Semiparametric Estimation of Drift, Rotation and Scaling in Sparse Sequential Dynamic Imaging:

Asymptotic theory and an application in nanoscale fluorescence microscopy

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"
der Georg-August-Universität Göttingen

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

vorgelegt von
Anne Hobert, geb. Hein
aus München

Göttingen, 2018

Betreuungsausschuss:

Prof. Dr. Axel Munk
Institut für Mathematische Stochastik, Universität Göttingen

Prof. Dr. Tatyana Krivobokova
Institut für Mathematische Stochastik, Universität Göttingen

Mitglieder der Prüfungskommission:

Referent:
Prof. Dr. Axel Munk
Institut für Mathematische Stochastik, Universität Göttingen

Korreferentin:
Prof. Dr. Tatyana Krivobokova
Institut für Mathematische Stochastik, Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Dorothea Bahns
Mathematisches Institut, Universität Göttingen

PD Dr. Alexander Egner
Laser-Laboratorium Göttingen

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

Prof. Dr. Anja Sturm
Institut für Mathematische Stochastik, Universität Göttingen

Tag der mündlichen Prüfung: 29. Januar 2019

Acknowledgments

After three years of working on this thesis I want to take this opportunity to express my deepest gratitude to the many people whom I am indebted to for making it possible that I now complete my doctoral project. I am incredibly grateful and feel tremendously lucky that you offered me your great company, support and guidance during this time!

- First of all I would like to thank my supervisor **Prof. Dr. Axel Munk** for giving me the chance to work in his group on this challenging topic. I greatly appreciate his assistance in the development of this work and the open and encouraging atmosphere he creates.

- I thank **Prof. Dr. Tatyana Krivobokova** for being my second reviewer.

- Thank you, **Prof. Dr. Dorothea Bahns**, **PD Dr. Alexander Egner**, **Prof. Dr. Gerlind Plonka-Hoch**, and **Prof. Dr. Anja Sturm** for being members of my examination committee.

- I gratefully acknowledge the financial support by the **German Research Foundation DFG** through subproject A06 of CRC 755.

- I am thankful to the members of the CRC 755 for our enjoyable collaboration. Special thanks go to **Dr. Oskar Laitenberger** for providing the SMS data, for his patient explanations of the experimental setup, and for giving his perspective from the applications side.

- I thank **Dr. Alexander Hartmann** for providing me with his materials and for his help even after he left academia.

- Advice given by **PD Dr. Timo Aspelmeier** and **Prof. Dr. Stephan Huckemann** has been a great help in finding the right model.

- I owe a great deal to **Dr. Benjamin Eltzner**, who spent a lot of time and effort discussing various aspects of this project with me, giving feedback and ideas on the modeling, the statistical theory and the implementation, and last but not least, proofreading this dissertation.

- I want to thank especially my colleagues and friends at the institute for making it such a pleasant place to work at. In particular, thank you **Carla Tameling** and **Dr. Katharina Proksch** (you are absolutely the best office mates I could have wished for), **Dr. Claudia**

König, **Dr. Merle Behr**, **Dr. Max Sommerfeld**, **Thomas Staudt**, **Laura Fee Schneider**, **Marcel Klatt** (vaccuna!), **Christian Boehm**, **Henning Höllwarth**, **Dr. Benjamin Eltzner** and **Robin Richter**!

- I thank my fellow Salseras and Salseros for their cordial welcome into the salsa community and for temporarily taking my mind off work. Special thanks to one of them, for finding the right words of motivation at exactly the right times.

- I am very grateful for the many friends who made and still make Göttingen my home. I want to mention especially **Luisa Huchting**, **Annelie von Trotha** and **Florian Berens**. I highly value your friendship and I am happy that you've accompanied me from the very beginning to this day and ongoing.

- Most of all, and more than I can express, I am grateful to my family, in particular my parents **Dr. Barbara Röbe-Oltmanns** and **Dr. Ekkehardt Hein**, my brothers **Fabian Hein** and **Jonathan Hein**, and my husband **Dr. Sebastian Hobert**. Without any of you, I would never have succeeded in finishing this PhD project. Infinite thanks to all of you for your moral support, encouragement, bluntness, honest advice, understanding and love.

Preface

Light microscopy is an important instrument in life sciences. Over the last two decades, superresolution fluorescence microscopy techniques have been established, breaking the Abbé diffraction barrier, which before had posed a resolution limitation for over a century. The fundamentally new idea of these approaches is to use optically switchable fluorophores in order to detect features within the resolution limit imposed by the diffraction barrier consecutively instead of simultaneously. However, the relatively long imaging times needed in many modern superresolution fluorescence microscopy techniques at the nanoscale, one of them being single marker switching (SMS) microscopy, come with their own drawbacks. The challenge lies in the correct alignment of long sequences of sparse but spatially and temporally highly resolved images. This alignment is necessary due to rigid motion of the displayed object of interest or its supporting area during the observation process. In this thesis, a semiparametric model for motion correction, including drift, rotation and scaling of the imaged specimen, is used to estimate the motion and correct for it, reconstructing thereby the true underlying structure of interest. This technique is also applicable in many other scenarios, where an aggregation of a collection of sparse images is employed to obtain a good reconstruction of the underlying structure, like, for example, in real time magnetic resonance imaging (MRI).

Further motivation and a more detailed description of the SMS imaging method are given in Chapter 1. In Chapter 2, a semiparametric model is developed and M-estimators for the parameters of the motion functions are derived, which are obtained by minimizing certain contrast functionals. The basic idea is to perform a two-step estimation, where the motion deformations are linearized by applying the Fourier-Mellin transform to the squared Fourier magnitudes of the obervations. This allows to estimate rotation and scaling in a first step, correct for it, and subsequently estimate translational drift. The main theoretical results, namely consistency as well as asymptotic normality of the motion parameter estimators are established in Chapter 3. Additionally, consistency of the final plug-in image estimator is obtained. The results of a simulation study and an application to real SMS microscopy data are presented in Chapter 4, demonstrating the practicability of this purely statistical method. It is shown to be competitive with state of the art calibration techniques which require to incorporate fiducial markers. Moreover, a simple bootstrap algorithm allows to quantify the precision of the motion estimate and visualize its effect on the final image estimation. A summary of the findings and outlook can be found in Chapter 5. We argue that purely statistical motion correction is even more robust than fiducial tracking rendering the latter superfluous in many applications. The proofs are presented separately in Chapter 6. Some auxiliary results are deferred to Appendix A

to avoid a distraction from the principle arguments. In Appendix B, well-known results from the literature, which are applied in the proofs, are collated for the readers' convenience. Appendix C holds additional figures with reconstruction results from our simulation study, which were excluded from the main text body in order to avoid lengthening it unnecessarily.
This thesis is an extension of previous work by Hartmann (2016) and constitutes a generalization of the developed method of pure drift estimation to more general motion types, namely any combination of drift, rotation and scaling. The theoretical results of the present document are joint work with Dr. Alexander Hartmann, who contributed equally to the demonstration of consistency. The elaboration of the proof of asymptotic normality, however, is an original result of the author of this dissertation. In addition, the derivation of the semiparametrical model has been revised, leading to a different approach which better represents the data acquisition process.
A publication together with the co-authors Dr. Alexander Hartmann, Dr. Benjamin Eltzner, Prof. Dr. Stephan Huckemann, Dr. Oskar Laitenberger, Dr. Claudia Geisler, PD Dr. Alexander Egner, and Prof. Dr. Axel Munk in a peer-reviewed journal, covering the main aspects of this thesis in a condensed format is in preparation. The programs and routines used in the application of the method to artificial and real data are based on code provided by Dr. Alexander Hartmann and have been modified and amended to fit the new model and the generalized setting.

Contents

List of Symbols

$\mathbb{N}$	set of positive integers
$\mathbb{Z}$	set of integers
$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of positive real numbers
$\mathbb{C}$	set of complex numbers
$\lvert\cdot\rvert$	absolute value
$\lVert\cdot\rVert$	Euclidean distance
$\langle\cdot,\cdot\rangle$	Euclidean inner product
$\mathrm{Ber}(p)$	Bernoulli random variable with success probability p
$\mathrm{PoiBin}\,(q)$	Poisson binomial random variable with probability vector $q = \left(q_j\right)_{j=1}^n$, for some $n \in \mathbb{N}$
$\mathrm{Poi}(\lambda)$	Poisson random variable with intensity λ
$\mathcal{N}(\mu, \Sigma)$	(Possibly multivariate) Gaussian random variable with expectation $\mu \in \mathbb{R}^d$ and covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$, for some $d \in \mathbb{N}$
B^C	the complement of the set B
$\mathcal{B}(X)$	Borel σ-field on the set X
$\Re(x)$	real part of a complex number $x \in \mathbb{C}$
$\Im(x)$	imaginary part of a complex number $x \in \mathbb{C}$
$\overline{x}$	complex conjugate of a complex number $x \in \mathbb{C}$
$\mathcal{O}, o$	Bachmann-Landau symbols
$\mathcal{O}_{\mathbb{P}}$	stochastic boundedness
$o_{\mathbb{P}}$	convergence to zero in probability
a.e.	almost everywhere, that is everywhere except on a set of measure 0
a.s.	almost surely, that is with probability one
i.i.d.	independent and identically distributed

CHAPTER 1

Introduction

Superresolution fluorescence microscopy is an important instrument for the investigation of the properties or the structure of biological molecules at the subcellular level. It enables the observation of active biological molecules at a resolution level down to 10-20 nm, giving rise to great advances in the understanding of signaling and transport processes within cells (see, e.g., Westphal et al., 2008; Berning et al., 2012; Jones et al., 2011; Huang et al., 2013). These imaging techniques have been developed and refined over the last two decades, overcoming the physical resolution limitation called the Abbé diffraction barrier, which before had posed a problem for all optical imaging methods for more than a century (see, e.g., Hell and Wichmann, 1994; Hell, 2007). The Abbé barrier describes the phenomenon that two features that are closer than a resolution limit of about 200 nm (approximately half the smallest wavelength of visible light) overlap and can not be distinguished (Abbe, 1873; Born and Wolf, 1999). The entirely new approach of superresolution imaging techniques is to register features within this resolution limit consecutively instead of simultaneously. This is achieved not by modifying the experimental setup, but by changing the appearance of the specimen over time. A variation of a fluorophore's ability to emit a fluorescence photon or of the properties of the emitted photon, like, for example, its color, allows for a much higher spatial resolution in fluorescence microscopy (Hell, 2009). The implementation of this approach in various methods (e.g., Hell, 2003; Betzig et al., 2006; Rust et al., 2006; Hess et al., 2006) has fundamentally enhanced the field of cell microscopy.

Two different categories of superresolution fluorescence microscopy techniques can be identified. The first group consists of deterministic imaging methods using a targeted mode. Here, fluorophores (markers) are switched on and off at predefined and precisely known coordinates. This group includes, among others, techniques such as *stimulated emission depletion* (STED) (Hell and Wichmann, 1994; Klar et al., 2000; Schmidt et al., 2008), *saturated patterned excitation microscopy* (SPEM) (Heintzmann et al., 2002), *saturated structured illumination microscopy* (SSIM) (Gustafsson, 2005), and *reversible saturable optical fluorescence transitions* (RESOLFT) (Hofmann et al., 2005; Hell, 2003). Because of the direct targeting, the specimen can usually be scanned in a relatively short time, and thus, movements are not a major source of blurring.

The second category comprises the techniques based on stochastic switching (single marker switching, SMS, or single molecule localization, SML), where the whole sample is illuminated simultaneously but with a low activation intensity. This leads to a random activation of very few markers, keeping all other markers in their non-fluorescent state. Since only a small proportion of all markers is visible in each image (or frame), the probability that two of them are closer than the diffraction barrier is negligible. Therefore, the deconvolution step needed in the first category can be replaced by a simple localization procedure. The drawback is that a large number of frames has to be recorded over a relatively long acquisition time to ensure that the whole structure of interest is registered with high enough precision, leading to a blurring of the final image due to movement of the imaged specimen (Laitenberger, 2018). This motion blur is the main source of distortion associated with SMS microscopy and dealing with this issue using a statistical approach is the focus of this thesis. Among the imaging techniques in this second category are *stochastic optical reconstruction microscopy (STORM)* (Rust et al., 2006; Holden et al., 2011), *photoactivated localization microscopy (PALM)* (Betzig et al., 2006), *fluorescence photoactivation localization microscopy (FPALM)* (Hess et al., 2006), and *PALM with independently running acquisition (PALMIRA)* (Geisler et al., 2007; Egner et al., 2007). See Hell (2007) or Sahl et al. (2017) for a survey and Aspelmeier et al. (2015) and references therein for a more detailed description of the underlying physical principles and methodology of techniques based on (superresolution) fluorescence microscopy.

1.1 Motion blur in SMS microscopy

As described in Aspelmeier et al. (2015), an SMS microscope is essentially a conventional fluorescence microscope with an additional activation laser (see Figure 1.1). In Figure 1.2, the imaging procedure is illustrated schematically. The data acquisition process in SMS microscopy is performed in two steps. The first step of the data acquisition is the transfer of a sparse random subset of all accessible markers to the active state by illumination of the whole sample with a low intensity. In the second step, the active markers are excited and then emit a random number of photons. This fluorescent signal is read out with a detector, and displayed as an image of well separated diffraction patterns. As mentioned in the above paragraph, active markers are sufficiently distant with high probability, and thus, any detected diffraction pattern can be assumed to originate from a single fluorescence marker. Hence, the unknown marker positions in each image are usually determined by calculating the centroid of their observed patterns. This way, spatial sparseness is physically enforced, and because of the known simple structure more sophisticated deconvolution methods are unnecessary. After this localization process, the markers are recorded in temporally and spatially highly resolved position histograms (see bottom row of Figure 1.2). The overlay of a large number of these frames gives the final SMS image (see Figure 1.2 on the right). Note that the localization precision in the single histograms will be $\sqrt{N}$ times better than the original resolution of the microscope, where N is the number of photons forming the pattern (Thompson et al., 2002). An exemplary single frame from the dataset we will analyze in Section 4.2, is displayed in Figure 1.3 as the result of the just described data acquisition procedure. For a more detailed description on the statistics of the activation, emission and detection processes, see Aspelmeier et al. (2015).

Due to the fact that only very few random markers are activated at any given time, each single

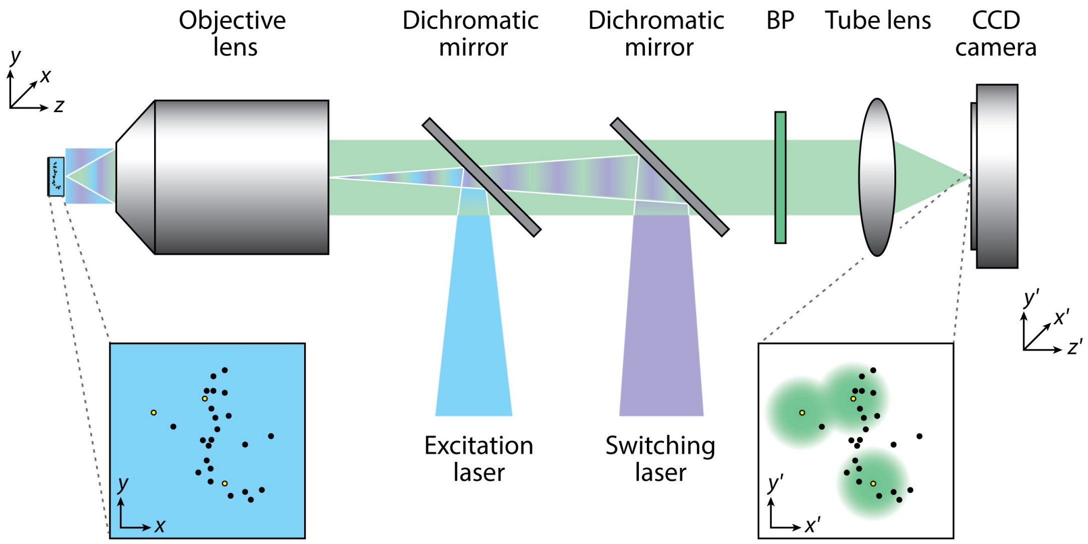

Figure 1.1: Experimental setup of SMS microscopy (Aspelmeier et al., 2015, Figure 10a). SMS microscope with additional activation laser (purple); few active markers (left inset) produce a image on the camera with well separated diffraction patterns (right inset).

image contains only little (but sparse) information. Consequently, a long sequence of images has to be recorded in order to ensure that each marker is observed at least once and the overlay of these frames represents the observed specimen. A comparison between the frames consisting of the detected diffraction patterns, their overlay forming the widefield image and the overlay of the localized data points is displayed in Figure 1.4. Usually, the number of recorded frames is in the range of tens of thousands with a temporal resolution of several milliseconds. Hence, the complete recording typically takes a few minutes. During this time, the specimen may move (see Geisler et al. (2012) and references therein), which leads to a blurring of the overlay forming the final SMS image, see Figure 1.4 on the right.

There are multiple reasons causing different types of movement during the measurement process. External systematic movements of the optical device may cause mechanical drift and rotation. Drift and rotation of the observed structure may further occur due to small vibrations coupled with a rigid specimen that is not perfectly adhesive to the object layer. A vertical movement of the specimen or the object layer can lead to a varying distance between the original focal plane and the ocular. This, or thermal expansion due to heating of the optical device may result in a scaled appearance of the image. Moreover, movement of the living specimen under the microscope, for example due to temperature variations, in horizontal direction (appearing as drift or rotation) or in vertical direction (appearing as scaling) may also contribute to motion blur.

The challenge is therefore to appropriately align the sparse frames correcting for this motion of the observed object. The current practice to tackle this problem is to incorporate fiducial markers (i.e., bright fluorescent microspheres) into the specimen, which can be tracked and used to correct for the motion, either during the measurement process or as a post-processing step after

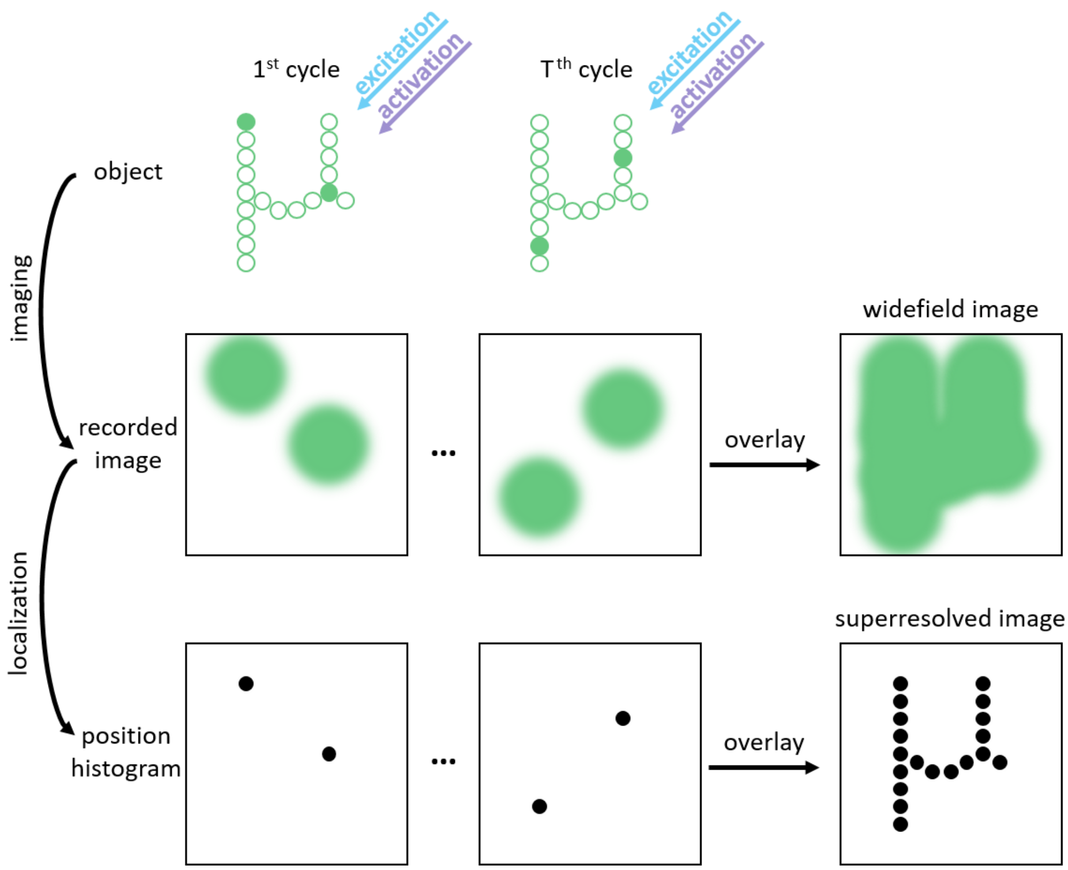

Figure 1.2: SMS imaging procedure (inspired by Aspelmeier et al. (2015, Figure 10b) and Laitenberger (2018, Figure 2.5.1)). In each cycle, a small number of activated fluorophores (top row) generates images on the detector with well separated diffraction patterns (middle row), the overlay of which forms the widefield image (middle row on the right); localization yields position histograms (bottom row), the overlay of which forms the final SMS image (bottom row on the right).

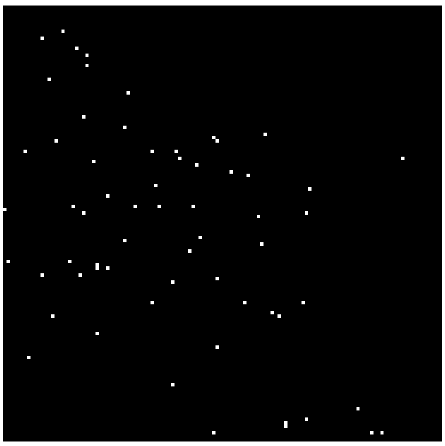

Figure 1.3: A single frame containing a sparse position histogram of the specimen of interest.

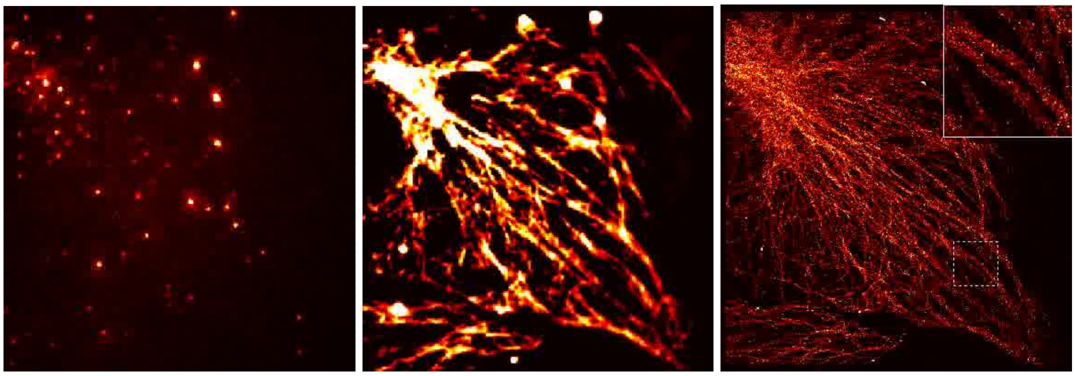

Figure 1.4: Single frames as they are recorded by the detector (left); widefield image (middle); overlay of all position histograms after localization (right). In the zoom-in in the upper right corner a motion blur is clearly visible.

the recording (see, e.g., Grover et al., 2015, and references therein). However, this approach has several disadvantages. Firstly, it is technically demanding and expensive to introduce the fiducial markers and attach them to the specimen. Secondly, often the fiducial markers outshine relevant parts of the image, making it impossible to identify the specimen's features in these areas (see, e.g., Geisler et al., 2012). The design of methods which enable the assessment of the drift, rotation, and scaling of the specimen without using fiducial markers is therefore a significant improvement.

1.2 Relation to the literature

A first attempt at estimating drift has been made by Geisler et al. (2012), who suggested a heuristic correlation method to align subsequent frames properly (see Deschout et al. (2014) for a survey on this issue). In Hartmann et al. (2015), the problem is addressed in a statistically rigorous way, focusing, however, exclusively on drift motion. Working with a parametric model for the drift function, they derived a consistent and asymptotically normally distributed M-estimator for the drift parameter. In this thesis, the M-estimation method is expanded to include also rotation and scaling of the observed specimen and any concatenation of the three, as initiated by Hartmann (2016). Similarly to before, we formulate a parametric model for drift, rotation and scaling functions. We obtain M-estimators for the motion function parameters, which are consistent and jointly asymptotically normally distributed as the acquisition time increases. We further prove consistency of the plug-in estimator for the image. Using these asymptotic results, we construct simple bootstrap confidence bands for the drift, rotation, and scaling functions yielding a measure to assess the statistical uncertainty of our reconstruction. With our generalization of the method, we are now able to handle all orientation preserving similarity transforms, i.e., all (sufficiently smooth) motion types that leave the object as such unchanged and only modify its position or the size in which it appears.

Like the preceeding paper on drift estimation (Hartmann et al., 2015) and the dissertation Hartmann (2016), the present work is closely related to Gamboa et al. (2007) and Bigot et al.

(2009). The former considers curves, which can be referred to as one-dimensional images, subjected to Gaussian noise and translation and the latter two-dimensional images with Gaussian noise, which have been transformed by translation, rotation and scaling. The idea of exploiting the shift-property of the Fourier transform and determine estimators as minimizers of certain contrast functionals stems from those papers. Furthermore, Bigot et al. (2009) describe already the two-step procedure based on the application of the Fourier-Mellin transform to the squared Fourier magnitudes of the data, which is used also here to combine estimation of drift with estimation of rotation and scaling.
Note however, that our asymptotics is substantially different to that underlying most other image alignment and registration methods, and in particular also to the setting used in the two articles just mentioned. Considering the number of recorded frames tending to infinity is specifically well applicable to the scenario of sparse single images and relatively long acquisition times which are inherent to techniques like SMS microscopy. In contrary, for other imaging methods usually the number of pixels is assumed to increase, and the full image is observed at each time step. The latter setting corresponds to an asymptotically ideal spatial resolution, whereas in our setting we assume an improving temporal resolution with a predefined spatial resolution. As a matter of fact, in both of the above mentioned works, the number of images is fixed and each single one is subjected to an unknown similarity transform. For each of these images, the transformation is estimated as an individual set of parameters, which means that the number of parameters is of the same scale as the number of observed images. They prove consistency for their estimators and asymptotic normality as the number of pixels tends to infinity. In contrast to that, here we work with parametrized motion functions, allowing for estimation of a time dependent motion using a fixed small number of parameters and sparse single frames, the number of which tends to infinity in our asymptotic setting.
Finally, we remark that even though our method is inspired by an application in SMS microscopy, it may be used in other scenarios, where the same setting applies, i.e., only a sequence of sparse registrations of an object is available, like for example in undersampled real time magnetic resonance imaging (Li et al., 2014).

CHAPTER 2

Modeling and estimation procedure

We continue in Section 2.1 with explaining our semiparametric model in detail and specifying the assumptions on the underlying gray scale image and the motion functions. In Section 2.2 we then elaborate on the estimation procedure, which enables us to perform motion correction in SMS microscopy data, where the motion can be drift, rotation, scaling or a combination of any of the three. The basic idea is to first correct for possible rotation and scaling, and subsequently estimate drift, obtaining then a final plug-in estimate for the SMS image displaying the specimen of interest.

2.1 The semiparametric model

We first derive a basic Bernoulli model explaining the data acquisition process well. Afterwards, we apply a binning procedure and several standard transformations and approximations leading to a Gaussian Fourier model, which is then used for the estimation of the motion parameters.
As described in the introduction, the measurement process involves the recording of a large number of frames consecutively. Each of these frames contains a collection of distinct diffraction patterns generated by the random sparse subset of fluorescent markers which are active during the recording of this frame. As a preprocessing step, the diffraction patterns are localized by calculating their centroid, which is only possible due to the known sparse structure. As mentioned in the introductory Chapter 1, the statistics of the activation, emission and detection processes generating the observed diffraction patterns in SMS microscopy and other superresolution imaging techniques will not be treated here. A detailed description can be found in Aspelmeier et al. (2015). For the purpose of this thesis, namely the motion correction of frames to obtain a deblurred final SMS image, it is favorable to focus on the preprocessed data, i.e., the localized position histograms.
To describe our model precisely, we introduce some notation. Our aim is to estimate the true unknown marker density $f^0 \colon \mathbb{R}^2 \to [0,1]$. For parameters $(\theta, \phi, \alpha) \in \Theta \times \Phi \times \mathrm{A} \subseteq \mathbb{R}^{d_1+d_2+d_3}$ in the compact parameter space $\Theta \times \Phi \times \mathrm{A}$ and time points $t \in [0,1]$, we consider

- drift vectors $\delta_t^\theta \in \mathbb{R}^2$,
- rotation angles $\rho_t^\phi \in (-\pi/2, \pi/2]$,
- and scaling factors $\sigma_t^\alpha \in [\sigma_{\min}, \sigma_{\max}]$, for some $\sigma_{\min}, \sigma_{\max} > 0$.

We further consider a fixed finite grid $\mathbb{X} := \{x_j \in \mathbb{R}^2, 1 \le j \le n\}$ of $n \in \mathbb{N}$ pixels and denote with

$$f_j^t = f^t(x_j) := f\left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}\left(x_j - \delta_t^{\theta_0}\right)\right),$$

the deformed gray scale image, where $(\theta_0, \phi_0, \alpha_0)$ are the unknown true parameters, and

$$R_\rho := \begin{pmatrix} \cos(\rho) & -\sin(\rho) \\ \sin(\rho) & \cos(\rho) \end{pmatrix}$$

is the rotation matrix with angle ρ. Here, f^t can be regarded as a shifted, rotated and scaled version of f^0 and f_j^t is its value at the pixel location $x_j \in \mathbb{X}$. Starting from the position histograms, we transfer each recorded marker position to the closest pixel position in $\mathbb{X}$. For reasonably large number of pixels, the error induced by this assignment is irrelevant compared to the motion blurring and can therefore be ignored - with one caveat: Rotation of small objects may be misclassified as drift. This can be dealt with by choosing an appropriate cutoff for the Fourier coefficients, see Section 4.1. The observed frames are now denoted as $O^t = \left(O_j^t\right)_{1\le j\le n}$, for time points $t \in \mathbb{T} := \{0, 1/T, \ldots, (T-1)/T\}$, where $T \in \mathbb{N}$ is the total number of frames. They consist of single observations $O_j^t = O^t(x_j)$ taking the values $O_j^t = 1$ if a signal was recorded at x_j, and $O_j^t = 0$ otherwise. Since the activation of fluorescent markers happens independently, the observations O_j^t can be modeled as independent realizations of Bernoulli random variables with some success probability p_j^t. This probability p_j^t is proportional to the marker density f_j^t at this pixel location $j \in \{1, \ldots, n\}$ at time $t \in \mathbb{T}$. It further depends on external influences given by the experimental setup, like the activation and excitation laser intensities or properties of the microscopy (e.g., its detection power). These external factors are collectively modeled as a contribution $p \in (0, 1)$, which is assumed to be fixed and known. Hence, we arrive at the following basic Bernoulli model for our independent observations:

$$O_j^t \sim \mathrm{Ber}(f_j^t \cdot p), \qquad 1 \le j \le n, \quad t \in \mathbb{T}. \tag{2.1}$$

Remark 2.1 (Bounds on the rotation angle ρ_t^ϕ and the scaling factor σ_t^α)**.** *As mentioned in the introduction, we will work with the squared Fourier magnitude of the marker density, which is invariant under rotation by an angle of multiples of* π *(Hartmann, 2016, Lemma 2.10). Therefore, we restrict ourselves to values of the rotation angle* ρ_t^ϕ *in an interval of length* ϕ *to ensure identifiability. We want the interval to contain* 0*, as we will assume that we have no rotation at the beginning, i.e.,* $\rho_0^\phi = 0$ *(see Assumption 2.14 (B1)). Hence, we choose the symmetric interval* $(-\pi/2, \pi/2]$*, which allows for clockwise and counter-clockwise rotation. The bounds on the scaling function* σ_t^α *are useful for technical reasons, but they are also induced by the setup, namely by the resolution of the microscope and the pixel size (*$\sigma_{\min}$*), and by the size of the observation window (*$\sigma_{\max}$*).*

2.1.1 Binning

In SMS microscopy, each frame $O^t = \left(O^t_j\right)_{1\leq j\leq n}$ typically contains very little information because the number of observed pixels is small, whereas the length T of the image sequence is comparatively large. The idea is to *bin* subsequent frames, i.e., take the point-wise average of them in order to increase the information per frame and reduce the noise level of the data. This represents a bias-variance trade-off in the following sense. Calculating the average over all observed frames gives an estimate for the true unknown image, which has a strongly reduced noise level due to the large number of single observations. However, as described in the introduction, the motion of the imaged object over time causes a large bias of the resulting superimposed image, which will be blurred very much. On the other hand, considering the single frames there is no motion of the object, since all observations on one specific frame have been obtained at the same time. Here, the issue is that the variance among the frames is high due to the extreme sparsity. The goal of the binning procedure is to strike a balance between both error sources. A suitable bin width is chosen small enough such that the binned frames are not blurred too strongly but also large enough to control the noise level and gain sufficient information about the observed specimen. This pre-averaging also has the benefit of reducing the memory needed to process reconstruction methods on the sequence.

Hence, for all $T \in \mathbb{N}$, we define a *bin size* $\beta_T \in \mathbb{N}$ such that $T/\beta_T \in \mathbb{N}$. We construct a new image sequence of length T/β_T by averaging over β_T subsequent frames,

$$\widetilde{O}^t_j := \frac{1}{\beta_T}\sum_{i=0}^{\beta_T-1} O_j^{t+i/T}, \quad t\in\tilde{\mathbb{T}}, j\in\{1,\dots,n\}, \tag{2.2}$$

where $\tilde{\mathbb{T}} := \{0, \beta_T/T, 2\beta_T/T, \dots, (T-\beta_T)/T\}$. The bin size β_T regulates the degree of sparsity of the binned frames. The scaled observations $\beta_T\widetilde{O}^t_j = \sum_{i=0}^{\beta_T-1} O_j^{t+i/T}$ follow a Poisson binomial distribution as sum of independent Bernoulli distributed random variables. Le Cam (1960) showed that they can be well approximated by a Poisson distribution with parameter $\sum_{i=0}^{\beta_T-1} f_j^{t+i/T}\cdot p = \beta_T\tilde{f}^t_j\cdot p$, where we denote the average marker density with

$$\tilde{f}^t_j = \frac{1}{\beta_T}\sum_{i=0}^{\beta_T-1} f_j^{t+i/T}, \quad t\in\tilde{\mathbb{T}}, j\in\{1,\dots,n\}.$$

The bin width β_T is chosen small enough, such that the maximal time step between two density values contributing to $\tilde{f}^t_j$, namely β_T/T, is very short and tends to zero as $T\to\infty$ (see Assumption 2.15). Since motion functions and marker density are smooth enough (see Assumptions 2.14 and 2.13), the average marker density $\tilde{f}^t_j$ is a good approximation to the single density values $f_j^{t+i/T}$ for $1\leq i\leq\beta_T-1$. Hence, we can reconstruct this average marker density instead, without inducing a significant error. The obtained Poisson distributed random variable $\mathrm{Poi}(\beta_T\tilde{f}^t_j\cdot p)$ equals in distribution the sum of β_T independent and identically Poisson distributed random variables with parameter $\tilde{f}^t_j\cdot p$. We perform a variance stabilizing transformation based on the Delta-method (Theorem B.7) applied to the i.i.d. $\mathrm{Poi}(\tilde{f}^t_j\cdot p)$ random variables using the function $g(x) = 2\sqrt{x}$. This yields

$$\sqrt{\beta_T}\left(2\sqrt{\frac{1}{\beta_T}\operatorname{Poi}\left(\beta_T \tilde{f}_j^t \cdot p\right)} - 2\sqrt{\tilde{f}_j^t \cdot p}\right) \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0,1).$$

Summarizing, we arrive at the following approximation:

$$\sqrt{\widetilde{O}_j^t} \approx \sqrt{\frac{1}{\beta_T}\operatorname{Poi}(\beta_T \tilde{f}_j^t \cdot p)} \approx \mathcal{N}\left(\sqrt{\tilde{f}_j^t \cdot p}, \frac{1}{4\beta_T}\right). \tag{2.3}$$

Remark 2.2 (Anscombe Transformation)**.** *In practice, instead of the exact variance stabilization transformation $x \mapsto 2\sqrt{x}$, often an Anscombe type transform, $x \mapsto 2\sqrt{x+c}$ for some constant $c > 0$ is used (see Anscombe (1948)), since they have a better finite sample size performance, depending on the choice of c. In our case we would primarily like to reduce the bias, keeping however an approximately constant variance. Following Chapter 2 of Brown et al. (2010), we therefore select $c = 1/4$ for the transformation of our real SMS data instead of $c = 3/8$, which would have optimal rates for the sole purpose of variance stabilization.*

Remark 2.3 (Justification of Gaussian approximation)**.** *Note that the approximation of the binned observations by the stated normal distribution can further be justified by the following argumentation. For any given $j \in \{1, \dots, n\}$, which describes an 'empty' pixel location x_j, i.e., a location without any markers present at time point t, the distribution of the observations O_j^t, $t \in \mathbb{T}$, degenerates to a dirac measure at 0. But so does the normal distribution on the right hand side of (2.3), since both mean and variance tend to zero in this case. For the remaining pixels containing signal, however, $\tilde{f}_j^t \cdot p$ are bounded away from 0 and 1, and therefore the corresponding observations O_j^t fulfill Lindeberg's condition (Billingsley, 1995, Theorem 27.2) for $T \to \infty$, implying the validity of the central limit theorem.*

In the following, we will only work with the binned observations $\widetilde{O}_j^t$. Therefore, we will omit the tilde and write again O_j^t for the binned observations as well as $\mathbb{T}$ for $\tilde{\mathbb{T}}$. As mentioned before, we can reconstruct the average marker density $\tilde{f}_j^t$ and will suppress the tilde here, too. Furthermore, assuming the detection probability p determined by the experimental setup to be known, we absorb it into the marker density f^0. For ease of notation, we leave out the square root emerging from the variance stabilization. This means, that we write f_j^t instead of $\sqrt{\tilde{f}_j^t \cdot p}$ and f for $\sqrt{f^0 \cdot p}$ in the remainder of this thesis, keeping in mind, that we need to invert the transform $x \mapsto \sqrt{x}$ in the end to obtain an estimator for the true (scaled) marker density $f^0 \cdot p$. Collating all these preliminary steps, we can now define the model on which our theory is based, and which approximates the actual data collection process sufficiently well.

Definition 2.4. *The approximate Gaussian model for SMS microscopy is given by*

$$O_j^t = f_j^t + \frac{1}{2\sqrt{\beta_T}}\epsilon_j^t, \quad t \in \mathbb{T}, j \in \{1, \dots, n\},$$

where $\epsilon \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ are standard Gaussian random variables.

2.1.2 The standard Fourier transform and its shift property for translation

In this subsection, we define the Fourier transform and in the following subsection the related (analytical) Fourier-Mellin transform, which are crucial for this work because of their (generalized) shift properties. First, we need the following definition of spaces of integrable functions.

Definition 2.5 (L^p-spaces)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. For $p \in [1, \infty]$, we define*

$$\mathcal{L}^p(\Omega, \mathcal{A}, \mu) := \left\{ g \colon \Omega \to \mathbb{C} \mid g \text{ is } \mu\text{-measurable and } \|g\|_{\mathcal{L}^p} < \infty \right\},$$

with the $\mathcal{L}^p$-seminorm

$$\|\cdot\|_{\mathcal{L}^p} : \mathcal{L}^p \to [0, \infty], \quad g \mapsto \left(\int_\Omega |g(x)|^p \, \mu(\mathrm{d}x) \right)^{1/p}, \text{ for } p \in [1, \infty), \text{ and}$$

$$\|\cdot\|_{\mathcal{L}^\infty} : \mathcal{L}^\infty \to [0, \infty], \quad g \mapsto \inf_{N \in \mathcal{A}, \mu(N)=0} \sup_{x \in \Omega \setminus N} |g(x)| \, .$$

Identifying functions that are equal μ-a.e. leads to the normed L^p-spaces*. To this end, let*

$$\mathcal{N}^p := \left\{ g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu) \mid g = 0 \ \mu\text{-a.e.} \right\}$$

be the set of functions which are 0 μ-a.e. Using this notation, we define the L^p-space $L^p(\Omega, \mathcal{A}, \mu) := \mathcal{L}^p(\Omega, \mathcal{A}, \mu)/\mathcal{N}^p$*, together with the* L^p-norm

$$\|\cdot\|_{L^p} : L^p(\Omega, \mathcal{A}, \mu) \to [0, \infty), \quad [g] \mapsto \|g\|_{\mathcal{L}^p} \, .$$

Remark 2.6. *We will often write $L^p(\mathbb{R}^2)$ for the L^p-space on $\mathbb{R}^2$ with the Borel σ-algebra and the Lebesgue-measure μ, $L^p(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu)$.*

Recall now the Fourier transform of a function $g \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$,

$$\mathcal{F}_g \colon \mathbb{R}^2 \to \mathbb{C}, \quad \omega \mapsto \int_{\mathbb{R}^2} \mathrm{e}^{-2\pi \mathrm{i} \langle \omega, x \rangle} g(x) \, \mathrm{d}x. \tag{2.4}$$

Let $(\delta, \rho, \sigma) \in \mathbb{R}^2 \times \mathbb{R} \times (0, \infty)$. As can be easily derived from the definition in (2.4), a shifted, rotated, and scaled version

$$\tilde{g} \colon \mathbb{R}^2 \to \mathbb{R}, \quad x \mapsto g\Big(1/\sigma \cdot R_{-\rho}\big(x - \delta\big)\Big)$$

of a function g has the Fourier transform

$$\mathcal{F}_{\tilde{g}}(\omega) = \sigma^2 \mathrm{e}^{-2\pi \mathrm{i} \langle \omega, \delta \rangle} \mathcal{F}_g\big(\sigma R_{-\rho} \omega\big), \tag{2.5}$$

transferring the rotation by $-\rho$ from the image domain into the Fourier domain while inverting the scaling factor $1/\sigma$. The drift term δ results in a phase shift. For $(\rho, \sigma) = (0, 1)$, (2.5) becomes the classical shift property of the Fourier transform for translation,

$$\mathcal{F}_{\tilde{g}}(\omega) = \mathrm{e}^{-2\pi \mathrm{i}\langle \omega, \delta\rangle} \mathcal{F}_g(\omega). \tag{2.6}$$

2.1.3 The analytical Fourier-Mellin transform and its shift property for rotation and scaling

To get a property similar to (2.6) for rotation and scaling, we consider a Fourier-type transform defined on the similarity group, which is called the Fourier-Mellin transform or FMT (see, e.g., Derrode and Ghorbel, 2004; Ghorbel, 1994; Lenz, 1990; Gauthier et al., 1991; Segman et al., 1992). Consider the locally compact groups $\big((0,\infty),\cdot\big)$ and $\big([0,2\pi),+\big)$, where the addition in the latter should be understood modulo 2π. Their direct product $G \coloneqq \big((0,\infty),\cdot\big) \times \big([0,2\pi),+\big)$ is also a locally compact group and can be equipped with the Haar measure $r^{-1}\,\mathrm{d}r\,\mathrm{d}\psi$, where $\mathrm{d}r$ and $\mathrm{d}\psi$ denote the standard Lebesgue measures on $(0,\infty)$ and on $[0,2\pi)$, respectively. The measure $r^{-1}\,\mathrm{d}r\,\mathrm{d}\psi$ is positive and invariant on G. Furthermore, G has the dual group $(\mathbb{R},+)\times(\mathbb{Z},+)$, representing the parameter space in the Fourier-Mellin domain. Hence, we can define a Fourier transform for functions on G (Rudin, 1990). To this end, for $p \in \{1,2\}$, let

$$L^p(G) \coloneqq \left\{ g\colon G \to \mathbb{R} \,\middle|\, \|g\|_{L^p(G)} < \infty \right\},$$

where

$$\|g\|_{L^p(G)} \coloneqq \left(\int_0^\infty \int_0^{2\pi} |g(r,\psi)|^p \,\mathrm{d}\psi\, \frac{\mathrm{d}r}{r} \right)^{1/p}.$$

The standard FMT of a function $g\colon G \to \mathbb{R}$ such that $g \in L^1(G)$ is given as

$$\widetilde{\mathcal{M}}_g\colon \mathbb{Z}\times\mathbb{R} \to \mathbb{C}, \quad (u,v) \mapsto \int_0^\infty \int_0^{2\pi} \mathrm{e}^{-2\pi \mathrm{i} u\psi} r^{-\mathrm{i}v} g(r,\psi)\,\mathrm{d}\psi\, \frac{\mathrm{d}r}{r}. \tag{2.7}$$

However, the FMT exists only for functions g that behave like r^γ in the vicinity of the origin (i.e., $r = 0$) for some $\gamma > 0$ (Derrode and Ghorbel, 2004), which usually does not hold for real grey value images or their Fourier transforms as their value would have to be zero for small r. To overcome this problem, Derrode and Ghorbel (2004) and Ghorbel (1994) have proposed to use $g_\gamma\colon (r,\psi) \mapsto r^\gamma g(r,\psi)$ instead of g in such contexts for some fixed $\gamma > 0$, which leads to the following definition of the so-called analytical Fourier-Mellin transform (AFMT) of g. If $g_\gamma \in L^1(G)$, let

$$\mathcal{M}_g\colon \mathbb{Z}\times\mathbb{R} \to \mathbb{C}, \quad (u,v) \mapsto \int_0^\infty \int_0^{2\pi} \mathrm{e}^{-2\pi \mathrm{i} u\psi} r^{\gamma-\mathrm{i}v} g(r,\psi)\,\mathrm{d}\psi\, \frac{\mathrm{d}r}{r}. \tag{2.8}$$

As stated in Rudin (1990), if $g_\gamma \in L^1(G) \cap L^2(G)$, the AFMT fulfills the following Parseval equation,

$$\left\|g_\gamma\right\|^2_{L^2(G)} = \int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_g(u,v)\right|^2 \,\mathrm{d}v. \tag{2.9}$$

Consider a rotated and scaled version $\tilde{g}(x) := g\big(1/\sigma \cdot R_{-\rho}x\big)$ of g, where $\rho \in [0, 2\pi)$ and $\sigma > 0$. Then,

$$\mathcal{M}_{\tilde{g}}(u, v) = \sigma^{\gamma - \mathrm{i}v} \mathrm{e}^{-2\pi \mathrm{i} u \rho} \mathcal{M}_g(u, v), \tag{2.10}$$

which can be interpreted as a shift property for the AFMT that converts rotation and scaling into a phase shift in the Fourier-Mellin domain as well as a multiplication of the magnitude with σ^γ. In order to be able to compute the Fourier-Mellin transform also for functions defined on $\mathbb{R}^2$ and not G, we will need the following coordinate transforms.

Definition 2.7 (Polar and log-polar coordinate transforms)**.** *We define the* polar coordinate transform $\mathcal{P}$ *and the* log-polar coordinate transform $\mathcal{LP}$ *as*

$$\begin{aligned} \mathcal{P} &: [0, \infty) \times [0, 2\pi) \to \mathbb{R}^2, \quad (r, \psi) \mapsto \big(r\cos(\psi), r\sin(\psi)\big), \\ \mathcal{LP} &: \mathbb{R} \times [0, 2\pi) \to \mathbb{R}^2, \quad (l, \psi) \mapsto \big(\mathrm{e}^l \cos(\psi), \mathrm{e}^l \sin(\psi)\big). \end{aligned}$$

Remark 2.8 (Connection between Fourier transform and Fourier-Mellin transform)**.**
Note that the analytical Fourier-Mellin transform is a Fourier-type transform from $\mathbb{R}_+ \times S^1$ onto the similarity group $\mathbb{R} \times \mathbb{Z}$. More specifically, for $g \in L^1\left(\mathbb{R}^2\right)$ we get $\mathcal{M}_g(u, v) = \mathcal{F}_{\tilde{g}}(u, v)$ by basic calculations, with

$$\tilde{g} : [0, \infty) \times [0, 2\pi), \quad (r, \psi) \mapsto r^\gamma (g \circ \mathcal{LP})(r, \psi).$$

This will allow us to use the Fast Fourier Transform algorithm (FFT, see, e.g., Cooley and Tukey (1965)) to efficiently compute the analytical Fourier-Mellin transform in the application to datasets.

2.1.4 Model assumptions

Before stating the formal assumptions that we make on the underlying image (Assumption 2.13), the motion functions (Assumption 2.14) and the binning and cutoff rates (Assumption 2.15), we introduce some terminology.

Definition 2.9 (Not translation, rotation, or scaling invariant)**.** *A function $g : \mathbb{R}^2 \to \mathbb{C}$ is called* not translation invariant*, if there is no $\delta \in \mathbb{R}^2 \setminus \{0\}$ such that $g(x) = g(x - \delta)$ for all $x \in \mathbb{R}^2$. Similarly, g is called* not rotation invariant*, if there is no $\rho \in (0, 2\pi)$ such that $g(x) = g\big(R_{-\rho}(x)\big)$ for all $x \in \mathbb{R}^2$. Moreover, g is called* not scaling invariant*, if there is no $\sigma \in \big((0, 1) \cup (1, \infty)\big)$ such that $g(x) = g\big(x/\sigma\big)$ for all $x \in \mathbb{R}^2$.*

Definition 2.10 (Identifiability)**.** *For some index set I, let $G_I = \left\{g^i : [0, 1] \to \mathbb{R} \mid i \in I\right\}$ a set of functions. We call G_I* identifiable*, if for all $i, j \in I$, the existence of a Borel set $B \subseteq [0, 1]$ of Lebesgue measure equal to 1 with $g^i(t) = g^j(t)$ for all $t \in B$ implies $i = j$.*

Definition 2.11 (Sobolev space)**.** *For $p > 0$, we call*

$$H^p(\mathbb{R}^2) := \left\{ g \in L^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu) \middle| \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right)^p \left|\mathcal{F}_g(\omega)\right|^2 \mathrm{d}\omega < \infty \right\}$$

the Sobolev space of order p, where μ is the Lebesgue measure.

Definition 2.12 (Total variation)**.** *Let $g\colon [0,1] \to \mathbb{C}$ and define the set of all finite partitions of $[0,1]$ as $P := \{\{t_0, \dots, t_k\} \mid k \in \mathbb{N}, 0 = t_0 < t_1 < \cdots < t_k = 1\}$. The* total variation *of g is*

$$\mathrm{TV}(g) := \sup_{\{t_0,\dots,t_k\} \in P} \sum_{i=0}^{k-1} |g(t_{i+1}) - g(t_i)| \,.$$

Assumption 2.13 (Assumptions on the image)**.**

(A1) The support of the marker density f is contained in a compact set, more specifically, there is a $C_f > 0$ such that $f(x) = 0$ for all $x \in \mathbb{R}^2$ with $\|x\| > C_f$. Furthermore, f is bounded, i.e., $\|f\|_\infty := \sup_{x \in \mathbb{R}^2} |f(x)| < \infty$.

(A2) The image f is not translation, rotation, or scaling invariant.

(A3) We have that $f \in L^2(\mathbb{R}^2) \cap H^{3+\kappa}(\mathbb{R}^2)$ for some $\kappa > 0$, where $L^2(\mathbb{R}^2)$ is the usual normed space of square integrable functions from Definition 2.5 and $H^{3+\kappa}(\mathbb{R}^2)$ is the Sobolev space defined in Defintion 2.11.

(A4) We have the following Sobolev-$(2 + \tilde{\kappa})$ condition for some $\tilde{\kappa} > 0$,

$$\int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left(1 + \|(u,v)\|^2\right)^{2+\tilde{\kappa}} \left| \mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \right|^2 \mathrm{d}v < \infty.$$

(A5) We have the following continuity condition: for any $\epsilon > 0$ there is a $\delta > 0$ such that for all $y \in \mathbb{R}^2$ with $\|y\| < \delta$

$$\int_{\mathbb{R}^2} |f(x+y) - f(x)| \,\mathrm{d}x < \epsilon.$$

Assumption 2.14 (Assumptions on the motion functions)**.**

(B1) Since we do not expect drift, rotation, or scaling at time $t = 0$, we assume that $\delta_0^\theta = 0$, $\rho_0^\phi = 0$, and $\sigma_0^\alpha = 1$ for all $(\theta, \phi, \alpha) \in \Theta \times \Phi \times \mathrm{A}$.

(B2) There are convex open neighborhoods $U \subseteq \Phi \times \mathrm{A}$ of (ϕ_0, α_0) and $U' \subseteq \Theta$ of θ_0 and there is a $C > 0$ such that for all $t \in [0,1]$, $\alpha \mapsto \sigma_t^\alpha$ and $\phi \mapsto \rho_t^\phi$ are twice differentiable on U, and $\theta \mapsto \delta_t^\theta$ is twice differentiable on U'. Moreover,

$$\left\|\mathrm{grad}_\theta(\delta_t^\theta)_1\right\|, \left\|\mathrm{grad}_\theta(\delta_t^\theta)_2\right\| \le C, \quad \left\|\mathrm{Hess}_\theta(\delta_t^\theta)_1\right\|_1, \left\|\mathrm{Hess}_\theta(\delta_t^\theta)_2\right\|_1 \le C^2,$$
$$\left\|\mathrm{grad}_\phi \rho_t^\phi\right\|, \left\|\mathrm{grad}_\alpha \sigma_t^\alpha\right\| \le C, \quad \left\|\mathrm{Hess}_\phi \rho_t^\phi\right\|_1, \left\|\mathrm{Hess}_\alpha \sigma_t^\alpha\right\|_1 \le C^2,$$

uniformly in θ, ϕ, α, and t.

(B3) The second partial derivatives

$$\theta \mapsto \frac{\partial^2 (\delta_t^\theta)_1}{\partial\theta_{m_1}\partial\theta_{m_1'}}, \quad \theta \mapsto \frac{\partial^2 (\delta_t^\theta)_2}{\partial\theta_{m_1}\partial\theta_{m_1'}}, \quad \phi \mapsto \frac{\partial^2 \rho_t^\phi}{\partial\phi_{m_2}\partial\phi_{m_2'}}, \quad \alpha \mapsto \frac{\partial^2 \sigma_t^\alpha}{\partial\alpha_{m_3}\partial\alpha_{m_3'}},$$

are continuous at the true parameters θ_0, ϕ_0 and α_0, respectively, for all $m_l, m_l' \in \{1, \ldots, d_l\}$, $l \in \{1, \ldots, 3\}$. Furthermore, the first partial derivatives at θ_0, ϕ_0, and α_0, as functions in t, are of bounded total variation, i.e., there is a $C' > 0$ such that

$$\mathrm{TV}\left(t \mapsto \frac{\partial(\delta_t^\theta)_1}{\partial\theta_{m_1}}\right), \mathrm{TV}\left(t \mapsto \frac{\partial(\delta_t^\theta)_2}{\partial\theta_{m_1}}\right), \mathrm{TV}\left(t \mapsto \frac{\partial\rho_t^\phi}{\partial\phi_{m_2}}\right), \mathrm{TV}\left(t \mapsto \frac{\partial\sigma_t^\alpha}{\partial\alpha_{m_3}}\right) < C',$$

for all $m_l \in \{1, \ldots, d_l\}$, $l \in \{1, \ldots, 3\}$.

(B4) The maps

$$\begin{aligned}
&\Theta \to L^1\big([0,1], \mathbb{R}^2\big), && \theta \mapsto \Big(\delta^\theta \colon t \mapsto \delta_t^\theta = \big((\delta_t^\theta)_1, (\delta_t^\theta)_2\big)\Big),\\
&\Phi \to L^1\big([0,1], (-\pi/2, \pi/2]\big), && \phi \mapsto \Big(\rho^\phi \colon t \mapsto \rho_t^\phi\Big),\\
&\mathrm{A} \to L^1\big([0,1], [\sigma_{\min}, \sigma_{\max}]\big), && \alpha \mapsto \Big(\sigma^\alpha \colon t \mapsto \sigma_t^\alpha\Big)
\end{aligned}$$

are continuous. Moreover, for each $(\theta, \phi, \alpha) \in \Theta \times \Phi \times \mathrm{A}$, the motion functions $t \mapsto \delta_t^\theta$, $t \mapsto \rho_t^\phi$, and $t \mapsto \sigma_t^\alpha$ are continuous.

(B5) The sets $\{t \mapsto \delta_t^\theta | \theta \in \Theta\}$, $\{t \mapsto \rho_t^\phi | \phi \in \Phi\}$, and $\{t \mapsto \sigma_t^\phi | \alpha \in \mathrm{A}\}$ are identifiable.

(B6) There are open neighborhoods $U_\delta \subseteq \Theta$ of θ_0 and $U_{\rho,\sigma} \subseteq \Phi \times \mathrm{A}$ of (ϕ_0, α_0) and constants $L_\delta, L_\rho, L_\sigma > 0$ such that the following local uniform Lipschitz conditions hold,

$$\sup_{t\in[0,1]} \left\| \delta_t^\theta - \delta_t^{\theta_0} \right\| \le L_\delta \|\theta - \theta_0\| \quad \text{for all } \theta \in U_\delta, \text{ as well as}$$

$$\sup_{t\in[0,1]} \left| \rho_t^\phi - \rho_t^{\phi_0} \right| \le L_\rho \|\phi - \phi_0\|, \text{ and } \sup_{t\in[0,1]} \left| \sigma_t^\alpha - \sigma_t^{\alpha_0} \right| \le L_\sigma \|\alpha - \alpha_0\|$$

$$\text{for all } (\phi, \alpha) \in U_{\rho,\sigma}.$$

(B7) There is a $C'' > 0$ such that uniformly in θ, ϕ, and α, respectively,

$$\mathrm{TV}\big(t \mapsto (\delta_t^\theta)_1\big) + \mathrm{TV}\big(t \mapsto (\delta_t^\theta)_2\big) < C'', \quad \mathrm{TV}(t \mapsto \rho_t^\phi) + \mathrm{TV}(t \mapsto \sigma_t^\alpha) < C''.$$

(B8) For each of the four gradients

$$\begin{aligned}
&[0,1] \to \mathbb{R}, \quad t \mapsto \mathrm{grad}_\theta \left(\delta_t^\theta\right)_1 \Big|_{\theta=\theta_0}, && [0,1] \to \mathbb{R}, \quad t \mapsto \mathrm{grad}_\theta \left(\delta_t^\theta\right)_2 \Big|_{\theta=\theta_0}\\
&[0,1] \to \mathbb{R}, \quad t \mapsto \mathrm{grad}_\phi \rho_t^\phi \Big|_{\phi=\phi_0}, && [0,1] \to \mathbb{R}, \quad t \mapsto \mathrm{grad}_\alpha \sigma_t^\alpha \Big|_{\alpha=\alpha_0}
\end{aligned}$$

the components are linearly independent functions in t.

Assumption 2.15 (Assumptions on the cutoff and binning rates)**.** *For the binning rate β_T and Fourier-cutoff rates r_T, u_T and v_T we assume the following asymptotic behavior*

(C1) $r_T, u_T, v_T, \beta_T \overset{T\to\infty}{\longrightarrow} \infty$, $\beta_T = o(T)$, $r_T = o(T^{1/6})$,

(C2) $\sqrt{T} r_T^4 = o(\beta_T)$, $\sqrt{T} r_T^{2+\gamma} = o(\beta_T)$, $\sqrt{T} u_T v_T \|(u_T, v_T)\|^2 r_T^{2\gamma} = o(\beta_T)$.

(C3) Let r_T, u_T and v_T be such that $\mathcal{M}_{|\mathcal{F}_f|^2}(u,v) = \mathcal{M}^T_{|\mathcal{F}_f|^2}(u,v) + o\left((u_T v_T)^{-1}\right)$.

(C4) consider only a subsequence of total number of frames $T \in \mathbb{N}$ such that (C1) and (C2) hold and $T/\beta_T \in \mathbb{N}$.

Remark 2.16 (Identifiability of the model)**.** *The Assumptions 2.13 (A2), 2.14 (B1, B5) are crucial to the identifiability of our model.*
If (A2) does not hold, for example, because f is invariant to rotations by some angle ρ', then the rotation function ρ^ϕ is only well defined modulo the period length ρ'. Similar problems arise for the drift and scaling functions.
If (B1) does not hold, we can choose arbitrary intercepts $(\delta_0, \rho_0, \sigma_0) \in \mathbb{R}^2 \times \mathbb{R} \times (0,\infty)$ and rewrite our model via $\tilde{\delta}_t^\theta := \delta_t^\theta + \delta_0$, $\tilde{\rho}_t^\phi := \rho_t^\phi + \rho_0$, $\tilde{\sigma}_t^\alpha := \sigma_t^\alpha \cdot \sigma_0$, and

$$\tilde{f}(x) := f\Big(\sigma_0 \cdot R_{\rho_0}\big(x + \delta_0\big)\Big), \quad x \in \mathbb{R}^2,$$

absorbing the intercepts into the function f.
Assumption 2.14 (B5) ensures that the motion functions can be identified by their respective parameters.

Example 2.17. *Clearly, an appropriate choice of the parametric model is crucial to obtain satisfactory results. As a very common example, consider polynomial models for the motion functions, i.e., for $t \in \mathbb{T}$ and some decomposition $d_1 = d_1' + d_1''$,*

$$\left(\delta_t^\theta\right)_1 = \sum_{m=0}^{d_1'} \theta_m t^m, \qquad \left(\delta_t^\theta\right)_2 = \sum_{m=0}^{d_1''} \theta_m t^m, \qquad \rho_t^\phi = \sum_{m=0}^{d_2} \phi_m t^m, \qquad \sigma_t^\alpha = \sum_{m=0}^{d_3} \alpha_m t^m.$$

To ensure identifiability, we need by Assumption (B1) that $\delta_0^\theta = 0_{d_1}$, $\rho_0^\phi = 0_{d_2}$ and $\sigma_0^\alpha = 1_{d_3}$, which is why we can restrict the above models to

$$\left(\delta_t^\theta\right)_1 = \sum_{m=1}^{d_1'} \theta_m t^m, \qquad \left(\delta_t^\theta\right)_2 = \sum_{m=1}^{d_1''} \theta_m t^m, \qquad \rho_t^\phi = \sum_{m=1}^{d_2} \phi_m t^m, \qquad \sigma_t^\alpha = 1 + \sum_{m=1}^{d_3} \alpha_m t^m.$$

The regularity conditions of Assumption 2.14 are trivially fulfilled by polynomial motion functions. Consider for example (B8). The components of $t \mapsto \operatorname{grad}_\theta \left(\delta_t^\theta\right)_1\Big|_{\theta=\theta_0}$ are just the monomials $\partial\left(\delta_t^\theta\right)_1/\partial\theta_m = t^m$, for $1 \le m \le d_1'$, and as such are linearly independent functions in t. The same is true for drift in y-direction, rotation and scaling.

2.2 Two-step estimation procedure for image registration

In the following, we describe a method for the estimation of the drift, rotation, and scaling parameters θ_0, ϕ_0, and α_0 based on M-estimation. This means that we define certain functions (called contrast functionals) depending on the data, which are small for parameter values close to the true parameters. To obtain estimators for the motion function parameters, we therefore minimize these empirical contrast functionals with respect to θ, ϕ, and α. To benefit from the (generalized) shift properties of the Fourier transform and the Fourier-Mellin transform, we transfer the model first to the Fourier domain and later to the Fourier-Mellin space to carry out the estimation of the motion function parameters. The Fourier transform of the binned observations $O^t : j \mapsto O^t_j$ is given by

$$\mathcal{F}_{O^t}(\omega) = \frac{1}{n}\sum_{j=1}^{n} \mathrm{e}^{-2\pi\mathrm{i}\langle\omega, x_j\rangle} O^t_j = \frac{1}{n}\sum_{j=1}^{n} \mathrm{e}^{-2\pi\mathrm{i}\langle\omega, x_j\rangle}\left(f^t_j + \frac{1}{2\sqrt{\beta_T}}\epsilon^t_j\right).$$

Denoting with

$$W^t(\omega) := \frac{1}{2n\sqrt{\beta_T}}\sum_{j=1}^{n} \mathrm{e}^{-2\pi\mathrm{i}\langle\omega, x_j\rangle}\epsilon^t_j \tag{2.11}$$

the Fourier transform of the Gaussian error term, we define the Fourier model for motion estimation in SMS microscopy data as follows.

