

# ON LOCALIZATION AND MULTISCALE IN DATA ASSIMILATION

Dissertation for the award of degree "Doctor rerum  
naturalium" (Dr.rer.nat) of the  
Georg-August-Universität-Göttingen within the doctoral  
program mathematics of the Georg-August University  
School of Science (GAUSS)

submitted by  
Aamir Nadeem

Gojra, Toba Tek Singh  
Pakistan

Göttingen, 2017

## Thesis Advisers

Prof. Dr. Roland Potthast  
Deutscher Wetterdienst (DWD), Germany

Prof. Dr. Russel Luke  
Institute for Numerical and Applied Mathematics,  
University of Goettingen

## Thesis Referees

Reviewer: Prof. Dr. Roland Potthast  
Deutscher Wetterdienst (DWD), Germany

Second Reviewer: Prof. Dr. Gert Lube  
Institute for Numerical and Applied Mathematics,  
University of Goettingen, Germany

Further members of the Examination Board:

Prof. Dr. Russel Luke  
Institute for Numerical and Applied Mathematics,  
University of Goettingen, Germany

Prof. Dr. Ingo Witt  
Mathematisches Institut,  
University of Goettingen, Germany

Jun.-Prof. Dr.-Ing. Marcus Baum  
Institute of Computer Science,  
University of Goettingen, Germany

Jun.-Prof. Dr Anja Fischer  
Institute for Numerical and Applied Mathematics,  
University of Goettingen, Germany

Tag der mündlichen Prüfung 22.05.2017

# Contents

<b>1</b>	<b>Introduction</b>	<b>10</b>
<b>2</b>	<b>Tools and Methods</b>	<b>15</b>
2.1	Basic Theory . . . . .	15
2.1.1	Integral Operators . . . . .	22
2.1.2	Regularisation Theory . . . . .	28
2.1.3	Cycling, Tikhonov regularisation and 3D-VAR . . . . .	33
2.1.4	Variational Assimilation 4D-VAR . . . . .	37
2.2	Kalman Smoother and Kalman Filter . . . . .	42
2.3	Ensemble Methods and Ensemble Kalman Filter . . . . .	49
<b>3</b>	<b>Transformed and Generalized Localization</b>	<b>56</b>
3.1	Issues Due to Small Number of Ensembles . . . . .	57
3.2	Remedy for undersampling . . . . .	60
3.2.1	Covariance Inflation . . . . .	61
3.2.2	Localization . . . . .	61
3.3	The EnKF and Localization . . . . .	64
3.4	Transformed Localization . . . . .	67
3.4.1	Generalized localization concept . . . . .	69
3.4.2	Transform based on Singular Value Decomposition . . . . .	77
3.5	Transformed Localization for Atmospheric Radiance Inversion . . . . .	78
3.6	Conclusions . . . . .	84
<b>4</b>	<b>On Multiscale Data Assimilation</b>	<b>86</b>
4.1	A Generic Multiscale Approach . . . . .	88
4.1.1	A Multiscale Step in State Space . . . . .	90
4.1.2	A Multiscale Step in Observation Space . . . . .	92
4.1.3	The Role of SVD for the Transformed Equation . . . . .	94
4.2	Non-Equivalence and Iterations of the Sequential Multiscale Approach . . . . .	95

4.2.1	On the Equivalence and Non-Equivalence of Sequential Multiscale . . . . .	95
4.2.2	General Form of Aliasing Problems . . . . .	99
4.2.3	Convergence for Iterated Sequential Multiscale . . . . .	99
4.2.4	Regularized Sequential Multiscale . . . . .	101
4.3	Numerical Examples . . . . .	107
4.4	Conclusions . . . . .	110

*TO*

*ABU, AMI ,UNCLE MUMTAZ  
&  
AYESHA*

# Acknowledgement

All praises be to Almighty Allah, the most beneficent, the most merciful who bestowed his mercy upon me and gave me enough vision, wisdom, courage and above all an opportunity. My whole hearted gratitude to my Holy Prophet Hazrat Muhammad (PBUH) who is a real source of knowledge and guidance for the whole world now and all the times to come.

I feel privilege while expressing my gratitude to my respected Supervisor Prof. Dr. Roland Potthast. A man who was an inspiration through out the course. I want to sincerely thank him for his sympathetic and remarkable guidance, continuous encouragement, valuable and noteworthy suggestions without which it would have difficult for me to complete the task. He enabled me to partially witness the unending vista of knowledge. No words to thank him much for giving me an opportunity to work in the institute of Numerical and Applied Mathematics of George August university of Göttingen and Deutscher wetterdienst(DWD).

Deep appreciation is due to my second Advisor Prof. Dr. Russel Luke. I am extremely and immensely obliged on all the support he had given to me. It is a matter of immense pleasure to express sincere thanks to Prof. Dr. Gert Lube for his continuous encouragement and tireless support. I thanks to all committee members for their support

Many thanks to all of my colleagues at the institute Matthew Tam, Yurii Malitskiy, Anna-Lena, Hieu Thao and Frederic Weidling and Arian Berdellima for their moral support. A special thank goes to Dr. Jochen Schulz for his help in software learning. I would like to thank all the technical and administrative staff of the institute for providing the environment and basis to make effective work possible.

I owe my deep gratitude and heartfelt thanks to my respected parents who supported me throughout my career and guided me at every step in my life. I felt privileged to have continous guidance, encouragement and profound prayers of my uncle Dr. Mumtaz Ahmed Khan. A thank you note to my brother for unconditional support and care.

It give me immense pleasure to express my most vigorously, deep gratitude and heartfelt wishes to my wife Ayesha Aamir who always cared and guided me in all

aspect of life whenever I needed her help. Without her tremendous help and support, I could not have achieved the hard task to complete this thesis.

My acknowledgement will remain incomplete if I do not mention the companionship of my all friends who keep me socially alive, always around to support. A very special thanks to Shahid Manzoor, Muhammad Umair, Dr. Shakeal, Dr. Sanaullah, Dr. Tariq Javed, Mohsin Shafeeq, Zaheer Abbas, Muhammad Awais, Abdul-ur-Rehman, Hussnain, Sufyan Rasheed, Abdul Sattar, Muhammad Zubair, Maaz. I am also thankful to Mr. and Mrs. Itzaz Malik for their affection and care

Lastly, I would like to thank the Federal Urdu University and Higher Education Commission of Pakistan (HEC) for the funding I have received. Without this funding I would not have been able to undertake this research.

# Abstract

The main goal of *inverse problems* is the reconstruction of unknown quantities, often they are located in an inaccessible part of space. *Data assimilation*, which is the integration of inverse problems techniques with dynamical systems with the task of forecasting the evolution of some quantity, has evolved strongly, with important applications in meteorology and atmospheric sciences.

*Localization* is an essential part of ensemble based assimilation schemes. The size of an ensemble is always much smaller than the dimension of the state space for real numerical predictions. It is necessary to ensure a sufficient number of degrees of freedom when generating the analysis ensemble and, thus, increase the rank of the set of analysis equations. Also, the small ensemble size gives an insufficient estimate of the background error correlations. Localization effectively eliminates spurious correlations in the background ensemble between distant state variables.

