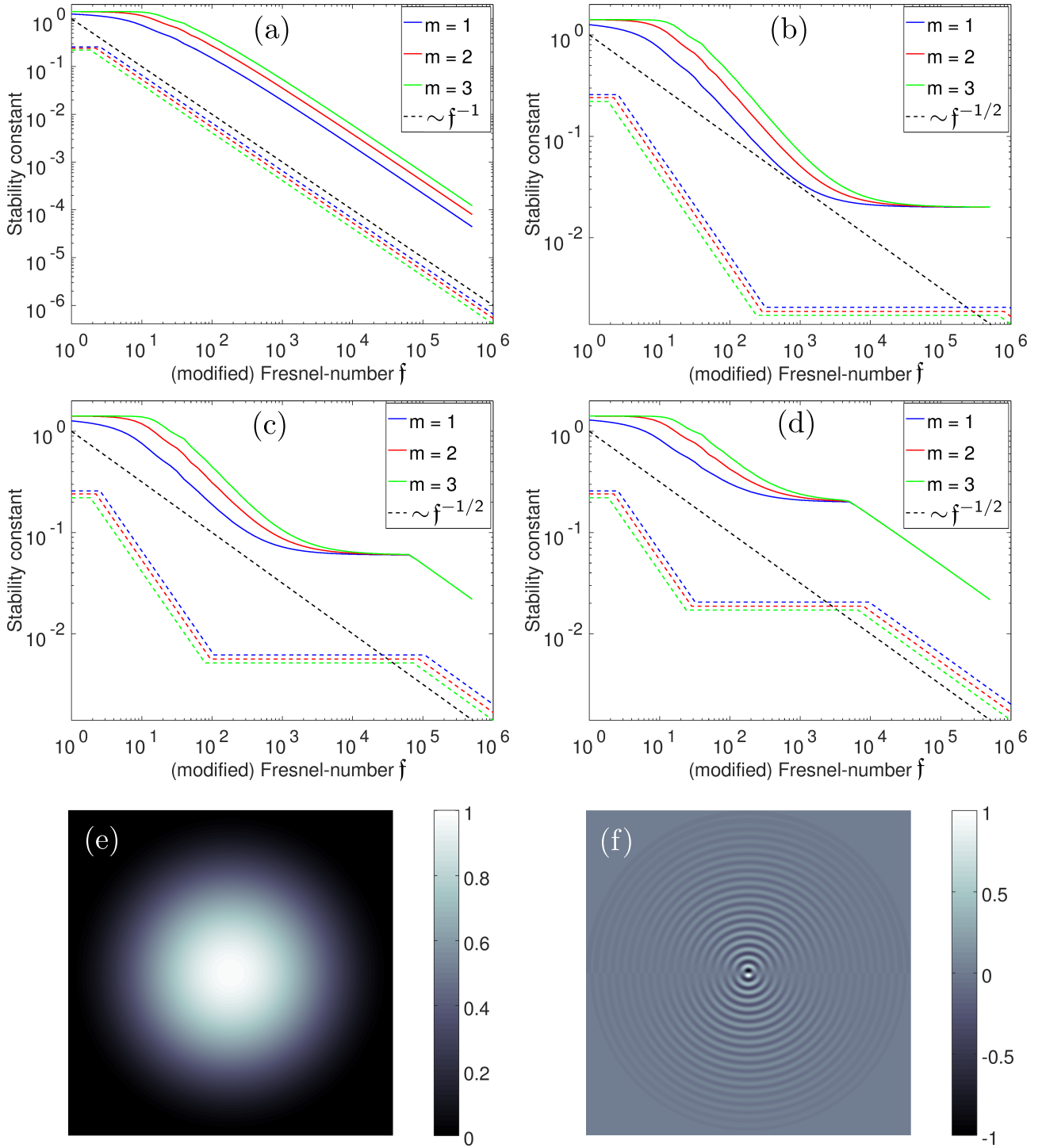


Figure 3.4. Numerical results for the stability constant $C_{IP2}(\Omega, f, \nu)$ for a ball-shaped support-domain $\Omega = B[0; 1/2]$. Subfigures (a)–(d) show the results for $\nu = \arctan(c_{\beta/\delta})$ corresponding to $c_{\beta/\delta} = 0\%$, 1% , 3% , 10% absorption, respectively. The solid lines plot the obtained numerical values for $C_{IP2}(\Omega, f, \nu)$ sampled at $f = 10^0, 10^{0.1}, 10^{0.2}, \dots, 10^5$. For comparison, the dashed lines of the same color plot the analytical bound from theorem 3.2. Subfigures (e) and (f) exemplarily show a contrast-minimizing mode for the setting $f = 10^4$, $m = 2$ for 0% and 10% absorption, respectively.

- For small \mathfrak{f} , the numerical results stabilize at some value $1 < C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu) < 1.5$ for all ν and m . By additional analysis, one may indeed show that is in agreement with theory, predicting that $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu) \rightarrow 2^{1/2}$ should hold for $\mathfrak{f} \rightarrow 0$.
- Dependence on m : the stability constants $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu)$ are found to be monotonically increasing with m , i.e. the higher the dimension the more stability. Yet, the quantitative differences between the values for $m = 1, 2, 3$ do not exceed a factor of ≈ 3 .
- For $\nu = 0$, i.e. pure phase contrast, there is a smooth transition from $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu) = \mathcal{O}(1)$ to decay rates $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu) = c_{1,m}\mathfrak{f}^{-1}$ for $\mathfrak{f} \gtrsim 10^2$, where $c_{1,1} \approx 22$, $c_{1,2} \approx 40$ and $c_{1,3} \approx 61$.
- For $\nu > 0$, i.e. non-vanishing absorption, the following modifications occur:
 - $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu)$ increases with ν , i.e. the more absorption the greater the stability.
 - For small \mathfrak{f} , the functional dependence $\mathfrak{f} \mapsto C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu)$ is similar as for $\nu = 0$.
 - For larger \mathfrak{f} , the stability constants tend to a plateau, whose value $2 \sin(\nu) = 2|s_\nu(0)|$ can be identified as the CTF-contrast at Fourier-frequency $\boldsymbol{\xi} = 0$.
 - For even larger \mathfrak{f} , the plateau ends with a “kink” beyond which the curves follow predicted asymptotic scaling $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu) \approx c_4\mathfrak{f}^{-1/2}$, where the value of the constant $c_4 \approx 15$ is notably quasi independent of m and ν . The kink coincides with the transition from unimodal- to oscillatory contrast-minimizing modes, see fig. 3.4(e),(f).

Although these observations are by far not exhaustive, they demonstrate how numerical stability computations may supplement the analytical predictions quantitatively. For clarity, let us emphasize the scope of the derived numerical stability methods:

Numerical Result 3.14 (stability computations). The stability constant $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu)$ from theorem 3.2 (and analogously C_{IP_1} and C_{IP_3}) may be approximated to high accuracy by numerical eigenvalue computations, which also provide *contrast-minimizing modes*, i.e. image-components that are least stable to reconstruct. Numerical methods cover the following regimes:

- (1) Arbitrary domains $\Omega \subset \mathbb{R}^m$ (of diameter 1) for Fresnel numbers $\mathfrak{f} \lesssim 10^{9/m}$, via §3.4.2.1.
- (2) Ball-domains $\Omega = B[\mathbf{a}; 1/2] \subset \mathbb{R}^m$ for Fresnel numbers $\mathfrak{f} \lesssim 10^6$, via §3.4.2.3.
- (3) Ball-domains $\Omega = B[\mathbf{a}; 1/2] \subset \mathbb{R}^m$ at *asymptotic* Fresnel numbers $\mathfrak{f} \rightarrow \infty$, via the fitted decay rates $C_{\text{IP}_2}(\Omega, \mathfrak{f}, 0) \sim c_{1,m}\mathfrak{f}^{-1}$ ($c_{1,m}$ as above) and $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu) \sim 15\mathfrak{f}^{-1/2}$ for $\nu > 0$.

It should be noted that $C_{\text{IP}_2}(\Omega, \mathfrak{f}, \nu)$ may usually be computed *within seconds* in setting (2) if $\mathfrak{f} \leq 10^5$. Furthermore, let us recall that the stability constants for geometrically complicated domains Ω may be approximated by such for ball-shaped supports, as discussed in §3.2.2.4.

3.5 Supplement: stability by non-negativity constraints

In the previous sections, stability has been analyzed under the assumption of support constraints on the unknown image. However, other types of a priori knowledge have been found to improve

image reconstruction in practical XPCI. In particular, *non-negativity* of phase- and absorption-image ϕ, μ are valuable constraints that are justified by the fundamental physics of X-ray interaction (§2.1.2), i.e. that “come for free”. In this supplement, we show that the obtained stability results also shed a light on the stabilizing effect of non-negativity constraints.

The analysis of non-negativity constraints is generally more difficult than that of support constraints due to the nonlinear nature of the former, pointed out in §2.4.6. To gain some insights on the effects, we start from an abstract variational reconstruction method for real-valued functions, $\phi^{\text{recon}} \in \operatorname{argmin}_{\phi \in L^2(\Omega, \mathbb{R})} J(\phi)$, with optional support constraint $\Omega \subseteq \mathbb{R}^m$ in analogy to the CTF-schemes applied in §3.1.1. A reconstruction $\phi_{\geq 0}^{\text{recon}}$ with additional non-negativity constraint corresponds to restricting the minimization to functions that satisfy $\phi \geq 0$ (almost everywhere). Under suitable assumptions on the functional J (convexity, coercivity and differentiability), the minimizers with and without non-negativity $\phi_{(\geq 0)}^{\text{recon}}$ are uniquely determined by the first-order optimality conditions, see e.g. [195, §2.5.1]: if we define $P_{\geq 0} : \phi \mapsto \max\{\phi, 0\}$ (pointwise max) and $P_A : \phi \mapsto \phi|_A$ (restriction to a measurable set $A \subset \mathbb{R}^m$), it holds that

$$\phi^{\text{recon}} \in \operatorname{argmin}_{\phi \in L^2(\Omega, \mathbb{R})} J(\phi) \Leftrightarrow \phi^{\text{recon}} - P_{\Omega}(\phi^{\text{recon}} - J'[\phi^{\text{recon}}]) = 0 \quad (3.5.1a)$$

$$\begin{aligned} \phi_{\geq 0}^{\text{recon}} \in \operatorname{argmin}_{\phi \in L^2(\Omega, \mathbb{R}), \phi \geq 0} J(\phi) &\Leftrightarrow \phi_{\geq 0}^{\text{recon}} - P_{\geq 0} P_{\Omega}(\phi_{\geq 0}^{\text{recon}} - J'[\phi_{\geq 0}^{\text{recon}}]) = 0 \\ &\Leftrightarrow \phi_{\geq 0}^{\text{recon}} - P_{\Omega_{\geq 0}}(\phi_{\geq 0}^{\text{recon}} - J'[\phi_{\geq 0}^{\text{recon}}]) = 0 \stackrel{(3.5.1a)}{\Leftrightarrow} \phi_{\geq 0}^{\text{recon}} \in \operatorname{argmin}_{\phi \in L^2(\Omega_{\geq 0}, \mathbb{R})} J(\phi) \end{aligned} \quad (3.5.1b)$$

with $\Omega_{\geq 0} := \{\mathbf{x} \in \Omega : \phi_{\geq 0}^{\text{recon}} - J'[\phi_{\geq 0}^{\text{recon}}](\mathbf{x}) \geq 0\}$. Importantly, the optimization on the r.h.s. of (3.5.1b) no longer *explicitly* imposes non-negativity but only asserts a (tightened) support constraint $\operatorname{supp}(\phi_{\geq 0}^{\text{recon}}) \subset \Omega_{\geq 0} \subset \Omega$. Hence, the equivalences in (3.5.1b) show the following:

Non-negativity constraints effectively act as automatic (tightened) support constraints.

The word “automatic” emphasizes that the effective support domain $\Omega_{\geq 0}$ is not manually set as in the case of normal support constraints, but is implicitly defined by details of J and typically hard to predict a priori*. In particular, it is not even clear whether a non-negativity constraint will be *active* at all: it may certainly happen that $\Omega_{\geq 0} = \Omega$ in (3.5.1b), in which case the constraint is without effect on the computed reconstruction. When active, however, the impact of imposing non-negativity may be readily understood in terms of the associated effective support constraint, according to (3.5.1b).

For the considered inverse problems of XPCI (and XPCT), the strong stabilizing effect of support-knowledge found in the preceding sections thus suggests similar benefits from exploiting the natural non-negativity of sought images ϕ, μ (or δ, β). We conclude with some remarks:

- *Complex-valued images:* While real-valued ϕ have been assumed above for simplicity, the statements generalize to settings where complex-valued functions $h \in L^2(\Omega)$ are reconstructed with non-negativity constraints in the real- and/or imaginary part. Yet, different effective supports $\Omega_{\geq 0}^{\text{Re}}, \Omega_{\geq 0}^{\text{Im}} \subset \Omega$ then arise for $\operatorname{Re}(h)$ and $\operatorname{Im}(h)$ in general.
- *General box constraints:* The results for non-negativity may be furthermore generalized to analyze the effect of arbitrary *box constraints* of the form $\phi_{\min} \leq \phi^{\text{recon}} \leq \phi_{\max}$.

*Incidentally, the idea of a certain class of reconstruction algorithms, so-called *active set methods* (see e.g. [94]), exactly boils down to imposing non-negativity by iteratively determining the effective support $\Omega_{\geq 0}$.

- *2D- versus 3D-non-negativity in XPCT*: In §3.3.4, it has been argued that 3D-support constraints provide a tighter description of the object-geometry in XPCT than confining the support its 2D-projections and therefore enable improved stability. Analogously, imposing non-negativity of a 3D-object is a strictly stronger (and thus more valuable) constraint than non-negativity of its 2D-projections: $f \geq 0$ implies that $\mathcal{P}_\theta(f) \geq 0$ for all incident directions $\theta \in \mathbb{S}^{m-1}$, yet the converse is *not* true.

3.6 Supplement: the completely holographic regime – a doubly sweet spot of XPCI

A remarkable common feature of all the derived stability results is that they *improve* as the (modified) Fresnel number \mathfrak{f} decreases, which corresponds to more and more *holographic* measurements as discussed §2.2.1.4. This makes the regime of low Fresnel numbers $\mathfrak{f} \lesssim 10$ highly interesting for applications. In this regime, all object-lengthscales up to the coarsest ones are holographic, i.e. substantially distorted by Fresnel diffraction, and therefore yield significant phase contrast. This is the physical explanation for the observed high stability of XPCI in this *completely holographic regime*, which is further investigated in the present supplement.

3.6.1 Asymptotic linearity of XPCI for low Fresnel numbers

Despite the apparent benefits of measuring at low Fresnel numbers, we recall that the considerations have so far been restricted to the *linearized* imaging model, while the behavior of nonlinear contributions has been neglected. In the following, it is shown that indeed also the nonlinear terms in the XPCI-forward model behave in a surprisingly favorable manner in the limit of small Fresnel numbers $\mathfrak{f} \rightarrow 0$. The principal observation is the following:

Theorem 3.15 (Band-limitation of the quadratic nonlinearity in XPCI). *Let $h \in L^1(\mathbb{R}^m) \cap L^2(\mathbb{R}^m)$. Then it holds that*

$$\mathcal{F}(|\mathcal{D}(h)|^2)(\boldsymbol{\xi}) = (2\pi)^{-\frac{m}{2}} ((n_{\mathfrak{f}} \cdot h) \star (n_{\mathfrak{f}} \cdot h))(-\boldsymbol{\xi}/\mathfrak{f}) \quad \text{for all } \boldsymbol{\xi} \in \mathbb{R}^m, \quad (3.6.1)$$

where “ \star ” denotes the correlation-operator, $f \star g(\mathbf{x}) := \int_{\mathbb{R}^m} f(\mathbf{x} + \mathbf{y}) \overline{g(\mathbf{y})} d\mathbf{y}$. In particular, if $h \in L^2(\Omega)$ is supported in some bounded domain $\Omega \subset \mathbb{R}^m$, then $|\mathcal{D}(h)|^2$ is band-limited with

$$\mathcal{F}(|\mathcal{D}(h)|^2)|_{(\mathfrak{f}(\Omega - \Omega))^c} = 0. \quad (3.6.2)$$

Proof. By the alternate form of the Fresnel propagator in (2.2.10), it holds that

$$\begin{aligned} |\mathcal{D}(h)|^2(\mathbf{x}) &= \mathfrak{f}^m |\mathcal{F}(n_{\mathfrak{f}} \cdot h)(\mathfrak{f} \cdot \mathbf{x})|^2 = (2\pi)^{-\frac{m}{2}} \mathfrak{f}^m \mathcal{F}((n_{\mathfrak{f}} \cdot h) \star (n_{\mathfrak{f}} \cdot h))(\mathfrak{f} \cdot \mathbf{x}) \\ &= (2\pi)^{-\frac{m}{2}} \mathcal{F}(((n_{\mathfrak{f}} \cdot h) \star (n_{\mathfrak{f}} \cdot h))((\cdot)/\mathfrak{f}))(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^m. \end{aligned} \quad (3.6.3)$$

Here, the convolution-theorem has been applied to rewrite the squared modulus of the Fourier transform in terms of an (auto-)correlation and the known behavior of the FT under coordinate-dilations has been exploited. Using the general identity $\mathcal{F}(f) = \mathcal{F}(f(-\cdot))$, we obtain

$$\mathcal{F}(|\mathcal{D}(h)|^2)(\boldsymbol{\xi}) = (2\pi)^{-\frac{m}{2}} ((n_{\mathfrak{f}} \cdot h) \star (n_{\mathfrak{f}} \cdot h))(-\boldsymbol{\xi}/\mathfrak{f}) \quad \text{for all } \boldsymbol{\xi} \in \mathbb{R}^m. \quad (3.6.4)$$

If the support of h is contained in Ω , then also $\text{supp}(n_{\mathfrak{f}} \cdot h) \subset \Omega$ and thus $\text{supp}((n_{\mathfrak{f}} \cdot h) \star (n_{\mathfrak{f}} \cdot h)) \subset \Omega - \Omega$ by standard results on the support of correlations. This implies the assertion. \square

According to theorem 3.15, the effects of the quadratic nonlinearity in XPCI are restricted to Fourier-components from within the sub-domain $\mathfrak{f} \cdot (\Omega - \Omega)$ of Fourier space. The restriction is typically to *low frequencies* as illustrated for the important special case of ball supports:

Example 3.16. Let $\Omega := B[\mathbf{a}, \frac{1}{2}]$ be given by a ball of diameter one around some $\mathbf{a} \in \mathbb{R}^m$. Then it holds that $\mathfrak{f} \cdot (\Omega - \Omega) = B[0, \mathfrak{f}]$ and hence, by theorem 3.15,

$$\mathcal{F}(|\mathcal{D}(h)|^2)|_{B[0, \mathfrak{f}]^c} = 0 \quad \text{for all } h \in L^2(\Omega). \quad (3.6.5)$$

Now let us consider the fully *nonlinear* XPCI-model $\mathcal{N}(h) = |\mathcal{D}(\exp(-h))|^2 - 1$ from §2.4.1. As used already in §2.4.7.2, the operator may be cast to the form

$$\mathcal{N}(h) = \mathcal{F}(o) + |\mathcal{D}(o)|^2 \quad \text{with } o := 1 - \exp(-h). \quad (3.6.6)$$

Hence, the *nonlinear* contrast $I - 1 = \mathcal{N}(h)$ in the hologram-data is linear w.r.t. the *incremental object transmission function* (iOTF) o , up to the remaining quadratic term $|\mathcal{D}(o)|^2$. Theorem 3.15 and example 3.16 imply that the latter may be eliminated by restricting to high frequencies, so that the remaining contrast is *completely* linear in the iOTF:

Corollary 3.17 (Linearity of XPCI-data in high Fourier-frequencies). *Let $h \in L^2(\Omega)$ for bounded $\Omega \subset \mathbb{R}^m$ with $\text{Re}(h(\mathbf{x})) \geq 0$ for almost all $\mathbf{x} \in \mathbb{R}^m$. Let \mathcal{N} be the nonlinear forward operator of XPCI. Then it holds that*

$$\mathcal{F}(\mathcal{N}(h))|_{(\mathfrak{f} \cdot (\Omega - \Omega))^c} = \mathcal{F}(\mathcal{F}(o))|_{(\mathfrak{f} \cdot (\Omega - \Omega))^c} \quad \text{with } o := 1 - \exp(-h) \quad (3.6.7)$$

i.e. nonlinear XPCI-data is linear in the iOTF o for all spatial frequencies outside the bounded domain $\mathfrak{f} \cdot (\Omega - \Omega)$. Moreover, o is uniquely determined by the data (3.6.7).

Proof. The assumptions $h \in L^2(\Omega)$ and $\text{Re}(h) \geq 0$ ensure that $o \in L^2(\Omega) \subset L^1(\mathbb{R}^m) \cap L^2(\mathbb{R}^m)$, as shown in §2.4.7.2. Hence, theorem 3.15 may be applied to $|\mathcal{D}(o)|^2$. By (3.6.6), this yields

$$\mathcal{F}(\mathcal{N}(h))|_{(\mathfrak{f} \cdot (\Omega - \Omega))^c} = \mathcal{F}(\mathcal{F}(o))|_{(\mathfrak{f} \cdot (\Omega - \Omega))^c} + \mathcal{F}(|\mathcal{D}(o)|^2)|_{(\mathfrak{f} \cdot (\Omega - \Omega))^c} = \mathcal{F}(\mathcal{F}(o))|_{(\mathfrak{f} \cdot (\Omega - \Omega))^c} \quad (3.6.8)$$

The uniqueness statement is a consequence of the principal uniqueness theorem 2.10. \square

The observation of linearity at high frequencies is not new but has already been used in a uniqueness proof for nonlinear XPCI, see for example [108, Theorem 3.3]. What has seemingly not been realized so far, however, is the strong dependence of the frequency cut-off on \mathfrak{f} and its implications on practical imaging. This is discussed in the following.

Corollary 3.17 shows that the nonlinearity of XPCI may theoretically be by-passed, up to a remaining reconstruction of h from $o = 1 - \exp(-h)$, which is a simple pointwise operation except for possible *phase-wrapping* ambiguities, see §2.5.1.1. This sounds too good to be true

and indeed there is a catch: for typical Fresnel numbers $f \gtrsim 10^3$ in experiments, the frequency-domain $f \cdot (\Omega - \Omega)$, in which nonlinearity is still present, is typically huge. Indeed, the cut-off frequency $\sim f$ often exceeds the maximal spatial frequency associated with the finite resolution of the acquired holograms. In this case corollary 3.17 is thus irrelevant for practical imaging.

Notably however, the situation changes dramatically for low Fresnel numbers $f \lesssim 10$, i.e. in the completely holographic regime, where the quadratic nonlinearity of XPCI only affects the coarsest relevant spatial frequencies. Most prominently, corollary 3.17 implies that the nonlinearity becomes *completely trivial* in the limit $f \rightarrow 0$. Accordingly, imaging at low Fresnel numbers is not only desirable from the point-of-view of stability for the linearized problem, but also because it decreases the nonlinearity of the image reconstruction problem.

3.6.2 Practical realization of completely holographic XPCI

The completely holographic setting, where the XPCI-data is holographic even for the coarsest object-scale, given by the sample-diameter, is not accessed by typical imaging setups. Accordingly, one may ask whether it is possible to do experiments in this regime at all with a typical synchrotron XPCI-setup like the GINIX from §1.2. The answer is yes but the path is slightly counter-intuitive: although small Fresnel numbers are typically associated with large geometrical magnifications, i.e. small source-to-sample-distances $d_{0,1}$, one has to make $d_{0,1}$ *large* to obtain completely holographic measurements, as will be seen in the following.

By the Fresnel-Scaling-Theorem (see §2.1.5.1), the modified Fresnel number associated with a *fixed* physical length b (e.g. the sample-diameter) in a divergent-beam setup is given by

$$f_{b,\text{eff}} = \frac{2\pi b^2}{\lambda d_{\text{eff}}} = M \cdot \frac{2\pi b^2}{\lambda d_{1,2}} \stackrel{d_{1,2} \approx d_{0,2}}{\approx} M \cdot \frac{2\pi b^2}{\lambda d_{0,2}} \quad (3.6.9)$$

where λ is the wavelength, $d_{0,2}$ denotes the source-to-detector-distance, $d_{1,2} = d_{0,2} - d_{0,1}$ the sample-to-detector-distance and $M = d_{0,2}/d_{0,1}$ the geometrical magnification. In the GINIX-setup, the distance $d_{0,2}$ is fixed and $d_{1,2} \approx d_{0,2}$ holds true, as assumed in (3.6.9), because the required magnifications $M \gg 1$ necessitate that $d_{0,1} \ll d_{0,2}$. According to (3.6.9), the Fresnel number $f_{b,\text{eff}}$ is thus essentially proportional to M . Hence, completely holographic XPCI-data may be acquired by imaging at relatively low magnification, as illustrated by example 3.18:

Example 3.18 (Fresnel number at large and small magnifications). Consider a sample of diameter $b = 10 \mu\text{m}$ imaged in the GINIX-setup (see §1.2) at source-to-detector-distance $d_{0,2} \approx 5 \text{ m}$ and wavelength $\lambda \approx 0.157 \text{ nm}$. Then it holds that

$$f_{b,\text{eff}} \approx 401 \quad \text{for} \quad d_{0,1} = 0.01 \text{ m} \quad (\text{high magnification } M = 500) \quad (3.6.10a)$$

$$f_{b,\text{eff}} \approx 20.8 \quad \text{for} \quad d_{0,1} = 0.2 \text{ m} \quad (\text{low magnification } M = 25) \quad (3.6.10b)$$

We emphasize that XPCI at the lower magnification $M = 25$ in example 3.18 is *extremely stable* – even if phase- and absorption are to be reconstructed independently from a single hologram, i.e. for the setting of inverse problem 3.1: using the numerical methods from §3.4, we find that a Fresnel number of $f = f_{b,\text{eff}} \approx 20.8$ associated with a 2D ball-support $\Omega = B[0; 1/2]$ corresponds to a stability-constant $C_{\text{IP1}}(\Omega, f) \approx 0.45$, which means that even the least stable

mode induces strong contrast in the data. Moreover, note that the wavelength λ and sample-diameter b in example 3.18 are similar as in the real-world XPCI-settings from fig. 1.3(a),(b). This shows that the completely holographic regime is indeed experimentally accessible.

Yet, we also see that measuring in this regime comes at a price: the object has to be moved far away from the source spot so that the magnification M is relatively low. This limits the achievable *resolution*: features of size smaller than $\Delta x_{\text{pix}}/M$, where Δx_{pix} is the physical detector-pixel size, cannot be resolved. The GINIX-data considered in this work has been acquired by a detector of pixel-size 6.5 micrometers so that the resolution at 25-times magnification would be limited to 260 nanometers, which is rather coarse compared to the total sample-size of 10 micrometers assumed in example 3.18. This limitation could be circumvented by using detectors with smaller pixels in principle, but this poses severe technological challenges.

3.7 Conclusions

To conclude this chapter, we discuss the findings of the presented article and corresponding supplements from a broader perspective, outlining conclusions for the design of XPCI-experiments and practical image reconstruction algorithms.

Significance of (support) constraints: The derived theorems reveal a significant beneficial effect of support constraints to image reconstruction in XPCI, in accordance with numerical reconstruction results, see [14, 13] and §3.1.1: inverse problems 3.1 to 3.3, that are ill-posed or even non-unique in the unconstrained case, become *well-posed* when the support of the image is known *a priori* to lie within a bounded domain $\Omega \subset \mathbb{R}^m$. Moreover, the tighter the constraints, i.e. the smaller the support-domain Ω (in terms of the associated Fresnel number f), the more stable image reconstruction becomes to errors in the measured hologram(s). The analysis also predicts a stabilizing effect of imposing *non-negativity* of the recovered phase- and absorption images ϕ, μ , as such constraints may be interpreted in terms of effective supports, see §3.5.

In order to take advantage of the theoretical predictions on stability, image reconstruction algorithms need to be capable of imposing support- and non-negativity constraints. This constitutes a major motivation for the numerical methods proposed in chapter 5.

3D constraints and tomographic consistency in XPCT: The extension of the results to phase contrast *tomography* in §3.3 revealed that stability of image reconstruction in XPCT depends on the *3D*-support $\Omega^{(3d)} \subset \mathbb{R}^3$ of the imaged object. As discussed in §3.3.4, this enables improved stability compared to the associated 2D-XPCI reconstruction problems of retrieving the projected phase-images (and absorption) from the holograms for each tomographic angle individually (owing to potentially sparser supports in 3D than in 2D (example 3.10) and generally larger stability constants in higher dimensions (§3.4.3)). Likewise, non-negativity constraints are more restrictive and thus more stabilizing when imposed on a 3D-object (§3.5). The origin of the improved stability lies in *tomographic consistency*, i.e. in the characteristic redundancies in tomographic data described by the Helgason-Ludwig-conditions (theorem 2.2).

The stability results for XPCT thus motivate *joint*- or *all-at-once* reconstruction methods of the kind introduced in §2.7.4: the 3D-nature of the stabilizing consistency constraints suggests

that these may only be fully exploited by algorithms that “bear in mind” the full 3D-geometry already in the phase reconstruction step by interweaving it with tomographic inversion.

From ill-posed to well-posed – a mathematical Curiosity? In terms of abstract Hilbert space geometry, the introduced support constraints correspond to a restriction $T|_V$ of a linear forward operator $T \in \{\mathcal{T}, \mathcal{S}_\nu, \mathcal{T}^{(f_1, f_2)}\} : X \rightarrow Y$ to an *infinite-dimensional* closed linear subspace $V \subset X$ with $X = L^2(\mathbb{R}^m, \mathbb{R})$ and $V = L^2(\Omega, \mathbb{R})$. It is quite unusual that such a “non-invasive” modification establishes well-posedness of a formerly ill-posed inverse problem.

However, there is a sound mathematical reason why this is possible for the XPCI-problem, unlike other settings: in many standard examples of inverse problems, the associated forward operator $T : X \rightarrow Y$ (or its Fréchet-derivative in the nonlinear case) is *compact*, i.e. maps bounded subsets of X to relatively compact subsets of Y . Compact linear operators on infinite-dimensional normed spaces X may *never* have a continuous inverse, their inversion is necessarily ill-posed. Importantly, also any restriction $T|_V : V \rightarrow Y$ to a linear subspace $V \subset X$ remains compact if T is compact and so ill-posedness persists whenever $V \subset X$ is infinite-dimensional. Accordingly, the stability approach of the presented article may not work for any inverse problem with a compact forward operator.

The somewhat unusual route to establishing well-posedness and stability is thus only possible because of the *non-compactness* of the XPCI-forward operators $T \in \{\mathcal{T}, \mathcal{S}_\nu, \mathcal{T}^{(f_1, f_2)}\}$ on their natural L^2 -domains. While this does not mean that similar stability results can be obtained for *all* inverse problems with a non-compact forward map, the analysis of the present work might still be adaptable to *some* problems of this kind.

How many holograms are needed? As noted in the motivation §3.1, typically several (e.g. four) holograms are measured at different Fresnel numbers to ensure stable image reconstruction in practical XPCI – even when homogeneity of the object may be assumed. Quantitative predictions for the governing Lipschitz-stability constant $C_{IP^*}(\dots)$ from theorems 3.1 to 3.3 or numerical computations via the methods from §3.4 allow to distinguish settings where multiple measurements are required from those where a single hologram is sufficient: if the available support constraint (possibly plus homogeneity) is sufficiently tight such that $C_{IP^*}(\dots) \gtrsim 0.1$, then the beneficial effect of acquiring additional holograms might be too weak to justify the experimental effort. Quantitatively, $C_{IP^*}(\dots) \gtrsim 0.1$ translates to the following requirements for the modified Fresnel number f associated with the support-size (compare fig. 3.4):

- *Significantly absorbing homogeneous objects*: $f \lesssim 10^3$
- *Weakly absorbing homogeneous objects* (including pure phase-objects): $f \lesssim 10^2$
- *General objects* (independent phase- and absorption-image): $f \lesssim 10$

As outlined in §3.6.2, f may be decreased (thus increasing stability) for a fixed specimen in a divergent-beam setup by increasing the source-to-sample-distance, at the cost of a lower geometric magnification. Combined with the identified stability-requirement $f \lesssim 10$, this observation in particular gives a detailed answer to a controversial question in XPCI, raised by the mismatch between uniqueness-theory and practical experience (compare §3.1):

Stably retrieving phase- and absorption-image from a single hologram without homogeneity constraint is possible – yet, only for small objects or at low geometric magnifications.

In how far a reduction of magnification to improve stability is reasonable will be further discussed in §4.3 in conjunction with the results from Article 2.

Whenever the above stable ranges of Fresnel numbers are not accessible, in particular when the sample is so large that it does not even fit into the imaged field-of-view (a *region-of-interest* imaging setting as in 1.3(c)), acquiring multiple holograms is still advisable from the point-of-view of the presented stability analysis.

Choice of Fresnel numbers in the case of multiple holograms: For the standard case of XPCI from multiple holograms under homogeneity constraints, the Fresnel numbers f_1, \dots, f_ℓ are often chosen such that the zeros of the associated CTFs $s_\nu^{(f_1)}, \dots, s_\nu^{(f_\ell)}$ mutually cancel out as much as possible to reduce ill-posedness [212], compare §2.4.4. Interestingly, the derived stability theorem 3.3 for the reconstruction of *general objects* (no homogeneity) from two holograms motivates a different setup: for the stability constant $C_{\text{IP3}}(\Omega, f_1, f_2)$ to be large, the two Fresnel numbers f_1, f_2 should be chosen such that the difference $f_1^{-1} - f_2^{-1}$ is as large as possible, corresponding to maximally distinct acquisition-setups for the two holograms.

Combining completely holographic with high-resolution measurements: Probably, the most important take-home message is the following:

A small Fresnel number is always desirable in view of stability of XPCI.

In particular, even the most challenging XPCI-setting become fully stable in the *completely holographic regime*, as outlined in §3.6. The analysis in §3.6.1 indicates that this even remains true when the full nonlinear XPCI-model is considered.

According to §3.6.2, only *low-resolution* images can be obtained from completely holographic measurements due to the required small geometrical magnifications and finite detector-resolutions. Notably, however, this is already of considerable value as *low-frequency* image-modes have been identified as the least stable ones to reconstruct in standard, less holographic, high-resolution XPCI-setups. This motivates to reconstruct from data given by one completely holographic hologram, stabilizing low-frequency components, and one or two additional holograms in the conventional high-magnification setting to circumvent the resolution-limits. In practice, image reconstruction may then be achieved via a scheme of the following form:

- (1) Image reconstruction from the completely holographic, low-magnification data.
- (2) Upsampling of the obtained low-resolution image by interpolation to the natural resolution of the high-magnification hologram(s).
- (3) Image reconstruction from the high-magnification data using the upsampled image as a prior, i.e. penalizing deviations from the latter.

In particular, such a procedure might also enable stable high-resolution imaging of phase and absorption as independent parameters, according to the presented stability analysis.

Chapter 4

Locality Estimates for Fresnel-wave-propagation and stability of X-ray phase contrast imaging with finite detectors

The second theoretical study, Article 2 [142], aims to understand XPCI slightly beyond the scope of the standard idealized models. Namely, the article considers the very practical limitation that real-world detectors have a finite extent and may thus never capture a “complete” hologram. From a practical point of view, the analysis sheds a light on the fundamental question of achievable *resolution* in XPCI. By comparing to real reconstruction examples in chapter 5, we will indeed see that the resolution limits arising from a finite detector play a significant role in practical imaging.

4.1 Motivation

The stability results presented in §3, appealingly strong though they are, are based on the linear CTF-model of XPCI, which involves several idealizations. The assumptions most frequently discussed (and relaxed) in the literature include ideal plane-wave illumination [83, 101, 170, 84], linearity (i.e. weak scattering) [194, 47, 87] as well as full coherence [166, 41, 86, 87], see also §2.1.5. An issue in real-world XPCI that is only rarely considered, however, is the simple fact that real-world detectors cover only a finitely large (typically square) area. This implies that only a *restriction* of the hologram intensity $I^{\text{obs}}|_K$ to some bounded *detection domain* $K \subset \mathbb{R}^2$ is measured. In the following, K is also referred to as the *field-of-view* (*FoV*) or simply *detector*.

It is important to note that, in principle, this experimental restriction is fundamentally incompatible with standard theory of linearized XPCI: restricting to a sub-domain $K \subsetneq \mathbb{R}^2$ breaks translational invariance, which is essential to formulating the CTF-model via Fourier-multipliers, i.e. convolutions. As a consequence, direct CTF-inversion (see §2.7.2.1), that correspond to a deconvolution operation, strictly loses its validity for data acquired by finite detectors. This breakdown of the CTF-model is easily missed as CTF-based reconstruction is typically implemented via FFTs, that implicitly replace the assumption of an infinite FoV with

(incorrect) periodic boundary conditions at the detector-edges, compare §2.6.2.3. Likewise, the stability theory from §3 does not simply carry over to a finite-detector-setting.

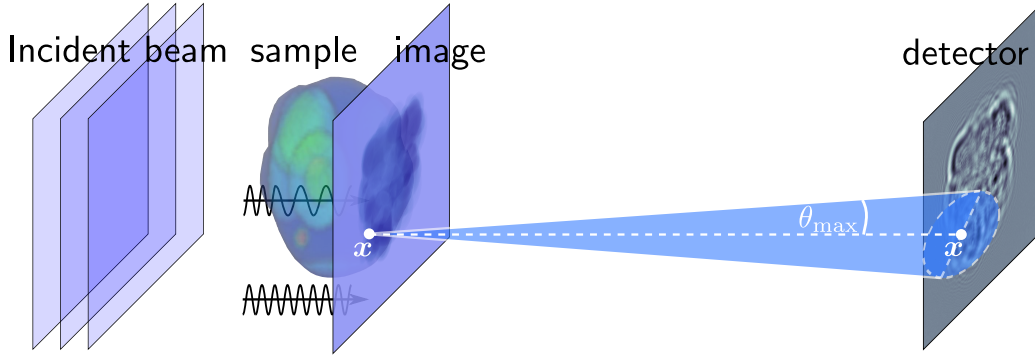


Figure 4.1. Numerical aperture NA in XPCI: As usual for lensless imaging systems, $NA = \sin(\theta_{\max})$ is defined via the maximum angle θ_{\max} such that all light scattered under θ_{\max} from a point x in the object-plane is still detected. Notably, this implies that NA varies within the field-of-view, indicating that the same holds for the achievable resolution.

What physics seems to tell us: When discussed in physical literature, the impact of a finite detector is often argued to be well-understood, see e.g. [158, 127, 126, 171]: for lensless imaging setups, the area of the FoV covered by the detector determines the *numerical aperture* $NA = \sin(\theta_{\max})$, where θ_{\max} is the maximal angle at which scattered light is still collected by the detector, as sketched in fig. 4.1. By Abbe’s diffraction limit, NA limits the *resolution* r , i.e. the size of the smallest object-features that may be faithfully imaged to

$$r_{\text{Abbe}} = \frac{\lambda}{2 \sin(\theta_{\max})} \quad (\lambda : \text{wavelength}) \quad (4.1.1)$$

Upon a second glance, however, this perspective on the finite-detector-problem can be argued to be over-simplifying for different reasons:

- Abbe’s criterion is strictly valid only for *far-field* optics, while *near-field* imaging is considered in the present work. Indeed, certain super-resolution imaging techniques [23, 89] exploit that the resolution-limit may be overcome in near-field optics.
- Even if a resolution limit arises, it remains unclear whether this is the *only* effect induced by a bounded field-of-view on image reconstruction in XPCI.
- How sharp is the resolution limit? For example, is it possible to improve the resolution by a factor of 1.5 by acquiring very low-noise holograms?
- Abbe’s diffraction limit may be *overly optimistic*: it is based on asking at what minimal distance a *very specific* object composed of two point sources is still faithfully imaged. On the contrary, to truly guarantee some resolution r , one would like *any* possible object-structures of lengthscales greater or equal to r to be accurately reconstructed.

Accordingly, it is necessary to study the effects of a finite FoV in XPCI in a rigorous manner beyond Abbe’s criterion. This is achieved in the present article, by deriving stability-estimates for image reconstruction in a finite-detector-setting.

Why not simply use a larger detector? One possible reason why the finite-detector-problem has received comparably little attention is that it seems easy to fix: if the current detector does not capture enough of the hologram, then why not just use a larger one? To understand this, one should recall from §2.2.1.5 that XPCI is a *holographic* imaging technique: the data-contrast mainly* arises from interference of the scattered wave-field induced by the imaged sample and the unscattered part of the probing X-ray beam. Importantly, this means that only data from detector-regions that are sufficiently *illuminated* by the primary beam provide strong contrast. On the contrary, the opening angle of the divergent X-ray beam in high-resolution XPCI-setups is typically very small (in the order of 0.1 degrees for the GINIX, see §1.2). As a consequence, the width of the produced beam-profile is typically hardly sufficient to ensure full illumination of a standard-size detector, as can be seen in fig. 2.3. The finite beam-divergence thus limits the (illuminated) FoV independently of the detector-size.

In principle, this limitation may be overcome by also using the *non-holographic* data-components for image reconstruction, given by intensity of the scattered wave-field alone ($|\Psi_s|^2$ in §2.2.1.5). This approach is known as *Fresnel coherent diffractive imaging* (FCDI) [208], as it resembles conventional CDI except that data is measured in the near-field, Fresnel-regime instead of the far-field regime. Importantly, FCDI also shares the drawbacks of CDI: challenges in phase reconstruction due to the highly non-linear problem structure and a dramatically worse fluence-resolution relationship compared to holographic XPCI [85]: much more X-ray photons are needed in (F)CDI to achieve the same resolution, which translates into higher exposure times and increased radiation damage. Accordingly, although XPCI-data also contains some information in the dark “outskirts” of (non-flat-field-corrected) holograms (see fig. 2.3(a)), its appealing advantages compared to CDI only apply within the illuminated detector-area. In this sense, we may conclude that a finite field-of-view, limited by the divergence of the probing X-ray beam, is fundamental to (holographic) XPCI. Understanding the effects of a finite FoV is thus vital to assess the potential and limits of the imaging technique.

4.2 Summary of the results

4.2.1 Basic setting

Holograms measured by a finite detector K provide data of the form $I|_K \approx 1 + \mathcal{N}(h)|_K$ or $I|_K \approx 1 + \mathcal{T}(h)|_K$, where $\mathcal{N}(h) = |\mathcal{D}(\exp(-h))|^2 - 1$ and $\mathcal{T}(h) = -2\text{Re}(\mathcal{D}(h))$ denote the principal forward maps of XPCI. The considered inverse problem 4.1 amounts to reconstructing the image h from such data. Notably, this problem combines difficulties arising from a finite FoV with those related to the phase-problem in XPCI. Therefore, the simplified setting of reconstructing h from *truncated Fresnel-data* $\mathcal{D}(h)|_K$ is studied in addition:

Inverse Problem 4.1 (Reconstruction of complex-valued images). For $\Omega, K \subset \mathbb{R}^m$, reconstruct a *complex-valued* image $h \in L^2(\Omega)$ from either of the following data:

(a) $g_{(a)}^{\text{obs}} \approx \mathcal{D}(h)|_K$

(b) $g_{(b)}^{\text{obs}} \approx \mathcal{T}(h)|_K$

*In fact, all contrast-components in XPCI-data that are linear in the object, being the only(!) contributions used in CTF-reconstruction for example, are holographic in the above sense.

$$(c) \quad g_{(c)}^{\text{obs}} \approx \mathcal{N}(h)|_K$$

As inverse problem 4.1(b),(c) correspond to reconstructions from *less data* (missing phase or imaginary part of $\mathcal{D}(\dots)$) compared to variant (a), any instability or ill-posedness in the model-problem (a) manifests also in the XPCI-problems (b) and (c). As studying (a) is simpler, this motivates to restrict the analysis to Fresnel-data, as is widely done in the article. Furthermore, note that we study inverse problem 4.1 in arbitrary dimensions $m \in \{1, 2, 3, \dots\}$ although XPCI corresponds to $m = 2$, for similar reasons as in the stability analysis of chapter 3.

Interestingly, it turns out to be highly relevant for the finite-detector-problem whether general complex-valued images are to be recovered or real-valuedness can be assumed. According to §2.4.3, the latter is true for image reconstruction in XPCI under *homogeneity constraints*, which renders the real-valued case highly relevant for practical imaging.

Inverse Problem 4.2 (Reconstruction of real-valued (homogeneous) images). For $\Omega, K \subset \mathbb{R}^m$ and $\nu \in \mathbb{R}$, reconstruct a *real-valued* image $\varphi \in L^2(\Omega, \mathbb{R})$ from either of the following data:

$$(a) \quad g_{(a)}^{\text{obs}} = \mathcal{D}(\varphi)|_K$$

$$(b) \quad g_{(b)}^{\text{obs}} = \mathcal{S}_\nu(\varphi)|_K$$

$$(c) \quad g_{(c)}^{\text{obs}} = \mathcal{N}_\nu(\varphi)|_K$$

4.2.2 Preliminary results and goal of the article

For a first assessment of the problem, some preliminary results are shown in the article:

- (1) *Existence of leakage*: For any compactly supported image $h \neq 0$, the Fresnel-data $\mathcal{D}(h)$ is always supported in the entire lateral domain, i.e. $\text{supp}(\mathcal{D}(h)) = \mathbb{R}^m$.
- (2) *Uniqueness*: If Ω is bounded and K contains an open set, all variants of inverse problems 4.1 and 4.2, are uniquely solvable according to theorem 2.10.
- (3) *Severe ill-posedness*: If K is bounded and Ω contains an open set, all variants of inverse problems 4.1 and 4.2 are severely ill-posed, i.e. the singular values associated with the forward operators decay super-algebraically.

Result (1) means that a finite FoV always has an effect on image reconstruction, no matter how large the detector is. Yet, statement (2) shows that these effects do not manifest in terms of non-uniqueness. Note that this seems contradictory to an anticipated resolution limit (compare §4.1), as it implies that *any*, i.e. also arbitrarily sharp structures may be recovered from finite detector data – *in theory*. However, (3) indicates that such reconstructions may still be too unstable to be feasible *in practice*: by the arising ill-posedness, there is always an infinite-dimensional subspace $X_{\text{unstab}} \subset L^2(\Omega, (\mathbb{R}))$ of image-modes which induce so little contrast in a finite-detector-setting that they are practically irretrievably lost for noisy data.

In combination, (2) and (3) show that the effects of a finite detector may only be understood by a *stability analysis*, elucidating the structure of the unstable modes in X_{unstab} and identifying subspaces $X_{\text{stab}} \subset L^2(\Omega, (\mathbb{R}))$ of images that may still be stably reconstructed. This is the principal aim of the following sections §4.2.3 to §4.2.7.

4.2.3 Assessment by Gaussian wave-packets:

Stability of inverse problems 4.1 and 4.2 is first assessed by studying Fresnel propagation of Gaussian wave-packets with width $\sigma > 0$, frequency-vector $\boldsymbol{\xi} \in \mathbb{R}^m$ and center $\mathbf{x}_0 \in K \subset \mathbb{R}^m$ within a square FoV $K = [-\frac{1}{2}; \frac{1}{2}]^m$:

$$h_{\boldsymbol{\xi}, \mathbf{x}_0}(\mathbf{x}) = \exp(i\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{x}_0)) \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_0\|^2}{2\sigma^2}\right), \quad \mathbf{x} \in \mathbb{R}^m. \quad (4.2.1)$$

Such functions constitute a special class of localized oscillatory signals, for which Fresnel propagation may be computed analytically: as visualized in fig. 4.2(a), $\mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{x}_0}) \propto \exp(-(\mathbf{x} - \mathbf{x}_0 - \boldsymbol{\xi}/f)^2/(2\tilde{\sigma}^2))$ is again a modulated Gaussian, but with shifted center at $\mathbf{x}_0 + \boldsymbol{\xi}/f$.

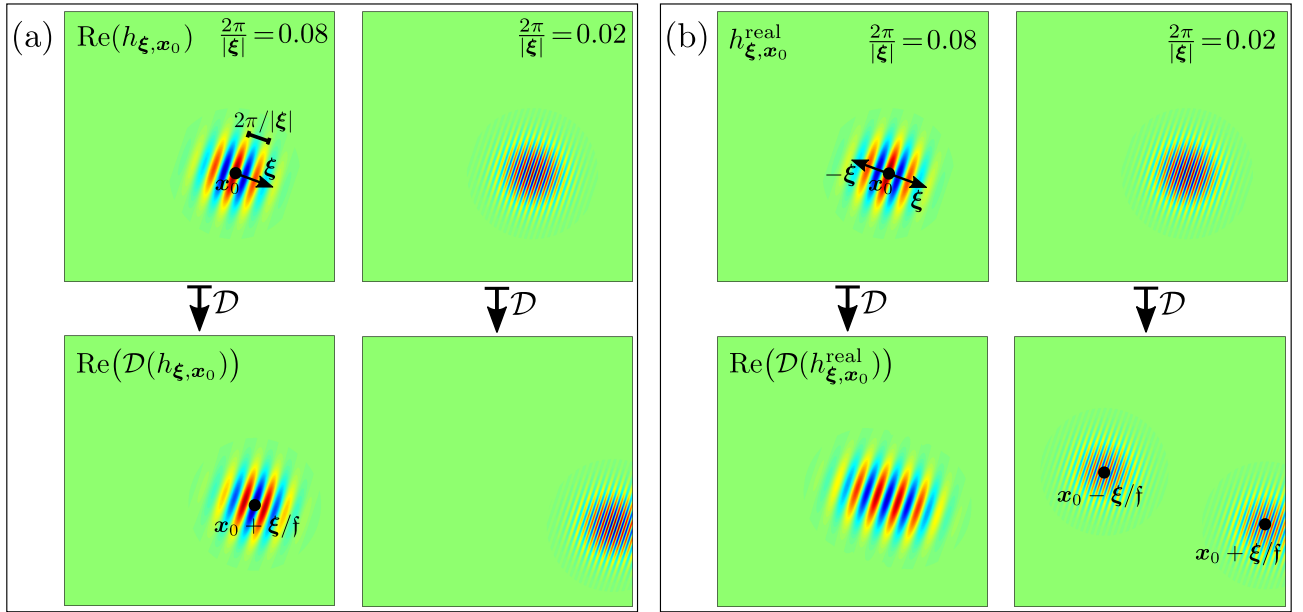


Figure 4.2. Propagation of Gaussian wave-packets within a square FoV $K = [-\frac{1}{2}; \frac{1}{2}]^2$ at $f = 10^3$. (a) Complex-valued wave-packets and their propagated versions $\mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{x}_0})|_K$, which are again Gaussian but with shifted center $\mathbf{x}_0 + \boldsymbol{\xi}/f$. (b) Real-valued case: wave-packets split upon Fresnel propagation into two Gaussians with centers $\mathbf{x}_0 \pm \boldsymbol{\xi}/f$. By determining the maximal value $\xi_r(\mathbf{x})$ such that all wave-packets of frequency $|\boldsymbol{\xi}| \leq \xi_r(\mathbf{x})$ with center \mathbf{x} remain visible within K upon Fresnel propagation, a local resolution $r(\mathbf{x}) = \pi/\xi_r(\mathbf{x})$ is estimated for each point in the FoV, yielding the plots in fig. 4.3.

The idea of the study is as follows: if image reconstruction is stable down to some resolution $r > 0$, then any perturbation of an object $h \rightarrow h + \tilde{h}$ by signals \tilde{h} varying on lengthscales greater or equal r should be “visible”, i.e. give rise to a significant increment $\mathcal{D}(h + \tilde{h})|_K - \mathcal{D}(h)|_K$ in the data. By linearity, this means that the induced relative contrast $\|\mathcal{D}(\tilde{h})|_K\|/\|\tilde{h}\|$ should be reasonably large. In particular, this should hold for all Gaussian wave-packets $\tilde{h} = h_{\boldsymbol{\xi}, \mathbf{x}_0}$ with frequency $|\boldsymbol{\xi}| \leq \pi/r$, as such constitute *special* perturbations “of resolution” $\pi/|\boldsymbol{\xi}|$ according to the half-wavelength criterion.

As its frequency $|\boldsymbol{\xi}|$ increases, any wave-packet $h_{\boldsymbol{\xi}, \mathbf{x}_0}$ will eventually propagate out of the FoV K , in which case it switches from being “visible” to “invisible”, see fig. 4.2(a). Accordingly, whether or not the center $\mathbf{x}_0 + \boldsymbol{\xi}/f$ of the propagated wave-packet $\mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{x}_0})|_K$ lies within K may serve as an (approximate) visibility-criterion. If $\text{dist}(\mathbf{x}, \partial K)$ denotes the distance of $\mathbf{x} \in K$ to the detector-boundary, the following holds true:

- (1) For all Gaussian wave-packets h_{ξ, \mathbf{x}_0} with frequency $|\xi| < \mathfrak{f} \operatorname{dist}(\mathbf{x}_0, \partial K)$, the propagated center $\mathbf{x}_0 + \xi/\mathfrak{f}$ lies within the FoV K , i.e. all such wave-packets are “visible”.
- (2) For all frequencies $\xi > \mathfrak{f} \operatorname{dist}(\mathbf{x}_0, \partial K)$, there exists a wave-packet h_{ξ, \mathbf{x}_0} with $|\xi| = \xi$ such that $\mathcal{D}(h_{\xi, \mathbf{x}_0})$ is centered outside K , i.e. which is “invisible”.

By the stability argument above, the local resolution at a point $\mathbf{x}_0 \in K$ based on the wave-packet analysis is therefore concluded it to be

$$r(\mathbf{x}_0) \approx \frac{\pi}{\mathfrak{f} \operatorname{dist}(\mathbf{x}_0, K)}. \quad (4.2.2)$$

Accordingly, the best resolution is achieved in the very center of the FoV with $\max_{\mathbf{x}_0 \in K} 1/r(\mathbf{x}_0) = 1/r(0) = \mathfrak{f}/2\pi = \mathfrak{f}$ and deteriorates towards the detector-edges, as visualized in fig. 4.3(a).

The resolution estimate (4.2.2) can be shown to coincide exactly with Abbe’s diffraction limit (4.1.1) associated with the spatially varying numerical aperture within the FoV, compare fig. 4.1. Yet, the Gaussian wave-packet analysis yields additional insights: if the width of the propagated wave-packet, $\tilde{\sigma} = (\sigma^2 + 1/(\mathfrak{f}\sigma)^2)^{1/2}$, is sufficiently small against $\operatorname{dist}(\mathbf{x}_0, K)$, the contrast $\zeta(\xi, \mathbf{x}_0) := \min_{|\xi|=\xi} \|\mathcal{D}(h_{\xi, \mathbf{x}_0})|_K\| / \|h_{\xi, \mathbf{x}_0}\|$ w.r.t. perturbations of frequency ξ drops rapidly near the resolution limit: according to mathematical details in the article, one has

$$\zeta(\xi, \mathbf{x}_0) \approx \begin{cases} 1 & \text{for } \xi \leq (1 - \varepsilon)\mathfrak{f} \operatorname{dist}(\mathbf{x}_0, K) \\ 0 & \text{for } \xi \geq (1 + \varepsilon)\mathfrak{f} \operatorname{dist}(\mathbf{x}_0, K) \end{cases} \quad \text{with transition } \varepsilon \sim \tilde{\sigma} / \operatorname{dist}(\mathbf{x}_0, K). \quad (4.2.3)$$

This implies that the resolution limit (4.2.2) is typically quite sharp – image reconstruction becomes severely unstable when aiming to resolve scales $\leq \pi/((1 + \varepsilon)\mathfrak{f} \operatorname{dist}(\mathbf{x}_0, K))$.

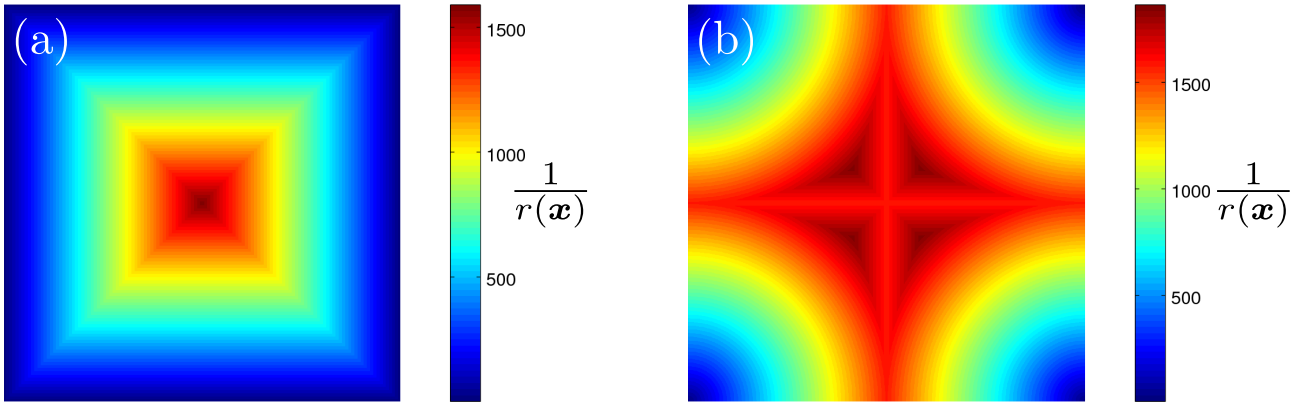


Figure 4.3. (a) Upper bound on the stably reconstructible local resolution $1/r(\mathbf{x})$ in inverse problem 4.1 (complex-valued images) for $m = 2$, $K = [-\frac{1}{2}; \frac{1}{2}]^2$, $\mathfrak{f} = 10^4$, based on the study of Gaussian wave-packets. (b) Same plot for inverse problem 4.2, i.e. for reconstruction of real-valued images.

The real-valued case: If the object to be reconstructed is known to be real-valued as in inverse problem 4.2, then only perturbations by *real-valued* wave-packets $h_{\xi, \mathbf{a}}^{\text{real}}(\mathbf{x}) := \cos(\xi \cdot (\mathbf{x} - \mathbf{a}) + \beta) \exp(-\|\mathbf{x} - \mathbf{x}_0\|^2 / (2\sigma^2))$ are admissible. As such can be represented as linear combinations of two wave-packets of the form (4.2.1) with the same center \mathbf{x}_0 but opposite frequency vectors $\pm \xi$, the Fresnel-data $\mathcal{D}(h_{\xi, \mathbf{a}}^{\text{real}})$ is formed by *two* Gaussians centered at $\mathbf{x}_0 \pm \xi/\mathfrak{f}$, as shown

fig. 4.2(b). A perturbation $h \mapsto h + h_{\xi, \alpha}^{\text{real}}$ is thus “visible” in the Fresnel-data $\mathcal{D}(h)|_K$ whenever *either* of these propagated centers still lies within the FoV, i.e. $\mathbf{x}_0 + \xi/f \in K$ or $\mathbf{x}_0 - \xi/f \in K$.

Based on this relaxed visibility-criterion for wave-packets compared to the complex-valued case, improved estimates of the local resolution are obtained, as plotted in fig. 4.3(b). Other than in fig. 4.3(a), the resolution in (b) is high throughout the FoV except for neighborhoods of the detector-*corners*. Note that this resolution-improvement is *not* predicted by Abbe’s criterion (4.1.1) and considerations of the numerical aperture. Yet, the resolution in the center of the FoV is unchanged, $r(0) = 2\pi/f = 1/f$ holds both in a real- and complex-valued setting.

4.2.4 General locality theory

In §4.2.3, stability has been studied only w.r.t. a *special* class of object-perturbations. Hence, the conclusions on the effect of a finite FoV, in particular the derived resolution limit, may be overly *optimistic*. Therefore, the Gaussian wave-packet study is complemented by *worst-case* estimates, bounding the leaked Fresnel-wave-field $\mathcal{D}(h)|_{K^c}$ for arbitrary object-signals h . While the required mathematical tools are technical in parts, the obtained results essentially reproduce the intuitive insights from §4.1 and §4.2.3 in a more rigorous form.

The analysis is based on the convolution-form (2.2.9) of the Fresnel propagator, $\mathcal{D}(h) = k_f * h$. Notably, the kernel $k_f(\mathbf{x}) \propto \exp(i\mathbf{f}\mathbf{x}/2)$ is not localized in any sense but has constant modulus throughout the entire \mathbb{R}^m . This non-locality means that Fresnel propagation may transport information over arbitrary lateral distances in principle, i.e. may arbitrarily *delocalize* object-structures. The principal idea to resolve this is to decompose the kernel into an “inner” and “outer” part, $k_f = k_f|_P + k_f|_{P^c}$ for some $P \subset \mathbb{R}^m$. Then the leaked wave-field outside the detection-domain K can be written in the form

$$\mathcal{D}(h)|_{K^c} = (k_f|_P * h)|_{K^c} + (k_f|_{P^c} * h)|_{K^c} \quad (4.2.4)$$

by linearity. Now, if $\text{supp}(h) \subset \Omega$ and if P is chosen such that $\bar{P} + \Omega \subset K$, it follows that $(k_f|_P * h)|_{K^c} = 0$ by standard results on the support of convolutions and thus

$$\mathcal{D}(h)|_{K^c} = (k_f|_{P^c} * h)|_{K^c} \quad \text{if} \quad \bar{P} + \Omega \subset K, \quad (4.2.5)$$

i.e. the leaked wave-field can be expressed as a convolution with only the “outer” propagating-kernel $k_f|_{P^c}$. This enables leakage-bounds in terms of the *filter-response* associated with $k_f|_{P^c}$:

$$\begin{aligned} \|\mathcal{D}(h)|_{K^c}\| &= \|(k_f|_{P^c} * h)|_{K^c}\| \leq \|k_f|_{P^c} * h\| = \|\hat{p}^{\text{leak}} \cdot \mathcal{F}(h)\|, \\ \hat{p}^{\text{leak}} &:= (2\pi)^{\frac{m}{2}} |\mathcal{F}(k_f|_{P^c})| \end{aligned} \quad (4.2.6)$$

where the convolution-theorem has been used. After casting the leakage filter-response \hat{p}^{leak} to a more explicit form, these considerations lead to principal bound for the leaked wave-field:

Theorem 4.1 (Principal leakage estimate). *Let $K, \Omega, P_{\text{leak}} \subset \mathbb{R}^m$ be measurable sets such that the boundary ∂K has Lebesgue-measure zero and $\Omega + P_{\text{leak}} = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in \Omega, \mathbf{y} \in P_{\text{leak}}\} \subset K$. Moreover, let $\mathcal{D}(\mathbf{1}_{P_{\text{leak}}^c}) \in L^\infty(\mathbb{R}^m)$. Then it holds for all $h \in L^2(\Omega)$*

$$\|\mathcal{D}(h)|_{K^c}\| \leq \|\hat{p}^{\text{leak}} \cdot \mathcal{F}(h)\| \quad \text{with} \quad p^{\text{leak}}(\xi) := |\mathcal{D}(\mathbf{1}_{P_{\text{leak}}^c})(\xi/f)|. \quad (4.2.7)$$

For the case of *rectangular* domains P , $\mathcal{D}(\mathbf{1}_{P^c})$ can be computed analytically via the known propagation-formula for the Heaviside-function $\theta(x) := \mathbf{1}_{\mathbb{R}_{\geq 0}}(x)$,

$$\mathcal{D}(\theta)(x) = \tilde{\theta}(\mathfrak{f}^{\frac{1}{2}}x) \quad \text{with} \quad \tilde{\theta}(x) := \frac{1}{2} \operatorname{erfc}\left(\frac{i-1}{2}x\right), \quad (4.2.8)$$

and using linearity, isotropy and separability of the Fresnel propagator. Via theorem 4.1, this enables explicit leakage estimates for the most relevant case of a square detector K :

Theorem 4.2 (Leakage bound for square domains). *Let $K = [-\frac{1}{2}; \frac{1}{2}]^m$ and $\Omega = [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta]^m$ for some $0 < \Delta < \frac{1}{2}$. Then it holds for all $h \in L^2(\Omega)$*

$$\|\mathcal{D}(h)|_{K^c}\| \leq \|\hat{p}_{\mathfrak{f}, \mathfrak{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(h)\|, \quad \hat{p}_{\mathfrak{f}, \mathfrak{f}_\Delta}^{\text{leak}}(\boldsymbol{\xi}) := \left(\sum_{j=1}^m |\tilde{\eta}_{\mathfrak{f}_\Delta}(\mathfrak{f}^{-\frac{1}{2}}\mathbf{e}_j \cdot \boldsymbol{\xi})|^2 \right)^{\frac{1}{2}} \quad (4.2.9)$$

$$\tilde{\eta}_{\mathfrak{f}_\Delta}(x) := (|\tilde{\theta}(x - \mathfrak{f}_\Delta^{1/2})|^2 + |\tilde{\theta}(-x - \mathfrak{f}_\Delta^{1/2})|^2)^{\frac{1}{2}} \quad (4.2.10)$$

where \mathbf{e}_j denotes the unit normal vector along the j th dimension and $\mathfrak{f}_\Delta := \Delta^2 \mathfrak{f}$ is the Fresnel number associated with the lengthscale Δ .

Figure 4.4(a) plots the filter-response $\hat{p}_{\mathfrak{f}, \mathfrak{f}_\Delta}^{\text{leak}}$ in $m = 1$ dimensions for $\mathfrak{f}_\Delta = 100$. The behavior is that of a *high-pass* filter, where the cut-off, i.e. the transition from $\hat{p}_{\mathfrak{f}, \mathfrak{f}_\Delta}^{\text{leak}} \ll 1$ to values around one, occurs at frequency $\xi_{\text{cut}} = (\mathfrak{f} \cdot \mathfrak{f}_\Delta)^{1/2} = \Delta \mathfrak{f}$. Thus, the bound (4.2.9) states that the fraction of the wave-field leaked outside the FoV is small for *low-frequency* objects, whereas variations on lengthscales smaller than π/ξ_{cut} may lead to a significant signal-loss. Since $\Delta = \min_{\mathbf{x} \in \Omega} \operatorname{dist}(\mathbf{x}, \partial K)$, this is in perfect agreement with the wave-packet study in §4.2.3.

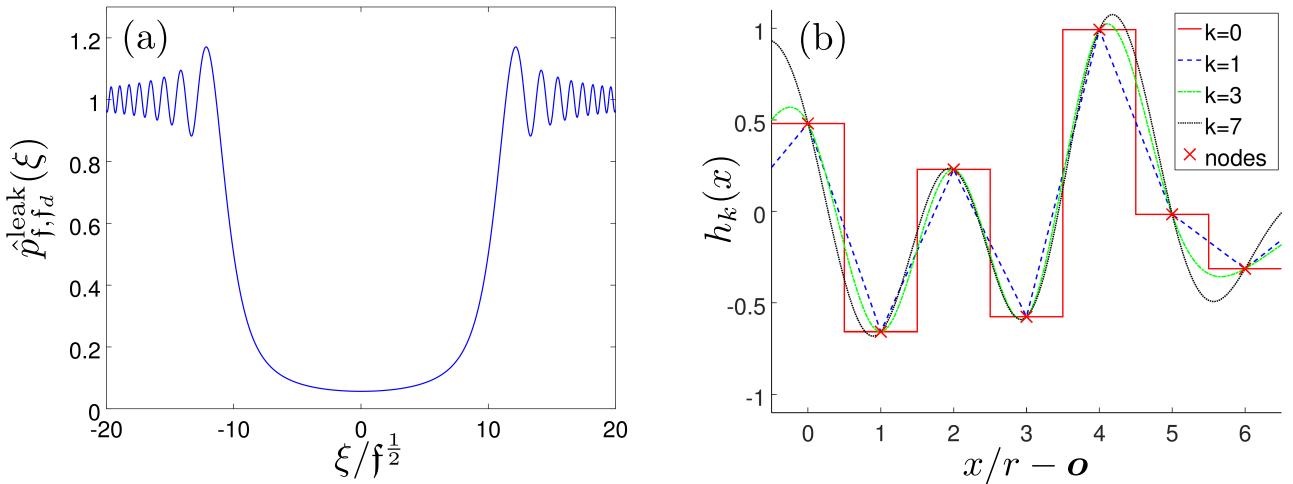


Figure 4.4. The principal ingredients of the analysis for the finite-FoV-problem. (a) Plot of the leakage-filter from theorem 4.2 for $\mathfrak{f}_\Delta = 100$ in one dimension. (b) 1D-examples of B-splines $h_k \in \mathfrak{B}_{k,r,\mathbf{o}}^m$ for different orders k as defined in §4.2.5, which are used to model objects of finite resolution r .

4.2.5 Stability estimates for spline-objects

The leakage-bound obtained so far guarantees that the information-loss due to a finite detector is small in some settings, but it is unclear whether it is small *enough* in the sense that the

wave-field retained within the FoV K is sufficient for stable image reconstruction. The next goal is therefore to derive stability estimates from the leakage-bound in theorem 4.2.

To this end, one has to restrict to a class of objects for which the right-hand side of (4.2.9) is uniformly small. By the high-pass nature of $\hat{p}_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}$, a canonical choice would be to restrict to *band-limited* objects h , whose Fourier transform $\mathcal{F}(h)(\boldsymbol{\xi})$ vanishes at frequencies above the cut-off of the filter $\hat{p}_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}$. However, this choice is not admissible here because, for theorem 4.2 to hold, h also has to be compactly supported (in real space) and thus cannot be bandlimited at the same time. This is a well-known result from Fourier analysis.

A more suitable class of objects is found in *multi-variate (B-)splines*:

$$h \in \mathfrak{B}_{k,r,\mathbf{o}}^m := \left\{ h : \mathbf{x} \mapsto \sum_{\mathbf{j} \in \mathbb{Z}^m} b_{\mathbf{j}} B_k^m(\mathbf{x}/r - \mathbf{j} - \mathbf{o}) : (b_{\mathbf{j}}) \in \ell^2(\mathbb{Z}^m) \right\} \quad (4.2.11a)$$

$$B_k^m(x_1, \dots, x_m) := \prod_{j=1}^m B_k(x_j), \quad B_k = \begin{cases} B_0 * B_{k-1} & \text{for } k \in \mathbb{N} \\ \mathbf{1}_{[-\frac{1}{2}; \frac{1}{2}]} & \text{for } k = 0 \end{cases} \quad (4.2.11b)$$

$k \in \mathbb{N}_0$ is the order of the splines, determining the smoothness of the interpolation between nodes that are uniformly arranged on the Cartesian grid $\mathbf{o} + r\mathbb{Z}^m$. The steplength r defines the smallest lengthscale of variations for elements $h \in \mathfrak{B}_{k,r,\mathbf{o}}^m$. Accordingly, splines model objects of *finite resolution* r . For illustration, fig. 4.4 shows examples of 1D-splines of different order k .

For the present work, the crucial property of the space $\mathfrak{B}_{k,r,\mathbf{o}}^m$ is that its elements are *quasi-bandlimited*, as is proven in the article (see theorem 9.14): for any $k \in \mathbb{N}_0$ and $\nu \geq 1$, there exists a constant $C_{\text{band}}(k, \nu) < 1$ such that, for $r > 0$ and $\Xi_r := [-\pi/r; \pi/r]$, it holds that

$$\|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\| \leq C_{\text{band}}(k, \nu) \|\mathcal{F}(h)\| \quad \text{for all } h \in \mathfrak{B}_{k,r,\mathbf{o}}^1. \quad (4.2.12)$$

The bound (4.2.12) states that the fraction of the Fourier transform concentrated outside the low-frequency band Ξ_r is uniformly bounded for splines of resolution r . For splines of high order k , the constant $C_{\text{band}}(k, \nu)$ decreases rapidly for values $\nu > 1$. The result may also be extended to splines in higher dimensions $m > 1$, compare theorem 9.15.

By combining the quasi-bandlimitation bound (4.2.12) with theorem 4.2, Lipschitz-stability estimates are proven for spline-objects. The principal idea is to decompose the expression for the leaked wave-field, $\hat{p}_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(h)$, into a low- and high-frequency parts:

$$\begin{aligned} \|\mathcal{D}(h)|_{K^c}\|^2 &\leq \|\hat{p}_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(h)\|^2 = \underbrace{\|\hat{p}_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(h)|_{\nu\Xi_r^m}\|^2}_{\leq C_1^2 \|h\|^2 \text{ by high-pass-form of } \hat{p}_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}} + \underbrace{\|\hat{p}_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(h)|_{(\nu\Xi_r^m)^c}\|^2}_{\leq C_2^2 \|h\|^2 \text{ according to (4.2.12)}} \quad (4.2.13) \\ &\leq C_1^2 \|h\|^2 + C_2^2 \|h\|^2 \end{aligned}$$

The constants are small $C_1, C_2 \ll 1$ for suitably chosen $\nu > 1$, provided that $r > \pi/(\Delta \mathbf{f})$ and k large enough, i.e. for sufficiently *smooth* splines. Since $\|\mathcal{D}(h)|_K\|^2 = \|\mathcal{D}(h)\|^2 - \|\mathcal{D}(h)|_{K^c}\|^2$ and $\|\mathcal{D}(h)\| = \|h\|$ by unitarity of \mathcal{D} , the leakage-bound (4.2.13) yields a stability estimate:

Theorem 4.3 (Stability estimate for spline-objects). *Let $K = [-\frac{1}{2}; \frac{1}{2}]^m$ and $\Omega = [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta]^m$ for $0 < \Delta < \frac{1}{2}$. Let $\mathbf{f}_\Delta := \Delta^2 \mathbf{f}$ and $\mathbf{f}_r := r^2 \mathbf{f}$ denote the Fresnel numbers of the scales Δ and $r > 0$. Furthermore, let $\nu \geq 1$ and $\Xi := [-\nu\pi/\mathbf{f}_r^{1/2}; \nu\pi/\mathbf{f}_r^{1/2}]$. Then it holds that*

$$\|\mathcal{D}(h)|_K\| \geq C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)^m \|h\| \quad \text{for all } h \in \mathfrak{B}_{k,r,\mathbf{o}}^m \cap L^2(\Omega). \quad (4.2.14)$$

With $\tilde{\eta}_{f_\Delta}$ as defined in theorem 4.2, the constant is given by

$$C_{\text{stab}}(f_\Delta, f_r, k, \nu) = \max \left\{ 1 - C_{\text{low}}^2 - C_{\text{band}}(k, \nu)^2 (C_{\text{tot}}^2 - C_{\text{low}}^2), 0 \right\}^{1/2} \quad (4.2.15)$$

$$C_{\text{low}} := \max_{x \in \Xi} \tilde{\eta}_{f_\Delta}(x), \quad C_{\text{tot}} := \max_{x \in \mathbb{R}} \tilde{\eta}_{f_\Delta}(x)$$

Theorem 4.3 yields the sought positive result: (spline-)objects of finite resolution with support sufficiently far away from the detector-boundary can be stably recovered from Fresnel-data acquired in a finite FoV. Examples in the article show that the spline-resolution r , down to which theorem 4.3 guarantees stability (i.e. $C_{\text{stab}} > 0$ and sufficiently large), can be as close as by a factor of ≈ 1.2 to the resolution-bound (4.2.2) from the wave-packet analysis. Upper and lower limits on the stably achievable resolution are thus in good agreement.

4.2.6 Improved bounds for real-valued objects

For Gaussian wave-packets, it has been found that real-valuedness induces a certain symmetry of the delocalizing action of Fresnel propagation. Indeed, this turns out to be a general principle beyond the considered class of signals, related to the fact that the Fourier transform of a real-valued function is *hermitean*, i.e. $\mathcal{F}(\varphi)(-\cdot) = \overline{\mathcal{F}(\varphi)}$. As a consequence, it holds for any filter-response $\hat{p} \in L^\infty(\mathbb{R}^m)$ that

$$\|\hat{p} \cdot \mathcal{F}(\varphi)\| = \|\text{sym}(\hat{p}) \cdot \mathcal{F}(\varphi)\| \quad \text{for all } \varphi \in L^2(\mathbb{R}^m, \mathbb{R}) \quad (\text{real-valued}), \quad (4.2.16)$$

where the symmetrized response is defined by $\text{sym}(\hat{p})(\boldsymbol{\xi}) := 2^{-1/2}(|\hat{p}(\boldsymbol{\xi})|^2 + |\hat{p}(-\boldsymbol{\xi})|^2)^{1/2}$.

The consequence of the above observations for Fresnel propagation is most prominent when considering a (hypothetical) setting where the detection domain is given by a half-space, w.l.o.g. $K = \mathbb{R}_{\geq 0} \times \mathbb{R}^{m-1}$. Then, for a real-valued signal $\varphi \in L^2(K, \mathbb{R})$ initially supported in K , the leaked wave-field $\mathcal{D}(\varphi)|_{K^c}$ may never exceed a certain fraction of the total intensity:

$$\|\mathcal{D}(\varphi)|_{K^c}\| \leq C_{\text{sym}} \|\varphi\| \quad \text{with} \quad C_{\text{sym}} \leq \max_{x \in \mathbb{R}} \text{sym}(\tilde{\theta})(x) \approx 0.837. \quad (4.2.17)$$

Notably, the bound (4.2.17) holds independently of any assumptions on the smoothness of φ or the distance of its support from the detector-boundary ∂K , but ceases to hold if real-valuedness is dropped. It states that only a limited fraction of a real-valued signal may propagate out of the FoV along a single direction and therefore denoted as *quasi-symmetric propagation principle*.

For the more practically relevant case of a square detector $K = [-\frac{1}{2}; \frac{1}{2}]^m$, object-information may leak outside the FoV along multiple directions. Hence, the *total* leakage along all $2m$ detector-edges may still be too large to ensure stability, despite the general bound (4.2.17). As a remedy, the quasi-symmetric propagation principle is combined with the leakage-estimates from theorems 4.1 and 4.2 in the complete article, compare §9.6. By this approach, it is proven that image reconstruction is stable at any point $\boldsymbol{x} \in K$ in the FoV that is close to no more than one detector-edge, where “how close” again depends on the finest object-scale r to be resolved, see §9.6.3 for details. In the natural 2D-setting $m = 2$, stability thus holds for objects of finite resolution with support sufficiently far from the *corners* of the detector (theorem 9.24). Accordingly, the stability analysis for real-valued objects qualitatively reproduces the picture from the wave-packet analysis in §4.2.3, visualized in fig. 4.3(b). Examples in the article also show *quantitative* agreement of the resulting resolution-estimates, though the deviations between upper and lower bounds are larger than in the complex-valued case. The latter indicates that the derived stability estimates for the real-valued setting are less sharp.

4.2.7 Extension to linearized XPCI

So far, the stability analysis of the finite-FoV-problem has been restricted to the simplest variant (a) of inverse problems 4.1 and 4.2, i.e. to image reconstruction from *phased* Fresnel-data $\mathcal{D}(h)|_K$. However, extending the results to the *phaseless* XPCI-setting is relatively simple as the hologram-data in variants (b) and (c) of inverse problems 4.1 and 4.2 differs from the Fresnel-data only by *pointwise* operations, $|\cdot|^2$ or $\text{Re}(\cdot)$, which commute with restrictions $f \mapsto f|_A$ to any set $A \subset \mathbb{R}^m$. In particular, for the linearized XPCI-setting in variant (b), this implies that the following bounds hold for any detection domain $K \subset \mathbb{R}^m$:

$$\frac{1}{2} \|\mathcal{T}(h)|_{K^c}\| = \|\text{Re}(\mathcal{D}(h))|_{K^c}\| = \|\text{Re}(\mathcal{D}(h)|_{K^c})\| \leq \|\mathcal{D}(h)|_{K^c}\| \quad \text{for } h \in L^2(\mathbb{R}^m) \quad (4.2.18a)$$

$$\frac{1}{2} \|\mathcal{S}_\nu(\varphi)|_{K^c}\| \stackrel{(2.4.6)}{=} \frac{1}{2} \|\mathcal{T}(ie^{-i\nu}\varphi)|_{K^c}\| \stackrel{(4.2.18a)}{\leq} \|\mathcal{D}(\varphi)|_{K^c}\| \quad \text{for } \varphi \in L^2(\mathbb{R}^m, \mathbb{R}). \quad (4.2.18b)$$

Intuitively, the estimates (4.2.18) express the simple fact that lost hologram-data outside the FoV may not contain more information than the corresponding fully phased Fresnel-data.

By virtue of (4.2.18), *all* of the derived leakage-estimates for Fresnel-data induce bounds for linearized XPCI. Stability estimates may also be obtained: for $T \in \{\mathcal{T}, \mathcal{S}_\nu\}$, it holds that

$$\|T(f)|_K\|^2 = \|T(f)\|^2 - \|T(f)|_{K^c}\|^2 \stackrel{(4.2.18)}{\geq} \|T(f)\|^2 - 4\|\mathcal{D}(f)|_{K^c}\|^2 \quad (4.2.19)$$

for all $f \in L^2(\mathbb{R}^m, (\mathbb{R}))$. The first summand on the r.h.s. of (4.2.19) is the expected contrast in the hypothetical case of an infinitely large detector $K = \mathbb{R}^m$. Accordingly, it may be bounded from below using the Lipschitz-stability estimates from article 1, see chapter 3, i.e. $\|T(f)\| \geq C_{\text{IP}^*}\|f\|$ for some $C_{\text{IP}^*} > 0$. The second, subtracted term in turn can be bounded from above via the previously derived leakage-estimates. In particular, for suitable spline-objects $f \in \mathfrak{B}_{k,r,\mathbf{o}}^m \cap L^2(\Omega, (\mathbb{R}))$, estimates of the form $\|\mathcal{D}(f)|_{K^c}\| \leq C_{\text{leak}}\|f\|$ with some $C_{\text{leak}} \ll 1$ hold by theorem 4.3. By combining the bounds, (4.2.19) implies

$$\|T(f)|_K\|^2 = (C_{\text{IP}^*}^2 - 4C_{\text{leak}}^2) \|f\|^2 \quad \text{for all } f \in \mathfrak{B}_{k,r,\mathbf{o}}^m \cap L^2(\Omega, (\mathbb{R})). \quad (4.2.20)$$

Although the estimate (4.2.20) is argued to be highly pessimistic (note in particular that the constant on the r.h.s. may become negative), it is demonstrated in the article that it indeed guarantees stability in certain practically relevant imaging regimes.

4.3 Conclusions

Verification of Abbe's diffraction limit – and beyond: The presented article shows that a finite field-of-view (FoV) in XPCI limits the stably achievable resolution r_{stab} in image reconstruction, in accordance with Abbe's diffraction limit associated with the numerical aperture of the setup (compare fig. 4.1): at some point $\mathbf{x} \in K$ within the FoV K , the smallest resolvable feature-size is $r_{\text{stab}}(\mathbf{x}) \approx \pi/(\mathfrak{f} \text{dist}(\mathbf{x}, \partial K))$, where \mathfrak{f} is the modified Fresnel number and $\text{dist}(\mathbf{x}, \partial K)$ the distance of \mathbf{x} to the detector-boundary ∂K . For a square detector $K = [-\frac{1}{2}; \frac{1}{2}]^m$, the highest possible resolution, attained in the center of the FoV, is thus $r_{\text{stab}}^{\max} \approx 2\pi/\mathfrak{f} = 1/\mathfrak{f}$.

However, the performed analysis also reveals valuable details beyond reproducing Abbe's limit: the wave-packet analysis in §4.2.3 indicates that instability of image reconstruction will arise *suddenly and violently* when disregarding the resolution limit $r_{\text{stab}}(\mathbf{x})$: already for

slightly smaller object-scales $r(\mathbf{x}) < (1 - \varepsilon)r_{\text{stab}}(\mathbf{x})$, there exist image-modes that induce quasi no contrast within the detection-domain K and which are thus practically irretrievably lost. Accordingly, the instability associated with the severe ill-posedness of the inverse problem (compare §4.2.2) is so strong that the resolution-limit is practically independent of the data-noise-level (recall that this was one of the open questions from §3.1) – even though uniqueness guarantees full reconstructability in the hypothetical case of exact data.

Another important insight from the analysis is that – at least for the case of *phased* Fresnel-data $\mathcal{D}(h)|_K$ – respecting Abbe’s diffraction limit is not only necessary but also sufficient for stable image reconstruction: if one restricts to images that comply with the resolution-limit, Lipschitz-stability is restored in the finite-FoV-setting according to theorem 4.3.

An unexpected benefit of homogeneity constraints: A surprising result of the analysis is the dependence of the effects of a finite FoV on whether real- or general complex-valued images are to be recovered: as seen from fig. 4.3, resolution degrades much less severely towards the edges of the FoV in the real-valued case. As explained in §4.2.3 and §4.2.6, the effect arises from the (quasi-)symmetric propagation-behavior of real-valued signals upon Fresnel propagation. For XPCI, the result means that image reconstruction under *homogeneity constraints*, where real-valued images are recovered, suffers from less severe resolution-limits than settings where phase and absorption are reconstructed as real- and imaginary parts of a general complex-valued image. Notably, this observation adds yet another difficulty to the latter “general-object-setting”, which is already particularly challenging due to known uniqueness- and stability-issues in the absence of a finite-detector-problem as detailed in chapter 3.

What about multiple holograms? For completeness, we note that the presented results may be readily extended to settings with multiple holograms. However, the additional data is not too useful in view of a finite FoV: from the results in §4.2.3, it can be seen that object-features, which are leaked outside the FoV in the hologram $I_1|_K$ corresponding to the largest Fresnel number $\mathbf{f}_1 > \mathbf{f}_2 > \dots > \mathbf{f}_\ell$, are also invisible in all other holograms $I_2|_K, \dots, I_\ell|_K$.

A (quasi) inevitable resolution-limit: All of the derived resolution estimates depend on the modified Fresnel number \mathbf{f} , where the resolution improves as \mathbf{f} increases. Accordingly, one might try to achieve higher resolutions by adapting the distance between sample and detector in an XPCI-setup. Unfortunately, however, this works only in a *parallel-beam* geometry. In real-world, *divergent-beam* XPCI-setups like the GINIX in fig. 1.2, the sample-to-detector-distance $d_{1,2}$ has to be almost as large as the *source*-to-detector-distance $d_{0,2} \approx d_{1,2}$ in order to obtain a significant geometrical magnification $M = d_{0,2}/(d_{0,2} - d_{1,2}) \gg 1$. As sketched in fig. 4.5, this implies that the maximum detectable angle θ_{max} of X-rays scattered by the sample (measured w.r.t. the optical axis, i.e. from the center of the FoV for simplicity) approximately coincides with the opening angle θ_{source} of the source-detector-ensemble, $\theta_{\text{max}} \approx \theta_{\text{source}}$. Hence, the numerical aperture of the imaging-system which determines the resolution is practically fixed, $\text{NA} = \sin(\theta_{\text{max}}) \approx \sin(\theta_{\text{source}})$. Recalling from §4.1 that θ_{source} is not only limited by the detector-size but also by the finite opening angle of the X-ray beam, we thus conclude:

The resolution of (holographic) XPCI in a strongly magnifying divergent-beam setup is fundamentally limited by the opening angle of the probing X-ray beam.

The fundamental resolution limit is given by $r_{\text{FoV}} \approx \lambda / (2 \sin(\theta_{\text{source}}))$. For typical GINIX-parameters ($\lambda = 0.157$ nm, $\theta_{\text{source}} \approx 0.08$ degrees), this translates to $r_{\text{FoV}} \approx 55$ nanometers.

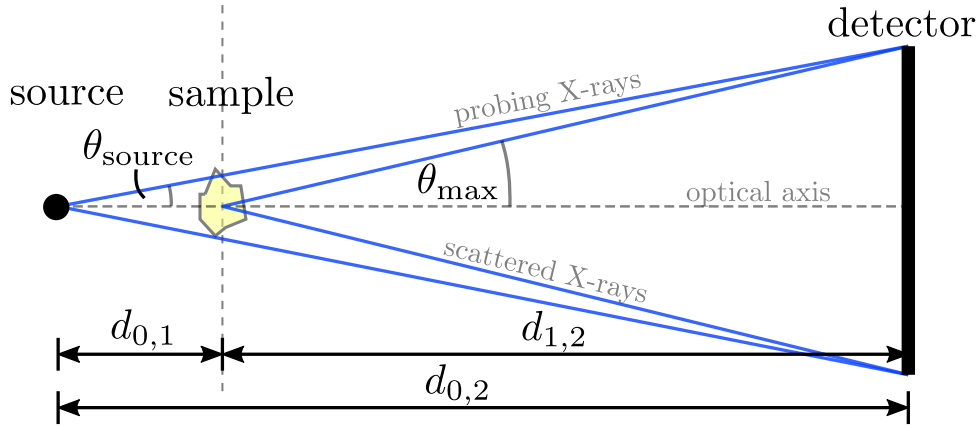


Figure 4.5. Numerical aperture of a divergent-beam XPCI-setup: due to the requirement of a large geometrical magnification $M = d_{0,2}/d_{0,1} \gg 1$, the maximum detectable scattering-angle θ_{max} always satisfies $\theta_{\text{max}} \approx \theta_{\text{source}}$ and is thus practically independent of $d_{0,1}$. Hence, so is the numerical aperture $\text{NA} = \sin(\theta_{\text{max}}) \approx \sin(\theta_{\text{source}})$ and the resulting resolution by Abbe's diffraction limit.

Optimal source-to-sample-distances – as large as possible, as small as necessary:

By combining the findings of the present article with the stability estimates from chapter 3, we may identify an optimal source-to-sample-distance $d_{0,1}$ in terms of resolution and stability of image reconstruction. Recall from §3.6.2 that choosing a large $d_{0,1}$ results in a smaller Fresnel number associated with the object-support and thus in improved stability. However, increasing $d_{0,1}$ also decreases the geometrical magnification M and therefore the effective detector-resolution $r_{\text{det}}^{\text{eff}}$: the pixelation of real-world detectors (see §2.1.5.4) defines some physical resolution r_{det} (possibly larger than the pixel-size Δx_{pix} due to a non-ideal point-spread-function) that limits the achievable object-resolution to $r_{\text{det}}^{\text{eff}} = r_{\text{det}}/M$, as explained in §2.1.5.1.

The idea is now to choose $d_{0,1}$ as the maximal value such that the resolution of the XPCI-setup is not yet bottlenecked by the lack of geometrical magnification, i.e. such that the resolution limit due to a finite FoV $r_{\text{FoV}} \approx \lambda / (2 \sin(\theta_{\text{source}}))$ and the effective detector resolution $r_{\text{det}}^{\text{eff}}$ are approximately balanced:

$$r_{\text{det}}^{\text{eff}} \approx r_{\text{FoV}} \Leftrightarrow d_{0,1} \approx \frac{\lambda d_{0,2}}{2 \sin(\theta_{\text{source}}) r_{\text{det}}} \quad (4.3.1)$$

For the GINIX-setup (see §1.2) with $d_{0,2} = 5$ m, an estimated detector-resolution of $r_{\text{det}} \approx 10$ μm and beam-divergence $\sin(\theta_{\text{source}}) \approx 6 \text{ mm} / d_{0,2} \approx 0.0012$ at photon-energy $E = 7.9$ keV ($\Rightarrow \lambda \approx 0.157$ nm), (4.3.1) gives $d_{0,1} \approx 33$ mm. Notably, this value exceeds the chosen distances $d_{0,1}$ for the experimental data considered in this thesis by factors between 2 and 4. According to the present analysis, the XPCI-experiments could thus have been conducted in a more holographic and thus more stable imaging-regime without deteriorating the resolution.

Open problems: The derived stability estimates for the finite-detector-problem with *phased* Fresnel-data $\mathcal{D}(h)|_K$ are quite sharp. Yet, note that this is not true for the extension to

the *phaseless* XPCI-setting in §4.2.7. The issue is that the approach widely disregards the possibly complicated interactions between instabilities due to a finite FoV on the one hand and those arising from the phase-problem on the other hand. The obtained stability estimates for linearized XPCI ((4.2.20) and further estimates in the article) are combinations of worst-case estimates for both types of instabilities and therefore highly non-optimal, as seen from the examples 9.27 to 9.29 in the complete article. A deeper analysis of the interaction between finite-detector-problem and phase-problem in XPCI is needed to rigorously prove that stable image reconstruction is possible up to Abbe's resolution limit also in the phaseless case.

A second open problem is an analysis of *region-of-interest-XPCI*, i.e. of settings where the imaged sample is too large to fit completely into the field-of-view, as is the case in fig. 1.3(c). In addition to the problem of information-leakage by the delocalizing action of Fresnel propagation considered in the present work, fringes from structures *outside* the FoV may then also propagate *into* the captured part of the hologram. The derived locality-estimates for Fresnel propagation provide a handle to analyze this more complicated problem in future work.

Chapter 5

Regularized Newton methods for X-ray phase contrast and general imaging problems

Bearing in mind the results from the theoretical analysis in Articles 1 and 2, we now proceed with Article 3 [143] to actual reconstruction algorithms for phase contrast imaging and -tomography. As a starting point, regularized Newton methods are proposed for image reconstruction, which may for once take advantage of available a priori constraints and, on the other hand, overcome the limitations of algorithms that rely on the linearized CTF-model of XPCI. In general, while the analysis of the preceding chapters 3 and 4 is mostly limited to the linear forward model, we will drop this restriction in the following algorithmic chapters 5 and 6.

5.1 Motivation

Significance of constraints: A major result of Article 1 (chapter 3) was that support constraints greatly improve stability of image reconstruction in XPCI – without such constraints, the inverse problem may be ill-posed or even non-unique. It has also been argued that imposing non-negativity may have a similar beneficial effect, see §3.7. Finally, the stability analysis of chapter 3 has also confirmed the potential gains of using *tomographic consistency* as a constraint in XPCT, by applying *all-at-once* methods in the sense of §2.7.4. In order to take advantage of these theoretical insights, it is necessary to apply image reconstruction algorithms that may flexibly incorporate such a priori knowledge.

Necessity of iterative methods: The most commonly used image reconstruction method in (holographic) XPCI is direct CTF-inversion. As detailed in §2.7.2.1, the approach boils down to applying quadratic Tikhonov regularization to the linearized XPCI-model, described by the forward map $\mathcal{S}_\nu^{(f_1, \dots, f_\ell)}$ (or $\mathcal{T}^{(f_1, \dots, f_\ell)}$). Owing to linearity and translation-invariance of these maps, i.e. by their *convolutional* structure, the reconstruction problem diagonalizes in Fourier space and thus may be implemented in a fast non-iterative manner using FFTs. Unfortunately, however, direct CTF-inversion is incompatible with the above constraints, compare §2.7.2.1:

- Support constraints $\text{supp}(h) \subset \Omega$ to $\Omega \subsetneq \mathbb{R}^m$ break the translation-invariance in real-space and thus the diagonality in Fourier space, which is required for fast implementation.
- Non-negativity is a *nonlinear* constraint: any algorithm that imposes it necessarily gives rise to a nonlinear dependence on the data – even for linear inverse problems!

In either case, the linear convolutional structure of the reconstruction problem is thus lost. Due to the large numerical problem-sizes encountered in XPCI (images with $n \gtrsim 10^6$ unknowns (pixels) to be reconstructed) and XPCT (3D-volumes with $n \gtrsim 10^9$ unknowns (voxels)), this rules out usage of direct inversion methods. In particular, note that that explicitly precomputing the (non-sparse!) system-matrices of size $n \times n$ is computationally prohibitive – not to speak of *inverting* these. However, *iterative* methods remain applicable since – contrary to a complete reconstruction – the individual iterations may usually still be implemented in an efficient matrix-free manner using FFTs. Hence, we arrive at the following conclusion:

Imposing support- and/or non-negativity constraints in XPCI/XPCT necessarily requires iterative reconstruction methods (to date).

Reconstructions beyond the linearized XPCI-model: In addition to the inability of incorporating certain a priori constraints, it should be emphasized that CTF-inversion is based on the *linearized* forward model of XPCI. Hence, the method is limited to the regime of validity of the underlying assumption of a weakly interacting object and severe artifacts may result when applied beyond this regime, as demonstrated e.g. in [87]. A major motivation to seek alternative reconstruction methods is to overcome these limitations by instead using the full *nonlinear* XPCI-model \mathcal{N}_ν from §2.4.1 or §2.4.3. While a typical argument against such an approach is the lack of known direct inversion formulas for \mathcal{N}_ν , note that this does not play a major role in the present context: after all, iterative schemes are already required to obtain the desired flexibility in imposing constraints according to the previous paragraph. Moreover, *forward* evaluations of a nonlinear map \mathcal{N}_ν (and of its derivatives), as used in iterative schemes, can typically be implemented as efficiently as for the linear models $\mathcal{T}, \mathcal{S}_\nu$. Hence, incorporating the nonlinearity of XPCI in iterative reconstruction methods comes widely “for free”.

Drawbacks of existing iterative methods: The prevailing class of iterative algorithms proposed for XPCI (see e.g. [76, 48, 14, 175, 165, 87]) follow the alternating-projection (AP-) approach detailed in §2.7.2.2: they reconstruct an image $f \in \bigcap_{i=1}^{\ell+k} C_i$ by iteratively projecting (or reflecting) onto constraint-sets defined by the measured holograms C_1, \dots, C_ℓ and additional a priori knowledge $C_{\ell+1}, \dots, C_{\ell+k}$. While such algorithms are relatively easy to implement, readily incorporate nonlinearity of the XPCI-model and offer great flexibility by the choice of the constraint sets C_i , they also have two main drawbacks: Firstly, while real-world measurements are never exact, the approach lacks a natural handle to account for data-errors. In particular, note that noise may cause the intersection $\bigcap_{i=1}^{\ell+k} C_i$ to be empty, in which case the aim of finding an element within becomes obsolete and the algorithmic behavior of AP is hard to predict. Secondly, (semi-)convergence of the algorithms in practice is often observed to be slow for XPCI, typically requiring $\mathcal{O}(10^3 \dots 10^4)$ iterations to (quasi-)stabilize at some solution [83, 87]. The latter effect can be understood by the simple algorithmic structure of AP-schemes: instead of considering the whole problem at once, the methods temporarily “forget” all the other

constraints (homogeneous object, support, non-negativity, ...) and data while projecting (or reflecting) onto a specific set C_i . This promotes inefficient oscillations of the iterates.

Potential benefits of Newton-type methods: The discussion motivates novel reconstruction methods that retain the flexibility of the AP-type algorithms but eliminate their drawbacks. The present Article 3 [143] proposes *iteratively regularized Gauss-Newton methods* (IRGNM) as a candidate for this task. The underlying Newton-like idea lies in computing constrained nonlinear reconstructions in XPCI by iteratively solving local linearizations of the problem. Thereby, the approach exploits differentiability of the XPCI-model (see §2.4.7.3) and accounts for the recurrent observation that the inverse problem of XPCI is only weakly nonlinear, as seen from the relative success of fully linear, CTF-based reconstructions. These properties of the IRGNM, combined with a partly more natural handling of constraints and data-errors, promise faster convergence than AP-methods. It will furthermore be seen that a variant of the IRGNM-approach allows for computationally efficient *all-at-once* reconstructions in XPCT.

5.2 Summary of the results

The article [143] is divided into three principal parts: the first reviews the general theory and implementation of IRGNM as well as of a *Kaczmarz-type* (or block-structured) variant of the approach. In the second part, the IRGNM is applied to XPCI in the two-dimensional, i.e. non-tomographic setting, exploring in particular the feasibility of jointly reconstructing phase- *and* absorption-image from a single hologram. The last part considers Newton-Kaczmarz-iterations as an approach to achieve *all-at-once* reconstructions in phase contrast tomography, i.e. for direct retrieval of the 3D-object density from the tomographic hologram series.

5.2.1 Iteratively regularized Gauss-Newton Method (IRGNM)

For the first part, a general ill-posed inverse problem of the form $g^{\text{obs}} = F(f^\dagger) + \epsilon$ is considered. f^\dagger denotes the unknown object and g^{obs} is the measured (intensity) data, that is related to f via a nonlinear forward operator $F : X \rightarrow Y$ between Hilbert spaces X (object space) and Y (data space), subject to additional data-errors ϵ .

The iteratively regularized Gauss-Newton method (IRGNM) proposed by BAKUSHINSKII [9] accounts for both nonlinearity and ill-posedness by performing iterations of the form

$$f_{k+1} = \underset{f \in X}{\operatorname{argmin}} \left\| F(f_k) + F'[f_k](f - f_k) - g^{\text{obs}} \right\|_Y^2 + \alpha_k \|f - f_0\|_X^2 \quad (5.2.1a)$$

$$= f_k + (F'[f_k]^* F'[f_k] + \alpha_k \operatorname{id}_X)^{-1} (F'[f_k]^* (g^{\text{obs}} - F(f_k)) + \alpha_k (f_0 - f_k)) \quad (5.2.1b)$$

for $k = 0, 1, \dots, k_{\text{stop}} - 1$ with some initial guess f_0 and regularization parameters $\alpha_k > 0$. The “Newton-like” idea involved in (5.2.1a) lies in replacing the nonlinear forward map F by its best local linear approximation $F(f) = F(f_k) + F'[f_k](f - f_k) + o(\|f - f_k\|)$, where the linear operator $F'[f_k]$ denotes the Fréchet-derivative (i.e. the Jacobian) of F at f_k .

As seen from the formula (5.2.1b), a benefit of this linearization is that the update can be evaluated relatively easily by a *forward*-evaluation of the nonlinear operator F , followed by inverting the selfadjoint positive-definite *linear* operator $F'[f_k]^* F'[f_k] + \alpha_k \operatorname{id}_X$ ($F'[f_k]^*$: adjoint

of $F'[f_k]$, id_X : identity map on X). The latter task can be achieved by generic iterative solvers such as conjugate-gradient (CG-) methods. An advantage of CG and related algorithms is that they can be implemented in a *matrix-free* manner: no matrix-representation of $F'[f_k]^*F'[f_k] + \alpha_k \text{id}_X$ is required (which would be computationally prohibitive for XPCI- and XPCCT-problems), but only (discretized) implementations of the evaluations $h \mapsto F'[f_k]h$ and $g \mapsto F'[f_k]^*g$.

Regularization property: The parameters $\alpha_k > 0$ ensure that the inverse of $F'[f_k]^*F'[f_k] + \alpha_k \text{id}_X$ is bounded so that the data-errors ϵ enter in the object-iterates f_{k+1} at most amplified by a factor $1/\alpha_k$ in each step. Accordingly, the IRGNM scheme is *regularizing*. It is indeed a regularization method in a technical sense if the nonlinearity of F is sufficiently well-behaved, locally converging to the true solution for vanishing data-errors $\epsilon \rightarrow 0$ given a suitable choice of the α_k and of the stop-index $k_{\text{stop}} \in \mathbb{N}$. A detailed convergence analysis of the IRGNM can be found e.g. in [9, 24, 97, 112]. In the present work, a geometric decrease $\alpha_{k+1} = \frac{2}{3}\alpha_k$ is used, where α_0 and k_{stop} are determined automatically via heuristics detailed in the article.

5.2.1.1 Choice of the norms

Data-space norm – the data-fidelity term: The choice of the norm $\|\cdot\|_Y$ in (5.2.1) allows to account for the expected statistics of the data-errors ϵ . Although noise in holograms is primarily due to *Poisson-noise*, see §2.1.3.3, the flat-field correction (§2.1.5.2) and systematic deviations from the ideal model give rise to a total error-statistics in XPCI that is hard to predict in general. We therefore resort to the standard choice of an L^2 -norm, i.e. $\|g\|_Y^2 := \int |g(x)|^2 dx$.

For completeness, we note that a quadratic *data-fidelity* term $\|g - g^{\text{obs}}\|_Y^2$, as employed in (5.2.1a), is not always the best choice to measure the proximity between the fitted data $g = F(f)$ and the observations g^{obs} . For example in the presence of large outliers in the data or for ϵ given by pure Poisson-noise, probabilistic models suggest that better reconstructions can be achieved if $\|g - g^{\text{obs}}\|_Y^2$ in (5.2.1a) is replaced by certain *non-quadratic* data-fidelity terms $\mathcal{S}(g^{\text{obs}}; g)$. Such generalizations of the IRGNM have been proposed e.g. in [98, 100, 99]. In a non-quadratic setting, however, the minimization problem in (5.2.1a) typically becomes harder to solve, leading to increased computational costs of the IRGNM-iterations. For this reason and due to the lack of an obvious statistical error model for XPCI, refinements of the chosen L^2 -data-fidelity are not considered here. Yet, the idea will be taken up again in chapter 6, where methods are presented that widely avoid a loss in computational performance.

Object-space norm – the regularization term: By the choice of the norm $\|\cdot\|_X$, desirable properties of the object-iterates f_k may be imposed via the regularization term $\alpha_k \|f - f_0\|_X^2$ in (5.2.1a) that penalizes deviations from the initial guess f_0 (typically: $f_0 = 0$). In the article, Sobolev-norms are considered, $\|f\|_X := \|(1 + |\xi|^2)^{s/2} \cdot \mathcal{F}(f)\|$, i.e. weighted L^2 -norms in Fourier space. The larger the parameter $s \geq 0$, the stronger the penalization in high spatial frequencies compared to low ones, which increases the smoothness of the reconstructed object $f_{k_{\text{stop}}}$ (or more generally of $f_{k_{\text{stop}}} - f_0$). The case $s = 0$ corresponds to the standard L^2 -norm.

Similarly as for the data-fidelity, the squared norm $\|f - f_0\|_X^2$ in (5.2.1a) may be replaced by more general *penalty functionals* $\mathcal{R}(f)$, increasing flexibility in imposing object-properties at the cost of a more complicated regularization-theory and often higher computational costs.

5.2.1.2 Incorporating constraints

Linear constraints: A major advantage of the IRGNM over the alternating-projection-type methods discussed in §5.1 is that it provides a very natural handle to incorporate *linear* constraints on the reconstructed object: provided that the set $C = \{f \in X : f \text{ satisfies constraints}\}$ is a closed linear subspace of the object-space X , the constraints are imposed *automatically* by the IRGNM-iterations (5.2.1a) if the full object-space X is replaced by C (which is again a Hilbert space) and the restriction $F|_C : C \rightarrow Y$ is employed as a forward operator. Implementation-wise, this does not change the computation of $F(f_k)$ or $F'[f_k]$, yet that of the adjoint $F'[f_k]^*$: it holds that $F|_C(f) = F(f)$ and $(F|_C)'[f]h = F'[f]h$ but $((F|_C)')^*[f]g = \mathcal{P}_C F'[f]^*g$, where $\mathcal{P} : X \rightarrow C$ is the orthogonal projection onto C .

As seen in §2.4.6, both *support-* and *homogeneity constraints* constitute linear constraints in XPCI in the above sense, so that these may be readily incorporated in the IRGNM.

Non-negativity constraints: In order to also impose *non-negativity* of the recovered object (in real- and/or imaginary part in the complex-valued case), which is a *nonlinear* constraint, a penalty-method is proposed in the article. For real-valued f , it amounts to appending the term $\mathcal{R}_{\geq 0}(f, f_k) := \gamma \|\max(0, -\text{sign}(f_k)) \cdot f\|_{L^2}^2$ ($\gamma > 0$) to the objective functional on the r.h.s. of (5.2.1a). This corresponds to an additional L^2 -penalty that selectively damps values of the next iterate $f_{k+1}(\mathbf{x})$ on those points \mathbf{x} where the current one is negative, $f_k(\mathbf{x}) < 0$. Thereby, negative values of the f_k are either corrected in the course of the IRGNM-iterations or damped to values close to zero, so that the final iterate $f_{k_{\text{stop}}}$ is approximately non-negative.

While the proposed penalty-method is somewhat heuristic, it is found to perform well for image reconstruction in XPCI and XPCT and it is favorable in terms of computational costs: contrary to more sophisticated approaches to imposing non-negativity such as semi-smooth Newton-methods [94], primal-dual- [36] or other splitting methods [70, 18], adding the penalty $\mathcal{R}_{\geq 0}(f, f_k)$ retains the simple quadratic structure of the optimization problem in (5.2.1a). Hence, the IRGNM-iterates remain efficiently computable via CG-methods.

5.2.2 IRGNM applied to XPCI

The general IRGNM-scheme is applied to the setting of XPCI by employing the principal forward map $F := \mathcal{N} : f \mapsto |\mathcal{D}(\exp(-f))|^2 - 1$ from §2.4.1, where $f = \mu + i\phi$ gives the sought phase- ϕ and absorption-images μ . The data g^{obs} is then given by the contrast in the measured hologram, $g^{\text{obs}} = I^{\text{obs}} - 1$, and the Fréchet-derivative of F can be evaluated via the formula

$$F'[f]h = -2\text{Re}(\overline{\mathcal{D}(\exp(-f))} \cdot \mathcal{D}(\exp(-f)h)), \quad (5.2.2)$$

as seen in §2.4.7.3. For the regularization-term in (5.2.1a), a Sobolev-norm with $s = \frac{1}{2}$ is used (see §5.2.1.1), as this is found to yield a reasonable compromise between noise-level and sharpness of the reconstructed images. The resulting IRGNM is discretized according to the general approach from §2.6.2, where the expression in (5.2.2) for the derivative carries over to the discrete setting. IRGNM for XPCI-reconstructions from multiple holograms and/or under homogeneity constraints may be constructed completely analogously by employing the associated forward operators from §2.4.3 and §2.4.4, setting $F \in \{\mathcal{N}^{(f_1, \dots, f_\ell)}, \mathcal{N}_\nu^{(f_1, \dots, f_\ell)}\}$.

The IRGNM-scheme is applied to experimental XPCI data, given by a single hologram of size 1920×1080 pixels (fig. 5.1(a)) acquired in the GINIX-setup from §1.2 at Fresnel number $f_{\text{pix}} \approx 1.77 \cdot 10^{-4}$. The object is a logo-shaped grating milled into a 200 nm-thick gold layer. Being composed of a single material, it is homogeneous with absorption-refraction-ratio $\beta/\delta = \mu/\phi = c_{\beta/\delta} \approx 0.105$ (gold at photon-energy 7.9 keV). The test case is named “*Logo-XPCI*”.

For validation, the images ϕ, μ are first reconstructed imposing homogeneity $\mu/\phi = c_{\beta/\delta}$ (but no other constraints) with the IRGNM-algorithm and with direct CTF-inversion, see §2.7.2.1. Both methods yield widely identical results in this setting. This is to be expected as the imaged object is rather weak and so the nonlinear capabilities of the IRGNM do not play a major role. Indeed, in the weak-object regime and without support- and non-negativity constraints, the action of IRGNM should reduce to CTF-inversion, which is confirmed by the results.

To challenge the capabilities of the IRGNM, a second reconstruction is computed *without* imposing a homogeneity constraint, so that phase-image ϕ and absorption-image μ have to be recovered as independent variables. As discussed in §2.5.1.2 and §3.1, such a reconstruction is severely non-unique in general and commonly considered as totally infeasible. However, the results from Article 1, notably theorem 3.1, show that support constraints enable stable image recovery – at least theoretically. Other than in direct CTF-inversion, such constraints may be readily incorporated in the IRGNM, as detailed in §5.2.1.2. The imposed circular support is marked by the blue-dashed lines in the plots of reconstructed images ϕ, μ in fig. 5.1(b),(c).