Definition 2.18 (Fourier Model). *For $t \in \mathbb{T}$ and $\omega \in \mathbb{R}^2$, with $W^t(\omega)$ from (2.11), we define*

$$Y^t(\omega) := \mathcal{F}_{O^t}(\omega) = \mathcal{F}_{f^t}(\omega) + W^t(\omega). \tag{2.12}$$

From the generalized shift property (2.5) we know that

$$\mathcal{F}_{f^t}(\omega) = (\sigma^{\alpha_0}_t)^2 \mathrm{e}^{-2\pi\mathrm{i}\left\langle\omega, \delta^{\theta_0}_t\right\rangle}\mathcal{F}_f(\sigma^{\alpha_0}_t R_{-\rho^{\phi_0}_t}\omega), \tag{2.13}$$

which implies

$$\left|\mathcal{F}_{f^t}(\omega)\right|^2 = (\sigma^{\alpha_0}_t)^4 \left|\mathcal{F}_f(\sigma^{\alpha_0}_t R_{-\rho^{\phi_0}_t}\omega)\right|^2. \tag{2.14}$$

Note that $\left|\mathcal{F}_{f^t}(\omega)\right|^2$ does not depend on the drift $\delta^{\theta_0}_t$. We aim to estimate the rotation parameter ϕ and the scaling parameter α from $\left\{\left|Y^t\right|^2\right\}_{t\in\mathbb{T}}$. Then, we can calibrate the images f^t with the estimated rotation and scaling, leaving only the drift to be estimated. Because of (2.12), the analytical Fourier-Mellin transform of $\left|Y^t\right|^2$ is

$$\begin{aligned}
\mathcal{M}_{|Y^t|^2}(u,v) &= \int_0^\infty\int_0^{2\pi} \mathrm{e}^{-2\pi\mathrm{i}u\psi} r^{-\mathrm{i}v} r^\gamma \left(\left|Y^t\right|^2 \circ \mathcal{P}\right)(r,\psi)\,\mathrm{d}\psi\,\frac{\mathrm{d}r}{r} \\
&= \int_0^\infty\int_0^{2\pi} \mathrm{e}^{-2\pi\mathrm{i}u\psi} r^{-\mathrm{i}v} r^\gamma \left(\left|\mathcal{F}_{f^t}\right|^2 \circ \mathcal{P} + \mathcal{W}^t \circ \mathcal{P}\right)(r,\psi)\,\mathrm{d}\psi\,\frac{\mathrm{d}r}{r} \\
&= \mathcal{M}_{\left|\mathcal{F}_{f^t}\right|^2}(u,v) + \mathcal{M}_{\mathcal{W}^t}(u,v),
\end{aligned} \tag{2.15}$$

where $\mathcal{P}$ is the polar coordinate transform and

$$\mathcal{W}^t(\omega) := \left|W^t(\omega)\right|^2 + 2\Re\left(\mathcal{F}_{f^t}(\omega)\overline{W^t(\omega)}\right). \tag{2.16}$$

We further define for a suitable cutoff $r_T \geq 1$ (see Assumption 2.15) the restricted version

$$\begin{aligned}\mathcal{M}^T_{|Y^t|^2}(u,v) :=& \int_0^{r_T}\int_0^{2\pi} \mathrm{e}^{-2\pi \mathrm{i} u\psi} r^{-\mathrm{i}v} r^{\gamma}\left(\left|Y^t\right|^2 \circ \mathcal{P}\right)(r,\psi)\,\mathrm{d}\psi\,\frac{\mathrm{d}r}{r}\\ =&\mathcal{M}^T_{\left|\mathcal{F}_{f^t}\right|^2}(u,v) + \mathcal{M}^T_{\mathcal{W}^t}(u,v).\end{aligned} \tag{2.17}$$

From the shift property of the analytical Fourier-Mellin transform (2.10) and (2.14), we get

$$\mathcal{M}_{\left|\mathcal{F}_{f^t}\right|^2}(u,v) = d_{u,v}(1/\sigma_t^{\alpha_0}, -\rho_t^{\phi_0})F^t(u,v), \tag{2.18}$$

where

$$d_{u,v}(\sigma,\rho) := \sigma^{-\mathrm{i}v}\mathrm{e}^{2\pi \mathrm{i} u\rho}, \quad \text{and} \quad F^t(u,v) := (\sigma_t^{\alpha_0})^{4-\gamma}\mathcal{M}_{|\mathcal{F}_f|^2}(u,v). \tag{2.19}$$

This, together with (2.15), gives the idea that (if the error terms get small)

$$d_{u,v}(\sigma_t^{\alpha_0},\rho_t^{\phi_0})\mathcal{M}_{|Y^t|^2}(u,v) \approx F^t(u,v).$$

Based on this, we define an estimator for the scaling and rotation parameters as a minimizer of a contrast functional, as defined below.

Definition 2.19 (Contrast functionals for rotation and scaling)**.** *For suitable Fourier cutoffs $u_T, v_T \geq 1$ (see Assumptions 2.15), we define the* empirical contrast functional for rotation and scaling*,*

$$\begin{aligned}&\widetilde{M}_T(\phi,\alpha)\\ &:= \int_{-v_T}^{v_T}\sum_{|u|\leq u_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\left|d_{u,v}(\sigma_t^{\alpha},\rho_t^{\phi})\mathcal{M}^T_{|Y^t|^2}(u,v) - \frac{\beta_T}{T}\sum_{t'\in\mathbb{T}} d_{u,v}\left(\sigma_{t'}^{\alpha},\rho_{t'}^{\phi}\right)\mathcal{M}^T_{|Y^{t'}|^2}(u,v)\right|^2\mathrm{d}v\\ &= M_T^0 + M_T(\phi,\alpha),\end{aligned}$$

with

$$\begin{aligned}M_T^0 &:= \int_{-v_T}^{v_T}\sum_{|u|\leq u_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\left|\mathcal{M}^T_{|Y^t|^2}(u,v)\right|^2\,\mathrm{d}v,\\ M_T(\phi,\alpha) &:= -\int_{-v_T}^{v_T}\sum_{|u|\leq u_T}\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}(\sigma_t^{\alpha},\rho_t^{\phi})\mathcal{M}^T_{|Y^t|^2}(u,v)\right|^2\,\mathrm{d}v,\end{aligned}$$

where we used $\left|d_{u,v}(\sigma,\rho)\right| = 1$.

Similarly, we define the population contrast functional for rotation and scaling,

$$
\begin{aligned}
\widetilde{M}(\phi,\alpha) &:= \int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \int_0^1 \Bigg| d_{u,v}\left(\frac{\sigma_t^{\alpha}}{\sigma_t^{\alpha_0}}, \rho_t^{\phi} - \rho_t^{\phi_0}\right) F^t(u,v) \\
&\qquad - \int_0^1 d_{u,v}\left(\frac{\sigma_{t'}^{\alpha}}{\sigma_{t'}^{\alpha_0}}, \rho_{t'}^{\phi} - \rho_{t'}^{\phi_0}\right) F^{t'}(u,v)\,\mathrm{d}t' \Bigg|^2 \mathrm{d}t\,\mathrm{d}v \\
&= M^0 + M(\phi,\alpha),
\end{aligned}
$$

with

$$
\begin{aligned}
M^0 &:= \int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \int_0^1 \left|F^t(u,v)\right|^2 \mathrm{d}t\,\mathrm{d}v, \\
M(\phi,\alpha) &:= -\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left| \int_0^1 d_{u,v}\left(\frac{\sigma_t^{\alpha}}{\sigma_t^{\alpha_0}}, \rho_t^{\phi} - \rho_t^{\phi_0}\right) F^t(u,v)\,\mathrm{d}t \right|^2 \mathrm{d}v.
\end{aligned}
$$

We now define the parameter estimators as minimizers of the empirical contrast functional. Since M_T^0 and M^0 are constant in (ϕ,α), we can equivalently minimize M_T or M, respectively.

Definition 2.20. *(Scaling and rotation parameter estimator) M-estimators for the rotation and scaling parameters ϕ and α are defined as*

$$
(\hat{\phi}_T, \hat{\alpha}_T) \in \operatorname*{argmin}_{(\phi,\alpha)\in \mathrm{A}\times\Phi} M_T(\phi,\alpha).
$$

The next step is to calibrate the Fourier data Y^t with those estimators, which leads to the following model. Note, that the following Definitions 2.21 and 2.22 are formulated for arbitrary $(\phi,\alpha) \in \Phi\times\mathrm{A}$, because we need to compute derivatives later to show asymptotic normality. However, we plug in $(\hat{\phi}_T, \hat{\alpha}_T)$ for the drift estimation (see Definition 2.23).

Definition 2.21 (Fourier model after rotation and scaling correction)**.** *Define the transformation (combining rotation and scaling)*

$$
\tau\colon \Phi\times\mathrm{A}\times[0,1] \to \mathbb{R}^{2\times 2}, \quad (\phi,\alpha,t) \mapsto \tau_t^{(\phi,\alpha)} := \frac{\sigma_t^{\alpha_0}}{\sigma_t^{\alpha}} R_{\rho_t^{\phi} - \rho_t^{\phi_0}}.
$$

Note that $\tau_t^{(\phi_0,\alpha_0)} = \mathrm{id}_{\mathbb{R}^2}$ for all $t\in[0,1]$. For $\phi\in\Phi$, $\alpha\in\mathrm{A}$, $\omega\in\Omega$, and $t\in[0,1]$, we define the drift correction term

$$
\begin{aligned}
h_{\omega}^{t,t'}&: \Theta\times\Phi\times\mathrm{A} \to \mathbb{C}, \\
&(\theta;\phi,\alpha) \mapsto \exp\left(2\pi\mathrm{i}\left(\left\langle (\sigma_t^{\alpha})^{-1} R_{\rho_t^{\phi}}\omega, \delta_t^{\theta} - \delta_t^{\theta_0}\right\rangle - \left\langle (\sigma_{t'}^{\alpha})^{-1} R_{\rho_{t'}^{\phi}}\omega, \delta_{t'}^{\theta}\right\rangle\right)\right), \qquad (2.20)
\end{aligned}
$$

the error term corrected for rotation and scaling

$$
V_T^t(\omega;\phi,\alpha) := (\sigma_t^{\alpha})^{-2} W^t\left(1/\sigma_t^{\alpha}\cdot R_{\rho_t^{\phi}}\omega\right)
$$

$$= \frac{1}{2n\sqrt{\beta_T}(\sigma_t^\alpha)^2} \sum_{j=1}^n \exp\left(-2\pi\mathrm{i}\left\langle 1/\sigma_t^\alpha \cdot R_{\rho_t^\phi}\omega, x_j\right\rangle\right)\epsilon_j^t, \tag{2.21}$$

and the Fourier data corrected for rotation and scaling,

$$\begin{aligned} Z_T^t(\omega;\phi,\alpha) &:= \left(\sigma_t^\alpha\right)^{-2} Y^t\left(\left(\sigma_t^\alpha\right)^{-1} R_{\rho_t^\phi}\omega\right) \\ &= h_\omega^{0,t}(\theta_0;\phi,\alpha)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha}\right)^2 \mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right) + V_T^t(\omega;\phi,\alpha). \end{aligned} \tag{2.22}$$

Similarly to the estimation of the rotation and scaling parameters, we minimize a contrast functional to estimate the true drift parameter θ_0.

Definition 2.22 (Contrast functionals for drift)**.** *For a suitable* cutoff $r_T \geq 1$ *as in (2.17) (see Assumption 2.15), let* $\Omega_T := \left\{\omega \in \mathbb{R}^2 \mid \|\omega\| < r_T\right\}$ *be the (open) Euclidean ball with center* $0 \in \mathbb{R}^2$ *and radius* r_T *and define the* empirical contrast functional (for drift)*,*

$$\begin{aligned} \widetilde{N}_T(\theta;\phi,\alpha) &:= \int_{\Omega_T} \frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\Bigg| h_\omega^{0,t}(\theta;\phi,\alpha)^{-1} Z_T^t(\omega;\phi,\alpha) \\ &\qquad - \frac{\beta_T}{T}\sum_{t'\in\mathbb{T}} h_\omega^{0,t'}(\theta;\phi,\alpha)^{-1} Z_T^{t'}(\omega;\phi,\alpha)\Bigg|^2 \,\mathrm{d}\omega \\ &= N_T^0(\phi,\alpha) + N_T(\theta;\phi,\alpha), \end{aligned}$$

with

$$\begin{aligned} N_T^0(\phi,\alpha) &:= \int_{\Omega_T} \frac{\beta_T}{T}\sum_{t\in\mathbb{T}} \left|Z_T^t(\omega;\phi,\alpha)\right|^2 \,\mathrm{d}\omega, \\ N_T(\theta;\phi,\alpha) &:= -\int_{\Omega_T}\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{0,t}(\theta;\phi,\alpha)^{-1} Z_T^t(\omega;\phi,\alpha)\right|^2 \,\mathrm{d}\omega, \end{aligned}$$

where we used $\left|h_\omega^{0,t}(\theta;\phi,\alpha)^{-1}\right| = 1$. *For notational purpose, let*

$$F_\omega^t(\theta;\phi,\alpha) := h_\omega^{t,0}(\theta;\phi,\alpha)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha}\right)^2 \mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right),$$

with $h_\omega^{t,0}$ *defined by (2.20). Note that*

$$h_\omega^{0,t}(\theta;\phi,\alpha)^{-1} Z_T^t(\omega;\phi,\alpha) = F_\omega^t(\theta;\phi,\alpha) + h_\omega^{0,t'}(\theta;\phi,\alpha)^{-1} V_T^{t'}(\omega;\phi,\alpha).$$

We define the population contrast functional (for drift)*,*

$$\begin{aligned} \widetilde{N}(\theta;\phi,\alpha) &:= \int_{\mathbb{R}^2}\int_0^1 \left|F_\omega^t(\theta;\phi,\alpha) - \int_0^1 F_\omega^{t'}(\theta;\phi,\alpha)\,\mathrm{d}t'\right|^2 \mathrm{d}t\,\mathrm{d}\omega \\ &= N^0(\phi,\alpha) + N(\theta;\phi,\alpha), \end{aligned}$$

where

$$N^0(\phi,\alpha) := \int_{\mathbb{R}^2}\int_0^1 \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\alpha}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right)\right|^2 \mathrm{d}t\,\mathrm{d}\omega,$$

$$N(\theta;\phi,\alpha) := -\int_{\mathbb{R}^2}\left|\int_0^1 F_\omega^t(\theta;\phi,\alpha)\,\mathrm{d}t\right|^2 \mathrm{d}\omega.$$

Similarly to the definition of $(\hat{\phi}_T, \hat{\alpha}_T)$, we ignore the parts of the contrast functional that are constant in θ. We will repeatedly use the following decompostion:

$$N_T(\theta;\phi,\alpha) = A_T(\theta;\phi,\alpha) + B_T(\theta;\phi,\alpha) + C_T(\theta;\phi,\alpha), \tag{2.23}$$

where

$$A_T(\theta;\phi,\alpha) := -\int_{\Omega_T}\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} F_\omega^t(\theta;\phi,\alpha)\right|^2 \mathrm{d}\omega$$

$$B_T(\theta;\phi,\alpha) := -\int_{\Omega_T} 2\Re\left[\left(\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} F_\omega^t(\theta;\phi,\alpha)\right) \cdot \left(\frac{\beta_T}{T}\sum_{t'\in\mathbb{T}} \overline{h_\omega^{0,t'}(\theta;\phi,\alpha)^{-1} V_T^{t'}(\omega;\phi,\alpha)}\right)\right]\mathrm{d}\omega$$

$$C_T(\theta;\phi,\alpha) := -\int_{\Omega_T}\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{0,t}(\theta;\phi,\alpha)^{-1} V_T^t(\omega;\phi,\alpha)\right|^2 \mathrm{d}\omega.$$

Now, we can define estimators for the drift parameter θ_0 and the unknown image f.

Definition 2.23 (Drift parameter estimator and image estimator)**.** *An M-estimator for the drift parameter θ is defined to be*

$$\hat{\theta}_T \in \operatorname*{argmin}_{\theta\in\Theta} N_T(\theta;\hat{\phi}_T,\hat{\alpha}_T).$$

Moreover, we define a preliminary estimator for f as the inverse Fourier transform of the calibrated Fourier data,

$$\hat{f}_T'(x_j) := \int_{\Omega_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} \mathrm{e}^{2\pi\mathrm{i}\langle\omega,x_j\rangle} h_\omega^{0,t}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right)^{-1} Z_T^t(\omega;\hat{\phi}_T,\hat{\alpha}_T)\,\mathrm{d}\omega, \quad j\in\{1,\ldots,n\}.$$

Recall that we still need to invert the variance stabilization transform in order to obtain an estimator for the actual marker density, leading to the following definition for the final image estimator:

$$\hat{f}_T(x_j) := \left(\hat{f}_T'(x_j)\right)^2, \quad j\in\{1,\ldots,n\}.$$

The two-step estimation method is summarized in Algorithm 1.

Algorithm 1 Motion correction using semiparametric M-estimation

Choose a bin size β_T, $\gamma > 0$, cutoffs $r_T, u_T, v_T \geq 1$ and parametric models for the motion functions δ^θ, ρ^ϕ and σ^α.

1. Given a sequence of observed frames average over β_T subsequent frames to obtain the binned frames $(O^t)_{t\in\mathbb{T}}$.
2. Apply the Anscombe transformation with constant $c = 1/4$, as described in Remark 2.2.
3. Approximate the squared Fourier magnitudes $\left|Y^t\right|^2$ by $|\mathcal{F}_{O^t}|^2$, $t \in \mathbb{T}$.
4. Calculate the Analytical Fourier-Mellin transform $\mathcal{M}_{|Y^t|^2}$, $t \in \mathbb{T}$.
5. Estimate the rotation and scaling parameters (ϕ_0, α_0) through minimizers $(\hat{\phi}_T, \hat{\alpha}_T)$ of the contrast functional $M_T(\phi, \alpha)$.
6. Correct Y^t for rotation and scaling and arrive at
$$Z_T^t(\omega; \hat{\phi}_T, \hat{\alpha}_T) = \left(\sigma_t^{\hat{\alpha}_T}\right)^{-2} Y^t\left(\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} R_{\rho_t^{\hat{\phi}_T}}\, \omega\right), t \in \mathbb{T}.$$
7. Estimate the drift parameter θ through a minimizer $\hat{\theta}_T$ of the contrast functional $N_T(\theta; \hat{\phi}_T, \hat{\alpha}_T)$.
8. Correct Z^t for drift.
9. Obtain an estimator $\hat{f}_T$ for the image by applying the inverse Fourier transform to the calibrated frames and inverting the Anscombe transform.

CHAPTER 3

Theoretical results

This chapter contains the main theoretical results including consistency of the motion parameter estimators and the final image estimator (Section 3.1) as well as joint asymptotic normality of the motion parameter estimators (Section 3.2). The results have been grouped thematically into statements on consistency and statements on distributional limits, since we proof a joint central limit theorem for all three motion function parameters. However, we need some of the outcomes of Section 3.2, namely asymptotic normality of the rotation and scaling parameters, already in Section 3.1 to show consistency of the drift parameter estimator. The implications are as follows:

Consistency of rotation and scaling parameters

$\Downarrow$

Asymptotic normality of rotation and scaling parameters

$\Downarrow$

Consistency of drift parameter

$\Downarrow$

Joint central limit theorem

For better readability only sketches of the proofs are included, and the full versions are moved to a separate final chapter, Chapter 6.

3.1 Consistency

Under the model assumptions formulated in 2.1.4 the estimators $(\hat{\theta}_T, \hat{\phi}_T, \hat{\alpha}_T)$ from Definitions 2.20 and 2.23 as well as the estimator $\hat{f}_T$ for f from Definition 2.23 are consistent.

Theorem 3.1 (Consistency of rotation and scaling parameters)**.** *Suppose that the Assumptions 2.13 (A2-A4), 2.14 (B1, B4-B5, and B7), and 2.15 hold. Then the rotation and scaling estimator $(\hat{\phi}_T, \hat{\alpha}_T)$ from Definition 2.20 is consistent, i.e.,*

$$(\hat{\phi}_T, \hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} (\phi_0, \alpha_0) \textit{ in probability.} \tag{3.1}$$

Sketch of proof. The proof follows a standard three step argument in M-estimation (e.g., van der

Vaart (2000); Gamboa et al. (2007); Bigot et al. (2009); Hartmann et al. (2015)). The three steps are:

1. Show the uniqueness of the population contrast minimizer (ϕ_0, α_0).
2. Show the continuity of the population contrast functional M.
3. Show that $M_T \overset{T\to\infty}{\longrightarrow} M$ in probability, uniformly over (ϕ, α).

Together with the compactness of $\Phi \times \mathrm{A}$, parts 1 and 2 ensure the condition that (ϕ_0, α_0) is a well separated point of minimum. Part 3 proves uniform convergence of the empirical contrast functional. Since $(\hat{\phi}_T, \hat{\alpha}_T)$ are defined as minimizers of M_T, the condition that $M_T(\hat{\phi}_T, \hat{\alpha}_T) \leq M_T(\phi_0, \alpha_0) + o_{\mathbb{P}}(1)$ is trivially fulfilled. Hence, the desired consistency follows directly from Theorem B.6. Note that in van der Vaart (2000), the theorem is formulated for a maximization problem. A detailed proof of the three steps can be found in Chapter 6. □

Theorem 3.2 (Consistency of the drift parameter)**.** *Suppose that the Assumptions 2.13, 2.14 (B1-B5, B7-B8) and 2.15 hold. If* $\sqrt{T}(\hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0)$ *is asymptotically centered normal, then the drift estimator* $\hat{\theta}_T$ *from Definition 2.23 is consistent, i.e.,*

$$\hat{\theta}_T \overset{T\to\infty}{\longrightarrow} \theta_0 \quad \textit{in probability.} \tag{3.2}$$

Sketch of proof. The proof of consistency of the drift parameter estimator follows the standard three step argument, which we used in the proof of Theorem 3.1, including the application of Theorem B.6. The three steps are here:

1. Show the uniqueness of the minimizer θ_0 of the population contrast $N(\,\cdot\,; \phi_0, \alpha_0)$.
2. Show the continuity in θ of the population contrast functional $N(\,\cdot\,; \phi_0, \alpha_0)$.
3. Show that $N_T(\theta; \hat{\phi}_T, \hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} N(\theta; \phi_0, \alpha_0)$ in probability, uniformly over θ.

The proofs of the three steps are very similar to the reasoning used in the proof of Theorem 3.1 and are deferred to Chapter 6, too. □

Theorem 3.3 (Consistency of the image estimator)**.** *Under the Assumptions 2.13, 2.14 and 2.15 the image estimator* $\hat{f}_T$ *from Definition 2.23 is consistent, i.e.,*

$$\left\|\hat{f}_T - f\right\|_{L^2} \overset{T\to\infty}{\longrightarrow} 0 \quad \textit{in probability.} \tag{3.3}$$

Sketch of proof. Using the Plancherel equality (Theorem B.2) we show that the difference $\left\|\hat{f}'_T(x_j) - f\right\|^2_{L^2}$ vanishes asymptotically, in probability, where f' denotes the transformed marker density with integrated square root (see the model derivation in Section 2.1). By the continuous mapping theorem (see, e.g., Theorem 2.3 in van der Vaart (2000)), we can conclude convergence of the final image estimator $\hat{f}_T$, for which the variance stabilizing transformation has been inverted, to the original (scaled) marker density $f^0 \cdot p$. A detailed argument can be found in Chapter 6. □

Theorem 3.4 (Consistency)**.** *Under the Assumptions 2.13, 2.14, and 2.15, we have that*

(i) $(\hat{\theta}_T, \hat{\phi}_T, \hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} (\theta_0, \phi_0, \alpha_0)$ *in probability,*

(ii) $\left\|\hat{f}_T - f^0 \cdot p\right\|_{L^2} \overset{T\to\infty}{\longrightarrow} 0$ *in probability.*

Outline of proof. The proof of part (i) is split up into the proof of consistency of the rotation and scaling parameter estimators (Theorem 3.1), and the consistency of the drift parameter estimator (Theorem 3.2). The proofs of these two theorems have the same structure and both rely mainly upon standard argumentation for M-estimators, as stated, e.g., in Theorem B.6. The proof of consistency of the drift parameter estimator further uses asymptotic normality of the rotation and scaling parameter estimators as given by Theorem 3.5 in the subsequent Section 3.2. The joint consistency of all three motion parameters then follows directly with Theorem 2.7 from van der Vaart (2000). The consistency of the image estimator, part (ii), is proven in Theorem 3.3. □

3.2 Asymptotic normality

Theorem 3.5 (Central limit theorem for rotation and scaling parameters)**.** *Suppose that Assumptions 2.13 (A2-A4), 2.14 (B2-B4, B8), and 2.15 hold. Then,*

$$\sqrt{T}\begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} = -\sum_{t'\in\mathbb{T}}\sum_{j\in J_T^{t'}} H_M^{-1} w_j^{t'} \epsilon_j^{t'} + o_{\mathbb{P}}(1),$$

with $H_M = \mathrm{Hess}_{(\phi,\alpha)} M(\phi_0, \alpha_0)$ *from Lemma 6.11 and weights* $w_j^{t'} \in \mathbb{R}^{d_2+d_3}$ *from Theorem 6.10. In particular,*

$$\sqrt{T}\begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0, H_M^{-1}\Sigma_{RS} H_M^{-1}) \quad \text{in distribution,}$$

with Σ_{RS} *given explicitly in Definition 6.9.*

Sketch of proof. For better readability, the detailed proof is again deferred to Chapter 6 and only a brief sketch of the single steps is given in the following. Likewise, the exact expression of the weights and the covariance matrix of the limiting distribution can be found in Chapter 6.
The first step is to show that the gradient of the empirical contrast functional converges in probability to a linear combination of the independent error terms. In particular, $\sqrt{T}\mathrm{grad}_{(\phi,\alpha)} M_T(\phi_0, \alpha_0)$ asymptotically follows a normal distribution, see Theorem 6.10. In Lemma 6.11 we prove that, under some assumptions, the Hessian of the population contrast functional at the true parameters is invertible, and in Theorem 6.12 we see that the Hessian of the empirical contrast functional converges in probability toward the Hessian of the population contrast functional, i.e., $\mathrm{Hess}_{(\phi,\alpha)} M_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) \overset{T\to\infty}{\longrightarrow} H_M$ in probability for all sequences such that $(\hat{\phi}_T^*, \hat{\alpha}_T^*) \overset{T\to\infty}{\longrightarrow} (\phi_0, \alpha_0)$ in probability. Finally, using the differentiability assumption on the contrast functionals in a neighborhood of the true parameters (as an implication of Assumption 2.14 (B3)), the mean

value theorem and Lemma A.12 on uniform tightness of sequences of random variables, we can combine the previous results to obtain asymptotic normality for the rotation and scaling parameter estimators. □

Theorem 3.6 (Central limit theorem for the drift estimator)**.** *Suppose that the Assumptions 2.13, 2.14, and 2.15 hold. Then,*

$$\sqrt{T}\left(\hat{\theta}_T - \theta_0\right) \overset{T\to\infty}{\longrightarrow} \mathcal{N}\left(0, H_N^{-1}\Sigma_D H_N^{-1}\right),$$

with H_N from Lemma 6.15, and some covariance matrix Σ_D, which can be expressed in terms of the covariance matrices Σ_{RS} and $\widetilde{\Sigma}$ appearing in Theorems 3.5 and 6.14.

Sketch of proof. We first show in Theorem 6.13 that the mixed derivatives of the empirical contrast functional for drift converge to the derivatives of the population contrast functional. Similarly to before, we prove in Theorem 6.14 that the gradient of the contrast functional for drift, $\mathrm{grad}_\theta N_T(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)$, is a linear combination of the error terms with an additional vanishing term. In Lemma 6.15 we show that the Hessian $H_N := \mathrm{Hess}_\theta N(\theta_0; \phi_0, \alpha_0)$ is invertible, and in Theorem 6.16 that it converges in probability: $\mathrm{Hess}_\theta N_T(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} H_N$ for any sequence $\hat{\theta}_T^* \overset{T\to\infty}{\longrightarrow} \theta_0$ converging in probability. We bring these results together and apply the mean value theorem to show that $\sqrt{T}(\hat{\theta}_T - \theta_0)$ can be written as a linear combination of the aforementioned independent Gaussian random variables with an additional vanishing term. The details of this proof are deferred to Chapter 6. □

Theorem 3.7 (Joint central limit theorem)**.** *Under the Assumptions 2.13, 2.14, and 2.15, we have with some covariance matrix Σ, which is explicitly given in Chapter 6, that*

$$\sqrt{T}\begin{pmatrix} \hat{\theta}_T - \theta_0 \\ \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0,\Sigma) \quad \textit{in distribution,}$$

i.e., the estimators for the motion function parameters are jointly asymptotically normally distributed.

Outline of proof. For the proof of consistency of the drift parameter estimators in Theorem 3.4, we already used the result from Theorem 3.5, that $\sqrt{T}(\hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0)$ can be written as a linear combination of the independent Gaussian errors ϵ_j^t and an additional vanishing term. The crucial step is then to show that $\sqrt{T}(\hat{\theta}_T - \theta_0)$ can be expressed as a linear combination of *the same* Gaussian variables and some additional vanishing term in Theorem 3.6. The linear combinations of the rotation and scaling parameter estimators and of the drift parameter estimators are then combined into one, establishing convergence of the joint distribution and thereby concluding the proof of Theorem 3.7. □

CHAPTER 4

Application: Simulation study and real SMS data

To illustrate the applicability of our method, we show the results of a simulation study using polynomial models for the motion functions in Section 4.1. Moreover, our reconstruction technique is applied to real SMS data and compared with calibration using fiducial markers in Section 4.2. We show that our method is competitive to this current approach revealing the incorporation of fiducials as redundant in the analysis and processing of SMS images. Finally, simple bootstrap confidence bands for the drift rotation and scaling estimators are constructed in Section 4.3, quantifying the statistical uncertainty.

4.1 Simulation study

In order to validate our method and to demonstrate its applicability, we conducted a simulation study. The image displayed in Figure 4.1 is used as the true underlying structure $f : [0,1]^2 \to [0,1]$. As we would like the simulations to be comparable to our motivating real data example, we chose the true motion function parameters such that the total amount of drift, rotation and scaling are of similar size to the ones observed in the SMS microscopy data, which is analyzed in Section 4.2. We used a 256×256 pixel grid for the gray scale images and $T = 200$ binned frames for the simulation runs, corresponding to a supposed binning size of about $\beta_T = 150$. This binning size is close to the square root of the typical total number of recorded frames, which we found to be a suitable size to balance the retained motion blur and the noise level in the averaged frames well. For comparison, we also included simulation runs using only $T = 100$ binned frames and ones with a 128×128 pixel grid.

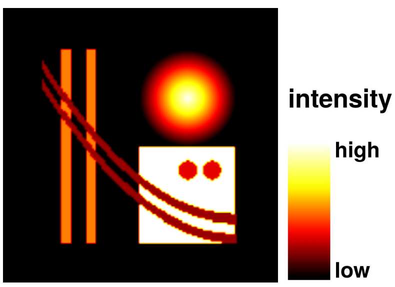

Figure 4.1: True underlying test image with intensities in [0, 1] displayed as colors ranging from black (0) over red and yellow to white (1).

Data generation. We simulated two different statistical models. Firstly, we considered observations generated as Gaussian random variables corresponding to our theoretical model, which is an approximate model for the SMS data. Secondly, we generated the binned frames as arrays of Poisson binomially distributed random variables. As described in Section 2.1, this is the model which is closest to the idealized data acquisition process. The results show that the estimated parameter values and the image reconstructions differ only slightly from the ones obtained using the Gaussian model. This empirically validates the theoretically justified normal approximation. To be more specific on how the artificial observations are constructed, for the Gaussian case we generate independently distributed observations

$$O_j^t \sim \mathcal{N}\left(f_j^t, \frac{1}{4\beta_T}\right), \text{ for } j = 1 \dots n,\ t \in \mathbb{T},$$

where $f_j^t = \sqrt{\tilde{f}_j^t \cdot p}$ for the detection probability p and the average marker density $\tilde{f}_j^t$ from Section 2.1. The square root is induced by the variance stabilizing transformation in the Gaussian approximation. In the binomial case, we have the independently distributed binned observations

$$O_j^t = \frac{1}{\beta_T} P_j^t, \quad \text{with} \quad P_j^t \sim \text{PoiBin}\left(\left(f_j^{t+i/T} \cdot p\right)_{i=0}^{\beta_T - 1}\right), \text{ for } j = 1 \dots n,\ t \in \mathbb{T}.$$

Here, $\text{PoiBin}\left(\left(f_j^{t+i/T} \cdot p\right)_{i=0}^{\beta_T - 1}\right)$ denotes the Poisson binomially distributed random variable with probability vector $\left(f_j^{t+i/T} \cdot p\right)_{i=0}^{\beta_T}$, i.e., the sum of β_T independent Bernoulli random variables having success probabilities $f_j^{t+i/T} \cdot p$, $0 \leq i \leq \beta_T - 1$, resp. We do not need the square root in this case, since the Anscombe type transformation will be performed during the reconstruction process in this model, just like in the application of the method to real data.

For each of the statistical models, we present the results of two different parametrical models for the motion functions. Both are polynomial models, as introduced in Example 2.17, namely using linear and quadratic motion functions. For the linear model, the drift vector at time $t \in \mathbb{T} = \{i/T : 0 \leq i \leq T-1\}$ is given by

$$\delta_t^{\theta_0} = \theta_0 \cdot t \in \mathbb{R}^2,$$

for a true drift parameter $\theta_0 = ((\theta_0)_1, (\theta_0)_2)^\top \in \Theta \subset \mathbb{R}^2$, and analogously we have the one-dimensional rotation angle,

$$\rho_t^{\phi_0} = \phi_0 \cdot t, \text{ with } t \in \mathbb{T},\ \phi_0 \in \Phi,$$

and scaling factor,

$$\sigma_t^{\alpha_0} = 1 + \alpha_0 \cdot t, \text{ with } t \in \mathbb{T}, \alpha_0 \in \mathrm{A}.$$

Here, the parameter spaces Θ, Φ and A are chosen in such a way, that we only consider sensible parameter choices, see also Remark 2.1. For the drift we chose $\Theta = [-1,1]^2$, ensuring that $-1 \leq \left(\delta_t^{\theta_0}\right)_1, \left(\delta_t^{\theta_0}\right)_2 \leq 1$ for all $t \in \mathbb{T}$, which means that the object of interest moves at most as far as the width and height of the observation window. We do not need to consider other parameter values for the drift, since for farther drift, the structure moves out of the imaged area and can not be registered any more. We have argued already in Remark 2.1, why it is reasonable that we restrict ourselves to rotation angles $\rho_t^{\phi} \in (-\pi/2, \pi/2)$ and scaling factors $\sigma_t^{\alpha} \in [\sigma_{\min}, \sigma_{\max}]$. For the linear model, we can simply choose $\Phi = (-\pi/2, \pi/2)$ and $\mathrm{A} = [\sigma_{\min} - 1, \sigma_{\max} - 1]$ in order to ensure that these constraints hold. Appropriate values for the boundaries for the scaling factor are $\sigma_{\min} = 1/256$ and $\sigma_{\max} = 2$, induced by the relative pixel size and the size of the observation window around the imaged structure.

In the quadratic setting, we get analogous expressions for the motion functions, namely for $\theta_0 \in \Theta \subset \mathbb{R}^4$, $\phi_0 \in \Phi \subset \mathbb{R}^2$, $\alpha_0 \in \mathrm{A} \subset \mathbb{R}^2$, and $t \in \mathbb{T}$,

$$\begin{aligned}
\delta_t^{\theta_0} &= \begin{pmatrix} (\theta_0)_1 \\ (\theta_0)_3 \end{pmatrix} \cdot t + \begin{pmatrix} (\theta_0)_2 \\ (\theta_0)_4 \end{pmatrix} \cdot t^2 \in [0,1]^2, \\
\rho_t^{\phi_0} &= (\phi_0)_1 \cdot t + (\phi_0)_2 \cdot t^2 \in (-\pi/2, \pi/2), \\
\sigma_t^{\alpha_0} &= 1 + (\alpha_0)_1 \cdot t + (\alpha_0)_2 \cdot t^2 \in [\sigma_{\min}, \sigma_{\max}].
\end{aligned}$$

To translate these conditions into conditions on the parameters, consider for example drift in x-direction. It is necessary that the endpoints are contained in $[-1,1]$. This is fulfilled if $(\theta_0)_1 + (\theta_0)_2 \in [-1,1]$. We further need that the value at the extreme point is contained in $[-1,1]$ if it is attained for some $t \in [0,1]$. Hence, we want the following implication to hold:

$$-\frac{(\theta_0)_1}{2\,(\theta_0)_2} \in (0,1) \quad \Longrightarrow \quad -\frac{(\theta_0)_1^2}{4\,(\theta_0)_2} \in [-1,1].$$

Conditions for the parameters of the other motion types can be obtained in an analogous way.

Estimation and reconstruction. After creating the observations in the described way, we applied our reconstruction method to the artificial dataset. As described in Section 2.2, estimators for the motion function parameters are obtained as minimizers of the empirical contrast functionals given there. These minimizers are determined by a standard Nelder-Mead-type algorithm, which is preimplemented in the statistical software R. As initial value for the optimization in drift estimation we used $0 \in \mathbb{R}^{d_1}$. For the estimation of the rotation and scaling parameters, we used the vector with components equal to 0.5, since we found that the built-in optimization routine of R tends to never leave $0 \in \mathbb{R}^{d_2+d_3}$ if this is used as starting value. The Fourier transform and Fourier-Mellin transform are computed using the fast Fourier transform algorithm (FFT, see, e.g., Cooley and Tukey, 1965). This is possible, since the Fourier-Mellin transform is a Fourier-type transform as explained in Remark 2.8. Due to discretization and the relatively small total rotation angle, we will be able to see the rotation in the Fourier domain only for rather large structures in the original image. The reason for this is that only for long objects, a rotation will result in a shift of parts of the object to a new pixel. For small features, rotation will appear as a translation of the whole feature. Hence, we choose a quite small cutoff parameter of 16 for the Fourier type coefficients to avoid misinterpreting rotation as drift, which would result in an underestimation of the rotation and overestimation of drift.

Results. The estimated values for the motion function parameters are summarized in Table 4.1. Table 4.2 shows the mean values for the parameters calculated from 100 independent simulation runs using the same underlying marker density and true motion function parameters.

To quantify how reliably our method is able to correct for the motion blur, we consider the square root of the mean squared error (RMSE) $\mathbb{E}\left(\left\|(\theta_0, \phi_0, \alpha_0) - \left(\hat{\theta}_T, \hat{\phi}_T, \hat{\alpha}_T\right)\right\|^2\right)$, calculated from the 100 simulation runs we used to compute the mean parameter values in Table 4.2. We further report the RMSE of the single motion types, $\mathbb{E}\left(\left\|\left(\theta_0 - \hat{\theta}_T\right)_x\right\|^2\right)$, and likewise for drift in y-direction, rotation and scaling in order to gain insight into what motion types are most difficult to estimate correctly. The corresponding results are displayed in Table 4.3.

In the linear model, the rotation angle seems to be the most difficult to estimate correctly. The scaling factor is reconstructed quite well already. Yet, for the drift parameter estimator we obtain even lower RMSE-values. This indicates, that translational movement can be detected and removed best with our method, although the drift correction is performed on data which have been corrected for rotation and scaling by the slightly less accurate estimators for these motion types. Despite the little differences in accuracy, it can be seen that we also obtain quite good results for the combined estimation of all three motion types. RMSE-values for the quadratic setting are generally slightly higher than for the linear setting. Here, the scaling factor and the drift in y-direction yield the highest RMSE values.

In Figure 4.2 we present the resulting final image estimator for linear motion functions, and in Figure 4.3 for quadratic motion functions. Both figures show the results obtained using $T = 200$ binned frames and a 256×256 pixel grid. On the left is the original image for comparison with the reconstructions. The middle column holds the results obtained in the Gaussian setting, the right column the ones for the Poisson binomial setting. We show a single binned frame each in the first row, i.e., a sparse subsample of the underlying marker density visualized in Figure 4.1. In the second row, the superpositions of all frames with a clearly visible motion blur is displayed. The third row contains the final reconstructions, which have greatly improved resolution compared to the original superimposed SMS image and capture the main features of the underlying structure. The reconstructions in the linear model are almost identical with the original image. In the quadratic setting some deformation remains, but the blurring is still significantly reduced compared to the superimposed images. The visual inspection confirms therefore, that the estimation works better for the linear model. The reconstructions in the quadratic setting are still satisfying, nonetheless. Even better results can be obtained in both motion models by calculating the average over the final images estimates from all 100 simulation runs, as shown in the last row.

However, the increase in quality by averaging over multiple simulation runs is costly in terms of runtime. Even though the reconstruction itself only takes between 20 seconds for a 128×128 pixel grid and $T = 100$ binned frames and 2 minutes for a 256×256 pixel grid and $T = 200$

binned frames on a standard laptop, the data generation is computationally intensive and may lead to a runtime of up to an hour for a single run using a 256×256 pixel grid. Moreover, decreasing the number of binned frames T reduces the time needed for the reconstruction, but has no effect on the run time of the data generation process, since we still have to compute the marker densities for all original time points in order to reproduce the small bias retained in the binned frames. As the data generation mainly governs the total run time of the simulation, we can hardly gain any speed by using a smaller T. To obtain results faster, the number of pixels may be reduced at the cost of a lower total resolution and possibly more difficulties in estimating the rotational motion. Appendix C contains reconstructions using $T = 100$ frames and a pixel grid of size 128×128.

			linear motion	quadratic motion
true parameter $(\theta_0; \phi_0; \alpha_0)$			(0.059, 0.041; 0.039; 0.001)	(0.029, 0.029, 0.016, 0.039; 0.026, 0.031; 0.001, 0.001)
statistical model	pixels	T	$(\hat{\theta}_T; \hat{\phi}_T; \hat{\alpha}_T)$	$(\hat{\theta}_T; \hat{\phi}_T; \hat{\alpha}_T)$
Gaussian	256	100	(0.059, 0.046; 0.026; -0.003)	(0.002, 0.053, 0.012, 0.024; 0.069, 0.083; 0.028, -0.028)
		200	(0.059, 0.041; 0.027; -0.001)	(0.015, 0.041, 0.015, 0.023; 0.082, 0.064; 0.037, -0.045)
	128	100	(0.059, 0.051; 0.011; -0.013)	(0.058, 0.005, 0.039, 0.017; 0.058, 0.016; -0.064, 0.070)
		200	(0.059, 0.041; 0.015; -0.001)	(0.032, 0.026, 0.034, 0.029; -0.005, 0.042; -0.051, 0.064)
Poisson binomial	256	100	(0.059, 0.032; 0.037; -0.001)	(0.003, 0.043, -0.011, 0.030; 0.082, 0.075; 0.027, -0.025)
		200	(0.059, 0.037; 0.007; -0.001)	(0.010, 0.041, -0.007, 0.025; 0.068, 0.085; 0.032, -0.034)
	128	100	(0.059, 0.041; 0.011; 0.001)	(0.034, 0.021, 0.019, 0.034; -0.050, 0.059; -0.068, 0.080)
		200	(0.059, 0.041; 0.021; -0.002)	(0.010, 0.041, 0.012, 0.029; 0.067, 0.004; 0.053, -0.053)

Table 4.1: Estimated parameter $(\hat{\theta}_T; \hat{\phi}_T; \hat{\alpha}_T)$ for one simulation run. Results of Gaussian and Poisson binomial models with linear and quadratic motion functions for a total number of binned frames of $T = 100$ and $T = 200$ on 128×128 and 256×256 pixel grids.

			linear motion	quadratic motion
true parameter $(\theta_0; \phi_0; \alpha_0)$			(0.059, 0.041; 0.031; 0.01)	(0.029, 0.029, 0.016, 0.039; 0.026, 0.031; 0.01, 0.01)
statistical model	pixels	T	mean of $(\hat{\theta}_T; \hat{\phi}_T; \hat{\alpha}_T)$	mean of $(\hat{\theta}_T; \hat{\phi}_T; \hat{\alpha}_T)$
Gaussian	256	100	(0.059, 0.042; 0.024; -0.001)	(0.011, 0.045, 0.012, 0.025; 0.063, 0.078; 0.034, -0.032)
		200	(0.059, 0.043; 0.024; -0.001)	(0.018, 0.039, 0.015, 0.028; 0.065, 0.056; 0.34, -0.042)
	128	100	(0.059, 0.050; 0.019; -0.012)	(0.037, 0.019, 0.025, 0.027; 0.002, 0.051; -0.007, 0.004)
		200	(0.059, 0.044; 0.038; -0.001)	(0.035, 0.022, 0.027, 0.027; 0.0312, 0.031; -0.045, 0.057)
Poisson binomial	256	100	(0.059, 0.039; 0.004; 0)	(0.006, 0.042, -0.008, 0.028; 0.074, 0.080; 0.073, 0.042)
		200	(0.059, 0.040; 0.003; 0)	(0.008, 0.041, -0.006, 0.027; 0.072, 0.071; 0.028, -0.028)
	128	100	(0.057, 0.038; 0.015; 0.002)	(0.031, 0.024, 0.020, 0.017; 0.035, 0.065; -0.034, 0.046)
		200	(0.058, 0.039; 0.018; 0.001)	(0.007, 0.041, 0.005, 0.027; 0.059, 0.033; 0.046, -0.045)

Table 4.2: Setting as in Table 4.1. Means of estimated parameter values $(\hat{\theta}_T; \hat{\phi}_T; \hat{\alpha}_T)$ from 100 simulation runs.

statistical model	pixels	T	linear motion RMSE$(\hat{\theta}_T; \hat{\phi}_T; \hat{\alpha}_T)$ (RMSE$(\hat{\theta}_T)$; RMSE$(\hat{\phi}_T)$; RMSE$(\hat{\alpha}_T)$)	quadratic motion RMSE$(\hat{\theta}_T; \hat{\phi}_T; \hat{\alpha}_T)$ (RMSE$(\hat{\theta}_T)$; RMSE$(\hat{\phi}_T)$; RMSE$(\hat{\alpha}_T)$)
Gaussian	256	100	1e-2 (4e-4, 2e-3; 1e-2; 2e-3)	8e-2 (2e-2, 6e-2; 2e-2; 6e-2)
		200	1e-2 (8e-4, 3e-3; 1e-2; 3e-3)	8e-2 (1e-2, 5e-2; 1e-2; 5e-2)
	128	100	3e-2 (6e-4, 1e-2; 2e-2; 1e-2)	9e-2 (2e-2, 6e-2; 2e-2; 6e-2)
		200	3e-2 (1e-3, 7e-3; 3e-2; 7e-3)	9e-2 (2e-2, 6e-2; 2e-2; 6e-2)
Poisson binomial	256	100	3e-2 (8e-4, 3e-3; 3e-2; 2e-3)	9e-2 (2e-2, 5e-2; 2e-2; 5e-2)
		200	3e-2 (7e-4, 2e-3; 3e-2; 2e-3)	1e-1 (2e-2, 6e-2; 2e-2; 6e-2)
	128	100	3e-2 (4e-3, 6e-3; 3e-2; 5e-3)	1e-1 (1e-2, 7e-2; 1e-2; 6e-2)
		200	3e-2 (3e-3, 7e-3; 3e-2; 5e-3)	1e-1 (2e-2, 8e-2; 2e-2; 8e-2)

Table 4.3: Setting as in Table 4.1. Empirical values of the RMSE of $\left(\hat{\theta}_T, \hat{\phi}_T, \hat{\alpha}_T\right)$ and of the components, $\hat{\theta}_T$, $\hat{\phi}_T$, and $\hat{\alpha}_T$ computed from 20 simulation runs are displayed.

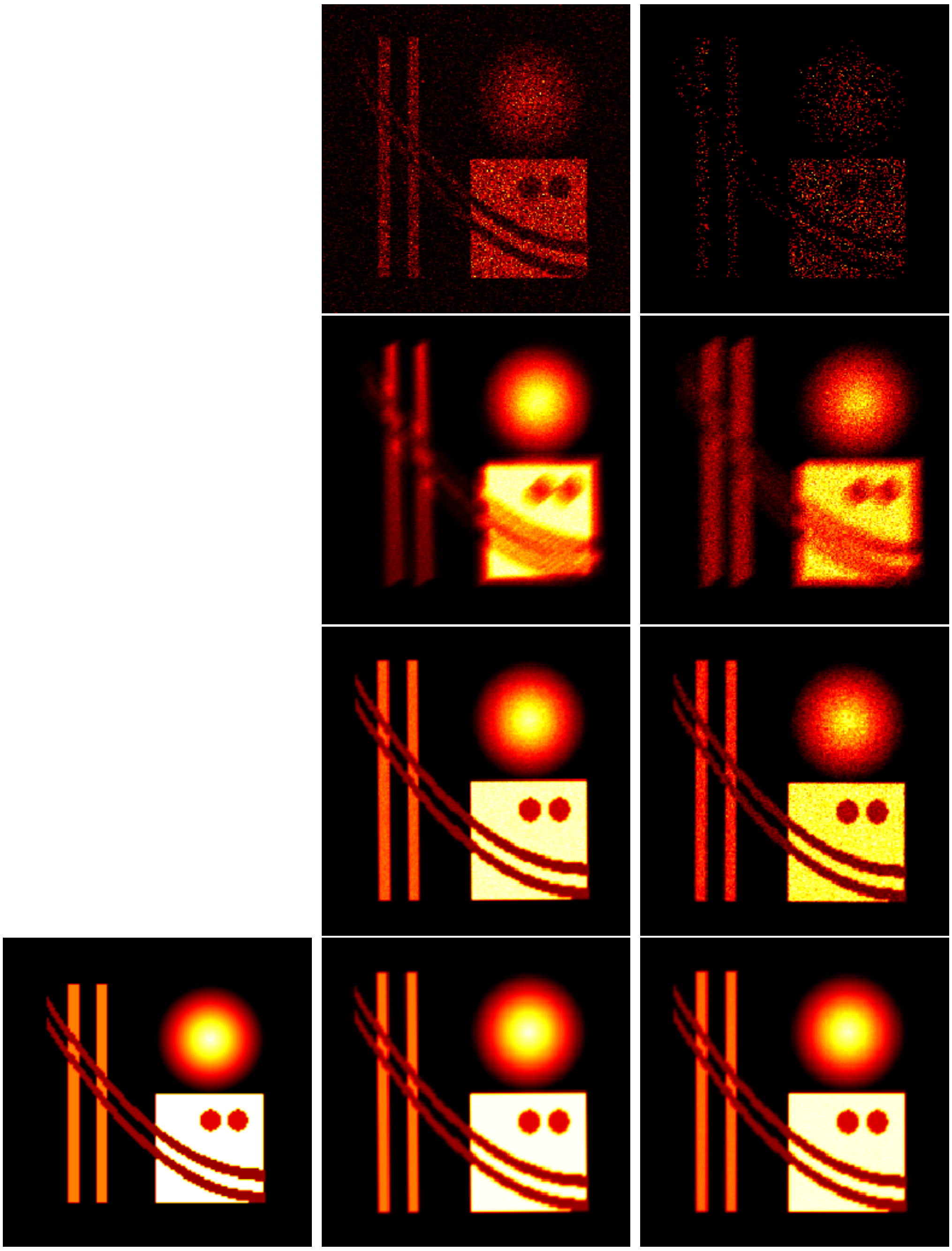

Figure 4.2: Image reconstructions of the simulation study for linear motion model with $T = 200$ binned frames on a 256×256 pixel grid: true underlying image (left); for Gaussian (middle column) and Poisson binomial (right column) model, a single binned frame (first row), the blurred superpositions of all frames (second row), final image estimates, which are corrected for rotation, scaling and translational drift (third row) and average over images estimates from 100 simulation runs (fourth row).

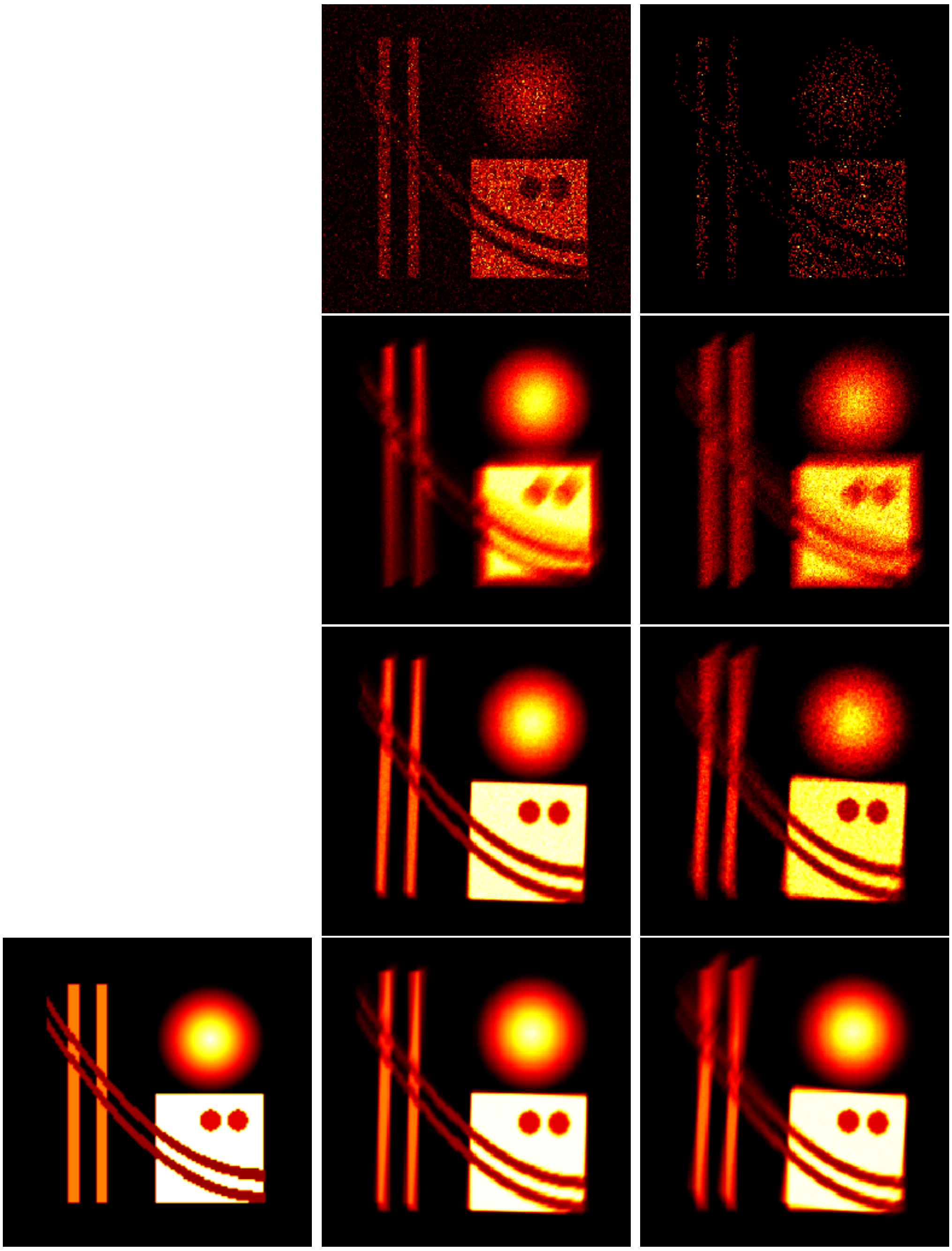

Figure 4.3: Image reconstructions of the simulation study for quadratic motion model with $T = 200$ binned frames on a 256×256 pixel grid: true underlying image (left); for Gaussian (middle column) and Poisson binomial (right column) model, a single binned frame (first row), the blurred superpositions of all frames (second row), final image estimates, which are corrected for rotation, scaling and translational drift (third row) and average over images estimates from 100 simulation runs (fourth row).

4.2 Application to SMS data

In this section, we demonstrate how our method can be applied to estimate and correct for the motion blur in SMS microscopy data. We first briefly describe the experimental setup in which the data have been obtained, as explained by Dr. Oskar Laitenberger, who conducted the measurements. Afterwards, we discuss suitable choices for the types of the motion function and other model parameters. Finally, we show the results of our reconstruction method and compare it to the ones obtained with conventional fiducial marker tracking (for a description of the fiducial marker tracking procedure see Algorithm 8.4 in Hartmann (2016)). Note that in order to enable this comparison, the data have been obtained with fiducial markers included into the sample. We want to demonstrate the validity of our method without the additional information stemming from these traceable markers. For this reason, we delete the signal originating from the fiducial markers and perform the reconstruction only on the remaining observations.

Setup. A modified Leica DMIRE2 body was used, which was equipped with an oil immersion objective (UPLSAPO 100XO of Olympus) and a self-constructed stable sample holder ensuring that the sample drift is well below the expected average localization accuracy. A dichroic mirror (545 DCXRUV reflection 360–535 nm > 90%, transmission 555–750 nm > 90% of AHF Analysentechnik AG) separated the fluorescence light of Rhodamine B (Belov et al., 2009) from the excitation light of wavelength 532 nm generated by a continuous wave laser (HB-Laser Germany, model LC-LS-532-1.2W). Furthermore, the fluorescence light passed a bandpass filter ET 560/40 (transmission 544-578 nm > 90% of AHF Analysentechnik AG) right in front of the EMCCD-camera (iXon X3 of Andor). To control the number of events per frame, a continuous wave UV laser generating 371 nm (Coherent Cube 371nm/16mW) was used. The named sample holder is mounted on top of a translation (SLC of SmarAct) and rotation stage (RVS80CC of Newport).
In the imaging process a series of $T = 29000$ single frames was taken over a time period of 600 s. This corresponds to an exposition time of about 20 ms for each frame. During the measurement, a controlled rotation with linear angle velocity and maximal displacement of 1.4° was applied to the sample, a stained β-tubulin in HeLa cells. Moreover, the sample was subjected to small uncontrolled translations caused by vibrations and possibly a small scaling due to heating. Usually, the rotation center is not exactly in the middle of the field of view. This introduces an additional translation with a trigonometrical component, which can be approximated by a linear model well enough because of the small total rotational displacement.

Model and parameter choices. Due to the relatively small total displacement, movement can be approximated reasonably well by linear functions and still produce very satisfactory reconstructions. However, other models, using for example splines, might lead to even better results (see Chapter 5). For the given data set a linear model is adequate. As explained in the previous section on the simulation study (Section 4.1) we chose a relatively small value of 16 for the Fourier cutoff to avoid misinterpretating rotation as drift. From the 29000 initial frames we obtain $T = 200$ binned frames, on which the calibration method is performed. With a value of 145, the corresponding bin size is not far from the square root of the total number of frames in order to balance reduction of noise level and retained motion blur well. We used a pixel grid of size 512×507, corresponding to pixel size of about 45 nm in both directions (about twice the localization precision).

Results. The obtained values for the motion function parameter estimators are

$$\hat{\theta}_T = (-0.025, -0.015)^\top$$
$$\hat{\phi}_T = 0.012$$
$$\hat{\alpha}_T = -0.0003,$$

which corresponds to a maximal drift of about 13 pixels in negative x-direction, 7 pixels in negative y-direction, a maximal rotational displacement of about 0.7° and an reduction in size of about 0.03 percent. These values are not too far from what we would expect. With the reference method of fiducial marker tracking, we obtained a maximal drift of about 13 pixels in negative x-direction, 8 pixels in negative y-direction, a maximal rotational displacement of about 1.3° and a size reduction of about 0.1 percent. The rotation angle for fiducial marker tracking is closer to the externally applied rotation movement of 1.4°. However, because of the small values of the total rotation angle and scaling factor, the slight discrepancies do not have a significant effect on the reconstructions.

In Figure 4.4, we show the reconstructed images obtained in our data analysis. We present the superposition of all single frames on the left, the final image estimator produced by our correction method in the middle and the reference image, where the motion blur was removed using fiducial marker tracking, on the right. The second row shows a zoom in for each of the above images. It can be seen very well that our reconstruction gives great improvement in resolution compared to the original superimposed SMS image. Comparing the zoomed in images shows particularly well that after our motion correction single filament strands can be distinguished, whereas in the original SMS image on the left only a large region containing signal can be identified. Furthermore, the reconstruction is of a quality at least as high as the one obtained using fiducial marker tracking. Note that there are different ways to implement fiducial marker tracking. Better results can be obtained, additional knowledge about the values of the parameters and motion function types is used (see Laitenberger (2018)).

Remark 4.1 (Variation of the detection probability p)**.** *In real data sets, the detection probability p from (2.1), which we assume to be constant in our model, might in fact vary over time, for example due to bleaching of some of the markers towards the end of the measurement process. Because of this bleaching, larger structures appear fragmented into smaller pieces even in the binned images. For this reason, the last part of the observations sometimes is discarded and the estimation is run only on the remaining subset of frames. We refrain from this adjustment, however, since the extrapolation of the parameters which were obtained in this way is problematic.*

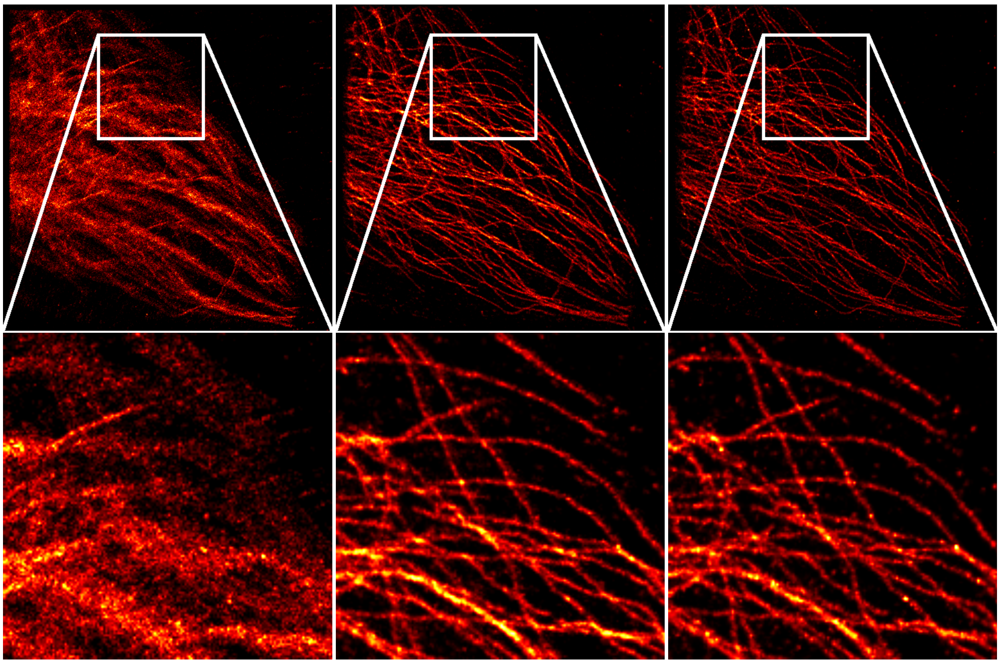

Figure 4.4: Image reconstructions of the SMS data analysis: overlay of the single frames with clearly visible motion blur (left); reconstruction obtained with our correction method (middle); results of a simple fiducial tracking procedure (right). The second row shows zoom-ins of the above images.