The choice of the localization radius needs to depend on the number of ensemble members for the short-range forecasts used to calculate the background error for the analysis step as well as on the number of observations and the observation error. However, a challenge arises when the observation operator under consideration is *non-local* (e.g. satellite radiance data), the localization which is applicable to the problem can be severely limited, with strong effects on the quality of the assimilation step.

We study a transformation approach to change non-local operators to local operators in transformed space, such that localization becomes applicable. We interpret this approach as a generalized localization and study its general algebraic formulation. Examples are provided for a compact integral operator and a non-local matrix observation operator to demonstrate the feasibility of the approach and study the quality of the assimilation by transformation.

In particular, we apply the approach to temperature profile reconstruction from infrared measurements given by the IASI Infrared Sounder and show that the approach is feasible for this important data type in atmospheric analysis and forecasting. We also believe that our derivations will work in a similar way for particle filters.



If we do localization we have to adjust the localization radius to the scale of the problem. In this respect choosing large radius mean the data we have is important for large scale and vice versa. Thus localization is linked to the scale of the meteorological processes under consideration. Inherently when we choose localization, we chose scale. This leads to the topic of scales, which usually comes under the umbrella of *multiscale methods*.

The broad idea of a multiscale approach is to decompose your problem into different scales or levels and to use these decompositions either for constructing appropriate approximations or to solve smaller problems on each of these levels, leading to increased stability or increased efficiency.

Our goal is to analyse the *sequential multiscale approach* applied to an inversion or state estimation problem. We work in a generic setup given by a Hilbert space environment. We work out the analysis both for an *unregularized* and a *regularized* sequential multiscale inversion. In general the sequential multiscale approach is not equivalent to a full solution, but we show that under appropriate assumptions we obtain convergence of an *iterative* sequential multiscale version of the method. For the *regularized case* we develop a strategy to appropriately adapt the regularization when an iterative approach is taken.

We demonstrate the validity of the iterative sequential multiscale approach by testing the method on an integral equation as it appears for atmospheric temperature retrieval from infrared satellite radiances.

# Chapter 1

## Introduction

*Data Assimilation* (DA) for *Numerical Weather Prediction* (NWP) systems is the collection of algorithms and techniques used to estimate the state of the atmosphere at a given moment in time (referred to as the *analysis*), thus providing initial conditions for NWP models. The strong nonlinear nature of atmospheric dynamics makes DA critical in the pursue of accurate weather forecasting.

The goal of DA techniques in NWP is therefore to use the information coming from the different observation systems to infer the model state vector variables as well as the uncertainty in its components and correlations between these. This is normally achieved by sequentially updating a prior state estimate generated through a short range forecast (*first guess* or *background*) using the NWP model (which therefore contains information from past observations), with the newly available observations since the last update. As model biases and DA can interact and give rise to undesired feedback effects, it is important to assess any DA technique in a cycling environment that simulates the periodic updating needed by any operational forecasting system.

The *Kalman Filter* (KF) equations (introduced in [92]) provide explicit update equations in terms of a *Kalman gain matrix*, which multiplied by the difference between observations and the background in observation space, yields the increments to be added to the first guess state vector to obtain the analysis. Its equations also give expressions for the propagation of the estimate state and its covariance to the next time step. It is the optimal (in the sense of minimum variance) solution to the linearized DA problem under the assumptions of Gaussian error distributions and observation operators (used to transform the model state into its observation space equivalent).

The *Extended Kalman Filter* (EKF) uses the full nonlinear model and observation operators, while keeping the linearizations of both in computing background

covariance and Kalman gain (see for example [90]). This technique is however computationally very expensive and currently unfeasible in an operational context. Current DA frameworks of operational NWP centers around the world are based almost invariably in what are approximations or simplifications in one way or another of these equations. Variational methods for example, give up (at least most of) the uncertainty propagation, as the time iteration in the Kalman filter is suppressed (3D-VAR) or limited to the assimilation time window (4D-VAR).

Another approach to approximating the EKF equations is through the use of *Ensemble Data Assimilation* (EDA) techniques. The Ensemble Kalman Filter (EnKF) introduced in [53] and [82] is a Monte Carlo approximation of the KF. The key idea consists in choosing an ensemble of initial conditions which characterize not only the analysis (given by a linear combination of the ensemble members), but also its uncertainty (analysis error covariance). The full nonlinear model can be then used to evolve each ensemble member. In doing so, the dimensionality of the problem is reduced to that of the subspace spanned by the ensemble (which is less than the number of ensemble members). While computationally very efficient, this has the drawback of only being able to describe uncertainty in this subspace.

On top of providing a flow dependent characterization of the system's uncertainty, ensemble data assimilation techniques provide a natural initialization for ensemble prediction systems (EPS) (see [14, 62, 71, 108, 139]). Further more, it can be used in combination with other algorithms to provide an evolving covariance for the latter in what are known as *hybrid methods* (see for example [55, 94]). The Local Ensemble Transform Kalman Filter (LETKF) introduces spatial localization in the EnKF, and amounts to allowing for different linear combinations of the ensemble members for different grid points. The global analysis is then no longer confined to the ensemble space and spurious correlations between distant locations in the background covariance matrix due to the limited ensemble size are avoided. Different definitions and extensions have been proposed and studied in the last years, see for example [15, 28, 56, 57, 136, 162, 171]. While focused in the context of limited area DA, [114, 137, 173] is a good summary of the state of the art and recent advances and challenges in the area. [87] describes a computationally efficient implementation of an LETKF.

Since the work of Evensen 1994 [53], in meteorology *ensemble filters* have been under heavy development [15, 28, 56, 82, 171, 136]. Important centers in Numerical Weather Prediction are now using ensemble filters as their operational algorithms for assimilation or are preparing such use, for example Canada, the US, Japan, UK, Germany etc., e.g. [35, 36, 120, 157, 158, 172]. The further development of different versions of the ensemble Kalman filter for either global or regional NWP is one of the key tasks of current research in the data assimilation community, see for example

[69, 84, 85, 86, 87, 112, 27, 144, 145, 177].

Ensemble filters approximate the covariance of the assimilation system by a stochastic estimator. They are local low-dimensional approximations to the Kalman filter which can be efficiently used in large-scale systems. Here, we study the square root filter as described in [62], which is a different way to write [87]. When a low number of ensemble members is employed, *localization* is a key ingredient of the ensemble filter. It means that we carry out the analysis locally and then construct a global solution from the local analysis ensembles. Localization increases the effective number of degrees of freedom. Different localization methods and their analysis can be found for example in [16, 30, 34, 38, 87, 88, 99, 131, 144, 130, 177]. For this work, we will think of *localization* as a method to restrict a problem defined on a set  $\Omega$  to a subset  $U \subset \Omega$ . Any data outside of  $U$  is ignored, as well as any function defined outside of  $U$ , see equation (3.4.22) for more details. Then, the problem is solved for several subsets  $U$  covering the whole domain  $\Omega$ , and the total solution is obtained by a combination of the localized solutions. We show in (3.4.23) that this can be understood as a *projection method*.