Despite the challenging setting, the structure of the logo-shaped grating is found to be well-recovered both in the recovered phase ϕ and in the absorption μ up to some low-frequency mode that causes distortions of the quantitative values especially in μ : the insets in fig. 5.1(b),(c) show artificial bright and dark spot in the center. Notably, these artifacts violate the physically motivated *non-positivity** of ϕ and μ for the considered sample. As a consequence, it is found that the low-frequency artifacts may be suppressed by imposing this as a constraint via the proposed penalty-approach, as revealed by the non-inset parts of fig. 5.1(b),(c).

Due to the homogeneity of the object, the real-data example is arguably not ideal to assess the capability of reconstructing phase and absorption independently. In the article, the reconstruction has therefore been repeated for simulated data in the same setting, except that the logo-structure is only present in the absorption-image μ , whereas the phase-image ϕ is just a homogeneous disc. It is found that the IRGNM-algorithm is indeed capable of correctly recovering the logo in μ , while keeping the phase-image free of traces of this structure, i.e. achieves truly independent reconstructions of ϕ and μ .

5.2.3 Regularizing Newton-Kaczmarz for XPCT

The IRGNM-scheme could be readily adapted also to phase contrast *tomography* (XPCT) by employing the associated forward map from §2.4.5, $G : f \mapsto (F(\mathcal{P}_{\theta_1}(f)), \dots, F(\mathcal{P}_{\theta_N}(f)))$. Here, $f = k\beta + ik\delta$ is the scaled contrast in refractive index, $\theta_1, \dots, \theta_N \in \mathbb{S}^2$ denote the tomographic incident directions and $F \in \{\mathcal{N}^{(f_1, \dots, f_\ell)}, \mathcal{N}_\nu^{(f_1, \dots, f_\ell)}\}$ is underlying XPCI-model (one or multiple holograms, with or without homogeneity constraint). However, evaluations of the

*Here, the contrast in ϕ and μ is *negative* as the logo-structure corresponds to *removed* material in a gold-layer of uniform thickness. Imposing non-positivity as a constraint works completely analogously as non-negativity.

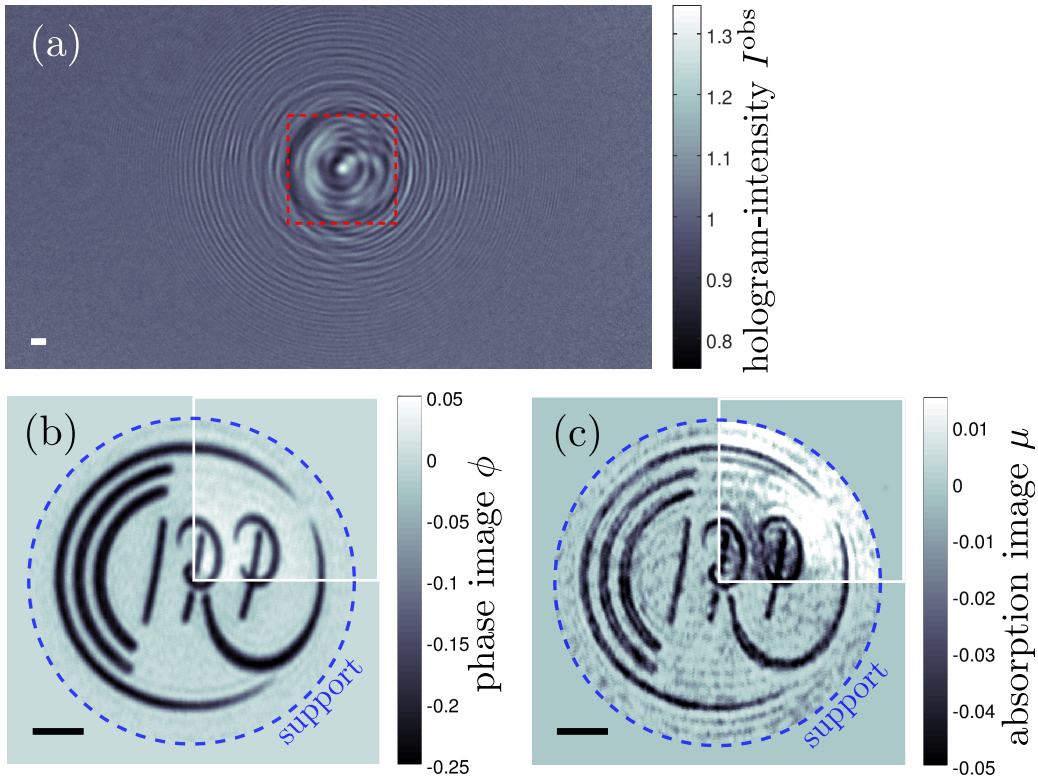


Figure 5.1. IRGNM reconstruction of the “Logo-XPCI” data set. Phase ϕ and absorption μ are recovered as independent images. Scale bars: $1 \mu\text{m}$.

(a) Hologram data. (b) Reconstructed phase-image ϕ . (c) Reconstructed absorption-image μ .

The red-dashed box in (a) marks the region shown in (b) and (c). A circular support (blue-dashed region in (b) and (c)) and values $\phi, \mu \leq 0$ have been imposed as constraints. Insets in the top-right corners of (b) and (c) show the respective parts for reconstructions *without* imposing $\phi, \mu \leq 0$.

map G and of its derivative are computationally expensive, which is why a Kaczmarz-type variant of the IRGNM, termed (regularizing) Newton-Kaczmarz [30], is proposed for XPCT.

Analogously to classical Kaczmarz-iterations for linear systems [109], the idea is to decompose the problem into smaller sub-problems and to update the iterates f_k by fitting a single sub-data at a time. For XPCT, it is natural to decompose the hologram-data $I_{\text{PCT}}^{\text{obs}} = (I_1^{\text{obs}}, \dots, I_N^{\text{obs}})$ into blocks corresponding to the different incident directions θ_j . If the forward map is decomposed accordingly, $G = (G_1, \dots, G_N)$, Newton-Kaczmarz iterations are of the form

$$f_{k+1} = \underset{f \in X}{\operatorname{argmin}} \left\| G_{j_k}(f_k) + G'_{j_k}[f_k](f - f_k) - g_{j_k}^{\text{obs}} \right\|_{L^2}^2 + \alpha_{k,1} \|f - f_k\|_X^2 + \alpha_{k,2} \|f - f_0\|_X^2 \quad (5.2.3)$$

with $g_{j_k}^{\text{obs}} := I_{j_k}^{\text{obs}} - 1$. The updates are thus given by IRGNM-steps on the sub-problems $G_{j_k}(f) \approx g_{j_k}^{\text{obs}}$ with an additional regularization term $\alpha_{k,1} \|f - f_k\|_X^2$ that penalizes deviations from the previous iterate f_k . The latter is vital because otherwise the algorithm could completely “forget” all the data fitted before, just to fit the current block as well as possible. Note that we employ a simple L^2 -norm for $\|\cdot\|_X$, for the simple reason that H^s -norms with $s > 0$ would require that FFTs and inverse FFTs of large 3D-arrays are computed in each iteration, which would severely increase the computational costs of the algorithm. The indices j_1, j_2, \dots in (5.2.3) determine the processing-order of the data-blocks. In the present work, a randomized order is chosen.

The Newton-Kaczmarz approach is applied to experimental XPCT data for a nano-crystal formed by polystyrene-colloids of diameter 400 nm. The data set is composed of a tomographic series of 249 holograms of size 1024×1024 pixels each, acquired for incident angles between 0 and 173° at a single Fresnel number $f_{\text{pix}} \approx 2.41 \cdot 10^{-4}$ (i.e. one hologram per incident angle). We term this test case “*Colloid-XPCT*”. Negligible absorption $\beta = 0$ and non-negativity of the refractive decrement $\delta \geq 0$ are imposed as constraints as well as a cube-shaped support of size 256^3 voxels, motivated by the small object-size inferred from the hologram-data in fig. 5.2(a). The Newton-Kaczmarz reconstruction is run with constant regularization-parameters $\alpha = \alpha_{k,1} = 10^3 \alpha_{k,2}$ (where α is chosen according to a heuristic detailed in the article) over two *cycles*, i.e. such that each hologram $I_{j_k}^{\text{obs}}$ is fitted exactly twice by iterations of the form (5.2.3).

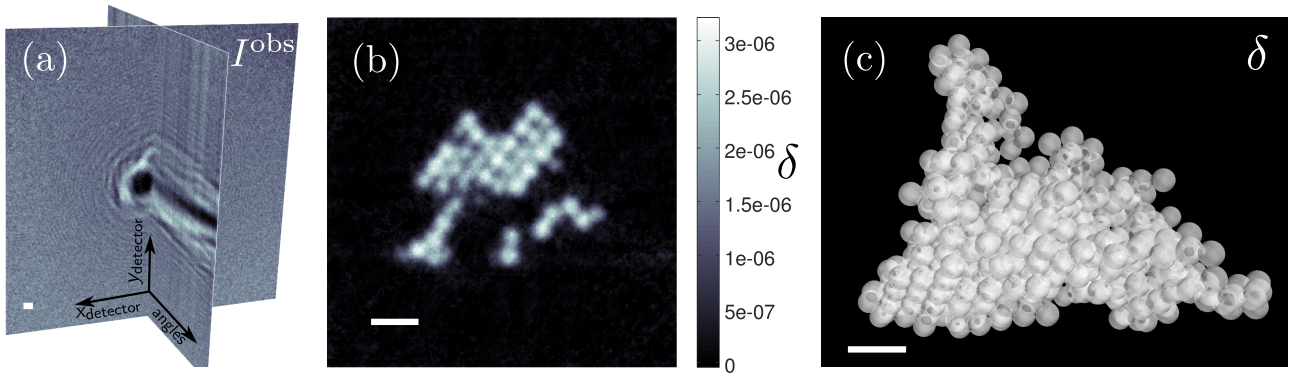


Figure 5.2. Newton-Kaczmarz reconstruction of the “Colloid-XPCT” data set. Scale bars: $1 \mu\text{m}$. (a) Orthoslices through the 3D tomographic data set (249 holograms of size 1024×1024 pixels). (b) Central orthoslice (perpendicular to the tomographic axis) through the reconstructed 3D-volume. (c) 3D-Rendering of the colloidal crystal based on the reconstructed refractive decrement δ .

A 2D-slice of the recovered 3D-structure of the refractive decrement δ is shown in fig. 5.2(b). The reconstruction is found to be in good agreement with the expectations for the imaged sample, essentially showing a binary object (constant value $\delta > 0$ within the polystyrene-colloids, $\delta = 0$ outside) up to noise and a slight blur. By computing the Fourier-shell-correlation [201], the resolution of the reconstruction is estimated as $\approx 95 \text{ nm}$, compared to $\approx 105 \text{ nm}$ for a second reconstruction without non-negativity constraint (not shown). Thus, non-negativity slightly enhances the resolution, presumably because the constraint suppresses noise in the object owing to its stabilizing effect (compare §3.5) and promotes a sharper and cleaner separation of the colloids from the background. In the article, it is furthermore demonstrated that the achieved reconstruction-quality is sufficient to accurately determine the colloid-sites, i.e. the crystal-structure, the result of which is shown as a 3D-rendering in fig. 5.2(c).

5.3 Supplement: comparison to theoretical predictions

The stability results from chapters 3 and 4 permit theoretical predictions on the achievable image quality that can be related to the practical reconstructions obtained in the presented article. This is done in the following supplement.

5.3.1 Comparison to stability theory from Article 1

As both the logo-structure and the colloidal crystal are *weak objects* to good approximation, the stability theory for *linearized* XPCI from chapter 3 applies to the considered settings.

Logo-XPCI: In the 2D test case, a circular support of diameter 280 pixels has been imposed. The modified Fresnel number associated with this lengthscale is $f_{\text{supp}} = 2\pi \cdot 280^2 f_{\text{pix}} \approx 87$. As phase-shifts and absorption are reconstructed as *independent* images from a single hologram, the test case corresponds to the most challenging XPCI-setting, stated as inverse problem 3.1. Accordingly, the stability constant $C_{\text{stab}} \approx C_{\text{IP1}}(B[0; \frac{1}{2}], f_{\text{supp}})$ may be estimated via the stability bound (3.2.3) from theorem 3.1 (approximating the circular support by an enclosing square). This yields that the least stable image-mode h_0 induces a relative contrast of

$$\|\mathcal{I}(h_0)\|/\|h_0\| = C_{\text{stab}} \geq 1.26 \cdot 10^{-4} + \mathcal{O}(10^{-8}). \quad (5.3.1)$$

This contrast is on the threshold of being detectable but arguably not sufficient to enable a stable reconstruction in practice. However, the more in-depth modal stability theory from Article 1, see §8.3.4, also predicts that *only a single mode* exhibits such a low contrast whereas the two next-to-least-stable modes induce a relative contrast of at least $> 7 \cdot 10^{-4}$. Accordingly, the contrast increases quite fast with the order of the modes so that only very few are unstable to reconstruct. Moreover, it is known from the stability analysis that these are *low-frequency* modes. The obtained numerical results are in agreement with this theoretical picture: in the reconstruction with support constraint only, shown in the insets in fig. 5.1(b),(c), low-frequency instability manifest as slowly varying perturbations of the background, whereas high-frequency structures such as edges are accurately recovered. By imposing the correct sign of the images ϕ and μ as an additional constraint, which effectively tightens the support constraint and thereby improves stability according to §3.5, the remaining low-frequency artifacts are suppressed.

Colloid-XPCT: For the XPCT test case, the recovered object has been restricted to a cube of size 256^3 voxels. This corresponds to a modified Fresnel number $f_{\text{supp}} \approx 2\pi \cdot 256^2 \cdot f_{\text{pix}} \approx 99$ of the cubic support. Moreover, a pure phase object ($\beta = 0$) has been assumed in the reconstruction. According to §3.3, the XPCT-setting thus corresponds to inverse problem 3.2 with $\nu = 0$ and $m = 3$ (neglecting the finite sampling in the tomographic angles). Using the numerical methods from §3.4, the corresponding stability constant is computed to be

$$C_{\text{stab}} = C_{\text{IP2}}([- \frac{1}{2}; \frac{1}{2}]^3, f_{\text{supp}}) \approx 0.32. \quad (5.3.2)$$

As the *maximal* relative contrast of an image in linearized XPCI is 2 (by theorem 2.3), this result means that even the least stable mode yields strong contrast in the XPCT-data. According to the stability theory from §3.3, the imaging-problem is thus practically as stable as a reconstruction from pure tomographic projection-data – no additional instability arises from the phase contrast setting. In particular, this means that only little gain was to be expected if holograms had been acquired at more than one Fresnel number for the considered data set.

The predicted high stability explains the achieved high reconstruction quality in fig. 5.2 (no low-frequency halos or residual fringe-artifacts can be identified in the plotted slice) despite the relatively high noise-level in the data.

5.3.2 Comparison to resolution estimates from Article 2

As seen in chapter 4, the limited field-of-view (FoV), for which the considered hologram-data has been recorded, gives rise to a fundamental resolution limit: if f_{FoV} denotes the Fresnel number associated with the detector's aspect length w (*effective* length in divergent-beam geometries, compare §2.1.5.1), image-features of size $< w/f_{\text{FoV}}$ cannot be resolved. The reconstruction results allow to compare the practically achieved resolution to this theoretical limit.

Logo-XPCI: In the logo-experiment, the detector was rectangular with 1920 pixels of effective size $\Delta x_{\text{pix}}^{\text{eff}} \approx 22$ nm horizontally and 1080 pixels along the vertical direction. This corresponds to effective aspect lengths of $w_x = 1920 \cdot 22$ nm ≈ 42 μm (horizontally) and $w_y \approx 24$ μm (vertically) and Fresnel numbers $f_{\text{FoV},x} = 1920^2 \cdot f_{\text{pix}} \approx 652$, $f_{\text{FoV},y} = 1080^2 \cdot f_{\text{pix}} \approx 206$. This anisotropic setting may be confusing: which aspect length or Fresnel number is now the one that determines the theoretical resolution? According to the analysis in Article 2, the answer is *both*: it is seen from the Gaussian-wave-packet study in §4.2.3 that that aspect length along an axis determines the resolution w.r.t. features varying exactly along this specific direction! Consequently, a significantly *anisotropic* resolution is to be expected, where the minimal resolvable feature sizes are $r_x = w_x/f_{\text{FoV},x} \approx 65$ nm horizontally and $r_y = w_y/f_{\text{FoV},y} \approx 116$ nm vertically. Indeed, this anisotropy can be identified in the obtained reconstructions: upon a closer look, the logo-edges in fig. 5.1(b) turn out to be slightly more smeared out vertically than horizontally. Moreover, the overall empirical resolution of ≈ 100 nm estimated via line-scans (not shown) is in good quantitative agreement with the theoretical predictions.

Colloid-XPCT: In the tomographic test case, the field-of-view is square with 1024 pixels of effective size $\Delta x_{\text{pix}}^{\text{eff}} \approx 29$ nm along each axis, resulting in an effective aspect length of $w \approx 30$ μm with associated Fresnel number $f_{\text{FoV}} = 1024^2 \cdot f_{\text{pix}} \approx 253$. The predicted resolution is thus $r = w/f_{\text{FoV}} \approx 119$ nm. Again, this is in good agreement with the empirical estimate ≈ 95 nm obtained via the FSC-method. The surprising observation that the achieved resolution in practice is slightly *better* than the theoretical limit may be explained by the nonlinear sharpening effect of the non-negativity constraint, compare §5.2.3, and again by anisotropy: while the theoretical prediction $r = w/f_{\text{FoV}}$ from Article 2 limits the *worst-case* resolution along all possible orientations of object-features, the FSC-estimate can be seen to provide a directionally *averaged* resolution in some sense.

5.4 Conclusions

As usual, we conclude the chapter by considering the findings of the presented article from a broader perspective, outlining benefits, drawbacks and possible improvements of the proposed Newton-type image reconstruction algorithms.

XPCI pushed to its theoretical limits – by exploiting a priori constraints: In the presented article, image reconstruction has been successfully demonstrated for both simulated and experimental data in what is considered one of the most challenging setting of XPCI: retrieving phase ϕ and absorption μ as *two* independent images from *one* hologram. According

to the analysis in chapter 3 and the discussion in §5.3.1, such a reconstruction is ambiguous in general – only the ability of the proposed IRGNM-algorithm to impose support constraints and non-negativity allows to stably disentangle ϕ and μ . In particular, this means that the obtained results are beyond reach for direct CTF-inversion methods as introduced in §2.7.2.1 (notably, the current standard in holographic XPCI!), due to their aforementioned inability of incorporating such a priori constraints, compare §5.1.

The comparison to the theoretical results from chapters 3 and 4 drawn in the supplement §5.3 indeed shows that the achieved reconstructions attain the theoretical limits of XPCI: the empirically found resolution reaches the predicted diffraction-limit arising from the finite detector-size in both the Logo-XPCI and the Colloid-XPCT test cases. Moreover, it is found that the independent reconstruction of ϕ and μ for the Logo-XPCI data set is at the threshold of what is feasible according to stability theory. This demonstrates the potential of the proposed regularized Newton-methods in making the most out of available hologram-data.

Newton-Kaczmarz for XPCT imposes tomographic consistency: We emphasize that the proposed Newton-Kaczmarz method for XPCT is a *joint* or *all-at-once* approach in the sense of §2.7.4: the 3D-object is directly recovered from the tomographic hologram series so that tomographic consistency is implicitly exploited as a constraint in the phase reconstruction of the holograms. In particular, note that the algorithm allows to set *3D*-support constraints, which promotes stability of image reconstruction according to the discussion in §3.7.

Flexibility w.r.t. model-refinements: Both IRGNM and Newton-Kaczmarz are *generic* reconstruction methods in the sense that they do not rely on details of the forward operator (except for differentiability). Consequently, the schemes may be readily adapted to modified XPCI-models, for instance accounting for a non-trivial profile of the probing X-ray beam (§2.1.5.2) or partial coherence (§2.1.5.3). Additionally, we recall from 5.2.1.1 that the proposed algorithms may be adapted to the expected statistics of the data-errors. In the long run, the flexibility of regularized Newton methods may thus enable improvements of the reconstruction-quality in XPCI and XPCT by incorporating such refined models of image-formation.

Nonlinear capabilities remain to be validated: While the versatility of the Newton-type approach in imposing constraints has been thoroughly exploited in the considered examples, this is *not* true for its ability to account for nonlinearity: though the proposed algorithms are based on the *nonlinear* XPCI-model in principle (as opposed to CTF-based methods, which are inherently linear), the objects in the two test cases are too weakly scattering for nonlinearities to have a significant impact. However, real-data XPCI-examples where IRGNM produce good results, whereas linear image reconstruction completely fails, do exist [87]. Moreover, an XPCT-data set that truly requires the nonlinear capabilities of the Newton-Kaczmarz approach will be reconstructed at the end of next chapter in §6.4.

Computational performance issues: Although efficiency-considerations have been taken into account in the design of the proposed algorithms, computational costs remain a significant downside of the IRGNM: typically, five to ten Newton-iterations of the form (5.2.1a) have to be computed, each of which involves an iterative solution of a quadratic minimization problem, which usually requires $\mathcal{O}(10)$ conjugate-gradient (CG-) iterations. Notably, each of the latter

inner iterations amounts to about as much computational effort as a complete reconstruction by direct CTF-inversion. Accordingly, typical computation times for an IRGNM-reconstruction (*minutes* for holograms with $\mathcal{O}(10^6)$ pixels on a single CPU) are in the order of 100-times longer than for standard CTF-methods (usually in the order of *one second*). Though the required number of iterations in IRGNM is typically smaller than for alternating-projection-type schemes (see §5.1), it thus needs to be further reduced to make the approach suitable for high-throughput imaging at synchrotrons. For example, this might be achievable by suitable preconditioning, e.g. based on multi-grid methods [29] as in solvers for partial differential equations.

Concerning the required number of iterations, the situation is better for the proposed Newton-Kaczmarz method for XPCT: as demonstrated in §5.2.3, already two cycles over the hologram series, i.e. effectively two Newton-steps on the total problem, may be sufficient to achieve a high-quality reconstruction. Moreover, the strong regularization w.r.t. the preceding iterate f_k in the Newton-Kaczmarz updates (5.2.3) renders the computation *well-conditioned*, so that only few (typically ≤ 5) CG-iterations have to be performed per Kaczmarz-step. This makes the algorithm computationally more efficient than a simple adaptation of the (bulk, non-Kaczmarz-type) IRGNM to the XPCT-setting as implemented in [139].

However, the Newton-Kaczmarz iterations in XPCT are much more expensive to compute than the IRGNM-steps in XPCI as the former operate on $3D$ -objects instead of $2D$ -images and involve costly evaluations of the tomographic projectors \mathcal{P}_{θ_j} . This raises performance issues that widely rule out application of the proposed method in its present form to large-scale XPCT-problems (note that the numerical problem size for the reconstruction in §5.2.3 with 256^3 dimensions in discrete object-space and $249 \cdot 1024^2$ in data-space was still comparably small). In the subsequent chapter 6, a general algorithmic scheme for the computation of (Newton-)Kaczmarz iterations will be presented that helps to overcome these limitations.

Chapter 6

Generalized SART-Methods for Tomographic Imaging

In Article 3, Newton-Kaczmarz iterations have been proposed and successfully tested as a reconstruction method for XPCT. Article 4 [141] takes this idea up by deriving efficient formulas, termed *generalized SART* (or *GenSART*-) schemes, for the computation of such- and much more general Kaczmarz-iterations. As the required theory will become technical in parts, it is important to bear in mind that the principal goal is a purely practical, computational one: reduce the computational costs of (Newton-)Kaczmarz methods in order to make them applicable for large-scale tomographic reconstruction problems as encountered in XPCT.

6.1 Motivation

6.1.1 The problem of complexity in tomographic reconstruction

A recurrent practical problem in reconstruction algorithms for XPCT and other tomographic imaging modalities lies in the high dimensionality: as the unknown object needs to be sampled in *three* spatial dimensions, the number of degrees-of-freedom in the discretized problems to be solved typically scales like N^3 with the number of sampling points N along each dimension. In a high-resolution imaging setting with $N = 2048$, this third-power-dependence implies that one real-valued object already occupies 32 or 64 Gigabytes of memory in single or double floating-point-precision, respectively. This makes efficient memory-management a highly relevant issue for tomographic reconstruction algorithms – even when run on current workstations.

For large-scale XPCT problems, it is standard to reconstruct via a two-step approach: firstly, 2D-projections are reconstructed from the acquired holograms for each tomographic incident angle individually (by CTF-inversion, §2.7.2.1), followed by filtered back-projection (FBP, §2.7.3.1) applied to the recovered projections. Note that both steps only need to process 2D-data as the FBP-step can be performed slice-by-slice. This is highly efficient in terms of memory-requirements. On the contrary, any *all-at-once* reconstruction approach (see §2.7.4), like the Newton-Kaczmarz scheme from Article 3, combines these two steps and thereby establishes a *fully three-dimensional* object-data-dependence. Indeed, the whole idea of such methods is to exploit three-dimensionality as a constraint to stabilize phase reconstruction. Therefore, it cannot be avoided that they manipulate large 3D-arrays in some non-trivial manner.

In addition to memory-issues, the large data-dimensions bring about considerable computational costs, in particular for the evaluation of the tomographic (back-)projectors $\mathcal{P}_\theta, \mathcal{P}_\theta^*$, see §2.3: for standard discretizations*, numerically computing $\mathcal{P}_\theta(f)$ or $\mathcal{P}_\theta^*(p)$ for *one* incident angle θ amounts to $\mathcal{O}(N^3)$ floating point operations (FLOPs). Hence, a total (back-)projection $\mathcal{P}_{\text{tot}} := (\mathcal{P}_{\theta_1}, \dots, \mathcal{P}_{\theta_{N_{\text{proj}}}})$ or $\mathcal{P}_{\text{tot}}^*$ for *all* angles requires $\mathcal{O}(N_{\text{proj}}N^3)$ FLOPs. As typically $N_{\text{proj}} \sim N$ holds true, this means that $\geq \mathcal{O}(10^{12})$ operations have to be performed for problem-sizes $N \sim 1000$, which amounts to computation times in the order of *hours* on a single CPU. A large number of evaluations of \mathcal{P}_{tot} and/or $\mathcal{P}_{\text{tot}}^*$ are thus computationally prohibitive.

In view of computational costs and memory-requirements, we see that efficient tomographic reconstruction algorithms should meet the following criteria:

- (1) Minimize the number of evaluations of tomographic (back-)projectors $\mathcal{P}_{\theta_j}, \mathcal{P}_{\theta_j}^*$.
- (2) Work as little as possible on 3D-arrays and store as few as possible copies of these.
- (3) Parallelizability: allow subdivision into small tasks that can be performed *in parallel* by multiple processors without requiring too much communication between these.

It can be readily seen that standard two-step CTF+FBP-reconstructions in XPCT satisfy all of these performance-criteria almost perfectly. Yet, this approach does not bring the same capabilities as the Newton-Kaczmarz method from Article 3 in terms of nonlinearity, constraints, etc, as discussed in §5.4. On the other hand, the generalized SART-schemes proposed in the present work will enable Kaczmarz-type methods to also meet the performance-requirements. Special emphasis will be laid on criterion (1). However, we note that GenSART-methods also satisfy the criteria (2) and (3), as will be detailed in §6.3.

6.1.2 Classical Kaczmarz-iterations and SART

As a motivating example, we consider the simplest case of *classical* Kaczmarz-iterations for reconstructing an object f from pure projection-data $p^{\text{obs}} \approx \mathcal{P}_{\text{tot}}(f) = (\mathcal{P}_1(f), \dots, \mathcal{P}_{N_{\text{proj}}}(f))$, where the $\mathcal{P}_j := \mathcal{P}_{\theta_j}$ are parallel-beam projectors as defined in §2.3. The idea is to cyclically fit the object-iterates f_k to the tomographic data for one incident direction $\theta_j \in \mathbb{S}^2$ at a time:

$$f_{k+1} = \underset{f \in B_k}{\operatorname{argmin}} \|f - f_k\|^2 \quad \text{with} \quad B_k := \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \|\mathcal{P}_{j_k}(f) - p_{j_k}^{\text{obs}}\|^2 \quad (6.1.1)$$

Importantly, although the optimization problem in (6.1.1) looks complicated, it has a relatively simple analytic solution, as can be seen e.g. from the analysis in [155, §V.4.3]:

$$f_{k+1} = f_k + \mathcal{P}_{j_k}^*(\Delta p_k) \quad \text{with} \quad \Delta p_k(\mathbf{x}) = \begin{cases} \frac{p_{j_k}^{\text{obs}}(\mathbf{x}) - \mathcal{P}_{j_k}(f_k)(\mathbf{x})}{u_{j_k}(\mathbf{x})} & \text{if } u_{j_k}(\mathbf{x}) \neq 0 \\ 0 & \text{else} \end{cases}, \quad (6.1.2)$$

There exist methods for the evaluation $\mathcal{P}_{\text{tot}}^$ in $\mathcal{O}(N^3 \log N) + \mathcal{O}(N_{\text{proj}}N^2)$ operations, based on FFTs and interpolations in Fourier space, see e.g. [69, 113, 8]. *In practice*, however, we find that these approaches cannot compete with high-performance implementations of standard discretizations (as provided by the ASTRA-toolbox [199, 198], for example), which is why they are not further discussed.

where $u_{j_k} := \mathcal{P}_{j_k}(\mathbf{1}_\Omega)$ is the projected indicator function of the object-domain $\Omega \subset \mathbb{R}^3$ (*unit-projection*). By replacing \mathcal{P}_{j_k} , f_k , $p_{j_k}^{\text{obs}}$, u_{j_k} with suitable discretizations $\mathbf{P}_{j_k} \in \mathbb{R}^{m \times n}$, $\mathbf{f}_k \in \mathbb{R}^n$ and $\mathbf{p}_{j_k}^{\text{obs}}, \mathbf{u}_{j_k} \in \mathbb{R}^m$, the update (6.1.2) can be implemented as a numerical scheme:

Scheme 6.1 (Classical SART (essentially)).

- (1) Forward-project the current iterate: $\mathbf{p}_k = \mathbf{P}_{j_k}(\mathbf{f}_k)$
- (2) Compute increment in projection space: $\Delta \mathbf{p}_k = (\mathbf{p}_{j_k}^{\text{obs}} - \mathbf{p}_k) \oslash_{\neq 0} \mathbf{u}_{j_k}$
- (3) Back-project and increment: $\mathbf{f}_{k+1} = \mathbf{f}_k + \mathbf{P}_{j_k}^*(\Delta \mathbf{p}_k)$

The operator $\oslash_{\neq 0}$ assigns the element-wise quotient of vectors whenever the divisor is non-zero and 0 otherwise. It is shown in the article that scheme 6.1 essentially coincides with iterations of the *Simultaneous Algebraic Reconstruction Technique* (SART), which has been proposed in [6] following a completely different reasoning.

Notably, the SART-scheme 6.1 computes the Kaczmarz-iterations from (6.1.1) in a highly efficient manner: it requires only a single forward- and back-projection (steps (1) and (3)), whereas the actual update is computed in step (2) purely by manipulating *2D-projections* instead of *3D-objects*. Step (2) thus involves only arrays of size $m = \mathcal{O}(N^2)$ (instead of $n = \mathcal{O}(N^3)$), which makes it computationally cheap. Indeed, scheme 6.1 meets the performance-requirements from §6.1.1: it performs as few as possible operations on 3D-data and involves only one evaluation of \mathbf{P}_{j_k} and $\mathbf{P}_{j_k}^*$ each, resulting in $\mathcal{O}(N^3)$ computational costs of the total SART-update. A drawback is that it only computes minimizers up to discretization-errors:

Remark 6.2 (Inexactness in the discrete setting). While (6.1.2) constitutes an exact solution to the Kaczmarz-update (6.1.1), the discretized SART-update typically does *not* exactly solve the associated discrete problem, i.e. if \mathbf{f}_{k+1} is defined by scheme 6.1, then in general

$$\mathbf{f}_{k+1} \notin \operatorname{argmin}_{\mathbf{f} \in B_k} \|\mathbf{f} - \mathbf{f}_k\|_2^2 \quad \text{with} \quad B_k := \operatorname{argmin}_{\mathbf{f} \in \mathbb{R}^n} \|\mathbf{P}_{j_k}(\mathbf{f}) - \mathbf{p}_{j_k}^{\text{obs}}\|_2^2 \quad (6.1.3)$$

In agreement with the Newton-Kaczmarz results in §5.2.3, it is a recurrent observation that Kaczmarz methods exhibit fast *semi-convergence* [57]: while not necessarily converging for $k \rightarrow \infty$, the iterates f_k typically provide a reasonable reconstruction already after one or two fitting-cycles over the data, i.e. after $k \sim N_{\text{proj}}$ iterations. Hence, a complete Kaczmarz-reconstruction via SART amounts to similar $\mathcal{O}(N_{\text{proj}}N^3)$ computational costs as FBP.

6.1.3 The idea of generalizing SART

According to §6.1.2, the seemingly complicated Kaczmarz-updates in (6.1.1) can be computed *non-iteratively* by an efficient SART-scheme. The Newton-Kaczmarz algorithm from Article 3 constitutes a generalization and adaptation of the classical Kaczmarz method to the XPCT-setting. As will be seen in §6.2.5.2, the corresponding updates implemented in §5.2.3 can be cast to the generic form of *generalized* Kaczmarz-iterations:

$$f_{k+1} \in \operatorname{argmin}_{f \in L^2(\Omega)} \mathcal{S}_k(\mathcal{P}_{j_k}(f)) + \mathcal{R}_k(f) \quad (6.1.4)$$

for suitably defined functionals \mathcal{S}_k and \mathcal{R}_k . Several other recent algorithms give rise to updates of the form (6.1.4) when applied to tomographic reconstruction problems, among those so-called Tikhonov-Kaczmarz iterations [50, 115], incremental proximal methods [21, 22, 7] and stochastic primal-dual methods [35]. Consequently, efficient solution-formulas for (6.1.4) are of interest beyond the specific setting of XPCT.

Analogously as classical SART provides an efficient computational scheme for classical Kaczmarz-iterations (the case $\mathcal{R}_k(f) \equiv 0$, $\mathcal{S}_k(p) = \|p - p_{j_k}^{\text{obs}}\|_{L^2}^2$ in (6.1.4)), the principal aim of the presented article is to derive *generalized* SART-formulas for the computation of the *generalized* Kaczmarz-update (6.1.4).

6.2 Summary of the results

6.2.1 General setting

The article considers general inverse problems in transmission tomography, where the acquired data is given by a finite number of tomographic projections mapped under additional image-formation operators F_j , that may for example model phase contrast:

$$g_{\text{tot}} = \begin{pmatrix} g_1 \\ \vdots \\ g_{N_{\text{proj}}} \end{pmatrix} = \begin{pmatrix} F_1(\mathcal{P}_1(f)) \\ \vdots \\ F_{N_{\text{proj}}}(\mathcal{P}_{N_{\text{proj}}}(f)) \end{pmatrix} = F_{\text{tot}}(\mathcal{P}_{\text{tot}}(f)) \quad (6.2.1)$$

Here and in the following, g_{tot} is the total data composed of the individual blocks g_j that give the data acquired under the j th tomographic incident direction. The \mathcal{P}_j are the corresponding tomographic projectors, that generate the j th tomographic projection for a given 3D-density $f \in L^2(\Omega)$ with support in $\Omega \subset \mathbb{R}^3$ bounded. The aim is to reconstruct f from the data g_{tot} .

In the complete article (chapter 11), both the cases of parallel-beam- and cone-beam (divergent-beam) projectors are considered. As the geometry in holographic XPCI is typically well-approximated by the simpler parallel-beam setting (compare §2.1.5.1), however, this summary restricts to the parallel-beam case for simplicity. Accordingly, the projectors $\mathcal{P}_j := \mathcal{P}_{\theta_j}$ are defined by (2.3.1). Notably, absolutely *no* assumption is made on the angular sampling, i.e. the incident directions $\theta_j \in \mathbb{S}^2$ are completely arbitrary throughout this chapter.

The article considers Kaczmarz-iterations for the setting (6.2.1) of general variational form:

$$f_{k+1} \in \underset{f \in L^2(\Omega)}{\text{argmin}} \mathcal{S}_k(g_{j_k}^{\text{obs}}; F_{j_k}(\mathcal{P}_{j_k}(f))) + \mathcal{R}_k(f), \quad j_k \in \{1, \dots, N_{\text{proj}}\} \quad (6.2.2)$$

Analogously as for the IRGNM in §5.2.1, the data-fidelity functionals \mathcal{S}_k allow to adapt the scheme to the expected statistics of the data-errors, whereas the penalty functionals \mathcal{R}_k enforce desirable object-properties and bound the increments $f_{k+1} - f_k$, i.e. control the *stepsize*.

Notational simplifications: We note that the operators F_j may be absorbed into the functionals \mathcal{S}_k by a suitable redefinition. Likewise, we omit the dependence of \mathcal{S}_k on the measured data-chunks g_j^{obs} for brevity. Moreover, as we will only analyze a *single* iteration of the form (6.2.2), all the indices k and j_k are suppressed in the following. Hence, we simplify (6.2.2) to

$$f_{\text{new}} \in \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \mathcal{S}(\mathcal{P}(f)) + \mathcal{R}(f). \quad (6.2.3)$$

Finally, it can be assumed w.l.o.g. that the incident direction $\boldsymbol{\theta} \in \mathbb{S}^2$ coincides with the z -axis, by the freedom in choosing the coordinate system. Hence, we may restrict to \mathcal{P} given by

$$\mathcal{P}(f)(\mathbf{x}_{\perp}) = \mathcal{P}_{\mathbf{e}_z}(f)(\mathbf{x}_{\perp}) = \int_{\mathbb{R}} f(\mathbf{x}_{\perp}, z) \, dz \quad \text{for all } \mathbf{x}_{\perp} = (x, y) \in \mathbb{R}^2. \quad (6.2.4)$$

6.2.2 Properties of tomographic projectors and notation

In order to devise efficient solution-schemes for (6.2.3), the presented article takes advantage of characteristic properties of tomographic projectors.

We recall from §2.3.2 that $\mathcal{P} : L^2(\Omega) \rightarrow L^2(\mathbb{R}^2)$ is a bounded linear operator. For the present study, it is important that this statement can be sharpened: if we define the *projection-domain* $\mathbb{D}_{\mathcal{P}} := \operatorname{supp}(u_{\mathcal{P}}) \subset \mathbb{R}^2$ as the support of the *unit-projection* $u_{\mathcal{P}} := \mathcal{P}(\mathbf{1}_{\Omega})$, it holds that $\operatorname{supp}(\mathcal{P}(f)) \subset \mathbb{D}_{\mathcal{P}}$ for all $f \in L^2(\Omega)$. Hence, $\mathcal{P} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_{\mathcal{P}}) \subset L^2(\mathbb{R}^2)$ is a well-defined, bounded linear operator. Its adjoint is given by a *back-projection* operator:

$$\mathcal{P}^* : L^2(\mathbb{D}_{\mathcal{P}}) \rightarrow L^2(\Omega); \quad \mathcal{P}^*(p)(\mathbf{x}_{\perp}, z) = \begin{cases} p(\mathbf{x}_{\perp}) & \text{if } (\mathbf{x}_{\perp}, z) \in \Omega \\ 0 & \text{else} \end{cases}. \quad (6.2.5)$$

According to (6.2.5), $\mathcal{P}^*(p)$ is always constant along the tomographic *ray-direction* (w.l.o.g. the z -direction here), along which the operator \mathcal{P} integrates. This implies that

$$\mathcal{P}(w \cdot \mathcal{P}^*(p)) = \mathcal{P}(w) \cdot p \quad (6.2.6)$$

for any p and weight-function w for which the expression is well-defined. In particular, the following relation holds true:

$$\mathcal{P} \mathcal{P}^*(p) = \mathcal{P}(\mathbf{1}_{\Omega} \cdot \mathcal{P}^*(p)) = u_{\mathcal{P}} \cdot p \quad \text{for all } p \in L^2(\mathbb{D}_{\mathcal{P}}). \quad (6.2.7)$$

The formulas in (6.2.6) and (6.2.7) are highly valuable from a computational perspective: while \mathcal{P} and \mathcal{P}^* *alone* map between 3D-objects and 2D-projections and thus typically require $\mathcal{O}(N^3)$ floating point operations to evaluate numerically, *compositions* of these maps, $\mathcal{P} \mathcal{P}^*$ and $p \mapsto \mathcal{P}(w \cdot \mathcal{P}^*(p))$, are simply multiplication operators on 2D-projections. Hence, the latter maps can be implemented numerically at a favorable computational complexity of $\mathcal{O}(N^2)$. One promising approach for an efficient computation of the Kaczmarz-iterations in (6.2.3) is therefore to rewrite the optimization problem in terms of $\mathcal{P} \mathcal{P}^*$ as far as possible.

In order to make sense of the results of the subsequent sections, it is furthermore important to note that we adopt the following notational convention:

Convention 6.3. All local operations $(+, -, \cdot, /, (\cdot)^\gamma, \nabla, \dots)$ on functions $f \in L^2(\Omega) \subset L^2(\mathbb{R}^3)$ and $p \in L^2(\mathbb{D}_{\mathcal{P}}) \subset L^2(\mathbb{R}^2)$ are implicitly understood to be evaluated only in the *interior* of Ω and $\mathbb{D}_{\mathcal{P}}$ ($\text{int}(A) := \bigcup_{U \subset A \text{ open}} U$). For example, a quotient of $p_1, p_2 \in L^2(\mathbb{D}_{\mathcal{P}})$ is to be read as

$$(p_1/p_2)(\mathbf{x}_\perp) = \begin{cases} p_1(\mathbf{x}_\perp)/p_2(\mathbf{x}_\perp) & \text{if } \mathbf{x}_\perp \in \text{int}(\mathbb{D}_{\mathcal{P}}) \\ 0 & \text{else} \end{cases} \quad \text{for all } \mathbf{x}_\perp \in \mathbb{R}^2. \quad (6.2.8)$$

6.2.3 The Generalized SART-principle

Casting the optimization to projection-space: We aim to generalize the classical SART-scheme 6.1 from §6.1.2 to compute Kaczmarz-iterations of the form (6.2.3). One important property of the classical update-formula (6.1.2) is that the increments $f_{k+1} - f_k$ are always *back-projections* $\mathcal{P}_{j_k}(\Delta p_k)$ of some projection-increment Δp_k . If a similar property is known to hold in (6.2.3), i.e. if $f_{\text{new}} - f_{\text{ref}} = \mathcal{P}^*(\Delta p)$ for $\Delta p \in L^2(\mathbb{D}_{\mathcal{P}})$ and some *reference-object* $f_{\text{ref}} \in L^2(\Omega)$, then it would follow that

$$\begin{aligned} \mathcal{P}^*(\Delta p) &= f_{\text{new}} - f_{\text{ref}} \in \underset{\Delta f \in L^2(\Omega)}{\text{argmin}} \mathcal{S}(\mathcal{P}(f_{\text{ref}} + \Delta f)) + \mathcal{R}(f_{\text{ref}} + \Delta f) \\ &\Rightarrow \Delta p \in \underset{p \in L^2(\mathbb{D}_{\mathcal{P}})}{\text{argmin}} \mathcal{S}(\mathcal{P}(f_{\text{ref}}) + \mathcal{P}\mathcal{P}^*(\Delta p)) + \mathcal{R}(f_{\text{ref}} + \mathcal{P}^*(\Delta p)) \end{aligned} \quad (6.2.9)$$

According to (6.2.9), Δp (and hence $f_{\text{new}} = f_{\text{ref}} + \mathcal{P}^*(\Delta p)$) can be computed by optimizing in *projection-space*. In a discretized setting, this reduces the dimension of the optimization problem to $\mathcal{O}(N^2)$ compared to $\mathcal{O}(N^3)$ for the original problem in *object-space*, potentially enabling massive gains in computational performance. Moreover, note that the operator $\mathcal{P}\mathcal{P}^*$ arises here, which can be implemented at low computational costs, as argued in §6.2.2.

When are Kaczmarz-updates back-projections? For (6.2.9) to apply, we need to ensure that the Kaczmarz-increment $\Delta f := f_{\text{new}} - f_{\text{ref}} \in L^2(\Omega)$ is a back-projection. By standard Hilbert-space geometry, any $\Delta f \in L^2(\Omega)$ can be (uniquely and orthogonally) decomposed as $\Delta f = \Delta f_{\mathcal{P}^*} + \Delta f_0$, where $\Delta f_{\mathcal{P}^*}$ is an element of (the closure of) $\text{range}(\mathcal{P}^*) := \mathcal{P}^*(L^2(\mathbb{D}_{\mathcal{P}}))$ and $\Delta f_0 \in \text{kern}(\mathcal{P}) := \{f \in L^2(\Omega) : \mathcal{P}(f) = 0\}$ is in the null-space of \mathcal{P} . Hence, Δf is a back-projection if and only if its component in $\text{kern}(\mathcal{P})$ vanishes. Importantly, it holds that

$$\mathcal{S}(\mathcal{P}(f_{\text{ref}} + \Delta f)) = \mathcal{S}\left(\mathcal{P}(f_{\text{ref}}) + \mathcal{P}(\Delta f_{\mathcal{P}^*}) + \underbrace{\mathcal{P}(\Delta f_0)}_{=0}\right) = \mathcal{S}(\mathcal{P}(f_{\text{ref}} + \Delta f_{\mathcal{P}^*})), \quad (6.2.10)$$

i.e. the data-fidelity term in (6.2.3) does not “see” the null-space component Δf_0 . Accordingly, this part of the update Δf must be solely determined by the penalty functional \mathcal{R} . As long as \mathcal{R} does not actively *promote* that non-zero null-space components emerge, there will always be a minimizer of (6.2.3) such that $f_{\text{new}} - f_{\text{ref}}$ is a back-projection for some $f_{\text{ref}} \in L^2(\Omega)$. Imposing this property of \mathcal{R} is the idea underlying to the following assumption:

Assumption 6.4. Let $\tilde{P} : X \rightarrow Y$ be a bounded linear operator on Hilbert spaces X, Y with null-space $\text{kern}(\tilde{P})$ such that $\text{range}(\tilde{P}) = \tilde{P}(X) \subset Y$ is closed and let $\mathcal{R} : X \rightarrow \mathbb{R} \cup \{\infty\}$ be a functional. Assume that there exists an $f_{\text{ref}} \in X$ such that

$$\mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p) + f_0) \geq \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) \quad \text{for all } p \in Y, f_0 \in \text{kern}(\tilde{P}) \quad (\text{A})$$

Assumption 6.4 is stated in terms of an abstract map \tilde{P} because $\mathcal{P} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_{\mathcal{P}})$ does not have a closed range so that, strictly, it may not satisfy the assumption. However, it is shown in the article (see theorem 11.4) that a slightly modified projector, $\mathcal{P}_{\text{iso}} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_{\mathcal{P}})$ with $\mathcal{P}_{\text{iso}}(f) := u_{\mathcal{P}}^{-1/2} \cdot \mathcal{P}(f)$ (well-defined by convention 6.3), is bounded and has closed range. As a consequence, the non-closed range of \mathcal{P} only renders the proofs of subsequent results more technical, but does not cause any substantial difficulties.

The general principle: The crucial part of assumption 6.4 is the inequality (A), that ensures that \mathcal{R} uniformly penalizes deviations from f_{ref} by null-space elements. If this is satisfied, the arguments from (6.2.9) and (6.2.10) lead to the principal theorem of the article:

Theorem 6.5 (Generalized SART-principle). *Let assumption 6.4 be satisfied and let $\tilde{\mathcal{S}} : Y \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Assume that there exists a minimizer*

$$f_{\text{new}} \in \operatorname{argmin}_{f \in X} \tilde{\mathcal{S}}(\tilde{P}(f)) + \mathcal{R}(f). \quad (6.2.11)$$

Then there is a (possibly distinct) minimizer $\tilde{f}_{\text{new}} \in X$ of (6.2.11) given by

$$p_{\text{ref}} = \tilde{P}(f_{\text{ref}}) \quad (6.2.12a)$$

$$\Delta p \in \operatorname{argmin}_{p \in Y} \tilde{\mathcal{S}}(p_{\text{ref}} + \tilde{P}\tilde{P}^*(p)) + \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) \quad (6.2.12b)$$

$$\tilde{f}_{\text{new}} = f_{\text{ref}} + \tilde{P}^*(\Delta p) \quad (6.2.12c)$$

Conversely, any \tilde{f}_{new} given by (6.2.12) minimizes (6.2.11). Furthermore, if strict inequality holds in (A) whenever $f_0 \neq 0$, then all minimizers of (6.2.11) are of the form (6.2.12).

Although formulated in an abstract setting, the result in theorem 6.5 is particularly well-suited for the tomographic Kaczmarz-iterations in (6.2.3) for the following reasons:

- The optimization problem in (6.2.11) is cast from the object-space X to the image-space Y (projection-space) in (6.2.12b). In the Kaczmarz-setting, Y is much smaller than X .
- For $\tilde{P} = \mathcal{P}_{(\text{iso})}$, the operator $\tilde{P}\tilde{P}^*$ in (6.2.12b) is easy to evaluate, compare §6.2.2.

When theorem 6.5 is applied to Kaczmarz-iterations of the form (6.2.2) (or (6.2.3)), we refer to the resulting solution-formulas (6.2.12) as *generalized SART* (or *GenSART*-) schemes, emphasizing their similarity to classical SART-updates (compare scheme 6.1):

Scheme 6.6 (Generalized SART (GenSART)).

- (1) Forward-project the current iterate: $p_k = \mathcal{P}_{j_k}(f_k)$
- (2) Optimize in projection-space: $\Delta p_k = \operatorname{argmin}_p J(p, p_k, u_{j_k}, \dots)$
- (3) Back-project and increment: $f_{k+1} = f_k + \mathcal{P}_{j_k}^*(\Delta p_k)$

Scheme 6.6 coincides with the classical SART-scheme 6.1, except that step (2) is replaced by an abstract optimization problem that depends on details of the Kaczmarz-iterations (6.2.2) to be computed. As will be seen in §6.2.4, the (back-)projectors $\mathcal{P}_{j_k}, \mathcal{P}_{j_k}^*$ may typically be completely eliminated from the optimization step (2) in scheme 6.6. In combination with the low dimension $\mathcal{O}(N^2)$ of the projection-space in discretized settings (see above), this renders step (2) of GenSART computationally inexpensive compared to the forward- and back-projection steps (1) and (3) even if the minimizer has to be computed iteratively. Hence:

Generalized SART permits the computation of general tomographic Kaczmarz-iterations essentially at the computational costs of classical SART-updates.

6.2.4 Admissible penalty-functionals

In order to apply theorem 6.5 to the Kaczmarz-iterations (6.2.3), inequality (A) has to be verified for the chosen penalty functional \mathcal{R} . The null-space of a tomographic projector $\text{kern}(\mathcal{P})$ contains all elements of $L^2(\Omega)$ that integrate to zero along all tomographic rays and thus correspond to additional *oscillations* along the ray-direction compared to an increment $f_{\text{new}} - f_{\text{ref}} = \mathcal{P}^*(p)$ that is a pure back-projection. Intuitively, this means that any penalties \mathcal{R} which tend to damp out oscillatory deviations from the reference object f_{ref} , i.e. such that are *smoothing*, are promising candidates to satisfy assumption 6.4.

6.2.4.1 L^2 -penalties and generalizations

L^2 -penalties $\mathcal{R}(f) \propto \|f - f_{\text{ref}}\|_{L^2}^2$ have a smoothing effect. Indeed, such penalties are found to satisfy assumption 6.4 in a more general setting:

Lemma 6.7 (Quadratic norm penalties). *Let X, Y be Hilbert spaces, $\tilde{P} : X \rightarrow Y$ linear and bounded with closed range and let $\mathcal{R}(f) := \|f - f_{\text{ref}}\|_X^2$ for some $f_{\text{ref}} \in X$. Then assumption 6.4 is satisfied with strict inequality in (A) for all $f_0 \neq 0$ and it holds that*

$$\mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) = \langle p, \tilde{P}\tilde{P}^*(p) \rangle_Y \quad \text{for all } p \in Y. \quad (6.2.13)$$

Via theorem 6.5, lemma 6.7 yields generalized SART-schemes for L^2 -penalized Kaczmarz-iterations. By simplifying terms involving $\mathcal{P}\mathcal{P}^*$ via (6.2.6), the following result is obtained:

Theorem 6.8 (Generalized SART with L^2 -penalty). *Let $f_{\text{ref}} \in L^2(\Omega)$, $\alpha > 0$ and let $\mathcal{S} : L^2(\mathbb{D}_{\mathcal{P}}) \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Then the minimizers of*

$$f_{\text{new}} \in \underset{f \in L^2(\Omega)}{\text{argmin}} \mathcal{S}(\mathcal{P}(f)) + \alpha \|f - f_{\text{ref}}\|_{L^2}^2 \quad (6.2.14)$$

are uniquely determined by the GenSART-scheme

$$p_{\text{ref}} = \mathcal{P}(f_{\text{ref}}) \quad (6.2.15a)$$

$$\Delta p \in \underset{p \in L^2(\mathbb{D}_{\mathcal{P}})}{\text{argmin}} \mathcal{S}(p_{\text{ref}} + u_{\mathcal{P}}^{1/2} \cdot p) + \alpha \|p\|_{L^2}^2 \quad (6.2.15b)$$

$$f_{\text{new}} = f_{\text{ref}} + \mathcal{P}^*(u_{\mathcal{P}}^{-1/2} \cdot \Delta p). \quad (6.2.15c)$$

Recall that the multiplication by $u_{\mathcal{P}}^{-1/2}$ in (6.2.15) is defined in the spirit of convention 6.3. Importantly, theorem 6.8 holds whenever the penalty \mathcal{R} is given by a squared L^2 -norm – completely independently of the choice of the data-functional \mathcal{S} ! Furthermore, generalizations of theorem 6.8 are proven in the complete article for different settings, compare §11.4.

- (1) *Weighted projectors* (theorem 11.11): $f_{\text{new}} \in \operatorname{argmin} \mathcal{S}(\mathcal{P}(\lambda \cdot f)) + \alpha \|f - f_{\text{ref}}\|_{L^2}^2$ for some bounded function λ with $\lambda(\mathbf{x}) \neq 0$ for almost all $\mathbf{x} \in \Omega$.
- (2) *Weighted L^2 -penalties* (corollary 11.12): $\mathcal{R}(f) = \alpha \|w^{-1/2} \cdot (f - f_{\text{ref}})\|_{L^2}^2$ for some bounded weight-function w with $w(\mathbf{x}) > 0$ for almost all $\mathbf{x} \in \Omega$.
- (3) *L^q -penalties* (theorem 11.17): $\mathcal{R}(f) \propto \|f - f_{\text{ref}}\|_{L^q}^q = \int_{\Omega} |f - f_{\text{ref}}|^q d\mathbf{x}$ for $1 \leq q < \infty$.

For all of these choices, the principal structure of the GenSART-scheme (6.2.15) is retained, enabling computations of the Kaczmarz-iterates at the efficiency of classical SART. For cases (1) and (2), the proof again uses lemma 6.7, yet taking \tilde{P} as a suitably *weighted* projector, whereas the *non-quadratic* penalty in (3) requires a different approach based on Jensen’s inequality.

6.2.4.2 Gradient penalties

As smoothing has been identified as an decisive property, it is natural to guess that assumption 6.4 is also satisfied for *gradient- L^2* -penalties of the form

$$\mathcal{R}(f) := \|\nabla(f - f_{\text{ref}})\|_{L^2} = \int_{\Omega} |\nabla(f - f_{\text{ref}})|^2 d\mathbf{x}. \quad (6.2.16)$$

Importantly, gradients in this section are meant to be computed only in the interiors of Ω and $\mathbb{D}_{\mathcal{P}}$ – not across boundaries, see convention 6.3. We analyze \mathcal{R} on the Sobolev space of weakly differentiable functions in Ω with gradient in L^2 , $W^{1,2}(\Omega) := \{f \in L^2(\Omega) : |\nabla f| \in L^2(\Omega)\}$.

With \mathcal{R} as in (6.2.16), we have for all sufficiently smooth $\mathcal{P}^*(p) \in \operatorname{range}(\mathcal{P}^*)$, $f_0 \in \operatorname{kern}(\mathcal{P})$

$$\mathcal{R}(f^{(\text{ref})}) + \mathcal{P}^*(p) + f_0 = \|\nabla \mathcal{P}^*(p)\|_{L^2}^2 + \|\nabla f_0\|_{L^2}^2 + 2\operatorname{Re} \int_{\Omega} \nabla \mathcal{P}^*(p) \cdot \overline{\nabla f_0} d\mathbf{x}. \quad (6.2.17)$$

From (6.2.17), it can be seen that $\mathcal{R}(f^{(\text{ref})}) + \mathcal{P}^*(p) + f_0 \geq \mathcal{R}(f^{(\text{ref})}) + \mathcal{P}^*(p)$ (inequality (A)) holds in general if and only if $\int_{\Omega} \nabla \mathcal{P}^*(p) \cdot \overline{\nabla f_0} d\mathbf{x} = 0$ for any admissible $\mathcal{P}^*(p)$ and f_0 . As detailed in the article, this turns out to be only *almost* the case, up to some boundary-terms that arise in manipulations of the integral due to thickness-variations of the domain Ω .

As a consequence, assumption 6.4 is only satisfied in slightly restricted settings. Two valid scenarios are identified in the article:

- (1) Ω is a cuboid aligned with the z -axis or, more generally, $\Omega = \Omega_{\perp} \times \Omega_{\parallel}$ for some $\Omega_{\perp} \subset \mathbb{R}^2$ and $\Omega_{\parallel} \subset \mathbb{R}$.
- (2) f_0 is additionally assumed to vanish on the boundary of Ω , i.e. $f_0 \in \operatorname{kern}(\mathcal{P}) \cap W_0^{1,2}(\Omega)$, where $W_0^{1,2}(\Omega) \subset W^{1,2}(\Omega)$ denotes the usual Sobolev space of trace-zero functions.

While setting **(1)** is geometrically restrictive, following the idea of **(2)** ultimately implies that generalized SART-schemes only hold for gradient-penalized Kaczmarz-iterations if the optimization in (6.2.3) is restricted to a slightly smaller search set:

$$f_{\text{new}} - f_{\text{ref}} \in W_{\mathcal{P}}^{1,2}(\Omega) := \left(\overline{\text{range}(\mathcal{P}^*)} \cap W^{1,2}(\Omega) \right) \oplus \left(\text{kern}(\mathcal{P}) \cap W_0^{1,2}(\Omega) \right). \quad (6.2.18)$$

It is furthermore shown in the article that the gradient-penalty can be simplified for arguments given by back-projections: it holds that

$$\mathcal{R}(f^{\text{ref}}) + \mathcal{P}^*(p) = \|\nabla \mathcal{P}^*(p)\|_{L^2} = \|u_{\mathcal{P}}^{1/2} \cdot \nabla(p)\|_{L^2} \quad \text{for all } \mathcal{P}^*(p) \in W^{1,2}(\Omega). \quad (6.2.19)$$

Using (6.2.19) and the restriction (6.2.18), generalized SART-schemes are obtained for arbitrary convex combinations of L^2 - and gradient- L^2 -penalties, i.e. for general Sobolev- $W^{1,2}$ -penalties:

Theorem 6.9 (Generalized SART with Sobolev- $W^{1,2}$ -penalties). *Let $f_{\text{ref}} \in L^2(\Omega)$, $\alpha > 0$, $0 \leq \gamma \leq 1$ and let $\mathcal{S} : L^2(\mathbb{D}_{\mathcal{P}}) \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Then the minimizers of*

$$f_{\text{new}} \in \underset{f \in f_{\text{ref}} + W_{\mathcal{P}}^{1,2}(\Omega)}{\text{argmin}} \mathcal{S}(\mathcal{P}(f)) + \alpha \left((1 - \gamma) \|f - f_{\text{ref}}\|_{L^2}^2 + \gamma \|\nabla(f - f_{\text{ref}})\|_{L^2}^2 \right). \quad (6.2.20)$$

are uniquely determined by the GenSART-scheme

$$p := \mathcal{P}(f_{\text{ref}}) \quad (6.2.21a)$$

$$\Delta p \in \underset{p \in L^2(\mathbb{D}_{\mathcal{P}})}{\text{argmin}} \mathcal{S}(p_{\text{ref}} + u_{\mathcal{P}}^{1/2} \cdot p) + \alpha(1 - \gamma) \|p\|_{L^2}^2 + \alpha\gamma \|u_{\mathcal{P}}^{1/2} \cdot \nabla(u_{\mathcal{P}}^{-1/2} \cdot p)\|_{L^2}^2 \quad (6.2.21b)$$

$$f_{\text{new}} = f_{\text{ref}} + \mathcal{P}^*(u_{\mathcal{P}}^{-1/2} \cdot \Delta p) \quad (6.2.21c)$$

Admittedly, theorem 6.9 is flawed by the required restriction of the natural search set $W^{1,2}(\Omega)$ (all functions for which the gradient-penalty is finite) in the optimization-problem (6.2.20) to increments $f_{\text{new}} - f_{\text{ref}} \in W_{\mathcal{P}}^{1,2}(\Omega)$. Yet, while $W_{\mathcal{P}}^{1,2}(\Omega) \subsetneq W^{1,2}(\Omega)$, the “missing space” V such that $W_{\mathcal{P}}^{1,2}(\Omega) \oplus V = W^{1,2}(\Omega)$ is relatively small, as discussed in the article. Moreover, by (6.2.18), $W_{\mathcal{P}}^{1,2}(\Omega)$ is *much larger* than the set $\overline{\text{range}(\mathcal{P}^*)} \cap W^{1,2}(\Omega)$, over which the scheme (6.2.21) trivially provides the optimum (by (6.2.9) and (6.2.19)). In this sense, theorem 6.9 is mathematically non-trivial and also valuable in practice: as $W_{\mathcal{P}}^{1,2}(\Omega)$ contains *almost* all admissible increments, the GenSART-scheme (6.2.21) is expected to yield *almost* optimal Kaczmarz-iterates in terms of an optimization of (6.2.20) over the total space $L^2(\Omega)$.

6.2.5 Applications: XPCT and beyond

The most appealing feature of the generalized SART principle is that the data-functional \mathcal{S} in (6.2.11) may be completely arbitrary. Recalling that we have set $\mathcal{S}(p) := \mathcal{S}_k(g_j^{\text{obs}}; F_j(p))$ in §6.2.1 for notational brevity, incorporating both the image-formation maps F_j and the dependence on the observed data g_j^{obs} , we see that the freedom of choosing \mathcal{S} paves the way for various interesting applications, including the Newton-Kaczmarz method for XPCT from §5.2.3.

6.2.5.1 Noise-model-adapted GenSART

Analogously as discussed for the IRGNM in §5.2.1.1, generalized Kaczmarz-iterations (6.2.2) may be adapted to the expected data-error statistics via the choice of the data-fidelities \mathcal{S}_k . A standard approach to achieve this is to choose \mathcal{S}_k as the *negative log-likelihood*, $\mathcal{S}_k(g_j^{\text{obs}}; g_j) = -\ln \mathbb{P}(g_j^{\text{obs}}|g_j)$, where $\mathbb{P}(g_j^{\text{obs}}|g_j)$ is the conditional probability(-density) of measuring g_j^{obs} given that the true data is g_j . In particular, the following examples are of interest:[†]

- *Gaussian noise models* lead to (weighted) L^2 -penalties, $\mathcal{S}_k(g_j^{\text{obs}}; g_j) = \|(g_j - g_j^{\text{obs}})/\sigma_j\|_{L^2}^2$, where σ_j is the (possibly spatially varying) standard deviation of the data-errors.
- *Heavy-tailed distributions (robust fitting)*: Gaussian error models do not account for the possibility of large outliers in the data and are therefore prone to over-fitting. In order to achieve more robust reconstructions, it is standard to assume more *heavy-tailed* probability-distributions for the data-errors such as (smoothed) Laplace- or Student's t-distributions, which lead to functionals of the form ($\nu > 0$: robustness-parameter)

$$\mathcal{S}_k(g_j^{\text{obs}}; g_j) = \int s_-(g_j(x) - g_j^{\text{obs}}(x)) dx, \quad s_- \in \{s_{L^1_{\text{H}}, \nu}, s_{\text{s-t}, \nu}\}, \quad (6.2.22a)$$

$$s_{L^1_{\text{H}}, \nu}(y) := \begin{cases} |y|^2 & \text{if } |y| \leq \nu \\ 2\nu|y| - \nu^2 & \text{else} \end{cases}, \quad s_{\text{s-t}, \nu}(y) := \nu^2 \ln(1 + |y|^2/\nu^2) \quad (6.2.22b)$$

- *Poisson noise*: Data-errors that arise due to the discreteness of the detected photons (or electrons or other quantum-mechanical particles) in an imaging-experiment follow a Poisson-distribution, as discussed in §2.1.5.4 for XPCI. If this is the only source of noise, the log-likelihood approach yields a data-fidelity $\mathcal{S}_k(g_j^{\text{obs}}; g_j)$ given by the *Kullback-Leibler-divergence*, as detailed in the article and other works (see e.g. [99, 206]):

$$\mathcal{S}_k(g_j^{\text{obs}}; g_j) = \int \text{kl}(tg_j^{\text{obs}}; tg_j) d\mu, \quad \text{kl}(b; a) := \begin{cases} a - b - b \ln\left(\frac{a}{b}\right) & a, b \geq 0 \\ \infty & \text{else} \end{cases} \quad (6.2.23)$$

with a suitable integral-measure μ and an exposure-time parameter $t > 0$.

For any of these choices of \mathcal{S}_k , Kaczmarz-iterations of the general form (6.2.2) may be computed for arbitrary image-formation models F_j and any of the penalty functionals from §6.2.4.

Efficient closed-form optimization for simple models: In general, the optimization in projection-space, step (2) in the principal GenSART-scheme 6.6, has to be performed iteratively, which may be computationally demanding. An exception is given by standard tomographic reconstruction problems with pure *projection-data* $g_j^{\text{obs}} \approx \mathcal{P}_j(f)$ (i.e. no phase contrast or other non-trivial image-formation maps F_j involved), as arising in classical CT and many other imaging modalities: as detailed in the article, the optimization-step may then be evaluated *analytically* for any of the above noise-model-adapted data-fidelities \mathcal{S}_k , owing to the fact that the total data-functional \mathcal{S} in the general Kaczmarz-step (6.2.3) assumes an *integral-form*:

$$\mathcal{S}(p) = \mathcal{S}_k(g_j^{\text{obs}}; F_j(p)) = \int_{\mathbb{R}^2} s(\mathbf{x}, p(\mathbf{x})) d\mathbf{x} \quad (+ \text{const}) \quad (6.2.24)$$

[†]As often done, we restrict to models that assume stochastic independence of errors at different points.

for some function $s : \mathbb{R}^2 \times \mathbb{C} \rightarrow \mathbb{R} \cup \{\infty\}$ that depends on the specific setting. For the L^2 -regularized GenSART-schemes from theorem 6.8 (as well as for the outlined generalizations), the optimization in projection-space then decouples into a family of *scalar* problems:

$$\begin{aligned} \Delta p &\in \operatorname{argmin}_{p \in L^2(\mathbb{D}_{\mathcal{P}})} \mathcal{S} \left(p_{\text{ref}} + u_{\mathcal{P}}^{1/2} \cdot p \right) + \alpha \|p\|_{L^2}^2 \\ \Leftrightarrow \Delta p(\mathbf{x}) &\in \operatorname{argmin}_{p_{\mathbf{x}} \in \mathbb{R}} s(\mathbf{x}, p_{\text{ref}}(\mathbf{x}) + u_{\mathcal{P}}(\mathbf{x})^{1/2} p_{\mathbf{x}}) + \alpha (p_{\mathbf{x}} - p_{\text{ref}}(\mathbf{x}))^2, \quad \mathbf{x} \in \mathbb{R}^2. \end{aligned} \quad (6.2.25)$$

The scalar optimizations in (6.2.25) can be performed to machine-precision in $\mathcal{O}(1)$ arithmetic operations per point \mathbf{x} for all of the data-fidelities \mathcal{S}_k presented above. In a discretized setting with N samples per spatial dimension, this translates to a total complexity of $\mathcal{O}(N^2)$ of the optimization step (2) in the GenSART-scheme 6.6, which is negligible compared to the $\mathcal{O}(N^3)$ -evaluations of \mathcal{P}_j and \mathcal{P}_j^* required in steps (1) and (3). Consequently, L^2 -regularized noise-level-adapted GenSART-iterations may be implemented at *exactly* the same runtime as classical SART (scheme 6.1) – while offering greatly increased flexibility.

6.2.5.2 Newton-Kaczmarz-GenSART for XPCT

The Newton-Kaczmarz method for XPCT from Article 3 (§5.2.3) lead to iterations of the form

$$\begin{aligned} f_{k+1} = \operatorname{argmin}_{f \in L^2(\Omega)} &\|F(\mathcal{P}_{j_k}(f_k)) + F'[\mathcal{P}_{j_k}(f_k)]\mathcal{P}_{j_k}(f - f_k) - g_{j_k}^{\text{obs}}\|_{L^2}^2 \\ &+ \alpha_{k,1} \|f - f_k\|_X^2 + \alpha_{k,2} \|f - f_0\|_X^2 \end{aligned} \quad (6.2.26)$$

where $F \in \{\mathcal{N}^{(f_1, \dots, f_\ell)}, \mathcal{N}_\alpha^{(f_1, \dots, f_\ell)}\}$ is one of the XPCI-forward maps from §2.4 (possibly for multiple holograms acquired per tomographic incident direction). Upon defining

$$\begin{aligned} p_k^{(\text{ref})} &:= \mathcal{P}_{j_k}(f_k^{(\text{ref})}) & \text{with } f_k^{\text{ref}} &:= (\alpha_{k,1} f_0 + \alpha_{k,2} f_k) / \alpha_k, \quad \alpha_k := \alpha_{k,1} + \alpha_{k,2} \\ \mathcal{S}(p) &:= \|F'[p_k](p - p_k^{\text{ref}}) - r_k\|_{L^2}^2 & \text{with } r_k &:= g_{j_k}^{\text{obs}} - F(p_k) - F'[p_k](p_k^{\text{ref}} - p_k) \\ \mathcal{R}(f) &:= \alpha_{k,1} \|f - f_k\|_X^2 + \alpha_{k,2} \|f - f_0\|_X^2 = \alpha_k \|f - f_k^{\text{ref}}\|_X^2 + \text{const}, \end{aligned} \quad (6.2.27)$$

the iterations (6.2.26) are found to match the general Kaczmarz-setting in (6.2.3). For Sobolev- $W^{1,2}$ -norms $\|f\|_X^2 := (1 - \gamma)\|f\|_{L^2}^2 + \gamma\|\nabla f\|_{L^2}^2$ (including L^2 -norms as the special case $\gamma = 0$), they may thus be computed via the GenSART-scheme from theorem 6.9 ($u_j = \mathcal{P}_j(\mathbf{1}_\Omega)$: unit-projection):

$$\begin{aligned} f_{k+1} \approx f_k^{\text{ref}} + \mathcal{P}_{j_k}^* &\left(u_{j_k}^{-1/2} \cdot \operatorname{argmin}_{p \in L^2(\mathbb{D}_{\mathcal{P}})} \|F'[p_k](u_{j_k}^{1/2} \cdot p) - r_k\|_{L^2}^2 \right. \\ &\left. + \alpha_k \left((1 - \gamma)\|p\|_{L^2}^2 + \gamma\|u_{j_k}^{1/2} \cdot \nabla(u_{j_k}^{-1/2} \cdot p)\|_{L^2}^2 \right) \right), \end{aligned} \quad (6.2.28)$$

where the “ \approx ” refers to the involved slight approximation discussed in §6.2.4.2.

The optimization in (6.2.28) does not admit an analytical solution but has to be solved iteratively, e.g. by a CG-method. Yet, contrary to the method initially proposed in Article 3, the optimization no longer involves any costly $\mathcal{O}(N^3)$ forward- or back-projections $\mathcal{P}_j, \mathcal{P}_j^*$.

6.2.5.3 Newton-Kaczmarz-GenSART for polychromatic CT

The generalized SART-approach may also be applied to a *nonlinear* model of classical computed tomography (CT) from [51, 104]. This model incorporates effects due to polychromaticity of X-ray sources in clinical applications and may thereby avoid so-called *beam-hardening* artifacts. Despite the nonlinearity, it is shown in the article that L^2 -regularized Newton-Kaczmarz-iterations for the corresponding forward model may be computed via GenSART-schemes in a *non-iterative* manner at the optimal $\mathcal{O}(N^3)$ computational complexity of classical SART-updates, compare §6.1.2. The result makes use of the mentioned generalization of theorem 6.8 for weighted projectors $f \mapsto \mathcal{P}(\lambda \cdot f)$.

6.2.5.4 Extensions

The scope of the presented generalized SART-schemes may be extended in several ways:

- *Box constraints* (e.g. *non-negativity*): As is standard in Kaczmarz methods, constraints on the object-values, $f_{\min} \leq f \leq f_{\max}$, may be imposed by interlacing the iterations with *projection-steps* $f_{k+1} \leftarrow \max \{ \min \{ f_{k+1}, f_{\max} \}, f_{\min} \}$ onto the admissible range.
- *Multiple quadratic regularizers*: As used in (6.2.27), penalties \mathcal{R} that impose quadratic regularization w.r.t. *multiple* priors, $\mathcal{R}(f) = \alpha_1 \|f - f_{\text{ref},1}\|^2 + \alpha_2 \|f - f_{\text{ref},2}\|^2 + \dots$, may be simplified to $\mathcal{R}(f) = (\alpha_1 + \alpha_2 + \dots) \|f - f_{\text{ref}}\|^2$ (+ constant). This permits an implementation of Kaczmarz-iterations with such penalties via the derived GenSART-schemes.
- *Kaczmarz-type splitting methods*: To incorporate penalties \mathcal{R} that do *not* satisfy assumption 6.4, one may combine Kaczmarz-iterations with a *splitting approach*, see e.g. [43, 164]: the idea is to alternate (*forward*) (sub-)gradient-descent- or (*backward*) proximal steps on the data-fidelity- and the penalty functionals ($\partial\mathcal{R}$: subdifferential, τ, σ : stepsizes):

$$f_{k+\frac{1}{2}} = \operatorname{argmin}_{f \in L^2(\Omega)} \mathcal{S}_k(\mathcal{P}_{j_k}(f)) + \frac{1}{2\tau_k} \|f - f_k\|_{L^2}^2 \quad (6.2.29a)$$

$$f_{k+1} = \begin{cases} \operatorname{argmin}_{f \in L^2(\Omega)} \mathcal{R}(f) + \frac{1}{2\sigma_k} \|f - f_{k+\frac{1}{2}}\|_{L^2}^2 & \text{(backward-backward)} \\ f_{k+\frac{1}{2}} - \sigma_k \partial\mathcal{R}(f_{k+\frac{1}{2}}) & \text{(backward-forward)} \end{cases} \quad (6.2.29b)$$

Such Kaczmarz-type splitting methods have been proposed recently in [21, 22, 7, 35]. By theorem 6.8, GenSART-schemes may be used to efficiently compute the proximal step in the tomographic data-term, (6.2.29a).

6.2.6 Remarks on discretization

While the abstract setting of the principal theorem 6.5 equally applies to continuous and discretized tomographic inverse problems, the specific GenSART-schemes of §6.2.4 are strictly valid only in continuous settings. The main reason is that the diagonality of the operator $\mathcal{P}\mathcal{P}^*$ for a tomographic projector \mathcal{P} (compare (6.2.7)) in general does *not* carry over to discretizations $\mathbf{P} \in \mathbb{R}^{m \times n}$ of \mathcal{P} : the matrix $\mathbf{P}\mathbf{P}^* \in \mathbb{R}^{m \times m}$ will typically not be diagonal, i.e. not simply an element-wise multiplication with some discrete unit-projection $\mathbf{u} \in \mathbb{R}^m$. This peculiarity makes the discretization of GenSART less straightforward than one might hope. In principle, there are two possible strategies:

- *Post-discretization*: Devise Kaczmarz-iterations and corresponding GenSART-formulas in a continuous setting, then exchange the continuous quantities $\mathcal{P}, f, \mathcal{S}, \mathcal{R}, u_{\mathcal{P}}$ by discretizations $\mathbf{P} \in \mathbb{R}^{m \times n}, \mathbf{f} \in \mathbb{R}^n, S, R, \mathbf{u} \in \mathbb{R}^m$. *Example*: (\odot : element-wise product)

$$f_{\text{new}} \in \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \mathcal{S}(\mathcal{P}(f)) + \alpha \|f - f_{\text{ref}}\|_{L^2}^2 \quad (6.2.30)$$

$$\overset{\text{(GenSART)}}{\rightsquigarrow} f_{\text{new}} \in f_{\text{ref}} + \mathcal{P}^* \left(u_{\mathcal{P}}^{-1/2} \cdot \underset{p \in L^2(\mathbb{D}_{\mathcal{P}})}{\operatorname{argmin}} \mathcal{S}(\mathcal{P}(f_{\text{ref}}) + u_{\mathcal{P}}^{1/2} \cdot p) + \alpha \|p\|_{L^2}^2 \right) \quad (6.2.31)$$

$$\overset{\text{(discretize)}}{\rightsquigarrow} \mathbf{f}_{\text{new}} \in \mathbf{f}_{\text{ref}} + \mathbf{P}^* \left(\mathbf{u}^{-1/2} \odot \underset{\mathbf{p} \in \mathbb{R}^m}{\operatorname{argmin}} S(\mathbf{P}(\mathbf{f}_{\text{ref}}) + \mathbf{u}^{1/2} \odot \mathbf{p}) + \alpha \|\mathbf{p}\|_2^2 \right) \quad (6.2.32)$$

Analogously to classical SART, see remark 6.2, a drawback of this approach is that the discretized GenSART-update (6.2.32) in general does not solve the discrete optimization problem associated with (6.2.30), i.e. $\mathbf{f}_{\text{new}} \notin \operatorname{argmin}_{\mathbf{f} \in \mathbb{R}^n} S(\mathbf{P}(\mathbf{f})) + \alpha \|\mathbf{f} - \mathbf{f}_{\text{ref}}\|_2^2$.

- *Pre-discretization*: Start from Kaczmarz-iterations in a discretized setting and restrict to increments $\mathbf{f}_{\text{new}} - \mathbf{f}_{\text{ref}} \in \operatorname{range}(\mathbf{P}^*)$ as motivated by the theory of GenSART. *Example*:

$$\overset{\text{(discretize)}}{\rightsquigarrow} \mathbf{f}_{\text{new}} \in \underset{\mathbf{f} \in \mathbb{R}^n}{\operatorname{argmin}} S(\mathbf{P}(\mathbf{f})) + \alpha \|\mathbf{f} - \mathbf{f}_{\text{ref}}\|_2^2 \quad (6.2.33)$$

$$\overset{\text{(GenSART)}}{\rightsquigarrow} \mathbf{f}_{\text{new}} \in \underset{\mathbf{f} \in \mathbf{f}_{\text{ref}} + \operatorname{range}(\mathbf{P}^*)}{\operatorname{argmin}} S(\mathbf{P}(\mathbf{f})) + \alpha \|\mathbf{f} - \mathbf{f}_{\text{ref}}\|_2^2 \quad (6.2.34)$$

$$= \mathbf{f}_{\text{ref}} + \mathbf{P}^* \left(\underset{\mathbf{p} \in \mathbb{R}^m}{\operatorname{argmin}} S(\mathbf{P}(\mathbf{f}_{\text{ref}}) + \mathbf{P}\mathbf{P}^*(\mathbf{p})) + \alpha \|\mathbf{P}^*\mathbf{p}\|_2^2 \right) \quad (6.2.35)$$

The benefit of the approach is that the iterates \mathbf{f}_{new} now solve the discrete optimization problem, though in general only within a smaller search set.

The pre-discretization approach is expected to be numerically more stable because the regularization by the L^2 -penalty is applied in discrete space and may thus also suppress *discretization errors*. In the post-discretization approach, on the contrary, the regularizing iterations are derived without reference to the final, discretized setting so that inaccuracies due to finite sampling are not controlled by the regularization. In practice, it is found that both methods typically yield results that are almost indistinguishable, whereas the second approach often amounts to more computational effort: for once, the (sparse) matrix $\mathbf{P}\mathbf{P}^*$ has to be assembled, which may be costly. Secondly, in settings where the optimization in (6.2.31) factorizes into scalar problems (compare §6.2.5), this carries over to (6.2.32), yet typically not to (6.2.35) due to the non-diagonal structure of $\mathbf{P}\mathbf{P}^*$. Therefore, post-discretization is usually preferable.

6.2.7 Numerical examples

6.2.7.1 Robust tomographic reconstruction

In the first numerical example, GenSART-schemes are considered for the setting of robust tomography, i.e. tomographic reconstruction from data with large outliers. The principal idea is to test the data-fidelity terms for “robust-fitting” proposed in §6.2.5.1.

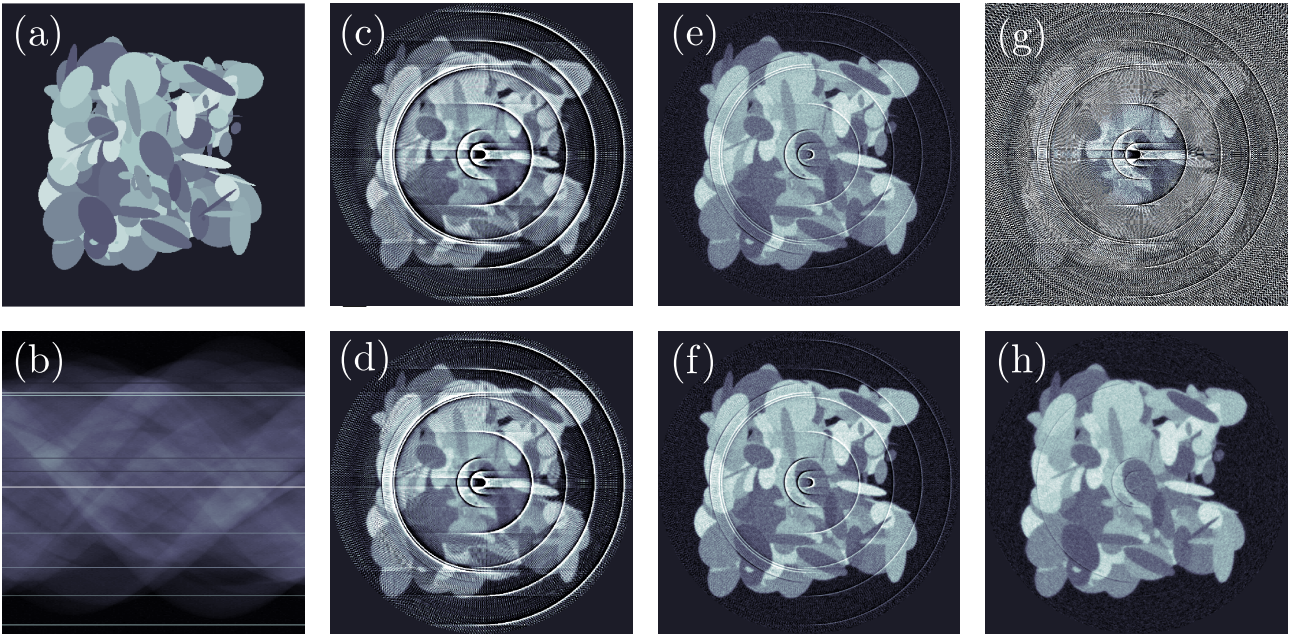


Figure 6.1. Robust tomography test case: (a) True simulated 2D-object. (b) Simulated projection-data: sinogram-plot showing the stripe-artifacts. (c)-(h) Reconstructions from the data in (b) using different methods: (g) filtered back-projection. (c),(e) Tikhonov-reconstruction with L^2 - and L^1 -Huber data-fidelity (6.2.36). (d),(f),(h) GenSART-reconstruction with L^2 - and L^1 -Huber- and Student's-t data fidelity (6.2.37). The linear color scale is the same in all subfigures except (b).

The study considers projection-data $(g_j^{\text{obs}})_{j=1}^{180} \approx (\mathcal{P}_j(f))_{j=1}^{180}$ for tomographic incident angles $\theta = 0^\circ, 1^\circ, \dots, 179^\circ$, simulated for the 2D-object[‡] f shown in fig. 6.1(a). In addition to noise, the tomographic data is corrupted by isolated *stripe-artifacts*, as visualized by the sinogram-plot in fig. 6.1(b). Figure 6.1(c)-(h) shows reconstructions computed by different methods, using data-fidelities of the form $\mathcal{S}(g_j^{\text{obs}}; g_j) = \int s_-(g_j(x) - g_j^{\text{obs}}(x)) dx$ as defined in (6.2.22):

- Figure 6.1(g): Standard filtered back-projection (FBP, see §2.7.3.1)
- Figure 6.1(c),(e): Tikhonov regularization with (c) L^2 - and (e) L^1 -Huber-data-fidelity:

$$f^{\text{Tikh}} = \operatorname{argmin}_{f \in L^2(\Omega)} \left(\sum_{j=1}^{N_{\text{proj}}} \mathcal{S}(g_j^{\text{obs}}; \mathcal{P}_j(f)) \right) + \alpha \|f\|_{L^2}^2, \quad s_- \in \{|\cdot|^2, s_{L^1_{\text{H}}, \nu}\} \quad (6.2.36)$$

- Figure 6.1(d),(f),(h): Kaczmarz-iterations with (d) L^2 -, (f) L^1 -Huber- and (h) Student's-t data-fidelity over one symmetric Kaczmarz-cycle (i.e. $k_{\text{stop}} = 2N_{\text{proj}}$ and $j_{k_{\text{stop}}-k} = j_{k-1}$), computed via GenSART-schemes (post-discretization strategy):

$$f_{k+1} = \operatorname{argmin}_{f \in L^2(\Omega)} \mathcal{S}(g_{j_k}^{\text{obs}}; \mathcal{P}_{j_k}(f)) + 2\alpha \|f - f_k\|_{L^2}^2, \quad s_- \in \{|\cdot|^2, s_{L^1_{\text{H}}, \nu}, s_{\text{s-t}, \nu}\} \quad (6.2.37)$$

It is found that the reconstructions with the L^1 -Huber and Student-t data-fidelities perform well up to slight ring-artifacts, whereas the L^2 -results and especially the FBP-reconstruction

[‡]Although the presented analysis has been carried out for a fully 3D tomographic setting, note that all results trivially generalize to the limiting case of 2D-objects and 1D-tomographic projections.

are strongly corrupted. Moreover, by comparing fig. 6.1(c),(d) and fig. 6.1(e),(f), respectively, it can be seen that the Tikhonov- and Kaczmarz methods with the same data-fidelities perform almost identically, in accordance with theoretical predictions detailed in the complete article. Notably, all of the GenSART-methods (including the one with the *non-convex* Student's-t-term!) require the same computational effort, which is essentially that of two evaluations of the full forward- and back-projectors $\mathcal{P}_{\text{tot}} = (\mathcal{P}_1, \dots, \mathcal{P}_{N_{\text{proj}}})$ and $\mathcal{P}_{\text{tot}}^*$. On the contrary, the minimization of the Tikhonov-functional in (6.2.37) has to be carried out iteratively, which requires several costly evaluations of \mathcal{P}_{tot} and $\mathcal{P}_{\text{tot}}^*$ to ensure convergence (more than 50 in the case of the L^1 -Huber-penalty). Hence, the numerical study shows the following:

Kaczmarz-type methods implemented via GenSART-schemes may serve as a computationally efficient surrogate for Tikhonov regularization.

6.2.7.2 Colloid-XPCT revisited

In the second numerical example, GenSART is applied to the XPCT-problem by implementing the Newton-Kaczmarz-update in (6.2.28) via the post-discretization approach and supplementing it with optional non-negativity constraints (see §6.2.5.4). This leads to algorithm 6.1, which is not limited to XPCT but is valid for arbitrary differentiable image-formation maps F .

The quadratic optimization problems in projection-space are solved by a conjugate-gradient (CG-) method applied to the associated normal equations, which are positive-definite linear systems and thus uniquely solvable. Importantly, the discrete gradient ∇ as well as the element-wise divisions by powers of the unit-projections, $\mathbf{U}_{j_k}^{-1/2}$ and $\mathbf{u}_{j_k}^{-1/2} \odot$, must be implemented in accordance with convention 6.3: the operations have to be evaluated only on those vector-entries that belong to the discrete analogue of the projection-domain $\mathbb{D}_{\mathcal{P}_{j_k}}$, defined by the indices of the non-zero entries of \mathbf{u}_{j_k} – for all vector-entries outside this discrete support, the value 0 must be assigned. Details are discussed in the complete article, see §11.6.3.

Algorithm 6.1 Newton-Kaczmarz-GenSART (for XPCT)

Input: initial guess $\mathbf{f}_0 \in \mathbb{C}^n$, discretized forward map $F : \mathbb{C}^{m_{\text{proj}}} \rightarrow \mathbb{R}^{m_{\text{data}}}$, discrete gradient $\nabla \in \mathbb{R}^{m_{\text{grad}} \times m_{\text{proj}}}$ (finite differences), regularization parameters $\{\alpha_{k,i}\} \subset \mathbb{R}_{\geq 0}$, $0 \leq \gamma \leq 1$

Input: (for $j = 1, \dots, N_{\text{proj}}$) data $\mathbf{g}_j^{\text{obs}} \in \mathbb{R}^{m_{\text{data}}}$, projectors $\mathbf{P}_j \in \mathbb{R}^{m_{\text{proj}} \times n}$, unit-projections $\mathbf{u}_j \in \mathbb{R}_{\geq 0}^{m_{\text{proj}}}$, $\mathbf{U}_j := \text{diag}(\mathbf{u}_j)$, $\mathbf{U}_{\nabla,j} \in \mathbb{R}^{m_{\text{grad}} \times m_{\text{grad}}}$: discretization of $\nabla(p) \mapsto \mathbf{u}_j \cdot \nabla(p)$.

for $k = 0, \dots, k_{\text{stop}} - 1$ **do**

$\alpha = \alpha_{k,1} + \alpha_{k,2}$, $\mathbf{f}_k^{\text{ref}} = (\alpha_{k,1}\mathbf{f}_k + \alpha_{k,2}\mathbf{f}_0)/\alpha$ % trivial step if $\alpha_{k,2} = 0$

$\mathbf{p}_k^{(\text{ref})} = \mathbf{P}_{j_k}\mathbf{f}_k^{(\text{ref})}$, $\mathbf{T}_k = F'[\mathbf{p}_k]$, $\mathbf{r}_k = \mathbf{g}_{j_k}^{\text{obs}} - F(\mathbf{p}_k) - \mathbf{T}_k(\mathbf{p}_k^{\text{ref}} - \mathbf{p}_k)$

% Optimize in projection-space by solving the normal equation with a CG-method:

$\Delta\mathbf{p}_k = \left(\mathbf{U}_{j_k}^{1/2}\mathbf{T}_k^*\mathbf{T}_k\mathbf{U}_{j_k}^{1/2} + \alpha(1-\gamma)\mathbf{I} + \alpha\gamma(\mathbf{U}_{j_k}^{-1/2}\nabla^*\mathbf{U}_{\nabla,j_k}\nabla\mathbf{U}_{j_k}^{-1/2}) \right)^{-1} \left(\mathbf{u}_{j_k}^{1/2} \odot \mathbf{T}_k^*(\mathbf{r}_k) \right)$

$\mathbf{f}_{k+1} = \mathbf{f}_k^{\text{ref}} + \mathbf{P}_{j_k}^* \left(\mathbf{u}_{j_k}^{-1/2} \odot \Delta\mathbf{p}_k \right)$

$[\mathbf{f}_{k+1} = \max\{0, \text{Re}(\mathbf{f}_{k+1})\} + i \max\{0, \text{Im}(\mathbf{f}_{k+1})\}]$ % optional non-negativity constraint

end for

Output: final object iterate \mathbf{f}_k

For validation of the algorithm, it is applied to the ‘‘Colloid XPCT’’ data set from Article 3 (using the same constraints: support, non-negativity and pure phase object) and the results are

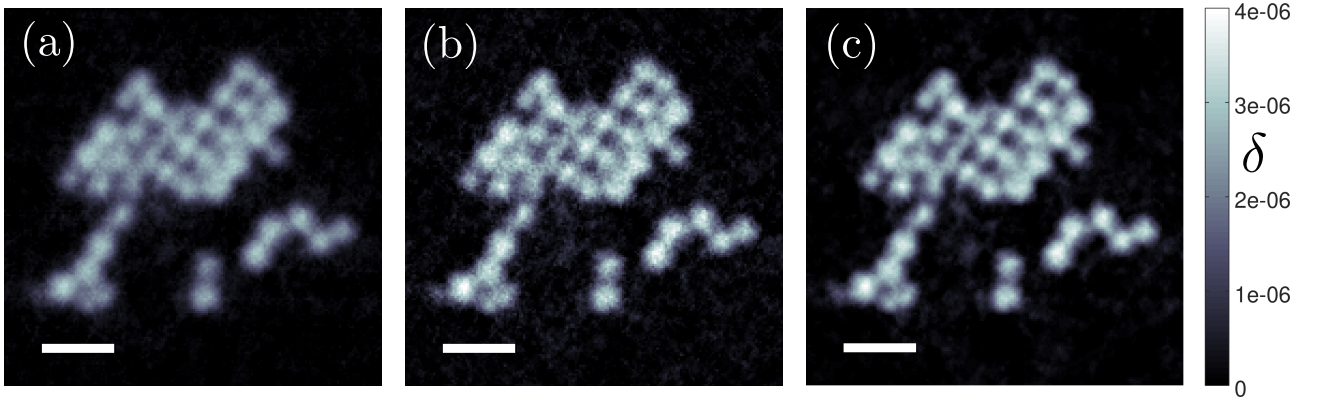


Figure 6.2. Plots of the central slice of the reconstructed object for the “Colloid XPCT” data set from Article 3 (see §5.2.3), obtained by different methods using the same support-, negativity- and pure-phase-object constraints. Scale bars are $1\ \mu\text{m}$.

(a) L^2 -regularized Newton-Kaczmarz reconstruction from §5.2.3 without GenSART.

(b) Newton-Kaczmarz-GenSART (algorithm 6.1) with $\gamma = 0$ (pure L^2 -penalty).

(c) Newton-Kaczmarz-GenSART (algorithm 6.1) with $\gamma = 0.8$ (smoothing $W^{1,2}$ -penalty).

compared to previous ones from §5.2.3. In order to assess the achievable reconstruction quality at minimum possible computational costs, algorithm 6.1 is run for only a single Kaczmarz-cycle, i.e. $k_{\text{stop}} = N_{\text{proj}}$, at relatively weak regularization, $\alpha_{k,1} = \alpha = 500$, $\alpha_{k,2} = 0$. Figure 6.2 plots central slices (zoomed to the relevant region) of the reconstructed objects from (a) Article 3 (L^2 -regularized Newton-Kaczmarz without GenSART, compare §5.2.3), (b) algorithm 6.1 with $\gamma = 0$ (pure L^2 -penalty) and (c) algorithm 6.1 with $\gamma = 0.8$ (smoothing Sobolev $W^{1,2}$ -penalty).

For the L^2 -regularized reconstructions, the results from Article 3 (fig. 6.2(a)) are less noisy but also exhibit weaker contrast and stronger blur than the results from algorithm 6.1 in fig. 6.2(b). Overall, the achieved image-quality is thus comparable, yet the level of regularization is different. On the contrary, the $W^{1,2}$ -regularized reconstruction in fig. 6.2(c) combines the contrast and edge-sharpness in (b) with the low noise-level in (a) and thus arguably outperforms both reconstructions with pure L^2 -penalty. An explanation is that $W^{1,2}$ -penalties enforce a strong damping of high spatial frequencies, where the impact of noise is typically most pronounced, whereas coarser, low-frequency object-structures are only negligibly affected. This selective high-frequency penalization is apparently more “to the point” for the present data set than L^2 -regularization, which *uniformly* penalizes in all Fourier-frequencies.

Sobolev-penalties could in principle also be applied within the original Newton-Kaczmarz scheme from Article 3. However, this would involve evaluations of a discrete gradient in the *high-dimensional* object-space, each amounting to at least $\mathcal{O}(N^3)$ floating-point operations, and thus greatly increase the computational costs[§]. On the contrary, in the GenSART-algorithm 6.1, only the *low-dimensional* linear system (corresponding to the optimization in projection-space in the general scheme 6.6) is affected when modifying the penalty-term, so that the computational costs increase only slightly. In a nutshell, this demonstrates the following:

GenSART renders advanced regularization in XPCT computationally affordable.

[§]Notably, not only would the computation of each CG-iteration become more expensive but gradient-penalties typically also deteriorate the condition of the problem, so that more iterations are needed for convergence.

6.3 Supplement: High-performance implementation

The main motivation for the generalized SART-schemes was to render the computation of (Newton-)Kaczmarz iterations efficient in terms of computational complexity, i.e. to minimize the number of required *floating-point operations* (FLOPs). This has been achieved. However, *FLOP-efficiency* of an algorithm alone does not guarantee that it also permits an efficient implementation in terms of actual *runtimes* on computers. This supplement therefore outlines implementational details that permit to translate the *theoretical* efficiency of GenSART to competitive computation times *in practice*. While the presentation will sometimes refer to XPCT in order to identify relevant numerical problem-dimensions, the outlined implementational issues and -solutions apply to GenSART-schemes for arbitrary tomographic reconstruction problems. Moreover, although the discussion of computational topics evidently only makes sense for *discretized* settings, note that we will nevertheless use continuous quantities $f_k, \mathcal{P}_j, p_k, u_j$ etc. as notation for easier reference to the GenSART-theory from §6.2.

6.3.1 General considerations

6.3.1.1 Computational- and data-complexity

For an overview of the possible bottlenecks in implementing GenSART it is illustrative to recall the general structure of scheme 6.6 and to assign for each sub-step: (a) the amount of data that needs to be accessed (data-complexity) and (b) the required number of FLOPs (computational complexity), both in $\mathcal{O}(N^\alpha)$ -notation, where N is the number of sampling points (pixels, voxels) along each dimension. Note that we also consider optional box constraints on the iterates f_{k+1} , compare §6.2.5.4. The identified complexities are summarized in table 6.1.

	FLOP-complexity	Data-complexity
(1) Forward-project: $p_k = \mathcal{P}_{j_k}(f_k)$	$\mathcal{O}(N^3)$	3D-volume: $\mathcal{O}(N^3)$
(2) Optimize: $\Delta p_k = \operatorname{argmin}_p J(p, p_k, u_{j_k}, g_{j_k}^{\text{obs}})$	$\geq \mathcal{O}(N^2)$	2D-images: $\mathcal{O}(N^2)$
(3) Back-project: $\tilde{f}_{k+1} = f_k + \mathcal{P}_{j_k}^*(\Delta p_k)$	$\mathcal{O}(N^3)$	3D-volume(s): $\mathcal{O}(N^3)$
((3b) $f_{k+1} = \max(\min(\tilde{f}_{k+1}, f_{\max}), f_{\min})$)	$\mathcal{O}(N^3)$	3D-volume(s): $\mathcal{O}(N^3)$

Table 6.1. Complexity of the steps in the GenSART-scheme 6.6 (with box constraints) in terms of arithmetic operations (FLOP-complexity) and memory-requirements due to the amount of involved data (data-complexity). N is the number of sampling points (pixels, voxels) along each dimension.

Computational bottlenecks: For the XPCT-problem, all sub-steps typically involve a comparable number of computations as the optimization in projection-space (2) has to be performed iteratively, compare algorithm 6.1, and each iteration involves 2D-FFTs amounting to $\mathcal{O}(N^2 \log N)$ FLOPs. Notably, this means that *all* rows in table 6.1 have to be implemented at a similar efficiency-level in order for none of the sub-steps (1) to (3b) to bottleneck the whole algorithm. In other words: all sub-steps are *performance-critical*. On the contrary, in the settings identified in §6.2.5.1 where the optimization-step (2) admits an analytical solution in $\mathcal{O}(N^2)$ operations, its efficiency has a negligible impact on the total computation time.

Memory bottlenecks: For typical problem sizes $N \sim 1000$, arrays of size $\mathcal{O}(N^2)$ in single- or double-floating point precision occupy (tens or hundreds of) *Megabytes* in main memory (RAM), whereas storing a 3D-object volume of size $\mathcal{O}(N^3)$ requires (tens or hundreds of) *Gigabytes* of RAM. As current PCs and workstations typically come with a few to a few hundred Gigabytes main-memory, this means that operations **(1)**, **(3)** and **(3b)** in table 6.1 are *memory-critical*, whereas physical memory-limits are practically irrelevant for the optimization-step **(2)**.

6.3.1.2 Minimizing memory-requirements

In-place arithmetics: Clearly, the current iterate f_k has to be stored in RAM while running a GenSART-algorithm. When naively implementing step **(3)** in table 6.1, one or more additional arrays of the same $\mathcal{O}(N^3)$ size will be created temporarily. In view of memory-limits, this should be avoided, which is possible by implementing the update $\tilde{f}_{k+1} = f_k + \mathcal{P}_{j_k}^*(\Delta p_k)$ as an *in-place* operation: the computed samples of the back-projection $\mathcal{P}_{j_k}^*(\Delta p_k)$ are directly added to the array storing f_k . After the execution, the array will then be overwritten (by a discretization of) the sought update \tilde{f}_{k+1} . Analogously, the optional box constraints in step **(3b)** can be implemented *in-place*, by element-wise overwriting \tilde{f}_{k+1} with f_{k+1} .

Minimal memory-requirements of GenSART-schemes: If the outlined in-place approach is pursued, GenSART-algorithms thus require only *one* 3D-object-array of size $\mathcal{O}(N^3)$ to be stored in RAM throughout the whole execution. From table 6.1 it is seen that all other data accessed in GenSART-iterations are of size $\mathcal{O}(N^2)$ and thus negligible. Two restrictions to these general observations might come to the reader's mind:

- *What about the measured data?* The total size of the tomographic data $(g_1^{\text{obs}}, \dots, g_{N_{\text{proj}}}^{\text{obs}})$ is typically also $\mathcal{O}(N^3)$ and may occupy even more memory than the object-iterates f_k in absolute terms. However, only small data-chunks $g_{j_k}^{\text{obs}}$ of size $\mathcal{O}(N^2)$ are processed in one iteration. As a benefit, the total data may be stored in *external memory*, from which the required data-chunks are loaded into RAM *on demand*, at the cost of negligibly increased runtimes. Note that this idea may well be put into the extreme, see remark 6.10 below.
- *Support constraints:* Complicated geometries of the support-domain $\Omega \supset \text{supp}(f_k)$, i.e. others than standard choices like cuboids, cylinders, balls etc., may require to retain an explicit description of it in RAM. This amounts to storing an additional 3D-array, whose $\mathcal{O}(N^3)$ boolean entries (true or false) tell whether or not a given voxel in the computational domain belongs to Ω . As boolean values occupy one Byte of memory, compared to four Bytes for each entry of the f_k , complicated support constraints thus induce additional memory-requirements of 1/4 the size of the object-iterate f_k .

All in all, the discussion shows that GenSART-algorithms may be implemented at appealingly low memory-requirements: typically, available RAM of 1.5-times the size of the object-iterate f_k in single-precision is sufficient without limiting computational performance.

Remark 6.10 (Online capabilities of GenSART). In GenSART-algorithms (and other Kaczmarz-type methods), newly acquired tomographic views $g_{j_k}^{\text{obs}}$ may be directly used to update the object-iterates f_k , without requiring a complete tomographic series to be available at any time. This makes the approach well-suited for *online* or *time-resolved* tomography.

6.3.2 Parallelization

As the clock rate of processors has practically reached its physical limits, high-performance-computations necessarily rely on *parallelization*: tasks and data are distributed to different processors, that perform the computations in parallel as far as possible. The significance of parallelization can be seen from a simple comparison: single cores of the main processor (CPU) on current computers perform $\mathcal{O}(10^9)$ FLOPs per second, whereas a mid-price GPU (*graphical processing unit*) can easily achieve $\mathcal{O}(10^{12})$ FLOPs per second via massive parallelization. Notably, this means that algorithmic tasks of computational complexity $\mathcal{O}(N^3)$ (such as evaluation of \mathcal{P}_j for *one* tomographic angle) may *in practice* take as long to execute as others requiring $\mathcal{O}(N^4)$ FLOPs (e.g. an evaluation of $\mathcal{P}_{\text{tot}} = (\mathcal{P}_1, \dots, \mathcal{P}_{N_{\text{proj}}})$ for *all* angles) even for large problems of size $N \sim 10^3$.

The example shows that parallelization is crucial to obtain *practically* efficient GenSART-algorithms. Issues and solutions in parallelizing GenSART are discussed in the following.

6.3.2.1 Parallelization of the optimization in projection-space

As discussed in §6.3.1.1, the optimization step (2) in table 6.1 typically requires amounts of RAM much smaller than one Gigabyte since only 2D-arrays (images) are processed. Owing to this, the step may be parallelized at relative ease, by simply executing the optimization algorithm of choice completely on a GPU. Thereby, optimized *off-the-shelf* parallelizations of the involved operations are used, as provided by available packages in `Matlab` or `Python`, for example. Notably, as both the input-data $p_k, g_{jk}^{\text{obs}}, u_{jk}$ and the output increment Δp_k are of size $\mathcal{O}(N^2)$, the required memory-transfers from- and to the GPU are typically negligible.

6.3.2.2 Parallelization of forward- and back-projections

At the first glance, off-the-shelf parallelizations seem available also for the required forward- and back-projections $\mathcal{P}_j, \mathcal{P}_j^*$. The ASTRA-toolbox [199, 198], for example, is a software-package that provides CPU- and GPU-based implementations of tomographic projectors and -back-projectors as well as of FBP and other basic reconstruction algorithms. Notably, the massively parallel GPU-algorithms can be *blazingly fast*: on a Nvidia GeForce GTX TITAN, it computes a full forward- or back-projections for a (real-valued) object of size 1000^3 voxels under 1000 incident angles ($\mathcal{O}(N^4)$ operations with $N = 10^3$) in about 15 seconds[¶]. Appealing though these results are, the timed operations are not quite what is needed in implementations of GenSART-schemes: here, forward- and back-projections $\mathcal{P}_j, \mathcal{P}_j^*$ need to be computed for only a *single* tomographic incident angle at a time. So how does ASTRA perform on this task? Runtime tests yield about 2 seconds for \mathcal{P}_j and ≈ 3 seconds for a \mathcal{P}_j^* (1000^3 -sized volume, 1000^2 -sized projection). Note that this is only 5- to 8-times faster than evaluations of \mathcal{P}_{tot} and $\mathcal{P}_{\text{tot}}^*$, although the latter amount to 1000-times more computations to be performed!

So what is going wrong here? Is it that ASTRA's parallelization of (back-)projection-algorithm is less efficient if it is run for only one tomographic angle at a time? Partly so, but a closer look reveals this not to be the main issue here. Instead, the main computational bottleneck turns out to be the necessary transfers between main-memory (RAM) and dedicated

[¶]All runtimes have been obtained by calling ASTRA from within `Matlab` R2017b on a workstation with current hardware. However, the identified tendencies in runtime are independent of the specific system.

	RAM-to-GPU	GPU-work	GPU-to-RAM	Runtime ($N=10^3$)
Full projection \mathcal{P}_{tot}	$\mathcal{O}(N^3)$	$\mathcal{O}(N^4)$	$\mathcal{O}(N^3)$	~ 15 s
Full back-projection $\mathcal{P}_{\text{tot}}^*$	$\mathcal{O}(N^3)$	$\mathcal{O}(N^4)$	$\mathcal{O}(N^3)$	~ 15 s
Single projection \mathcal{P}_j	$\mathcal{O}(N^3)$	$\mathcal{O}(N^3)$	$\mathcal{O}(N^2)$	~ 2 s
Single back-projection \mathcal{P}_j^*	$\mathcal{O}(N^2)$	$\mathcal{O}(N^3)$	$\mathcal{O}(N^3)$	~ 3 s

Table 6.2. Complexity (in terms of the resolution N along each spatial dimension) and runtimes of GPU-based tomographic (back-)projection algorithms provided by the ASTRA-toolbox [199, 198], for a full set of N tomographic angles ($\mathcal{P}_{\text{tot}}, \mathcal{P}_{\text{tot}}^*$) and for a single angle ($\mathcal{P}_j, \mathcal{P}_j^*$), respectively. Complexity is analyzed w.r.t. FLOPs to be performed on the GPU (“GPU-work”) and w.r.t. required transfers between main- and GPU-memory (“RAM-to-GPU” and “GPU-to-RAM”). Runtimes have been tested on a current workstation with a Nvidia GeForce GTX TITAN GPU.

GPU-memory: evaluations of $\mathcal{P}_j, \mathcal{P}_j^*$ always require a full 3D object-volume to be transferred to or from the GPU. As seen from a detailed comparison of the operations in table 6.2, this amounts to the same $\mathcal{O}(N^3)$ of data-transfers as in the case of $\mathcal{P}_{\text{tot}}, \mathcal{P}_{\text{tot}}^*$ (up to a factor of 2) – only the GPU-work increases in the latter case ($\mathcal{O}(N^4)$ instead of $\mathcal{O}(N^3)$). As RAM-to-GPU- and especially GPU-to-RAM-transfers are typically much slower than GPU-computations, the former constitute a severe bottleneck for evaluations of \mathcal{P}_j and \mathcal{P}_j^* .

Avoiding memory-transfers: Besides ASTRA’s default behavior where object-volume and projections are copied to and from the GPU-memory prior to and after each (back-)projector-evaluation, the package also provides (experimental) support of *persistent GPU-memory*. Accordingly, the object-volume can be retained in dedicated GPU-RAM, avoiding the bottlenecking $\mathcal{O}(N^3)$ memory-transfers. Upon inserting typical problem-sizes, however, this option is seen to be a purely theoretical one for high-resolution XPCT-reconstructions: storing one 2048^3 -sized real-valued object volume in single-precision already requires 32 Gigabytes, whereas current commercially available GPUs come with at most 16 Gigabytes of dedicated memory.