4.3 Bootstrap confidence bands

We would like to have some kind of assessment of how accurate our estimation of the parameters in the real data example is. Our goal in this chapter is therefore to construct confidence bands for the estimated motion functions that give regions around the estimator such that if the true parameter lies outside these regions, a data set resulting in the estimated parameter would only arise with a predefined low probability under the assumed model. Ideally, these bands are rather narrow around our estimate. From Section 3.2 we know that our M-estimators are asymptotically normally distributed. We have, however, no results on the actual distribution of the estimators. Nevertheless, the convergence towards a Gaussian distribution justifies the application of a bootstrap procedure in order to estimate confidence bands. Like in the previous paper (Hartmann et al., 2015) on pure drift estimation, we do this using the method described in Hall and Pittelkow (1990), which was already applied in a similar way in Hartmann (2016) based on the asymptotic normality of the parameters, which was only conjectured then. To illustrate the procedure, we briefly outline how it is applied to the estimator $\left(\delta_t^{\hat{\theta}_T}\right)_1$ of drift in x-direction. The confidence bands for the other motion types are computed analogously.
We determine the residuals

$$r_j^t := O_j^t - \hat{f}_T\left(1/\sigma_t^{\hat{\alpha}_T} R_{-\rho_t^{\hat{\phi}_T}}\left(x_j - \delta_t^{\hat{\theta}_T}\right)\right), \qquad t \in \mathbb{T}, 1 \leq j \leq n, \tag{4.1}$$

which are estimators for the Gaussian errors $\frac{1}{2\sqrt{\beta_T}}\epsilon_j^t = O_j^t - f\left(\frac{1}{\sigma_t^{\alpha_0}} R_{-\rho_t^{\phi_0}}\left(x_j - \delta_t^{\theta_0}\right)\right)$, and define the standardized difference

$$\Delta_t := \left(\left(\delta_t^{\hat{\theta}_T} \right)_1 - \left(\delta_t^{\theta_0} \right)_1 \right) / \lambda_{\delta,1}, \tag{4.2}$$

where $\lambda_{\delta,1}$ is the unknown standard deviation of the estimator of drift in x-direction. We decide upon a shape for the confidence band by choosing two template functions $g_+, g_- \colon [0,1] \to [0,\infty)$. The size of the band and thereby also the coverage rate is governed by two scaling factors $u_+, u_- \in [0,\infty)$. The lower and upper boundaries of the confidence band for drift in x-direction, $\left(\delta_t^{\theta_0} \right)_1$, are then given by

$$\left(\delta_t^{\hat{\theta}_T} \right)_1 - \lambda_{\delta,1} u_- g_-(t), \text{ and } \left(\delta_t^{\hat{\theta}_T} \right)_1 + \lambda_{\delta,1} u_+ g_+(t),$$

or equivalently by $-u_+ g_+(t)$ and $u_- g_-(t)$ for the standardized difference Δ_t. For a given confidence level $\eta \in (0,1)$ we want to minimize the width $u_+ + u_-$ of the confidence band under the constraint that

$$\mathbb{P}_{\theta_0} \left(\left(\delta_t^{\theta_0} \right)_1 \in \left[\left(\delta_t^{\hat{\theta}_T} \right)_1 - \lambda_{\delta,1} u_- g_-(t), \left(\delta_t^{\hat{\theta}_T} \right)_1 + \lambda_{\delta,1} u_+ g_+(t) \right] \text{ for all } t \in [0,1] \right) \geq 1 - \eta,$$

or equivalently

$$\mathbb{P}_{\theta_0} \left(\Delta_t \in \left[-u_+ g_+(t), u_- g_-(t) \right] \text{ for all } t \in [0,1] \right) \geq 1 - \eta.$$

Since the distribution of Δ_t is unknown, we apply a bootstrap procedure in order to approximate the quantiles. For every pixel location $1 \leq j \leq n$ and every time point $t \in \mathbb{T}$, we independently draw $B \in \mathbb{N}$ times with replacement from the set of all residuals $\{ r_{j'}^{t'} | 1 \leq j' \leq n, t' \in \mathbb{T} \}$ with $r_{j'}^{t'}$ from (4.1). These give rise to the resampled errors $\left(\epsilon_j^t \right)^{(b)}$ and observations

$$\left(O_j^t \right)^{(b)} = \hat{f}_T \left(1/\sigma_t^{\hat{\alpha}_T} R_{-\rho_t^{\hat{\phi}_T}} \left(x_j - \delta_t^{\hat{\theta}_T} \right) \right) + \left(\epsilon_j^t \right)^{(b)}, \qquad 1 \leq j \leq n,\ t \in \mathbb{T},\ 1 \leq b \leq B.$$

For every $1 \leq b \leq B$ we run our estimation method on the replicate observations given by $\left\{ \left(O_j^t \right)^{(b)} | 1 \leq j \leq n, t \in \mathbb{T} \right\}$ and thereby produce bootstrap replicates $\hat{\theta}_T^{(b)}$, $\hat{\phi}_T^{(b)}$ and $\hat{\alpha}_T^{(b)}$ of the parameter estimators and hence, replicates $\hat{f}_T^{(b)}$ of the image estimator as well. We can now compute replicates of the standardized difference,

$$\Delta_t^{(b)} = \left(\left(\delta_t^{\hat{\theta}_T^{(b)}} \right)_1 - \left(\delta_t^{\hat{\theta}_T} \right)_1 \right) / \hat{\lambda}_{\delta,1},$$

where

$$\hat{\lambda}_{\delta,1} = \sqrt{ \frac{1}{TB} \sum_{b=1}^{B} \sum_{t \in \mathbb{T}} \left(\left(\delta_t^{\hat{\theta}_T^{(b)}} \right)_1 - \left(\delta_t^{\hat{\theta}_T} \right)_1 \right)^2 }$$

is the empirical standard deviation. Using these replicates, we minimize the sum $u_+ + u_-$ such that

$$\#\left\{1 \leq b \leq B \middle| \Delta_t^{(b)} \in \left[-u_+ g_+(t), u_- g_-(t)\right] \text{ for all } t \in [0,1]\right\} \geq (1-\eta)B.$$

In our application we only use polynomial models for the motion function, as described in Example 2.17, focusing on linear models in the SMS data analysis. Furthermore, by Assumption 2.14 (B1) the motion functions have predefined values $\delta_0^\theta = 0, \rho_0^\phi = 0$ and $\sigma_0^\alpha = 1$ at time $t = 0$. Now, the fact that for polynomials on $[0,1]$ the linear part dominates the others ($t \geq t^p$ for all $t \in [0,1], p > 1$) justifies the use of linear template functions $g_+(t) = g_-(t) = t$.
In Figure 4.5, the confidence bands for the motion functions obtained using $B = 100$ bootstrap replicates and a confidence level of $\eta = 0.05$ are displayed. The insets in the plots of the estimated drift curves show a zoom-in on the confidence bands for the last minute of the imaging process. The motion paths estimated using fiducial marker tracking are given in red. The values obtained with both methods are of similar order of magnitudes. Generally, the estimation with fiducial marker tracking is much more variable. This is partly because in practice, the linear motion enforced by our parametric model can only be an approximation to the real movement. However, other effects leading to a more unstable fiducial estimation also play a role like movement of the fiducial markers relative to the sample structure or mistakes in the classification of observations into signal by the sample and signal by a fiducial marker. Still, for the estimation of rotation and scaling, the fiducial curves are mostly covered by the confidence bands around our estimators. For the drift estimation this is not the case anymore, mainly because of the very small width of the corresponding confidence bands. However, since the maximal translational offsets are almost identical for drift in x-direction and very close for drift in y-direction, the narrowness of the confidence bands mainly indicates that our estimation is able to estimate the best possible linear approximation with extremely high accuracy.
In order to illustrate the result on the reliability of our estimator, we use the bootstrap replicates $\hat{f}_T^{(b)}$ of the image estimator and compute their average. The result is displayed in Figure 4.6. As is to be expected, the bootstrap average (in the middle) is a little blurrier than the original estimate (on the left). However, due to the small widths of the confidence bands, they are remarkably close. If there was a high variability in the resulting estimator, the average image would be blurred much more or it would contain off-set features that originate from parameter outliers. This is not the case, which is a strong indicator that our reconstruction method works well and is able to reliably perform a motion correction. Moreover, in comparison with the reconstruction obtained by a simple fiducial marker tracking implementation (on the right), the bootstrap average is only slightly more blurred and still captures all the features and filament strands that can be identified by fiducial marker tracking.

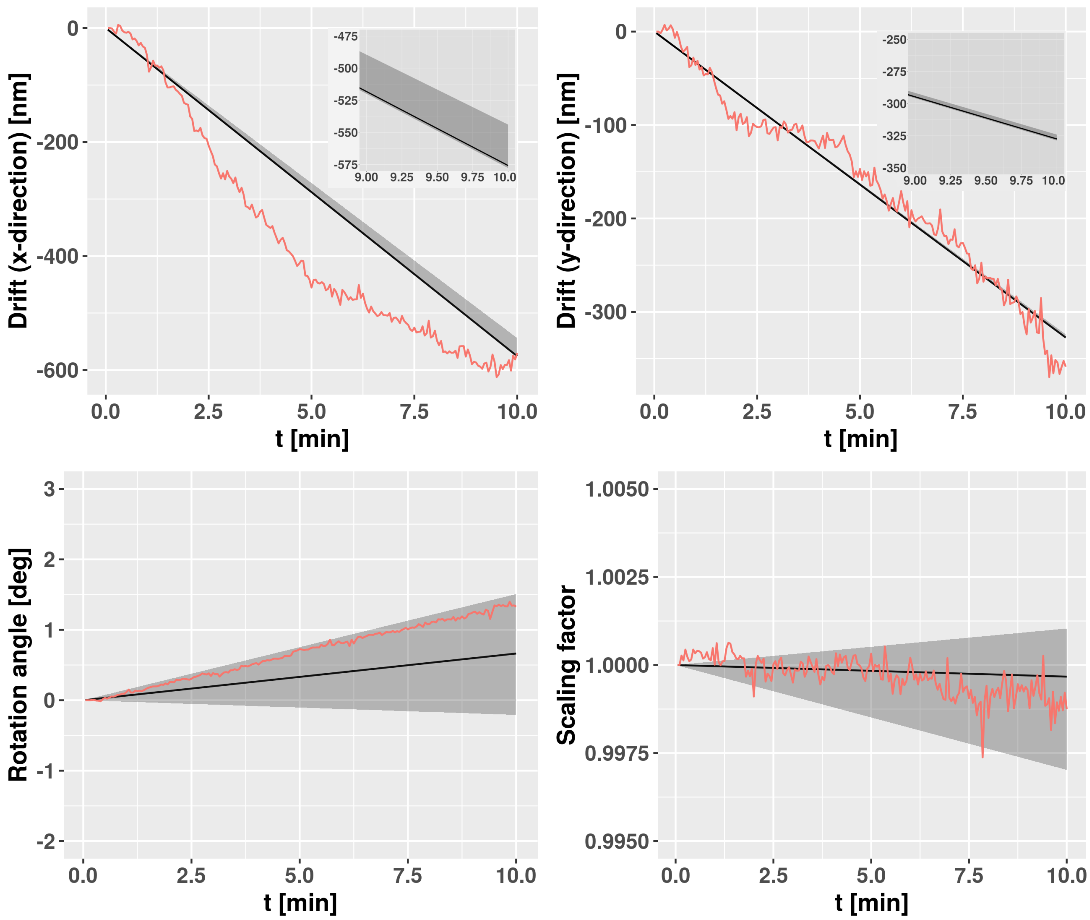

Figure 4.5: Bootstrap confidence bands (grey area) around the motion functions with the estimated parameters (black) and fiducial marker tracking paths (red); for the drift parameters (upper row), insets show zoom-ins on the last minute.

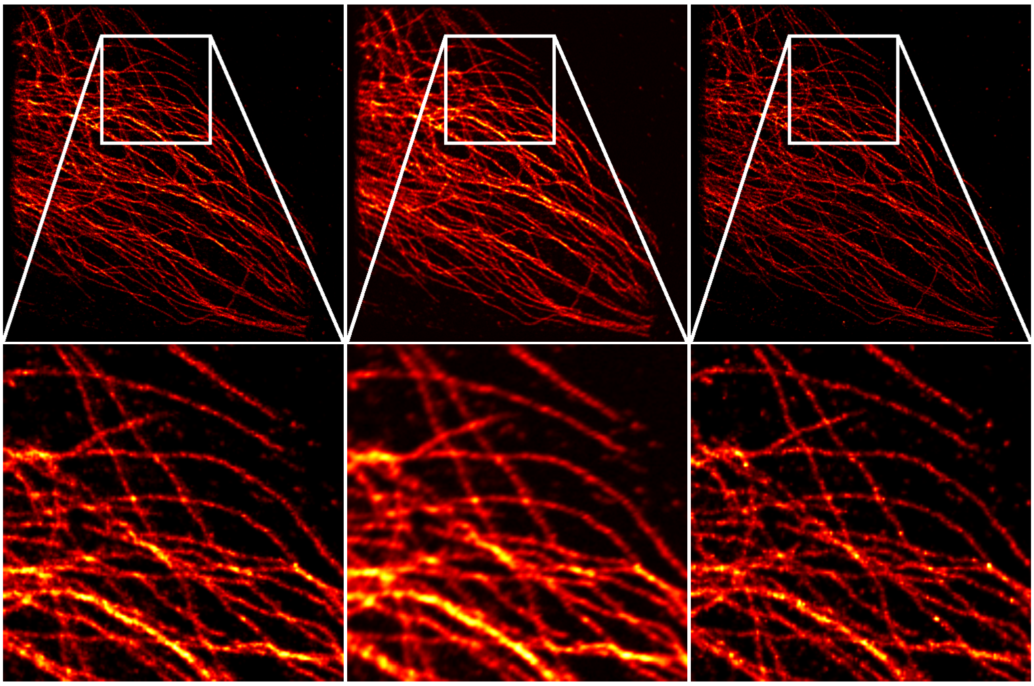

Figure 4.6: Results of the Bootstrap confidence analysis: image estimator from the previous section (left); average of the bootstrap replicates of the image estimate (middle); results of a simple fiducial tracking procedure (right). The second row shows zoom-ins of the above pictures.

CHAPTER 5

Summary and Outlook

In this thesis, the method of drift correction for SMS data or other application scenarios featuring sparse sequential dynamic imaging from Hartmann et al. (2015) was extended to incorporate also rotational and scaling movement as proposed in Hartmann (2016), using the two step estimation procedure described in Bigot et al. (2009).

The sparsity of the single frames in SMS microscopy data enables a significant improvement in resolution compared to conventional imaging techniques. At the same time, the induced need to record a large number of frames imposes a major blurring due to movement of the structure of interest. We presented a semiparametric estimation approach to correct for this motion, and examined the asymptotic properties of our estimators, which are defined as minimizers of certain contrast functionals. We were able to prove consistency for all motion function parameter estimators and for the final image estimator. Furthermore, we showed that the parameter estimators are asymptotically normally distributed, which enabled us to make confidence statements using a bootstrap procedure. We conducted a simulation study and performed a reconstruction of real SMS data. In the latter we used fiducial marker tracking as a reference. Our findings indicate that the current practice of fiducial marker tracking is not needed in most cases. SMS data can be corrected for motion by our purely statistical approach, which makes the incorporation of bright fiducial markers obsolete in future measurements. Moreover, our results are relatively stable with respect to the parameter choices, like, for example, the bin size β_T, the Fourier thresholds and the size of the pixel grid. More care should be taken in the choice of the starting values for the minimization.

It remains to investigate whether a varying detection probability can be included in the statistical model. As mentioned before, in real data sets the detection probability usually decreases due to bleaching of the fluorescent markers. This can only partly be made up for by adjusting the laser intensity used for reading out the signal. Secondly, for some application scenarios it might be of interest to implement further models for the motion functions apart from polynomials in the reconstruction software package. Moreover, an extension of the estimation method using non-parametric approaches could be studied, for example, an estimation of the motion functions using splines. Another approach would be to perform a reconstruction by directly estimating the marker density - without inference on any motion functions - using barycenters based on optimal transport distances. However, in that case it is not immediately clear, how the fact that the structure moves only very little between two consecutive frames can be used to make up for

the sparsity of the single frames. Finally, it would be desirable to loosen the assumption that the whole imaged structure undergoes the same motion deformation and allow also for local distortions of the observed specimen.
The presented method serves as a prototype for motion correction in SMS microscopy and many other typical imaging techniques where sparse observations with high temporal resolution are blurred by relative motion of the object to be reconstructed.

CHAPTER 6

Proofs

In this chapter we present detailed proofs for our theoretical results on the asymptotic properties of the estimators of the motion function parameters. We start in Section 6.1 by showing some properties of the error terms and the motion correction terms. Some statements are taken from Hartmann (2016) together with their proofs, whereas others are inspired by similar results in the same document or present generalized results, which are adapted to the setting of estimation of all three motion types, drift, rotation and scaling in this work. These preliminary results will be used in Section 6.2 for the demonstrations of consistency and joint asymptotic normality.

6.1 Properties of the correction functions and error terms

In the following, some basic properties of the error terms $W^t(\omega)$ from (2.11), $\mathcal{W}^t(\omega)$ from (2.16), and $V_T^t(\omega; \hat{\phi}_T, \hat{\alpha}_T)$ from (2.21) (Lemmas 6.1, 6.2, and 6.3), and of the derivatives of the motion correction terms $d_{u,v}(\sigma_t^\alpha, \rho_t^\phi)$ defined in (2.19), and $h_\omega^{t,t'}(\theta; \phi, \alpha)$ defined in (2.20) (Lemmas 6.4 and 6.7) are collated. Most of them can be obtained by basic calculations, which are similar to the corresponding proofs in Chapter 5 of Hartmann (2016). However, since we adjusted our model to better fit the data registration process, we arrive at slightly different values and rates. Therefore, we present the proofs, which are adapted from the ones found in Hartmann (2016).

Lemma 6.1 (Properties of $W^t(\omega)$, see also Lemma 5.1 in Hartmann (2016))**.** *Recall* $W^t(\omega) = \frac{1}{2n\sqrt{\beta_T}} \sum_{j=1}^n \mathrm{e}^{-2\pi \mathrm{i} \langle \omega, x_j \rangle} \epsilon_j^t$ *from (2.11). The following properties hold.*

1. *The real and imaginary parts of* $W^t(\omega)$ *are centered normal,*

$$\Re\left(W^t(\omega)\right) \sim \mathcal{N}\left(0, \frac{1}{4n^2\beta_T} \sum_{j=1}^n \left(\cos\left(2\pi \left\langle \omega, x_j \right\rangle\right)\right)^2\right),$$

$$\Im\left(W^t(\omega)\right) \sim \mathcal{N}\left(0, \frac{1}{4n^2\beta_T} \sum_{j=1}^n \left(\sin\left(2\pi \left\langle \omega, x_j \right\rangle\right)\right)^2\right).$$

2. *We have* $\mathbb{E}\left(\left|W^t(\omega)\right|^2\right) = \frac{1}{4n\beta_T}$.

3. $W^t(\omega)$ *and* $W^{t'}(\omega')$ *are independent unless* $t = t'$.

4. *For $j, j' \in \{1, \dots, n\}$, and $\omega \in \mathbb{R}^2$, define $\cos_{\omega}^{j,j'} := \cos\left(2\pi \left\langle \omega, x_j - x_{j'} \right\rangle\right)$. With this, we have the following expressions for the expectation of the mixed error terms:*

$$\begin{aligned}&\mathbb{E}(\Re\left(\mathcal{F}_{f^t}(\omega)\overline{W^t(\omega)}\right)\Re\left(\mathcal{F}_{f^t}(\omega')\overline{W^t(\omega')}\right))\\ =&\frac{1}{4n^4\beta_T}\sum_{j,j',j''=1}^{n}\cos_{\omega}^{j,j'}\cos_{\omega'}^{j'',j'} f^t(x_j)f^t(x_{j''}),\end{aligned}$$

$$\begin{aligned}\mathbb{E}(\Re\left(\mathcal{F}_{f^t}(\omega)\overline{W^t(\omega)}\right)\left|W^t(\omega')\right|^2) &= 0,\\ \textit{and } \mathbb{E}(\left|W^t(\omega)\right|^2\left|W^t(\omega')\right|^2) &= \frac{1}{16n^4\beta_T^2}\left[3n + \sum_{j\neq j'}\left(1 + 2\cos_{\omega}^{j,j'}\cos_{\omega'}^{j,j'}\right)\right].\end{aligned}$$

Proof.

1. Since they are linear combinations of independent centered Gaussian random variables $\left\{\epsilon_j^t \mid j \in J_T^t\right\}$, $\Re\left(W^t(\omega)\right)$ and $\Im\left(W^t(\omega)\right)$ are also centered Gaussian. Because the ϵ_j^t are standard normal and independent, we get

$$\begin{aligned}\operatorname{Var}\left(\Re\left(W^t(\omega)\right)\right) &= \frac{1}{4n^2\beta_T}\sum_{j=1}^{n}\operatorname{Var}\Big(\cos\left(2\pi\left\langle\omega, x_j\right\rangle\right)\epsilon_j^t\Big)\\ &= \frac{1}{4n^2\beta_T}\sum_{j=1}^{n}\Big(\cos\left(2\pi\left\langle\omega, x_j\right\rangle\right)\Big)^2,\end{aligned}$$

and, similarly,

$$\begin{aligned}\operatorname{Var}\left(\Im\left(W^t(\omega)\right)\right) &= \frac{1}{4n^2\beta_T}\sum_{j=1}^{n}\operatorname{Var}\Big(-\sin\left(2\pi\left\langle\omega, x_j\right\rangle\right)\epsilon_j^t\Big)\\ &= \frac{1}{4n^2\beta_T}\sum_{j=1}^{n}\Big(\sin\left(2\pi\left\langle\omega, x_j\right\rangle\right)\Big)^2,\end{aligned}$$

2. This follows from

$$\begin{aligned}\mathbb{E}\Big(\left|W^t(\omega)\right|^2\Big) &= \mathbb{E}\Big(\Re\left(W^t(\omega)\right)^2\Big) + \mathbb{E}\Big(\Im\left(W^t(\omega)\right)^2\Big)\\ &= \operatorname{Var}\Big(\Re\left(W^t(\omega)\right)\Big) + \operatorname{Var}\Big(\Im\left(W^t(\omega)\right)\Big)\\ &= \frac{1}{4n^2\beta_T}\sum_{j=1}^{n}\Big(\left(\cos\left(2\pi\left\langle\omega, x_j\right\rangle\right)\right)^2 + \left(\sin\left(2\pi\left\langle\omega, x_j\right\rangle\right)\right)^2\Big)\\ &= \frac{1}{4n^2\beta_T}\sum_{j=1}^{n}1 = \frac{1}{4n\beta_T}.\end{aligned}$$

3. Since $\left\{\epsilon_j^t \,\middle|\, j \in \{1\dots,n\}\right\}$ and $\left\{\epsilon_{j'}^{t'} \,\middle|\, j' \in \{1\dots,n\}\right\}$ are independent for $t \neq t'$, so are $W^t(\omega)$ and $W^{t'}(\omega')$.

4. First, note that

$$\begin{aligned}\Re\left(\mathcal{F}_{f^t}(\omega)\overline{W^t(\omega)}\right) &= \Re\left(n^{-2}\beta_T^{-1/2}\sum_{j,j'=1}^{n}\mathrm{e}^{-2\pi\mathrm{i}\langle\omega,x_j\rangle}f^t(x_j)\mathrm{e}^{2\pi\mathrm{i}\langle\omega,x_{j'}\rangle}\epsilon_{j'}^t\right)\\ &= \frac{1}{2n^2\sqrt{\beta_T}}\sum_{j,j'=1}^{n}\cos_\omega^{j,j'}f^t(x_j)\epsilon_{j'}^t.\end{aligned}$$

Since $\mathbb{E}(\epsilon_{j'}^t\epsilon_{j'''}^t) = 0$ unless $j' = j'''$ and $\mathbb{E}\left((\epsilon_{j'}^t)^2\right) = 1$, it follows that

$$\mathbb{E}\Big[\Re\left(\mathcal{F}_{f^t}(\omega)\overline{W^t(\omega)}\right)\Re\left(\mathcal{F}_{f^t}(\omega')\overline{W^t(\omega')}\right)\Big] = \frac{1}{4n^4\beta_T}\sum_{j,j',j''=1}^{n}\cos_\omega^{j,j'}\cos_{\omega'}^{j'',j'}f^t(x_j)f^t(x_{j''}).$$

Furthermore, using the trigonometrical addition theorems,

$$\begin{aligned}\left|W^t(\omega)\right|^2 &= \left(\frac{1}{2n\sqrt{\beta_T}}\sum_{j=1}^{n}\cos\left(2\pi\left\langle\omega,x_j\right\rangle\right)\epsilon_j^t\right)^2 + \left(\frac{1}{2n\sqrt{\beta_T}}\sum_{j=1}^{n}\sin\left(2\pi\left\langle\omega,x_j\right\rangle\right)\epsilon_j^t\right)^2\\ &= \frac{1}{4n^2\beta_T}\left(\sum_{j,j'=1}^{n}\cos\left(2\pi\left\langle\omega,x_j\right\rangle\right)\cos\left(2\pi\left\langle\omega,x_{j'}\right\rangle\right)\epsilon_j^t\epsilon_{j'}^t\right.\\ &\qquad\left.+\sum_{j,j'=1}^{n}\sin\left(2\pi\left\langle\omega,x_j\right\rangle\right)\sin\left(2\pi\left\langle\omega,x_{j'}\right\rangle\right)\epsilon_j^t\epsilon_{j'}^t\right)\\ &= \frac{1}{4n^2\beta_T}\sum_{j,j'=1}^{n}\cos_\omega^{j,j'}\epsilon_j^t\epsilon_{j'}^t.\end{aligned}$$

Because $\mathbb{E}(\epsilon_{j'}^t\epsilon_{j''}^t\epsilon_{j'''}^t) = 0$ even if $j' = j'' = j'''$, we get

$$\begin{aligned}&\mathbb{E}\Big[\Re\left(\mathcal{F}_{f^t}(\omega)\overline{W^t(\omega)}\right)\left|W^t(\omega')\right|^2\Big]\\ &= \frac{1}{8n^4\beta_T^{3/2}}\sum_{j,j',j'',j'''=1}^{n}\cos_\omega^{j,j'}\cos_{\omega'}^{j'',j'''}f^t(x_j)\mathbb{E}\left[\epsilon_{j'}^t\epsilon_{j''}^t\epsilon_{j'''}^t\right] = 0.\end{aligned}$$

Finally, because $\mathbb{E}\left((\epsilon_j^t)^4\right) = 3$,

$$\begin{aligned}\mathbb{E}\left(\left|W^t(\omega)\right|^2\left|W^t(\omega')\right|^2\right) &= \frac{1}{16n^4\beta_T^2}\sum_{j,j',j'',j'''=1}^{n}\cos_\omega^{j,j'}\cos_{\omega'}^{j'',j'''}\mathbb{E}\left(\epsilon_j^t\epsilon_{j'}^t\epsilon_{j''}^t\epsilon_{j'''}^t\right)\\ &= \frac{1}{16n^4\beta_T^2}\left(\sum_{j=j'=j''=j'''}+\sum_{j=j'\neq j''=j'''}+\sum_{j=j''\neq j'=j'''}+\sum_{j=j'''\neq j'=j''}\right)\\ &\qquad\left(\cos_\omega^{j,j'}\cos_{\omega'}^{j'',j'''}\mathbb{E}\left(\epsilon_j^t\epsilon_{j'}^t\epsilon_{j''}^t\epsilon_{j'''}^t\right)\right)\\ &= \frac{1}{16n^4\beta_T^2}\left(3n+\sum_{j\neq j'}\left(1+2\cos_\omega^{j,j'}\cos_{\omega'}^{j,j'}\right)\right).\end{aligned}$$

□

Lemma 6.2 (Properties of $\mathcal{M}^T_{\mathcal{W}^t}(u,v)$, see also Lemma 5.2 in Hartmann (2016))**.** *Recall* $\mathcal{W}^t = \left|W^t(\omega)\right|^2 + 2\Re\left(\mathcal{F}_{f^t}(\omega)\overline{W^t(\omega)}\right)$ *from (2.16). We have*

$$\mathbb{E}\left[\left|\mathcal{M}^T_{\mathcal{W}^t}(u,v)\right|^2\right] = O\left(\frac{r_T^{2\gamma}}{\beta_T}\right). \tag{6.1}$$

Proof. With Lemma 6.1, we get

$$\begin{aligned}
&\mathbb{E}\left[\left|\mathcal{M}^T_{\mathcal{W}^t}(u,v)\right|^2\right]\\
=&\mathbb{E}\left[\left|\int_0^{r_T}\int_0^{2\pi} \mathrm{e}^{-2\pi\mathrm{i}u\psi} r^{\gamma-\mathrm{i}v}\left(\mathcal{W}^t\circ\mathcal{P}\right)(r,\psi)\,\mathrm{d}\psi\,\frac{\mathrm{d}r}{r}\right|^2\right]\\
=&\int_0^{r_T}\int_0^{r_T}\int_0^{2\pi}\int_0^{2\pi} \mathrm{e}^{-2\pi\mathrm{i}u(\psi-\psi')} r^{\gamma-\mathrm{i}v}(r')^{\gamma+\mathrm{i}v}\\
&\cdot\mathbb{E}\left[\left(\left|W^t\left(\mathcal{P}(r,\psi)\right)\right|^2 + 2\Re\left(\left(\mathcal{F}_{f^t}\overline{W^t}\right)\left(\mathcal{P}(r,\psi)\right)\right)\right)\right.\\
&\cdot\left.\left(\left|W^t\left(\mathcal{P}(r',\psi')\right)\right|^2 + 2\Re\left(\left(\mathcal{F}_{f^t}\overline{W^t}\right)\left(\mathcal{P}(r',\psi')\right)\right)\right)\right]\mathrm{d}\psi\,\mathrm{d}\psi'\,\frac{\mathrm{d}r}{r}\,\frac{\mathrm{d}r'}{r'}\\
=&\int_0^{r_T}\int_0^{r_T}\int_0^{2\pi}\int_0^{2\pi} \mathrm{e}^{-2\pi\mathrm{i}u(\psi-\psi')} r^{\gamma-\mathrm{i}v}(r')^{\gamma+\mathrm{i}v}\left\{\mathbb{E}\left[\left|W^t\left(\mathcal{P}(r,\psi)\right)\right|^2\left|W^t\left(\mathcal{P}(r',\psi')\right)\right|^2\right]\right.\\
&+4\mathbb{E}\left[\Re\left(\left(\mathcal{F}_{f^t}\overline{W^t}\right)\left(\mathcal{P}(r,\psi)\right)\right)\Re\left(\left(\mathcal{F}_{f^t}\overline{W^t}\right)\left(\mathcal{P}(r',\psi')\right)\right)\right]\left.\right\}\mathrm{d}\psi\,\mathrm{d}\psi'\,\frac{\mathrm{d}r}{r}\,\frac{\mathrm{d}r'}{r'}\\
=&\int_0^{r_T}\int_0^{r_T}\int_0^{2\pi}\int_0^{2\pi} \mathrm{e}^{-2\pi\mathrm{i}u(\psi-\psi')} r^{\gamma-\mathrm{i}v}(r')^{\gamma+\mathrm{i}v}\left\{\frac{1}{16n^4\beta_T^2}\left[3n\right.\right.\\
&+\sum_{j\neq j'}\left(1+2\cos\left(2\pi\left\langle\mathcal{P}(r,\psi),x_j-x_{j'}\right\rangle\right)\cos\left(2\pi\left\langle\mathcal{P}(r',\psi'),x_j-x_{j'}\right\rangle\right)\right)\left.\right]\\
&+\frac{1}{n^4\beta_T}\sum_{j,j',j''=1}^{n}\left[\cos\left(2\pi\left\langle\mathcal{P}(r,\psi),x_j^t-x_{j'}\right\rangle\right)\cos\left(2\pi\left\langle\mathcal{P}(r',\psi'),x_{j''}^t-x_{j'}\right\rangle\right)\right.\\
&\cdot\left.\left.f^t(x_j^t)f^t(x_{j''}^t)\right]\right\}\mathrm{d}\psi\,\mathrm{d}\psi'\,\frac{\mathrm{d}r}{r}\,\frac{\mathrm{d}r'}{r'}.
\end{aligned}$$

In particular, because of the triangle inequality and $|\cos(x)| \leq 1$ for all $x \in \mathbb{R}$,

$$\begin{aligned}
&\mathbb{E}\left[\left|\mathcal{M}^T_{\mathcal{W}^t}(u,v)\right|^2\right]\\
\leq&\int_0^{r_T}\int_0^{r_T}\int_0^{2\pi}\int_0^{2\pi} r^\gamma(r')^\gamma\left(\frac{3}{16n^2\beta_T^2}+\frac{1}{n\beta_T}\|f\|_\infty^2\right)\mathrm{d}\psi\,\mathrm{d}\psi'\,\frac{\mathrm{d}r}{r}\,\frac{\mathrm{d}r'}{r'}\\
=&4\pi^2\left(\frac{3}{16n^2\beta_T^2}+\frac{1}{n\beta_T}\|f\|_\infty^2\right)\left(\int_0^{r_T} r^{\gamma-1}\,\mathrm{d}r\right)^2\\
=&4\pi^2 r_T^{2\gamma}\gamma^{-2}\left(\frac{3}{16n^2\beta_T^2}+\frac{1}{n\beta_T}\|f\|_\infty^2\right) = O\left(\frac{r_T^{2\gamma}}{\beta_T}\right),
\end{aligned}$$

where we used that f is bounded by Assumption 2.13 (A1). □

Lemma 6.3 (Properties of $V_T^t(\omega;\hat\phi_T,\hat\alpha_T)$, see also Lemma 5.3 in Hartmann (2016))**.** *Recall* $V_T^t(\omega;\hat\phi_T,\hat\alpha_T) = \frac{1}{2n\sqrt{\beta_T(\sigma_t^\alpha)^2}}\sum_{j=1}^n \exp\left(-2\pi\mathrm{i}\left\langle 1/\sigma_t^\alpha \cdot R_{\rho_t^\phi}\omega, x_j\right\rangle\right)\epsilon_j^t$ *from (2.21). For all* $T\in\mathbb{N}$, $t,t'\in\mathbb{T}$, *and* $\omega\in\mathbb{R}^2$, *we have*

$$\mathbb{E}\left|V_T^t(\omega;\hat\phi_T,\hat\alpha_T)\overline{V_T^{t'}(\omega;\hat\phi_T,\hat\alpha_T)}\right| = O\left(\frac{1}{\beta_T}\right).$$

Proof. Since the ϵ_j^t are independent standard normal random variables, we get

$$\begin{aligned}
&\mathbb{E}\left|V_T^t(\omega;\hat\phi_T,\hat\alpha_T)\overline{V_T^{t'}(\omega;\hat\phi_T,\hat\alpha_T)}\right|\\
=\ &\mathbb{E}\Bigg|\frac{1}{4n^2\beta_T(\sigma_t^{\hat\alpha_T})^2(\sigma_{t'}^{\hat\alpha_T})^2}\sum_{j,j'=1}^n \exp\left(-2\pi\mathrm{i}\left\langle 1/\sigma_t^{\hat\alpha_T}\cdot R_{\rho_t^{\hat\phi_T}}\omega, x_j\right\rangle\right)\\
&\cdot\exp\left(2\pi\mathrm{i}\left\langle 1/\sigma_{t'}^{\hat\alpha_T}\cdot R_{\rho_{t'}^{\hat\phi_T}}\omega, x_{j'}\right\rangle\right)\epsilon_j^t\epsilon_{j'}^{t'}\Bigg|\\
\le\ &\frac{1}{4n^2\beta_T\sigma_{\min}^4}\sum_{j,j'=1}^n\mathbb{E}\left|\epsilon_j^t\epsilon_{j'}^{t'}\right| = \frac{1}{4\beta_T\sigma_{\min}^4} = O\left(\frac{1}{\beta_T}\right).
\end{aligned}$$

□

The following three lemmas on the motion correction terms and their derivatives are generalized and reviewed versions of similar statements in Hartmann (2016), which are modified to the setting of three motion types and proved rigorously here. Note that in contrast to Hartmann (2016), a dependency of the drift correction error terms on the rotation and scaling parameters is included. This is necessary in order to be able to calculate the mixed derivatives of the drift contrast functional in the proof of asymptotic normality of the drift parameter.

Lemma 6.4 (Derivatives of $d_{u,v}(\sigma_t^\alpha,\rho_t^\phi)$, see also Lemma 5.7 in Hartmann (2016))**.** *Under Assumption 2.14 (B2), for* $u\in\mathbb{Z}$, $v\in\mathbb{R}$, $t,t'\in[0,1]$, *and* $(\phi,\alpha)\in U$, *we define*

$$d_{u,v}^{t,t'}(\phi,\alpha) := d_{u,v}(\sigma_t^\alpha/(\sigma_{t'}^\alpha\sigma_t^{\alpha_0}),\rho_t^\phi-\rho_{t'}^\phi-\rho_t^{\phi_0})\in\mathbb{C}, \tag{6.2}$$

$$\boldsymbol{a}_{u,v}^{t,t'}(\phi,\alpha) := \left(2\pi u\,\mathrm{grad}_\phi^\top(\rho_t^\phi-\rho_{t'}^\phi), -v\sigma_{t'}^\alpha\sigma_t^{\alpha_0}/\sigma_t^\alpha\,\mathrm{grad}_\alpha^\top(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right)^\top\in\mathbb{R}^{d_2+d_3}, \tag{6.3}$$

$$\boldsymbol{H}_{u,v}^{t,t'}(\phi,\alpha) := \begin{pmatrix}\boldsymbol{H}_1 & 0\\ 0 & \boldsymbol{H}_2\end{pmatrix}\in\mathbb{R}^{(d_2+d_3)\times(d_2+d_3)}, \tag{6.4}$$

where

$$\begin{aligned}
\boldsymbol{H}_1 &:= 2\pi u\,\mathrm{Hess}_\phi(\rho_t^\phi-\rho_{t'}^\phi),\\
\textit{and }\boldsymbol{H}_2 &:= v\frac{\sigma_{t'}^\alpha}{\sigma_t^\alpha}\left(\frac{\sigma_{t'}^\alpha}{\sigma_t^\alpha}\mathrm{grad}_\alpha\left(\frac{\sigma_t^\alpha}{\sigma_{t'}^\alpha}\right)\mathrm{grad}_\alpha^\top\left(\frac{\sigma_t^\alpha}{\sigma_{t'}^\alpha}\right)-\mathrm{Hess}_\alpha\left(\frac{\sigma_t^\alpha}{\sigma_{t'}^\alpha}\right)\right).
\end{aligned}$$

Note, that $d_{u,v}^{t,0}(\phi_0,\alpha_0) = d_{u,v}(1,0) = 1$. *There is a constant* $\widetilde{C}>0$ *(independent from u, v, t, t′, φ,* *and α) such that*

$$
\begin{aligned}
\operatorname{grad}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha) &= \mathrm{i}\boldsymbol{a}_{u,v}^{t,t'}(\phi,\alpha) d_{u,v}^{t,t'}(\phi,\alpha), && (6.5)\\
\operatorname{Hess}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha) &= \left(\mathrm{i}\boldsymbol{H}_{u,v}^{t,t'}(\phi,\alpha) - \boldsymbol{a}_{u,v}^{t,t'}(\phi,\alpha)\boldsymbol{a}_{u,v}^{t,t'}(\phi,\alpha)^\top\right) d_{u,v}^{t,t'}(\phi,\alpha), && (6.6)\\
\left\|\boldsymbol{a}_{u,v}^{t,t'}(\phi,\alpha)\right\| &\le \widetilde{C}\,\|(u,v)\|, && (6.7)\\
\left\|\operatorname{grad}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha)\right\| &\le \widetilde{C}\,\|(u,v)\|, && (6.8)\\
\left\|\operatorname{Hess}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha)\right\|_1 &\le \widetilde{C}\,\|(u,v)\| + \widetilde{C}^2\,\|(u,v)\|^2. && (6.9)
\end{aligned}
$$

Proof. First of all,

$$
\begin{aligned}
& \operatorname{grad}_{(\phi,\alpha)} d_{u,v}(\sigma_t^\alpha/(\sigma_{t'}^\alpha\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0})\\
={}& \operatorname{grad}_{(\phi,\alpha)} \exp\big(2\pi\mathrm{i}u(\rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0}) - \mathrm{i}v\log(\sigma_t^\alpha/(\sigma_{t'}^\alpha\sigma_t^{\alpha_0}))\big)\\
={}& d_{u,v}(\sigma_t^\alpha/(\sigma_{t'}^\alpha\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0})\\
& \cdot\mathrm{i}\left(2\pi u\operatorname{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v\sigma_{t'}^\alpha/\sigma_t^\alpha\operatorname{grad}_\alpha^\top(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right)^\top,
\end{aligned}
$$

which proves (6.5). It follows that

$$
\begin{aligned}
& \operatorname{Hess}_{(\phi,\alpha)} d_{u,v}(\sigma_t^\alpha/(\sigma_{t'}^\alpha\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0})\\
={}& \operatorname{grad}_{(\phi,\alpha)}^\top \Big[d_{u,v}(\sigma_t^\alpha/(\sigma_{t'}^\alpha\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0})\\
& \cdot\mathrm{i}\left(2\pi u\operatorname{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v\sigma_{t'}^\alpha/\sigma_t^\alpha\operatorname{grad}_\alpha^\top(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right)^\top\Big]\\
={}& d_{u,v}(\sigma_t^\alpha/(\sigma_{t'}^\alpha\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_{t'}^\phi - \rho_t^{\phi_0})\\
& \cdot\Big[-\left(2\pi u\operatorname{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v\sigma_{t'}^\alpha/\sigma_t^\alpha\operatorname{grad}_\alpha^\top(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right)^\top\\
& \cdot\left(2\pi u\operatorname{grad}_\phi^\top(\rho_t^\phi - \rho_{t'}^\phi), -v\sigma_{t'}^\alpha/\sigma_t^\alpha\operatorname{grad}_\alpha^\top(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right)\\
& +\mathrm{i}\begin{pmatrix} 2\pi u\operatorname{Hess}_\phi(\rho_t^\phi - \rho_{t'}^\phi) & 0\\ 0 & v\frac{\sigma_{t'}^\alpha}{\sigma_t^\alpha}\left(\frac{\sigma_{t'}^\alpha}{\sigma_t^\alpha}\operatorname{grad}_\alpha\left(\frac{\sigma_t^\alpha}{\sigma_{t'}^\alpha}\right)\operatorname{grad}_\alpha^\top\left(\frac{\sigma_t^\alpha}{\sigma_{t'}^\alpha}\right) - \operatorname{Hess}_\alpha\left(\frac{\sigma_t^\alpha}{\sigma_{t'}^\alpha}\right)\right)\end{pmatrix}\Big],
\end{aligned}
$$

proving (6.6). Now, let $\widetilde{C}_1 := \max\{4\pi C, 2C\sigma_{\max}/\sigma_{\min}^2\}$ with $C > 0$ from Assumption 2.14 (B2). Then,

$$
\begin{aligned}
\left\|\boldsymbol{a}_{u,v}^{t,t'}(\phi,\alpha)\right\|^2 &\le 4\pi^2u^2\left\|\operatorname{grad}_\phi(\rho_t^\phi - \rho_{t'}^\phi)\right\|^2 + v^2\left(\frac{\sigma_{t'}^\alpha}{\sigma_t^\alpha}\right)^2\left\|\operatorname{grad}_\alpha(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right\|^2\\
&\le 4\pi^2u^2\left(\left\|\operatorname{grad}_\phi\rho_t^\phi\right\| + \left\|\operatorname{grad}_\phi\rho_{t'}^\phi\right\|\right)^2\\
&\quad + v^2\left(\frac{1}{\sigma_t^\alpha\sigma_{t'}^\alpha}\right)^2\left(\sigma_{t'}^\alpha\left\|\operatorname{grad}_\alpha\sigma_t^\alpha\right\| + \sigma_t^\alpha\left\|\operatorname{grad}_\alpha\sigma_{t'}^\alpha\right\|\right)^2\\
&\le 16\pi^2C^2u^2 + 4\frac{\sigma_{\max}^2}{\sigma_{\min}^4}C^2v^2 \le \widetilde{C}_1^2\,\|(u,v)\|^2,
\end{aligned}
$$

which implies (6.7). Hence, (6.8) holds because, by (6.5),

$$\left\|\operatorname{grad}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha)\right\| = \left\|\boldsymbol{a}_{u,v}^{t,t'}(\phi,\alpha)\right\|.$$

Furthermore, by Assumption 2.14 (B2), and because of

$$\left\|xy^\top\right\|_1 = \|x\|_1 \|y\|_1 \le d_3 \|x\| \|y\| \quad \text{for all } x, y \in \mathbb{R}^{d_3},$$

we have that

$$\begin{aligned}
& \left\| \operatorname{Hess}_\alpha(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right\|_1 \\
= \; & \left\| \left(\sigma_{t'}^\alpha \operatorname{Hess}_\alpha \sigma_t^\alpha + \operatorname{grad}_\alpha \sigma_t^\alpha \operatorname{grad}_\alpha^\top \sigma_{t'}^\alpha - \sigma_t^\alpha \operatorname{Hess}_\alpha \sigma_{t'}^\alpha - \operatorname{grad}_\alpha \sigma_{t'}^\alpha \operatorname{grad}_\alpha^\top \sigma_t^\alpha\right)(\sigma_{t'}^\alpha)^{-2} \right. \\
& \left. -2(\sigma_{t'}^\alpha)^{-3}\left(\sigma_{t'}^\alpha \operatorname{grad}_\alpha \sigma_t^\alpha - \sigma_t^\alpha \operatorname{grad}_\alpha \sigma_{t'}^\alpha\right) \operatorname{grad}_\alpha^\top \sigma_{t'}^\alpha \right\|_1 \\
\le \; & \frac{2C^2(\sigma_{\max}+1)}{\sigma_{\min}^2} + \frac{4d_3C^2\sigma_{\max}}{\sigma_{\min}^3} \le \frac{8d_3C^2\sigma_{\max}}{\sigma_{\min}^3}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\left\|\boldsymbol{H}_{u,v}^{t,t'}(\phi,\alpha)\right\|_1 \; = \; & 2\pi|u| \left\| \operatorname{Hess}_\phi(\rho_t^\phi - \rho_{t'}^\phi)\right\|_1 \\
& + |v| \frac{\sigma_{t'}^\alpha}{\sigma_t^\alpha}\left(\frac{\sigma_{t'}^\alpha}{\sigma_t^\alpha}\left\|\operatorname{grad}_\alpha(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right\|_1^2 + \left\|\operatorname{Hess}_\alpha(\sigma_t^\alpha/\sigma_{t'}^\alpha)\right\|_1\right) \\
\le \; & 4\pi C^2|u| + |v|C^2\frac{\sigma_{\max}^2}{\sigma_{\min}^2} + 8d_3C^2\frac{\sigma_{\max}^2}{\sigma_{\min}^4}|v| \le \widetilde{C}_2\|(u,v)\|,
\end{aligned}$$

where $\widetilde{C}_2 := C^2 \max\{4\pi, \sigma_{\max}^2\sigma_{\min}^{-2}, 8d_3C^2\sigma_{\max}^2\sigma_{\min}^{-4}\}$, which implies

$$\left\| \operatorname{Hess}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha)\right\|_1 \le \left\|\boldsymbol{H}_{u,v}^{t,t'}(\phi,\alpha)\right\|_1 + d_3\left\|\boldsymbol{a}_{u,v}^{t,t'}(\phi,\alpha)\right\|^2 \le \widetilde{C}_2\|(u,v)\| + d_3\widetilde{C}_1^2\|(u,v)\|^2.$$

Then, (6.9) holds with $\widetilde{C} := \max\left\{\sqrt{d_3}\widetilde{C}_1, \widetilde{C}_2\right\}$. □

Lemma 6.5 (Properties of $h_\omega^{t,t'}(\theta;\phi,\alpha)$, see also Lemma 5.8 in Hartmann (2016))**.** *Recall from (2.20) that*

$$h_\omega^{t,t'}(\theta;\phi,\alpha) = \exp\left(2\pi \mathrm{i}\left(\left\langle(\sigma_t^\alpha)^{-1}R_{\rho_t^\phi}\omega, \delta_t^\theta - \delta_t^{\theta_0}\right\rangle - \left\langle(\sigma_{t'}^\alpha)^{-1}R_{\rho_{t'}^\phi}\omega, \delta_{t'}^\theta\right\rangle\right)\right).$$

We have

$$\left|h_\omega^{t,t'}(\theta;\phi,\alpha)\right| \equiv 1, \tag{6.10}$$

$$\overline{h_\omega^{t,t'}(\theta;\phi,\alpha)} = h_\omega^{t,t'}(\theta;\phi,\alpha)^{-1}, \tag{6.11}$$

$$h_\omega^{t,0}(\theta_0;\phi_0,\alpha_0) = 1, \tag{6.12}$$

$$h_\omega^{t,0}(\theta;\phi,\alpha)h_\omega^{0,t'}(\theta;\phi,\alpha) = h_\omega^{t,t'}(\theta;\phi,\alpha). \tag{6.13}$$

Moreover, the following terms can be expressed with the help of $h_\omega^{t,t'}(\theta;\phi,\alpha)$*:*

$$\exp\left(2\pi\mathrm{i}\left\langle (\sigma_t^\alpha)^{-1}R_{\rho_t^\phi}\omega,\delta_t^\theta\right\rangle\right) = h_\omega^{0,t}(\theta;\phi,\alpha)^{-1}, \tag{6.14}$$

$$\exp\left(2\pi\mathrm{i}\left\langle (\sigma_t^\alpha)^{-1}R_{\rho_t^\phi}\omega,\delta_t^\theta-\delta_t^{\theta_0}\right\rangle\right)\overline{\exp\left(2\pi\mathrm{i}\left\langle (\sigma_{t'}^\alpha)^{-1}R_{\rho_{t'}^\phi}\omega,\delta_{t'}^\theta-\delta_{t'}^{0}\right\rangle\right)}$$
$$= h_\omega^{t,0}(\theta;\phi,\alpha)/h_\omega^{t',0}(\theta;\phi,\alpha). \tag{6.15}$$

Proof. The fact that $\left|\mathrm{e}^{\mathrm{i}x}\right| = 1$ and $\overline{\mathrm{e}^{\mathrm{i}x}} = \mathrm{e}^{-\mathrm{i}x}$ for all $x \in \mathbb{R}$ implies (6.10) and (6.11). The properties (6.12) and (6.13) follow because $\mathrm{e}^0 = 1$ and $\mathrm{e}^x\mathrm{e}^y = \mathrm{e}^{x+y}$ for all $x, y \in \mathbb{C}$. (6.14) and (6.15) hold by definition of $h_\omega^{t,t'}(\theta;\phi,\alpha)$. □

Before formulating the Lemma on the derivatives of $h_\omega^{t,t'}$, some definitions to shorten the notation are in order.

Definition 6.6. *Let* $\omega = (\omega_1,\omega_2) \in \mathbb{R}^2$*,* $t,t' \in [0,1]$ *and* $(\theta,\phi,\alpha) \in \mathbb{R}^{d_1+d_2+d_3}$*. Let further* $\boldsymbol{e}^{t,t'}(\theta;\phi,\alpha) := \left(\boldsymbol{e}_1^{t,t'}(\theta;\phi,\alpha), \boldsymbol{e}_2^{t,t'}(\theta;\phi,\alpha)\right)$*, where for* $j = 1,2$

$$\boldsymbol{e}_j^{t,t'}(\theta;\phi,\alpha) := \left(\left(\sigma_t^\alpha\right)^{-1}R_{-\rho_t^\phi}\left(\delta_t^\theta-\delta_t^{\theta_0}\right) - \left(\sigma_{t'}^\alpha\right)^{-1}R_{-\rho_{t'}^\phi}\delta_{t'}^\theta\right)_j.$$

Note that

$$h_\omega^{t,t'}(\theta;\phi,\alpha) = \exp\left(2\pi\mathrm{i}\left\langle\omega,\boldsymbol{e}^{t,t'}(\theta;\phi,\alpha)\right\rangle\right). \tag{6.16}$$

We further define

$$\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) := 2\pi\left(\omega_1\mathrm{grad}_\theta\left(\boldsymbol{e}_1^{t,t'}(\theta;\phi,\alpha)\right)+\omega_2\mathrm{grad}_\theta\left(\boldsymbol{e}_2^{t,t'}(\theta;\phi,\alpha)\right)\right). \tag{6.17}$$

Similarly,

$$\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha) := 2\pi\left(\left(\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right)_\phi^\top, \left(\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right)_\alpha^\top\right)^\top \in \mathbb{R}^{d_2+d_3}, \tag{6.18}$$

where

$$\left(\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right)_\phi := \omega_1\mathrm{grad}_\phi\left(\boldsymbol{e}_1^{t,t'}(\theta;\phi,\alpha)\right)+\omega_2\mathrm{grad}_\phi\left(\boldsymbol{e}_2^{t,t'}(\theta;\phi,\alpha)\right)$$
$$\left(\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right)_\alpha := \omega_1\mathrm{grad}_\alpha\left(\boldsymbol{e}_1^{t,t'}(\theta;\phi,\alpha)\right)+\omega_2\mathrm{grad}_\alpha\left(\boldsymbol{e}_2^{t,t'}(\theta;\phi,\alpha)\right).$$

In order to examine also the second derivatives we further define

$$\boldsymbol{H}_\omega^{t,t'}(\theta;\phi,\alpha) := 2\pi\left(\omega_1\,\mathrm{Hess}_\theta\left(\boldsymbol{e}_1^{t,t'}(\theta;\phi,\alpha)\right)+\omega_2\,\mathrm{Hess}_\theta\left(\boldsymbol{e}_2^{t,t'}(\theta;\phi,\alpha)\right)\right). \tag{6.19}$$

To tackle the mixed derivatives we will need

$$\boldsymbol{G}_\omega^{t,t'}(\theta;\phi,\alpha) := \left(\mathrm{grad}_\phi\left(\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right),\ \mathrm{grad}_\alpha\left(\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right)\right) \in \mathbb{R}^{d_1\times(d_2+d_3)}, \tag{6.20}$$

where we write $\mathrm{grad}_\phi\left(\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right)$ *for the Jacobian of* $\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)$*, and similarly for the second block, in a slight abuse of notation.*

Lemma 6.7 (Derivatives of $h_\omega^{t,t'}(\theta;\phi,\alpha)$, generalization of Lemma 5.9 in Hartmann (2016))**.** *Under Assumptions 2.14 (B2, B3) it holds that*

$$\mathrm{grad}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha) = \mathrm{i}\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)h_\omega^{t,t'}(\theta;\phi,\alpha), \tag{6.21}$$

$$\mathrm{grad}_{(\phi,\alpha)} h_\omega^{t,t'}(\theta;\phi,\alpha) = \mathrm{i}\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)h_\omega^{t,t'}(\theta;\phi,\alpha), \tag{6.22}$$

$$\mathrm{Hess}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha) = \left(\mathrm{i}\boldsymbol{H}_\omega^{t,t'}(\theta;\phi,\alpha) - \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)^\top\right)h_\omega^{t,t'}(\theta;\phi,\alpha), \tag{6.23}$$

$$\mathrm{grad}_{(\phi,\alpha)}\mathrm{grad}_\theta^\top h_\omega^{t,t'}(\theta;\phi,\alpha) = \left(\mathrm{i}\boldsymbol{G}_\omega^{t,t'}(\theta;\phi,\alpha) - \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)^\top\right) h_\omega^{t,t'}(\theta;\phi,\alpha), \tag{6.24}$$

where we write $\mathrm{grad}_{(\phi,\alpha)}\mathrm{grad}_\theta^\top h_\omega^{t,t'}(\theta;\phi,\alpha)$ *for the Jacobian of* $\mathrm{grad}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha)$ *in a slight abuse of notation. Moreover, there is a constant* $\widetilde{C} > 0$ *(independent of* ω*,* t*,* t'*, and the parameters* θ*,* ϕ *and* α*) such that*

$$\left\|\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|, \left\|\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right\| \le \widetilde{C}\,\|\omega\|, \tag{6.25}$$

$$\left\|\mathrm{grad}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha)\right\|, \left\|\mathrm{grad}_{(\phi,\alpha)} h_\omega^{t,t'}(\theta;\phi,\alpha)\right\| \le \widetilde{C}\,\|\omega\|, \tag{6.26}$$

$$\left\|\boldsymbol{H}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 \le \widetilde{C}\,\|\omega\|, \tag{6.27}$$

$$\left\|\mathrm{Hess}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 \le \widetilde{C}\,\|\omega\| + \widetilde{C}^2\,\|\omega\|^2 \tag{6.28}$$

$$\left\|\boldsymbol{G}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 \le \widetilde{C}\,\|\omega\| \tag{6.29}$$

$$\left\|\mathrm{grad}_{(\phi,\alpha)}\mathrm{grad}_\theta^\top h_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 \le \widetilde{C}\,\|\omega\| + \widetilde{C}^2\,\|\omega\|^2 \tag{6.30}$$

Proof. Using (6.16) we get for the gradient with respect to θ that

$$\begin{aligned}
\mathrm{grad}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha) &= \mathrm{grad}_\theta \exp\left(2\pi\mathrm{i}\left\langle\omega, \boldsymbol{e}^{t,t'}(\theta;\phi,\alpha)\right\rangle\right)\\
&= h_\omega^{t,t'}(\theta;\phi,\alpha)2\pi\mathrm{i}\,\mathrm{grad}_\theta\left\langle\omega, \boldsymbol{e}^{t,t'}(\theta;\phi,\alpha)\right\rangle\\
&= h_\omega^{t,t'}(\theta;\phi,\alpha)2\pi\mathrm{i}\left(\omega_1\mathrm{grad}_\theta\left(\boldsymbol{e}_1^{t,t'}(\theta;\phi,\alpha)\right) + \omega_2\mathrm{grad}_\theta\left(\boldsymbol{e}_2^{t,t'}(\theta;\phi,\alpha)\right)\right)\\
&= \mathrm{i}h_\omega^{t,t'}(\theta;\phi,\alpha)\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha),
\end{aligned}$$

proving (6.21). Similarly, for the gradient of $h_\omega^{t,t'}$ with respect to ϕ and α it holds that

$$\begin{aligned}
&\mathrm{grad}_{(\phi,\alpha)} h_\omega^{t,t'}(\theta;\phi,\alpha)\\
=&\mathrm{grad}_{(\phi,\alpha)}\exp\left(2\pi\mathrm{i}\left(\left\langle\omega, \boldsymbol{e}^{t,t'}(\theta;\phi,\alpha)\right\rangle\right)\right)\\
=&h_\omega^{t,t'}(\theta;\phi,\alpha)2\pi\mathrm{i}\,\mathrm{grad}_{(\phi,\alpha)}\left\langle\omega, \boldsymbol{e}^{t,t'}(\theta;\phi,\alpha)\right\rangle\\
=&h_\omega^{t,t'}(\theta;\phi,\alpha)2\pi\mathrm{i}\left(\omega_1\mathrm{grad}_{(\phi,\alpha)}\left(\boldsymbol{e}_1^{t,t'}(\theta;\phi,\alpha)\right) + \omega_2\mathrm{grad}_{(\phi,\alpha)}\left(\boldsymbol{e}_2^{t,t'}(\theta;\phi,\alpha)\right)\right)\\
=&\mathrm{i}h_\omega^{t,t'}(\theta;\phi,\alpha)\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha),
\end{aligned}$$

implying (6.22). Since $\operatorname{grad}_\theta \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) = \boldsymbol{H}_\omega^{t,t'}(\theta;\phi,\alpha)$, where we use again a somewhat sloppy notation for the Jacobian of $\boldsymbol{b}_\omega^{t,t'}$, we have

$$\begin{aligned}
&\operatorname{Hess}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha)\\
&= \operatorname{grad}_\theta \left(\mathrm{i} h_\omega^{t,t'}(\theta;\phi,\alpha) \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \right)\\
&= \mathrm{i}^2 h_\omega^{t,t'}(\theta;\phi,\alpha) \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)^\top + \mathrm{i} h_\omega^{t,t'}(\theta;\phi,\alpha) \boldsymbol{H}_\omega^{t,t'}(\theta;\phi,\alpha)\\
&= \left(\mathrm{i} \boldsymbol{H}_\omega^{t,t'}(\theta;\phi,\alpha) - \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)^\top \right) h_\omega^{t,t'}(\theta;\phi,\alpha),
\end{aligned}$$

which proves (6.23). Similarly, we have that

$$\operatorname{grad}_{(\phi,\alpha)} \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) = \left(\operatorname{grad}_\phi \left(\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \right),\ \operatorname{grad}_\alpha \left(\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \right) \right) = \boldsymbol{G}_\omega^{t,t'}(\theta;\phi,\alpha),$$

and therefore,

$$\begin{aligned}
&\operatorname{grad}_{(\phi,\alpha)} \operatorname{grad}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha)\\
&= \mathrm{i} \left(\operatorname{grad}_{(\phi,\alpha)} \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) h_\omega^{t,t'}(\theta;\phi,\alpha) + \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \operatorname{grad}_{(\phi,\alpha)}^\top h_\omega^{t,t'}(\theta;\phi,\alpha) \right)\\
&= \left(\mathrm{i} \boldsymbol{G}_\omega^{t,t'}(\theta;\phi,\alpha) - \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)^\top \right) h_\omega^{t,t'}(\theta;\phi,\alpha),
\end{aligned}$$

which shows (6.24).