As the first main contribution of this thesis, in chapter 3 we suggest a general algorithm which allows localization in the case of non-local observation operators. The basic idea is to employ a *transformation* of the spaces under consideration, i.e. the state space and the observations space, to make the observation operator more *local* in the transformed space. Then, localization is used in the transformed space and the solution is transformed back afterwards. The method is called *transformed* or *generalized localization*.

We test the *applicability* of transformed localization by studying the effect of the transformed localization when applied to a IASI retrieval problem [32, 148], which with broad sensitivity functions over the large parts of each atmospheric column for most of its channels is known to have a strong non-locality and is of high interest to both the research community and operational centers of weather prediction.

The choice of the localization radius is an important and difficult topic in data assimilation. A large localization radius means that the observations are relevant for the large scales of the dynamic problem under consideration. A small localization radius means that the observation is important for local or high-resolution phenomena. This leads to the topic of *scales*, which usually comes under the umbrella of *multiscale methods*. The basic scientific question to explore here is how the assimilation of data is influencing the different scales of the dynamical system. We understand localization as a special topic of the larger field of *multiscale data assimilation*. This leads us to the second main topic of this thesis.

In chapter 4 we study a *multiscale approach* to data assimilation. Multiscale approaches are very popular for e.g. partial differential equations and for many

applied fields. The idea of a multiscale approach is to decompose your problem into different scales or levels and to solve smaller problems on each of these levels. A generic approach is the *sequential multiscale approach* to an inversion step where you first solve the problem in a *large-scale subspace* and then successively move to *finer scale* spaces. Here, we restrict our analysis to *one* inversion or assimilation step. In this case an example could be chosen by taking an *integral equation of the first kind* (as the IASI retrieval problem above) for some inverse problem. It also applies to one single data assimilation cycling step, where such covariances naturally appear as a core ingredient of data assimilation algorithms. We will work in a generic setup given by a Hilbert space environment. Our main goal is the *understanding* of the properties of sequential multiscale inversion and the analysis of *convergence* when the sequential multiscale approach is *iterated*.

We will study the sequential and iterative sequential approaches by an analysis either in *state space* or *observation space*. Our analysis will study typical *aliasing problems* which appear in a multiscale framework. It is shown that in general the sequential multiscale approach is not equivalent to a full solution, but that under appropriate assumptions we obtain convergence of an *iterative* sequential multiscale version of the method. We first study the *unregularized case* and then treat the *regularized problem*. We provide a mathematical analysis of the situation and demonstrate its validity by numerical examples.

The part on generalized localization of this thesis has been published in journal *Mathematical Methods in Applied Sciences* [126] in 2016. The content of multiscale is under review in *Journal of Computational and Applied Mathematics* [127].

## Abbreviations

DWD	Deutscher Wetterdienst
DA	Data Assimilation
NWP	Numerical Weather Prediction
AIRS	Atmospheric Infrared Sounder
IASI	Infrared atmospheric sounding interferometer
AMSU	Advanced Microwave Sounding Unit
RTTOV	Radiative Transfer for TOVS
TOVS	Television operational vertical sounder
HIRS	High resolution infrared sounder
ATMS	Advanced technology microwave sounder
ICON	Icosahedral nonhydrostatic general circulation model
3D-VAR	Three-dimensional variational data assimilation scheme
4D-VAR	Four-dimensional variational data assimilation scheme
KF	Kalman Filter
SRF	Square Root Filter
EKF	Extended Kalman Filter
EDA	Ensemble Data Assimilation
EnKF	Ensemble Kalman Filter
EPS	Ensemble prediction systems
LETKF	Local Ensemble Transform Kalman Filter
EKSRF	Ensemble Kalman Square Root Filter
SVD	Singular value decomposition

# Chapter 2

## Tools and Methods

### 2.1 Basic Theory

The study of many classical mathematical modelling problems is heavily dependent on their boundary conditions and input data. These problems continuously exhibit dependence on the input and approximation error decreases in numerical approximation with the increase of grid points.

However, often we do not observe such kind of regular behaviour in solving the integral equations of first kind, or when we go to the inverse or reconstruction problems. Often, the solutions do not continuously depend on data – such that noise is amplified. Also, often they show non-uniqueness.

The rule for numerical algorithms that more efforts will lead to better results does not hold for solving inverse or reconstruction problems. Increasing the number of grid points for approximation often leads to more instability. Such solution inherit more and more noise which is carried on from the measured data. We use the term *ill-posed* to describe such kind of behaviour in inverse problems.

Many methods have been concocted and developed to deal with ill-posed problems. They make ill-posed problem more regular, that is they approximate them by stable, bounded and continuous algorithms or operators. These approaches or methods are called *regularisation* methods. For a selection from the large body of literature see [9, 37, 44, 50, 62, 68, 70, 77, 78, 95, 100, 103, 129, 159, 170, 133, 69].

Keller [98] proposed very general definition of an inverse problem which is usually used as standard in literature:

**DEFINITION 2.1.1 (Inverse Problem).** *We call two problems inverse of one another if the formulation of each involves all or part of the solution of the other. Often, for historical reasons, one of the two problems has been studied extensively for some*

time, while the other is newer and not so well understood. In such the cases, the former problem is called the *direct problem*, while the later is called the *inverse problem*.

Following Keller two problems are inverse to each other if the formation of each of them needs the partial or full knowledge of the other. The problem which is studied earlier and in more detail is usually called *direct* and the other is *inverse problem*. However, there is another more important difference between these two problems. It can be understand by the concept of *well-posedness* introduced by Hadamard. In his lectures [70] he argued that the mathematical model of some physical problem (he was thinking for boundary value problems for partial differential equation) has to be properly posed or well-posed.

Hadamard in 1923 postulated three properties which needed to be satisfied for what he called a *properly posed* or *well-posed* problem.

DEFINITION 2.1.2 (Well-posedness). *A problem is called well-posed if we have*

1. (*existence*) *there exists a solution to the problem,*
2. (*uniqueness*) *there is at most one solution to the problem,*
3. (*stability*) *the solution depends continuously on the data.*

So existence, uniqueness and stability makes problem well-posed. If any of the postulates is violated then the problem is not well-posed. A problem which is not well-posed is called *ill-posed* or *improperly-posed*. If the two problems are inverse of each other then the one which is ill-posed is called *inverse problem* and other is called *direct problem* or *forward problem*. Mostly, the direct problems are well-posed problems.

As inverse problem is set in a way that it is inverse of a direct problem. Often, the direct problem is to find some quantity  $f$  from some known causes and conditions  $\varphi$ , under the physical or mathematical model  $H$  and described by the relation (or operator equation)

$$H(\varphi) = f \tag{2.1.1}$$

The operator  $H$  is generally nonlinear and describes the equations which connect the model variables to observed data. Hence, solving the inverse problem we infer knowledge from some known measurement  $f$  about the unknown quantity  $\varphi$ , which may be a specific cause, condition, state or parameter of the mathematical model. The solution of inverse problem is called the *reconstruction* or *retrieval* of  $\varphi$  from data  $f$ . We study specific types of inverse problems which are to be solved in the analysis step of data assimilation. They usually come with some dynamic aspect.