6.3.3 Efficient implementation by distributed design

According to the preceding section, we have the following computational deadlock-situation:

- Massively parallel GPU-based computations promise massive performance gains
- Memory-transfers between RAM and GPU in each GenSART-iteration are prohibitive
- Typical object-sizes in XPCT do not fit the dedicated memory of a *single* GPU

As suggested by the final point, the deadlock may be resolved by a *distributed* algorithmic design: the object-iterates f_k are stored and managed as separate chunks by *multiple* GPUs^{||} and each of these GPUs computes the forward and back-projections from- and onto its dedicated object-chunk. Implementing GenSART-algorithms in a distributed manner is quite simple as only projections and back-projections (up to optional pointwise operations) are performed on

^{||}Note that the approach is equally applicable for multiple CPU-cores or any other machines.

the 3D-object, whereas the complicated part of the algorithm works completely in projection-space and is thus unchanged. All of the operations on 3D-objects turn out to be well-suited for distributed implementation as they require quasi no communication between the machines and thus induce practically no computational overhead. This is detailed in the following.

6.3.3.1 Distributed operations

The principal idea for distributed GenSART lies in a disjoint decomposition of the object-domain $\Omega = \bigcup_{j=1}^{n_{\text{chunk}}} \Omega_i$, where each of the GPUs manages the object-iterates f_k restricted to a specific subdomain Ω_i , i.e. (a discretized version of) $f_{k,i} := f_k|_{\Omega_i}$. The objects may then be reassembled by summation, $f_k = \sum_{i=1}^{n_{\text{chunk}}} f_{k,i}$. Importantly, however, none of the steps in the GenSART-scheme 6.6 requires *explicit* assembly of the object-chunks, as argued below.

Distributed forward-projections: By linearity, forward-projections may be rewritten as

$$\mathcal{P}_j(f_k) = \mathcal{P}_j\left(\sum_{i=1}^{n_{\text{chunk}}} f_{k,i}\right) = \sum_{i=1}^{n_{\text{chunk}}} \mathcal{P}_j(f_{k,i}) \quad (6.3.1)$$

According to (6.3.1), the complete object f_k may be forward-projected by summing projections $\mathcal{P}_j(f_{k,i})$ of the individual chunks, which can be computed completely independently by the different GPUs. The computational costs of the final summation are at most $\mathcal{O}(n_{\text{chunk}}N^2)$.

Distributed back-projections: The back-projection step in GenSART may be rewritten as

$$\begin{aligned} \sum_{i=1}^{n_{\text{chunk}}} f_{k+1,i} &= f_{k+1} = f_k + \mathcal{P}_j^*(\Delta p_k) = \sum_{i=1}^{n_{\text{chunk}}} f_{k,i} + \mathcal{P}_j^*(\Delta p_k)|_{\Omega_i} \\ \stackrel{\text{"}|_{\Omega_i}"}{\Leftrightarrow} f_{k+1,i} &= f_{k,i} + \mathcal{P}_j^*(\Delta p_k)|_{\Omega_i} \quad \text{for all } i = 1, \dots, n_{\text{chunk}}. \end{aligned} \quad (6.3.2)$$

Hence, the update is performed implicitly if each GPU adds $\mathcal{P}_j^*(\Delta p_k)|_{\Omega_i}$ to its object-chunk, i.e. computes an in-place back-projection onto its subdomain Ω_i . This requires that the (relevant part of the) increment Δp_k is copied to each GPU, i.e. $\leq \mathcal{O}(n_{\text{chunk}}N^2)$ memory-transfers.

Distributed pointwise manipulations: For completeness, let us also note that *pointwise* operations on the object-iterates of the form $f_k(\mathbf{x}) = r(\mathbf{x}, \tilde{f}_k(\mathbf{x}))$, as required when imposing box constraints for example, are fully compatible with the distributed design:

$$\begin{aligned} f_k(\mathbf{x}) &= r(\mathbf{x}, \tilde{f}_k(\mathbf{x})) \quad \text{for all } \mathbf{x} \in \Omega \\ \Leftrightarrow f_{k,i}(\mathbf{x}) &= r(\mathbf{x}, \tilde{f}_{k,i}(\mathbf{x})) \quad \text{for all } \mathbf{x} \in \Omega, i = 1, \dots, n_{\text{chunk}}. \end{aligned} \quad (6.3.3)$$

Slice-wise distribution – the ideal setting: The general approach to distributed GenSART works for arbitrary object- and tomographic acquisition-geometries (including cone-beam!). The standard setting in XPCT and in many other imaging modalities, however, are parallel-beam projections acquired for rotations around a fixed tomographic axis. In such a setting, the most straightforward and also the most efficient distributed design lies in choosing

the sub-volumes Ω_i as *slabs* (i.e. stacks of multiple slices) of the object-volume perpendicular to the tomographic axis, i.e. if the latter is along the y -dimension, then

$$\Omega_i := \{(x, y, z) \in \Omega : y_{i-1} \leq y < y_i\} \quad \text{for some} \quad y_0 < y_1 < \dots < y_{n_{\text{chunk}}}. \quad (6.3.4)$$

One benefit is that the assembly of the forward-projections in (6.3.1) reduces to a mere *concatenation* of the segments computed by the distinct GPUs: for all $(x, y) \in \mathbb{R}^2$ it holds that

$$\mathcal{P}_j(f_k)(x, y) = \sum_{i=1}^{n_{\text{chunk}}} \mathcal{P}_j(f_{k,i})(x, y) = \begin{cases} \mathcal{P}_j(f_{k,i_y})(x, y) & \text{if } y \in [y_{i_y-1}; y_{i_y}) \\ 0 & \text{else} \end{cases}, \quad (6.3.5)$$

which reduces the costs of the step to $\mathcal{O}(N^2)$ memory-transfers. An analogous reduction occurs in the back-projection step, as only the segment $\Delta p_k|_{\{y_{i-1} \leq y < y_i\}}$ has to be copied to the i th GPU.

6.3.3.2 Complete distributed tomographic reconstruction

§6.3.3.1 shows that GenSART-iterations may be implemented in a distributed manner at relative ease and at the expense of a computational overhead of at most $\mathcal{O}(n_{\text{chunk}}N^2)$ – which is negligible compared to the $\mathcal{O}(N^3)$ computational costs (provided that $n_{\text{chunk}} = \mathcal{O}(1)$). In a nutshell, the underlying idea is to carry out the required communication between the individual GPUs completely in the (low-dimensional) projection-space – very much in the spirit of the basic principle of GenSART. For a complete tomographic reconstruction, this means that the individual object-chunks may be stored in the dedicated memory of the managing GPUs (or CPUs) over all GenSART-iterations. Only upon *initializing* and *finalizing*, the initial guess f_0 needs to be copied as chunks $f_{0,i} = f_0|_{\Omega_i}$ to the GPUs and the final iterate needs to be assembled in RAM, $f_{k_{\text{stop}}} = \sum_{i=1}^{n_{\text{chunk}}} f_{k_{\text{stop},i}}$, respectively. Both steps amount to $\mathcal{O}(N^3)$ of additional communication, but notably need to be performed only *once*.

All in all, we thus find that a total distributed GenSART-reconstruction requires at most $\mathcal{O}(N^3) + \mathcal{O}(n_{\text{chunk}}k_{\text{stop}}N^2)$ of memory-transfers compared to $\geq \mathcal{O}(k_{\text{stop}}N^3)$ FLOPs of parallelized (GPU-)computations. Hence, contrary to a naive ASTRA-based implementation, see §6.3.2.2, communication no longer bottlenecks the overall computational performance.

6.3.3.3 Realization on current hardware

Finally, we bring the ideas of this section together to argue that an efficient, parallelized version of GenSART can actually be implemented on current hardware. From §6.3.1.2, we know that about $1.5 \cdot 32 = 48$ Gigabytes of *total* RAM suffice to reconstruct real-valued object-volumes of the typical XPCT-size 2048^3 . Within the distributed design, each machine thus needs $48/n_{\text{chunk}}$ Gigabytes of dedicated RAM, which permits the following implementations:

- (1) *Fully GPU-based* ($4(+1)$ GPUs): The operations on the 3D-volume (\mathcal{P}_j , \mathcal{P}_j^* and point-wise manipulations) are handled by four GPUs, each with ≥ 12 Gigabytes of dedicated RAM. One of these GPUs (or an extra one) performs the optimization in projection-space.
- (2) *Mixed CPU-GPU-based* (X CPU-kernels + 1 GPU): The distributed design is instead implemented on CPU-kernels, which perform all required operations on 3D-volumes in parallel. The optimization problem in projection-space is solved by a single GPU.

Both options can be realized on commercially available hardware. **(1)** is potentially faster, but also (financially) more expensive due to the high costs of dedicated GPU-RAM. On the other hand, the hardware-requirements of **(2)** are typically met by standard workstations.

6.4 Supplement: Application to a large-scale data set

Although the Newton-Kaczmarz-GenSART algorithm 6.1 has already been successfully applied to the “colloid-XPCT” data set in §6.2.7.2, it could be argued that this example corresponds to an artificially simple XPCT-setting for the following reasons:

- *Small problem size:* With only 249 holograms of size 1024^2 pixels, the data set is unusually small. Owing to the small support of the imaged object, the reconstructed object-volume could furthermore be restricted to a computational domain of only 256^3 voxels.
- *Weak object:* As the imaged colloidal crystal induces phase-shifts $|\phi| \leq 0.4$, it can be treated as a weak object to good approximation. Consequently, accurate reconstructions can also be achieved with *linear* CTF-based methods instead of a Newton-type scheme.

The aim of the following supplement is therefore two-fold: for once, we demonstrate that GenSART-reconstructions may be computed in acceptable runtimes for typical large-scale synchrotron XPCT data via the high-performance implementation from §6.3. On the other hand, we will consider a data set where linear CTF-based reconstruction truly reaches its limits and show how our *nonlinear* Newton-Kaczmarz approach helps to overcome these.

6.4.1 Description of the data set

We consider XPCT-data of a catalyst-particle, as used in fluid-catalytic cracking in petroleum-refineries, see e.g. [202]. The data has been acquired in the GINIX setup at a photon-energy of $E = 13.8$ keV. According to the researchers involved in the experiment, the main constituents of the particle are clay, zeolith Y and amorphous alumina silica. Additionally, it is expected to contain the heavy metals iron, nickel and vanadium as contaminants.

The data set consists of 4000 holograms of image-size 2048×2048 , acquired under 1000 incident angles θ that uniformly sample the interval $\theta \in [0; 2\pi]$. For each view-angle, holograms have been measured under four different source-to-sample-distances, corresponding to Fresnel numbers $f_{\text{pix}} \in \{9.22 \cdot 10^{-4}, 9.04 \cdot 10^{-4}, 8.53 \cdot 10^{-4}, 7.68 \cdot 10^{-4}\}$ associated with the effective pixel size $\Delta x_{\text{pix}}^{\text{eff}} = 64.4$ nm of the divergent-beam setup (compare §2.1.5.1). The considered hologram-data has been flat-field-corrected, see §2.1.5.2. The four holograms for each tomographic angle have furthermore been rescaled to account for the different geometric magnifications and mutually aligned to correct for undesired motions in the imaging-setup during acquisition. An exemplary set of (pre-processed) holograms for a single view angle is shown in fig. 6.3.

6.4.2 Reconstruction methods

The main constituents of the sample are different ceramic materials, which all have roughly the same absorption-refraction-ratio $0.003 \leq \beta/\delta \leq 0.004$ at the considered photon-energy $E =$

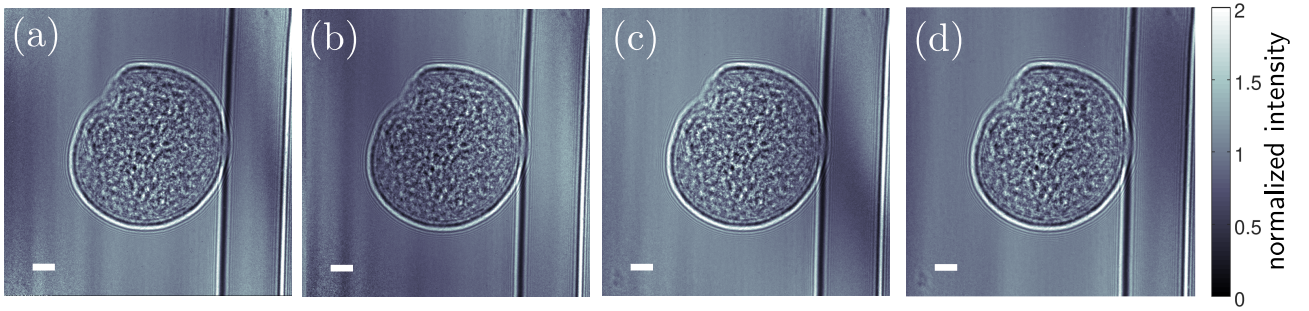


Figure 6.3. Exemplary set of (flat-field-corrected) holograms of the imaged catalyst-particle for one tomographic angle acquired under four different source-to-sample-distances (i.e. Fresnel numbers). The total XPCT-data set consists of four holograms for all 1000 incident angles. Scale bars: 10 μm .

13.8 keV, according to literature values [92]. Therefore, the catalyst-particle is reconstructed as a *homogeneous object* with constant $c_{\beta/\delta} = 0.0035$ and $\nu = \arctan(c_{\beta/\delta})$, compare §2.4.3. From the holograms in fig. 6.3, it can furthermore be seen that the particle is only about half of the field-of-view in diameter, which may be exploited by imposing a *support constraint*. However, this constraint only holds *approximately* as the sample-holder induces some signal (bright and dark vertical lines on the r.h.s. of the holograms in fig. 6.3) that is not consistent with it.

For comparison, the XPCT-data is reconstructed using different methods:

- (1) (*CTF+FBP*) Direct CTF-inversion (parameters: $\alpha_1 = 10^{-4}$ and $\alpha_2 = 10^{-2}$), followed by filtered back-projection of the recovered projection-data, see §2.7.2.1 and §2.7.3.1.
- (2) (*Newton-Kaczmarz*) Reconstruction via algorithm 6.1 with $F = \mathcal{N}_\nu^{(f_1, \dots, f_4)}$, imposing non-negativity as well as a relatively tight support constraint. Further parameters are $\alpha_{k,1} = \alpha = 500$, $\alpha_{k,2} = 0$, $\gamma = 0.9$ and $k_{\text{stop}} = N_{\text{proj}} = 1000$ (single Kaczmarz-cycle).
- (3) (*Linear Kaczmarz*) Same as (2), except that the *nonlinear* XPCI-forward map $\mathcal{N}_\nu^{(f_1, \dots, f_4)}$ is replaced by the corresponding *linear* CTF-forward model, setting $F = \mathcal{L}_\nu^{(f_1, \dots, f_4)}$.

The Kaczmarz-reconstructions are computed via GenSART-schemes, using the parallel implementation proposed in §6.3 in the “Mixed CPU-GPU-based” variant (see §6.3.3.3) on six cores of an Intel Xeon E5-2609 CPU and a Nvidia GeForce GTX TITAN GPU. The CTF+FBP-reconstruction is computed on the same machine via FFT-based CTF-inversion on a single CPU, followed by GPU-based filtered back-projection using the implementation in the ASTRA-toolbox [199, 198] with default parameters. Total computation times are about 30 minutes for CTF+FBP (1) and about five hours for the GenSART-reconstructions (2) and (3).

6.4.3 Results

The 3D-structures of the refractive decrement δ obtained via the different methods are plotted in fig. 6.4(a)–(i) as central orthoslices along the different coordinate-planes ((a)–(c): CTF+FBP, (d)–(f): linear Kaczmarz, (g)–(i): Newton-Kaczmarz). For a detailed comparison of the achieved image-quality, fig. 6.4(j)–(l) additionally plot zooms of the regions marked by the dashed boxes in subfigures (b), (e) and (h), respectively.

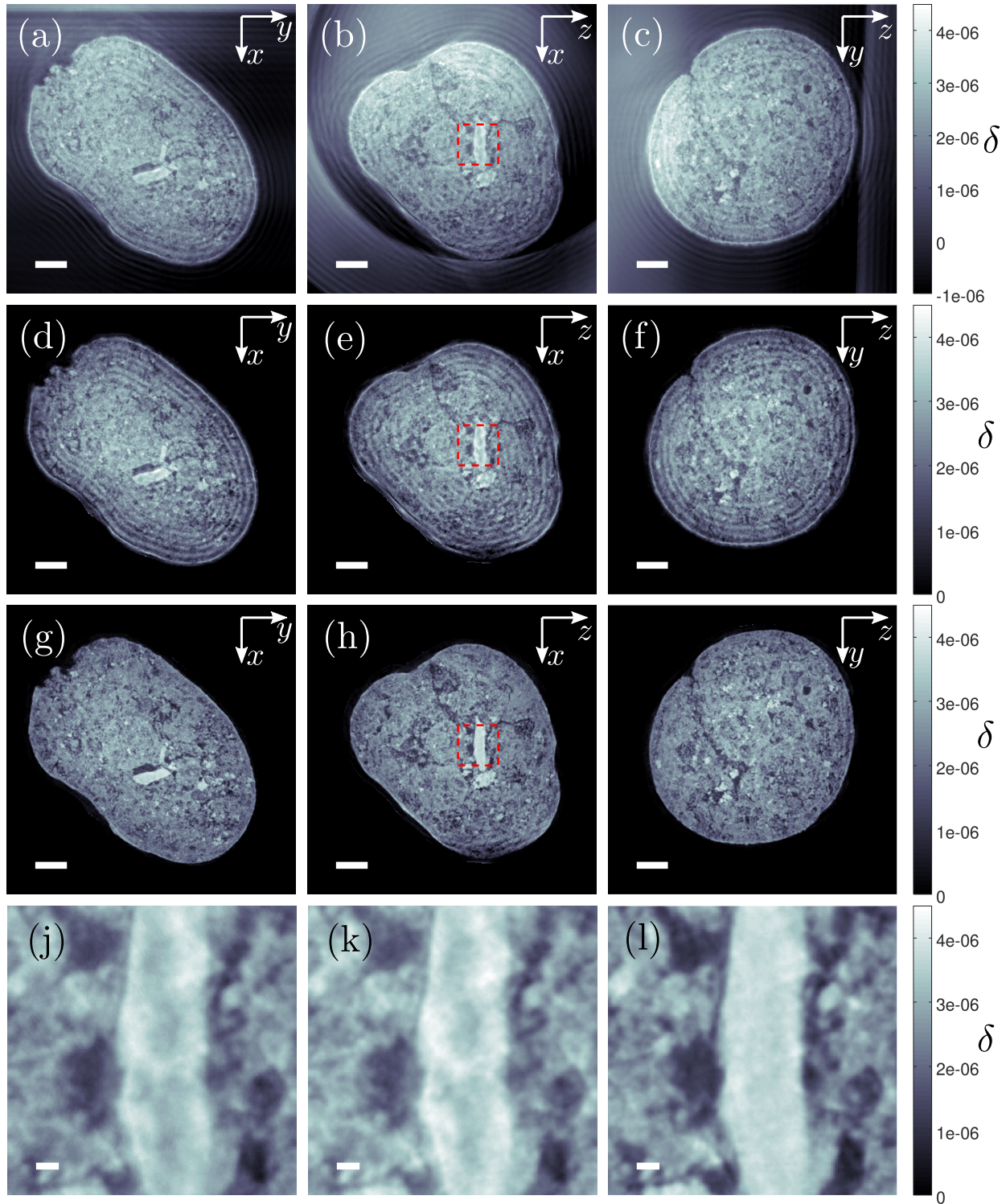


Figure 6.4. XPCT-reconstructions of the catalyst-particle. Subfigures (a)–(i) plot central orthoslices of the recovered 3D-volumes of the refractive decrement δ , where columns correspond to slices along different coordinate planes and rows to different reconstruction methods described in §6.4.2: (a)–(c) CTF+FBP. (d)–(f) Linear Kaczmarz. (g)–(i) Newton-Kaczmarz. Subfigures (j)–(l) show zooms of red-dashed regions in the images in (b), (e) and (h), respectively. Scale bars are $10\ \mu\text{m}$ in (a)–(i) and $1\ \mu\text{m}$ in subfigures (j)–(l).

By comparing subfigures (a)–(c) and (d)–(f), it can be seen that the *linear* methods yield similar results for the considered XPCT-data. The main difference between those two is that the CTF+FBP results show some *low-frequency* variations, both in the background and within the object: for example, the bottom-side of the object in fig. 6.4(b) is overall darker than the upside, but not so in fig. 6.4(e). This difference can be understood by the stability analysis for linearized XPCI (and XPCT) from Article 1: as holograms for multiple Fresnel numbers f_1, f_2, f_3, f_4 are used in the reconstructions, zeros of the individual CTFs $s_\nu^{(f_j)}$ at larger Fourier-frequencies are eliminated as explained in §2.4.4. Yet, the characteristic instability at low Fourier-frequencies $\xi \approx 0$ persists. As seen in chapter 3, the 3D-support constraint and non-negativity of δ , which are imposed by the Kaczmarz method, stabilize especially this low-frequency regime. Since this stabilization is missing for the CTF+FBP reconstruction, *quantitative* perturbations of the recovered δ -values arise from flat-field-correction-related low-frequency errors in the holograms (compare §2.1.5.2 and fig. 6.3). On the contrary, *qualitative*, high-frequency image-structures such as edges come out widely identically as by the linear Kaczmarz method.

In comparison to the *nonlinear* Newton-Kaczmarz reconstruction in fig. 6.4(g)–(i), the most prominent feature of the *linear* results in fig. 6.4(a)–(f) is given by an additional concentric fringe-pattern in the images. The latter structures constitute a severe artifact that can be clearly attributed to nonlinearity: after all, the only difference between the reconstructions in subfigures (d)–(f) and (g)–(i) is that the former are based on the linear (CTF-) forward model, whereas the latter also models nonlinear parts of the contrast in the measured holograms. Notably, the nonlinearity-artifacts are found to distort the linear reconstructions in a qualitative, *structural* manner, whereas the quantitative values of δ are (on average) almost identical in fig. 6.4(d)–(f) and (g)–(i). Indeed, although the fringe-artifacts are most pronounced near the edges, the zooms in fig. 6.4(j)–(l) reveal that nonlinearity corrupts the linear reconstructions even in the very center of the object: the images in (j) and (k) have a more “washed-out” appearance, with lower contrast than the Newton-Kaczmarz result in (l). In particular, note that the bright, vertically elongated structure in the center of the zoom-images, that is characterized by an almost uniform density in the nonlinear reconstruction, shows strong local variations of the δ -values for the linear results in (j) and (k). This example demonstrates that the nonlinearity-artifacts in the recovered images might lead to wrong physical interpretations, emphasizing the necessity of a nonlinear reconstruction method for the considered data set.

While there is no proof that the Newton-Kaczmarz result yields the “true” image of the catalyst-particle, the reconstructed 3D-structure of δ is arguably the one that is most consistent with physical expectation and free of clearly identifiable artifacts. In particular, it is interesting to note that the recovered object shows no traces of *phase-wrapping ambiguities* (see §2.5.1.1), even though the object induces phase-shifts ϕ of magnitude up to $\approx 4\pi$. All in all, the considered reconstructions from real-data thus reveal the following:

Nonlinearity may cause severe artifacts in real-world XPCT if a linear CTF-model is assumed in image reconstruction. Newton-Kaczmarz-GenSART methods can avoid these artifacts and remain computationally feasible even for large-scale data sets.

Further remarks: Interestingly, the artifact-causing nonlinearity of the imaged catalyst-particle was originally overcome not by adjustments on the *algorithmic*- but rather on the *experimental* side: the particle was embedded in wax in order to reduce the jump of the

refractive decrement δ between the medium (formerly air, $\delta = 0$, then wax with $\delta > 0$) and the object. New holograms were acquired for the embedded sample, from which, finally, the particle was reconstructed using CTF+FBP. In other words, *contrast was artificially reduced* in order to permit linear reconstructions. In view of the present results, we see that the usage of nonlinear reconstruction algorithms may help to avoid such “experimental detours”, which are for once time-consuming and, on the other hand, need not always be applicable.

6.5 Conclusions

To conclude this chapter, we highlight the scope of the findings from the presented article and its supplements, outlining some future prospects for applications in XPCT and beyond.

GenSART enables Kaczmarz-reconstructions at minimal complexity: Let us clarify once again that the goal of the presented article was *not* to design novel image reconstruction methods, but to devise efficient schemes for the implementation of existing ones – namely for the large and well-established class of (regularized) Kaczmarz methods, including the Newton-Kaczmarz algorithm for XPCT from chapter 5. Notably, this has been achieved to a maximum possible degree in some sense: in most settings, the proposed generalized SART-scheme 6.6 computes Kaczmarz-iterations (of a very general form!) essentially at the costs of *one* forward- and back-projection, $f_k \mapsto \mathcal{P}_j(f_k)$ and $\Delta p_k \mapsto \mathcal{P}_j^*(\Delta p_k)$. Both amount to $\mathcal{O}(N^3)$ arithmetic operations (see §6.1) that are quasi inevitable: the forward-projection is required for comparing the current iterate to the measured tomographic data and the back-projection simply applies the computed update to the 3D-object. By minimizing the number of (back-)projector-evaluations, GenSART enables Kaczmarz-reconstructions at an overall computational complexity of $\mathcal{O}(N^4)$, analogously as classical SART (see §6.1.2). In principle, GenSART-schemes thus reduce the computational costs of Kaczmarz methods to an amount comparable to filtered back-projection (FBP). Yet, we also recall from §6.3 that translating their favorable *theoretical* complexity to low algorithmic runtimes *in practice* is more challenging than for FBP.

Efficiency-gains by exploiting the tomographic problem-structure: Importantly, the proposed GenSART-approach makes explicit use of the specific problem-structure of Kaczmarz-iterations for *tomographic* reconstruction problems. In a nutshell, GenSART exploits the simple *projection*-geometry that defines the principal relation between the unknown 3D-object and the measured 2D-tomographic images. The reduction of computational costs compared to *generic* implementations of Kaczmarz methods, as considered in §5.2.3 for the XPCT-problem, is thus enabled by adopting a *more specific* perspective on the problem: while the Newton-Kaczmarz method from §5.2.3 is applicable for arbitrary block-structured inverse problems $(G_1(f), \dots, G_{N_{\text{proj}}}(f)) \approx (g_1^{\text{obs}}, \dots, g_{N_{\text{proj}}}^{\text{obs}})$, computation of the iterates via GenSART-schemes, as implemented in algorithm 6.1, explicitly requires the operator-blocks G_j to be of the form $G_j(f) = F_j(\mathcal{P}_j(f))$ with a tomographic projector \mathcal{P}_j . This explicit restriction to tomographic problems is what buys the efficiency-gains achieved in the present work.

Constrained nonlinear all-at-once XPCT-reconstructions at tolerable runtimes: GenSART-methods allow for highly efficient parallelized implementations, compare §6.3. As

demonstrated in §6.4, the Newton-Kaczmarz-GenSART algorithm 6.1 may thereby achieve XPCT-reconstructions of large-scale data sets within a few hours on available hardware. While this is still significantly slower than direct CTF-inversion combined with filtered back-projection (CTF+FBP), such a computation-time is nevertheless arguably *tolerable* as it is comparable to the experimental *acquisition-time* of the tomographic data set. Hence, image reconstruction would at least not bottleneck the overall throughput of the XPCT imaging-workflow. At the same time, Newton-Kaczmarz methods bring more capabilities in terms of accounting for non-linearity and additional constraints (3D-supports, non-negativity and tomographic consistency, compare §5.4), which have been shown to be vital in §6.4 and in the previous chapters. In view of these benefits, the remaining disadvantage of Newton-Kaczmarz-GenSART compared to CTF+FBP in terms of runtimes may be regarded as a reasonable price to pay.

Ready for refined imaging-models: We recall that *literally nothing* has to be assumed on the data-fidelity \mathcal{S} in the considered general Kaczmarz-iterations (6.2.3) for the GenSART-principle to be applicable. In XPCT-reconstructions, the freedom in choosing \mathcal{S} allows to adapt the Newton-Kaczmarz-GenSART-algorithm 6.1 by incorporating model-refinements:

- *Image-formation:* The operator F in algorithm 6.1 (that was absorbed in \mathcal{S} for the theoretical analysis) is not limited to any specific XPCI-forward map but may be completely arbitrary. Hence, it may be adapted to account for finite coherence or other possible real-world imperfections of the imaging system, compare §2.1.5.
- *Data-errors:* As discussed in §6.2.5.1, \mathcal{S} may be additionally tuned to account for the expected statistics of the errors in the data. In particular, this allows to adapt the Newton-Kaczmarz method to *Poisson-noise* or to increase its *robustness to large outliers*. The potential benefit of such adaptations has been demonstrated in §6.2.7.1.

Importantly, the principal GenSART-scheme 6.6 remains unchanged under these modifications, only the optimization step in projection-space is affected. As a consequence, their impact on the algorithmic runtime is expected to be relatively small, even if the number of required iterations in the numerical optimization increases. This is in strong contrast to the original Newton-Kaczmarz method from chapter 5, for which more iterations per Kaczmarz-step directly translate to an increase in the number of costly evaluations of \mathcal{P}_j and \mathcal{P}_j^* , compare §5.4.

In addition to the flexibility associated with the choice of \mathcal{S} , we note that the GenSART-theory remains valid for *cone-beam* tomography, as is shown in the complete version of the article (chapter 11). If required, this permits for example a more accurate treatment of the natural divergent-beam geometry of real-world XPCT-setups, compare §2.1.5.1.

Flexible in the data-model – restrictive in the penalty-term: Contrary to the freedom in adapting the data-functional \mathcal{S} to a specific imaging modality (XPCT or practically any other transmission-tomography setting for which the projection-approximation from §2.1.3.2 is satisfied), the choice of the penalty-functional \mathcal{R} in the Kaczmarz-iterations (6.2.3) is strongly restricted by assumption 6.4, which is needed for the principal theorem 6.5 to apply. As a consequence, GenSART-schemes are only known to hold for the classes of penalties treated in §6.2.4, limiting the possibilities to impose prior knowledge on the object-iterates beyond support constraints and basic smoothness. This is probably the main drawback of the approach.

However, the considered choices of \mathcal{R} , for which the GenSART-principle was shown to apply, are probably not exhaustive. GenSART-schemes for other penalties might be proven to hold in future work (possibly in an approximate manner, similarly as in [6.2.4.2](#)). Additionally, usage of GenSART in conjunction with proximal splitting-strategies (see [§6.2.5.4](#)) may greatly extend the range of imposable object-priors, including total variation- or other sparsity-promoting penalties for example. The potential of such Kaczmarz-type splitting methods for tomographic problems has been demonstrated e.g. in [\[35\]](#). GenSART-schemes might help to further reduce their computational costs and thereby contribute to making reconstructions with advanced priors feasible also for large-scale 3D tomographic data sets – in XPCT and beyond.

Chapter 7

Discussion and Outlook

Rapid progress in experimental control of X-rays continues to erode former frontiers of coherent imaging techniques like (propagation-based) X-ray phase contrast imaging (XPCI) and -tomography (XPCT). The overall goal of this thesis was to catch up with the experimental capabilities by developing mathematical theory and algorithms for image-reconstruction in XPCI and XPCT beyond the well-established but limited realm of direct CTF-inversion. Contributions have been made in a number of different directions, as detailed in the conclusion-sections to chapters 3 to 6. Probably the most significant one on the theoretical side lies in quantifying the surprisingly strong stabilizing effect of support constraints on the inverse problems of XPCI and XPCT, ensuring well-posedness of the involved, generally ill-posed and non-unique phase retrieval step: provided a sufficiently small Fresnel number of the support, i.e. a sufficiently holographic setting, already a single hologram may be sufficient to stably retrieve both phase- and absorption-image, which was considered impossible prior to this work. The principal achievement on the algorithmic side lies in devising numerical reconstruction methods that outperform direct CTF-inversion by incorporating a priori constraints and nonlinearity but still remain computationally feasible even for “big-data”-problems in XPCT.

At the same time, the mathematical treatment in this work is far from exhaustive. Most importantly, it should be emphasized that the derived theoretical results are widely limited to the scope of the *linearized* forward models of XPCI and XPCT. Hence, the achieved numerical reconstructions in §6.4, based on the full *nonlinear* model, are so far – highly promising though they seem – not backed by any deeper theoretical understanding. A thorough analysis of fully nonlinear XPCI in future research is required to put the proposed Newton-type reconstruction methods on solid grounds and may help to further improve these. Beyond linearity, the obtained results also based on other idealizing assumptions, such as full coherence of the X-ray beam and an ideal plane-wave profile of its wavefronts, which might be relaxed in future studies. Notably, refined XPCI-models that account for such non-idealities may be flexibly incorporated in the regularized Newton(-Kaczmarz) algorithms of this work, owing to their generic nature.

On the other hand, many of the results in this thesis may be of interest in a far more general scope than the specific imaging modalities of XPCI and XPCT. From a broader perspective, the stability theorems from chapters 3 and 4 can be interpreted as general statements on the reconstructability of a local perturbation (the image in the specific case of XPCI) to a known wave-field (the wave-fronts of the illuminating X-ray beam) from phaseless measurements of the propagated total wave-field. Consequently, the findings may be relevant to a wide range of

related settings, for example including wavefront-sensing and the large class of inverse scattering problems where an object is to be recovered from the induced perturbation to a monochromatic wave. On the contrary, it should be emphasized once again that the principal results of this work do *not* simply carry over to coherent diffractive imaging (CDI) [146, 147], even though this technique can be interpreted as the far-field limit of XPCI and thus seems highly related. The reason is that the developed mathematical theory and algorithms heavily rely on *holographic* contrast, arising from interference of the known incident wave with the secondary wave induced by scattering on the unknown object, whereas such contrast is not natively present in CDI-data. Recently however, variants of CDI, where holographic contrast is established by introducing additional reference-signals, have been discussed in the literature [169, 128, 19, 182]. The phase retrieval problems arising in such settings are structurally similar to that of XPCI so that the methodology of the present work might be adaptable. Furthermore, the reconstruction algorithms proposed in chapters 5 and 6, in particular the Newton-Kaczmarz method for XPCT and its efficient implementation via generalized SART schemes, may be adapted to many other inverse problems in transmission-tomography.

An immediate and fruitful application of this thesis' findings beyond X-ray imaging might be found in the emerging technique of near-field holography with *electrons* [67, 132]: owing to progress in sample-preparation and -fixation, it is now possible to record holograms of single proteins illuminated by a highly coherent, divergent beam of low-energy electrons. The acquisition geometry is thus identical as in XPCI – only a different form of radiation is used. Moreover, also the principal image-formation model essentially carries over from X-rays to electrons (see e.g. [159] for details), so that the theory and algorithms of the present work are applicable. Judging from the electron-holograms in [132], it indeed seems that the insights gained in this thesis may be highly relevant for the novel imaging technique for the following reasons:

1. Both phase- and absorption contrast may be induced by the specimens and no trivial coupling between the two contrast-forms is expected to hold.
2. Due to the targeted subnanometer-resolutions, acquiring holograms at more than one source-to-sample-distance seems undesirable in view of mechanical stability requirements.
3. The imaged protein-molecules are small compared to the captured field-of-view so that strong support constraints hold true.
4. The Fresnel number associated with the molecule-diameter is low, i.e. the hologram-data is *deeply-* or even *completely holographic* (see §3.6), so that image-reconstruction (with support constraints!) is expected to be highly stable – despite the points 1. and 2.

In the original study [132], however, the images are recovered by simple Fresnel-back-propagation of the hologram, a simple, merely *qualitative* image-reconstruction method that already dates back to the pioneering work on holography by DENNIS GABOR [72]. Significant improvements in image-quality might be achievable with the algorithms proposed in the present work.

The example shows that the field of coherent (near-field) imaging methods continues to evolve dynamically, where novel experimental possibilities necessitate mathematics beyond established standards to be fully exploited. The theoretical insights and practical algorithms of this thesis may contribute to meeting current and future challenges in XPCI and related imaging modalities, thereby helping to push these fascinating techniques to their full potential.

Part III

Complete Articles

Chapter 8

Complete Article 1: Stability estimates for linearized near-field phase retrieval in X-ray phase contrast imaging

Reproduced from the published* article [144]:

Maretzke S, Hohage T. *SIAM Journal on Applied Mathematics*, 77(2):384-408, 2017.

DOI: [10.1137/16M1086170](https://doi.org/10.1137/16M1086170)

(subject to slight adjustments of the notation and minor corrections)

Abstract: Propagation-based X-ray phase contrast enables nanoscale imaging of biological tissue by probing not only the attenuation, but also the real part of the refractive index of the sample. Since only intensities of diffracted waves can be measured, the main mathematical challenge consists in a phase-retrieval problem in the near-field regime. We treat an often used linearized version of this problem known as contract transfer function model. Surprisingly, this inverse problem turns out to be well-posed assuming only a compact support of the imaged object. Moreover, we establish bounds on the Lipschitz stability constant. In general this constant grows exponentially with the Fresnel number of the imaging setup. However, both for homogeneous objects, characterized by a fixed ratio of the induced refractive phase shifts and attenuation, and in the case of measurements at two distances, a much more favorable algebraic dependence on the Fresnel number can be shown. In some cases we establish order optimality of our estimates.

8.1 Introduction

Over the past two decades, the dramatic increase in coherence and brightness of large-scale X-ray sources, such as third generation synchrotrons and free-electron-lasers, has paved the way for X-ray phase contrast imaging [158]. Classical X-ray radiography is limited to measuring the attenuation experienced by radiation traversing the probed object. Writing the refractive index in the X-ray physics notation $n = 1 - \delta + i\beta$ with $0 \leq \beta, \delta \ll 1$, this amounts to imaging β .

*Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Phase contrast techniques additionally probe the real-valued decrement δ of n , which induces phase shifts in the transmitted X-ray wave field. This enables imaging of biological cells and other micro-scale light-element specimen, for which $\beta \ll \delta$ holds in the hard X-ray regime [207, 162, 42, 145]. Owing to the small wavelength of X-rays, nano-scale spatial resolutions can be achieved as has been demonstrated down to 20 nanometers [13]. Moreover, phase contrast imaging can be combined with tomography, capable of resolving the refractive index of an unknown object in 3D [108, 14, 119, 122, 176, 143].

Unfortunately, the refractive phase shifts of the X-ray field cannot be observed directly by common CCD detectors due to their physical limitation to measuring wave intensities, i.e. the squared modulus of the wave field. In propagation-based phase contrast imaging, also known as *inline holography*, the required phase-sensitivity is achieved simply by free-space propagation without any optical elements: if the detector is placed in some finite distance down-stream of the sample, the imprinted phase shifts in the object’s exit plane ($z = 0$ in fig. 8.1(b)) are partially encoded into measurable intensities by *diffraction*, i.e. self-interference of the wave field. We assume that the diffraction pattern or *hologram* is recorded in the optical near-field of the sample so that propagation is described by the Fresnel propagator [160]. In particular, we do not consider the corresponding far-field setup (coherent diffractive imaging, see e.g. [167, 147]), where the data is given by Fourier magnitudes.

In this work, we are thus concerned with the reconstruction of the (generally complex-valued) wave-field perturbation h induced by the object from measured near-field intensities I . As this implicitly amounts to recovering the lost phase information in the data, i.e. to solving a *phase retrieval* problem, the question immediately arises whether the image recovery is actually *unique* and *stable*. Indeed, it is commonly argued [108, 157, 32] that diffraction patterns from at least two different sample-detector distances are required for a unique reconstruction of the imprinted phase shifts and attenuation. Assuming a *support constraint*, however, i.e. under the often physically reasonable assumption that the image h is non-zero only in compact subdomain of the field of view, we could show uniqueness of the reconstruction from a *single* hologram in [140] – even if h is complex-valued (see also [117] for a uniqueness result for the Helmholtz equation with real-valued n from phaseless near-field data in an interval of frequencies). In [143], such a reconstruction of a compactly supported complex image is demonstrated for simulated and experimental data, which turns out to be feasible, yet susceptible to low-frequency artifacts. Support constraints have also been found to stabilize image reconstruction in simpler settings where the probed object can be assumed to be completely non-absorbing [76, 14].

These observations call for a better understanding of the *stability* of the considered near-field phase retrieval problem, which is the goal of this paper. We do so within a linearization of the relation between image h and the resulting intensity data I , valid for sufficiently “small” h , i.e. for weakly interacting objects similarly as in a recent stability analysis of domain reconstructions in phaseless inverse scattering [5]. The linearization is known as *contrast transfer function* model [79, 194] and frequently applied in X-ray phase contrast imaging [42, 81, 96, 125, 122]. In this work, we analyze the arising linear forward operator \mathcal{T} under the assumption that h has compact support. We prove that the associated inverse problem is not only unique but even *well-posed* in the sense that the recovered image h depends continuously on the measured data $\mathcal{T}(h)$, i.e. finite data errors lead to bounded deviations in the recovered h . This result, which is quite surprising for an inverse problem with remote measurements, is achieved by relating the setting to a reconstruction from incomplete Fourier data. By the same technique we also

derive explicit *stability estimates*, bounding the reconstruction error that results from a given data noise level in terms of the dimensionless Fresnel number.

In general, we find that the stability constant decays exponentially with the Fresnel number f and thus hardly gives any useful stability bounds for many experimental X-ray phase contrast setups. However, we establish much more favorable $\mathcal{O}(f^{-1})$ - and even $\mathcal{O}(f^{-1/2})$ -decay rates in two relevant situations: The first assumes proportionality of the real and the imaginary part of the image h , which occurs e.g. for single-material and non-absorbing samples. The second situation concerns general objects, but two measurements at different distances. It is well-known that the forward operators in these cases are Fourier multipliers with so-called contrast transfer functions (CTF), and the zeros of the CTFs are responsible for general ill-posedness of the inverse problem. Our analysis exploits the regularizing effect on these zeros of the smoothness in Fourier space that results from the assumed support constraints.

The remainder of this paper is organized as follows: in §8.2, the mathematical model of X-ray phase contrast imaging is introduced and the considered inverse reconstruction problems, corresponding to different constraints and measurement setups, are motivated. Our principal stability results are stated in §8.2.4. §8.3, §8.4 and §8.5 contain the analysis for each of the inverse problems, including the proofs of the main results. In §8.6, we discuss implications of our findings and possible extensions.

8.2 Imaging problems and main results

8.2.1 Physical model

An exemplary experimental setup for X-ray phase contrast imaging on a third generation synchrotron source (GINIX setup [178] at P10-beamline, DESY) is shown in fig. 8.1(a). We describe this imaging system by a standard wave-optical model as schematically visualized in fig. 8.1(b) [166, 39]: an unknown sample is illuminated by an incident plane electromagnetic wave $\Psi_i(\mathbf{x}, z) = \exp(ikz)$, where $\mathbf{x} \in \mathbb{R}^2$ and $z \in \mathbb{R}$ denote the lateral- and axial coordinates, respectively. By scattering interaction, object information is encoded as a perturbation of the wave field $\Psi = \Psi_i + \Psi_s$ within the exit plane $z = 0$ of the probed sample. A detector measures the resulting near-field diffraction pattern (or hologram), given by the intensity $I(\mathbf{x}) = |\Psi(\mathbf{x}, d)|^2$ of the propagated wave fronts in some plane at finite distance $d > 0$ behind the object. The phase of the complex-valued field Ψ cannot be observed directly, yet *diffraction* partially encodes phase variations in the exit plane $z = 0$ into measurable intensities I at the detector.

In general, the above physical model is governed by the Helmholtz equation $\Delta\Psi + n^2k^2\Psi = 0$ in \mathbb{R}^3 , where the object gives rise to a spatially varying refractive index $n = 1 - \delta + i\beta$. In the hard X-ray regime of very large wavenumbers k , typical samples such as biological cells are often sufficiently thin and weakly interacting for the scattering to be well-approximated by geometrical optics [160, 108]. Within this approximation, the wave fronts $\Psi_z := \Psi(\cdot, z)$ in the exit plane are given by

$$\Psi_0 = \exp(-h) \quad \text{with} \quad h = i\phi + \mu = ik \int_{\mathbb{R}} (\delta - i\beta) dz \quad (8.2.1)$$

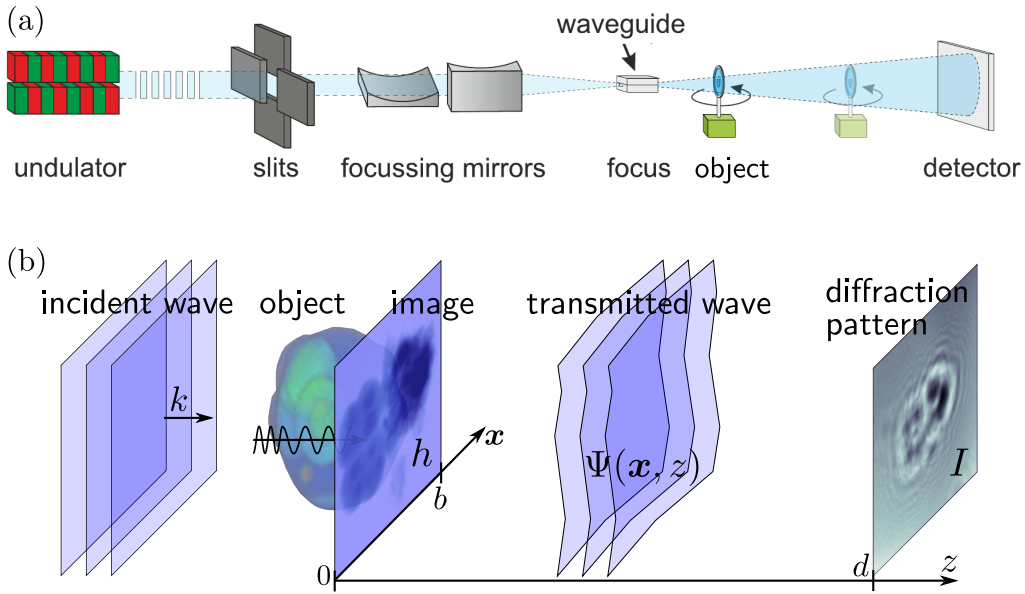


Figure 8.1. Setup of propagation-based X-ray phase contrast imaging (inline holography).

(a) Sketch of an experimental realization [143] (GINIX [178] at P10-beamline, DESY)

(b) Physical model [140]: incident plane waves are scattered by an unknown sample, imprinting a phase- and absorption image $h = -i\phi - \mu$ upon the transmitted wave fronts $\Psi(\cdot, 0)$. The resulting near-field diffraction pattern (hologram) $I = |\Psi(\cdot, d)|^2$ is recorded at some distance $z = d$ behind the object. (plotted experimental data: hologram I and reconstructed image h of *d. radiodurans* bacteria).

Accordingly, the perturbed wave yields line integrals over δ and β along the incident z -direction, corresponding to a *projection image* of the sample in the form of phase shifts ϕ and attenuation μ . In particular, (8.2.1) implies that the imprinted image h satisfies a *support constraint* $\text{supp}(h) \subset \Omega$ whenever the object is laterally finite, i.e. if $\delta(\mathbf{x}, z) = \beta(\mathbf{x}, z) = 0$ for all $z \in \mathbb{R}$, $\mathbf{x} \in \Omega$ outside some bounded domain $\Omega \subset \mathbb{R}^2$.

If $|\boldsymbol{\xi}| \ll k$ for all relevant spatial frequencies $\boldsymbol{\xi}$ of h , the total wave field will be of the form $\Psi(\mathbf{x}, z) = e^{ikz}\tilde{\Psi}(\mathbf{x}, z)$ with a slowly varying envelope $\tilde{\Psi}$ such that $e^{-ikz}(\partial_z^2 + k^2)\Psi \approx 2ik\partial_z\tilde{\Psi}$ by neglecting $\partial_z^2\tilde{\Psi}$. This yields the one-way-, Schrödinger- or paraxial approximation

$$2ki\partial_z\tilde{\Psi} + \Delta_{\mathbf{x}}\tilde{\Psi} \approx 0$$

to the Helmholtz equation. Within this commonly used model, the free-space propagation of the wave fronts to the detector is described by the *Fresnel propagator* [160]:

$$\mathcal{D}(\Psi_0) := \exp(-ikd)\Psi_d = \mathcal{F}^{-1}(m_{\mathfrak{f}} \cdot \mathcal{F}(\Psi_0)) \quad m_{\mathfrak{f}}(\boldsymbol{\xi}) := \exp\left(\frac{-i|\boldsymbol{\xi}|^2}{2\mathfrak{f}}\right) \quad (8.2.2)$$

Here, \mathcal{F} is the Fourier transform and

$$\mathfrak{f} := \frac{kb^2}{d} \quad \text{or} \quad \mathfrak{f} := \frac{\mathfrak{f}}{2\pi} \quad (8.2.3)$$

denote the dimensionless *Fresnel number* of the setup. b is a physical length that corresponds to length 1 in dimensionless coordinates and will be chosen as the support diameter of the

image h , see fig. 8.1(b). Usually \mathfrak{f} is referred to as Fresnel number, but this convention would lead to an abundance of 2π factors in our computations. Therefore, we will mostly use \mathfrak{f} with the notation (8.2.3) chosen in analogy to Planck's constant. \mathfrak{f} governs the impact of diffraction in the imaging setup, where smaller values correspond to stronger diffractive distortion of the propagating wave field. Typical values in experimental X-ray phase contrast setups are in the range $10 \leq \mathfrak{f} \leq 10^4$.

By combination of (8.2.1) and (8.2.2), we find that the unknown object image h is related to the observable intensity data $I = |\Psi_d|^2$ by the nonlinear forward operator

$$I = F(h) := |\mathcal{D}(\exp(-h))|^2 \quad \text{with} \quad h = i\phi + \mu \quad (8.2.4)$$

We note that similar models apply to imaging with *electrons* [62, 205, 129] owing to the mathematical equivalence of the time-independent Schrödinger equation and the Helmholtz equation.

8.2.2 Weak object limit and principal inverse problem

By (8.2.4), the image h is in general complex-valued, whereas the intensity data I is real-valued. This suggests that the data is insufficient for unique and stable recovery of h [157, 32]. We analyze this question of ill-posedness within the commonly used *weak-object-approximation* [166, 42, 161, 81]: in the case of weak absorption $\mu \ll 1$ and slowly varying phase shifts ϕ , nonlinear terms in h can be neglected in (8.2.4) [194], giving

$$F(h) = 1 + \mathcal{T}(h) + \mathcal{O}(h^2) \quad \text{with} \quad \mathcal{T}(h) := -2\text{Re}(\mathcal{D}(h)). \quad (8.2.5)$$

Here, Re denotes the pointwise real-part and we have used that $\mathcal{D}(1) = 1$. Note that \mathcal{D} is unitary on $L^2(\mathbb{R}^m)$, so \mathcal{T} defines a bounded \mathbb{R} -linear operator on $L^2(\mathbb{R}^m)$. Rather than by (8.2.5), \mathcal{T} is more commonly written in terms of the phase shifts ϕ and absorption μ via sinusoidal *contrast transfer functions* (CTF) [79]:

$$\mathcal{T}(i\phi + \mu) = -2\mathcal{F}^{-1} \left(\sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathfrak{f}} \right) \mathcal{F}(\phi) + \cos \left(\frac{|\boldsymbol{\xi}|^2}{2\mathfrak{f}} \right) \mathcal{F}(\mu) \right). \quad (8.2.6)$$

Although the physical model of §8.2.1 leads to two-dimensional images h and holograms I , we will study the operator \mathcal{T} in a more general \mathbb{R}^m -setting. This might allow application of our results to situations described by (quasi-) 1D-models and to *phase contrast tomography*, which can be interpreted as a $\mathcal{3D}$ -imaging modality [119, 176]. As physically motivated in §8.2.1, we impose support constraints by assuming

$$h \in L^2(\Omega) := \{h \in L^2(\mathbb{R}^m) : h|_{\mathbb{R}^m \setminus \Omega} = 0\} \quad \text{for some} \quad \Omega \subset \mathbb{R}^m. \quad (8.2.7)$$

Moreover, we denote by $\|h\| := (\int_{\mathbb{R}^m} |h|^2 \, d\mathbf{x})^{1/2}$ the standard L^2 -norm in \mathbb{R}^m . The principal image reconstruction problem of this work then reads as follows:

Inverse Problem 8.1 (Phase contrast imaging of weak objects). For a given support $\Omega \subset \mathbb{R}^m$, recover a complex-valued image $h \in L^2(\Omega)$ from noisy intensity data

$$I^\epsilon = 1 + \mathcal{T}(h) + \epsilon \quad \text{with} \quad \|\epsilon\| \leq \epsilon.$$

8.2.3 Homogeneous and non-absorbing objects

It is often legitimate to assume that the object is *homogeneous* in the sense that phase shifts ϕ and attenuation μ are proportional, i.e.

$$h = \mu + i\phi = ie^{-i\nu}\varphi \quad (8.2.8)$$

for some $\nu \in [0; \pi)$ and a *real-valued* function $\varphi \in L^2(\mathbb{R}^m)$. Note that this includes the special case $\nu = 0$ which corresponds to $\mu = 0$ and thus to a purely phase shifting, i.e. non-absorbing object, providing an excellent model for hard X-ray imaging of light-element samples. By plugging (8.2.8) into (8.2.6) and rearranging by trigonometric identities, we obtain a forward operator incorporating the *homogeneity constraint*:

$$\begin{aligned} \mathcal{S}_\nu : L^2(\mathbb{R}^m) &\rightarrow L^2(\mathbb{R}^m); \varphi \mapsto 2\mathcal{F}^{-1}(s_\nu \cdot \mathcal{F}(\varphi)), & s_\nu(\boldsymbol{\xi}) &:= \sin\left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}} + \nu\right) \\ \mathcal{T}(ie^{-i\nu}\varphi) &= \mathcal{S}_\nu(\varphi) & \text{for all real-valued } \varphi &\in L^2(\mathbb{R}^m). \end{aligned} \quad (8.2.9)$$

Accordingly, the forward model reduces to a multiplication with the contrast transfer function s_ν (CTF) in Fourier space [79, 194]. This makes the inversion of \mathcal{S}_ν significantly easier than that of \mathcal{T} , which is why we state it as a second inverse problem:

Inverse Problem 8.2 (Phase contrast imaging of weak homogeneous objects). For given $\Omega \subset \mathbb{R}^m$, recover a real-valued image $\varphi \in L^2(\Omega)$ from noisy intensity data

$$I^\epsilon = 1 + \mathcal{S}_\nu(\varphi) + \epsilon \quad \text{with} \quad \|\epsilon\| \leq \epsilon.$$

8.2.4 Stability estimates

The statement of inverse problems 8.1 and 8.2 immediately raises the question whether these are uniquely solvable and whether the solution is *stable* with respect to noise ϵ . In order to illustrate the significance of this problem, we first recall some well-known facts on the derived inverse problems without assuming a support constraint, i.e. for $\Omega = \mathbb{R}^m$: as the null-space $\text{kern}(\mathcal{T}) = \{i\mathcal{D}^{-1}(f) : f \in L^2(\mathbb{R}^m) \text{ real-valued}\}$ of the forward map $\mathcal{T} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is huge, inverse problem 8.1 is heavily non-unique in this setting. The operator \mathcal{S}_ν , on the other hand, is indeed injective so that inverse problem 8.2 is uniquely solvable. However, the inversion of \mathcal{S}_ν is *ill-posed* since noise in Fourier-frequencies near the zeros of the CTF s_ν is amplified by arbitrary factors in the reconstruction.

To our great surprise, imposing a support constraint with bounded Ω does not only rule out non-uniqueness (as proven in [140]), but even turns inverse problem 8.1 into a *well-posed* problem: every admissible image $h \in L^2(\Omega)$ gives rise to finite *contrast* $\|\mathcal{T}(h)\| \geq C_{\text{IP1}}\|h\|$ in the observable data with some lower bound $C_{\text{IP1}} > 0$:

Theorem 8.1 (Well-posedness and stability estimate for inverse problem 8.1). *Let the support-domain Ω be given by a stripe of width 1, without loss of generality $\Omega := [-1/2; 1/2] \times \mathbb{R}^{m-1}$. Then there exists a constant $C_{\text{IP1}}(\Omega, \mathbf{f}) > 0$ such that*

$$\|\mathcal{T}(h)\| \geq C_{\text{IP1}}(\Omega, \mathbf{f})\|h\| \quad \text{for all } h \in L^2(\Omega), \quad (8.2.10)$$

i.e. inverse problem 8.1 is well-posed. The stability constant satisfies the estimate

$$C_{\text{IP1}}(\Omega, \mathfrak{f}) \geq (2\pi\mathfrak{f})^{\frac{1}{4}} \left(1 - \frac{3}{8\mathfrak{f}} + \mathcal{O}(\mathfrak{f}^{-2}) \right) \exp(-\mathfrak{f}/8). \quad (8.2.11)$$

theorem 8.1 is proven in §8.3 along with a characterization of the least stable modes, i.e. of the images h that induce least contrast under \mathcal{I} . Notably, (8.2.10) implies $\|h\| \leq C_{\text{IP1}}(\Omega, \mathfrak{f})^{-1} \|\mathcal{I}(h)\|$, ensuring finite amplification of data errors $\leq C_{\text{IP1}}(\Omega, \mathfrak{f})^{-1} \|\epsilon\|$ upon inversion of \mathcal{I} and thus *stability* of the reconstruction of h from I^ϵ . By (8.2.11), however, the constant $C_{\text{IP1}}(\Omega, \mathfrak{f})$ decays (nearly) exponentially with increasing \mathfrak{f} so that theorem 8.1 hardly guarantees stability in any practical sense for Fresnel numbers $\mathfrak{f} \geq 100$. Fortunately, the stability estimate can be improved to algebraic decay with \mathfrak{f} in the case of inverse problem 8.2, as shown in §8.4:

Theorem 8.2 (Well-posedness and stability estimate for inverse problem 8.2). *Let the support-domain $\Omega := \{\mathbf{x} \in \mathbb{R}^m : |\mathbf{x}| \leq \frac{1}{2}\}$ be a ball of diameter 1. Then the stability constant $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) := \inf_{\varphi \in L^2(\Omega), \|\varphi\|=1} \|\mathcal{S}_\nu(\varphi)\|$ of inverse problem 8.2 is bounded by*

$$C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) \geq \max \left\{ \min \{c_1, c_2\mathfrak{f}^{-1}\}, \min \left\{ c_3\nu, c_4\mathfrak{f}^{-\frac{1}{2}} \right\} \right\} \quad (8.2.12)$$

for some constants $c_j > 0$ that depend only on the dimension m . In particular, $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) = \mathcal{O}(\mathfrak{f}^{-1})$ for $\nu = 0$ and $C_{\text{IP2}}(\Omega, \mathfrak{f}, \nu) = \mathcal{O}(\mathfrak{f}^{-1/2})$ for $\nu > 0$ as $\mathfrak{f} \rightarrow \infty$.

We recall that the physical lengthscale b underlying to the Fresnel number \mathfrak{f} in theorems 8.1 and 8.2 is the diameter of the support-domain Ω as the latter is taken to be unit length. Accordingly, the resulting stability constants $C_{\text{IP1}}, C_{\text{IP2}}$ are much smaller than 1 for typical values $10 \leq \mathfrak{f} \leq 10^4$. Moreover, we emphasize that the stated results for inverse problems 8.1 and 8.2 are both for reconstructions from a *single* diffraction pattern. Image recovery from *two* holograms recorded at different distances is treated in §8.5 as a corollary of the stability analysis of inverse problem 8.2.

8.3 Stability analysis of inverse problem 8.1

8.3.1 Principal approach

We start our analysis with inverse problem 8.1, corresponding to the recovery of general complex-valued images from a single hologram without homogeneity constraint. In order to understand its mathematical structure, it is instructive to rewrite the forward operator \mathcal{I} in the form

$$\begin{aligned} -\mathcal{I}(h) &= 2\text{Re}(\mathcal{D}(h)) = \mathcal{D}(h) + \overline{\mathcal{D}(h)} = \mathcal{D}(h) + \mathcal{F}^{-1}(m_{\mathfrak{f}}^{-1} \cdot \mathcal{F}(\overline{h})) \\ &= \mathcal{D}(h) + \mathcal{D}^{-1}(\overline{h}) \quad \text{for } h \in L^2(\mathbb{R}^m). \end{aligned} \quad (8.3.1)$$

Here, the overbar denotes complex conjugation and we have used that the Fresnel propagation factor $m_{\mathfrak{f}}$ is unitary. By (8.3.1), the linearized contrast $\mathcal{I}(h)$ in the intensity data is given by a superposition of the propagated image $\mathcal{D}(h)$ and the *back-propagated twin-image* $\mathcal{D}^{-1}(\overline{h})$. Since

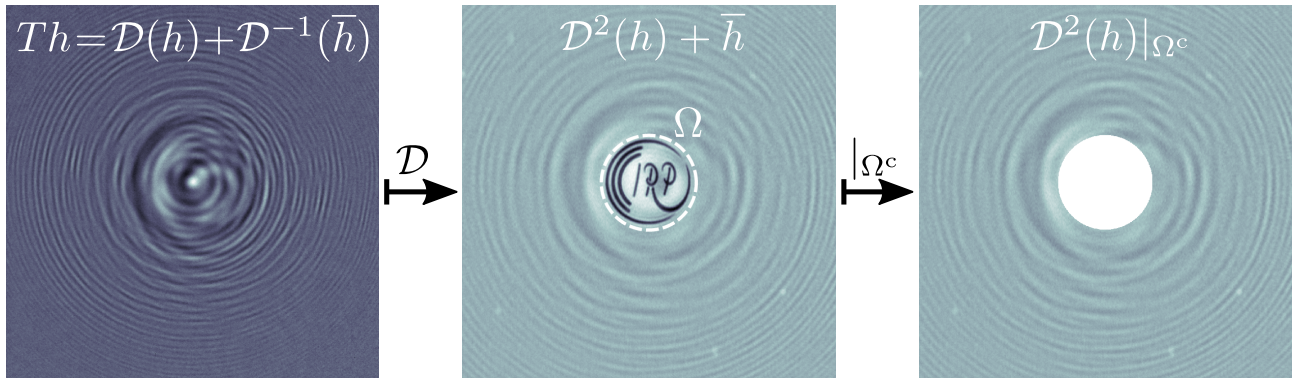


Figure 8.2. Illustration of the principal idea for the stability analysis of inverse problem 8.1. By applying the Fresnel propagator \mathcal{D} to data $\mathcal{I}(h)$, the *twin-image* \bar{h} becomes sharp (*Gabor holography*, see logo in central panel). By restricting to the complement Ω^c of the support-domain $\Omega \supset \text{supp}(h)$, \bar{h} is eliminated and incomplete Fresnel data $\mathcal{D}^2(h)|_{\Omega^c}$ is obtained (right panel). Images show real parts of numerically computed fields from a hologram (left panel) recorded at GINIX [178], P10-beamline, DESY.

\mathcal{D} is unitary, reconstructing h from one of these components *alone* would be straightforward. Solving inverse problem 8.1 thus amounts to disentangling image and twin-image.

In order to separate these components, we propagate the near-field hologram $\mathcal{I}(h)$ by application of \mathcal{D} . By (8.3.1), this yields

$$-\mathcal{D}\mathcal{I}(h) = \mathcal{D}^2(h) + \bar{h}. \quad (8.3.2)$$

Accordingly, we recover the sharp twin-image \bar{h} up to perturbations originating from the doubly propagated image $\mathcal{D}^2(h)$. This is the principle of *Gabor holography* [72], which can be used as a *qualitative* image reconstruction technique [160] as illustrated in fig. 8.2. Here, we follow this approach in a converse manner: rather than contenting ourselves with the perturbed twin-image \bar{h} , we exploit a *support constraint*

$$\text{supp}(h) = \overline{\{\mathbf{x} \in \mathbb{R}^m : h(\mathbf{x}) \neq 0\}} \subset \Omega \quad (8.3.3)$$

by restricting (8.3.2) to the complement of Ω . This yields

$$-\mathcal{D}\mathcal{I}(h)|_{\Omega^c} = \mathcal{D}^2(h)|_{\Omega^c} + \bar{h}|_{\Omega^c} = \mathcal{D}^2(h)|_{\Omega^c} \quad \text{for any } h \in L^2(\Omega). \quad (8.3.4)$$

By the proposed propagation-and-restriction procedure, the twin-image is thus completely eliminated from the data as sketched in fig. 8.2. Note that the map $\mathcal{I}(h) \mapsto -\mathcal{D}(\mathcal{I}(h))|_{\Omega^c} = \mathcal{D}^2(h)|_{\Omega^c}$ is norm-decreasing in $L^2(\mathbb{R}^m)$. Hence, (8.3.4) implies that the solution of inverse problem 8.1 is at most as ill-posed as the reconstruction from incomplete Fresnel data $\mathcal{D}^2(h)|_{\Omega^c}$. This reduction to a data completion problem is the principal idea of our stability analysis for inverse problem 8.1.

8.3.2 Reduction to Fourier completion problem

In order to gain a simpler expression for the reduced data $\mathcal{D}^2(h)|_{\Omega^c}$ obtained in §8.3.1, we use an alternate form of the Fresnel propagator \mathcal{D} . By application of the convolution theorem to

(8.2.2), the following representation can be obtained (see e.g. [160]):

$$\mathcal{D}(h)(\mathbf{x}) = e^{-im\pi/4} \mathfrak{f}^{\frac{m}{2}} n_{\mathfrak{f}}(\mathbf{x}) \cdot \mathcal{F}(n_{\mathfrak{f}} \cdot h)(\mathfrak{f}\mathbf{x}), \quad n_{\mathfrak{f}}(\mathbf{x}) := \exp\left(\frac{i}{2}\mathfrak{f}|\mathbf{x}|^2\right) \quad (8.3.5)$$

for all $\mathbf{x} \in \mathbb{R}^m$. (8.3.5) reveals that - up to pointwise multiplications with the unitary factor $n_{\mathfrak{f}}$ and rescaling - \mathcal{D} may be written as a Fourier transform. Combined with the approach outlined in §8.3.1, this allows to identify stability of inverse problem 8.1 with the reconstruction of a function from *incomplete Fourier data*:

Theorem 8.3 (Reduction to Fourier data completion problem). *Within the setting of inverse problem 8.1, let $\Omega_{\mathfrak{f}} := \{(\mathfrak{f}/2)\mathbf{x} : \mathbf{x} \in \Omega\}$ for arbitrary $\Omega \subset \mathbb{R}^m$. Then*

$$\|\mathcal{T}(h)\| \geq \|\mathcal{F}(n_{\mathfrak{f}}^{1/2} \cdot h)|_{\Omega_{\mathfrak{f}}^c}\| \quad \text{for all } h \in L^2(\Omega) \quad (8.3.6)$$

where $n_{\mathfrak{f}}^{1/2}(\mathbf{x}) = \exp(i\mathfrak{f}|\mathbf{x}|^2/4)$. In particular, we have the relative stability estimate

$$C_{\text{IP1}}(\Omega, \mathfrak{f}) = \inf_{h \in L^2(\Omega), \|h\|=1} \|\mathcal{T}(h)\| \geq \inf_{h \in L^2(\Omega), \|h\|=1} \|\mathcal{F}(h)|_{\Omega_{\mathfrak{f}}^c}\| \quad (8.3.7)$$

Proof. \mathcal{D} is unitary and the restriction to $\Omega^c \subset \mathbb{R}^m$ defines an orthogonal projection in $L^2(\mathbb{R}^m)$. Hence, (8.3.4) implies the estimate

$$\|\mathcal{T}(h)\| = \|\mathcal{D}\mathcal{T}(h)\| \geq \|\mathcal{D}\mathcal{T}(h)|_{\Omega^c}\| = \|\mathcal{D}^2(h)|_{\Omega^c}\| \quad \text{for any } h \in L^2(\Omega). \quad (8.3.8)$$

By (8.2.2), \mathcal{D}^2 is again a Fresnel propagator, yet to the Fresnel number $\mathfrak{f}/2$. Accordingly, employing the alternative form (8.3.5) and exploiting that $|e^{-im\pi/4} n_{\mathfrak{f}}^{1/2}| \equiv 1$ gives

$$\|\mathcal{D}^2(h)|_{\Omega^c}\| = \left(\frac{1}{2}\mathfrak{f}\right)^{\frac{m}{2}} \|\mathcal{F}(n_{\mathfrak{f}}^{1/2} \cdot h)\left(\frac{1}{2}\mathfrak{f}\cdot\right)|_{\Omega^c}\|. \quad (8.3.9)$$

Introducing new coordinates $\boldsymbol{\xi} := (\mathfrak{f}/2)\mathbf{x}$ and using that $\Omega_{\mathfrak{f}}^c = (\mathfrak{f}/2) \cdot \Omega^c$ holds by definition, this expression can be simplified to

$$\begin{aligned} \|\mathcal{D}^2(h)|_{\Omega^c}\|^2 &= \left(\frac{1}{2}\mathfrak{f}\right)^m \int_{\Omega^c} |\mathcal{F}(n_{\mathfrak{f}}^{1/2} \cdot h)\left(\frac{1}{2}\mathfrak{f}\mathbf{x}\right)|^2 d\mathbf{x} \\ &= \int_{\Omega_{\mathfrak{f}}^c} |\mathcal{F}(n_{\mathfrak{f}}^{1/2} \cdot h)(\boldsymbol{\xi})|^2 d\boldsymbol{\xi} = \|\mathcal{F}(n_{\mathfrak{f}}^{1/2} \cdot h)|_{\Omega_{\mathfrak{f}}^c}\|^2. \end{aligned} \quad (8.3.10)$$

Combining (8.3.8) and (8.3.10) yields the first assertion (8.3.6). The second estimate (8.3.7) then follows from the fact that the map $h \mapsto n_{\mathfrak{f}}^{1/2} \cdot h$ is isometric and bijective on $L^2(\Omega)$. \square

theorem 8.3 states that the solution of inverse problem 8.1 is *at least as stable* as the reconstruction of an L^2 -function f with support in $\Omega \subset \mathbb{R}^m$ from incomplete Fourier measurements $\mathcal{F}(f)|_{\Omega_{\mathfrak{f}}^c}$. For compact Ω , the latter Fourier completion problem can be shown to be unique and even *well-posed* by employing an uncertainty principle for the 1D-Fourier transform derived by Nazarov [156] (see [90] for an English proof and [106] for its multidimensional generalization). Rather than following this approach for general support shapes, however, we will restrict to the special case of rectangular support-domains Ω . This will enable a more explicit characterization of the dependence of C_{IP1} on the Fresnel number \mathfrak{f} via theorem 8.3 as well as additional insights concerning the nature of the least stable modes of the forward operator \mathcal{T} .

8.3.3 Stability result for stripe-shaped supports

In the following, we restrict to the simple case of a stripe-shaped support-domain $\Omega \subset \mathbb{R}^m$ as considered in theorem 8.1. Note that the forward operator \mathcal{S} is rotationally- and translationally invariant as is the Fresnel propagator \mathcal{D} . Hence, it is sufficient to consider domains of the form $\Omega := [-b/2; b/2] \times \mathbb{R}^{m-1}$. Moreover, b can be set to 1 which means that we define the Fresnel number \mathfrak{f} with respect to the support diameter. Accordingly, we may indeed restrict to the special case $\Omega := [-1/2; 1/2] \times \mathbb{R}^{m-1}$ as done in theorem 8.1 without loss of generality.

Now we can employ the characterization of the stability constant in theorem 8.3 for this special domain. We define $I := [-1/2; 1/2]$ and $I_{\mathfrak{f}} := [-\mathfrak{f}/4; \mathfrak{f}/4]$ for notational convenience. Recall that the m -dimensional Fourier transform $\mathcal{F}^{(m)}$ is a tensor product of one-dimensional Fourier transforms $\mathcal{F}^{(1)}$ applied along the different coordinate dimensions, i.e. $\mathcal{F}^{(m)} = \mathcal{F}^{(1)} \otimes \dots \otimes \mathcal{F}^{(1)}$. Owing to the cartesian product structure of $\Omega = I \times \mathbb{R}^{m-1}$ and $\Omega_{\mathfrak{f}}^c = I_{\mathfrak{f}}^c \times \mathbb{R}^{m-1}$, the restricted Fourier transforms

$$\mathcal{F}_{\Omega, \Omega_{\mathfrak{f}}^c}^{(m)} : L_{\Omega}^2 \rightarrow L_{\Omega_{\mathfrak{f}}^c}^2; h \mapsto (\mathcal{F}^{(m)}h)|_{\Omega_{\mathfrak{f}}^c}, \quad \mathcal{F}_{I, I_{\mathfrak{f}}^c}^{(1)} : L_I^2 \rightarrow L_{I_{\mathfrak{f}}^c}^2; h \mapsto (\mathcal{F}^{(1)}h)|_{I_{\mathfrak{f}}^c}, \quad (8.3.11)$$

are likewise related by $\mathcal{F}_{\Omega, \Omega_{\mathfrak{f}}^c}^{(m)} = \mathcal{F}_{I, I_{\mathfrak{f}}^c}^{(1)} \otimes \mathcal{F}^{(1)} \otimes \dots \otimes \mathcal{F}^{(1)}$. Applying this relation to the stability estimate in (8.3.7) and exploiting unitarity of the Fourier transform yields

$$\begin{aligned} C_{\text{IP1}}(\Omega, \mathfrak{f}) &= \inf_{h \in L^2(\Omega), \|h\|=1} \|\mathcal{F}_{\Omega, \Omega_{\mathfrak{f}}^c}^{(m)} h\| = \inf_{h \in L_I^2, \|h\|=1} \|\mathcal{F}_{I, I_{\mathfrak{f}}^c}^{(1)} h\| \\ &= \left(1 - \sup_{h \in L_I^2, \|h\|=1} \|\mathcal{F}^{(1)}(h)|_{I_{\mathfrak{f}}}\|^2 \right)^{1/2}. \end{aligned} \quad (8.3.12)$$

We thus need to estimate the norm of the 1D-operator $\mathcal{F}_{\mathfrak{f}} : L_I^2 \rightarrow L_{I_{\mathfrak{f}}}^2; h \mapsto \mathcal{F}^{(1)}(h)|_{I_{\mathfrak{f}}}$. This is achieved by explicit computation of the operator $\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}$ ($\mathcal{F}_{\mathfrak{f}}^*$: adjoint of $\mathcal{F}_{\mathfrak{f}}$), which turns out to be part of a well-studied family of compact and self-adjoint integral operators. Their eigenfunctions are known as prolate spheroidal wave functions, and the eigenpairs have been studied for example in [185, 186]. By applying these known results we obtain the following theorem:

Theorem 8.4 (Spectral characterization of $\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}$). *Let $\mathcal{F}_{\mathfrak{f}}$ defined by (8.3.11). Then*

$$\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}(f) \left(\frac{x}{2} \right) = \int_1^1 \frac{\sin(c(x-y))}{\pi(x-y)} f \left(\frac{y}{2} \right) dy \quad \text{with} \quad c := \mathfrak{f}/8. \quad (8.3.13)$$

for all $h \in L_I^2$ and $x \in [-1; 1]$, and $\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}$ is compact. The eigenvalues $\{\lambda_{c,j}\}_{j \in \mathbb{N}_0} \subset \mathbb{R}_+$ and associated eigenfunctions $\{\psi_{c,j}\}_{j \in \mathbb{N}_0} \subset L_I^2$ of $\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}$ thus coincide with those in [185, 186]. In particular, all eigenvalues $\lambda_{c,0} > \lambda_{c,1} > \dots$ have multiplicity one, and the $\psi_{c,j}$ may be chosen to form an orthonormal basis of L_I^2 . Moreover, $\lambda_{c,0} < 1$ holds true and, for fixed $j \in \mathbb{N}_0$ and $\mathfrak{f} \rightarrow \infty$, $\lambda_{c,j}$ has the asymptotic expansion

$$1 - \lambda_{c,j} = \frac{(2\pi)^{\frac{1}{2}} \mathfrak{f}^{j+\frac{1}{2}}}{j!} \left(1 - \frac{6j^2 - 2j + 3}{4\mathfrak{f}} + \mathcal{O}\left(\frac{1}{\mathfrak{f}^2}\right) \right) \exp(-\mathfrak{f}/4). \quad (8.3.14)$$

Proof. The restriction to the interval $I_{\mathfrak{f}} = [-\mathfrak{f}/4; \mathfrak{f}/4]$ can be written in the form of a multiplication with its indicator function $\mathbf{1}_{I_{\mathfrak{f}}}$. By the convolution theorem, we thus obtain for all $h \in L^2_{I_{\mathfrak{f}}}$, $x \in [-1; 1]$

$$\begin{aligned} \mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}(h) \left(\frac{x}{2} \right) &= \mathcal{F}^{-1} (\mathbf{1}_{I_{\mathfrak{f}}} \cdot \mathcal{F}(h)) \left(\frac{x}{2} \right) = (2\pi)^{-\frac{1}{2}} \mathcal{F}^{-1} (\mathbf{1}_{I_{\mathfrak{f}}}) * h \left(\frac{x}{2} \right) \\ &= \frac{1}{2\pi} \int_{-1/2}^{1/2} \frac{2 \sin \left(\frac{\mathfrak{f}}{4} \left(\frac{x}{2} - y \right) \right)}{\frac{1}{2}x - y} f(y) \, dy = \int_1^1 \frac{\sin(c(x-y))}{\pi(x-y)} f \left(\frac{y}{2} \right) \, dy. \end{aligned}$$

The spectral characterization of the resulting integral operators in [185, 186] directly yields the claimed properties of the eigensystem $\{(\lambda_{c,j}, \psi_{c,j})\}_{j \in \mathbb{N}_0}$ of $\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}$. In particular, the asymptotic expansion (8.3.14) is an analogue of the formula [185, eq. (2)].

Since $\mathcal{F}_{\mathfrak{f}}$ is a restriction of the Fourier transform, we have $\|\mathcal{F}_{\mathfrak{f}}\| \leq \|\mathcal{F}\| = 1$. Hence, the principal eigenvalue of $\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}$ must satisfy $\lambda_{c,0} \leq 1$. If $\lambda_{c,0} = 1$ then

$$\|\mathcal{F}(\psi_{c,0})|_{I_{\mathfrak{f}}^c}\|^2 = \|\mathcal{F}(\psi_{c,0})\|^2 - \|\mathcal{F}(\psi_{c,0})|_{I_{\mathfrak{f}}}\|^2 = \|\psi_{c,0}\|^2 - \lambda_{c,0} \|\psi_{c,0}\|^2 = 0,$$

i.e. $\mathcal{F}(\psi_{c,0})$ would have to vanish outside the interval $I_{\mathfrak{f}}$. However, as $\psi_{c,0}$ is compactly supported, $\mathcal{F}(\psi_{c,0})$ is an entire function and thus vanishes identically if $\mathcal{F}(\psi_{c,0})|_{I_{\mathfrak{f}}^c} = 0$. This is impossible since $\psi_{c,0}$ is an eigenfunction. Hence, $\lambda_{c,0} < 1$ must hold true. \square

We emphasize the nontrivial dependence of both the eigenvalues $\lambda_{c,j}$ and the eigenfunctions $\psi_{c,j}$ on the parameter $c = \mathfrak{f}/8$. For convenience, however, we will suppress the subscript c in the following. theorem 8.4 constitutes the final ingredient which is needed to prove the sought stability result for inverse problem 8.1:

Proof of theorem 8.1. According to the characterization of the stability constant in (8.3.12), $C_{\text{IP1}}(\Omega, \mathfrak{f})$ can be expressed in terms of the operator norm of $\mathcal{F}_{\mathfrak{f}}$. Since $\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}$ is compact with principal eigenvalue $\lambda_0 < 1$ and orthonormal eigenfunctions $\{\psi_j\}_{j \in \mathbb{N}_0}$ as characterized in theorem 8.4, we have

$$\|\mathcal{F}_{\mathfrak{f}}\|^2 = \|\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}}\| = \sup_{j \in \mathbb{N}_0} \|\mathcal{F}_{\mathfrak{f}}^* \mathcal{F}_{\mathfrak{f}} \psi_j\| = \sup_{j \in \mathbb{N}_0} \lambda_j = \lambda_0$$

By (8.3.12), this implies $C_{\text{IP1}}(\Omega, \mathfrak{f})^2 = 1 - \lambda_0 > 0$, i.e. *well-posedness* of inverse problem 8.1. Setting $j = 0$ in (8.3.14) yields the asymptotic characterization (8.2.11). \square

8.3.4 Characterization of the least stable modes

So far, we have not exploited the full potential of the reduction to a Fourier data completion problem in theorem 8.3 yet: only the characterization of the stability constant (8.3.7), i.e. of the *worst-case-stability*, has been used in the proof of theorem 8.1. Notably however, the more general estimate (8.3.6) even bounds the contrast $\|\mathcal{F}(h)\|$ attained by *individual images* h with respect to the corresponding incomplete Fourier data $\mathcal{F}(n_{\mathfrak{f}}^{1/2} \cdot h)|_{\Omega_{\mathfrak{f}}^c}$. This enables a precise prediction of the reconstruction stability for different image modes in inverse problem 8.1 beyond the universal lower bound proven in theorem 8.1.

In order to avoid notational difficulties in the argument, we do the analysis for a box-shaped support-domain $\Omega := [-1/2; 1/2]^m = I^m$. Owing to the simple Cartesian geometry, the results obtained for a stripe support are easily generalized to this case, including a characterization of the stability of individual modes. We define m -dimensional prolate spheroidal wave functions as the tensor product

$$\psi_{\mathbf{j}}(\mathbf{x}) := (\psi_{j_1} \otimes \dots \otimes \psi_{j_m})(\mathbf{x}) := \prod_{j=1}^m \psi_j(x_j) \quad \text{for } \mathbf{j} = (j_1, \dots, j_m) \in \mathbb{N}_0^m. \quad (8.3.15)$$

Moreover, let $\langle f, g \rangle := \int_{\Omega} f(\mathbf{x}) \overline{g(\mathbf{x})} \, d\mathbf{x}$ for $h_1, h_2 \in L^2(\Omega)$ denote usual L^2 -inner product. With this notation, we obtain the following modal stability estimates:

Theorem 8.5 (Stability of individual modes in inverse problem 8.1). *Let $\Omega = I^m$ and let $\{(\lambda_j, \psi_j)\}_{j \in \mathbb{N}_0}$ denote the eigenvalue decomposition of $\mathcal{F}_{\dagger}^* \mathcal{F}_{\dagger}$ in theorem 8.4. Moreover, define $\phi_{\mathbf{j}} := n_{\dagger}^{-1/2} \cdot \psi_{\mathbf{j}}$ for all multi-indices $\mathbf{j} \in \mathbb{N}_0^m$. Then $\{\phi_{\mathbf{j}}\}_{\mathbf{j} \in \mathbb{N}_0^m}$ is an orthonormal basis of $L^2(\Omega)$, and with $c_{\text{IP1}, \mathbf{j}} := (1 - \prod_{l=1}^m \lambda_{j_l})^{1/2}$ we have*

$$\|\mathcal{F}(h)\|^2 \geq \sum_{\mathbf{j} \in \mathbb{N}_0^m} c_{\text{IP1}, \mathbf{j}}^2 |\langle h, \phi_{\mathbf{j}} \rangle|^2 \quad \text{for any } h = \sum_{\mathbf{j} \in \mathbb{N}_0^m} \langle h, \phi_{\mathbf{j}} \rangle \phi_{\mathbf{j}} \in L^2(\Omega). \quad (8.3.16)$$

Proof. As $\{\psi_{\mathbf{j}}\}_{\mathbf{j} \in \mathbb{N}_0}$ is an orthonormal basis of L_I^2 , the tensor products $\psi_{\mathbf{j}} = \psi_{j_1} \otimes \dots \otimes \psi_{j_m}$ form an orthonormal basis of $L^2(\Omega)$ with $\Omega = I^m$. Since $n_{\dagger}^{-1/2}$ is unitary, the same is true for $\{\phi_{\mathbf{j}}\}_{\mathbf{j} \in \mathbb{N}_0^m}$ so that any $h \in L^2(\Omega)$ can be written as $h = \sum_{\mathbf{j} \in \mathbb{N}_0^m} a_{\mathbf{j}} \phi_{\mathbf{j}}$ with $a_{\mathbf{j}} := \langle h, \phi_{\mathbf{j}} \rangle$. By the tensor product structure of $\psi_{\mathbf{j}}$ and $\mathcal{F}^{(m)}$, we further have

$$\begin{aligned} \mathcal{F}^{(m)}(n_{\dagger}^{1/2} \cdot \phi_{\mathbf{j}})|_{\Omega_{\dagger}} &= \mathcal{F}^{(m)}(\psi_{\mathbf{j}})|_{I^m} = (\mathcal{F}^{(1)}\psi_{j_1})|_{I_1} \otimes \dots \otimes (\mathcal{F}^{(1)}\psi_{j_m})|_{I_m} \\ &= \mathcal{F}_{\dagger}\psi_{j_1} \otimes \dots \otimes \mathcal{F}_{\dagger}\psi_{j_m} \end{aligned}$$

for all $\mathbf{j} = (j_1, \dots, j_m) \in \mathbb{N}_0$. As the ψ_l are eigenfunctions of the 1D-map $\mathcal{F}_{\dagger}^* \mathcal{F}_{\dagger}$ to the eigenvalues λ_l (see theorem 8.4), the above relation implies

$$\begin{aligned} &\left\langle \mathcal{F}(n_{\dagger}^{1/2} \cdot \phi_{\mathbf{j}})|_{\Omega_{\dagger}}, \mathcal{F}(n_{\dagger}^{1/2} \cdot \phi_{\mathbf{k}})|_{\Omega_{\dagger}} \right\rangle \\ &= \prod_{l=1}^m \langle \mathcal{F}_{\dagger}^* \mathcal{F}_{\dagger}(\psi_{j_l}), \psi_{k_l} \rangle = \prod_{l=1}^m \lambda_{j_l} \langle \psi_{j_l}, \psi_{k_l} \rangle = \delta_{\mathbf{j}\mathbf{k}} \prod_{l=1}^m \lambda_{j_l} = \delta_{\mathbf{j}\mathbf{k}} (1 - c_{\text{IP1}, \mathbf{j}}^2) \end{aligned} \quad (8.3.17)$$

for all $\mathbf{j}, \mathbf{k} \in \mathbb{N}_0^m$ where $\delta_{\mathbf{j}\mathbf{k}} \in \{0, 1\}$ is the multidimensional Kronecker-symbol. Using the principal bound (8.3.6) from theorem 8.3, unitarity of \mathcal{F} and (8.3.17) finally gives

$$\begin{aligned} \|\mathcal{F}(h)\|^2 &\geq \|\mathcal{F}(n_{\dagger}^{1/2} \cdot h)|_{\Omega_{\dagger}}\|^2 = \|h\|^2 - \|\mathcal{F}(n_{\dagger}^{1/2} \cdot h)|_{\Omega_{\dagger}}\|^2 \\ &= \sum_{\mathbf{j} \in \mathbb{N}_0^m} |a_{\mathbf{j}}|^2 - \sum_{\mathbf{j}, \mathbf{k} \in \mathbb{N}_0^m} a_{\mathbf{j}} \overline{a_{\mathbf{k}}} \left\langle \mathcal{F}(n_{\dagger}^{1/2} \cdot \phi_{\mathbf{j}})|_{\Omega_{\dagger}}, \mathcal{F}(n_{\dagger}^{1/2} \cdot \phi_{\mathbf{k}})|_{\Omega_{\dagger}} \right\rangle = \sum_{\mathbf{j} \in \mathbb{N}_0^m} c_{\text{IP1}, \mathbf{j}}^2 |a_{\mathbf{j}}|^2. \end{aligned}$$

□

In theorem 8.5 we obtain *individual* stability constants $c_{\text{IP1},\mathbf{j}}$ such that each of the constructed orthonormal basis modes $\phi_{\mathbf{j}}$ attains data contrast $\|\mathcal{S}(\phi_{\mathbf{j}})\| \geq c_{\text{IP1},\mathbf{j}}$. As the sequence of eigenvalues $\{\lambda_j\}_{k \in \mathbb{N}_0}$ in theorem 8.4 is strictly decreasing, it is readily seen that stability *increases* with the multi-index \mathbf{j} , i.e.

$$c_{\text{IP1},\mathbf{k}} > c_{\text{IP1},\mathbf{j}} \quad \text{if} \quad \mathbf{k} \neq \mathbf{j}, k_l \geq j_l \quad \text{for all} \quad l = 1, \dots, m. \quad (8.3.18)$$

According to theorem 8.5, the least stable modes are thus exactly the basis functions $\phi_{\mathbf{j}}$ of (componentwise) small multi-index \mathbf{j} , which are given by prolate spheroidal wave functions $\psi_{\mathbf{j}} = \psi_{j_1} \otimes \dots \otimes \psi_{j_m}$, modulated by the Fresnel factor $n_{\mathbf{f}}^{-1/2}$. Details on the shape of the ψ_j in turn are readily available, see e.g. [185]. In particular, the index j can be interpreted as a *frequency* since ψ_j is smooth and real-valued with exactly j zeros and $j+1$ extrema within the interval $(-1/2; 1/2)$. By extending this observation to the $\phi_{\mathbf{j}}$ we deduce the nature of the least stable modes in inverse problem 8.1:

Corollary 8.6 (Least stable modes in inverse problem 8.1). *The least stable modes in inverse problem 8.1 for $\Omega = [-1/2; 1/2]^m$ are low-frequency prolate spheroidal wave functions modulated by the Fresnel-factor $n_{\mathbf{f}}^{-1/2}(\mathbf{x}) = \exp(-i\mathbf{f}|\mathbf{x}|^2/4)$.*

8.3.5 Numerical Validation

We return to the starting point of our stability analysis, namely the reduction to a Fourier data completion problem via the principal estimate (8.3.6). The question whether or not the derived stability bounds in theorem 8.1 and theorem 8.5 are *optimal* (or at least close to) crucially depends on the sharpness of this inequality. In the following we investigate this remaining issue numerically.

To this end, we approximate the stability bound $C_{\text{IP1}}(\Omega, \mathbf{f})$ by computing the smallest singular value of a discretized forward operator \mathcal{S} in $m = 1$ dimensions. The Fresnel propagator is approximated by fast Fourier transforms on a 1D-grid of 512^2 equidistant points, where the central 512 grid points form the supporting interval $\Omega = [-1/2; 1/2]$. We compute the smallest singular value of \mathcal{S} via a power method for Fresnel numbers $\mathbf{f} \in [1; 10]$. These numerical results for the complete, yet discretized forward operator, are compared to the asymptotic stability bound (8.2.11) in theorem 8.1, neglecting the $\mathcal{O}(\mathbf{f}^{-2})$ -contributions. The different predictions for the stability constant $C_{\text{IP1}}(\Omega, \mathbf{f})$ are plotted in fig. 8.3(a).

The semilogarithmic plot shows excellent agreement between the analytical bound (8.2.11) and the numerical approximation in the asymptotic limit $\mathbf{f} \rightarrow \infty$. This indicates that our stability analysis, based on the potentially lossy estimate (8.3.6), is surprisingly sharp. Accordingly, one might expect that also the corresponding least stable modes $\phi_{\mathbf{j}}$ are well-characterized by theorem 8.5. This is supported by the simulation results: fig. 8.3(b) exemplarily plots the numerically computed mode ϕ_0 to the minimum singular value of \mathcal{S} for $\mathbf{f} = \mathbf{f}/(2\pi) = 10$. According to theorem 8.1, the plotted product with the factor $n_{\mathbf{f}}^{1/2}$ should yield the zeroth order prolate spheroidal wave function ψ_0 . This is confirmed by the smooth unimodal profiles obtained in fig. 8.3(b).

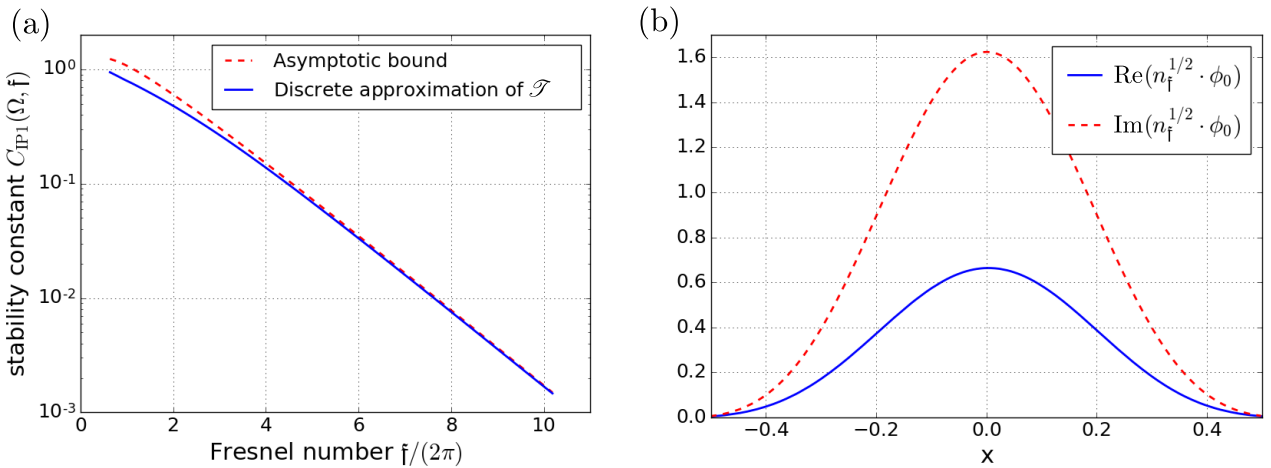


Figure 8.3. Numerical validation of the stability analysis for inverse problem 8.1.

(a) Comparison of the analytical bound (8.2.11) for the stability constant C_{IP1} (red-dashed line) to numerical computations of the smallest singular value of \mathcal{T} (blue-solid).

(b) Numerically computed least stable singular mode ϕ_0 for $\mathfrak{f} = \mathfrak{f}/(2\pi) = 10$. The plotted modulation with $n_{\mathfrak{f}}^{1/2}$ reveals a unimodal structure as predicted by theorem 8.5.

8.4 Stability analysis of inverse problem 8.2

The aim of this section is to prove theorem 8.2, establishing *algebraic* rates of the stability constant $C_{\text{IP2}}(\Omega, \mathfrak{f}) \gtrsim \mathfrak{f}^{-1}$ under a homogeneity constraint for the imaged object.

8.4.1 Preparations and Fourier domain splitting

Throughout this section, let $\varphi \in L^2(\mathbb{R}^m)$ and let $\hat{\varphi} := \mathcal{F}(\varphi)$ denote its Fourier transform. Recall from definition (8.2.9) that the operator \mathcal{S}_ν multiplies with $-2s_\nu(\boldsymbol{\xi}) = -2 \sin(|\boldsymbol{\xi}|^2/(2\mathfrak{f}) + \nu)$ in the Fourier domain, and hence

$$\|\mathcal{S}_\nu(\varphi)\| = \|\mathcal{F}\mathcal{S}_\nu(\varphi)\| = 2\|s_\nu \cdot \hat{\varphi}\|. \quad (8.4.1)$$

According to (8.4.1), the images φ attaining low contrast, i.e. small $\|\mathcal{S}_\nu(\varphi)\|$, are exactly those for which $\hat{\varphi}$ is concentrated about the zero-manifolds of the CTF s_ν . As signals $\hat{\varphi} \in L^2(\mathbb{R}^m)$ may be arbitrarily sharply peaked about these manifolds of zero contrast, inverse problem 8.2 is ill-posed for general images $\varphi \in L^2(\mathbb{R}^m)$.

Now, we additionally assume a support constraint $\varphi \in L^2(\Omega)$ for $\Omega := B[0; R]$. Why does this constraint ensure well-posedness in the light of the problematic CTF-zeros? The explanation lies in the well-known fact that a compact support in real-space implies \mathcal{C}^∞ -smoothness (indeed analyticity) of the Fourier transform with norm-bounds on the derivatives in terms of the support size R . Owing to this regularity, $\hat{\varphi}$ may not be arbitrarily concentrated about the zero-manifolds of s_ν , which enables stability as illustrated in fig. 8.4. The following lemma quantifies the smoothness of $\hat{\varphi}$ in a suitable form for the subsequent analysis:

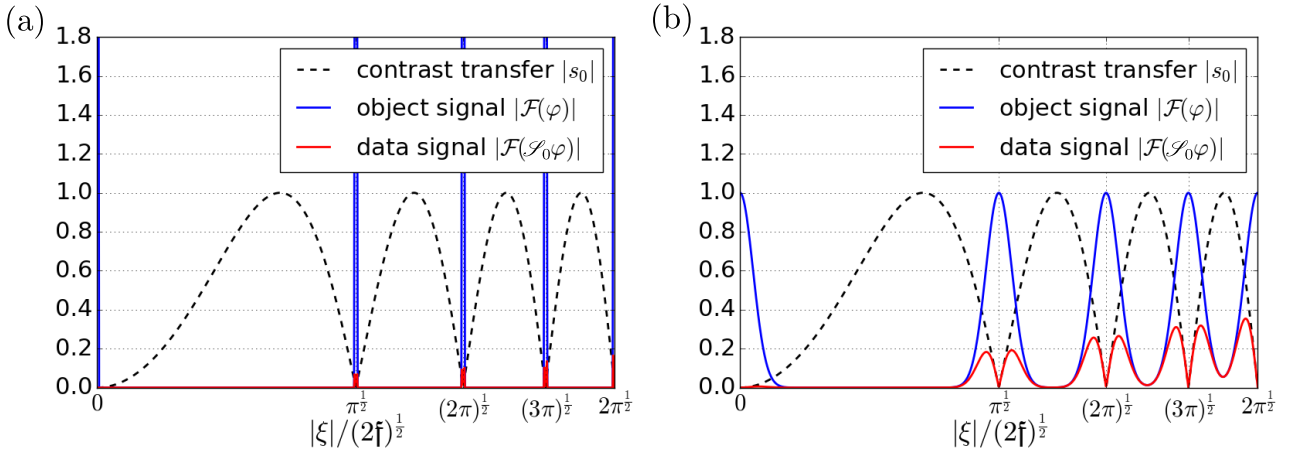


Figure 8.4. Illustration of the principal argument for stability of inverse problem 8.2: the black-dashed line plots the radially symmetric contrast transfer function s_ν for $\nu = 0$. Blue and red solid lines show examples of Fourier space signals of objects φ and their images under the forward operator \mathcal{S}_ν , respectively.

(a) General images $\varphi \in L^2(\mathbb{R}^m)$: $\mathcal{F}(\varphi)$ may be arbitrarily peaked at the zeros of s_ν .

(b) Support constraint $\varphi \in L^2(\Omega)$: $\mathcal{F}(\varphi)$ is smooth and has finitely sharp peaks, which ensures that the total contrast is bounded from below: $\|\mathcal{F}(\mathcal{S}_\nu(\varphi))\| \geq C_{IP2} \|\mathcal{F}(\varphi)\|$.

Lemma 8.7. Let $g \in L^2(\Omega)$ with support in $\Omega = B[0; R]$ and Fourier transform $\hat{g} := \mathcal{F}(g)$. Let Δ be the Laplacian on \mathbb{R}^m . Then we have for any measurable set $B \subset \mathbb{R}^m$

$$\int_B -\Delta |\hat{g}|^2 \, d\xi \leq 2R^2 \|g\|^2. \quad (8.4.2)$$

Proof. The compact support of g implies infinite smoothness of $\hat{g} = \mathcal{F}(g)$. Using the identity $-\Delta \mathcal{F}(g) = \mathcal{F}(|\mathbf{x}|^2 g)$ and Cauchy-Schwarz's inequality we obtain

$$\begin{aligned} \int_B -\Delta |\hat{g}|^2 \, d\xi &= -2 \int_B \left(|\nabla \hat{g}|^2 + \operatorname{Re}(\bar{\hat{g}} \cdot \Delta \hat{g}) \right) \, d\xi \leq 2 \left| \int_B \bar{\hat{g}} \cdot \mathcal{F}(|\mathbf{x}|^2 g) \, d\xi \right| \\ &\leq 2 \|\hat{g}\|_B \|\mathcal{F}(|\mathbf{x}|^2 g)\|_B \leq 2 \|\hat{g}\| \|\mathcal{F}(|\mathbf{x}|^2 g)\| \leq 2 \|g\| \|\mathbf{x}|^2 g\|. \end{aligned}$$

As g vanishes outside $B[0; R]$, we further have $\|\mathbf{x}|^2 g\| \leq R^2 \|g\|$. \square

Note that lemma 8.7 can be interpreted as an *uncertainty principle*: a bound for the derivative $-\Delta |\mathcal{F}(g)|^2$, limiting the sharpness of features in Fourier space, arises from the confinement $\operatorname{supp}(g) \subset B[0; R]$ of the corresponding real-space signal g .

For a quantitative analysis, we decompose the norm on the right-hand side of (8.4.1) into stable bulk integrals and potentially unstable parts about the CTF-zeros by cutting out the subdomain $D_\varepsilon := \{\xi \in \mathbb{R}^m : |s_\nu(\xi)| \geq \sin(\varepsilon)\}$ for some $0 < \varepsilon \leq \pi/6$:

$$\|s_\nu \cdot \hat{\varphi}\|^2 = \underbrace{\int_{D_\varepsilon} s_\nu(\xi)^2 |\hat{\varphi}(\xi)|^2 \, d\xi}_{=: J_\varepsilon} + \sum_{j=0}^{\infty} \underbrace{\int_{B_j} s_\nu(\xi)^2 |\hat{\varphi}(\xi)|^2 \, d\xi}_{=: J_j} \quad (8.4.3)$$

Here, the B_j denote the annular connected components of $\mathbb{R}^m \setminus D_\varepsilon$ about the j -th zero-manifold of s_ν at radius $\xi_j := (2\mathfrak{f})^{\frac{1}{2}}(j\pi - \nu)^{\frac{1}{2}}$, i.e.

$$\begin{aligned} B_j &= \{\boldsymbol{\xi} \in \mathbb{R}^m : |\boldsymbol{\xi}| \in (b_{j-}; b_{j+})\}, & b_{j\pm} &:= (\xi_j^2 \pm 2\mathfrak{f}\varepsilon)^{\frac{1}{2}} \quad \text{for } j \in \mathbb{N} \\ B_0 &= \{\boldsymbol{\xi} \in \mathbb{R}^m : |\boldsymbol{\xi}| < b_0\}, & b_0 &:= (2\mathfrak{f})^{\frac{1}{2}} \max(\varepsilon - \nu, 0)^{\frac{1}{2}}. \end{aligned} \quad (8.4.4)$$

Note that the constructed Fourier domain splitting is disjoint, i.e. $\mathbb{R}^m = (\bigsqcup_{j=0}^{\infty} B_j) \sqcup D_\varepsilon$. From the definition of the domain D_ε , it immediately follows that

$$J_\varepsilon \geq \sin(\varepsilon) \|\hat{\varphi}|_{D_\varepsilon}\|. \quad (8.4.5)$$

Hence, what remains to be done is to derive bounds for the sub-integrals J_j around the zero-manifolds of s_ν and to choose ε to balance the contributions in (8.4.3).

8.4.2 Estimate for the central CTF-minimum

We first consider the ball-shaped domain B_0 , i.e. the low frequency part of the Fourier domain splitting. Note that B_0 does not contain a zero of s_ν if $\nu > 0$ but still a (possibly small) local minimum of s_ν^2 at $\boldsymbol{\xi} = 0$. Since $B_0 = \emptyset$ for $\nu \geq \varepsilon$, we may restrict to the case $\nu < \varepsilon$.