The bounds on the derivatives can be obtained as follows. With C from Assumption 2.14 (B2) and using that $\|x\| \le \|x\|_1 \le \sqrt{d}\,\|x\|$ for any $x \in \mathbb{R}^d$, we have for $\tilde{C}_1 := 4C\sqrt{2d_1}\sigma_{\min}^{-1}$ that

$$\begin{aligned}
&\left\| \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \right\|\\
&= \Bigg\| \omega_1 \left[\frac{1}{\sigma_t^\alpha} \left(\cos\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_1 + \sin\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_2 \right) \right.\\
&\qquad \left. - \frac{1}{\sigma_{t'}^\alpha} \left(\cos\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1 + \sin\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2 \right) \right]\\
&\quad + \omega_2 \left[\frac{1}{\sigma_t^\alpha} \left(-\sin\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_1 + \cos\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_2 \right) \right.\\
&\qquad \left. - \frac{1}{\sigma_{t'}^\alpha} \left(-\sin\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1 + \cos\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2 \right) \right] \Bigg\|_1\\
&\le \frac{|\omega_1| + |\omega_2|}{\sigma_{\min}} \left(\left\| \operatorname{grad}_\theta \left(\delta_t^\theta\right)_1 \right\|_1 + \left\| \operatorname{grad}_\theta \left(\delta_t^\theta\right)_2 \right\|_1 + \left\| \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1 \right\|_1 + \left\| \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2 \right\|_1 \right)\\
&\le 4C\sqrt{2d_1}\sigma_{\min}^{-1} \|\omega\| = \tilde{C}_1 \|\omega\|.
\end{aligned}$$

Since, by Assumption 2.14 (B7), the drift function δ^θ has bounded variation over [0, 1] uniformly in θ, it is bounded as function in t, uniformly in θ. Hence, there is a constant C_2 such that

$$\left\| \delta_t^\theta \right\| \le C_2 \text{ uniformly in } \theta \text{ and } t. \tag{6.31}$$

We get with $C > 0$ from Assumption 2.14 (B2) and using again the equivalence of the norms that

$$
\begin{aligned}
&\left\|\left(\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right)_\phi\right\|_1 \\
&= \Big\| \left(\sigma_t^\alpha\right)^{-1}\Big[\omega_1\left(-\sin\left(\rho_t^\phi\right)\left(\delta_t^\theta-\delta_t^{\theta_0}\right)_1+\cos\left(\rho_t^\phi\right)\left(\delta_t^\theta-\delta_t^{\theta_0}\right)_2\right) \\
&\qquad +\omega_2\left(-\cos\left(\rho_t^\phi\right)\left(\delta_t^\theta-\delta_t^{\theta_0}\right)_1-\sin\left(\rho_t^\phi\right)\left(\delta_t^\theta-\delta_t^{\theta_0}\right)_2\right)\Big]\operatorname{grad}_\phi\rho_t^\phi \\
&\quad +\left(\sigma_{t'}^\alpha\right)^{-1}\Big[\omega_1\left(-\sin\left(\rho_{t'}^\phi\right)\left(\delta_{t'}^\theta\right)_1+\cos\left(\rho_{t'}^\phi\right)\left(\delta_{t'}^\theta\right)_2\right) \\
&\qquad +\omega_2\left(-\cos\left(\rho_{t'}^\phi\right)\left(\delta_{t'}^\theta\right)_1-\sin\left(\rho_{t'}^\phi\right)\left(\delta_{t'}^\theta\right)_2\right)\Big]\operatorname{grad}_\phi\rho_{t'}^\phi\Big\|_1 \\
&\le \sigma_{\min}^{-1}\left(|\omega_1|+|\omega_2|\right)\left\|\delta_t^\theta-\delta_t^{\theta_0}\right\|_1\left\|\operatorname{grad}_\phi\rho_t^\phi\right\|_1+\sigma_{\min}^{-1}\left(|\omega_1|+|\omega_2|\right)\left\|\delta_{t'}^\theta\right\|_1\left\|\operatorname{grad}_\phi\rho_{t'}^\phi\right\|_1 \\
&\le 6\sqrt{d_2}\,\sigma_{\min}^{-1}\|\omega\|\,C_2C,
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
&\left\|\left(\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right)_\alpha\right\|_1 \\
&= \Big\| \left(\sigma_t^\alpha\right)^{-2}\operatorname{grad}_\alpha\sigma_t^\alpha\Big[\omega_1\left(\cos\left(\rho_t^\phi\right)\left(\delta_t^\theta-\delta_t^{\theta_0}\right)_1+\sin\left(\rho_t^\phi\right)\left(\delta_t^\theta-\delta_t^{\theta_0}\right)_2\right) \\
&\qquad +\omega_2\left(-\sin\left(\rho_t^\phi\right)\left(\delta_t^\theta-\delta_t^{\theta_0}\right)_1+\cos\left(\rho_t^\phi\right)\left(\delta_t^\theta-\delta_t^{\theta_0}\right)_2\right)\Big] \\
&\quad +\left(\sigma_{t'}^\alpha\right)^{-2}\operatorname{grad}_\alpha\sigma_{t'}^\alpha\Big[\omega_1\left(\cos\left(\rho_{t'}^\phi\right)\left(\delta_{t'}^\theta\right)_1+\sin\left(\rho_{t'}^\phi\right)\left(\delta_{t'}^\theta\right)_2\right) \\
&\qquad +\omega_2\left(-\sin\left(\rho_{t'}^\phi\right)\left(\delta_{t'}^\theta\right)_1+\cos\left(\rho_{t'}^\phi\right)\left(\delta_{t'}^\theta\right)_2\right)\Big]\Big\|_1 \\
&\le 6\sqrt{d_3}\,\sigma_{\min}^{-2}\|\omega\|\,C_2C.
\end{aligned}
$$

Thus, for $\tilde{C}_2 \coloneqq 12\pi C C_2\left(\sqrt{d_2}\sigma_{\min}^{-1}+\sqrt{d_3}\sigma_{\min}^{-2}\right)$

$$
\begin{aligned}
\left\|\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right\| &\le 2\pi\left(\left\|\left(\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right)_\phi\right\|_1+\left\|\left(\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right)_\alpha\right\|_1\right) \\
&\le 12\pi C_2 C\left(\sqrt{d_2}\sigma_{\min}^{-1}+\sqrt{d_3}\sigma_{\min}^{-2}\right)\|\omega\| = \tilde{C}_2\|\omega\|,
\end{aligned}
$$

proving (6.25). Combining this with (6.21) and (6.10), it follows that

$$
\left\|\operatorname{grad}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha)\right\| = \left\|\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|\left|h_\omega^{t,t'}(\theta;\phi,\alpha)\right| \le \tilde{C}_1\|\omega\|,
$$

and, using (6.22) that

$$
\left\|\operatorname{grad}_{(\phi,\alpha)} h_\omega^{t,t'}(\theta;\phi,\alpha)\right\| = \left\|\boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|\left|h_\omega^{t,t'}(\theta;\phi,\alpha)\right| \le \tilde{C}_2\|\omega\|,
$$

showing (6.26).

Next, we derive bounds for the mixed derivatives. For the first block we get

$$
\begin{aligned}
&\left\|\operatorname{grad}_\phi \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 \\
&= 2\pi \Big\| \left(\sigma_t^\alpha\right)^{-1} \Big[\omega_1 \left(-\sin\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_1 + \cos\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_2\right) \\
&\qquad + \omega_2 \left(-\cos\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_1 - \sin\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_2\right)\Big] \operatorname{grad}_\phi^\top \rho_t^\phi \\
&\quad - \left(\sigma_{t'}^\alpha\right)^{-1} \Big[\omega_1 \left(-\sin\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1 + \cos\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2\right) \\
&\qquad + \omega_2 \left(-\cos\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1 - \sin\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2\right)\Big] \operatorname{grad}_\phi^\top \rho_{t'}^\phi \Big\|_1 \\
&\le 2\pi\sigma_{\min}^{-1} \left(|\omega_1| + |\omega_2|\right) \Big(\left(\left\|\operatorname{grad}_\theta \left(\delta_t^\theta\right)_1\right\|_1 + \left\|\operatorname{grad}_\theta \left(\delta_t^\theta\right)_2\right\|_1\right) \left\|\operatorname{grad}_\phi^\top \rho_t^\phi\right\|_1 \\
&\qquad + \left(\left\|\operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1\right\|_1 + \left\|\operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2\right\|_1\right) \left\|\operatorname{grad}_\phi^\top \rho_{t'}^\phi\right\|_1 \Big) \\
&\le 4\pi\sqrt{2d_1d_2}\,\sigma_{\min}^{-1} C^2 \|\omega\| .
\end{aligned}
$$

For the second block we obtain

$$
\begin{aligned}
&\left\|\operatorname{grad}_\alpha \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 \\
&= 2\pi \Big\| -\left(\sigma_t^\alpha\right)^{-2} \Big[\omega_1 \left(\cos\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_1 + \sin\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_2\right) \\
&\qquad + \omega_2 \left(-\sin\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_1 + \cos\left(\rho_t^\phi\right) \operatorname{grad}_\theta \left(\delta_t^\theta\right)_2\right)\Big] \operatorname{grad}_\alpha^\top \sigma_t^\alpha \\
&\quad + \left(\sigma_{t'}^\alpha\right)^{-2} \Big[\omega_1 \left(\cos\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1 + \sin\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2\right) \\
&\qquad + \omega_2 \left(-\sin\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1 + \cos\left(\rho_{t'}^\phi\right) \operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2\right)\Big] \operatorname{grad}_\alpha^\top \sigma_{t'}^\alpha \Big\|_1 \\
&\le 2\pi\sigma_{\min}^{-2} \left(|\omega_1| + |\omega_2|\right) \Big(\left(\left\|\operatorname{grad}_\theta \left(\delta_t^\theta\right)_1\right\|_1 + \left\|\operatorname{grad}_\theta \left(\delta_t^\theta\right)_2\right\|_1\right) \left\|\operatorname{grad}_\alpha^\top \sigma_t^\alpha\right\|_1 \\
&\qquad + \left(\left\|\operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_1\right\|_1 + \left\|\operatorname{grad}_\theta \left(\delta_{t'}^\theta\right)_2\right\|_1\right) \left\|\operatorname{grad}_\alpha^\top \sigma_{t'}^\alpha\right\|_1 \Big) \\
&\le 4\pi\sqrt{2d_1d_3}\,\sigma_{\min}^{-2} C^2 \|\omega\| .
\end{aligned}
$$

Hence, for $\tilde{C}_3 \coloneqq 4\pi\sqrt{2d_1}C^2\left(\sqrt{d_2}\sigma_{\min}^{-1} + \sqrt{d_3}\sigma_{\min}^{-2}\right)$, we get (6.29) by

$$
\begin{aligned}
\left\|\boldsymbol{G}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 &= \left\|\operatorname{grad}_\phi \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 + \left\|\operatorname{grad}_\alpha \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 \\
&\le 4\pi\sqrt{2d_1}C^2\left(\sqrt{d_2}\sigma_{\min}^{-1} + \sqrt{d_3}\sigma_{\min}^{-2}\right)\|\omega\| = \tilde{C}_3\|\omega\| .
\end{aligned}
$$

Since also the second derivatives of the drift functions are bounded by Assumption 2.14 (B2), we get with C from that Assumption that for $\tilde{C}_4 = \tilde{C}_1/\sqrt{d_1}$

$$
\begin{aligned}
&\left\|\boldsymbol{H}_\omega^{t,t'}(\theta;\phi,\alpha)\right\|_1 \\
&= \Big\| \Big(\omega_1 \Big[\frac{1}{\sigma_t^\alpha}\left(\cos\left(\rho_t^\phi\right) \operatorname{Hess}_\theta \left(\delta_t^\theta\right)_1 - \sin\left(\rho_t^\phi\right) \operatorname{Hess}_\theta \left(\delta_t^\theta\right)_2\right) \\
&\qquad - \frac{1}{\sigma_{t'}^\alpha}\left(\cos\left(\rho_{t'}^\phi\right) \operatorname{Hess}_\theta \left(\delta_{t'}^\theta\right)_1 - \sin\left(\rho_{t'}^\phi\right) \operatorname{Hess}_\theta \left(\delta_{t'}^\theta\right)_2\right)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+ \omega_2 \left[\frac{1}{\sigma_t^\alpha} \left(\sin\left(\rho_t^\phi\right) \text{ Hess}_\theta \left(\delta_t^\theta\right)_1 + \cos\left(\rho_t^\phi\right) \text{ Hess}_\theta \left(\delta_t^\theta\right)_2 \right) \right.\\
&\left.\left. - \frac{1}{\sigma_{t'}^\alpha} \left(\sin\left(\rho_{t'}^\phi\right) \text{ Hess}_\theta \left(\delta_{t'}^\theta\right)_1 + \cos\left(\rho_{t'}^\phi\right) \text{ Hess}_\theta \left(\delta_{t'}^\theta\right)_2 \right) \right] \right) \Big\|_1 \\
&\leq \sigma_{\min}^{-1} \left(|\omega_1| + |\omega_2|\right) \left(\left\| \text{Hess}_\theta \left(\delta_t^\theta\right)_1 \right\|_1 + \left\| \text{Hess}_\theta \left(\delta_t^\theta\right)_2 \right\|_1 \right.\\
&\qquad \left. + \left\| \text{Hess}_\theta \left(\delta_{t'}^\theta\right)_1 \right\|_1 + \left\| \text{Hess}_\theta \left(\delta_{t'}^\theta\right)_2 \right\|_1 \right) \\
&\leq 4\sqrt{2}\sigma_{\min}^{-1} C \|\omega\| = \tilde{C}_4 \|\omega\|,
\end{aligned}
$$

proving (6.27).
Plugging this and (6.25) into (6.23), we obtain for $\tilde{C}_5 := \max\left\{ \sqrt{d_1}\tilde{C}_1, \tilde{C}_4 \right\}$

$$
\begin{aligned}
\left\| \text{Hess}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha) \right\|_1 &\leq \left(\left\| \boldsymbol{H}_\omega^{t,t'}(\theta;\phi,\alpha) \right\|_1 + d_1 \left\| \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \right\|^2 \right) \left| h_\omega^{t,t'}(\theta;\phi,\alpha) \right| \\
&\leq \tilde{C}_5 \|\omega\| + \tilde{C}_5^2 \|\omega\|^2,
\end{aligned}
$$

which shows (6.28). Likewise, for $\tilde{C}_6 = \max\{ \sqrt{d_1}\tilde{C}_1, \sqrt{d_2 + d_3}\tilde{C}_2, \tilde{C}_3\}$

$$
\begin{aligned}
&\left\| \text{grad}_{(\phi,\alpha)} \text{grad}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha) \right\|_1 \\
&\leq \left(\left\| \boldsymbol{G}_\omega^{t,t'}(\theta;\phi,\alpha) \right\|_1 + \sqrt{d_1}\sqrt{d_2 + d_3} \left\| \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha) \right\| \left\| \boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha) \right\| \right) \left| h_\omega^{t,t'}(\theta;\phi,\alpha) \right| \\
&\leq \tilde{C}_6 \|\omega\| + \tilde{C}_6^2 \|\omega\|^2,
\end{aligned}
$$

which proves (6.30). Taking $\tilde{C} = \max\{\tilde{C}_1, \tilde{C}_2, \tilde{C}_3, \tilde{C}_4, \tilde{C}_5, \tilde{C}_6\}$ yields the claim and finishes the proof of the Lemma. □

6.2 Details of the proofs

In this section, we provide proofs of our theoretical results on the asymptotic properties of the estimators of the motion function parameters using the preparatory results from the previous Section 6.1 and the auxiliary results from the Appendix, Section A.

6.2.1 Proof of Theorem 3.1 (Consistency of the rotation and scaling parameter estimators)

In this subsection we give the detailed proofs of steps 1 to 3 in order to complete the proof of consistency of the rotation and scaling parameter estimators. It is very similar to the proof of Theorem 6.13 in Hartmann (2016), but is modified here to fit the new model, which better describes the data acquisition process. Furthermore, some technical issues have been resolved.

Step 1: uniqueness of the contrast minimizer (ϕ_0, α_0). First, note that because $F^t(u,v) = (\sigma_t^{\alpha_0})^{4-\gamma}\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)$ and $\left|d_{u,v}(\sigma,\rho)\right| = 1$, we have

$$\begin{aligned}
M(\phi,\alpha) &= -\int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\int_0^1 d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0})(\sigma_t^{\alpha_0})^{4-\gamma}\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\,\mathrm{d}t\right|^2 \mathrm{d}v\\
&= -\int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\left|\int_0^1 d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0})(\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t\right|^2 \mathrm{d}v\\
&\geq -\int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\left(\int_0^1 \left|d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0})(\sigma_t^{\alpha_0})^{4-\gamma}\right|\,\mathrm{d}t\right)^2 \mathrm{d}v\\
&= -\int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\left(\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t\right)^2 \mathrm{d}v
\end{aligned}\tag{6.32}$$

for all (ϕ,α) with equality if $(\phi,\alpha) = (\phi_0,\alpha_0)$. Let $(\phi,\alpha) \in \Phi\times \mathrm{A}$ such that equality holds. Since f is not scaling invariant by Assumption 2.13 (A2), by Lemma A.7 there are $u \in \mathbb{Z}$ and an open Borel set $B \subseteq \mathbb{R}$ with positive Lebesgue-measure such that $\mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \neq 0$ for all $v \in B$. Then, for equality in (6.32) to hold, we must have

$$\left|\int_0^1 d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0})(\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t\right| = \int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t$$

for this u and all $v \in B$. By Lemma A.1, $t \mapsto d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0})$ is constant a.e. on $[0,1]$. Because of the identifiability constraint (Assumption 2.14 (B1)), $(\sigma_0^{\alpha},\rho_0^{\phi}) = (1,0)$ for all (ϕ,α) and by the continuity of σ^{α} and ρ^{ϕ} as functions in t (Assumption 2.14 (B4)), this constant has to be 1, i.e.,

$$1 = d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0}) = \exp\Big(-\mathrm{i}v\big(\log(\sigma_t^{\alpha})-\log(\sigma_t^{\alpha_0})\big) + 2\pi\mathrm{i}u\big(\rho_t^{\phi}-\rho_t^{\phi_0}\big)\Big)\tag{6.33}$$

a.e. on $[0,1]$ for u as above and all $v \in B$. Choose $v_1, v_2 \in B \setminus \{v\}$ such that

$$v - v_1 \in \mathbb{Q}, \quad v - v_2 \in \mathbb{R} \setminus \mathbb{Q}. \tag{6.34}$$

From (6.33),

$$\exp\Big(-\mathrm{i}(v - v_1)\big(\log(\sigma_t^{\alpha}) - \log(\sigma_t^{\alpha_0})\big)\Big) = \frac{d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})}{d_{u,v_1}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})} = 1,$$

and similarly, $\exp\Big(-\mathrm{i}(v - v_2)\big(\log(\sigma_t^{\alpha}) - \log(\sigma_t^{\alpha_0})\big)\Big) = 1$, which implies that

$$(v - v_1)\big(\log(\sigma_t^{\alpha}) - \log(\sigma_t^{\alpha_0})\big)/(2\pi) \in \mathbb{Z},$$
$$(v - v_2)\big(\log(\sigma_t^{\alpha}) - \log(\sigma_t^{\alpha_0})\big)/(2\pi) \in \mathbb{Z}.$$

Because of (6.34), this means that $\log(\sigma_t^{\alpha}) - \log(\sigma_t^{\alpha_0}) = 0$ a.e. on $[0,1]$. Since the logarithm is bijective and the scaling functions are identifiable (Assumption 2.14 (B5)), we get $\alpha = \alpha_0$. Since f is not rotation invariant by Assumption 2.13 (A2), by Lemma A.8 there are $u \in \mathbb{Z} \setminus \{0\}$ and $v \in \mathbb{R}$ such that $\mathcal{M}_{|\mathcal{F}_f|^2}(u, v) \neq 0$. Then, for equality in (6.32) to hold, we must have

$$\left| \int_0^1 d_{u'',v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})(\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t \right| = \int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t.$$

Like above, we conclude that (6.33) holds for these u and v. It follows that

$$\exp\Big(2\pi\mathrm{i}u\big(\rho_t^{\phi} - \rho_t^{\phi_0}\big)\Big) = 1, \quad \text{a.e. on } [0,1],$$

which means that $u(\rho_t^{\phi} - \rho_t^{\phi_0}) \in \mathbb{Z}$ a.e. on $[0,1]$. The function $t \mapsto \rho_t^{\phi} - \rho_t^{\phi_0}$ is continuous by Assumption 2.14 (B4) and takes the value 0 at $t = 0$ (Assumption 2.14 (B1)). Hence, so does the function $t \mapsto u(\rho_t^{\phi} - \rho_t^{\phi_0})$. Together, this yields that $u(\rho_t^{\phi} - \rho_t^{\phi_0}) = 0$ a.e. on $[0,1]$. As $u \neq 0$, we obtain $\rho_t^{\phi} = \rho_t^{\phi_0}$ and the identifiability of the rotation functions (Assumption 2.14 (B5)) yields $\phi = \phi_0$.

Step 2: Continuity of the population contrast functional M**.** The second step follows essentially from Theorem B.3 on the continuity of parameter integrals. The measurability, continuity, and integrability conditions of Assumptions 2.13 (A3, A4) and 2.14 (B4) ensure the applicability of the mentioned result. More precisely, by Assumption 2.14 (B4), the functions $t \mapsto \sigma_t^{\alpha}$ and $t \mapsto \rho_t^{\alpha}$ are measurable for all $\alpha \in \mathrm{A}$, $\phi \in \Phi$. Therefore, the functions

$$t \mapsto d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})F^t(u, v) = d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})(\sigma_t^{\alpha_0})^{4-\gamma}\mathcal{M}_{|\mathcal{F}_f|^2}(u, v)$$

are measurable for all $u \in \mathbb{Z}$, $v \in \mathbb{R}$, $\alpha \in \mathrm{A}$, $\phi \in \Phi$, as they are concatenations of measurable functions. By the same Assumption, the functions $\alpha \mapsto \sigma_t^{\alpha}$ and $\phi \mapsto \rho_t^{\phi}$ are continuous for all $t \in [0,1]$. As a concatenation of continuous functions, $(\phi, \alpha) \mapsto d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})F^t(u, v)$

is also continuous for all $u \in \mathbb{Z}$, $v \in \mathbb{R}$, and $t \in [0,1]$. Furthermore, the constant function

$$t \mapsto \tilde{g}_{u,v} := \max\left\{\sigma_{\max}^{4-\gamma}, \sigma_{\min}^{4-\gamma}\right\} \mathcal{M}_{|\mathcal{F}_f|^2}(u,v)$$

is an integrable majorant for $t \mapsto d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})F^t(u,v)$. Thus, we can apply Theorem B.3 on the continuity of parameter integrals, to get that

$$(\phi,\alpha) \mapsto \int_0^1 d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})F^t(u,v)\,\mathrm{d}t$$

is continuous for all $u \in \mathbb{Z}$, $v \in \mathbb{R}$. Because $x \mapsto |x|^2$ is continuous, so is

$$(\phi,\alpha) \mapsto g_{u,v}^{\alpha,\phi} := \left|\int_0^1 d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})F^t(u,v)\,\mathrm{d}t\right|^2.$$

By Assumption 2.13 (A3) and Lemma A.3, $\mathcal{M}_{|\mathcal{F}_f|^2}$ is continuous. Since the function $(u,v) \mapsto d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})$ is continuous, too, we get that $(u,v) \mapsto d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})F^t(u,v)$ is continuous. Hence, by the same chain of arguments as above, the function $(u,v) \mapsto g_{u,v}^{\alpha,\phi}$ is continuous and as such Lebesgue measurable for all $\alpha \in \mathrm{A}$, $\phi \in \Phi$. Since $g_{u,v}^{\alpha,\phi} \leq \left|\tilde{g}_{u,v}\right|^2$ for all $\alpha \in \mathrm{A}$, $\phi \in \Phi$, and by Assumption 2.13 (A4),

$$\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\tilde{g}_{u,v}\right|^2 \mathrm{d}v = \max\left\{\sigma_{\max}^{8-2\gamma}, \sigma_{\min}^{8-2\gamma}\right\} \int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \mathrm{d}v < \infty,$$

the function $(u,v) \mapsto \left|\tilde{g}_{u,v}\right|^2$ is an integrable majorant for $(u,v) \mapsto g_{u,v}^{\alpha,\phi}$. Applying Theorem B.3 again yields the continuity of M.

Step 3: Convergence of the empirical contrast functional $M_T \overset{T\to\infty}{\longrightarrow} M$ in probability uniformly in (ϕ,α). By Assumption 2.15 (C3), $\mathcal{M}_{|\mathcal{F}_f|^2}(u,v) = \mathcal{M}^T_{|\mathcal{F}_f|^2}(u,v) + o\left((u_T v_T)^{-1}\right)$. Hence, from (2.15) and (2.18), we get

$$\mathcal{M}^T_{|Y^t|^2}(u,v) = d_{u,v}(1/\sigma_t^{\alpha_0}, -\rho_t^{\phi_0})F^t(u,v) + \mathcal{M}^T_{W^t}(u,v) + o\left((u_T v_T)^{-1}\right). \tag{6.35}$$

Therefore, the following decomposition is justified:

$$\begin{aligned} M_T(\phi,\alpha) &= -\int_{-v_T}^{v_T} \sum_{|u|\leq u_T} \left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}(\sigma_t^{\alpha},\rho_t^{\phi})\mathcal{M}^T_{|Y^t|^2}(u,v)\right|^2 \mathrm{d}v \\ &= -\int_{-v_T}^{v_T} \sum_{|u|\leq u_T} \left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0}, \rho_t^{\phi} - \rho_t^{\phi_0})F^t(u,v) \right. \\ &\qquad \left. + d_{u,v}(\sigma_t^{\alpha},\rho_t^{\phi})\mathcal{M}^T_{W^t}(u,v)\right|^2 \mathrm{d}v + o(1) \\ &= A_T(\phi,\alpha) + B_T(\phi,\alpha) + C_T(\phi,\alpha) + o(1), \end{aligned} \tag{6.36}$$

with

$$\begin{aligned}
A_T(\phi,\alpha) &:= -\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0})F^t(u,v)\right|^2 \mathrm{d}v,\\
B_T(\phi,\alpha) &:= -\int_{-v_T}^{v_T}\sum_{|u|\le u_T} 2\Re\Bigg[\left(\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0})F^t(u,v)\right)\\
&\qquad\cdot\left(\frac{\beta_T}{T}\sum_{t'\in\mathbb{T}}\overline{d_{u,v}(\sigma_{t'}^{\alpha},\rho_{t'}^{\phi})\mathcal{M}^T_{\mathcal{W}^{t'}}(u,v)}\right)\Bigg]\mathrm{d}v,\\
C_T(\phi,\alpha) &:= -\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}(\sigma_t^{\alpha},\rho_t^{\phi})\mathcal{M}^T_{W^t}(u,v)\right|^2 \mathrm{d}v,
\end{aligned}$$

where we used that $|a\pm b|^2 = |a|^2 \pm 2\Re(a\bar{b}) + |b|^2$ for $a,b\in\mathbb{C}$. The idea is to show the convergence of the deterministic part $A_T \overset{T\to\infty}{\longrightarrow} M$ uniformly in (ϕ,α) and of the random part $B_T + C_T \overset{T\to\infty}{\longrightarrow} 0$ in probability uniformly in (ϕ,α).
Recall that with $d_{u,v}^{t,t'}(\phi,\alpha)$ from (6.2), we have $d_{u,v}(\sigma_t^{\alpha}/\sigma_t^{\alpha_0},\rho_t^{\phi}-\rho_t^{\phi_0}) = d_{u,v}^{t,0}(\phi,\alpha)$. It holds that

$$\begin{aligned}
&|A_T(\phi,\alpha)-M(\phi,\alpha)|\\
={}& \Bigg|\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\left(\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}^{t,0}(\phi,\alpha)F^t(u,v)\right|^2 - \left|\int_0^1 d_{u,v}^{t,0}(\phi,\alpha)F^t(u,v)\,\mathrm{d}t\right|^2\right)\mathrm{d}v\\
&-\left(\int_{-\infty}^{-v_T}\sum_{|u|\le u_T}+\int_{v_T}^{\infty}\sum_{|u|\le u_T}+\int_{\mathbb{R}}\sum_{|u|>u_T}\right)\left|\int_0^1 d_{u,v}^{t,0}(\phi,\alpha)F^t(u,v)\,\mathrm{d}t\right|^2\mathrm{d}v\Bigg|\\
={}& \Bigg|\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\left(\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma}\right|^2 - \left|\int_0^1 d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t\right|^2\right)\mathrm{d}v\\
&-\left(\int_{-\infty}^{-v_T}\sum_{|u|\le u_T}+\int_{v_T}^{\infty}\sum_{|u|\le u_T}+\int_{\mathbb{R}}\sum_{|u|>u_T}\right)\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\left|\int_0^1 d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t\right|^2\mathrm{d}v\Bigg|\\
\le{}& 2C_\gamma\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma} - \int_0^1 d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma}\,\mathrm{d}t\right|\mathrm{d}v\\
&+\left(\int_{-\infty}^{-v_T}\sum_{|u|\le u_T}+\int_{v_T}^{\infty}\sum_{|u|\le u_T}+\int_{\mathbb{R}}\sum_{|u|>u_T}\right)\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\int_0^1\left|d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma}\right|^2\mathrm{d}t\,\mathrm{d}v\\
\le{}& \frac{2C_\gamma\beta_T}{T}\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\mathrm{TV}\left(t\mapsto d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma}\right)\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\mathrm{d}v\\
&+C_\gamma\left(\int_{-\infty}^{-v_T}\sum_{|u|\le u_T}+\int_{v_T}^{\infty}\sum_{|u|\le u_T}+\int_{\mathbb{R}}\sum_{|u|>u_T}\right)\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\mathrm{d}v, \qquad (6.37)
\end{aligned}$$

where for the first inequality we used $|a|^2 - |b|^2 \leq 2C\,|a-b|$ for $a, b \in \mathbb{C}$ such that $|a|, |b| \leq C$ with the constant

$$C_\gamma := \max\left\{\sigma_{\min}^{8-2\gamma}, \sigma_{\max}^{8-2\gamma}, \sigma_{\min}^{4-\gamma}, \sigma_{\max}^{4-\gamma}\right\}, \tag{6.38}$$

and for the second inequality we used part 2 of Lemma A.4. Because

$$\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \, \mathrm{d}v < \infty \tag{6.39}$$

by Assumption 2.13 (A4), and because $u_T, v_T \overset{T\to\infty}{\longrightarrow} \infty$ by Assumption 2.15 (C1), the second term in (6.37) vanishes for $T \to \infty$. To tackle the first term, we show that

$$\sup_{(\phi,\alpha)\in\Phi\times\mathrm{A}} \mathrm{TV}\left(t \mapsto d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma}\right) \leq C_4\,\|(v,u)\| + C_3, \tag{6.40}$$

with some constants $C_3, C_4 > 0$. First of all, $\sigma^\alpha\colon t \mapsto \sigma_t^\alpha$ and $\rho^\phi\colon t \mapsto \rho_t^\phi$ are of bounded variation uniformly in $\alpha \in \mathrm{A}$, $\phi \in \Phi$, by Assumption 2.14 (B7). It holds that $\sigma \mapsto \sigma^{4-\gamma}$ is Lipschitz-continuous on $[\sigma_{\min}, \sigma_{\max}]$ with constant in $\left\{(4-\gamma)\sigma_{\max}^{3-\gamma}, (4-\gamma)\sigma_{\min}^{-3+\gamma}, 0, (-4+\gamma)\sigma_{\min}^{-3+\gamma}\right\}$ (for $\gamma < 3$, $3 < \gamma < 4$, $\gamma = 4$, and $\gamma > 4$, respectively). Furthermore, the logarithm, restricted to the interval $[\sigma_{\min}, \sigma_{\max}]$, is differentiable with derivative bounded by $\log'(\sigma_{\min}) = 1/\sigma_{\min}$. Hence, $\log_{|[\sigma_{\min},\sigma_{\max}]}$ is Lipschitz-continuous with Lipschitz-constant $1/\sigma_{\min}$. Since $t \mapsto \sigma_t^\alpha$ is of bounded variation uniformly in $\alpha \in \mathrm{A}$, so is $t \mapsto \log(\sigma_t^\alpha)$. Together with part 4 of Lemma A.4 on the total variation of linear combinations of functions, this implies that there are $C_1, C_2, C_3 > 0$ such that

$$\mathrm{TV}\big(\log(\sigma^\alpha) - \log(\sigma^{\alpha_0})\big) \leq C_1, \quad \mathrm{TV}\big(\rho^\phi - \rho^{\phi_0}\big) \leq C_2, \quad \mathrm{TV}\big((\sigma^{\alpha_0})^{4-\gamma}\big) \leq C_3, \tag{6.41}$$

uniformly in (ϕ, α). Since $\|x\| \leq \|x\|_1$ for all $x \in \mathbb{R}^2$, we have for all $t, t' \in [0,1]$, that

$$\begin{aligned}
& \Big|\exp\Big(\mathrm{i}\,\Big\langle\big(-v, 2\pi u\big), \big(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_t^{\phi_0}\big)\Big\rangle\Big) \\
& - \exp\Big(\mathrm{i}\,\Big\langle\big(-v, 2\pi u\big), \big(\log(\sigma_{t'}^\alpha) - \log(\sigma_{t'}^{\alpha_0}), \rho_{t'}^\phi - \rho_{t'}^{\phi_0}\big)\Big\rangle\Big)\Big| \\
\leq\; & \sqrt{2}\,\big\|\big(2\pi u, -v\big)\big\| \\
& \cdot \Big\|\big(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_t^{\phi_0}\big) - \big(\log(\sigma_{t'}^\alpha) - \log(\sigma_{t'}^{\alpha_0}), \rho_{t'}^\phi - \rho_{t'}^{\phi_0}\big)\Big\|_1 \\
\leq\; & 2\pi\sqrt{2}\,\|(u,v)\|\,\Big(\Big|\big(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0})\big) - \big(\log(\sigma_{t'}^\alpha) - \log(\sigma_{t'}^{\alpha_0})\big)\Big| \\
& + \Big|\big(\rho_t^\phi - \rho_t^{\phi_0}\big) - \big(\rho_{t'}^\phi - \rho_{t'}^{\phi_0}\big)\Big|\Big),
\end{aligned}$$

where for the first inequality we used that $x \mapsto \mathrm{e}^{\mathrm{i}\langle a,x\rangle}$, is Lipschitz-continuous with Lipschitz-constant $\sqrt{2}\,\|a\|$ for $a \in \mathbb{R}^2$. Hence,

$$\begin{aligned}
\mathrm{TV}\Big(t \mapsto d_{u,v}^{t,0}(\phi,\alpha)\Big) &= \mathrm{TV}\Big(t \mapsto \exp\Big(\mathrm{i}\,\Big\langle\big(-v, 2\pi u\big), \big(\log(\sigma_t^\alpha) - \log(\sigma_t^{\alpha_0}), \rho_t^\phi - \rho_t^{\phi_0}\big)\Big\rangle\Big)\Big) \\
&\leq 2\pi\sqrt{2}(C_1 + C_2)\,\|(u,v)\|\,.
\end{aligned}$$

Now, part 1 of Lemma A.4 yields

$$\begin{aligned}
&\operatorname{TV}\big(t \mapsto d_{u,v}^{t,0}(\phi,\alpha)(\sigma_t^{\alpha_0})^{4-\gamma}\big)\\
\leq\ &\big\|t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma}\big\|_\infty \operatorname{TV}\big(t \mapsto d_{u,v}^{t,0}(\phi,\alpha)\big) + \big\|t \mapsto d_{u,v}^{t,0}(\phi,\alpha)\big\|_\infty \operatorname{TV}\big(t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma}\big)\\
\leq\ &C_4 \|(u,v)\| + C_3,
\end{aligned}$$

uniformly in (ϕ,α), where $C_4 := 2\pi\sqrt{2C_\gamma}(C_1 + C_2)$ with C_γ from (6.38), proving (6.40). From (6.37) and (6.40), we get

$$\begin{aligned}
|A_T(\phi,\alpha) - M(\phi,\alpha)| &\leq \frac{2C_\gamma\beta_T}{T}\int_{-v_T}^{v_T}\sum_{|u|\leq u_T}\big(C_4\|(u,v)\| + C_3\big)\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \mathrm{d}v + o(1)\\
&\leq \frac{2C_\gamma C_5\beta_T}{T} + o(1) = o(1),
\end{aligned}$$

where the integral is bounded by some constant $C_5 > 0$ because of Assumption 2.13 (A4) and the Sobolev Imbedding Theorem (see Theorem B.11). Since C_γ and C_5 do not depend on (ϕ,α),

$$A_T \overset{T\to\infty}{\longrightarrow} M \quad \text{uniformly in } (\phi,\alpha). \tag{6.42}$$

Next, we show that $\mathbb{E}C_T \overset{T\to\infty}{\longrightarrow} 0$ uniformly in (ϕ,α), which implies uniform convergence $C_T \overset{T\to\infty}{\longrightarrow} 0$ in probability due to Markov's inequality (Theorem B.10). With the Cauchy-Schwarz inequality (Theorem B.1) and Lemma 6.2, we get that

$$\begin{aligned}
0 &\geq \inf_{(\phi,\alpha)\in\Phi\times\mathrm{A}} \mathbb{E}C_T(\phi,\alpha)\\
&= \inf_{(\phi,\alpha)\in\Phi\times\mathrm{A}} \mathbb{E}\left(-\int_{-v_T}^{v_T}\sum_{|u|\leq u_T}\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} d_{u,v}(\sigma_t^\alpha,\rho_t^\phi)\mathcal{M}_{W^t}^T(u,v)\right|^2 \mathrm{d}v\right)\\
&\geq \inf_{(\phi,\alpha)\in\Phi\times\mathrm{A}} \mathbb{E}\left(-\int_{-v_T}^{v_T}\sum_{|u|\leq u_T}\left(\frac{\beta_T}{T}\sum_{t'\in\mathbb{T}}\left|d_{u,v}(\sigma_{t'}^\alpha,\rho_{t'}^\phi)\right|^2\right)\left(\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\left|\mathcal{M}_{W^t}^T(u,v)\right|^2\right)\mathrm{d}v\right)\\
&= -\int_{-v_T}^{v_T}\sum_{|u|\leq u_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\mathbb{E}\left(\left|\mathcal{M}_{W^t}^T(u,v)\right|^2\right)\mathrm{d}v\\
&\geq -\int_{-v_T}^{v_T}\sum_{|u|\leq u_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\mathcal{O}\left(\frac{r_T^{2\gamma}}{\beta_T}\right)\mathrm{d}v\\
&= \mathcal{O}\left(\frac{r_T^{2\gamma}u_T v_T}{\beta_T}\right).
\end{aligned}$$

Since $r_T^{2\gamma}u_T v_T\beta_T^{-1} \overset{T\to\infty}{\longrightarrow} 0$ by Assumption 2.15 (C2), $\mathbb{E}C_T(\phi,\alpha) \overset{T\to\infty}{\longrightarrow} 0$ uniformly in (ϕ,α), and thus,

$$C_T(\phi,\alpha) \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability uniformly in } (\phi,\alpha). \tag{6.43}$$

Finally, the Cauchy-Schwarz inequality (Theorem B.1) and Slutzky's Lemma (Theorem B.9) imply that

$$\left(B_T(\phi,\alpha)\right)^2 \leq 4A_T(\phi,\alpha)C_T(\phi,\alpha) \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability uniformly in } (\phi,\alpha). \tag{6.44}$$

Applying Slutzky's Lemma (Theorem B.9) again, it then follows from (6.42), (6.43), and (6.44) that

$$M_T(\phi,\alpha) = A_T(\phi,\alpha) + o_{\mathbb{P}}(1) \overset{T\to\infty}{\longrightarrow} M(\phi,\alpha)$$

in probability uniformly in (ϕ,α), finishing the proof of Theorem 3.1. □

6.2.2 Proof of Theorem 3.2 (Consistency of the drift parameter estimator)

As already indicated in the sketch of the proof from Section 3.1, we use asymptotic normality of the rotation and drift parameter estimators as given by Theorem 3.5 to show consistency of the drift parameter estimator. More specifically, we will need some implications of this theorem, which are stated in Lemma 6.8 below. Using that, we give the thorough demonstration of Theorem 3.2 (Steps 1 to 3), proving consistency of the drift parameter estimator. The argument is close to the proof of Theorem 5.16 in Hartmann (2016). However, we adjusted the details using our generalized expression for the drift correction error term $h_\omega^{t,t'}(\theta;\phi,\alpha)$ depending on all three motion function parameters.

Lemma 6.8. *Under the Assumption 2.14 (B3), if $\sqrt{T}(\hat\phi_T-\phi_0,\hat\alpha_T-\alpha_0)$ is asymptotically centered normal, we have for all $\omega\in\mathbb{R}^2\setminus\{0\}$ and $t\in[0,1]$, that for $\tau_t^{(\phi,\alpha)} = \frac{\sigma_t^{\alpha_0}}{\sigma_t^{\alpha}} R_{\rho_t^{\phi}-\rho_t^{\phi_0}}$ from Definition 2.21,*

$$\sqrt{T}\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega - \omega\right)$$

is asymptotically centered normal. Furthermore, it holds that

$$\left|\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right) - \mathcal{F}_f(\omega)\right| = O_{\mathbb{P}}\left(T^{-1/2}\right), \tag{6.45}$$

implying

$$\left|\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\right| = O_{\mathbb{P}}\left(T^{-1/2} + \left|\mathcal{F}_f(\omega)\right|\right), \tag{6.46}$$

and

$$\left|\Im\left[\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat\phi_T,\hat\alpha_T)}\omega\right)}\right]\right| \leq O_{\mathbb{P}}\left(\frac{1}{T}\right) + 2\left|\mathcal{F}_f(\omega)\right| O_{\mathbb{P}}\left(\frac{1}{\sqrt{T}}\right). \tag{6.47}$$

Proof. For $\omega\in\mathbb{R}^2\setminus\{0\}$ and $t\in[0,1]$, let

$$g_\omega^t\colon \mathbb{R}^{d_2+d_3}\to\mathbb{R}^2, \quad (\phi,\alpha)\mapsto \frac{\sigma_t^{\alpha_0}}{\sigma_t^{\alpha}} R_{\rho_t^{\phi}-\rho_t^{\phi_0}}\omega.$$

By Assumption 2.14 (B3), there is an open neighborhood $U\subseteq\Phi\times\mathrm{A}$ of (ϕ_0,α_0) such that $\phi\mapsto\rho_t^{\phi}$ and $\alpha\mapsto\sigma_t^{\alpha}$ are continuously differentiable on U. Hence, g_ω^t is continuously differentiable on U. Because $g_\omega^t(\phi_0,\alpha_0)=\omega$, applying the Delta method (Theorem B.7) yields the first assertion. From this, the second line follows using Assumptions 2.13 (A1) and (A3), since the Fourier transform $\mathcal{F}_g\colon \mathbb{R}^2 \to \mathbb{C}$ of functions $g\colon \mathbb{R}^2 \to \mathbb{C}$, which fulfill the condition that $L_g := 2\pi\sqrt{2}\int_{\mathbb{R}^2}\|x\|\,|g(x)|\,\mathrm{d}x < \infty$, is Lipschitz-continuous with Lipschitz-constant L_g:

$$\left|\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right) - \mathcal{F}_f(\omega)\right| \le L_f \left\|\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega - \omega\right\| = O_{\mathbb{P}}\left(\frac{1}{\sqrt{T}}\right).$$

The third statement is a direct implication of this:

$$\left|\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\right| \le \left|\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right) - \mathcal{F}_f(\omega)\right| + \left|\mathcal{F}_f(\omega)\right| = O_{\mathbb{P}}\left(\frac{1}{\sqrt{T}} + \left|\mathcal{F}_f(\omega)\right|\right).$$

Moreover, since $\Im\left[\mathcal{F}_f(\omega)\overline{\mathcal{F}_f(\omega)}\right] = \Im\left[\left|\mathcal{F}_f(\omega)\right|^2\right] = 0$, it follows that

$$\begin{aligned}
&\left|\Im\left[\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat\phi_T,\hat\alpha_T)}\omega\right)}\right]\right| \\
= &\left|\Im\left[\left\{\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right) - \mathcal{F}_f(\omega)\right\}\overline{\left\{\mathcal{F}_f\left(\tau_{t'}^{(\hat\phi_T,\hat\alpha_T)}\omega\right) - \mathcal{F}_f(\omega)\right\}}\right.\right. \\
&\left.\left. + \left\{\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right) - \mathcal{F}_f(\omega)\right\}\overline{\mathcal{F}_f(\omega)} + \mathcal{F}_f(\omega)\overline{\left\{\mathcal{F}_f\left(\tau_{t'}^{(\hat\phi_T,\hat\alpha_T)}\omega\right) - \mathcal{F}_f(\omega)\right\}}\right]\right| \\
\le\ & O_{\mathbb{P}}\left(\frac{1}{T}\right) + 2\left|\mathcal{F}_f(\omega)\right| O_{\mathbb{P}}\left(\frac{1}{\sqrt{T}}\right). \quad \square
\end{aligned}$$

Step 1: Uniqueness of contrast minimizer θ_0. Using that $\left|h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\right| = 1$, we have that

$$\begin{aligned}
N(\theta;\phi_0,\alpha_0) &= -\int_{\mathbb{R}^2}\left|\mathcal{F}_f(\omega)\right|^2\left|\int_0^1 h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\,\mathrm{d}t\right|^2 \mathrm{d}\omega \\
&\ge -\int_{\mathbb{R}^2}\left|\mathcal{F}_f(\omega)\right|^2\left(\int_0^1\left|h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\right|\mathrm{d}t\right)^2 \mathrm{d}\omega \\
&\ge -\int_{\mathbb{R}^2}\left|\mathcal{F}_f(\omega)\right|^2 \mathrm{d}\omega
\end{aligned} \tag{6.48}$$

for all $\theta \in \Theta$ with equality if $\theta = \theta_0$. Let $\theta \in \Theta$ such that equality in (6.48) holds. Since f is not translation invariant by Assumption 2.13 (A2), by Lemma A.9, there is an open Borel set $B \subseteq \mathbb{R}^2$ with positive Lebesgue measure such that $\mathcal{F}_f(\omega) \neq 0$ for all $\omega \in B$. Since equality in (6.48) holds, we have

$$\left|\int_0^1 h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\,\mathrm{d}t\right| = 1 \quad \text{for all } \omega \in B.$$

By Lemma A.1, the function $t \mapsto h_\omega^{t,0}(\theta;\phi_0,\alpha_0)$ is constant a.e. on $[0,1]$. Because of the identifiability constraint $\delta_0^\theta = 0$ for all $\theta \in \Theta$ (see Assumption 2.14 (B1)), the continuity of δ^θ at $t = 0$ (Assumption 2.14 (B4)), and the continuous mapping theorem (Theorem B.8), this constant has to be 1, i.e., for all $\omega \in B$,

$$1 = h_\omega^{t,0}(\theta;\phi_0,\alpha_0) = \exp\left(2\pi\mathrm{i}\left\langle\omega, 1/\sigma_t^{\alpha_0}\cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right\rangle\right), \tag{6.49}$$

a.e. on $[0,1]$. Fix $\omega = (\omega_1, \omega_2) \in B$ and choose $\omega_1^{(1)}, \omega_1^{(2)} \in \mathbb{R} \setminus \{\omega_1\}$ and $\omega_2^{(1)}, \omega_2^{(2)} \in \mathbb{R} \setminus \{\omega_2\}$ such that $(\omega_1^{(1)}, \omega_2), (\omega_1^{(2)}, \omega_2), (\omega_1, \omega_2^{(1)}), (\omega_1, \omega_2^{(2)}) \in B$ and

$$\omega_1 - \omega_1^{(1)}, \omega_2 - \omega_2^{(1)} \in \mathbb{Q}, \quad \omega_1 - \omega_1^{(2)}, \omega_2 - \omega_2^{(2)} \in \mathbb{R} \setminus \mathbb{Q}. \tag{6.50}$$

From (6.49),

$$\begin{aligned} 1 &= \frac{\exp\left(2\pi \mathrm{i} \left\langle \omega, 1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right\rangle\right)}{\exp\left(2\pi \mathrm{i} \left\langle (\omega_1^{(1)}, \omega_2), 1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) \right\rangle\right)} \\ &= \exp\left(2\pi \mathrm{i}(\omega_1 - \omega_1^{(1)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right)_1\right), \end{aligned}$$

and similarly,

$$\begin{aligned} 1 &= \exp\left(2\pi \mathrm{i}(\omega_1 - \omega_1^{(2)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right)_1\right), \\ 1 &= \exp\left(2\pi \mathrm{i}(\omega_2 - \omega_2^{(1)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right)_2\right), \\ \text{and } 1 &= \exp\left(2\pi \mathrm{i}(\omega_2 - \omega_2^{(2)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right)_2\right). \end{aligned}$$

This implies that

$$\begin{aligned} (\omega_1 - \omega_1^{(1)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right)_1 &\in \mathbb{Z}, \\ (\omega_1 - \omega_1^{(2)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right)_1 &\in \mathbb{Z}, \\ (\omega_2 - \omega_2^{(1)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right)_2 &\in \mathbb{Z}, \\ \text{and } (\omega_2 - \omega_2^{(2)}) \left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0})\right)_2 &\in \mathbb{Z}. \end{aligned}$$

Because of (6.50), we get that

$$1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}(\delta_t^\theta - \delta_t^{\theta_0}) = (0,0)^\top \quad \text{a.e. on } [0,1]. \tag{6.51}$$

Since $R_{-\rho_t^{\phi_0}}$ is a rotation matrix,

$$\det\left(1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}}\right) = (\sigma_t^{\alpha_0})^{-2} \geq (\sigma_{\max})^{-2} > 0 \quad \text{for all } t \in [0,1],$$

which, together with (6.51), implies that $\delta_t^\theta = \delta_t^{\theta_0}$ a.e. on $[0,1]$. Since the drift function is identifiable by Assumption 2.14 (B5), we conclude that $\theta = \theta_0$.

Step 2: Continuity of the population contrast functional for drift $N(\,\cdot\,; \phi_0, \alpha_0)$.

By Assumption 2.14 (B4), the motion functions δ^θ, ρ^ϕ, and σ^α are measurable for all $\theta \in \Theta$, $\phi \in \Phi$ and $\alpha \in \mathrm{A}$. Hence, $t \mapsto h_\omega^{t,0}(\theta; \phi, \alpha)$ is measurable for all $\theta \in \Theta$, $\phi \in \Phi$, $\alpha \in \mathrm{A}$, and $\omega \in \mathbb{R}^2$, as a concatenation of measurable functions. By the same assumption, the functions $\theta \mapsto \delta_t^\theta$, $\phi \mapsto \rho_t^\phi$, and $\alpha \mapsto \sigma_t^\alpha$ are continuous for all $t \in [0,1]$, giving that $(\theta, \phi, \alpha) \mapsto h_\omega^{t,0}(\theta; \phi, \alpha)$ is continuous for all $t \in [0,1]$ and $\omega \in \mathbb{R}^2$, as a concatenation of continuous functions. Furthermore, $t \mapsto 1$ is an integrable majorant for $t \mapsto h_\omega^{t,0}(\theta;\phi,\alpha)$. Consequently, Theorem B.3 on the continuity of parameter integrals yields that

$$(\theta,\phi,\alpha) \mapsto \int_0^1 h_\omega^{t,0}(\theta;\phi,\alpha)\,\mathrm{d}t$$

is continuous for all $\omega \in \mathbb{R}^2$. Since $x \mapsto |x|^2$ is continuous, we get that

$$(\theta,\phi,\alpha) \mapsto g_\omega^{\theta,\phi,\alpha} := \left|\int_0^1 h_\omega^{t,0}(\theta;\phi,\alpha)\,\mathrm{d}t\right|^2$$

is continuous for all $\omega \in \mathbb{R}^2$. By the same argument, because $\omega \mapsto h_\omega^{t,0}(\theta;\phi,\alpha)$ is continuous for all $(\theta,\phi,\alpha,t) \in \Theta \times \Phi \times \mathrm{A} \times [0,1]$, $\omega \mapsto g_\omega^{\theta,\phi,\alpha}$ is continuous, too, and hence, Lebesgue-measurable. Since $\left|\mathcal{F}_f(\omega)\right|^2$ is continuous in ω as Fourier transform of a function that satisfies $\int_{\mathbb{R}^2} \|x\|\,|f(x)|\,\mathrm{d}x < \infty$ (by Assumptions 2.13 (A1, A3)) and since $\left|\mathcal{F}_f(\omega)\right|^2$ is constant in (θ,ϕ,α), the product $\left|\mathcal{F}_f(\omega)\right|^2 g_\omega^{\theta,\phi,\alpha}$ is also continuous in ω as well as in (θ,ϕ,α) as a concatenation of continuous functions. In particular, $\omega \mapsto \left|\mathcal{F}_f(\omega)\right|^2 g_\omega^{\theta,\phi,\alpha}$ is Lebesgue-measurable. Furthermore, $\omega \mapsto \left|\mathcal{F}_f(\omega)\right|^2$ is an integrable majorant for $\omega \mapsto \left|\mathcal{F}_f(\omega)\right|^2 g_\omega^{\theta,\phi,\alpha}$ because of Assumption 2.13 (A3) and $g_\omega^{\theta,\phi,\alpha} \le 1$. Hence, we can apply Theorem B.3 on the continuity of parameter integrals again to get the continuity of $N(\,\cdot\,;\phi_0,\alpha_0)$.

Step 3: Convergence of $N_T(\theta;\hat\phi_T,\hat\alpha_T) \overset{T\to\infty}{\longrightarrow} N(\theta;\phi_0,\alpha_0)$ in probability uniformly in θ.

Recall the decomposition $N_T(\theta;\hat\phi_T,\hat\alpha_T) = A_T(\theta;\hat\phi_T,\hat\alpha_T) + B_T(\theta;\hat\phi_T,\hat\alpha_T) + C_T(\theta;\hat\phi_T,\hat\alpha_T)$ from (2.23). We will show convergence of the first term $A_T(\theta;\hat\phi_T,\hat\alpha_T) \overset{T\to\infty}{\longrightarrow} N(\theta;\phi_0,\alpha_0)$ in probability uniformly in θ, and of the other two, $B_T(\theta;\hat\phi_T,\hat\alpha_T) + C_T(\theta;\hat\phi_T,\hat\alpha_T) \overset{T\to\infty}{\longrightarrow} 0$ in probability uniformly in θ. Because of Assumption 2.13 (A3) and $\Omega_T \nearrow \mathbb{R}^2$ as $T \to \infty$, we have that

$$\int_{\mathbb{R}^2\setminus\Omega_T} \left|\mathcal{F}_f(\omega)\right|^2 \left|\int_0^1 h_\omega^{t,0}\left(\theta;\hat\phi_T,\hat\alpha_T\right)\,\mathrm{d}t\right|^2 \mathrm{d}\omega \le \int_{\mathbb{R}^2\setminus\Omega_T} \left|\mathcal{F}_f(\omega)\right|^2 \mathrm{d}\omega \overset{T\to\infty}{\longrightarrow} 0.$$

Hence,

$$\begin{aligned}
&\left|A_T(\theta;\hat\phi_T,\hat\alpha_T) - N(\theta;\phi_0,\alpha_0)\right| \\
=\;& \left|\int_{\Omega_T} \left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{t,0}\left(\theta;\hat\phi_T,\hat\alpha_T\right)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat\alpha_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\right|^2 \mathrm{d}\omega \right.\\
&\left. -\left(\int_{\Omega_T} + \int_{\mathbb{R}^2\setminus\Omega_T}\right)\left|\int_0^1 h_\omega^{t,0}\left(\theta;\phi_0,\alpha_0\right)\mathcal{F}_f\left(\omega\right)\,\mathrm{d}t\right|^2 \mathrm{d}\omega\right| \\
=\;& \left|\int_{\Omega_T} \left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{t,0}\left(\theta;\hat\phi_T,\hat\alpha_T\right)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat\alpha_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\right|^2 \right.\\
&\left. -\left|\int_0^1 h_\omega^{t,0}\left(\theta;\phi_0,\alpha_0\right)\mathcal{F}_f\left(\omega\right)\,\mathrm{d}t\right|^2 \mathrm{d}\omega\right| + o(1).
\end{aligned}$$

Since $|a|^2 - |b|^2 \leq 2C\,|a-b|$ for $a, b \in \mathbb{C}$ and $C > 0$ such that $|a|, |b| \leq C$ holds, it follows that

$$\begin{aligned}
& \left|A_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) - N(\theta;\phi_0,\alpha_0)\right| \\
\leq\ & \int_{\Omega_T} 2\left\|\mathcal{F}_f\right\|_\infty \left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^2 \left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{t,0}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)\right. \\
& \left. - \int_0^1 h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\,\mathcal{F}_f(\omega)\,\mathrm{d}t\right| \mathrm{d}\omega + o(1) \\
=\ & 2\left\|\mathcal{F}_f\right\|_\infty \left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^2 \int_{\Omega_T}\left|\sum_{t\in\mathbb{T}}\int_0^{\beta_T/T} h_\omega^{t,0}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)\right. \\
& \left. - h_\omega^{t+t',0}(\theta;\phi_0,\alpha_0)\,\mathcal{F}_f(\omega)\,\mathrm{d}t'\right| \mathrm{d}\omega + o(1) \\
\leq\ & 2\left\|\mathcal{F}_f\right\|_\infty \left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^2 \int_{\Omega_T}\sum_{t\in\mathbb{T}}\int_0^{\beta_T/T} \qquad (6.52) \\
& \left(\left|h_\omega^{t,0}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) - h_\omega^{t,0}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right)\mathcal{F}_f(\omega)\right|\right. \\
& \left. + \left|h_\omega^{t,0}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right)\mathcal{F}_f(\omega) - h_\omega^{t+t',0}(\theta;\phi_0,\alpha_0)\,\mathcal{F}_f(\omega)\right|\right)\mathrm{d}t'\,\mathrm{d}\omega + o(1) \\
=\ & 2\left\|\mathcal{F}_f\right\|_\infty \left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^2 \int_{\Omega_T}\sum_{t\in\mathbb{T}}\int_0^{\beta_T/T}\left(\left|\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) - \mathcal{F}_f(\omega)\right|\right. \\
& \left. + \left|h_\omega^{t,0}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right) - h_\omega^{t+t',0}(\theta;\phi_0,\alpha_0)\right|\left|\mathcal{F}_f(\omega)\right|\right)\mathrm{d}t'\,\mathrm{d}\omega + o(1) \qquad (6.53)
\end{aligned}$$

First, we consider the second part of (6.53). Because $x \mapsto \mathrm{e}^{\mathrm{i}\langle a,x\rangle}$, is Lipschitz-continuous with Lipschitz-constant $\sqrt{2}\,\|a\|$ for $a \in \mathbb{R}^2$, we get

$$\begin{aligned}
& \int_{\Omega_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\int_0^{\beta_T/T}\left|h_\omega^{t,0}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right) - h_\omega^{t+t',0}(\theta;\phi_0,\alpha_0)\right|\left|\mathcal{F}_f(\omega)\right|\,\mathrm{d}t'\,\mathrm{d}\omega \\
=\ & \int_{\Omega_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\int_0^{\beta_T/T}\left|\exp\left(2\pi\mathrm{i}\left\langle 1/\sigma_t^{\hat{\alpha}_T}\cdot R_{\rho_t^{\hat{\phi}_T}}\omega, \delta_t^\theta - \delta_t^{\theta_0}\right\rangle\right)\right. \\
& \left. - \exp\left(2\pi\mathrm{i}\left\langle 1/\sigma_{t+t'}^{\alpha_0}\cdot R_{\rho_{t+t'}^{\phi_0}}\omega, \delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}\right\rangle\right)\right|\left|\mathcal{F}_f(\omega)\right|\,\mathrm{d}t'\,\mathrm{d}\omega \\
=\ & \int_{\Omega_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\int_0^{\beta_T/T}\left|\exp\left(2\pi\mathrm{i}\left\langle \omega, 1/\sigma_t^{\hat{\alpha}_T}\cdot R_{-\rho_t^{\hat{\phi}_T}}(\delta_t^\theta - \delta_t^{\theta_0})\right\rangle\right)\right. \\
& \left. - \exp\left(2\pi\mathrm{i}\left\langle \omega, 1/\sigma_{t+t'}^{\alpha_0}\cdot R_{-\rho_{t+t'}^{\phi_0}}(\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0})\right\rangle\right)\right|\left|\mathcal{F}_f(\omega)\right|\,\mathrm{d}t'\,\mathrm{d}\omega \\
\leq\ & 2^{3/2}\pi\int_{\Omega_T}\|\omega\|\left|\mathcal{F}_f(\omega)\right|\,\mathrm{d}\omega\sum_{t\in\mathbb{T}}\int_0^{\beta_T/T} \\
& \left\|1/\sigma_t^{\hat{\alpha}_T}\cdot R_{-\rho_t^{\hat{\phi}_T}}(\delta_t^\theta - \delta_t^{\theta_0}) - 1/\sigma_{t+t'}^{\alpha_0}\cdot R_{-\rho_{t+t'}^{\phi_0}}(\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0})\right\|\,\mathrm{d}t' \\
\leq\ & 2^{3/2}\pi\int_{\Omega_T}\|\omega\|\left|\mathcal{F}_f(\omega)\right|\,\mathrm{d}\omega\cdot\sum_{t\in\mathbb{T}}\int_0^{\beta_T/T}\left(1/\sigma_t^{\hat{\alpha}_T}\left\|(\delta_t^\theta - \delta_t^{\theta_0}) - (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0})\right\|\right.
\end{aligned}$$

$$+ \left\| \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} - 1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} \right) (\delta_{t+t'}^{\theta} - \delta_{t+t'}^{\theta_0}) \right\| \Big) \, \mathrm{d}t', \tag{6.54}$$

where we used the fact that $\|R\delta\| = \|\delta\|$ for any rotation matrix $R \in \mathbb{R}^{2\times 2}$ and any $\delta \in \mathbb{R}^2$. Since the drift function δ^θ is of bounded variation uniformly in θ (Assumption 2.14 (B7)), there is a $C > 0$ such that

$$\begin{aligned}
&\sum_{t\in\mathbb{T}} \int_0^{\beta_T/T} 1/\sigma_t^{\hat{\alpha}_T} \left\| (\delta_t^\theta - \delta_t^{\theta_0}) - (\delta_{t+t'}^\theta - \delta_{t+t'}^{\theta_0}) \right\| \mathrm{d}t' \\
\leq\ & \sigma_{\min}^{-1} \int_0^{\beta_T/T} \sum_{t\in\mathbb{T}} \left(\left\| \delta_t^\theta - \delta_{t+t'}^\theta \right\| + \left\| \delta_t^{\theta_0} - \delta_{t+t'}^{\theta_0} \right\| \right) \mathrm{d}t' \\
\leq\ & \frac{2C\beta_T}{\sigma_{\min} T} = o(1)
\end{aligned} \tag{6.55}$$

because of $\beta_T = o(T)$ (Assumption 2.15 (C1)). Together with Assumption 2.13 (A3) and Lemma A.10, it follows that the first part of (6.54) converges to zero in probability. To see that the second part of (6.54) also vanishes, recall that the drift function δ^θ is bounded as function in t, uniformly in θ, as argued in the previous Section 6.1. Hence, with C_2 from (6.31), it holds that

$$\begin{aligned}
&\sum_{t\in\mathbb{T}} \int_0^{\beta_T/T} \left\| \left(1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} - 1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} \right) (\delta_{t+t'}^{\theta} - \delta_{t+t'}^{\theta_0}) \right\| \mathrm{d}t' \\
\leq\ & 2C_2 \sum_{t\in\mathbb{T}} \int_0^{\beta_T/T} \left\| 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} - 1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} \right\| \mathrm{d}t' \\
\leq\ & \frac{2C_2\beta_T}{T} \sum_{t\in\mathbb{T}} \left\| 1/\sigma_t^{\hat{\alpha}_T} \cdot R_{-\rho_t^{\hat{\phi}_T}} - 1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} \right\| \\
& + 2C_2 \int_0^{\beta_T/T} \sum_{t\in\mathbb{T}} \left\| 1/\sigma_t^{\alpha_0} \cdot R_{-\rho_t^{\phi_0}} - 1/\sigma_{t+t'}^{\alpha_0} \cdot R_{-\rho_{t+t'}^{\phi_0}} \right\| \mathrm{d}t' \\
=\ & \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{T}} + \frac{\beta_T}{T} \right) = o_{\mathbb{P}}(1),
\end{aligned} \tag{6.56}$$

due to the Delta method (Theorem B.7), applied to the consistent rotation and scaling parameter estimators, and the bounded total variation of the rotation and scaling functions. This means that (6.54) converges to zero in probability and hence, so does the second part of (6.53).