The phrase *to assimilate* conjures up a variety of meanings depending on its use in biological, social and physical sciences. With all of its uses and meanings the word embraces the concept of incorporation. For our case we expand its meaning by affixing the words dynamic and data. Here, *dynamic* implies use of some law or a set of equations, typically physical laws. Now, we have dynamic laws, we have data and we assimilate. Hence, it is the incorporation of data into laws that capture the meaning of *dynamic data assimilation*. So the aim of data assimilation (and the inverse problem) is to blend different information sources in order to get the best possible estimate of the state. The sources generally rely on observations of the system and, further, the physical laws describing its behaviour and, more specifically, a particular numerical model.

We cannot take observations only, because often observations are too sparse or partial in geophysics; so some additional information is needed to interpolate the information in observations to uncover the function or quantity. We need the numerical model to perform this task. Second, observations contain errors (they are noisy). Combining (by means of the model) several noisy observations can be an efficient way to filter out at least part of the noise and to provide a more accurate estimate [7, 69, 108].

Environmental phenomena can be realistically described by mathematical and numerical models of the system dynamics. These models have ability to predict the future behaviour of the system, provided the *initial state* of the system is known. As mentioned above, the availability of the data that describe complete states of system at specific time is very rare. The inaccuracies in the models and in given prior data along with random noise can drag the predicted states away from actual states of the system. To overcome this case, observations are incorporated with model equations over the time to obtain improved estimates along with the uncertainty in the estimates.

The problem of state-estimation is an inverse problem and can be treated using *observers* and/or *filters* derived by *feedback design techniques* [24, 133], however several traditional technique for state-estimation are not practicable for large non-linear systems which arise in environmental science. So new data assimilation schemes have been developed to carry out accurate state estimation. The purpose of these data assimilation schemes is to take measured observations in combination with the model dynamics in order to derive accurate estimation of present and future state of the system, alone with estimation of uncertainty in estimated state.

The most important aspect of *data assimilation problems* is that we need to deal with *large non-linear models* with a large number of state variables of order  $O(10^7 - 10^9)$ . The dynamics are multi-scale and often unstable and/or chaotic. The observations are also very large of order  $O(10^5 - 10^7)$  over a specific time

window. Usually, the distribution of data is not even in space or time, often there are areas where measurements are missing. With all these side conditions, the data assimilation problem is ill-posed in the sense of both non-uniqueness and non-stability, i.e. the state estimate is sensitive to errors.

For this problem the literature describes two basic approaches. The first works with 'dynamic observations', it is formulated as a *sequential data assimilation scheme*, while the second uses 'direct observations' and formulates a *four-dimensional data assimilation scheme*. In the first approach the observations are passed to the model dynamically at each time these are available, they produce best estimate and are utilized to predict future states. In the second a trajectory is found that best fits the observation over the time window. The state at the end of the window is used to calculate the forecast. Both solve same optimal state estimation problem under certain mathematical assumptions.

In operational systems it is difficult to solve full optimization problems in real time, therefore different approximations are employed. Further tools such as covariance inflation and covariance localization are important approximation methods within modern ensemble data assimilation data assimilation schemes. Ensemble assimilation schemes can be employed for many different models. Next we describe the mathematical formulation of data assimilation problem with techniques described in [69, 133, 129].

We use the letters  $X$  and  $Y$  for the state and observation space. We assume that  $X$  and  $Y$  are Hilbert spaces unless otherwise stated. Let  $\varphi \in X$  be the state vector containing all the state variables and  $\varphi_k \in X$  be the state at time  $t_k$ . Furthermore, the model operator at time  $t_k$  is denoted as  $M_k : X \rightarrow X$ . It evolves the state from time  $t_k$  to  $t_{k+1}$ , that is

$$\varphi_{k+1} = M_k(\varphi_k).$$

The model evolution from time  $t_l$  to  $t_k$  is denoted as

$$M_{k,\ell} = M_{k-1}M_{k-2} \dots, M_{\ell+1}M_\ell, \quad k > \ell \in \mathbb{N}_0 \quad (2.1.2)$$

Let  $f_k \in Y_k$  is the observation vector containing all observation at time  $t_k$  and  $H_k : X_k \rightarrow Y_k$  be the observation operator which maps the state space to the observation space at time  $t_k$ . Then, the data assimilation problem is defined as

DEFINITION 2.1.3 (Data assimilation problem). *Given observations  $f_k \in Y_k$  at time  $t_k$  to determine the state  $\varphi_k \in X$  from the operator equations*

$$H_k(\varphi_k) = f_k, \quad k = 0, 1, 2, \dots \quad (2.1.3)$$

*subject to model dynamics  $M_k : X \rightarrow X$  given by  $\varphi_{k+1} = M_k(\varphi_k)$ ,  $k = 0, 1, 2, \dots$ ,*

In NWP the model operator  $M_k$  involves the nonsteady solution of nonlinear partial differential equation. The observation operator  $H_k$  changes at each time step, but for simplicity here we consider the case where  $H$  is constant over time, i.e.  $H_k := H$ . The observation operator  $H_k$  and measurement  $f_k$  are contaminated. The model  $M_k$  does not represent true dynamic because of model errors also nonlinear models are very chaotic. So in context of data assimilation we need extra information from some known *prior distributions* (also known as *first guess distributions* or *background distributions*) about the state variable  $\varphi_k \in X$ .

The operator equation (2.1.3) (also 2.1.1) is ill-posed in the sense that one or more than one Hadamard conditions in Definition 2.1.2 are not satisfied. The above Definition 2.1.2 of well-posedness in the setting of mathematical notation for operator equations is as follows.

DEFINITION 2.1.4. *Let  $X$  and  $Y$  be normed spaces and  $H : X \rightarrow Y$  a nonlinear mapping. Then, the operator equation  $H(\varphi) = f$  is well-posed if the following holds*

- *Existence: For every  $f \in Y$  there exists at least one  $\varphi \in X$  such that  $H(\varphi) = f$ , that is the operator  $H$  is surjective.*
- *Uniqueness: The solution  $\varphi$  from  $H(\varphi) = f$  is unique, that is the operator  $H$  is injective.*
- *Stability: The solution  $\varphi$  depends continuously on the data  $f$ , that is, it is stable with respect to perturbations in  $f$ .*

Both the existence and uniqueness in the case of a general nonlinear operator  $H$  need not be satisfied. If the existence is violated, then the possibility is that  $f \in \mathcal{R}(H)$  (range of  $H$ ) and perturbed data or noisy data  $f^\delta \notin \mathcal{R}(H)$ . Then, replacing the original equation by a minimisation problem will always lead to a solution which is a best fit to the data, calculating the so-called analysis  $\varphi^{(a)}$  by

$$\varphi^{(a)} := \arg \min_{\varphi} \|f - H(\varphi)\|_Y^2 \quad (2.1.4)$$

for  $f \in \mathcal{R}(H)$ , where the norm  $\|\cdot\|_Y$  is the standard norm of the space  $Y$ . The second postulate in Definition 2.1.4 is about the uniqueness of the solution, which means that inverse operator  $H^{-1} : \mathcal{R}(H) \subseteq Y \rightarrow X$  with  $H^{-1}(f) = \varphi$  exists.