By (8.4.4) and the assumption $\varepsilon \leq \pi/6$, we have $|\boldsymbol{\xi}|^2/(2\mathfrak{f}) + \nu \in [0; \pi/6)$ and hence $s_\nu(\boldsymbol{\xi})^2 \geq C_0(|\boldsymbol{\xi}|^2/(2\mathfrak{f}) + \nu)^2$ for all $\boldsymbol{\xi} \in B_0$ with $C_0 := \sin(\pi/6)^2/(\pi/6)^2$. This implies

$$J_0 = \int_{B_0} s_\nu(\boldsymbol{\xi})^2 |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \geq C_0 \int_{B_0} \left(\frac{|\boldsymbol{\xi}|^4}{4\mathfrak{f}^2} + \frac{\nu|\boldsymbol{\xi}|^2}{\mathfrak{f}} + \nu^2 \right) |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \quad (8.4.6)$$

From (8.4.6), it can be seen that J_0 is bounded from below by $\nu^2 \|\hat{\varphi}|_{B_0}\|^2$ and thus indeed stable if $\nu \neq 0$. However, since ν is typically (almost) zero in hard X-ray imaging, we have to resort to the integral summands that involve powers of $|\boldsymbol{\xi}|^2$ in order to achieve robust estimates. These integrands have a zero at $\boldsymbol{\xi} = 0 \in B_0$, which is why smoothness of $|\hat{\varphi}|^2$ has to be exploited to obtain reasonable estimates. This is achieved by the following lemma:

Lemma 8.8. *Let $D := \{\boldsymbol{x} \in \mathbb{R}^m : |\boldsymbol{x}| \leq a\}$ be a closed concentric ball of radius a in \mathbb{R}^m . Let $g, w \in \mathcal{C}^2(D)$ where w is radially symmetric, i.e. $w(\boldsymbol{x}) = w_0(|\boldsymbol{x}|)$ for all $\boldsymbol{x} \in D$ and some function $w_0 : [0; a] \rightarrow \mathbb{R}$. Then*

$$\int_D \Delta w \cdot g \, d\boldsymbol{x} = \frac{mw_0'(a)}{a} \int_D g \, d\boldsymbol{x} + \int_D \left(\frac{w_0'(a)}{2a} (a^2 - |\boldsymbol{x}|^2) - (w_0(a) - w) \right) \Delta g \, d\boldsymbol{x}.$$

Proof. By Green's second identity we have

$$\begin{aligned} \int_D \Delta w \cdot g \, d\boldsymbol{x} &= \int_D w \cdot \Delta g \, d\boldsymbol{x} + \int_{\partial D} \left(g \frac{\partial w}{\partial \boldsymbol{n}} - w \frac{\partial g}{\partial \boldsymbol{n}} \right) \, dS(x) \\ &= \int_D w \cdot \Delta g \, d\boldsymbol{x} + w_0'(a) \int_{\partial D} g \, dS(x) - w_0(a) \int_{\partial D} \frac{\partial g}{\partial \boldsymbol{n}} \, dS(x). \end{aligned} \quad (8.4.7)$$

Here, $\partial/(\partial\mathbf{n})$ denotes the derivative along the unit normal vector pointing to the outside of ∂D . The boundary terms can be eliminated via the relations

$$\begin{aligned} \int_{\partial D} \frac{\partial g}{\partial \mathbf{n}} \, dS(x) &= \int_D \Delta g \, d\mathbf{x} \\ \int_{\partial D} g \, dS(x) &= \frac{1}{2a} \int_{\partial D} g \frac{\partial(|\mathbf{x}|^2)}{\partial \mathbf{n}} \, dS(x) \\ &= \frac{1}{2a} \left(\int_{\partial D} |\mathbf{x}|^2 \frac{\partial g}{\partial \mathbf{n}} \, dS(x) + 2m \int_D g \, d\mathbf{x} - \int_D |\mathbf{x}|^2 \Delta g \, d\mathbf{x} \right) \\ &= \frac{1}{2a} \left(2m \int_D g \, d\mathbf{x} + \int_D (a^2 - |\mathbf{x}|^2) \Delta g \, d\mathbf{x} \right), \end{aligned}$$

which again follow from Green's second identity. Plugging this into (8.4.7) yields the claimed identity. \square

As the functions $\boldsymbol{\xi} \mapsto |\boldsymbol{\xi}|^2$ and $\boldsymbol{\xi} \mapsto |\boldsymbol{\xi}|^4$ are radially symmetric and since B_0 is a concentric ball of radius b_0 , we find that lemma 8.8 can be applied to the integrals in (8.4.6). This yields the following estimate for the considered low-frequency subdomain:

Lemma 8.9. *Let $\nu < \varepsilon \leq \pi/6$ and $B_0^- := \{\boldsymbol{\xi} \in B_0 : \Delta|\hat{\varphi}|^2(\boldsymbol{\xi}) \leq 0\}$. Then*

$$\begin{aligned} J_0 &\geq C_0 \left(\left(\frac{m}{m+4}(\varepsilon - \nu)^2 + \frac{2m}{m+2}\nu(\varepsilon - \nu) + \nu^2 \right) \|\hat{\varphi}|_{B_0}\|^2 \right. \\ &\quad \left. + \mathfrak{f} \left(\frac{2}{3(m+4)}(\varepsilon - \nu)^3 + \frac{1}{m+2}\nu(\varepsilon - \nu)^2 \right) \int_{B_0^-} \Delta|\hat{\varphi}|^2 \, d\boldsymbol{\xi} \right). \end{aligned} \quad (8.4.8)$$

Proof. For the functions $w_1(\boldsymbol{\xi}) := |\boldsymbol{\xi}|^6/(6(m+4))$ and $w_2(\boldsymbol{\xi}) := |\boldsymbol{\xi}|^4/(4(m+2))$, we have that $\Delta w_1(\boldsymbol{\xi}) = |\boldsymbol{\xi}|^4$ and $\Delta w_2(\boldsymbol{\xi}) = |\boldsymbol{\xi}|^2$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$. Accordingly, the integral in (8.4.6) matches the setting of lemma 8.8. An application gives

$$\begin{aligned} \int_{B_0} |\boldsymbol{\xi}|^4 |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} &= \int_{B_0} \Delta w_1(\boldsymbol{\xi}) |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \\ &= \frac{mb_0^4}{m+4} \int_{B_0} |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} + \frac{1}{m+4} \int_{B_0} \left(\frac{b_0^4}{2}(b_0^2 - |\boldsymbol{\xi}|^2) - \frac{1}{6}(b_0^6 - |\boldsymbol{\xi}|^6) \right) \Delta|\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \\ &\geq \frac{mb_0^4}{m+4} \int_{B_0} |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} + \frac{b_0^6}{3(m+4)} \int_{B_0^-} \Delta|\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \end{aligned}$$

Here, the integral over $\Delta|\hat{\varphi}|^2$ has been bounded from below by the integral within the subdomain $B_0^- = \{\boldsymbol{x} \in B_0 : \Delta|\hat{\varphi}|^2(\boldsymbol{x}) \leq 0\}$ multiplied by the maximum of the non-negative factor $(b_0^4/2)(b_0^2 - |\boldsymbol{\xi}|^2) - \frac{1}{6}(b_0^6 - |\boldsymbol{\xi}|^6)$. Analogously we obtain

$$\begin{aligned} \int_{B_0} |\boldsymbol{\xi}|^2 |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} &= \int_{B_0} \Delta w_2(\boldsymbol{\xi}) |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \\ &= \frac{mb_0^2}{m+2} \int_{B_0} |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} + \frac{1}{m+2} \int_{B_0} \left(\frac{b_0^2}{2}(b_0^2 - |\boldsymbol{\xi}|^2) - \frac{1}{4}(b_0^4 - |\boldsymbol{\xi}|^4) \right) \Delta|\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \\ &\geq \frac{mb_0^2}{m+2} \int_{B_0} |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} + \frac{b_0^4}{4(m+2)} \int_{B_0^-} \Delta|\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \end{aligned}$$

Inserting these relations into (8.4.6) yields

$$\begin{aligned}
 J_0 &\geq C_0 \int_{B_0} \left(\frac{|\boldsymbol{\xi}|^4}{4\mathfrak{f}^2} + \frac{\nu|\boldsymbol{\xi}|^2}{\mathfrak{f}} + \nu^2 \right) |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \\
 &\geq C_0 \left(\frac{mb_0^4}{4\mathfrak{f}^2(m+4)} + \frac{m\nu b_0^2}{\mathfrak{f}(m+2)} + \nu^2 \right) \int_{B_0} |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \\
 &\quad + C_0 \left(\frac{b_0^6}{12\mathfrak{f}^2(m+4)} + \frac{b_0^4\nu}{4\mathfrak{f}(m+2)} \right) \int_{B_0^-} \Delta |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi}.
 \end{aligned}$$

Re-substituting $b_0 = (2\mathfrak{f})^{\frac{1}{2}}(\epsilon - \nu)^{\frac{1}{2}}$ into this estimate gives the assertion. \square

8.4.3 Estimate for the first order CTF-zeros

Before we can derive a global stability estimate from (8.4.8), we have to consider the remaining integrals J_j , $j \geq 1$ in the Fourier domain splitting (8.4.3). The integration domains B_j are annular shells around the first order zero-manifolds $Z_j = \{\boldsymbol{\xi} \in \mathbb{R}^m : |\boldsymbol{\xi}| = \xi_j\}$, $j \in \mathbb{N}$ of the CTF s_ν . According to (8.4.4), we have $|\boldsymbol{\xi}|^2/(2\mathfrak{f}) + \nu - j\pi \in (-\epsilon; \epsilon)$ for all $\boldsymbol{\xi} \in B_j$ and thus

$$s_\nu(\boldsymbol{\xi})^2 = \sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathfrak{f}} + \nu + j\pi \right)^2 \geq \frac{\sin(\epsilon)^2}{\epsilon^2} \left(\frac{|\boldsymbol{\xi}|^2}{2\mathfrak{f}} + \nu + j\pi \right)^2 \stackrel{\epsilon \leq \frac{\pi}{6}}{\geq} \frac{C_0}{4\mathfrak{f}^2} (|\boldsymbol{\xi}|^2 - \xi_j^2)^2. \quad (8.4.9)$$

Using that $(|\boldsymbol{\xi}|^2 - \xi_j^2)^2 = (|\boldsymbol{\xi}| + \xi_j)^2(|\boldsymbol{\xi}| - \xi_j)^2 \geq (b_{j-} + \xi_j)^2(|\boldsymbol{\xi}| - \xi_j)^2$, this yields

$$J_j = \int_{B_j} s_\nu(\boldsymbol{\xi})^2 |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \geq \frac{C_0(b_{j-} + \xi_j)^2}{4\mathfrak{f}^2} \tilde{J}_j, \quad \tilde{J}_j := \int_{B_j} (|\boldsymbol{\xi}| - \xi_j)^2 |\hat{\varphi}(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi}. \quad (8.4.10)$$

for $j \in \mathbb{N}$. Accordingly, we have to estimate polynomially weighted integrals, similar to the preceding section, yet within the shell-shaped domains B_j instead of a concentric ball. We achieve this by introducing polar coordinates and applying lemma 8.8 to the resulting one-dimensional radial integrals:

Lemma 8.10 (Stability estimate for the first order CTF-zeros). *Let $0 < \epsilon \leq \pi/6$ and $j \in \mathbb{N}$. Then there exists a measurable subset $B_j^- \subset B_j$ and a constant $0 < C_1 \leq 1/4$, which depends only on m , such that*

$$J_j \geq C_0 \left(C_1 \epsilon^2 \|\hat{\varphi}|_{B_j}\|^2 + \frac{\epsilon^4 \mathfrak{f}}{32(j\pi - \nu)} \int_{B_j^-} \Delta |\hat{\varphi}|^2(\boldsymbol{\xi}) \, d\boldsymbol{\xi} \right). \quad (8.4.11)$$

Proof. Let $j \in \mathbb{N}$ arbitrary. In order to make lemma 8.8 applicable to the present setting, we rewrite the integral \tilde{J}_j in (8.4.10) by introducing polar coordinates $\boldsymbol{\xi} = \xi \boldsymbol{\theta}$, $\xi \geq 0$, $\boldsymbol{\theta} \in \mathbb{S}^{m-1}$ (\mathbb{S}^{m-1} : unit sphere in \mathbb{R}^m). With the notation in (8.4.4), this yields

$$\tilde{J}_j = \int_{b_{j-}}^{b_{j+}} (\xi - \xi_j)^2 \xi^{m-1} \varphi_{\text{rad}}(\xi) \, d\xi, \quad \varphi_{\text{rad}}(\xi) := \int_{\mathbb{S}^{m-1}} |\hat{\varphi}(\xi \boldsymbol{\theta})|^2 \, d\boldsymbol{\theta}. \quad (8.4.12)$$

Setting $a_j := b_{j+} - \xi_j$, $\eta := \xi - \xi_j$, $\varphi_j(\xi - \xi_j) := \xi^{(m-1)/2} \varphi_{\text{rad}}(\xi)$, we obtain the bound

$$\begin{aligned} \tilde{J}_j &= \int_{\xi_j - a_j}^{\xi_j + a_j} (\xi - \xi_j)^2 \xi^{\frac{m-1}{2}} \varphi_j(\xi - \xi_j) \, d\xi + \int_{b_{j-}}^{\xi_j - a_j} (\xi - \xi_j)^2 \xi^{m-1} \varphi_{\text{rad}}(\xi) \, d\xi \\ &\geq (\xi_j - a_j)^{\frac{m-1}{2}} \int_{-a_j}^{a_j} \eta^2 \varphi_j(\eta) \, d\eta + a_j^2 \int_{b_{j-}}^{\xi_j - a_j} \xi^{m-1} \varphi_{\text{rad}}(\xi) \, d\xi. \end{aligned} \quad (8.4.13)$$

The first integral on the right hand side of (8.4.13) matches the setting of lemma 8.8 in $m = 1$ dimensions with weight function $w(\eta) := \eta^4/12$. This yields

$$\begin{aligned} \int_{-a_j}^{a_j} \eta^2 \varphi_j(\eta) \, d\eta &= \frac{a_j^2}{3} \int_{-a_j}^{a_j} \varphi_j(\eta) \, d\eta + \int_{-a_j}^{a_j} \left(\frac{a_j^2(a_j^2 - \eta^2)}{6} - \frac{a_j^4 - \eta^4}{12} \right) \varphi_j''(\eta) \, d\eta \\ &\geq \frac{a_j^2}{3} \int_{-a_j}^{a_j} \varphi_j(\eta) \, d\eta + \frac{a_j^4}{12} \int_{I_j^-} \varphi_j''(\eta) \, d\eta, \end{aligned} \quad (8.4.14)$$

where we have set $I_j^- := \{\eta \in [-a_j; a_j] : \varphi_j''(\eta) \leq 0\}$ and used that $0 \leq (a_j^2/6)(a_j^2 - \eta^2) - (a_j^4 - \eta^4)/12 \leq a_j^4/12$. Setting $B_j^- := \{\boldsymbol{\xi} \in B_j : |\boldsymbol{\xi}| \in I_j^- + \xi_j\}$ and re-substituting φ_{rad} and $|\hat{\varphi}|^2$ for φ_j , we estimate the second integral in (8.4.14) by

$$\begin{aligned} &\underbrace{(\xi_j - a_j)^{\frac{m-1}{2}}}_{\leq \xi^{(m-1)/2} \text{ for } \xi \in I_j^- + \xi_j} \int_{I_j^-} \underbrace{\varphi_j''(\eta)}_{\leq 0} \, d\eta \geq \int_{I_j^- + \xi_j} \xi^{\frac{m-1}{2}} \partial_\xi^2 \left(\xi^{\frac{m-1}{2}} \varphi_{\text{rad}}(\xi) \right) \, d\xi \\ &= \int_{I_j^- + \xi_j} \xi^{m-1} \int_{\mathbb{S}^{m-1}} \left(\partial_\xi^2 + \frac{m-1}{\xi} \partial_\xi + \frac{(m-1)(m-3)}{4\xi^2} \right) |\hat{\varphi}(\xi\boldsymbol{\theta})|^2 \, d\boldsymbol{\theta} \, d\xi \\ &= \int_{B_j^-} \left(\Delta + \frac{(m-1)(m-3)}{4|\boldsymbol{\xi}|^2} \right) |\hat{\varphi}|^2(\xi\boldsymbol{\theta}) \, d\boldsymbol{\xi} \\ &\geq \int_{B_j^-} \Delta |\hat{\varphi}|^2(\boldsymbol{\xi}) \, d\boldsymbol{\xi} - \frac{\delta_{m2}}{4b_{j-}^2} \int_{B_j^-} |\hat{\varphi}|^2(\boldsymbol{\xi}) \, d\boldsymbol{\xi}. \end{aligned} \quad (8.4.15)$$

Here we have identified the radial part of the Laplacian in polar coordinates and implicitly added the angular part, which does not contribute to the \mathbb{S}^{m-1} -integrals. The inequality in the last line follows from the fact that $(m-1)(m-3)/(4|\boldsymbol{\xi}|^2)$ is negative only for $m = 2$, in which case it is bounded by the given term (δ_{ij} : Kronecker-Delta). Re-substituting $|\hat{\varphi}|^2$ in the remaining integrals in (8.4.13) and (8.4.14) gives

$$\begin{aligned} &(\xi_j - a_j)^{\frac{m-1}{2}} \frac{a_j^2}{3} \int_{-a_j}^{a_j} \varphi_j(\eta) \, d\eta + a_j^2 \int_{b_{j-}}^{\xi_j - a_j} \xi^{m-1} \varphi_{\text{rad}}(\xi) \, d\xi \\ &\geq \left(\frac{\xi_j - a_j}{\xi_j + a_j} \right)^{\frac{m-1}{2}} \frac{a_j^2}{3} \int_{\xi_j - a_j}^{\xi_j + a_j} \xi^{m-1} \varphi_{\text{rad}}(\xi) \, d\xi + a_j^2 \int_{b_{j-}}^{\xi_j - a_j} \xi^{m-1} \varphi_{\text{rad}}(\xi) \, d\xi \\ &\geq \left(\frac{\xi_j - a_j}{\xi_j + a_j} \right)^{\frac{m-1}{2}} \frac{a_j^2}{3} \int_{b_{j-}}^{b_{j+}} \xi^{m-1} \int_{\mathbb{S}^{m-1}} |\hat{\varphi}|^2(\xi\boldsymbol{\theta}) \, d\boldsymbol{\theta} \, d\xi = \left(\frac{\xi_j - a_j}{\xi_j + a_j} \right)^{\frac{m-1}{2}} \frac{a_j^2}{3} \|\hat{\varphi}\|_{B_j}^2. \end{aligned} \quad (8.4.16)$$

Combining the estimates (8.4.13), (8.4.14), (8.4.15) and (8.4.16), we thus obtain for all $j \in \mathbb{N}$

$$\tilde{J}_j \geq \underbrace{\frac{a_j^2}{3} \left(\left(\frac{\xi_j - a_j}{\xi_j + a_j} \right)^{\frac{m-1}{2}} - \frac{a_j^2}{16b_{j-}^2} \right)}_{=: \lambda} \|\hat{\varphi}\|_{B_j}^2 + \underbrace{\frac{a_j^4}{12}}_{=: \mu} \int_{B_j^-} \Delta |\hat{\varphi}|^2(\boldsymbol{\xi}) \, d\boldsymbol{\xi} \quad (8.4.17)$$

What remains to be done is to derive uniform bounds for the constants in (8.4.17) within the assumed parameter range $j \in \mathbb{N}$, $\nu \leq \pi/2$ and $0 < \varepsilon \leq \pi/6$. First, we have $\xi_j^2 = 2\mathfrak{f}(\pi j - \nu) \geq \pi\mathfrak{f}$ and hence by Taylor-expansion of the square-root

$$a_j = (\xi_j^2 + 2\mathfrak{f}\varepsilon)^{\frac{1}{2}} - \xi_j \in \frac{\varepsilon\mathfrak{f}}{\xi_j} \cdot \left[1 - \frac{\varepsilon\mathfrak{f}}{2\xi_j^2}; 1\right] \subset \frac{\varepsilon\mathfrak{f}}{\xi_j} \cdot \left[1 - \frac{\varepsilon}{2\pi}; 1\right] \subset \frac{\varepsilon\mathfrak{f}}{\xi_j} \cdot \left[\frac{11}{12}; 1\right]. \quad (8.4.18)$$

Using (8.4.18) along with $b_{j-}^2 = \xi_j^2 - 2\mathfrak{f}\varepsilon \stackrel{\varepsilon \leq \pi/6}{\geq} \xi_j^2 - \pi\mathfrak{f}/3 \stackrel{\xi_j^2 \geq \pi\mathfrak{f}}{\geq} 2\pi\mathfrak{f}/3$ furthermore gives

$$\frac{a_j^2}{16b_{j-}^2} \stackrel{(8.4.18)}{\leq} \frac{\varepsilon^2\mathfrak{f}^2}{16\xi_j^2(\xi_j^2 - 2\mathfrak{f}\varepsilon)} \leq \frac{3\varepsilon^2}{32\pi^2} \leq \frac{1}{384} \quad (8.4.19a)$$

$$\left(\frac{\xi_j - a_j}{\xi_j + a_j}\right)^{\frac{1}{2}} \geq 1 - \frac{a_j}{\xi_j} \stackrel{(8.4.18)}{\geq} 1 - \frac{\varepsilon\mathfrak{f}}{\xi_j^2} \geq 1 - \frac{\varepsilon}{\pi} \geq \frac{5}{6} \quad (8.4.19b)$$

Combining (8.4.18) and (8.4.19) gives $\lambda \geq (4/3)C_1\varepsilon^2\mathfrak{f}^2/\xi_j^2$ with $C_1 := (121/576)((5/6)^{m-1} - \delta_{m2}/384)$ and $\mu \leq \varepsilon^4\mathfrak{f}^4/(12\xi_j^4)$. Substituting these bounds into (8.4.17), we obtain

$$\tilde{J}_j \geq \frac{4\mathfrak{f}^2}{3\xi_j^2} \left(C_1 \|\hat{\varphi}|_{B_j}\|^2 + \frac{\varepsilon^4\mathfrak{f}^2}{16\xi_j^2} \int_{B_j^-} \Delta|\hat{\varphi}|^2(\boldsymbol{\xi}) \, d\boldsymbol{\xi} \right). \quad (8.4.20)$$

Now the assertion follows by inserting (8.4.17) into (8.4.10), using that $\mathfrak{f}/\xi_j^2 = 1/(2(j\pi - \nu))$ and $(\xi_j + b_{j-})^2 > 3\xi_j^2$ according to the estimate $b_{j-}^2 \geq 2\xi_j^2/3$. \square

8.4.4 Global stability results

With the bounds for different subdomains of the Fourier space obtained in lemmas 8.9 and 8.10, we are now in a position to prove a global stability result for inverse problem 8.2. For $\varepsilon > \nu$, an application of the estimates (8.4.5), (8.4.8) and (8.4.11) to the sub-integrals in (8.4.3) and lemma 8.7 with $\Omega = B[0; 1/2]$ yields

$$\begin{aligned} \|s_\nu \cdot \hat{\varphi}\|^2 &\geq C_0 \left(\zeta(\varepsilon, \nu) \|\varphi\|^2 + \mathfrak{f}\eta(\varepsilon, \nu) \int_{B^-} \Delta|\hat{\varphi}|^2 \, d\boldsymbol{\xi} \right) \\ &\geq C_0 (\zeta(\varepsilon, \nu) - \frac{\mathfrak{f}}{2}\eta(\varepsilon, \nu)) \|\varphi\|^2 \end{aligned} \quad (8.4.21)$$

Here, we have used that $\|\hat{\varphi}|_{D_\varepsilon}\|^2 + \sum_{j=0}^{\infty} \|\hat{\varphi}|_{B_j}\|^2 = \|\hat{\varphi}\|^2 = \|\varphi\|^2$ by the disjoint Fourier domain splitting in §8.4.1 and defined $B^- := \bigcup_{j=0}^{\infty} B_j^-$ as well as

$$\begin{aligned} \zeta(\varepsilon, \nu) &:= \min \left\{ \frac{m}{m+4}(\varepsilon - \nu)^2 + \frac{2m}{m+2}\nu(\varepsilon - \nu) + \nu^2, C_1\varepsilon^2 \right\}, \\ \eta(\varepsilon, \nu) &:= \max \left\{ \frac{2}{3(m+4)}(\varepsilon - \nu)^3 + \frac{1}{m+2}\nu(\varepsilon - \nu)^2, \frac{1}{32(\pi-\nu)}\varepsilon^4 \right\} \quad \text{for } \varepsilon > \nu. \end{aligned} \quad (8.4.22a)$$

Note that the bounding constant $\sin(\varepsilon)$ for the integral J_ε does not appear in $\zeta(\varepsilon, \nu)$ as $\sin(\varepsilon) \geq C_1\varepsilon^2$ holds true within the entire parameter range $0 \leq \varepsilon \leq \pi/6$.

In the case $\varepsilon \leq \nu$, the subdomains $B_0^- \subset B_0 = \emptyset$ are empty so that contributions from the sub-integral J_0 can be suppressed in the estimate (8.4.21). Hence, we may set

$$\zeta(\varepsilon, \nu) := C_1 \varepsilon^2, \quad \eta(\varepsilon, \nu) := \frac{\varepsilon^4}{32(\pi - \nu)} \quad \text{for } \varepsilon \leq \nu. \quad (8.4.22b)$$

Since $\eta(\varepsilon, \nu)$ is of higher-order in ε than $\zeta(\varepsilon, \nu)$ according to (8.4.22a) and (8.4.22b), it is possible to find an optimal value $\varepsilon = \varepsilon_{\text{opt}}(\nu, \mathbf{f})$ such that the constant on the right-hand side of (8.4.21) is maximal and in particular positive. This idea leads to the sought estimates for the stability constant $C_{\text{IP2}}(\Omega, \mathbf{f}, \nu)$ in theorem 8.2:

Proof of theorem 8.2. We first study the regime $\varepsilon \geq \nu$, as is necessary in particular if $\nu = 0$. Since

$$\begin{aligned} & \frac{m}{m+4}(\varepsilon - \nu)^2 + \frac{2m}{m+2}\nu(\varepsilon - \nu) + \nu^2 \\ &= \frac{m}{m+4}\varepsilon^2 + \left(\frac{2m}{m+2} - \frac{2m}{m+4}\right)\nu(\varepsilon - \nu) + \left(\frac{m}{m+4} - \frac{2m}{m+2} + 1\right)\nu^2 \geq \frac{m}{m+4}\varepsilon^2, \end{aligned}$$

(8.4.22a) yields the bound

$$\zeta(\varepsilon, \nu) \geq \zeta_0 \varepsilon^2, \quad \zeta_0 := \min\left\{C_1, \frac{m}{m+4}\right\}. \quad (8.4.23)$$

To bound η in (8.4.22a) we observe that $\nu(\varepsilon - \nu)^2 \leq 4\varepsilon^3/27$ for $0 \leq \nu < \varepsilon$ to obtain

$$\frac{2}{3(m+4)}(\varepsilon - \nu)^3 + \frac{1}{m+2}\nu(\varepsilon - \nu)^2 \leq \left(\frac{2}{3(m+4)} + \frac{4}{27(m+2)}\right)\varepsilon^3 \leq \left(\frac{2}{15} + \frac{4}{81}\right)\varepsilon^3 < \frac{1}{5}\varepsilon^3.$$

As $\frac{\varepsilon}{32(\pi - \nu)} < \frac{1}{5}$ for $0 \leq \nu \leq \varepsilon \leq \frac{\pi}{6}$ this implies by (8.4.22a)

$$\eta(\varepsilon, \nu) \leq \frac{1}{5}\varepsilon^3. \quad (8.4.24)$$

A comparison with (8.4.22b) shows that the estimates (8.4.23) and (8.4.24) remain valid for $\varepsilon < \nu$ and thus hold for all $0 \leq \varepsilon \leq \pi/6$.

Now we apply the derived bounds on $\zeta(\varepsilon, \nu)$ and $\eta(\varepsilon, \nu)$ in (8.4.21) to estimate the stability constant of inverse problem 8.2:

$$\begin{aligned} C_{\text{IP2}}(\Omega, \mathbf{f}, \nu)^2 &= \inf_{\varphi \in L^2(\Omega), \|\varphi\|=1} \|S\varphi\|^2 \stackrel{(8.4.1)}{=} \inf_{\varphi \in L^2(\Omega), \|\varphi\|=1} 4\|s_\nu \cdot \hat{\varphi}\|^2 \\ &\geq \sup_{0 \leq \varepsilon \leq \pi/6} 4C_0 \varepsilon^2 \left(\zeta_0 - \frac{1}{10}\mathbf{f}\varepsilon\right) = \begin{cases} \frac{1600C_0\zeta_0^3}{27\mathbf{f}^2} & \text{if } \mathbf{f} \geq \frac{40\zeta_0}{\pi} \\ \frac{\pi^2 C_0}{9} \left(\zeta_0 - \frac{\pi}{60}\mathbf{f}\right) & \text{if } \mathbf{f} < \frac{40\zeta_0}{\pi} \end{cases} \quad (8.4.25) \\ &\geq \min\{c_1^2, c_2^2 \mathbf{f}^{-2}\} \end{aligned}$$

The supremum is attained at $\varepsilon = \varepsilon_{\text{opt}} := \min\{\pi/6, 20\zeta_0/(3\mathbf{f})\}$, and we have defined $c_1 := (\pi^2 C_0 \zeta_0 / 27)^{1/2}$ and $c_2 := (1600 C_0 \zeta_0^3 / 27)^{1/2}$.

It remains to show improved bounds for $\nu > 0$ and large \mathbf{f} . Here we choose $0 \leq \varepsilon \leq \nu/3$ to satisfy $\varepsilon \leq \pi/6$ for all $\nu \in (0, \pi/2]$. Moreover, we make use of the simplified expressions for

ζ and η in (8.4.22b), that apply in this parameter range. Inserting (8.4.22b) into (8.4.21) and using that $\pi - \nu \geq \pi/2$ gives

$$\begin{aligned} 4\|s_\nu \cdot \hat{\varphi}\|^2 &\geq \sup_{0 < \varepsilon \leq \nu/3} 4C_0\varepsilon^2 \left(C_1 - \frac{1}{64(\pi-\nu)} \mathfrak{f}\varepsilon^2 \right) \geq \sup_{0 < \varepsilon \leq \nu/3} 4C_0\varepsilon^2 \left(C_1 - \frac{1}{32\pi} \mathfrak{f}\varepsilon^2 \right) \\ &= \begin{cases} \frac{32\pi C_0 C_1^2}{\mathfrak{f}} & \text{if } \mathfrak{f} \geq \frac{144\pi C_1}{\nu^2} \\ \frac{4C_0\nu^2}{9} \left(C_1 - \frac{\nu^2 \mathfrak{f}}{288\pi} \right) & \text{if } \mathfrak{f} < \frac{144\pi C_1}{\nu^2} \end{cases} \quad (8.4.26) \\ &\geq \min \{ c_3^2 \nu^2, c_4^2 \mathfrak{f}^{-1} \} \end{aligned}$$

for all $\varphi \in L^2_\Omega$ with $\|\varphi\| = 1$. Here, the constants are chosen as $c_3 := (2C_0C_1/9)^{1/2}$ and $c_4 := (32\pi C_0)^{1/2} C_1$ and the supremum is attained at $\varepsilon_{\text{opt}} = \min\{\nu/3, (16\pi C_1/\mathfrak{f})^{1/2}\}$. Combining (8.4.25) and (8.4.26) yields the claimed stability estimate (8.2.12). \square

8.4.5 Optimality

As opposed to the analysis of inverse problem 8.1, the bounds on the stability constant $C_{\text{IP}2}$ derived in this section can be expected to be highly non-optimal: the principal derivative-bound (8.4.2) is clearly not sharp and various additional estimates have been made in lemma 8.9, lemma 8.10 and in the proof of theorem 8.2 just to simplify the arising terms. Accordingly, the obtained numerical values for the constants c_j in (8.2.12) can be expected to be *overly pessimistic*. Sharper characterizations may be achieved by numerically computing the minimal singular value of the forward operator S by a similar approach as in §8.3.5.

Independently of such numerical refinements, it is of interest whether the achieved analytical bounds $C_{\text{IP}2}(\Omega, \mathfrak{f}, \nu) \gtrsim \mathfrak{f}^{-\nu}$ are at least of *optimal order* ν in the Fresnel number \mathfrak{f} for $\mathfrak{f} \rightarrow \infty$. A positive answer is given by the following theorem:

Theorem 8.11 (Order-optimality of the stability bounds for inverse problem 8.2). *Within the setting of theorem 8.2, let $c'_1, c'_2, \nu > 0$ be constants such that*

$$C_{\text{IP}2}(\Omega, \mathfrak{f}, 0) \geq \min \{ c'_1, c'_2 \mathfrak{f}^{-\nu} \} \quad \text{for all } \mathfrak{f} > 0. \quad (8.4.27)$$

Then $\nu \geq 1$, i.e. the bound (8.2.12) is of optimal asymptotic order in \mathfrak{f} for $\nu = 0$.

Proof. Assume $\nu < 1$. Let $\nu = 0$ and $\varphi \in L^2(\Omega) \cap \mathcal{C}^2(\mathbb{R}^m)$ with $\|\varphi\| = 1$. Then we have by the definition of the forward operator in (8.2.9)

$$\begin{aligned} \|\mathcal{S}_0\varphi\| &= 2\|s_\nu \cdot \mathcal{F}(\varphi)\| = 2 \left(\int_{\mathbb{R}^m} \sin\left(\frac{|\xi|^2}{2\mathfrak{f}}\right)^2 |\mathcal{F}(\varphi)(\xi)|^2 d\xi \right)^{1/2} \\ &\leq \frac{\sin(x)^2 \leq x^2}{\mathfrak{f}} \left(\int_{\mathbb{R}^m} |\xi|^2 |\mathcal{F}(\varphi)(\xi)|^2 d\xi \right)^{1/2} = \mathfrak{f}^{-1} \|\Delta\varphi\|. \end{aligned} \quad (8.4.28)$$

Note that $\|\Delta\varphi\| \neq 0$ by the maximum principle since Ω is bounded. Now let $\mathfrak{f} > \max\{(c'_1/c'_2)^{-1/\nu}, (2\pi c'_2/\|\Delta\varphi\|)^{1/(1-\nu)}\}$. Then (8.4.27) and (8.4.28) imply

$$C_{\text{IP}2}(\Omega, \mathfrak{f}, 0) \stackrel{c'_2 \mathfrak{f}^{-\nu} < c'_1}{=} c'_2 \mathfrak{f}^{-\nu} = \frac{2\pi c'_2 \mathfrak{f}^{1-\nu}}{\|\Delta\varphi\|} \mathfrak{f}^{-1} \|\Delta\varphi\| \stackrel{(8.4.28)}{\geq} \frac{2\pi c'_2 \mathfrak{f}^{1-\nu}}{\|\Delta\varphi\|} \|\mathcal{S}_0\varphi\| > \|\mathcal{S}_0\varphi\|.$$

Since $\|\varphi\| = 1$, this contradicts the definition of the stability constant C_{IP2} . Hence, $\nu \geq 1$ must hold by contradiction. \square

Intuitively, the upper bound $-\nu \leq -1$ for the achievable order in theorem 8.11 is related to the *second* order zero of the CTF s_ν at $\boldsymbol{\xi} = 0$ for $\nu = 0$, which allows to bound $|s_\nu|$ by a quadratic function. In the case $\nu > 0$, this zero no longer exists which enables higher order behavior $C_{\text{IP2}} \gtrsim \mathfrak{f}^{-1/2}$ as shown in theorem 8.2. As the CTF still has *first order* zero manifolds, we conjecture that this rate is likewise of optimal order. However, simple arguments as in the proof of theorem 8.11 suffer from technical difficulties arising from the spherical geometry and the \mathfrak{f} -dependence of the first order zero-manifolds. A treatment of the case $\nu > 0$ is therefore omitted here.

8.5 Image reconstruction from two measurements

By axial translation of the object in fig. 8.1(a), holograms I_1, I_2 may be recorded for different sample-detector distances and thus at different Fresnel numbers $\mathfrak{f}_1 \neq \mathfrak{f}_2$. It is often stated [42, 108, 32, 122] that the acquired additional data permits a more stable phase retrieval in this setting - in particular if phase shifts ϕ and attenuation μ are to be recovered as independent parameters. Within the weak object approximation (see §8.2.2), this setting amounts to reconstructing $h = \mu + i\phi$ from measurements $(\mathcal{T}^{(\mathfrak{f}_1)}(h), \mathcal{T}^{(\mathfrak{f}_2)}(h))$, where $\mathcal{T}^{(\mathfrak{f}_j)}$ denotes the linearized forward operator in (8.2.5) to the Fresnel number $\mathfrak{f} = \mathfrak{f}_j$. Adopting the CTF-formulation in (8.2.6), the two-hologram setting is modeled by the following forward map:

$$S^{(\mathfrak{f}_1, \mathfrak{f}_2)} : L^2(\mathbb{R}^m)^2 \rightarrow L^2(\mathbb{R}^m)^2; \begin{pmatrix} \phi \\ \mu \end{pmatrix} \mapsto -2\mathcal{F}^{-1}(\mathbf{S} \cdot \mathcal{F} \begin{pmatrix} \phi \\ \mu \end{pmatrix}), \quad \mathbf{S}(\boldsymbol{\xi}) := \begin{pmatrix} \sin\left(\frac{|\boldsymbol{\xi}|^2}{2\mathfrak{f}_1}\right) & \cos\left(\frac{|\boldsymbol{\xi}|^2}{2\mathfrak{f}_1}\right) \\ \sin\left(\frac{|\boldsymbol{\xi}|^2}{2\mathfrak{f}_2}\right) & \cos\left(\frac{|\boldsymbol{\xi}|^2}{2\mathfrak{f}_2}\right) \end{pmatrix}$$

$$S^{(\mathfrak{f}_1, \mathfrak{f}_2)} \begin{pmatrix} \phi \\ \mu \end{pmatrix} = \begin{pmatrix} \mathcal{T}^{(\mathfrak{f}_1)}(\mu + i\phi) \\ \mathcal{T}^{(\mathfrak{f}_2)}(\mu + i\phi) \end{pmatrix} \quad \text{for all real-valued } \mu, \phi \in L^2(\mathbb{R}^m). \quad (8.5.1)$$

Here, the matrix-vector-product is to be understood as a point-wise product of the vector- and matrix-valued function values and the (inverse) Fourier transforms $\mathcal{F}, \mathcal{F}^{-1}$ are meant component-wise. Furthermore, we denote by $\|f\| := (\|f_1\|^2 + \|f_2\|^2)^{1/2}$ for $f = (f_1, f_2) \in L^2(\mathbb{R}^m)^2$ the usual norm on $L^2(\mathbb{R}^m)^2$. With the forward model defined in (8.5.1), image recovery from two diffraction patterns may then be stated as:

Inverse Problem 8.3 (Phase contrast imaging from two measurements). For given support $\Omega \subset \mathbb{R}^m$, recover the images $\phi, \mu \in L^2(\Omega)$ from two noisy holograms

$$\begin{pmatrix} I_1^\epsilon - 1 \\ I_2^\epsilon - 1 \end{pmatrix} = S^{(\mathfrak{f}_1, \mathfrak{f}_2)} \begin{pmatrix} \phi \\ \mu \end{pmatrix} + \boldsymbol{\epsilon} \quad \text{with} \quad \|\boldsymbol{\epsilon}\| \leq \epsilon.$$

As the same amount of object information is to be recovered from a larger data set, it is evident that inverse problem 8.3 cannot be more ill-posed or ill-conditioned than inverse problem 8.1. In order to gain deeper insight, it is illustrative to write down an explicit inverse

of the forward operator $S^{(\mathbf{f}_1, \mathbf{f}_2)}$ defined in (8.5.1) by point-wise inversion of the contrast-transfer-matrix \mathbf{S} as done in [81]: setting $\mathbf{f}_- := (\mathbf{f}_1^{-1} - \mathbf{f}_2^{-1})^{-1}$ this yields for any $(\phi, \mu) \in L^2(\mathbb{R}^m)$

$$\begin{pmatrix} \phi \\ \mu \end{pmatrix} = \mathcal{F}^{-1} \left(\frac{1}{2 \sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_-} \right)} \begin{pmatrix} \cos \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_2} \right) & -\cos \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_1} \right) \\ -\sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_2} \right) & \sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_1} \right) \end{pmatrix} \mathcal{F} S^{(\mathbf{f}_1, \mathbf{f}_2)} \begin{pmatrix} \phi \\ \mu \end{pmatrix} \right). \quad (8.5.2)$$

All of the operations on the right-hand side of in (8.5.2) are bounded except for the point-wise division by the factor $2 \sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_-} \right)$. According to (8.2.9), we find that the latter operation exactly corresponds to solving inverse problem 8.2 for a pure phase object (i.e. $\nu = 0$) at an effective Fresnel number of $\mathbf{f} = \mathbf{f}_-$. In particular, this implies that the solution of inverse problem 8.3 is unique but ill-posed without a support constraint, i.e. for general $\phi, \mu \in L^2(\mathbb{R}^m)$. By a more rigorous analysis of the analogy to inverse problem 8.2, we obtain stability estimates:

Theorem 8.12 (Stability estimate for inverse problem 8.3). *Let $\mathcal{S}_0^{(\mathbf{f}^-)}$ denote the forward operator of inverse problem 8.2 for $\nu = 0$ and $\mathbf{f} = \mathbf{f}_- := (\mathbf{f}_1^{-1} - \mathbf{f}_2^{-1})^{-1}$. Then*

$$\|S^{(\mathbf{f}_1, \mathbf{f}_2)}(\phi, \mu)\| \geq 2^{-\frac{1}{2}} \|\mathcal{S}_0^{(\mathbf{f}^-)}(\phi + i\mu)\| \quad \text{for all real-valued } \phi, \mu \in L^2(\mathbb{R}^m) \quad (8.5.3)$$

In particular, for any support-domain $\Omega \subset \mathbb{R}^m$, we have the relative stability estimate

$$\|S^{(\mathbf{f}_1, \mathbf{f}_2)}(\phi, \mu)\| \geq 2^{-\frac{1}{2}} C_{\text{IP2}}(\Omega, \mathbf{f}_-, 0) \|\phi, \mu\| \quad \text{for all real-valued } \phi, \mu \in L^2(\Omega). \quad (8.5.4)$$

Proof. Let $\phi, \mu \in L^2(\mathbb{R}^m)$ be real-valued. Setting $\hat{\mathbf{h}} := (\mathcal{F}(\phi), \mathcal{F}(\mu))$ and exploiting that \mathcal{F} is unitary, we obtain by (8.5.1)

$$\|S^{(\mathbf{f}_1, \mathbf{f}_2)}(\phi, \mu)\|^2 = 4 \|\mathbf{S} \hat{\mathbf{h}}\|^2 = 4 \int_{\mathbb{R}^m} |\mathbf{S}(\boldsymbol{\xi}) \hat{\mathbf{h}}(\boldsymbol{\xi})|^2 d\boldsymbol{\xi}.$$

The integrand $|\mathbf{S}(\boldsymbol{\xi}) \hat{\mathbf{h}}(\boldsymbol{\xi})|^2$ is bounded from below by $\sigma_0^2(\boldsymbol{\xi}) |\hat{\mathbf{h}}(\boldsymbol{\xi})|^2$ with the smallest singular value $\sigma_0(\boldsymbol{\xi})$ of the 2-by-2 matrix $\mathbf{S}(\boldsymbol{\xi})$. Direct computation shows

$$\sigma_0^2(\boldsymbol{\xi}) = 1 - \left| \cos \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_-} \right) \right| \geq \frac{1}{2} \sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_-} \right)^2 \quad \text{for all } \boldsymbol{\xi} \in \mathbb{R}^m.$$

By inserting this result into the previous equation and comparing to (8.2.9), we obtain

$$\begin{aligned} \|S^{(\mathbf{f}_1, \mathbf{f}_2)}(\phi, \mu)\|^2 &\geq 2 \int_{\mathbb{R}^m} \sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_-} \right)^2 |\hat{\mathbf{h}}(\boldsymbol{\xi})|^2 d\boldsymbol{\xi} \\ &= 2 \int_{\mathbb{R}^m} \sin \left(\frac{|\boldsymbol{\xi}|^2}{2\mathbf{f}_-} \right)^2 |\mathcal{F}(\phi)(\boldsymbol{\xi}) + i\mathcal{F}(\mu)(\boldsymbol{\xi})|^2 d\boldsymbol{\xi} = \frac{1}{2} \|\mathcal{S}_0^{(\mathbf{f}^-)}(\phi + i\mu)\|^2. \end{aligned}$$

This proves (8.5.3). The second inequality (8.5.4) follows by bounding the right-hand side of (8.5.3) with the stability constant of inverse problem 8.2. \square

We emphasize that theorem 8.12 does not only relate the worst case stability of inverse problems 8.2 and 8.3 but (8.5.3) identifies the contrast attained by individual modes under the forward operators $\mathcal{S}_0^{(\mathbf{f}^-)}$ and $S^{(\mathbf{f}_1, \mathbf{f}_2)}$. Specifically, whenever $\|\mathcal{S}_0^{(\mathbf{f}^-)}(\phi + i\mu)\|$ is large, the

mode $(\phi, \mu)^T$ also attains high contrast under the forward map $S^{(f_1, f_2)}$ of inverse problem 8.3. It should furthermore be noted that the derived stability estimates are in terms of the *difference Fresnel number* $f_-^{-1} = f_1^{-1} - f_2^{-1}$. Hence, if the two holograms are recorded at similar Fresnel numbers $f_1 \approx f_2$, i.e. for only slightly varied setup parameters, f_- is very large so that the stability bounds in theorem 8.12 will hardly be better than for the *single* measurement case. This is quite intuitive as the second measurement provides only little additional information in such a setting.

8.6 Discussion and conclusions

In this paper we have studied the stability of phase retrieval in propagation-based phase contrast imaging within the linear contrast-transfer-function model (CTF) [79, 194], valid for weakly interacting objects. While the image reconstruction problem is generally ill-posed and even severely non-unique if both phase-shifts ϕ and attenuation μ are to be recovered from a single (phaseless) intensity measurement (inverse problem 8.1), we could show *well-posedness* under a support constraint: if the image $h = -\mu - i\phi$ is known to be supported in a bounded domain $\Omega \subset \mathbb{R}^m$, then the attained contrast $\|\mathcal{T}(h)\| \geq C_{\text{IP1}}(\Omega, \mathbf{f})\|h\|$ is bounded from below by a positive stability constant $C_{\text{IP1}}(\Omega, \mathbf{f}) > 0$ according to theorem 8.1.

Numerical simulations (see §8.3.5) suggest that the derived bound $C_{\text{IP1}}(\Omega, \mathbf{f}) \gtrsim f^{1/4} \exp(-f/8)$ is quite sharp. Accordingly, C_{IP1} unfortunately decays nearly exponentially with the Fresnel number \mathbf{f} of the imaging setup (computed w.r.t. the support diameter!), which translates into partly ridiculously small values within the typical range $10 \leq \mathbf{f} \leq 10^4$ in X-ray phase contrast experiments. For instance, the estimate (8.2.11) gives $C_{\text{IP1}}(\Omega, 2\pi \cdot 100) \leq 10^{-33}$ at a Fresnel number $\mathbf{f} = 100$. In other words: although well-posed, the inversion of the forward operator \mathcal{T} is typically *heavily ill-conditioned*, admitting amplification of data errors by bounded, yet *huge* factors C_{IP1}^{-1} in the image reconstruction. It is thus not surprising that the independent recovery of both ϕ and μ from a single hologram is typically severely corrupted by artifacts in practice, see e.g. [175, 139], and commonly considered as not feasible [108, 157, 32].

On the other hand, such a reconstruction has been successfully demonstrated in [143] up to slight low-frequency artifacts. In view of the present work, the key ingredient to this demonstration can be identified as the comparably small Fresnel number $\mathbf{f} \approx 14$ of the support in the considered setup. By the characterization of the least stable modes in theorem 8.5, this corresponds to a minimum contrast

$$\|\mathcal{T}(\phi_{(0,0)})\| \geq 10^{-4}\|\phi_{(0,0)}\| \quad \text{and} \quad \|\mathcal{T}(\phi_{(0,1)})\| \geq 7 \cdot 10^{-4}\|\phi_{(0,1)}\| \quad (8.6.1)$$

for the least and second least stable image modes $\phi_{(0,0)}$ and $\phi_{(0,1)}$. The reconstruction of these low-frequency image components is thus feasible in principle, yet numerically cumbersome, which explains the residual artifacts in the reconstruction. Higher-order image modes attain contrast in the order of the data noise level or above and may thus be recovered with reasonable accuracy. Our stability estimates thus predict that the joint recovery of phase shifts ϕ and absorption μ is feasible if and only if $\mathbf{f} \lesssim 10$, i.e. in the *deeply holographic regime* - in excellent agreement with numerical reconstructions.

According to the improved stability estimate for the settings of inverse problems 8.2 and 8.3 (see theorems 8.2 and 8.12) image recovery can be performed without additional regularization

provided a sufficiently strong support constraint. This is consistent with the observed high performance of iterative projection algorithms in phase retrieval (see [17, 138, 66] and references therein, mostly for the far-field case and e.g. [76, 14] for near-field phase contrast), which make use of support constraints.

Let us discuss some possible extensions of our results. First of all we have assumed that diffraction patterns can be measured in a whole detector plane, whereas real-world X-ray detectors necessarily cover only a finite area. If intensities are measured only in a bounded subdomain, the singular values of the forward operator \mathcal{T} will eventually decay exponentially since \mathcal{T} can be written as an integral operator with an analytic (though highly oscillatory) kernel. On the other hand, Fresnel propagation is mathematically equivalent to time-evolution within the free Schrödinger equation (if the z -coordinate in fig. 8.1(b) is identified with time). Localization properties of the latter model suggest that the finiteness of the detection domain has little influence if it is chosen sufficiently large – in agreement with experimental and numerical experience. A better understanding and analysis of the impact of a finite field of view will be an interesting goal for future research. Further possible extensions include a treatment of the nonlinear imaging model (8.2.4), non-plane wave illumination and partial coherence. Moreover, in *region-of-interest* imaging of extended objects, support constraints do not hold true so that stability needs to be established by other means.

Finally, we emphasize that – although seemingly specific to the considered imaging setup – our analysis treats a fairly general physical problem: the reconstruction of a (compactly supported) perturbation to a plane wave from intensities of the propagated field under the paraxial Helmholtz equation or – equivalently – within time-dependent Schrödinger dynamics. The results may thus be relevant for several related wave-optical and quantum-mechanical inverse problems.

Acknowledgements

We thank Tim Salditt, Aike Ruhlandt and Martin Krenkel from the Institute for X-ray Physics at the University of Göttingen for enlightening discussions on phase contrast setups, a priori constraints and stability of image reconstruction, as well as for providing the experimental data used in figs. 8.1 and 8.2 for visualization of the theory in this work.

Chapter 9

Complete Article 2: Locality estimates for Fresnel-wave-propagation and stability of X-ray phase contrast imaging with finite detectors

Reproduced from the published article [144]:*

Maretzke S. *Inverse Problems*, **34**(12):124004, 2018. DOI: [10.1088/1361-6420/aae78f](https://doi.org/10.1088/1361-6420/aae78f)

(subject to slight adjustments of the notation and minor corrections)

Abstract: Coherent wave-propagation in the near-field Fresnel-regime is the underlying contrast-mechanism to (propagation-based) X-ray phase contrast imaging (XPCI), an emerging lensless technique that enables 2D- and 3D-imaging of biological soft tissues and other light-element samples down to nanometer-resolutions. Mathematically, propagation is described by the Fresnel-propagator, a convolution with an arbitrarily non-local kernel. As real-world detectors may only capture a finite field-of-view, this non-locality implies that the recorded diffraction-patterns are necessarily incomplete. This raises the question of stability of image reconstruction from the truncated data – even if the complex-valued wave-field, and not just its modulus, could be measured. Contrary to the latter restriction of the acquisition, known as the phase-problem, the finite-detector-problem has not received much attention in literature. The present work therefore analyzes locality of Fresnel-propagation in order to establish stability of XPCI with finite detectors. Image reconstruction is shown to be severely ill-posed in this setting – even without a phase-problem. However, quantitative estimates of the leaked wave-field reveal that Lipschitz-stability holds down to a sharp resolution limit that depends on the detector-size and varies within the field-of-view. The smallest resolvable lengthscale is found to be $\approx 1/\mathfrak{f}$ times the detector’s aspect length, where \mathfrak{f} is the Fresnel number associated with the latter scale. The stability results are extended to phaseless imaging in the linear contrast-transfer-function regime.

*Copyright © 2018 IOP Publishing Ltd

9.1 Introduction

State-of-the-art high-resolution imaging techniques are a driving force behind current biomedical- and material science. Among such, (propagation-based) X-ray phase contrast imaging (XPCI), also known as near-field holography, stands out as it yields two- or three-dimensional images down to nanometer-resolutions with high penetration-depths at relatively low radiation-dose and sample-preparation requirements [207, 166, 162, 42, 13, 143, 121, 85].

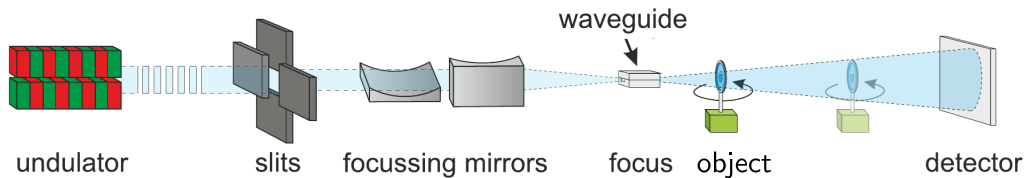


Figure 9.1. Exemplary experimental setup for (propagation-based) X-ray phase contrast imaging (XPCI) at synchrotrons (sketch of the GINIX-experiment [178] at P10-beamline, DESY).

The setup of XPCI is appealingly simple, see the example sketched in fig. 9.1: essentially, it boils down to a coherent X-ray beam illuminating an unknown object and a detector that records the resulting near-field diffraction pattern, also termed *hologram*, at a finite distance behind the sample. The coherent wave-propagation from the sample to the detector, described by the Helmholtz equation in the *paraxial-* or *Fresnel-*approximation [108, 160], is essential as it enables *phase contrast*: it partially encodes phase-shifts in the complex-valued X-ray wave-field Ψ induced by refraction within the sample into measurable wave intensities $\propto |\Psi|^2$, thereby circumventing the well-known *phase problem*, i.e. the inability to measure the phase of Ψ directly. This permits imaging of biological soft tissues and other light-element samples, for which the absorption of X-rays – but not refraction – is negligible [166].

To obtain an interpretable image of the sample, the induced phase-shifts (and absorption) have to be reconstructed from the measured hologram(s), i.e. an inverse problem has to be solved. By the limitation of the data to the squared modulus $|\Psi|^2$, this requires to recover the missing phase-information. For the present setting, however, this task is comparably well-understood by now and routinely solved using data from multiple sample-detector-distances along with a linearization of the contrast known as the contrast-transfer-function (CTF) model [42, 194, 122, 121] and/or additional a priori knowledge on the recovered images [14, 13, 165, 143]. Indeed, it is shown in previous work that the mild assumption of a known compact support of the image ensures well-posedness of the reconstruction in the linear CTF-regime [144].

What is typically tacitly ignored, however, is the data-incompleteness arising from the finiteness of the field-of-view captured by the detector due to the de-localizing action of (Fresnel-)wave-propagation: existing theory mostly assumes data within the complete infinite detector-plane and most reconstruction methods implicitly assume periodic detector-boundaries, possibly combined with artificial extension of the data by padding. While this produces reasonable results in practice, theoretical understanding for the effects of a finite detector and of the associated heuristic corrections is lacking.

This work aims to close this gap of theory by deriving rigorous estimates on the locality of information-transport by wave-propagation in the Fresnel-regime with the ultimate goal of

extending existing stability estimates for XPCI to settings with finite detectors. In particular, the focus is on the question of *resolution*:

Given an XPCI setup, what is the size of the smallest sample-features that can be stably reconstructed from the measured data?

In physics literature [158, 127], a finite detector in XPCI is typically argued to limit the numerical aperture of the imaging setup, which translates to a quantitative bound on the achievable resolution according to Abbe's diffraction limit, see e.g. [26, 131]. However, the underlying reasoning is strictly valid only for *far-field* optics, whereas XPCI is a *near-field* imaging technique. Rigorous theory is thus necessary to supplement physical intuition.

The manuscript is organized as follows: §9.2 introduces the mathematical setting and notation as well as some preliminary insights on the finite-detector problem. In §9.3, the relation between resolution and detector-size is assessed by the study of Gaussian wave-packets, yielding *best-case* estimates in some sense. These are then complemented by *worst-case* estimates on stability of image reconstruction derived in §9.4, §9.5 and §9.6 under different a priori assumptions on the unknown objects. Having derived all of these results under the simplifying assumption that also the phase of the data is measured, the obtained locality- and stability estimates are then extended to the phaseless case of linearized XPCI in §9.7. §9.8 concludes this work.

Despite the focus on XPCI, note that the derived estimates may be extended to a wide range of wave-propagation problems from classical physics and quantum-mechanics.

9.2 Background

9.2.1 Basic setting

9.2.1.1 Fresnel-propagation

We consider the problem of reconstructing a function $h : \mathbb{R}^m \rightarrow \mathbb{C}$ from partial knowledge of Fresnel-data:

$$\mathcal{D}(h) := \mathcal{F}(m_f \cdot \mathcal{F}^{-1}(h)) \quad \text{with} \quad m_f(\boldsymbol{\xi}) := \exp(-i\boldsymbol{\xi}^2/(2f)), \quad f > 0, \boldsymbol{\xi} \in \mathbb{R}^m. \quad (9.2.1)$$

$\mathcal{F}(f)(\boldsymbol{\xi}) := (2\pi)^{-m/2} \int_{\mathbb{R}^m} \exp(-i\boldsymbol{\xi} \cdot \mathbf{x}) f(\mathbf{x}) d\mathbf{x}$ denotes the m -dimensional Fourier transform. The *Fresnel-propagator* \mathcal{D} models the free-space propagation of time-harmonic wave-fields $\Psi(\mathbf{x}, z) = \tilde{\Psi}(\mathbf{x}, z) \exp(iz)$ with slowly varying envelope $\tilde{\Psi}$ within the regime of the paraxial Helmholtz-equation (see e.g. [160] for details):

$$\begin{cases} (2if\partial_z + \Delta_{\mathbf{x}}) \tilde{\Psi}(\mathbf{x}, z) = 0 & (\mathbf{x}, z) \in \mathbb{R}^m \times (0; 1) \\ \tilde{\Psi}(\mathbf{x}, 0) = h(\mathbf{x}) & \mathbf{x} \in \mathbb{R}^m \end{cases} \Rightarrow \tilde{\Psi}(\cdot, 1) = \mathcal{D}(\tilde{\Psi}(\cdot, 0)) \quad (9.2.2)$$

9.2.1.2 Forward models

As detailed in [166, 108, 160, 144], Fresnel-diffraction data arises in X-ray phase contrast imaging (XPCI): if the incident beam in fig. 9.1 is modeled by a plane wave as sketched in

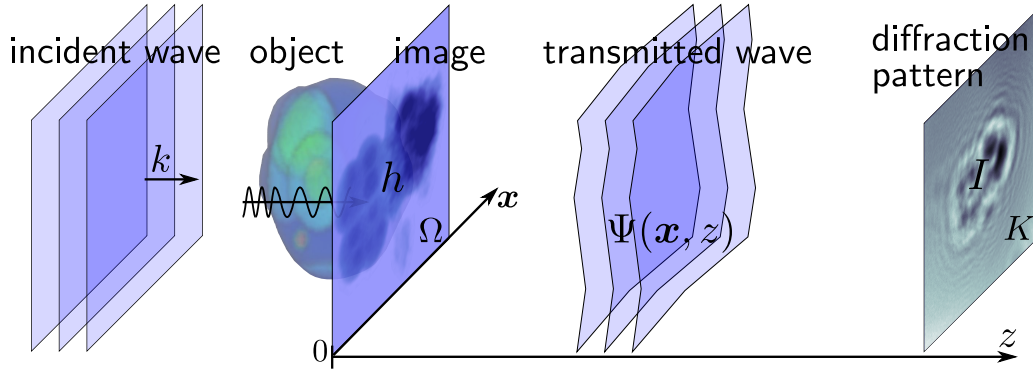


Figure 9.2. Schematic model of propagation-based XPCI: incident plane waves are scattered by a sample, imprinting phase-shifts and absorption $h = -i\phi - \mu$ upon the transmitted wave-field $\Psi(\cdot, 0)$ within the *object-domain* Ω . The intensity of the resulting near-field diffraction pattern $I = |\Psi(\cdot, d)|^2$ is recorded within the *detection-domain* K at some distance behind the object.

fig. 9.2, the wave-field in the object’s exit-plane $z = 0$ is

$$\tilde{\Psi}(\cdot, 0) = \exp(-h) \quad \text{with} \quad h = i\phi + \mu = ik \int_{\mathbb{R}} (\delta - i\beta) dz, \quad (9.2.3)$$

(within some standard approximations of X-ray optics [108]), where $n(\mathbf{x}, z) = 1 - \delta(\mathbf{x}, z) + i\beta(\mathbf{x}, z)$ is the spatially varying refractive index of the sample. The complex-valued *image* h is thus a projection of the sample-characterizing quantities δ, β . As the wave-field in the detector-plane relates to $\tilde{\Psi}(\cdot, 0)$ via Fresnel-propagation, the detected intensities are given by

$$I(\mathbf{x}) = |\mathcal{D}(\tilde{\Psi}(\cdot, 0))(\mathbf{x})|^2 = |\mathcal{D}(\exp(-h))(\mathbf{x})|^2. \quad (9.2.4)$$

Under the additional assumption that the object is sufficiently weakly scattering for the image to be “small” in a suitable sense, (9.2.4) may be linearized:

$$I = 1 + \mathcal{I}(h) + \mathcal{O}(h^2) \quad \text{with} \quad \mathcal{I}(h) = -2\text{Re}(\mathcal{D}(h)), \quad (9.2.5)$$

where Re is the pointwise real part. In Fourier-space, the contrast in the phase-shifts ϕ and attenuation μ is then described by oscillatory *contrast-transfer-functions* (CTF):

$$\mathcal{F}(\mathcal{I}(i\phi + \mu))(\boldsymbol{\xi}) = -2 \sin(|\boldsymbol{\xi}|^2/(2f)) \mathcal{F}(\phi)(\boldsymbol{\xi}) - 2 \cos(|\boldsymbol{\xi}|^2/(2f)) \mathcal{F}(\mu)(\boldsymbol{\xi}) \quad (9.2.6)$$

for all $\boldsymbol{\xi} \in \mathbb{R}^m$. Therefore, the linearized XPCI-model is also termed CTF-model.

Furthermore, it is often assumed [161, 194] that the object is *homogeneous*, in the sense that refraction and absorption are proportional: $h = \mu + i\phi = ie^{-i\nu}\varphi$ for some $\nu \in [0; \pi)$ and a *real-valued* function φ . In the linearized case, this yields a modified CTF-model:

$$\mathcal{S}_\nu(\varphi) := \mathcal{I}(h) = -2\mathcal{F}^{-1}(s_\nu \cdot \mathcal{F}(\varphi)), \quad s_\nu(\boldsymbol{\xi}) := \sin(|\boldsymbol{\xi}|^2/(2f) + \nu) \quad (9.2.7)$$

The case $\nu = 0$ corresponds to *pure phase objects* which induce negligible absorption $\mu \approx 0$.

9.2.1.3 Object- and detection-domains

We assume that the approximate size of the imaged object is known a priori. Then there exists a bounded *object-domain* $\Omega \subset \mathbb{R}^m$ such that the unknown image h satisfies $\text{supp}(h) = \overline{\{\mathbf{x} \in \mathbb{R}^m : h(\mathbf{x}) \neq 0\}} \subset \Omega$, where the overbar denotes set-closure. We consider the L^2 -functions satisfying this *support constraint*:

$$h \in L^2(\Omega) := \{h : \mathbb{R}^m \rightarrow \mathbb{C} : \text{supp}(h) \subset \Omega, \int_{\mathbb{R}^m} |h|^2 d\mathbf{x} < \infty\}. \quad (9.2.8)$$

Throughout this work, $\langle f, g \rangle := \int_{\mathbb{R}^m} f(\mathbf{x})g(\mathbf{x})^* d\mathbf{x}$ and $\|h\| := \langle h, h \rangle^{1/2}$ refer to the inner-product and norm in the space of square-integrable functions $L^2(\mathbb{R}^m)$.

Contrary to most previous work, we account for the fact that real-world detectors may only record data within a bounded *detection-domain* $K \subsetneq \mathbb{R}^m$, also referred to as *field-of-view* (*FoV*) or simply *detector*. Thus, only *restrictions* $I|_K$ of the intensity-data in (9.2.3), defined by $I|_K(\mathbf{x}) = I(\mathbf{x})$ for $\mathbf{x} \in K$ and $I|_K(\mathbf{x}) = 0$ otherwise, are available. By considering continuous measurements, however, we neglect that detectors are composed of *discrete* pixels.

For the XPCI-setting, a two-dimensional square detector $K = [-\frac{1}{2}; \frac{1}{2}]^2$ is certainly of highest practical relevance. By the analysis in [119, 176], however, also the case $m = 3$ is of interest as it arises in a linearized model of *tomographic* imaging. Moreover, $m = 3$ is also the natural dimension for an alternate application from quantum-mechanics:

Remark 9.1 (Application in quantum mechanics). The paraxial Helmholtz equation in (9.2.2) is equivalent to the time-dependent Schrödinger-equation for a free electron if z is identified with the time-dimension. Accordingly, all results of this work can be interpreted in view of the question how much probability-mass of a quantum-mechanical wave-function, initially localized in $\Omega \subset \mathbb{R}^m$, leaks out of some domain $K \subset \mathbb{R}^m$ upon time-propagation.

Therefore, the analysis is carried out independently of the dimension m as far as possible.

9.2.1.4 Fresnel number(s)

The dimensionless parameter \mathfrak{f} in (9.2.1) is the (modified) Fresnel number of the imaging setup (related to the classically defined Fresnel number \mathfrak{f} by a convenient 2π -factor: $\mathfrak{f} = 2\pi\mathfrak{f}$). It is defined as $\mathfrak{f} = kb^2/d$, where k is the wavenumber of the incident plane-wave in fig. 9.2, d is the distance between object- and detector-plane and b is the physical length that corresponds to unity in the dimensionless coordinates \mathbf{x} . The value of \mathfrak{f} determines how strongly structures of lengthscale 1 in an object h are distorted upon Fresnel-propagation $h \mapsto \mathcal{D}(h)$: for $\mathfrak{f} \gg 1$, structures are essentially preserved whereas $\mathfrak{f} \ll 1$ corresponds to full far-field diffraction.

The FoV will typically be taken as the unit-square, $K = [-\frac{1}{2}; \frac{1}{2}]^m$. Thereby, the Fresnel number \mathfrak{f} is implicitly defined with b as the detector's physical aspect length. Typical values are then in the range $10^3 \lesssim \mathfrak{f} \lesssim 10^5$ for high-resolution XPCI-experiments at synchrotrons. By the freedom in choosing b , however, one can also associate a Fresnel number with any other lateral scale: if σ is a dimensionless length, $\mathfrak{f}_\sigma := \sigma^2\mathfrak{f}$ is the Fresnel number that describes diffraction on the physical scale corresponding to σ .

9.2.1.5 Inverse problems

In order to study XPCI with a finite FoV, we consider image reconstruction problems both with complex- and phaseless Fresnel-data (ϵ : data-errors):

Inverse Problem 9.1 (Reconstruction of complex-valued images). For $\Omega, K \subset \mathbb{R}^m$, reconstruct a *complex-valued* $h \in L^2(\Omega)$ from either of the following data:

- (a) $g_{(a)}^{\text{obs}} = \mathcal{D}(h)|_K + \epsilon$
- (b) $g_{(b)}^{\text{obs}} = \mathcal{F}(h)|_K + \epsilon$
- (c) $g_{(c)}^{\text{obs}} = |\mathcal{D}(\exp(h))|_K^2 + \epsilon$

Inverse Problem 9.2 (Reconstruction of real-valued (homogeneous) images). For $\Omega, K \subset \mathbb{R}^m$ and $\nu \in [0; \pi)$, reconstruct a *real-valued* $\varphi \in L^2(\Omega, \mathbb{R})$ from either of the following data:

- (a) $g_{(a)}^{\text{obs}} = \mathcal{D}(\varphi)|_K + \epsilon$
- (b) $g_{(b)}^{\text{obs}} = \mathcal{S}_\nu(\varphi)|_K + \epsilon$
- (c) $g_{(c)}^{\text{obs}} = |\mathcal{D}(\exp(-ie^{-i\nu}\varphi))|_K^2 + \epsilon$

It should be emphasized that reconstructions in the setting of inverse problem 9.2(b),(c), are currently standard in XPCI, whereas solving inverse problem 9.1(b),(c) is typically considered too unstable due to the larger number of unknowns to be recovered. Yet, it is not at all obvious that inverse problem 9.1 and inverse problem 9.2 also exhibit different effects due to a finite FoV, i.e. that real-valuedness¹ is relevant for the present study. Surprisingly, however, this indeed turns out to be the case.

To identify the effects of a finite FoV, we will mostly consider the non-phaseless problems inverse problem 9.1(a) and inverse problem 9.2(a). By the richness of measured data, however, the problems (b) and (c) are clearly harder to solve than the variants (a) and inverse problem 9.2(a) is easier to solve than any of the others. In particular, this “hierarchy-of-difficulties” means that any instabilities in inverse problem 9.1(a) and inverse problem 9.2(a) will necessarily also be present in the phaseless problems.

9.2.2 Properties of the Fresnel propagator

As a preparation for the subsequent analysis, we summarize some basic properties of the Fresnel propagator, see also [160, 130, 100]:

¹Although we will refer to “real-valued” signals throughout the work, note that all the results obtained for such trivially extend to signals given by real-functions multiplied by a global complex phase.

(P1) *Unitary operator:* The map $\mathcal{D} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ defines a linear *isometry*:

$$\|\mathcal{D}(f)\| = \|f\| \quad \text{for all } f \in L^2(\mathbb{R}^m). \quad (\text{P1})$$

This also implies that $\mathcal{T}, \mathcal{S}_\nu : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ are bounded with $\|\mathcal{T}\|, \|\mathcal{S}_\nu\| \leq 2$.

(P2) *Convolution form:* As a Fourier-multiplier, \mathcal{D} can be alternatively written as a convolution: for all $f \in L^1(\mathbb{R}^m) \cap L^2(\mathbb{R}^m)$, it holds that

$$\begin{aligned} \mathcal{D}(f)(\mathbf{x}) &= (n_f * f)(\mathbf{x}) = \int_{\mathbb{R}^m} k_f(\mathbf{x} - \mathbf{y}) f(\mathbf{y}) d\mathbf{y} \quad \text{for all } \mathbf{x} \in \mathbb{R}^m \\ k_f &= u_0 (f/(2\pi))^{\frac{m}{2}} \cdot n_f, \quad n_f(\mathbf{x}) = \exp(i\mathbf{f}\mathbf{x}^2/2), \quad u_0 = \exp(-im\pi/4) \end{aligned} \quad (\text{P2})$$

(P3) *Alternate form:* By rearranging the convolution-formulation (P2), the following alternate form of the Fresnel propagator can be obtained:

$$\mathcal{D}(f)(\mathbf{x}) = u_0 \mathbf{f}^{\frac{m}{2}} n_f(\mathbf{x}) \cdot \mathcal{F}(n_f \cdot f)(\mathbf{f}\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^m \quad (\text{P3})$$

(P4) *Separability:* m_f factorizes into a product of functions of a single coordinate:

$$\begin{aligned} m_f(\boldsymbol{\xi}) &= \exp\left(-\frac{i\boldsymbol{\xi}^2}{2\mathbf{f}}\right) = \exp\left(-\frac{i}{2\mathbf{f}} \sum_{j=1}^m \xi_j^2\right) = \prod_{j=1}^m \exp\left(-\frac{i\xi_j^2}{2\mathbf{f}}\right) = \prod_{j=1}^m m_{f,j}(\boldsymbol{\xi}), \\ m_{f,j}(\boldsymbol{\xi}) &:= m_f(\xi_j) \quad \text{for all } \boldsymbol{\xi} = (\xi_1, \dots, \xi_m) \in \mathbb{R}^m. \end{aligned}$$

Consequently, \mathcal{D} factorizes into a commuting product of quasi-1D Fresnel-propagators acting along the different dimensions:

$$\begin{aligned} \mathcal{D}(f) &= \mathcal{D}_1 \dots \mathcal{D}_m(f) = \mathcal{D}_m \dots \mathcal{D}_1(f) \\ \mathcal{D}_j(f) &:= \mathcal{F}^{-1}(m_{f,j} \cdot \mathcal{F}(f)) = \mathcal{F}_j^{-1}(m_{f,j} \cdot \mathcal{F}_j(f)). \end{aligned} \quad (\text{P4})$$

$\mathcal{F}_j : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$ is the 1D-Fourier transform along the j th dimension.

(P5) *Isotropy and translation invariance:* As a convolution operator, \mathcal{D} is translation invariant, i.e. commutes with coordinate-shifts. As m_f is invariant under orthogonal transformations, i.e. $m_f(A\boldsymbol{\xi}) = m_f(\boldsymbol{\xi})$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$, $A \in O(m)$, \mathcal{D} also commutes with orthogonal coordinate transforms, i.e. acts *isotropically* along all dimensions:

$$\begin{aligned} \mathcal{D}\mathcal{A} &= \mathcal{A}\mathcal{D} \quad \text{for all } \mathcal{A} : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m); \mathcal{A}(f)(\mathbf{x}) = f(A\mathbf{x} + \mathbf{a}) \\ & \quad A \in O(m), \mathbf{a} \in \mathbb{R}^m. \end{aligned} \quad (\text{P5})$$

(P6) *Extension to distributions:* \mathcal{D} can be extended to tempered distributions $\mathcal{S}'(\mathbb{R}^m)$, i.e. to the dual space of smooth and rapidly decaying Schwartz-functions $\mathcal{S}(\mathbb{R}^m)$:

$$(\mathcal{D}(T))(u) := T(\mathcal{D}(u)) \quad \text{for all } T \in \mathcal{S}'(\mathbb{R}^m), u \in \mathcal{S}(\mathbb{R}^m). \quad (\text{P6})$$

In particular, one has $\mathcal{D}(1) = 1$ for the constant 1-function. Moreover, by continuity of $\mathcal{D}, \mathcal{F} : \mathcal{S}'(\mathbb{R}^m) \rightarrow \mathcal{S}'(\mathbb{R}^m)$, (P3) remains valid in a distributional sense.

9.2.3 Preliminary results

We aim to characterize the ill-posedness of inverse problems inverse problem 9.1 and inverse problem 9.2. Let us first note that Fresnel-propagation is, in principle, arbitrarily non-local:

Theorem 9.2 (Arbitrary non-locality of Fresnel-propagation). *Let $0 \neq h \in L^2(\mathbb{R}^m)$ have compact support. Then $\mathcal{D}(h)$ is supported within the whole \mathbb{R}^m , $\text{supp}(\mathcal{D}(h)) = \mathbb{R}^m$.*

Proof. By (P3) and the Paley-Wiener-Schwartz-theorem, $\mathcal{D}(h)$ is an entire analytic function. Thus, $\mathcal{D}(h)$ is non-zero almost everywhere in \mathbb{R}^m . \square

Accordingly, measuring diffraction-data only within a finite FoV will always result in some information-*leakage*. One might think that this ultimately introduces *non-uniqueness* of the reconstruction. This is however not the case, as has been shown in previous work:

Theorem 9.3 (Uniqueness [140]). *Let $\Omega \subset \mathbb{R}^m$ bounded and let $K \subset \mathbb{R}^m$ contain an open set. Then inverse problem 9.1 and inverse problem 9.2 are uniquely solvable (up to periodicity of the exponential in (c)).*

Theorem 9.3 means that the question, whether a small detection-domain K raises issues, admits no simple yes-no-answer. Indeed, it implies that the effects of the size of K can only be understood by studying *stability*. We recall that – for infinite detectors – the *linear* inverse problems inverse problem 9.1(a),(b) and inverse problem 9.2(a),(b) are Lipschitz-stable, i.e. *well-posed*:

Theorem 9.4 (Well-posedness for infinite detectors and compact supports [144]). *Let $\Omega \subset \mathbb{R}^m$ be bounded and let $K = \mathbb{R}^m$. Then inverse problem 9.1(a),(b) and inverse problem 9.2(a),(b) are well-posed, i.e. if $T : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$ denotes the corresponding forward operator, then there exist constants $C_{\text{stab}}^{\text{IP}^*} > 0$, depending on \mathfrak{f}, m, Ω (and ν), such that*

$$\|T(h)\| \geq C_{\text{stab}}^{\text{IP}^*} \|h\| \quad \text{for all } h \in L^2(\Omega). \quad (9.2.9)$$

Proof. For inverse problem 9.1(a) and inverse problem 9.2(a), the result is due to the unitarity of the Fresnel propagator (P1) and one has $C_{\text{stab}}^{\text{IP}1(\text{a})} = C_{\text{stab}}^{\text{IP}2(\text{a})} = 1$. For inverse problem 9.1(b) and inverse problem 9.2(b), the general statement along with estimates of the constants $C_{\text{stab}}^{\text{IP}^*}$ is proven in [144]. \square

The point of Lipschitz-stability estimates of the form (9.2.9) is that they are necessary and sufficient for the operator T to have a bounded (pseudo-)inverse T^\dagger and thereby ensure that data-errors ϵ induce only *bounded* deviations $\leq (C_{\text{stab}}^{\text{IP}^*})^{-1} \|\epsilon\|$ in the reconstructions. Clearly, one would like to have similar results for finite detectors $K \subsetneq \mathbb{R}^m$. However, the following theorem shows that stability may deteriorate dramatically due to a finite FoV:

Theorem 9.5 (Severe ill-posedness for bounded detectors). *Let $\Omega, K \subset \mathbb{R}^m$ be bounded with non-empty interior. Then inverse problem 9.1 and inverse problem 9.2 are severely ill-posed.*

Proof. By the hierarchy-of-difficulty discussed in §9.2.1, it is sufficient to prove the claim for inverse problem 9.2(a). Accordingly, we have to consider the singular values of the forward operator $T : L^2(\Omega, \mathbb{R}) \rightarrow L^2(K)$; $h \mapsto \mathcal{D}(h)|_K$. Thus, we compute T^*T . Using the convolution-form (P2), it can be shown that, for arbitrary $h \in L^2(\Omega, \mathbb{R})$,

$$T^*T(\mathbf{x}) = \int_{\Omega} \left(\underbrace{\int_K \operatorname{Re} \left(k_{\mathfrak{f}}(\mathbf{x} - \mathbf{y}) \cdot \overline{k_{\mathfrak{f}}(\mathbf{y} - \mathbf{y}')} \right) d\mathbf{y}}_{k(\mathbf{x}, \mathbf{y}')} \right) h(\mathbf{y}') d\mathbf{y}' \quad \text{for all } \mathbf{x} \in \Omega.$$

Accordingly, T^*T is given by an integral-operator with kernel k . Since $k_{\mathfrak{f}}$ is bounded and infinitely smooth, so is k and $k \in L^2(\Omega \times \Omega)$ by boundedness of Ω . In total, this implies that T^*T is an infinitely smoothing compact integral-operator so that its eigenvalues, the squared singular values of T , decay super-algebraically. This shows that inverse problem 9.2(a) and hence all considered inverse problems are severely ill-posed. \square

Importantly, the severe ill-posedness arises *independently* of the phase-problem, i.e. also for reconstructions from seemingly complete Fresnel-data $\mathcal{D}(h)|_K$. In practice, the result means that there will always be a large number of image-modes that cannot be recovered from finite detector data at any realistically achievable noise-levels. This prediction is in contradiction to the stable reconstructions achieved in practical XPCI and thus necessitates a deeper analysis of the nature of the found ill-posedness.

9.3 Assessment by Gaussian wave-packets

In the following, we aim to assess stability of inverse problem 9.1 and inverse problem 9.2 by considering Gaussian wave-packets as a special class of object-signals h , for which Fresnel-propagation may be computed analytically. The theory is completely analogous to the textbook-example of wave-packets for the time-dependent Schrödinger-equation.

9.3.1 The Gaussian-beam solution

We consider centered Gaussians of width $\sigma > 0$:

$$p_{\sigma}(\mathbf{x}) = (2\pi\sigma^2)^{-m/2} \exp\left(-\frac{\mathbf{x}^2}{2\sigma^2}\right) \quad \text{for all } \mathbf{x} \in \mathbb{R}^m \quad (9.3.1)$$

Owing to the Gaussian form, $\mathcal{D}(p_{\sigma})$ can be computed explicitly. It constitutes an exact solution to the paraxial Helmholtz equation (9.2.2) known as the *Gaussian beam*, see e.g. [190, Sec. 3.1]. With a certain unitary factor c_0 , it can be written in the form

$$\mathcal{D}(p_{\sigma})(\mathbf{x}) = \frac{\tilde{\sigma}^{m/2} c_0}{\sigma^{m/2}} \exp\left(\frac{i\mathbf{x}^2}{2\eta^2}\right) p_{\tilde{\sigma}}(\mathbf{x}), \quad \eta^2 := \frac{1 + \sigma^4 \mathfrak{f}^2}{\mathfrak{f}}, \quad \tilde{\sigma}^2 := \frac{\eta^2}{\sigma^2 \mathfrak{f}}. \quad (9.3.2)$$

Accordingly, $\mathcal{D}(p_{\sigma})$ is again of Gaussian shape, yet modulated by a unitary oscillatory factor.

Consider the limit $\sigma \rightarrow 0$ of a more and more localized peak. Then the *propagated width* $\tilde{\sigma}$ tends to infinity according to (9.3.2), i.e. the propagated Gaussian $\mathcal{D}(p_\sigma)$ becomes arbitrarily *delocalized*. Indeed, it holds that

$$\lim_{\sigma \rightarrow 0} \|\mathcal{D}(p_\sigma)|_K\|/\|p_\sigma\| = 0 \quad (9.3.3)$$

for any bounded detection-domain $K \subset \mathbb{R}^m$. The example indicates that, asymptotically, the sharper a feature in the object the less contrast it induces in the diffraction data on a finite detector K . Accordingly, a finite FoV limits the achievable *resolution*.

9.3.2 Gaussian wave-packets

In order to further investigate the relation between the detection-domain K and resolution, we study the propagation of Gaussian wave-packets, given by a Gaussian peak that is modulated by a sinusoidal oscillation:

$$h_{\boldsymbol{\xi}, \mathbf{a}}(\mathbf{x}) := \exp(i\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{a})) p_\sigma(\mathbf{x} - \mathbf{a}), \quad \boldsymbol{\xi}, \mathbf{a} \in \mathbb{R}^m. \quad (9.3.4)$$

Analytical propagation of such signals is enabled by the following lemma:

Lemma 9.6 (Fresnel propagation under frequency shifts). *For $\boldsymbol{\xi}, \mathbf{b} \in \mathbb{R}^m$, $\mathbf{e}_\xi(\mathbf{x}) := \exp(i\boldsymbol{\xi} \cdot \mathbf{x})$ denote the Fourier mode to the frequency $\boldsymbol{\xi}$ and $T_{\mathbf{b}} : f \mapsto f(\cdot + \mathbf{b})$ the translation by \mathbf{b} . Then it holds for all $f \in L^2(\mathbb{R}^m)$ that*

$$\mathcal{D}(\mathbf{e}_\xi \cdot f) = m_f(\boldsymbol{\xi}) \cdot \mathbf{e}_\xi \cdot T_{-\boldsymbol{\xi}/f}(\mathcal{D}(f)) \quad (9.3.5)$$

where m_f is the Fresnel factor from (9.2.1).

Lemma 9.6 is proven in appendix 9.A. It states that the Fresnel propagator partly translates frequency-shifts into spatial shifts. By applying (9.3.5) to the Gaussian-beam (9.3.2), we obtain an analytical formula for the propagation of Gaussian wave-packets:

$$\mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{a}})(\mathbf{x}) = \nu_\xi(\mathbf{x} - \mathbf{a} - \boldsymbol{\xi}/f) p_{\tilde{\sigma}}(\mathbf{x} - \mathbf{a} - \boldsymbol{\xi}/f) \quad (9.3.6a)$$

$$\nu_\xi(\mathbf{x}) = \frac{\tilde{\sigma}^{m/2} c_0}{\sigma^{m/2}} \exp\left(i\left(\frac{\mathbf{x}^2}{2\eta^2} + \boldsymbol{\xi} \cdot \mathbf{x} + \frac{\boldsymbol{\xi}^2}{2f}\right)\right). \quad (9.3.6b)$$

The oscillatory factor ν_ξ has constant modulus. Hence, the envelope $|\mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{a}})| \propto p_{\tilde{\sigma}}(\cdot - \mathbf{a} - \boldsymbol{\xi}/f)$ is again a Gaussian of width $\tilde{\sigma}$, whose center is *shifted* by $\boldsymbol{\xi}/f$ with respect to that of the original wave-packet $h_{\boldsymbol{\xi}, \mathbf{a}}$. Accordingly, wave-packets propagate *laterally* within the field-of-view upon action of the Fresnel-propagator.

9.3.3 Resolution estimates via Gaussian wave-packets

We aim to use the analytical propagation formula (9.3.6) for Gaussian wave-packets to derive *upper* bounds the achievable resolution in the reconstruction for inverse problem 9.1 and inverse problem 9.2. Since uniqueness always holds, see theorem 9.3, the only reasonable way to define

resolution is via *stability*: if we claim that the reconstruction has a resolution $1/r$, i.e. that features of the object down to a size $r > 0$ are faithfully recovered, then the reconstruction should be *stable* to perturbations of the object h by any function \tilde{h} that varies on lengthscales r , i.e. the induced contrast in the data should be sufficiently large compared to $\|\tilde{h}\|$. By the hierarchy-of-difficulty of the considered inverse problems and linearity of \mathcal{D} , a necessary condition for this to hold is that $\|\mathcal{D}(h+\tilde{h})|_K - \mathcal{D}(h)|_K\|/\|\tilde{h}\| = \|\mathcal{D}(\tilde{h})|_K\|/\|\tilde{h}\|$ is non-negligible.

Gaussian wave-packets $\tilde{h} = h_{\boldsymbol{\xi}, \mathbf{x}_0}$ of frequency $|\boldsymbol{\xi}| \leq \pi/r$ constitute *special* perturbations varying on lengthscales $\gtrsim r$. Thus, we can derive *upper*, i.e. possibly optimistic bounds on the achievable resolution $1/r$ by identifying parameter-regimes, for which $\|\mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{x}_0})|_K\|/\|h_{\boldsymbol{\xi}, \mathbf{x}_0}\|$ is negligibly small.

9.3.3.1 Resolution for complex-valued images

We study inverse problem 9.1(a) for a square detection-domain $K := [-\frac{1}{2}; \frac{1}{2}]^m$, $\Omega \subset K$. In this setting, the unknown image $h \in L^2(\Omega)$ is complex-valued so that Gaussian wave-packets $\tilde{h} = h_{\boldsymbol{\xi}, \mathbf{x}_0}$ of the form (9.3.4) centered at some point $\mathbf{x}_0 \in \Omega$ constitute admissible perturbations². As seen from (9.3.6), the center of the Gaussian is then shifted to the point $\mathbf{x}_{\text{prop}} := \mathbf{x}_0 + \boldsymbol{\xi}/\mathfrak{f}$ upon Fresnel-propagation. Accordingly, if we consider wave-packets of larger and larger frequency $|\boldsymbol{\xi}|$, then the propagated wave-packet will eventually leave the detection domain, as visualized in fig. 9.3. More quantitatively, upon defining the path-length from a point \mathbf{x} to the detector-boundary ∂K along a direction \mathbf{n} ,

$$\text{dist}_{\mathbf{n}}(\mathbf{x}, \partial K) = \inf \{y \geq 0 : \mathbf{x} + y\mathbf{n} \notin K\} \quad \text{for } \mathbf{x}, \mathbf{n} \in \mathbb{R}^m : |\mathbf{n}| = 1, \quad (9.3.7)$$

the propagated center \mathbf{x}_{prop} is inside K if and only if $\text{dist}_{\boldsymbol{\xi}/|\boldsymbol{\xi}|}(\mathbf{x}_0, \partial K) \leq |\boldsymbol{\xi}|/\mathfrak{f}$. If $\mathbf{x}_{\text{prop}} \in K$, then the induced data-contrast is non-negligible:

$$\|\mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{x}_0})|_K\| \geq 2^{-m/2} \|h_{\boldsymbol{\xi}, \mathbf{x}_0}\| \quad \text{if } \mathbf{x}_{\text{prop}} \in K. \quad (9.3.8)$$

On the contrary, if $\mathbf{x}_{\text{prop}} \notin K$ with distance $\text{dist}(\mathbf{x}_{\text{prop}}, K) \gtrsim \tilde{\sigma}$ greater than the propagated width $\tilde{\sigma}$ of the wave-packet, then the contrast may be quite small:

$$\|\mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{x}_0})|_K\| \leq \frac{1}{2} \text{erfc} \left(\frac{\text{dist}(\mathbf{x}_{\text{prop}}, K)}{\tilde{\sigma}} \right)^{\frac{1}{2}} \|h_{\boldsymbol{\xi}, \mathbf{x}_0}\|. \quad (9.3.9)$$

As the complementary error function $\text{erfc}(x)$ decays very fast for $x \gtrsim 1$, (9.3.9) shows that the perturbation $h \mapsto h + h_{\boldsymbol{\xi}, \mathbf{x}_0}$ is practically *invisible* in the data $\mathcal{D}(h + h_{\boldsymbol{\xi}, \mathbf{x}_0})|_K$ if $|\boldsymbol{\xi}|$ is sufficiently large. In other words, oscillations at \mathbf{x}_0 above a certain cutoff-frequency cannot be *resolved*.

The construction reveals that the local resolution $1/r(\mathbf{x}_0)$ at a point \mathbf{x}_0 is closely related to the distance to the detector-boundary $\text{dist}(\mathbf{x}, \partial K) = \min_{|\mathbf{n}|=1} \text{dist}_{\mathbf{n}}(\mathbf{x}, \partial K)$:

- 1 For all Gaussian wave-packets $h_{\boldsymbol{\xi}, \mathbf{x}_0}$ with $|\boldsymbol{\xi}| < \mathfrak{f} \text{dist}(\mathbf{x}_0, \partial K)$, the propagated center \mathbf{x}_{prop} lies within the detection-domain K

²We ignore that the Gaussian wave-packet is technically not compactly supported and thus $h + h_{\boldsymbol{\xi}, \mathbf{x}_0} \notin L^2(\Omega)$. Note, however, that $h_{\boldsymbol{\xi}, \mathbf{x}_0}|_{\Omega} \approx h_{\boldsymbol{\xi}, \mathbf{x}_0}$ up to a very small L^2 -error given that \mathbf{x}_0 is sufficiently far from the boundary of Ω in units of the Gaussian's width σ .

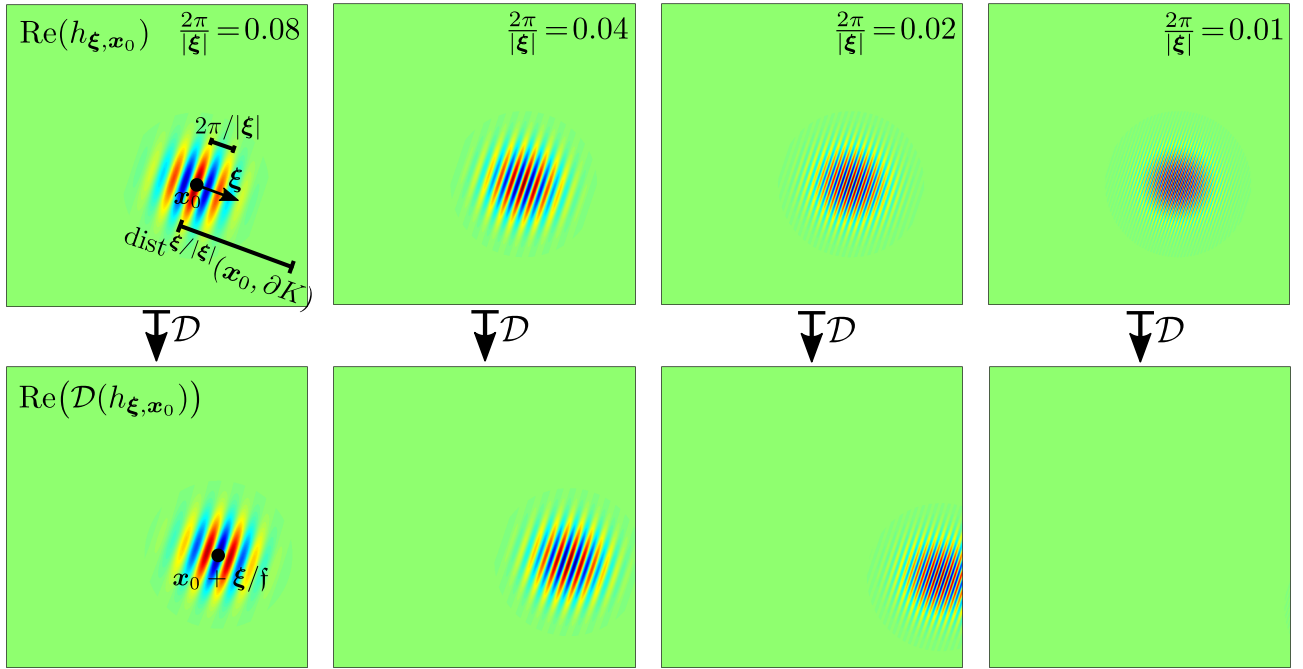


Figure 9.3. Propagation of a Gaussian wave-packet h_{ξ, x_0} for $m = 2$, $K = [-\frac{1}{2}; \frac{1}{2}]^2$, $\mathfrak{f} = 10^3$, $\sigma = 0.08$. Plotted are the real-parts of the complex-valued wave-packet (top row) and its propagated version $\mathcal{D}(h_{\xi, x_0})|_K$ (bottom row) computed via (9.3.6) for different frequencies $|\xi|$. As $|\xi|$ increases from left to right, the propagated wave-packet $\mathcal{D}(h_{\xi, x_0})$ is more and more shifted with respect to h_{ξ, x_0} until it leaves the field of view K (right-most column) and is thus practically invisible to the considered imaging setup. The linear colorscale is identical in all images.

- 2 For all frequencies $\xi > \mathfrak{f} \text{dist}(\mathbf{x}, \partial K)$, there exists a wave-packet $h_{\xi, \mathbf{x}}$ with $|\xi| = \xi$, such that the propagated center \mathbf{x}_{prop} lies outside K

As wave-packets leaving the field-of-view K correspond to non-resolvable lengthscales, these observations translate into a resolution estimate:

Result 9.7 (Resolution limit for complex-valued image reconstruction). For K convex and $\Omega \subset K$, stable reconstruction in inverse problem 9.1 can only be achieved down to a *local resolution limit*

$$1/r(\mathbf{x}) \lesssim \frac{\mathfrak{f} \text{dist}(\mathbf{x}, \partial K)}{\pi} \quad \text{for all } \mathbf{x} \in \Omega, \quad (9.3.10)$$

where $r(\mathbf{x})$ denotes the smallest resolvable feature-size of the image h at position \mathbf{x} .

In particular, for $K = [-\frac{1}{2}; \frac{1}{2}]^m$, the global maximum resolution is bounded by the *classical Fresnel number* of the imaging setup (see §9.2.1.4): $\max_{\mathbf{x} \in K} 1/r(\mathbf{x}) \lesssim \mathfrak{f}/(2\pi) = \mathfrak{f}$.

The resolution limit stated in result 9.7 is *isotropic* – the resolution for features along a specific direction may be higher. Figure 9.4(a) shows the spatially varying resolution according to the estimate (9.3.10) for the exemplary setting $m = 2$, $K = [-\frac{1}{2}; \frac{1}{2}]^2$, $\mathfrak{f} = 10^4$. Note that the maximum resolution $\max_{\mathbf{x} \in K} 1/r(\mathbf{x}) = \mathfrak{f}$ coincides with predictions according to Abbe’s diffraction limit if the detector-size defines the numerical aperture, compare [158, 127]. Interestingly,

however, the resolution only attains this optimum in the very center of the FoV as it decreases towards the detector-edges.

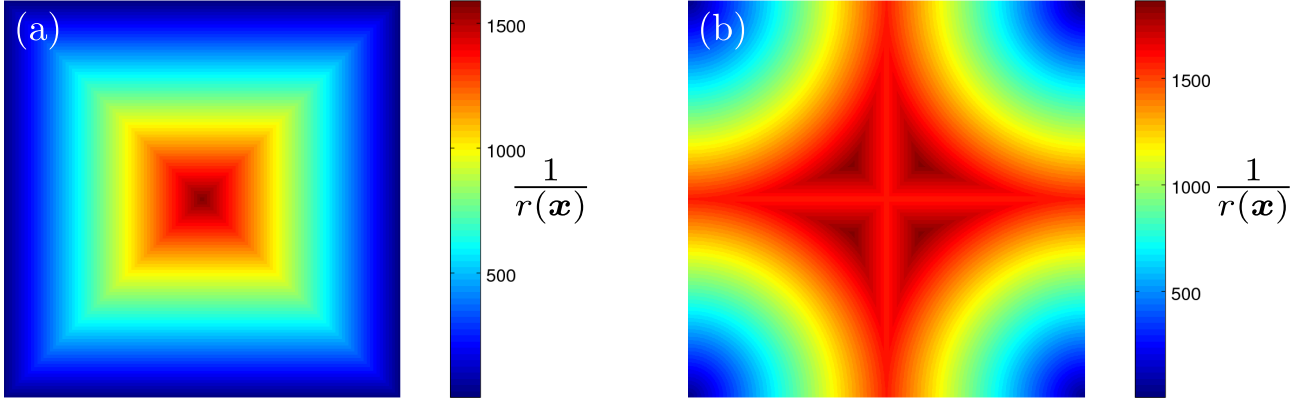


Figure 9.4. (a) Upper bound on the stably reconstructible local resolution $1/r(\mathbf{x})$ in inverse problem 9.1 (complex-valued images) for $m = 2$, $K = [-\frac{1}{2}; \frac{1}{2}]^2$, $f = 10^4$ according to the estimate (9.3.10). (b) Same plot for inverse problem 9.2, i.e. for real-valued image reconstruction, according to the estimate (9.3.15).

9.3.3.2 Resolution for real-valued images

In the case of inverse problem 9.2(a), real-valued images are to be reconstructed so that complex-valued Gaussian wave-packets are no longer admissible perturbations. Accordingly, we study *real-valued* wave-packets. Such signals are given by a superposition of two Gaussian wave-packets with wavevectors $\boldsymbol{\xi}$ and $-\boldsymbol{\xi}$:

$$\begin{aligned} h_{\boldsymbol{\xi}, \mathbf{a}}^{\text{real}}(\mathbf{x}) &:= \cos(\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{a}) + \beta) p_{\sigma}(\mathbf{x} - \mathbf{a}) \\ &= \text{Re}(e^{i\beta} h_{\boldsymbol{\xi}, \mathbf{a}}(\mathbf{x})) = \frac{1}{2} (e^{i\beta} h_{\boldsymbol{\xi}, \mathbf{a}}(\mathbf{x}) + e^{-i\beta} h_{-\boldsymbol{\xi}, \mathbf{a}}(\mathbf{x})) \end{aligned} \quad (9.3.11)$$

for $\mathbf{x}, \boldsymbol{\xi}, \mathbf{a} \in \mathbb{R}^m$, $\beta \in [0; 2\pi)$. Using (9.3.6) and linearity of the Fresnel-propagator, an analytical propagation formula is obtained for $h_{\boldsymbol{\xi}, \mathbf{a}}^{\text{real}}$:

$$\begin{aligned} \mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{a}}^{\text{real}})(\mathbf{x}) &= \frac{e^{i\beta}}{2} \mathcal{D}(h_{\boldsymbol{\xi}, \mathbf{a}})(\mathbf{x}) + \frac{e^{-i\beta}}{2} \mathcal{D}(h_{-\boldsymbol{\xi}, \mathbf{a}})(\mathbf{x}) \\ &= \frac{e^{i\beta}}{2} \nu_{\boldsymbol{\xi}}(\mathbf{x} - \mathbf{a} - \boldsymbol{\xi}/f) p_{\bar{\sigma}}(\mathbf{x} - \mathbf{a} - \boldsymbol{\xi}/f) + \frac{e^{-i\beta}}{2} \nu_{-\boldsymbol{\xi}}(\mathbf{x} - \mathbf{a} + \boldsymbol{\xi}/f) p_{\bar{\sigma}}(\mathbf{x} - \mathbf{a} + \boldsymbol{\xi}/f) \end{aligned} \quad (9.3.12)$$

The analytical solution (9.3.12) reveals surprising features of the propagated signal: upon propagation, the wave-packet splits up into two packets propagating into opposite directions $\pm \boldsymbol{\xi}$ as visualized in fig. 9.5. This has important consequences in terms of stability: if an object $h \in L^2(\Omega, \mathbb{R})$ is perturbed by a real-valued wave-packet $h_{\boldsymbol{\xi}, \mathbf{x}}^{\text{real}}$ at some point $\mathbf{x} \in K$, then this perturbation manifests non-negligibly in the data $\mathcal{D}(h + h_{\boldsymbol{\xi}, \mathbf{x}}^{\text{real}})|_K$ as long as *either* of the two wave-packets remains within the field-of-view K . For a point $\mathbf{x} \in K$ and a direction \mathbf{n} , we

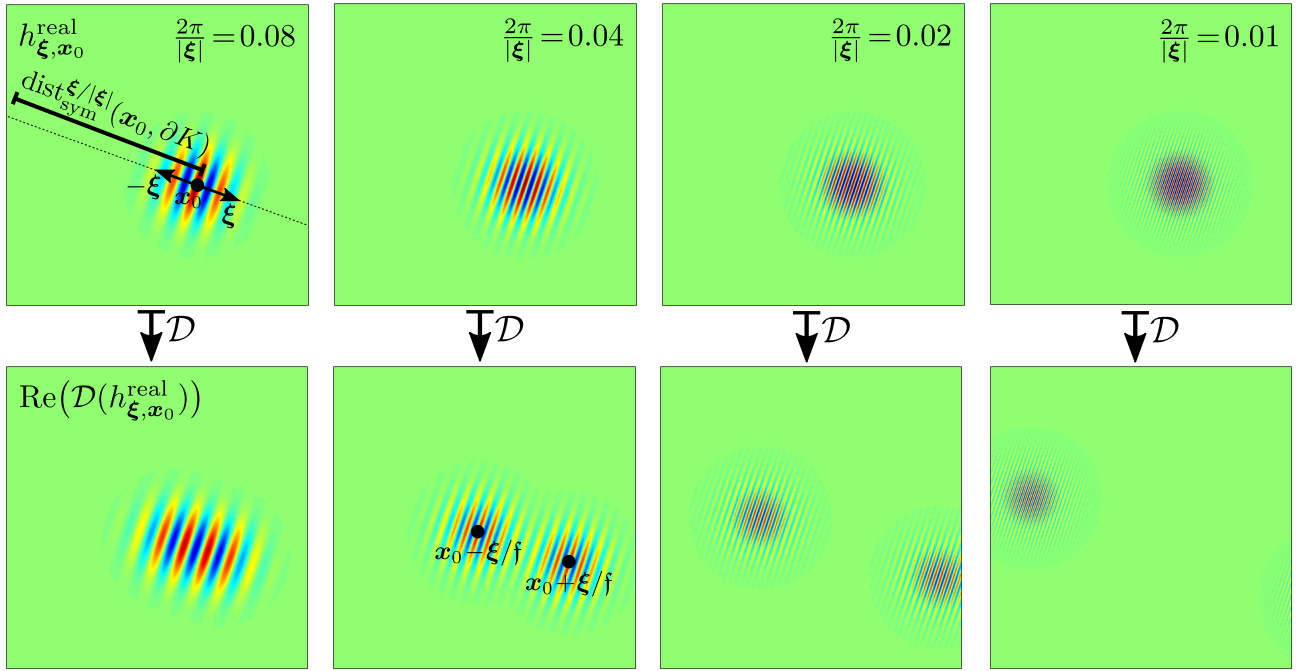


Figure 9.5. Analogue of fig. 9.3 for a real-valued wave-packet $h_{\xi, x_0}^{\text{real}}$. Upon Fresnel-propagation, such split up into two wave-packets that laterally propagate along opposite directions $\pm \xi$. Consequently, the induced data contrast $\mathcal{D}(h_{\xi, x_0}^{\text{real}})|_K$ is non-negligible under milder conditions than in the complex-valued case. For details, see text.

therefore introduce the following distance-measure:

$$\text{dist}_{\text{sym}}^n(\mathbf{x}, \partial K) = \inf \{y \geq 0 : \mathbf{x} + yn \notin K \text{ and } \mathbf{x} - yn \notin K\} \quad (9.3.13)$$

$$\text{dist}_{\text{sym}}(\mathbf{x}, \partial K) = \inf_{|n|=1} \text{dist}_{\text{sym}}^n(\mathbf{x}, \partial K). \quad (9.3.14)$$

$\text{dist}_{\text{sym}}^n(\mathbf{x}, \partial K)$ gives the larger length of the two line-segments $\{\mathbf{x} \pm yn : y \geq 0\} \cap K$, which connect \mathbf{x} with the boundary of ∂K along \mathbf{n} . In view of wave-packets, the interpretation is simple: for $\mathbf{x}, \xi \in \mathbb{R}^m$, the centers of *both* propagating wave-packets forming $\mathcal{D}(h_{\xi, x}^{\text{real}})$ lie outside K if and only if $\text{dist}_{\text{sym}}^{\xi/|\xi|}(\mathbf{x}, \partial K) < |\xi|/f$. Hence, the following relations hold true:

- 1 For all wave-packets $h_{\xi, x}^{\text{real}}$ with $|\xi| < f \text{dist}_{\text{sym}}(\mathbf{x}, \partial K)$, the center of one of the propagating wave-packets lies within K .
- 2 For all frequencies $\xi > f \text{dist}_{\text{sym}}(\mathbf{x}, \partial K)$, there exists a wave-packet $h_{\xi, x}^{\text{real}}$ with $|\xi| = \xi$, such that the center of both wave-packets lie outside of K .

Accordingly, the quantity $\text{dist}_{\text{sym}}(\mathbf{x}, \partial K)$ yields an upper bound for the local resolution in the real-valued setting:

Result 9.8 (Resolution limit for real-valued image reconstruction). For K convex and $\Omega \subset K$, stable reconstruction in inverse problem 9.2 can only be achieved down to a *local resolution limit*

$$1/r(\mathbf{x}) \lesssim \frac{f \text{dist}_{\text{sym}}(\mathbf{x}, \partial K)}{\pi} \quad \text{for all } \mathbf{x} \in \Omega, \quad (9.3.15)$$

where $r(\mathbf{x})$ denotes the smallest resolvable feature-size of the image φ at position \mathbf{x} .

For $m = 2$, $K = [-\frac{1}{2}; \frac{1}{2}]^2$, dist_{sym} can be evaluated analytically:

$$\text{dist}_{\text{sym}}((x_1, x_2), \partial K) = \min \left\{ \left(\left(\frac{1}{2} - |x_1| \right)^2 + \left(\frac{1}{2} - |x_2| \right)^2 \right)^{\frac{1}{2}}, \min_{j \in \{1, 2\}} \max_{\pm} \frac{1}{2} \pm x_j \right\} \quad (9.3.16)$$

The resulting spatially varying resolution for $\mathfrak{f} = 10^4$ is plotted in fig. 9.4(b). Notably, the maximum resolution is attained slightly off-center and is higher than in complex-valued case, compare fig. 9.4(a). Moreover, a high resolution $1/r \geq 1000$ is obtained within a much larger subdomain of the field of view K . Most prominently, the resolution in fig. 9.4(b) even remains large near the detector boundary – except for the corners of K . Yet, the maximum resolution remains essentially bounded by $\max_{\mathbf{x} \in K} 1/r(\mathbf{x}) \lesssim \mathfrak{f}$.

9.4 Locality estimates for complex-valued objects

The goal of the subsequent sections is to complement the (potentially) *optimistic* resolution estimates from §9.3 with *worst-case* bounds. Accordingly, we aim to prove that stable image reconstruction can indeed be achieved down to a certain resolution. Note that this is necessarily more involved than the preceding analysis because stability has to be proven with respect to *general* perturbations instead of considering just a special class like Gaussian wave-packets.

9.4.1 Basic idea and preliminaries

The principal difficulty in proving stability-estimates for bounded detection domains $K \subset \mathbb{R}^m$ lies in the pronounced *non-locality* of the Fresnel-propagator: according to (P2), it is given by a convolution with a kernel $k_{\mathfrak{f}}(\mathbf{x}) \propto \exp(i\mathfrak{f}\mathbf{x}/2)$ that shows *no spatial decay whatsoever!* Hence, Fresnel-propagation may transport object-information over arbitrary lateral distances in principle, i.e. features of the imaged object $h \in L^2(\Omega)$ with $\Omega \subset K$ may manifest far outside the field-of-view K in the diffraction data $\mathcal{D}(h)$. In addition to this non-locality in real-space, any restriction to $K \subsetneq \mathbb{R}^m$ breaks the translational invariance of \mathcal{D} and thus its diagonality, i.e. locality, in Fourier-space.

On the other hand, it has been seen in §9.3 that the distance, by which object-information is transported laterally, depends on the spatial frequencies of the signal. Accordingly, locality might be established by restricting to lower frequencies, i.e. to sufficiently smooth objects.

The principal idea of the subsequent analysis is to decompose the convolution kernel $k_{\mathfrak{f}}$ into an inner, local part, and an outer non-local part:

$$k_{\mathfrak{f}} = k_{\mathfrak{f}}|_P + k_{\mathfrak{f}}|_{P^c} \quad \text{for some } P \subset \mathbb{R}^m. \quad (9.4.1)$$

For an object $h \in L^2(\Omega)$ supported in $\Omega \subset K \subset \mathbb{R}^m$ and a suitably chosen P , the wave-field leaked outside K depends only on the outer part: $\mathcal{D}(h)|_{K^c} = (k_{\mathfrak{f}} * h)|_{K^c} = (k_{\mathfrak{f}}|_{P^c} * h)|_{K^c}$. This implies estimates of the form $\|\mathcal{D}(h)|_{K^c}\| \leq \|k_{\mathfrak{f}}|_{P^c} * h\|$, which are diagonal in Fourier-space and thus simple to interpret as the norm of a *filtered* object.

Notation: indicator functions For a set $A \subset \mathbb{R}^m$, let $\mathbf{1}_A : \mathbb{R}^m \rightarrow \mathbb{R}$ be defined by $\mathbf{1}_A(\mathbf{x}) = 1$ if $\mathbf{x} \in A$ and $\mathbf{1}_A(\mathbf{x}) = 0$ otherwise.

9.4.2 Principal leakage estimates

Our principal leakage estimate is based on the insight that the frequency response of a restricted Fresnel-kernel $k_{\mathfrak{f}}|_P$ is readily computable:

Lemma 9.9 (Frequency response of a restricted Fresnel-kernels). *Let $P \subset \mathbb{R}^m$ be a measurable set such that $\mathcal{D}(\mathbf{1}_P) \in L^\infty(\mathbb{R}^m)$ is well-defined and bounded. Let $k_{\mathfrak{f}}$ denote the convolution-kernel of the Fresnel-propagator. Then it holds for all $h \in L^2(\mathbb{R}^m)$ that*

$$\mathcal{F}(k_{\mathfrak{f}}|_P * h)(\boldsymbol{\xi}) = m_{\mathfrak{f}}(\boldsymbol{\xi}) \cdot \mathcal{D}(\mathbf{1}_P)(\boldsymbol{\xi}/\mathfrak{f}) \cdot \mathcal{F}(h)(\boldsymbol{\xi}) \quad \text{for almost all } \boldsymbol{\xi} \in \mathbb{R}^m \quad (9.4.2)$$

and in particular for any measurable set $A \subset \mathbb{R}^m$:

$$\|(k_{\mathfrak{f}}|_P * h)|_A\| \leq \|\mathcal{D}(\mathbf{1}_P)(\cdot/\mathfrak{f}) \cdot \mathcal{F}(h)\|. \quad (9.4.3)$$

Proof. By the assumption $\mathcal{D}(\mathbf{1}_P) \in L^\infty(\mathbb{R}^m)$, both sides of the equation (9.4.2) are continuous in h with respect to the L^2 -norm. Hence, it is sufficient to prove the claim for Schwartz-functions $h \in \mathcal{S}(\mathbb{R}^m)$ by denseness of these in $L^2(\mathbb{R}^m)$.

For $h \in \mathcal{S}(\mathbb{R}^m)$, the convolution $k_{\mathfrak{f}}|_P * h$ is well-defined in a pointwise sense but can also be regarded as convolution between a Schwartz-function and a tempered distribution $k_{\mathfrak{f}}|_P \in \mathcal{S}'(\mathbb{R}^m)$. Accordingly, the convolution theorem holds, i.e.

$$\mathcal{F}((\mathbf{1}_P \cdot k_{\mathfrak{f}}) * h) = (2\pi)^{-m/2} \mathcal{F}(\mathbf{1}_P \cdot k_{\mathfrak{f}}) \cdot \mathcal{F}(h) \quad (9.4.4)$$

in a distributional sense. Recalling that the alternate form of the Fresnel-propagator (P3) remains valid for tempered distributions, we get

$$\begin{aligned} \mathcal{F}(\mathbf{1}_P \cdot k_{\mathfrak{f}}) &= u_0 \mathfrak{f}^{\frac{m}{2}} \mathcal{F}(\mathbf{1}_P \cdot n_{\mathfrak{f}}) = \frac{1}{n_{\mathfrak{f}}(\cdot/\mathfrak{f})} (u_0 \mathfrak{f}^{\frac{m}{2}} n_{\mathfrak{f}}(\cdot/\mathfrak{f}) \cdot \mathcal{F}(\mathbf{1}_P \cdot n_{\mathfrak{f}})(\mathfrak{f}(\cdot/\mathfrak{f}))) \\ &\stackrel{\text{(P3)}}{=} m_{\mathfrak{f}} \cdot \mathcal{D}(\mathbf{1}_P)(\cdot/\mathfrak{f}) \end{aligned} \quad (9.4.5)$$

Inserting (9.4.5) into (9.4.4) yields (9.4.2). The inequality (9.4.3) now follows by using unitarity of the Fourier transform along with the observations that $m_{\mathfrak{f}}$ has constant modulus 1 and that the restriction-operation $f \mapsto f|_A$ is non-increasing in the L^2 -norm:

$$\begin{aligned} \|(k_{\mathfrak{f}}|_P * h)|_A\| &\leq \|k_{\mathfrak{f}}|_P * h\| = \|\mathcal{F}(k_{\mathfrak{f}}|_P * h)\| \stackrel{(9.4.2)}{=} \|m_{\mathfrak{f}} \cdot \mathcal{D}(\mathbf{1}_P)(\cdot/\mathfrak{f}) \cdot \mathcal{F}(h)\| \\ &= \|\mathcal{D}(\mathbf{1}_P)(\cdot/\mathfrak{f}) \cdot \mathcal{F}(h)\|. \end{aligned} \quad (9.4.6)$$

□

A surprising feature of lemma 9.9 is that $\mathcal{D}(\dots)$ occurs as a factor in *Fourier-space*. Similar as lemma 9.6, this reveals an interesting real-space-Fourier-space-duality of the Fresnel-propagator. Using lemma 9.9, we may derive leakage estimates as outlined in §9.4.1:

Theorem 9.10 (Principal leakage estimate). *Let $K, \Omega, P_{\text{leak}} \subset \mathbb{R}^m$ be measurable sets such that the boundary ∂K has Lebesgue-measure zero and $\Omega + P_{\text{leak}} = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in \Omega, \mathbf{y} \in P_{\text{leak}}\} \subset K$. Moreover, let $\mathcal{D}(\mathbf{1}_{P_{\text{leak}}^c}) \in L^\infty(\mathbb{R}^m)$. Then it holds for all $h \in L^2(\Omega)$*

$$\|\mathcal{D}(h)|_{K^c}\| \leq \|\hat{p}^{\text{leak}} \cdot \mathcal{F}(h)\|, \quad p^{\text{leak}}(\boldsymbol{\xi}) := |\mathcal{D}(\mathbf{1}_{P_{\text{leak}}^c})(\boldsymbol{\xi}/f)|. \quad (9.4.7)$$

Proof. By a similar continuity argument as in lemma 9.9 it is sufficient to prove the claim for Schwartz-functions $h \in L^2(\Omega) \cap \mathcal{S}(\mathbb{R}^m)$. Then the convolution-form (P2) of the Fresnel-propagator may be used. Hence, we have

$$\mathcal{D}(h)|_{K^c} = (k_f * h)|_{K^c} = ((\mathbf{1}_{P_{\text{leak}}} \cdot k_f) * h)|_{K^c} + ((\mathbf{1}_{P_{\text{leak}}^c} \cdot k_f) * h)|_{K^c} \quad (9.4.8)$$

According to standard results on the support of convolutions, it holds that

$$\text{supp}((\mathbf{1}_{P_{\text{leak}}} \cdot k_f) * h) \subset \overline{\text{supp}(\mathbf{1}_{P_{\text{leak}}} \cdot k_f) + \text{supp}(h)} \subset \overline{P_{\text{leak}} + \Omega} \subset \overline{K}. \quad (9.4.9)$$

(9.4.9) implies that $((\mathbf{1}_{P_{\text{leak}}} \cdot k_f) * h)|_{K^c}$ vanishes except for possibly the boundary ∂K . As ∂K is a set of measure zero, $((\mathbf{1}_{P_{\text{leak}}} \cdot k_f) * h)|_{K^c} = 0$ holds in an L^2 -sense. Thus, (9.4.8) yields

$$\|\mathcal{D}(h)|_{K^c}\| = \|((\mathbf{1}_{P_{\text{leak}}^c} \cdot k_f) * h)|_{K^c}\| \quad (9.4.10)$$

Applying the bound (9.4.3) to the right-hand side of (9.4.10) now yields the assertion. \square

Theorem 9.10 bounds the leaked wave-field $\mathcal{D}(h)|_{K^c}$ in terms of a *filtering-operation*. In order to predict in which cases leakage is small or large, we need to understand the nature of the filter-response \hat{p}^{leak} that weights the Fourier-components of h in (9.4.7). If $\Omega \subset K \subset \mathbb{R}^m$, then the largest admissible set P in theorem 9.10 is some bounded domain containing 0, where the exact size of P depends on the distance between Ω and ∂K . Let us assume that the size of P is much larger than $1/f^{\frac{1}{2}}$ as will be the typical case in the following. Then the indicator-function $\mathbf{1}_{P^c}$ is essentially preserved upon Fresnel-propagation, i.e. $\mathcal{D}(\mathbf{1}_{P^c}) \approx \mathbf{1}_{P^c} = 1 - \mathbf{1}_P$ up to some oscillations near the boundary of P . Accordingly, $\hat{p}^{\text{leak}} = \mathcal{D}(\mathbf{1}_{P^c})(\cdot/f)$ acts as a *high-pass filter*, essentially damping out all Fourier-frequencies within the domain $f \cdot \mathbf{1}_P$. Thus, the right-hand side of (9.4.7) is small for sufficiently *smooth* objects h .