Next, we show that the first part of (6.53) vanishes, too. We have that

$$\begin{aligned}
&\frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)} \omega \right) - \mathcal{F}_f(\omega) \right| \\
\leq\ & \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left(\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 \left| \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)} \omega \right) - \mathcal{F}_f(\omega) \right| + \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}} \right)^2 - 1 \right| \mathcal{F}_f(\omega) \right) \\
\leq\ & \left(\frac{\max\{1, \sigma_{\max}\}}{\sigma_{\min}} \right)^2 \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left(L_f \left\| \tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)} \omega - \omega \right\| + \left| (\sigma_t^{\alpha_0})^2 - (\sigma_t^{\hat{\alpha}_T})^2 \right| \left\| \mathcal{F}_f \right\|_\infty \right) \\
=\ & \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{T}} \right),
\end{aligned} \tag{6.57}$$

where we used the Lipschitz continuity of the Fourier transform (see Lemma A.2), Lemma 6.8, and the fact that $\sqrt{T}\left((\sigma_t^{\hat{\alpha}_T})^2 - (\sigma_t^{\alpha_0})^2\right)$ is asymptotically centered normal for $T \to \infty$ by the Delta-method (Theorem B.7) and Theorem 3.5. Hence, with Assumption 2.15 (C1) it follows that

$$\int_{\Omega_T} \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) - \mathcal{F}_f(\omega)\right| \mathrm{d}\omega = O_{\mathbb{P}}\left(\frac{r_T^2}{\sqrt{T}}\right) = o_{\mathbb{P}}(1). \tag{6.58}$$

Collecting (6.53), (6.54), and (6.58), we obtain

$$\left|A_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) - N(\theta;\phi_0,\alpha_0)\right| \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability uniformly in } \theta, . \tag{6.59}$$

Now we prove that $C_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} 0$ in probability uniformly in θ. With the Cauchy-Schwarz inequality (Theorem B.1) and Lemma 6.3, we get

$$\begin{aligned}
0 &\geq \mathbb{E}\left(C_T(\theta;\hat{\phi}_T,\hat{\alpha}_T)\right) \\
&= -\mathbb{E}\left(\int_{\Omega_T}\left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{0,t}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right)^{-1} V_T^t(\omega;\hat{\phi}_T,\hat{\alpha}_T)\right|^2 \mathrm{d}\omega\right) \\
&\geq -\mathbb{E}\left(\int_{\Omega_T}\left(\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\left|h_\omega^{0,t}\left(\theta;\hat{\phi}_T,\hat{\alpha}_T\right)^{-1}\right|^2\right)\left(\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\left|V_T^t(\omega;\hat{\phi}_T,\hat{\alpha}_T)\right|^2\right)\mathrm{d}\omega\right) \\
&= -\int_{\Omega_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\mathbb{E}\left(\left|V_T^t(\omega;\hat{\phi}_T,\hat{\alpha}_T)\right|^2\right)\mathrm{d}\omega \\
&= -\int_{\Omega_T} O\left(\frac{1}{\beta_T}\right)\mathrm{d}\omega \\
&= O\left(\frac{r_T^2}{\beta_T}\right).
\end{aligned}$$

Since $r_T^2/\beta_T \overset{T\to\infty}{\longrightarrow} 0$ by Assumption 2.15 (C1), we get $\mathbb{E}\left(C_T(\theta;\hat{\phi}_T,\hat{\alpha}_T)\right) \overset{T\to\infty}{\longrightarrow} 0$ uniformly in θ, and thus,

$$C_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability uniformly in } \theta. \tag{6.60}$$

Finally, the Cauchy-Schwarz inequality (Theorem B.1) implies

$$\left(B_T(\theta;\hat{\phi}_T,\hat{\alpha}_T)\right)^2 \leq 4A_T(\theta;\hat{\phi}_T,\hat{\alpha}_T)C_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability} \tag{6.61}$$

uniformly in θ. From (6.59), (6.60), and (6.61), we get $N_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} N(\theta;\phi_0,\alpha_0)$ in probability uniformly in θ, finishing the proof of Theorem 3.2. □

6.2.3 Proof of Theorem 3.3: Consistency of the image estimator

As an implication of Theorems 3.1 and 3.2, in this subsection we obtain consistency of the final image estimator. Using the Plancherel equality (Theorem B.2) and (2.22), we have

$$
\begin{aligned}
&\left\|\hat{f}_T' - f\right\|_{L^2}^2 \\
=\ & \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} h_\omega^{0,t}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right)^{-1} Z_T^t(\omega;\hat{\phi}_T,\hat{\alpha}_T) - \mathcal{F}_f(\omega)\right|^2 \mathrm{d}\omega + \int_{\mathbb{R}^2\setminus\Omega_T} \left|\mathcal{F}_f(\omega)\right|^2 \mathrm{d}\omega \\
=\ & \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} h_\omega^{t,0}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) - \mathcal{F}_f(\omega)\right. \\
& \left. + \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} h_\omega^{0,t}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right)^{-1} V_T^t(\omega;\hat{\phi}_T,\hat{\alpha}_T)\right|^2 \mathrm{d}\omega + o(1) \\
=\ & D_T + E_T + C_T(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T) + o(1),
\end{aligned}
$$

with C_T from (2.23) and

$$
\begin{aligned}
D_T &:= \int_{\Omega_T} \left| \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} h_\omega^{t,0}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) - \mathcal{F}_f(\omega)\right|^2 \mathrm{d}\omega, \\
E_T &:= \int_{\Omega_T} 2\Re\left[\left(\frac{\beta_T}{T} \sum_{t\in\mathbb{T}} h_\omega^{t,0}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) - \mathcal{F}_f(\omega)\right)\right. \\
&\quad \left. \cdot \left(\frac{\beta_T}{T} \sum_{t'\in\mathbb{T}} \overline{h_\omega^{0,t}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right)^{-1} V_T^{t'}(\omega;\hat{\phi}_T,\hat{\alpha}_T)}\right)\right] \mathrm{d}\omega,
\end{aligned}
$$

and $C_T(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T) \le \sup_{\theta\in\Theta} C_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} 0$ in probability as shown in the proof of step 3 of Theorem 3.2. Because of (6.57), we have

$$
\begin{aligned}
D_T \le & \int_{\Omega_T} \left(\frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left| h_\omega^{t,0}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) - h_\omega^{t,0}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right)\mathcal{F}_f(\omega)\right|\right. \\
& \left. + \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left| h_\omega^{t,0}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right)\mathcal{F}_f(\omega) - \mathcal{F}_f(\omega)\right|\right)^2 \mathrm{d}\omega \\
= & \int_{\Omega_T} \left(\frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left| \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) - \mathcal{F}_f(\omega)\right|\right. \\
& \left. + \left|\mathcal{F}_f(\omega)\right| \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left| h_\omega^{t,0}\left(\hat{\theta}_T;\hat{\phi}_T,\hat{\alpha}_T\right) - 1\right| \right)^2 \mathrm{d}\omega \\
\le & \int_{\Omega_T} \left(O_{\mathbb{P}}\left(\frac{1}{\sqrt{T}}\right) + 2\sqrt{2}\pi\sigma_{\min}^{-1} \|\omega\| \left|\mathcal{F}_f(\omega)\right| \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \left\| \delta_t^{\hat{\theta}_T} - \delta_t^{\theta_0}\right\| \right)^2 \mathrm{d}\omega \\
\le & \int_{\Omega_T} \left(O_{\mathbb{P}}\left(\frac{1}{\sqrt{T}}\right) + 2\sqrt{2}\pi\sigma_{\min}^{-1} \|\omega\| \left|\mathcal{F}_f(\omega)\right| O_{\mathbb{P}}\left(L_\delta \left\|\hat{\theta}_T - \theta_0\right\|\right) \right)^2 \mathrm{d}\omega
\end{aligned}
$$

$$
\begin{aligned}
&= O_{\mathbb{P}}\left(\frac{r_T^2}{T}\right) + O_{\mathbb{P}}\left(\frac{1}{\sqrt{T}}\left\|\hat{\theta}_T - \theta_0\right\| \int_{\Omega_T} \|\omega\| \left|\mathcal{F}_f(\omega)\right| \,\mathrm{d}\omega\right) \\
&\quad + O_{\mathbb{P}}\left(\left\|\hat{\theta}_T - \theta_0\right\|^2 \int_{\Omega_T} \|\omega\|^2 \left|\mathcal{F}_f(\omega)\right|^2 \,\mathrm{d}\omega\right),
\end{aligned}
$$

where we used that $x \mapsto \mathrm{e}^{\mathrm{i}\langle a,x\rangle}$, $a \in \mathbb{R}^2$, is Lipschitz-continuous with Lipschitz-constant $\sqrt{2}\,\|a\|$ and that $\|R\omega\| = \|\omega\|$ for any Rotation matrix R for the second inequality, as well as Assumption 2.14 (B6) for the third. Since $r_T^2/T \overset{T\to\infty}{\longrightarrow} 0$ by Assumption 2.15 (C1), it follows by the consistency of the drift parameter estimator (Theorem 3.2) and Assumption 2.13 (A3) together with Lemma A.10, that $D_T \overset{T\to\infty}{\longrightarrow} 0$ in probability.
Finally, by the Cauchy-Schwarz inequality (Theorem B.1),

$$E_T^2 \leq 4 D_T C_T(\hat{\theta}_T; \hat{\phi}_T, \hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} 0$$

in probability. Since $\hat{f}_T$ is defined as the squared value of $\hat{f}_T'$, the continuous mapping theorem (Theorem B.8) immediately gives consistency also for the final image estimator $\hat{f}_T$, completing the proof of (3.3).

6.2.4 Proof of Theorem 3.5: Central limit theorem for the rotation and scaling parameter estimators

We first prove three results on the derivatives of the empirical contrast functional M_T in Theorem 6.10, Lemma 6.11, and Theorem 6.12, and then give the detailed proof of the central limit theorem for rotation and scaling, Theorem 3.5. The first of the mentioned results is a central limit theorem for $\operatorname{grad}_{(\phi,\alpha)} M_T(\phi_0, \alpha_0)$, formulated in the next theorem. Prior to it, we introduce some notation. The second result then describes properties of the Hessian of the population contrast functional M, and the last shows the convergence of the Hessian of the empirical contrast functional M_T to the Hessian of M. Combining all three we then conclude the proof of Theorem 3.5. The main structure is analogous to the proof of asymptotic normality of the drift parameter estimator in Hartmann et al. (2015) and of asymptotic normality of the rotation and scaling parameter estimators in Hartmann (2016). Here, the argument is modified to fit the revised model and the details are worked out in a mathematically rigorous way.

Definition 6.9. *With $d_{u,v}^{t,t'}(\phi, \alpha)$ from (6.2) let*

$$G_j^{t'} := \Im\left[\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\overline{\mathcal{M}_{q_j^t}(u,v)}\boldsymbol{a}_{u,v}^{t,t'}(\phi_0,\alpha_0) d_{u,v}^{0,t'}(\phi_0,\alpha_0)\,\mathrm{d}v\,\mathrm{d}t\right],$$

where $q_j^t(\omega) := \Re\left(\mathrm{e}^{2\pi\mathrm{i}\langle\omega,x_j\rangle}\mathcal{F}_{f^t}(\omega)\right)$. We define

$$\Sigma_{RS} := 4\int_0^1 \frac{1}{n}\sum_{j=1}^n G_j^{t'}(G_j^{t'})^\top\,\mathrm{d}t'. \tag{6.62}$$

Theorem 6.10 (Central limit theorem for $\operatorname{grad}_{(\phi,\alpha)} M_T(\phi_0,\alpha_0)$)**.** *Under the Assumptions 2.13 (A4), 2.14 (B2), and 2.15, we have that*

$$\sqrt{T}\operatorname{grad}_{(\phi,\alpha)} M_T(\phi_0,\alpha_0) = \sum_{t'\in\mathbb{T}}\sum_{j=1}^{n} w_j^{t'}\epsilon_j^{t'} + o_{\mathbb{P}}(1), \tag{6.63}$$

with weights

$$w_j^{t'} := -\frac{2\beta_T^{3/2}}{nT^{3/2}}\sum_{t\in\mathbb{T}}\sum_{|u|\le u_T}\int_{-v_T}^{v_T}\Re\left[\operatorname{grad}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha)\Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)} F^t(u,v)\overline{\mathcal{M}^T_{q_j^{t'}}(u,v)}\right]\mathrm{d}v,$$

$t'\in\mathbb{T}$, $1\le j\le n$, with $q_j^t(\omega) := \Re\left(\mathrm{e}^{2\pi\mathrm{i}\langle\omega,x_j\rangle}\mathcal{F}_{f^t}(\omega)\right)$ as in Definition 6.9. In particular,

$$\sqrt{T}\operatorname{grad}_{(\phi,\alpha)} M_T(\phi_0,\alpha_0) \xrightarrow{T\to\infty} \mathcal{N}(0,\Sigma_{RS}) \quad \textit{in distribution,}$$

with covariance matrix Σ_{RS} from (6.62).

Proof. We will use the decomposition $M_T = A_T + B_T + C_T + o(1)$ from (6.36) and show that $\sqrt{T}\operatorname{grad}_{(\phi,\alpha)} A_T(\phi_0,\alpha_0) = 0$, $\sqrt{T}\operatorname{grad}_{(\phi,\alpha)} C_T(\phi_0,\alpha_0) \xrightarrow{T\to\infty} 0$ in probability, while

$$\sqrt{T}\operatorname{grad}_{(\phi,\alpha)} B_T(\phi_0,\alpha_0) = \sum_{t\in\mathbb{T}}\sum_{j\in J_t^t} w_j^t\epsilon_j^t + o_{\mathbb{P}}(1).$$

Since the error terms ϵ_j^t are mutually independent it is then easy to prove the asymptotic normality

$$\sqrt{T}\operatorname{grad}_{(\phi,\alpha)} B_T(\phi_0,\alpha_0) \xrightarrow{T\to\infty} \mathcal{N}(0,\Sigma_{RS}) \quad \text{in distribution.}$$

First, consider the gradient of the integrand of A_T. By Lemma A.5 and Lemma 6.4, we get

$$\begin{aligned}
&\operatorname{grad}_{(\phi,\alpha)}\left(-\sum_{t,t'\in\mathbb{T}} d_{u,v}^{t,0}(\phi,\alpha)F^t(u,v)\overline{d_{u,v}^{t',0}(\phi,\alpha)F^{t'}(u,v)}\right)\\
=&-2\sum_{t,t'\in\mathbb{T}}\Re\left(\operatorname{grad}_{(\phi,\alpha)} d_{u,v}^{t,0}(\phi,\alpha)F^t(u,v)\overline{d_{u,v}^{t',0}(\phi,\alpha)F^{t'}(u,v)}\right)\\
=&2\sum_{t,t'\in\mathbb{T}}\left(2\pi u\,(\operatorname{grad}_\phi\rho_t^\phi)^\top, -v\sigma_t^{\alpha_0}/\sigma_t^\alpha(\operatorname{grad}_\alpha\sigma_t^\alpha)^\top\right)^\top\\
&\cdot\Im\left(d_{u,v}^{t,0}(\phi,\alpha)\overline{d_{u,v}^{t',0}(\phi,\alpha)}F^t(u,v)\overline{F^{t'}(u,v)}\right),
\end{aligned}\tag{6.64}$$

Because of $d_{u,v}^{t,0}(\alpha_0,\phi_0) = 1$ and

$$\Im\left(F^t(u,v)\overline{F^{t'}(u,v)}\right) = (\sigma_t^{\alpha_0})^{4-\gamma}(\sigma_{t'}^{\alpha_0})^{4-\gamma}\Im\left(\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\right) = 0,$$

(6.64) vanishes for $(\phi,\alpha) = (\phi_0,\alpha_0)$, implying that

$$\sqrt{T}\operatorname{grad}_{(\phi,\alpha)} A_T(\phi_0,\alpha_0) = 0. \tag{6.65}$$

Next, we consider the asymptotic behaviour of $\sqrt{T}\mathrm{grad}_{(\phi,\alpha)}C_T(\phi_0,\alpha_0)$. By Lemma A.5 and Lemma 6.4, we get for parameters in the neighborhood U from Assumption 2.14 (B2) around the true parameters, that

$$
\begin{aligned}
& \mathrm{grad}_{(\phi,\alpha)}\left(-\sum_{t,t'\in\mathbb{T}} d_{u,v}(\sigma_t^\alpha,\rho_t^\phi)\mathcal{M}^T_{\mathcal{W}^t}(u,v)\overline{d_{u,v}(\sigma_{t'}^\alpha,\rho_{t'}^\phi)\mathcal{M}^T_{\mathcal{W}^{t'}}(u,v)}\right)\\
= \; & -2\sum_{t,t'\in\mathbb{T}}\Re\left(\mathrm{grad}_{(\phi,\alpha)}d_{u,v}(\sigma_t^\alpha,\rho_t^\phi)\mathcal{M}^T_{\mathcal{W}^t}(u,v)\overline{d_{u,v}(\sigma_{t'}^\alpha,\rho_{t'}^\phi)\mathcal{M}^T_{\mathcal{W}^{t'}}(u,v)}\right)\\
= \; & 2\sum_{t\in\mathbb{T}}\sum_{t'\in\mathbb{T}\setminus\{t\}}\left(2\pi u(\mathrm{grad}_\phi\rho_t^\phi)^\top, -v/\sigma_t^\alpha\cdot(\mathrm{grad}_\alpha\sigma_t^\alpha)^\top\right)^\top\\
& \cdot\Im\left(d_{u,v}(\sigma_t^\alpha,\rho_t^\phi)\mathcal{M}^T_{\mathcal{W}^t}(u,v)\overline{d_{u,v}(\sigma_{t'}^\alpha,\rho_{t'}^\phi)\mathcal{M}^T_{\mathcal{W}^{t'}}(u,v)}\right),
\end{aligned}
$$

where the terms with $t' = t$ vanish due to

$$
\Im\left(\left|d_{u,v}(\sigma_t^\alpha,\rho_t^\phi)\mathcal{M}^T_{\mathcal{W}^t}(u,v)\right|^2\right) = 0.
$$

With Lemma 6.4, it follows that

$$
\begin{aligned}
& \mathbb{E}\left\|\sqrt{T}\mathrm{grad}_{(\phi,\alpha)}C_T(\phi_0,\alpha_0)\right\|\\
= \; & \mathbb{E}\Bigg\|\sqrt{T}\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\frac{2\beta_T^2}{T^2}\sum_{t\in\mathbb{T}}\sum_{t'\in\mathbb{T}\setminus\{t\}}\left(2\pi u(\mathrm{grad}_\phi\rho_t^\phi)^\top, -v/\sigma_t^\alpha\cdot(\mathrm{grad}_\alpha\sigma_t^\alpha)^\top\right)^\top\\
& \cdot\Im\left(d_{u,v}(\sigma_t^\alpha,\rho_t^\phi)\mathcal{M}^T_{\mathcal{W}^t}(u,v)\overline{d_{u,v}(\sigma_{t'}^\alpha,\rho_{t'}^\phi)\mathcal{M}^T_{\mathcal{W}^{t'}}(u,v)}\right)\,\mathrm{d}v\Bigg\|\\
\le \; & 2\widetilde{C}\sqrt{T}\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\|(u,v)\|\frac{\beta_T^2}{T^2}\sum_{t\in\mathbb{T}}\sum_{t'\in\mathbb{T}\setminus\{t\}}\mathbb{E}\left|\mathcal{M}^T_{\mathcal{W}^t}(u,v)\right|\mathbb{E}\left|\mathcal{M}^T_{\mathcal{W}^{t'}}(u,v)\right|\,\mathrm{d}v\\
\le \; & 2\widetilde{C}\sqrt{T}\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\|(u_T,v_T)\|\frac{\beta_T^2}{T^2}\sum_{t\in\mathbb{T}}\sum_{t'\in\mathbb{T}\setminus\{t\}}O\left(\frac{r_T^{2\gamma}}{\beta_T}\right)\,\mathrm{d}v\\
= \; & O\left(\frac{\sqrt{T}u_Tv_T\,\|(u_T,v_T)\|\,r_T^{2\gamma}}{\beta_T}\right),
\end{aligned}
$$

where we used that, due to Lemma 6.2,

$$
\mathbb{E}\left|\mathcal{M}^T_{\mathcal{W}^t}(u,v)\right| \le \sqrt{\mathbb{E}\left(\left|\mathcal{M}^T_{\mathcal{W}^t}(u,v)\right|^2\right)} = O\left(\frac{r_T^\gamma}{\sqrt{\beta_T}}\right).
$$

Since $\sqrt{T}u_Tv_T\,\|(u_T,v_T)\|\,r_T^{2\gamma}/\beta_T \stackrel{T\to\infty}{\longrightarrow} 0$ by Assumption 2.15 (C2), we have

$$
\sqrt{T}\mathrm{grad}_{(\phi,\alpha)}C_T(\phi_0,\alpha_0) \stackrel{T\to\infty}{\longrightarrow} 0 \quad \text{in probability.} \tag{6.66}
$$

Finally, we tackle $\sqrt{T}\mathrm{grad}_{(\phi,\alpha)}B_T(\phi_0,\alpha_0)$. We write $B_T = B_T^{(1)} + B_T^{(2)}$ with

$$B_T^{(1)}(\phi,\alpha) := -\int_{-v_T}^{v_T} \sum_{|u|\le u_T} 2\Re\left[\left(\frac{\beta_T}{T} \sum_{t\in\mathbb{T}} d_{u,v}^{t,0}(\phi,\alpha) F^t(u,v) \right) \cdot \left(\frac{\beta_T}{T} \sum_{t'\in\mathbb{T}} d_{u,v}^{0,t'}(\phi,\alpha) \overline{\mathcal{M}^T_{|W^{t'}|^2}(u,v)} \right) \right] \mathrm{d}v,$$

$$B_T^{(2)}(\phi,\alpha) := -\int_{-v_T}^{v_T} \sum_{|u|\le u_T} 2\Re\left[\left(\frac{\beta_T}{T} \sum_{t\in\mathbb{T}} d_{u,v}^{t,0}(\phi,\alpha) F^t(u,v) \right) \cdot \left(\frac{\beta_T}{T} \sum_{t'\in\mathbb{T}} d_{u,v}^{0,t'}(\phi,\alpha) \overline{\mathcal{M}^T_{2\Re\left(\mathcal{F}_{f^{t'}}\overline{W^{t'}}\right)}(u,v)} \right) \right] \mathrm{d}v.$$

With Lemma 6.1, we have

$$\begin{aligned}
\mathbb{E}\left|\mathcal{M}^T_{|W^{t}|^2}(u,v)\right| &\le \int_{\Omega_T} \|\omega\|^\gamma\, \mathbb{E}\left(\left|W^t(\omega)\right|^2\right) \mathrm{d}\omega \\
&= \int_{\Omega_T} \|\omega\|^\gamma \frac{1}{4n\beta_T} \mathrm{d}\omega \\
&= O\left(\frac{r_T^{2+\gamma}}{\beta_T}\right).
\end{aligned}$$

Hence, with Lemma 6.4,

$$\begin{aligned}
& \mathbb{E}\left\| \sqrt{T}\mathrm{grad}_{\phi,\alpha} B_T^{(1)}(\phi_0,\alpha_0)\right\| \\
= \;& \mathbb{E}\Bigg\| 2\sqrt{T} \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \frac{\beta_T}{T} \sum_{t'\in\mathbb{T}} \\
& \Re\left[\mathrm{grad}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha)\Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)} F^t(u,v) \overline{\mathcal{M}^T_{|W^{t'}|^2}(u,v)} \right] \mathrm{d}v \Bigg\| \\
\le \;& 2\widetilde{C}\sqrt{T} \frac{\beta_T}{T} \sum_{t\in\mathbb{T}} \frac{\beta_T}{T} \sum_{t'\in\mathbb{T}} \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \|(u,v)\| \left|F^t(u,v)\right| \mathbb{E}\left|\mathcal{M}^T_{|W^{t'}|^2}(u,v)\right| \mathrm{d}v \\
= \;& O\left(\frac{\sqrt{T} r_T^{2+\gamma}}{\beta_T} \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \|(u,v)\| \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right| \mathrm{d}v \right) \\
= \;& O\left(\frac{\sqrt{T} r_T^{2+\gamma}}{\beta_T} \right),
\end{aligned}$$

due to Assumption 2.13 (A4) and Lemma A.11. Since $\sqrt{T} r_T^{2+\gamma}/\beta_T \overset{T\to\infty}{\longrightarrow} 0$ by Assumption 2.15 (C2), we get using Markov's inequality (Theorem B.10) that

$$\sqrt{T}\mathrm{grad}_{\phi,\alpha} B_T^{(1)}(\phi_0,\alpha_0) \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability.} \tag{6.67}$$

It remains to show the asymptotic normality of $\sqrt{T}\operatorname{grad}_{\phi,\alpha} B_T^{(2)}(\phi_0,\alpha_0)$. We have

$$
\begin{aligned}
&\mathcal{M}^T_{2\Re\left(\mathcal{F}_{f^t}\overline{W^t}\right)}(u,v)\\
=\ & 2\int_0^{r_T}\int_0^{2\pi} \mathrm{e}^{-2\pi \mathrm{i} u\psi} r^{-\mathrm{i}v} r^{\gamma}\Big(\Re\left(\mathcal{F}_{f^t}\overline{W^t}\right)\circ\mathcal{P}\Big)(r,\psi)\,\mathrm{d}\psi\,\frac{\mathrm{d}r}{r}\\
=\ & 2\int_0^{r_T}\int_0^{2\pi} \mathrm{e}^{-2\pi \mathrm{i} u\psi} r^{-\mathrm{i}v} r^{\gamma}\Re\left[\Big(\mathcal{F}_{f^t}\circ\mathcal{P}\Big)(r,\psi)\left(\frac{1}{2n\sqrt{\beta_T}}\sum_{j=1}^n \mathrm{e}^{2\pi\mathrm{i}\langle\mathcal{P}(r,\psi),x_j\rangle}\epsilon_j^t\right)\right]\mathrm{d}\psi\,\frac{\mathrm{d}r}{r}\\
=\ & \frac{1}{n\sqrt{\beta_T}}\sum_{j=1}^n \epsilon_j^t\int_0^{r_T}\int_0^{2\pi} \mathrm{e}^{-2\pi \mathrm{i} u\psi} r^{-\mathrm{i}v} r^{\gamma}\Re\left[\mathrm{e}^{2\pi\mathrm{i}\langle\mathcal{P}(r,\psi),x_j\rangle}\Big(\mathcal{F}_{f^t}\circ\mathcal{P}\Big)(r,\psi)\right]\mathrm{d}\psi\,\frac{\mathrm{d}r}{r}\\
=\ & \frac{1}{n\sqrt{\beta_T}}\sum_{j=1}^n \mathcal{M}^T_{q_j^t}(u,v)\epsilon_j^t,
\end{aligned}
$$

with $q_j^t(\omega) \coloneqq \Re\left(\mathrm{e}^{2\pi\mathrm{i}\langle\omega,x_j\rangle}\mathcal{F}_{f^t}(\omega)\right)$ from Definition 6.9. Hence,

$$
\begin{aligned}
B_T^{(2)}(\phi,\alpha) &= -\frac{2\beta_T}{T}\sum_{t'\in\mathbb{T}}\Re\left[\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\int_{-v_T}^{v_T}\sum_{|u|\le u_T} d_{u,v}^{t,t'}(\phi,\alpha)F^t(u,v)\overline{\mathcal{M}^T_{2\Re\left(\mathcal{F}_{f^{t'}}\overline{W^{t'}}\right)}(u,v)}\right]\mathrm{d}v\\
&= -\frac{2\sqrt{\beta_T}}{T}\sum_{t'\in\mathbb{T}}\frac{1}{n}\sum_{j=1}^n\epsilon_j^{t'}\Re\left[\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\int_{-v_T}^{v_T}\sum_{|u|\le u_T} d_{u,v}^{t,t'}(\phi,\alpha)F^t(u,v)\overline{\mathcal{M}^T_{q_j^{t'}}(u,v)}\,\mathrm{d}v\right].
\end{aligned}
\tag{6.68}
$$

Now, let $\xi\in\mathbb{R}^{d_2+d_3}$. From (6.5) and (6.68), we get that at $(\phi,\alpha)=(\phi_0,\alpha_0)$

$$
\begin{aligned}
&\left\langle \xi,\ \sqrt{T}\operatorname{grad}_{(\phi,\alpha)}B_T^{(2)}(\phi_0,\alpha_0)\right\rangle\\
=\ & -2\sqrt{\frac{\beta_T}{T}}\sum_{t'\in\mathbb{T}}\frac{1}{n}\sum_{j=1}^n\epsilon_j^{t'}\Re\left[\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\right.\\
&\left.\cdot\left\langle\xi,\operatorname{grad}_{(\phi,\alpha)}d_{u,v}^{t,t'}(\phi,\alpha)\Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)}\right\rangle F^t(u,v)\overline{\mathcal{M}^T_{q_j^{t'}}(u,v)}\,\mathrm{d}v\right]\\
=\ & 2\sqrt{\frac{\beta_T}{T}}\sum_{t'\in\mathbb{T}}\frac{1}{n}\sum_{j=1}^n\epsilon_j^{t'}\Im\left[\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\right.\\
&\left.\cdot\left\langle\xi,\boldsymbol{a}_{u,v}^{t,t'}(\phi_0,\alpha_0)d_{u,v}^{0,t'}(\phi_0,\alpha_0)\right\rangle F^t(u,v)\overline{\mathcal{M}^T_{q_j^{t'}}(u,v)}\,\mathrm{d}v\right]
\end{aligned}
$$

is a linear combination of independent standard-normal random variables $\epsilon_j^{t'}$ and therefore a centered Gaussian random variable with variance

$$
\begin{aligned}
&\frac{4\beta_T}{T}\sum_{t'\in\mathbb{T}}\frac{1}{n}\sum_{j=1}^n\Im\left[\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\int_{-v_T}^{v_T}\sum_{|u|\le u_T}F^t(u,v)\overline{\mathcal{M}^T_{q_j^{t'}}(u,v)}\right.\\
&\qquad\qquad\left.\cdot\left\langle\xi,\boldsymbol{a}_{u,v}^{t,t'}(\phi_0,\alpha_0)d_{u,v}^{0,t'}(\phi_0,\alpha_0)\right\rangle\,\mathrm{d}v\right]^2
\end{aligned}
$$

$$\begin{aligned}
&\overset{T\to\infty}{\longrightarrow} \quad 4\int_0^1 \frac{1}{n}\sum_{j=1}^{n} \Im\Bigg[\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma}\int_{\mathbb{R}}\sum_{u\in\mathbb{Z}} \mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\overline{\mathcal{M}_{q_j^{t'}}(u,v)} \\
&\qquad\qquad \cdot\Big\langle \xi, \boldsymbol{a}_{u,v}^{t,t'}(\phi_0,\alpha_0) d_{u,v}^{0,t'}(\phi_0,\alpha_0)\Big\rangle\,\mathrm{d}v\,\mathrm{d}t\Bigg]^2 \mathrm{d}t' \\
&= \quad \xi^{\top}\Sigma_{RS}\,\xi, \qquad (6.69)
\end{aligned}$$

with Σ_{RS} from (6.62). Note that Σ_{RS} has finite operatornorm, since by Lemma 6.4, the Cauchy-Schwarz-inequality (Theorem B.1), and the Parseval equation (2.9),

$$\begin{aligned}
&\left|\int_{\mathbb{R}}\sum_{u\in\mathbb{Z}} \mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\overline{\mathcal{M}_{q_j^{t'}}(u,v)}\boldsymbol{a}_{u,v}^{t,t'}(\phi_0,\alpha_0) d_{u,v}^{0,t'}(\phi_0,\alpha_0)\,\mathrm{d}v\right| \\
&\le C\int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|\left|\mathcal{M}_{q_j^{t'}}(u,v)\right| \|(u,v)\|\,\mathrm{d}v \\
&\le C\int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \|(u,v)\|^2\,\mathrm{d}v \int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\mathcal{M}_{q_j^{t'}}(u,v)\right|^2\mathrm{d}v \\
&= \int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \|(u,v)\|^2\,\mathrm{d}v \int_{\mathbb{R}^2}\left|q_j^{t'}(\omega)\right|^2\mathrm{d}\omega \\
&\le \int_{\mathbb{R}}\sum_{u\in\mathbb{Z}}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \|(u,v)\|^2\,\mathrm{d}v \int_{\mathbb{R}^2}\left|\mathcal{F}_f(\omega)\right|^2\mathrm{d}\omega < \infty,
\end{aligned}$$

where the first integral is finite because of Assumption 2.13 (A4), and $\mathcal{F}_f$ is square integrable by the Plancherel Theorem (Theorem B.2) using that $f \in L^2(\mathbb{R}^2)$. Because the upper bound is independent of t', the norm of Σ_{RS} is bounded, as well. By Lemma A.13 and the Cramér-Wold Device (Theorem B.5), we have

$$\sqrt{T}\,\mathrm{grad}_{(\phi,\alpha)} B_T^{(2)}(\phi_0,\alpha_0) \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0,\Sigma_{RS}) \quad \text{in distribution} \qquad (6.70)$$

as an application of Theorem 2.13 in van der Vaart (2000). From (6.65), (6.66), (6.67), (6.70), and Slutzky's Lemma (Theorem B.9), we deduce

$$\sqrt{T}\,\mathrm{grad}_{(\phi,\alpha)} M_T(\phi_0,\alpha_0) = \sqrt{T}\,\mathrm{grad}_{(\phi,\alpha)} B_T^{(2)}(\phi_0,\alpha_0) + o_{\mathbb{P}}(1) \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0,\Sigma_{RS})$$

in distribution, completing the proof. □

Lemma 6.11. *Under the Assumptions 2.13 (A4) and 2.14 (B2-B3),* $\mathrm{Hess}_{(\phi,\alpha)} M(\phi,\alpha)$ *has finite operator norm for all* $(\phi,\alpha) \in U$ *with* $U \subseteq \Phi \times \mathrm{A}$ *from Assumption 2.14 (B2). Furthermore, the matrix*

$$H_M := \mathrm{Hess}_{(\phi,\alpha)} M(\phi_0,\alpha_0) \qquad (6.71)$$

is symmetric. If the Assumptions 2.13 (A2, A3) and 2.14 (B1, B4, and B8) hold, H_M *is also positive definite and hence, invertible.*

Proof. By Assumptions 2.14 (B2-B3), Lemma 6.4, Lemma A.6, and Theorem B.4 on the differentiability of parameter integrals,

$$
\begin{aligned}
&\mathrm{Hess}_{(\phi,\alpha)} M(\phi,\alpha)\\
=&-\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \mathrm{Hess}_{(\phi,\alpha)}\left(\int_0^1\int_0^1 d_{u,v}^{t,0}(\phi,\alpha) F^t(u,v)\overline{d_{u,v}^{t',0}(\phi,\alpha)F^{t'}(u,v)}\,\mathrm{d}t\,\mathrm{d}t'\right)\mathrm{d}v\\
=&-2\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \int_0^1\int_0^1 (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma}\\
&\cdot\Re\left(\overline{d_{u,v}^{t',0}(\phi,\alpha)}\,\mathrm{Hess}_{(\phi,\alpha)} d_{u,v}^{t,0}(\phi,\alpha) + \mathrm{grad}_{(\phi,\alpha)} d_{u,v}^{t,0}(\phi,\alpha)\overline{\mathrm{grad}_{(\phi,\alpha)}^\top d_{u,v}^{t',0}(\phi,\alpha)}\right)\mathrm{d}t\,\mathrm{d}t'\,\mathrm{d}v\\
=&-2\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \int_0^1\int_0^1 (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma}\Re\left(d_{u,v}^{t,0}(\phi,\alpha)\overline{d_{u,v}^{t',0}(\phi,\alpha)}\right.\\
&\left.\cdot\left[\mathrm{i}\boldsymbol{H}_{u,v}^{t,0}(\phi,\alpha) - \boldsymbol{a}_{u,v}^{t,0}(\phi,\alpha)\left(\boldsymbol{a}_{u,v}^{t,0}(\phi,\alpha)-\boldsymbol{a}_{u,v}^{t',0}(\phi,\alpha)\right)^\top\right]\right)\mathrm{d}t\,\mathrm{d}t'\,\mathrm{d}v. \qquad (6.72)
\end{aligned}
$$

Let $\xi \in \mathbb{R}^{d_2+d_3}$ with $\|\xi\| = 1$. By Assumption 2.13 (A4) and the Sobolev embedding theorem (Theorem B.11), we have

$$
\begin{aligned}
&\left\|\mathrm{Hess}_{(\phi,\alpha)} M(\phi,\alpha)\xi\right\|\\
\le\ & 2\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \int_0^1\int_0^1 (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma}\\
&\cdot\left[\left\|\boldsymbol{H}_{u,v}^{t,0}(\phi,\alpha)\xi\right\| + \left\|\boldsymbol{a}_{u,v}^{t,0}(\phi,\alpha)\left(\boldsymbol{a}_{u,v}^{t,0}(\phi,\alpha)-\boldsymbol{a}_{u,v}^{t',0}(\phi,\alpha)\right)^\top\xi\right\|\right]\mathrm{d}v\\
\le\ & 2\max\left\{\sigma_{\max}^{4-\gamma},\sigma_{\min}^{4-\gamma}\right\}\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\left[\widetilde{C}\,\|(u,v)\| + 2\widetilde{C}^2\,\|(u,v)\|^2\right]\mathrm{d}v < \infty,
\end{aligned}
$$

with $\widetilde{C} > 0$ from Lemma 6.4. Hence, $\mathrm{Hess}_{(\phi,\alpha)} M(\phi,\alpha)$ has finite operator norm. From (6.72), we get using $d_{u,v}^{t,0}(\alpha_0,\phi_0) = 1$, that at $(\phi,\alpha) = (\phi_0,\alpha_0)$

$$
\begin{aligned}
H_M &= \mathrm{Hess}_{(\phi,\alpha)} M(\phi_0,\alpha_0)\\
&= 2\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\left[\int_0^1\int_0^1 (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma}\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)^\top\,\mathrm{d}t\,\mathrm{d}t'\right.\\
&\quad\left.-\left(\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma}\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\,\mathrm{d}t\right)\left(\int_0^1 (\sigma_{t'}^{\alpha_0})^{4-\gamma}\boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\,\mathrm{d}t'\right)^\top\right]\mathrm{d}v. \qquad (6.73)
\end{aligned}
$$

Since matrices of the form $xx^\top$ with $x \in \mathbb{R}^{d_2+d_3}$ are always symmetric, it follows that H_M is symmetric.

Now, let $\xi^{(2)} \in \mathbb{R}^{d_2}$ and $\xi^{(3)} \in \mathbb{R}^{d_3}$ such that $\xi := \left((\xi^{(2)})^\top, (\xi^{(3)})^\top\right)^\top \neq 0$. By Assumptions 2.13 (A2, A3) and Lemma A.7, there are $u \in \mathbb{Z}\setminus\{0\}$ and an open Borel set $B \subseteq \mathbb{R}$ with positive Lebesgue-measure such that $\mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \neq 0$ for all $v \in B$. The goal is now to show that if Assumptions 2.14 (B4, B8) hold, there is another Borel set $B' \subseteq [0,1]$ with positive

Lebesgue-measure such that

$$B' \to \mathbb{R}, \quad t \mapsto \left\langle \xi, \boldsymbol{a}^{t,0}_{u,v}(\phi_0, \alpha_0) \right\rangle \tag{6.74}$$

is not constant for all $v \in \mathbb{R} \setminus V^0$ with some Lebesgue null-set $V^0 \subseteq \mathbb{R}$. To this end, define $S_1(t) := \sum_{m=1}^{d_2} \xi^{(2)}_m \frac{\partial \rho^\phi_t}{\partial \phi_m}\Big|_{\phi=\phi_0}$ and $S_2(t) := \sum_{m'=1}^{d_3} \xi^{(3)}_{m'} \frac{\partial \sigma^\alpha_t}{\partial \alpha_{m'}}\Big|_{\alpha=\alpha_0}$. Note that if there was a constant c such that $S_1(t) = c$ a.e. this constant would have to be $c = 0$, since by Assumption 2.14 (B1)

$$S_1(0) = \sum_{m=1}^{d_2} \xi^{(2)}_m \frac{\partial \rho^\phi_0}{\partial \phi_m}\Big|_{\phi=\phi_0} = 0,$$

and S_1 is continuous at $t = 0$ as linear function of $t \mapsto \rho^\phi_t$, which is continuous at $t = 0$ by Assumption 2.14 (B4). By the same argument, applied to $t \mapsto \sigma^\alpha_t$, we know that S_2 has to be zero if it is constant. This property carries over to the scalar product $t \mapsto \left\langle \xi, \boldsymbol{a}^{t,0}_{u,v}(\phi_0, \alpha_0) \right\rangle$. Suppose now, it were constant (that is, equal to zero):

$$\left\langle \xi, \boldsymbol{a}^{t,0}_{u,v}(\phi_0, \alpha_0) \right\rangle = 2\pi u S_1(t) - v S_2(t) = 0 \text{ a.e.}, \tag{6.75}$$

which is equivalent to $S_1(t) = (2\pi u)^{-1} v S_2(t)$ a.e. For this to hold true, either both $S_1 = 0$ a.e. and $S_2 = 0$ a.e. or the value of v is determined by u, implying $v \in V_0 \subset \mathbb{R}$ for some Lebesgue null-set V_0. However, by Assumption 2.14 (B8) the components of $\operatorname{grad}_\phi \rho^\phi_t\big|_{\phi=\phi_0}$ are linearly independent, which is a contradiction to $S_1(t) = 0$ a.e. By the same argument, applied to $\operatorname{grad}_\alpha \sigma^\alpha_t\big|_{\alpha=\alpha_0}$, we get that S_2 cannot be constant a.e. We can now conclude that (6.75) can only be satisfied for $v \in V_0$, where V_0 has Lebesgue measure zero. Hence, there is a Borel set B' of positive Lebesgue measure such that $t \mapsto \left\langle \xi, \boldsymbol{a}^{t,0}_{u,v}(\phi_0, \alpha_0) \right\rangle$ is non-constant on B' for almost all $v \in \mathbb{R}$.

From the Cauchy-Schwarz inequality (Theorem B.1), we have that

$$\left(\int_0^1 g_1(t) g_2(t)\, \mathrm{d}t \right)^2 \le \int_0^1 g_1(t)^2\, \mathrm{d}t \int_0^1 g_2(t')^2\, \mathrm{d}t'$$

for all integrable functions $g_1, g_2 \colon [0,1] \to \mathbb{R}$, with equality if and only if g_1 and g_2 are linearly dependent a.e. Let

$$g^{u,v}_1(t) := (\sigma^{\alpha_0}_t)^{2-\gamma/2} \left\langle \xi, \boldsymbol{a}^{t,0}_{u,v}(\phi_0, \alpha_0) \right\rangle, \quad g^{u,v}_2(t) := (\sigma^{\alpha_0}_t)^{2-\gamma/2}.$$

For all $v \in \mathbb{R} \setminus V^0$, these are linearly independent, since $t \mapsto \left\langle \xi, \boldsymbol{a}^{t,0}_{u,v}(\phi_0, \alpha_0) \right\rangle$ is not constant. Hence,

$$\left(\int_0^1 (\sigma^{\alpha_0}_t)^{4-\gamma} \left\langle \xi, \boldsymbol{a}^{t,0}_{u,v}(\phi_0, \alpha_0) \right\rangle \mathrm{d}t \right)^2 < \int_0^1 (\sigma^{\alpha_0}_t)^{4-\gamma} \left\langle \xi, \boldsymbol{a}^{t,0}_{u,v}(\phi_0, \alpha_0) \right\rangle^2 \mathrm{d}t \int_0^1 (\sigma^{\alpha_0}_{t'})^{4-\gamma}\, \mathrm{d}t'$$

for all $v \in \mathbb{R} \setminus V^0$. It follows that

$$\begin{aligned}\xi^\top H_M \xi \;\geq\; & 2 \int_{\mathbb{R}\setminus V^0} \left| \mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \right|^2 \left[\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \left\langle \xi, \boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0) \right\rangle^2 \,\mathrm{d}t \int_0^1 (\sigma_{t'}^{\alpha_0})^{4-\gamma} \,\mathrm{d}t' \right. \\ & \left. - \left(\int_0^1 (\sigma_t^{\alpha_0})^{4-\gamma} \left\langle \xi, \boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0) \right\rangle \,\mathrm{d}t \right)^2 \right] \mathrm{d}v \\ >\; & 0, \end{aligned} \tag{6.76}$$

as the integrand (as a function in v) is strictly positive. We conclude that H_M is symmetric and positive definite and thus, invertible. □

Theorem 6.12. *Under the Assumptions 2.13 (A4), 2.14 (B2-B4), and 2.15 let $(\hat{\phi}_T^*, \hat{\alpha}_T^*)_{T\in\mathbb{N}}$ a sequence of random vectors with values in U, such that $(\hat{\phi}_T^*, \hat{\alpha}_T^*) \xrightarrow{T\to\infty} (\phi_0, \alpha_0)$ in probability. Then, with H_M from (6.71),*

$$\left\| \operatorname{Hess}_{(\phi,\alpha)} M_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 \xrightarrow{T\to\infty} 0 \quad \textit{in probability,}$$

Proof. We will use the decomposition $M_T = A_T + B_T + C_T + o(1)$ from (6.36) and show that

$$\begin{aligned} \left\| \operatorname{Hess}_{(\phi,\alpha)} A_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 &\xrightarrow{T\to\infty} 0 \quad \text{in probability,} \\ \sup_{(\phi,\alpha)\in U} \left\| \operatorname{Hess}_{(\phi,\alpha)} B_T(\phi,\alpha) + \operatorname{Hess}_{(\phi,\alpha)} C_T(\phi,\alpha) \right\|_1 &\xrightarrow{T\to\infty} 0 \quad \text{in probability.} \end{aligned}$$

By Lemma 6.4 and Lemma A.5, we get

$$\begin{aligned} & \operatorname{Hess}_{(\phi,\alpha)} A_T(\phi,\alpha) \\ &= - \int_{-v_T}^{v_T} \sum_{|u|\leq u_T} \frac{2\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} \Re\left[F^t(u,v) \overline{F^{t'}(u,v)} \right. \\ &\quad \left. \cdot \left(\overline{d_{u,v}^{t',0}(\phi,\alpha)} \operatorname{Hess}_{(\phi,\alpha)} d_{u,v}^{t,0}(\phi,\alpha) + \operatorname{grad}_{(\phi,\alpha)} d_{u,v}^{t,0}(\phi,\alpha) \overline{\operatorname{grad}_{(\phi,\alpha)}^T d_{u,v}^{t',0}(\phi,\alpha)} \right) \right] \mathrm{d}v \\ &= - 2 \int_{-v_T}^{v_T} \sum_{|u|\leq u_T} \left| \mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \right|^2 \frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} (\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0})^{4-\gamma} \\ &\quad \cdot \Re\left[d_{u,v}^{t,0}(\phi,\alpha) \overline{d_{u,v}^{t',0}(\phi,\alpha)} \left(\mathrm{i} \boldsymbol{H}_{u,v}^{t,0}(\phi,\alpha) - \boldsymbol{a}_{u,v}^{t,0}(\phi,\alpha) \left(\boldsymbol{a}_{u,v}^{t,0}(\phi,\alpha) - \boldsymbol{a}_{u,v}^{t',0}(\phi,\alpha) \right)^\top \right) \right] \mathrm{d}v. \end{aligned} \tag{6.77}$$

Since $\phi \mapsto \rho_t^\phi$ and $\alpha \mapsto \sigma_t^\alpha$ are continuous by Assumption 2.14 (B4) and $d_{u,v}$ is Lipschitz-continuous, as $x \mapsto \mathrm{e}^{\mathrm{i}\langle a,x\rangle}$ is Lipschitz-continuous with Lipschitz-constant $\sqrt{2}\,\|a\|$ for $a \in \mathbb{R}^2$, the continuous mapping theorem (Theorem B.8) yields that

$$\left| d_{u,v}^{t,0}(\hat{\alpha}_T^*, \hat{\phi}_T^*) - 1 \right| \leq \sqrt{2}\,\|(2\pi u, -v)\| \left\| \begin{pmatrix} \rho_t^{\hat{\phi}_T^*} - \rho_t^{\phi_0} \\ \log(\sigma_t^{\hat{\alpha}_T^*}) - \log(\sigma_t^{\alpha_0}) \end{pmatrix} \right\| = o_{\mathbb{P}}\left(\|(u,v)\| \right),$$

where the logarithm is Lipschitz-continuous on $[\sigma_{\min}, \sigma_{\max}]$ because it has bounded derivative on this compact interval. Hence,

$$d_{u,v}^{t,0}(\hat{\alpha}_T^*, \hat{\phi}_T^*)\overline{d_{u,v}^{t',0}(\hat{\alpha}_T^*, \hat{\phi}_T^*)} = 1 + o_{\mathbb{P}}\left(\|(u,v)\|^2 + \|(u,v)\|\right).$$

In particular, the imaginary part vanishes asymptotically. With Lemma 6.4 and Assumption 2.13 (A4), it follows that

$$\begin{aligned}
&\Bigg\| \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \\
&\cdot\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}} (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma}\,\Re\left[d_{u,v}^{t,0}(\phi,\alpha)\overline{d_{u,v}^{t',0}(\phi,\alpha)}\mathrm{i}\boldsymbol{H}_{u,v}^{t,0}(\phi,\alpha)\right]\mathrm{d}v\Bigg\|_1 \\
\le\;& \max\left\{\sigma_{\min}^{8-2\gamma}, \sigma_{\max}^{8-2\gamma}\right\}\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \\
&\cdot\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left|\Im\left[d_{u,v}^{t,0}(\phi,\alpha)\overline{d_{u,v}^{t',0}(\phi,\alpha)}\right]\right|\left\|\boldsymbol{H}_{u,v}^{t,0}(\phi,\alpha)\right\|_1\mathrm{d}v \\
=\;& o_{\mathbb{P}}\left(\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\left(\|(u,v)\|^3+\|(u,v)\|^2\right)\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\mathrm{d}v\right) \\
=\;& o_{\mathbb{P}}(1).
\end{aligned} \tag{6.78}$$

Moreover,

$$\int_{-v_T}^{v_T}\sum_{|u|\le u_T}\left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}(\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma}\,\Re\left[\mathrm{i}\boldsymbol{H}_{u,v}^{t,0}(\phi_0,\alpha_0)\right]\mathrm{d}v = 0. \tag{6.79}$$

From Assumption 2.14 (B2), Lemma 6.4, and the continuous mapping theorem (Theorem B.8), we get that

$$\begin{aligned}
&\Big\|\boldsymbol{a}_{u,v}^{t,0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)\left(\boldsymbol{a}_{u,v}^{t,0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)-\boldsymbol{a}_{u,v}^{t',0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)\right)^\top \\
&-\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\left(\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)-\boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\right)^\top\Big\|_1 \\
\le\;& \left\|\boldsymbol{a}_{u,v}^{t,0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)\left[\left(\boldsymbol{a}_{u,v}^{t,0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)-\boldsymbol{a}_{u,v}^{t',0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)\right)^\top-\left(\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)-\boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\right)^\top\right]\right\|_1 \\
&+\left\|\left(\boldsymbol{a}_{u,v}^{t,0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)-\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right)\left(\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)-\boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\right)^\top\right\|_1 \\
\le\;& \left\|\boldsymbol{a}_{u,v}^{t,0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)\right\|_1\left(\left\|\boldsymbol{a}_{u,v}^{t,0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)-\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right\|_1+\left\|\boldsymbol{a}_{u,v}^{t',0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)-\boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\right\|_1\right) \\
&+\left\|\boldsymbol{a}_{u,v}^{t,0}(\hat{\phi}_T^*,\hat{\alpha}_T^*)-\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right\|_1\left(\left\|\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right\|_1+\left\|\boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\right\|_1\right) \\
\le\;& 4(d_2+d_3)\widetilde{C}\,\|(u,v)\|\left(2\pi|u|\left\|\mathrm{grad}_\phi(\rho_t^{\hat{\phi}_T^*})-\mathrm{grad}_\phi(\rho_t^{\phi_0})\right\|\right. \\
&\left.+|v|\left\|\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T^*}}\mathrm{grad}_\alpha(\sigma_t^{\hat{\alpha}_T^*})-\mathrm{grad}_\alpha(\sigma_t^{\alpha_0})\right\|\right) \\
=\;& o_{\mathbb{P}}\left(\|(u,v)\|^2\right).
\end{aligned}$$

This, together with (6.77), (6.78), (6.79), and Assumption 2.14 (B2), gives

$$
\begin{aligned}
&\left\| \operatorname{Hess}_{(\phi,\alpha)} A_T(\hat\phi_T^*, \hat\alpha_T^*) - \operatorname{Hess}_{(\phi,\alpha)} A_T(\phi_0,\alpha_0)\right\|_1 \\
=\;& 2\left\| \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma} \Big\{ \Re\left[d_{u,v}^{t,0}(\hat\alpha_T^*, \hat\phi_T^*) \overline{d_{u,v}^{t',0}(\hat\alpha_T^*, \hat\phi_T^*)}\right] \right. \\
&\cdot \boldsymbol{a}_{u,v}^{t,0}(\hat\phi_T^*, \hat\alpha_T^*)\big(\boldsymbol{a}_{u,v}^{t,0}(\hat\phi_T^*, \hat\alpha_T^*) - \boldsymbol{a}_{u,v}^{t',0}(\hat\phi_T^*, \hat\alpha_T^*)\big)^\top \\
&\left. - \boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\big(\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0) - \boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\big)^\top \Big\} \,\mathrm{d}v \right\|_1 + o_{\mathbb{P}}(1) \\
\le\;& o_{\mathbb{P}}\left(\int_{-v_T}^{v_T} \sum_{|u|\le u_T} \|(u,v)\|^4 \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \mathrm{d}v \right) + o_{\mathbb{P}}(1) \qquad (6.80) \\
=\;& o_{\mathbb{P}}(1). \qquad (6.81)
\end{aligned}
$$

By the first part of Lemma 6.11, H_M has finite operator norm. In particular, the components of H_M are finite. Hence,

$$
\left\| 2\int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \cdot \int_0^1\int_0^1 (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma} \boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\big(\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0) - \boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\big)^\top \mathrm{d}t\,\mathrm{d}t'\,\mathrm{d}v \right\|_1 = \mathcal{O}(1),
$$

as $T\to\infty$. and therefore

$$
\left\| 2\int_{\mathbb{R}\setminus[-v_T,v_T]} \sum_{u\in\mathbb{Z}:|u|>u_T} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \cdot \int_0^1\int_0^1 (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma} \boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\big(\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0) - \boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)\big)^\top \mathrm{d}t\,\mathrm{d}t'\,\mathrm{d}v \right\|_1 = o(1).
$$

With (6.77) and (6.81), this implies that

$$
\begin{aligned}
&\left\| \operatorname{Hess}_{(\phi,\alpha)} A_T(\hat\phi_T^*, \hat\alpha_T^*) - H_M\right\|_1 \\
=\;& \left\| \operatorname{Hess}_{(\phi,\alpha)} A_T(\phi_0,\alpha_0) - H_M\right\|_1 + o_{\mathbb{P}}(1) \\
\le\;& 2\int_{-v_T}^{v_T} \sum_{|u|\le u_T} \left|\mathcal{M}_{|\mathcal{F}_f|^2}(u,v)\right|^2 \left(\left\| \frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}} \boldsymbol{A}_{u,v}^{t,t'} - \int_0^1\int_0^1 \boldsymbol{A}_{u,v}^{t,t'}\,\mathrm{d}t\,\mathrm{d}t' \right\|_1 \right. \\
&\left. + \left\| \frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}} \boldsymbol{B}_{u,v}^{t,t'} - \int_0^1\int_0^1 \boldsymbol{B}_{u,v}^{t,t'}\,\mathrm{d}t\,\mathrm{d}t' \right\|_1 \right) \mathrm{d}v + o_{\mathbb{P}}(1), \qquad (6.82)
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{A}_{u,v}^{t,t'} &:= (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma} \boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)^\top, \\
\boldsymbol{B}_{u,v}^{t,t'} &:= (\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0})^{4-\gamma} \boldsymbol{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\boldsymbol{a}_{u,v}^{t',0}(\phi_0,\alpha_0)^\top.
\end{aligned}
$$

Note that functions on $\mathbb{R}$ with bounded derivative on a compact interval are Lipschitz, restricted to that interval. Using parts 1, 3 and 4 of Lemma A.4, we have by Assumption 2.14 (B3) that

$$
\begin{aligned}
&\left\| \frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} \boldsymbol{A}_{u,v}^{t,t'} - \int_0^1 \int_0^1 \boldsymbol{A}_{u,v}^{t,t'} \,\mathrm{d}t\,\mathrm{d}t' \right\|_1 \\
= \;& \sum_{m,m'=1}^{d_2+d_3} \left| \frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} \left(\boldsymbol{A}_{u,v}^{t,t'}\right)_{m,m'} - \int_0^1 \int_0^1 \left(\boldsymbol{A}_{u,v}^{t,t'}\right)_{m,m'} \,\mathrm{d}t\,\mathrm{d}t' \right| \\
\leq \;& \max\left\{C_\gamma \widetilde{C}^2 \|(u,v)\|^2, C_\gamma\right\} \frac{\beta_T}{T} \\
&\cdot \sum_{m,m'=1}^{d_2+d_3} \left\{ \mathrm{TV}\left[t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma} \left(\mathbf{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right)_m \left(\mathbf{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right)_{m'}\right] + \mathrm{TV}\left[t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma}\right] \right\} \\
\leq \;& \max\left\{C_\gamma \widetilde{C}^2 \|(u,v)\|^2, C_\gamma\right\} \frac{\beta_T}{T} \left\{ 2(d_2+d_3) C_\gamma \widetilde{C} \|(u,v)\| \sum_{m=1}^{d_2+d_3} \mathrm{TV}\left[t \mapsto \left(\mathbf{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right)_m\right] \right. \\
&\left. + (d_2+d_3)^2 C_3 \widetilde{C}^2 \|(u,v)\|^2 + (d_2+d_3)^2 C_3 \right\} \\
\leq \;& (d_2+d_3) \max\left\{C_\gamma \widetilde{C}^2 \|(u,v)\|^2, C_\gamma\right\} \frac{\beta_T}{T} \left\{ 2 C_\gamma \widetilde{C} \|(u,v)\| \left(2\pi |u| \sum_{m=1}^{d_2} \mathrm{TV}\left[t \mapsto \left.\frac{\partial \rho_t^\phi}{\partial \phi_m}\right|_{\phi=\phi_0} \right] \right.\right. \\
&\left.\left. + |v| \sum_{m'=1}^{d_3} \mathrm{TV}\left[t \mapsto \left.\frac{\partial \sigma_t^\alpha}{\partial \alpha_{m'}}\right|_{\alpha=\alpha_0} \right] \right) + (d_2+d_3) C_3 \left(\widetilde{C}^2 \|(u,v)\|^2 + 1\right) \right\} \\
= \;& O_{\mathbb{P}}\left(\|(u,v)\|^4 \frac{\beta_T}{T} \right),
\end{aligned}
\tag{6.83}
$$

with $\widetilde{C} > 0$ from Lemma 6.4, $C_\gamma \coloneqq \max\left\{\sigma_{\min}^{4-\gamma}, \sigma_{\max}^{4-\gamma}\right\}$, and $C_3 > 0$ from (6.41). Similarly,

$$
\begin{aligned}
&\left\| \frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} \boldsymbol{B}_{u,v}^{t,t'} - \int_0^1 \int_0^1 \boldsymbol{B}_{u,v}^{t,t'} \,\mathrm{d}t\,\mathrm{d}t' \right\|_1 \\
\leq \;& C_\gamma \widetilde{C} \|(u,v)\| \frac{\beta_T}{T} \\
&\cdot \sum_{m,m'=1}^{d_2+d_3} \left\{ \mathrm{TV}\left[t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma} \left(\mathbf{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right)_m\right] + \mathrm{TV}\left[t \mapsto (\sigma_t^{\alpha_0})^{4-\gamma} \left(\mathbf{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right)_{m'}\right] \right\} \\
\leq \;& 2(d_2+d_3) C_\gamma \widetilde{C} \|(u,v)\| \frac{\beta_T}{T} \\
&\cdot \left\{ C_\gamma \sum_{m=1}^{d_2+d_3} \mathrm{TV}\left[t \mapsto \left(\mathbf{a}_{u,v}^{t,0}(\phi_0,\alpha_0)\right)_m\right] + (d_2+d_3) C_3 \widetilde{C} \|(u,v)\| \right\} \\
= \;& O_{\mathbb{P}}\left(\|(u,v)\|^2 \frac{\beta_T}{T} \right).
\end{aligned}
\tag{6.84}
$$

From (6.82), (6.83), and (6.84), we get that

$$
\left\| \mathrm{Hess}_{(\phi,\alpha)} A_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 O_{\mathbb{P}}\left(\frac{\beta_T}{T} \int_{-v_T}^{v_T} \sum_{|u|\leq u_T} \|(u,v)\|^4 \left| \mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \right|^2 \mathrm{d}v \right),
\tag{6.85}
$$

which converges to 0 due to Assumptions 2.13 (A4) and 2.15 (C1).

Next, we show that

$$\sup_{(\phi,\alpha)\in U} \left\| \operatorname{Hess}_{(\phi,\alpha)} B_T(\phi,\alpha) \right\|_1 \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability.} \tag{6.86}$$

From Lemma 6.4, we have

$$\sup_{(\phi,\alpha)\in U} \left\| \operatorname{Hess}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha) \right\|_1 = O\big(\|(u,v)\|^2\big). \tag{6.87}$$

By Lemma 6.2 and (6.87), we get that

$$\begin{aligned}
&\sup_{(\phi,\alpha)\in U} \left\| \operatorname{Hess}_{(\phi,\alpha)} B_T(\phi,\alpha) \right\|_1 \\
=\;& \sup_{(\phi,\alpha)\in U} \left\| \int_{-v_T}^{v_T} \sum_{|u|\le u_T} 2\Re\left[\frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} F^t(u,v) \overline{\mathcal{M}_{\mathcal{W}^{t'}}^T(u,v)} \operatorname{Hess}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha) \right] \mathrm{d}v \right\|_1 \\
\le\;& 2C_\gamma \sup_{(\phi,\alpha)\in U} \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \left| \mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \right| \frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} \left| \mathcal{M}_{\mathcal{W}^{t'}}^T(u,v) \right| \left\| \operatorname{Hess}_{(\phi,\alpha)} d_{u,v}^{t,t'}(\phi,\alpha) \right\|_1 \mathrm{d}v \\
=\;& \mathcal{O}_{\mathbb{P}}\left(\frac{r_T^\gamma}{\sqrt{\beta_T}} \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \|(u,v)\|^2 \left| \mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \right| \mathrm{d}v \right) \\
\le\;& \mathcal{O}_{\mathbb{P}}\left(\frac{r_T^\gamma \|(u_T,v_T)\|}{\sqrt{\beta_T}} \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \|(u,v)\| \left| \mathcal{M}_{|\mathcal{F}_f|^2}(u,v) \right| \mathrm{d}v \right).
\end{aligned}$$

Because of Assumption 2.13 (A4), Lemma A.11 and since $r_T^{2\gamma} \|(u_T,v_T)\|^2 / \beta_T \overset{T\to\infty}{\longrightarrow} 0$ as a consequence of Assumption 2.15 (C2), the above converges to 0 and we conclude that (6.86) holds.

Finally, we consider $\sup_{(\phi,\alpha)\in U} \left\| \operatorname{Hess}_{(\phi,\alpha)} C_T(\phi,\alpha) \right\|_1$. By Lemma 6.2 and (6.87),

$$\begin{aligned}
&\sup_{(\phi,\alpha)\in U} \left\| \operatorname{Hess}_{(\phi,\alpha)} C_T(\phi,\alpha) \right\|_1 \\
=\;& \sup_{(\phi,\alpha)\in U} \left\| \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} \operatorname{Hess}_{(\phi,\alpha)} \frac{d_{u,v}^{0,t'}(\phi,\alpha)}{d_{u,v}^{0,t}(\phi,\alpha)} \mathcal{M}_{\mathcal{W}^t}^T(u,v) \overline{\mathcal{M}_{\mathcal{W}^{t'}}^T(u,v)} \mathrm{d}v \right\|_1 \\
\le\;& \sup_{(\phi,\alpha)\in U} \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \frac{\beta_T^2}{T^2} \sum_{t,t'\in\mathbb{T}} \left\| \operatorname{Hess}_{(\phi,\alpha)} \frac{d_{u,v}^{0,t'}(\phi,\alpha)}{d_{u,v}^{0,t}(\phi,\alpha)} \right\|_1 \left| \mathcal{M}_{\mathcal{W}^t}^T(u,v) \right| \left| \mathcal{M}_{\mathcal{W}^{t'}}^T(u,v) \right| \mathrm{d}v \\
=\;& \mathcal{O}_{\mathbb{P}}\left(\frac{r_T^{2\gamma}}{\beta_T} \int_{-v_T}^{v_T} \sum_{|u|\le u_T} \|(u,v)\|^2 \,\mathrm{d}v \right) \\
=\;& \mathcal{O}_{\mathbb{P}}\left(\frac{\|(u_T,v_T)\|^2 u_T v_T r_T^{2\gamma}}{\beta_T} \right).
\end{aligned}$$

Since $\|(u_T, v_T)\|^2 u_T v_T r_T^{2\gamma}/\beta_T \overset{T\to\infty}{\longrightarrow} 0$ by Assumption 2.15 (C2), we get

$$\sup_{(\phi,\alpha)\in U} \left\| \text{Hess}_{(\phi,\alpha)} C_T(\phi,\alpha) \right\|_1 \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability.} \tag{6.88}$$

From (6.85), (6.86), and (6.88), we conclude that

$$\left\| \text{Hess}_{(\phi,\alpha)} M_T(\hat{\phi}_T^*, \hat{\alpha}_T^*) - H_M \right\|_1 \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability.} \qquad \square$$

Proof of Theorem 3.5 (Central limit theorem for the rotation and scaling parameter estimators).