If uniqueness is violated, then either *additional data* have to be observed or the set of *admissible solutions* has to be restricted using a-priori information on the solution. Uniqueness can be ensured by considering a *minimum norm solution* or taking special solutions that are closest to a reference element  $\varphi^* \in X$ . Hence, uniqueness can be ensured by

$$\|\varphi_{uni} - \varphi^*\|_X = \min_{\varphi \in X, H(\varphi)=f} \left( \|\varphi - \varphi^*\|_X \right).$$

The third condition in Definition 2.1.4 states that the inverse operator  $H^{-1} : \mathcal{R}(H) \subseteq Y \rightarrow X$  is continuous. A failure of this third one is most delicate to deal with because a small contamination of the right hand side  $f \in Y$  leads to large errors in the solution  $\varphi \in X$ . The problem needs some *regularisation*. We will look at this in Section 2.1.2.

Thus from above we conclude that the operator equation (2.1.3) is well-posed if the operator  $H_k$  is bijective and the well defined inverse  $H_k^{-1}$  is continuous. By solving a minimization problem a least square solution can be obtained

$$\arg \min_{\varphi_k \in X} (\|f_k - H_k(\varphi_k)\|_Y^2), \quad k = 0, 1, 2, \dots \quad (2.1.5)$$

Solving (2.1.5) for each time step  $k$  we solve a sequential data assimilation problem. If we consider a nonlinear dynamic constraint  $\varphi_{k+1} = M_k(\varphi_k)$  where  $M_k : X \rightarrow X$  for time steps  $t_k$ ,  $k = 0, \dots, K$  and acquire the sum in each time interval  $[t_0, t_K]$ , the minimisation problem takes the form

$$\min_{\varphi_0 \in X} \sum_{k=0}^K (\|f_k - H_k(\varphi_k)\|_Y^2) = \min_{\varphi_0 \in X} \left( \sum_{k=0}^K \|f_k - H_k(M_{k,0}(\varphi_0))\|_Y^2 \right). \quad (2.1.6)$$

In the above equation  $M_{k,0}$  represents the model from time  $t_0$  to time  $t_k$ . By using the evolution of system dynamics (2.1.2) that is  $M_{k,0} = M_{k-1}M_{k-2} \dots M_0$  and  $M_{k,k} = I$  now with an appropriate operator  $\bar{H}$  both the data assimilation system that is *sequential* as in (2.1.5) and *consecutive* (2.1.6) can take the form

$$\min_{\varphi \in X} (\|\bar{f} - \bar{H}(\varphi)\|^2). \quad (2.1.7)$$

If  $\bar{f} \in R(\bar{H})$  then, (2.1.7) is equivalent to  $\bar{H}(\varphi) = \bar{f}$  (cf. (2.1.1)). For a sequential data assimilation system (2.1.5) we choose  $\bar{H} := H_k$ ,  $\bar{f} := f_k$  and  $\varphi := \varphi_k$  at every time step  $k = 0, 1, \dots$ . For the consecutive data assimilate system (2.1.6), the choice needs to be  $\bar{\varphi} := \varphi_0$

$$\bar{H} := \begin{pmatrix} H_0 \\ H_1 M_{1,0} \\ H_2 M_{2,0} \\ \vdots \\ H_K M_{K,0} \end{pmatrix}, \quad \text{and} \quad \bar{f} := \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_K \end{pmatrix}. \quad (2.1.8)$$

Since both the observation operator  $H_k$  and model dynamics  $M_k$  are nonlinear, the operator  $\bar{H}$  is nonlinear. In case the equation  $\bar{H}(\varphi) = \bar{f}$  is well-posed, then  $\bar{H}$  has a well-defined continuous inverse operator  $\bar{H}^{-1}$  and  $\mathcal{R}(\bar{H}) = Y$ .

Now, if state space  $X$  and observation space  $Y$  are Banach spaces, it is well known that the well-posedness can be established by the first two postulates of Definition 2.1.4, that is  $\mathcal{R}(\bar{H}) = Y$  and  $\mathcal{N}(\bar{H}) = \{0\}$  for the range  $\mathcal{R}(\bar{H})$  and null space  $\mathcal{N}(\bar{H})$  of  $\bar{H}$ .

In the case where  $\bar{H}$  is linear over the finite dimensional Hilbert space (particularly  $\mathcal{R}(\bar{H})$  is of finite dimension) then the stability in Definition 2.1.4 holds automatically, so the well-posedness can be obeyed by first two condition of definition. (The last condition follows from the compactness of the unit ball in finite dimensions [100]). The uniqueness  $\mathcal{N}(\bar{H}) = \{0\}$  is obviously satisfied for a *linear finite-dimensional* operator  $\bar{H}$  if the observation matrix  $\bar{H}$  is full row rank. Then, for this case from the system output the behaviour of the whole system can be determined. In other words we can say system is *observable*.

Finally, we need to talk about the stability of the operator equation  $\bar{H}(\varphi) = \bar{f}$  (or  $H\varphi = H(\varphi) = f$ , a notation we use further on) with a *compact* linear operator  $H : X \rightarrow Y$  in infinite dimensions. As in infinite dimension a compact linear operator is always ill-posed (where  $\mathcal{R}(H)$  is not closed), the solution needs regularisation.

The *discretization* of an infinite dimensional unstable ill-posed problem leads to a finite dimensional well-posed problem, but the discretized problem is *ill-posed* in the sense that small errors in the input data lead to large errors in the solution. Thus, again some regularisation is necessary for this kind of finite dimensional problem arising from infinite dimensional ill-posed operator equations.

Next, we consider the *singular values decomposition* for compact linear operators between Hilbert spaces as in [100]).

LEMMA 2.1.5 (Singular system of compact linear operators.). *Let  $H : X \rightarrow Y$  be a compact linear operator. Then there exist sets of indices  $J = \{1, \dots, m\}$  for  $\dim(\mathcal{R}(H)) = m$  and  $J = \mathbb{N}$  for  $\dim(\mathcal{R}(H)) = \infty$  orthonormal system  $\{u_j\}_{j \in J}$  in  $X$  and  $\{v_j\}_{j \in J}$  in  $Y$  and a sequence  $\{\sigma_j\}_{j \in J}$  of positive real numbers with the following properties:*

$$\{\sigma_j\}_{j \in J} \text{ is non-increasing and } \lim_{j \rightarrow \infty} \sigma_j = 0 \text{ for } J = \mathbb{N}, \quad (2.1.9)$$

$$Hu_j = \sigma_j v_j, \quad (j \in J) \text{ and } H^* v_j = \sigma_j u_j, \quad (j \in J). \quad (2.1.10)$$

For all  $\varphi \in X$  there exists an element  $\varphi_0 \in u_j$  with

$$\varphi = \varphi_0 + \sum_{j \in J} \langle \varphi, u_j \rangle_X u_j \text{ and } H\varphi = \sum_{j \in J} \sigma_j \langle \varphi, u_j \rangle_X v_j \quad (2.1.11)$$

Furthermore

$$H^*f = \sum_{j \in J} \sigma_j \langle f, v_j \rangle_Y u_j \quad (2.1.12)$$

holds for all  $f \in Y$ . The countable set of triples  $\{\sigma_j, u_j, v_j\}_{j \in J}$  is called *singular system*,  $\{\sigma_j\}_{j \in J}$  are called *singular values*,  $\{u_j\}_{j \in J}$  are *right singular vectors* and form an orthonormal basis for  $\mathcal{N}(H)^\perp$  and  $\{v_j\}_{j \in J}$  are *left singular vectors* and form an orthonormal basis for  $\overline{\mathcal{R}(H)}$ .