9.4.3 Explicit leakage bounds for rectangular domains

For general sets P_{leak} , \hat{p}^{leak} from theorem 9.10 cannot be computed explicitly. An exception is given by rectangular domains owing to the known Fresnel-transform of the Heaviside-function $\theta = \mathbf{1}_{\mathbb{R}_{\geq 0}}$ in terms of *Fresnel-integrals* [130]:

$$\begin{aligned} \mathcal{D}(\theta)(x) &= \frac{1}{2} - \frac{1-i}{2} \left(\text{C} \left(\frac{-f^{\frac{1}{2}}x}{\pi^{\frac{1}{2}}} \right) + i \text{S} \left(\frac{-f^{\frac{1}{2}}x}{\pi^{\frac{1}{2}}} \right) \right) =: \tilde{\theta}(f^{\frac{1}{2}}x) \\ \text{C}(x) &:= \int_0^x \cos\left(\frac{\pi}{2}t^2\right) dt, \quad \text{S}(x) := \int_0^x \sin\left(\frac{\pi}{2}t^2\right) dt \quad \text{for all } x \in \mathbb{R} \end{aligned} \quad (9.4.11)$$

Note that $\tilde{\theta}$ is an entire analytic function and bounded with $\max_{x \in \mathbb{R}} |\tilde{\theta}(x)| \leq 1.171$.

By the separability- and isotropy-properties of the Fresnel-propagator, (P4) and (P5), the explicit solution generalizes to half-spaces $H_{a,\mathbf{n}} := \{\mathbf{x} \in \mathbb{R}^m : \mathbf{n} \cdot \mathbf{x} \geq a\}$ in arbitrary dimensions with $a \in \mathbb{R}$ and $\mathbf{n} \in \mathbb{S}^{m-1} = \{\mathbf{x} \in \mathbb{R}^m : |\mathbf{n}| = 1\}$:

$$\mathcal{D}(\mathbf{1}_{H_{a,\mathbf{n}}})(\mathbf{x}) = \tilde{\theta}(\mathfrak{f}^{\frac{1}{2}}(\mathbf{n} \cdot \mathbf{x} - a)) \quad \text{for all } \mathbf{x} \in \mathbb{R}^m. \quad (9.4.12)$$

Using linearity of \mathcal{D} , (9.4.11) furthermore yields the Fresnel-transform of intervals:

$$\begin{aligned} \mathcal{D}(\mathbf{1}_{[-\Delta;\Delta]})(x) &= \mathcal{D}(\mathbf{1}_{[\Delta;\infty)})(x) - \mathcal{D}(\mathbf{1}_{[-\Delta;\infty)})(x) = \tilde{\theta}\left(\mathfrak{f}^{\frac{1}{2}}(x - \Delta)\right) - \tilde{\theta}\left(\mathfrak{f}^{\frac{1}{2}}(x + \Delta)\right) \\ &= \tilde{\theta}\left(\mathfrak{f}^{\frac{1}{2}}x - \mathfrak{f}_{\Delta}^{\frac{1}{2}}\right) - \tilde{\theta}\left(\mathfrak{f}^{\frac{1}{2}}x + \mathfrak{f}_{\Delta}^{\frac{1}{2}}\right) =: \tilde{t}_{\mathfrak{f}_{\Delta}}(\mathfrak{f}^{\frac{1}{2}}x) \end{aligned} \quad (9.4.13)$$

Here, we introduced the Fresnel number associated with the lateral lengthscale $\Delta > 0$, $\mathfrak{f}_{\Delta} := \Delta^2 \mathfrak{f}$, compare §9.2.1.4. Again by the separability of the Fresnel-propagator, this generalizes to stripe-shaped domains $S_{\Delta,\mathbf{n}} := \{\mathbf{x} \in \mathbb{R}^m : -\Delta \leq \mathbf{n} \cdot \mathbf{x} \leq \Delta\}$ and squares:

$$\mathcal{D}(S_{\Delta,\mathbf{n}})(\mathbf{x}) = \tilde{t}_{\mathfrak{f}_{\Delta}}\left(\mathfrak{f}^{\frac{1}{2}}(\mathbf{n} \cdot \mathbf{x})\right) =: \tilde{t}_{\mathfrak{f}_{\Delta},\mathbf{n}}(\mathbf{x}), \quad \tilde{t}_{\mathfrak{f}_{\Delta},j} := \tilde{t}_{\mathfrak{f}_{\Delta},\mathbf{e}_j} \quad (9.4.14)$$

$$\mathcal{D}(\mathbf{1}_{[-\Delta;\Delta]^m})(\mathbf{x}) = \prod_{j=1}^m \mathcal{D}(\mathbf{1}_{[-\Delta;\Delta]})(x_j) = \prod_{j=1}^m \tilde{t}_{\mathfrak{f}_{\Delta}}(\mathfrak{f}^{\frac{1}{2}}x_j) = \prod_{j=1}^m \tilde{t}_{\mathfrak{f}_{\Delta},j}(\mathfrak{f}^{\frac{1}{2}}\mathbf{x}) \quad (9.4.15)$$

for all $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^m$, where $\mathbf{e}_j \in \mathbb{R}^m$ denotes the j -th unit normal vector. Finally, indicator functions of *complements* are simple to propagate using linearity and $\mathcal{D}(1) = 1$:

$$\mathcal{D}(\mathbf{1}_{A^c}) = \mathcal{D}(1 - \mathbf{1}_A) = \mathcal{D}(1) - \mathcal{D}(\mathbf{1}_A) = 1 - \mathcal{D}(\mathbf{1}_A) \quad \text{for } A \subset \mathbb{R}^m. \quad (9.4.16)$$

Using the formulas derived above, we can explicitly write down the filter from theorem 9.10 for the important special case of square domains:

Corollary 9.11 (Leakage bound for square domains). *Let $K = [-\frac{1}{2}; \frac{1}{2}]^m$ and $\Omega = [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta]^m$ for some $0 < \Delta < \frac{1}{2}$. Then it holds for all $h \in L^2(\Omega)$*

$$\|\mathcal{D}(h)|_{K^c}\| \leq \|\hat{p}_{\square,\mathfrak{f},\mathfrak{f}_{\Delta}}^{\text{leak}} \cdot \mathcal{F}(h)\|, \quad \hat{p}_{\square,\mathfrak{f},\mathfrak{f}_{\Delta}}^{\text{leak}}(\boldsymbol{\xi}) := \left| 1 - \prod_{j=1}^m \tilde{t}_{\mathfrak{f}_{\Delta},j}(\boldsymbol{\xi}/\mathfrak{f}^{\frac{1}{2}}) \right|, \quad \boldsymbol{\xi} \in \mathbb{R}^m \quad (9.4.17)$$

Proof. If we set $P_{\text{leak}} := [-\Delta; \Delta]^m$, the assumptions of theorem 9.10 are satisfied. The expression for the filter in (9.4.17) follows by using (9.4.15) along with (9.4.16). \square

The filter-response $\hat{p}_{\square,\mathfrak{f},\mathfrak{f}_{\Delta}}^{\text{leak}}$ from corollary 9.11 is plotted in fig. 9.6(a) for an exemplary 1D-setting. It can be seen to be a *high-pass* filter of width $\approx 2(\mathfrak{f} \cdot \mathfrak{f}_{\Delta})^{1/2} = 2\mathfrak{f}\Delta$ in Fourier-space. Note that this width is in perfect agreement with the expected cut-off frequency from the Gaussian wave-packet analysis in §9.3.3.1 for the considered distance $\text{dist}(\Omega, \partial K) = \Delta$ between object-domain and detector-boundary. However, $\hat{p}_{\square,\mathfrak{f},\mathfrak{f}_{\Delta}}^{\text{leak}}$ is heavily oscillatory on fine scales and is *not* everywhere ≤ 1 , although this would be reasonable by unitarity of the Fresnel-propagator. Another drawback of the filter-response in (9.4.17) is that it varies in a non-trivial manner in higher dimensions.

Both the oscillatory behavior and the complicated high-dimensional structure can be resolved by exploiting the simple rectangular geometry to obtain an alternative filter:

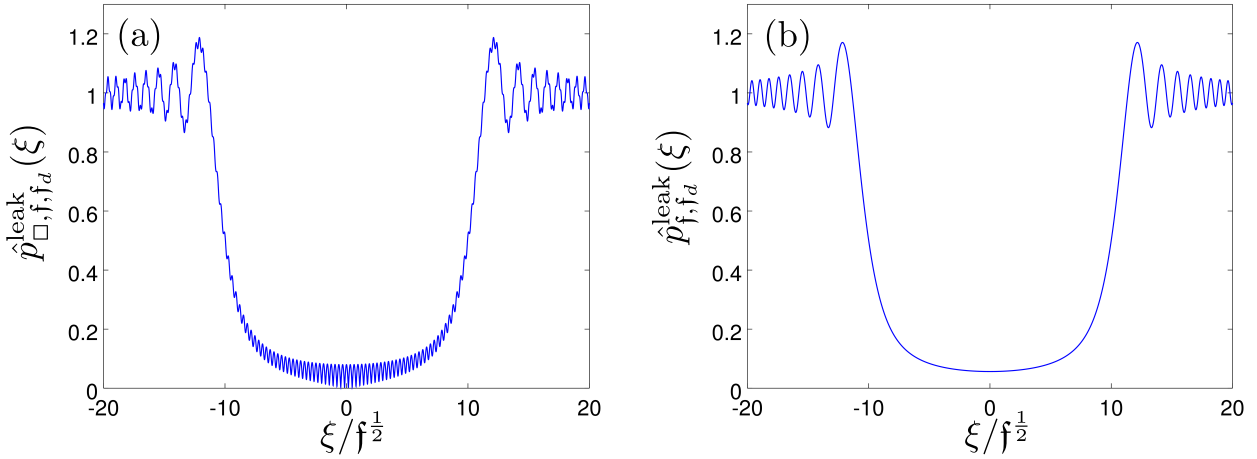


Figure 9.6. Plot of the leakage-filters for square-domains from (a) corollary 9.11 and (b) theorem 9.12 for $m = 1$ dimensions and $f_\Delta = 100$.

Theorem 9.12 (Leakage bound for square domains with simplified filter). *Within the setting of corollary 9.11, it holds for all $h \in L^2(\Omega)$*

$$\|\mathcal{D}(h)|_{K^c}\| \leq \|\hat{p}_{f, f_\Delta}^{\text{leak}} \cdot \mathcal{F}(h)\|, \quad \hat{p}_{f, f_\Delta}^{\text{leak}}(\boldsymbol{\xi}) := \left(\sum_{j=1}^m \left| \tilde{\eta}_{f_\Delta} \left(f^{-\frac{1}{2}} \mathbf{e}_j \cdot \boldsymbol{\xi} \right) \right|^2 \right)^{\frac{1}{2}} \quad (9.4.18)$$

$$\tilde{\eta}_{f_\Delta}(x) := \left(|\tilde{\theta}(x - f_\Delta^{\frac{1}{2}})|^2 + |\tilde{\theta}(-x - f_\Delta^{\frac{1}{2}})|^2 \right)^{\frac{1}{2}} \quad (9.4.19)$$

where \mathbf{e}_j denotes the unit normal vector along the j -th dimension.

Proof. Let $h \in L^2(\Omega)$. If we define the half-spaces $H_{j,\pm} := \mathbb{R}^{j-1} \times \pm[-\frac{1}{2}; \infty) \times \mathbb{R}^{m-j}$, then it holds that $K^c = \bigcup_{j=1, \pm}^m H_{j,\pm}$. Thus,

$$\|\mathcal{D}(h)|_{K^c}\|^2 \leq \sum_{\pm, j=1}^m \|\mathcal{D}(h)|_{H_{j,\pm}}\|^2 \quad (9.4.20)$$

Upon setting $\tilde{K} := H_{j,\pm}$ and $P_{\text{leak}} := \mathbb{R}^{j-1} \times \pm[-\Delta; \infty) \times \mathbb{R}^{m-j}$, theorem 9.10 is applicable so that each of the squared norms on the right-hand side (9.4.20) can be bounded via (9.4.7):

$$\begin{aligned} \|\mathcal{D}(h)|_{H_{j,\pm}}\|^2 &\leq \langle \mathcal{F}(h), |\hat{p}_{f, f_\Delta, j, \pm}^{\text{leak}}|^2 \cdot \mathcal{F}(h) \rangle \\ \hat{p}_{f, f_\Delta, j, \pm}^{\text{leak}}(\boldsymbol{\xi}) &= \mathcal{D}(\mathbf{1}_{P_{\text{leak}}^c})(\boldsymbol{\xi}/f) \stackrel{(9.4.12)}{=} \tilde{\theta}(\pm f^{-\frac{1}{2}} \mathbf{e}_j \cdot \boldsymbol{\xi} - f_\Delta^{\frac{1}{2}}). \end{aligned} \quad (9.4.21)$$

Substituting (9.4.21) into (9.4.20) and using sesqui-linearity yields the assertion. \square

Figure 9.6(b) plots the alternate filter-response $\hat{p}_{f, f_\Delta}^{\text{leak}}$ for the same 1D-setting as in fig. 9.6(a). The plot reveals that the filter-profiles are almost identical except that the oscillations are eliminated from the low-frequency regime. This makes it easier to derive bounds for $|\hat{p}_{f, f_\Delta}^{\text{leak}}|$ in a given frequency-interval compared to the original filter $\hat{p}_{\square, f, f_\Delta}^{\text{leak}}$.

9.4.4 Stability estimates

By unitarity of the Fresnel-propagator \mathcal{D} , *upper* bounds on the wave-field leaked into K^c induce *lower* bounds on the contrast within the field-of-view K , i.e. stability estimates for the reconstruction of h from data $\mathcal{D}(h)|_K$:

Corollary 9.13 (Stability estimates for square domains). *Let $K = [-\frac{1}{2}; \frac{1}{2}]^m$ and $\Omega = [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta]^m$ for some $0 < \Delta < \frac{1}{2}$. Then it holds for all $h \in L^2(\Omega)$*

$$\|\mathcal{D}(h)|_K\|^2 \geq \langle \mathcal{F}(h), (1 - |\hat{p}^{\text{leak}}|^2) \cdot \mathcal{F}(h) \rangle \quad (9.4.22)$$

with $\hat{p}^{\text{leak}} \in \{\hat{p}_{\text{f},\text{f}\Delta}^{\text{leak}}, \hat{p}_{\text{f},\text{f}\Delta}^{\text{leak}}\}$ as defined in theorem 9.12 and corollary 9.11.

Proof. The claim follows from theorem 9.12 and corollary 9.11 as $\|\mathcal{D}(h)|_K\|^2 = \|\mathcal{D}(h)\|^2 - \|\mathcal{D}(h)|_{K^c}\|^2$ and $\|\mathcal{D}(h)\|^2 = \langle \mathcal{F}(h), \mathcal{F}(h) \rangle$ as \mathcal{D} and \mathcal{F} are unitary. \square

Corollary 9.13 gives lower- and upper bounds on the contrast on a square detector in terms of *filtering* operations with explicitly known profile in Fourier-space. It is tempting to interpret the bound as the norm of a low-pass-filtered version of h :

$$\langle \mathcal{F}(h), (1 - |\hat{p}^{\text{leak}}|^2) \cdot \mathcal{F}(h) \rangle \stackrel{!}{=} \|(1 - |\hat{p}^{\text{leak}}|^2)^{\frac{1}{2}} \cdot \mathcal{F}(h)\|^2 \quad (9.4.23)$$

However, this is technically not correct because $|\hat{p}^{\text{leak}}(\boldsymbol{\xi})|$ typically attains values greater than 1 at frequencies above the cut-off $|\boldsymbol{\xi}| \geq \text{f}\Delta$, see fig. 9.6. This means that the bound (9.4.22) indeed permits *negative* contrast in certain Fourier-frequencies. While this is certainly counter-intuitive from a physical point-of-view, one has to cope with this peculiarity in order to make sense of the stability estimates.

Since $|\hat{p}^{\text{leak}}|$ is typically much smaller than 1 for low frequencies (compare fig. 9.6), the right-hand side of (9.4.22) will be positive for objects h whose Fourier-transform $\mathcal{F}(h)$ is sufficiently localized at low frequencies. Accordingly, a natural candidate for a class of functions that can be stably recovered from $\mathcal{D}(h)|_K$ would be *band-limited* ones, such that $\mathcal{F}(h)$ vanishes above a certain maximum frequency, including all parts of the Fourier domain where $1 - |\hat{p}^{\text{leak}}|^2$ is negative. Importantly, however, corollary 9.13 also assumes h to have compact support in real-space so that $\mathcal{F}(h)$ is an entire function and thus cannot vanish in any open set $U \subset \mathbb{R}^m$ unless $\mathcal{F}(h)$, and hence h , is identically zero.

In general, we see that determining a stable class of objects naturally involves the classical problem of finding functions that are well-localized in real-space and Fourier-space at the same time, governed by so-called *uncertainty principles*. See e.g. [185, 68] for reviews on this topic. As a solution, we will restrict to objects given by B-splines, which may have a compact support *and* will be shown to be *essentially* band-limited in a suitable quantitative manner.

9.5 Stability estimates for spline objects

In the following, we derive stability results for objects given by multi-variate B-splines, which can be regarded as images of *finite resolution*. Such a restriction also makes sense from an

experimentalist's point-of-view as the finite number of detector-pixels introduce a natural discretization in any real-world XPCI setup.

9.5.1 Multi-variate B-splines

As a model for discretized, i.e. *pixelated* images, we consider spaces of j -th order multi-variate B-splines: for a fixed *resolution* $1/r$ with $r > 0$ and origin $\mathbf{o} \in [0; 1)^m$, we arrange nodes on a uniform Cartesian grid in $\mathbb{G}_{r,\mathbf{o}}^m := \{\mathbf{o} + r\mathbf{j} : \mathbf{j} \in \mathbb{Z}^m\} \subset \mathbb{R}^m$: Now we define objects as linear combinations of basis-splines centered at these nodes:

$$\mathfrak{B}_{k,r,\mathbf{o}}^m := \left\{ h : \mathbf{x} \mapsto \sum_{\mathbf{j} \in \mathbb{Z}^m} b_{\mathbf{j}} B_k^m(\mathbf{x}/r - \mathbf{j} - \mathbf{o}) : (b_{\mathbf{j}}) \in \ell^2(\mathbb{Z}^m) \right\} \quad (9.5.1a)$$

$$B_k^m(x_1, \dots, x_m) := \prod_{j=1}^m B_k(x_j), \quad B_k = \begin{cases} B_0 * B_{k-1} & \text{for } k \in \mathbb{N} \\ \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2})} & \text{for } k = 0 \end{cases} \quad (9.5.1b)$$

For details and explicit formulas of B-splines, see for example [197, 196]. For our purposes here it is sufficient to note that $\text{supp}(B_k^m) = [-\frac{k+1}{2}, \frac{k+1}{2}]^m$ and $B_k^m \in \mathcal{C}^{k-1}(\mathbb{R}^m)$ for $k \geq 1$.

9.5.1.1 Approximation properties

Splines interpolate values assigned on the grid nodes: for any sequence $(y_{\mathbf{j}}) \in \ell^2(\mathbb{Z}^m)$ and $k \in \mathbb{N}_0$, there exists a unique spline $h \in \mathfrak{B}_{k,r,\mathbf{o}}^m$ such that $h(r\mathbf{j} + \mathbf{o}) = y_{\mathbf{j}}$ for all $\mathbf{j} \in \mathbb{Z}^m$ and the map $(y_{\mathbf{j}}) \mapsto h$ is continuous from $\ell^2(\mathbb{Z}^m)$ to $L^2(\mathbb{R}^m)$. This is related to the fact that B-splines form a *Riesz sequence* [38]:

$$r^{m/2} C_{\text{Riesz},k}^m \|(b_{\mathbf{j}})\|_{\ell^2(\mathbb{Z}^m)} \leq \|h\| \leq r^{m/2} \|(b_{\mathbf{j}})\|_{\ell^2(\mathbb{Z}^m)} \quad (9.5.2)$$

for some constants $C_{\text{Riesz},k} > 0$ and all $h = \sum_{\mathbf{j} \in \mathbb{Z}^m} b_{\mathbf{j}} B_k^m(\cdot/r - \mathbf{j} - \mathbf{o}) \in \mathfrak{B}_{k,r,\mathbf{o}}^m$. The Riesz-sequence-property ensures stability of the approximation of functions by B-splines.

9.5.1.2 Separability

According to their definition in (9.5.1), B-splines exhibit a separable structure: for any $h \in \mathfrak{B}_{k,r,\mathbf{o}}^m$, $1 \leq j \leq m$ and $\mathbf{x}_{<j} \in \mathbb{R}^{j-1}$, $\mathbf{x}_{>j} \in \mathbb{R}^{m-j}$ fixed, it holds that $h(\mathbf{x}_{<j}, \mathbf{x}_{>j}) : x_j \mapsto h(\mathbf{x}_{<j}, x_j, \mathbf{x}_{>j}) \in \mathfrak{B}_{k,r,\mathbf{o}}^1$. In other words, multi-variate B-splines are one-dimensional B-splines along each coordinate dimension.

9.5.2 Quasi-band-limitation of B-splines

Our interest in B-splines is mainly due to their property of being quasi band-limited. As the following estimate of this quasi-band-limitation is slightly off-topic and lengthy to derive, its proof is given in appendix 9.B.

Theorem 9.14 (Quasi-band-limitation of univariate B-splines). *Let $k \in \mathbb{N}_0$, $r > 0$, $\Xi_r := [-\frac{\pi}{r}; \frac{\pi}{r}]$ and $\nu \geq 1$. Then it holds that*

$$\|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\| \leq C_{\text{band}}(k, \nu) \|\mathcal{F}(h)\| \quad \text{for all } h \in \mathfrak{B}_{k,r,\mathbf{o}}^1, \quad (9.5.3)$$

where the constant $C_{\text{band}}(k, \nu) < 1$ is given by

$$C_{\text{band}}(k, \nu)^2 = \frac{c_{\text{band}}(k, \nu)}{1 + c_{\text{band}}(k, \nu)}, \quad c_{\text{band}}(k, \nu) = c_{\text{band},0}(k, \nu) + \sum_{n=\lceil(\nu-1)/2\rceil}^{\infty} \frac{2}{(2n+1)^{2(k+1)}} \\ c_{\text{band},0}(k, \nu) = \max \left\{ \frac{\max\{\tilde{\nu}, 0\}^{2(k+1)}}{(\nu + 2\tilde{\nu})^{2(k+1)}} + \frac{\max\{\tilde{\nu}, 0\}^{2(k+1)}}{\nu^{2(k+1)}} - \frac{1}{\bar{\nu}^{2(k+1)}}, 0 \right\} \quad (9.5.4)$$

where $\lceil \cdot \rceil$ is the ‘‘round up’’-operation, $\bar{\nu} := 1 + 2\lceil(\nu - 1)/2\rceil$ and $\tilde{\nu} := \bar{\nu} - \nu - 1$.

Conversely, for any $\nu < 1$, there exists an $h \in \mathfrak{B}_{k,r,\mathbf{o}}^1$ such that $\mathcal{F}(h)|_{(\nu\Xi_r)^c} = \mathcal{F}(h)$, i.e. no estimate of the form (9.5.3) can hold true for any constant $C_{\text{band}}(k, \nu) < 1$.

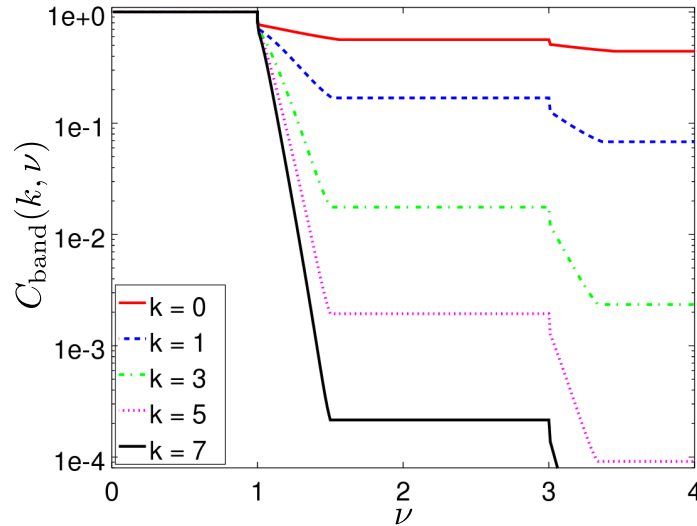


Figure 9.7. Semi-logarithmic plot of the quasi-band-limitation constant $C_{\text{band}}(k, \nu)$ from theorem 9.14 for different spline-orders k .

The constant $C_{\text{band}}(k, \nu)$ in (9.5.3) may be readily evaluated by computing the infinite series in (9.5.4) via known analytical formulas. In fig. 9.7, $C_{\text{band}}(k, \nu)$ is plotted against ν for different spline-orders $k = 0, 1, 3, 5, 7$. It can be seen that the bound $C_{\text{band}}(k, \nu)$ drops *discontinuously* from 1 to $\approx 2^{-1/2}$ at $\nu = 1$ and then decreases exponentially until $\nu \approx 1.5$, where the decrease is sharper for higher spline-orders k . For $\nu \in [1.5; 3]$, the value of $C_{\text{band}}(k, \nu)$ stagnates before it continues to decrease within the interval $[3; 3.5]$ and so on.

By exploiting the separable structure of B-splines discussed in §9.5.1.2, the 1D-result in theorem 9.14 may be easily generalized to higher dimensions:

Theorem 9.15 (Quasi-band-limitation of multivariate B-splines). *Let $k \in \mathbb{N}_0$, $r > 0$, $\nu \geq 1$, $\Xi_r := [-\frac{\pi}{r}; \frac{\pi}{r}]^m$ and $\Xi_{r,j} := \mathbb{R}^{j-1} \times [-\frac{\pi}{r}; \frac{\pi}{r}] \times \mathbb{R}^{m-j}$. Then it holds that*

$$\|\mathcal{F}(h)|_{(\nu\Xi_{r,j})^c}\| \leq C_{\text{band}}(k, \nu) \|\mathcal{F}(h)\| \quad \text{for all } 1 \leq j \leq m, \quad (9.5.5)$$

$$\|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\| \leq (1 - (1 - C_{\text{band}}(k, \nu)^2)^m)^{\frac{1}{2}} \|\mathcal{F}(h)\| \quad \text{for all } h \in \mathfrak{B}_{k,r,\mathbf{o}}^m \quad (9.5.6)$$

9.5.3 Stability estimates

In the language of regularization theory, the transition to finitely sampled B-spline objects corresponds to imposing a (very strong) *source condition*. Similarly as proven in [2, 20] for other severely ill-posed problems, such a “finite-resolution” source condition enables Lipschitz-stability estimates for image reconstruction from truncated Fresnel-data. This is seen by combining the quasi-band-limitation results from §9.5.2 with the leakage estimates from §9.4.3:

Theorem 9.16 (Stability estimate for spline-objects). *Let $K = [-\frac{1}{2}; \frac{1}{2}]^m$ and $\Omega = [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta]^m$ for $0 < \Delta < \frac{1}{2}$. Let $\mathbf{f}_\Delta := \Delta^2 \mathbf{f}$ and $\mathbf{f}_r := r^2 \mathbf{f}$ denote the Fresnel numbers associated with the length-scales Δ and r , respectively (compare §9.2.1.4). Furthermore, let $\nu \geq 1$ and $\Xi := [-\nu\pi/\mathbf{f}_r^{1/2}; \nu\pi/\mathbf{f}_r^{1/2}]$. Then it holds that*

$$\|\mathcal{D}(h)|_K\| \geq C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)^m \|h\| \quad \text{for all } h \in \mathfrak{B}_{k,r,\mathbf{o}}^m \cap L^2(\Omega). \quad (9.5.7)$$

With $\tilde{\eta}_{\mathbf{f}_\Delta}$ as defined in theorem 9.12, the constant is given by

$$\begin{aligned} C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)^2 &= 1 - C_{\text{low}}^2 - C_{\text{band}}(k, \nu)^2 (C_{\text{tot}}^2 - C_{\text{low}}^2) \\ C_{\text{low}} &:= \max_{x \in \Xi} \tilde{\eta}_{\mathbf{f}_\Delta}(x), \quad C_{\text{tot}} := \max_{x \in \mathbb{R}} \tilde{\eta}_{\mathbf{f}_\Delta}(x) \end{aligned} \quad (9.5.8)$$

Proof. We first prove the claim for $m = 1$, i.e. let $h \in \mathfrak{B}_{k,r,\mathbf{o}}^1 \cap L^2(\Omega)$ for $K = [-\frac{1}{2}; \frac{1}{2}]$ and $\Omega = [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta]$. By theorem 9.12, it then holds that

$$\|\mathcal{D}(h)|_{K^c}\|^2 \leq \langle \mathcal{F}(h), |p_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}|^2 \cdot \mathcal{F}(h) \rangle, \quad \hat{p}_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}(\xi) = \tilde{\eta}_{\mathbf{f}_\Delta}(\xi/\mathbf{f}^{\frac{1}{2}}), \quad \xi \in \mathbb{R} \quad (9.5.9)$$

Let $\Xi_r := (\mathbf{f}^{\frac{1}{2}}/\nu) \cdot \Xi = [-\pi/r; \pi/r]$. Then we have by definition of the constants $C_{\text{low}}, C_{\text{tot}}$

$$\max_{\xi \in (\nu\Xi_r)} |p_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}(\xi)| = \max_{\xi \in (\nu\Xi_r)} \tilde{\eta}_{\mathbf{f}_\Delta}(\xi/\mathbf{f}^{\frac{1}{2}}) = \max_{x \in \Xi} \tilde{\eta}_{\mathbf{f}_\Delta}(x) = C_{\text{low}} \quad (9.5.10)$$

$$\max_{\xi \in \mathbb{R}} |p_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}(\xi)| = \max_{x \in \mathbb{R}} \tilde{\eta}_{\mathbf{f}_\Delta}(x) = C_{\text{tot}} \quad (9.5.11)$$

By combining these bounds with the estimate (9.5.9), we obtain

$$\begin{aligned} \|\mathcal{D}(h)|_{K^c}\|^2 - C_{\text{low}}^2 \|\mathcal{F}(h)\|^2 &\leq \langle \mathcal{F}(h), \underbrace{(|p_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}|^2 - C_{\text{low}}^2)}_{\leq 0 \text{ in } \nu\Xi_r} \cdot \mathcal{F}(h) \rangle \\ &\leq \langle \mathcal{F}(h)|_{(\nu\Xi_r)^c}, (|p_{\mathbf{f}, \mathbf{f}_\Delta}^{\text{leak}}|^2 - C_{\text{low}}^2) \cdot \mathcal{F}(h)|_{(\nu\Xi_r)^c} \rangle \leq (C_{\text{tot}}^2 - C_{\text{low}}^2) \|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\|^2 \end{aligned} \quad (9.5.12)$$

Since $h \in \mathfrak{B}_{k,r,\mathbf{o}}^1$ and $\nu \geq 1$, $\|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\|$ can be bounded via the quasi-band-limitation theorem 9.14: $\|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\| \leq C_{\text{band}}(k, \nu) \|\mathcal{F}(h)\|$. Inserting this bound into (9.5.12) yields

$$\begin{aligned} \|\mathcal{D}(h)|_K\|^2 &= \|\mathcal{D}(h)\|^2 - \|\mathcal{D}(h)|_{K^c}\|^2 = \|\mathcal{F}(h)\|^2 - \|\mathcal{D}(h)|_{K^c}\|^2 \\ &= \|\mathcal{F}(h)\|^2 - C_{\text{low}}^2 \|\mathcal{F}(h)\|^2 - (\|\mathcal{D}(h)|_{K^c}\|^2 - C_{\text{low}}^2 \|\mathcal{F}(h)\|^2) \\ &\geq (1 - C_{\text{low}}^2 - C_{\text{band}}(k, \nu)^2 (C_{\text{tot}}^2 - C_{\text{low}}^2)) \|\mathcal{F}(h)\|^2 = C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)^2 \|h\|^2. \end{aligned} \quad (9.5.13)$$

Extension to $m > 1$: The result may be generalized to higher dimensions by exploiting the separability of the Fresnel-propagator (P4), of multi-variate B-splines and of the domains Ω and K : if we set $K_j := \mathbb{R}^{j-1} \times [-\frac{1}{2}; \frac{1}{2}] \times \mathbb{R}^{m-j}$ and $\Omega_j := \mathbb{R}^{j-1} \times [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta] \times \mathbb{R}^{m-j}$ and factorize the Fresnel propagator $\mathcal{D} = \mathcal{D}_m \dots \mathcal{D}_1$, then we have $K = \bigcap_{j=1}^m K_j$, $\Omega = \bigcap_{j=1}^m \Omega_j$ and the restriction to K_j commutes with \mathcal{D}_i for any $i \neq j$. Thus,

$$\begin{aligned} \mathcal{D}(h)|_K &= (\dots (\mathcal{D}_m \dots \mathcal{D}_1(h)|_{K_1})|_{K_2} \dots)|_{K_m} = \mathcal{D}_m(\dots \mathcal{D}_2(\mathcal{D}_1(h)|_{K_1})|_{K_2} \dots)|_{K_m} \\ &= \mathcal{D}_m(h_m)|_{K_m} \quad \text{with} \quad h_j := \mathcal{D}_{j-1}(\dots \mathcal{D}_2(\mathcal{D}_1(h)|_{K_1})|_{K_2} \dots)|_{K_{j-1}}. \end{aligned} \quad (9.5.14)$$

Moreover, as the operators $T_j : f \mapsto \mathcal{D}_j(f)|_{K_j}$ act only along the j -th coordinate dimension and since $h \in \mathfrak{B}_{k,r,\mathbf{o}}^m$ with $\text{supp}(h) \subset \Omega$, it holds that $\text{supp}(h_j) \subset \Omega_j$ for all $1 \leq j \leq m$ and h_j is a 1D B-spline when restricted to the j -th coordinate dimension (compare §9.5.1.2). This implies that we may bound expressions of the form $\mathcal{D}_j(h_j)|_{K_j}$ using the bound for $m = 1$ dimensions derived above. By recursive application of this argument, we arrive at

$$\begin{aligned} \|\mathcal{D}(h)|_K\| &= \|\mathcal{D}_m(h_m)|_{K_m}\| \geq C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu) \|h_m\| \\ &= C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu) \|\mathcal{D}_{m-1}(h_{m-1})|_{K_{m-1}}\| \geq \dots \geq C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)^m \|h\|. \end{aligned} \quad (9.5.15)$$

□

9.5.4 Application: resolution estimates

The stability estimate in theorem 9.16 can be used to verify that an imaging setup allows for a certain resolution at a realistic noise level within the setting of inverse problem inverse problem 9.1(a). We can address to types of questions:

- 1 For a fixed (spline-)resolution $1/r$, how stable is the reconstruction within a square object-domain Ω depending on its distance Δ to the detector boundary ∂K ?
- 2 If we require a stability estimate $\|\mathcal{D}(h)|_K\| \geq C\|h\|$ with some minimal contrast $C \in (0; 1)$, what resolution $1/r$ can be achieved depending on d ?

We illustrate this for an exemplary setting in $m = 2$ dimensions with square detector $K = [-\frac{1}{2}; \frac{1}{2}]^2$, Fresnel number $\mathbf{f} = 10^4$ and splines of order $k = 7$.

For setting 1, let us examine how stably features of size $r = 1/500$ can be reconstructed. We compute values of the stability-constant $C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)$ for different $0 < \Delta < \frac{1}{2}$, $\mathbf{f}_\Delta = \Delta^2 \mathbf{f}$ and a suitable ν (here, we choose $\nu = 1.2$ fixed but note that, in principle, one could optimize over this parameter as the bound (9.5.7) holds for *all* $\nu \geq 1$). For any point $\mathbf{x} \in K$, we can then express the *local* stability of the reconstruction at a point \mathbf{x} as

$$\begin{aligned} c_{\text{stab},r}(\mathbf{x}) &:= \sup \{ C_{\text{stab}}(\mathbf{f}_\Delta, r^2 \mathbf{f}, k, \nu) : \mathbf{x} \in [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta]^m \} \\ &= C_{\text{stab}}(\text{dist}(\mathbf{x}, \partial K)^2 \mathbf{f}, r^2 \mathbf{f}, k, \nu) \quad (\partial K: \text{detector-boundary}) \end{aligned} \quad (9.5.16)$$

The resulting values of $c_{\text{stab},r}(\mathbf{x})$ are plotted in fig. 9.8(a). It can be seen that $c_{\text{stab},r}(\mathbf{x}) = 0$, indicating instability, holds true up to $\text{dist}(\mathbf{x}, \partial K) \gtrsim \pi/(\mathbf{f}r)$ and then increases very quickly to

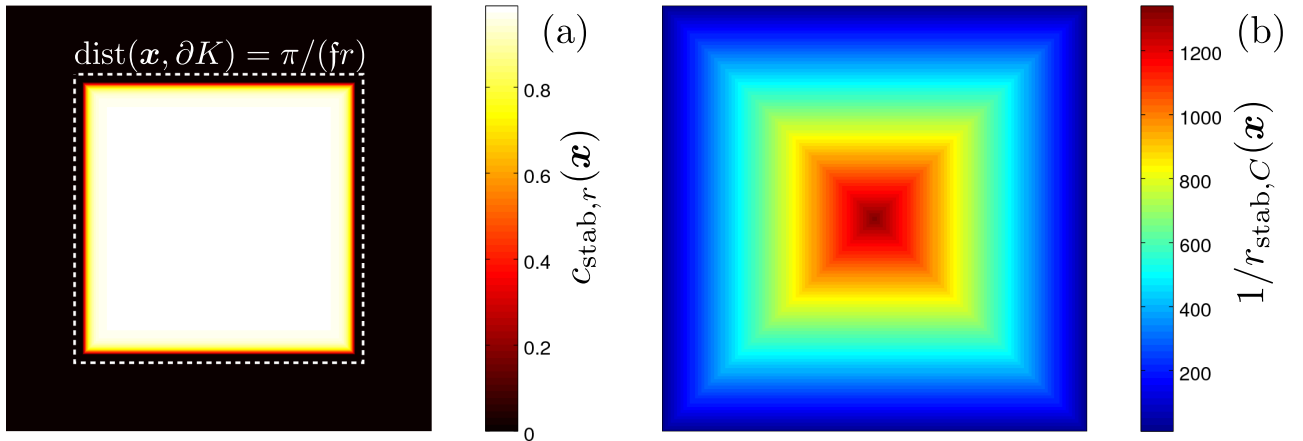


Figure 9.8. (a) Local stability of the reconstruction in inverse problem 9.1(a) for a square FoV $K = [-\frac{1}{2}; \frac{1}{2}]^2$ and $\mathfrak{f} = 10^4$, according to (9.5.16) and theorem 9.16 for $k = 7$, $\nu = 1.2$. The dashed line bounds the region that is expected to be stable according to the wave-packet-analysis in §9.3.3.1. (b) Plot of the stably achievable local resolution computed via (9.5.17) for $C = 1/4$.

a value close 1 for larger distances to the detector-boundary. These results are in very good agreement with the resolution estimates from the analysis of Gaussian wave-packets in §9.3.

For problems of the type 2, we can use (9.5.16) to express the stably achievable resolution:

$$r_{\text{stab},C}(\mathbf{x}) := \inf \{r > 0 : c_{\text{stab},r}(\mathbf{x}) \geq C\} \quad (9.5.17)$$

Numerically computed values of $1/r_{\text{stab},C}(\mathbf{x})$ for $C = 1/4$ are plotted in fig. 9.8(b). The plot turns out to be practically identical to 9.4(a), up to slightly lower resolutions by a global factor of about 1.2. In other words, the *worst-case* resolution estimates of the present section are very close to the possibly optimistic bounds derived in §9.3.3.1.

9.6 Improved estimates for real-valued objects

9.6.1 Quasi-symmetric propagation principle

In the preceding sections, we have derived locality- and stability estimates for Fresnel-propagation in terms of essentially two ingredients: smoothness, i.e. a finite resolution, and distance to the detector-boundary. Moreover, as both best-case- and worst-case-stability has been considered, these ingredients have been shown to be both *necessary and sufficient*! In §9.3.3.2, however, it has been found that the reconstruction of *real-valued* images is subject to much less severe resolution limits, based on the observation that real-valued Gaussian wave-packets propagate *symmetrically* upon Fresnel propagation.

Clearly, the observed behavior of wave-packets could be just a peculiarity of the considered, very special class of functions. Yet, *quasi*-symmetric propagation of real-valued signals turns out to be a general principle, that is closely related to the characteristic symmetry-properties of their Fourier transforms: for any $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}$, $\mathcal{F}(\varphi)$ is a *Hermitian* function, i.e. it holds that $\mathcal{F}(\varphi)(-\boldsymbol{\xi}) = \overline{\mathcal{F}(\varphi)(\boldsymbol{\xi})}$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$. We use this property via the following lemma:

Lemma 9.17. *Let $\varphi \in L^2(\mathbb{R}^m, \mathbb{R})$ be real-valued and $\hat{p} \in L^\infty(\mathbb{R}^m)$. Then it holds that*

$$\|\hat{p} \cdot \mathcal{F}(\varphi)\| = \|\text{sym}(\hat{p}) \cdot \mathcal{F}(\varphi)\|, \quad (9.6.1)$$

where $\text{sym}(\hat{p})$ is defined by $\text{sym}(\hat{p})(\boldsymbol{\xi}) := 2^{-1/2}(|\hat{p}(\boldsymbol{\xi})|^2 + |\hat{p}(-\boldsymbol{\xi})|^2)^{1/2}$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$.

Proof. As φ is real-valued, $|\mathcal{F}(\varphi)(-\boldsymbol{\xi})| = |\overline{\mathcal{F}(\varphi)(\boldsymbol{\xi})}| = |\mathcal{F}(\varphi)(\boldsymbol{\xi})|$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$. Thus,

$$\begin{aligned} 2\|\hat{p} \cdot \mathcal{F}(\varphi)\|^2 &= 2 \int_{\mathbb{R}^m} |\hat{p}(\boldsymbol{\xi})|^2 |\mathcal{F}(\varphi)(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} = \int_{\mathbb{R}^m} |\hat{p}(\boldsymbol{\xi})|^2 (|\mathcal{F}(\varphi)(\boldsymbol{\xi})|^2 + |\mathcal{F}(\varphi)(-\boldsymbol{\xi})|^2) \, d\boldsymbol{\xi} \\ &= \int_{\mathbb{R}^m} |\hat{p}(\boldsymbol{\xi})|^2 |\mathcal{F}(\varphi)(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} + \int_{\mathbb{R}^m} |\hat{p}(\boldsymbol{\xi})|^2 |\mathcal{F}(\varphi)(-\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \\ &= \int_{\mathbb{R}^m} (|\hat{p}(\boldsymbol{\xi})|^2 + |\hat{p}(-\boldsymbol{\xi})|^2) |\mathcal{F}(\varphi)(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} = 2 \int_{\mathbb{R}^m} |\text{sym}(\hat{p})(\boldsymbol{\xi})|^2 |\mathcal{F}(\varphi)(\boldsymbol{\xi})|^2 \, d\boldsymbol{\xi} \\ &= 2\|\text{sym}(\hat{p}) \cdot \mathcal{F}(\varphi)\|^2. \end{aligned} \quad (9.6.2)$$

□

Despite its simplicity, lemma 9.17 enables us to prove a surprisingly general result on the propagation of real-valued signals:

Theorem 9.18 (Quasi-symmetric propagation of real-valued signals). *For $a \in \mathbb{R}$ and $\mathbf{n} \in \mathbb{S}^{m-1}$, let $K := \{\mathbf{x} \in \mathbb{R}^m : \mathbf{n} \cdot \mathbf{x} \geq a\} \subset \mathbb{R}^m$ be a half-space. Then it holds that*

$$\|\mathcal{D}(\varphi)|_{K^c}\| \leq C_{\text{sym}} \|\varphi\| \quad \text{for all real-valued } \varphi \in L^2(K, \mathbb{R}) \quad (9.6.3)$$

with a universal constant $C_{\text{sym}} < 1$, independent of \mathbf{f} , m , a , \mathbf{n} and φ , that is bounded by

$$C_{\text{sym}} \leq \max_{x \in \mathbb{R}} \text{sym}(\tilde{\theta})(x) \leq 0.837. \quad (9.6.4)$$

On the contrary, for general, complex-valued signals $\varphi \in L^2(K, \mathbb{C})$, no bound of the form (9.6.3) may hold true for any $C_{\text{sym}} < 1$.

Proof. Let $\varphi \in L^2(K, \mathbb{R})$ be real-valued. If we set $P_{\text{leak}} := \{\mathbf{x} \in \mathbb{R}^m : \mathbf{n} \cdot \mathbf{x} \geq 0\}$, theorem 9.10 is applicable and we obtain by (9.4.7)

$$\|\mathcal{D}(\varphi)|_{K^c}\| \leq \|\hat{p}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| \quad (9.6.5)$$

where the filter is given by $\hat{p}^{\text{leak}}(\boldsymbol{\xi}) = \mathcal{D}(\mathbf{1}_{P_{\text{leak}}^c})(\boldsymbol{\xi}/\mathbf{f}) = \tilde{\theta}(-\mathbf{n} \cdot \boldsymbol{\xi}/\mathbf{f}^{1/2})$ by the analytical propagation-formula (9.4.12) for half-spaces. By lemma 9.17, it follows that

$$\|\hat{p}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| = \|\text{sym}(\hat{p}^{\text{leak}}) \cdot \mathcal{F}(\varphi)\| \leq \left(\max_{\boldsymbol{\xi} \in \mathbb{R}^m} |\text{sym}(\hat{p}^{\text{leak}})(\boldsymbol{\xi})| \right) \cdot \|\mathcal{F}(\varphi)\| \quad (9.6.6)$$

The result can be further simplified by using that \hat{p}^{leak} varies only along the axis \mathbf{n} :

$$\begin{aligned} \max_{\boldsymbol{\xi} \in \mathbb{R}^m} |\text{sym}(\hat{p}^{\text{leak}})(\boldsymbol{\xi})| &= 2^{-\frac{1}{2}} \max_{\boldsymbol{\xi} \in \mathbb{R}^m} (|\tilde{\theta}(-\mathbf{n} \cdot \boldsymbol{\xi}/f^{1/2})|^2 + |\tilde{\theta}(\mathbf{n} \cdot \boldsymbol{\xi}/f^{1/2})|^2)^{\frac{1}{2}} \\ &= 2^{-\frac{1}{2}} \max_{x \in \mathbb{R}} (|\tilde{\theta}(-x)|^2 + |\tilde{\theta}(x)|^2)^{\frac{1}{2}} = \max_{x \in \mathbb{R}} \text{sym}(\tilde{\theta})(x). \end{aligned} \quad (9.6.7)$$

Combining (9.6.5), (9.6.6) and (9.6.7) yields the first assertion.

Now let us drop the assumption of real-valuedness, i.e. let $0 \neq h \in L^2(K, \mathbb{C})$ be arbitrary. By lemma 9.6, the propagated intensity $|\mathcal{D}(h)|^2$ may be shifted in arbitrary directions and arbitrarily far by replacing h with $\tilde{h} : \mathbf{x} \mapsto \exp(i\boldsymbol{\xi} \cdot \mathbf{x})h(\mathbf{x})$ for a suitable $\boldsymbol{\xi} \in \mathbb{R}^m$, while one still has $\tilde{h} \in L^2(K)$ with $\|\tilde{h}\| = \|h\|$. By this shifting-mechanism, one may thus construct \tilde{h} for which $|\mathcal{D}(\tilde{h})|^2$ is arbitrarily concentrated in K^c , i.e. $\|\mathcal{D}(\tilde{h})|_{K^c}\|/\|\tilde{h}\|$ may be arbitrarily close to 1. Hence, no non-trivial bound of the form (9.6.3) may hold for complex-valued signals. \square

Theorem 9.18 states that – independent of any smoothness constraints (!) – only a limited fraction of a real-valued signal may propagate out of its support in a single direction. As is also stated in the theorem, this situation is unique to the real-valued case. We note that the analytical estimate for the constant C_{sym} is not optimal:

Remark 9.19 (Optimal value of C_{sym}). Numerical eigenvalue computations (not shown) indicate that the optimal value of the symmetric-propagation constant is given by $C_{\text{sym}} \approx 0.721$. Accordingly, at most a fraction of $0.721^2 \approx 0.52$ of the intensity of a real-valued signal may leak out of the field-of-view along a single direction.

9.6.2 Construction of improved leakage bounds

Next, we extend the quasi-symmetric propagation bound in theorem 9.18 from half-spaces to the more practically relevant case of square FoVs. In such a setting, the propagated signal may always leak out of the detection domain along two opposite directions so that (quasi-)symmetric propagation alone may not guarantee finite leakage. Instead, we have to combine the latter principle with the detector-distance-based leakage estimates of the preceding sections. The idea is simple: along each direction, we can decompose an object-signal into a part with support close to the detector-boundary ∂K , to be bounded by exploiting quasi-symmetric propagation, and another part that is concentrated far away from ∂K and which thus can be bounded using the theory from §9.4. We first prove such a bound for half-spaces:

Lemma 9.20. *For $\mathbf{n} \in \mathbb{S}^{m-1}$, $a \in \mathbb{R}$ and $\Delta > 0$, let $K := \{\mathbf{x} \in \mathbb{R}^m : \mathbf{n} \cdot \mathbf{x} \geq a\}$, $\Omega = K$ and $\Omega_{\leq \Delta} := \Omega \cap \{\mathbf{x} \in \mathbb{R}^m : \mathbf{n} \cdot \mathbf{x} \leq a + \Delta\}$. Then it holds that*

$$\|\mathcal{D}(\varphi)|_{K^c}\| \leq 2^{-\frac{1}{2}} \|\hat{p}_{f, f_{\Delta}}^{\text{leak}}(\mathbf{n} \cdot (\cdot)) \cdot \mathcal{F}(\varphi)\| + C_{\text{sym}, f_{\Delta}} \|\varphi|_{\Omega_{\leq \Delta}}\| \quad \text{for all } \varphi \in L^2(\Omega, \mathbb{R}) \quad (9.6.8)$$

with $f_{\Delta} = \Delta^2 f$ and $\hat{p}_{f, f_{\Delta}}^{\text{leak}}$ as in theorem 9.12 and the constant $C_{\text{sym}, f_{\Delta}}$ is given by

$$C_{\text{sym}, f_{\Delta}} := \max_{x \in \mathbb{R}} \text{sym}(\tilde{\theta}_{f_{\Delta}})(x), \quad \tilde{\theta}_{f_{\Delta}}(x) := \tilde{\theta}(x) - \tilde{\theta}(x - f_{\Delta}^{\frac{1}{2}}). \quad (9.6.9)$$

Proof. By separability (P4) and isotropy (P5), it is sufficient to prove the claim for the 1D-setting $m = 1$, $\mathbf{n} = 1$, $a = 0$, $K = \Omega = [0; \infty)$ and $\Omega_{\leq \Delta} = [0; \Delta]$.

Thus, let $\varphi \in L^2(K, \mathbb{R})$ be arbitrary. We follow a similar approach as in theorem 9.10: using the convolution-form of the Fresnel-propagator (P2), we obtain

$$\begin{aligned} \mathcal{D}(\varphi)|_{K^c} &= (k_{\mathfrak{f}} * \varphi)|_{K^c} = ((k_{\mathfrak{f}}|_{(-\infty; -\Delta)} + k_{\mathfrak{f}}|_{[-\Delta; 0]} + k_{\mathfrak{f}}|_K) * \varphi)|_{K^c} \\ &= (k_{\mathfrak{f}}|_{(-\infty; -\Delta)} * \varphi)|_{K^c} + (k_{\mathfrak{f}}|_{[-\Delta; 0]} * \varphi)|_{K^c} + (k_{\mathfrak{f}}|_K * \varphi)|_{K^c} \end{aligned} \quad (9.6.10)$$

Now we decompose φ into left-hand- and right-hand parts, $\varphi = \varphi_\ell + \varphi_r$ with $\varphi_\ell := \varphi|_{\Omega_{\leq \Delta}}$, $\varphi_r := \varphi|_{\Omega_{> \Delta}^c}$. By standard results on the support of convolutions, we then have

$$\begin{aligned} \text{supp}(k_{\mathfrak{f}}|_K * \varphi) &\subset \text{supp}(k_{\mathfrak{f}}|_K) + \text{supp}(\varphi) \subset K + K = K \\ \text{supp}(k_{\mathfrak{f}}|_{[-\Delta; 0]} * \varphi_r) &\subset \text{supp}(k_{\mathfrak{f}}|_{[-\Delta; 0]}) + \text{supp}(\varphi_r) \subset [-\Delta; 0] + [\Delta; \infty) = K \end{aligned}$$

Together with (9.6.10), this implies that $\mathcal{D}(\varphi)|_{K^c} = (k_{\mathfrak{f}}|_{(-\infty; -\Delta)} * \varphi)|_{K^c} + (k_{\mathfrak{f}}|_{[-\Delta; 0]} * \varphi_\ell)|_{K^c}$. An application of the triangle inequality and lemma 9.9 thus yields

$$\begin{aligned} \|\mathcal{D}(\varphi)|_{K^c}\| &\leq \|(k_{\mathfrak{f}}|_{(-\infty; -\Delta)} * \varphi)|_{K^c}\| + \|(k_{\mathfrak{f}}|_{[-\Delta; 0]} * \varphi_\ell)|_{K^c}\| \\ &\leq \|\mathcal{D}(\mathbf{1}_{(-\infty; -\Delta)})(\cdot/\mathfrak{f}) \cdot \mathcal{F}(\varphi)\| + \|\mathcal{D}(\mathbf{1}_{[-\Delta; 0]})(\cdot/\mathfrak{f}) \cdot \mathcal{F}(\varphi_\ell)\| \end{aligned} \quad (9.6.11)$$

Using the exact propagation-formulas from §9.4.3, we get

$$\begin{aligned} \mathcal{D}(\mathbf{1}_{(-\infty; -\Delta)})(\xi/\mathfrak{f}) &= \tilde{\theta}(-\xi/\mathfrak{f}^{\frac{1}{2}} - \mathfrak{f}_\Delta^{\frac{1}{2}}) \\ \mathcal{D}(\mathbf{1}_{[-\Delta; 0]})(\xi/\mathfrak{f}) &= \mathcal{D}(\mathbf{1}_{(-\infty; 0]})(\xi/\mathfrak{f}) - \mathcal{D}(\mathbf{1}_{(-\infty; -\Delta)})(\xi/\mathfrak{f}) \\ &= \tilde{\theta}(-\xi/\mathfrak{f}^{\frac{1}{2}}) - \tilde{\theta}(-\xi/\mathfrak{f}^{\frac{1}{2}} - \mathfrak{f}_\Delta^{\frac{1}{2}}) = \tilde{\theta}_{\mathfrak{f}_\Delta}(-\xi/\mathfrak{f}^{\frac{1}{2}}) \end{aligned} \quad (9.6.12)$$

Moreover, since φ and thus φ_ℓ are real-valued, lemma 9.17 is applicable. Thus,

$$\begin{aligned} \|\mathcal{D}(\mathbf{1}_{(-\infty; -\Delta)})(\cdot/\mathfrak{f}) \cdot \mathcal{F}(\varphi)\| &= \|\text{sym}(\tilde{\theta}((-\cdot)/\mathfrak{f}^{\frac{1}{2}} - \mathfrak{f}_\Delta^{\frac{1}{2}})) \cdot \mathcal{F}(\varphi)\| \stackrel{(9.4.18)}{=} 2^{-\frac{1}{2}} \|\hat{p}_{\mathfrak{f}, \mathfrak{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| \\ \|\mathcal{D}(\mathbf{1}_{[-\Delta; 0]})(\cdot/\mathfrak{f}) \cdot \mathcal{F}(\varphi_\ell)\| &= \|\tilde{\theta}_{\mathfrak{f}_\Delta}(-\cdot/\mathfrak{f}^{\frac{1}{2}}) \cdot \mathcal{F}(\varphi_\ell)\| = \|\text{sym}(\tilde{\theta}_{\mathfrak{f}_\Delta})(\cdot/\mathfrak{f}^{\frac{1}{2}}) \cdot \mathcal{F}(\varphi_\ell)\| \\ &\leq \left(\max_{x \in \mathbb{R}} \text{sym}(\tilde{\theta}_{\mathfrak{f}_\Delta})(x) \right) \|\varphi_\ell\| = C_{\text{sym}, \mathfrak{f}_\Delta} \|\varphi_\ell\| \end{aligned} \quad (9.6.13)$$

Substituting (9.6.13) into (9.6.11) now yields the assertion. \square

Note that the constant $C_{\text{sym}, \mathfrak{f}_\Delta}$ in (9.6.8) attains almost the same values as C_{sym} within the relevant regime $\mathfrak{f}_\Delta \gg 1$. Next, we extend lemma 9.20 to square detectors $K = [-\frac{1}{2}; \frac{1}{2}]^m$ by decomposing K^c into half-spaces. By far the strongest result is obtained for a 1D-case:

Theorem 9.21 (Leakage estimate for real-valued objects in 1D-intervals). *Let $m = 1$, and $\Omega = K = [-\frac{1}{2}; \frac{1}{2}]$. Then it holds that*

$$\|\mathcal{D}(\varphi)|_{K^c}\| \leq \|\hat{p}_{\mathfrak{f}, \mathfrak{f}/4}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| + C_{\text{sym}, \mathfrak{f}/4} \|\varphi\| \quad \text{for all } \varphi \in L^2(\Omega, \mathbb{R}). \quad (9.6.14)$$

Proof. Let $\varphi \in L^2(\Omega, \mathbb{R})$ be arbitrary. We decompose K^c into a left-hand and a right-hand part: $K^c = L \cup R$ with $L = (-\infty; -\frac{1}{2})$, $R = -L$. Then it holds that

$$\|\mathcal{D}(\varphi)|_{K^c}\|^2 = \|\mathcal{D}(\varphi)|_L\|^2 + \|\mathcal{D}(\varphi)|_R\|^2 \quad (9.6.15)$$

If we set $a = -\frac{1}{2}$, $d = \frac{1}{2}$, $\mathbf{n} = 1$, $K = L^c$ and $\Omega_{\leq \Delta} = [-\frac{1}{2}; 0]$, it can be seen that the assumptions of lemma 9.20 are satisfied. Thus, we obtain

$$\|\mathcal{D}(\varphi)|_L\| \leq 2^{-\frac{1}{2}} \|\hat{p}_{\mathfrak{f}, \mathfrak{f}/4}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| + C_{\text{sym}, \mathfrak{f}/4} \|\varphi_\ell\|, \quad (9.6.16)$$

where $\varphi_\ell := \varphi|_{[-\frac{1}{2}; 0]}$ denotes the left-hand part of φ . If we define $\varphi_r := \varphi|_{[0; \frac{1}{2}]}$, an analogous estimate can be obtained for the right-hand domain R :

$$\begin{aligned} \|\mathcal{D}(\varphi)|_R\| &\leq 2^{-\frac{1}{2}} \|\hat{p}_{\mathfrak{f}, \mathfrak{f}/4}^{\text{leak}}(-\cdot) \cdot \mathcal{F}(\varphi)\| + C_{\text{sym}, \mathfrak{f}/4} \|\varphi_r\| \\ &= 2^{-\frac{1}{2}} \|\hat{p}_{\mathfrak{f}, \mathfrak{f}/4}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| + C_{\text{sym}, \mathfrak{f}/4} \|\varphi_r\|, \end{aligned} \quad (9.6.17)$$

where it has been exploited that $\hat{p}_{\text{leak}, \mathfrak{f}/4}^{\text{real}}$ is an even function by definition.

Now we apply the estimates (9.6.16) and (9.6.17) to (9.6.15) and exploit that φ_ℓ and φ_r are L^2 -orthogonal so that $\|\varphi_\ell\|^2 + \|\varphi_r\|^2 = \|\varphi\|^2$ and $\|\varphi_\ell\| + \|\varphi_r\| \leq 2^{1/2} \|\varphi\|$:

$$\begin{aligned} \|\mathcal{D}(\varphi)|_{K^c}\|^2 &\leq \|\hat{p}_{\mathfrak{f}, \mathfrak{f}/4}^{\text{leak}} \cdot \mathcal{F}(\varphi)\|^2 + C_{\text{sym}, \mathfrak{f}/4}^2 \left(\|\varphi_\ell\|^2 + \|\varphi_r\|^2 \right) \\ &\quad + 2^{\frac{1}{2}} C_{\text{sym}, \mathfrak{f}/4} \|\hat{p}_{\mathfrak{f}, \mathfrak{f}/4}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| \left(\|\varphi_\ell\| + \|\varphi_r\| \right) \\ &\leq \left(\|\hat{p}_{\mathfrak{f}, \mathfrak{f}/4}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| + C_{\text{sym}, \mathfrak{f}/4} \|\varphi\| \right)^2. \end{aligned}$$

□

As seen in §9.4.3, $\hat{p}_{\mathfrak{f}, \mathfrak{f}/4}^{\text{leak}}$ acts as high-pass filter with cutoff-frequency $|\xi| \approx \mathfrak{f}/2$. Provided that $C_{\text{sym}, \mathfrak{f}/4} < 1$ is small enough, (9.6.14) thus guarantees positive contrast $\|\mathcal{D}(\varphi)|_K\| \geq (1 - C_{\text{sym}, \mathfrak{f}/4}^2 - \varepsilon^2)^{1/2} \|\varphi\|$ for some small $\varepsilon > 0$ if $\mathcal{F}(\varphi)$ is concentrated within the interval $[-\mathfrak{f}/2; \mathfrak{f}/2]$. This indicates that image reconstruction is stable down to features of size $r \gtrsim 2\pi/\mathfrak{f}$, which is already the *upper* limit for the achievable resolution by §9.3.3.1. Moreover, as the object-domain is $\Omega = K$ in theorem 9.21, this optimal resolution can be obtained in the entire FoV!

However, the surprisingly strong 1D-result does *not* carry over to higher dimensions because square detectors $K = [-\frac{1}{2}; \frac{1}{2}]^m$ for $m > 1$ have *corners*, close to which image reconstruction is unstable down to low spatial frequencies as found in §9.3.3.2. We have to exclude the considered objects from having support in these unstable regions:

Theorem 9.22 (Leakage estimate for real-valued objects in square domains). *Let $K = [-\frac{1}{2}; \frac{1}{2}]^m$ and $\Omega := \bigcup_{j=1}^m S_{\Delta, j}$ with $S_{\Delta, j} := \mathbb{R}^{j-1} \times [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta] \times \mathbb{R}^{m-j}$ for $0 < \Delta < \frac{1}{2}$. Then it holds that*

$$\|\mathcal{D}(\varphi)|_{K^c}\| \leq \|\hat{p}_{\mathfrak{f}, \mathfrak{f}\Delta}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| + C_{\text{sym}, \mathfrak{f}\Delta} \|\varphi|_{\Omega_{\leq \Delta}}\| \quad \text{for all } \varphi \in L^2(\Omega, \mathbb{R}), \quad (9.6.18)$$

where $\Omega_{\leq \Delta} := \Omega \setminus (-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta)^m$ denotes the part of Ω with distance less than Δ to ∂K .

Proof. Let $\varphi \in L^2(\Omega, \mathbb{R})$. If we define the half-spaces $K_{j,\pm} := \mathbb{R}^{j-1} \times \pm[-\frac{1}{2}; \infty) \times \mathbb{R}^{m-j}$, then it holds that $K^c = \bigcup_{j=1,\pm}^m K_{j,\pm}^c$. Thus, we have

$$\|\mathcal{D}(\varphi)|_{K^c}\|^2 \leq \sum_{j=1,\pm}^m \|\mathcal{D}(\varphi)|_{K_{j,\pm}^c}\|^2. \quad (9.6.19)$$

By construction, each of the squared norms on the right-hand side can be estimated via lemma 9.20 (with parameters $a = -\frac{1}{2}$, $\mathbf{n} = \pm \mathbf{e}_j$), yielding

$$\|\mathcal{D}(\varphi)|_{K_{j,\pm}^c}\| \leq 2^{-\frac{1}{2}} \|\hat{p}_{\mathbf{f},\mathbf{f}_\Delta,j}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| + C_{\text{sym},\mathbf{f}_\Delta} \|\varphi_{j,\pm}\|, \quad (9.6.20)$$

where we have defined $\hat{p}_{\mathbf{f},\mathbf{f}_\Delta,j}^{\text{leak}}(\boldsymbol{\xi}) := \hat{p}_{\mathbf{f},\mathbf{f}_\Delta}^{\text{leak}}(\mathbf{e}_j \cdot \boldsymbol{\xi}) = \hat{p}_{\mathbf{f},\mathbf{f}_\Delta}^{\text{leak}}(-\mathbf{e}_j \cdot \boldsymbol{\xi})$ and $\varphi_{j,\pm} := \varphi|_{\Omega_{j,\pm}}$ with $\Omega_{j,\pm} := \Omega \cap (\mathbb{R}^{j-1} \times \pm[-\frac{1}{2}; -\frac{1}{2} + \Delta] \times \mathbb{R}^{m-j})$. Inserting (9.6.20) into (9.6.19) yields

$$\begin{aligned} \|\mathcal{D}(\varphi)|_{K^c}\|^2 &\leq \frac{1}{2} \sum_{j=1,\pm}^m \|\hat{p}_{\mathbf{f},\mathbf{f}_\Delta,j}^{\text{leak}} \cdot \mathcal{F}(\varphi)\|^2 + C_{\text{sym},\mathbf{f}_\Delta}^2 \left(\sum_{j=1,\pm}^m \|\varphi_{j,\pm}\|^2 \right) \\ &\quad + 2^{\frac{1}{2}} C_{\text{sym},\mathbf{f}_\Delta} \sum_{j=1,\pm}^m \|\hat{p}_{\mathbf{f},\mathbf{f}_\Delta,j}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| \|\varphi_{j,\pm}\|. \end{aligned} \quad (9.6.21)$$

The last summand on the right-hand side of (9.6.21) can be regarded as a euclidean inner product in \mathbb{R}^{2m} . By applying Cauchy-Schwarz' inequality to this term and using that $\sum_{j=1,\pm}^m \|\hat{p}_{\mathbf{f},\mathbf{f}_\Delta,j}^{\text{leak}} \cdot \mathcal{F}(\varphi)\|^2 = 2\|\hat{p}_{\mathbf{f},\mathbf{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(\varphi)\|^2$ by (9.4.18), (9.6.21) becomes

$$\|\mathcal{D}(\varphi)|_{K^c}\|^2 \leq \left(\|\hat{p}_{\mathbf{f},\mathbf{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| + C_{\text{sym},\mathbf{f}_\Delta} \left(\sum_{j=1,\pm}^m \|\varphi_{j,\pm}\|^2 \right)^{\frac{1}{2}} \right)^2 \quad (9.6.22)$$

Now the choice of Ω ensures that the sub-domains $\{\Omega_{j,\pm}\}$ are mutually disjoint (up to intersections of measure zero). Hence, the $\{\varphi_{j,\pm}\}$ are mutually L^2 -orthogonal, which implies

$$\left(\sum_{j=1,\pm}^m \|\varphi_{j,\pm}\|^2 \right)^{\frac{1}{2}} = \|\sum_{j=1,\pm}^m \varphi_{j,\pm}\| = \|\varphi|_{(\bigcup_{j,\pm} \Omega_{j,\pm})}\| = \|\varphi|_{\Omega_{\leq \Delta}}\|. \quad (9.6.23)$$

□

9.6.3 Stability estimates for spline objects

The leakage estimates from the preceding section may be used to derive stability estimates for spline objects analogously as in §9.5.3.

Theorem 9.23 (Stability estimate for real-valued splines in intervals). *Let $\Omega = K = [-\frac{1}{2}; \frac{1}{2}]$, $k \in \mathbb{N}_0$, $r > 0$, $\nu \geq 1$, $\mathbf{f}_r = r^2 \mathbf{f}$ and $\Xi := [-\nu\pi/\mathbf{f}_r^{1/2}; \nu\pi/\mathbf{f}_r^{1/2}]$. Then it holds that*

$$\|\mathcal{D}(\varphi)|_K\| \geq C_{\text{stab}}^{\text{real},(1d)}(\mathbf{f}, \mathbf{f}_r, k, \nu) \|\varphi\| \quad \text{for all } \varphi \in \mathfrak{B}_{k,r,\mathbf{o}}^1 \cap L^2(\Omega, \mathbb{R}) \quad (9.6.24)$$

where the constant $C_{\text{stab}}^{\text{real,(1d)}}(\mathbf{f}, \mathbf{f}_r, k, \nu)$ is given by

$$\begin{aligned} C_{\text{stab}}^{\text{real,(1d)}}(\mathbf{f}, \mathbf{f}_r, k, \nu)^2 &= 1 - \left(C_{\text{sym},\mathbf{f}/4} + (C_{\text{low}}^2 + C_{\text{band}}(k, \nu)^2 (C_{\text{tot}}^2 - C_{\text{low}}^2))^{\frac{1}{2}} \right)^2 \\ C_{\text{low}} &:= \max_{x \in \Xi} \tilde{\eta}_{\mathbf{f}/4}(x) = \max_{\xi \in \mathbf{f}^{1/2}\Xi} \hat{p}_{\mathbf{f},\mathbf{f}/4}^{\text{leak}}(\xi), \quad C_{\text{tot}} := \max_{x \in \mathbb{R}} \tilde{\eta}_{\mathbf{f}/4}(x) = \max_{\xi \in \mathbb{R}} \hat{p}_{\mathbf{f},\mathbf{f}/4}^{\text{leak}}(\xi) \end{aligned} \quad (9.6.25)$$

Proof. The proof is similar to that of theorem 9.16: the setting matches the assumptions of theorem 9.21. With $\Xi_r := (\mathbf{f}^{1/2}/\nu) \cdot \Xi = [-\pi/r; \pi/r]$, the leakage bound (9.6.14) yields

$$\begin{aligned} (\|\mathcal{D}(\varphi)|_{K^c}\| - C_{\text{sym},\mathbf{f}/4}\|\varphi\|)^2 - C_{\text{low}}^2\|\varphi\|^2 &\leq \langle \mathcal{F}(\varphi), (|\hat{p}_{\mathbf{f},\mathbf{f}/4}^{\text{leak}}|^2 - C_{\text{low}}^2) \cdot \mathcal{F}(\varphi) \rangle \\ &\leq \langle \mathcal{F}(\varphi)|_{(\mathbf{f}^{1/2}\Xi)^c}, (|\hat{p}_{\mathbf{f},\mathbf{f}/4}^{\text{leak}}|^2 - C_{\text{low}}^2) \cdot \mathcal{F}(\varphi)|_{(\mathbf{f}^{1/2}\Xi)^c} \rangle \\ &\leq (C_{\text{tot}}^2 - C_{\text{low}}^2) \|\mathcal{F}(\varphi)|_{(\nu\Xi_r)^c}\|^2 \leq C_{\text{band}}(k, \nu)^2 (C_{\text{tot}}^2 - C_{\text{low}}^2) \|\varphi\|^2 \end{aligned} \quad (9.6.26)$$

for any $\varphi \in \mathfrak{B}_{k,r,\mathbf{o}}^1 \cap L^2(\Omega, \mathbb{R})$, where the quasi-band-limitation theorem 9.14 has been applied in the final step. Rearranging (9.6.26) yields

$$\begin{aligned} \|\mathcal{D}(\varphi)|_{K^c}\| &= C_{\text{sym},\mathbf{f}/4}\|\varphi\| + (\|\mathcal{D}(\varphi)|_{K^c}\| - C_{\text{sym},\mathbf{f}/4}\|\varphi\|) \\ &\leq \left(C_{\text{sym},\mathbf{f}/4} + (C_{\text{low}}^2 + C_{\text{band}}(k, \nu)^2 (C_{\text{tot}}^2 - C_{\text{low}}^2))^{\frac{1}{2}} \right) \|\varphi\| \\ &= (1 - C_{\text{stab}}^{\text{real,(1d)}}(\mathbf{f}, \mathbf{f}_r, k, \nu)^2)^{\frac{1}{2}} \|\varphi\| \end{aligned} \quad (9.6.27)$$

Since $\|\mathcal{D}(\varphi)|_K\| = (\|\varphi\|^2 - \|\mathcal{D}(\varphi)|_{K^c}\|^2)^{1/2}$, (9.6.27) proves the assertion. \square

Once more, the remarkable aspect of the 1D stability result in theorem 9.23 is that does not require any distance between the object-domain Ω and the boundary of K . Analogously, we can obtain a stability estimate for the higher-dimensional case:

Theorem 9.24 (Stability estimate for real-valued splines in square domains). *Within the setting of theorem 9.22, let $\mathbf{f}_\Delta = \Delta^2 \mathbf{f}$, $\mathbf{f}_r = r^2 \mathbf{f}$ and $\Xi := [-\nu\pi/\mathbf{f}_r^{1/2}; \nu\pi/\mathbf{f}_r^{1/2}]$. Then it holds that*

$$\|\mathcal{D}(\varphi)|_K\| \geq C_{\text{stab}}^{\text{real},m}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu) \|\varphi\| \quad \text{for all } \varphi \in \mathfrak{B}_{k,r,\mathbf{o}}^1 \cap L^2(\Omega, \mathbb{R}) \quad (9.6.28)$$

where the constant $C_{\text{stab}}^{\text{real},m}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)$ is given by

$$\begin{aligned} C_{\text{stab}}^{\text{real},m}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)^2 &= 1 - \left(C_{\text{sym},\mathbf{f}_\Delta} + m^{\frac{1}{2}} (C_{\text{low}}^2 + C_{\text{band}}(k, \nu)^2 (C_{\text{tot}}^2 - C_{\text{low}}^2))^{\frac{1}{2}} \right)^2 \\ C_{\text{low}} &:= \max_{\xi \in \Xi} \tilde{\eta}_{\mathbf{f}_\Delta}(x), \quad C_{\text{tot}} := \max_{\xi \in \mathbb{R}} \tilde{\eta}_{\mathbf{f}_\Delta}(x) \end{aligned} \quad (9.6.29)$$

Proof. Let $\varphi \in \mathfrak{B}_{k,r,\mathbf{o}}^1 \cap L^2(\Omega, \mathbb{R})$. Since $\|\varphi|_{\Omega_{<d}}\| \leq \|\varphi\|^2$, we then have by theorem 9.22:

$$\|\mathcal{D}(\varphi)|_{K^c}\| - C_{\text{sym},\mathbf{f}_\Delta}\|\varphi\| \leq \|\hat{p}_{\mathbf{f},\mathbf{f}_\Delta}^{\text{leak}} \cdot \mathcal{F}(\varphi)\| = \left(\sum_{j=1}^m \langle \mathcal{F}(\varphi), |\hat{p}_{\mathbf{f},\mathbf{f}_\Delta,j}^{\text{leak}}|^2 \cdot \mathcal{F}(\varphi) \rangle \right)^{\frac{1}{2}} \quad (9.6.30)$$

with quasi-1D functions $\hat{p}_{\mathbf{f}, \mathbf{f}_{\Delta}, j}^{\text{leak}}(\boldsymbol{\xi}) = \hat{p}_{\mathbf{f}, \mathbf{f}_{\Delta}}^{\text{leak}}(\mathbf{e}_j \cdot \boldsymbol{\xi})$ as defined in the proof of theorem 9.22. Let $\Xi_j := \mathbb{R}^{j-1} \times \Xi \times \mathbb{R}^{m-j}$ and $\Xi_{r,j} := (\mathbf{f}^{1/2}/\nu) \cdot \Xi_j = \mathbb{R}^{j-1} \times [-\pi/r; \pi/r] \times \mathbb{R}^{m-j}$ for $1 \leq j \leq m$. Then it holds that $\max_{\boldsymbol{\xi} \in (\nu \Xi_{r,j})} |\hat{p}_{\mathbf{f}, \mathbf{f}_{\Delta}, j}^{\text{leak}}(\boldsymbol{\xi})| = C_{\text{low}}$ and $\max_{\boldsymbol{\xi} \in \mathbb{R}^m} |\hat{p}_{\mathbf{f}, \mathbf{f}_{\Delta}, j}^{\text{leak}}(\boldsymbol{\xi})| = C_{\text{tot}}$ and hence, by a derivation completely analogously as in (9.6.26),

$$\begin{aligned} \langle \mathcal{F}(\varphi), |\hat{p}_{\mathbf{f}, \mathbf{f}_{\Delta}, j}^{\text{leak}}|^2 \cdot \mathcal{F}(\varphi) \rangle - C_{\text{low}}^2 \|\varphi\|^2 &\leq (C_{\text{tot}}^2 - C_{\text{low}}^2) \|\mathcal{F}(\varphi)|_{(\nu \Xi_{r,j})^c}\|^2 \\ &\leq C_{\text{band}}(k, \nu)^2 (C_{\text{tot}}^2 - C_{\text{low}}^2) \|\varphi\|^2 \end{aligned} \quad (9.6.31)$$

for all $1 \leq j \leq m$, where theorem 9.15 has been applied. Bounding the right-hand side of (9.6.30) via (9.6.31) and rearranging as in the proof of theorem 9.23 yields the assertion. \square

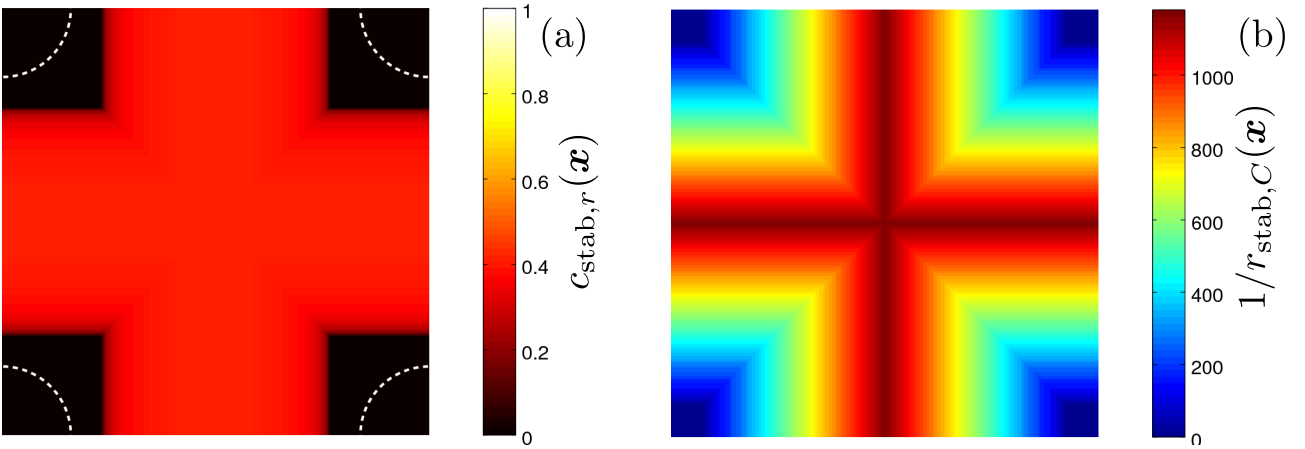


Figure 9.9. Same plot as fig. 9.8, yet for the *real-valued* setting of inverse problem 9.2(a). Local stability constant (a) and -resolution (b) have been computed according to theorem 9.24 via (9.6.32) and (9.6.33).

9.6.4 Application: resolution estimates

Analogously as for the complex-valued case in §9.5.4, we can use theorem 9.24 to assess the resolution within the *real-valued* setting of inverse problem 9.2(a).

For illustration, we consider exactly the same setting as in §9.5.4, but express the local stability constant and resolution via the improved bound (9.6.24), exploiting real-valuedness:

$$c_{\text{stab},r}(\mathbf{x}) := \sup \{ C_{\text{stab}}^{\text{real},m}(\mathbf{f}_{\Delta}, \mathbf{f}_r, k, \nu) : \mathbf{x} \in \bigcup_{j=1}^m S_{d,j} \} \quad (9.6.32)$$

$$r_{\text{stab},C}(\mathbf{x}) := \inf \{ r > 0 : c_{\text{stab},r}(\mathbf{x}) \geq C \} \quad (9.6.33)$$

with $S_{d,j}$ as defined in theorem 9.22. $c_{\text{stab},r}$ and $1/r_{\text{stab},C}$ are plotted in fig. 9.9(a),(b).

According to fig. 9.9(a), stable reconstruction is guaranteed within the entire FoV except for square-shaped neighborhoods around the corners of K . The width of the unstable region is about 1.5 times $\pi/(\mathbf{f}r)$ – the value that is to be expected from the analysis §9.3.3.2. Likewise, the local resolutions in fig. 9.9(b) are qualitatively in good agreement with the results from the wave-packet-analysis in §9.3.3.2, compare fig. 9.4(b).

9.7 Extension to the phaseless case: application to linearized XPCI

So far, the analysis has been limited to the case where the full complex-valued propagated wave field $\mathcal{D}(h)(\mathbf{x})$ – including the phase – is measured at each point $\mathbf{x} \in K$ of the FoV. In the following, we outline how the results can be extended to the case of phaseless data. We consider the inverse problems inverse problem 9.1(b) and inverse problem 9.2(b) that model image reconstruction in XPCI within the linear CTF-regime. On the contrary, analyzing the *nonlinear* problems inverse problem 9.1(c) and inverse problem 9.2(c) is beyond reach as stability is an open problem for these even in the case of a full FoV $K = \mathbb{R}^m$.

9.7.1 Leakage estimates

As a first step, we aim to bound the amount of data that is leaked outside a square field-of-view within the setting of inverse problem 9.1(b) and inverse problem 9.2(b). This is fairly simple as the measured data, $\sim 2\text{Re}(\mathcal{D}(h))$, relates to Fresnel-propagation simply by the *pointwise* real-part and $|\text{Re}(z)| \leq |z|$ for all $z \in \mathbb{C}$. This yields the following bound:

Theorem 9.25 (Leakage bound for linearized XPCI data). *Let $K \subset \mathbb{R}^m$ be measurable and $\mathcal{T}, \mathcal{S}_\nu$ be the forward maps from inverse problem 9.1(b) and inverse problem 9.2(b). Then it holds that*

$$\|\mathcal{T}(h)|_{K^c}\| \leq 2 \|\mathcal{D}(h)|_{K^c}\| \quad \text{for all } h \in L^2(\mathbb{R}^m), \quad (9.7.1)$$

$$\|\mathcal{S}_\nu(\varphi)|_{K^c}\| \leq 2 \|\mathcal{D}(\varphi)|_{K^c}\| \quad \text{for all } \varphi \in L^2(\mathbb{R}^m, \mathbb{R}). \quad (9.7.2)$$

The gist of theorem 9.25 is simple: it states that the leaked part of XPCI data, $\mathcal{T}(h)|_{K^c}$, cannot contain more information than the corresponding phased Fresnel-data $\mathcal{D}(h)|_{K^c}$. Despite its simplicity, however, this result has important consequences: by theorem 9.25, literally *any* of the leakage estimate of the preceding sections induces a bound for the phaseless case.

9.7.2 Stability estimates

Using the simple insight from theorem 9.25, we may derive stability estimates for phase contrast imaging with finite detectors. To this end, we combine leakage estimates with the stability results for XPCI with *infinite* FoVs from theorem 9.4:

Theorem 9.26 (Stability estimate for linearized XPCI with square detector). *Let $K = [-\frac{1}{2}; \frac{1}{2}]^m$ and $\Omega \subset [-\frac{1}{2} + \Delta; \frac{1}{2} - \Delta]^m$ for some $0 < \Delta < \frac{1}{2}$. Let $T \in \{\mathcal{T}, \mathcal{S}_\nu\}$ and $h \in L^2(\Omega)$, where h is assumed to be real-valued if $T = \mathcal{S}_\nu$. Furthermore, let $C_{\text{stab}}^{\text{IP}^*}(\Omega, \mathfrak{f}, (\nu))$ denote the stability constant of T for a full FoV from theorem 9.4. Then it holds that*

$$\|T(h)|_K\|^2 \geq \|T(h)\|^2 - 4 \|\hat{p}_{\mathfrak{f}, \Delta}^{\text{leak}} \cdot \mathcal{F}(h)\|^2 \quad (9.7.3)$$

$$\geq C_{\text{stab}}^{\text{IP}^*}(\Omega, \mathfrak{f}, (\nu))^2 \|h\|^2 - 4 \|\hat{p}_{\mathfrak{f}, \Delta}^{\text{leak}} \cdot \mathcal{F}(h)\|^2. \quad (9.7.4)$$

If $h \in \mathfrak{B}_{k,r,\mathbf{o}}^m$ is moreover a B-spline and $\nu \geq 1$, then (9.7.4) further implies that

$$\|T(h)|_K\|^2 \geq (C_{\text{stab}}^{\text{IP}^*}(\Omega, \mathbf{f}, (\nu))^2 - 4(1 - C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, \nu)^{2m})) \|h\|^2 \quad (9.7.5)$$

where the notation is as in theorem 9.16.

Proof. The first inequality, (9.7.3), is obtained by bounding $\|T(h)|_{K^c}\|$ via theorems 9.12 and 9.25 and using that $\|T(h)|_K\|^2 = \|T(h)\|^2 - \|T(h)|_{K^c}\|^2$. (9.7.4) then follows from (9.7.3) by estimating $\|T(h)\|$ via theorem 9.4. The bound (9.7.5) is obtained analogously if $\|\mathcal{D}(h)|_{K^c}\|^2 = \|h\|^2 - \|\mathcal{D}(h)|_K\|^2$ is estimated via theorem 9.16 instead of theorem 9.12. \square

While the right-hand side of (9.7.5) is clearly the simplest of all bounds in theorem 9.26, it is also the most pessimistic. The reason is that both the full-FoV-contrast $\|T(h)\|$ and the leaked part $\|T(h)|_{K^c}\|$ are bounded via *worst-case* estimates. Hence, the bound (9.7.5) is likely to be far from sharp since, otherwise, some $h \in \mathfrak{B}_{k,r,\mathbf{o}}^m \cap L^2(\Omega)$ would have to both minimize $\|T(h)\|$ and maximize $\|T(h)|_{K^c}\|$. However, as shown in [144], $\|T(h)\|$ is minimized by *low-frequency* modes, whereas the leakage estimates are in terms of *high-pass* filters.

Despite its lossiness, we demonstrate that the bound (9.7.5) may indeed guarantee stability in practically relevant settings. To this end, the required stability constant for an infinite FoV $C_{\text{stab}}^{\text{IP}^*}$ is approximated numerically, which can be done to high accuracy for ball- or square-domains Ω . Let us first consider inverse problem 9.1(b). This problem is excessively ill-conditioned [144] even for a full FoV, except for settings with very small object-domains. For such a case, we show that stability also holds with finite detectors:

Example 9.27 (Stability estimate for XPCI of weak objects (inverse problem 9.1(b))). Let $\mathbf{f} = 2 \cdot 10^3$ and $K = [-\frac{1}{2}; \frac{1}{2}]^2$. Let $h \in \mathfrak{B}_{k,r,\mathbf{o}}^m \cap L^2(\Omega)$ with support $\Omega = [-\frac{1}{20}; \frac{1}{20}]^2$, resolution $1/r = 190$ and spline order $k = 7$. Then the bound (9.7.5) guarantees stability with

$$\|\mathcal{T}(h)|_K\| \geq 0.12\|h\| \quad (C_{\text{stab}}^{\text{IP}1(\text{b})}(\Omega, \mathbf{f}) \geq 0.328, C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, 1.2) \geq 0.988). \quad (9.7.6)$$

By result 9.7, an *upper* bound for the resolution is given by $1/r \lesssim 0.45\mathbf{f}/\pi \approx 290$.

Unfortunately, as $C_{\text{stab}}^{\text{IP}1(\text{b})}(\Omega, \mathbf{f})$ decays exponentially with the Fresnel number associated with the size of Ω [144], stability cannot be guaranteed for larger object-domains Ω or \mathbf{f} .

The situation is better for inverse problem 9.2(b), i.e. for the reconstruction of *homogeneous objects* as introduced in §9.2.1.2. Of particular relevance are non-absorbing, pure phase objects:

Example 9.28 (Stability estimate for XPCI of weak phase objects (inverse problem 9.2(b): $\nu = 0$)). Let $\mathbf{f} = 5 \cdot 10^3$ and $K = [-\frac{1}{2}; \frac{1}{2}]^2$. Let $\varphi \in \mathfrak{B}_{k,r,\mathbf{o}}^m \cap L^2(\Omega, \mathbb{R})$ with $\Omega = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x}| \leq \frac{1}{10}\}$, resolution $1/r = 350$ and $k = 7$. Then the bound (9.7.5) guarantees stability with

$$\|\mathcal{S}_0(\varphi)|_K\| \geq 0.05\|\varphi\| \quad (C_{\text{stab}}^{\text{IP}2(\text{b})}(\Omega, \mathbf{f}, 0) \geq 0.151, C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, 1.25) \geq 0.997). \quad (9.7.7)$$

By result 9.8, an *upper* bound for the resolution is given by $1/r \lesssim 0.5\mathbf{f}/\pi \approx 800$.

Yet, the full-FoV stability constant $C_{\text{stab}}^{\text{IP2(b)}}(\Omega, \mathbf{f}, 0)$ decays like \mathbf{f}^{-1} for $\mathbf{f} \rightarrow \infty$, which is still too fast for (9.7.5) to guarantee stability at larger Fresnel numbers. This is different when the imaged sample is also known to be slightly absorbing, in which case the asymptotics improve to $C_{\text{stab}}^{\text{IP2(b)}}(\Omega, \mathbf{f}, \nu) \gtrsim \mathbf{f}^{-1/2}$ [144]. This enables stability guarantees for reconstructions at *optical* resolutions as fine as the native resolution of typical detectors. In such a setting, the finite FoV is no longer a limiting factor for the performance of the imaging setup. We consider an example for a sample satisfying $\mu = 0.1\phi$, i.e. for 10% absorption (see §9.2.1.2):

Example 9.29 (Stability estimate for XPCI of homogeneous objects (inverse problem 9.2(b): $\nu = \arctan(\frac{1}{10})$)). Let $\mathbf{f} = 4 \cdot 10^4$ and $K = [-\frac{1}{2}; \frac{1}{2}]^2$. Let $\varphi \in \mathfrak{B}_{k,r,\mathbf{o}}^m \cap L^2(\Omega, \mathbb{R})$ with $\Omega = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x}| \leq \frac{1}{4}\}$, resolution $1/r = 2000$ and $k = 7$. Then the bound (9.7.5) guarantees stability with

$$\|\mathcal{L}_\nu(\varphi)|_K\| \geq 0.08\|\varphi\| \quad (C_{\text{stab}}^{\text{IP2(b)}}(\Omega, \mathbf{f}, \nu) \geq 0.147, C_{\text{stab}}(\mathbf{f}_\Delta, \mathbf{f}_r, k, 1.25) \geq 0.998). \quad (9.7.8)$$

By result 9.8, an *upper* bound for the resolution is given by $1/r \lesssim 2^{-3/2}\mathbf{f}/\pi \approx 4500$.

9.7.3 Improved estimates for real-valued objects

In principle, the improved leakage bounds for the real-valued setting from §9.6 apply to the CTF-based reconstruction of homogeneous objects, inverse problem 9.2(b). Unfortunately, the derived bounds are too pessimistic in this setting to enable stability estimates for practically relevant Fresnel numbers. However, note that numerical simulations (not shown) indicate that the larger stability regions for the real-valued case, shown in figs. 9.4 and 9.9, indeed seem to carry over to the phaseless XPCI-setting.

9.8 Conclusions

We have studied locality of wave-propagation in the Fresnel- (or paraxial) regime in order to quantify the effects of a finite detector on the stability of X-ray phase contrast imaging (XPCI). The analysis shows that locality depends on spatial frequencies, i.e. the finer the features of some object h the more delocalized it is upon Fresnel-propagation $h \mapsto \mathcal{D}(h)$. As a consequence, *truncated* diffraction-data, as measured by any real-world detector, introduces a spatially varying resolution limit within the field-of-view: features of the imaged object finer than some limiting length-scale r_{stab} may induce a signal in the diffraction-pattern that essentially *leaks* out the detection-domain K upon propagation and thus cannot be stably reconstructed from the data. On the contrary, Lipschitz-stability estimates hold for images that comply with the resolution limit, as has been proven for multi-variate B-splines. The decisive property of B-splines for this result is that they are *quasi band-limited* functions. Notably, the obtained estimates on their concentration in Fourier-space (theorems 9.14 and 9.15) may be of interest beyond the specific inverse problems considered this work.

The stability results do not only hold for the (hypothetical) case where full complex-valued Fresnel-data $\mathcal{D}(h)|_K$ is measured, but have also been extended to the *phaseless* setting of

XPCI in the linear CTF-regime. However, as the (possibly complicated) interplay between the instabilities due to a finite FoV and those due to the missing phase is not taken into account, the derived estimates for the phaseless case are expected to be highly non-optimal.

The maximum resolution for a square detector is found to be $1/r_{\text{stab}} \approx \mathfrak{f}$, in accordance with the numerical aperture of the lensless imaging setup [158, 127], where $\mathfrak{f} = b^2/(\lambda d)$ is the Fresnel number associated with the detector's aspect-length b (λ : wavelength, d : propagation-distance). Hence, if \mathfrak{f} is smaller than the number of detector-pixels along one dimension, the finite FoV bottlenecks the achievable resolution. For complex-valued images to be reconstructed, the optimal resolution is moreover attained only in the very center of the FoV. Interestingly, this situation is much worse than for the standard XPCI case of homogeneous objects, that boils down to reconstructing a *real-valued* image. In the latter case, maximum resolution $\approx \mathfrak{f}$ can be achieved in large parts of the FoV, except for the detector-corners.

The analysis of this work may be readily extended. For once, all results can be adjusted to non-square object- and detection-domains at the cost of a more involved notation. Moreover, it is straightforward to extend the derived locality-bounds to multiple diffraction-patterns acquired at different Fresnel numbers $\mathfrak{f}_1, \mathfrak{f}_2, \dots$, which is a typical setting in XPCI. However, the larger amount of data is not too useful in view of a finite detector because, according to this work's analysis, features that leak outside the FoV for the largest Fresnel number are lost in *all* diffraction patterns. Finally, the estimates obtained within the Fresnel-regime may be generalized to propagation within the full Helmholtz equation, by combining them with bounds on the deviation from the paraxial limit. Thereby, the results might be applied to a large range of scattering experiments that give rise to *approximately* paraxial wave-fields.

Acknowledgments

The author thanks Johannes Hagemann, Malte Vassholz, Thorsten Hohage and Tim Salditt for inspiring discussions. Financial support by Deutsche Forschungsgemeinschaft DFG through SFB 755 - Nanoscale Photonic Imaging is gratefully acknowledged.

Appendix

9.A Fresnel-propagation and frequency shifts

Proof of lemma 9.6. By the alternate form of the Fresnel propagator (P3), we have

$$e^{im\pi/4\mathfrak{f}^{-\frac{m}{2}}}\mathcal{D}(\mathbf{e}_a \cdot f) = n_{\mathfrak{f}} \cdot \mathcal{F}(n_{\mathfrak{f}} \cdot \mathbf{e}_a \cdot f)(\mathfrak{f}(\cdot)). \quad (9.A.1)$$

Moreover, it holds for all $\mathbf{x} \in \mathbb{R}^m$

$$\begin{aligned} n_{\mathfrak{f}} \cdot \mathbf{e}_a(\mathbf{x}) &= \exp\left(i\left(\frac{\mathfrak{f}\mathbf{x}^2}{2} + \mathbf{a} \cdot \mathbf{x}\right)\right) = \exp\left(-\frac{i\mathbf{a}^2}{2\mathfrak{f}}\right) \exp\left(\frac{i\mathfrak{f}(\mathbf{x} + \mathbf{a}/\mathfrak{f})^2}{2}\right) \\ &= m_{\mathfrak{f}}(\mathbf{a}) \cdot n_{\mathfrak{f}}(\mathbf{x} + \mathbf{a}/\mathfrak{f}) = m_{\mathfrak{f}}(\mathbf{a}) \cdot T_{\mathbf{a}/\mathfrak{f}}(n_{\mathfrak{f}})(\mathbf{x}), \end{aligned} \quad (9.A.2)$$

Since $(T_{\mathbf{t}})^{-1} = T_{-\mathbf{t}}$ and $\mathcal{F}(T_{\mathbf{t}}(g)) = \mathbf{e}_{\mathbf{t}} \cdot \mathcal{F}(g)$ for any $\mathbf{t} \in \mathbb{R}^m$, $g \in L^2(\mathbb{R}^m)$, we thus have

$$\begin{aligned}
n_{\mathfrak{f}} \cdot \mathcal{F}(n_{\mathfrak{f}} \cdot \mathbf{e}_{\mathbf{a}} \cdot f)(\mathfrak{f}(\cdot)) &= m_{\mathfrak{f}}(\mathbf{a}) \cdot n_{\mathfrak{f}} \cdot \mathcal{F}(T_{\mathbf{a}/\mathfrak{f}}(n_{\mathfrak{f}}) \cdot f)(\mathfrak{f}(\cdot)) \\
&= m_{\mathfrak{f}}(\mathbf{a}) \cdot n_{\mathfrak{f}} \cdot \mathcal{F}(T_{\mathbf{a}/\mathfrak{f}}(n_{\mathfrak{f}} \cdot T_{-\mathbf{a}/\mathfrak{f}}(f)))(\mathfrak{f}(\cdot)) \\
&= m_{\mathfrak{f}}(\mathbf{a}) \cdot n_{\mathfrak{f}} \cdot \mathbf{e}_{\mathbf{a}/\mathfrak{f}}(\mathfrak{f}(\cdot)) \cdot \mathcal{F}(n_{\mathfrak{f}} \cdot T_{-\mathbf{a}/\mathfrak{f}}(f))(\mathfrak{f}(\cdot)) \\
&= m_{\mathfrak{f}}(\mathbf{a}) \cdot \mathbf{e}_{\mathbf{a}} \cdot (n_{\mathfrak{f}} \cdot \mathcal{F}(n_{\mathfrak{f}} \cdot T_{-\mathbf{a}/\mathfrak{f}}(f))(\mathfrak{f}(\cdot))) \\
&= m_{\mathfrak{f}}(\mathbf{a}) \cdot \mathbf{e}_{\mathbf{a}} \cdot e^{i\pi/4} \mathfrak{f}^{-\frac{m}{2}} \mathcal{D}(T_{-\mathbf{a}/\mathfrak{f}}(f))
\end{aligned} \tag{9.A.3}$$

By comparing to (9.A.1) and exploiting that \mathcal{D} commutes with translations as a convolution operator, we finally obtain

$$\mathcal{D}(\mathbf{e}_{\mathbf{a}} \cdot f) = m_{\mathfrak{f}}(\mathbf{a}) \cdot \mathbf{e}_{\mathbf{a}} \cdot \mathcal{D}(T_{-\mathbf{a}/\mathfrak{f}}(f)) = m_{\mathfrak{f}}(\mathbf{a}) \cdot \mathbf{e}_{\mathbf{a}} \cdot T_{-\mathbf{a}/\mathfrak{f}}(\mathcal{D}(f)). \quad \square$$

9.B Quasi-band-limitation of B-splines

Proof of theorem 9.14. We prove the estimate (9.5.3) for $h = \sum_{j \in \mathbb{Z}} b_j B_k(\cdot/r - j - \mathbf{o}) \in \mathfrak{B}_{k,r,\mathbf{o}}^1$ with coefficients that vanish for all but finitely many entries, i.e. $(b_j) \in \ell^0(\mathbb{Z}) := \{(c_j)_{j \in \mathbb{Z}} \subset \mathbb{C} : \exists J \subset \mathbb{Z} \text{ finite s.t. } c_l = 0 \text{ for } l \in \mathbb{Z} \setminus J\}$. This is sufficient since such splines form an L^2 -dense subspace of $\mathfrak{B}_{k,r,\mathbf{o}}^1$ (by denseness of $\ell^0(\mathbb{Z})$ in $\ell^2(\mathbb{Z})$ and the Riesz-sequence property (9.5.2)) and both sides of (9.5.3) are L^2 -continuous in h .

For the considered h , all sums of the form $\sum_{j \in \mathbb{Z}} b_j(\dots)$ are finite. By linearity and the behavior of the Fourier-transform under translations and dilations, this implies that

$$\begin{aligned}
\mathcal{F}(h)(\xi) &= \mathcal{F}\left(\sum_{j \in \mathbb{Z}} b_j B_k(\cdot/r - j - \mathbf{o})\right)(\xi) = \sum_{j \in \mathbb{Z}} b_j \mathcal{F}(B_k(\cdot/r - j - \mathbf{o}))(\xi) \\
&= \underbrace{\left(\exp(-ir\xi\mathbf{o}) \sum_{j \in \mathbb{Z}} b_j \exp(-ir\xi j)\right)}_{=: \hat{h}_{\text{per}}(r\xi)} r \mathcal{F}(B_k)(r\xi) \quad \text{for all } \xi \in \mathbb{R}.
\end{aligned} \tag{9.B.1}$$

From (9.B.1), it can be readily seen that the function \hat{h}_{per} is 2π -periodic, i.e. $\hat{h}_{\text{per}}(\xi + 2\pi l) = \hat{h}_{\text{per}}(\xi)$ for all $\xi \in \mathbb{R}$, $l \in \mathbb{Z}$.

In order to prove the estimate (9.5.3), we decompose the Fourier-domain: with $\bar{\nu} := 1 + 2\lceil(\nu - 1)/2\rceil$ as defined in the assumptions, it holds that

$$(\nu \Xi_r)^c = ((\bar{\nu} \Xi_r) \setminus (\nu \Xi_r)) \cup \bigcup_{n=1+\lceil(\nu-1)/2\rceil}^{\infty} \left(\Xi_r + \frac{2\pi}{r} n \right) \cup \left(\Xi_r - \frac{2\pi}{r} n \right), \tag{9.B.2}$$

where the union is mutually disjoint except for intersections of Lebesgue-measure zero. Accordingly, the squared L^2 -norm over $(\nu \Xi_r)^c$ can be written as a sum

$$\|\mathcal{F}(h)|_{(\nu \Xi_r)^c}\|^2 = \|\mathcal{F}(h)|_{(\bar{\nu} \Xi_r) \setminus (\nu \Xi_r)}\|^2 + \sum_{n=1+\lceil(\nu-1)/2\rceil}^{\infty} \left(\|\mathcal{F}(h)|_{\Xi_r + \frac{2\pi}{r} n}\|^2 + \|\mathcal{F}(h)|_{\Xi_r - \frac{2\pi}{r} n}\|^2 \right) \tag{9.B.3}$$

We first consider the squared norms in the second summand on the right-hand-side of (9.B.3). By the 2π -periodicity of \hat{h}_{per} , we have

$$\begin{aligned} r^{-1} \|\mathcal{F}(h)|_{\Xi_r + \frac{2\pi}{r}l}\|^2 &= r \int_{\Xi_r + \frac{2\pi}{r}l} \left| \hat{h}_{\text{per}}(r\xi) \right|^2 |\mathcal{F}(B_k)(r\xi)|^2 d\xi = \int_{(2\pi l - 1)\pi}^{(2\pi l + 1)\pi} \left| \hat{h}_{\text{per}}(\xi) \right|^2 |\mathcal{F}(B_k)(\xi)|^2 d\xi \\ &= \int_{-\pi}^{\pi} \left| \hat{h}_{\text{per}}(\xi) \right|^2 |\mathcal{F}(B_k)(\xi + 2\pi l)|^2 d\xi \end{aligned} \quad (9.B.4)$$

for all $l \in \mathbb{Z}$. Hence, we obtain for all $n \in \mathbb{N}$

$$\begin{aligned} \|\mathcal{F}(h)|_{(\Xi_r + \frac{2\pi n}{r}) \cup (\Xi_r - \frac{2\pi n}{r})}\|^2 &= r \int_{-\pi}^{\pi} \left| \hat{h}_{\text{per}}(\xi) \right|^2 (|\mathcal{F}(B_k)(\xi + 2\pi n)|^2 + |\mathcal{F}(B_k)(\xi - 2\pi n)|^2) d\xi \\ &\leq r c_{k,n} \int_{-\pi}^{\pi} \left| \hat{h}_{\text{per}}(\xi) \right|^2 |\mathcal{F}(B_k)(\xi)|^2 d\xi = c_{k,n} \|\mathcal{F}(h)|_{\Xi_r}\|^2 \end{aligned} \quad (9.B.5)$$

$$c_{k,n} := \sup_{\xi \in [-\pi; \pi]} w_{k,n}(\xi), \quad w_{k,n}(\xi) := \frac{|\mathcal{F}(B_k)(\xi + 2\pi n)|^2 + |\mathcal{F}(B_k)(\xi - 2\pi n)|^2}{|\mathcal{F}(B_k)(\xi)|^2}.$$

We aim to explicitly compute the coefficients $c_{k,n}$. To this end, we use the known Fourier transform of B_k , $\mathcal{F}(B_k)(\xi) = (2\pi)^{-1/2} \text{sinc}(\xi/2)^{k+1}$ for all $\xi \in \mathbb{R}$ where $\text{sinc}(x) := \sin(x)/x$. As the function \sin^2 is π -periodic, it holds that

$$\frac{|\mathcal{F}(B_k)(\xi + 2\pi l)|^2}{|\mathcal{F}(B_k)(\xi)|^2} = \frac{\sin(\xi/2 + \pi l)^{2(k+1)}}{\sin(\xi/2)^{2(k+1)}} \cdot \frac{(\xi/2)^{2(k+1)}}{(\xi/2 + \pi l)^{2(k+1)}} = \frac{\xi^{2(k+1)}}{(\xi + 2\pi l)^{2(k+1)}} \quad (9.B.6)$$

for all $\xi \in [-\pi; \pi]$, $l \in \mathbb{Z}$. Accordingly, the weight-function $w_{k,n}$ is given by

$$\begin{aligned} w_{k,n}(\xi) &= \frac{\xi^{2(k+1)}}{(\xi + 2\pi n)^{2(k+1)}} + \frac{\xi^{2(k+1)}}{(\xi - 2\pi n)^{2(k+1)}} \\ \Rightarrow c_{k,n} &= \sup_{\xi \in [-\pi; \pi]} w_{k,n}(\xi) = w_{k,n}(\pm\pi) = \frac{1}{(2n - 1)^{2(k+1)}} + \frac{1}{(2n + 1)^{2(k+1)}}, \end{aligned} \quad (9.B.7)$$

for all $n \in \mathbb{N}$, where the second line follows from the fact that $w_{k,n} : [-\pi; \pi] \rightarrow \mathbb{R}$ is even and attains its maximum at the boundary as a convex function.

Now it remains to bound the first term on the right-hand side of (9.B.3). By definition, it holds that $\bar{\nu} \geq \nu$, where equality holds if and only if $\nu \in 2\mathbb{N} - 1$, in which case the considered term vanishes. Hence, we restrict to $\bar{\nu} > \nu$. By transforming the integration variable and

exploiting periodicity analogously as in (9.B.4), we obtain

$$\begin{aligned}
 r^{-1} \left\| \mathcal{F}(h)|_{(\bar{\nu}\Xi_r) \setminus (\nu\Xi_r)} \right\|^2 &= \left(\int_{-\bar{\nu}\pi}^{-\nu\pi} + \int_{\nu\pi}^{\bar{\nu}\pi} \right) \left| \hat{h}_{\text{per}}(\xi) \right|^2 |\mathcal{F}(B_k)(\xi)|^2 d\xi \\
 &= \int_{-\pi}^{\bar{\nu}\pi} \left| \hat{h}_{\text{per}}(\xi - (\bar{\nu} - 1)\pi) \right|^2 |\mathcal{F}(B_k)(\xi)|^2 d\xi \\
 &\quad + \int_{-\bar{\nu}\pi}^{\pi} \left| \hat{h}_{\text{per}}(\xi) \right|^2 |\mathcal{F}(B_k)(\xi + (\bar{\nu} - 1)\pi)|^2 d\xi \\
 &\leq c_{k,0} r^{-1} \left\| \mathcal{F}(h)|_{\Xi_r} \right\|^2, \quad c_{k,0} := \sup_{\xi \in [-\pi; \pi]} w_{k,0}(\xi) \tag{9.B.8}
 \end{aligned}$$

$$w_{k,0}(\xi) := \begin{cases} \frac{|\mathcal{F}(B_k)(\xi - (\bar{\nu} - 1)\pi)|^2}{|\mathcal{F}(B_k)(\xi)|^2} & \text{for } \xi \leq -|\tilde{\nu}|\pi \\ \frac{|\mathcal{F}(B_k)(\xi - (\bar{\nu} - 1)\pi)|^2}{|\mathcal{F}(B_k)(\xi)|^2} + \frac{|\mathcal{F}(B_k)(\xi + (\bar{\nu} - 1)\pi)|^2}{|\mathcal{F}(B_k)(\xi)|^2} & \text{for } -\tilde{\nu} < \xi < \tilde{\nu} \\ \frac{|\mathcal{F}(B_k)(\xi + (\bar{\nu} - 1)\pi)|^2}{|\mathcal{F}(B_k)(\xi)|^2} & \text{for } \xi > |\tilde{\nu}|\pi \\ 0 & \text{else} \end{cases},$$

where $\tilde{\nu} = \bar{\nu} - \nu - 1 \in (-1; 1)$ has been inserted. Since $(\bar{\nu} - 1)\pi$ is necessarily an integer-multiple of 2π , we may again use the relation (9.B.6) to simplify $w_{k,0}$:

$$w_{k,0}(\xi) := \begin{cases} \frac{\xi^{2(k+1)}}{(\xi - (\bar{\nu} - 1)\pi)^{2(k+1)}} & \text{for } \xi \leq -|\tilde{\nu}|\pi \\ \frac{\xi^{2(k+1)}}{(\xi - (\bar{\nu} - 1)\pi)^{2(k+1)}} + \frac{\xi^{2(k+1)}}{(\xi + (\bar{\nu} - 1)\pi)^{2(k+1)}} & \text{for } -\tilde{\nu} < \xi < \tilde{\nu} \\ \frac{\xi^{2(k+1)}}{(\xi + (\bar{\nu} - 1)\pi)^{2(k+1)}} & \text{for } \xi > |\tilde{\nu}|\pi \\ 0 & \text{else} \end{cases}, \tag{9.B.9}$$

The function $w_{k,0}$ can be readily seen to be smooth and convex on each of the intervals $[-\pi; -|\tilde{\nu}|\pi]$, $(-\tilde{\nu}\pi; \tilde{\nu}\pi)$ and $(|\tilde{\nu}|\pi; \pi]$. Consequently, the supremum over $[-\pi; \pi]$ is attained at one of the six boundary points of these intervals. By the symmetry $w_{k,0}(-\xi) = w_{k,0}(\xi)$, it is furthermore sufficient to consider non-negative values of ξ .