We now bring together the results of Theorem 6.10, Lemma 6.11 and Theorem 6.12 to proof asymptotic normality for the rotation and scaling parameter estimators.

By Assumption 2.14 (B3), M_T is twice continuously differentiable in a convex open neighborhood $U \subseteq \Phi \times \mathrm{A}$ of (ϕ_0, α_0). In particular, if M_T has a minimum at some $(\phi, \alpha) \in U$, then $\text{grad}_{(\phi,\alpha)} M_T(\phi, \alpha) = 0$. Let

$$G_T(\phi,\alpha) := \begin{cases} \text{grad}_{(\phi,\alpha)} M_T(\phi,\alpha), & \text{if } (\phi,\alpha) \in U, \\ \text{grad}_{(\phi,\alpha)} M_T(\phi_0,\alpha_0), & \text{if } (\phi,\alpha) \in (\Phi \times \mathrm{A}) \setminus U. \end{cases}$$

Since $(\hat{\phi}_T, \hat{\alpha}_T)$ is defined as a minimizer of M_T (i.e., $\text{grad}_{(\phi,\alpha)} M_T(\hat{\phi}_T, \hat{\alpha}_T) = 0$) and because $(\hat{\phi}_T, \hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} (\phi_0, \alpha_0) \in U$ in probability, we have for all $\epsilon > 0$ that

$$\begin{aligned} \mathbb{P}\left(\sqrt{T} G_T(\hat{\phi}_T, \hat{\alpha}_T) > \epsilon \right) &= \mathbb{P}\left(\sqrt{T} \text{grad}_{(\phi,\alpha)} M_T(\hat{\phi}_T, \hat{\alpha}_T) > \epsilon, (\hat{\phi}_T, \hat{\alpha}_T) \in U \right) \\ &\quad + \mathbb{P}\left(\sqrt{T} \text{grad}_{(\phi,\alpha)} M_T(\phi_0, \alpha_0) > \epsilon, (\hat{\phi}_T, \hat{\alpha}_T) \notin U \right) \\ &\leq \mathbb{P}\left((\hat{\phi}_T, \hat{\alpha}_T) \notin U \right) \overset{T\to\infty}{\longrightarrow} 0, \end{aligned}$$

which means that

$$\sqrt{T} G_T(\hat{\phi}_T, \hat{\alpha}_T) = o_{\mathbb{P}}(1). \tag{6.89}$$

For $(\phi, \alpha) \in U$, we can apply the mean value theorem for real functions of multiple variables to each component of $\text{grad}_{(\phi,\alpha)} M_T(\phi, \alpha)$ to get that

$$\text{grad}_{(\phi,\alpha)} M_T(\phi,\alpha) = \text{grad}_{(\phi,\alpha)} M_T(\phi_0,\alpha_0) + \text{Hess}_{(\phi,\alpha)} M_T(\phi^\dagger, \alpha^\dagger) \begin{pmatrix} \phi - \phi_0 \\ \alpha - \alpha_0 \end{pmatrix}, \tag{6.90}$$

where $(\phi^\dagger, \alpha^\dagger) \in U$ such that its components are convex combinations of the respective components of (ϕ, α) and (ϕ_0, α_0). By (6.90), on the event $\{(\hat{\phi}_T, \hat{\alpha}_T) \in U\}$, we can find $(\hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger) \in U$ between $(\hat{\phi}_T, \hat{\alpha}_T)$ and (ϕ_0, α_0) such that

$$\text{grad}_{(\phi,\alpha)} M_T(\hat{\phi}_T, \hat{\alpha}_T) = \text{grad}_{(\phi,\alpha)} M_T(\phi_0,\alpha_0) + \text{Hess}_{(\phi,\alpha)} M_T(\hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger) \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix}. \tag{6.91}$$

With the definitions

$$(\hat\phi_T^*, \hat\alpha_T^*) := \begin{cases} (\hat\phi_T^\dagger, \hat\alpha_T^\dagger), & \text{if } (\hat\phi_T, \hat\alpha_T) \in U, \\ (\phi_0, \alpha_0), & \text{if } (\hat\phi_T, \hat\alpha_T) \in (\Phi \times \mathrm{A}) \setminus U, \end{cases}$$

and

$$H_T := \begin{cases} \mathrm{Hess}_{(\phi,\alpha)} M_T(\hat\phi_T^*, \hat\alpha_T^*) = \mathrm{Hess}_{(\phi,\alpha)} M_T(\hat\phi_T^\dagger, \hat\alpha_T^\dagger), & \text{if } (\hat\phi_T, \hat\alpha_T) \in U, \\ 0, & \text{if } (\hat\phi_T, \hat\alpha_T) \in (\Phi \times \mathrm{A}) \setminus U, \end{cases}$$

and using (6.89), we get that

$$\sqrt{T} \mathrm{grad}_{(\phi,\alpha)} M_T(\phi_0, \alpha_0) + H_T \sqrt{T} \begin{pmatrix} \hat\phi_T - \phi_0 \\ \hat\alpha_T - \alpha_0 \end{pmatrix} = \sqrt{T} G_T(\hat\phi_T, \hat\alpha_T) = o_{\mathbb{P}}(1), \tag{6.92}$$

which holds on $\{(\hat\phi_T, \hat\alpha_T) \notin U\}$ by design of G_T and H_T and on $\{(\hat\phi_T, \hat\alpha_T) \in U\}$ due to (6.91). Equation (6.92) and Theorem 6.10 yield that

$$H_T \sqrt{T} \begin{pmatrix} \hat\phi_T - \phi_0 \\ \hat\alpha_T - \alpha_0 \end{pmatrix} = -\sqrt{T} \mathrm{grad}_{(\phi,\alpha)} M_T(\phi_0, \alpha_0) + o_{\mathbb{P}}(1) = O_{\mathbb{P}}(1). \tag{6.93}$$

Since $(\hat\phi_T^*, \hat\alpha_T^*) \in U$ is between $(\hat\phi_T, \hat\alpha_T)$ and (ϕ_0, α_0) and $(\hat\phi_T, \hat\alpha_T)$ is a consistent estimator, we have that $(\hat\phi_T^*, \hat\alpha_T^*) \xrightarrow{T\to\infty} (\phi_0, \alpha_0)$ in probability. Because of Assumption 2.14 (B8) and Lemma 6.11, H_M is invertible, and by Theorem 6.12, $H_T \xrightarrow{T\to\infty} H_M$ in probability. Together with (6.93) and Lemma A.12, this implies that

$$\sqrt{T} \begin{pmatrix} \hat\phi_T - \phi_0 \\ \hat\alpha_T - \alpha_0 \end{pmatrix} = O_{\mathbb{P}}(1). \tag{6.94}$$

Hence, again with Theorem 6.12,

$$(H_T - H_M) \sqrt{T} \begin{pmatrix} \hat\phi_T - \phi_0 \\ \hat\alpha_T - \alpha_0 \end{pmatrix} = o_{\mathbb{P}}(1) O_{\mathbb{P}}(1) = o_{\mathbb{P}}(1).$$

From this, (6.93), and Theorem 6.10, it follows that

$$\begin{aligned} H_M \sqrt{T} \begin{pmatrix} \hat\phi_T - \phi_0 \\ \hat\alpha_T - \alpha_0 \end{pmatrix} &= -\sqrt{T} \mathrm{grad}_{(\phi,\alpha)} M_T(\phi_0, \alpha_0) + o_{\mathbb{P}}(1) \\ &\xrightarrow{T\to\infty} \mathcal{N}(0, \Sigma_{RS}) \quad \text{in distribution,} \end{aligned}$$

where we used that, for all centered normal random vectors X, X and $-X$ have the same distribution. Finally, multiplication with H_M^{-1} yields the assertion. □

6.2.5 Proofs of Theorem 3.6 and Theorem 3.7: Central limit theorem for the drift parameter estimator and joint central limit theorem

The first step of the proof of asymptotic normality of the drift parameter estimator is to show the convergence of the mixed derivatives of the empirical contrast functional N_T, see Theorem 6.13. Following that, we proceed in a way similar to the proof of asymptotic normality of the rotation and scaling parameter estimators. In Theorem 6.14, we show the asymptotic normality of the gradient of the empirical contrast functional N_T at the true parameters, in Lemma 6.15 we derive properties of the Hessian of the population contrast functional N, and in Theorem 6.16 we then show the convergence of the Hessian of the empirical contrast functional, N_T, to the population contrast functional, N. Afterwards, we collect the results to prove asymptotic normality of the drift parameter estimator and combine this with the asymptotic normality of the rotation and scaling parameter estimators (see Theorem 3.5) in order to obtain the desired joint central limit theorem of all three motion function parameters. Similarly to before, the proofs of the preparatory Theorems 6.13, 6.14, 6.16 and Lemma 6.15 are based on the corresponding Theorems 6.35, 6.41 and 6.42 as well as Lemma 6.39 in Hartmann (2016), but are edited to fit the revised model and reworked to provide for more comprehensive demonstrations. Moreover, instead of uniform tightness of the gradient of the empirical contrast functional and the drift parameter estimator, here a central limit theorem is proven, which together with the asymptotic normality of the rotation and scaling parameter estimators enables the derivation of a joint central limit theorem of all three motion parameter estimators. This justifies the bootstrap procedure, which is applied in Section 4.3 in order to assess the statistical uncertainty of the estimation.

Theorem 6.13. *Suppose that Assumptions 2.13 (A1,A3), 2.14 (B2-B3) and Assumption 2.15 hold and that* $(\hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger) \overset{T\to\infty}{\longrightarrow} (\phi_0, \alpha_0)$ *in probability. Then, we have that*

$$\left\|\operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top N_T(\theta_0; \hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger) - \operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top N(\theta_0; \phi_0, \alpha_0)\right\|_1 \overset{T\to\infty}{\longrightarrow} 0 \text{ in probability.}$$

Proof. Again, we proceed in two steps, using the decomposition (2.23). For U from Assumption 2.14 (B2) we show that

$$\left\|\operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top A_T(\theta_0; \hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger) - \operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top N(\theta_0; \phi_0, \alpha_0)\right\|_1 \overset{T\to\infty}{\longrightarrow} 0 \text{ in probability,}$$

and

$$\sup_{(\phi,\alpha)\in U} \left\|\operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top B_T(\theta_0; \phi, \alpha) + \operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top C_T(\theta_0; \phi, \alpha)\right\|_1 \overset{T\to\infty}{\longrightarrow} 0 \text{ in probability.}$$

First, show the convergence of the mixed derivatives of A_T. Using Lemma A.6 and Lemma 6.7 we see that

$$\begin{aligned}
&\operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top N(\theta; \phi, \alpha)\\
&= -\operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top \int_{\mathbb{R}^2} \left| \int_0^1 F_\omega^t(\theta; \phi, \alpha)\,\mathrm{d}t \right|^2 \mathrm{d}\omega
\end{aligned}$$

$$= - \operatorname{grad}_{(\phi,\alpha)} \int_{R^2} \int_0^1 \int_0^1 2\Re \left(\operatorname{grad}_\theta^\top F_\omega^t(\theta;\phi,\alpha) \overline{F_\omega^{t'}(\theta;\phi,\alpha)} \right) \mathrm{d}t\mathrm{d}t'\mathrm{d}\omega$$
$$= - 2 \int_{R^2} \int_0^1 \int_0^1 \operatorname{grad}_{(\phi,\alpha)} \Re \left(\mathrm{i} \boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha)^\top F_\omega^t(\theta;\phi,\alpha) \overline{F_\omega^{t'}(\theta;\phi,\alpha)} \right) \mathrm{d}t\mathrm{d}t'\mathrm{d}\omega$$
$$= 2 \int_{R^2} \int_0^1 \int_0^1 \operatorname{grad}_{(\phi,\alpha)} \Im \left(\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha)^\top F_\omega^t(\theta;\phi,\alpha) \overline{F_\omega^{t'}(\theta;\phi,\alpha)} \right) \mathrm{d}t\mathrm{d}t'\mathrm{d}\omega$$

and in particular

$$\operatorname{grad}_{(\phi,\alpha)} \operatorname{grad}_\theta^\top N(\theta_0;\phi_0,\alpha_0)$$
$$= 2 \int_{R^2} \int_0^1 \int_0^1 \operatorname{grad}_{(\phi,\alpha)} \Im \left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi,\alpha)^\top F_\omega^t(\theta_0;\phi,\alpha) \overline{F_\omega^{t'}(\theta_0;\phi,\alpha)} \right) \Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)} \mathrm{d}t\mathrm{d}t'\mathrm{d}\omega.$$

Now, we calculate the mixed derivatives of A_T. For the gradient of the integrand of $A_T(\theta;\phi,\alpha)$ with respect to θ we get

$$\operatorname{grad}_\theta \left[- \sum_{t,t' \in \mathbb{T}} F_\omega^t(\theta_0;\phi,\alpha) \overline{F_\omega^{t'}(\theta_0;\phi,\alpha)} \right]$$
$$= -2 \sum_{t,t' \in \mathbb{T}} \Re \left[\operatorname{grad}_\theta \left(h_\omega^{t,0}(\theta;\phi,\alpha) \right) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha} \right)^2 \mathcal{F}_f \left(\tau_t^{(\phi,\alpha)} \omega \right) \right.$$
$$\left. \cdot \overline{h_\omega^{t',0}(\theta;\phi,\alpha) \left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^\alpha} \right)^2 \mathcal{F}_f \left(\tau_{t'}^{(\phi,\alpha)} \omega \right)} \right]$$
$$= 2 \sum_{t,t' \in \mathbb{T}} \Im \left[\left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^\alpha \sigma_{t'}^\alpha} \right)^2 \boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha) \frac{h_\omega^{t,0}(\theta;\phi,\alpha)}{h_\omega^{t',0}(\theta;\phi,\alpha)} \mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega) \overline{\mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)} \right]$$
$$= 2 \sum_{t,t' \in \mathbb{T}} \Im \left[\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha) F_\omega^t(\theta;\phi,\alpha) \overline{F_\omega^{t'}(\theta;\phi,\alpha)} \right].$$

In a second step, we see that

$$\operatorname{grad}_{(\phi,\alpha)} \left\{ \Im \left[\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha)^\top F_\omega^t(\theta;\phi,\alpha) \overline{F_\omega^{t'}(\theta;\phi,\alpha)} \right] \right\} \Big|_{(\phi,\alpha)=(\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger)}$$
$$= \operatorname{grad}_{(\phi,\alpha)} \left\{ \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^\alpha \sigma_{t'}^\alpha} \right)^2 \Im \left[\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha)^\top h_\omega^{t,0}(\theta;\phi,\alpha) \mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega) \right. \right.$$
$$\left. \left. \cdot \overline{h_\omega^{t',0}(\theta;\phi,\alpha) \mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)} \right] \right\} \Big|_{(\phi,\alpha)=(\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger)}$$
$$= \left[\operatorname{grad}_{(\phi,\alpha)} \left\{ \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^\alpha \sigma_{t'}^\alpha} \right)^2 \right\} \Im \left[\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha)^\top h_\omega^{t,0}(\theta;\phi,\alpha) \mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega) \right. \right.$$
$$\left. \left. \cdot \overline{h_\omega^{t',0}(\theta;\phi,\alpha) \mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)} \right] \right] \Big|_{(\phi,\alpha)=(\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger)}$$
$$+ \left[\left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^\alpha \sigma_{t'}^\alpha} \right)^2 \operatorname{grad}_{(\phi,\alpha)} \Im \left[\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha)^\top h_\omega^{t,0}(\theta;\phi,\alpha) \mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega) \right. \right.$$

$$\cdot \overline{h_\omega^{t',0}(\theta;\phi,\alpha)\mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)}\Big]\Big]\Big|_{(\phi,\alpha)=(\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger)}$$

$$\overset{T\to\infty}{\longrightarrow} \Im\left[\boldsymbol{b}_\omega^{t,0}(\theta;\phi_0,\alpha_0)^\top \frac{h_\omega^{t,0}(\theta;\phi_0,\alpha_0)}{h_\omega^{t',0}(\theta;\phi_0,\alpha_0)}\left|\mathcal{F}_f(\omega)\right|^2\right] \operatorname{grad}_{(\phi,\alpha)}\left\{\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\alpha}\sigma_{t'}^{\alpha}}\right)^2\right\}\Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)}$$
$$+ \operatorname{grad}_{(\phi,\alpha)}\Im\left[\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha)^\top F_\omega^t(\theta;\phi,\alpha)\overline{F_\omega^{t'}(\theta;\phi,\alpha)}\right]\Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)}, \tag{6.95}$$

with convergence in probability, and therefore the mixed derivatives of A_T can be obtained as

$$\operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top A_T(\theta_0;\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger)$$
$$= 2\sum_{t,t'\in\mathbb{T}}\left(\frac{\beta_T}{T}\right)^2\int_{\Omega_T}\operatorname{grad}_{(\phi,\alpha)}\Im\left[\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi,\alpha)^\top F_\omega^t(\theta_0;\phi,\alpha)\overline{F_\omega^{t'}(\theta_0;\phi,\alpha)}\right]\Big|_{(\phi,\alpha)=(\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger)}\mathrm{d}\omega$$
$$\overset{T\to\infty}{\longrightarrow} 2\int_{\mathbb{R}^2}\int_0^1\int_0^1 \Im\left[\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)^\top\left|\mathcal{F}_f(\omega)\right|^2\right]\operatorname{grad}_{(\phi,\alpha)}\left\{\left(\frac{\sigma_{t'}^{\alpha_0}\sigma_t^{\alpha_0}}{\sigma_{t'}^{\alpha}\sigma_t^{\alpha}}\right)^2\right\}\Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)}$$
$$+ \operatorname{grad}_{(\phi,\alpha)}\Im\left[\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi,\alpha)^\top\frac{h_\omega^{t,0}(\theta_0;\phi,\alpha)}{h_\omega^{t',0}(\theta_0;\phi,\alpha)}\mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega)\overline{\mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)}\right]\Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)}\mathrm{d}t\mathrm{d}t'\mathrm{d}\omega$$
$$= 2\int_{\mathbb{R}^2}\int_0^1\int_0^1\operatorname{grad}_{(\phi,\alpha)}\Im\left[\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi,\alpha)^\top\right.$$
$$\left.\cdot\frac{h_\omega^{t,0}(\theta_0;\phi,\alpha)}{h_\omega^{t',0}(\theta_0;\phi,\alpha)}\mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega)\overline{\mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)}\right]\Big|_{(\phi,\alpha)=(\phi_0,\alpha_0)}\mathrm{d}t\mathrm{d}t'\mathrm{d}\omega$$
$$= \operatorname{grad}_{(\phi,\alpha)}\operatorname{grad}_\theta^\top N(\theta_0;\phi_0,\alpha_0)$$

in probability, where we used (6.10) and the fact that the imaginary part of reals is zero in the second and third steps. The integrals converge by the dominated convergence theorem (Amann and Escher, 2001, Theorem 3.12), as shown below. We need to give an upper bound for the norm of the appearing gradient. This can be done as follows. First, we need some preparatory steps. Note that with $\widetilde{C}$ from Lemma 6.7 and using $\left|h_\omega^{t,t'}(\theta;\phi,\alpha)\right| = 1$,

$$\left\|\operatorname{grad}_{(\phi,\alpha)}\left(h_\omega^{t,0}(\theta;\phi,\alpha)/h_\omega^{t',0}(\theta;\phi,\alpha)\right)\right\|$$
$$= \left\|\frac{\boldsymbol{c}_\omega^{t,0}(\theta;\phi,\alpha)h_\omega^{t,0}(\theta;\phi,\alpha)h_\omega^{t',0}(\theta;\phi,\alpha) - \boldsymbol{c}_\omega^{t',0}(\theta;\phi,\alpha)h_\omega^{t,0}(\theta;\phi,\alpha)h_\omega^{t',0}(\theta;\phi,\alpha)}{h_\omega^{t',0}(\theta;\phi,\alpha)^2}\right\|$$
$$\le 2\widetilde{C}\left\|\omega\right\|. \tag{6.96}$$

Second, since the norms of the gradients of the motions functions are bounded by C from Assumption 2.14 (B2), we have for $x \in \mathbb{R}^2$ with $C^\tau := 2\sigma_{\max}\left(\sigma_{min}^{-1} + \sigma_{\min}^{-2}\right)C$, using the norm equivalence that

$$\left\|\operatorname{grad}_{(\phi,\alpha)}\left\langle\tau_t^{\phi,\alpha}\omega, x\right\rangle\right\| \le \left\|\operatorname{grad}_{(\phi,\alpha)}\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\alpha}}\left[\left(\cos\left(\rho_t^\phi - \rho_t^{\phi_0}\right)\omega_1 - \sin\left(\rho_t^\phi - \rho_t^{\phi_0}\right)\omega_2\right)x_1\right.\right.\right.$$
$$\left.\left.\left. + \left(\sin\left(\rho_t^\phi - \rho_t^{\phi_0}\right)\omega_1 + \cos\left(\rho_t^\phi - \rho_t^{\phi_0}\right)\omega_2\right)x_2\right]\right)\right\|_1$$
$$\le \frac{\sigma_{\max}}{\sigma_{\min}^2}C\left\|\omega\right\|_1\left\|x\right\|_1 + \frac{\sigma_{\max}}{\sigma_{\min}}C\left\|\omega\right\|_1\left\|x\right\|_1 \le C^\tau\left\|\omega\right\|\left\|x\right\|.$$

We only need to consider $x \in \mathbb{R}^2$ with $\|x\| \le C_f$ for C_f from Assumption 2.13 (A1). Since f is integrable as a consequence of Assumption 2.13 (A1), we have that the integral over its absolute value is bounded by some constant $\widetilde{C}_f$. It now follows that

$$\begin{aligned}
&\left\|\operatorname{grad}_{(\phi,\alpha)}\mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega)\right\| = \left\|\operatorname{grad}_{(\phi,\alpha)}\int_{\mathbb{R}^2}\exp\left(-2\pi\mathrm{i}\left\langle\tau_t^{(\phi,\alpha)}\omega, x\right\rangle\right)f(x)\mathrm{d}x\right\| \\
&= 2\pi\left\|\int_{\mathbb{R}^2}\exp\left(-2\pi\mathrm{i}\left\langle\tau_t^{(\phi,\alpha)}\omega, x\right\rangle\right)f(x)\operatorname{grad}_{(\phi,\alpha)}\left\langle\tau_t^{(\phi,\alpha)}\omega, x\right\rangle\mathrm{d}x\right\| \\
&\le 2\pi C^\tau\|\omega\|\,C_f\int_{\mathbb{R}^2}|f(x)|\,\mathrm{d}x \le 2\pi C^\tau C_f\widetilde{C}_f\|\omega\|. \qquad (6.97)
\end{aligned}$$

Hence, using Lemma 6.8, we get that

$$\begin{aligned}
&\left\|\operatorname{grad}_{(\phi,\alpha)}\left(\mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega)\overline{\mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)}\right)\Big|_{(\phi,\alpha)=(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\right\| \\
&=\left\|\operatorname{grad}_{(\phi,\alpha)}\mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega)\Big|_{(\phi,\alpha)=(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\overline{\mathcal{F}_f(\tau_{t'}^{(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\omega)}\right. \\
&\quad\left.+\mathcal{F}_f(\tau_t^{(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\omega)\operatorname{grad}_{(\phi,\alpha)}\overline{\mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)}\Big|_{(\phi,\alpha)=(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\right\| \\
&\le 4\pi C^\tau C_f\widetilde{C}_f\|\omega\|\left(\left|\mathcal{F}_f(\omega)\right| + O_{\mathbb{P}}\left(T^{-\frac{1}{2}}\right)\right). \qquad (6.98)
\end{aligned}$$

We further obtain for C from Assumption 2.14 (B2)

$$\begin{aligned}
\left\|\operatorname{grad}_{(\phi,\alpha)}\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\alpha}\sigma_{t'}^{\alpha}}\right)^2\right\| &= \left\|-2\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\alpha}\sigma_{t'}^{\alpha}}\right)^2\frac{\sigma_{t'}^{\alpha}\operatorname{grad}_{(\phi,\alpha)}\sigma_t^{\alpha}+\sigma_t^{\alpha}\operatorname{grad}_{(\phi,\alpha)}\sigma_{t'}^{\alpha}}{\sigma_t^{\alpha}\sigma_{t'}^{\alpha}}\right\| \\
&\le 4\sigma_{\max}^4\sigma_{\min}^{-5}C. \qquad (6.99)
\end{aligned}$$

Collecting (6.96), (6.98) and (6.99) and using $\left|h_\omega^{t,t'}(\theta;\phi,\alpha)\right| = 1$ and Lemma 6.7, we obtain

$$\begin{aligned}
&\left\|\operatorname{grad}_{(\phi,\alpha)}\left[F_\omega^t(\theta_0;\phi,\alpha)\overline{F_\omega^{t'}(\theta_0;\phi,\alpha)}\right]\Big|_{(\phi,\alpha)=(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\right\|_1 \\
&=\left\|\operatorname{grad}_{(\phi,\alpha)}\frac{h_\omega^{t,0}(\theta;\phi,\alpha)}{h_\omega^{t',0}(\theta;\phi,\alpha)}\Big|_{(\phi,\alpha)=(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat\alpha_T^\dagger}\sigma_{t'}^{\hat\alpha_T^\dagger}}\right)^2\mathcal{F}_f(\tau_t^{(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\omega)\overline{\mathcal{F}_f(\tau_{t'}^{(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\omega)}\right. \\
&\quad+\frac{h_\omega^{t,0}(\theta;\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}{h_\omega^{t',0}(\theta;\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\operatorname{grad}_{(\phi,\alpha)}\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\alpha}\sigma_{t'}^{\alpha}}\right)^2\Big|_{(\phi,\alpha)=(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\mathcal{F}_f(\tau_t^{(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\omega)\overline{\mathcal{F}_f(\tau_{t'}^{(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\omega)} \\
&\quad\left.+\frac{h_\omega^{t,0}(\theta;\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}{h_\omega^{t',0}(\theta;\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat\alpha_T^\dagger}\sigma_{t'}^{\hat\alpha_T^\dagger}}\right)^2\operatorname{grad}_{(\phi,\alpha)}\left(\mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega)\overline{\mathcal{F}_f(\tau_{t'}^{(\phi,\alpha)}\omega)}\right)\Big|_{(\phi,\alpha)=(\hat\phi_T^\dagger,\hat\alpha_T^\dagger)}\right\|_1 \\
&\le 2\widetilde{C}\|\omega\|\sqrt{d_2+d_3}\,\sigma_{\max}^4\sigma_{\min}^{-4}\left(\left|\mathcal{F}_f(\omega)\right|+O_{\mathbb{P}}\left(T^{-\frac{1}{2}}\right)\right)^2 \\
&\quad+4\sigma_{\max}^4\sigma_{\min}^{-5}C\left(\left|\mathcal{F}_f(\omega)\right|+O_{\mathbb{P}}\left(T^{-\frac{1}{2}}\right)\right)^2 \\
&\quad+4\pi\|\omega\|\sqrt{d_2+d_3}\,\sigma_{\max}^4\sigma_{\min}^{-4}C^\tau C_f\widetilde{C}_f\left(\left|\mathcal{F}_f(\omega)\right|+O_{\mathbb{P}}\left(T^{-\frac{1}{2}}\right)\right)
\end{aligned}$$

Finally, we arrive at the integrable upper bound

$$
\begin{aligned}
&\left\| \left\| \operatorname{grad}_{(\phi,\alpha)} \Im \left[\boldsymbol{b}_{\omega}^{t,0}(\theta_0;\phi,\alpha) F_{\omega}^{t}(\theta_0;\phi,\alpha) \overline{F_{\omega}^{t'}(\theta_0;\phi,\alpha)} \right] \Big|_{(\phi,\alpha)=(\hat{\phi}_T^{\dagger},\hat{\alpha}_T^{\dagger})} \right\| \right\|_1 \\
= &\left\| \left\| \Im \left[\boldsymbol{G}_{\omega}^{t,0}(\theta_0;\hat{\phi}_T^{\dagger},\hat{\alpha}_T^{\dagger}) F_{\omega}^{t}(\theta_0;\phi,\alpha) \overline{F_{\omega}^{t'}(\theta_0;\phi,\alpha)} \right. \right. \right. \\
&\left. \left. \left. + \boldsymbol{b}_{\omega}^{t,0}(\theta_0;\hat{\phi}_T^{\dagger},\hat{\alpha}_T^{\dagger}) \operatorname{grad}_{(\phi,\alpha)} \left(F_{\omega}^{t}(\theta_0;\phi,\alpha) \overline{F_{\omega}^{t'}(\theta_0;\phi,\alpha)} \right) \Big|_{(\phi,\alpha)=(\hat{\phi}_T^{\dagger},\hat{\alpha}_T^{\dagger})} \right] \right\| \right\|_1 \\
\leq &\left(\widetilde{C} \|\omega\| + \widetilde{C}^2 \|\omega\|^2 \right) \sigma_{\max}^4 \sigma_{\min}^{-4} \left(\left| \mathcal{F}_f(\omega) \right| + O_{\mathbb{P}}\left(T^{-\frac{1}{2}} \right) \right)^2 \\
&+ 2\widetilde{C}^2 \|\omega\|^2 \sqrt{d_1(d_2+d_3)} \sigma_{\max}^4 \sigma_{\min}^{-4} \left(\left| \mathcal{F}_f(\omega) \right| + O_{\mathbb{P}}\left(T^{-\frac{1}{2}} \right) \right)^2 \\
&+ 4\widetilde{C} \|\omega\| \sqrt{d_1} \sigma_{\max}^4 \sigma_{\min}^{-5} C \left(\left| \mathcal{F}_f(\omega) \right| + O_{\mathbb{P}}\left(T^{-\frac{1}{2}} \right) \right)^2 \\
&+ 4\pi\widetilde{C} \|\omega\|^2 \sqrt{d_1(d_2+d_3)} \sigma_{\max}^4 \sigma_{\min}^{-4} C^{\tau} C_f \widetilde{C}_f \left(\left| \mathcal{F}_f(\omega) \right| + O_{\mathbb{P}}\left(T^{-\frac{1}{2}} \right) \right) \\
= &\, C_1 \|\omega\| \left(\left| \mathcal{F}_f(\omega) \right| + O_{\mathbb{P}}\left(T^{-\frac{1}{2}} \right) \right)^2 + C_2 \|\omega\|^2 \left(\left| \mathcal{F}_f(\omega) \right| + O_{\mathbb{P}}\left(T^{-\frac{1}{2}} \right) \right)^2 \\
&+ C_3 \|\omega\|^2 \left(\left| \mathcal{F}_f(\omega) \right| + O_{\mathbb{P}}\left(T^{-\frac{1}{2}} \right) \right),
\end{aligned}
$$

for some constants $C_1, C_2, C_3 > 0$. By Assumption 2.13 (A3) and Lemma A.10, $\|\omega\|^2 \left|\mathcal{F}_f(\omega)\right|$ is integrable. Since $H^2(\mathbb{R}^2) \subset H^1(\mathbb{R}^2) \subset H^{1/2}(\mathbb{R}^2)$ by the Sobolev embedding theorem (Theorem B.11), $\|\omega\|^2 \left|\mathcal{F}_f(\omega)\right|^2$ and $\|\omega\| \left|\mathcal{F}_f(\omega)\right|^2$ are integrable as well. Since $r_T^4/T = o(1)$ by Assumption 2.15 (C1), the remaining terms vanish. This yields the integrability of the upper bound needed to apply the dominated convergence theorem (Amann and Escher, 2001, Theorem 3.12) and hence, it finishes the proof of the convergence of the mixed derivatives of A_T.

Next, we tackle the mixed derivatives of C_T. Similarly to before, we get for $\widetilde{C}$ from Lemma 6.7, using Lemma A.5, that

$$
\begin{aligned}
&\left\| \operatorname{grad}_{(\phi,\alpha)} \operatorname{grad}_{\theta}^{\top} C_T(\theta_0;\phi,\alpha) \right\| \\
= &\left\| \operatorname{grad}_{(\phi,\alpha)} \operatorname{grad}_{\theta}^{\top} \int_{\Omega} \left(\frac{\beta_T}{T} \right)^2 \sum_{t,t' \in \mathbb{T}} h_{\omega}^{0,t}(\theta;\phi,\alpha)^{-1} V_T^{t}(\omega;\phi,\alpha) \right. \\
&\left. \qquad \cdot \overline{h_{\omega}^{0,t'}(\theta;\phi,\alpha)^{-1} V_T^{t'}(\omega;\phi,\alpha)} \mathrm{d}\omega \Big|_{\theta=\theta_0} \right\| \\
= &\left\| 2 \int_{\Omega_T} \left(\frac{\beta_T}{T} \right)^2 \sum_{t,t' \in \mathbb{T}} \operatorname{grad}_{(\phi,\alpha)} \Im \left[\boldsymbol{b}_{\omega}^{t,0}(\theta_0;\phi,\alpha)^{\top} h_{\omega}^{0,t'}(\theta_0;\phi,\alpha) V_T^{t}(\omega;\phi,\alpha) \right. \right. \\
&\left. \left. \qquad \cdot \overline{h_{\omega}^{0,t}(\theta_0;\phi,\alpha) V_T^{t'}(\omega;\phi,\alpha)} \right] \mathrm{d}\omega \right\| \\
\leq &\, 2 \int_{\Omega_T} \left(\frac{\beta_T}{T} \right)^2 \sum_{t,t' \in \mathbb{T}} \left\| \operatorname{grad}_{(\phi,\alpha)} \Im \left[\boldsymbol{b}_{\omega}^{t,0}(\theta_0;\phi,\alpha)^{\top} h_{\omega}^{0,t'}(\theta_0;\phi,\alpha) V_T^{t}(\omega;\phi,\alpha) \right. \right. \\
&\left. \left. \qquad \cdot \overline{h_{\omega}^{0,t}(\theta_0;\phi,\alpha) V_T^{t'}(\omega;\phi,\alpha)} \right] \right\| \mathrm{d}\omega.
\end{aligned}
$$

It remains to show that uniformly in (ϕ, α), ω and in t and t'

$$\left\| \operatorname{grad}_{(\phi,\alpha)} \Im \left[\boldsymbol{b}_{\omega}^{t,0}(\theta_0; \phi, \alpha)^\top \frac{h_{\omega}^{0,t'}(\theta_0; \phi, \alpha)}{h_{\omega}^{0,t}(\theta_0; \phi, \alpha)} V_T^t(\omega; \phi, \alpha) \overline{V_T^{t'}(\omega; \phi, \alpha)} \right] \right\| \xrightarrow{T\to\infty} 0 \text{ in probability.} \tag{6.100}$$

Since the convergence is uniformly in ω, the integral will vanish asymptotically as well. First, we note that

$$\begin{aligned} \left| V_T^t(\omega; \phi, \alpha) \right| &= \left| \frac{1}{2n\sqrt{\beta_T}(\sigma_t^\alpha)^2} \sum_{j=1}^n \exp\left(-2\pi \mathrm{i} \left\langle 1/\sigma_t^\alpha \cdot R_{\rho_t^\phi} \omega, x_j \right\rangle \right) \epsilon_j^t \right| \\ &\le \frac{1}{2\sigma_{\min}^2 \sqrt{\beta_T}} \sum_{j=1}^n \left| \epsilon_j^t \right| = O_{\mathbb{P}}\left(\beta_T^{-\frac{1}{2}} \right) \end{aligned} \tag{6.101}$$

With C from Assumption 2.14 (B2) and r_T from Definition 2.22 we have using the norm equivalence that

$$\begin{aligned} &\left\| \operatorname{grad}_{(\phi,\alpha)} \left\langle 1/\sigma_t^\alpha \cdot R_{\rho_t^\phi} \omega, x_j \right\rangle \right\| \\ =& \left\| \operatorname{grad}_{(\phi,\alpha)} \left(\frac{1}{\sigma_t^\alpha} \left(\left(\cos\left(\rho_t^\phi\right) \omega_1 - \sin\left(\rho_t^\phi\right) \omega_2 \right) x_{j,1} \right.\right.\right. \\ & \qquad \left.\left.\left. + \left(\sin\left(\rho_t^\phi\right) \omega_1 + \cos\left(\rho_t^\phi\right) \omega_2 \right) x_{j,2} \right) \right) \right\| \\ \le& \frac{1}{\sigma_t^\alpha} \left\| \operatorname{grad}_\phi \rho_t^\phi \right\|_1 \left| \left(-\sin\left(\rho_t^\phi\right) \omega_1 - \cos\left(\rho_t^\phi\right) \omega_2 \right) x_{j,1} + \left(\cos\left(\rho_t^\phi\right) \omega_1 - \sin\left(\rho_t^\phi\right) \right) x_{j,2} \right| \\ &+ \frac{1}{\left(\sigma_t^\alpha\right)^2} \left\| \operatorname{grad}_\alpha \sigma_t^\alpha \right\|_1 \left| \left(\cos\left(\rho_t^\phi\right) \omega_1 - \sin\left(\rho_t^\phi\right) \omega_2 \right) x_{j,1} \right. \\ & \qquad \left. + \left(\sin\left(\rho_t^\phi\right) \omega_1 + \cos\left(\rho_t^\phi\right) \omega_2 \right) x_{j,2} \right| \\ \le& \sqrt{2} \sqrt{d_2 + d_3 C} \left\| \omega \right\| \left\| x_j \right\|_1 \left(\sigma_{\min}^{-2} + \sigma_{\min}^{-1} \right) \\ =& O_{\mathbb{P}}\left(r_T \right), \end{aligned}$$

since $\left\| x_j \right\|_1 \le 2$ for all $j = 1, \dots, n$. From this, it follows that, using (6.101) we get

$$\begin{aligned} &\left\| \operatorname{grad}_{(\phi,\alpha)} V_T^t(\omega; \phi, \alpha) \right\| \\ =& \left\| \frac{-2\operatorname{grad}_{(\phi,\alpha)} \sigma_t^\alpha}{\sigma_t^\alpha} V_T^t(\omega; \phi, \alpha) + V_T^t(\omega; \phi, \alpha) \operatorname{grad}_{(\phi,\alpha)} \left(-2\pi \mathrm{i} \left\langle (\sigma_t^\alpha)^{-1} R_{\rho_t^\phi} \omega, x_j \right\rangle \right) \right\| \\ \le& \frac{2}{\sigma_t^\alpha} \left\| \operatorname{grad}_\alpha \sigma_t^\alpha \right\|_1 \left| V_T^t(\omega; \phi, \alpha) \right| + \frac{2\pi\lambda}{\sqrt{\beta_T} \left(\sigma_t^\alpha\right)^2} \sum_{j \in J_T} \left| \epsilon_j^t \right| \left\| \operatorname{grad}_{(\phi,\alpha)} \left\langle \left(\sigma_t^\alpha\right)^{-1} R_{\rho_t^\phi} \omega, x_j \right\rangle \right\| \\ \le& \frac{2\sqrt{d_2 + d_3} C\lambda}{\sigma_{\min}^3 \sqrt{\beta_T}} \sum_{j=1}^n \left| \epsilon_j^t \right| + \frac{2\pi\lambda C' r_T}{\sigma_{\min}^2 \sqrt{\beta_T}} \sum_{j=1}^n \left| \epsilon_j^t \right| \\ \le& \frac{2\lambda r_T}{\sigma_{\min}^2 \sqrt{\beta_T}} \left(\frac{\sqrt{d_2 + d_3} C}{\sigma_{\min}} + \pi C' \right) \sum_{j=1}^n \left| \epsilon_j^t \right| = O_{\mathbb{P}}\left(r_T \beta_T^{-\frac{1}{2}} \right), \end{aligned} \tag{6.102}$$

which implies

$$\begin{aligned}\left\|\mathrm{grad}_{(\phi,\alpha)}\left(V_T^t(\omega;\phi,\alpha)\overline{V_T^{t'}(\omega;\phi,\alpha)}\right)\right\| &\leq 2\left\|\mathrm{grad}_{(\phi,\alpha)}V_T^t(\omega;\phi,\alpha)\right\|\left|V_T^t(\omega;\phi,\alpha)\right| \\ &= O_{\mathbb{P}}\left(r_T\beta_T^{-1}\right) \qquad (6.103)\end{aligned}$$

Hence, using again Lemma 6.7, $\left|h_\omega^{t,t'}(\theta;\phi,\alpha)\right| = 1$ and the bounds obtained in (6.101) and (6.103), we can conclude that for $\omega \in \Omega_T$,

$$\begin{aligned}&\left\|\mathrm{grad}_{(\phi,\alpha)}\mathfrak{I}\left[\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi,\alpha)^\top h_\omega^{0,t'}(\theta_0;\phi,\alpha)h_\omega^{0,t}(\theta_0;\phi,\alpha)^{-1}V_T^t(\omega;\phi,\alpha)\overline{V_T^{t'}(\omega;\phi,\alpha)}\right]\right\|_1 \\ =&\left\|\mathfrak{I}\left[\boldsymbol{G}_\omega^{t,0}(\theta_0;\phi,\alpha)^\top h_\omega^{0,t'}(\theta_0;\phi,\alpha)h_\omega^{0,t}(\theta_0;\phi,\alpha)^{-1}V_T^t(\omega;\phi,\alpha)\overline{V_T^{t'}(\omega;\phi,\alpha)}\right.\right. \\ &+\mathrm{i}\boldsymbol{c}_\omega^{t,t'}(\theta_0;\phi,\alpha)\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi,\alpha)^\top h_\omega^{0,t'}(\theta_0;\phi,\alpha)h_\omega^{0,t}(\theta_0;\phi,\alpha)^{-1}V_T^t(\omega;\phi,\alpha)\overline{V_T^{t'}(\omega;\phi,\alpha)} \\ &\left.\left.+\ \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi,\alpha)^\top h_\omega^{0,t'}(\theta_0;\phi,\alpha)h_\omega^{0,t}(\theta_0;\phi,\alpha)^{-1}\mathrm{grad}_{(\phi,\alpha)}\left(V_T^t(\omega;\phi,\alpha)\overline{V_T^{t'}(\omega;\phi,\alpha)}\right)\right]\right\|_1 \\ =&O_{\mathbb{P}}\left(r_T\beta_T^{-1}\right)+2O_{\mathbb{P}}\left(r_T^2\beta_T^{-1}\right)\overset{T\to\infty}{\longrightarrow}0 \text{ in probability,}\end{aligned}$$

since $r_T^2 = o(\beta_T)$ by Assumption 2.15 (C1).

Finally, consider the mixed derivatives of B_T. Note that for $\widetilde{C}_f$ from (6.97)

$$\begin{aligned}\left|\mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}(\omega)\right)\right| &\leq \int_{\mathbb{R}^2}\left|\exp\left(-2\pi\mathrm{i}\left\langle\tau_t^{(\phi,\alpha)}(\omega),x\right\rangle\right)f(x)\right|\mathrm{d}x \\ &= \int_{\mathbb{R}^2}|f(x)|\,\mathrm{d}x \leq \widetilde{C}_f. \qquad (6.104)\end{aligned}$$

By the same arguments as above, the first factor of B_T converges to some deterministic term while the second factor converges to zero. More specifically, we have for the gradient of the integrand of B_T that

$$\begin{aligned}&\mathrm{grad}_\theta\mathfrak{R}\left[\left(\sum_{t\in\mathbb{T}}F_\omega^t(\theta;\phi,\alpha)\right)\left(\sum_{t'\in\mathbb{T}}\overline{h_\omega^{0,t'}(\theta;\phi,\alpha)^{-1}V_T^{t'}(\omega;\phi,\alpha)}\right)\right] \\ &=\mathfrak{R}\left[\sum_{t,t'\in\mathbb{T}}\mathrm{grad}_\theta^\top h_\omega^{t,t'}(\theta;\phi,\alpha)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha}\right)^2\mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega)\overline{V_T^{t'}(\omega;\phi,\alpha)}\right] \\ &=\mathfrak{R}\left[\sum_{t,t'\in\mathbb{T}}\mathrm{i}\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)^\top h_\omega^{t,t'}(\theta;\phi,\alpha)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha}\right)^2\mathcal{F}_f(\tau_t^{(\phi,\alpha)}\omega)\overline{V_T^{t'}(\omega;\phi,\alpha)}\right].\end{aligned}$$

Applying also the gradient with respect to (ϕ,α), we obtain for $\omega \in \Omega_T$

$$\begin{aligned}&\mathrm{grad}_{(\phi,\alpha)}\mathfrak{R}\left[\left(\frac{\beta_T}{T}\right)^2\sum_{t,t'\in\mathbb{T}}\mathrm{i}\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)^\top h_\omega^{t,t'}(\theta;\phi,\alpha)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha}\right)^2\mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right)\overline{V_T^{t'}(\omega;\phi,\alpha)}\right] \\ =&\mathfrak{R}\left[\left(\frac{\beta_T}{T}\right)^2\sum_{t,t'\in\mathbb{T}}\mathrm{i}\boldsymbol{G}_\omega^{t,t'}(\theta;\phi,\alpha)^\top h_\omega^{t,t'}(\theta;\phi,\alpha)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha}\right)^2\mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right)\overline{V_T^{t'}(\omega;\phi,\alpha)}\right. \\ &-\ \boldsymbol{c}_\omega^{t,t'}(\theta;\phi,\alpha)\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)^\top h_\omega^{t,t'}(\theta;\phi,\alpha)\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^\alpha}\right)^2\mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right)\overline{V_T^{t'}(\omega;\phi,\alpha)}\end{aligned}$$

$$
\begin{aligned}
&+ \mathrm{i} b_\omega^{t,t'}(\theta;\phi,\alpha)^\top h_\omega^{t,t'}(\theta;\phi,\alpha) \mathrm{grad}_{(\phi,\alpha)} \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\alpha}}\right)^2 \mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right) \overline{V_T^{t'}(\omega;\phi,\alpha)} \\
&+ \mathrm{i} b_\omega^{t,t'}(\theta;\phi,\alpha)^\top h_\omega^{t,t'}(\theta;\phi,\alpha) \left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\alpha}}\right)^2 \mathrm{grad}_{(\phi,\alpha)} \mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right) \overline{V_T^{t'}(\omega;\phi,\alpha)} \\
&+ \mathrm{i} b_\omega^{t,t'}(\theta;\phi,\alpha)^\top h_\omega^{t,t'}(\theta;\phi,\alpha) \mathcal{F}_f\left(\tau_t^{(\phi,\alpha)}\omega\right) \overline{\mathrm{grad}_{(\phi,\alpha)} V_T^{t'}(\omega;\phi,\alpha)}\Bigg] \\
&= O_{\mathbb{P}}\left(r_T\beta_T^{-\frac{1}{2}}\right) + O_{\mathbb{P}}\left(r_T^2\beta_T^{-\frac{1}{2}}\right) + O_{\mathbb{P}}\left(r_T\beta_T^{-\frac{1}{2}}\right) + O_{\mathbb{P}}\left(r_T^2\beta_T^{-\frac{1}{2}}\right) + O_{\mathbb{P}}\left(r_T^2\beta_T^{-1}\right) \\
&= O_{\mathbb{P}}\left(r_T^2\beta_T^{-\frac{1}{2}}\right),
\end{aligned} \tag{6.105}
$$

using $\left|h_\omega^{t,t'}(\theta;\phi,\alpha)\right| = 1$ and the bounds from Assumption 2.14 (B2), Lemma 6.7, (6.104) and (6.97), as well as (6.101) and (6.102). Thus, the norm of the mixed derivatives of B_T is bounded by

$$
\left\|\mathrm{grad}_{(\phi,\alpha)}\mathrm{grad}_\theta B_T(\theta;\phi,\alpha)\right\|_1 = \int_{\Omega_T} O_{\mathbb{P}}\left(r_T^2\beta_T^{-\frac{1}{2}}\right) \mathrm{d}\omega = O_{\mathbb{P}}\left(r_T^4\beta_T^{-\frac{1}{2}}\right),
$$

which converges to zero because $r_T^4 = o(\sqrt{\beta_T})$ as a consequence of Assumption 2.15 (C1). This concludes the proof of the second statement and finishes the proof of the convergence of the mixed derivatives of the empirical contrast functional,

$$
\mathrm{grad}_{(\phi,\alpha)}\mathrm{grad}_\theta^\top N_T(\theta_0;\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger) \overset{T\to\infty}{\longrightarrow} \mathrm{grad}_{(\phi,\alpha)}\mathrm{grad}_\theta^\top N(\theta_0;\phi_0,\alpha_0) \text{ in probability.} \qquad \square
$$

Theorem 6.14. *Suppose that Assumptions 2.13 (A3), 2.14 (B2-B3) and 2.15 hold. Then* $\mathrm{grad}_\theta N_T(\theta_0;\phi_0,\alpha_0)$ *is asymptotically normally distributed:*

$$
\sqrt{T}\mathrm{grad}_\theta N_T(\theta_0;\phi_0,\alpha_0) \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0,\widetilde{\Sigma}),
$$

for some covariance matrix $\widetilde{\Sigma}$ *stated explicitly in the proof below.*

Proof. As before, we use the decompostion $N_T(\theta) := N_T(\theta;\phi_0,\alpha_0) = A_T(\theta)+B_T(\theta)+C_T(\theta)$ from (2.23), where plugging in ϕ_0 and α_0, using $\tau_t^{(\phi_0,\alpha_0)}\omega = \omega$ and denoting $V_T^t(\omega) := V_T^t(\omega;\phi_0,\alpha_0)$ yields the following simplified expressions for the three summands:

$$
\begin{aligned}
A_T(\theta) &:= -\int_{\Omega_T} \left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} F_\omega^t(\theta;\phi_0,\alpha_0)\right|^2 \mathrm{d}\omega = -\int_{\Omega_T} \left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\mathcal{F}_f(\omega)\right|^2 \mathrm{d}\omega, \\
B_T(\theta) &:= -\int_{\Omega_T} 2\Re\left[\left(\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} F_\omega^t(\theta;\phi_0,\alpha_0)\right)\left(\frac{\beta_T}{T}\sum_{t'\in\mathbb{T}} \overline{h_\omega^{0,t'}(\theta;\phi_0,\alpha_0)^{-1}V_T^{t'}(\omega)}\right)\right] \mathrm{d}\omega \\
&= -\int_{\Omega_T} 2\Re\left[\left(\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\mathcal{F}_f(\omega)\right)\left(\frac{\beta_T}{T}\sum_{t'\in\mathbb{T}} \overline{h_\omega^{0,t'}(\theta;\phi_0,\alpha_0)^{-1}V_T^{t'}(\omega)}\right)\right] \mathrm{d}\omega, \\
C_T(\theta) &:= -\int_{\Omega_T} \left|\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} h_\omega^{0,t}(\theta;\phi_0,\alpha_0)^{-1}V_T^t(\omega)\right|^2 \mathrm{d}\omega.
\end{aligned}
$$

Similarly to before we devide the proof into several steps. The aim is to show that

$$\sqrt{T}\mathrm{grad}_\theta A_T(\theta_0) = 0, \tag{6.106}$$

$$\sqrt{T}\mathrm{grad}_\theta C_T(\theta_0) \overset{T\to\infty}{\longrightarrow} 0 \text{ in probability}, \tag{6.107}$$

$$\text{and } \sqrt{T}\mathrm{grad}_\theta B_T(\theta_0) \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0,\widetilde{\Sigma}). \tag{6.108}$$

Let

$$S_j^{t'} = 2\Im\left[\int_{\mathbb{R}^2}\int_0^1 G_j^{t'}(\omega)\boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)h_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\mathrm{d}t\,\mathrm{d}\omega\right],$$

with $G_j^{t'}(\omega)$ defined by

$$G_j^{t'}(\omega) \coloneqq \left(\sigma_{t'}^{\alpha_0}\right)^{-2}\exp\left(2\pi\mathrm{i}\left\langle \frac{1}{\sigma_{t'}^{\alpha_0}}R_{\rho_{t'}^{\phi_0}}\omega, x_j\right\rangle\right)\mathcal{F}_f(\omega). \tag{6.109}$$

The covariance matrix $\widetilde{\Sigma}$ is then given by

$$\widetilde{\Sigma} = \int_0^1 \frac{1}{n^2}\sum_{j=1}^n S_j^{t'}(S_j^{t'})^T\,\mathrm{d}t'. \tag{6.110}$$

Note that $\widetilde{\Sigma}$ has finite operatornorm, since by Lemma 6.7 and Lemma A.10,

$$\begin{aligned}&\left|\int_{\mathbb{R}^2}\int_0^1 G_j^{t'}(\omega)\boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)h_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\mathrm{d}t\,\mathrm{d}\omega\right|\\ \le&\int_{\mathbb{R}^2}\left|\mathcal{F}_f(\omega)\right|\|\omega\|\,\mathrm{d}\omega<\infty.\end{aligned}$$

First, consider the gradient of A_T. By Lemma A.5, Lemma 6.7, and since $h_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)=1$, it holds that

$$\begin{aligned}&\sqrt{T}\mathrm{grad}_\theta A_T(\theta)\big|_{\theta=\theta_0}\\ =&-\sqrt{T}\int_{\Omega_T}\left(\frac{\beta_T}{T}\right)^2\mathrm{grad}_\theta\left[\sum_{t,t'\in\mathbb{T}}h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\mathcal{F}_f(\omega)\overline{h_\omega^{t',0}(\theta;\phi_0,\alpha_0)\mathcal{F}_f(\omega)}\right]\Bigg|_{\theta=\theta_0}\mathrm{d}\omega\\ =&-\sqrt{T}\int_{\Omega_T}\left(\frac{\beta_T}{T}\right)^2 2\Re\left[\sum_{t,t'\in\mathbb{T}}\mathrm{grad}_\theta h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\big|_{\theta=\theta_0}\left|\mathcal{F}_f(\omega)\right|^2\right]\mathrm{d}\omega\\ =&-\sqrt{T}\int_{\Omega_T}\left(\frac{\beta_T}{T}\right)^2 2\Re\left[\sum_{t,t'\in\mathbb{T}}\mathrm{i}\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\left|\mathcal{F}_f(\omega)\right|^2\right]\mathrm{d}\omega\\ =&\sqrt{T}\int_{\Omega_T}\left(\frac{\beta_T}{T}\right)^2 2\sum_{t,t'\in\mathbb{T}}\Im\left[\boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\left|\mathcal{F}_f(\omega)\right|^2\right]\mathrm{d}\omega\\ =&\,0,\end{aligned}$$

proving (6.106). Next, we show the convergence of the gradient of C_T. For the integrand we can calculate using Lemma 6.7 and Lemma 6.3

$$
\begin{aligned}
&\mathbb{E}\left\| -\mathrm{grad}_\theta \left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} h^{0,t}_\omega(\theta;\phi_0,\alpha_0)^{-1}\, V^t_T(\omega)\overline{h^{0,t'}_\omega(\theta;\phi_0,\alpha_0)^{-1}\, V^{t'}_T(\omega)}\bigg|_{\theta=\theta_0} \right\| \\
=&\mathbb{E}\left\| 2\Re\left[\left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} \mathrm{i}\boldsymbol{b}^{0,t}_\omega(\theta_0;\phi_0,\alpha_0) h^{0,t}_\omega(\theta_0;\phi_0,\alpha_0)^{-1}\, V^t_T(\omega)\overline{h^{0,t'}_\omega(\theta_0;\phi_0,\alpha_0)^{-1}\, V^{t'}_T(\omega)}\right]\right\| \\
\le& 2\widetilde{C}\,\|\omega\| \left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} \mathbb{E}\left| V^t_T(\omega)\overline{V^{t'}_T(\omega)}\right| \\
=& 2\widetilde{C}\,\|\omega\| \left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} O\left(\frac{1}{\beta_T}\right) \\
=& O\left(\frac{r_T}{\beta_T}\right),
\end{aligned}
$$

for $\omega \in \Omega_T$, with $\widetilde{C}$ from Lemma 6.7. Hence, for the gradient of C_T it holds that

$$
\mathbb{E}\left\| \sqrt{T}\mathrm{grad}_\theta C_T(\theta_0)\right\| \le \sqrt{T}\int_{\Omega_T} O\left(r_T\beta_T^{-1}\right)\mathrm{d}\omega = O\left(\sqrt{T} r_T^3 \beta_T^{-1}\right).
$$

By Assumption 2.15 (C2), $\sqrt{T} r_T^3\beta_T^{-1} \to 0$ as $T\to\infty$ and hence, (6.107) follows.

Finally, we tackle the asymptotic behavior of the remaining term $\sqrt{T}\mathrm{grad}_\theta B_T(\theta_0)$. For the gradient we obtain

$$
\begin{aligned}
&\mathrm{grad}_\theta B_T(\theta_0) \\
&= -\int_{\Omega_T} 2\Re\left[\left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} \mathrm{grad}_\theta h^{t,t'}_\omega(\theta;\phi_0,\alpha_0)\bigg|_{\theta=\theta_0} \mathcal{F}_f(\omega)\,\overline{V^{t'}_T(\omega)}\right]\mathrm{d}\omega \\
&= -\int_{\Omega_T} 2\Re\left[\left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} \mathrm{i}\boldsymbol{b}^{t,t'}_\omega(\theta_0;\phi_0,\alpha_0) h^{t,t'}_\omega(\theta_0;\phi_0,\alpha_0)\,\mathcal{F}_f(\omega)\,\overline{V^{t'}_T(\omega)}\right]\mathrm{d}\omega \\
&= \int_{\Omega_T} 2\left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} \Im\left[\boldsymbol{b}^{t,t'}_\omega(\theta_0;\phi_0,\alpha_0) h^{t,t'}_\omega(\theta_0;\phi_0,\alpha_0)\,\mathcal{F}_f(\omega)\,\overline{V^{t'}_T(\omega)}\right]\mathrm{d}\omega \\
&= \int_{\Omega_T} 2\left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} \Im\Bigg[\boldsymbol{b}^{t,t'}_\omega(\theta_0;\phi_0,\alpha_0) h^{t,t'}_\omega(\theta_0;\phi_0,\alpha_0)\,\mathcal{F}_f(\omega) \\
&\qquad \cdot\frac{1}{2n\sqrt{\beta_T}(\sigma^{\alpha_0}_{t'})^2}\sum_{j=1}^n \exp\left(2\pi\mathrm{i}\left\langle 1/\sigma^{\alpha_0}_{t'}\cdot R_{\rho^{\phi_0}_{t'}}\omega, x_j\right\rangle\right)\epsilon^{t'}_j\Bigg]\mathrm{d}\omega \\
&= \frac{\sqrt{\beta_T}}{T}\sum_{t'\in\mathbb{T}}\frac{1}{n}\sum_{j=1}^n \epsilon^{t'}_j\,\Im\left[\int_{\Omega_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}} \boldsymbol{b}^{t,t'}_\omega(\theta_0;\phi_0,\alpha_0) h^{t,t'}_\omega(\theta_0;\phi_0,\alpha_0)\,G^{t'}_j(\omega)\,\mathrm{d}\omega\right],
\end{aligned}
$$

with $G^{t'}_j$ from 6.109.