*Proof.* See [62], Lemma 1.3. □

We can extend the concept of ill-posedness to the nonlinear operators [50, 100] by taking a linearization of the nonlinear problem, for example the *Fréchet derivative* of the nonlinear operator. The Fréchet derivative of a compact linear operator is also compact. This leads to the idea of locally ill-posed problems for nonlinear operator equations.

Linearization is an essential ingredient for solving nonlinear problems computationally. Hence, for the iterative solutions of a nonlinear problem most of the results from linear problems can be employed. We solve the nonlinear problem by solving a linear problem in each iteration or data assimilation step.

### 2.1.1 Integral Operators

An *integral equation* is an equation which involves an unknown function  $\varphi$  under the integral sign. Integral equations are very useful mathematical tool for both pure and applied mathematics. With the development of science many physical laws and principles have been discovered which, when restated in mathematical language, often are differential equations. Many of these boundary and initial value problems associated with partial (PDE) and ordinary (ODE) differential equations can be transformed into integral equations.

A typical type of an integral equation in  $\varphi$  is of the form

$$\varphi(x) = f(x) + \lambda \int_{\alpha(x)}^{\beta(x)} K(x, t) \varphi(t) dt \quad (2.1.13)$$

where  $K(x, t)$  is called the kernel of the integral equation (2.1.13),  $\alpha(x)$  and  $\beta(x)$  are lower and upper limit of integration. Here both the kernel  $K(x, t)$  and the function  $f(x)$  in equation (2.1.13) are given. The unknown function  $\varphi$  need to determined. The variable  $\lambda$  is a constant parameter.

If the unknown function  $\varphi$  appears linearly then integral equation is called *linear*, otherwise *non-linear*. If the unknown function  $\varphi$  appears only under the integral sign

then it is called integral equation of *first kind* and if unknown function  $\varphi$  appears both outside and inside then it is called integral equation of *second kind*. If  $\varphi = 0$  is a solution of an integral equation the equation is called homogeneous integral equation otherwise non-homogeneous.

The most commonly used integral equations fall under two major classes, namely *Volterra* and *Fredholm* integral equations. If the limits of the integral are fixed, then it is called a *Fredholm* integral equation otherwise it is known as a *Volterra integral equation*. The equations

$$\int_a^x K(x, t)\varphi(t)dt = f(x) \quad (2.1.14)$$

and

$$\varphi(x) - \int_a^x K(x, t)\varphi(t)dt = f(x) \quad (2.1.15)$$

are the typical examples of *Volterra* integral equations of first and second kind, respectively. These integral equation arise in physics, biology, chemistry and engineering problems. Many initial and boundary values problems associated with ordinary and partial differential equations can be cast into integral equations of Volterra and Fredholm types, respectively.

An integral equation with infinite limit or when its kernel becomes infinity at a certain point of the integration interval is called *singular integral equation*. When the unknown function  $\varphi$  appears as combination of its derivative and under the integral sign, the resulting equation is called an *Integro-differential equation*.

The above integral equation of the first and second kind is often written in operator notation

$$A\varphi = f \quad (2.1.16)$$

and

$$\varphi - A\varphi = f. \quad (2.1.17)$$

With the following definition and statement we follow Theorem 2.8 of Kress [103].

**DEFINITION 2.1.6** (Integral operator). *The linear operator  $A : C(U) \rightarrow C(U)$  defined by*

$$(A\varphi)(x) := \int_U K(x, y)\varphi(y)dy \quad x \in U, \quad (2.1.18)$$

*is called integral operator with continuous kernel  $K : U \times U \rightarrow \mathbb{D}$ , if  $U \subset \mathbb{R}^m$  is nonempty compact and Jordan measurable set that coincides with the closure of its interior and if  $K$  is a continous function. It is a bounded operator with*

$$\|A\|_\infty = \max_{x \in U} \int_U |K(x, y)|dy. \quad (2.1.19)$$

EXAMPLE 2.1.7. Consider the integral operator

$$(A\varphi)(x) := \int_a^b k(x, y)\varphi(y)dy, \quad x \in [a, b] \quad (2.1.20)$$

with some continuous kernel  $k : [a, b] \times [a, b] \rightarrow \mathbb{R}$ . Then we can estimate

$$\begin{aligned} |A\varphi(x)| &= \left| \int_a^b k(x, y)\varphi(y)dy \right| \\ &\leq \int_a^b |k(x, y)| \cdot |\varphi(y)| dy \\ &\leq \int_a^b |k(x, y)| \cdot \sup_{y \in [a, b]} |\varphi(y)| dy \\ &= C\|\varphi\|_\infty \end{aligned} \quad (2.1.21)$$

with

$$C := \sup_{x \in [a, b]} \int_a^b |k(x, y)| dy. \quad (2.1.22)$$

This proves that the operator  $A : C([a, b]) \rightarrow C([a, b])$  is bounded when  $C[a, b]$  is equipped with the maximum norm.

The study of *existence* and *uniqueness* of solutions to integral equations is very important. There are two different approaches for the existence and uniqueness of solutions to integral equations of the second kind. First, for a bounded linear operator  $A$  with sufficiently small norm (i.e.,  $\|A\| < 1$ .) such that  $A$  is a contraction, existence can be established by the *Neumann series*. The second approach is using *Riesz theory* when  $A$  is compact.

THEOREM 2.1.8. Consider a bounded linear operator  $A : X \rightarrow X$  on a Banach space  $X$ . If the norm of  $A$  satisfies  $\|A\| < 1$ , then the operator  $I - A$  has a bounded inverse on  $X$  which is given by Neumann series

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k. \quad (2.1.23)$$

It satisfies the estimates

$$\|(I - A)^{-1}\| = \frac{1}{1 - \|A\|} \quad (2.1.24)$$



In this case for each  $f \in X$  the integral equation of the second kind

$$(I - A)\varphi = f \quad (2.1.25)$$

does have a unique solution  $\varphi \in X$  which in norm depends continuously on the right-hand side  $f$ .

*Proof.* See [129], Theorem 2.3.12. □

The constructive calculation of solutions is possible by *successive approximations* as follows.