We first consider the case $\tilde{\nu} \in (-1; 0]$. Then the interval $(-\tilde{\nu}\pi; \tilde{\nu}\pi)$ is empty and the definition of $w_{k,0}$ simplifies accordingly so that $c_{k,0}$ can be computed as

$$c_{k,0} = \max \left\{ \lim_{\xi \searrow |\tilde{\nu}|\pi} w_{k,0}(\xi), w_{k,0}(\pi) \right\} = \max \left\{ \frac{|\tilde{\nu}|}{((\bar{\nu} - 1) + |\tilde{\nu}|)}, \frac{1}{\bar{\nu}} \right\}^{2(k+1)} = \frac{1}{\bar{\nu}^{2(k+1)}} \tag{9.B.10}$$

for all $\bar{\nu} \in 2\mathbb{N} - 1$. On the other hand, if $\tilde{\nu} \in (0; 1)$, then also the interior domain-part $(-\tilde{\nu}\pi; \tilde{\nu}\pi)$ has to be considered in the computation of the supremum:

$$\begin{aligned}
 c_{k,0} &= \max \left\{ \lim_{\xi \nearrow \tilde{\nu}\pi} w_{k,0}(\xi), \lim_{\xi \searrow \tilde{\nu}\pi} w_{k,0}(\xi), w_{k,0}(\pi) \right\} \\
 &= \max \left\{ \frac{\tilde{\nu}^{2(k+1)}}{((\bar{\nu} - 1) + \tilde{\nu})^{2(k+1)}} + \frac{\tilde{\nu}^{2(k+1)}}{((\bar{\nu} - 1) - \tilde{\nu})^{2(k+1)}}, \frac{\tilde{\nu}^{2(k+1)}}{((\bar{\nu} - 1) + \tilde{\nu})^{2(k+1)}}, \frac{1}{\bar{\nu}^{2(k+1)}} \right\} \\
 &\stackrel{(\nu = \bar{\nu} - 1 - \tilde{\nu})}{=} \frac{1}{\bar{\nu}^{2(k+1)}} + \underbrace{\max \left\{ \frac{\max\{\tilde{\nu}, 0\}^{2(k+1)}}{(\nu + 2\tilde{\nu})^{2(k+1)}} + \frac{\max\{\tilde{\nu}, 0\}^{2(k+1)}}{\nu^{2(k+1)}} - \frac{1}{\bar{\nu}^{2(k+1)}}, 0 \right\}}_{c_{\text{band},0}(k,\nu)}. \tag{9.B.11}
 \end{aligned}$$

By comparing to (9.B.10), it can be seen that equality between the left-hand side and the bottom line of (9.B.11) remains valid for $\tilde{\nu} \in (-1; 0]$, i.e. holds true in general.

By inserting (9.B.5), (9.B.7), (9.B.8) and (9.B.11) into (9.B.3), we finally arrive at

$$\begin{aligned}
 \|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\|^2 &\leq \left(c_{\text{band},0}(k, \nu) + \frac{1}{\bar{\nu}^{2(k+1)}} + \sum_{n=1+\lceil(\nu-1)/2\rceil}^{\infty} c_{k,n} \right) \|\mathcal{F}(h)|_{\Xi_r}\|^2 \\
 &= \left(c_{\text{band},0}(k, \nu) + \sum_{n=\lceil(\nu-1)/2\rceil}^{\infty} \frac{2}{(2n+1)^{2(k+1)}} \right) \|\mathcal{F}(h)|_{\Xi_r}\|^2 \\
 &= c_{\text{band}}(k, \nu) \|\mathcal{F}(h)|_{\Xi_r}\|^2 \stackrel{\Xi_r \subset \nu\Xi_r}{\leq} c_{\text{band}}(k, \nu)^2 \|\mathcal{F}(h)|_{\nu\Xi_r}\|^2 \quad (9.B.12)
 \end{aligned}$$

The assertion now follows by exploiting that $\|\mathcal{F}(h)\|^2 = \|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\|^2 + \|\mathcal{F}(h)|_{\nu\Xi_r}\|^2$:

$$\begin{aligned}
 C_{\text{band}}(k, \nu)^2 \|\mathcal{F}(h)\|^2 &= C_{\text{band}}(k, \nu)^2 \|\mathcal{F}(h)|_{\nu\Xi_r}\|^2 + C_{\text{band}}(k, \nu)^2 \|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\|^2 \\
 &\stackrel{(9.B.12)}{\geq} C_{\text{band}}(k, \nu)^2 \left(\frac{1}{c_{\text{band}}(k, \nu)} + 1 \right) \|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\|^2 = \|\mathcal{F}(h)|_{(\nu\Xi_r)^c}\|^2. \quad (9.B.13)
 \end{aligned}$$

Negative result for $\nu < 1$: Now let $\nu < 1$. Then, by the theory of Fourier series, there exists a sequence $(b_j)_{j \in \mathbb{Z}}$ such that

$$\hat{b}_\nu(\xi) := \exp(-ir\xi\mathbf{o}) \sum_{j \in \mathbb{Z}} b_j \exp(-ir\xi j) = \begin{cases} 1 & \text{if } \xi \in \Xi_r \setminus (\nu\Xi_r) \\ 0 & \text{if } \xi \in \nu\Xi_r \end{cases}, \quad \xi \in [-\pi; \pi] \quad (9.B.14)$$

in an L^2 -sense. If we define $h := \sum_{j \in \mathbb{Z}} b_j B_k(\cdot/r - j - \mathbf{o})$ as the corresponding B-spline, then the periodic part of $\mathcal{F}(h)$ in (9.4.7) is given by $\hat{h}_{\text{per}} = \hat{b}_\nu$. Hence, it follows that $\text{supp}(\mathcal{F}(h)) \cap [-\pi; \pi] = \text{supp}(\hat{b}_\nu) \subset (\nu\Xi_r)^c$, i.e. $\mathcal{F}(h) = \mathcal{F}(h)|_{(\nu\Xi_r)^c}$. The constructed example shows that no non-trivial bound of the form (9.5.3) may hold true for $\nu < 1$. \square

Chapter 10

Complete Article 3: Regularized Newton methods for X-ray phase contrast and general imaging problems

Reproduced from the published article [143]:*

Maretzke S, Bartels M, Krenkel M, Salditt T, Hohage T. *Optics Express*, **24**(6):6490–6506, 2016.

DOI: [10.1364/OE.24.006490](https://doi.org/10.1364/OE.24.006490)

(subject to slight adjustments of the notation and minor corrections)

Abstract: Like many other advanced imaging methods, x-ray phase contrast imaging and tomography require mathematical inversion of the observed data to obtain real-space information. While an accurate forward model describing the generally nonlinear image formation from a given object to the observations is often available, explicit inversion formulas are typically not known. Moreover, the measured data might be insufficient for stable image reconstruction, in which case it has to be complemented by suitable *a priori* information. In this work, regularized Newton methods are presented as a general framework for the solution of such ill-posed nonlinear imaging problems. For a proof of principle, the approach is applied to x-ray phase contrast imaging in the near-field propagation regime. Simultaneous recovery of the phase- and amplitude from a single near-field diffraction pattern without homogeneity constraints is demonstrated for the first time. The presented methods further permit *all-at-once* phase contrast tomography, i.e. simultaneous phase retrieval and tomographic inversion. We demonstrate the potential of this approach by three-dimensional imaging of a colloidal crystal at 95 nm isotropic resolution.

*Copyright © 2016 Optical Society of America. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modifications of the content of this paper are prohibited.

10.1 Introduction

Lensless coherent diffractive x-ray imaging (CDI) has opened up a new field of high resolution structure analysis beyond the ensemble averaging of conventional x-ray diffraction [158, 167, 147]. Typically, lensless x-ray imaging setups are closer to diffraction experiments than to a classical microscope setup with lenses, except that they require a sufficiently coherent probing wavefront or beam as well as sufficient sampling of the diffraction pattern. In other words, the imaging systems is essentially based on free space propagation between object and detector. Depending on whether the data is recorded in the optical near-field or far-field, the propagation is modeled by the Fresnel propagator or a Fourier transform (Fraunhofer far-field regime), respectively. As in conventional diffraction, each detector pixel carries information about all object pixels. Therefore, if the data is modeled in the detector plane, high spatial resolution can only be achieved in terms of the average structure. Contrarily, if the data is *inverted* to reconstruct the object, the individual real space configuration is depicted, beyond an ensemble average (or more precisely an average over the entire illuminated volume). Importantly, the image formation is still very similar to a plain diffraction experiment, while the data analysis is not. Data Modeling for example by least-square fitting is replaced by image reconstruction. Model formulation is replaced by the formulation of *a priori* knowledge (constraints), which are required to compensate for the missing information on the phase of the diffraction field, and hence to achieve a unique solution. Accordingly, the difference is brought about by the inversion of the diffraction process: by solving the inverse problem of (non-crystallographic) diffraction, we obtain access to the individual configuration instead of the average sizes and correlation lengths in the object. It is for this reason that iterative algorithms [146, 17, 59, 160, 138], being to date the engine of CDI, have received so much attention. Iteratively cycling between the detector and object planes, they feed in both measured data and additional *a priori* information on the solution.

CDI uses *a priori* knowledge, for example related to the object support or its optical constants (positivity, pure phase contrast). Diversity in the data generated by illuminating the same object pixels by different wavefronts may be exploited by ptychographic algorithms [63, 172, 191, 137, 192]. In general, however, one is interested in finding experimental settings allowing for robust phase retrieval with least invasive constraints, dose and accumulation times. In many applications for example, one is interested in reconstructing from single recordings, without scanning or multiple exposures. At the same time, the object may be composed of several materials, both with phase and amplitude contrast. It is therefore often of advantage to perform the imaging in the optical near-field rather than far-field. Here, phase information is much more directly encoded in the diffraction pattern, by ways of interference between diffracted and reference waves. Recently, improved uniqueness results have been presented for near-field imaging using the theory of entire functions [140]. In fact, contrary to a common belief, see e.g. [108, 32], measurements at only one distance are sufficient to determine both the real and the imaginary part of any compactly supported refractive index distribution. These results are much stronger than those for the Fraunhofer regime, as obtained e.g. by [116]. Loosly speaking, the near-field case can thus be expected to require less information to complement the measurements than the far-field case. In practice, however, many iterative algorithms which have been adapted to the near-field case [76, 188, 170] require similar data diversity or constraints.

Notwithstanding the merits of many highly performing iterative algorithms, it is therefore necessary to broaden the perspective of the phase retrieval field. Here, we present iteratively regularized Gauss-Newton methods (IRGNM) [9] as an alternative approach to phase retrieval and other imaging problems. In this method, each iterate is computed to provide an optimal compromise between agreement with the measured data and additional constraints on the basis of a local linearization of contrast formation, as we discuss further below. The approach is related to the regularized gradient descent methods for phase reconstruction proposed in [47], but promises improved convergence owing to the Newton-like solution of linearized subproblems. We apply the general IRGNM framework to near-field phase contrast with x-rays, for the reasons mentioned above, and since recent progress in propagation imaging has narrowed the gap in resolution compared to CDI. In fact, using highly divergent and coherent quasi-spherical wavefronts, x-ray imaging in the optical near-field has been recently demonstrated down to 20 nm resolution [13].

The scope of the present work is two-fold. Firstly, we give a concise overview of iteratively regularized Gauss-Newton methods in view of x-ray imaging, summarizing and explaining recent mathematical literature in this field for an applied audience. Secondly, we demonstrate the performance of this approach in solving the phase problem on the level of real state-of-the-art experimental data. In particular, we apply an IRGNM approach to three-dimensional (3d) imaging, i.e. we show how the method can be used to perform phase retrieval and tomographic reconstruction simultaneously. This strategy has been argued to enable improved accuracy compared to splitting the reconstruction into phase retrieval problems for each angle to recover the fields in the object plane and a subsequent inversion of the Radon transform [37, 16, 175, 139]. We show that IRGNM approaches offer significant flexibility in treating various experimental setups and different *a priori* information. In the long run, we also expect advantages owing to the fact that the noise characteristics of the data can be suitably accounted for in this framework. For the important example of Poisson data, Newton-type regularization methods with Kullback-Leibler-type data fidelity terms have already been proposed [98].

The manuscript is organized as follows: §10.2 introduces IRGNM in view of image reconstruction problems from the principal idea to practical implementation. In §10.3, the framework is applied to (2d) near-field phase contrast with hard x-rays, imaging the phase shifts and absorption induced by a nano-structured test pattern. A Kaczmarz-type IRGNM suitable for tomographic imaging is presented in §10.4 and applied to resolve the structure of a colloidal micro-crystal.

10.2 Regularized Newton methods for imaging

10.2.1 Basic approach

We consider an abstract imaging system of the form

$$I^{\text{obs}} = F(f^\dagger) + \epsilon. \quad (10.2.1)$$

Here $I^{\text{obs}} \in Y$ denotes some observable intensity data, given by the image of the unknown object $f^\dagger \in X$ (e.g. a spatially varying refractive index or scatterer positions) under a known forward operator F and superimposed measurement errors ϵ . Note that ϵ may depend on $F(f^\dagger)$ as is the case for Poisson data. For many models, an explicit inversion formula for F is not

known. Moreover, even if the inverse F^{-1} is available, it is often not continuous and would thus amplify errors ϵ by large magnitudes if applied directly to the data I^{obs} . Such *ill-posedness* of the reconstruction despite uniqueness is well-established e.g. for computed tomography [155]. We seek a method to stably recover f^\dagger from an ill-posed problem of the form (10.2.1).

Notably, the operator $F : X \rightarrow Y$ modeling the imaging system is nonlinear in general, for example whenever a phase retrieval problem described by a squared modulus operation is involved. Nevertheless, reasonable results can often be achieved using a linearization of contrast formation, as given for instance by the contrast transfer function in electron microscopy or coherent x-ray imaging [62, 205, 79, 42].

Mathematically, such first order approximations are justified by the Fréchet differentiability of the forward operator F , i.e. for any f there exists a bounded linear map $F'[f]$ such that $\lim_{h \rightarrow 0} (F(f+h) - F(f) - F'[f]h) / \|h\| = 0$. A natural approach to solve (10.2.1) is then given by Newton-type iterations

$$f_{k+1} = f_k + F'[f_k]^{-1} (I^{\text{obs}} - F(f_k)). \quad (10.2.2)$$

As opposed to methods based on a *static* linearization of contrast formation, the linearizations in (10.2.2) are computed about the current iterate f_k and thereby account for moderate nonlinearity.

However, iterations of the form (10.2.2) are often neither feasible nor desirable for imaging because the linearized problems - just like the nonlinear equation (10.2.1) - are typically ill-posed. Hence, the solution of (10.2.2) is unstable and may not even exist. Again, one may think of ambiguities in phase retrieval problems. A remedy is given by iteratively regularized Gauss-Newton methods (IRGNM) as first proposed by [9], corresponding to Tikhonov regularization of the Newton steps:

$$f_{k+1} = \underset{f \in X}{\operatorname{argmin}} \left\| F(f_k) + F'[f_k](f - f_k) - I^{\text{obs}} \right\|_Y^2 + \alpha_k \|f - f_0\|_X^2 \quad (10.2.3)$$

Here, $\|\cdot\|_Y$ and $\|\cdot\|_X$ denote the norms in Hilbert spaces X and Y , $f_0 \in X$ is the initial guess and $\alpha_k > 0$ is a regularization parameter. In this setting, it can be shown that (10.2.3) always has a unique solution given by [60]

$$f_{k+1} = f_k + (F'[f_k]^* F'[f_k] + \alpha_k)^{-1} \left(F'[f_k]^* (I^{\text{obs}} - F(f_k)) + \alpha_k (f_0 - f_k) \right). \quad (10.2.4)$$

$F'[f_k]^*$ denotes the adjoint of the linear map $F'[f_k] : X \rightarrow Y$. In (10.2.4), only the inverse of the selfadjoint positive-definite operator $F'[f_k]^* F'[f_k] + \alpha_k$ has to be computed, which is *bounded* according to the estimate

$$\| (F'[f_k]^* F'[f_k] + \alpha_k)^{-1} \| \leq \frac{1}{\alpha_k}. \quad (10.2.5)$$

Accordingly, the iterate f_{k+1} depends continuously on I^{obs} , i.e. the impact of data errors on the reconstruction is *regularized*.

As the IRGNM is based on linearizations of the imaging operator, yet iteratively updated, the approach is best suited for weakly or moderately nonlinear problems. Formally, convergence of the method to f^\dagger for $\epsilon \rightarrow 0$ can indeed be shown given bounds on the nonlinearity of F along with suitably chosen α_k and f_0 [24, 112].

10.2.2 Parameter choice and constraints

The first term on the right hand side of (10.2.3) measures the agreement of the object with the observed data I^{obs} based on the current linearization. To achieve competitive results, the choice of the norm $\|\cdot\|_Y$ should reflect the statistical properties of the data errors ϵ , e.g. by taking the negative log-likelihood of the measured signal I^{obs} . For additive Gaussian white noise, this consideration leads to the choice of the standard L^2 -norm, i.e.

$$\|I - I^{\text{obs}}\|_Y := \int |I - I^{\text{obs}}|^2 dx. \quad (10.2.6)$$

For Poisson noise, the resulting data fidelity term is the Kullback-Leibler-divergence. This distance measure can be implemented in the framework of generalized Newton methods as demonstrated by [98]. Within the IRGNM, a quadratic approximation about its minimum may be used as a norm

$$\|I - I^{\text{obs}}\|_Y := \left\| \frac{I - I^{\text{obs}}}{\max(I_0, I^{\text{obs}})^{\frac{1}{2}}} \right\|_{L^2}, \quad (10.2.7)$$

where $I_0 > 0$ is a regularizing parameter.

On the right hand side of (10.2.3), the data residual is balanced with the regularization term $\alpha_k \|f - f_0\|_X^2$ bounding the deviation from the initial guess f_0 . The regularization parameter weights the different contributions: if α_k is very small, there is essentially no regularization and the norm bound (10.2.5) diverges, allowing for large amplifications of the data error. If α_k is chosen too large on the other hand, the Newton iterate computed via (10.2.4) need not have much to do with the actual image to be reconstructed. A good choice of α_k must thus balance data- and approximation errors. One possible strategy is the following:

- Choose α_0 to approximately balance the norms in (10.2.3), for example by setting $\alpha_0 \sim \|F'[f_0]F'[f_0]^*(I^{\text{obs}})\|_Y^2 / \|F'[f_0]^*(I^{\text{obs}})\|_X^2$.
- Reduce α_k by a constant factor, e.g. $\alpha_{k+1}/\alpha_k = \frac{2}{3}$.
- Stop at the first k s.t. $s_k := \|F(f_k) - I^{\text{obs}}\|_Y \leq \tau \|\epsilon\|_Y$, i.e. when the residual attains $\tau \geq 1$ times the noise level.

The stopping criterion, known as Morozov's discrepancy principle [153], requires good knowledge of the magnitude of data errors. When this is not available, noise level-free stopping rules need to be applied, see e.g. [60, 88, 114]. Here, the principal idea is to make s_k as small as possible while limiting heuristic measures for the impact of data errors such as the object norm $\|f_k - f_0\|_X$ or the inverse regularization parameter $\frac{1}{\alpha_k}$.

By the choice of the norm $\|\cdot\|_X$, we may define desirable properties of the object f to be reconstructed. Choosing the standard L^2 -norm here prevents isolated spikes, promoting more evenly distributed values. Typical images are expected to be of higher regularity, for example being composed of smoothly varying areas bounded by sharp edges. This may be exploited to obtain higher robustness to high-frequency errors by regularizing with Sobolev norms

$$\|f\|_X := \|(1 + \xi^2)^{\frac{\xi}{2}} \mathcal{F}(f)(\xi)\|_{L^2} \quad (10.2.8)$$

where \mathcal{F} denotes the Fourier transform and $\boldsymbol{\xi}$ the frequency coordinates. The exponent $s \geq 0$ tunes the degree of smoothness.

Beyond smoothness it is often desirable to impose additional constraints, which corresponds to restricting the set of admissible solutions $f_k \in C \subset X$. Prominent examples are real-valuedness, support constraints or positivity. Geometrically, the former two types are represented by linear subspaces $C \subset X$. Imposing these constraints in the IRGNM simply amounts to substituting $F'[f_k]^*$ with $\mathcal{P}F'[f_k]^*$ in (10.2.4), where $\mathcal{P} : X \rightarrow C$ is the orthogonal projection onto C . Positivity, on the other hand, is a nonlinear convex constraint. It may be included within a generalized Newton framework via a nonsmooth regularization term. In practice, this amounts to solving the minimization problem (10.2.3) restricted to $f \in C$, as can be done using semismooth Newton methods [94]. One approach to approximate sign constraints within the IRGNM framework of the present work lies in supplementing (10.2.3) with the penalty term

$$\gamma \|\min(0, f_k) - \min(0, \text{sign}(f_k))(f - f_k)\|_{L^2}^2, \quad (10.2.9)$$

which tends to correct negative values of f_k in the subsequent iterate. The coefficient $\gamma > 0$ determines the weight of the constraint and should be comparable to α_0 . To achieve strict positivity, one may let $\gamma \rightarrow \infty$ at constant α_k in the final iterates.

For numerical implementation of the IRGNM, all that needs to be done is to exchange the imaging operator F and its derivative F' in (10.2.3) by suitable discrete approximations. The norms $\|\cdot\|_{X_{\text{dis}}}$ and $\|\cdot\|_{Y_{\text{dis}}}$ in the discretized object- and image spaces $X_{\text{dis}} = \mathbb{R}^{N_X}$, $Y_{\text{dis}} = \mathbb{R}^{N_Y}$ are characterized by their Gramians $\mathbf{G}_X, \mathbf{G}_Y$ with respect to the Euclidean norm, i.e.

$$\|\mathbf{f}\|_{X_{\text{dis}}}^2 = \mathbf{f}^T \mathbf{G}_X \mathbf{f}, \quad \|\mathbf{I}\|_{Y_{\text{dis}}}^2 = \mathbf{I}^T \mathbf{G}_Y \mathbf{I} \quad (10.2.10)$$

where \mathbf{f}^T denotes the transpose. The adjoint of the Fréchet derivative $F'_{\text{dis}}[\mathbf{f}] : X_{\text{dis}} \rightarrow Y_{\text{dis}}$ can be implemented via

$$F'_{\text{dis}}[\mathbf{f}]^* = \mathbf{G}_X^{-1} F'_{\text{dis}}[\mathbf{f}]^T \mathbf{G}_Y. \quad (10.2.11)$$

The Hermitean positive-definite linear problem in the Newton step (10.2.3) can be solved efficiently by a conjugate gradients (CG) method. A major advantage of this IRGNM-CG scheme is that only the *forward* maps $F_{\text{dis}}, F'_{\text{dis}}[\mathbf{f}]$ and $F'_{\text{dis}}[\mathbf{f}]^T$ need to be evaluated, which are often easy to implement. Moreover, note that the matrices $F'_{\text{dis}}[\mathbf{f}]$ usually need not (and should not) be set up explicitly to compute the matrix-vector products $\mathbf{h} \mapsto F'_{\text{dis}}[\mathbf{f}]^{(\text{T})} \mathbf{h}$.

10.2.3 Newton-Kaczmarz methods

The Newton-type method presented so far is an *all-at-once* approach to image reconstruction: linearizations of the problem (10.2.1) are solved in each iteration, incorporating all constraints and the complete measured data I^{obs} . In view of 3d or even 4d tomographic or time-resolved imaging data, this results in computations with huge arrays, posing numerical challenges. For *linear* problems, a well-known remedy is given by Kaczmarz-type methods [109] such as the Algebraic Reconstruction technique in computed tomography [77, 155], which cyclically solve small under-determined subproblems. Indeed, many nonlinear imaging problems also lend themselves to a natural separation into different subproblems so that (10.2.1) becomes

$$(I_1^{\text{obs}}, \dots, I_p^{\text{obs}}) = (F_1, \dots, F_p)(f) + \boldsymbol{\epsilon}. \quad (10.2.12)$$

For problems of this form, [30] proposed regularized Newton-Kaczmarz methods, where in each iteration only one of the linearized subproblems $F_{j_k}(f_k) + F'_{j_k}[f_k](f_{k+1} - f_k) = I_{j_k}^{\text{obs}}$ is considered. A Kaczmarz-type equivalent of the IRGNM-update (10.2.3) takes the form

$$\begin{aligned} f_{k+1} = \operatorname{argmin}_{f \in X} & \left\| F_{j_k}(f_k) + F'_{j_k}[f_k](f - f_k) - I_{j_k}^{\text{obs}} \right\|_{Y_{j_k}}^2 \\ & + \alpha_k (\beta_k \|f - f_0\|_X^2 + (1 - \beta_k) \|f - f_k\|_X^2) \end{aligned} \quad (10.2.13)$$

The additional regularization term bounds the deviations from the preceding iterate f_k according to the weights $\beta_k \in [0; 1]$, ensuring that the reconstruction does not change too much within one Newton iteration. The sequence (j_k) determines the processing order of the different subproblems which may be adjusted to the requirements of the particular imaging problem.

10.3 Application to propagation-based phase contrast

Next, we apply the regularized Newton framework presented in section 10.2 to propagation-based phase contrast x-ray imaging. An exemplary experimental setup is sketched in fig. 10.1: quasi-monochromatic undulator radiation is focused by a pair of elliptical mirrors onto an x-ray waveguide. The exit of the waveguide serves as a quasi point source illuminating an object, which is placed at a distance of several mm behind the waveguide. The resulting diffraction patterns (holograms) are recorded by a detector at about 5m distance behind the focal plane, capturing the entire cone beam emanating from the waveguide with the sample induced interference pattern. The imaging data presented in this work has been recorded at the GINIX endstation at P10 beamline, DESY, Hamburg, described in [111, 178]. For details on the waveguide system and comparable high-resolution imaging results, see also [13].

For the imaging model, a fully coherent illumination of the sample by a plane wave is assumed. Note that we consider an effective parallel beam geometry equivalent to the experimental cone-beam setup. As in previous studies of cone beam x-ray propagation imaging [166, 145], we make use of the simple variable transformation, accounting for the geometric magnification, mapping the spherical beam illumination to an effective plane wave case. For the present imaging system, the data treatment is detailed by [13]. Moreover, we assume the object to be sufficiently thin and weak for the wave field to remain spherical and for the transmission to be well described by geometrical optics (projection approximation). This is typically well-satisfied for hard x-ray imaging [145, 160]. Under these assumptions, the intensity at the detector can be modeled as

$$I = F(f) := |\mathcal{D}(\exp(-f))|^2, \quad (10.3.1)$$

where $f = \mu + i\phi$ parametrizes the phase shifts ϕ and attenuation μ induced by the specimen, i.e. its image. \mathcal{D} denotes the Fresnel propagator, implementing free-space propagation of the transmitted parabolic wave field from sample onto the detector:

$$\mathcal{D}(\psi) := \mathcal{F}^{-1} \left(\exp \left(-\frac{i\xi^2}{4\pi\mathfrak{f}} \right) \mathcal{F}(\psi)(\xi) \right) \quad (10.3.2)$$

$\mathfrak{f} = b^2 k / (2\pi d)$ is the Fresnel number of the setup, where b is the size of a feature corresponding to unit length in the frequency vector ξ (typically the size of a pixel), d is the defocus distance and k is the x-ray wavenumber.

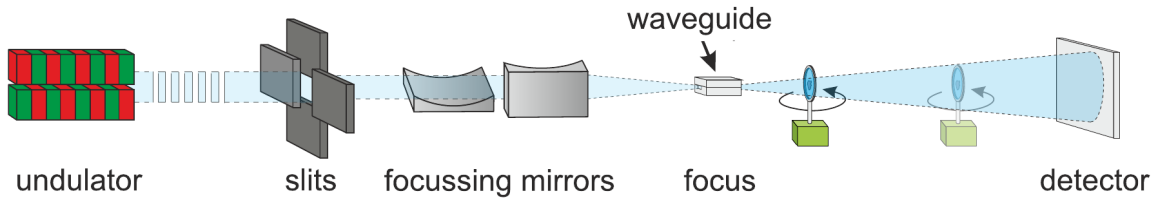


Figure 10.1. Schematic setup for propagation-based phase contrast imaging with hard x-rays: GINIX at P10 beamline, DESY [122, 178]. Quasi-monochromatic x-rays are focused onto a waveguide, illuminating a downstream object by a cone beam emanating from this coherent quasi-point source. Rotation of the sample allows for tomographic measurements.

From (10.3.1), we see that phase contrast imaging is nonlinear in general. Nevertheless, the model can be linearized for weak phase shifts and attenuation ϕ, μ , yielding the contrast transfer function (CTF) [79, 194]

$$\mathcal{F}(F(f) - 1) \approx -2 \sin\left(\frac{\xi^2}{4\pi f}\right) \mathcal{F}(\phi)(\xi) - 2 \cos\left(\frac{\xi^2}{4\pi f}\right) \mathcal{F}(\mu)(\xi) \quad (10.3.3)$$

Assuming vanishing attenuation $\mu \approx 0$ or single-material objects $\mu \propto \phi$, this formula can be directly inverted for image reconstruction. This approach yields satisfactory reconstructions especially if data from multiple defocus distances is available [42, 122].

From the CTF solution (10.3.3), it can be seen that regularized Newton methods are well-suited for phase contrast imaging in two respects: firstly, the zeros of the oscillatory prefactors make the image reconstruction problem ill-posed even in the simplest case of a weak non-absorbing object so that regularization is required. Secondly, the relative success of linear CTF-based methods indicate that the nonlinearity of the imaging method is sufficiently weak in a large regime of experimental setups.

For suitable spaces X, Y , the imaging operator $F : X \rightarrow Y$ defined by (10.3.1) can be shown to be Fréchet differentiable. Using linearity of \mathcal{D} and the expansions $\exp(x + h) = \exp(x)(1 + h) + \mathcal{O}(h^2)$ and $|x + h|^2 = |x|^2 + 2\text{Re}(\bar{x} \cdot h) + \mathcal{O}(h^2)$, we obtain the first-order Taylor approximation

$$F(f + h) = F(f) + -2\text{Re}\left(\overline{\mathcal{D}(\exp(-f))} \cdot \mathcal{D}(\exp(-f)h)\right) + \mathcal{O}(h^2). \quad (10.3.4)$$

This allows to read off the Fréchet derivative given by the linear term in h :

$$F'[f]h = -2\text{Re}\left(\overline{\mathcal{D}(\exp(-f))} \cdot \mathcal{D}(\exp(-f)h)\right). \quad (10.3.5)$$

Here the overbar denotes complex conjugation and Re the pointwise real part. The special case $f = 0$ in (10.3.5) reproduces the CTF. Implementing (10.3.1) and (10.3.5) for discrete images in the spaces $X := \mathbb{C}^{M_X \times N_X}$ and $Y := \mathbb{R}^{M_Y \times N_Y}$, using suitably padded fast Fourier transforms for the discrete Fresnel propagator, the reconstruction problem of propagation-based phase contrast imaging can be solved by the IRGNM of section 10.2.

As a proof of concept, we consider the diffraction pattern of an IRP logo, engraved into a thin gold film. The data has been recorded on the aforementioned GINIX setup at an energy of $E = 7.9 \text{ keV}$ is shown in fig. 10.2(a). The plotted intensity data has been normalized by the corresponding flat-field, i.e. the image of the empty beam, in order to meet the assumption of

plane wave illumination as justified by the analysis of [101]. The Fresnel number is $\mathfrak{f} = 1.77 \cdot 10^{-4}$ at an effective pixel size of $b = 21.7$ nm of the 1080×1920 -sized images and a maximum flux per pixel of ≈ 3400 photons.

For the Newton reconstruction, we choose the Sobolev norm in (10.2.8) with an exponent $s = 0.5$ as a regularization term and an L^2 -norm for the data residual. The regularization parameters α_k are chosen according to the procedure outlined in section 10.2.2. The algorithm is stopped as soon as the reduction of the residual $s_k - s_{k-1}$ achieved in the k -th iteration falls below 1% of the maximum decrease $\max_{j < k} s_j - s_{j-1}$. For the initial guess, we simply take $f_0 = 0$. We exploit that the specimen is made of gold by prescribing its characteristic between absorption μ and phase shifts ϕ of $c_{\mu/\phi} \approx 0.105$ [92]. The constraint is imposed by substituting $f = (i + c_{\mu/\phi})\phi$ and reconstructing the real-valued ϕ . Apart from this, neither positivity nor a support constraint is assumed. The latter means that the phase shifts ϕ are computed within the entire field of view of 1080×1920 pixels without any oversampling.

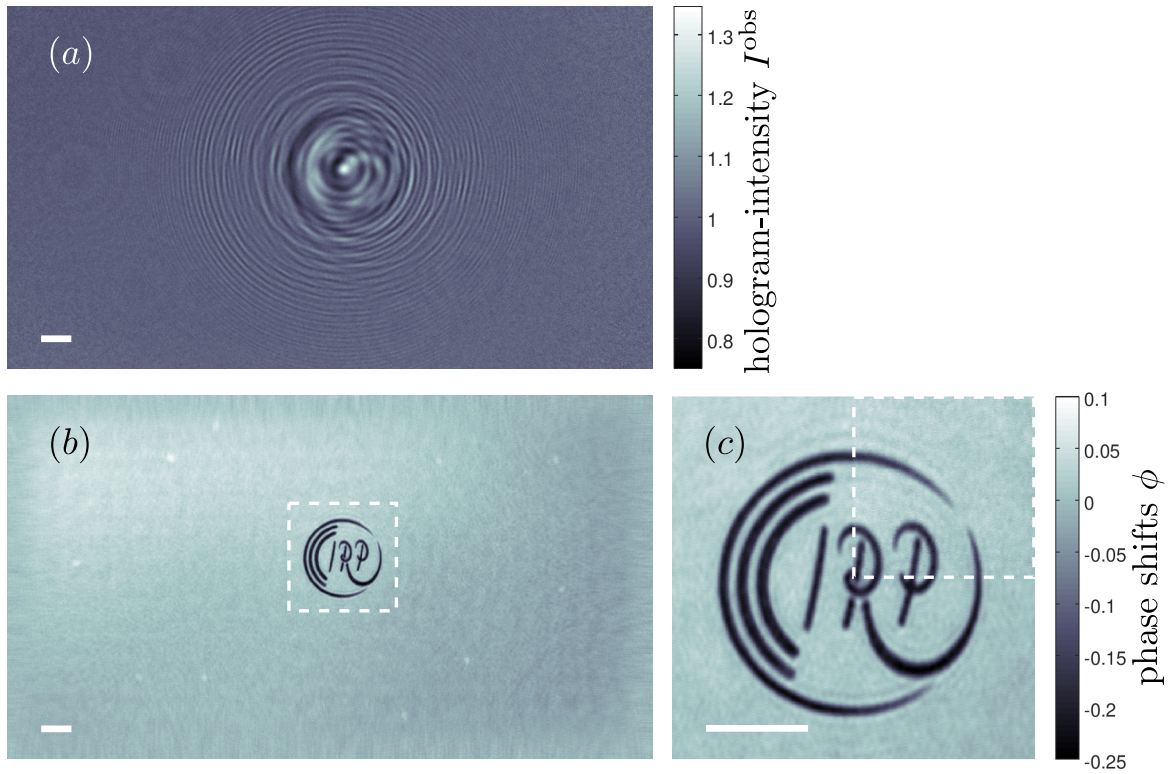


Figure 10.2. X-ray phase contrast imaging of a nano-structured object using the IRGNM reconstruction algorithm. The test pattern (institute logo) was defined in a thin gold film by focused ion beam milling. (a) Diffraction pattern recorded at GINIX endstation, P10 beamline, DESY [178] (flat-field corrected, maximum photon flux per pixel ≈ 3400). (b) Reconstructed phase image of the entire field of view, assuming a fixed ratio of 0.105 between absorption μ and phase shifts ϕ . (c) Magnification of (b) in the framed region around the logo. For comparison, the dashed inset shows the corresponding part of the phase map reconstructed by direct inversion of the CTF (10.3.3) via the methods of [122]. Scale bars: $2 \mu\text{m}$ in the effective geometry. Fringe artifacts on the upside of the logo are due to diffraction fringes leaving the field of view in (a). The phase shifts $\phi \approx 0.20$ induced by the object correspond to a thickness of the gold layer of about 100 nm.

Convergence of the IRGNM is reached after 12 Newton steps, corresponding to a total of 242 CG-iterations. The resulting phase map is plotted in fig. 10.2(b). Despite the lack

of oversampling, it can be seen that both the object and the background come out quite clean except for some stripes near the boundary due to the applied smooth replicate padding. Occasional white spots and the smooth background variations, on the other hand, may be attributed to some dirt in the imaging optics and physical variations of the thickness of the gold layer, respectively. The magnification of the reconstruction in a region around the IRP-logo depicted in fig. 10.2(c) confirms the high uniformity of the recovered phase shifts in the regions within and without the logo, reproducing the binary test pattern. Slight fringe artifacts can be identified in the vertical direction, especially on the upper edge of the logo. These may be explained by incompleteness of the data. Indeed, it can be seen from fig. 10.2(a) that parts of the diffraction pattern lie outside of the recorded field of view. The reconstruction of the missing object information is based on *a priori* constraints only, which naturally gives rise to artifacts.

It should be noted that the imaging setting considered so far still lies well within the regime of applicability of reconstruction method based on the CTF (10.3.3): for once, the phase shifts within the logo of ≈ 0.20 rad, induced by the 100 nm-thick gold layer of refractivity $k\delta \approx 1.96 \mu\text{m}^{-1}$ [92], are sufficiently weak for a global linearization of contrast formation to be reasonably accurate. Secondly, the assumption of a single material, i.e. of proportional fields ϕ and μ , allows for a direct inversion of (10.3.3). In the considered example, the latter approach indeed yields results of comparable quality as shown by the dashed inset in fig. 10.2(c). If ϕ and μ are considered as independent variables and only a single diffraction pattern is available, directly inverting (10.3.3) is impossible. This suggests that the imaging problem is non-unique in this case. On the contrary, theoretical investigation has recently shown that a unique recovery is indeed feasible if μ and ϕ are compactly supported [140]. The mathematical reason is that compactness of the support translates into correlations in Fourier space, which can be used to disentangle ϕ and μ in (10.3.3). By its ability to incorporate *a priori* constraints, the IRGNM approach may thus enable reconstructions beyond the limitations of direct inversion methods.

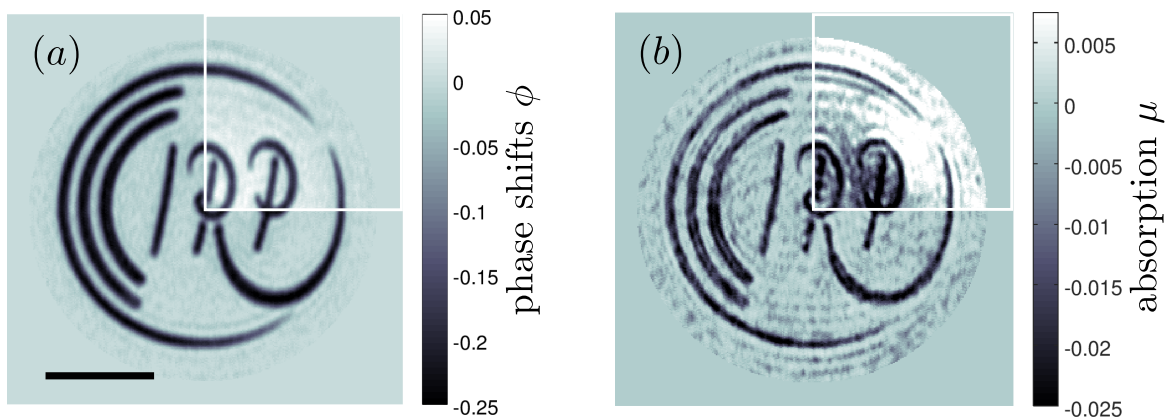


Figure 10.3. Simultaneous IRGNM reconstruction of phase shifts ϕ and absorption μ from the data in fig. 10.2(a) without assuming a fixed ratio μ/ϕ . Negative values of μ and ϕ indicate missing material in the gold film. The circular support visible in the phase- and absorption maps and negativity of μ and ϕ has been imposed as a constraint. All other parameters are retained as in the computation of fig. 10.2(b). The dashed inset in the absorption image shows the reconstruction without negativity constraint in the IRGNM and demonstrates that simultaneous recovery tends to introduce low-frequency artifacts. These are effectively suppressed by exploiting physical *a priori* knowledge on the sign of phase shifts and absorption. Scale bar: $2 \mu\text{m}$.

For a first experimental verification of a simultaneous recovery of both phase and absorption from a single hologram, we repeat the IRGNM reconstruction using the same parameters but without imposing a coupling of ϕ and μ . Following the uniqueness analysis, we impose a loose circular support constraint around the logo. Note that this induces a strong oversampling in the data by a factor of ≈ 17 , owing to the considerably reduced number of object pixels on which ϕ and μ have to be reconstructed. Moreover, negativity of μ and ϕ (the test pattern represents *missing* material, so that the contrast is inverted!) is imposed in the IRGNM via the penalty term approach outlined in section 10.2.2. The recovered phase and absorption obtained after 10 Newton steps (199 CG-iterations) is plotted in fig. 10.3. For comparison, dashed insets show the corresponding image parts obtained from a second reconstruction without negativity constraints.

Both the phase shifts ϕ in 10.3(a) and the attenuation coefficient μ plotted in 10.3(b) clearly represent the object shape, where the magnitudes roughly reproduce the material-specific ratio of $\mu/\phi \approx 0.1$. The only visible artifacts are those attributed to the missing fringes, which have been observed previously. In the IRGNM-reconstruction without negativity constraints, additional low-frequency errors appear especially in the recovered attenuation, as can be seen from the central dark spot surrounded by a bright halo in the inset of fig. 10.3(b). Simulations as well as preliminary analytical studies indicate that the susceptibility to these halo artifacts is characteristic of simultaneous retrieval of phase and absorption. As observed in the present reconstruction, however, they seem to be suppressed very effectively by imposing suitable sign constraints. Owing to the flexibility of exploiting physical *a priori* knowledge in the IRGNM, we thus indeed achieve an almost perfect simultaneous recovery of μ and ϕ from a single hologram - which has to date been considered impossible [157, 32].

Although proportionality of μ and ϕ is not imposed in the reconstruction in Fig. 10.3, the surprising quality might still be owing to some implicit preference of the IRGNM for homogeneous objects. We test this by repeating the reconstruction for a simulated object that is composed of two different materials. To this end, a binarized version of the recovered IRP logo with constant $\phi = 0.2$ and $\mu = 0.02$ is embedded into a purely phase-shifting disc with $\phi = 0.2$ and $\mu = 0$. All setup- and reconstruction parameters are chosen exactly as in Fig. 10.3. A single synthetic hologram (not shown) is simulated by mapping the exact object with the forward operator and superimposing Gaussian white noise with standard deviation $\sigma = 0.02$. The IRGNM reconstruction of ϕ and μ from this data is shown in Fig. 10.4 along with the exact solutions.

Good agreement with the exact object is found both in phase and absorption. The different object features are correctly attributed to the ϕ - and μ -components in Figs. 10.4(c) and (d) and no artifacts except for the known slight halo and fringe structures are visible. In particular, the absorption contrast enabled by the simultaneous recovery clearly reveals the logo structure despite the low signal-to-noise-ratio in μ . The logo is not represented in the phase image and thus likely invisible to any reconstruction method that assumes proportionality of μ and ϕ .

10.4 Phase contrast tomography of a colloidal crystal

By rotating the specimen within the incident beam in the imaging setup of fig. 10.1, propagation-based phase contrast can be extended to a tomographic imaging method, capable of resolving three-dimensional variations of the complex refractive index $n = 1 - \delta + i\beta$. Within the geo-

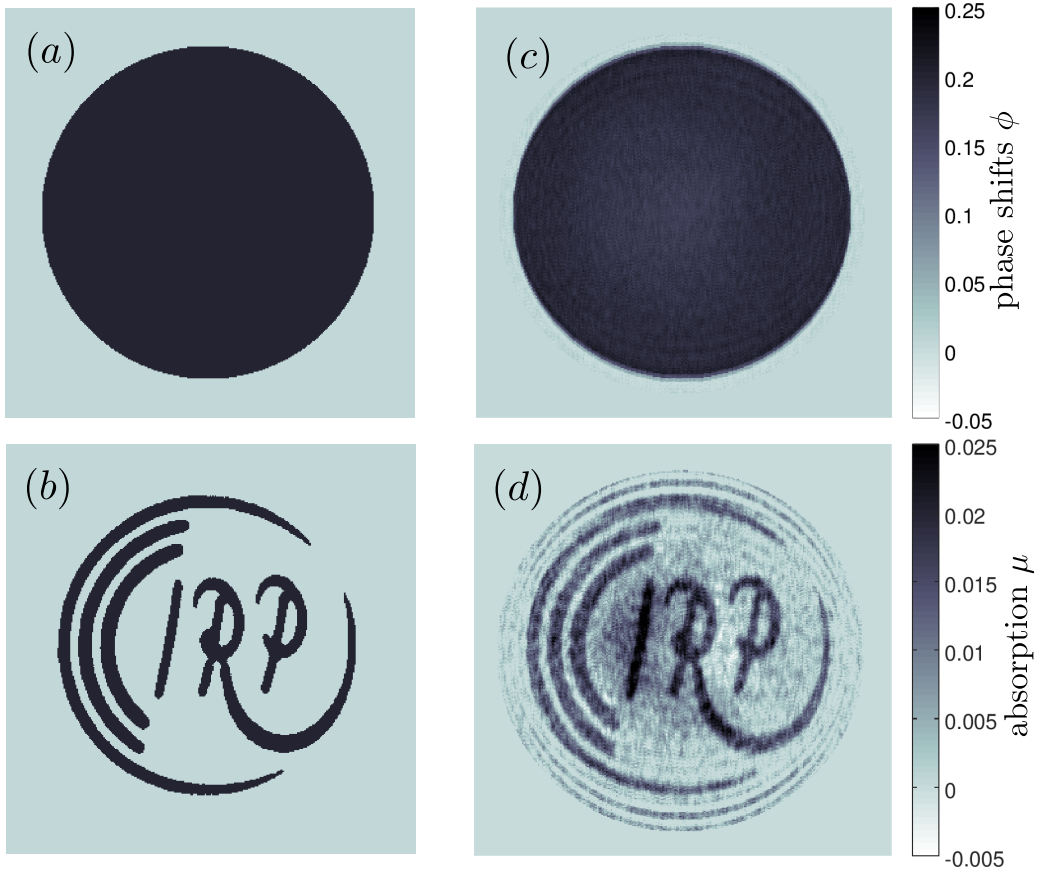


Figure 10.4. Supplementary IRGNM reconstruction of a non-homogeneous object from a simulated near-field hologram with 2% Gaussian white noise. Constraints, reconstruction- and setup parameters are chosen as for the experimental data in fig. 10.3. (a) and (b): Exact object composed of an absorbing logo structure with $\phi = 0.2$, $\mu = 0.02$ (gold) embedded into a non-absorbing disc that induces the same phase shifts $\phi = 0.2$ but with $\mu = 0$. (c) and (d): Recovered phase and absorption images. Except for a slight halo, features in ϕ and μ are correctly identified, revealing the hidden logo structure by the enabled absorption contrast.

metrical optics approximation of section 10.3, the imprinted phase- and absorption images ϕ and μ are proportional to the projection of δ and β , respectively, along the optical axis. For a tomographic incident angle θ , we have

$$\mu_\theta + i\phi_\theta = k \int (\beta_\theta + i\delta_\theta) dz = k\mathcal{P}_\theta(\beta_\theta + i\delta_\theta) \quad (10.4.1)$$

where δ_θ and β_θ denote the fields δ and β in a rotated coordinate system and \mathcal{P}_θ is the parallel-beam projector (X-ray transform) mapping onto the line integrals along θ . Combining (10.3.1) and (10.4.1), we obtain a forward operator F mapping the object function $f = k\beta + ik\delta$ onto the corresponding intensity data I under *all* of the $N \in \mathbb{N}$ different incident angles θ_j :

$$I = F(f) := |\mathcal{D}(\exp(-\mathcal{P}_{\text{tot}}(f)))|^2, \quad \mathcal{P}_{\text{tot}} := (\mathcal{P}_{\theta_1}, \mathcal{P}_{\theta_2}, \dots, \mathcal{P}_{\theta_N}) \quad (10.4.2)$$

Since \mathcal{P}_{tot} is linear and bounded on suitable function-spaces, F is Fréchet differentiable with a derivative of similar form as (10.3.5).

Implementing (10.4.2) in a regularized Newton-type framework corresponds to an *all-at-once* approach to phase contrast tomography, as phase- and tomographic reconstruction are not carried out in subsequent steps but simultaneously. In particular, this implies that tomographic correlations between images under different incident angles are incorporated already in the phase retrieval step. Similarly as in coherent diffractive imaging [37, 16], this has been shown to improve stability and accuracy of the reconstruction [175, 139]. Unfortunately, a numerical inversion of F by an IRGNM is not viable for high resolution data sets due to memory constraints and high computational costs of the evaluation of the Radon transform.

As a remedy, we propose to solve the reconstruction problem of phase contrast tomography by a regularized Newton-Kaczmarz methods as presented in section 10.2.3. The tomographic setting suggests a decomposition of the forward operator $F = (F_1, \dots, F_p)$, corresponding to small sub-data sets of diffraction patterns from only very few incident angles. Each Newton step (10.2.13) then only requires evaluations of comparably small Radon transforms and processing of feasible chunks of data. At the same time, tomographic consistency is exploited by directly reconstructing a three-dimensional object instead of single projections. Furthermore, the F_j may be chosen such that the most strongly correlated images for only slightly differing incident angles are processed simultaneously.

We implement the Newton-Kaczmarz step (10.2.13) for the imaging operator (10.4.2), using the simple parameter choice $\alpha_k = \alpha_0$ and $\beta_k = \beta_0 \ll 1$, where α_0 is estimated according to section 10.2.2. Firstly, this regularization imposes the current reconstruction as a strong prior for the next iteration. Secondly, it also yields a favorable condition number for the linear Newton steps so that CG-methods typically converge after 3-5 iterations.

For a proof of concept, we consider phase contrast images of a colloidal crystal composed of 415 nm-sized polystyrene beads, imaged at an energy $E = 7.9$ keV in the experimental setup of fig. 10.1. The tomographic data set is composed of 249 diffraction patterns of size 1024×1024 acquired at a maximum photon flux of 770 per pixel under incident angles θ between 0 and 173° . The raw data correction included an alignment of the tomographic projections to the common center of mass and iterative reprojections, following the approach of [210]. The photon flux corresponds to an upper bound of 110 kGy for the absorbed dose over the total tomographic data acquisition. The flat-field corrected intensities are visualized in fig. 10.5(a). In the effective geometry, the Fresnel number is $\mathfrak{f} = 2.41 \cdot 10^{-4}$ at an effective pixel size of 29.3 nm.

From the fringes in the data set in fig. 10.5(a) as well as preliminary reconstructions, one can infer the object location in the center of the field of view. This is exploited by restricting the reconstruction to a central 256^3 voxel cube, imposing a loose 3d-support constraint. Moreover, the hydrocarbon composition of the polystyrene spheres and the photon energy allow using $\beta = 0$ as a constraint (non-absorbing object), i.e. to reconstruct only the refractive increment δ . Unfortunately, a Sobolev-type regularization term (10.2.8) would bottleneck the proposed Newton-Kaczmarz-scheme as the required FFTs in the Gramian would constitute the only $\mathcal{O}(N^3 \log N)$ operations in the Newton updates. We therefore recur to simple L^2 -regularization and L^2 -data fidelity terms, corresponding to identity Gramians $\mathbf{G}_X, \mathbf{G}_Y$ in (10.2.10). However, positivity of δ is imposed by supplementing the Kaczmarz-Newton step (10.2.13) with the penalty term in (10.2.9). The weights of the penalty terms are set to $\gamma = \alpha_0$ and $\beta_k = 0.001$. For the initial guess, we simply choose $f_0 = 0$. The forward operator F is decomposed such that angular “wedges” of six adjacent diffraction patterns are processed simultaneously in each Newton step. The number of iterations is determined such that each hologram is visited

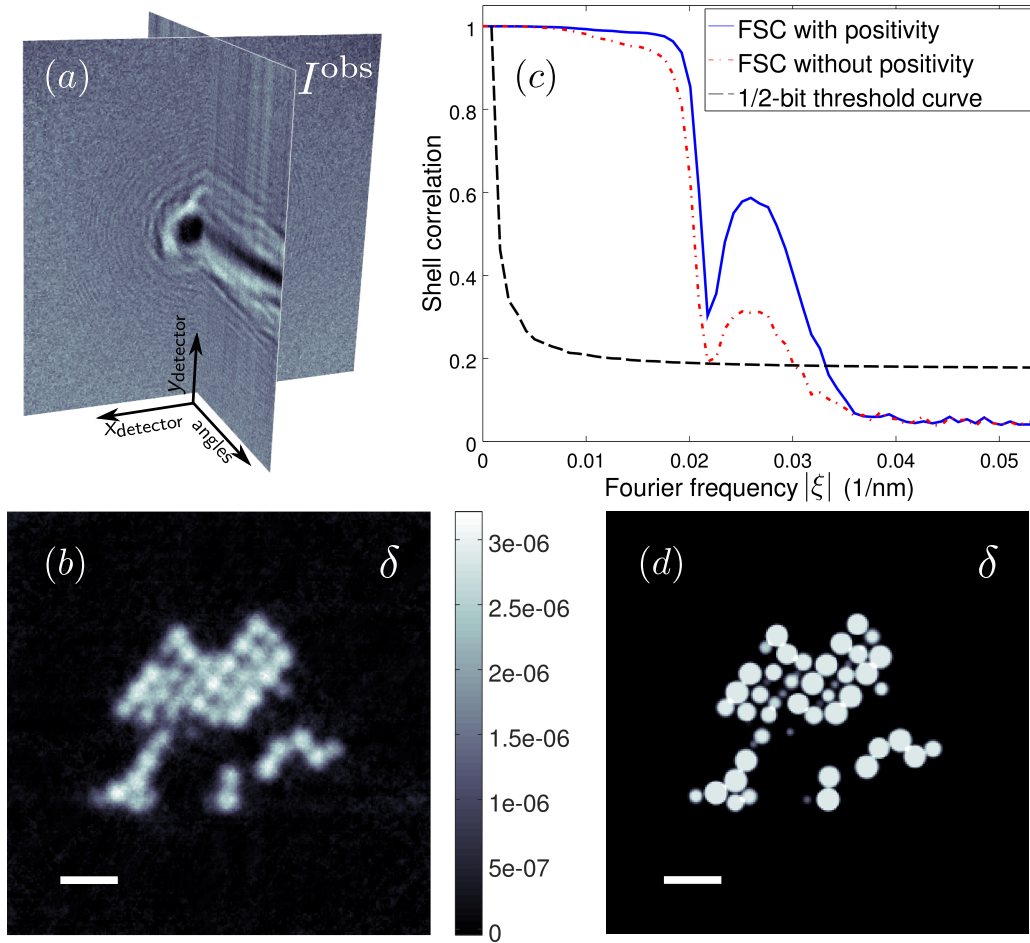


Figure 10.5. X-ray phase contrast tomography of a colloidal micro-crystal composed of 415 nm-size polystyrene beads, reconstructed by a regularized Newton-Kaczmarz method. The maximum photon flux is 770 per pixel corresponding to an absorbed dose ≤ 110 kGy. (a) Slice visualization of the tomographic data given by 1024×1024 flat-field corrected holograms measured under 249 incident angles between 0 and 173° at GINIX endstation, P10 beamline, DESY [178]. (b) Central slice of the reconstructed 256^3 voxel volume, plotting the increment δ of the refractive index $n = 1 - \delta + i\beta$ which is proportional to electron density. *A priori* constraints were positivity $\delta \geq 0$ and vanishing absorption $\beta = 0$. The scale bar is $1 \mu\text{m}$. (c) Fourier shell correlation (FSC) computed for reconstructions from complementary data sets of 125 and 124 incident angles (green curve). For comparison, the blue curve shows the result without positivity constraint and the red dashed one plots the 1/2-bit threshold curve [201]. The intersection between the green- and the red curve indicates an achieved resolution of 95 nm. (d) Binary representation of the slice in (b), determined by deconvolving the reconstructed δ with the form factor of a 415 nm-sized homogeneous sphere after Gaussian filtering.

twice where the processing order of the wedges is random to minimize directional bias in the reconstructed object.

The reconstruction runs on a simple laptop with Intel i7 CPU and 8 Gigabytes of RAM, computing for about 20 minutes. Each of the Newton steps takes four CG-iterations, which means that the total solution - processing every incident angle twice - requires only 10 forward and adjoint operations of the full imaging operator F . Figure 10.5(b) shows the central slice of the reconstructed δ perpendicular to the tomographic axis. It can be seen that the binary

character of the specimen and the spherical shapes of the beads are well-resolved, being clearly distinguishable both from one another and from the background. Notably, no artifacts caused by angular undersampling or due to the missing wedge of 6° can be identified.

We estimate the resolution by computing the Fourier shell correlation (FSC) [200] of two auxiliary reconstructions from complementary sets of 125 and 124 incident angles, respectively. In fig. 10.5(c), the resulting correlation is compared to the 1/2-bit threshold curve as proposed by [201]. For comparison, the FSC is also computed for a reconstruction without positivity constraint. The beneficial effect of the latter can be seen from the gap between the blue and green curves in fig. 10.5(c), showing a significantly higher correlation especially in the resolution-critical part around the intersections with the threshold curve. The intersections indicate a resolution of about 95 nm and 105 nm (half of the corresponding Fourier wavelengths) for the reconstructions with and without positivity constraint, respectively. Notably, the local minima of the FSC curves at $|\xi| \approx 0.022 \text{ nm}^{-1}$ are no artifacts but neatly coincide with the first order zero of the form factor of 415 nm-sized spheres. This emphasizes the sensitivity of the imaging method to structural features.

According to the slice plotted in fig. 10.5(b), the positions of the polystyrene spheres and thus the crystal structure could be determined in principle by visual inspection of the reconstructed δ . In order to determine the colloid locations numerically and independent of an operator, we smoothen the recovered object by a Gaussian of 95 nm FWHM to reduce the impact of noise before deconvolving with the form factor of a homogeneous 415 nm-sized sphere (with regularization around the zeros in Fourier space). This procedure results in Gaussian peaks centered at the positions of polystyrene beads. By computing the maxima of this field using quadratic interpolation about the peaks, we thus obtain a representation of the crystal structure.

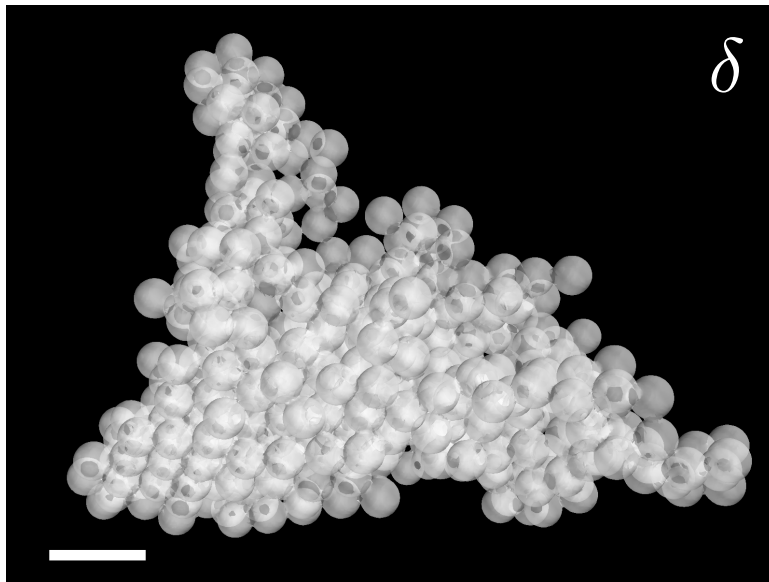


Figure 10.6. 3d-rendering of a colloidal crystal of polystyrene nano-beads reconstructed by Newton-based phase contrast tomography. The underlying binary object, a slice of which is shown in fig. 10.5(d), has been determined by deconvolving the smoothened reconstruction of the refraction parameter δ with the form factor of a homogeneous sphere of diameter 415 nm. The coordinates of the colloid sites are provided in Data File 1. Scale bar: $1 \mu\text{m}$.

By the simple deconvolution procedure outlined above, 448 colloid positions are identified. Convoluting the obtained Dirac delta-array of bead locations with the ideal 415 nm-sized sphere yields a binary representation of the colloidal crystal. A slice of this binarization corresponding to fig. 10.5(b) is plotted in 10.5(d). Figure 10.6 shows a 3d-rendering of the determined colloidal crystal. The corresponding site coordinates of the colloids, i.e. the centers of the spherical beads, are provided in Data File 1. The imaged micro-crystal contains regions of (approximately) hexagonal close-packing as well as cubic and amorphous regions, induced by an interplay of bulk- and surface effects.

10.5 Conclusions

In this work, we have presented iteratively regularized Gauss-Newton methods (IRGNM) as a generic approach to solve nonlinear ill-posed image reconstruction problems. The principal idea is to reconstruct an unknown object by iteratively inverting linearizations of a known imaging model, which describes contrast formation from object to observable data. In order to compensate for missing information, the iterates are computed to provide an optimal compromise between the measurements and additional *a priori* information on the unknown object. The IRGNM approach differs from well-known alternating-projection-type algorithms typically used in CDI in that it exploits differentiability and *simultaneously* processes constraints and observed data. This promises improved convergence.

By applying regularized Newton methods to near-field phase contrast imaging, both in 2d and in a tomographic 3d setting, we have demonstrated their flexibility in treating different experimental setups. The reconstruction of a nano-structured test pattern shows that IRGNM constitute a reasonable generalization of direct inversion methods based on a *global* linearization of contrast formation. Owing to its potential to incorporate moderate nonlinearity and *a priori* knowledge, e.g. on support or positivity, the approach permits faithful reconstructions beyond the scope of such direct methods. This has been demonstrated by the simultaneous recovery of magnitude and phase of the test object from a single diffraction pattern without assuming proportionality of phase shifts and absorption. The validity of the reconstruction is supported by results for a simulated object composed of two materials, which are correctly identified by the algorithm. However, numerical simulations also indicate that the stability of the simultaneous recovery deteriorates significantly with increasing Fresnel number. Thus, it is possibly only feasible for deeply holographic near field data as considered in this work.

A further benefit of the presented regularized Newton framework is that it requires only a *forward* model for the imaging setup and no explicit knowledge of (approximate) inverses. This enables IRGNM reconstructions also for complicated multi-staged imaging setups, possibly including the influence of various experimental parameters. In the considered example of phase contrast tomography this flexibility permits to directly recover a probed 3d-object from the complete tomogram of diffraction patterns instead of subsequently performing (2d) phase retrieval and tomographic backprojection. We have demonstrated the potential of this approach by imaging a colloidal crystal composed of 415 nm-sized polystyrene beads at 95 nm resolution, reconstructed by an efficient regularized Newton-Kaczmarz method.

X-ray imaging techniques currently reach out for novel frontiers, ranging from new contrast modalities in material science and dose reduction in biomedical analysis to in-operando studies and investigations of highest spatial and temporal resolution by FEL pulses. To enable

these goals, the mathematical modeling underlying image reconstruction has to become increasingly accurate in order not to bottleneck the achievable contrast and resolution: apart from nonlinearity of image formation and Poissonian noise statistics, future imaging models may need to better account for partial coherence, non-uniform illumination, mechanical vibrations and movements, misalignment or geometrical aberrations, just to name a few. Owing to their flexibility, we are convinced that regularized Newton methods, as presented in this work, may greatly foster such developments and thus contribute to a bring lensless x-ray techniques as well as other cutting-edge imaging methods to the level of quantitative structural measurements.

Acknowledgments

We thank Janne-Mieke Meijer and Andrei Petukhov for providing us with a colloidal crystal film some time ago, from which we could later prepare the micro-crystal used here as a test object. MB acknowledges John Miao for helpful discussions regarding the alignment of the colloidal data set. Support by the German Research Foundation DFG through project C02 of the Collaborative Research Center SFB 755 - Nanoscale Photonic Imaging is gratefully acknowledged.

Chapter 11

Complete Article 4: Generalized SART-Methods for Tomographic Imaging

Reproduced from the manuscript [141]:

Maretzke S. *Inverse Problems* (in review), 2018.

Preprint: [arXiv:1803.04726](https://arxiv.org/abs/1803.04726)

Abstract: Nowadays, the field computed tomography (CT) encompasses a large variety of settings, ranging from nanoscale to meter-sized objects imaged by different kinds of radiation in various acquisition modes. This experimental diversity challenges the flexibility of tomographic reconstruction methods. Kaczmarz-type methods, which exploit the natural block-structure of tomographic inverse problems, are a promising candidate to provide the required versatility in a computationally efficient manner. In the present work, it is shown that indeed a surprisingly general class of tomographic Kaczmarz-iterations may be efficiently evaluated via computational schemes of a similar structure as updates of the so-called simultaneous algebraic reconstruction technique (SART). This enables regularized reconstructions with non-trivial image-formation models as well as non-quadratic or even non-convex data-fidelity terms at low computational costs. Moreover, the proposed generalized SART schemes are equally applicable in parallel- and cone-beam settings and regardless of the choice of tomographic incident directions. Their potential is illustrated by outlining applications in several non-standard tomographic settings, including polychromatic CT and X-ray phase contrast tomography.

11.1 Introduction

Since the pioneering works of Cormack [44, 45] and Hounsfield [103], the field of computed tomography (CT) has broadened considerably. While classical CT based on the partial attenuation of X-rays by matter continues to be a principal workhorse of medical diagnosis, several other applications have emerged over the past decades: for instance, state-of-the-art transmission-electron-microscopes (TEM) may resolve unknown structures in three dimensions

down to sub-nanometer resolutions by acquiring a series TEM-images under different incident directions of the electron-beam [148, 159]. Moreover, the advent of coherent X-ray sources has enabled *phase contrast* techniques and thereby extended the scope of X-ray tomography to quasi nonabsorbing micro- and nanoscale objects such as single biological cells or viruses [42, 122, 80, 56, 13]. Other CT applications rely on gamma-rays or even cosmic myons to image strongly attenuating objects such as oil pipelines [107] or ancient pyramids [152].

All of the above settings come with their own peculiarities: In electron tomography for instance, tomographic views may be typically only acquired for a few incident directions in a limited range due to radiation damage and geometrical restrictions. In phase contrast tomography, the acquired data is given by diffraction patterns, that relate to the actual projections of the object’s refractivity in a highly non-trivial manner. Moreover, even the seemingly standard application of medical CT involves a generally nonlinear inverse problem. Disregarding this nonlinearity may lead to severe artifacts associated with so-called *beam-hardening* or *photon-starvation* effects [10, 151].

To address this variety of different settings, there exist for once a large number of different preprocessing methods that aim to make the data “ready” for reconstruction by standard algorithms such as filtered backprojection (FBP) or its approximate analogue for cone-beam tomography, the FDK-algorithm [64]. While this enables fast reconstruction, the applied pre-corrections are often heuristic and limited in their effectivity. For this reason, variational reconstruction methods have received increased attention recently, see e.g. [184, 18, 211, 181, 183, 119, 31]. These start from an image-formation model for the considered setting, supplement it with additional *a priori* knowledge on the unknown object and expected data-errors and reconstruct by minimizing a cost functional that incorporates all of this information. While this approach is much more flexible in accounting for the peculiarities of a specific imaging modality, it suffers from computational complexity: typically, the optimization has to be carried out iteratively where each of the many iterations requires $\mathcal{O}(N^4)$ arithmetic operations – as much as a full FBP- or FDK-reconstruction. In current high-resolution CT applications with $N \gtrsim 10^3$ sampling-points along each dimension, this constitutes a major bottleneck.

Kaczmarz-type- or block-iterative methods like the (simultaneous) algebraic reconstruction technique (S)ART [77, 6] may provide a compromise between the flexibility of variational methods and the favorable complexity of FBP. By fitting only small *data-blocks* in each iteration, Kaczmarz-updates can be computed at lower computational costs than bulk-iterations on the complete data. SART, for example, updates the reconstructed object by matching its projection to the measured data under one tomographic incident direction per step. Moreover, Kaczmarz-methods are often observed to exhibit fast *semi-convergence* [155, 57], arriving at accurate reconstructions already after $\mathcal{O}(1)$ fitting-cycles over the tomographic data set. If the iterations are sufficiently cheap to compute, this enables image recovery at overall computational costs comparable to FBP or FDK. Recent variants of Kaczmarz-type schemes may also incorporate advanced priors such as total-variation-penalties via interlacing gradient-descent- or proximal iterations [21, 7, 52] or block-structured primal-dual-methods [35].

In the present work, it is shown that a very general class of regularized Kaczmarz-iterations (also called “Tikhonov-Kaczmarz” [50, 115] or “incremental proximal” iterations [22, 7] by other authors) for tomographic problems may be evaluated via efficient computational schemes of a similar structure as classical SART-steps. The approach is therefore named *generalized SART* (GenSART). The underlying idea is that – as long as the object is fitted to exactly *one* tomo-

graphic view per iteration – the computed updates will typically be uniform along the direction of the tomographic rays, i.e. a *back-projection* of some increment in the lower-dimensional projection-space. This reasoning holds true for both parallel- and cone-beam settings, a large variety of data-fidelity functionals, complicated image-formation models such as diffraction operators and different regularization terms. The potential of GenSART is demonstrated by outlining applications in Poisson-noise-adapted and outlier-robust tomographic reconstruction, X-ray phase contrast tomography as well as for a polychromatic CT model.

This manuscript is organized as follows: §11.2 introduces the basic tomographic imaging model and gives some background on Kaczmarz-type reconstruction methods. In §11.3, the generalized SART principle is presented in an abstract setting, which is shown to be applicable to the considered Kaczmarz-iterations in §11.4. §11.5 outlines several applications of the proposed GenSART-schemes, for some of which numerical examples are presented in §11.6.

11.2 Background

11.2.1 Tomographic imaging model

We consider general tomographic inverse problems, for which the dependence of the data g_{tot} from the sought object f can be modeled as

$$g_{\text{tot}} = \begin{pmatrix} g_1 \\ \vdots \\ g_{N_{\text{proj}}} \end{pmatrix} = \begin{pmatrix} F_1(P_1(f)) \\ \vdots \\ F_{N_{\text{proj}}}(P_{N_{\text{proj}}}(f)) \end{pmatrix} = F_{\text{tot}}(P_{\text{tot}}(f)) \quad (11.2.1)$$

P_j are parallel-beam- or cone-beam projectors acting on a 3D-object density $f \in L^2(\Omega)$ ($L^2(\Omega) := \{f : \mathbb{R}^3 \rightarrow \mathbb{R} : \|f\|_{L^2} < \infty, \text{supp}(f) \subset \Omega\}$, $\|h\|_{L^2}^2 := \int |h(x)|^2 dx$) with support $\text{supp}(f)$ in a bounded open domain $\Omega \subset \mathbb{R}^3$. The setting is sketched in fig. 11.1. Note that (11.2.1) constitutes a *semi-continuous* tomography model in that f and the $g_1, \dots, g_{N_{\text{proj}}}$ are modeled as continuous images, but the number of tomographic views $N_{\text{proj}} \in \mathbb{N}$ is finite.

In a *parallel-beam* tomography setup, each $P_j = \mathcal{P}_{\theta^{(j)}}$ maps f onto its line integrals along a certain incident direction $\theta \in \mathbb{S}^2 := \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = 1\}$:

$$\mathcal{P}_{\theta}(f)(\mathbf{x}_{\perp}) := \int_{\mathbb{R}} f(x\mathbf{n}_x + y\mathbf{n}_y + z\theta) dz \quad \text{for all } \mathbf{x}_{\perp} = (x, y) \in \mathbb{R}^2, \quad (11.2.2)$$

where $\mathbf{n}_x, \mathbf{n}_y \in \mathbb{S}^2$ are chosen such that $\{\theta, \mathbf{n}_x, \mathbf{n}_y\}$ forms an orthonormal basis of \mathbb{R}^3 .

In a *cone-beam* tomography setup modeled by $P_j = \mathcal{D}_{\mathbf{s}^{(j)}}$, a single projection is given by line integrals along rays emanating from a point-source $\mathbf{s} \in \mathbb{R}^3 \setminus \bar{\Omega}$ along all directions:

$$\mathcal{D}_{\mathbf{s}}(f)(\varphi) := \int_0^{\infty} f(\mathbf{s} + t\varphi) dt \quad \text{for all } \varphi \in \mathbb{S}^2. \quad (11.2.3)$$

The maps F_j in (11.2.1) denote (possibly nonlinear) *image-formation operators*, modeling the relation between the raw projections $p_j = P_j(f)$ and the data g_j that is actually detected. For instance, the choice $F_j(p_j) = I_0 \cdot \exp(-p_j)$ models absorption-based X-ray tomography with monochromatic illumination of known intensity distribution I_0 . More complicated operators can be derived to model polychromatic CT or phase contrast tomography, for example.

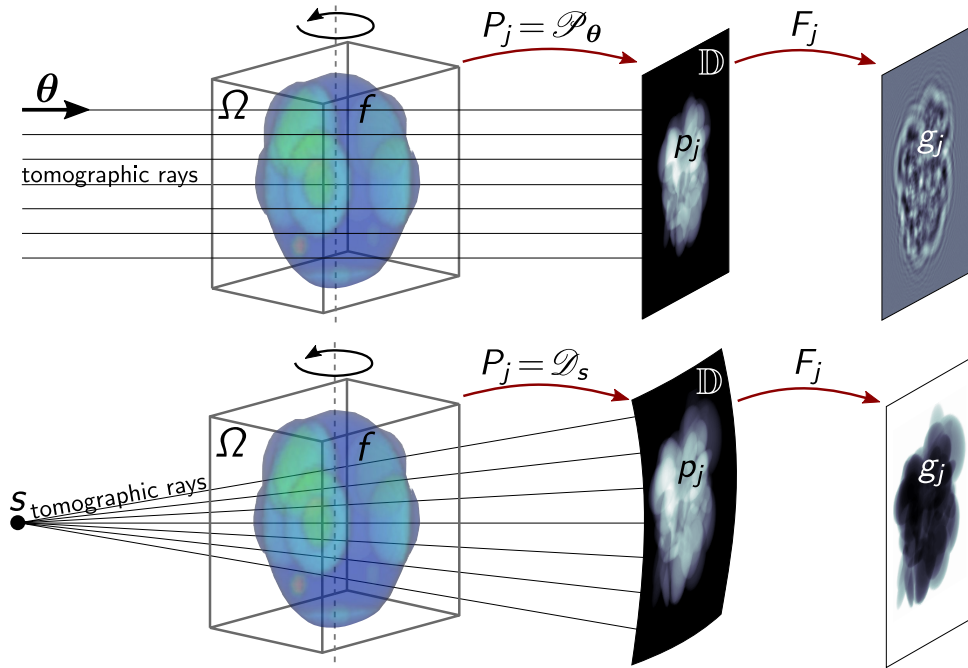


Figure 11.1. Sketch of the considered tomographic model: the measured data g_j is given by a finite number of tomographic projections $p_j = P_j(f)$ of an unknown 3D-object f , mapped under additional image-formation operators F_j . Top: parallel-beam example with diffractive image-formation (phase contrast). Bottom: cone-beam setup with absorption-contrast as in conventional CT.

11.2.2 Inverse problem and a priori constraints

This work is concerned with reconstructing the an unknown object f from tomographic data g_{tot} modeled by (11.2.1), i.e. in solving the following inverse problem:

Inverse Problem 11.1 (Tomographic reconstruction). For a given domain $\Omega \subset \mathbb{R}^3$, parallel- or cone-beam projectors $P_1, \dots, P_{N_{\text{proj}}}$ and image-formation operators $F_1, \dots, F_{N_{\text{proj}}}$, reconstruct an object $f \in L^2(\Omega)$ from noisy data

$$g_j^{\text{obs}} = F_j(P_j(f)) + \epsilon_j \quad \text{for } j = 1, \dots, N_{\text{proj}}. \tag{11.2.4}$$

As Inverse Problem 11.1 is typically ill-posed even in the trivial case where the F_j are just given by the identity, one typically aims to exploit available *a priori* knowledge on the unknown density f to be reconstructed. Common a priori constraints include:

- *Support constraints:* f is known to be identically zero outside $\Omega \subset \mathbb{R}^3$. Note that such constraints are already incorporated in the tomographic imaging model from §11.2.1.
- *Box constraints:* the admissible values of f may be bounded from below and/or above. For example, it is often physically justified to assume *non-negativity*, $f \geq 0$.
- *Regularity constraints:* A certain smoothness of f is assumed in the reconstruction, for example by imposing that the total variation seminorm $\int_{\Omega} |\nabla f| dx$ is small.

In addition to incorporating such constraints, one often seeks to account for an (approximately) known statistics of the data errors ϵ .

11.2.3 Reconstruction methods

Variational methods: The widely used filtered back-projection- (FBP) and Feldkamp-Davis-Kress (FDK) reconstruction algorithms often lack flexibility to accurately account for specific tomographic settings and available a priori knowledge. As a remedy, variational methods have been proposed. The idea is to minimize a *generalized Tikhonov functional*:

$$f^{\text{rec}} \in \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \mathcal{S}_{\text{tot}}(g_{\text{tot}}^{\text{obs}}; F_{\text{tot}}(P_{\text{tot}}f)) + \mathcal{R}_{\text{tot}}(f). \quad (11.2.5)$$

The data-fidelity functional $\mathcal{S}_{\text{tot}}(g_{\text{tot}}; \cdot)$ may account for the expected data-error-statistics, while the penalty functional \mathcal{R}_{tot} allows to impose a priori knowledge on the object.

The main drawback of variational methods lies in their computational costs. If the minimizer in (11.2.5) is computed by generic (convex) optimization methods, such as primal-dual methods [36, 183], fast iterative soft shrinkage [18] or the alternating direction method of multipliers [70, 181], many iterations are typically needed for convergence. Moreover, each iteration usually involves an evaluation of the full tomographic projector P_{tot} and its adjoint P_{tot}^* and thus amounts to similar computational costs as a *complete reconstruction* by FBP or FDK.

Kaczmarz-type methods: One approach to decrease the number of expensive evaluations of P_{tot} and P_{tot}^* compared to (bulk) variational methods is to exploit the block-structure of Inverse Problem 11.1 by performing cyclic iterations on the sub-problems $g_j^{\text{obs}} = F_j(P_j(f)) + \epsilon_j$. We consider such *Kaczmarz-* or *block-iterative-*methods in a general variational form

$$f_{k+1} \in \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \mathcal{S}_k(g_{j_k}^{\text{obs}}; F_{j_k}(P_{j_k}(f))) + \mathcal{R}_k(f), \quad j_k \in \{1, \dots, N_{\text{proj}}\}. \quad (11.2.6)$$

The processing order $\{j_k\}$ is typically chosen as an integer number of *cycles* over the whole data, given by permutations of the indices $\{1, \dots, N_{\text{proj}}\}$. Provided that the functionals $\mathcal{S}_k, \mathcal{R}_k$ in (11.2.6) relate to $\mathcal{S}_{\text{tot}}, \mathcal{R}_{\text{tot}}$ in (11.2.5) in a suitable manner, Kaczmarz-schemes may yield similar reconstructions as the associated bulk variational methods, as will be seen in §11.2.4.

The formulation (11.2.6) generalizes classical Kaczmarz-iterations, see e.g. [155, §V.3]:

$$\begin{aligned} f_{k+1} &= \underset{f \in B_k}{\operatorname{argmin}} \|f - f_k\|_{L^2}^2 \quad \text{with} \quad B_k := \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \|P_{j_k}(f) - g_{j_k}\|_{L^2}^2 \\ &= \lim_{\alpha \rightarrow 0} \left(\underset{f \in L^2(\Omega)}{\operatorname{argmin}} \|P_{j_k}(f) - g_{j_k}\|_{L^2}^2 + \alpha \|f - f_k\|_{L^2}^2 \right) \end{aligned} \quad (11.2.7)$$

The well-known *algebraic reconstruction technique* (ART) [109, 77] is an analogue of this approach applied to a fully forward model, where the projectors P_{j_k} are replaced by single rows of a projector matrix. On the other hand, *simultaneous ART* (SART), proposed in [6] as a heuristic approach to stabilize ART by simultaneously updating all matrix-rows corresponding to a single tomographic projection, can be interpreted as the discretization of an analytical formula for the Kaczmarz-iterate in (11.2.7). This will be seen in §11.3.2.

11.2.4 Convergence of Kaczmarz-iterations and relation to Tikhonov regularization

Classical Kaczmarz-iterations of the form (11.2.7) are known to exhibit fast *semi-convergence*, typically yielding a regularized reconstruction within $\mathcal{O}(1)$ cycles while increasingly amplifying

data-noise if more iterations are preformed, see [57] and references therein. Recently, Kindermann and Leitão [115] obtained a much more concise characterization of convergence for *regularized* (Tikhonov-)Kaczmarz-iterations, based on previous work by Elfving and Nikazad [58]. One of their principal results (cf. [115, Theorem 3]) can be cast to the following form:

Theorem 11.1 (Kindermann & Leitão [115]). *Let X, Y_1, \dots, Y_N be Hilbert spaces, $f_0 \in X$ and $g^{\text{obs}} = (g_1^{\text{obs}}, \dots, g_N^{\text{obs}}) \in Y_{\text{tot}} := \prod_j Y_j$. Moreover, let $A_j : X \rightarrow Y_j$ linear-bounded and $A_{\text{tot}} := (A_1, \dots, A_N)^T : X \rightarrow Y_{\text{tot}}$. Then, after a symmetric Kaczmarz-cycle,*

$$f_{k+1} = \begin{cases} \operatorname{argmin}_{f \in X} \|A_{k+1}f - g_{k+1}^{\text{obs}}\|_{Y_j}^2 + \alpha \|f - f_k\|_X^2 & \text{for } k < N \\ \operatorname{argmin}_{f \in X} \|A_{2N-k}f - g_{2N-k}^{\text{obs}}\|_{Y_j}^2 + \alpha \|f - f_k\|_X^2 & \text{for } k \geq N \end{cases} \quad (11.2.8)$$

for $k = 0, 1, \dots, 2N - 1$, the final iterate minimizes a bulk Tikhonov-functional:

$$f_{2N} = \operatorname{argmin}_{f \in X} \|W \cdot (A_{\text{tot}}f - g^{\text{obs}})\|_Y^2 + \frac{\alpha}{2} \|f - f_0\|_X^2 \quad (11.2.9)$$

$$W = \begin{pmatrix} \operatorname{id} + \frac{1}{2\alpha} A_1 A_1^* & & & 0 \\ & \ddots & & \\ 0 & & \operatorname{id} + \frac{1}{2\alpha} A_N A_N^* & \end{pmatrix}^{1/2} \begin{pmatrix} \operatorname{id} & -\frac{1}{\alpha} A_1 A_2^* & \dots & -\frac{1}{\alpha} A_1 A_N^* \\ & \ddots & \ddots & \vdots \\ 0 & & \ddots & -\frac{1}{\alpha} A_{N-1} A_N^* \\ & & & \operatorname{id} \end{pmatrix}^{-1}$$

Theorem 11.1 is applicable to the considered Kaczmarz-iterations in (11.2.6) if the data-fidelities \mathcal{S}_k are quadratic, the maps F_j are linear and $\mathcal{R}_k(f) = \alpha \|f - f_k\|_X^2$ for some Hilbert-space-norm $\|\cdot\|_X$. For this case, it states that already *one* symmetric Kaczmarz-cycle essentially yields a minimizer of the associated bulk Tikhonov functional, up to a slight modification of the data-fidelity term by the operator W in (11.2.9). Notably, W deviates from the identity only by contributions of order α^{-1} . This implies that the result is not applicable to the limit $\alpha \rightarrow 0$ of the classical Kaczmarz-method (11.2.7). On the contrary, for sufficiently large α , theorem 11.1 shows that regularized Kaczmarz-iterations may be used to approximate Tikhonov-minimizers, i.e. to *emulate* bulk variational methods – at least in the considered quadratic setting. Other convergence results pointing in a similar direction are given in [21, 22, 7].

11.2.5 Contribution

In order for Kaczmarz-methods to provide a truly efficient alternative to bulk variational approaches, essentially two conditions have to be satisfied:

1. The iterates f_k arrive at a reasonable reconstruction after few cycles (ideally one).
2. The individual Kaczmarz-iterations may be evaluated at low computational costs.

The convergence theory reviewed in §11.2.4 indicates that condition 1 is often fulfilled, in agreement with empirical observations for Kaczmarz-reconstructions in practice. The present article, on the other hand, does *not* focus on convergence but contributes to meeting condition 2 by proposing computationally efficient solution-formulas, termed generalized SART-schemes, for Kaczmarz-iterations of the general variational form (11.2.6).

11.3 The Generalized SART-Principle

11.3.1 Preparations: notation and analysis of the projectors

In order to analyze the Kaczmarz-iterations (11.2.6), we need some properties of the projection operators $P \in \{\mathcal{P}_\theta, \mathcal{D}_s\}$ for a single incident direction θ or source position s , see §11.2.1. First of all, we note that different values of θ and s merely correspond to a rotation and shift of the coordinate system, respectively. We may therefore restrict our analysis without loss of generality to the cases $\theta = (0, 0, 1)^\top$ and $s = \mathbf{0}$, i.e. to the operators

$$\mathcal{P}(f)(\mathbf{x}_\perp) := \int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} f(\mathbf{x}_\perp, z) dz, \quad \mathcal{D}(f)(\varphi) := \int_{L_\varphi^{\mathcal{D}}} f(t\varphi) dt \quad \text{for } \mathbf{x}_\perp \in \mathbb{R}^2, \varphi \in \mathbb{S}^2 \quad (11.3.1)$$

where we have defined the *ray-segments* intersecting Ω in the parallel- or cone-beam setting:

$$L_{\mathbf{x}_\perp}^{\mathcal{P}} := \{z \in \mathbb{R} : (\mathbf{x}_\perp, z) \in \Omega\}, \quad \mathbb{D}_{\mathcal{P}} := \{\mathbf{x}_\perp \in \mathbb{R}^2 : L_{\mathbf{x}_\perp}^{\mathcal{P}} \neq \emptyset\} \quad (11.3.2a)$$

$$L_\varphi^{\mathcal{D}} := \{t \in [0; \infty) : t\varphi \in \Omega\}, \quad \mathbb{D}_{\mathcal{D}} := \{\varphi \in \mathbb{S}^2 : L_\varphi^{\mathcal{D}} \neq \emptyset\} \quad (11.3.2b)$$

By definition, $\mathcal{P}(f)$ and $\mathcal{D}(f)$ vanish outside the *projection-domains* $\mathbb{D}_{\mathcal{P}} \subset \mathbb{R}^2, \mathbb{D}_{\mathcal{D}} \subset \mathbb{S}^2$ for any $f \in L^2(\Omega)$. Moreover, as Ω is open, so are the sets $L_{\mathbf{x}_\perp}^{\mathcal{P}}, L_\varphi^{\mathcal{D}} \subset \mathbb{R}$ and $\mathbb{D}_{\mathcal{P}} \subset \mathbb{R}^2, \mathbb{D}_{\mathcal{D}} \subset \mathbb{S}^2$. To enable a unified treatment of parallel- and cone-beam settings, we set $\mathbb{D} := \mathbb{R}^2$ if $P = \mathcal{P}$ and $\mathbb{D} := \mathbb{S}^2$ if $P = \mathcal{D}$.

We define the *ray-density functions* w_P and (*weighted*) *unit-projections* u_P, \tilde{u}_P :

$$w_P : \Omega \rightarrow \mathbb{R}_{>0}; \mathbf{x} \mapsto \begin{cases} 1 & \text{for } P = \mathcal{P} \\ |\mathbf{x}|^{-2} & \text{for } P = \mathcal{D} \end{cases}, \quad u_P := P(\mathbf{1}_\Omega), \quad \tilde{u}_P := P(w_P) \quad (11.3.3)$$

where $\mathbf{1}_\Omega \in L^2(\Omega)$ is the constant 1-function in Ω . Note that $\tilde{u}_P = u_P$ for $P = \mathcal{P}$, whereas $\tilde{u}_P \neq u_P$ for $P = \mathcal{D}$, and that $\mathbb{D}_P = \{x \in \mathbb{D} : u_P(x) > 0\} = \{x \in \mathbb{D} : \tilde{u}_P(x) > 0\}$.

Back-projections: We introduce the *back-projections* corresponding to \mathcal{P} and \mathcal{D} :

$$\mathcal{P}^B(p)(\mathbf{x}_\perp, z) := \begin{cases} p(\mathbf{x}_\perp) & \text{if } (\mathbf{x}_\perp, z) \in \Omega \\ 0 & \text{else} \end{cases}, \quad \mathcal{D}^B(p)(\mathbf{x}) := \begin{cases} p(\mathbf{x}/|\mathbf{x}|) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{else} \end{cases} \quad (11.3.4)$$

$\mathcal{P}^B(p)$ and $\mathcal{D}^B(p)$ are constant along all rays $L_{\mathbf{x}_\perp}^{\mathcal{P}}$ or $L_\varphi^{\mathcal{D}}$, which are exactly the integration-domains in the definition (11.3.1) of \mathcal{P} and \mathcal{D} , respectively. In particular, this implies that

$$P(w \cdot P^B(p)) = P(w) \cdot p \quad \text{for } P \in \{\mathcal{P}, \mathcal{D}\} \quad (11.3.5)$$

and all w, p for which this expression makes sense.

Spaces: We study the maps \mathcal{P}, \mathcal{D} as operators between spaces of square-integrable functions $L^2(\Omega) := \{f \in L^2(\mathbb{R}^3) : \text{supp}(f) \subset \Omega\}$ and $L^2(\mathbb{D}_P) := \{p \in L^2(\mathbb{D}) : \text{supp}(p) \subset \mathbb{D}_P\}$ ($\mathbb{D}_P \subset \mathbb{D}$ as defined in (11.3.2)), with the usual inner product $\langle f, g \rangle_{L^2} = \int_\Omega f(x)g(x) dx$ and corresponding norm $\|f\|_{L^2} := \langle f, f \rangle_{L^2}^{1/2}$. For simplicity, we restrict to real-valued functions throughout this work, although all results naturally carry over to the complex-valued case. To avoid confusion, it is crucial to note that our notation will follow convention 11.2:

Convention 11.2. All local operations $(+, -, \cdot, /, (\cdot)^\gamma, \nabla, \dots)$ on functions $f \in L^2(\Omega) \subset L^2(\mathbb{R}^3)$ or $p \in L^2(\mathbb{D}_P) \subset L^2(\mathbb{D})$ are implicitly understood to be performed only within the open domains Ω or \mathbb{D}_P . For example, a quotient of $p_1, p_2 \in L^2(\mathbb{D}_P)$ is to be read as

$$(p_1/p_2)(x) = \begin{cases} p_1(x)/p_2(x) & \text{for } x \in \mathbb{D}_P \\ 0 & \text{else} \end{cases} \quad \text{for all } x \in \mathbb{D}. \quad (11.3.6)$$

Adjoints: For $P \in \{\mathcal{P}, \mathcal{D}\}$, it can be shown that $P : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$ is a bounded linear operator. The adjoints are given by (weighted) back-projections (see e.g. [155, 133]):

$$P^* : L^2(\mathbb{D}_P) \rightarrow L^2(\Omega); P^*(p) = w_P \cdot P^B(p) \quad \text{for } P \in \{\mathcal{P}, \mathcal{D}\}. \quad (11.3.7)$$

Notably, the relation (11.3.5) applied to (11.3.7) yields

$$PP^*(p) = P(w_P \cdot P^B(p)) = \tilde{u}_P \cdot p \quad \text{for all } p \in L^2(\mathbb{D}_P). \quad (11.3.8)$$

Remark 11.3. The formula (11.3.8) is of high computational value: it states that, while evaluations of the operators P and P^* *alone* may be complicated and costly to compute, the composition PP^* can be implemented as a simple pointwise multiplication of projections. This observation is a key ingredient to an efficient computation of Kaczmarz-iterates.

Geometric characterization: In the analysis, we will need to consider *weighted* projectors:

$$P_{\text{iso}} : f \mapsto \tilde{u}_P^{-1/2} \cdot P(f) \quad \text{for } P \in \{\mathcal{P}, \mathcal{D}\}. \quad (11.3.9)$$

Note that the expression is well-defined by convention 11.2 since $\tilde{u}_P(x) > 0$ for all $x \in \mathbb{D}_P$. At the first glance, the definition of P_{iso} may still seem artificial. Yet, it enables a compact geometric characterization of the projectors, which is probably common knowledge at least for $P = \mathcal{P}$, see e.g. [155, § V.4.3]. The proof of the following result is given in appendix 11.A:

Theorem 11.4 (Geometry of the projectors). *For $P \in \{\mathcal{P}, \mathcal{D}\}$, $P_{\text{iso}} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$ is bounded with norm 1 and the adjoint $P_{\text{iso}}^* : L^2(\mathbb{D}_P) \rightarrow L^2(\Omega); p \mapsto w_P \cdot P^B(\tilde{u}_P^{-1/2} \cdot p)$ is isometric. In particular, P_{iso} and P_{iso}^* have closed range, $P_{\text{iso}}^* P_{\text{iso}}$ is the orthogonal projection onto $\text{range}(P_{\text{iso}}^*)$ and $P_{\text{iso}} P_{\text{iso}}^* = \text{id}_{L^2(\mathbb{D}_P)}$ is the identity on $L^2(\mathbb{D}_P)$.*

11.3.2 L^2 -SART: a promising example

As a motivation for the general result presented in the subsequent section, we consider an example of Kaczmarz-iterations, that turn out to be computable via a simple analytical formula. Let the iterates be defined by

$$f_{k+1} \in \underset{f \in L^2(\Omega)}{\text{argmin}} \|P_{j_k}(f) - g_{j_k}^{\text{obs}}\|_{L^2}^2 + \alpha \|f - f_k\|_{L^2}^2 \quad (11.3.10)$$

with regularization parameter $\alpha > 0$ and $g_{j_k}^{\text{obs}} \in L^2(\mathbb{D}_P)$. The unique minimizer of this variational problem is given by the solution to the corresponding normal equation:

$$\begin{aligned} f_{k+1} - f_k &= (P_{j_k}^* P_{j_k} + \alpha)^{-1} P_{j_k}^* (g_{j_k}^{\text{obs}} - P_{j_k}(f_k)) \\ &= P_{j_k}^* (P_{j_k} P_{j_k}^* + \alpha)^{-1} (g_{j_k}^{\text{obs}} - P_{j_k}(f_k)). \end{aligned} \quad (11.3.11)$$

The second equality in (11.3.11) uses an identity from functional calculus, see e.g. [60]. According to (11.3.8), $(P_{j_k} P_{j_k}^* + \alpha)^{-1}$ is a division by $\tilde{u}_{j_k} + \alpha$ (with $\tilde{u}_{j_k} := \tilde{u}_{P_{j_k}}$) so that (11.3.11) yields

$$f_{k+1} = f_k + P_{j_k}^* \left(\frac{g_{j_k}^{\text{obs}} - P_{j_k}(f_k)}{\tilde{u}_{j_k} + \alpha} \right), \quad (11.3.12)$$

where all arithmetic operations are understood to be pointwise.

Relation to SART: By exchanging the continuous object density f_k , projection data $g_{j_k}^{\text{obs}}$, projector P_{j_k} and unit-projection \tilde{u}_{j_k} in (11.3.12) with suitable discretizations $\mathbf{P}_{j_k} \in \mathbb{R}^{m \times n}$, $\mathbf{f}_k \in \mathbb{R}^n$, $\tilde{\mathbf{u}}_{j_k}, \mathbf{g}_{j_k}^{\text{obs}} \in \mathbb{R}^m$, a numerically implementable update-formula is obtained:

$$\mathbf{f}_{k+1} = \mathbf{f}_k + \mathbf{P}_{j_k}^* \left((\mathbf{g}_{j_k}^{\text{obs}} - \mathbf{P}_{j_k} \mathbf{f}_k) \oslash (\tilde{\mathbf{u}}_{j_k} + \alpha) \right) \quad (11.3.13)$$

where \oslash denotes element-wise division of vectors. Notably, the iterate \mathbf{f}_{k+1} from (11.3.13) is in general *not* a solution to the discrete analogue of (11.3.10), i.e. typically

$$\mathbf{f}_{k+1} \notin \underset{\mathbf{f} \in \mathbb{R}^n}{\text{argmin}} \left\| \mathbf{P}_{j_k} \mathbf{f} - \mathbf{g}_{j_k}^{\text{obs}} \right\|_2^2 + \alpha \left\| \mathbf{f} - \mathbf{f}_k \right\|_2^2, \quad (11.3.14)$$

because $\mathbf{P}_{j_k} \mathbf{P}_{j_k}^*$ – unlike $P_{j_k} P_{j_k}^*$ – is not diagonal for standard discretizations. Interestingly however, the update (11.3.13) is closely related to SART, that was originally derived without reference to the continuous model. The classical SART-update [6] can be written as

$$\mathbf{f}_{k+1} = \mathbf{f}_k + \mathbf{P}_{j_k}^* \left((\mathbf{g}_{j_k}^{\text{obs}} - \mathbf{P}_{j_k} \mathbf{f}_k) \oslash \mathbf{P}_{j_k}(\mathbf{1}) \right) \oslash \mathbf{P}_{j_k}^*(\mathbf{1}), \quad (11.3.15)$$

where $\mathbf{1} = (1, 1, \dots, 1)^T$ denotes one-vectors of suitable length. In the parallel-beam case $P_{j_k} = \mathcal{P}$, one has $\tilde{u}_{j_k} = P_{j_k}(\mathbf{1}_\Omega)$ and $P_{j_k}^*(\mathbf{1}_{\mathbb{D}_P}) = \mathbf{1}_\Omega$ by (11.3.3) and (11.3.4). The term $\mathbf{P}_{j_k}(\mathbf{1})$ in (11.3.15) can thus be identified with $\tilde{\mathbf{u}}_{j_k}$ in (11.3.13) and the division by $\mathbf{P}_{j_k}^*(\mathbf{1})$ is just redundant from the perspective of the continuous model. Accordingly, the SART-update (11.3.15) essentially corresponds to the limit $\alpha \rightarrow 0$ of (11.3.13). Since (11.3.13) has been derived as a discretization of (11.3.10), SART (11.3.15) may thus be interpreted as a formula to compute the classical Kaczmarz-iterations in (11.2.7). Conversely, (11.3.13) can be seen as an L^2 -regularized SART-variant.

The SART-Scheme: Analogously to classical SART, the update-formula (11.3.13) computes the Kaczmarz-iterate in (11.3.10) via a highly efficient *non-iterative* scheme, that requires only a single evaluation of \mathbf{P}_{j_k} and $\mathbf{P}_{j_k}^*$ each:

Scheme 11.5 (L^2 -regularized SART).

- (1) Forward-project the current iterate: $\mathbf{p}_k = \mathbf{P}_{j_k}(\mathbf{f}_k)$
- (2) Compute increment in projection space: $\Delta \mathbf{p}_k = (\mathbf{g}_{j_k} - \mathbf{p}_k) \oslash (\tilde{\mathbf{u}}_{j_k} + \alpha)$
- (3) Back-project and increment: $\mathbf{f}_{k+1} = \mathbf{f}_k + \mathbf{P}_{j_k}^*(\Delta \mathbf{p}_k)$

Notably, the actual data-fitting step **(2)** works exclusively on *2D-projections* $\mathbf{g}_{j_k}, \mathbf{p}_k, \tilde{\mathbf{u}}_{j_k} \in \mathbb{R}^m$, which are typically discretized by $m = \mathcal{O}(N^2)$ pixels and thus *low-dimensional* compared to the discrete *3D-objects* $\mathbf{f}_k \in \mathbb{R}^n$ with $n = \mathcal{O}(N^3)$ (N sampling-points per spatial dimension). This renders step **(2)** cheap to compute, so that the total computational costs of scheme 11.5 are essentially that of the evaluations of \mathbf{P}_{j_k} and $\mathbf{P}_{j_k}^*$ in steps **(1)** and **(3)**. Combined with the fast (semi-)convergence of Kaczmarz-methods in $\mathcal{O}(1)$ cycles (see §11.2.4), this implies that L^2 -SART-schemes allow to compute reconstructions at a favorable computational complexity of $\mathcal{O}(1)$ evaluations of the full projector $\mathbf{P}_{\text{tot}} = (\mathbf{P}_1, \dots, \mathbf{P}_{N_{\text{proj}}})$ and $\mathbf{P}_{\text{tot}}^*$, compare §11.2.3.

11.3.3 Generalized SART framework

In §11.3.2, it has been shown that L^2 -Kaczmarz-iterations (11.3.10) can be evaluated in a simple and efficient manner although they involve a seemingly complicated optimization problem. Motivated by this result, we explore in how far more general Kaczmarz-iterations permit an efficient computation analogous to the SART-like scheme 11.5. For notational convenience, we drop the subscripts in (11.2.6) at this point and abbreviate the data-fidelity as $\mathcal{S}(p) := \mathcal{S}_k(g_{j_k}^{\text{obs}}; F_{j_k}(p))$, absorbing the maps F_{j_k} into the functional. Thereby, the considered Kaczmarz-iterations are cast to the generic form

$$f_{\text{new}} \in \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \mathcal{S}(P(f)) + \mathcal{R}(f) = \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \tilde{\mathcal{S}}(\tilde{P}(f)) + \mathcal{R}(f). \quad (11.3.16)$$

On the right-hand side of (11.3.16), it has been used that (P, \mathcal{S}) may be replaced by any pair $(\tilde{P}, \tilde{\mathcal{S}})$ of modified projectors and data-fidelities, provided that $\tilde{\mathcal{S}}(\tilde{P}(f)) = \mathcal{S}(P(f))$ for all $f \in L^2(\Omega)$. We have to exploit this freedom since our mathematical framework will require \tilde{P} to have *closed range*, which is not satisfied for $\tilde{P} = P \in \{\mathcal{P}, \mathcal{D}\} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$, but only for suitably weighted versions like $\tilde{P} = P_{\text{iso}}$, compare theorem 11.4. We consider the optimization problem in (11.3.16) for a general bounded linear operator $\tilde{P} : X \rightarrow Y$ on Hilbert spaces X, Y . In this setting, we find that the key ingredient to computing (11.3.16) via a SART-like scheme is the following assumption on the geometrical compatibility of \tilde{P} and \mathcal{R} :

Assumption 11.6. Let $\tilde{P} : X \rightarrow Y$ be a bounded linear operator on Hilbert spaces X, Y with null-space $\operatorname{kern}(\tilde{P}) = \{f \in X : \tilde{P}(f) = 0\}$ such that $\operatorname{range}(\tilde{P}) = \tilde{P}(X) \subset Y$ is closed and let $\mathcal{R} : X \rightarrow \mathbb{R} \cup \{\infty\}$ be a functional. Assume that there exists an $f_{\text{ref}} \in X$ such that

$$\mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p) + f_0) \geq \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) \quad \text{for all } p \in Y, f_0 \in \operatorname{kern}(\tilde{P}). \quad (\text{A})$$

Assumption 11.6 ensures that the penalty functional \mathcal{R} uniformly penalizes deviations from a certain reference element f_{ref} by elements from the null-space of \tilde{P} . If this condition is satisfied, (11.3.16) can be evaluated according to our principal theorem:

Theorem 11.7 (Generalized SART-principle). *Let assumption 11.6 be satisfied and let $\tilde{\mathcal{S}} : Y \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Assume that there exists a minimizer*

$$f_{\text{new}} \in \underset{f \in X}{\operatorname{argmin}} \tilde{\mathcal{S}}(\tilde{P}(f)) + \mathcal{R}(f). \quad (11.3.17)$$

Then there is a (possibly distinct) minimizer $\tilde{f}_{\text{new}} \in X$ of (11.3.17) given by

$$p_{\text{ref}} = \tilde{P}(f_{\text{ref}}) \quad (11.3.18a)$$

$$\Delta p \in \operatorname{argmin}_{p \in Y} \tilde{\mathcal{S}}(p_{\text{ref}} + \tilde{P}\tilde{P}^*(p)) + \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) \quad (11.3.18b)$$

$$\tilde{f}_{\text{new}} = f_{\text{ref}} + \tilde{P}^*(\Delta p) \quad (11.3.18c)$$

Conversely, any \tilde{f}_{new} given by (11.3.18) minimizes (11.3.17). Furthermore, if strict inequality holds in (A) whenever $f_0 \neq 0$, then all minimizers of (11.3.17) are of the form (11.3.18).

Proof. Let $\Delta f := f_{\text{new}} - f_{\text{ref}}$. Since \tilde{P} has closed range, the same holds true for the adjoint \tilde{P}^* by the closed-range-theorem. Hence, there is an orthogonal decomposition $X = \operatorname{range}(\tilde{P}^*) \oplus_{\perp} \operatorname{kern}(\tilde{P})$ so that, in particular, there exist $f_0 \in \operatorname{kern}(\tilde{P})$ and $\Delta p \in Y$ such that $\Delta f = \tilde{P}^*(\Delta p) + f_0$. Now define $\tilde{f}_{\text{new}} := f_{\text{ref}} + \tilde{P}^*(\Delta p)$. Then (A) implies that

$$\mathcal{R}(\tilde{f}_{\text{new}}) = \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) \leq \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p) + f_0) = \mathcal{R}(f_{\text{new}}). \quad (11.3.19)$$

Moreover, it holds that $\tilde{\mathcal{S}}(\tilde{P}(\tilde{f}_{\text{new}})) = \tilde{\mathcal{S}}(\tilde{P}(f_{\text{new}}))$ since $\tilde{f}_{\text{new}} - f_{\text{new}} \in \operatorname{kern}(\tilde{P})$. Thus,

$$\tilde{\mathcal{S}}(\tilde{P}(\tilde{f}_{\text{new}})) + \mathcal{R}(\tilde{f}_{\text{new}}) \leq \tilde{\mathcal{S}}(\tilde{P}(f_{\text{new}})) + \mathcal{R}(f_{\text{new}}) \quad (11.3.20)$$

where the left-hand side is strictly smaller if $\tilde{f}_{\text{new}} \neq f_{\text{new}}$ under the additional assumption that strict inequality holds in (A) for $f_0 \neq 0$. This proves that \tilde{f}_{new} is a minimizer and that $\tilde{f}_{\text{new}} = f_{\text{new}}$ must hold in the case of strict inequality. Moreover, Δp satisfies

$$\begin{aligned} \tilde{P}^*(\Delta p) &\in \operatorname{argmin}_{\Delta f \in X} \tilde{\mathcal{S}}(\tilde{P}(f_{\text{ref}} + \Delta f)) + \mathcal{R}(f_{\text{ref}} + \Delta f) \\ \Rightarrow \Delta p &\in \operatorname{argmin}_{p \in Y} \tilde{\mathcal{S}}(\tilde{P}(f_{\text{ref}}) + \tilde{P}\tilde{P}^*(p)) + \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)), \end{aligned} \quad (11.3.21)$$

which proves (11.3.18). For the converse statement, let \tilde{f}_{new} be given by (11.3.18). Then \tilde{f}_{new} minimizes the cost-functional $\mathcal{C}(f) := \tilde{\mathcal{S}}(\tilde{P}(f)) + \mathcal{R}(f)$ over all $f \in A := f_{\text{ref}} + \operatorname{range}(\tilde{P}^*)$. Since assumption 11.6 implies that $\mathcal{C}(f_{\text{new}} + f_0) \geq \mathcal{C}(\tilde{f}_{\text{new}})$ for all $f_0 \in \operatorname{kern}(\tilde{P})$, \tilde{f}_{new} also minimizes \mathcal{C} over $X = A + \operatorname{kern}(\tilde{P})$ and hence solves (11.3.17). \square

We aim to apply theorem 11.7 to the tomographic Kaczmarz-iterations in (11.3.16). Although stated in an abstract manner, the result is particularly well-suited for this application:

- (a) The optimization problem (11.3.18b) is on the image-space Y of \tilde{P} (*projection-space*), which is much lower-dimensional than the object-space X in the considered setting.
- (b) The operator $\tilde{P}\tilde{P}^*$ appearing in (11.3.18b) is trivial to evaluate for tomographic projectors $\tilde{P} \in \{\mathcal{P}_{\text{iso}}, \mathcal{D}_{\text{iso}}\}$, contrary to \tilde{P} or \tilde{P}^* alone (see remark 11.3 and theorem 11.4).

To apply theorem 11.7, *literally nothing* has to be assumed on the functional $\tilde{\mathcal{S}}$, so that the result offers complete freedom in the choice of the data-fidelity \mathcal{S} in (11.3.16). What remains is to verify assumption 11.6 for $\tilde{P} \in \{\mathcal{P}_{\text{iso}}, \mathcal{D}_{\text{iso}}\}$ (or related choices) and suitable penalties \mathcal{R} . This is established in §11.4. We refer to the resulting formulas of the kind (11.3.18) as *generalized SART-* or *GenSART-schemes*, aluding to their structural similarity to scheme 11.5.

11.4 Admissible penalty functionals

The aim of this section is to identify penalty functionals $\mathcal{R} = \mathcal{R}_k$ that satisfy assumption 11.6, in which case the Kaczmarz-iterations in (11.2.6) can be computed via GenSART-schemes by virtue of theorem 11.7.