For $\xi \in \mathbb{R}^{d_1}$, we get that

$$\begin{aligned}&\left\langle \xi,\ \sqrt{T}\mathrm{grad}_\theta B_T(\theta_0)\right\rangle\\ &= \sqrt{\frac{\beta_T}{T}}\sum_{t'\in\mathbb{T}}\frac{1}{n}\sum_{j=1}^n \epsilon_j^{t'}\,\Im\left[\int_{\Omega_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\left\langle \xi, \boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\right\rangle h_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\,G_j^{t'}(\omega)\,\mathrm{d}\omega\right]\end{aligned}$$

is a linear combination of the independent standard normal random variables $\epsilon_j^{t'}$ and thus, itself a centered Gaussian random variable with variance

$$\frac{\beta_T}{T}\sum_{t'\in\mathbb{T}}\frac{1}{n^2}\sum_{j=1}^n \Im\left[\int_{\Omega_T}\frac{\beta_T}{T}\sum_{t\in\mathbb{T}}\left\langle \xi, \boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\right\rangle h_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\,G_j^{t'}(\omega)\,\mathrm{d}\omega\right]^2,$$

converging to

$$\int_0^1\frac{1}{n^2}\sum_{j=1}^n \Im\left[\int_{\mathbb{R}^2}\int_0^1\left\langle \xi, \boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\right\rangle h_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)\,G_j^{t'}(\omega)\,\mathrm{d}t\,\mathrm{d}\omega\right]^2\mathrm{d}t',$$

as $T\to\infty$. This can be written as $\xi^T\widetilde{\Sigma}\xi$ with $\widetilde{\Sigma}$ from (6.110). Cramér-Wold's Theorem (Theorem B.5) yields the claimed convergence in distribution of the gradient of B_T. Now, (6.106), (6.107) and (6.108) together with Slutzky's Lemma (Theorem B.9) prove the third statement, namely that $\mathrm{grad}_\theta N_T(\theta_0;\phi_0,\alpha_0)$ can be written as a linear combination of the errors ϵ_j^t plus some $o_{\mathbb{P}}(1)$ term. □

Lemma 6.15. *Under Assumptions 2.13 (A2-A3) and 2.14 (B2-B3),* $\mathrm{Hess}_\theta N(\theta;\phi_0,\alpha_0)$ *has finite operator norm for all* $\theta\in U'$ *with* $U'\subset\Theta$ *from Assumption 2.14 (B2). Moreover,*

$$H_N := \mathrm{Hess}_\theta N(\theta_0;\phi_0,\alpha_0) \tag{6.111}$$

is symmetric. If the Assumptions 2.14 (B1, B4, and B8) hold, H_N *is also positive definite and thus, invertible.*

Proof. By Lemma 6.7, Lemma A.6, and Theorem B.4 on the differentiability of parameter integrals, we get

$$\begin{aligned}&\mathrm{Hess}_\theta N(\theta;\phi_0,\alpha_0)\\ &= -\mathrm{Hess}_\theta\int_{\mathbb{R}^2}\left|\mathcal{F}_f(\omega)\right|^2\int_0^1\int_0^1 h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\overline{h_\omega^{t',0}(\theta;\phi_0,\alpha_0)}\mathrm{d}t\mathrm{d}t'\mathrm{d}\omega\\ &= -2\int_{\mathbb{R}^2}\left|\mathcal{F}_f(\omega)\right|^2\int_0^1\int_0^1\Re\left(\overline{h_\omega^{t',0}(\theta;\phi_0,\alpha_0)}\,\mathrm{Hess}_\theta h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\right.\\ &\quad\left.+\ \mathrm{grad}_\theta^\top h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\overline{\mathrm{grad}_\theta h_\omega^{t',0}(\theta;\phi_0,\alpha_0)}\right)\mathrm{d}t\mathrm{d}t'\mathrm{d}\omega\\ &= -2\int_{\mathbb{R}^2}\left|\mathcal{F}_f(\omega)\right|^2\int_0^1\int_0^1\Re\left(h_\omega^{t,0}(\theta;\phi_0,\alpha_0)\overline{h_\omega^{t',0}(\theta;\phi_0,\alpha_0)}\right.\end{aligned}$$

$$\cdot \left[\mathrm{i}\boldsymbol{H}_\omega^{t,0}(\theta;\phi,\alpha) - \boldsymbol{b}_\omega^{t,0}(\theta;\phi_0,\alpha_0)\left(\boldsymbol{b}_\omega^{t,0}(\theta;\phi_0,\alpha_0) - \boldsymbol{b}_\omega^{t',0}(\theta;\phi_0,\alpha_0)\right)^\top\right]\Bigg)\,\mathrm{d}t\mathrm{d}t'\mathrm{d}\omega. \quad (6.112)$$

Let $\xi \in \mathbb{R}^{d_1}$ with $\|\xi\| = 1$. Then, because $f \in H^2(\mathbb{R}^2)$ by Assumption 2.13 (A3), using the equivalence of norms and the Sobolev embedding theorem, we have with $\widetilde{C}$ from Lemma 6.7,

$$\begin{aligned}
&\|\operatorname{Hess}_\theta N(\theta;\phi_0,\alpha_0)\xi\| \\
&\le 2\int_{\mathbb{R}^2} \left|\mathcal{F}_f(\omega)\right|^2 \int_0^1\int_0^1 \left\|\boldsymbol{H}_\omega^{t,0}(\theta;\phi,\alpha)\right\|_1 + \left\|\boldsymbol{b}_\omega^{t,0}(\theta;\phi_0,\alpha_0)\right\| \\
&\qquad \cdot \left(\left\|\boldsymbol{b}_\omega^{t,0}(\theta;\phi_0,\alpha_0)\right\| + \left\|\boldsymbol{b}_\omega^{t',0}(\theta;\phi_0,\alpha_0)\right\|\right)\mathrm{d}t\mathrm{d}t'\mathrm{d}\omega \\
&\le 2\int_{\mathbb{R}^2} \left|\mathcal{F}_f(\omega)\right|^2 \left(\widetilde{C}\,\|\omega\| + 2\widetilde{C}^2\,\|\omega\|^2\right)\mathrm{d}\omega < \infty.
\end{aligned}$$

Thus, $\operatorname{Hess}_\theta(\theta_0;\phi_0,\alpha_0)$ has finite operator norm. From (6.112), we have for $\theta = \theta_0$ that

$$\begin{aligned}
H_N &= \operatorname{Hess}_\theta N(\theta_0;\phi_0,\alpha_0) \\
&= 2\int_{\mathbb{R}^2} \left|\mathcal{F}_f(\omega)\right|^2 \int_0^1\int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0) - \boldsymbol{b}_\omega^{t',0}(\theta_0;\phi_0,\alpha_0)\right)^\top \mathrm{d}t\mathrm{d}t'\mathrm{d}\omega \\
&= 2\int_{\mathbb{R}^2} \left|\mathcal{F}_f(\omega)\right|^2 \Bigg(\int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)^\top \mathrm{d}t \\
&\qquad - \left(\int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\mathrm{d}t\right)\left(\int_0^1 \boldsymbol{b}_\omega^{t',0}(\theta_0;\phi_0,\alpha_0)\mathrm{d}t'\right)^\top\Bigg)\mathrm{d}\omega. \quad (6.113)
\end{aligned}$$

Since matrices of the form $xx^\top$ with $x \in \mathbb{R}^{d_1}$ are symmetric, it follows that H_N is symmetric. Now, let $\xi \in \mathbb{R}^{d_1} \setminus \{0\}$. By Assumption 2.13 (A2) and Lemma A.9, there is an open Borel set $B \subseteq \mathbb{R}^2$ with positive Lebesgue measure such that $\mathcal{F}_f(\omega) \neq 0$ for all $\omega \in B$. Similarly to (6.74), the goal is now to show that if Assumptions 2.14 (B4, B8) hold, there is another Borel set $B' \subseteq [0,1]$ with positive Lebesgue measure such that

$$B' \to \mathbb{R}, \quad t \mapsto \left\langle \xi, \mathbf{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right\rangle$$

is not constant for all $\omega \in \mathbb{R}^2 \setminus \Omega^0$ with some Lebesgue null-set $\Omega^0 \subseteq \mathbb{R}^2$. To this end we define $S_1(t) := \sum_{m=1}^{d_1} \xi_m \frac{\partial(\delta_t^\theta)_1}{\partial\theta_m}\Big|_{\theta=\theta_0}$ and $S_2(t) := \sum_{m=1}^{d_1} \xi_m \frac{\partial(\delta_t^\theta)_2}{\partial\theta_m}\Big|_{\theta=\theta_0}$. Similarly to before we know that if S_i, $i = 1,2$ were constant a.e. they would have to be zero a.e. as $\left(\delta_t^\theta\right)_i$, $i = 1,2$, are continuous at $t = 0$ by Assumption 2.14 (B4) and take the value $\left(\delta_0^\theta\right)_i = 0$ there by Assumption 2.14 (B1). However, by Assumption 2.14 (B8) the components of the gradient of the drift function are linearly independent, which is why there has to be some Borel set B' such that S_1 is non-constant on this set, and the same holds for S_2. Let now

$$\begin{aligned}
a(t) &:= \left(\sigma_t^{\alpha_0}\right)^{-1}\left(\cos\left(\rho_t^{\phi_0}\right)S_1(t) + \sin\left(\rho_t^{\phi_0}\right)S_2(t)\right), \\
b(t) &:= \left(\sigma_t^{\alpha_0}\right)^{-1}\left(-\sin\left(\rho_t^{\phi_0}\right)S_1(t) + \cos\left(\rho_t^{\phi_0}\right)S_2(t)\right),
\end{aligned}$$

and suppose the scalar product is constant, i.e.,

$$\left\langle \xi, \mathbf{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right\rangle = 2\pi\left(\omega_1 a(t) + \omega_2 b(t)\right) = c \text{ a.e.}$$

For the same reasons as before, this can only hold for $c = 0$. For that to be fulfilled either we need $a(t) = 0$ a.e. and $b(t) = 0$ a.e. or ω_1 is determined by ω_2, which implies that $\omega \in \Omega^0$ for some Lebesgue null-set $\Omega^0 \subset \mathbb{R}$. Assume both a and b are zero a.e. Since $\sigma_t^{\alpha_0} \in [\sigma_{\min}, \sigma_{\max}]$, from $a(t) = 0$ a.e. we get that $S_2(t) = -\left(\sin\left(\rho_t^{\phi_0}\right)\right)^{-1}\cos\left(\rho_t^{\phi_0}\right) S_1(t)$ a.e. Together with $b(t) = 0$ a.e. this yields that

$$-\sin\left(\rho_t^{\phi_0}\right) S_1(t) - \cos\left(\rho_t^{\phi_0}\right)\left(\sin\left(\rho_t^{\phi_0}\right)\right)^{-1}\cos\left(\rho_t^{\phi_0}\right) S_1(t) = \left(\sin\left(\rho_t^{\phi_0}\right)\right)^{-1} S_1(t) = 0$$

a.e., leading to a contradiction since $S_1(t)$ is non-constant on B'. Hence, it follows that

$$B' \to \mathbb{R}, \quad t \mapsto \left\langle \xi, \mathbf{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right\rangle$$

is not constant for all $\omega \in \mathbb{R}^2 \setminus \Omega^0$ with some Lebesgue null-set $\Omega^0 \subseteq \mathbb{R}^2$.
Since the Cauchy-Schwarz inequality (Theorem B.1) implies that $\left(\int_0^1 g(t)\,\mathrm{d}t\right)^2 \leq \int_0^1 g(t)^2\,\mathrm{d}t$ for all integrable functions $g\colon [0,1] \to \mathbb{R}$, with equality if and only if g is constant a.e., we get that for almost all ω

$$\int_0^1 \left\langle \xi, \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right\rangle^2 \mathrm{d}t - \left(\int_0^1 \left\langle \xi, \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right\rangle \mathrm{d}t\right)^2 > 0.$$

Hence,

$$\xi^\top H_N \xi \geq 2\int_B \left|\mathcal{F}_f(\omega)\right|^2 \left[\int_0^1 \left\langle \xi, \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right\rangle^2 \mathrm{d}t - \left(\int_0^1 \left\langle \xi, \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right\rangle \mathrm{d}t\right)^2\right] \mathrm{d}\omega > 0,$$

since the integrand (as a function in ω) is strictly positive on B. We conclude that H_N is symmetric and positive definite and thus, invertible. □

Theorem 6.16. *Under the Assumption 2.14 (B2), let $\left(\hat{\theta}_T^*\right)_{T\in\mathbb{N}}$ be a sequence of random vectors with values in U' from that Assumption, such that $\hat{\theta}_T^* \xrightarrow{T\to\infty} \theta_0$ in probability. Suppose, that the Assumptions 2.13 (A3) and 2.14 (B2 -B4, B6) hold. Assume further that Assumption 2.15 is fulfilled. Then, with H_N from (6.111),*

$$\left\| \operatorname{Hess}_\theta N_T(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T) - H_N \right\|_1 \xrightarrow{T\to\infty} 0 \quad \textit{in probability.}$$

Proof. The idea is to consider again the decomposition (2.23) and show that

$$\left\| \operatorname{Hess}_\theta A_T(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T) - H_N \right\| \xrightarrow{T\to\infty} 0 \text{ in probability, and}$$

$$\sup_{\theta\in U'} \left\| \operatorname{Hess}_\theta B_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) + \operatorname{Hess}_\theta C_T(\theta;\hat{\phi}_T,\hat{\alpha}_T) \right\| \xrightarrow{T\to\infty} 0 \text{ in probability,}$$

with $U' \subset \Theta$ from Assumption 2.14 (B2).

First we tackle the Hessian of A_T. With Lemma 6.7, Lemma A.5, and $\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}$ from Definition 2.21, we get

$$
\begin{aligned}
&\mathrm{Hess}_\theta A_T(\theta;\hat{\phi}_T,\hat{\alpha}_T)\\
&=\mathrm{Hess}_\theta\Bigg[-\int_{\Omega_T}\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}}\right)^2 h_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)\left(\frac{\sigma_{t'}^{\alpha_0}}{\sigma_{t'}^{\hat{\alpha}_T}}\right)^2\\
&\quad\cdot\overline{h_\omega^{t',0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)}\,\mathrm{d}\omega\Bigg]\\
&=-2\int_{\Omega_T}\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}\sigma_{t'}^{\hat{\alpha}_T}}\right)^2\Re\Bigg\{\mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)\overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)}\\
&\quad\cdot\ \mathrm{Hess}_\theta h_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_\omega^{t',0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)}+\mathrm{grad}_\theta h_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\overline{\mathrm{grad}_\theta^\top h_\omega^{t',0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)}\Bigg\}\,\mathrm{d}\omega\\
&=-2\int_{\Omega_T}\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}\sigma_{t'}^{\hat{\alpha}_T}}\right)^2\Re\Bigg\{\mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)\overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)}\\
&\quad\cdot h_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_\omega^{t',0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)}\Big[\mathrm{i}\boldsymbol{H}_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)-\boldsymbol{b}_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\boldsymbol{b}_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)^\top\\
&\quad+\boldsymbol{b}_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\boldsymbol{b}_\omega^{t',0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)^\top\Big]\Bigg\}\,\mathrm{d}\omega\\
&=-2\int_{\Omega_T}\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}\sigma_{t'}^{\hat{\alpha}_T}}\right)^2\Re\Bigg\{\mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)\overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)}h_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\\
&\quad\cdot\overline{h_\omega^{t',0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)}\Big[\mathrm{i}\boldsymbol{H}_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)-\boldsymbol{b}_\omega^{t,0}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\boldsymbol{b}_\omega^{t,t'}(\theta;\hat{\phi}_T,\hat{\alpha}_T)^\top\Big]\Bigg\}\,\mathrm{d}\omega. \qquad (6.114)
\end{aligned}
$$

where we used that $\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha)-\boldsymbol{b}_\omega^{t',0}(\theta;\phi,\alpha)=\boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)$. By the Lipschitz-continuity of $h_\omega^{t,t'}$ in δ_t^θ as a function of the type $x\mapsto \mathrm{e}^{\mathrm{i}\langle a,x\rangle}$ having Lipschitz constant $\sqrt{2}\,\|a\|$ we have for $\hat{\theta}_T^*\in U_\delta$ with $U_\delta\subset\Theta$ from Assumption 2.14 (B6) that

$$
\begin{aligned}
\left|h_\omega^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)-1\right| &\le 2\pi\sqrt{2}\left\|\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1}R_{\rho_t^{\hat{\phi}_T}}\omega\right\|\left\|\delta_t^{\hat{\theta}_T^*}-\delta_t^{\theta_0}\right\|\\
&\le 2\pi\sqrt{2}\sigma_{\min}^{-1}\|\omega\|\,L_\delta\left\|\hat{\theta}_T^*-\theta_0\right\|\\
&= o_{\mathbb{P}}\big(\|\omega\|\big), \qquad (6.115)
\end{aligned}
$$

since $\left\|R_{\rho_t^{\hat{\phi}_T}}\omega\right\|=\|\omega\|$, as any rotation is isometric. Because of this and $\left|h_\omega^{t,t'}(\theta;\phi,\alpha)\right|=1$ we have

$$
\begin{aligned}
&\left|h_\omega^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_\omega^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)}-1\right|\\
\le\ &\left|h_\omega^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)-1\right|\left|\overline{h_\omega^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)}\right|+\left|\overline{h_\omega^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)}-1\right|\\
=\ &o_{\mathbb{P}}\big(\|\omega\|\big). \qquad (6.116)
\end{aligned}
$$

In particular, the imaginary part converges to 0. With (6.47), (6.46), and (6.116), we have using Lemma 6.7 and again $\left|h_{\omega}^{t,t'}(\theta;\phi,\alpha)\right| = 1$, that

$$\begin{aligned}
& \Big\| \Re \left\{ \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)} h_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_{\omega}^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)} \right. \\
& \left. \cdot \mathrm{i}\boldsymbol{H}_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T) \right\} \Big\|_1 \\
= & \left| \Im \left\{ \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)} h_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_{\omega}^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)} \right\} \right| \\
& \cdot \left\| \boldsymbol{H}_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T) \right\|_1 \\
= & \Big| \Im \left\{ \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)} \right\} \Re \left\{ h_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_{\omega}^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)} \right\} \\
+ & \Re \left\{ \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)} \right\} \Im \left\{ h_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_{\omega}^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)} \right\} \Big| \\
& \cdot \left\| \boldsymbol{H}_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T) \right\|_1 \\
\leq & \left| \Im \left\{ \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)} \right\} \right| \left\| \boldsymbol{H}_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T) \right\|_1 \\
& + \left| \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \right| \left| \mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \right| \left| \Im \left\{ h_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_{\omega}^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)} \right\} \right| \\
& \cdot \left\| \boldsymbol{H}_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T) \right\|_1 \\
= & \mathcal{O}_{\mathbb{P}}\left(\frac{\|\omega\|}{T} + \frac{\|\omega\| \left|\mathcal{F}_f(\omega)\right|}{\sqrt{T}} \right) + o_{\mathbb{P}}\left(\|\omega\|^2 \left|\mathcal{F}_f(\omega)\right|^2 \right).
\end{aligned}$$

It follows that

$$\begin{aligned}
& \Big\| \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T}\sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re \Big\{ \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)} \\
& \cdot h_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)\overline{h_{\omega}^{t',0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T)} \cdot \mathrm{i}\boldsymbol{H}_{\omega}^{t,0}(\hat{\theta}_T^*;\hat{\phi}_T,\hat{\alpha}_T) \Big\} \,\mathrm{d}\omega \Big\|_1 \\
\leq & \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^4 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \mathcal{O}_{\mathbb{P}}\left(\frac{\|\omega\|}{T} + \frac{\|\omega\| \left|\mathcal{F}_f(\omega)\right|}{\sqrt{T}} \right) + o_{\mathbb{P}}\left(\|\omega\|^2 \left|\mathcal{F}_f(\omega)\right|^2 \right) \,\mathrm{d}\omega \\
= & \mathcal{O}_{\mathbb{P}}\left(\frac{r_T^3}{T} + \frac{1}{\sqrt{T}} \int_{\Omega_T} \|\omega\| \left|\mathcal{F}_f(\omega)\right| \,\mathrm{d}\omega \right) + o_{\mathbb{P}}\left(\int_{\Omega_T} \|\omega\|^2 \left|\mathcal{F}_f(\omega)\right|^2 \,\mathrm{d}\omega \right) \\
= & o_{\mathbb{P}}(1),
\end{aligned} \tag{6.117}$$

because of $r_T^3/T \overset{T\to\infty}{\longrightarrow} 0$ (Assumption 2.15 (C1)), Assumption 2.13 (A3) and Lemma A.10. Similarly,

$$\begin{aligned}
& \left\| \Re \left\{ \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)} \cdot \mathrm{i}\boldsymbol{H}_{\omega}^{t,0}(\theta_0;\hat{\phi}_T,\hat{\alpha}_T) \right\} \right\|_1 \\
= & \left| \Im \left\{ \mathcal{F}_f\left(\tau_t^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right) \overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat{\phi}_T,\hat{\alpha}_T)}\omega\right)} \right\} \right| \left\| \boldsymbol{H}_{\omega}^{t,0}(\theta_0;\hat{\phi}_T,\hat{\alpha}_T) \right\|_1 \\
= & \mathcal{O}_{\mathbb{P}}\left(\frac{\|\omega\|}{T} + \frac{\|\omega\| \left|\mathcal{F}_f(\omega)\right|}{\sqrt{T}} \right),
\end{aligned}$$

which implies that

$$\left\| \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \quad \Re\left\{ \mathcal{F}_f \left(\tau_t^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right) \overline{\mathcal{F}_f \left(\tau_{t'}^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right)} \cdot \mathrm{i} \boldsymbol{H}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \right\} \mathrm{d}\omega \right\|_1 = o_{\mathbb{P}}(1). \tag{6.118}$$

For $i = 1, 2$ we define the notation

$$g\delta_t^i := \mathrm{grad}_\theta \left(\delta_t^\theta\right)_i \big|_{\theta = \hat{\theta}_T^*} - \mathrm{grad}_\theta \left(\delta_t^\theta\right)_i \big|_{\theta=\theta_0}.$$

Note that since $\hat{\theta}_T^* \overset{T\to\infty}{\longrightarrow} \theta_0$ and δ_t^θ is differentiable by Assumption 2.14 (B2), we have that $\left\|g\delta_t^i\right\| \overset{T\to\infty}{\longrightarrow} 0$ by the continuous mapping theorem (Theorem B.8). Using this, we obtain

$$\begin{aligned}
&\left\| \left(\boldsymbol{b}_\omega^{t,t'}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \right)_\theta - \left(\boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \right\|_1 \right. \\
&= 2\pi \left\| \omega_1 \left[\left(\sigma_t^{\hat{\alpha}_T} \right)^{-1} \left(\cos\left(\rho_t^{\hat{\phi}_T}\right) g\delta_t^1 + \sin\left(\rho_t^{\hat{\phi}_T}\right) g\delta_t^2 \right) \right. \right. \\
&\qquad \left. - \left(\sigma_{t'}^{\hat{\alpha}_T} \right)^{-1} \left(\cos\left(\rho_{t'}^{\hat{\phi}_T}\right) g\delta_{t'}^1 + \sin\left(\rho_{t'}^{\hat{\phi}_T}\right) g\delta_{t'}^2 \right) \right] \\
&\qquad + \omega_2 \left[\left(\sigma_t^{\hat{\alpha}_T} \right)^{-1} \left(-\sin\left(\rho_t^{\hat{\phi}_T}\right) g\delta_t^1 + \cos\left(\rho_t^{\hat{\phi}_T}\right) g\delta_t^2 \right) \right. \\
&\qquad \left. \left. - \left(\sigma_{t'}^{\hat{\alpha}_T} \right)^{-1} \left(-\sin\left(\rho_{t'}^{\hat{\phi}_T}\right) g\delta_{t'}^1 + \cos\left(\rho_{t'}^{\hat{\phi}_T}\right) g\delta_{t'}^2 \right) \right] \right\|_1 \\
&\leq \|\omega\|_1 \, \sigma_{min}^{-1} \left(\left\|g\delta_t^1\right\|_1 + \left\|g\delta_t^2\right\|_1 + \left\|g\delta_{t'}^1\right\|_1 + \left\|g\delta_{t'}^2\right\|_1 \right) \\
&= \|\omega\|_1 \, o_{\mathbb{P}}(1) = o_{\mathbb{P}}\left(\|\omega\|\right).
\end{aligned} \tag{6.119}$$

Hence, for $\widetilde{C}$ from Lemma 6.7 we have using that $\left|h_\omega^{t,t'}(\theta; \phi, \alpha)\right| = 1$, as well as (6.116) and (6.119), that

$$\begin{aligned}
&\left\| h_\omega^{t,0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \overline{h_\omega^{t',0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T)} \boldsymbol{b}_\omega^{t,0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \boldsymbol{b}_\omega^{t,t'}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T)^\top \right. \\
&\left. - \boldsymbol{b}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top \right\|_1 \\
\leq \quad & \left| h_\omega^{t,0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \overline{h_\omega^{t',0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T)} - 1 \right| \left\| \boldsymbol{b}_\omega^{t,0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \boldsymbol{b}_\omega^{t,t'}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T)^\top \right\|_1 \\
+ \quad & \left\| \boldsymbol{b}_\omega^{t,0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \left[\boldsymbol{b}_\omega^{t,t'}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) - \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \right]^\top \right\|_1 \\
+ \quad & \left\| \left[\boldsymbol{b}_\omega^{t,0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) - \boldsymbol{b}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \right] \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top \right\|_1 \\
\leq \quad & d_1 \widetilde{C}^2 \|\omega\|^2 \, o_{\mathbb{P}}\left(\|\omega\|\right) + \sqrt{d_1} \widetilde{C} \|\omega\| \, o_{\mathbb{P}}\left(\|\omega\|\right) \\
= \quad & o_{\mathbb{P}}\left(\|\omega\|^3\right).
\end{aligned} \tag{6.120}$$

With (6.46), (6.114), (6.117), (6.118), and (6.120) it follows that

$$\left\| \mathrm{Hess}_\theta A_T(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) - \mathrm{Hess}_\theta A_T(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \right\|_1$$

$$
\begin{aligned}
&= \Bigg\| 2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re\left\{ \mathcal{F}_f \left(\tau_t^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right) \overline{\mathcal{F}_f \left(\tau_{t'}^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right)} \right. \\
&\quad \cdot \Big(h_\omega^{t,0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \overline{h_\omega^{t',0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T)} \boldsymbol{b}_\omega^{t,0}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) \boldsymbol{b}_\omega^{t,t'}(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T)^\top \\
&\quad \left. - \boldsymbol{b}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top \Big) \right\} \mathrm{d}\omega \Bigg\|_1 + o_{\mathbb{P}}(1) \\
&\leq 2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^4 \left| \mathcal{F}_f \left(\tau_t^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right) \right| \left| \mathcal{F}_f \left(\tau_{t'}^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right) \right| o_{\mathbb{P}}\left(\|\omega\|^3 \right) \mathrm{d}\omega + o_{\mathbb{P}}(1) \\
&= o_{\mathbb{P}} \left(\int_{\Omega_T} \|\omega\|^3 \left(\frac{1}{T} + \frac{\left| \mathcal{F}_f(\omega) \right|}{\sqrt{T}} + \left| \mathcal{F}_f(\omega) \right|^2 \right) \mathrm{d}\omega \right) + o_{\mathbb{P}}(1) \\
&= o_{\mathbb{P}} \left(\frac{r_T^5}{T} + \frac{r_T^2}{\sqrt{T}} \int_{\Omega_T} \|\omega\| \left| \mathcal{F}_f(\omega) \right| \mathrm{d}\omega + \int_{\Omega_T} \|\omega\|^3 \left| \mathcal{F}_f(\omega) \right|^2 \mathrm{d}\omega \right) + o_{\mathbb{P}}(1) \\
&= o_{\mathbb{P}}(1), \qquad (6.121)
\end{aligned}
$$

because of $r_T^5/T \overset{T\to\infty}{\longrightarrow} 0$ (Assumption 2.15 (C1)), Assumption 2.13 (A3) and Lemma A.10. By Lemma 6.15, H_N has finite operator norm. In particular, the components of H_N are finite. Hence, with (6.113)

$$
\left\| 2 \int_{\mathbb{R}^2 \setminus \Omega_T} \left| \mathcal{F}_f(\omega) \right|^2 \int_0^1 \int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0; \phi_0, \alpha_0) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \mathrm{d}t \, \mathrm{d}t' \, \mathrm{d}\omega \right\|_1 = o(1),
$$

as $T \to \infty$. With (6.114), (6.117) and (6.121), it follows that

$$
\begin{aligned}
&\left\| \operatorname{Hess}_\theta A_T(\hat{\theta}_T^*; \hat{\phi}_T, \hat{\alpha}_T) - H_N \right\|_1 = \left\| \operatorname{Hess}_\theta A_T(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) - H_N \right\|_1 + o_{\mathbb{P}}(1) \\
&\leq 2 \int_{\Omega_T} \Bigg\| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \Bigg[\left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re\left\{ \mathcal{F}_f \left(\tau_t^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right) \overline{\mathcal{F}_f \left(\tau_{t'}^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right)} \right\} \\
&\quad \cdot \boldsymbol{b}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top \Bigg] \\
&\quad - \left| \mathcal{F}_f(\omega) \right|^2 \int_0^1 \int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0; \phi_0, \alpha_0) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \mathrm{d}t \, \mathrm{d}t' \Bigg\|_1 \mathrm{d}\omega + o_{\mathbb{P}}(1) \\
&\leq 2 \int_{\Omega_T} \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \Bigg[\left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \left| \mathcal{F}_f \left(\tau_t^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right) \overline{\mathcal{F}_f \left(\tau_{t'}^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right)} \right| \\
&\quad \cdot \left\| \boldsymbol{b}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) - \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top \boldsymbol{b}_\omega^{t,0}(\theta_0; \phi_0, \alpha_0) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \right\|_1 \\
&\quad + \left| \left(\frac{\sigma_t^{\alpha_0} \sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat{\alpha}_T} \sigma_{t'}^{\hat{\alpha}_T}} \right)^2 \Re\left\{ \mathcal{F}_f \left(\tau_t^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right) \overline{\mathcal{F}_f \left(\tau_{t'}^{(\hat{\phi}_T, \hat{\alpha}_T)} \omega \right)} \right\} - \left| \mathcal{F}_f(\omega) \right|^2 \right| \\
&\quad \cdot \left\| \boldsymbol{b}_\omega^{t,0}(\theta_0; \phi_0, \alpha_0) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \right\|_1 \Bigg] \\
&\quad + \left| \mathcal{F}_f(\omega) \right|^2 \Bigg\| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \boldsymbol{b}_\omega^{t,0}(\theta_0; \phi_0, \alpha_0) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top
\end{aligned}
$$

$$- \int_0^1 \int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0; \phi_0, \alpha_0) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \, \mathrm{d}t \, \mathrm{d}t' \Bigg\|_1 \mathrm{d}\omega + o_{\mathbb{P}}(1) \tag{6.122}$$

To show convergence of the first of the three summands, note that by Assumption 2.14 (B4) and the continuous mapping theorem (Theorem B.8)

$$\begin{aligned}
&\left| \left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \cos\left(\rho_t^{\hat{\phi}_T}\right) - \left(\sigma_t^{\alpha_0}\right)^{-1} \cos\left(\rho_t^{\phi_0}\right) \right| \\
&= \left| \left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \left(\cos\left(\rho_t^{\hat{\phi}_T}\right) - \cos\left(\rho_t^{\phi_0}\right)\right) \right| + \left| \left(\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} - \left(\sigma_t^{\alpha_0}\right)^{-1} \right) \cos\left(\rho_t^{\phi_0}\right) \right| \\
&\leq \sigma_{\min}^{-1} o_{\mathbb{P}}(1) + o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1).
\end{aligned}$$

The same holds true if we substitute the cosine by the sine function. Hence, it follows that $\left\| \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top - \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \right\|_1 = o_{\mathbb{P}}(\|\omega\|)$, as well, from the definition of $\boldsymbol{b}_\omega^{t,t'}$ and Assumption 2.14 (B2). From this it follows for C from Assumption 2.14 (B2), that

$$\begin{aligned}
&\left\| \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top - \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \right\|_1 \\
&= \Bigg\| \omega_1 \Bigg[\left(\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \cos\left(\rho_t^{\hat{\phi}_T}\right) - \left(\sigma_t^{\alpha_0}\right)^{-1} \cos\left(\rho_t^{\phi_0}\right) \right) \operatorname{grad}_\theta \left(\delta_t^{\theta_0}\right)_1 \\
&\qquad + \left(\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \sin\left(\rho_t^{\hat{\phi}_T}\right) - \left(\sigma_t^{\alpha_0}\right)^{-1} \sin\left(\rho_t^{\phi_0}\right) \right) \operatorname{grad}_\theta \left(\delta_t^{\theta_0}\right)_2 \\
&\qquad - \left(\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \cos\left(\rho_{t'}^{\hat{\phi}_T}\right) - \left(\sigma_t^{\alpha_0}\right)^{-1} \cos\left(\rho_{t'}^{\phi_0}\right) \right) \operatorname{grad}_\theta \left(\delta_{t'}^{\theta_0}\right)_1 \\
&\qquad - \left(\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \sin\left(\rho_{t'}^{\hat{\phi}_T}\right) - \left(\sigma_t^{\alpha_0}\right)^{-1} \sin\left(\rho_{t'}^{\phi_0}\right) \right) \operatorname{grad}_\theta \left(\delta_{t'}^{\theta_0}\right)_2 \Bigg] \\
&\quad + \omega_2 \Bigg[\left(- \left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \sin\left(\rho_t^{\hat{\phi}_T}\right) + \left(\sigma_t^{\alpha_0}\right)^{-1} \sin\left(\rho_t^{\phi_0}\right) \right) \operatorname{grad}_\theta \left(\delta_t^{\theta_0}\right)_1 \\
&\qquad + \left(\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \cos\left(\rho_t^{\hat{\phi}_T}\right) - \left(\sigma_t^{\alpha_0}\right)^{-1} \cos\left(\rho_t^{\phi_0}\right) \right) \operatorname{grad}_\theta \left(\delta_t^{\theta_0}\right)_2 \\
&\qquad - \left(- \left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \sin\left(\rho_{t'}^{\hat{\phi}_T}\right) + \left(\sigma_t^{\alpha_0}\right)^{-1} \sin\left(\rho_{t'}^{\phi_0}\right) \right) \operatorname{grad}_\theta \left(\delta_{t'}^{\theta_0}\right)_1 \\
&\qquad - \left(\left(\sigma_t^{\hat{\alpha}_T}\right)^{-1} \cos\left(\rho_{t'}^{\hat{\phi}_T}\right) - \left(\sigma_t^{\alpha_0}\right)^{-1} \cos\left(\rho_{t'}^{\phi_0}\right) \right) \operatorname{grad}_\theta \left(\delta_{t'}^{\theta_0}\right)_2 \Bigg] \Bigg\|_1 \\
&= 4C \|\omega\|_1 \, o_{\mathbb{P}}(1) = o_{\mathbb{P}}(\|\omega\|).
\end{aligned}$$

Thus, for $\widetilde{C}$ from Lemma 6.7

$$\begin{aligned}
&\left\| \boldsymbol{b}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top - \boldsymbol{b}_\omega^{t,0}(\theta_0; \phi_0, \alpha_0) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \right\|_1 \\
&\leq \left\| \boldsymbol{b}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) \left(\boldsymbol{b}_\omega^{t,t'}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T)^\top - \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \right) \right\|_1 \\
&\quad + \left\| \left(\boldsymbol{b}_\omega^{t,0}(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) - \boldsymbol{b}_\omega^{t,0}(\theta_0; \phi_0, \alpha_0) \right) \boldsymbol{b}_\omega^{t,t'}(\theta_0; \phi_0, \alpha_0)^\top \right\|_1 \\
&\leq 2 \sqrt{d_1} \widetilde{C} \|\omega\| \, o_{\mathbb{P}}(\|\omega\|) = o_{\mathbb{P}}\left(\|\omega\|^2\right).
\end{aligned} \tag{6.123}$$

Using (6.46) we conclude that for the first summand in (6.122)

$$\begin{aligned}
& 2\int_{\Omega_T}\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat\alpha_T}\sigma_{t'}^{\hat\alpha_T}}\right)^2\left|\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat\phi_T,\hat\alpha_T)}\omega\right)}\right| \\
& \cdot\left\|\boldsymbol{b}_\omega^{t,0}(\theta_0;\hat\phi_T,\hat\alpha_T)\boldsymbol{b}_\omega^{t,t'}(\theta_0;\hat\phi_T,\hat\alpha_T)^\top-\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)^\top\right\|_1\,\mathrm{d}\omega \\
\leq\ & 2\int_{\Omega_T}\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^4\left|\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\right|\left|\mathcal{F}_f\left(\tau_{t'}^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\right|o_{\mathbb{P}}\left(\|\omega\|^2\right)\mathrm{d}\omega \\
=\ & o_{\mathbb{P}}\left(\int_{\Omega_T}\|\omega\|^2\left(\frac{1}{T}+\frac{\left|\mathcal{F}_f(\omega)\right|}{\sqrt{T}}+\left|\mathcal{F}_f(\omega)\right|^2\right)\mathrm{d}\omega\right) \\
=\ & o_{\mathbb{P}}(1),
\end{aligned} \tag{6.124}$$

analogously to before, by Assumption 2.15 and Assumption 2.13 (A3) together with Lemma A.10. To tackle the second summand observe that by the continuous mapping theorem (Theorem B.8) we have

$$\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat\alpha_T}\sigma_{t'}^{\hat\alpha_T}}\right)^2\xrightarrow{T\to\infty}1\quad\text{in probability.}$$

Using again (6.46) and $\widetilde{C}$ from Lemma 6.7 we can show that

$$\begin{aligned}
& \int_{\Omega_T}\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\left|\left(\frac{\sigma_t^{\alpha_0}\sigma_{t'}^{\alpha_0}}{\sigma_t^{\hat\alpha_T}\sigma_{t'}^{\hat\alpha_T}}\right)^2\Re\left\{\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\overline{\mathcal{F}_f\left(\tau_{t'}^{(\hat\phi_T,\hat\alpha_T)}\omega\right)}\right\}-\left|\mathcal{F}_f(\omega)\right|^2\right| \\
& \cdot\left\|\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)^\top\right\|_1\,\mathrm{d}\omega \\
\leq\ & \int_{\Omega_T}d_1\widetilde{C}^2\|\omega\|^2\left|\big(1+o_{\mathbb{P}}(1)\big)O_{\mathbb{P}}\left(\frac{1}{T}+\frac{\left|\mathcal{F}_f(\omega)\right|}{\sqrt{T}}+\left|\mathcal{F}_f(\omega)\right|^2\right)-\left|\mathcal{F}_f(\omega)\right|^2\right|\mathrm{d}\omega \\
=\ & o_{\mathbb{P}}(1),
\end{aligned} \tag{6.125}$$

by Assumption 2.13 (A3), Lemma A.10 and Assumption 2.15. Note that under Assumption 2.14 (B3), using that $x\mapsto x^{-1}$ restricted to $[\sigma_{\min},\sigma_{\max}]$ as well as $x\mapsto\cos(x)$ and $x\mapsto\sin(x)$ are Lipschitz continuous functions preserving the bounded variation property, we get from part 1 of Lemma A.4 that $\mathrm{TV}\left(t\mapsto\left(\sigma_t^{\alpha_0}\right)^{-1}\sin\left(\rho_t^{\phi_0}\right)\right)$ and $\mathrm{TV}\left(t\mapsto\left(\sigma_t^{\alpha_0}\right)^{-1}\cos\left(\rho_t^{\phi_0}\right)\right)$ are bounded. Using part 4 and again part 1 of Lemma A.4, as well as Assumption 2.14 (B3), we obtain for $1\leq m\leq d_1$ with C from Assumption 2.14 (B2) that

$$\begin{aligned}
& \mathrm{TV}\left(t\mapsto\left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right)_m\right) \\
=&\mathrm{TV}\left(t\mapsto\left.\frac{\partial\left(\delta_t^\theta\right)_1}{\partial\theta_m}\right|_{\theta=\theta_0}\left(\omega_1\frac{\cos\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}+\omega_2\frac{-\sin\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right)\right. \\
&\left.+\left.\frac{\partial\left(\delta_t^\theta\right)_2}{\partial\theta_m}\right|_{\theta=\theta_0}\left(\omega_1\frac{\sin\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}+\omega_2\frac{\cos\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right)\right)
\end{aligned}$$

$$
\begin{aligned}
\leq & \mathrm{TV}\left(t \mapsto \left.\frac{\partial\left(\delta_t^\theta\right)_1}{\partial\theta_m}\right|_{\theta=\theta_0}\right) \cdot \left\|\left(\omega_1 \frac{\cos\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}} + \omega_2 \frac{-\sin\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right)\right\|_\infty \\
& + \left\|\left.\frac{\partial\left(\delta_t^\theta\right)_1}{\partial\theta_m}\right|_{\theta=\theta_0}\right\|_\infty \cdot \mathrm{TV}\left(t \mapsto \omega_1 \frac{\cos\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}} + \omega_2 \frac{-\sin\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right) \\
+ & \mathrm{TV}\left(t \mapsto \left.\frac{\partial\left(\delta_t^\theta\right)_2}{\partial\theta_m}\right|_{\theta=\theta_0}\right) \cdot \left\|\omega_1 \frac{\sin\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}} + \omega_2 \frac{\cos\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right\|_\infty \\
& + \left\|\left.\frac{\partial\left(\delta_t^\theta\right)_2}{\partial\theta_m}\right|_{\theta=\theta_0}\right\|_\infty \cdot \mathrm{TV}\left(t \mapsto \omega_1 \frac{\sin\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}} + \omega_2 \frac{\cos\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right) \\
\leq & \frac{2\|\omega\|_1}{\sigma_{min}} \mathrm{TV}\left(t \mapsto \left.\frac{\partial\left(\delta_t^\theta\right)_1}{\partial\theta_m}\right|_{\theta=\theta_0}\right) \\
& + C\left(|\omega_1| \mathrm{TV}\left(t \mapsto \frac{\cos\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right) + |\omega_2| \mathrm{TV}\left(t \mapsto \frac{\sin\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right)\right) \\
+ & \frac{2\|\omega\|_1}{\sigma_{min}} \mathrm{TV}\left(t \mapsto \left.\frac{\partial\left(\delta_t^\theta\right)_2}{\partial\theta_m}\right|_{\theta=\theta_0}\right) \\
& + C\left(|\omega_1| \mathrm{TV}\left(t \mapsto \frac{\sin\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right) + |\omega_2| \mathrm{TV}\left(t \mapsto \frac{\cos\left(\rho_t^{\phi_0}\right)}{\sigma_t^{\alpha_0}}\right)\right) \\
= & O(\|\omega\|).
\end{aligned}
\tag{6.126}
$$

By parts 3 and 4 of Lemma A.4 and Assumption 2.14 (B3), we get that

$$
\begin{aligned}
& \left\|\frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0) \boldsymbol{b}_\omega^{t',0}(\theta_0;\phi_0,\alpha_0)^\top - \int_0^1 \int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0) \boldsymbol{b}_\omega^{t',0}(\theta_0;\phi_0,\alpha_0)^\top \,\mathrm{d}t\,\mathrm{d}t'\right\|_1 \\
= & \sum_{m,m'=1}^{d_1} \left| \frac{\beta_T^2}{T^2} \sum_{t,t' \in \mathbb{T}} \left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right)_m \left(\boldsymbol{b}_\omega^{t',0}(\theta_0;\phi_0,\alpha_0)\right)_{m'} \right. \\
& \left. - \int_0^1 \int_0^1 \left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right)_m \left(\boldsymbol{b}_\omega^{t',0}(\theta_0;\phi_0,\alpha_0)_{m'}\right) \,\mathrm{d}t\,\mathrm{d}t' \right| \\
\leq & \widetilde{C}\,\|\omega\| \frac{\beta_T}{T} \sum_{m,m'=1}^{d_1} \left[\mathrm{TV}\left(t \mapsto \left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right)_m\right) + \mathrm{TV}\left(t' \mapsto \left(\boldsymbol{b}_\omega^{t',0}(\theta_0;\phi_0,\alpha_0)_{m'}\right)\right)\right] \\
= & O\left(\|\omega\|^2 \frac{\beta_T}{T}\right),
\end{aligned}
\tag{6.127}
$$

with $\widetilde{C} > 0$ from Lemma 6.7. Similarly, with parts 1 and 2 of Lemma A.4,

$$
\left\|\frac{\beta_T}{T} \sum_{t \in \mathbb{T}} \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0) \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)^\top - \int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0) \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)^\top \,\mathrm{d}t\right\|_1
$$

$$
\begin{aligned}
&\le \frac{\beta_T}{T}\sum_{m,m'=1}^{d_1} \mathrm{TV}\left(t \mapsto \left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right)_m \left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right)_{m'}\right)\\
&\le 2d_1\widetilde{C}\,\|\omega\|\,\frac{\beta_T}{T}\sum_{m=1}^{d_1}\mathrm{TV}\left(t\mapsto \left(\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\right)_m\right)\\
&= O\left(\|\omega\|^2\frac{\beta_T}{T}\right). && (6.128)
\end{aligned}
$$

Now, (6.127) and (6.128) together with $\boldsymbol{b}_\omega^{t,0}(\theta;\phi,\alpha) - \boldsymbol{b}_\omega^{t',0}(\theta;\phi,\alpha) = \boldsymbol{b}_\omega^{t,t'}(\theta;\phi,\alpha)$ yield that

$$
\begin{aligned}
&\left\|\frac{\beta_T^2}{T^2}\sum_{t,t'\in\mathbb{T}}\boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)^\top - \int_0^1\int_0^1 \boldsymbol{b}_\omega^{t,0}(\theta_0;\phi_0,\alpha_0)\boldsymbol{b}_\omega^{t,t'}(\theta_0;\phi_0,\alpha_0)^\top\,\mathrm{d}t\,\mathrm{d}t'\right\|_1\\
&= O\left(\|\omega\|^2\frac{\beta_T}{T}\right). \qquad (6.129)
\end{aligned}
$$

Plugging (6.124), (6.125) and (6.129) into (6.122), we obtain that

$$
\begin{aligned}
\left\|\operatorname{Hess}_\theta A_T(\hat\theta_T^*;\hat\phi_T,\hat\alpha_T) - H_N\right\|_1 &= 3o_{\mathbb{P}}(1) + O\left(\frac{\beta_T}{T}\int_{\Omega_T}\|\omega\|^2\left|\mathcal{F}_f(\omega)\right|^2\,\mathrm{d}\omega\right)\\
&= o_{\mathbb{P}}(1). \qquad (6.130)
\end{aligned}
$$

Next, show that $\sup_{\theta\in U'}\left\|\operatorname{Hess}_\theta B_T(\theta;\hat\phi_T,\hat\alpha_T)\right\| \overset{T\to\infty}{\longrightarrow} 0$ in probability. By Lemma 6.7,

$$
\sup_{\theta\in U'}\left\|\operatorname{Hess}_\theta h_\omega^{t,t'}(\theta;\phi,\alpha)\right\| = O_{\mathbb{P}}(\|\omega\|^2).
$$

Using further (6.46), the fact that $\left|V_T^t(\omega;\hat\phi_T,\hat\alpha_T)\right| = O_{\mathbb{P}}(\beta_T^{-1/2})$ by Lemma 6.3 and Markov's inequality (Theorem B.9), we obtain

$$
\begin{aligned}
&\sup_{\theta\in U'}\left\|\operatorname{Hess}_\theta B_T(\theta;\hat\phi_T,\hat\alpha_T)\right\|\\
&= \sup_{\theta\in U'}\left\|\int_{\Omega_T} 2\Re\left[\left(\frac{\beta_T}{T}\right)^2\sum_{t,t'\in\mathbb{T}}\operatorname{Hess}_\theta h_\omega^{t,t'}(\theta;\hat\phi_T,\hat\alpha_T)\right.\right.\\
&\qquad\left.\left.\cdot\left(\frac{\sigma_t^{\alpha_0}}{\sigma_t^{\hat\alpha_T}}\right)^2\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\overline{V_T^{t'}(\omega;\hat\phi_T,\hat\alpha_T)}\right]\mathrm{d}\omega\right\|\\
&\le \sup_{\theta\in U'}\int_{\Omega_T} 2\left(\frac{\beta_T}{T}\right)^2\sum_{t,t'\in\mathbb{T}}\left\|\operatorname{Hess}_\theta h_\omega^{t,t'}(\theta;\hat\phi_T,\hat\alpha_T)\right\|\\
&\qquad\cdot\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^2\left|\mathcal{F}_f\left(\tau_t^{(\hat\phi_T,\hat\alpha_T)}\omega\right)\right|\left|V_T^{t'}(\omega;\hat\phi_T,\hat\alpha_T)\right|\mathrm{d}\omega\\
&= O_{\mathbb{P}}\left(\int_{\Omega_T}\frac{\|\omega\|^2}{\sqrt{\beta_T}}\left(\frac{1}{\sqrt{T}} + \left|\mathcal{F}_f(\omega)\right|\right)\mathrm{d}\omega\right)\\
&= O_{\mathbb{P}}\left(\frac{r_T^4}{\sqrt{T\beta_T}} + \frac{r_T}{\sqrt{\beta_T}}\int_{\Omega_T}\|\omega\|\left|\mathcal{F}_f(\omega)\right|\mathrm{d}\omega\right).
\end{aligned}
$$

By Assumption 2.15, $\frac{r_T^4}{\sqrt{T\beta_T}} \overset{T\to\infty}{\longrightarrow} 0$ and $\frac{r_T}{\sqrt{\beta_T}} \overset{T\to\infty}{\longrightarrow} 0$. The integral is bounded by the Sobolev condition (Assumption 2.13 (A3)) and Lemma A.10 and hence, we get the desired convergence of the Hessian of B_T:

$$\sup_{\theta\in U'} \left\| \mathrm{Hess}_\theta B_T(\theta; \hat{\phi}_T, \hat{\alpha}_T)\right\| \overset{T\to\infty}{\longrightarrow} 0 \text{ in probability.} \tag{6.131}$$

Finally, consider the Hessian of C_T. Note that

$$h_\omega^{0,t}(\theta;\phi,\alpha)^{-1}\overline{h_\omega^{0,t'}(\theta;\phi,\alpha)^{-1}} = h_\omega^{t,t'}(\theta;\phi,\alpha)h_\omega^{0,t}(\theta_0;\phi_0,\alpha_0)^{-1}.$$

Using this together with Lemmas 6.7 and 6.3, we have similarly to before that

$$\begin{aligned}
&\sup_{\theta\in U'} \left\| \mathrm{Hess}_\theta C_T(\theta; \hat{\phi}_T, \hat{\alpha}_T)\right\| \\
&= \sup_{\theta\in U'} \Bigg\| \int_{\Omega_T} \left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} \mathrm{Hess}_\theta h_\omega^{t,t'}(\theta;\hat{\phi}_T,\hat{\alpha}_T) h_\omega^{0,t}(\theta_0;\phi_0,\alpha_0)^{-1} \\
&\qquad \cdot V_T^t(\omega;\hat{\phi}_T,\hat{\alpha}_T)\overline{V_T^{t'}(\omega;\hat{\phi}_T,\hat{\alpha}_T)} \mathrm{d}\omega \Bigg\| \\
&\le \sup_{\theta\in U'} \int_{\Omega_T} \left(\frac{\beta_T}{T}\right)^2 \sum_{t,t'\in\mathbb{T}} \left\| \mathrm{Hess}_\theta h_\omega^{t,t'}(\theta;\hat{\phi}_T,\hat{\alpha}_T)\right\| \left| V_T^t(\omega;\hat{\phi}_T,\hat{\alpha}_T)\overline{V_T^{t'}(\omega;\hat{\phi}_T,\hat{\alpha}_T)}\right| \mathrm{d}\omega \\
&= \int_{\Omega_T} O_{\mathbb{P}}\left(\frac{\|\omega\|^2}{\beta_T}\right) \mathrm{d}\omega \\
&= O_{\mathbb{P}}\left(\frac{r_T^4}{\beta_T}\right).
\end{aligned}$$

Since $\frac{r_T^4}{\beta_T} \overset{T\to\infty}{\longrightarrow} 0$, we have

$$\sup_{\theta\in U'} \left\| \mathrm{Hess}_\theta C_T(\theta; \hat{\phi}_T, \hat{\alpha}_T)\right\| \overset{T\to\infty}{\longrightarrow} 0 \text{ in probability.} \tag{6.132}$$

Together, (6.130), (6.131) and (6.132) yield the claimed convergence of the Hessian of the contrast functional from the first statement:

$$\left\| \mathrm{Hess}_\theta N_T(\hat{\theta}_T^\dagger; \hat{\phi}_T, \hat{\alpha}_T) - \mathrm{Hess}_\theta N(\theta_0;\phi_0,\alpha_0)\right\| \overset{T\to\infty}{\longrightarrow} 0 \text{ in probability.} \qquad \square$$

Proof of Theorem 3.6 (Central limit theorem for the drift parameter estimator).
Combining the results, we obtain a central limit theorem for the estimator of the drift parameter. As a consequence of Assumption 2.14 (B4), N_T is twice continuously differentiable in θ in a convex open neighborhood $U \subset \Theta$ of θ_0. In particular, if N_T has a minimum at some $\theta \in U$ for some (ϕ,α), then $\mathrm{grad}_\theta N_T(\theta;\phi,\alpha) = 0$. Let $U' \subset \Phi \times \mathrm{A}$ be some convex open neighborhood of

(ϕ_0, α_0) and define

$$G_T(\theta) := \begin{cases} \mathrm{grad}_\theta N_T(\theta; \hat{\phi}_T, \hat{\alpha}_T), & \text{if } (\theta, \hat{\phi}_T, \hat{\alpha}_T) \in U \times U', \\ \mathrm{grad}_\theta N_T(\theta_0; \phi_0, \alpha_0), & \text{if } \theta \in \Theta \setminus U \text{ or } (\hat{\phi}_T, \hat{\alpha}_T) \in (\Phi \times \mathrm{A}) \setminus U'. \end{cases}$$

As $\hat{\theta}_T$ is defined as a minimizer of $N_T(\theta; \hat{\phi}_T, \hat{\alpha}_T)$ and so $\mathrm{grad}_\theta N_T(\hat{\theta}_T; \hat{\phi}_T, \hat{\alpha}_T) = 0$, and because $\hat{\theta}_T \overset{T\to\infty}{\longrightarrow} \theta_0 \in U$, $(\hat{\phi}_T, \hat{\alpha}_T) \overset{T\to\infty}{\longrightarrow} (\phi_0, \alpha_0) \in U'$ in probability, we have that for all $\epsilon > 0$

$$\begin{aligned} & \mathbb{P}\left(\sqrt{T} G_T(\hat{\theta}_T) > \epsilon\right) \\ \leq\ & \mathbb{P}\left(\sqrt{T}\mathrm{grad}_\theta N_T(\hat{\theta}_T; \hat{\phi}_T, \hat{\alpha}_T) > \epsilon, \hat{\theta}_T \in U, (\hat{\phi}_T, \hat{\alpha}_T) \in U'\right) \\ +\ & \mathbb{P}\left(\sqrt{T}\mathrm{grad}_\theta N_T(\theta_0; \phi_0, \alpha_0) > \epsilon, \hat{\theta}_T \notin U\right) \\ +\ & \mathbb{P}\left(\sqrt{T}\mathrm{grad}_\theta N_T(\theta_0; \phi_0, \alpha_0) > \epsilon, (\hat{\phi}_T, \hat{\alpha}_T) \notin U'\right) \\ \leq\ & \mathbb{P}\left(\hat{\theta}_T \notin U\right) + \mathbb{P}\left((\hat{\phi}_T, \hat{\alpha}_T) \notin U'\right) \overset{T\to\infty}{\longrightarrow} 0, \end{aligned}$$

which means that

$$\sqrt{T} G_T(\hat{\theta}_T) = o_{\mathbb{P}}(1). \tag{6.133}$$

On the event that $\hat{\theta}_T \in U$, and that $(\hat{\phi}_T, \hat{\alpha}_T) \in U'$ we can apply the mean value theorem for real functions of multiple variables to each component of $\mathrm{grad}_\theta N_T(\theta; \hat{\phi}_T, \hat{\alpha}_T)$ as a function of θ and of $\mathrm{grad}_\theta N_T(\theta_0; \phi, \alpha)$ as a function of (ϕ, α), resp. to get that

$$\begin{aligned} & \mathrm{grad}_\theta N_T(\hat{\theta}_T; \hat{\phi}_T, \hat{\alpha}_T) \\ &= \mathrm{grad}_\theta N_T(\theta_0; \hat{\phi}_T, \hat{\alpha}_T) + \mathrm{Hess}_\theta N_T(\hat{\theta}_T^\dagger; \hat{\phi}_T, \hat{\alpha}_T)\left(\hat{\theta}_T - \theta_0\right) \\ &= \mathrm{grad}_\theta N_T(\theta_0; \phi_0, \alpha_0) + \left(\mathrm{grad}_{(\phi,\alpha)} \mathrm{grad}_\theta^\top N_T(\theta_0; \hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger)\right)^\top \begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} \\ &+ \mathrm{Hess}_\theta N_T(\hat{\theta}_T^\dagger; \hat{\phi}_T, \hat{\alpha}_T)\left(\hat{\theta}_T - \theta_0\right), \end{aligned} \tag{6.134}$$

for some intermediate values denoted by the superscript †. With (6.133), (6.134), and the definitions

$$H_T := \begin{cases} \mathrm{Hess}_\theta N_T(\hat{\theta}_T^\dagger; \hat{\phi}_T, \hat{\alpha}_T), & \text{if } (\hat{\theta}_T, \hat{\phi}_T, \hat{\alpha}_T) \in U \times U', \\ 0, & \text{if } \hat{\theta}_T \in \Theta \setminus U \text{ or } (\hat{\phi}_T, \hat{\alpha}_T) \in (\Phi \times \mathrm{A}) \setminus U', \end{cases}$$

and

$$\widetilde{D}_T := \begin{cases} \left(\mathrm{grad}_{(\phi,\alpha)} \mathrm{grad}_\theta^\top N_T(\theta_0; \hat{\phi}_T^\dagger, \hat{\alpha}_T^\dagger)\right)^\top, & \text{if } (\hat{\theta}_T, \hat{\phi}_T, \hat{\alpha}_T) \in U \times U', \\ 0, & \text{if } \hat{\theta}_T \in \Theta \setminus U \text{ or } (\hat{\phi}_T, \hat{\alpha}_T) \in (\Phi \times \mathrm{A}) \setminus U', \end{cases}$$

we get that

$$\sqrt{T}\mathrm{grad}_\theta N_T(\theta_0;\phi_0,\alpha_0) + \widetilde{D}_T\sqrt{T}\begin{pmatrix}\hat{\phi}_T-\phi_0\\ \hat{\alpha}_T-\alpha_0\end{pmatrix} + H_T\sqrt{T}\left(\hat{\theta}_T-\theta_0\right) = \sqrt{T}G_T(\hat{\theta}_T) = o_{\mathbb{P}}(1), \tag{6.135}$$

which holds on $\{\hat{\theta}_T \notin U\}$ and $\{(\hat{\phi}_T,\hat{\alpha}_T)\notin U'\}$ by design of G_T, H_T and $\widetilde{D}_T$, and on $\{(\hat{\theta}_T,\hat{\phi}_T,\hat{\alpha}_T)\in U\times U'\}$ because of (6.134). Equation (6.135) yields that

$$H_T\sqrt{T}\left(\hat{\theta}_T-\theta_0\right) = -\widetilde{D}_T\sqrt{T}\begin{pmatrix}\hat{\phi}_T-\phi_0\\ \hat{\alpha}_T-\alpha_0\end{pmatrix} - \sqrt{T}\mathrm{grad}_\theta N_T(\theta_0;\phi_0,\alpha_0) + o_{\mathbb{P}}(1). \tag{6.136}$$

By Theorem 3.5 there are weights $w_j^{t'}$ such that

$$\sqrt{T}\begin{pmatrix}\hat{\phi}_T-\phi_0\\ \hat{\alpha}_T-\alpha_0\end{pmatrix} = -\sum_{t'\in\mathbb{T}}\sum_{j=1}^{n} H_M^{-1}w_j^{t'}\epsilon_j^{t'} + o_{\mathbb{P}}(1) = O_{\mathbb{P}}(1).$$

It holds that $\widetilde{D}_T \overset{T\to\infty}{\longrightarrow} D_N := \left(\mathrm{grad}_{(\phi,\alpha)}\mathrm{grad}_\theta^\top N(\theta_0;\phi_0,\alpha_0)\right)^\top$ in probability, since by the assumptions formulated in Theorem 6.13 $(\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger) \overset{T\to\infty}{\longrightarrow} (\phi_0,\alpha_0)$ in probability and the estimators $\hat{\theta}_T$, $\hat{\phi}_T$, and $\hat{\alpha}_T$ are consistent by Theorems 3.2 and 3.1, resp. Hence,

$$\left(\widetilde{D}_T - D_N\right)\sqrt{T}\begin{pmatrix}\hat{\phi}_T-\phi_0\\ \hat{\alpha}_T-\alpha_0\end{pmatrix} = o_{\mathbb{P}}(1)O_{\mathbb{P}}(1) = o_{\mathbb{P}}(1).$$

Similarly,

$$\sqrt{T}\mathrm{grad}_\theta N_T(\theta_0;\phi_0,\alpha_0) = \sum_{t'\in\mathbb{T}}\sum_{j=1}^{n}\tilde{w}_j^{t'}\epsilon_j^{t'} + o_{\mathbb{P}}(1)$$

for some weights $\tilde{w}_j^{t'}$ by Theorem 6.14. Plugging these results into (6.136) we obtain for $\breve{w}_j^{t'} := -D_N H_M^{-1}w_j^{t'} + \tilde{w}_j^{t'}$

$$\begin{aligned} H_T\sqrt{T}\left(\hat{\theta}_T-\theta_0\right) &= -D_N\sqrt{T}\begin{pmatrix}\hat{\phi}_T-\phi_0\\ \hat{\alpha}_T-\alpha_0\end{pmatrix} - \sqrt{T}\mathrm{grad}_\theta N_T(\theta_0;\phi_0,\alpha_0) + o_{\mathbb{P}}(1) \\ &= -\sum_{t'\in\mathbb{T}}\sum_{j=1}^{n}\breve{w}_j^{t'}\epsilon_j^{t'} + o_{\mathbb{P}}(1) = O_{\mathbb{P}}(1). \end{aligned} \tag{6.137}$$

Since $\hat{\theta}_T^\dagger$ is between $\hat{\theta}_T$ and θ_0 and $\hat{\theta}_T$ is a consistent estimator, we have that $\hat{\theta}_T^\dagger \overset{T\to\infty}{\longrightarrow} \theta_0$ in probability, and similarly $(\hat{\phi}_T^\dagger,\hat{\alpha}_T^\dagger) \overset{T\to\infty}{\longrightarrow} (\phi_0,\alpha_0)$. Because of Lemma 6.15, H_N is invertible, and by Theorem 6.16, $H_T \overset{T\to\infty}{\longrightarrow} H_N$ in probability. Together with (6.136) and Lemma A.12, this implies that

$$\sqrt{T}\left(\hat{\theta}_T-\theta_0\right) = O_{\mathbb{P}}(1).$$

Hence, again with Theorem 6.16,

$$(H_T - H_N)\,\sqrt{T}\left(\hat{\theta}_T - \theta_0\right) = o_{\mathbb{P}}(1)O_{\mathbb{P}}(1) = o_{\mathbb{P}}(1).$$

From this and (6.137) it follows that for $\Sigma_D := D_N H_M^{-1}\Sigma_{RS} H_M^{-1} D_N + \widetilde{\Sigma}$ with Σ_{RS} from Theorem 3.5 and $\widetilde{\Sigma}$ from Theorem 6.14 we have

$$H_N\sqrt{T}\left(\hat{\theta}_T - \theta_0\right) = -\sum_{t'\in\mathbb{T}}\sum_{j\in J^{t'}} \check{w}_j^{t'}\epsilon_j^{t'} + o_{\mathbb{P}}(1) \overset{T\to\infty}{\longrightarrow} \mathbb{N}(0,\Sigma_D) \quad \text{in distribution,}$$

where we used that, for all centered normal random vectors X, X and $-X$ have the same distribution. Finally, multiplication with H_N^{-1} yields the assertions and concludes the proof of asymptotic normality of the drift parameter estimator. □

Proof of Theorem 3.7 (Central limit theorem for the joint distribution of the motion function parameters).
As a consequence of Theorems 3.5 and 3.6, we establish joint asymptotic normality of all three motion function parameter estimators, finishing thereby the proof of Theorem 3.7. By Theorem 3.5 there are weights $w_j^{t'}$ such that

$$\sqrt{T}\begin{pmatrix} \hat{\phi}_T - \phi_0 \\ \hat{\alpha}_T - \alpha_0 \end{pmatrix} = -\sum_{t'\in\mathbb{T}}\sum_{j=1}^{n} H_M^{-1} w_j^{t'}\epsilon_j^{t'} + o_{\mathbb{P}}(1).$$

Moreover, from the proof of Theorem 3.6 we know that there are weights $\check{w}_j^{t'}$ such that

$$\sqrt{T}\left(\hat{\theta}_T - \theta_0\right) = -\sum_{t'\in\mathbb{T}}\sum_{j=1}^{n} H_N^{-1}\check{w}_j^{t'}\epsilon_j^{t'} + o_{\mathbb{P}}(1).$$

Let now $\xi \in \mathbb{R}^{d_1+d_2+d_3}\setminus\{0\}$. Writing $\xi = (\xi^1,\xi^2) \in \mathbb{R}^{d_1}\times\mathbb{R}^{d_2+d_3}$ we obtain

$$\begin{aligned}
&\left\langle \xi,\ \sqrt{T}\left(\hat{\theta}_T - \theta_0, \hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0\right)\right\rangle \\
&= \left\langle \xi^1,\ \sqrt{T}(\hat{\theta}_T - \theta_0)\right\rangle + \left\langle \xi^2,\ \sqrt{T}\left(\hat{\phi}_T - \phi_0, \hat{\alpha}_T - \alpha_0\right)\right\rangle \\
&= -\sum_{t'\in\mathbb{T}}\sum_{j=1}^{n}\left(\left\langle \xi^1, H_N^{-1}\check{w}_j^{t'}\right\rangle + \left\langle \xi^2, H_M^{-1}w_j^{t'}\right\rangle\right)\epsilon_j^{t'} + o_{\mathbb{P}}(1).
\end{aligned}$$

The expression $-\sum_{t'\in\mathbb{T}}\sum_{j=1}^{n}\left(\left\langle \xi^1, H_N^{-1}\check{w}_j^{t'}\right\rangle + \left\langle \xi^2, H_M^{-1}w_j^{t'}\right\rangle\right)\epsilon_j^{t'}$ is a linear combination of the error terms $\epsilon_j^{t'}$ and as such is centered normally distributed with variance given by

$$\operatorname{Var}\left(-\sum_{t'\in\mathbb{T}}\sum_{j=1}^{n}\left(\left\langle \xi^1, H_N^{-1}\check{w}_j^{t'}\right\rangle + \left\langle \xi^2, H_M^{-1}w_j^{t'}\right\rangle\right)\epsilon_j^{t'}\right)$$

$$
\begin{aligned}
&= \sum_{t' \in \mathbb{T}} \sum_{j=1}^{n} \left(\left\langle \xi^1, H_N^{-1} \check{w}_j^{t'} \right\rangle + \left\langle \xi^2, H_M^{-1} w_j^{t'} \right\rangle \right)^2 \mathrm{Var}\left(\epsilon_j^{t'} \right) \\
&= \sum_{t' \in \mathbb{T}} \sum_{j=1}^{n} \left\langle \xi^1, H_N^{-1} \check{w}_j^{t'} \right\rangle^2 + \sum_{t' \in \mathbb{T}} \sum_{j=1}^{n} \left\langle \xi^2, H_M^{-1} w_j^{t'} \right\rangle^2 \\
&\quad + 2 \sum_{t' \in \mathbb{T}} \sum_{j=1}^{n} \left\langle \xi^1, H_N^{-1} \check{w}_j^{t'} \right\rangle \left\langle \xi^2, H_M^{-1} w_j^{t'} \right\rangle \\
&= \left(\xi^1\right)^\top H_N^{-1} \Sigma_D H_N^{-1} \xi^1 + \left(\xi^2\right)^\top H_M^{-1} \Sigma_{RS} H_M^{-1} \xi^2 \\
&\quad + 2 \left(\xi^1\right)^\top \left(H_N^{-1} \Sigma_D H_N^{-1}\right)^{1/2} \left(H_M^{-1} \Sigma_{RS} H_M^{-1}\right)^{1/2} \xi^2 \\
&= \xi \top \Sigma \xi,
\end{aligned}
$$

where

$$
\Sigma := \begin{pmatrix} H_N^{-1} \Sigma_D H_N^{-1} & \left(H_N^{-1} \Sigma_D H_N^{-1}\right)^{1/2} \left(H_M^{-1} \Sigma_{RS} H_M^{-1}\right)^{1/2} \\ \left(H_N^{-1} \Sigma_D H_N^{-1}\right)^{1/2} \left(H_M^{-1} \Sigma_{RS} H_M^{-1}\right)^{1/2} & H_M^{-1} \Sigma_{RS} H_M^{-1} \end{pmatrix},
$$

with Σ_{RS} from Theorem 3.5 and Σ_D from Theorem 3.6. This concludes the proof of Theorem 3.7. □

APPENDIX A

Auxiliary Results

In this first appendix, we list some general mathematical results. Some of them have already been shown and used in Hartmann (2016), others are inspired by similar statements in the same document, but are proved here for the first time in a mathematically rigorous way. For each result the corresponding statement in Hartmann (2016) is given, whenever applicable. The following six lemmas can be found in Hartmann (2016) together with detailed proofs.