**THEOREM 2.1.9.** *Consider a bounded linear operator  $A : X \rightarrow X$  on a Banach space  $X$ . If the norm of  $A$  satisfies  $\|A\| \leq 1$ , then the sequence  $\varphi_n$  of successive approximations*

$$\varphi_{n+1} := A\varphi_n + f, \quad n = 0, 1, 2, \dots \quad (2.1.26)$$

*with starting value  $\varphi_0 = 0$  converges to the unique solution  $\varphi$  of  $(I - A)\varphi = f$ .*

*Proof.* See [103], Theorem 2.10. □

Successive approximations and the Neumann series are very good tools used often in mathematics. They also have drawbacks: *firstly*, generally it cannot be summed up in closed form and *secondly*, the theory is applicable only for  $\|A\| < 1$ . This condition is not satisfied in many integral equations which arise from applications. That is why these have been considered as a very important problem until 1900, when Hilbert, Fredholm and Riesz made significant contribution to the treatment of such integral equations. This leads us to the compact operators.

**DEFINITION 2.1.10 (Compact Operator).** *The linear operator  $A : X \rightarrow Y$  between norm space  $X$  and  $Y$  is compact, if it maps from bounded set in  $X$  to a relatively compact set in  $Y$ .*

This is a topological definition, which has the ability to provide a flexible and deep analysis. For normed spaces we can define compactness equivalently based on sequences, which is often used in the framework of functional analysis as follows.

**THEOREM 2.1.11.** *A linear operator  $A : X \rightarrow Y$  between normed spaces  $X$  and  $Y$  is compact if and only if for each bounded sequence  $(\varphi_n) \subset X$  the image sequence  $(A\varphi_n)$  contains a convergent subsequence in  $Y$ .*

*Proof.* See [129], Theorem 2.3.15. □

THEOREM 2.1.12. *The integral operator (2.1.6) with continuous kernel is a compact operator on  $C(U)$ .*

*Proof.* See [103], Theorem 2.21. □

THEOREM 2.1.13. *Let  $X$  be a normed space and  $Y$  a Banach space. If a sequence  $A_n : X \rightarrow Y$  of compact linear operators is norm convergent towards an operator  $A : X \rightarrow Y$ , then  $A$  is compact.*

*Proof.* See [129], Theorem 2.3.17. □

THEOREM 2.1.14. *A bounded operator  $A : X \rightarrow Y$  with finite dimensional range  $A(X)$  is compact.*

*Proof.* See [129], Theorem 2.3.18. □

THEOREM 2.1.15. *The identity operator  $A : X \rightarrow X$  is compact if and only if  $X$  has finite dimension.*

*Proof.* See [129], Theorem 2.3.20. □

In the light of above theorems the integral equations of first (2.1.16) and second kind (2.1.25) employ compact operators  $A : X \rightarrow X$ . The compactness of finite dimensional operators can be obtained through finite dimensional approximations [43, 103] or by Weierstrass approximation theorem [129]. The compact operator  $A : X \rightarrow Y$  is not boundedly invertible on an infinite dimensional space  $X$ , otherwise  $A^{-1} \circ A = I$  would also be compact for an infinite dimension. Hence, the integral equations of first kind with continuous kernel are not boundedly invertible. This means they inherit ill-posedness. We will discuss this later in more detail.

Second kind integral equations with a compact operator  $A$  appear in many application from fluid dynamics, acoustics, electromagnetic waves, potential theory and quantum mechanics. For the solution of such equations for a long period of 150 years there did not exist a solution theory – until Fredholm and Riesz came with their contributions.

For the following results, we define  $L = I - A$ , then an integral equation of the second kind can be written as  $L\varphi = f$ .

THEOREM 2.1.16 (First Riesz Theorem). *The null space of the operator  $L$  is a finite dimensional subspace of  $X$ .*

*Proof.* See [129], Theorem 2.3.22. □

THEOREM 2.1.17 (Second Riesz Theorem). *The range  $L(X)$  of the operator  $L$  is a closed linear subspace.*

*Proof.* See [129], Theorem 2.3.23. □

THEOREM 2.1.18 (Third Riesz Theorem). *There exists a unique integer  $r \in \mathbb{N}_0$ , such that*

$$\begin{aligned} (i) \quad \{0\} = N(L^0) \subsetneq N(L^1) \subsetneq N(L^r) = N(L^{r+1}) = \dots \\ (ii) \quad X = L^0(X) \supsetneq L^1(X) \supsetneq L^r(X) = L^{r+1}(X) = \dots \end{aligned} \quad (2.1.27)$$

*The integer  $r$  is called **Riesz number** of the operator  $A$  and it is important to know that either (i) and (ii) can determine  $r$ . Furthermore, we have*

$$X = N(L^r) \oplus L^r(X) \quad (2.1.28)$$

*Proof.* See [129], Theorem 2.3.24. □

Riesz theory was an important break through for the solution of (2.1.25).

THEOREM 2.1.19. *Let  $X$  be the normed space and  $A : X \rightarrow X$  is a compact linear operator. Then the operator  $I - A$  is injective if and only if it is surjective. Also the inverse operator  $(I - A)^{-1} : X \rightarrow X$  is bounded provided that  $I - A$  is injective.*

*Proof.* See [129], Theorem 2.3.25. □

COROLLARY 2.1.20. *Let  $A : X \rightarrow X$  be a compact linear operator on a normed space  $X$ . If the homogeneous equation*

$$\varphi - A\varphi = 0$$

*only has the trivial solution  $\varphi = 0$ , then for each  $f \in X$  the corresponding inhomogeneous equation (2.1.25) has a unique solution  $\varphi \in X$  and this solution depends continuously on  $f$ .*

*Proof.* See [103], Theorem 3.5. □

THEOREM 2.1.21. *Let  $X$  and  $Y$  be two normed spaces and let  $A : X \rightarrow Y$  be a compact linear operator. Then the integral equation of the first kind*

$$A\varphi = f \tag{2.1.29}$$

*is ill-posed provided that the normed space  $X$  is infinite dimensional*

*Proof.* See [129], Theorem 3.1.3. Suppose that  $A^{-1}$  is bounded. Then the product  $A^{-1}A = I$  is compact. Thus  $I$  is compact on the infinite dimensional space  $X$ , which is a contradiction since by Theorem 2.1.15 the identity  $I$  is compact only if  $X$  is of *finite* dimension. So our supposition is wrong, hence  $A$  has no bounded inverse. Thus an integral equation of the first kind with infinite dimensional space is ill-posed.  $\square$

From above we can conclude that integral equations of the first kind with continuous kernel or with weak singularity exhibit ill-posed problems. Hadamard's third condition that a solution depends continuously on data is violated in case of first kind integral equations due to the fact that first kind of integral operator (2.1.16) are not boundedly invertible. So small changes in measurements lead to large changes in solutions, i.e. solutions are unstable. The remedy against this instability is to find a bounded stable approximation  $R_\alpha$  depending on some parameter  $\alpha$ . This leads us to regularisation theory.

## 2.1.2 Regularisation Theory

Consider the problem  $H\varphi = f$  with a compact operator  $H$ , it is ill-posed in the infinite dimensional case. It has no bounded inverse that is  $\|H^{-1}\| = \infty$ , so the third condition in Hadamard's Definition 2.1.4 is violated. Thus in order to solve the equation  $H\varphi = f$  (or, the equivalent minimisation problem  $\min \|H\varphi - f\|^2$  for  $f \notin \mathcal{R}(H)$ ), regularisation is required.