11.4.1 Preliminary insights

In order to gain an intuition for the admissible penalties, let us first discuss the meaning of the condition (A). It asserts that – relative to a certain reference object f_{ref} – any deviation by an element from the null-space $\text{kern}(\tilde{P})$ is penalized or at least does not decrease the value of \mathcal{R} . For (weighted) parallel-beam projectors $\tilde{P} = \mathcal{P}_{\text{iso}}$, elements in $\text{kern}(\tilde{P}) = \text{kern}(\mathcal{P}_{\text{iso}}) = \text{kern}(\mathcal{P})$ are exactly functions that have zero mean along each tomographic ray, whereas an element $f \in L^2(\Omega)$ is in $\text{range}(\tilde{P}^*) = \text{range}(\mathcal{P}^*)$ if and only if f is constant in Ω along the ray-direction, see §11.3.1. The gist of theorem 11.7 for parallel-beam geometries can thus be stated as follows: if \mathcal{R} uniformly penalizes oscillations along the rays, the increment computed in (11.2.6) is constant in ray-direction and thus $f_{\text{new}} - f_{\text{ref}} \in \overline{\text{range}(\mathcal{P}^*)}$. In cone-beam settings $\tilde{P} = \mathcal{D}_{\text{iso}}$, the characterization of $\text{kern}(\tilde{P})$ is identical as in the parallel-beam case. However, elements in $\text{range}(\tilde{P}^*) = \text{range}(\mathcal{D}^*)$ are only constant along the rays up to a scaling with the ray-density, see (11.3.7).

In a nutshell, the above indicates that *smoothing* penalties, that tend to damp out variations of $f_{\text{new}} - f_{\text{ref}}$ wherever possible, are promising candidates to satisfy assumption 11.6. Indeed, it will be verified for (weighted) L^2 -, quadratic gradient- and general L^q -penalties in the following sections. Before proceeding to the analysis of these settings, let us observe that convex combinations of admissible penalty functionals still satisfy assumption 11.6. This enables an application of theorem 11.7 to Kaczmarz-iterations with mixed penalties:

Lemma 11.8. *Let $\alpha_1, \alpha_2 \geq 0$ and let $\mathcal{R}_1, \mathcal{R}_2 : X \rightarrow \mathbb{R} \cup \{\infty\}$ be functionals satisfying (A) for the same $f_{\text{ref}} \in X$ and $\tilde{P} : X \rightarrow Y$. Then $\mathcal{R} := \alpha_1 \mathcal{R}_1 + \alpha_2 \mathcal{R}_2$ satisfies (A).*

Proof. This follows by summing scaled versions of (A) for \mathcal{R}_1 and \mathcal{R}_2 . □

11.4.2 (Weighted) L^2 -penalties

As a simple candidate for admissible penalty functionals \mathcal{R} within the framework of theorem 11.7, we consider quadratic penalties of the form

$$\mathcal{R}(f) := \|f - f_{\text{ref}}\|_X^2, \tag{11.4.1}$$

where $\|\cdot\|_X$ denotes the norm of the Hilbert space X . Owing to the geometric nature of the condition (A), it is straightforward to show that such a choice satisfies assumption 11.6. Note that the result is not limited to tomographic projectors $P \in \{\mathcal{P}, \mathcal{D}\}$, but holds within the general abstract setting of assumption 11.6:

Lemma 11.9 (Quadratic norm penalties). *Let X, Y be Hilbert spaces, $\tilde{P} : X \rightarrow Y$ linear-bounded with closed range and let $\mathcal{R}(f) := \|f - f_{\text{ref}}\|_X^2$ for $f_{\text{ref}} \in X$. Then assumption 11.6 is satisfied with strict inequality in (A) for all $f_0 \neq 0$ and it holds that*

$$\mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) = \langle p, \tilde{P}\tilde{P}^*(p) \rangle_Y \quad \text{for all } p \in Y. \quad (\langle \cdot, \cdot \rangle_Y: \text{ inner product in } Y) \quad (11.4.2)$$

Proof. In the considered setting, the condition (A) follows simply by the orthogonality $\text{kern}(\tilde{P}) \perp \text{range}(\tilde{P}^*)$: for all $p \in \text{range}(\tilde{P}^*)$, $f_0 \in \text{kern}(\tilde{P})$, it holds that

$$\begin{aligned} \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p) + f_0) &= \|\tilde{P}^*(p) + f_0\|_X^2 \stackrel{\tilde{P}^*(p) \perp f_0}{=} \|\tilde{P}^*(p)\|_X^2 + \|f_0\|_X^2 \\ &\geq \|\tilde{P}^*(p)\|_X^2 = \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) \\ &= \langle \tilde{P}^*(p), \tilde{P}^*(p) \rangle_X = \langle p, \tilde{P}\tilde{P}^*(p) \rangle_Y \end{aligned} \quad (11.4.3)$$

Moreover, since $\|f_0\|_X^2 > 0$ whenever $f_0 \neq 0$, strict inequality holds in this case. The equality in the third line simply follows by the defining property of the adjoint. \square

The abstract result from lemma 11.9 can be applied to establish GenSART-schemes for general Kaczmarz-iterations with L^2 -penalty:

Theorem 11.10 (Generalized SART with L^2 -penalty). *Let $P \in \{\mathcal{P}, \mathcal{D}\} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$, $f_{\text{ref}} \in L^2(\Omega)$, $\alpha > 0$ and let $\mathcal{S} : L^2(\mathbb{D}_P) \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Then the minimizers of*

$$f_{\text{new}} \in \underset{f \in L^2(\Omega)}{\text{argmin}} \mathcal{S}(P(f)) + \alpha \|f - f_{\text{ref}}\|_{L^2}^2 \quad (11.4.4)$$

are uniquely determined by the GenSART-scheme

$$p_{\text{ref}} = P(f_{\text{ref}}) \quad (11.4.5a)$$

$$\Delta p \in \underset{p \in L^2(\mathbb{D}_P)}{\text{argmin}} \mathcal{S}\left(p_{\text{ref}} + \tilde{u}_P^{1/2} \cdot p\right) + \alpha \|p\|_{L^2}^2 \quad (11.4.5b)$$

$$f_{\text{new}} = f_{\text{ref}} + P^*(\tilde{u}_P^{-1/2} \cdot \Delta p). \quad (11.4.5c)$$

The result (11.4.5) constitutes a generalization of the L^2 -SART-scheme derived in §11.3.2 for arbitrary data-fidelities \mathcal{S} . Indeed, the choice $\mathcal{S}(p) = \|p - g^{\text{obs}}\|_{L^2}^2$ in (11.4.5) reproduces the formula (11.3.12). Importantly, the optimization problem (11.4.5b) in projection-space no longer contains evaluations of P or P^* and is thus easy to solve given that \mathcal{S} is sufficiently simple. We omit the proof of theorem 11.10 as it is just a special case of the following result:

Theorem 11.11 (Generalized SART with L^2 -penalty and weighted projector). *Let $P \in \{\mathcal{P}, \mathcal{D}\} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$, $f_{\text{ref}} \in L^2(\Omega)$, $\alpha > 0$ and let $\mathcal{S} : L^2(\mathbb{D}_P) \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Moreover, let $\lambda : \Omega \rightarrow \mathbb{R}$ with $\lambda_{\min} \leq |\lambda(\mathbf{x})| \leq \lambda_{\max}$ for almost all $\mathbf{x} \in \Omega$ and some constants $0 < \lambda_{\min} \leq \lambda_{\max} < \infty$. Then the minimizers of*

$$f_{\text{new}} \in \underset{f \in L^2(\Omega)}{\text{argmin}} \mathcal{S}(P(\lambda \cdot f)) + \alpha \|f - f_{\text{ref}}\|_{L^2}^2. \quad (11.4.6)$$

are uniquely determined by the GenSART-scheme

$$p_{\text{ref}} = P(\lambda \cdot f_{\text{ref}}) \quad (11.4.7a)$$

$$\Delta p \in \underset{p \in L^2(\mathbb{D}_P)}{\text{argmin}} \mathcal{S}(p_{\text{ref}} + \lambda_P \cdot \tilde{u}_P^{-1/2} \cdot p) + \alpha \|\lambda_P^{1/2} \cdot p\|_{L^2}^2 \quad (11.4.7b)$$

$$f_{\text{new}} = f_{\text{ref}} + \lambda \cdot P^*(\tilde{u}_P^{-1/2} \cdot \Delta p) \quad (11.4.7c)$$

with $\lambda_P = P(w_P \cdot |\lambda|^2) / \tilde{u}_P$, w_P denoting the ray-density introduced in §11.3.1.

Proof. The bounds for λ ensure that $M_\lambda : L^2(\Omega) \rightarrow L^2(\Omega)$; $f \mapsto \lambda \cdot f$ is an isomorphism. We set $X := L^2(\Omega)$, $Y := L^2(\mathbb{D}_P)$ and $\tilde{P} := P_{\text{iso}} \circ M_\lambda : X \rightarrow Y$ with $P_{\text{iso}}(p) = \tilde{u}_P^{-1/2} \cdot P(p)$ as defined in (11.3.9). Then \tilde{P} has closed range by theorem 11.4 and M_λ being an isomorphism. By setting $\tilde{\mathcal{S}}(p) := \mathcal{S}(\tilde{u}_P^{1/2} \cdot p)$ and $\tilde{\mathcal{R}}(f) := \alpha \|f - f_{\text{ref}}\|_{L^2}^2$, (11.4.6) is cast to the form (11.3.17).

According to lemmas 11.8 and 11.9, assumption 11.6 is satisfied in the strict-inequality-version so that the GenSART-theorem 11.7 is applicable. Hence, the minimizers f_{new} of (11.4.6) can be computed via (11.3.18). Substituting the expressions for $\tilde{\mathcal{S}}$, \tilde{P} , $\tilde{\mathcal{R}}$ and exploiting that $\tilde{\mathcal{R}}(f_{\text{ref}} + \tilde{P}^*(p)) = \langle p, \tilde{P}\tilde{P}^*(p) \rangle_Y$ according to (11.4.2) yields

$$\tilde{p}_{\text{ref}} = \tilde{P}(f_{\text{ref}}) = \tilde{u}_P^{-1/2} \cdot \tilde{P}(\lambda \cdot f_{\text{ref}}) \quad (11.4.8a)$$

$$\Delta p \in \underset{p \in L^2(\mathbb{D}_P)}{\text{argmin}} \mathcal{S}(\tilde{u}_P^{1/2} \cdot \tilde{p}_{\text{ref}} + \tilde{u}_P^{1/2} \cdot \tilde{P}\tilde{P}^*(p)) + \alpha \langle p, \tilde{P}\tilde{P}^*(p) \rangle_{L^2} \quad (11.4.8b)$$

$$f_{\text{new}} = f_{\text{ref}} + \lambda \cdot \tilde{P}^*(\tilde{u}_P^{-1/2} \cdot p). \quad (11.4.8c)$$

From this expression, the formula (11.4.7) is obtained by setting $p_{\text{ref}} := \tilde{u}_P^{1/2} \cdot \tilde{p}_{\text{ref}}$ and using that, according to (11.3.5) and (11.3.7),

$$\begin{aligned} \tilde{P}\tilde{P}^*(p) &= \tilde{u}_P^{-1/2} \cdot P(|\lambda|^2 \cdot P^*(\tilde{u}_P^{-1/2} \cdot p)) = \tilde{u}_P^{-1/2} \cdot P((w_P \cdot |\lambda|^2) \cdot P^B(\tilde{u}_P^{-1/2} \cdot p)) \\ &= \tilde{u}_P^{-1} \cdot P(w_P \cdot |\lambda|^2) \cdot p = \lambda_P \cdot p \quad \text{for all } p \in L^2(\mathbb{D}_P). \quad \square \end{aligned}$$

Note that, once again, the optimization problem (11.4.7b) does not involve P or P^* except in the form of the precomputable functions λ_P . Theorem 11.10 is obtained from theorem 11.11 by choosing $\lambda := 1$. Moreover, the theorem enables GenSART for *weighted* L^2 -penalties:

Corollary 11.12 (Generalized SART with weighted L^2 -penalties). *Let $P \in \{\mathcal{P}, \mathcal{D}\} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$, $f_{\text{ref}} \in L^2(\Omega)$, $\alpha > 0$ and let $\mathcal{S} : L^2(\mathbb{D}_P) \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Moreover, let $w : \Omega \rightarrow \mathbb{R}_{>0}$ with $w_{\min} \leq w(\mathbf{x}) \leq w_{\max}$ for almost all $\mathbf{x} \in \Omega$ and some constants $0 < w_{\min} \leq w_{\max} < \infty$. Then the minimizers of*

$$f_{\text{new}} \in \underset{f \in L^2(\Omega)}{\text{argmin}} \mathcal{S}(P(f)) + \alpha \|w^{-1/2} \cdot (f - f_{\text{ref}})\|_{L^2}^2. \quad (11.4.9)$$

are uniquely determined by the GenSART-scheme

$$p_{\text{ref}} = P(f_{\text{ref}}) \quad (11.4.10a)$$

$$\Delta p \in \underset{p \in L^2(\mathbb{D}_P)}{\text{argmin}} \mathcal{S}(p_{\text{ref}} + v_P \cdot \tilde{u}_P^{1/2} \cdot p) + \alpha \|v_P^{1/2} \cdot p\|_{L^2}^2 \quad (11.4.10b)$$

$$f_{\text{new}} = f_{\text{ref}} + w \cdot P^*(\tilde{u}_P^{-1/2} \cdot \Delta p) \quad (11.4.10c)$$

where $v_P = P(w \cdot w_P) / \tilde{u}_P$, w_P denoting the ray-density introduced in §11.3.1.

Proof. f_{new} is a minimizer of (11.4.9) if and only if $\tilde{f}_{\text{new}} := w^{-1/2} \cdot f_{\text{new}}$ solves the optimization problem (11.4.6) for $\lambda := w^{1/2}$ and f_{ref} replaced by $\tilde{f}_{\text{ref}} := w^{-1/2} \cdot f_{\text{ref}}$. The resulting GenSART-scheme (11.4.7) for \tilde{f}_{new} yields (11.4.10). \square

The more general, weighted GenSART-schemes from theorem 11.11 and corollary 11.12 enable interesting applications, as will be seen in §11.5.2.2. Moreover, note that the assumption that λ or w are bounded from below can be dropped at the cost of a more technical proof. Indeed, the formulas (11.4.7), (11.4.10) still make sense if λ or w vanishes in parts of Ω .

11.4.3 Gradient-penalties

In order to enforce a certain smoothness of the reconstructed object, variational methods often use penalties that involve derivatives. Similarly, we seek to extend the above results to Kaczmarz-iterations (11.2.6) with quadratic gradient penalties:

$$\mathcal{R} : L^2(\Omega) \rightarrow \mathbb{R} \cup \{\infty\}; \quad \mathcal{R}(f) := \|\nabla(f - f_{\text{ref}})\|_{L^2}^2 = \int_{\Omega} |\nabla(f - f_{\text{ref}})|^2 \, d\mathbf{x} \quad (11.4.11)$$

with the convention that $\|\nabla h\| = \infty$ for any $h \in L^2(U) \setminus W^{1,2}(U)$, where $W^{1,2}(U)$ is the usual Sobolev-space $W^{1,2}(U) := \{f \in L^2(U) : f \text{ weakly differentiable in } U \text{ with } |\nabla f| \in L^2(\Omega)\}$ on an open domain $U \subset \mathbb{R}^m$, with norm $\|f\|_{W^{1,2}}^2 := \|f\|_{L^2}^2 + \|\nabla f\|_{L^2}^2$. Moreover, we recall that gradients of functions $f \in L^2(\Omega)$, $p \in L^2(\mathbb{D}_P)$ are understood to be computed only *within* $\Omega \subset \mathbb{R}^3$ and $\mathbb{D}_P \subset \mathbb{D}$, i.e. in particular *not* across boundaries (convention 11.2).

By §11.4.1, one might hope that the gradient penalty in (11.4.11) satisfies assumption 11.6 owing to its smoothing effect. Indeed, this turns out to be *almost* true, up to some complications arising from thickness-variations of the domain Ω . This is explained in the following.

For simplicity, we study the parallel-beam case, $\tilde{P} = \mathcal{P}_{\text{iso}}$, and continuously differentiable functions $f_0 \in \mathcal{C}^1(\Omega) \cap \ker(\tilde{P})$ and $p \in L^2(\mathbb{D}_P)$ s.t. $\tilde{P}^*(p) \in \mathcal{C}^1(\Omega)$. Then we have that

$$\begin{aligned} \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p) + f_0) &= \int_{\Omega} |\nabla(\mathcal{P}_{\text{iso}}^*(p) + f_0)|^2 \, d\mathbf{x} \\ &= \|\nabla \mathcal{P}_{\text{iso}}^*(p)\|_{L^2}^2 + \|\nabla f_0\|_{L^2}^2 + 2 \int_{\Omega} \nabla \mathcal{P}_{\text{iso}}^*(p) \cdot \nabla f_0 \, d\mathbf{x}. \end{aligned} \quad (11.4.12)$$

By (11.4.12), the inequality (A) holds true if and only if the mixed integral on the right-hand side vanishes for all admissible p and f_0 .

Now let us assume that Ω is convex with \mathcal{C}^1 -boundary. Then we may write it in the form $\Omega = \{(\mathbf{x}_{\perp}, z) \in \mathbb{R}^3 : a(\mathbf{x}_{\perp}) < z < b(\mathbf{x}_{\perp})\}$ for continuous functions $a, b : \mathbb{R}^2 \rightarrow \mathbb{R}$, $a \leq b$, that give the entrance- and exit points of the tomographic rays into Ω . Hence, it holds that

$$\begin{aligned} \int_{\Omega} \nabla \mathcal{P}_{\text{iso}}^*(p) \cdot \nabla f_0 \, d\mathbf{x} &= \int_{\mathbb{R}^2} \left(\int_{a(\mathbf{x}_{\perp})}^{b(\mathbf{x}_{\perp})} \nabla \mathcal{P}_{\text{iso}}^*(p)(\mathbf{x}_{\perp}, z) \cdot \nabla f_0(\mathbf{x}_{\perp}, z) \, dz \right) d\mathbf{x}_{\perp} \\ &= \int_{\mathbb{R}^2} \nabla(p/u_{\mathcal{P}}^{1/2})(\mathbf{x}_{\perp}) \cdot \left(\int_{a(\mathbf{x}_{\perp})}^{b(\mathbf{x}_{\perp})} \nabla_{\perp} f_0(\mathbf{x}_{\perp}, z) \, dz \right) d\mathbf{x}_{\perp}. \end{aligned} \quad (11.4.13)$$

In the second step, we exploited that $\mathcal{P}_{\text{iso}}^*(p)(\mathbf{x}_{\perp}, z) = \mathcal{P}^{\text{B}}(p/u_{\mathcal{P}}^{1/2})(\mathbf{x}_{\perp}, z) = p(\mathbf{x}_{\perp})$ is constant in z so that the z -component of the gradient $\nabla = (\nabla_{\perp}, \partial_z)$ vanishes and $\nabla \mathcal{P}_{\text{iso}}^*(p)(\mathbf{x}_{\perp}, z) =$

$\nabla(p/u_{\mathcal{P}}^{1/2})(\mathbf{x}_{\perp})$ can be pulled out of the inner integral. Since a and b are continuously differentiable within the open set $U := \{\mathbf{x}_{\perp} \in \mathbb{R}^2 : a(\mathbf{x}_{\perp}) < b(\mathbf{x}_{\perp})\}$ by the assumptions on Ω , we can apply Leibniz' rule to the inner integrals in (11.4.13):

$$\int_{a(\mathbf{x}_{\perp})}^{b(\mathbf{x}_{\perp})} \nabla_{\perp} f_0(\mathbf{x}_{\perp}, z) dz = \nabla_{\perp} \int_{a(\mathbf{x}_{\perp})}^{b(\mathbf{x}_{\perp})} f_0(\mathbf{x}_{\perp}, z) dz - (f_0(\mathbf{x}_{\perp}, b(\mathbf{x}_{\perp}))\nabla b(\mathbf{x}_{\perp}) - f_0(\mathbf{x}_{\perp}, a(\mathbf{x}_{\perp}))\nabla a(\mathbf{x}_{\perp})) \quad \text{for all } \mathbf{x}_{\perp} \in U. \quad (11.4.14)$$

The first integral on the right-hand side of (11.4.14) can be identified as $\nabla \mathcal{P}(f_0)(\mathbf{x}_{\perp})$ and vanishes since $f_0 \in \text{kern}(\mathcal{P}_{\text{iso}}) = \text{kern}(\mathcal{P})$. The second term, arising from variations of the domain-boundary-functions a and b , does *not* vanish in general. Yet, it vanishes if *either*

- a, b are constant within U . This is only the case for domains of the form $\Omega = \Omega_{\perp} \times \Omega_{\parallel}$ for some $\Omega_{\perp} \subset \mathbb{R}^2$ and $\Omega_{\parallel} \subset \mathbb{R}$, such as cuboids aligned with the ray-direction.
- f_0 vanishes on the boundary of Ω .

(11.4.14), (11.4.13) and (11.4.12) indicate that either of the above restrictions may ensure that the inequality (A) is satisfied. We focus on the second one in order to retain geometrical flexibility of the object-domain Ω . In order to verify assumption 11.6 in a Sobolev-space setting, we consider the subspace of functions with trace zero, $W_0^{1,2}(\Omega) := \overline{\mathcal{C}_c^{\infty}(\Omega)} \subset W^{1,2}(\Omega)$, defined as the closure of the smooth and compactly supported functions $\mathcal{C}_c^{\infty}(\Omega)$ in the $W^{1,2}$ -topology. Firstly, we need to analyze the mutual behavior of the projectors and the gradient operator, as established by the following lemma, which is proven in appendix 11.B:

Lemma 11.13. *Let $P \in \{\mathcal{P}, \mathcal{D}\}$, $f \in W_0^{1,2}(\Omega)$ and $p \in L^2(\mathbb{D}_P)$. Denote by $\nabla_{\mathbb{D}}$ the gradient on the detection domain $\mathbb{D} \in \{\mathbb{R}^2, \mathbb{S}^2\}$ and by ∇_P the component of gradient in \mathbb{R}^3 perpendicular to the local ray-direction of the projector P . Then it holds that*

$$P(w_P^{-1/2} \cdot \nabla_P f) = \nabla_{\mathbb{D}} P(f) \quad (11.4.15)$$

$$\|\tilde{u}_P^{-1/2} \cdot \nabla_{\mathbb{D}} P(f)\|_{L^2} \leq \|w_P^{-1/2} \cdot \nabla_P f\|_{L^2}, \quad (11.4.16)$$

where w_P is the ray-density defined in §11.3.1. Moreover, $P^{\text{B}}(p) \in W^{1,2}(\Omega)$ holds true if and only if $u_P^{1/2} \cdot \nabla_{\mathbb{D}} p \in L^2(\mathbb{D}_P)$ and in this case

$$\|\nabla P^{\text{B}}(p)\|_{L^2} = \|u_P^{1/2} \cdot \nabla_{\mathbb{D}} p\|_{L^2} \quad (11.4.17)$$

$$\langle \nabla P^{\text{B}}(p), \nabla f \rangle_{L^2} = \langle u_P^{1/2} \cdot \nabla_{\mathbb{D}} p, u_P^{-1/2} \cdot \nabla_{\mathbb{D}} P(f) \rangle_{L^2}. \quad (11.4.18)$$

It should be emphasized that the occurrence of both \tilde{u}_P and u_P in lemma 11.13 is *not* a typo, but is indeed necessary to correctly cover the cone-beam case $P = \mathcal{D}$. In a parallel-beam setting $P = \mathcal{P}$, on the other hand, one has $\tilde{u}_P = u_P$ and $w_P = 1$, so that the expressions can be simplified. Lemma 11.13 permits to prove the admissibility of gradient-penalties, yet in a slightly restricted setting due to the ‘‘geometrical complications’’ explained above:

Lemma 11.14 (Admissibility of L^2 -gradient-penalties). *Let $P \in \{\mathcal{P}, \mathcal{D}\}$, $f_{\text{ref}} \in L^2(\Omega)$ and let \mathcal{R} be defined by (11.4.11). Let $X := (L^2(\Omega), \langle \cdot, \cdot \rangle_P)$ be equipped with the inner product $\langle f_1, f_2 \rangle_P := \langle w_P \cdot f_1, f_2 \rangle_{L^2}$. and define $\tilde{P} : X \rightarrow L^2(\mathbb{D}_P)$; $f \mapsto u_P^{-1/2} \cdot P(f)$. Then assumption 11.6 is satisfied if we restrict to elements $f_0 \in \ker(\tilde{P}) \cap W_0^{1,2}(\Omega)$ in (A) and strict inequality holds for all $f_0 \neq 0$. Moreover,*

$$\mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) = \|u_P^{1/2} \cdot \nabla_{\mathbb{D}}(u_P^{-1/2} \cdot p)\|_{L^2}^2 \quad (11.4.19)$$

for all $p \in L^2(\mathbb{D}_P)$ where possibly both sides may attain the value ∞ .

Proof. By definition, it holds that $\tilde{P} = M \circ P_{\text{iso}} \circ \iota_X$, where P_{iso} denotes the weighted projector from theorem 11.4, $M : L^2(\mathbb{D}_P) \rightarrow L^2(\mathbb{D}_P)$; $p \mapsto (\tilde{u}_P/u_P)^{1/2} \cdot p$ and $\iota_X : X \hookrightarrow L^2(\Omega)$; $f \mapsto f$ is the canonical embedding. The norms of X and $L^2(\Omega)$ are equivalent since $c \leq w_P \leq C$ is bounded from below and above by constants $0 < c \leq C$. The bounds also imply that $c^{1/2} \leq (\tilde{u}_P/u_P)^{1/2} \leq C^{1/2}$ by (11.3.3). Hence, ι_X and M are isomorphisms so that $\tilde{P} = M \circ P_{\text{iso}} \circ \iota_X$ has closed range by theorem 11.4. Moreover, it holds that $\iota_X^*(f) = w_P^{-1} \cdot f$ for all $f \in L^2(\Omega)$. Hence, we have for all $p \in L^2(\mathbb{D}_P)$

$$\tilde{P}^*(p) = \iota_X^*(P^*(u_P^{-1/2} \cdot p)) = w_P^{-1} \cdot P^*(u_P^{-1/2} \cdot p) \stackrel{(11.3.7)}{=} P^{\text{B}}(u_P^{-1/2} \cdot p). \quad (11.4.20)$$

Now let $p \in L^2(\mathbb{D}_P)$ and $f_0 \in \ker(\tilde{P}) \cap W_0^{1,2}(\Omega)$ be arbitrary. If $\tilde{P}^*(p) \notin W^{1,2}(\Omega)$, then also $\tilde{P}^*(p) + f_0 \notin W^{1,2}(\Omega)$ and so $\mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p) + f_0) = \infty$ by (11.4.11), in which case (A) is trivially satisfied. Hence, we assume that $\tilde{P}^*(p) \in W^{1,2}(\Omega)$ in the following. Then it holds that

$$\begin{aligned} \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p) + f_0) &= \|\nabla(\tilde{P}^*(p) + f_0)\|_{L^2}^2 \\ &= \|\nabla \tilde{P}^*(p)\|_{L^2}^2 + 2\langle \nabla \tilde{P}^*(p), \nabla f_0 \rangle_{L^2} + \|\nabla f_0\|_{L^2}^2 \\ &\geq \|\nabla \tilde{P}^*(p)\|_{L^2}^2 + 2\langle \nabla \tilde{P}^*(p), \nabla f_0 \rangle_{L^2} \\ &= \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) + 2\langle \nabla \tilde{P}^*(p), \nabla f_0 \rangle_{L^2}. \end{aligned} \quad (11.4.21)$$

We need to show that the mixed term on the right-hand side of (11.4.21) vanishes. However, this is a simple consequence of lemma 11.13: since $\tilde{P}^*(p) \in W^{1,2}(\Omega)$ and

$$\tilde{P}^*(p) = \iota_X^* \circ P_{\text{iso}}^* \circ M^*(p) = w_P^{-1} \cdot w_P \cdot P^{\text{B}}(\tilde{u}_P^{-1/2} \cdot (\tilde{u}_P/u_P)^{1/2} \cdot p) = P^{\text{B}}(u_P^{-1/2} \cdot p), \quad (11.4.22)$$

it follows that $u_P^{1/2} \cdot \nabla_{\mathbb{D}}(u_P^{-1/2} \cdot p) \in L^2(\mathbb{D}_P)$ and we obtain by application of (11.4.18)

$$\begin{aligned} \langle \nabla \tilde{P}^*(p), \nabla f_0 \rangle_{L^2} &= \langle \nabla P^{\text{B}}(u_P^{-1/2} \cdot p), \nabla f_0 \rangle_{L^2} \\ &= \langle u_P^{1/2} \cdot \nabla_{\mathbb{D}}(u_P^{-1/2} \cdot p), u_P^{-1/2} \cdot \nabla_{\mathbb{D}} P(f_0) \rangle_{L^2} \stackrel{f_0 \in \ker(\tilde{P}) = \ker(P)}{=} 0. \end{aligned} \quad (11.4.23)$$

Inserting this result into (11.4.21) shows that the inequality (A) is satisfied for all $f_0 \in \ker(\tilde{P}) \cap W_0^{1,2}(\Omega)$. Moreover, if $f_0 \neq 0$, then f_0 is necessarily *non-constrant* in Ω so that $\|\nabla f_0\|_{L^2} > 0$. By the second line in (11.4.21), this implies that strict inequality holds in (A) if $f_0 \neq 0$. Finally, combining (11.4.20) and (11.4.17) yields

$$\mathcal{R}(f_{\text{ref}} + P^*(p)) = \|\nabla(P^{\text{B}}(u_P^{-1/2} \cdot p))\|_{L^2}^2 = \|u_P^{1/2} \cdot \nabla_{\mathbb{D}}(u_P^{-1/2} \cdot p)\|_{L^2}^2, \quad (11.4.24)$$

which, by lemma 11.13, remains valid if the expressions attain the value ∞ . \square

Lemma 11.14 enables GenSART-schemes for Kaczmarz-iterations of the form (11.3.17) with $\mathcal{R}(f) = \|\nabla(f - f_{\text{ref}})\|_{L^2}^2$ – if the optimization is restricted to a slightly smaller choice set:

$$f - f_{\text{ref}} \in W_P^{1,2}(\Omega) := (\overline{\text{range}(P^B)} \cap W^{1,2}(\Omega)) \oplus (\text{kern}(P) \cap W_0^{1,2}(\Omega)) \subsetneq W^{1,2}(\Omega) \quad (11.4.25)$$

By lemmas 11.8 and 11.9, this remains true for more general Sobolev- $W^{1,2}$ -like penalties,

$$\mathcal{R}(f) := \alpha \|f - f_{\text{ref}}\|_{W_{\gamma,P}^{1,2}}^2 \quad \text{with} \quad \|h\|_{W_{\gamma,P}^{1,2}}^2 := (1 - \gamma) \|w_P^{1/2} \cdot h\|_{L^2}^2 + \gamma \|\nabla h\|_{L^2}^2 \quad (11.4.26)$$

for some $0 \leq \gamma \leq 1$. Note that $h \mapsto \|w_P^{1/2} \cdot h\|_{L^2}$ is the Hilbert-space-norm of X in lemma 11.14, which is identical to the L^2 -norm in the parallel-beam case $P = \mathcal{P}$ but *not* in a cone-beam setting $P = \mathcal{D}$, since $w_{\mathcal{D}} = \mathbf{1}_\Omega$ but $w_{\mathcal{D}} \neq \mathbf{1}_\Omega$. The general result reads as follows:

Theorem 11.15 (Generalized SART with $W^{1,2}$ -penalties). *Let $P \in \{\mathcal{P}, \mathcal{D}\} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$, $f_{\text{ref}} \in L^2(\Omega)$, $\alpha > 0$ and let $\mathcal{S} : L^2(\mathbb{D}_P) \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Then the minimizers of*

$$f_{\text{new}} \in \underset{f \in f_{\text{ref}} + W_P^{1,2}(\Omega)}{\text{argmin}} \mathcal{S}(P(f)) + \alpha \|f - f_{\text{ref}}\|_{W_{\gamma,P}^{1,2}}^2. \quad (11.4.27)$$

are uniquely determined by the GenSART-scheme

$$p_{\text{ref}} = P(f_{\text{ref}}) \quad (11.4.28a)$$

$$\Delta p \in \underset{p \in L^2(\mathbb{D}_P)}{\text{argmin}} \mathcal{S}(p_{\text{ref}} + u_P^{1/2} \cdot p) + \alpha(1 - \gamma) \|p\|_{L^2}^2 + \alpha\gamma \|u_P^{1/2} \cdot \nabla_{\mathbb{D}}(u_P^{-1/2} \cdot p)\|_{L^2}^2 \quad (11.4.28b)$$

$$f_{\text{new}} = f_{\text{ref}} + P^B(u_P^{-1/2} \cdot \Delta p). \quad (11.4.28c)$$

Proof. Let $X := (L^2(\Omega), \langle \cdot, \cdot \rangle_P)$ and $P : X \rightarrow L^2(\mathbb{D}_P)$ be defined as in lemma 11.14 and $\tilde{\mathcal{S}}(p) := \mathcal{S}(u_P^{1/2} \cdot p)$. Then the optimization problem (11.4.27) can be written in the form

$$f_{\text{new}} \in \underset{f \in X}{\text{argmin}} \tilde{\mathcal{S}}(P(f)) + \alpha(1 - \gamma)\mathcal{R}_1(f) + \alpha\gamma\mathcal{R}_2(f). \quad (11.4.29)$$

where the functionals $\mathcal{R}_1, \mathcal{R}_2 : X \rightarrow \mathbb{R} \cup \{\infty\}$ are given by

$$\mathcal{R}_1(f) := \|f - f_{\text{ref}}\|_X^2, \quad \mathcal{R}_2(f) := \|\nabla(f - f_{\text{ref}})\|_{L^2}^2 + \chi_{\tilde{X}}(f - f_{\text{ref}}). \quad (11.4.30)$$

$\chi_{\tilde{X}} : X \rightarrow \mathbb{R} \cup \{\infty\}$ is the indicator functional of $\tilde{X} := W_P^{1,2}(\Omega)$, defined by $\chi(h) = 0$ if $h \in \tilde{X}$ and $\chi(h) = \infty$ otherwise. As $\text{range}(P^B) = \text{range}(\tilde{P}^*)$ and $\text{kern}(P) = \text{kern}(\tilde{P})$, it holds that $\tilde{X} = (\text{range}(\tilde{P}) \cap W^{1,2}(\Omega)) \oplus (\text{kern}(\tilde{P}) \cap W_0^{1,2}(\Omega))$. By lemmas 11.9 and 11.14, assumption 11.6 is thus satisfied for $\tilde{P} : X \rightarrow L^2(\mathbb{D}_P)$ and $\mathcal{R} = \mathcal{R}_j$ for $j = 1, 2$. According lemma 11.8, the same is true for $\mathcal{R} := \alpha(1 - \gamma)\mathcal{R}_1 + \alpha\gamma\mathcal{R}_2$. Hence, the GenSART-theorem 11.7 is applicable to (11.4.29) so that a minimizers f_{new} can be found via the scheme (11.3.18):

$$\tilde{p}_{\text{ref}} = \tilde{P}(f_{\text{ref}}) = u_P^{-1/2} \cdot P(f_{\text{ref}}) \quad (11.4.31a)$$

$$\Delta p \in \underset{p \in L^2(\mathbb{D}_P)}{\text{argmin}} \tilde{\mathcal{S}}(\tilde{p}_{\text{ref}} + \tilde{P}\tilde{P}^*(p)) + \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) \quad (11.4.31b)$$

$$f_{\text{new}} = f_{\text{ref}} + \tilde{P}^*(\Delta p) \quad (11.4.31c)$$

By lemmas 11.9 and 11.14, the penalty term can be rewritten to

$$\begin{aligned} \mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) &= \alpha(1 - \gamma)\mathcal{R}_1(f_{\text{ref}} + \tilde{P}^*(p)) + \alpha\gamma\mathcal{R}_2(f_{\text{ref}} + \tilde{P}^*(p)) \\ &= \alpha(1 - \gamma)\langle p, \tilde{P}\tilde{P}^*(p) \rangle_{L^2} + \alpha\gamma\|u_P^{1/2} \cdot \nabla_{\mathbb{D}}(u_P^{-1/2} \cdot p)\|_{L^2}^2. \end{aligned} \quad (11.4.32)$$

Moreover, as seen in the proof of lemma 11.14, it holds that $\tilde{P}^*(p) = P^{\text{B}}(u_P^{-1/2} \cdot p)$ and thus $\tilde{P}\tilde{P}^*(p) = u_P^{-1/2} \cdot P P^{\text{B}}(u_P^{-1/2} \cdot p) = p$ for all $p \in L^2(\mathbb{D}_P)$ by (11.3.5). Substituting these expressions into (11.4.31) along with $p_{\text{ref}} = u_P^{1/2} \cdot \tilde{p}_{\text{ref}}$ and $\mathcal{S}(p) = \tilde{\mathcal{S}}(u_P^{-1/2} \cdot p)$ yields (11.4.28). \square

To conclude this section, we make a few remarks on peculiarities of theorem 11.15:

- The derived SART formula (11.4.28) only involves the *unweighted* unit-projection u_P and back-projector P^{B} and *not* the L^2 -adjoint $P^* : p \mapsto w_P \cdot P^{\text{B}}(p)$. Accordingly, the ray-density-weighting of the back-projection in the cone-beam case is omitted. This is quite intuitive since back-projecting uniformly along the rays results in smaller values of the gradient-penalty functional (11.4.11) and thus a non-weighted back-projection can be regarded as the natural one in the considered setting.
- In (11.4.27), the increment $\Delta f := f_{\text{new}} - f_{\text{ref}}$ is not optimized over the whole feasible set $W^{1,2}(\Omega)$, but only within the closed subspace $W_P^{1,2}(\Omega) \subsetneq W^{1,2}(\Omega)$. This may seem like a fundamental flaw of the result. Notably, however, $W_P^{1,2}(\Omega)$ is *much larger* than the mere space of back-projections $\overline{\text{range}(P^{\text{B}})} \cap W^{1,2}(\Omega) \subsetneq W_P^{1,2}(\Omega)$, over which the scheme (11.4.28) trivially computes the optimal increment $\Delta f = P^{\text{B}}(u_P^{-1/2} \cdot \Delta p)$. In this sense, theorem 11.15 still provides a non-trivial simplification of the optimization problem in (11.4.27). Indeed, $W_P^{1,2}(\Omega)$ is arguably *almost* as large as $W^{1,2}(\Omega)$: $\overline{\text{range}(P^{\text{B}})} \cap W^{1,2}(\Omega)$ contains all functions that are constant along the rays and $\text{kern}(P) \cap W_0^{1,2}(\Omega)$ all those, which have zero mean along these and vanish on the boundary of Ω . Accordingly, the “missing subspace”, i.e. the orthogonal complement of $W_P^{1,2}(\Omega)$ within $W^{1,2}(\Omega)$, must be composed solely of functions in $\text{kern}(P)$ that are *linear* along all ray-segments. In practice, the GenSART-iterates defined by (11.4.28) are thus expected to provide *almost* the optimum of the objective in (11.4.27) over all $f \in W^{1,2}(\Omega)$ (or $f \in L^2(\Omega)$).

11.4.4 L^q -penalties

The aim of this section is to demonstrate that assumption 11.6 does not restrict the choice of penalty functionals \mathcal{R} to quadratic ones. We consider L^q -penalties

$$\mathcal{R}(f) := \|f - f_{\text{ref}}\|_{L^q}^q \quad \text{with} \quad \|h\|_{L^q}^q := \int_{\mathbb{R}^m} |h(\mathbf{x})|^q d\mathbf{x} \in \mathbb{R} \cup \{\infty\}, \quad 1 \leq q < \infty \quad (11.4.33)$$

and define $L^q(\Omega) := \{f : \|f\|_{L^q} < \infty\}$ as usual. We prove the admissibility of such penalties only for a parallel-beam setting $P = \mathcal{P}$. An extension to the cone-beam case may be possible.

Lemma 11.16 (Admissibility of L^q -penalties). *Let $\tilde{P} = \mathcal{P}_{\text{iso}} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$ and let $\mathcal{R} : L^2(\Omega) \rightarrow \mathbb{R} \cup \{\infty\}$ be defined by (11.4.33). Then assumption 11.6 is satisfied and*

$$\mathcal{R}(f_{\text{ref}} + \tilde{P}^*(p)) = \begin{cases} \|u_{\mathcal{P}}^{1/q-1/2} \cdot p\|_{L^q}^q & \text{if } \tilde{P}^*(p) \in L^q(\Omega) \\ \infty & \text{else} \end{cases} \quad \text{for all } p \in L^2(\mathbb{D}_P). \quad (11.4.34)$$

The proof of lemma 11.16 is given in appendix 11.C. For completeness, we state the GenSART-scheme that is obtained by applying theorem 11.7 to the setting in lemma 11.16:

Theorem 11.17 (Generalized SART with L^q -penalties). *Let $f_{\text{ref}} \in L^2(\Omega)$, $\alpha > 0$ and let $\mathcal{S} : L^2(\mathbb{D}_P) \rightarrow \mathbb{R} \cup \{\infty\}$ be any functional. Assume that there exists a minimizer*

$$f_{\text{new}} \in \operatorname{argmin}_{f \in L^2(\Omega)} \mathcal{S}(\mathcal{P}(f)) + \alpha \|f - f_{\text{ref}}\|_{L^q}^q. \quad (11.4.35)$$

Then any \tilde{f}_{new} determined by the GenSART-scheme

$$p_{\text{ref}} = \mathcal{P}(f_{\text{ref}}) \quad (11.4.36a)$$

$$\Delta p \in \operatorname{argmin}_{p \in L^2(\mathbb{D}_P)} \mathcal{S}(p_{\text{ref}} + u_{\mathcal{P}}^{1/2} \cdot p) + \alpha \|u_{\mathcal{P}}^{1/q-1/2} \cdot p\|_{L^q}^q \quad (11.4.36b)$$

$$\tilde{f}_{\text{new}} = f_{\text{ref}} + \mathcal{P}^*(u_{\mathcal{P}}^{-1/2} \cdot \Delta p) \quad (11.4.36c)$$

also minimizes (11.4.35). If $q > 1$, then any minimizer f_{new} of (11.4.35) is of the form (11.4.36).

11.5 Applications

In the preceding sections, it has been analyzed in which abstract situations Kaczmarz-iterations of the form (11.2.6) can be computed via a generalized SART-scheme. In the following, the principal theory is applied to design tailored methods for various settings of tomographic imaging. Specifically, the aim is to exploit the extraordinary freedom that the GenSART-theorem 11.7 offers in choosing the data-fidelity functionals \mathcal{S}_k and image-formation operators F_j . Despite differences between the specific applications, it should be emphasized that all of the proposed methods are applicable for both parallel- and cone-beam acquisition-geometries, without any requirements on the incident directions or source positions.

11.5.1 Noise-model-adapted GenSART

As outlined in §11.2.3, variational- and Kaczmarz-type reconstruction methods may account for the expected statistics of the data errors ϵ in Inverse Problem 11.1 by suitably choosing the data-fidelity functionals \mathcal{S}_k in (11.2.6). We illustrate this for Kaczmarz-iterations with a simple L^2 -penalty and fixed $\mathcal{S}_k = \mathcal{S}$:

$$f_{k+1} \in \operatorname{argmin}_{f \in L^2(\Omega)} \mathcal{S}(g_{j_k}^{\text{obs}}; F_{j_k}(P_{j_k}(f))) + \alpha \|f - f_k\|_{L^2}^2 \quad (11.5.1)$$

By theorem 11.10, the minimizer can be computed via the SART-like scheme

$$p_k = P_{j_k}(f_k) \quad (11.5.2a)$$

$$\Delta p_k \in \underset{p \in L^2(\mathbb{D}_{j_k})}{\operatorname{argmin}} \mathcal{S} \left(g_{j_k}^{\text{obs}}; F_{j_k} \left(p_k + \tilde{u}_{j_k}^{1/2} \cdot p \right) \right) + \alpha \|p\|_{L^2}^2 \quad (11.5.2b)$$

$$f_{k+1} = f_k + P_{j_k}^* \left(\tilde{u}_{j_k}^{-1/2} \cdot \Delta p_k \right) \quad (11.5.2c)$$

where \mathbb{D}_j and \tilde{u}_j denote the projection-domain and weighted unit-projection of the projector P_j , respectively, see §11.3.1. According to Bayesian theory, the Kaczmarz-iterations in (11.5.1) can be tailored for a specific (probabilistic) model of the data errors ϵ by choosing \mathcal{S} as the negative log-likelihood of the fitted data $g_{j_k} = F_{j_k}(P_{j_k}(f))$ given the observations $g_{j_k}^{\text{obs}}$:

$$\mathcal{S}(g_j^{\text{obs}}; g_j) := -\ln \mathbb{P}(g_j^{\text{obs}} | g_j), \quad (11.5.3)$$

where $\mathbb{P}(g_j^{\text{obs}} | g_j)$ denotes the probability of measuring g_j^{obs} given that the true data is g_j . By inserting (11.5.3) into (11.5.2), generic noise-model-adapted GenSART-schemes are obtained.

Efficient closed-form optimization in projection-space: For general noise-models and image-formation operators F_j , the optimization problem in (11.5.2b) could still be hard to solve, in spite of being cast to the low-dimensional projection-space via the GenSART-approach. In the following, we therefore outline practically relevant settings where the optimization-step in the GenSART-scheme (11.5.2) may be performed at negligible computational costs.

Often, the data-errors ϵ at different spatial positions can be assumed to be stochastically independent. If the F_j are pointwise operators, i.e. “ $F_{j_k}(p)(x) = F_{j_k}(p(x))$ ” for all p and x (in particular if $F_j = \text{id}$), the data-fidelity obtained via (11.5.3) is then of *integral-form*:

$$\mathcal{S}(g_j^{\text{obs}}; F_j(p)) = \int_{\mathbb{D}_j} s_j(x, p(x)) \, dx + c \quad \text{for some} \quad s_j : \mathbb{D}_j \times \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\} \quad (11.5.4)$$

with some additive constant $c \in \mathbb{R}$ that does not affect the minimizer in (11.5.1). By substituting (11.5.4) into the the objective-functional to be minimized in (11.5.2), we obtain

$$\begin{aligned} & \mathcal{S} \left(g_{j_k}^{\text{obs}}; F_{j_k} \left(p_k + \tilde{u}_{j_k}^{1/2} \cdot p \right) \right) + \alpha \|p\|_{L^2}^2 \\ &= \int_{\mathbb{D}_{j_k}} \left(s_{j_k}(x, p_k(x) + \tilde{u}_{j_k}(x)^{1/2} p(x)) + \alpha |p(x)|^2 \right) \, dx \end{aligned} \quad (11.5.5)$$

The integrand in (11.5.5) depends only on point-evaluations of p , i.e. only on the local function value. As a consequence, the optimization in (11.5.2b) is equivalent to a family of *scalar* problems:

$$\begin{aligned} & \Delta p_k \in \underset{p \in L^2(\mathbb{D}_{j_k})}{\operatorname{argmin}} \mathcal{S}(g_{j_k}^{\text{obs}}; F_{j_k}(p_k + \tilde{u}_{j_k}^{1/2} \cdot p)) + \alpha \|p\|_{L^2}^2 \\ \Leftrightarrow & \Delta p_k(x) \in \underset{y \in \mathbb{R}}{\operatorname{argmin}} s_{j_k}(x, p_k(x) + \tilde{u}_{j_k}(x)^{1/2} y) + \alpha y^2 \\ &= \tilde{u}_{j_k}(x)^{-1/2} \cdot \left(\operatorname{prox}(s_{j_k}(x, \cdot))(p_k(x), 2\tilde{u}_{j_k}(x)^{1/2}/\alpha) - p_k(x) \right) \end{aligned} \quad (11.5.6)$$

for almost all $x \in \mathbb{D}_{j_k}$. Here, the usual *proximal map* of a functional has been introduced:

$$\text{prox}(s)(y, \tau) := \underset{x \in \mathbb{R}}{\text{argmin}} s(x) + \frac{(x - y)^2}{2\tau} \quad (11.5.7)$$

The prox in (11.5.6) can be typically evaluated numerically in $\mathcal{O}(1)$ floating-point operations. Hence, a discretized form of the optimization problem in (11.5.2) can be solved in $\mathcal{O}(m_{\text{proj}})$, where m_{proj} is the number of degrees-of-freedom of a discretized projection. This enables evaluations of (11.5.2) at literally the same computational costs as classical SART-iterations, compare §11.3.2 – even for highly non-trivial choices of \mathcal{S} , as demonstrated by the subsequent examples.

11.5.1.1 (Weighted) L^2 -fidelities

For completeness, we mention the case of (weighted) L^2 -data fidelities $\mathcal{S}(g_j^{\text{obs}}; F_j(p)) = \|(F_j(p) - g_j^{\text{obs}})/\sigma_j\|_{L^2}^2$, which are adapted to data errors caused by Gaussian white noise of possibly spatially varying variance σ_j^2 . For $F_j = \text{id}$, this choice of \mathcal{S} is of the integral form (11.5.4) with $s_j(x, y) := (y - g_j^{\text{obs}}(x))^2/\sigma_j(x)^2$ and a simple proximal map, $\text{prox}(s_j(x, \cdot))(y, \tau) = (2\sigma_j(x)^2 y + \tau g_j^{\text{obs}}(x))/(2\sigma_j(x)^2 + \tau)$.

11.5.1.2 Robust GenSART

As a first non-standard application, we consider the problem of robust tomographic reconstruction: systematic errors in the acquisition geometry or modeling-inaccuracies due to nonlinear effects, as arising from metal-inclusions in soft tissue for example [10], tend to produce large outliers in the data, i.e. errors with highly non-Gaussian statistics. In such a setting, an L^2 -data-fidelity is far too greedy in fitting the data.

The problem has been addressed by employing a more robust (Huber-) L^1 -term, see e.g. [7], or even non-convex data-fidelity functionals such as the negative log-likelihood of the Student's t-distribution, as proposed in [25]. For trivial image-formation maps $F_j = \text{id}$, these choices correspond to a data-term $\mathcal{S}(g_j^{\text{obs}}; F_j(p))$ of the integral-form (11.5.4) with

$$s_j(x, y) = s_-(y - g_j^{\text{obs}}), \quad s_- \in \{s_{L^1_{\text{H}}, \nu}, s_{\text{s-t}, \nu}\} \quad (11.5.8a)$$

$$s_{L^1_{\text{H}}, \nu}(y) := \begin{cases} |y|^2 & \text{if } |y| \leq \nu \\ 2\nu|y| - \nu^2 & \text{else} \end{cases}, \quad s_{\text{s-t}, \nu}(y) := \nu^2 \ln(1 + |y|^2/\nu^2) \quad (11.5.8b)$$

Both functions in (11.5.8b) show the same quadratic growth behavior as $y \mapsto |y|^2$ for $|y| \ll \nu$ whereas, for $|y| \geq \nu$, the growth is only linear in the L^1 -Huber- and logarithmic in the Student's-t-case. Accordingly, the resulting data fidelities \mathcal{S} defined by (11.5.8a) and (11.5.4) behave like an L^2 -term for small deviations between the measured and fitted data, but penalize much less strongly when the deviations are larger, thereby yielding an increased robustness towards outliers. The proximal maps of $s_{L^1_{\text{H}}, \nu}$, $s_{\text{s-t}, \nu}$ are given by

$$\text{prox}(s_{L^1_{\text{H}}, \nu})(y, \tau) = y - \frac{2\nu\tau y}{\max\{|y|, 2\nu\tau + 1\}} \quad (11.5.9a)$$

$$\text{prox}(s_{\text{s-t}, \nu})(y, \tau) = \underset{x_0 \in \mathbb{R}}{\text{argmin}} \{s_{\text{s-t}, \nu}(x_0) : x_0 \in \mathbb{R}, (x_0^2 + \nu^2)(x_0 - y) + 2\tau\nu^2 x_0 = 0\} \quad (11.5.9b)$$

The cubic equation in (11.5.9b) has one or three real roots x_0 , which can be computed analytically in $\mathcal{O}(1)$. In the case of three solutions, $\operatorname{argmin} \{s_{s-t,\nu}(x_0)\}$ may be determined among these by simple trial-and-error. Consequently, the GenSART-scheme (11.5.2) can be evaluated efficiently via (11.5.6) for the considered robust data-fidelity terms. Notably, the simplicity of the L^1 -Huber-prox has already been exploited in [7] to construct efficient robust *ART*-iterations.

11.5.1.3 Poisson-noise-adapted GenSART

In many practical applications of X-ray- or electron tomography, the data errors are primarily due to the *Poisson-statistics* of the detection process: detector pixels actually count a *discrete* number of incident photons or electrons over some exposure time $t > 0$, where the counts follow a Poisson-distribution. Disregarding this effect may lead to severe anisotropic noise in the reconstruction, for example when a sample in CT is so strongly absorbing along certain incident directions that only very few counts are detected (*photon starvation*, see e.g. [151]).

If the detector is composed of $m_{\text{proj}} \in \mathbb{N}$ pixels with spatially varying detection sensitivities $\omega_i : \mathbb{D} \rightarrow \mathbb{R}_{\geq 0}$, the measured data is given by a vector

$$g_j^{\text{obs}} = (g_{ji}^{\text{obs}})_{i=1}^{m_{\text{proj}}} \in \mathbb{R}^{m_{\text{proj}}}, \quad g_{ji}^{\text{obs}} \sim \text{Poi}(tg_{ji}), \quad g_{ji} = \mathcal{M}_i(g_j) := \int_{\mathbb{D}} \omega_i g_j \, dx \quad (11.5.10)$$

where g_j denotes the exact continuous data and $X \sim \text{Poi}(\lambda)$ means that X is a Poisson-distributed random variable of intensity $\lambda \geq 0$. In this setting, the log-likelihood in (11.5.3) leads to the discrete Kullback-Leibler-divergence, see [99] for details:

$$\mathcal{S}^{\text{Poi}}(g_j^{\text{obs}}; g_j) := \sum_{i=1}^{m_{\text{proj}}} \text{KL}(g_{ji}^{\text{obs}}; t\mathcal{M}_i(g_j)), \quad \text{KL}(b; a) := \begin{cases} a - b - b \ln\left(\frac{a}{b}\right) & a, b \geq 0 \\ \infty & \text{else} \end{cases} \quad (11.5.11)$$

with the conventions that $\ln(0) = -\infty$ and $0 \cdot \ln(a/0) = 0$ for all $a \geq 0$.

Under the assumption that variations of the true data g_j are negligible within the support of ω_i (i.e. within a single pixel!), \mathcal{S}^{Poi} can be approximated in the form

$$\mathcal{S}^{\text{Poi}}(g_j^{\text{obs}}; g_j) \approx \int_{\mathbb{D}} \text{KL}(g_{j,\text{cont}}^{\text{obs}}(x); tg_j(x)) \cdot \omega(x) \, dx + c =: \mathcal{S}(g_j^{\text{obs}}; g_j) \quad (11.5.12)$$

where $\omega := \sum_{i=1}^{m_{\text{proj}}} \omega_i$, $g_{j,\text{cont}}^{\text{obs}}(x) := (\sum_{i=1}^{m_{\text{proj}}} g_{ji}^{\text{obs}} \omega_i(x) / \int_{\mathbb{D}} \omega_i \, dx) / \omega(x)$ (with the convention that $0/0 = 0$) and c is independent of g_j . The derivation is given in appendix 11.D.

So far, we have not specified the image-formation operators F_j , relating the data g_j to the tomographic projections $P_j(f)$. We consider two different acquisition modes:

- *Dark-field imaging*: The exact data is directly proportional to the projections, i.e. $F_j(p_j) = I_j \cdot p_j$ where I_j is the illumination intensity. This applies for example to HAADF-STEM, a state-of-the-art electron tomography technique, see e.g. [148, 159].
- *Bright-field imaging*: The tomographic data gives the relative attenuation experienced by the illuminating beam, i.e. $F_j(p_j) = I_j \cdot \exp(-p_j)$. This is the model for classical (monochromatic) X-ray computed tomography.

Inserting these models into (11.5.12), it can be seen that the resulting data-term $\mathcal{S}(g_j^{\text{obs}}; F_j(p))$ is of the integral-form (11.5.4) with

$$s_j(x, y) = \begin{cases} \text{KL}(g_{j,\text{cont}}^{\text{obs}}(x); tI_j(x) \cdot y) & \text{(dark-field)} \\ tI_j(x) \exp(-y) + g_{j,\text{cont}}^{\text{obs}}(x) \cdot y & \text{(bright-field)} \end{cases}. \quad (11.5.13)$$

In the dark-field case, $\text{prox}(s_j(x, \cdot))$ has a well-known closed form, see e.g. [99]. In the bright-field case, $\text{prox}(s_j(x, \cdot))$ may be evaluated numerically by a few iterations of Newton's method. This admits an efficient implementation of Poisson-noise-adapted GenSART-schemes.

11.5.2 Regularized Newton-Kaczmarz-GenSART

Regularized Newton-Kaczmarz methods have been proposed in [30] for the solution of general block-structured inverse problems $G(f) = (G_1(f), \dots, G_N(f)) = (g_1^{\text{obs}}, \dots, g_N^{\text{obs}})$ with *nonlinear* forward operators $G_j : X \rightarrow Y_j$ between Hilbert spaces X, Y_1, \dots, Y_N . In its simplest form, the approach boils down to performing Levenberg-Marquardt iterations on the different sub-problems $G_j(f) = g_j^{\text{obs}}$:

$$f_{k+1} = \underset{f \in X}{\text{argmin}} \left\| G_{j_k}(f_k) + G'_{j_k}[f_k](f - f_k) - g_{j_k}^{\text{obs}} \right\|_{Y_j}^2 + \alpha \|f - f_k\|_X^2 \quad (11.5.14)$$

where $G'_j[f_k] : X \rightarrow Y_j$ denotes the Fréchet-derivative of the operator at f_k . In the following, two nonlinear tomographic reconstruction problems are presented for which the iterations (11.5.14) can be computed efficiently via generalized SART-schemes.

11.5.2.1 Propagation-based X-ray phase contrast tomography

We consider the setting of (propagation-based) X-ray phase contrast tomography (XPCT), see e.g. [42, 13, 122, 176, 143]. In this experimental setup, the recorded data is given by near-field diffraction patterns, that relate to tomographic projections of the object density via a highly non-trivial image-formation operator $F_j = F$: under the standard assumptions of an ideal, fully coherent X-ray beam and negligible absorption (often a very good approximation at high X-ray energies), the measured parallel-beam(!) data under the j th tomographic incident direction is modeled by

$$g_j := F(\mathcal{P}_j(f)) \quad \text{with} \quad F(p) := |\mathcal{D}(\exp(-ip))|^2 - 1. \quad (11.5.15)$$

Here, $|\cdot|^2$ denotes the pointwise squared modulus of a complex-valued field and \mathcal{D} is the *Fresnel propagator*, which is given by a unitary Fourier-multiplier (\mathcal{F} : Fourier transform):

$$\mathcal{D}(\psi) := \mathcal{F}^{-1}(m_{\mathfrak{f}} \cdot \mathcal{F}(\psi)), \quad m_{\mathfrak{f}}(\boldsymbol{\xi}) := \exp(-i\boldsymbol{\xi}^2/(4\pi\mathfrak{f})) \quad \text{for} \quad \boldsymbol{\xi} \in \mathbb{R}^m, \quad (11.5.16)$$

where \mathfrak{f} is the Fresnel number of the imaging setup. The example matches the Newton-Kaczmarz setting if we define $G_j(f) := F(\mathcal{P}_j(f))$. The Fréchet-derivative is given by

$$G_j[f]h = F'[\mathcal{P}_j(f)]\mathcal{P}_j(h), \quad F'[p]h_p = 2\text{Im} \left(\overline{\mathcal{D}(\exp(-ip))} \cdot \mathcal{D}(\exp(-ip) \cdot h_p) \right) \quad (11.5.17)$$

where the overbar denotes complex conjugation and Im the pointwise imaginary part. Newton-Kaczmarz iterations for this problem with L^2 -data-fidelity and Sobolev- $W^{1,2}$ -penalty, as first proposed in [143], are of the form

$$f_{k+1} = \underset{f \in L^2(\Omega)}{\text{argmin}} \left\| F(\mathcal{P}_{j_k}(f_k)) + F'[\mathcal{P}_{j_k}(f_k)](\mathcal{P}_{j_k}(f - f_k)) - g_{j_k}^{\text{obs}} \right\|_{L^2}^2 + \alpha \left((1 - \gamma) \|f - f_k\|_{L^2}^2 + \gamma \|\nabla(f - f_k)\|_{L^2}^2 \right). \quad (11.5.18)$$

If we take $\mathcal{S}(p) := \|F(\mathcal{P}_{j_k}(f_k)) + F'[\mathcal{P}_{j_k}(f_k)](p - \mathcal{P}_{j_k}(f_k)) - g_{j_k}^{\text{obs}}\|_{L^2}^2$, (11.5.18) matches the setting of theorem 11.15. Hence, within the minor approximation discussed in the end of §11.4.3, the minimizer may be computed via the GenSART-scheme

$$f_{k+1} \approx f_k + \mathcal{P}_{j_k}^* \left(u_{j_k}^{-1/2} \cdot \underset{p \in L^2(\mathbb{D}_{j_k})}{\text{argmin}} \left\| F'[\mathcal{P}_{j_k}(f_k)](u_{j_k}^{1/2} \cdot p) - r_{j_k} \right\|_{L^2}^2 + \alpha_k (1 - \gamma) \|p\|_{L^2}^2 + \alpha_k \gamma \|u_{j_k}^{1/2} \cdot \nabla_{\mathbb{D}}(u_{j_k}^{-1/2} \cdot p)\|_{L^2}^2 \right) \quad (11.5.19)$$

with residual $r_{j_k} := g_{j_k}^{\text{obs}} - F(\mathcal{P}_{j_k}(f_k))$. The quadratic optimization problem in (11.5.19) can be solved for example by a conjugate-gradient method applied to the normal equation.

11.5.2.2 Polychromatic CT

If the polychromatic nature of the X-rays in conventional CT-scanners is neglected, so called *beam-hardening* artifacts may arise [10]. In [51, 104], a simplified model for polychromatic CT has been proposed, which partially accounts for the arising nonlinear effects. Within this model, the detected intensity data g_j for the j th tomographic projection is predicted as

$$g_j = \int \underbrace{I_0(\varepsilon) \exp \left(- \Phi(\varepsilon) P_j(\phi(f)) - \Theta(\varepsilon) P_j(\theta(f)) \right)}_{=: G_{j,\varepsilon}(f)} d\varepsilon =: G_j(f). \quad (11.5.20)$$

Here, $I_0(\varepsilon)$ is the emitted intensity of the X-ray source at photon-energy ε and f is the spatially varying attenuation at some reference-energy ε_0 . $\Phi(\varepsilon)\phi(f)$ and $\Theta(\varepsilon)\theta(f)$ model the attenuation's photo-electric- and Compton-scattering-components, respectively. The main approximation is that these functions are assumed to be representable as a function of f multiplied by energy-dependent scaling factors:

$$\Phi(\varepsilon) = \frac{\varepsilon_0^3}{\varepsilon^3} \quad \text{and} \quad \Theta(\varepsilon) = \frac{f_{\text{KN}}(\varepsilon)}{f_{\text{KN}}(\varepsilon_0)}, \quad f_{\text{KN}}: \text{Klein-Nishina function} \quad (11.5.21)$$

Note that the expressions $\phi(f)$ and $\theta(f)$ are to be understood pointwise, i.e. $\phi(f)(x) = \phi(f(x))$ and $\theta(f)(x) = \theta(f(x))$ (with a slight abuse of notation). The scalar functions $\phi, \theta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ interpolate known value-pairs $\{(f_m, \phi_m)\}$ and $\{(f_m, \theta_m)\}$ for different materials $m = 1, 2, \dots$ in the imaged object, such as water, fat and bone, see [51, 104] for details. We assume that ϕ and θ are continuously differentiable.

Notably, the nonlinearity in (11.5.20) *cannot* be described by a nonlinear map acting on the projections, i.e. $G_j(f) \neq F_j(P_j(f))$, so that the setting does not seem to match our tomographic

model (11.2.1). However, we may still compute the Fréchet-derivative:

$$G'_j[f]h_f = -G_j^\Phi(f) \cdot P_j(\phi'(f) \cdot h_f) - G_j^\Theta(f) \cdot P_j(\theta'(f) \cdot h_f) \quad (11.5.22a)$$

$$G_j^\Phi(f) := \int \Phi(\varepsilon)G_{j,\varepsilon}(f) \, d\varepsilon, \quad G_j^\Theta(f) := \int \Theta(\varepsilon)G_{j,\varepsilon}(f) \, d\varepsilon \quad (11.5.22b)$$

By (11.3.5), it holds that $G_j^\Phi(f) \cdot P_j(\phi'(f) \cdot h_f) = P_j(\phi'(f) \cdot h_f \cdot P_j^B(G_j^\Phi(f)))$, which yields

$$G'_j[f]h_f = P_j(\lambda_j(f) \cdot h_f), \quad \lambda_j(f) = \phi'(f) \cdot P_j^B(G_j^\Phi(f)) + \theta'(f) \cdot P_j^B(G_j^\Theta(f)) \quad (11.5.23a)$$

Using (11.5.23), Newton-Kaczmarz iterations (11.5.14) for the considered problem with an L^2 -penalty can be written in the form

$$f_{k+1} = \operatorname{argmin}_{f \in L^2(\Omega)} \|P_{j_k}(\lambda_{j_k}(f_k) \cdot (f - f_k)) - r_{j_k}\|_{L^2}^2 + \alpha_k \|f - f_k\|_{L^2}^2 \quad (11.5.24)$$

with residual $r_{j_k} = g_{j_k}^{\text{obs}} - G_{j_k}(f_k)$. (11.5.24) matches the setting of theorem 11.11. By rearranging the resulting GenSART-scheme (11.4.7) for (11.5.24), we obtain the update-formula

$$f_{k+1} = f_k + \lambda_{j_k}(f_k) \cdot P_{j_k}^* \left(\frac{r_{j_k}}{P_{j_k}(w_{j_k} \cdot |\lambda_{j_k}(f_k)|^2) + \alpha_k} \right) \quad (11.5.25)$$

where w_{j_k} denotes the ray-density to the projector P_{j_k} . Including the necessary computations of $G_{j_k}(f_k)$ and $\lambda_{j_k}(f_k)$, evaluating (11.5.25) requires three evaluations of the forward- and back-projectors P_{j_k} and $P_{j_k}^*$, plus computationally inexpensive pointwise operations.

Similarly efficient formulas may be obtained if the L^2 -data-fidelity in (11.5.24) is replaced by the Poisson-noise-adapted Kullback-Leibler term from §11.5.1.3.

11.5.3 Extensions

The following section outlines different ideas on how to extend the generalized SART-approach to an even more versatile tool for devising efficient Kaczmarz-type reconstruction methods.

11.5.3.1 Box constraints

Analogously as in other Kaczmarz-methods, box constraints $f_{\min} \leq f \leq f_{\max}$ on the admissible values of the object f may be incorporated in GenSART-schemes simply by setting

$$f_{k+1} \leftarrow \max \{ \min \{ f_{k+1}, f_{\max} \}, f_{\min} \}. \quad (11.5.26)$$

after each iteration. This approach can be interpreted as interlacing *projections* onto the convex set $\{f \in L^2(\Omega) : f_{\min} \leq f \leq f_{\max}\}$ and is standard in Kaczmarz-type tomographic reconstructions ever since the introduction of ART [77].

11.5.3.2 Additional quadratic regularizer

So far, the penalization in the considered Kaczmarz-iterations was always with respect to the preceding iterate f_k . In addition, it might be desirable to impose a static regularizer such that the total penalty is given by

$$\mathcal{R}_k(f) = \alpha_1 \|T(f - f_k)\|_Z^2 + \alpha_2 \|T(f - f_{\text{ref}})\|_Z^2, \quad (11.5.27)$$

for $\alpha_1, \alpha_2 > 0$, Hilbert spaces X, Z and a bounded linear operator $T : X \rightarrow Z$. By writing the squared norms as inner products, the expression can be cast to the form

$$\mathcal{R}_k(f) = (\alpha_1 + \alpha_2) \|T(f - f_{k,\text{ref}})\|_Z^2 + c \quad (11.5.28)$$

where $f_{k,\text{ref}} := (\alpha_1 f_k + \alpha_2 f_{\text{ref}}) / (\alpha_2 + \alpha_1)$ and the constant $c \in \mathbb{R}$ is independent of f and thus irrelevant for the computation of the minimizer. Hence, the resulting Kaczmarz-iterations are of the same form as before, up to a modified reference solution $f_k \rightarrow f_{k,\text{ref}}$ and regularization parameter $\alpha \rightarrow \alpha_1 + \alpha_2$. This means that also the derived GenSART-formulas only have to be modified by exchanging these parameters.

11.5.3.3 Kaczmarz-type splitting and primal-dual methods

Variational reconstruction methods seek to minimize terms of the form $\mathcal{S}_{\text{tot}}(P_{\text{tot}}(f)) + \mathcal{R}(f)$ with $\mathcal{S}_{\text{tot}}(p_1, \dots, p_{N_{\text{proj}}}) = \sum_{j=1}^{N_{\text{proj}}} \mathcal{S}(p_j)$, compare §11.2.3. Often, this is achieved by some *splitting* method, alternating (sub-)gradient-descent- (“forward-”) or proximal (“backward-”) iterations with respect to the data-fidelity \mathcal{S}_{tot} and the penalty functional \mathcal{R} (or their duals), see e.g. [43]. It is straightforward to combine such an approach with a Kaczmarz-type strategy that exploits the block-structure of \mathcal{S}_{tot} to reduce computational costs of the individual iterations. Examples of (primal) Kaczmarz-type splitting methods are given by iterations ($\partial\mathcal{R}$: subdifferential)

$$f_{k+\frac{1}{2}} \in \operatorname{argmin}_{f \in L^2(\Omega)} \mathcal{S}_{j_k}(P_{j_k}(f)) + \frac{1}{2\tau_k} \|f - f_k\|_{L^2}^2 \quad (11.5.29a)$$

$$f_{k+1} \in \begin{cases} \operatorname{argmin}_{f \in L^2(\Omega)} \mathcal{R}(f) + \frac{1}{2\sigma_k} \|f - f_{k+\frac{1}{2}}\|_{L^2}^2 & \text{(backward-backward)} \\ f_{k+\frac{1}{2}} - \sigma_k \partial\mathcal{R}(f_{k+\frac{1}{2}}) & \text{(backward-forward)} \end{cases} \quad (11.5.29b)$$

with stepsize-parameters τ_k, σ_k . Algorithms of this kind have been proposed and analyzed in [21, 22, 7]. Importantly, the proximal iteration for the data-fidelity, (11.5.29a), can be computed efficiently via the GenSART-scheme from theorem 11.10. Similarly, GenSART-formulas may be used to compute proximal steps in block-primal-dual methods as considered in [35].

11.6 Numerical examples

All of the GenSART-schemes from §11.5 have been successfully implemented as numerical algorithms. In the following, exemplary results are presented.

11.6.1 Implementation

In previous studies, Kaczmarz-type reconstruction methods have usually been derived for a *discretized* tomographic model. On the contrary, the theory in this work relies on properties of the parallel- or cone-beam projectors $P \in \{\mathcal{P}, \mathcal{D}\}$ that are valid only in *continuous space*. In particular, while the generalized SART theorem 11.7 is equally valid in finite dimensions, the tentative simplicity of PP^* does typically *not* carry over to discretizations of these operators. Moreover, discrete analogues of the penalty functionals considered in §11.4 will in general no longer satisfy assumption 11.6.

For this reason, we pursue a *post-discretization* strategy: we propose Kaczmarz-iterations within a *continuous* tomographic model, then devise GenSART-schemes for their computation using the results from the preceding sections and *finally* discretize these schemes to obtain numerically implementable SART-like iterations. If $\mathbf{P} \in \mathbb{R}^{m \times n}$, $\mathbf{f} \in \mathbb{R}^n$, S , $\mathbf{u} \in \mathbb{R}^m$ are suitable discretizations of P, f, \mathcal{S}, u_P , then this would look as follows for L^2 -penalized Kaczmarz:

$$f_{\text{new}} \in \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \mathcal{S}(P(f)) + \alpha \|f - f_{\text{ref}}\|_{L^2}^2 \quad (11.6.1)$$

$$\stackrel{(\text{GenSART})}{\rightsquigarrow} f_{\text{new}} \in f_{\text{ref}} + P^* \left(\tilde{u}_P^{-1/2} \cdot \underset{p \in L^2(\mathbb{D}_P)}{\operatorname{argmin}} \mathcal{S}(P(f_{\text{ref}}) + \tilde{u}_P^{1/2} \cdot p) + \alpha \|p\|_{L^2}^2 \right) \quad (11.6.2)$$

$$\stackrel{(\text{discretize})}{\rightsquigarrow} \mathbf{f}_{\text{new}} \in \mathbf{f}_{\text{ref}} + \mathbf{P}^* \left(\tilde{\mathbf{u}}^{-1/2} \odot \underset{\mathbf{p} \in \mathbb{R}^m}{\operatorname{argmin}} S(\mathbf{P}(\mathbf{f}_{\text{ref}}) + \tilde{\mathbf{u}}^{1/2} \odot \mathbf{p}) + \alpha \|\mathbf{p}\|_2^2 \right) \quad (11.6.3)$$

A drawback of the approach is that the discrete GenSART-update (11.6.3) in general does not *exactly* solve a discretized version of (11.6.1), i.e. $\mathbf{f}_{\text{new}} \notin \underset{\mathbf{f} \in \mathbb{R}^n}{\operatorname{argmin}} S(\mathbf{P}(\mathbf{f})) + \alpha \|\mathbf{f} - \mathbf{f}_{\text{ref}}\|_2^2$, but only up to *discretization errors*. While this inexactness might lead to numerical instabilities in principle, no such effects are observed for the examples in §11.6.2 and §11.6.3.

As usual, we assume discrete objects $\mathbf{f} \in \mathbb{R}^n$ and data $\mathbf{g}_j \in \mathbb{R}^{m_{\text{proj}}}$ to provide samples of the continuous quantities $f \in L^2(\Omega)$ and g_j on equidistant Cartesian grids. Pointwise operations on functions are then represented by element-wise operations of vectors. Integrals in continuous space can be approximated by summation over the entries of the corresponding vectors in the discretized model and derivatives can be implemented via finite differences. For example, L^q -norms are then identified with q -norms in \mathbb{R}^n , i.e.

$$\int_{\Omega} |f(\mathbf{x})|^q d\mathbf{x} = \|f\|_{L^q}^q \sim \|\mathbf{f}\|_q^q = \sum_{i=1}^n |\mathbf{f}_i|^q \quad \text{if } \mathbf{f} = (\mathbf{f}_i)_{i=1}^n \text{ discretizes } f. \quad (11.6.4)$$

Discrete and continuous quantities can be related via sampling operators,

$$S_{\text{O}} : L^2(\Omega) \rightarrow \mathbb{R}^n; f \mapsto \left(\int_{\mathbb{V}_i} f d\mathbf{x} \right)_{i=1}^n \quad (11.6.5)$$

for voxels $\mathbb{V}_i \subset \Omega$ disjointly covering the object-domain Ω , and an analogous map $S_{\text{D}} : L^2(\mathbb{D}) \rightarrow \mathbb{R}^{m_{\text{proj}}}$ in the projection-domain. Projectors $\mathbf{P}_j \in \{\mathcal{P}_{\theta}, \mathcal{D}_s\}$ may then be naturally discretized via $\mathbf{P}_j := S_{\text{D}} P_j S_{\text{O}}^*$. See also [209] for alternative discretizations.

Unit-projections and precomputations: Recall that all of the generalized SART formulas in §11.4 and §11.5 involve (weighted) unit-projections u_j or \tilde{u}_j . Clearly, discrete approximations $\mathbf{u}_j, \tilde{\mathbf{u}}_j$ of these are needed for numerical computations. For general object-domains Ω , these may be precomputed via $\mathbf{u}_j = \mathbf{P}_j(\mathbf{1}_{\mathbb{R}^n})$ and $\tilde{\mathbf{u}}_j = \mathbf{P}_j(\mathbf{w}_j)$, where \mathbf{w}_j is a suitable discretization of the ray-density. In this case, GenSART-methods thus require one additional evaluation of the full projector $(\mathbf{P}_1, \dots, \mathbf{P}_{N_{\text{proj}}})$ prior to the actual iterations. However, for geometrically simple domains Ω such as boxes, cylinders and balls, $u_j = P_j(\mathbf{1}_{\Omega})$, $\tilde{u}_j = P_j(w_j)$ can also be computed analytically. Thereby, costly precomputations may be avoided.

Furthermore, discretized GenSART-schemes typically involve element-wise *divisions* by powers of \mathbf{u}_j or $\tilde{\mathbf{u}}_j$. In accordance with convention 11.2 for the underlying continuous-space model, such operations should only be performed for *non-vanishing* entries of $\mathbf{u}_j, \tilde{\mathbf{u}}_j$. For all other entries, the result of expressions of the form $\mathbf{p} \oslash \mathbf{u}_j^{\mu}$ may simply be taken to be zero.

11.6.2 Robust tomography test case

As a first numerical example, we consider the application of robust reconstruction from tomographic projection data, as introduced in 11.5.1.2. To this end, we compare L^2 -regularized Tikhonov-regularization and Kaczmarz-iterations

$$f^{\text{Tik}} = \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \mathcal{S}_{\text{tot}}(g_{\text{tot}}^{\text{obs}}; P_{\text{tot}}(f)) + \alpha_{\text{Tik}} \|f\|_{L^2}^2 \quad (11.6.6a)$$

$$\begin{aligned} f_{k+1} &= \underset{f \in L^2(\Omega)}{\operatorname{argmin}} \mathcal{S}(g_{j_k}^{\text{obs}}; P_{j_k}(f)) + \alpha \|f - f_k\|_{L^2}^2 \\ &= f_k + P_{j_k}^* \left(\tilde{u}_{j_k}^{-1/2} \cdot \underset{p \in L^2(\mathbb{D}_{j_k})}{\operatorname{argmin}} \mathcal{S}(g_{j_k}^{\text{obs}}; P_{j_k}(f_k) + \tilde{u}_{j_k}^{1/2} \cdot p) + \alpha \|p\|_{L^2}^2 \right). \end{aligned} \quad (11.6.6b)$$

with data-fidelities given by $\mathcal{S}_{\text{tot}}(g_{\text{tot}}^{\text{obs}}; P_{\text{tot}}(f)) = \sum_{j=1}^{N_{\text{proj}}} \mathcal{S}(g_j^{\text{obs}}; P_j(f))$ where

$$\mathcal{S}(g_j^{\text{obs}}; p) = \int_{\mathbb{D}_j} s(p(\mathbf{x}) - g_j^{\text{obs}}(\mathbf{x})) \, d\mathbf{x}, \quad s \in \{s_{L^2}, s_{L_{\text{H}}^1, \nu}, s_{\text{s-t}, \nu}\}. \quad (11.6.7)$$

The L^1 -Huber- and Student's t-functions $s_{L_{\text{H}}^1, \nu}$ and $s_{\text{s-t}, \nu}$ are defined in (11.5.8) and $s_{L^2}(y) := |y|^2$ simply corresponds to an L^2 -data-fidelity. s_{L^2} , $s_{L_{\text{H}}^1, \nu}$ and $s_{\text{s-t}, \nu}$ are plotted in fig. 11.2(c).

As numerical reconstruction methods, we consider for once a discretized version of the Tikhonov regularization in (11.6.6a), stated as algorithm 11.1. On the other hand, we design a generalized SART-analogue, algorithm 11.2, by discretizing the update-formula (11.6.6b). Note that the discrete analogue of the optimization problem in (11.6.6b) factorizes into a family of scalar problems just like in the continuous setting, see §11.5.1. This enables a highly efficient implementation of this step regardless of the choice $s \in \{s_{L^2}, s_{L_{\text{H}}^1, \nu}, s_{\text{s-t}, \nu}\}$.

Algorithm 11.1 Robust Tikhonov reconstruction

Input: Data $\mathbf{g}_{\text{tot}}^{\text{obs}} \in \mathbb{R}^{m_{\text{proj}} N_{\text{proj}}}$, projector $\mathbf{P}_{\text{tot}} \in \mathbb{R}^{(m_{\text{proj}} N_{\text{proj}}) \times n}$, regularization parameter $\alpha_{\text{Tik}} > 0$, data-fidelity $s \in \{s_{L^2}, s_{L_{\text{H}}^1, \nu}, s_{\text{s-t}, \nu}\}$, $\mathcal{S}_{\text{tot}}(\mathbf{g}; \mathbf{p}) := \sum_{i=1}^{m_{\text{proj}} N_{\text{proj}}} s(p_i - g_i)$

Output: $\mathbf{f}^{\text{Tik}} \in \underset{\mathbf{f} \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{S}_{\text{tot}}(\mathbf{g}_{\text{tot}}^{\text{obs}}; \mathbf{P}_{\text{tot}} \mathbf{f}) + \alpha_{\text{Tik}} \|\mathbf{f}\|_2^2$

Algorithm 11.2 Robust SART reconstruction

Input: Data $\mathbf{g}_j^{\text{obs}} \in \mathbb{R}^{m_{\text{proj}}}$, projectors $\mathbf{P}_j \in \mathbb{R}^{m_{\text{proj}} \times n}$, regularization parameter $\alpha > 0$, initial guess $\mathbf{f}_0 \in \mathbb{R}^n$, data-fidelity $s \in \{s_{L^2}, s_{L_{\text{H}}^1, \nu}, s_{\text{s-t}, \nu}\}$, $\mathcal{S}(\mathbf{g}; \mathbf{p}) := \sum_{i=1}^{m_{\text{proj}}} s(p_i - g_i)$, weighted unit-projections $\tilde{\mathbf{u}}_j$.

for $k = 0, \dots, k_{\text{stop}} - 1$ **do**

$$\mathbf{p}_k = \mathbf{P}_{j_k}(\mathbf{f}_k)$$

$$\Delta \mathbf{p}_k \in \underset{\mathbf{p} \in \mathbb{R}^{m_{\text{proj}}}}{\operatorname{argmin}} \mathcal{S}(g_{j_k}^{\text{obs}}; \mathbf{p}_k + \tilde{\mathbf{u}}_{j_k}^{1/2} \odot \mathbf{p}) + \alpha \|\mathbf{p}\|_2^2$$

$$\mathbf{f}_{k+1} = \mathbf{f}_k + \mathbf{P}_{j_k}^* \left(\Delta \mathbf{p}_k \odot \tilde{\mathbf{u}}_{j_k}^{1/2} \right)$$

end for

Output: final object-iterate $\mathbf{f}_{k_{\text{stop}}}$

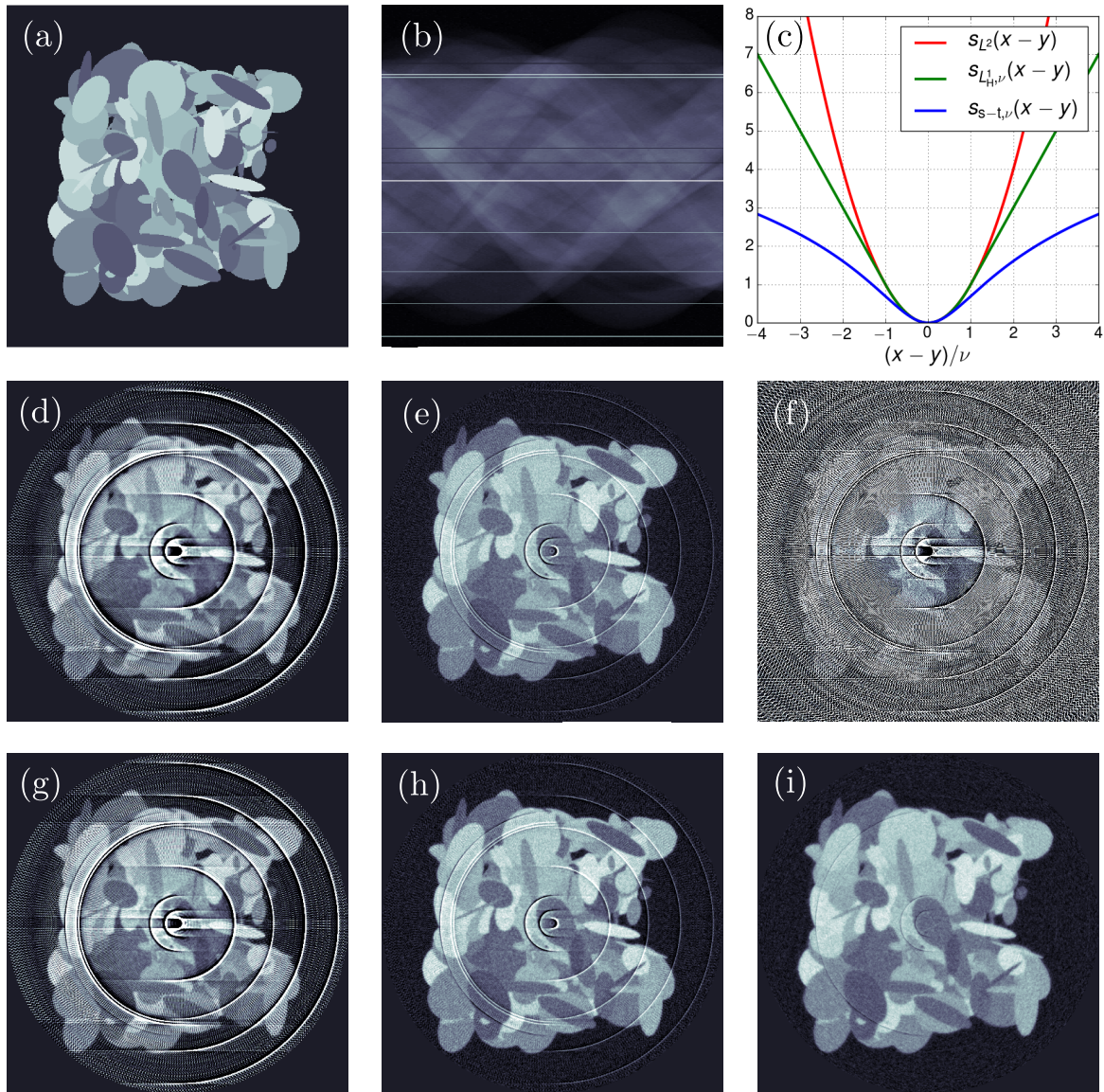


Figure 11.2. Numerical robust tomography test case: (a) 512×512 -sized object phantom. (b) Simulated parallel-beam data. (c) Plot of the different data-fidelity functions. (f) FBP-reconstruction. (d),(e) Tikhonov-reconstruction (algorithm 11.1) with L^2 - and L^1 -Huber-data-fidelity. (g)–(i) SART-reconstruction (algorithm 11.2) with L^2 - and L^1 -Huber- and Student’s-t data-fidelity. The linear color scale in all object-plots (subfigures (a) and (d)–(i)) is identical. See text for details.

We test the different tomographic reconstruction algorithms for a 2D-phantom f^{exact} of image-size 512×512 pixels composed of random ellipses, shown in fig. 11.2(a). Note that such a 2D-setting can be treated as a limiting case of the 3D-geometries studied in this work. Parallel-beam data is simulated from the phantom under tomographic incident angles $\theta = 0^\circ, 1^\circ, \dots, 179^\circ$. To avoid “inverse crime”, different discretizations of the tomographic projectors are used in the data-simulation and the reconstruction. The simulated exact data is corrupted by Gaussian white noise of relative magnitude 2% in 2-norm and additionally by systematic outliers: following [25], we simulate 2% randomly chosen dead detector pixels that measure

a constant incorrect value in all projections. Such measurement errors manifest as horizontal stripes in the sinogram-plot of the tomographic data shown in fig. 11.2(b).

The data is reconstructed with algorithms 11.1 and 11.2 using the data-fidelities $s_{L^2}, s_{L^1_{\text{H}}, \nu}, s_{\text{s-t}, \nu}$ with regularization parameters $\alpha_{\text{Tik}} = 300$ and $\alpha = 2\alpha_{\text{Tik}}$, the latter choice being motivated by §11.2.4. The robustness-parameter ν (compare §11.5.1.2) is taken as 20% of the standard-deviation of the data. Algorithm 11.2 is run for one symmetric Kaczmarz-cycle, i.e. $k_{\text{stop}} = 2N_{\text{proj}}$ and $j_k = j_{k_{\text{stop}}-1-k}$ for all k . The processing order $j_0, \dots, j_{N_{\text{proj}}-1}$ of the tomographic projections is chosen according to a multi-level-scheme from [78].

The Tikhonov results for the data-fidelities $s \in \{s_{L^2}, s_{L^1_{\text{H}}, \nu}\}$ (computed by a conjugate-gradient method and a linearly convergent primal-dual algorithm from [36], respectively) are plotted in fig. 11.2(d) and (e). It is not clear how to reliably minimize the Tikhonov-functional for the non-convex Student's-t functional $s = s_{\text{s-t}, \nu}$ so that this case is omitted here. The GenSART-results for all(!) data-fidelities are plotted in fig. 11.2(f)-(h). For comparison, fig. 11.2(c) also shows a reconstruction by filtered back-projection (FBP), computed using an implementation from the ASTRA-toolbox [199, 198] with default-parameters.

As expected, the results for L^2 -data fidelities (fig. 11.2(d),(f)) show pronounced ring- and pattern-artifacts arising from the dead pixels – not to speak of the FBP-solution in fig. 11.2(c). The artifacts are significantly reduced for the L^1 -Huber-reconstructions (fig. 11.2(e),(g)). Finally, the reconstruction with the non-convex and even more robust Student's-t data-fidelity in fig. 11.2(h) is almost completely free of artifacts.

Notably, for both the L^2 - and the non-quadratic(!) L^1 -Huber-term, the Tikhonov- and GenSART-reconstructions turn out to be qualitatively indistinguishable. This suggests that variational methods can indeed be emulated by Kaczmarz-iterations – even in settings where this is not predicted by the convergence theory from §11.2.4. At the same time, it should be emphasized that the computational costs are practically identical for all three GenSART-methods, essentially amounting to *two* evaluations of the full projector \mathbf{P}_{tot} and its adjoint $\mathbf{P}_{\text{tot}}^*$, which is significantly less than for the Tikhonov-reconstructions: for the non-quadratic L^1 -Huber-case the primal-dual algorithm requires 73 iterations, and thus 73(!) evaluations of $\mathbf{P}_{\text{tot}}, \mathbf{P}_{\text{tot}}^*$, to converge to a prescribed accuracy of 1% according to a criterion from [36].

11.6.3 Newton-Kaczmarz-GenSART for experimental XPCT-data

For a second and somewhat more involved numerical test case, we implement the Newton-Kaczmarz-iterations for X-ray phase contrast tomography (XPCT) from §11.5.2.1. To this end, the obtained GenSART-formula (11.5.19) is discretized, where the gradients ∇ are replaced by finite-difference operators $\nabla \in \mathbb{R}^{M_{\text{grad}} \times m_{\text{proj}}}$, $(\nabla((p_i)_{i=1}^{m_{\text{proj}}}))_m = p_{k_m} - p_{i_m}$ for index pairs (i_m, k_m) of neighboring pixels. Importantly, as implied by the treatment of the continuous-space setting in §11.4.3, the set $\mathcal{N}_{\nabla} = \{(i_m, k_m)\}_{m=1}^{M_{\nabla}}$ must be restricted to such neighbor-index-pairs, for which the unit-projection $\mathbf{u}_j = (u_{ji})_{i=1}^{m_{\text{proj}}}$ does not vanish for either of the indices, i.e. $u_{ji_m} \neq 0$ and $u_{jk_m} \neq 0$. In a nutshell, this means that the discrete gradient ∇ must only be computed *within* the support of the unit-projection in accordance with convention 11.2. Upon discretizing (11.5.19), the optimization problem in projection-space assumes the form

$$\Delta \mathbf{p}_k = \underset{\mathbf{p} \in \mathbb{R}^{m_{\text{proj}}}}{\text{argmin}} \left\| \mathbf{T}_k(\mathbf{u}_{j_k}^{1/2} \odot \mathbf{p}) - \mathbf{r}_k \right\|_2^2 + \alpha(1 - \gamma) \|\mathbf{p}\|_2^2 + \alpha\gamma \left\| \mathbf{U}_{\nabla, j_k}^{1/2} \nabla(\mathbf{p} \odot \mathbf{u}_{j_k}^{1/2}) \right\|_2^2 \quad (11.6.8)$$

where \mathbf{r}_k is the residual and $\mathbf{T}_k = F'[\mathbf{P}_{j_k}(\mathbf{f}_k)] \in \mathbb{R}^{m_{\text{data}} \times m_{\text{proj}}}$ is the Fréchet-derivative of the discretized image-formation map $F : \mathbb{R}^{m_{\text{proj}}} \rightarrow \mathbb{R}^{m_{\text{data}}}$. The $\mathbf{U}_{\nabla,j} \in \mathbb{R}^{M_{\text{grad}} \times M_{\text{grad}}}$ are diagonal, positive-semidefinite matrices that implement a discrete analogue $\nabla(\mathbf{p}) \mapsto \mathbf{U}_{\nabla,j} \nabla(\mathbf{p})$ of the multiplication $\nabla(p) \mapsto u_j \cdot \nabla(p)$ for gradients of continuous projections p .

With $\mathbf{U}_j := \text{diag}(\mathbf{u}_j) \in \mathbb{R}^{m_{\text{proj}} \times m_{\text{proj}}}$ and $\mathbf{I} \in \mathbb{R}^{m_{\text{proj}} \times m_{\text{proj}}}$ denoting the identity, (11.6.8) then amounts to solving the associated normal equation:

$$\Delta \mathbf{p}_k = \left(\mathbf{U}_{j_k}^{1/2} \mathbf{T}_k^* \mathbf{T}_k \mathbf{U}_{j_k}^{1/2} + \alpha(1 - \gamma) \mathbf{I} + \alpha \gamma \mathbf{U}_{j_k}^{-1/2} \nabla^* \mathbf{U}_{\nabla,j_k} \nabla \mathbf{U}_{j_k}^{-1/2} \right)^{-1} \mathbf{T}_k^* (\mathbf{u}_j^{1/2} \odot \mathbf{r}_k) \quad (11.6.9)$$

The map F as well as its derivative $F'[\mathbf{p}]$ can be implemented using only fast Fourier transforms and pointwise operations, leading to a computational complexity of $\mathcal{O}(m_{\text{data}} \log(m_{\text{data}}))$. Accordingly, the above symmetric positive-definite problem can be solved efficiently by a matrix-free conjugate-gradient (CG) method. The obtained Newton-Kaczmarz-GenSART method is summarized in algorithm 11.3, which is notably not limited to XPCT but may be adapted for a wide range of other image-formation operators F .

Algorithm 11.3 Newton-Kaczmarz-GenSART (for X-ray phase contrast tomography)

Input: Data $\mathbf{g}_j^{\text{obs}} \in \mathbb{R}^{m_{\text{data}}}$, parallel-beam projectors $\mathbf{P}_j \in \mathbb{R}^{n \times m_{\text{proj}}}$, initial guess $\mathbf{f}_0 \in \mathbb{R}^n$, regularization $\alpha > 0$, $\gamma \in [0; 1]$, unit-projections \mathbf{u}_j , $\mathbf{U}_j := \text{diag}(\mathbf{u}_j)$ and $\mathbf{U}_{\nabla,j} \in \mathbb{R}^{m_{\text{grad}} \times m_{\text{grad}}}$, image-formation operator $F : \mathbb{R}^{m_{\text{proj}}} \rightarrow \mathbb{R}^{m_{\text{data}}}$, Fréchet-derivative F' .

for $k = 0, \dots, k_{\text{stop}} - 1$ **do**

$$\mathbf{p}_k = \mathbf{P}_{j_k} \mathbf{f}_k, \mathbf{r}_k = \mathbf{g}_{j_k}^{\text{obs}} - F(\mathbf{p}_k), \mathbf{T}_k = F'[\mathbf{p}_k]$$

$$\Delta \mathbf{p}_k \stackrel{\text{CG}}{=} \left(\mathbf{U}_{j_k}^{1/2} \mathbf{T}_k^* \mathbf{T}_k \mathbf{U}_{j_k}^{1/2} + \alpha(1 - \gamma) \mathbf{I} + \alpha \gamma \mathbf{U}_{j_k}^{-1/2} \nabla^* \mathbf{U}_{\nabla,j_k} \nabla \mathbf{U}_{j_k}^{-1/2} \right)^{-1} \mathbf{T}_k^* (\mathbf{u}_j^{1/2} \odot \mathbf{r}_k)$$

$$\mathbf{f}_{k+1} = \mathbf{f}_k + \mathbf{P}_j^* (\mathbf{u}_j^{-1/2} \odot \Delta \mathbf{p}_k)$$

$$[\mathbf{f}_{k+1} = \max\{0, \mathbf{f}_{k+1}\}] \quad \% \text{ optional non-negativity constraint}$$

end for

Output: final object-iterate $\mathbf{f}_{k_{\text{stop}}}$

We apply algorithm 11.3 to experimental XPCT-data $\mathbf{g}_{\text{tot}}^{\text{obs}} \in \mathbb{R}^{N_{\text{proj}} m_{\text{data}}}$ from [143], composed of $N_{\text{proj}} = 249$ near-field diffraction patterns of size $m_{\text{data}} = 1024^2$ pixels each, acquired in an (effective) parallel-beam setting under tomographic incident angles between 0° and 173° . The 3D tomographic data, measured at the synchrotron light-source PETRAIII (see [178] for experimental details), is visualized by orthoslices in fig. 11.3(a).

It is known that the object \mathbf{f} – a colloidal nano-crystal – is contained in a centered cube of size 256^3 voxels, which we take as the reconstruction volume Ω (*support constraint*). Moreover, the true values of \mathbf{f} are known to be non-negative for physical reasons. We compute reconstructions using algorithm 11.3 with $\mathbf{f}_0 = 0$ and regularization parameter $\alpha = 500$ in four different setups: with $\gamma = 0$ (i.e. pure L^2 -penalty) and $\gamma = 0.8$ (primarily gradient-penalization), with and without a non-negativity constraint, respectively. For each of these settings, *one* (non-symmetric) Kaczmarz-cycle is run, again following the MLS-scheme from [78] for the processing order of the data. Figure 11.3(c)-(f) plots the central slices of the reconstructed 256^3 -voxel volumes perpendicular to the tomographic axis. For comparison, fig. 11.3(b) shows the same view for the reconstruction from [143], obtained by a similar L^2 -regularized Newton-Kaczmarz method, but without using a GenSART-scheme.

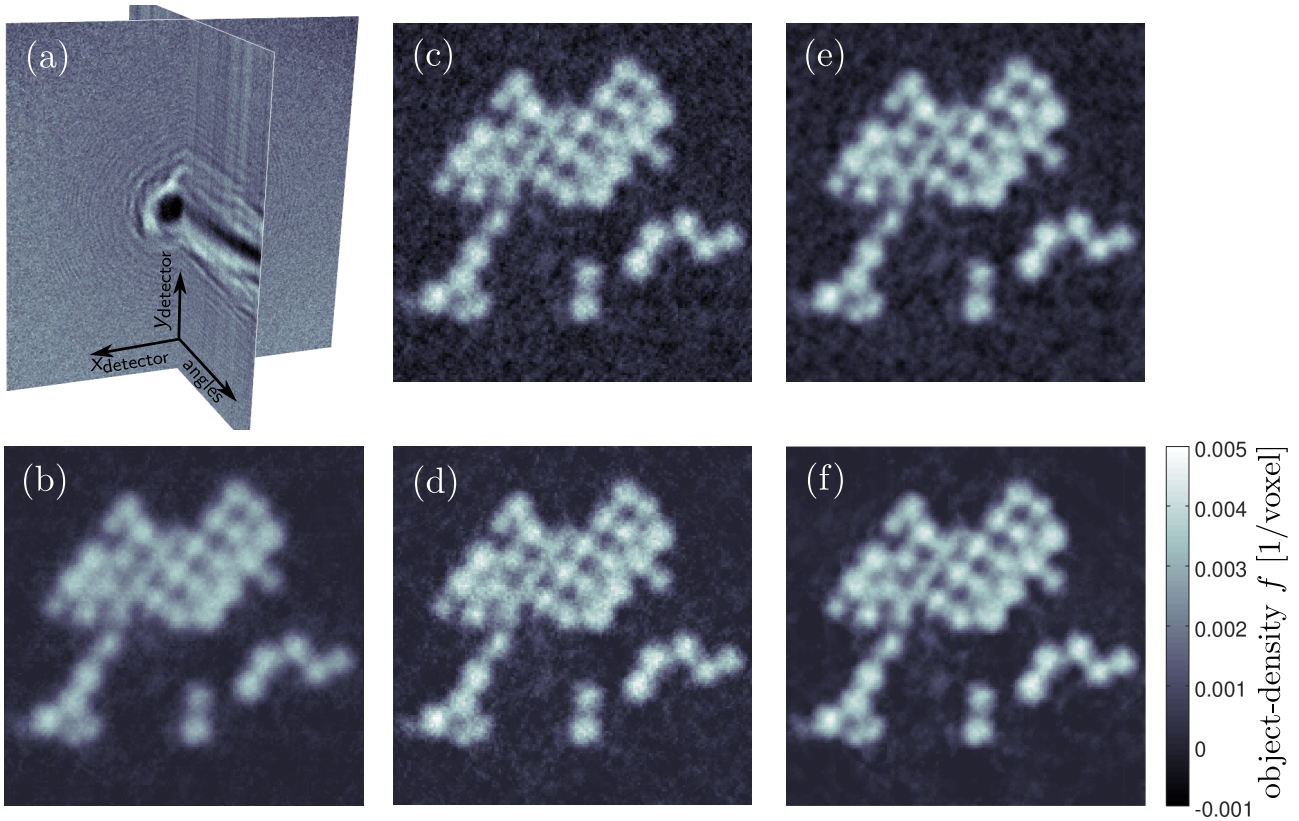


Figure 11.3. X-ray phase contrast tomography test case: (a) Orthoslice-plot of the 3D tomographic data $\mathbf{g}^{\text{obs}} = (\mathbf{g}_1^{\text{obs}}, \dots, \mathbf{g}_{N_{\text{proj}}}^{\text{obs}})$, composed of 1024×1024 -sized diffraction patterns for 249 incident angles. (b) Reconstruction from [143]: central slice of the 256^3 -voxel volume perpendicular to the tomographic axis (zoomed to the object-containing region). (c)–(f) Same slice of the reconstructed object after one cycle of algorithm 11.3: (c) $\gamma = 0$, no non-negativity constraint. (d) $\gamma = 0$, with non-negativity. (e) $\gamma = 0.8$, no non-negativity. (f) $\gamma = 0.8$, with non-negativity. The color scale in (b)–(f) is identical.

By nature of the imaged sample, the slices should show circular cross-sections of the colloids of constant density within a zero background, i.e. a binary image. It can be seen that all methods perform widely similarly in terms of contrast and resolution where the four reconstructions with algorithm 11.3 (fig. 11.3(c)–(f)) seem to yield somewhat less blurry colloids with a higher contrast to the background than the reference-result in fig. 11.3(b). Moreover, comparing fig. 11.3(c),(d) to (e),(f) respectively, reveals the anticipated noise-suppressing effect of the gradient-penalty compared to pure L^2 -regularization, at essentially the same sharpness and contrast. Indeed, the reconstruction in fig. 11.3(f) using both a gradient-penalty and non-negativity constraint seems closest to the expected ideal binary object and is arguably best suited for further image-segmentation and -analysis – despite the slight approximation involved in the computation via GenSART-schemes (compare §11.4.3).

11.7 Conclusions

In this work, efficient solution formulas have been proposed for the computation of regularized Kaczmarz-iterations (also known as “Tikhonov-Kaczmarz” or “incremental proximal iterations”) for tomographic reconstruction. By their structural analogy and similar computational efficiency to classical SART-iterations [6], the derived schemes are termed generalized SART (GenSART). Notably, the approach strongly exploits mathematical structures specific to tomographic inverse problems, namely the well-known fact that the acquired data for a single tomographic view only contains information on the *projection* of the three-dimensional object to a 2D-manifold. As a consequence, Kaczmarz-updates may often be computed in the *lower-dimensional* projection-space, reducing the computational costs.

As demonstrated in the manuscript, this enables efficient Kaczmarz-methods for various non-standard tomographic settings such as robust reconstruction in the presence of large outliers (§11.5.1.2, §11.6.2), photon-starvation- (§11.5.1.3) and beam-hardening-resistant tomography (§11.5.2.2) or phase contrast tomography (§11.5.2.1, §11.6.3). The key feature of the GenSART-approach here lies in its flexibility in the choice of the data-fidelity and image-formation model (including even *non-convex* choices) – essentially without effect on the computational costs! Note that this constitutes a crucial difference to (bulk-) variational reconstruction methods, where the number of required iterations typically grows dramatically as soon as non-quadratic functionals are to be minimized. The proposed methods are furthermore applicable to both parallel-beam- and cone-beam-tomography, absolutely regardless of the sampling in the tomographic incident directions and naturally incorporate support constraints on the unknown object, thus retaining the geometrical flexibility of SART.

In addition to (weighted) L^2 -penalties and Sobolev- $W^{1,2}$ -regularizers that can be directly incorporated into GenSART-schemes (yet, the latter with the flaw of a slight approximation, see §11.4.3), the approach may be combined with interlacing proximal- or gradient-descent iterations to enable advanced regularizers, as outlined in §11.5.3. Likewise, GenSART-formulas might serve to efficiently evaluate proximal maps in Kaczmarz-type primal-dual methods of the kind proposed in [35]. Such applications constitute promising prospects for future work.

Importantly, this work is *not* concerned with convergence of Kaczmarz-methods beyond known results (e.g. [115]) but focusses on an efficient computation of the iterates. Yet, it clearly motivates a further analysis of Kaczmarz-convergence with non-quadratic data-fidelities.

Another important goal is *parallelization*: while GenSART-iterations are efficient in terms of computational complexity, Kaczmarz-methods are more *sequential* in nature than bulk-variational methods (though, by far not as sequential as ART). This complicates accelerating the approach by parallel computing, which for bulk methods can be readily achieved by using massively parallelized implementations of tomographic (back-)projections from the ASTRA-toolbox [199, 198] for example. If the issue of parallelization is resolved, the results of this work show that Kaczmarz-methods computed via GenSART-schemes may provide a highly efficient alternative to variational reconstruction methods – at almost the same flexibility.

Appendix

11.A Geometry of the Projectors

Proof of Theorem 11.4: We show that $B_{\text{iso}} : L^2(\mathbb{D}_P) \rightarrow L^2(\Omega)$; $p \mapsto w_P \cdot \tilde{P}^{\text{B}}(\tilde{u}_P^{-1/2} \cdot p)$ is well-defined and isometric. Let $p \in L^2(\mathbb{D}_P)$ be arbitrary. For $P = \mathcal{P}$, we have

$$\begin{aligned} \|B_{\text{iso}}(p)\|_{L^2}^2 &= \int_{\Omega} \left| P^{\text{B}}(\tilde{u}_P^{-1/2} \cdot p)(\mathbf{x}) \right|^2 d\mathbf{x} \stackrel{(11.3.4)}{=} \int_{\mathbb{D}_{\mathcal{P}}} \left| (\tilde{u}_P^{-1/2} \cdot p)(\mathbf{x}_{\perp}) \right|^2 \cdot \underbrace{\left(\int_{L_{\mathbf{x}_{\perp}}^{\mathcal{P}}} 1 dz \right)}_{=P(\mathbf{1}_{\Omega})(\mathbf{x}_{\perp})=u_P(\mathbf{x}_{\perp})} d\mathbf{x}_{\perp} \\ &= \int_{\Omega_P} (u_P(\mathbf{x}_{\perp})/\tilde{u}_P(\mathbf{x}_{\perp})) \cdot |p(\mathbf{x}_{\perp})|^2 d\mathbf{x}_{\perp} \stackrel{\tilde{u}_P \equiv u_P}{=} \|p\|_{L^2(\mathbb{D}_{\mathcal{P}})}^2. \end{aligned} \quad (11.A.1)$$

By introducing polar coordinates, we obtain the same for the cone-beam case $P = \mathcal{D}$:

$$\begin{aligned} \|B_{\text{iso}}(p)\|_{L^2}^2 &= \int_{\Omega} \left| w_{\mathcal{D}} \cdot \mathcal{D}^{\text{B}}(\tilde{u}_{\mathcal{D}}^{-1/2} \cdot p)(\mathbf{x}) \right|^2 d\mathbf{x} = \int_{\mathbb{D}_{\mathcal{D}}} |(\tilde{u}_{\mathcal{D}}^{-1/2} \cdot p)(\varphi)|^2 \int_{L_{\varphi}^{\mathcal{D}}} |w_{\mathcal{D}}(t\varphi)|^2 t^2 dt d\varphi \\ &\stackrel{w_{\mathcal{D}}(t\varphi) \equiv t^{-2}}{=} \int_{\Omega_{\mathcal{D}}} \tilde{u}_{\mathcal{D}}(\varphi)^{-1} \cdot |p(\varphi)|^2 \cdot \underbrace{\left(\int_{L_{\varphi}^{\mathcal{D}}} t^{-2} dt \right)}_{\mathcal{D}(w_{\mathcal{D}})(\varphi) = \tilde{u}_{\mathcal{D}}(\varphi)} d\varphi = \|p\|_{L^2(\mathbb{D}_{\mathcal{D}})}^2. \end{aligned} \quad (11.A.2)$$

Accordingly, $B_{\text{iso}} : L^2(\mathbb{D}_P) \rightarrow L^2(\Omega)$ is well-defined and isometric for $P \in \{\mathcal{P}, \mathcal{D}\}$. Since $B_{\text{iso}}(p) = P^*(\tilde{u}_P^{-1/2} \cdot p)$, its adjoint is given by $B_{\text{iso}}^*(f) = \tilde{u}_P^{-1/2} \cdot P(f) = P_{\text{iso}}(f)$. Hence, $P_{\text{iso}} = B_{\text{iso}}^* : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$ is bounded with norm $\|P_{\text{iso}}\| = \|B_{\text{iso}}\| = 1$. By the isometry-property, $P_{\text{iso}}^* = B_{\text{iso}}$ has closed range. According to the closed range theorem, the same thus holds true for P_{iso} . As P_{iso}^* is isometric, $P_{\text{iso}}^* : L^2(\mathbb{D}_P) \rightarrow \text{range}(P_{\text{iso}}^*)$ defines a unitary operator, which implies the characterizations of $P_{\text{iso}}^* P_{\text{iso}}$ and $P_{\text{iso}} P_{\text{iso}}^* = \text{id}_{L^2(\mathbb{D}_P)}$. \square

11.B Projectors and Gradients

Proof of Lemma 11.13: Let $f \in \mathcal{C}_c^{\infty}(\Omega)$ be smooth and compactly supported. Then f can be identified with a function in $\mathcal{C}_c^{\infty}(\mathbb{R}^3)$ by simply extending it with 0 outside Ω (notably, this would *not* be true if we only assumed $f \in \mathcal{C}^{\infty}(\Omega)$).