Lemma A.1 (Lemma A.11 in Hartmann (2016))**.** *Let $g_1\colon [0,1] \to \mathbb{C}$ and $g_2\colon [0,1] \to (0,\infty)$ integrable such that $|g_1| \le 1$ and*

$$\left| \int_0^1 g_1(t) g_2(t)\,\mathrm{d}t \right| = \int_0^1 g_2(t)\,\mathrm{d}t. \tag{A.1}$$

Then, there is $c \in \mathbb{C}$ with $|c| = 1$ such that $g_1(t) = c$ a.e. on $[0,1]$.

Lemma A.2 (Lemma A.12 in Hartmann (2016))**.** *Let $g\colon \mathbb{R}^2 \to \mathbb{C}$ such that*

$$L_g \coloneqq 2\pi\sqrt{2} \int_{\mathbb{R}^2} \|x\|\,|g(x)|\,\mathrm{d}x < \infty.$$

Then, the Fourier transform $\mathcal{F}_g\colon \mathbb{R}^2 \to \mathbb{C}$ is Lipschitz-continuous with Lipschitz-constant L_g.

Lemma A.3 (Lemma A.13 in Hartmann (2016))**.** *Let $g\colon \mathbb{R}^2 \to \mathbb{C}$ such that*

$$L'_g \coloneqq \sqrt{2} \int_0^\infty \int_0^{2\pi} \left\| \left(2\pi\psi, \log(r)\right) \right\| r^\gamma \,|(g \circ \mathcal{P})(r,\psi)|\,\mathrm{d}\psi\,\frac{\mathrm{d}r}{r} < \infty.$$

Then, the analytical Fourier-Mellin transform $\mathcal{M}_g\colon \mathbb{R}^2 \to \mathbb{C}$ is Lipschitz-continuous with Lipschitz-constant L'_g.

Lemma A.4 (Lemmas A.5 (part 2), A.6, A.7 and A.9 in Hartmann (2016))**.** *Consider $g, g_1, g_2\colon [0,1] \to \mathbb{C}$ and $C > 0$ such that $|g_1(t)| \le C$ and $|g_2(t)| \le C$ for all $t \in [0,1]$. Let $T \in \mathbb{N}$ and $t_i \coloneqq i/T$ for $i \in \{0,1,\dots,T\}$. Then,*

1. $\mathrm{TV}(g_1 \cdot g_2) \le \|g_2\|_\infty \mathrm{TV}(g_1) + \|g_1\|_\infty \mathrm{TV}(g_2)$,
2. $\left| \frac{1}{T} \sum_{i=0}^{T-1} g(t_i) - \int_0^1 g(t)\,\mathrm{d}t \right| \le \frac{\mathrm{TV}(g)}{T}$, *and*
3. $\left| \left(\frac{1}{T} \sum_{i=0}^{T-1} g_1(t_i) \right) \left(\frac{1}{T} \sum_{i'=0}^{T-1} g_2(t_{i'}) \right) - \int_0^1 g_1(t)\,\mathrm{d}t \int_0^1 g_2(t')\,\mathrm{d}t' \right| \le \frac{C(\mathrm{TV}(g_1)+\mathrm{TV}(g_2))}{T}$.

4. *If additionally* $g = ag_1 + bg_2$ *for* $a, b \in \mathbb{C}$, *then* $\mathrm{TV}(g) \le |a|\,\mathrm{TV}(g_1) + |b|\,\mathrm{TV}(g_2)$.

Lemma A.5 (Lemma A.14 in Hartmann (2016))**.** *Let I a finite index set,* $d \in \mathbb{N}$, *and* $g_i : \mathbb{R}^d \to \mathbb{C}$ *for* $i \in I$.

1. *If* g_i *is differentiable for all* $i \in I$, *then, for all* $m \in \{1, \dots, d\}$,

$$\frac{\partial}{\partial x_m} \sum_{i,i' \in I} g_i(x)\overline{g_{i'}(x)} = 2 \sum_{i,i' \in I} \Re\left(\frac{\partial g_i}{\partial x_m}(x)\overline{g_{i'}(x)} \right).$$

2. *If* g_i *is twice differentiable for all* $i \in I$, *then, for all* $k, m \in \{1, \dots, d\}$,

$$\frac{\partial^2}{\partial x_k \partial x_m} \sum_{i,i' \in I} g_i(x)\overline{g_{i'}(x)} = 2 \sum_{i,i' \in I} \Re\left(\frac{\partial^2 g_i}{\partial x_k \partial x_m}(x)\overline{g_{i'}(x)} + \frac{\partial g_i}{\partial x_m}(x)\overline{\frac{\partial g_{i'}}{\partial x_k}(x)} \right).$$

Lemma A.6 (Lemma A.15 in Hartmann (2016))**.** *Let* $d \in \mathbb{N}$ *and let* $g : \mathbb{R}^d \times [0, 1] \to \mathbb{C}$ *such that*

1. $t \mapsto g(x, t)$ *is integrable for all* $x \in \mathbb{R}^d$,
2. $x \mapsto g(x, t)$ *is continuously differentiable a.e. on* $[0, 1]$,
3. *there is an integrable* $h : [0, 1] \to [0, \infty)$ *such that, for all* $(x, t) \in \mathbb{R}^d \times [0, 1]$ *and for all* $m \in \{1, \dots, d\}$,

$$\left| \frac{\partial g}{\partial x_m}(x, t) \right| \le h(t).$$

Then,

$$\frac{\partial}{\partial x_m} \int_0^1 \int_0^1 g(x, t)\overline{g(x, t')}\,\mathrm{d}t\,\mathrm{d}t' = 2 \int_0^1 \int_0^1 \Re\left(\frac{\partial g}{\partial x_m}(x, t)\overline{g(x, t')} \right) \mathrm{d}t\,\mathrm{d}t'.$$

The following two lemmas provide characterization of rotation and scaling invariance properties, linking them to the existence of regions where the analytical Fourier-Mellin transform does not vanish. They are inspired by similar statements in Hartmann (2016).

Lemma A.7 (inspired by Lemma 6.5 in Hartmann (2016))**.** *Let* $g \in L^2(\mathbb{R}^2)$ *and* $\gamma > 0$ *such that* $\left(\omega \mapsto \|\omega\|^\gamma \left|\mathcal{F}_g(\omega)\right|^2\right) \in L^1(\mathbb{R}^2)$. *The function g is not scaling invariant if and only if there are* $u \in \mathbb{Z}$ *and an open Borel set* $B \subseteq \mathbb{R}$ *with positive Lebesgue measure such that* $\mathcal{M}_{|\mathcal{F}_g|^2}(u, v) \neq 0$ *for all* $v \in B$.

Proof. Let $\sigma \in \mathbb{R}^2 \times (0, \infty)$ such that $g(x) = g\left(1/\sigma \cdot x\right)$ for all $x \in \mathbb{R}^2$. Because of the generalized shift property (2.5), this implies that

$$\left|\mathcal{F}_g(\omega)\right|^2 = \sigma^4 \left|\mathcal{F}_g(\sigma\omega)\right|^2 \quad \text{for all } \omega \in \mathbb{R}^2.$$

Hence, by (2.10),

$$\mathcal{M}_{|\mathcal{F}_g|^2}(u, v) = \sigma^{4-\gamma+\mathrm{i}v} \mathcal{M}_{|\mathcal{F}_g|^2}(u, v) \quad \text{for all } (u, v) \in \mathbb{Z} \times \mathbb{R}.$$

For $\mathcal{M}_{|\mathcal{F}_g|^2}(u,v) \neq 0$, this implies that $\sigma^{4-\gamma+\mathrm{i}v} = 1$. In the case of $\gamma \neq 4$, taking the absolute value yields $\sigma = 1$. For $\gamma = 4$, it follows that

$$\mathrm{e}^{\mathrm{i}v\log(\sigma)} = 1. \tag{A.2}$$

Assume that this holds for all v in an open Borel set $B \subseteq \mathbb{R}$ with positive Lebesgue-measure. Because both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense in $\mathbb{R}$, we can therefore fix $v \in B$ and choose $v_1, v_2 \in B \setminus \{v\}$ such that

$$v - v_1 \in \mathbb{Q}, \quad \text{while} \quad v - v_2 \in \mathbb{R} \setminus \mathbb{Q}. \tag{A.3}$$

By (A.2), we have that

$$\mathrm{e}^{\mathrm{i}(v-v_1)\log(\sigma)} = \frac{\mathrm{e}^{\mathrm{i}v\log(\sigma)}}{\mathrm{e}^{\mathrm{i}v_1\log(\sigma)}} = 1,$$

which implies $(v - v_1)\log(\sigma)/2\pi \in \mathbb{Z}$. Similarly, we get $(v - v_2)\log(\sigma)/2\pi \in \mathbb{Z}$. Because of (A.3), this implies $\log(\sigma) = 0$ and thus, $\sigma = 1$ for all possible values of γ.
On the other hand, if for all $u \in \mathbb{Z}$ and all open Borel sets $B \subseteq \mathbb{R}$ with positive Lebesgue-measure there is $v \in B$ such that $\mathcal{M}_{|\mathcal{F}_g|^2}(u,v) = 0$, then $\mathcal{M}_{|\mathcal{F}_g|^2} = 0$ a.e. With the inverse analytical Fourier-Mellin transform, it follows that

$$r^\gamma \left(\left|\mathcal{F}_g\right|^2 \circ \mathcal{P}\right)(r,\psi) = \int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \mathrm{e}^{2\pi\mathrm{i}u\psi} r^{\mathrm{i}v} \mathcal{M}_{|\mathcal{F}_g|^2}(u,v)\,\mathrm{d}v = 0 \quad \text{for all } (r,\psi) \in [0,\infty) \times [0,2\pi).$$

Hence, $\mathcal{F}_g = 0$, and thus,

$$g(x) = \int_{\mathbb{R}^2} \mathrm{e}^{2\pi\mathrm{i}\langle x,\omega\rangle} \mathcal{F}_g(\omega)\,\mathrm{d}\omega = 0 \quad \text{for all } x \in \mathbb{R}^2,$$

i.e., g is constant and, in particular, $g(x) = g\big(1/\sigma \cdot (x)\big)$ for all $x \in \mathbb{R}^2$ and $\sigma \in \mathbb{R}^2 \times (0,\infty)$. □

Lemma A.8 (inspired by Lemma 6.4 in Hartmann (2016))**.** *Let $g \in L^2(\mathbb{R}^2)$ and $\gamma > 0$ such that $\left(\omega \mapsto \|\omega\|^\gamma \left|\mathcal{F}_g(\omega)\right|^2\right) \in L^1(\mathbb{R}^2)$. If the function g is not rotation invariant, there are $u \in \mathbb{Z} \setminus \{0\}$ and $v \in \mathbb{R}$ such that $\mathcal{M}_{|\mathcal{F}_g|^2}(u,v) \neq 0$.*

Proof. Assume that for all $u \in \mathbb{Z} \setminus \{0\}$ and $v \in \mathbb{R}$ we have that $\mathcal{M}_{|\mathcal{F}_g|^2}(u,v) = 0$. With the inverse analytical Fourier-Mellin transform, we get that

$$r^\gamma \left(\left|\mathcal{F}_g\right|^2 \circ \mathcal{P}\right)(r,\psi) = \int_{\mathbb{R}} \sum_{u\in\mathbb{Z}} \mathrm{e}^{2\pi\mathrm{i}u\psi} r^{\mathrm{i}v} \mathcal{M}_{|\mathcal{F}_g|^2}(u,v)\,\mathrm{d}v = 0 \quad \text{for all } (r,\psi) \in [0,\infty) \times [0,2\pi),$$

implying that $\mathcal{F}_g = 0$. Hence,

$$g(x) = \int_{\mathbb{R}^2} \mathrm{e}^{2\pi\mathrm{i}\langle x,\omega\rangle} \mathcal{F}_g(\omega)\,\mathrm{d}\omega = 0 \quad \text{for all } x \in \mathbb{R}^2,$$

i.e., g is constant and, in particular, $g(x) = g\big(R_{-\rho}(x)\big)$ for all $x \in \mathbb{R}^2$ and $\rho \in \mathbb{R}^2 \times [0,2\pi)$ □

The following three results are again taken from Hartmann (2016) together with their proofs. The first one is a characterization of translation invariance by Fourier transform properties and is analogous to the previous two characterizations of rotation and scaling invariance. The second one is a result on uniform tightness of sequences of random variables. The last one is a result on the convergence of normally distributed random variables.

Lemma A.9 (Lemma 6.27 in Hartmann (2016)). *A function $g \in L^2(\mathbb{R}^2)$ is not translation invariant if and only if there is an open Borel set $B \subseteq \mathbb{R}^2$ with positive Lebesgue-measure such that $\mathcal{F}_g(\omega) \neq 0$ for all $\omega \in B$.*

Lemma A.10 (Integrability of the Fourier transform). *Let $g \in L^2(\mathbb{R}^2) \cap H^{3+\kappa}(\mathbb{R}^2)$ for some $\kappa > 0$. Then it holds that*

$$\int_{\mathbb{R}^2} \|\omega\|^j \left|\mathcal{F}_g(\omega)\right| \mathrm{d}\omega < \infty, \quad j = 1, 2.$$

Proof. Using the Cauchy-Schwarz-inequality (Theorem B.1) in the second step and a transform to polar coordinates in the third, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^2} \|\omega\|^2 \left|\mathcal{F}_g(\omega)\right| \mathrm{d}\omega \\
&= \int_{\mathbb{R}^2} \|\omega\|^2 \left(1 + \|\omega\|^2\right)^{(1+\kappa)/2} \left|\mathcal{F}_g(\omega)\right| \left(1 + \|\omega\|^2\right)^{-(1+\kappa)/2} \mathrm{d}\omega \\
&\leq \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right)^2 \left(1 + \|\omega\|^2\right)^{(1+\kappa)} \left|\mathcal{F}_g(\omega)\right|^2 \mathrm{d}\omega \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right)^{-(1+\kappa)} \mathrm{d}\omega \\
&= \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right)^{(3+\kappa)} \left|\mathcal{F}_g(\omega)\right|^2 \mathrm{d}\omega \cdot 2\pi \int_0^\infty \frac{r}{\left(1 + r^2\right)^{1+\kappa}} \mathrm{d}r \\
&= \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right)^{(3+\kappa)} \left|\mathcal{F}_g(\omega)\right|^2 \mathrm{d}\omega \cdot \frac{\pi}{\kappa} < \infty,
\end{aligned}$$

since $g \in H^{3+\kappa}(\mathbb{R}^2)$. Similarly,

$$\begin{aligned}
&\int_{\mathbb{R}^2} \|\omega\| \left|\mathcal{F}_g(\omega)\right| \mathrm{d}\omega \\
&= \int_{\mathbb{R}^2} \|\omega\| \left(1 + \|\omega\|^2\right)^{(1+\kappa)/2} \left|\mathcal{F}_g(\omega)\right| \left(1 + \|\omega\|^2\right)^{-(1+\kappa)/2} \mathrm{d}\omega \\
&\leq \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right) \left(1 + \|\omega\|^2\right)^{(1+\kappa)} \left|\mathcal{F}_g(\omega)\right|^2 \mathrm{d}\omega \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right)^{-(1+\kappa)} \mathrm{d}\omega \\
&= \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right)^{(2+\kappa)} \left|\mathcal{F}_g(\omega)\right|^2 \mathrm{d}\omega \cdot 2\pi \int_0^\infty \frac{r}{\left(1 + r^2\right)^{1+\kappa}} \mathrm{d}r \\
&= \int_{\mathbb{R}^2} \left(1 + \|\omega\|^2\right)^{(2+\kappa)} \left|\mathcal{F}_g(\omega)\right|^2 \mathrm{d}\omega \cdot \frac{\pi}{\kappa} < \infty,
\end{aligned}$$

since $g \in H^{3+\kappa}(\mathbb{R}^2) \subset H^{2+\kappa}(\mathbb{R}^2)$ by the Sobolev Embedding Theorem, Theorem B.11. □

Lemma A.11 (Integrability of the analytical Fourier-Mellin transform)**.** *Assume that* $g \in L^1(\mathbb{R} *_{>0} \times S^1)$ *and that*

$$\int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left(1 + \|(u,v)\|^2\right)^{2+\kappa} \left|\mathcal{M}_g(u,v)\right|^2 \, \mathrm{d}v < \infty$$

holds for some $\kappa > 0$. *Then we have*

$$\int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \|(u,v)\| \left|\mathcal{M}_g(u,v)\right| \, \mathrm{d}v < \infty.$$

Proof. Analogously to the previous Lemma A.10, we get

$$\begin{aligned}
&\int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \|(u,v)\| \left|\mathcal{M}_g(u,v)\right| \, \mathrm{d}v \\
&= \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \|(u,v)\| \left(1 + \|(u,v)\|^2\right)^{(1+\kappa)/2} \left|\mathcal{M}_g(u,v)\right| \left(1 + \|(u,v)\|^2\right)^{-(1+\kappa)/2} \, \mathrm{d}v \\
&\leq \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \|(u,v)\|^2 \left(1 + \|(u,v)\|^2\right)^{(1+\kappa)} \left|\mathcal{M}_g(u,v)\right|^2 \, \mathrm{d}v \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left(1 + \|(u,v)\|^2\right)^{-(1+\kappa)} \, \mathrm{d}v \\
&= \int_{\mathbb{R}} \sum_{u \in \mathbb{Z}} \left(1 + \|(u,v)\|^2\right)^{(2+\kappa)} \left|\mathcal{M}_g(u,v)\right|^2 \, \mathrm{d}v \cdot \frac{\pi}{\kappa} < \infty.
\end{aligned}$$

□

Lemma A.12 (Lemma A.17 in Hartmann (2016))**.** *For* $d \in \mathbb{N}$ *and a probability space* $(\Omega, \mathcal{A}, \mathbb{P})$*, let* $(X_T)_{T \in \mathbb{N}}$ *a sequence of random matrices* $X_T : \Omega \to \mathbb{R}^{d \times d}$ *and* $X \in \mathbb{R}^{d \times d}$ *such that* X *is invertible and* $X_T \overset{T \to \infty}{\longrightarrow} X$ *in probability. Furthermore, let* $(Y_T)_{T \in \mathbb{N}}$ *a sequence of random vectors in* $\mathbb{R}^d$*, such that* $(X_T Y_T)_{T \in \mathbb{N}}$ *is uniformly tight. Then,* $(Y_T)_{T \in \mathbb{N}}$ *is uniformly tight.*

Lemma A.13 (Lemma A.16 in Hartmann (2016))**.** *Let* $(\lambda_T)_{T \in \mathbb{N}}$ *a sequence in* $(0, \infty)$ *and* $\lambda \in (0, \infty)$ *such that* $\lambda_T \overset{T \to \infty}{\longrightarrow} \lambda$*. Furthermore, let* $X_T \sim \mathcal{N}(0, \lambda_T)$ *and* $X \sim \mathcal{N}(0, \lambda)$*. Then,* $X_T \overset{T \to \infty}{\longrightarrow} X$ *in distribution.*

APPENDIX B

Known theorems from the literature

In this second appendix, for the readers convenience, some of the theorems from the literature, which are applied in this thesis, are stated explicitly.

Theorem B.1 (Cauchy-Schwarz inequality, Theorem 2.2.7 in Hassani (2013))**.**
Let $g_1, g_2 \in L^2(\mathbb{R}^2)$. Then their product is integrable and

$$\left| \int_{\mathbb{R}^2} g_1(x) g_2(x) \mathrm{d}x \right|^2 \leq \int_{\mathbb{R}^2} |g_1(x)|^2 \, \mathrm{d}x \int_{\mathbb{R}^2} |g_2(x)|^2 \, \mathrm{d}x,$$

with equality if and only if g_1 and g_2 are not linearly independent.

Theorem B.2 (Plancherel, Theorem 1.6.1 in Rudin (1990))**.** *Let $d \in \mathbb{N}$. There is an isometry $\Psi\colon L^2\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda\big) \to L^2\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda\big)$ which is unitary (i.e., for all $g_1, g_2 \in L^2\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda\big)$ we have that $\langle \Psi g_1, \Psi g_2 \rangle = \langle g_1, g_2 \rangle$) and uniquely defined by $\Psi(g) = \mathcal{F}_g$ for all $g \in \mathcal{S}(\mathbb{R}^d)$, where the* Schwartz space *$\mathcal{S}(\mathbb{R}^d)$ is a certain subset of the set of smooth functions that is dense in $L^p\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda\big)$ for all $p \in [0, \infty)$.*

Theorem B.3 (Continuity of parameter integrals, Theorem 128.1 in Heuser (1995))**.** *Let X a metric space, E a Banach space, and $(\Omega, \mathcal{A}, \mu)$ a measure space. Let $g\colon X \times \Omega \to E$ with*

(i) $\omega \mapsto g(x, \omega)$ is in μ-measurable for all $x \in X$,

(ii) $x \mapsto g(x, \omega)$ is continuous μ-a.e. on Ω,

(iii) there is an $h \in \mathcal{L}^1(\Omega, \mu, E)$ such that $|g(x, \omega)| \leq h(\omega)$ for all $(x, \omega) \in X \times \Omega$.

Then, $G\colon X \to E$, $x \mapsto \int_\Omega g(x, \omega)\, \mu(\mathrm{d}\omega)$ is well-defined and continuous.

Theorem B.4 (Differentiability of parameter integrals, Theorem 128.2 in Heuser (1995))**.** *Let $d \in \mathbb{N}$, $U \subseteq \mathbb{R}^d$ open, E a Banach space, and $(\Omega, \mathcal{A}, \mu)$ a measure space. Let $g\colon U \times \Omega \to E$ with*

(i) $\omega \mapsto g(x, \omega)$ is in $\mathcal{L}^1(\Omega, \mu, E)$ for all $x \in U$,

(ii) $x \mapsto g(x, \omega)$ is continuously differentiable μ-a.e. on Ω,

(iii) there is an $h \in \mathcal{L}^1(\Omega, \mu, E)$ such that for all $(x, \omega) \in U \times \Omega$, $i \in \{1, \ldots, d\}$,

$$\left| \frac{\partial g}{\partial x_i}(x, \omega) \right| \leq h(\omega).$$

Then, $G\colon U \to E$, $x \mapsto \int_\Omega g(x,\omega)\,\mu(\mathrm{d}\omega)$ is well-defined and continuously differentiable such that for all $(x,\omega) \in U \times \Omega$, $i \in \{1,\dots,d\}$,

$$\frac{\partial G}{\partial x_i}(x) = \int_\Omega \frac{\partial g}{\partial x_j}(x,\omega)\,\mu(\mathrm{d}\omega).$$

Corollary B.5 (Cramér-Wold Device, p. 16 in van der Vaart (2000))**.** *Let for $d \in \mathbb{N}$, $(X_T)_{T\in\mathbb{N}}$ be a sequence of random vectors in $\mathbb{R}^d$, and X a random vector in $\mathbb{R}^d$. Then, $X_T \xrightarrow{\mathcal{D}} X$ if and only if $\langle \xi, X_T\rangle \xrightarrow{\mathcal{D}} \langle \xi, X\rangle$ for all $\xi \in \mathbb{R}^d$.*

Theorem B.6 (Theorem 5.7 in van der Vaart (2000))**.** *Let $\Theta \subseteq \mathbb{R}^d$, $\theta_0 \in \Theta$, and consider the function $M\colon \Theta \to \mathbb{R}$. Furthermore, let $(Y_T)_{T\in\mathbb{N}}$ a sequence of random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in a measure space $(\Omega', \mathcal{A}')$. Let $m\colon \Theta \times \Omega' \to \mathbb{R}$ a function such that $y \mapsto m(\theta, y)$ is measurable for all $\theta \in \Theta$ and $\theta \mapsto m(\theta, y)$ is continuous for all $y \in \Omega'$. For all $T \in \mathbb{N}$, define*

$$M_T\colon \Theta \to \mathbb{R}, \quad \theta \mapsto m(\theta, Y_T).$$

Let $(\hat\theta_T)_{T\in\mathbb{N}}$ a sequence of estimators for θ. Assume that

$$\sup_{\theta\in\Theta} |M_T(\theta) - M(\theta)| \overset{T\to\infty}{\longrightarrow} 0 \quad \text{in probability}, \tag{B.1}$$

$$\inf\{M(\theta) \mid \theta \in \Theta, \|\theta - \theta_0\| \geq \epsilon\} > M(\theta_0) \quad \text{for all } \epsilon > 0, \tag{B.2}$$

$$\limsup_{T\to\infty}\big(M_T(\hat\theta_T) - M_T(\theta_0)\big) \leq 0. \tag{B.3}$$

Then, $\hat\theta_T \overset{T\to\infty}{\longrightarrow} \theta_0$ in probability.

Theorem B.7 (Delta method, Theorem 3.8 in van der Vaart (2000))**.** *Let $(\hat\mu_T)_{T\in\mathbb{N}}$ a sequence of random vectors in $\mathbb{R}^d$, $\mu_0 \in \mathbb{R}^d$, and $\Sigma \in \mathbb{R}^{d\times d}$, such that*

$$\sqrt{T}(\hat\mu_T - \mu_0) \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0, \Sigma) \quad \text{in distribution.}$$

Let $k \in \mathbb{N}$ and let $g\colon \mathbb{R}^d \to \mathbb{R}^k$ continuously differentiable. Then

$$\sqrt{T}\big(g(\hat\mu_T) - g(\mu_0)\big) \overset{T\to\infty}{\longrightarrow} \mathcal{N}(0, J_g(\mu_0)\Sigma J_g(\mu_0)^\top) \quad \text{in distribution,}$$

where $J_g(\mu_0) \in \mathbb{R}^{k\times d}$ is the Jacobi matrix of g at μ_0.

Theorem B.8 (Continuous mapping theorem, Theorem 2.3 in van der Vaart (2000))**.**
Let $d, k \in \mathbb{N}$ and let $g\colon \mathbb{R}^d \to \mathbb{R}^k$ continuous a.e. on $\mathbb{R}^d$. Furthermore, let $(X_T)_{T\in\mathbb{N}}$ a sequence of random vectors in R^d and X a random vector in $\mathbb{R}^d$.

1. *If $X_T \overset{T\to\infty}{\longrightarrow} X$ in distribution, then $g(X_T) \overset{T\to\infty}{\longrightarrow} g(X)$ in distribution.*
2. *If $X_T \overset{T\to\infty}{\longrightarrow} X$ in probability, then $g(X_T) \overset{T\to\infty}{\longrightarrow} g(X)$ in probability.*
3. *If $X_T \overset{T\to\infty}{\longrightarrow} X$ almost surely, then $g(X_T) \overset{T\to\infty}{\longrightarrow} g(X)$ almost surely.*

Theorem B.9 (Slutzky's Lemma, Theorem 2.8 in van der Vaart (2000))**.** *For $d \in \mathbb{N}$, let $(X_T)_{T\in\mathbb{N}}$ and $(Y_T)_{T\in\mathbb{N}}$ sequences of random vectors in R^d, $(Z_T)_{T\in\mathbb{N}}$ a sequence of random variables in $\mathbb{R}$, X a random vector in $\mathbb{R}^d$, $c \in \mathbb{R}^d$, and $c' \in \mathbb{R}$. Assume that $X_T \overset{T\to\infty}{\longrightarrow} X$ in distribution, $Y_T \overset{T\to\infty}{\longrightarrow} c$ in probability, and $Z_T \overset{T\to\infty}{\longrightarrow} c'$ in probability. Then,*

1. *$X_T + Y_T \overset{T\to\infty}{\longrightarrow} X + c$ in distribution,*
2. *$Z_T X_T \overset{T\to\infty}{\longrightarrow} c'X$ in distribution,*
3. *if $Z_T \neq 0$ almost surely and $c' \neq 0$, $Z_T^{-1} X_T \overset{T\to\infty}{\longrightarrow} (c')^{-1} X$ in distribution.*

Theorem B.10 (Markov's inequality, Example 2.6 in van der Vaart (2000))**.** *Let $d \in \mathbb{N}$, $m, p > 0$, and let $(X_T)_{T\in\mathbb{N}}$ a sequence of random vectors in $\mathbb{R}^d$. Then,*

$$\mathbb{P}\big(\|X_T\| > m\big) \leq \frac{\mathbb{E}\big(\|X_T\|^p\big)}{m^p}.$$

Theorem B.11 (Sobolev Embedding, Proposition 2 of Chapter 2.3 in Triebel (1983))**.** *For $s \in \mathbb{R}$ and $p \in (1, \infty)$, let $H_p^s(\mathbb{R}^n)$ be the Bessel potential spaces (which coincide with the Sobolev spaces as defined in this thesis for $p = 2$). Then, for $\epsilon > 0$,*

$$H_p^{s+\epsilon}(\mathbb{R}^n) \subset H_p^s(\mathbb{R}^n)$$

is a continuous embedding.

APPENDIX C

Additional results of the simulation study

This section contains the reconstruction results obtained in our simulation study, which were not shown in the main text body to avoid lengthening it unnecessarily. They vary in the number of binned frames ($T = 100$ or $T = 200$), the size of the pixel grid (128×128 or 256×256), and the polynomial degree of the motion functions (linear or quadratic). Each figure is build as the corresponding figures in Section 4.1, showing the original image on the left, results for the Gaussian model in the middle column, and for the Poisson binomial model in the right column. The first row shows a single binned frame, the second row the overlay of all frames, the third row contains the final image estimator and the last row displays the average over all image estimators from 100 simulation runs.

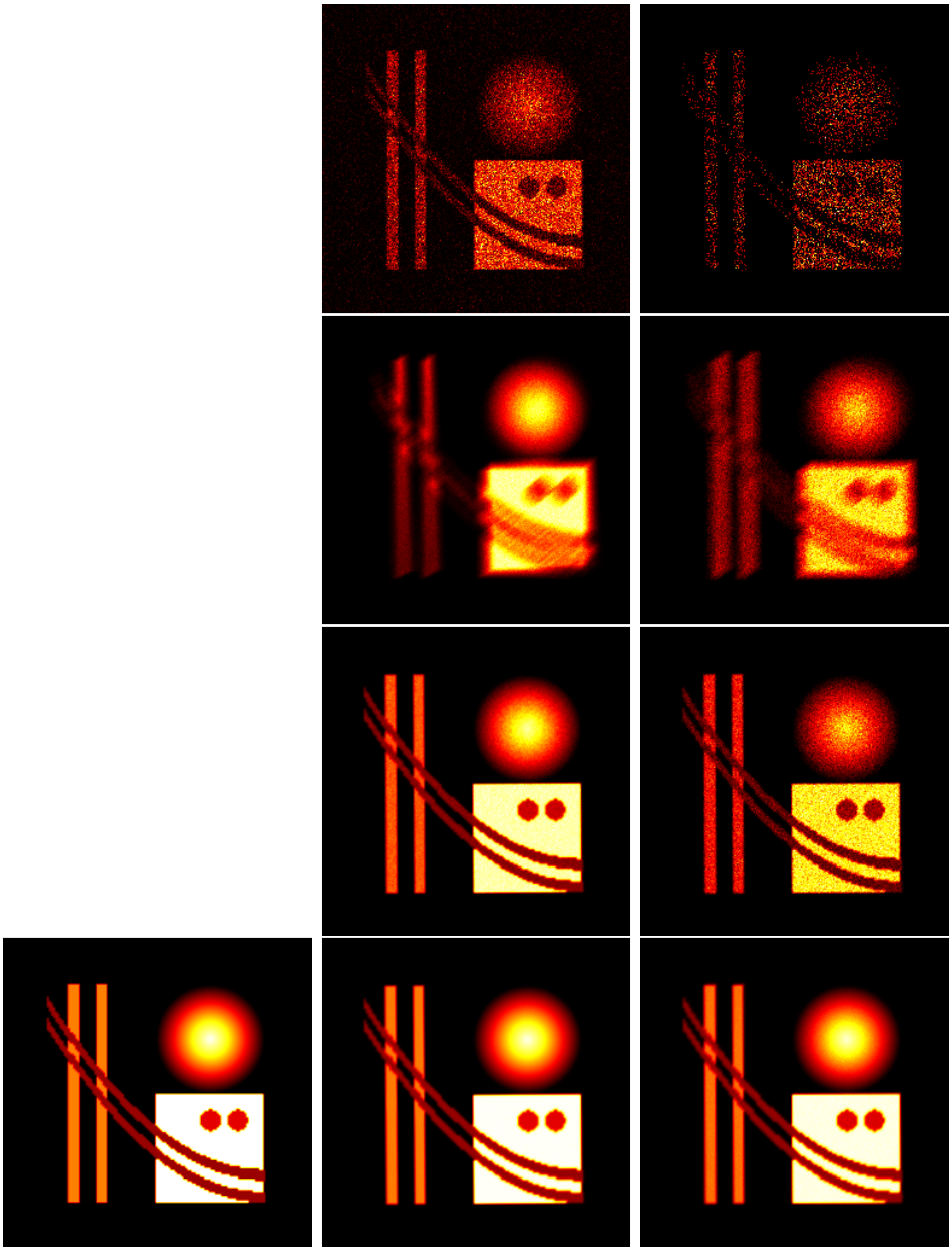

Figure C.1: Image reconstructions of the simulation study for linear motion model with $T = 100$ binned frames on a 256×256 pixel grid: true underlying image (left); for Gaussian (middle column) and Poisson binomial (right column) model, a single binned frame (first row), the blurred superpositions of all frames (second row), final image estimates, which are corrected for rotation, scaling and translational drift (third row) and average over images estimates from 100 simulation runs (fourth row).

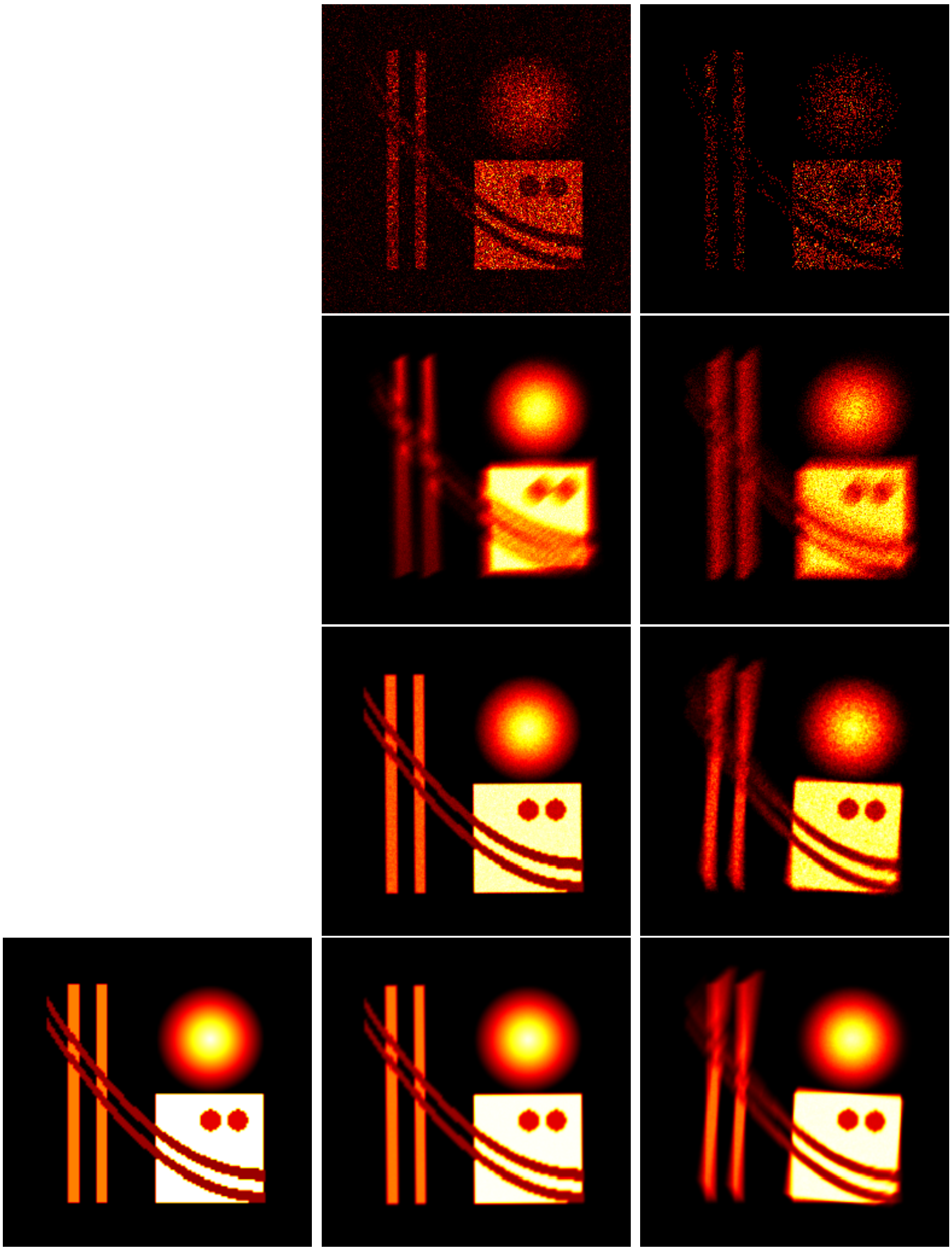

Figure C.2: Image reconstructions of the simulation study for quadratic motion model with $T = 100$ binned frames on a 256×256 pixel grid: true underlying image (left); for Gaussian (middle column) and Poisson binomial (right column) model, a single binned frame (first row), the blurred superpositions of all frames (second row), final image estimates, which are corrected for rotation, scaling and translational drift (third row) and average over images estimates from 100 simulation runs (fourth row).

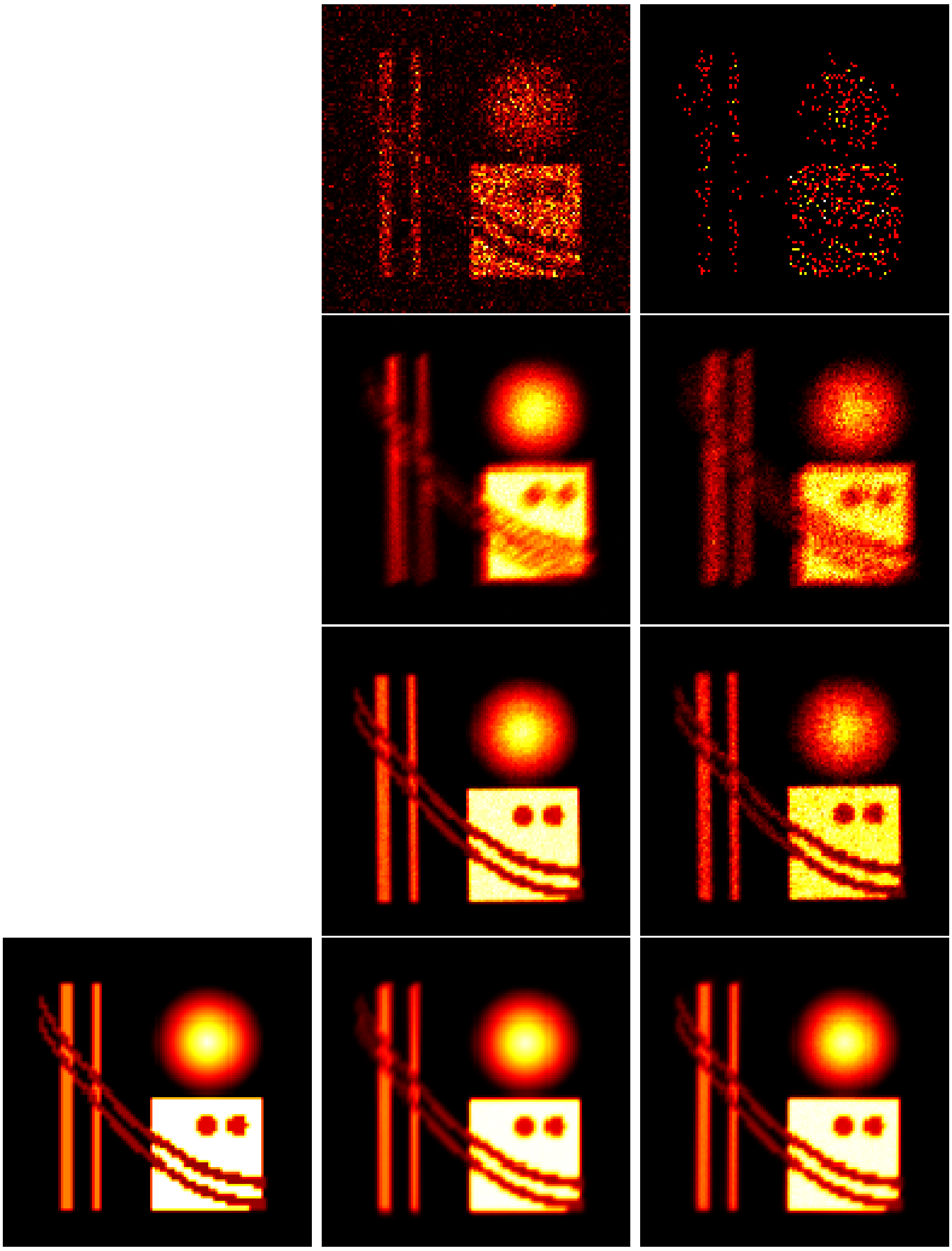

Figure C.3: Image reconstructions of the simulation study for linear motion model with $T = 200$ binned frames on a 128×128 pixel grid: true underlying image (left); for Gaussian (middle column) and Poisson binomial (right column) model, a single binned frame (first row), the blurred superpositions of all frames (second row), final image estimates, which are corrected for rotation, scaling and translational drift (third row) and average over images estimates from 100 simulation runs (fourth row).

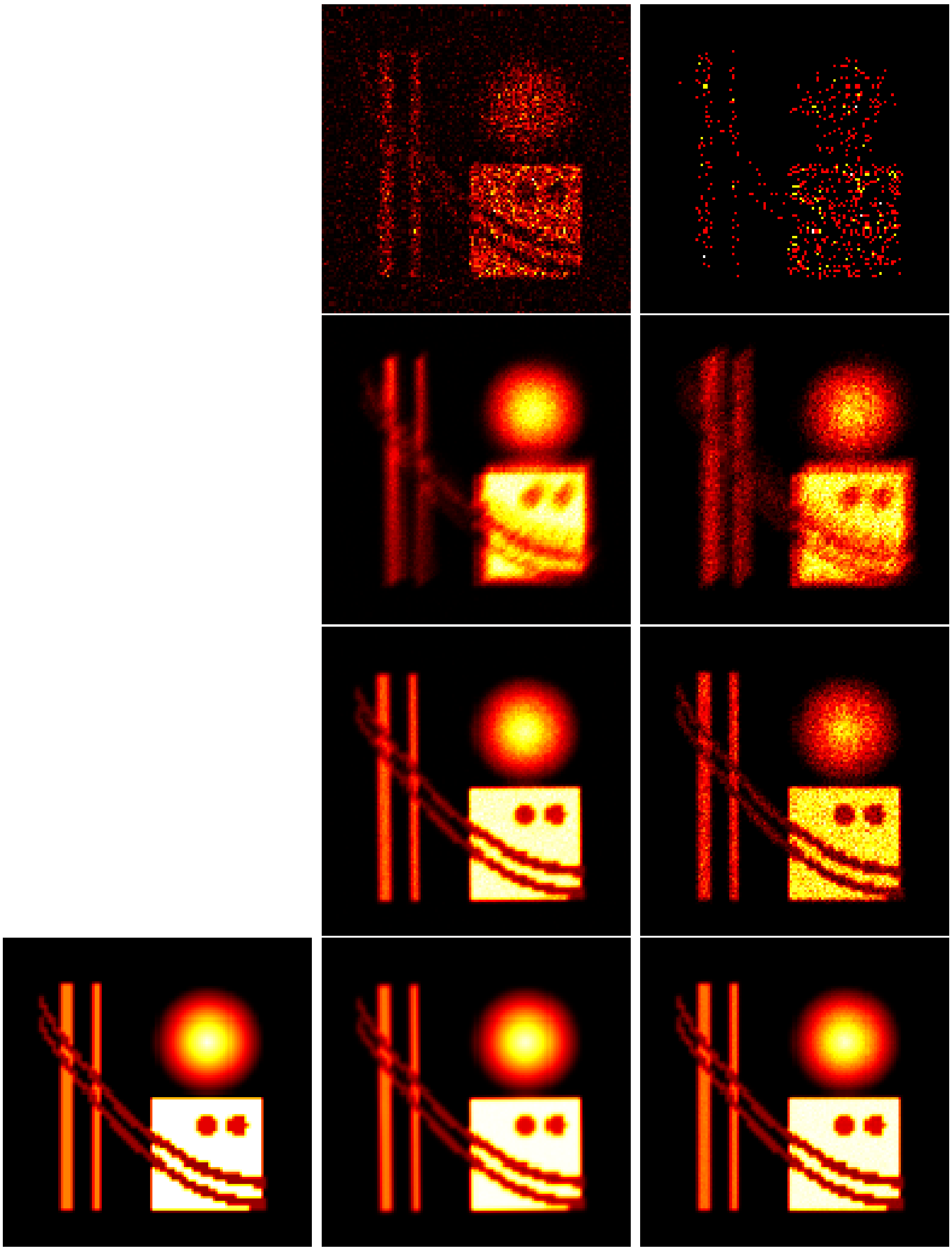

Figure C.4: Image reconstructions of the simulation study for quadratic motion model with $T = 200$ binned frames on a 128×128 pixel grid: true underlying image (left); for Gaussian (middle column) and Poisson binomial (right column) model, a single binned frame (first row), the blurred superpositions of all frames (second row), final image estimates, which are corrected for rotation, scaling and translational drift (third row) and average over images estimates from 100 simulation runs (fourth row).

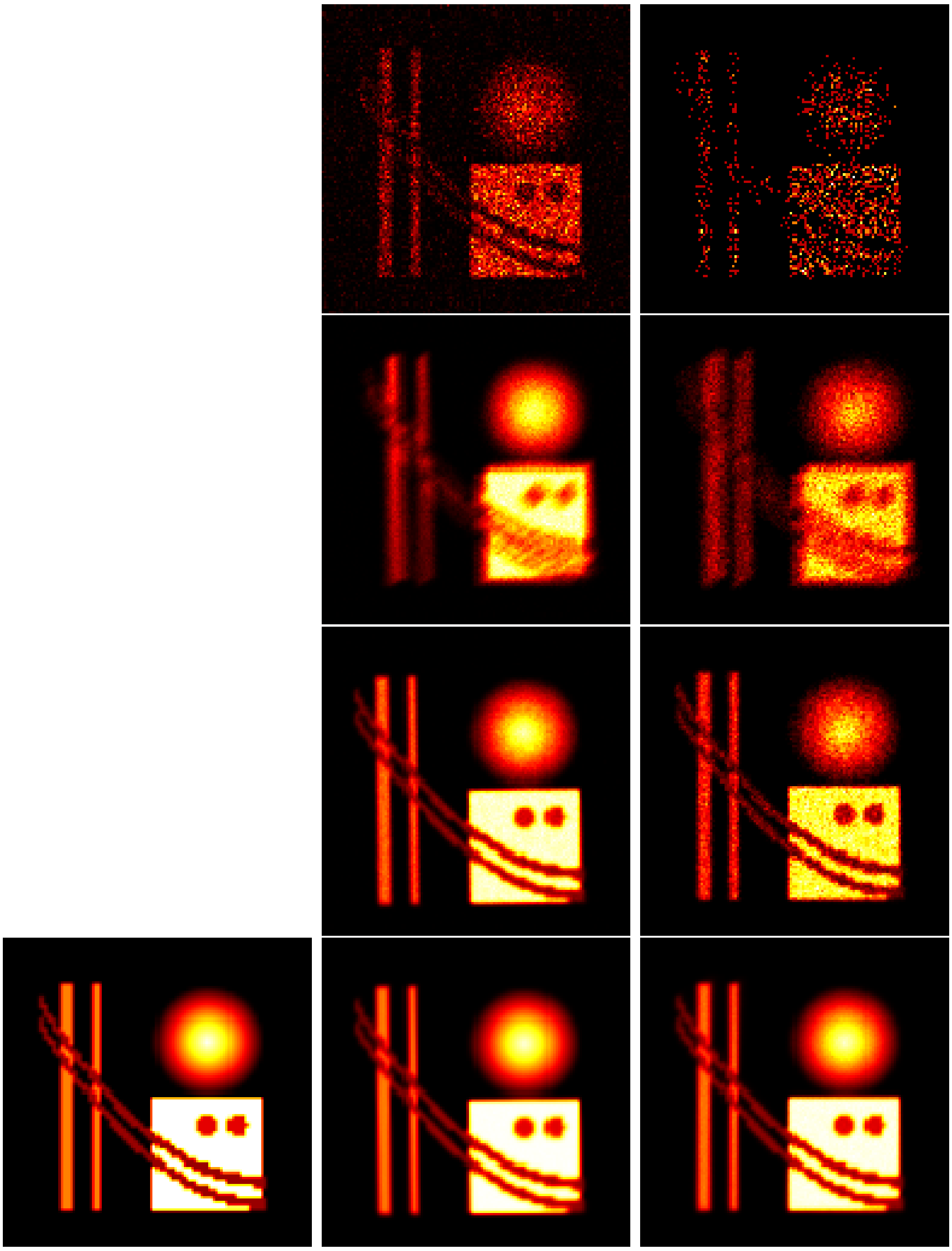

Figure C.5: Image reconstructions of the simulation study for linear motion model with $T = 100$ binned frames on a 128×128 pixel grid: true underlying image (left); for Gaussian (middle column) and Poisson binomial (right column) model, a single binned frame (first row), the blurred superpositions of all frames (second row), final image estimates, which are corrected for rotation, scaling and translational drift (third row) and average over images estimates from 100 simulation runs (fourth row).

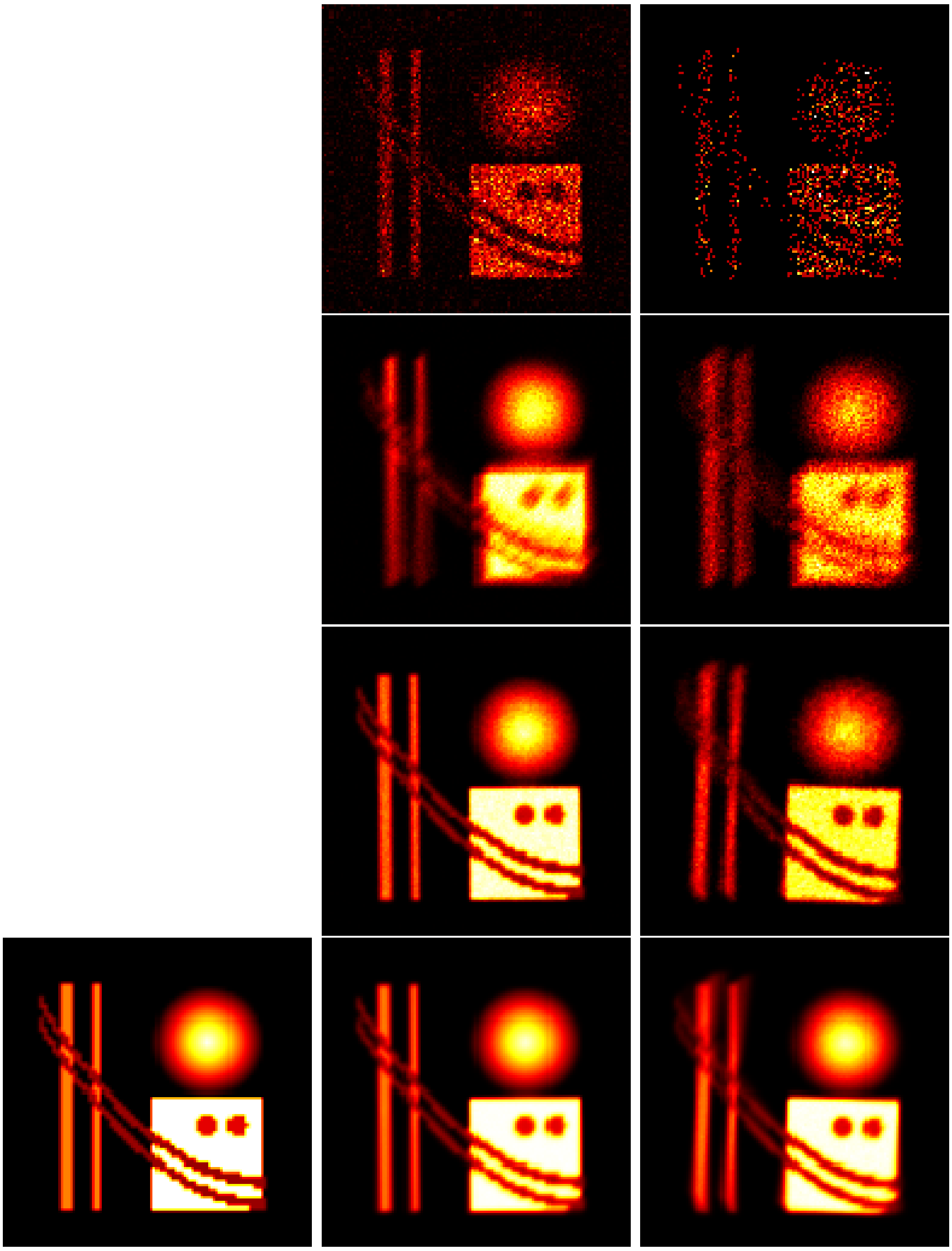

Figure C.6: Image reconstructions of the simulation study for quadratic motion model with $T = 100$ binned frames on a 128×128 pixel grid: true underlying image (left); for Gaussian (middle column) and Poisson binomial (right column) model, a single binned frame (first row), the blurred superpositions of all frames (second row), final image estimates, which are corrected for rotation, scaling and translational drift (third row) and average over images estimates from 100 simulation runs (fourth row).

Bibliography

Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für mikroskopische Anatomie*, 9(1):413–468.

Amann, H. and Escher, J. (2001). *Analysis III*. Grundstudium Mathematik, Analysis. Birkhäuser Basel, 1 edition.

Anscombe, F. J. (1948). The Transformation of Poisson, Binomial and Negative-Binomial Data. *Biometrika*, 35(3/4):246–254.

Aspelmeier, T., Egner, A., and Munk, A. (2015). Modern statistical challenges in high-resolution fluorescence microscopy. *Annu. Rev. Statist. Appl.*, 2:163–202.

Belov, V. N., Bossi, M. L., Fölling, J., Boyarskiy, V. P., and Hell, S. W. (2009). Rhodamine spiroamides for multicolor single-molecule switching fluorescent nanoscopy. *Chemistry - A European Journal*, 15:10762–10776.

Berning, S., Willig, K. I., Steffens, H., Dibaj, P., and Hell, S. W. (2012). Nanoscopy in a living mouse brain. *Science*, 335(6068):551.

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645.

Bigot, J., Gamboa, F., and Vimond, M. (2009). Estimation of translation, rotation, and scaling between noisy images using the Fourier-Mellin transform. *SIAM J. Imaging Sci.*, 2(2):614–645.

Billingsley, P. (1995). *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 3 edition.

Born, M. and Wolf, E. (1999). *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Cambridge University Press, Cambridge ; New York, 7th expanded ed edition.

Brown, L., Cai, T., Zhang, R., Zhao, L., and Zhou, H. (2010). The root–unroot algorithm for density estimation as implemented via wavelet block thresholding. *Probability Theory and Related Fields*, 146(3-4):401–433.

Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. *Math. Comput.*, 19:297–301.

Derrode, S. and Ghorbel, F. (2004). Shape analysis and symmetry detection in gray-level objects using the analytical fourier-mellin representation. *Signal Process.*, 84:25–39.

Deschout, H., Zanacchi, F. C., Mlodzianoski, M., Diaspro, A., Bewersdorf, J., Hess, S. T., and Braeckmans, K. (2014). Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat. Methods*, 11(3):253–266.

Egner, A., Geisler, C., von Middendorff, C., Bock, H., Wenzel, D., Medda, R., Andresen, M., Stiel, A., Jakobs, S., Eggeling, C., Schönle, A., and Hell, S. W. (2007). Fluorescence nanoscopy in whole cells by asynchronous localization of photoswitching emitters. *Biophys. J.*, 93(9):3285–3290.

Gamboa, F., Loubes, J.-M., and Maza, E. (2007). Semi-parametric estimation of shifts. *Electron. J. Statist.*, 1:616–640.

Gauthier, J. P., Bornard, G., and Silbermann, M. (1991). Motion and pattern analysis: Harmonic analysis on motion groups an their homogeneous spaces. *IEEE Trans. Syst., Man, Cyber.*, 21(1):159–172.

Geisler, C., Hotz, T., Schönle, A., Hell, S. W., Munk, A., and Egner, A. (2012). Drift estimation for single marker switching based imaging schemes. *Opt. Express*, 20(7):7274–7289.

Geisler, C., Schönle, A., von Middendorff, C., Bock, H., Eggeling, C., Egner, A., and Hell, S. W. (2007). Resolution of $\lambda/10$ in fluorescence microscopy using fast single molecule photo-switching. *Appl. Phys. A*, 88(2):223–226.

Ghorbel, F. (1994). A complete invariant description for gray-level images by the harmonic analysis approach. *Pattern Recogn. Lett.*, 15(10):1043–1051.

Grover, G., Mohrman, W., and Piestun, R. (2015). Real-time adaptive drift correction for super-resolution localization microscopy. *Optics Express*, 23(18):23887–23898.

Gustafsson, M. G. L. (2005). Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci. U.S.A.*, 102(37):13081–13086.

Hall, P. and Pittelkow, Y. E. (1990). Simultaneous bootstrap confidence bands in regression. *J. Statist. Comput. Simul.*, 37(1–2):99–113.

Hartmann, A. (2016). *Estimating rigid motion in sparse sequential dynamic imaging: with application to nanoscale fluorescence microscopy*. PhD thesis, University of Goettingen.

Hartmann, A., Huckemann, S., Dannemann, J., Laitenberger, O., Geisler, C., Egner, A., and Munk, A. (2015). Drift estimation in sparse sequential dynamic imaging, with application to nanoscale fluorescence microscopy. *J. R. Statist. Soc. B*, 78:563–587.

Hassani, S. (2013). *Mathematical Physics: A Modern Introduction to Its Foundations*. Springer International Publishing, 2 edition.

Heintzmann, R., Jovin, T. M., and Cremer, C. (2002). Saturated patterned excitation microscopy–a concept for optical resolution improvement. *J. Opt. Soc. Am. A, Opt. Image Sci. Vis.*, 19(8):1599–1609.

Hell, S. W. (2003). Toward fluorescence nanoscopy. *Nat. Biotechnol.*, 21(11):1347–1355.

Hell, S. W. (2007). Far-field optical nanoscopy. *Science*, 316(5828):1153–1158.

Hell, S. W. (2009). Microscopy and its focal switch. *Nat. Methods*, 6(1):24–32.

Hell, S. W. and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.*, 19(11):780–782.

Hess, S. T., Girirajan, T. P. K., and Mason, M. D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.*, 91(11):4258–4272.

Heuser, H. (1995). *Analysis 2*. Teubner, 9th edition.

Hofmann, M., Eggeling, C., Jakobs, S., and Hell, S. W. (2005). Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 102(49):17565–17569.

Holden, S. J., Uphoff, S., and Kapanides, A. N. (2011). Daostorm: An algorithm for high-density super-resolution microscopy. *Nat. Methods*, 8(4):279–280.

Huang, F., Hartwich, T. M. P., Rivera-Molina, F. E., Lin, Y., Duim, W. C., Long, J. J., Uchil, P. D., Myers, J. R., Baird, M. A., Mothes, W., Davidson, M. W., Toomre, D., and Bewersdorf, J. (2013). Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nat. Methods*, 10(7):653–658.

Jones, S. A., Shim, S.-H., He, J., and Zhuang, X. (2011). Fast, three-dimensional super-resolution imaging of live cells. *Nat. Methods*, 8(6):499–505.

Klar, T. A., Jakobs, S., Dyba, M., Egner, A., and Hell, S. W. (2000). Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc. Natl. Acad. Sci. U.S.A.*, 97(15):8206–8210.

Laitenberger, O. (2018). *Statistical analysis methods for time varying nanoscale imaging problems*. PhD thesis, University of Goettingen.

Le Cam, L. (1960). An approximation theorem for the Poisson binomial distribution. *Pac. J. Math.*, 10(4):1181–1197.

Lenz, R. (1990). *Group Theoretical Methods in Image Processing*. Lecture Notes in Comput. Sci. 413. Springer-Verlag, Berlin.

Li, H., Haltmeier, M., Zhang, S., Frahm, J., and Munk, A. (2014). Aggregated motion estimation for real-time MRI reconstruction. *Magn. Reson. Med.*, 72:1039–1048.

Rudin, W. (1990). *Fourier Analysis on Groups*. Wiley Classics Library Edition. John Wiley & Sons, New York. (Reprint of the 1962 original).

Rust, M. J., Bates, M., and Zhuang, X. W. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods*, 3(10):793–795.

Sahl, S. J., Hell, S. W., and Jakobs, S. (2017). Fluorescence nanoscopy in cell biology. *Nature Reviews Molecular Cell Biology*, 18(11):685–701.

Schmidt, R., Wurm, C. A., Jakobs, S., Engelhardt, J., Egner, A., and Hell, S. W. (2008). Spherical nanosized focal spot unravels the interior of cells. *Nat. Methods*, 5(6):539–544.

Segman, J. S., Rubinstein, J., and Zeevi, Y. Y. (1992). The canonical coordinates method for pattern deformation: Theoretical and computational considerations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(12):1171–1183.

Thompson, R., Larson, D., and Webb, W. (2002). Precise nanometer localization ana-lysis for individual fluorescent probes. *Biophys. J.*, 82(5):2775—-2783.

Triebel, H. (1983). *Theory of Function Spaces*. Modern Birkhäuser Classics. Birkhäuser Basel.

van der Vaart, A. W. (2000). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Westphal, V., Rizzoli, S. O., Lauterbach, M. A., Kamin, D., Jahn, R., and Hell, S. W. (2008). Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. *Science*, 320(5873):246–249.