We decompose the gradient into mutually orthogonal components along- and perpendicular to the rays: in the parallel-beam case $P = \mathcal{P}$, the ray-direction is constant and points along the z -axis (without loss of generality). Accordingly, ∇_P is simply the gradient with respect to the remaining Cartesian coordinates \mathbf{x}_{\perp} , i.e. $\nabla_P = \nabla_{\perp}$. Since Ω is bounded, there exist constants $a_P < b_P$ such that Ω is contained in the stripe $\mathbb{R}^2 \times (a_P; b_P) \supset \Omega$. Accordingly, we obtain by Leibniz' integral rule

$$\begin{aligned} P(w_P^{-1/2} \cdot \nabla_P f)(\mathbf{x}_{\perp}) &= P(\nabla_P f)(\mathbf{x}_{\perp}) = \int_{a_P}^{b_P} \nabla_{\perp} f(\mathbf{x}_{\perp}, z) dz \\ &= \nabla_{\mathbb{R}^2} \int_{a_P}^{b_P} f(\mathbf{x}_{\perp}, z) dz = \nabla_{\mathbb{R}^2} P(f)(\mathbf{x}_{\perp}) = \nabla_{\mathbb{D}} P(f)(\mathbf{x}_{\perp}) \end{aligned} \quad (11.B.1)$$

for all $\mathbf{x}_\perp \in \mathbb{R}^2$. Here, it was used that the ray-density is constant in the considered parallel-beam case, $w_P = 1$ for $P = \mathcal{P}$. For the cone-beam case $P = \mathcal{D}$, the gradient can be expressed in polar coordinates: if $f^{(p)}$ is defined by $f^{(p)}(\boldsymbol{\varphi}, t) := f(t\boldsymbol{\varphi})$, then

$$\begin{aligned} \nabla f(t\boldsymbol{\varphi}) &= t^{-1} \nabla_{\boldsymbol{\varphi}} f^{(p)}(\boldsymbol{\varphi}, t) + \boldsymbol{\varphi} \partial_t f^{(p)}(\boldsymbol{\varphi}, t) \\ &= w_{\mathcal{D}}(t\boldsymbol{\varphi})^{1/2} \nabla_{\boldsymbol{\varphi}} f^{(p)}(\boldsymbol{\varphi}, t) + \boldsymbol{\varphi} \partial_t f^{(p)}(\boldsymbol{\varphi}, t) \quad \text{for all } t > 0, \boldsymbol{\varphi} \in \mathbb{S}^2. \end{aligned} \quad (11.B.2)$$

where the ray-density $w_{\mathcal{D}}(\mathbf{x}) = |\mathbf{x}|^{-2}$ has been inserted. The gradient-component $\nabla_P f$ perpendicular to the ray-direction $\boldsymbol{\varphi}$ is given by the first term on the right-hand side. Since Ω is bounded and $0 \notin \bar{\Omega}$, the domain is contained in some annulus, $\Omega \subset \{t\boldsymbol{\varphi} : \boldsymbol{\varphi} \in \mathbb{S}^2, a_{\mathcal{D}} < t < b_{\mathcal{D}}\}$ for $0 < a_{\mathcal{D}} < b_{\mathcal{D}}$. By another application of Leibniz' rule, this implies

$$\begin{aligned} P(w_P^{-1/2} \cdot \nabla_P f)(\boldsymbol{\varphi}) &= \int_{a_{\mathcal{D}}}^{b_{\mathcal{D}}} (w_P^{-1/2} \cdot \nabla_P f)(t\boldsymbol{\varphi}) dt = \int_{a_{\mathcal{D}}}^{b_{\mathcal{D}}} \nabla_{\boldsymbol{\varphi}} f^{(p)}(\boldsymbol{\varphi}, t) dt \\ &= \nabla_{\mathbb{S}^2} \int_{a_{\mathcal{D}}}^{b_{\mathcal{D}}} f^{(p)}(\boldsymbol{\varphi}, t) dt = \nabla_{\mathbb{S}^2} \mathcal{D}(f)(\boldsymbol{\varphi}) = \nabla_{\mathbb{D}} P(f)(\boldsymbol{\varphi}) \end{aligned} \quad (11.B.3)$$

for all $\boldsymbol{\varphi} \in \mathbb{S}^2$. Equations (11.B.1) and (11.B.3) show that (11.4.15) holds for all $f \in \mathcal{C}_c^\infty(\Omega)$ and $P \in \{\mathcal{P}, \mathcal{D}\}$. Since $P_{\text{iso}} : L^2(\Omega) \rightarrow L^2(\mathbb{D}_P)$; $p \mapsto \tilde{u}_P^{-1/2} \cdot P(p)$ is bounded with norm 1 according to Theorem 11.4, this furthermore implies that

$$\begin{aligned} \|\tilde{u}_P^{-1/2} \cdot \nabla_{\mathbb{D}} P(f)\|_{L^2} &= \|\tilde{u}_P^{-1/2} \cdot P(w_P^{-1/2} \cdot \nabla_P f)\|_{L^2} \leq \|w_P^{-1/2} \cdot \nabla_P f\|_{L^2} \\ &\leq c^{-1/2} \|\nabla_P f\|_{L^2} \leq c^{-1/2} \|\nabla f\|_{L^2} \leq c^{-1/2} \|f\|_{W^{1,2}} \end{aligned} \quad (11.B.4)$$

for all $f \in \mathcal{C}_c^\infty(\Omega)$ and $c := \min_{\mathbf{x} \in \Omega} w_P(\mathbf{x}) > 0$. As $\mathcal{C}_c^\infty(\Omega)$ is dense in $W_0^{1,2}(\Omega)$ with respect to the $W^{1,2}$ -norm, (11.B.4) shows that $p \mapsto \tilde{u}_P^{-1/2} \cdot \nabla_{\mathbb{D}} P(p)$ defines a bounded linear map $W_0^{1,2}(\Omega) \rightarrow L^2(\mathbb{D}_P)$. Hence, (11.4.15) and (11.4.16) remain valid for all $f \in W_0^{1,2}(\Omega)$.

Now let $P^{\text{B}}(p) \in W^{1,2}(\Omega)$. According to (11.3.4), $P^{\text{B}}(p)$ is constant within Ω along the ray-direction. Hence, the corresponding gradient-components vanish in an L^2 -sense so that $\nabla(P^{\text{B}}(p)) = \nabla_P(P^{\text{B}}(p))$. By inserting the expressions for ∇_P in the parallel- and cone-beam geometry, this yields for almost all $(\mathbf{x}_\perp, z), t\boldsymbol{\varphi} \in \Omega$

$$\nabla(\mathcal{P}^{\text{B}}(p))(\mathbf{x}_\perp, z) = \nabla_P(\mathcal{P}^{\text{B}}(p))(\mathbf{x}_\perp, z) \stackrel{(11.3.4)}{=} \nabla_{\mathbb{R}^2}(p)(\mathbf{x}_\perp) \quad (11.B.5)$$

$$\nabla(\mathcal{D}^{\text{B}}(p))(t\boldsymbol{\varphi}) = \nabla_P(\mathcal{D}^{\text{B}}(p))(t\boldsymbol{\varphi}) \stackrel{(11.3.4)}{=} t^{-1} \nabla_{\mathbb{S}^2}(p)(\boldsymbol{\varphi}). \quad (11.B.6)$$

By decomposing the norm-integral into integrals perpendicular and along the ray-direction, respectively, we may now compute the L^2 -norm of $\nabla(P^{\text{B}}(p))$:

$$\begin{aligned} \|\nabla(P^{\text{B}}(p))\|_{L^2}^2 &= \begin{cases} \int_{\mathbb{D}_{\mathcal{P}}} \int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} |\nabla_{\mathbb{R}^2}(p)(\mathbf{x}_\perp)|^2 dz d\mathbf{x}_\perp & \text{for } P = \mathcal{P} \\ \int_{\mathbb{D}_{\mathcal{D}}} \int_{L_{\boldsymbol{\varphi}}^{\mathcal{D}}} \left| \frac{1}{t} \nabla_{\mathbb{S}^2}(p)(\boldsymbol{\varphi}) \right|^2 t^2 dt d\boldsymbol{\varphi} & \text{for } P = \mathcal{D} \end{cases} \\ &= \begin{cases} \int_{\mathbb{D}_{\mathcal{P}}} |\nabla_{\mathbb{R}^2}(p)(\mathbf{x}_\perp)|^2 \left(\int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} dz \right) d\mathbf{x}_\perp & \text{for } P = \mathcal{P} \\ \int_{\mathbb{D}_{\mathcal{D}}} |\nabla_{\mathbb{S}^2}(p)(\boldsymbol{\varphi})|^2 \left(\int_{L_{\boldsymbol{\varphi}}^{\mathcal{D}}} dt \right) d\boldsymbol{\varphi} & \text{for } P = \mathcal{D} \end{cases} \\ &= \|\tilde{u}_P^{1/2} \cdot \nabla_{\mathbb{D}}(p)\|_{L^2}^2, \end{aligned} \quad (11.B.7)$$

which proves (11.4.17). The equality (11.B.7) furthermore shows the equivalences

$$P^B(p) \in W^{1,2}(\Omega) \Leftrightarrow \nabla(P^B(p)) \in L^2(\Omega) \Leftrightarrow u_P^{1/2} \cdot \nabla_{\mathbb{D}}(p) \in L^2(\mathbb{D}_P). \quad (11.B.8)$$

By continuity of $\nabla : W^{1,2}(\Omega) \rightarrow L^2(\Omega)$ and $\nabla_{\mathbb{D}}P$, it is sufficient to show (11.4.18) for $f \in \mathcal{C}_c^\infty(\Omega)$. Using the expressions for $\nabla(P^B(p))$ derived above, we obtain

$$\begin{aligned} \langle \nabla(P^B(p)), \nabla f \rangle_{L^2} &= \langle \nabla_P(P^B(p)), \nabla f \rangle_{L^2} = \langle \nabla_P(P^B(p)), \nabla_P f \rangle_{L^2} \\ &= \begin{cases} \int_{\mathbb{D}_{\mathcal{P}}} \nabla_{\mathbb{R}^2}(p)(\mathbf{x}_\perp) \cdot \int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} \nabla_\perp f(\mathbf{x}_\perp, z) \, dz \, d\mathbf{x}_\perp & \text{for } P = \mathcal{P} \\ \int_{\mathbb{D}_{\mathcal{D}}} \nabla_{\mathbb{S}^2}(p)(\varphi) \cdot \int_{L_{\varphi}^{\mathcal{D}}} \frac{1}{t^2} \nabla_\varphi f^{(p)}(\varphi, t) t^2 \, dt \, d\varphi & \text{for } P = \mathcal{D} \end{cases} \\ &\stackrel{(11.B.1), (11.B.3)}{=} \langle \nabla_{\mathbb{D}}(p), P(w_P^{-1/2} \cdot \nabla_P(f)) \rangle_{L^2} \\ &\stackrel{(11.4.15)}{=} \langle u_P^{1/2} \cdot \nabla_{\mathbb{D}}(p), u_P^{-1/2} \cdot \nabla_{\mathbb{D}}P(f) \rangle_{L^2} \quad \text{for all } f \in \mathcal{C}_c^\infty(\Omega). \end{aligned} \quad (11.B.9)$$

□

11.C Admissibility of L^q -Penalties

Proof of Lemma 11.16: Let $p \in L^2(\mathbb{D}_P)$ and $f_0 \in \ker(\mathcal{P}_{\text{iso}})$ be arbitrary. If $\mathcal{R}(f_{\text{ref}} + \mathcal{P}_{\text{iso}}^*(p) + f_0) = \|\mathcal{P}_{\text{iso}}^*(p) + f_0\|_{L^q}^q = \infty$, then (A) trivially holds true. Hence, we may assume that $f := \mathcal{P}_{\text{iso}}^*(p) + f_0 \in L^q(\Omega)$. With the rays $L_{\mathbf{x}_\perp}^{\mathcal{P}}$ defined as in (11.3.2), we then have

$$\mathcal{R}(f_{\text{ref}} + \mathcal{P}_{\text{iso}}^*(p) + f_0) = \int_{\Omega} |f|^q \, dx = \int_{\mathbb{R}^2} \left(\int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} |f(\mathbf{x}_\perp, z)|^q \, dz \right) d\mathbf{x}_\perp \quad (11.C.1)$$

where the inner integrals are finite for almost all $\mathbf{x}_\perp \in \mathbb{R}^2$. An application of Jensen's inequality to the convex function $\mathbb{R} \rightarrow \mathbb{R}; x \mapsto |x|^q$ in these shows that, for almost all $\mathbf{x}_\perp \in \mathbb{R}^2$,

$$\int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} |f(\mathbf{x}_\perp, z)|^q \, dz \geq \underbrace{\left(\int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} 1 \, dz \right)^{1-q}}_{=u_{\mathcal{P}}(\mathbf{x}_\perp)^{1-q}} \left| \int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} f(\mathbf{x}_\perp, z) \, dz \right|^q. \quad (11.C.2)$$

Comparing to (11.2.2), we find that the integral within the modulus exactly corresponds to an evaluation of $\mathcal{P} = u_{\mathcal{P}}^{1/2} \cdot \mathcal{P}_{\text{iso}}$. Since $f = \mathcal{P}_{\text{iso}}^*(p) + f_0$ with $f_0 \in \ker(\mathcal{P}_{\text{iso}})$, this implies

$$\begin{aligned} \int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} |f(\mathbf{x}_\perp, z)|^q \, dz &\geq u_{\mathcal{P}}(\mathbf{x}_\perp)^{1-q} |\mathcal{P}(f)(\mathbf{x}_\perp)|^q = u_{\mathcal{P}}(\mathbf{x}_\perp)^{1-q} |\mathcal{P} \mathcal{P}_{\text{iso}}^*(p)(\mathbf{x}_\perp)|^q \\ &= \left(\int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} 1 \, dz \right)^{1-q} \left| \int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} \underbrace{\mathcal{P}_{\text{iso}}^*(p)(\mathbf{x}_\perp, z)}_{\text{constant in } z} \, dz \right|^q = \int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} |\mathcal{P}_{\text{iso}}^*(p)(\mathbf{x}_\perp, z)|^q \, dz \end{aligned} \quad (11.C.3)$$

for almost all $\mathbf{x}_\perp \in \mathbb{R}^2$. Substituting the estimate (11.C.3) into (11.C.1) proves (A):

$$\mathcal{R}(f_{\text{ref}} + \mathcal{P}_{\text{iso}}^*(p) + f_0) \geq \int_{\mathbb{R}^2} \left(\int_{L_{\mathbf{x}_\perp}^{\mathcal{P}}} |\mathcal{P}_{\text{iso}}^*(p)(\mathbf{x}_\perp, z)|^q \, dz \right) d\mathbf{x}_\perp = \mathcal{R}(f_{\text{ref}} + \mathcal{P}_{\text{iso}}^*(p)). \quad (11.C.4)$$

In particular, (11.C.4) shows that $\|\mathcal{P}_{\text{iso}}^*(p)\|_{L^q}^q = \mathcal{R}(f_{\text{ref}} + \mathcal{P}_{\text{iso}}^*(p)) < \infty$, i.e. $\mathcal{P}_{\text{iso}}^*(p) \in L^q(\Omega)$. Since $\mathcal{P}\mathcal{P}_{\text{iso}}^*(p) = u_{\mathcal{P}}^{1/2} \cdot \mathcal{P}_{\text{iso}}\mathcal{P}_{\text{iso}}^*(p) = u_{\mathcal{P}}^{1/2} \cdot p$ by Theorem 11.4, (11.C.3) furthermore yields

$$\begin{aligned} \mathcal{R}(f_{\text{ref}} + \mathcal{P}_{\text{iso}}^*(p)) &= \int_{\mathbb{R}^2} \int_{L_{\mathbf{x}_{\perp}}^{\mathcal{P}}} |\mathcal{P}_{\text{iso}}^*(p)(\mathbf{x}_{\perp}, z)|^q dz \\ &= \int_{\mathbb{R}^2} u_{\mathcal{P}}(\mathbf{x}_{\perp})^{1-q/2} |p(\mathbf{x}_{\perp})|^q d\mathbf{x}_{\perp} = \|u_{\mathcal{P}}^{1/q-1/2} \cdot p\|_{L^q}^q. \end{aligned} \quad (11.C.5)$$

Now let $q > 1$ and $f_0 \neq 0$. Then $x \mapsto |x|^q$ is strictly convex so that strict inequality holds in (11.C.2) whenever the integrand $z \mapsto |f(\mathbf{x}_{\perp}, z)|$ with $f = \mathcal{P}_{\text{iso}}^*(p) + f_0$ is non-constant. The latter is clearly the case for some $\mathbf{x}_{\perp} \in \mathbb{R}^2$, as elements in $\text{kern}(\mathcal{P}_{\text{iso}})$ are, on each ray $L_{\mathbf{x}_{\perp}}^{\mathcal{P}}$, either non-constant or identically zero. Hence, strict inequality holds in (11.C.4) if $f_0 \neq 0$. \square

11.D Poisson-noise-adapted data fidelity

In the following, it is shown that the log-likelihood for Poisson-noisy data given in (11.5.11) can be approximated by the functional in (11.5.12) if variations of the true data g_j within the supports of the ω_i are negligible. Specifically, we assume that g_j is “constant enough” in $\text{supp}(\omega_i)$ such that it may be pulled in and out of the integrals, i.e. for some $x_{ji} \in \mathbb{D}$

$$\begin{aligned} \ln(t\mathcal{M}_i(g_j)) &= \ln\left(\int_{\mathbb{D}} t\omega_i g_j dx\right) \approx \ln\left(g_j(x_{ji}) \int_{\mathbb{D}} t\omega_i dx\right) = \ln(g_j(x_{ji})) + \tilde{c}_{ij} \\ &\approx \frac{\int_{\mathbb{D}} \ln(g_j) t\omega_i dx}{\int_{\mathbb{D}} t\omega_i dx} + \tilde{c}_{ij} \end{aligned} \quad (11.D.1)$$

where the \tilde{c}_{ij} are independent of g_j . Inserting this approximation into (11.5.11) yields

$$\begin{aligned} \mathcal{S}^{\text{Poi}}(g_j^{\text{obs}}; g_j) &\approx \sum_{i=1}^{m_{\text{proj}}} \int_{\mathbb{D}} tg_j \omega_i dx - g_{ji}^{\text{obs}} \cdot \left(\frac{\int_{\mathbb{D}} \ln(g_j) t\omega_i dx}{\int_{\mathbb{D}} t\omega_i dx} + \tilde{c}_{ij}\right) - g_{ji}^{\text{obs}} \ln(g_{ji}^{\text{obs}}) \\ &= \int_{\mathbb{D}} \left(\left(\sum_{i=1}^{m_{\text{proj}}} \omega_i\right) tg_j - \left(\sum_{i=1}^{m_{\text{proj}}} \frac{g_{ji}^{\text{obs}} \omega_i}{\int_{\mathbb{D}} \omega_i dx}\right) \cdot \ln(tg_j)\right) dx + \tilde{c} \\ &= \int_{\mathbb{D}} \text{KL}(g_{j,\text{cont}}^{\text{obs}}(x); tg_j(x)) \cdot \omega(x) dx + c \end{aligned} \quad (11.D.2)$$

with $\omega(x) := \sum_{i=1}^{m_{\text{proj}}} \omega_i(x)$, $g_{j,\text{cont}}^{\text{obs}}(x) := (\sum_{i=1}^{m_{\text{proj}}} g_{ji}^{\text{obs}} \omega_i(x) / \int_{\mathbb{D}} \omega_i dx) / \omega(x)$ and some c that is independent of g_j .

Appendix A

Supplementary proofs for numerical stability computations

The following appendix contains the proofs of the theorems in §3.4, that have been used to derive a numerical method for stability computations in linearized XPCI. For convenience, the proven theorems are also restated here.

A.1 Characterization of contrast-minimizing modes

The first section contains the left-out proofs from §3.4.1, concerning existence and properties of image-modes $\varphi_0 \in L^2(\Omega)$ that attain minimal contrast $\|\mathcal{S}_\nu(\varphi_0)\|/\|\varphi_0\|$.

Lemma (3.11). *Let $\Omega \subset \mathbb{R}^m$ be bounded, $\mathfrak{f} > 0$ and $\alpha \in [0; \pi)$. Then it holds that*

$$\|\mathcal{S}_\nu(\varphi)\|^2 = \langle \varphi, 2\varphi - 2\mathcal{K}_{\mathfrak{f},\alpha}(\varphi) \rangle \quad \text{for all } \varphi \in L^2(\Omega), \quad (\text{A.1.1})$$

where $\mathcal{K}_{\mathfrak{f},\alpha} : L^2(\Omega) \rightarrow L^2(\Omega)$ is a selfadjoint and compact linear operator defined by

$$\mathcal{K}_{\mathfrak{f},\alpha}(\varphi) = \mathcal{F}^{-1} \left(\hat{k}_{\mathfrak{f},\alpha} \cdot \mathcal{F}(\varphi) \right) |_\Omega = (k_{\mathfrak{f},\alpha} * \varphi) |_\Omega = \int_\Omega k_{\mathfrak{f},\alpha}(\cdot - \mathbf{y}) \varphi(\mathbf{y}) \, d\mathbf{y} \quad (\text{A.1.2})$$

$$\hat{k}_{\mathfrak{f},\alpha}(\boldsymbol{\xi}) = \cos \left(\frac{\boldsymbol{\xi}^2}{\mathfrak{f}} + 2\alpha \right), \quad k_{\mathfrak{f},\alpha}(\mathbf{x}) = (\mathfrak{f}/2)^{m/2} \cos \left(\frac{\mathfrak{f}\mathbf{x}^2}{4} - (2\alpha + m\pi/4) \right) \quad (\text{A.1.3})$$

Proof. Let $S_\Omega : L^2(\Omega) \rightarrow L^2(\mathbb{R}^m)$; $\varphi \mapsto \mathcal{S}_\nu(\varphi)$ denote the CTF-operator restricted to functions in $L^2(\Omega)$. By definition of the adjoint, it then holds for all $\varphi \in L^2(\Omega)$ that

$$\|\mathcal{S}_\nu(\varphi)\|^2 = \langle S_\Omega(\varphi), S_\Omega(\varphi) \rangle = \langle \varphi, S_\Omega^* S_\Omega(\varphi) \rangle \quad (\langle \cdot, \cdot \rangle: L^2 \text{ inner product}) \quad (\text{A.1.4})$$

S_Ω^* is the adjoint of \mathcal{S}_ν , composed with a restriction $f \mapsto f|_\Omega$. As $\mathcal{S}_\nu : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m)$; $\varphi \mapsto \mathcal{F}^{-1}(s_\alpha \cdot \mathcal{F}(\varphi))$ is a self-adjoint Fourier multiplier, it thus follows that

$$S_\Omega^* S_\Omega(\varphi) = (\mathcal{S}_\nu^2(\varphi)) |_\Omega = (\mathcal{F}^{-1}(s_\alpha^2 \cdot \mathcal{F}(\varphi))) |_\Omega \quad \text{for all } \varphi \in L^2(\Omega) \quad (\text{A.1.5})$$

By trigonometric identities, the squared CTF-factor s_α^2 can be cast to the form

$$\begin{aligned} s_\alpha^2(\boldsymbol{\xi}) &= 4 \sin^2 \left(\frac{\boldsymbol{\xi}^2}{2\mathfrak{f}} + \alpha \right) = 4 - 4 \cos^2 \left(\frac{\boldsymbol{\xi}^2}{2\mathfrak{f}} + \alpha \right) = 2 - 2 \cos \left(\frac{\boldsymbol{\xi}^2}{\mathfrak{f}} + 2\alpha \right) \\ &\stackrel{\text{(A.1.3)}}{=} 2 - 2\hat{k}_{\mathfrak{f},\alpha}(\boldsymbol{\xi}) \quad \text{for all } \boldsymbol{\xi} \in \mathbb{R}^m. \end{aligned} \quad (\text{A.1.6})$$

Substituting the r.h.s. of (A.1.6) into (A.1.5) yields $S_\Omega^* S_\Omega(\varphi) = (\mathcal{F}^{-1}((2 - 2\hat{k}_{\mathfrak{f},\alpha}) \cdot \mathcal{F}(\varphi)))|_\Omega = 2\varphi - 2\mathcal{K}_{\mathfrak{f},\alpha}(\varphi)$ and thus (A.1.1) according to (A.1.4)

As $\mathcal{K}_{\mathfrak{f},\alpha} : L^2(\Omega) \rightarrow L^2(\Omega)$ is essentially a Fourier-multiplier with a bounded real-valued factor $\hat{k}_{\mathfrak{f},\alpha}$, it is linear-bounded and selfadjoint. By the convolution theorem, the Fourier-multiplication may be rewritten as a convolution. Again by trigonometric identities, it follows that

$$\hat{k}_{\mathfrak{f},\alpha}(\boldsymbol{\xi}) = \frac{1}{2} e^{2i\alpha} \exp \left(\frac{i\boldsymbol{\xi}^2}{\mathfrak{f}} \right) + \frac{1}{2} e^{-2i\alpha} \exp \left(\frac{-i\boldsymbol{\xi}^2}{\mathfrak{f}} \right) = \frac{1}{2} e^{2i\alpha} m_{\mathfrak{f}/2}(\boldsymbol{\xi})^{-1} + \frac{1}{2} e^{-2i\alpha} m_{\mathfrak{f}/2}(\boldsymbol{\xi}). \quad (\text{A.1.7})$$

The identified factor $m_{\mathfrak{f}/2}$ is the Fourier-multiplier of \mathcal{D}^2 . Accordingly, (A.1.7) implies that

$$2\mathcal{K}_\alpha(\varphi) = e^{2i\alpha} \mathcal{D}_{\mathfrak{f}/2}^{-2}(\varphi)|_\Omega + e^{2i\alpha} \mathcal{D}^2(\varphi)|_\Omega = ((e^{2i\alpha} \overline{k_{\mathfrak{f}/2}} + e^{-2i\alpha} k_{\mathfrak{f}/2}) * \varphi)|_\Omega = (k_{\mathfrak{f},\alpha} * \varphi)|_\Omega, \quad (\text{A.1.8})$$

where the convolution-form of the Fresnel propagator (2.2.9) has been inserted to obtain the second equality. By employing the known representation of $k_{\mathfrak{f}/2}$, it follows that the convolution kernel $k_{\mathfrak{f},\alpha}$ on the r.h.s. of (A.1.8) is of the claimed form (A.1.3).

Finally, the convolution-form of $\mathcal{K}_{\mathfrak{f},\alpha}$ can be written as an integral operator with kernel $k(\mathbf{x}, \mathbf{y}) := k_{\mathfrak{f},\alpha}(\mathbf{x} - \mathbf{y})$, see right-most equality in (A.1.2). As Ω and $k_{\mathfrak{f},\alpha}$ are bounded, the kernel is square-integrable, $k \in L^2(\Omega \times \Omega)$, which implies that $\mathcal{K}_{\mathfrak{f},\alpha}$ is a compact operator. \square

Theorem (3.12). *Let $\Omega \subset \mathbb{R}^m$ be bounded, $\mathfrak{f} > 0$ and $\alpha \in \mathbb{R}$ and let $\mathcal{K}_{\mathfrak{f},\alpha}$ be the operator from lemma 3.11. Then either of the following cases hold true:*

- (1) $\mathcal{K}_{\mathfrak{f},\alpha}$ has a positive eigenvalue. Then there exist contrast-minimizing modes $\varphi_0 \in L^2(\Omega, \mathbb{R})$ and these are exactly the eigenfunctions of $\mathcal{K}_{\mathfrak{f},\alpha}$ to its maximal eigenvalue $\lambda_{\mathfrak{f},\alpha} > 0$, i.e. $\mathcal{K}_{\mathfrak{f},\alpha}(\varphi_0) = \lambda_{\mathfrak{f},\alpha} \varphi_0$, and

$$C_{\text{IP2}}(\Omega, \mathfrak{f}, \alpha) = (2 - 2\lambda_{\mathfrak{f},\alpha})^{1/2}. \quad (\text{A.1.9})$$

Moreover, there are finitely many linearly independent contrast-minimizing modes and any such mode is infinitely smooth in the interior of Ω .

- (2) $\mathcal{K}_{\mathfrak{f},\alpha}$ does not have positive eigenvalues, Then $C_{\text{IP2}}(\Omega, \mathfrak{f}, \alpha) = 2^{1/2}$.

Proof. Let $\varphi_0 \in L^2(\Omega, \mathbb{R})$ such that $\|\varphi_0\| = 1$. Then, by lemma 3.11, it holds that

$$\|\mathcal{S}_\nu(\varphi_0)\|^2 = \langle \varphi_0, 2\varphi_0 - 2\mathcal{K}_{\mathfrak{f},\alpha}(\varphi_0) \rangle = 2 - 2\langle \varphi_0, \mathcal{K}_{\mathfrak{f},\alpha}(\varphi_0) \rangle \quad (\text{A.1.10})$$

Accordingly, $\|\mathcal{S}_\nu(\varphi_0)\|^2$ is minimal exactly for those φ_0 that maximize the quadratic functional $J_{\mathfrak{f},\alpha} : \varphi \mapsto \langle \varphi, \mathcal{K}_{\mathfrak{f},\alpha}(\varphi) \rangle$. Since $\mathcal{K}_{\mathfrak{f},\alpha}$ is a self-adjoint compact operator on a complex Hilbert-space $L^2(\Omega)$, either of the following alternatives hold according to spectral theory:

- (1) The spectrum $\sigma(\mathcal{K}_{f,\alpha})$ of $\mathcal{K}_{f,\alpha}$ contains positive numbers. Then all of these numbers are discrete eigenvalues of $\mathcal{K}_{f,\alpha}$ of finite multiplicity and φ_0 maximizes $J_{f,\alpha}(\varphi_0)$ if and only if φ_0 is an eigenfunction to the maximal eigenvalue $\lambda_{f,\alpha}$, in which case $J_{f,\alpha}(\varphi_0) = \lambda_{f,\alpha}$ holds true. Since the operator $\mathcal{K}_{f,\alpha}$ preserves real-valuedness, the φ_0 can be chosen to be real-valued. By (A.1.10), this implies existence and finite multiplicity of contrast-minimizing modes in the sense of (3.4.1) as well as the optimal value $C_{\text{IP2}}(\Omega, \mathbf{f}, \alpha) = 2^{1/2}$ of the stability constant. As eigenfunctions of an integral operator with infinitely smooth kernel, contrast-minimizing modes are furthermore infinitely smooth: for any indices $l_1, \dots, l_m \in \mathbb{N}_0$, the derivative $\partial_1^{l_1} \dots \partial_m^{l_m} \varphi_0$ exists at any point \mathbf{x} in the interior of Ω :

$$\partial_1^{l_1} \dots \partial_m^{l_m} \varphi_0(\mathbf{x}) = \partial_1^{l_1} \dots \partial_m^{l_m} \frac{1}{\lambda_{f,\alpha}} \mathcal{K}_{f,\alpha} \varphi_0(\mathbf{x}) \stackrel{(3.4.3)}{=} \int_{\Omega} \left(\partial_1^{l_1} \dots \partial_m^{l_m} k_{f,\alpha}(\mathbf{x} - \mathbf{y}) \right) \varphi(\mathbf{y}) \, d\mathbf{y}$$

- (2) $\sigma(\mathcal{K}_{f,\alpha}) \subset \mathbb{R}_{\leq 0}$. Then $J_{f,\alpha}(\varphi_0) \leq 0$ for any choice of φ_0 , where the value may be arbitrarily close to zero (even if 0 is not an eigenvalue of $\mathcal{K}_{f,\alpha}$). By (A.1.10), this implies that

$$\|\mathcal{S}_{\nu}(\varphi_0)\| \geq C_{\text{IP2}}(\Omega, \mathbf{f}, \alpha) \quad \text{for all } \varphi_0 \in L^2(\Omega) \text{ s.t. } \|\varphi_0\| = 1$$

for $C_{\text{IP2}}(\Omega, \mathbf{f}, \alpha) = 2^{1/2}$ and for no larger value. \square

A.2 Symmetries of contrast-minimizing modes

The next topic to be addressed are the symmetries of contrast-minimizing modes for reflectionally or rotationally symmetric domains $\Omega \subset \mathbb{R}^m$, that have been exploited in 3.4.2 to construct efficient numerical schemes for stability computations in higher dimensions. As the principal mathematical results in theorem 3.13 are not restricted to the specific forward maps of XPCI, we prove them for general Fourier-multipliers:

Theorem A.1 (Contrast-minimizing modes under geometrical symmetries). *Let $\Omega \subset \mathbb{R}^m$ be measurable, $\mathbf{m}_{\text{rad}} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{C}$ a bounded function and let $T : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^m); \varphi \mapsto \mathcal{F}^{-1}(\mathbf{m} \cdot \mathcal{F}(\varphi))$ with $\mathbf{m}(\boldsymbol{\xi}) := \mathbf{m}_{\text{rad}}(|\boldsymbol{\xi}|)$ denote the associated isotropic Fourier-multiplier. Assume that there exists a contrast-minimizing mode*

$$\varphi_0 \in \underset{\varphi \in L^2(\Omega): \|\varphi\|=1}{\operatorname{argmin}} \|T(\varphi)\|^2. \quad (\text{A.2.1})$$

Then φ_0 can be chosen to be real-valued and in this case, the following holds true:

- (a) (reflectional symmetry) *If $\Omega = -\Omega$, then there exists either an even or an odd contrast-minimizing mode, i.e. φ_0 can be chosen such that*

$$\varphi_0(-\cdot) = \varphi_0 \quad \text{or} \quad \varphi_0(-\cdot) = -\varphi_0. \quad (\text{A.2.2})$$

- (b) (radial symmetry) *If $\Omega = B[0; r]$ for $r > 0$, then there exists either a radially symmetric or radially anti-symmetric contrast-minimizing mode, i.e. φ_0 can be chosen such that for some function $\varphi_{0,\text{rad}} \in L^2([0; r], \mathbb{R})$, $\boldsymbol{\theta}_0 \in \mathbb{S}^{m-1}$ and parity-parameter $\nu \in \{0, 1\}$*

$$\varphi_0(\mathbf{x}) = (\boldsymbol{\theta}_0 \cdot \frac{\mathbf{x}}{|\mathbf{x}|})^{\nu} |\mathbf{x}|^{-(m-1)/2} \varphi_{0,\text{rad}}(|\mathbf{x}|) \quad \text{for almost all } \mathbf{x} \in \mathbb{R}^m. \quad (\text{A.2.3})$$

Proof. Real-valuedness: Since $T^*T : \varphi \mapsto \mathcal{F}^{-1}(|\mathbf{m}|^2 \cdot \mathcal{F}(\varphi)(\boldsymbol{\xi}))$ is a Fourier-multiplier with a Hermitean factor ($|\mathbf{m}(-\boldsymbol{\xi})|^2 = |\mathbf{m}(\boldsymbol{\xi})|^2 = (|\mathbf{m}(\boldsymbol{\xi})|^2)$), it preserves real- and imaginary parts, i.e. $\text{Re}(T^*T(\varphi)) = T^*T(\text{Re}(\varphi))$ and $\text{Im}(T^*T(\varphi)) = T^*T(\text{Im}(\varphi))$ for all $\varphi \in \mathbb{R}^m$. Thus,

$$\begin{aligned} \|T(\varphi_0)\|^2 &= \langle T(\varphi_0), T(\varphi_0) \rangle = \langle \varphi_0, T^*T(\varphi_0) \rangle \\ &= \langle \text{Re}(\varphi_0), \text{Re}(T^*T(\varphi_0)) \rangle + \langle \text{Im}(\varphi_0), \text{Im}(T^*T(\varphi_0)) \rangle \\ &= \langle \text{Re}(\varphi_0), T^*T(\text{Re}(\varphi_0)) \rangle + \langle \text{Im}(\varphi_0), T^*T(\text{Im}(\varphi_0)) \rangle \\ &= \|T(\text{Re}(\varphi_0))\|^2 + \|T(\text{Im}(\varphi_0))\|^2 \end{aligned} \quad (\text{A.2.4})$$

Combined with the general identity $\|\varphi_0\|^2 = \|\text{Re}(\varphi_0)\|^2 + \|\text{Im}(\varphi_0)\|^2$, (A.2.4) implies that $P(\varphi_0)/\|P(\varphi_0)\|$ for $P \in \{\text{Re}, \text{Im}\}$ must also be minimizer of (A.2.1) whenever $P(\varphi_0) \neq 0$. Since $0 \neq \varphi_0 = \text{Re}(\varphi_0) + i\text{Im}(\varphi_0)$, there exists such a $P \in \{\text{Re}, \text{Im}\}$. The replacement $\varphi_0 \leftarrow P(\varphi_0)/\|P(\varphi_0)\|$ provides the sought real-valued contrast-minimizing mode.

(a) Reflectional symmetry: As $\mathbf{m}(-\boldsymbol{\xi}) = \mathbf{m}(\boldsymbol{\xi})$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$, the reflection $R : \varphi \mapsto \varphi(-\cdot)$ commutes with T , $TR = RT$ (and likewise T^* and T^*T). The condition $\Omega = -\Omega$ guarantees that R maps $L^2(\Omega)$ onto $L^2(\Omega)$. Hence, the same two properties hold for the (anti-)symmetrization operators $S : \varphi \mapsto \frac{1}{2}(\varphi + R(\varphi))$ and $A : \varphi \mapsto \frac{1}{2}(\varphi - R(\varphi))$. Using that the decomposition $\varphi = S(\varphi) + A(\varphi)$ is orthogonal for all *real-valued* $\varphi \in L^2(\mathbb{R}^m, \mathbb{R})$ and that T^*T preserves the real-valuedness of φ_0 , we thus have

$$\begin{aligned} \|\varphi_0\|^2 &= \|S(\varphi_0)\|^2 + \|A(\varphi_0)\|^2 \\ \|T(\varphi_0)\|^2 &= \langle \varphi_0, T^*T(\varphi_0) \rangle = \langle S(\varphi_0) + A(\varphi_0), T^*T(S(\varphi_0) + A(\varphi_0)) \rangle \\ &= \langle S(\varphi_0) + A(\varphi_0), S(T^*T(\varphi_0)) + A(T^*T(\varphi_0)) \rangle \\ &= \langle S(\varphi_0), S(T^*T(\varphi_0)) \rangle + \langle A(\varphi_0), A(T^*T(\varphi_0)) \rangle \\ &= \|T(S(\varphi_0))\|^2 + \|T(A(\varphi_0))\|^2 \end{aligned} \quad (\text{A.2.5})$$

By an analogous argument as for the real-valuedness, (A.2.5) implies that $P(\varphi_0)/\|P(\varphi_0)\|$ for $P \in \{S, A\}$ provides the sought (anti-)symmetric (even or odd) contrast-minimizing mode.

(b) Existence of radially (anti-)symmetric modes: We prove existence of radially (anti-)symmetric modes φ_0 by explicit construction. Since $\Omega = -\Omega$ for $\Omega = B[0; 1/2]$, there exists a real-valued contrast-minimizing mode $\varphi \in L^2(\Omega)$ that either even or odd. We set $\nu := 0$ if φ is even and $\nu := 1$ if φ is odd and consider the (scaled) Fourier transform $\hat{\varphi} := i^\nu \mathcal{F}(\varphi)$. $\hat{\varphi}$ is then real-valued and even/odd depending on whether φ is even/odd. Since $\text{supp}(\varphi) \subset \Omega = B[0; r]$ is compact, $\hat{\varphi}$ is an entire function by the Paley-Wiener theorem, i.e. has a globally and absolutely convergent Taylor-series (extendable to \mathbb{C}^m): with multi-index notation, $\boldsymbol{\xi}^\alpha := \xi_1^{\alpha_1} \cdot \dots \cdot \xi_m^{\alpha_m}$ and $|\alpha| := \sum_{i=1}^m \alpha_i$ for $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m) \in \mathbb{C}^m$, $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}_0$,

$$\hat{\varphi}(\boldsymbol{\xi}) = \sum_{k=0}^{\infty} \sum_{\alpha \in \mathbb{N}_0^m : |\alpha|=2k+\nu} c_\alpha \boldsymbol{\xi}^\alpha \quad \text{for all } \boldsymbol{\xi} \in \mathbb{R}^m \quad (\text{A.2.6})$$

for some coefficients $\{c_\alpha\} \subset \mathbb{R}$ (since $\hat{\varphi}$ is real-valued). The restriction to exponents $\alpha \in \mathbb{N}_0$ with $|\alpha| \in 2\mathbb{N}_0 + \nu$ follows from the even or odd parity of $\hat{\varphi}$.

We write $\hat{\varphi}$ in polar coordinates, $(\boldsymbol{\theta}, \xi) \mapsto \hat{\varphi}(\xi\boldsymbol{\theta})$, for $\xi \in \mathbb{R}$, $\boldsymbol{\theta} \in \mathbb{S}^{m-1}$ and express the dependence on $\boldsymbol{\theta}$ by expansion in *real-valued* m -dimensional spherical harmonics $\{Y_l\}_{l \in \mathbb{N}} \subset$

$L^2(\mathbb{S}^{m-1}, \mathbb{R})$, which is possible by the orthonormal-basis-property of $\{Y_l\}_{l \in \mathbb{N}}$:

$$\hat{\varphi}(\xi \boldsymbol{\theta}) = \sum_{l \in I} 2^{1/2} |\xi|^{-(m-1)/2} a_l(\xi) Y_l(\boldsymbol{\theta}) \quad \text{for all } \boldsymbol{\theta} \in \mathbb{S}^{m-1}, \xi \in \mathbb{R} \quad (\text{A.2.7})$$

for some nonempty set $\emptyset \neq I \subset \mathbb{N}$ and non-vanishing functions $\{a_l\}_{l \in I} \subset L^2(\mathbb{R}, \mathbb{R}) \setminus \{0\}$. The expansion (A.2.7). As $\hat{\varphi}$ is infinitely smooth, the expansion in (A.2.7) converges absolutely, both pointwise and in an L^2 -sense [34], which justifies the manipulations made below. The factor $2^{1/2} |\xi|^{-(m-1)/2}$ accounts for the weighting of integrals in polar coordinates, see (A.2.8).

Now we express the L^2 -norms of φ and $T(\varphi)$ in terms of this expansion. Using unitarity of \mathcal{F} , isotropy of the Fourier-multiplier $\mathbf{m} = \mathbf{m}_{\text{rad}}(|\cdot|)$ of T and orthonormality of the Y_l , $\int_{\mathbb{S}^{m-1}} \overline{Y_l(\boldsymbol{\theta})} Y_k(\boldsymbol{\theta}) \, d\boldsymbol{\theta} = \delta_{lk}$ (Kronecker-delta), we obtain

$$\begin{aligned} \|\varphi\|^2 &= \|\hat{\varphi}\|^2 = \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{S}^{m-1}} |\hat{\varphi}(\xi \boldsymbol{\theta})|^2 |\xi|^{m-1} \, d\xi \, d\boldsymbol{\theta} \\ &= \sum_{k, l \in I} \left(\int_{\mathbb{S}^{m-1}} \overline{Y_l(\boldsymbol{\theta})} Y_k(\boldsymbol{\theta}) \, d\boldsymbol{\theta} \right) \left(\int_{\mathbb{R}} a_l(\xi) a_k(\xi) \, d\xi \right) = \sum_{l \in I} \|a_l\|^2 \end{aligned} \quad (\text{A.2.8})$$

$$\|T(\varphi)\|^2 = \|\mathbf{m} \cdot \hat{\varphi}\|^2 = \frac{1}{2} \int_{\mathbb{R}} |\mathbf{m}_{\text{rad}}(|\xi|)|^2 \int_{\mathbb{S}^{m-1}} |\hat{\varphi}(\xi \boldsymbol{\theta})|^2 |\xi|^{m-1} \, d\xi \, d\boldsymbol{\theta} = \sum_{l \in I} \|\mathbf{m}_{\text{rad}} \cdot a_l\|^2. \quad (\text{A.2.9})$$

Here, we have extended \mathbf{m}_{rad} to an even function on \mathbb{R} by setting $\mathbf{m}_{\text{rad}}(\xi) := \mathbf{m}_{\text{rad}}(-\xi)$ for $\xi < 0$. As can be seen from elementary algebra, there must exist an $\ell \in I$ such that

$$\|\mathbf{m}_{\text{rad}} \cdot a_\ell\|^2 / \|a_\ell\|^2 \leq \left(\sum_{l \in I} \|\mathbf{m}_{\text{rad}} \cdot a_l\|^2 \right) / \left(\sum_{l \in I} \|a_l\|^2 \right). \quad (\text{A.2.10})$$

Let $\ell \in I$ such that (A.2.10) is satisfied and define $\tilde{a}_\ell(\xi) := 2^{1/2} |\xi|^{-(m-1)/2} a_\ell(\xi)$. By orthonormality of the Y_l , it then holds for almost all $\xi \in \mathbb{R}$ that

$$\tilde{a}_\ell(\xi) = \int_{\mathbb{S}^{m-1}} \overline{Y_\ell(\boldsymbol{\theta})} \hat{\varphi}(\xi \boldsymbol{\theta}) \, d\boldsymbol{\theta} = \underbrace{\sum_{k=0}^{\infty} \sum_{\alpha \in \mathbb{N}_0^m: |\alpha|=2k+\nu} c_\alpha \left(\int_{\mathbb{S}^{m-1}} \overline{Y_\ell(\boldsymbol{\theta})} \boldsymbol{\theta}^\alpha \, d\boldsymbol{\theta} \right)}_{=: d_{2k+\nu}} \xi^{2k+\nu} \quad (\text{A.2.11})$$

Global convergence of the Taylor-series on the r.h.s. of (A.2.11) follows from that in (A.2.6). Hence, (A.2.11) shows that \tilde{a}_ℓ is an entire function, which is even/odd if $\hat{\varphi}$ is even/odd. By the same argument, the following Taylor-series are globally convergent ($\boldsymbol{\theta}_0 \in \mathbb{S}^{m-1}$ arbitrary):

$$\hat{\varphi}_0(\boldsymbol{\xi}) := \begin{cases} \sum_{k=0}^{\infty} d_{2k} (\boldsymbol{\xi} \cdot \boldsymbol{\xi})^k & \text{if } \nu = 0 \text{ (even case)} \\ (\boldsymbol{\theta}_0 \cdot \boldsymbol{\xi}) \sum_{k=0}^{\infty} d_{2k+\nu} (\boldsymbol{\xi} \cdot \boldsymbol{\xi})^k & \text{if } \nu = 1 \text{ (odd case)} \end{cases} \quad \text{for all } \boldsymbol{\xi} \in \mathbb{R}^m \quad (\text{A.2.12})$$

and show that $\varphi_0 := i^{-\nu} \mathcal{F}^{-1}(\hat{\varphi}_0)$ is a radially (anti-)symmetric contrast-minimizing mode:

- *Square-integrability*: By comparing (A.2.12) and (A.2.11), we find that $\hat{\varphi}_0(\xi \boldsymbol{\theta}) = (\boldsymbol{\theta}_0 \cdot \boldsymbol{\theta})^\nu \tilde{a}_\ell(\xi)$ almost everywhere. With $c_\nu := (\int_{\mathbb{S}^{m-1}} (\boldsymbol{\theta}_0 \cdot \boldsymbol{\theta})^{2\nu} \, d\boldsymbol{\theta})^{1/2}$, it thus holds that

$$\|\hat{\varphi}_0\|^2 = \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{S}^{m-1}} |\hat{\varphi}_0(\xi \boldsymbol{\theta})|^2 |\xi|^{m-1} \, d\xi \, d\boldsymbol{\theta} = \frac{c_\nu^2}{2} \int_{\mathbb{R}} |\tilde{a}_\ell(\xi)|^2 |\xi|^{m-1} \, d\xi = c_\nu^2 \|a_\ell\|^2. \quad (\text{A.2.13})$$

Since $\|a_\ell\| < \infty$, this proves that $\hat{\varphi}_0 \in L^2(\mathbb{R}^m)$ and hence $\varphi_0 = i^{-\nu} \mathcal{F}^{-1}(\hat{\varphi}_0) \in L^2(\mathbb{R}^m)$.

- *Support in Ω* : Since the original contrast-minimizing mode φ is in $L^2(\Omega)$ for $\Omega = B[0; r]$, a detailed version of the Paley-Wiener theorem given in [102, p. 181] implies that $\hat{\varphi}$ (as an entire function in \mathbb{C}^m) satisfies

$$|\hat{\varphi}(\boldsymbol{\xi})| \leq C \exp(r|\operatorname{Im}(\boldsymbol{\xi})|) \quad \text{for all } \boldsymbol{\xi} \in \mathbb{C}^m. \quad (\text{A.2.14})$$

By definition, $\hat{\varphi}_0$ is also an entire function by (A.2.12) and in $L^2(\mathbb{R}^m)$. According to the converse statement in the Paley-Wiener theorem from [102, p. 181], it is thus sufficient to prove that the bound (A.2.14) also holds for $\hat{\varphi}_0$ with a possibly modified constant $C_0 > 0$ in order to verify that $\varphi_0 \in L^2(\Omega)$.

The argument to be made is widely analogous as for classical support-theorems for the m -dimensional Radon-transform¹ [134, Theorem 2.2]. Let $\boldsymbol{\xi} = \boldsymbol{\eta} + i\boldsymbol{\zeta} \in \mathbb{C}^m \setminus \{0\}$ with $\boldsymbol{\eta}, \boldsymbol{\zeta} \in \mathbb{R}^m$ and $\xi \in \mathbb{C}$ such that $\xi^2 = \boldsymbol{\xi} \cdot \boldsymbol{\xi}$. Then explicit computation shows that

$$\begin{aligned} \operatorname{Im}(\xi)^2 &= r(|\boldsymbol{\zeta}|^2 - |\boldsymbol{\eta}|^2 + ((|\boldsymbol{\eta}|^2 - |\boldsymbol{\zeta}|^2)^2 + 4(\boldsymbol{\eta} \cdot \boldsymbol{\zeta})^2)^{1/2}) \\ &\leq r(|\boldsymbol{\zeta}|^2 - |\boldsymbol{\eta}|^2 + ((|\boldsymbol{\eta}|^2 - |\boldsymbol{\zeta}|^2)^2 + 4|\boldsymbol{\eta}|^2|\boldsymbol{\zeta}|^2)^{1/2}) = |\boldsymbol{\zeta}|^2 = |\operatorname{Im}(\boldsymbol{\xi})|^2 \end{aligned} \quad (\text{A.2.15})$$

By (A.2.11) and (A.2.12), it holds that $\hat{\varphi}_0(\boldsymbol{\xi}) = (\boldsymbol{\theta}_0 \cdot \boldsymbol{\xi})^\nu \tilde{a}_\ell(\xi) / \xi^\nu$ and thus $|\hat{\varphi}_0(\boldsymbol{\xi})| \leq |\boldsymbol{\theta}_0|^\nu (|\boldsymbol{\eta}|^2 + |\boldsymbol{\zeta}|^2)^{\nu/2} |\xi|^{-\nu} |\tilde{a}_\ell(\xi)| \leq |\tilde{a}_\ell(\xi)|$ (the bound remains true for $\xi = 0 \Leftrightarrow \boldsymbol{\xi} = 0$). Letting $C_0 := C \int_{\mathbb{S}^{m-1}} |Y_\ell(\boldsymbol{\theta})| d\boldsymbol{\theta}$, we obtain the sought analogue of (A.2.14) for $\hat{\varphi}_0$:

$$\begin{aligned} |\hat{\varphi}_0(\boldsymbol{\xi})| &\leq |\tilde{a}_\ell(\xi)| \stackrel{(\text{A.2.11})}{\leq} \int_{\mathbb{S}^{m-1}} |Y_\ell(\boldsymbol{\theta})| |\hat{\varphi}(\xi\boldsymbol{\theta})| d\boldsymbol{\theta} \stackrel{(\text{A.2.14})}{\leq} C_0 \exp(r|\operatorname{Im}(\xi\boldsymbol{\theta})|) \\ &= C_0 \exp(r|\operatorname{Im}(\xi)|) \stackrel{(\text{A.2.15})}{\leq} C_0 \exp(r|\operatorname{Im}(\boldsymbol{\xi})|). \end{aligned} \quad (\text{A.2.16})$$

- *Contrast-minimizing property*: Analogously to (A.2.13), it follows that

$$\|T(\varphi_0)\|^2 = \frac{c_\nu^2}{2} \int_{\mathbb{R}} |\mathbf{m}_{\text{rad}}(|\xi|)|^2 |\tilde{a}_\ell(\xi)|^2 |\xi|^{m-1} d\xi = c_\nu^2 \|\mathbf{m}_{\text{rad}} \cdot a_\ell\|^2. \quad (\text{A.2.17})$$

By (A.2.9) and (A.2.10), this implies $\|T(\varphi_0)\|/\|\varphi_0\| = \|\mathbf{m}_{\text{rad}} \cdot a_\ell\|/\|a_\ell\| \leq \|T(\varphi)\|/\|\varphi\|$. As φ is a minimizer of (A.2.1), the same must thus hold for $\varphi_0/\|\varphi_0\|$.

- *Radial (anti-)symmetry*: First consider the even case $\nu = 0$. Then $\hat{\varphi}_0$ is radially symmetric by (A.2.12), i.e. $\hat{\varphi}_0(A\boldsymbol{\xi}) = \hat{\varphi}_0(\boldsymbol{\xi})$ for all $\boldsymbol{\xi} \in \mathbb{R}^m$ and arbitrary orthogonal matrices $A \in O(m) := \{B \in \mathbb{R}^{m \times m} : BB^T = B^T B = \text{id}\}$. As the \mathcal{F}^{-1} commutes with orthogonal transforms of the coordinate system, this implies radial symmetry of φ_0 as well:

$$\varphi_0(A\mathbf{x}) = \mathcal{F}^{-1}(\hat{\varphi}_0)(A\mathbf{x}) = \mathcal{F}^{-1}(\hat{\varphi}_0(A(\cdot)))(\mathbf{x}) = \mathcal{F}^{-1}(\hat{\varphi}_0)(\mathbf{x}) = \varphi_0(\mathbf{x}) \quad (\text{A.2.18})$$

for all $\mathbf{x} \in \mathbb{R}^m$, $A \in O(m)$. Thus, $\varphi_0(\mathbf{x}) = |\mathbf{x}|^{-(m-1)/2} \varphi_0^{\text{rad}}(|\mathbf{x}|)$ for some function $\varphi_0^{\text{rad}} : [0; r] \rightarrow \mathbb{R}$ and almost all $\mathbf{x} \in B[0; r]$. By writing the L^2 -integral in polar coordinates, it follows that $\|\varphi_0^{\text{rad}}\| = (\int_{\mathbb{S}^{m-1}} d\boldsymbol{\theta})^{1/2} \|\varphi_0\| < \infty$, i.e. $\varphi_0^{\text{rad}} \in L^2([0; r], \mathbb{R})$.

¹Indeed, there is a strong relation between the m -d Radon-transform $\mathcal{R}^{(m\text{d})}(f)(\boldsymbol{\theta}, x) := \int_{\mathbf{x} \cdot \boldsymbol{\theta} = 0} f(\mathbf{x}) d\mathbf{x}$ and the construction of φ_0 : By the Fourier-Slice-Theorem $\mathcal{F}(\mathcal{R}^{(m\text{d})}(f))(\boldsymbol{\theta}, \xi) \propto \mathcal{F}(f)(\xi\boldsymbol{\theta})$ (up to 2π -factors), one has $\mathcal{R}^{(m\text{d})}(\varphi_0)(\boldsymbol{\theta}, x) \propto (\boldsymbol{\theta}_0 \cdot \boldsymbol{\theta})^\nu \mathcal{F}^{-1}(\tilde{a}_\ell)(x)$ and $\mathcal{F}^{-1}(\tilde{a}_\ell)(x) \propto \int_{\mathbb{S}^{m-1}} Y_\ell(\boldsymbol{\theta}) \mathcal{R}^{(m\text{d})}(\varphi)(\boldsymbol{\theta}, x) d\boldsymbol{\theta}$ by (A.2.11).

For the odd case $\nu = 1$, note that, by (A.2.12), $\hat{\varphi}_0$ is of the form $\hat{\varphi}_0(\boldsymbol{\xi}) = (\boldsymbol{\theta}_0 \cdot \boldsymbol{\xi}) \tilde{\varphi}_0(\boldsymbol{\xi})$ with a radially symmetric $\tilde{\varphi}_0$. By differentiation-rules for the Fourier transform, this implies

$$\varphi_0 = -i\mathcal{F}^{-1}(\hat{\varphi}_0) = \mathcal{F}^{-1}(-i(\boldsymbol{\theta}_0 \cdot \boldsymbol{\xi})\tilde{\varphi}_0) = -(\boldsymbol{\theta}_0 \cdot \nabla)\mathcal{F}^{-1}(\tilde{\varphi}_0). \quad (\text{A.2.19})$$

By the argument made above, $\mathcal{F}^{-1}(\tilde{\varphi}_0)$ is radially symmetric as well, i.e. $\mathcal{F}^{-1}(\tilde{\varphi}_0)(\mathbf{x}) = \tilde{\varphi}_0^{\text{rad}}(|\mathbf{x}|)$ for all $\mathbf{x} \in \mathbb{R}^m$ and some function $\tilde{\varphi}_0^{\text{rad}} : [0; r] \rightarrow \mathbb{R}$. Hence, (A.2.19) yields

$$\varphi_0(\mathbf{x}) = -(\boldsymbol{\theta}_0 \cdot \nabla)\tilde{\varphi}_0^{\text{rad}}(|\mathbf{x}|) = (\boldsymbol{\theta}_0 \cdot \frac{\mathbf{x}}{|\mathbf{x}|})|\mathbf{x}|^{-(m-1)/2}\varphi_0^{\text{rad}}(|\mathbf{x}|) \quad (\text{A.2.20})$$

for almost all $\mathbf{x} \in B[0; r]$ with $\varphi_0^{\text{rad}}(x) := -|\mathbf{x}|^{(m-1)/2} \frac{d}{dx} \tilde{\varphi}_0^{\text{rad}}$. Hence, φ_0 is of the claimed radially anti-symmetric form. Square-integrability of φ_0^{rad} follows analogously as in the radially symmetric case. \square

The proof of theorem 3.13 from §3.4.2 essentially amounts to applying the general result from theorem A.1 to the specific Fourier-multiplier \mathcal{S}_ν .

Theorem (3.13). *Let $\mathfrak{f} > 0$, $\alpha \in \mathbb{R}$ and $\Omega \subset \mathbb{R}^m$ bounded. Then the following holds true:*

- (a) (point-symmetry) *If $\Omega = -\Omega$, then there exists either an even or an odd contrast-minimizing mode of \mathcal{S}_ν , i.e. φ_0 from theorem 3.12 may be chosen such that*

$$\varphi_0(-\cdot) = \varphi_0 \quad \text{or} \quad \varphi_0(-\cdot) = -\varphi_0. \quad (\text{A.2.21})$$

- (b) (radial symmetry) *If $\Omega = B[0; \frac{1}{2}]$, then there exists either a radially symmetric or a radially anti-symmetric contrast-minimizing mode, i.e. φ_0 from theorem 3.12 may be chosen such that, for some function $\varphi_0^{\text{rad}} \in L^2([0; \frac{1}{2}])$, either of the following applies:*

- (1) $\varphi_0(\mathbf{x}) = |\mathbf{x}|^{-\frac{m-1}{2}} \varphi_0^{\text{rad}}(|\mathbf{x}|)$ for all $\mathbf{x} \in \mathbb{R}^m \setminus \{0\}$. Then φ_0^{rad} is an eigenfunction of a selfadjoint compact integral operator $\mathcal{K}_{\mathfrak{f}, \alpha}^{\text{rad}, s} : L^2([0; \frac{1}{2}]) \rightarrow L^2([0; \frac{1}{2}])$ defined by

$$\begin{aligned} \mathcal{K}_{\mathfrak{f}, \alpha}^{\text{rad}, s}(\varphi_0^{\text{rad}})(r) &= \int_0^{\frac{1}{2}} K_{\mathfrak{f}, \alpha}^{\text{rad}, s}(r, \rho) \varphi_0^{\text{rad}}(\rho) d\rho \\ K_{\mathfrak{f}, \alpha}^{\text{rad}, s}(r, \rho) &= \left(\frac{\mathfrak{f}}{2}\right)^{\frac{m}{2}} (r\rho)^{\frac{m-1}{2}} \cos\left(\frac{\mathfrak{f}(r^2 + \rho^2)}{4} - \gamma\right) \begin{cases} 2 \cos(\frac{\mathfrak{f}}{2}r\rho) & (m=1) \\ 2\pi J_0(\frac{\mathfrak{f}}{2}r\rho) & (m=2) \\ 4\pi \text{sinc}(\frac{\mathfrak{f}}{2}r\rho) & (m=3) \end{cases} \end{aligned} \quad (\text{A.2.22})$$

to its maximal eigenvalue, which is $\lambda_{\mathfrak{f}, \nu}$. Then $\tilde{\varphi}_0 : \mathbf{x} \mapsto |\mathbf{x}|^{-\frac{m-1}{2}} \tilde{\varphi}_0^{\text{rad}}(|\mathbf{x}|)$ for any $\tilde{\varphi}_0^{\text{rad}} \in L^2([0; \frac{1}{2}])$ with $\mathcal{K}_{\mathfrak{f}, \nu}^{\text{rad}, s}(\tilde{\varphi}_0^{\text{rad}}) = \lambda_{\mathfrak{f}, \nu} \tilde{\varphi}_0^{\text{rad}}$ is also a contrast-minimizing mode.

- (2) $\varphi_0(\mathbf{x}) = (\boldsymbol{\theta}_0 \cdot \frac{\mathbf{x}}{|\mathbf{x}|})|\mathbf{x}|^{-\frac{m-1}{2}} \varphi_0^{\text{rad}}(|\mathbf{x}|)$ for all $\mathbf{x} \in \mathbb{R}^m \setminus \{0\}$ and some $\boldsymbol{\theta}_0 \in \mathbb{S}^{m-1}$. Then analogous statements to (1) hold for a modified integral operator $\mathcal{K}_{\mathfrak{f}, \alpha}^{\text{rad}, a}$ with kernel

$$K_{\mathfrak{f}, \alpha}^{\text{rad}, a}(r, \rho) = \left(\frac{\mathfrak{f}}{2}\right)^{\frac{m}{2}} (r\rho)^{\frac{m-1}{2}} \sin\left(\frac{\mathfrak{f}(r^2 + \rho^2)}{4} - \gamma\right) \begin{cases} 2 \sin(\frac{\mathfrak{f}}{2}r\rho) & (m=1) \\ 2\pi J_1(\frac{\mathfrak{f}}{2}r\rho) & (m=2) \\ 4\pi \text{sinc}_2(\frac{\mathfrak{f}}{2}r\rho) & (m=3) \end{cases} \quad (\text{A.2.23})$$

Here, the abbreviation $\gamma := 2\alpha - m\pi/4$ is used, $J_n : x \mapsto \frac{1}{\pi} \int_0^\pi \cos(nt - x \sin(t)) dt$ denote n th order Bessel functions of the first kind and $\text{sinc}_2 : x \mapsto (\sin(x) - x \cos(x))/x^2$.

Proof. Part (a) and existence of radially (anti-)symmetric contrast-minimizing modes follows from theorem A.1. Hence, what remains to be shown is the characterization in terms of the operators $\mathcal{K}_{f,\alpha}^{\text{rad,s}}$ and $\mathcal{K}_{f,\alpha}^{\text{rad,a}}$. We consider the case (1) where φ_0 is radially symmetric. By writing the integral operator $\mathcal{K}_{f,\alpha}$ in polar coordinates, we have for all $\mathbf{x} \in \Omega \setminus \{0\}$ ($\Omega = B[0; \frac{1}{2}]$)

$$\begin{aligned} \lambda_{f,\alpha} \varphi_0^{\text{rad}}(|\mathbf{x}|) &= |\mathbf{x}|^{\frac{m-1}{2}} \lambda_{f,\alpha} \varphi_0(\mathbf{x}) = |\mathbf{x}|^{\frac{m-1}{2}} \mathcal{K}_{f,\alpha} \varphi_0(\mathbf{x}) = \int_{\Omega} |\mathbf{x}|^{\frac{m-1}{2}} k_{f,\alpha}(\mathbf{x} - \mathbf{y}) |\mathbf{y}|^{-\frac{m-1}{2}} \varphi_0^{\text{rad}}(|\mathbf{y}|) d\mathbf{y} \\ &= |\mathbf{x}|^{\frac{m-1}{2}} \underbrace{\left(\int_{\mathbb{S}^{m-1}} k_{f,\alpha}(\mathbf{x} - \rho \boldsymbol{\theta}) d\boldsymbol{\theta} \right)}_{=K_{f,\alpha}^{\text{rad,s}}(|\mathbf{x}|,\rho)} \rho^{\frac{m-1}{2}} \varphi_0^{\text{rad}}(\rho) d\rho = \mathcal{K}_{f,\alpha}^{\text{rad,s}}(\varphi_0^{\text{rad}})(|\mathbf{x}|). \end{aligned} \quad (\text{A.2.24})$$

The fact that the underbraced expression only depends on the euclidean norm $|\mathbf{x}|$ is a consequence of the radial symmetry of the kernel $k_{f,\alpha}$. Explicit computations of the integral over the unit sphere \mathbb{S}^{m-1} show that the kernel $K_{f,\alpha}^{\text{rad,s}}$ of the radially integrated operator is form in (3.4.8). Accordingly, (A.2.24) shows that φ_0^{rad} is an eigenfunction of $\mathcal{K}_{f,\alpha}^{\text{rad,s}}$ to the eigenvalue $\lambda_{f,\alpha}$.

By writing the L^2 -norm in polar coordinates, it can be seen that the embedding

$$\iota_{\text{rad,s}} : L^2([0; \frac{1}{2}]) \rightarrow L^2(\Omega); \varphi^{\text{rad}} \mapsto |\cdot|^{-\frac{m-1}{2}} \varphi^{\text{rad}}(|\cdot|) \quad (\text{A.2.25})$$

is well-defined, linear and bounded (in fact isometric). Moreover, it can be readily seen that $\mathcal{K}_{f,\alpha}^{\text{rad,s}} = \iota_{\text{rad,s}}^* \mathcal{K}_{f,\alpha} \iota_{\text{rad,s}}$. As $\mathcal{K}_{f,\alpha} : L^2(\Omega) \rightarrow L^2(\Omega)$ is bounded, compact and selfadjoint, the same must thus hold for $\mathcal{K}_{f,\alpha}^{\text{rad,s}} : L^2([0; \frac{1}{2}]) \rightarrow L^2([0; \frac{1}{2}])$.

By reading (A.2.24) in a reverse manner, it is furthermore seen that any eigenpair $\lambda \in \mathbb{R}$, $\varphi_\lambda^{\text{rad}} \in L^2([0; \frac{1}{2}])$ ($\mathcal{K}_{f,\alpha}^{\text{rad,s}}$ has only real eigenvalues by selfadjointness) s.t. $\mathcal{K}_{f,\alpha}^{\text{rad,s}} \varphi_\lambda^{\text{rad}} = \lambda \varphi_\lambda^{\text{rad}}$ induces a radially symmetric eigenfunction $\varphi_\lambda = \iota_{\text{rad,s}}(\varphi_\lambda^{\text{rad}})$ of $\mathcal{K}_{f,\alpha}$ to the same eigenvalue. This also implies that $\lambda_{f,\alpha}$ has to be the maximal eigenvalue of $\mathcal{K}_{f,\alpha}^{\text{rad,s}}$.

The proof for the case of radially *anti*-symmetric modes works completely analogously. \square

Bibliography

- [1] J. Adler, H. Kohr, M. J. Ehrhardt, O. Öktem, et al. odlgroup/odl: ODL 0.7.0. *Zenodo*, <https://doi.org/10.5281/zenodo.1442734>, September 2018.
- [2] G. Alessandrini and S. Vessella. Lipschitz stability for the inverse conductivity problem. *Advances in Applied Mathematics*, 35(2):207–241, 2005.
- [3] J. Als-Nielsen and D. McMorrow. *Elements of modern X-ray physics*. John Wiley & Sons, 2011.
- [4] M. Altarelli, R. Brinkmann, M. Chergui, W. Decking, B. Dobson, S. Düsterer, G. Grübel, W. Graeff, H. Graafsma, J. Hajdu, et al. The european X-ray free-electron laser. *Technical design report, DESY*, 97:1–26, 2006.
- [5] H. Ammari, Y. T. Chow, and J. Zou. Phased and phaseless domain reconstructions in the inverse scattering problem via scattering coefficients. *SIAM Journal on Applied Mathematics*, 76(3):1000–1030, 2016.
- [6] A. H. Andersen and A. C. Kak. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. *Ultrasonic imaging*, 6(1):81–94, 1984.
- [7] M. S. Andersen and P. C. Hansen. Generalized row-action methods for tomographic imaging. *Numerical Algorithms*, 67(1):121–144, 2014.
- [8] A. Averbuch, R. R. Coifman, D. L. Donoho, M. Elad, and M. Israeli. Fast and accurate polar Fourier transform. *Applied and Computational Harmonic Analysis*, 21(2):145–167, 2006.
- [9] A. B. Bakushinskii. The problem of the convergence of the iteratively regularized Gauss-Newton method. *Zhurnal Vychislitel’noi Matematiki i Matematicheskoi Fiziki*, 32(9):1503–1509, 1992.
- [10] J. F. Barrett and N. Keat. Artifacts in CT: recognition and avoidance. *Radiographics*, 24(6):1679–1691, 2004.
- [11] R. Barrett, R. Baker, P. Cloetens, C. Morawe, R. Tucoulou, and A. Vivo. Reflective optics for hard X-ray nanofocusing applications at the ESRF. *Synchrotron Radiation News*, 29(4):10–15, 2016.
- [12] M. Bartels. *Cone-beam X-ray phase contrast tomography of biological samples: Optimization of contrast, resolution and field of view*, volume 13 of *Göttingen Series in X-ray Physics*. Universitätsverlag Göttingen, 2013.

- [13] M. Bartels, M. Krenkel, J. Haber, R. Wilke, and T. Salditt. X-ray holographic imaging of hydrated biological cells in solution. *Physical review letters*, 114(4):048103, 2015.
- [14] M. Bartels, M. Priebe, R. N. Wilke, S. P. Krüger, K. Giewekemeyer, S. Kalbfleisch, C. Olendrowitz, M. Sprung, and T. Salditt. Low-dose three-dimensional hard X-ray imaging of bacterial cells. *Optical Nanoscopy*, 1(1):1–7, 2012.
- [15] A. Barty, J. Küpper, and H. N. Chapman. Molecular imaging using X-ray free-electron lasers. *Annual review of physical chemistry*, 64:415–435, 2013.
- [16] A. Barty, S. Marchesini, H. Chapman, C. Cui, M. Howells, D. Shapiro, A. Minor, J. Spence, U. Weierstall, J. Ilavsky, et al. Three-dimensional coherent X-ray diffraction imaging of a ceramic nanofoam: Determination of structural deformation mechanisms. *Physical review letters*, 101(5):055501, 2008.
- [17] H. Bauschke, P. L. Combettes, and D. R. Luke. Phase retrieval, error reduction algorithm, and fienup variants: a view from convex optimization. *JOSA A*, 19(7):1334–1345, 2002.
- [18] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.
- [19] T. Bendory, R. Beinert, and Y. C. Eldar. Fourier phase retrieval: Uniqueness and algorithms. In *Compressed Sensing and its Applications*, pages 55–91. Springer, 2017.
- [20] E. Beretta, M. V. De Hoop, F. Faucher, and O. Scherzer. Inverse boundary value problem for the helmholtz equation: quantitative conditional lipschitz stability estimates. *SIAM Journal on Mathematical Analysis*, 48(6):3962–3983, 2016.
- [21] D. P. Bertsekas. Incremental gradient, subgradient, and proximal methods for convex optimization: A survey. *Optimization for Machine Learning*, 2010:1–38, 2011.
- [22] D. P. Bertsekas. Incremental proximal methods for large scale convex optimization. *Mathematical programming*, 129(2):163, 2011.
- [23] E. Betzig, A. Lewis, A. Harootunian, M. Isaacson, and E. Kratschmer. Near field scanning optical microscopy (NSOM): development and biophysical applications. *Biophysical journal*, 49(1):269–279, 1986.
- [24] B. Blaschke, A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauss-Newton method. *IMA Journal of Numerical Analysis*, 17(3):421–436, 1997.
- [25] F. Bleichrodt, T. van Leeuwen, and K. J. Batenburg. Robust artefact reduction in tomography using Student’s-t data fitting. In *The 13th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pages 395–398, 2015.
- [26] M. Born and E. Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

- [27] L. Bragg. Nobel lecture: The diffraction of X-rays by crystals. *The Nobel Prize in Physics 1915*, 2018. <https://www.nobelprize.org/prizes/physics/1915/wl-bragg/lecture> (accessed 2018-11-20).
- [28] A. Bravin, P. Coan, and P. Suortti. X-ray phase-contrast imaging: from pre-clinical applications towards clinics. *Physics in Medicine and Biology*, 58(1):R1, 2012.
- [29] W. L. Briggs, S. F. McCormick, et al. *A multigrid tutorial*. SIAM, 2000.
- [30] M. Burger and B. Kaltenbacher. Regularizing Newton–Kaczmarz methods for nonlinear ill-posed problems. *SIAM Journal on Numerical Analysis*, 44(1):153–182, 2006.
- [31] M. Burger, A. Sawatzky, and G. Steidl. First order algorithms in variational image processing. In *Splitting Methods in Communication, Imaging, Science, and Engineering*, pages 345–407. Springer, 2016.
- [32] A. Burvall, U. Lundström, P. A. Takman, D. H. Larsson, and H. M. Hertz. Phase retrieval in X-ray phase-contrast imaging suitable for tomography. *Optics express*, 19(11):10359–10376, 2011.
- [33] T. M. Buzug. *Computed tomography: from photon statistics to modern cone-beam CT*. Springer, 2008.
- [34] A. P. Calderón and A. Zygmund. Singular integral operators and differential equations. *American Journal of Mathematics*, 79(4):901–921, 1957.
- [35] A. Chambolle, M. J. Ehrhardt, P. Richtárik, and C.-B. Schönlieb. Stochastic primal-dual hybrid gradient algorithm with arbitrary sampling and imaging applications. *SIAM Journal on Optimization*, 28(4):2783–2808, 2018.
- [36] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.
- [37] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. Spence, et al. High-resolution ab initio three-dimensional X-ray diffraction microscopy. *JOSA A*, 23(5):1179–1200, 2006.
- [38] O. Christensen. B-spline generated frames. In *Four Short Courses on Harmonic Analysis*, pages 51–86. Springer, 2010.
- [39] P. Cloetens. *Contribution to phase contrast imaging, reconstruction and tomography with hard synchrotron radiation: principles, implementation and applications*. PhD thesis, VUB (Faculteit Toegepaste Wetenschappen, Vakgroep Natuurkunde en Fotonica), 1999.
- [40] P. Cloetens, R. Barrett, J. Baruchel, J.-P. Guigay, and M. Schlenker. Phase objects in synchrotron radiation hard X-ray imaging. *Journal of Physics D: Applied Physics*, 29(1):133, 1996.

- [41] P. Cloetens, W. Ludwig, J. Baruchel, J.-P. Guigay, P. Pernot-Rejmánková, M. Salomé-Pateyron, M. Schlenker, J.-Y. Buffière, E. Maire, and G. Peix. Hard X-ray phase imaging using simple propagation of a coherent synchrotron radiation beam. *Journal of Physics D: Applied Physics*, 32(10A):A145, 1999.
- [42] P. Cloetens, W. Ludwig, J. Baruchel, D. Van Dyck, J. Van Landuyt, J. Guigay, and M. Schlenker. Holotomography: Quantitative phase tomography with micrometer resolution using hard synchrotron radiation X-rays. *Applied Physics Letters*, 75(19):2912–2914, 1999.
- [43] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011.
- [44] A. M. Cormack. Representation of a function by its line integrals, with some radiological applications. *Journal of Applied Physics*, 34(9):2722–2727, 1963.
- [45] A. M. Cormack. Representation of a function by its line integrals, with some radiological applications. ii. *Journal of Applied Physics*, 35(10):2908–2913, 1964.
- [46] J. M. Cowley. *Diffraction Physics*. North Holland Amsterdam, 1995.
- [47] V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin. Non-linear iterative phase retrieval based on frechet derivative. *Optics express*, 19(23):22809–22819, 2011.
- [48] V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin. Nonlinear phase retrieval using projection operator and iterative wavelet thresholding. *IEEE Signal Processing Letters*, 19(9):579–582, 2012.
- [49] V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin. Nonlinear approaches for the single-distance phase retrieval problem involving regularizations with sparsity constraints. *Applied Optics*, 52(17):3977–3986, 2013.
- [50] A. De Cezaro, J. Baumeister, and A. Leitao. Modified iterated Tikhonov methods for solving systems of nonlinear ill-posed equations. *Inverse Problems and Imaging*, 5(1):1–17, 2011.
- [51] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens. An iterative maximum-likelihood polychromatic algorithm for CT. *IEEE Transactions on Medical Imaging*, 20(10):999–1008, 2001.
- [52] M. Defrise, C. Vanhove, and X. Liu. An algorithm for total variation regularization in high-dimensional linear problems. *Inverse Problems*, 27(6):065002, 2011.
- [53] V. Dicken. A new approach towards simultaneous activity and attenuation reconstruction in emission tomography. *Inverse Problems*, 15(4):931, 1999.
- [54] C. Eberl, F. Döring, T. Liese, F. Schlenkrich, B. Roos, M. Hahn, T. Hoinkes, A. Rauschenbeutel, M. Osterhoff, T. Salditt, et al. Fabrication of laser deposited high-quality multilayer zone plates for hard X-ray nanofocusing. *Applied surface science*, 307:638–644, 2014.

- [55] E. Eggl, S. Schleede, M. Bech, K. Achterhold, R. Loewen, R. D. Ruth, and F. Pfeiffer. X-ray phase-contrast tomography with a compact laser-driven synchrotron source. *Proceedings of the National Academy of Sciences*, page 201500938, 2015.
- [56] T. Ekeberg, M. Svenda, C. Abergel, F. R. Maia, V. Seltzer, J.-M. Claverie, M. Hantke, O. Jönsson, C. Nettelblad, G. Van Der Schot, et al. Three-dimensional reconstruction of the giant mimivirus particle with an X-ray free-electron laser. *Physical review letters*, 114(9):098102, 2015.
- [57] T. Elfving, P. C. Hansen, and T. Nikazad. Semi-convergence properties of Kaczmarz’s method. *Inverse Problems*, 30(5):055007, 2014.
- [58] T. Elfving and T. Nikazad. Properties of a class of block-iterative methods. *Inverse Problems*, 25(11):115011, 2009.
- [59] V. Elser. Phase retrieval by iterated projections. *JOSA A*, 20(1):40–55, 2003.
- [60] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*. Springer, 1996.
- [61] C. L. Epstein. How well does the finite Fourier transform approximate the Fourier transform? *Communications on Pure and Applied Mathematics*, 58(10):1421–1435, 2005.
- [62] H. Erickson and A. Klug. Measurement and compensation of defocusing and aberrations by Fourier processing of electron micrographs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 261(837):105–118, 1971.
- [63] H. Faulkner and J. Rodenburg. Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm. *Physical review letters*, 93(2):023903, 2004.
- [64] L. Feldkamp, L. Davis, and J. Kress. Practical cone-beam algorithm. *JOSA A*, 1(6):612–619, 1984.
- [65] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21(15):2758–2769, 1982.
- [66] J. R. Fienup. Phase retrieval algorithms: a personal tour. *Applied Optics*, 52(1):45–56, 2013.
- [67] H.-W. Fink, W. Stocker, and H. Schmid. Holography with low-energy electrons. *Physical review letters*, 65(10):1204, 1990.
- [68] G. B. Folland and A. Sitaram. The uncertainty principle: a mathematical survey. *Journal of Fourier Analysis and Applications*, 3(3):207–238, 1997.
- [69] K. Fourmont. Non-equispaced fast Fourier transforms with applications to tomography. *Journal of Fourier Analysis and Applications*, 9(5):431–450, 2003.
- [70] D. Gabay. Chapter IX: Applications of the method of multipliers to variational inequalities. In *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, pages 299 – 331. Elsevier, 1983.

- [71] D. Gabor. Nobel lecture: Holography, 1948-1971. *The Nobel Prize in Physics 1971*, 1972.
- [72] D. Gabor et al. A new microscopic principle. *Nature*, 161(4098):777–778, 1948.
- [73] D. Gabor et al. Microscopy by reconstructed wave-fronts. *Proceedings of the Royal Society of London A*, 197(1051):454–487, 1949.
- [74] R. Gerchberg and W. Saxton. Phase retrieval by iterated projections. *Optik*, 35:237, 1972.
- [75] K. Giewekemeyer. *A study on new approaches in coherent X-ray microscopy of biological specimens*, volume 5 of *Göttingen Series in X-ray Physics*. Universitätsverlag Göttingen, 2011.
- [76] K. Giewekemeyer, S. Krüger, S. Kalbfleisch, M. Bartels, C. Beta, and T. Salditt. X-ray propagation microscopy of biological cells using waveguides as a quasipoint source. *Physical Review A*, 83(2):023804, 2011.
- [77] R. Gordon, R. Bender, and G. T. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of theoretical Biology*, 29(3):471–481, 1970.
- [78] H. Guan and R. Gordon. A projection access order for speedy convergence of ART (algebraic reconstruction technique): a multilevel scheme for computed tomography. *Physics in Medicine and Biology*, 39(11):2005, 1994.
- [79] J. Guigay. Fourier-transform analysis of Fresnel diffraction patterns and in-line holograms. *Optik*, 49(1):121–125, 1977.
- [80] M. Guizar-Sicairos, J. J. Boon, K. Mader, A. Diaz, A. Menzel, and O. Bunk. Quantitative interior X-ray nanotomography by a hybrid imaging technique. *Optica*, 2(3):259–266, 2015.
- [81] T. E. Gureyev, T. J. Davis, A. Pogany, S. C. Mayo, and S. W. Wilkins. Optical phase retrieval by use of first Born-and Rytov-type approximations. *Applied Optics*, 43(12):2418–2430, 2004.
- [82] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, 13(49-52):28, 1902.
- [83] J. Hagemann, A.-L. Robisch, D. Luke, C. Homann, T. Hohage, P. Cloetens, H. Suhonen, and T. Salditt. Reconstruction of wave front and object for inline holography from a set of detection planes. *Optics express*, 22(10):11552–11569, 2014.
- [84] J. Hagemann, A.-L. Robisch, M. Osterhoff, and T. Salditt. Probe reconstruction for holographic X-ray imaging. *Journal of synchrotron radiation*, 24(2):498–505, 2017.
- [85] J. Hagemann and T. Salditt. The fluence–resolution relationship in holographic and coherent diffractive imaging. *Journal of applied crystallography*, 50(2):531–538, 2017.
- [86] J. Hagemann and T. Salditt. Reconstructing mode mixtures in the optical near-field. *Optics express*, 25(13):13973–13989, 2017.

- [87] J. Hagemann, M. Töpperwien, and T. Salditt. Phase retrieval for near-field X-ray imaging beyond linearisation or compact support. *Applied Physics Letters*, 113(4):041109, 2018.
- [88] P. C. Hansen. *Discrete inverse problems: insight and algorithms*. SIAM, 2010.
- [89] A. Harootunian, E. Betzig, M. Isaacson, and A. Lewis. Super-resolution fluorescence near-field scanning optical microscopy. *Applied Physics Letters*, 49(11):674–676, 1986.
- [90] V. Havin and B. Jöricke. *The Uncertainty Principle in Harmonic Analysis*. Springer, 1994.
- [91] S. Helgason. The Radon transform on Euclidean spaces, compact two-point homogeneous spaces and Grassmann manifolds. *Acta Mathematica*, 113(1):153–180, 1965.
- [92] B. L. Henke, E. M. Gullikson, and J. C. Davis. X-ray interactions: Photoabsorption, scattering, transmission, and reflection at E=50–30 000 eV, Z=1–92. *Atomic data and nuclear data tables*, 54(2):181–342, 1993.
- [93] A. Hertle. Continuity of the Radon transform and its inverse on Euclidean space. *Mathematische Zeitschrift*, 184(2):165–192, 1983.
- [94] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2002.
- [95] S. Hoffmann-Urlaub, P. Höhne, M. Kanbach, and T. Salditt. Advances in fabrication of X-ray waveguides. *Microelectronic Engineering*, 164:135–138, 2016.
- [96] R. Hofmann, J. Moosmann, and T. Baumbach. Criticality in single-distance phase retrieval. *Optics express*, 19(27):25881–25890, 2011.
- [97] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, 13(5):1279, 1997.
- [98] T. Hohage and F. Werner. Iteratively regularized Newton-type methods for general data misfit functionals and applications to Poisson data. *Numerische Mathematik*, 123(4):745–779, 2013.
- [99] T. Hohage and F. Werner. Inverse problems with poisson data: statistical regularization theory, applications and algorithms. *Inverse Problems*, 32(9):093001, 2016.
- [100] C. Homann. *Phase retrieval problems in X-ray physics. From modeling to efficient algorithms*, volume 16 of *Göttingen Series in X-ray Physics*. Universitätsverlag Göttingen, 2015.
- [101] C. Homann, T. Hohage, J. Hagemann, A.-L. Robisch, and T. Salditt. Validity of the empty-beam correction in near-field imaging. *Physical Review A*, 91(1):013821, 2015.
- [102] L. Hörmander. *The analysis of linear partial differential operators I*. Springer, 2003.
- [103] G. N. Hounsfield. Computerized transverse axial scanning (tomography): Part 1. description of system. *The British journal of radiology*, 46(552):1016–1022, 1973.

- [104] T. Humphries and A. Faidani. Segmentation-free quasi-Newton method for polyenergetic CT reconstruction. In *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pages 1–5. IEEE, 2014.
- [105] V. Isakov. *Inverse problems for partial differential equations*, volume 127 of *Applied Mathematical Sciences*. Springer, third edition, 2017.
- [106] P. Jaming. Nazarov’s uncertainty principles in higher dimension. *Journal of Approximation Theory*, 149(1):30–41, 2007.
- [107] G. Johansen. Gamma-ray tomography. In *Industrial Tomography*, pages 197–222. Elsevier, 2015.
- [108] P. Jonas and A. Louis. Phase contrast tomography using holographic measurements. *Inverse Problems*, 20(1):75, 2003.
- [109] S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin International de l’Academie Polonaise des Sciences et des Lettres*, 35:355–357, 1937.
- [110] A. C. Kak and M. Slaney. *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics, 2001.
- [111] S. Kalbfleisch, H. Neubauer, S. Krüger, M. Bartels, M. Osterhoff, K. Giewekemeyer, B. Hartmann, M. Sprung, and T. Salditt. The Göttingen holography endstation of beamline P10 at PETRA III/DESY. *AIP Conference Proceedings*, 1365(1):96–99, 2011.
- [112] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative regularization methods for nonlinear ill-posed problems*. Walter de Gruyter, 2008.
- [113] J. Keiner, S. Kunis, and D. Potts. Using NFFT 3—a software library for various nonequid-spaced fast Fourier transforms. *ACM Transactions on Mathematical Software (TOMS)*, 36(4):19, 2009.
- [114] S. Kindermann. Convergence analysis of minimization-based noise level-free parameter choice rules for linear ill-posed problems. *Electronic Transactions on Numerical Analysis*, 38:233–257, 2011.
- [115] S. Kindermann and A. Leitao. Convergence rates for Kaczmarz-type regularization methods. *Inverse Problems and Imaging*, 8(1):149–172, 2014.
- [116] M. V. Klibanov. On the recovery of a 2-d function from the modulus of its Fourier transform. *Journal of mathematical analysis and applications*, 323(2):818–843, 2006.
- [117] M. V. Klibanov. Phaseless inverse scattering problems in three dimensions. *SIAM Journal on Applied Mathematics*, 74(2):392–410, 2014.
- [118] M. V. Klibanov, P. E. Sacks, and A. V. Tikhonravov. The phase retrieval problem. *Inverse problems*, 11(1):1, 1995.
- [119] A. Kostenko, K. J. Batenburg, A. King, S. E. Offerman, and L. J. van Vliet. Total variation minimization approach in in-line X-ray phase-contrast tomography. *Optics express*, 21(10):12185–12196, 2013.

- [120] M. Krenkel. *Cone-beam X-ray phase-contrast tomography for the observation of single cells in whole organs*, volume 17 of *Göttingen Series in X-ray Physics*. Universitätsverlag Göttingen, 2015.
- [121] M. Krenkel, M. Toepperwien, F. Alves, and T. Salditt. Three-dimensional single-cell imaging with X-ray waveguides in the holographic regime. *Acta Crystallographica Section A: Foundations and Advances*, 73(4):282–292, 2017.
- [122] M. Krenkel, M. Töpperwien, M. Bartels, P. Lingor, D. Schild, and T. Salditt. X-ray phase contrast tomography from whole organ down to single cells. In *Developments in X-ray Tomography IX*, volume 9212, page 92120R. International Society for Optics and Photonics, 2014.
- [123] S. Lagomarsino, A. Cedola, P. Cloetens, S. Di Fonzo, W. Jark, G. Soullie, and C. Riekkel. Phase contrast hard X-ray microscopy with submicron resolution. *Applied physics letters*, 71(18):2557–2559, 1997.
- [124] M. Langer, P. Cloetens, J.-P. Guigay, and F. Peyrin. Quantitative comparison of direct phase retrieval algorithms in in-line phase tomography. *Medical physics*, 35(10):4556–4566, 2008.
- [125] M. Langer, P. Cloetens, A. Pacureanu, and F. Peyrin. X-ray in-line phase tomography of multimaterial objects. *Optics letters*, 37(11):2151–2153, 2012.
- [126] T. Latychevskaia and H.-W. Fink. Practical algorithms for simulation and reconstruction of digital in-line holograms. *Applied Optics*, 54(9):2424–2434, 2015.
- [127] T. Latychevskaia, J.-N. Longchamp, and H.-W. Fink. When holography meets coherent diffraction imaging. *Optics express*, 20(27):28871–28892, 2012.
- [128] B. Leshem, R. Xu, Y. Dallal, J. Miao, B. Nadler, D. Oron, N. Dudovich, and O. Raz. Direct single-shot phase retrieval from the diffraction pattern of separated objects. *Nature communications*, 7:10820, 2016.
- [129] H. Lichte and M. Lehmann. Electron holography—basics and applications. *Reports on Progress in Physics*, 71(1):016102, 2007.
- [130] M. Lieblich, T. Blu, and M. Unser. Fresnelets: new multiresolution wavelet bases for digital holography. *IEEE Transactions on image processing*, 12(1):29–43, 2003.
- [131] A. Lipson, S. G. Lipson, and H. Lipson. *Optical physics*. Cambridge University Press, 2010.
- [132] J.-N. Longchamp, S. Rauschenbach, S. Abb, C. Escher, T. Latychevskaia, K. Kern, and H.-W. Fink. Imaging proteins at the single-molecule level. *Proceedings of the National Academy of Sciences*, 114(7):1474–1479, 2017.
- [133] A. K. Louis. Exact cone beam reconstruction formulae for functions and their gradients for spherical and flat detectors. *Inverse Problems*, 32(11):115005, 2016.

- [134] D. Ludwig. The Radon transform on Euclidean space. *Communications on Pure and Applied Mathematics*, 19(1):49–81, 1966.
- [135] D. R. Luke. Relaxed averaged alternating reflections for diffraction imaging. *Inverse Problems*, 21(1):37, 2005.
- [136] D. R. Luke. Phase retrieval, what’s new. *SIAG/OPT Views and News*, 25(1):1–5, 2017.
- [137] A. M. Maiden and J. M. Rodenburg. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 109(10):1256–1262, 2009.
- [138] S. Marchesini. Invited article: A unified evaluation of iterative projection algorithms for phase retrieval. *Review of Scientific Instruments*, 78(1):011301, 2007.
- [139] S. Maretzke. Regularized Newton methods for simultaneous Radon inversion and phase retrieval in phase contrast tomography. *arXiv preprint*, page arXiv:1502.05073, 2015.
- [140] S. Maretzke. A uniqueness result for propagation-based phase contrast imaging from a single measurement. *Inverse Problems*, 31:065003, 2015.
- [141] S. Maretzke. Generalized SART-methods for tomographic imaging. *arXiv preprint*, page arXiv:1803.04726, 2018. submitted to *Inverse Problems*.
- [142] S. Maretzke. Locality estimates for Fresnel-wave-propagation and stability of near-field X-ray propagation imaging with finite detectors. *Inverse Problems*, 34(12):124004, 2018.
- [143] S. Maretzke, M. Bartels, M. Krenkel, T. Salditt, and T. Hohage. Regularized Newton methods for X-ray phase contrast and general imaging problems. *Optics express*, 24(6):6490–6506, 2016.
- [144] S. Maretzke and T. Hohage. Stability estimates for linearized near-field phase retrieval in X-ray phase contrast imaging. *SIAM Journal on Applied Mathematics*, 77(2):384–408, 2017.
- [145] S. Mayo, P. Miller, S. Wilkins, T. Davis, D. Gao, T. Gureyev, D. Paganin, D. Parry, A. Pogany, and A. Stevenson. Quantitative X-ray projection microscopy: phase-contrast and multi-spectral imaging. *Journal of microscopy*, 207(2):79–96, 2002.
- [146] J. Miao, P. Charalambous, J. Kirz, and D. Sayre. Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. *Nature*, 400(6742):342–344, 1999.
- [147] J. Miao, T. Ishikawa, I. K. Robinson, and M. M. Murnane. Beyond crystallography: Diffractive imaging using coherent X-ray light sources. *Science*, 348(6234):530–535, 2015.
- [148] P. Midgley and M. Weyland. 3d electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography. *Ultramicroscopy*, 96(3-4):413–431, 2003.
- [149] R. Millane. Phase retrieval in crystallography and optics. *JOSA A*, 7(3):394–411, 1990.
- [150] J. Moosmann, R. Hofmann, and T. Baumbach. Single-distance phase retrieval at large phase shifts. *Optics express*, 19(13):12066–12073, 2011.

- [151] I. Mori, Y. Machida, M. Osanai, and K. Iinuma. Photon starvation artifacts of X-ray CT: their true cause and a solution. *Radiological physics and technology*, 6(1):130–141, 2013.
- [152] K. Morishima, M. Kuno, A. Nishio, N. Kitagawa, Y. Manabe, M. Moto, F. Takasaki, H. Fujii, K. Satoh, H. Kodama, et al. Discovery of a big void in khufu’s pyramid by observation of cosmic-ray muons. *Nature*, 552(7685):386, 2017.
- [153] V. A. Morozov. On the solution of functional equations by the method of regularization. *Soviet Mathematics Doklady*, 7(1):414–417, 1966.
- [154] J. L. Mueller and S. Siltanen. *Linear and nonlinear inverse problems with practical applications*. SIAM, 2012.
- [155] F. Natterer. *The mathematics of computerized tomography*, volume 32 of *Classics of Applied Mathematics*. Society for Industrial and Applied Mathematics, 2001.
- [156] F. L. Nazarov. Local estimates for exponential polynomials and their applications to inequalities of the uncertainty principle type. *Algebra i Analiz*, 5(4):3–66, 1993.
- [157] K. A. Nugent. X-ray noninterferometric phase imaging: a unified picture. *JOSA A*, 24(2):536–547, 2007.
- [158] K. A. Nugent. Coherent methods in the X-ray sciences. *Advances in Physics*, 59(1):1–99, 2010.
- [159] O. Öktem. Mathematics of electron tomography. In *Handbook of Mathematical Methods in Imaging*, pages 937–1031. Springer, 2015.
- [160] D. Paganin. *Coherent X-ray optics*. Oxford University Press, 2006.
- [161] D. Paganin, S. Mayo, T. E. Gureyev, P. R. Miller, and S. W. Wilkins. Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object. *Journal of microscopy*, 206(1):33–40, 2002.
- [162] D. Paganin and K. A. Nugent. Noninterferometric phase imaging with partially coherent light. *Physical review letters*, 80(12):2586, 1998.
- [163] X. Pan, E. Y. Sidky, and M. Vannier. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse Problems*, 25(12):123009, 2009.
- [164] N. Parikh, S. Boyd, et al. Proximal algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239, 2014.
- [165] A. Pein, S. Loock, G. Plonka, and T. Salditt. Using sparsity information for iterative phase retrieval in X-ray propagation imaging. *Optics express*, 24(8):8332–8343, 2016.
- [166] A. Pogany, D. Gao, and S. Wilkins. Contrast and resolution in imaging with a microfocus X-ray source. *Review of Scientific Instruments*, 68(7):2774–2782, 1997.
- [167] H. Quiney. Coherent diffractive imaging using short wavelength light sources. *Journal of modern optics*, 57(13):1109–1149, 2010.

- [168] J. Radon. über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Classic papers in modern diagnostic radiology*, page 5, 1917.
- [169] O. Raz, B. Leshem, J. Miao, B. Nadler, D. Oron, and N. Dudovich. Direct phase retrieval in double blind Fourier holography. *Optics express*, 22(21):24935–24950, 2014.
- [170] A. Robisch, K. Kröger, A. Rack, and T. Salditt. Near-field ptychography using lateral and longitudinal shifts. *New Journal of Physics*, 17(7):073033, 2015.
- [171] A.-L. Robisch. *Phase retrieval for object and probe in the optical near-field*, volume 16 of *Göttingen Series in X-ray Physics*. Universitätsverlag Göttingen, 2016.
- [172] J. Rodenburg, A. Hurst, A. Cullis, B. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson. Hard-X-ray lensless imaging of extended objects. *Physical review letters*, 98(3):034801, 2007.
- [173] W. Röntgen. Über eine neue Art von Strahlen. *Sitzungsberichte der Physikalisch-medizinischen Gesellschaft zu Würzburg*, 132, 1895.
- [174] A. Ruhlandt. *Time-resolved X-ray phase-contrast tomography*. eDiss Uni Göttingen, 2018.
- [175] A. Ruhlandt, M. Krenkel, M. Bartels, and T. Salditt. Three-dimensional phase retrieval in propagation-based phase-contrast imaging. *Physical Review A*, 89(3):033847, 2014.
- [176] A. Ruhlandt and T. Salditt. Three-dimensional propagation in near-field tomographic X-ray phase retrieval. *Acta Crystallographica Section A: Foundations and Advances*, 72(2):215–221, 2016.
- [177] A. Ruhlandt, M. Töpperwien, M. Krenkel, R. Mokso, and T. Salditt. Four dimensional material movies: High speed phase-contrast tomography by backprojection along dynamically curved paths. *Scientific reports*, 7(1):6487, 2017.
- [178] T. Salditt, M. Osterhoff, M. Krenkel, R. N. Wilke, M. Priebe, M. Bartels, S. Kalbfleisch, and M. Sprung. Compound focusing mirror and X-ray waveguide optics for coherent imaging and nano-diffraction. *Journal of synchrotron radiation*, 22(4):867–878, 2015.
- [179] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*. Springer, 2009.
- [180] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization methods in Banach spaces*. Walter de Gruyter, 2012.
- [181] S. Setzer. Operator splittings, Bregman methods and frame shrinkage in image processing. *International Journal of Computer Vision*, 92(3):265–280, 2011.
- [182] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev. Phase retrieval with application to optical imaging: a contemporary overview. *IEEE Signal Processing Magazine*, 32(3):87–109, 2015.
- [183] E. Y. Sidky, J. H. Jørgensen, and X. Pan. Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle–Pock algorithm. *Physics in Medicine and Biology*, 57(10):3065, 2012.

- [184] E. Y. Sidky and X. Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in Medicine and Biology*, 53(17):4777, 2008.
- [185] D. Slepian. Some comments on Fourier analysis, uncertainty and modeling. *SIAM review*, 25(3):379–393, 1983.
- [186] D. Slepian and E. Sonnenblick. Eigenvalues associated with prolate spheroidal wave functions of zero order. *Bell System Technical Journal*, 44(8):1745–1759, 1965.
- [187] A. Snigirev, I. Snigireva, V. Kohn, S. Kuznetsov, and I. Schelokov. On the possibilities of X-ray phase contrast microimaging by coherent high-energy synchrotron radiation. *Review of Scientific Instruments*, 66(12):5486–5492, 1995.
- [188] M. Stockmar, P. Cloetens, I. Zanette, B. Enders, M. Dierolf, F. Pfeiffer, and P. Thibault. Near-field ptychography: phase retrieval for inline holography using a structured illumination. *Scientific reports*, 3, 2013.
- [189] P. F. Tavares, S. C. Leemann, M. Sjöström, and Å. Andersson. The MAX IV storage ring project. *Journal of synchrotron radiation*, 21(5):862–877, 2014.
- [190] M. C. Teich and B. Saleh. *Fundamentals of photonics*. John Wiley & Sons, 1991.
- [191] P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer. High-resolution scanning X-ray diffraction microscopy. *Science*, 321(5887):379–382, 2008.
- [192] P. Thibault and A. Menzel. Reconstructing state mixtures from diffraction measurements. *Nature*, 494(7435):68–71, 2013.
- [193] M. Töpperwien, F. van der Meer, C. Stadelmann, and T. Salditt. Three-dimensional virtual histology of human cerebellum by X-ray phase-contrast tomography. *Proceedings of the National Academy of Sciences*, page 201801678, 2018.
- [194] L. Turner, B. Dhal, J. Hayes, A. Mancuso, K. Nugent, D. Paterson, R. Scholten, C. Tran, and A. Peele. X-ray phase imaging: Demonstration of extended conditions for homogeneous objects. *Optics express*, 12(13):2960–2965, 2004.
- [195] M. Ulbrich. Optimization methods in Banach spaces. In *Optimization with PDE Constraints*, pages 97–156. Springer, 2009.
- [196] M. Unser, A. Aldroubi, and M. Eden. On the asymptotic convergence of B-spline wavelets to Gabor functions. *IEEE Transactions on Information Theory*, 38(2):864–872, 1992.
- [197] M. Unser, A. Aldroubi, M. Eden, et al. Fast B-spline transforms for continuous image representation and interpolation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(3):277–285, 1991.
- [198] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K. J. Batenburg, and J. Sijbers. Fast and flexible X-ray tomography using the ASTRA toolbox. *Optics express*, 24(22):25129–25147, 2016.

- [199] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers. The ASTRA toolbox: A platform for advanced algorithm development in electron tomography. *Ultramicroscopy*, 157:35–47, 2015.
- [200] M. van Heel. Similarity measures between images. *Ultramicroscopy*, 21(1):95–100, 1987.
- [201] M. van Heel and M. Schatz. Fourier shell correlation threshold criteria. *Journal of structural biology*, 151(3):250–262, 2005.
- [202] P. B. Venuto and E. T. Habib Jr. *Fluid catalytic cracking with zeolite catalysts*. Marcel Dekker Inc., 1979.
- [203] D. G. Voelz and M. C. Roggemann. Digital simulation of scalar optical diffraction: revisiting chirp function sampling criteria and consequences. *Applied Optics*, 48(32):6132–6142, 2009.
- [204] M. von Laue. Nobel lecture: Concerning the detection of X-ray interferences. *The Nobel Prize in Physics 1914*, 1920.
- [205] R. Wade. A brief look at imaging and contrast transfer. *Ultramicroscopy*, 46(1):145–156, 1992.
- [206] F. Werner. *Inverse problems with Poisson data: Tikhonov-type regularization and iteratively regularized Newton methods*. Der Andere Verlag, 2012.
- [207] S. Wilkins, T. Gureyev, D. Gao, A. Pogany, and A. Stevenson. Phase-contrast imaging using polychromatic hard X-rays. *Nature*, 384(6607):335–338, 1996.
- [208] G. Williams, H. Quiney, B. Dhal, C. Tran, K. A. Nugent, A. Peele, D. Paterson, and M. De Jonge. Fresnel coherent diffractive imaging. *Physical review letters*, 97(2):025506, 2006.
- [209] F. Xu and K. Mueller. A comparative study of popular interpolation and integration methods for use in computed tomography. In *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro*, pages 1252–1255, 2006.
- [210] R. Xu, C.-C. Chen, L. Wu, M. C. Scott, W. Theis, C. Ophus, M. Bartels, Y. Yang, H. Ramezani-Dakhel, M. R. Sawaya, H. Heinz, L. D. Marks, P. Ercius, and J. Miao. Three-dimensional coordinates of individual atoms in materials revealed by electron tomography. *Nature Materials*, 14:1099–1103, 2015.
- [211] J. Yang, H. Yu, M. Jiang, and G. Wang. High-order total variation minimization for interior tomography. *Inverse Problems*, 26(3):035013, 2010.
- [212] S. Zabler, P. Cloetens, J.-P. Guigay, J. Baruchel, and M. Schlenker. Optimization of phase contrast imaging using hard X-rays. *Review of Scientific Instruments*, 76(7):073705, 2005.
- [213] E. Zeidler. *Nonlinear functional analysis and its applications. I*. Springer-Verlag, New York, 1986. Fixed-point theorems, Translated from the German by Peter R. Wadsack.

List of own publications

S. Maretzke. Generalized SART-methods for tomographic imaging. *Inverse problems* (in review), arXiv:1803.04726.

S. Maretzke. Locality estimates for Fresnel-wave-propagation and stability of near-field X-ray propagation imaging with finite detectors. *Inverse problems*, 34(12):124004, 2018.

G. Zickert and S. Maretzke. Cryogenic electron tomography reconstructions from phaseless data. *Inverse problems*, 34(12):124001, 2018.

T. van Leeuwen, S. Maretzke, and K. J. Batenburg. Automatic alignment for three-dimensional tomographic reconstruction. *Inverse Problems*, 34(2):024004, 2018.

S. Maretzke and T. Hohage. Stability estimates for linearized near-field phase retrieval in X-ray phase contrast imaging. *SIAM Journal on Applied Mathematics*, 77(2):384–408, 2017.

S. Maretzke, M. Bartels, M. Krenkel, T. Salditt, and T. Hohage. Regularized Newton methods for X-ray phase contrast and general imaging problems. *Optics Express*, 24(6):6490–6506, 2016.

S. Maretzke. A uniqueness result for propagation-based phase contrast imaging from a single measurement. *Inverse Problems*, 31:065003, 2015.

S. Maretzke, B. Hof, and M. Avila. Transient growth in linearly stable Taylor–Couette flows. *Journal of Fluid Mechanics*, 742:254–290, 2014.

Statement of author’s contributions

The four journal articles reprinted in chapters 8 to 11 fulfill the requirements to be considered in a “kumulative Dissertation” according to the regulations of the doctoral program “Mathematical Sciences” in the Georg-August University School of Science (GAUSS) Göttingen:

- (1) S. Maretzke and T. Hohage. Stability estimates for linearized near-field phase retrieval in X-ray phase contrast imaging. *SIAM Journal on Applied Mathematics*, 77(2):384–408, 2017.
- (2) S. Maretzke. Locality estimates for Fresnel-wave-propagation and stability of near-field X-ray propagation imaging with finite detectors. *Inverse problems*, 34(12):124004, 2018.
- (3) S. Maretzke, M. Bartels, M. Krenkel, T. Salditt, and T. Hohage. Regularized Newton methods for X-ray phase contrast and general imaging problems. *Optics Express*, 24(6):6490–6506, 2016.
- (4) S. Maretzke. Generalized SART-methods for tomographic imaging. *Inverse problems* (in review), arXiv:1803.04726.

The contributions of the thesis’ author (SM) to the articles (1) and (3), which are joint work with additional co-authors, are detailed in the following.

The idea of analyzing stability of image-reconstruction in XPCI, which initiated the study in article (1), is due to TH. The manuscript was almost completely written by SM, except for short passages in the background- and conclusions-sections (§8.2 and §8.6) that were contributed by TH. The ideas for most of proofs were likewise developed by SM, supported by fruitful discussions with TH, whereas TH drew the crucial connection of the analyzed stability problem to the eigenvalue-computations in [185, 186] exploited in §8.3.3 and §8.3.4. Detailed mathematical proofs as well as the numerical computations in §8.3.5 were worked out by SM alone. Conclusions from the stability results were jointly discussed by both authors.

The concept of the article (3) was proposed by TS, who also wrote the introduction (§10.1) up to some amendments by SM as well as parts of the conclusions-section §10.5. The rest of the manuscript was written by SM, supported by suggestions from the other authors. The principal ideas of using regularized Newton algorithms in XPCI and Newton-Kaczmarz methods in XPCT stem from TH. TH also helped SM in the description of the underlying theory in §10.2. The experiments, that provided the XPCI- and XPCT-data sets considered in the article, were conducted jointly by MB, MK and TS, who also contributed information on the experimental background. The considered Newton-type image-reconstruction methods were implemented by SM, who also designed and ran the numerical test cases and visualized the results. Conclusions from the results were jointly discussed by all authors.

Danksagung

Während meiner Promotionszeit und der Erarbeitung der vorliegenden Dissertation war ich im Angesicht dieses Mammutprojekts mehr als nur einmal kurz davor abzubrechen. Dass ich mich dennoch immer wieder aufraffen und neu für meine Arbeit begeistern konnte, verdanke ich zum großen Teil den wertvollen Menschen um mich herum. Auch wenn ich aus Sorge jemand wichtigen zu vergessen hier niemanden namentlich hervorheben möchte, bin ich dankbar für jede Unterstützung, die ich durch meine Betreuer, Kollegen, Freunde und Familie erfahren habe – sei es in Form von fachlicher Expertise, geteilten Forschungsergebnissen, inspirierenden Diskussionen, aufmunternden Worten oder einfach nur ein wenig Zerstreuung. Ohne diese Hilfe hätte ich meine Promotion wohl nie zum Abschluss gebracht. Dankeschön! ☺

Funding: Das dieser Arbeit zugrundeliegende Promotionsprojekt wurde im Rahmen des Projekts C02 im Sonderforschungsbereichs 755 “Photonische Abbildungen auf der Nanometerskala” durch die Deutsche Forschungsgemeinschaft (DFG) finanziert. Für die finanzielle Unterstützung möchte ich mich herzlich bedanken.