Consider an operator  $H : X \rightarrow Y$  and its adjoint operator  $H^* : Y \rightarrow X$  and let  $\varphi$  be the solution to the least squares minimisation problem  $\min \|H\varphi - f\|^2$ . Then, the solution to the minimisation problem is equivalent to the solution of the *normal equations*

$$H^*H\varphi = H^*f. \tag{2.1.30}$$

If  $H$  is compact, then  $H^*H$  is compact, thus the normal equation (2.1.30) remain ill-posed because of the ill-posedness of the operator  $H^*H$ . Now, if we replace the normal equations by

$$(\alpha I + H^*H)\varphi_\alpha = (\alpha\varphi_\alpha + H^*H\varphi_\alpha) = H^*f. \tag{2.1.31}$$

for  $\alpha > 0$ , the operator  $(\alpha I + H^*H)$  is boundedly invertible. The the above equation is solved approximately by

$$\varphi_\alpha = (\alpha I + H^*H)^{-1}H^*f. \quad (2.1.32)$$

This equation is called *Tikhonov regularisation*. The parameter  $\alpha$  is called regularisation parameter [44, 103, 81, 132, 168, 165]

**THEOREM 2.1.22** (Tikhonov regularisation). *Let  $H : X \rightarrow Y$  be a compact linear operator. Then, the operator  $(\alpha I + H^*H)$  has a bounded inverse and the problem (2.1.31) is well-posed for  $\alpha > 0$  and  $(\alpha I + H^*H)^{-1}H^*f$  is the Tikhonov approximation of a minimum-norm least square solution  $\varphi$  of (2.1.30). Furthermore, the solution  $\varphi_\alpha$  is equivalent to the unique solution of the minimisation problem*

$$\operatorname{argmin}_{\varphi \in X} J_\alpha(\varphi)$$

with

$$J_\alpha(\varphi) := \|H\varphi - f\|_Y^2 + \alpha\|\varphi\|_X^2, \quad (2.1.33)$$

the so-called Tikhonov functional.

*Proof.* See [173], Theorem 2.1. □

If we replace term  $\|\varphi\|_X^2$  by  $\|\varphi - \varphi^b\|_X^2$ , where  $\varphi^b$  is called *back ground* or *prior knowledge*, then it is called *generalized Tikhonov regularisation*. We will discuss generalized Tikhonov regularisation in our next section.

Now the general linear regularisation scheme is defined as

**DEFINITION 2.1.23** (Regularisation scheme). *A family of bounded linear operators  $\{R_\alpha\}_{\alpha>0}$ ,  $R_\alpha : Y \rightarrow X$  is a linear regularisation scheme for the compact bounded linear injective operator  $H$  if*

$$\lim_{\alpha \rightarrow 0} R_\alpha H\varphi = \varphi \quad \forall \varphi \in X. \quad (2.1.34)$$

In other word  $R_\alpha$  is point wise convergent to  $H^{-1}$  that is

$$R_\alpha H\varphi \rightarrow \varphi, \quad \alpha \rightarrow 0 \quad \forall \varphi \in X. \quad (2.1.35)$$

It can also be written as  $R_\alpha \xrightarrow{p} H^{-1}$ ,  $\alpha \rightarrow 0$ . The convergence of  $R_\alpha \xrightarrow{p} H^{-1}$  is not norm convergence for ill-posed equations.

THEOREM 2.1.24. Consider normed spaces  $X, Y$ ,  $\dim X = \infty$ , a compact linear operator  $H : X \rightarrow Y$  and a regularization scheme  $R_\alpha$  for  $H^{-1}$ . Then the family  $R_\alpha$ ,  $\alpha > 0$  of bounded operators cannot be uniformly bounded with respect to  $\alpha > 0$  and the operators  $R_\alpha$  cannot be norm convergent.

*Proof.* See [129], Theorem 3.1.5 □

By the above Theorem 2.1.24 applied to the family of approximate inverses  $R_\alpha = (\alpha I + H^*H)^{-1}$  where the range  $\mathcal{R}(H)$  is not closed, we obtain

$$\lim_{\alpha \rightarrow 0} \|R_\alpha\| = \infty. \quad (2.1.36)$$

When this regularisation operator  $R_\alpha$  is applied to noisy data  $f^\delta$  we get regularised solutions

$$\varphi_\alpha^\delta = R_\alpha f^\delta, \quad (2.1.37)$$

where we assume a noise level  $\delta$  such that  $\|f^\delta - f\| \leq \delta$ .

Using Lemma 2.1.5 on the singular system of a compact operator we can also write the regularised solution of Tikhonov regularisation (2.1.33) in the form

$$\varphi_\alpha^\delta = \sum_{j \in J} \frac{\sigma_j}{\sigma_j^2 + \alpha} \langle f^\delta, v_j \rangle_Y u_j. \quad (2.1.38)$$

We can note that at  $\alpha = 0$  the solution  $\varphi_\alpha^\delta$  amplifies the noise in data  $f^\delta$  where for compact operators the singular values obey

$$\lim_{j \rightarrow \infty} \sigma_j = 0.$$

An approximate unique solution can be chosen to be given by  $\varphi = H^\dagger f$ , where  $H^\dagger = (H^*H)^{-1}H^*$  is called generalized or Moore-Penrose inverse (c.f. [9]) and  $H^\dagger : R(H) + R(H)^\perp \rightarrow X$  is continuous when  $\mathcal{R}(H)$  is closed. The Moore-Penrose inverse coincides with Tikhonov regularization for  $\alpha = 0$ .

We can estimate the so-called *total regularisation error* of the solution  $\varphi_\alpha^\delta$  defined by

$$\varphi_\alpha^\delta = R_\alpha f^\delta \text{ with } \|f^\delta - f\| \leq \delta \quad (2.1.39)$$

as follows. We first calculate

$$\begin{aligned} \varphi_\alpha^\delta - \varphi &= (\alpha I + H^*H)^{-1} H^* f^\delta \\ &\quad - (\alpha I + H^*H)^{-1} H^* f + (\alpha I + H^*H)^{-1} H^* f - \varphi \end{aligned}$$

The above leads to the equality and estimate

$$\begin{aligned} \|\varphi_\alpha^\delta - \varphi\| &= \|R_\alpha f^\delta - R_\alpha f + R_\alpha f - \varphi\| \\ \|\varphi_\alpha^\delta - \varphi\| &\leq \|R_\alpha\|\delta + \|R_\alpha f - \varphi\| \end{aligned}$$

and with the true solution  $\varphi$  and true data  $f = H\varphi$

$$\|\varphi_\alpha^\delta - \varphi\| \leq \|R_\alpha\|\delta + \|R_\alpha H\varphi - \varphi\|. \quad (2.1.40)$$

It shows that the total regularisation error is a sum of the *stability component*  $\|R_\alpha\|\delta$  and the *approximation error*  $\|R_\alpha H\varphi - \varphi\|_X$ . As first term on the right hand side the stability component describes the influence of data noise error of size  $\delta$ . It is amplified by  $R_\alpha$ , but due to boundedness of  $R_\alpha$  for  $\alpha > 0$  it can be controlled. The first term is also called *data error* of the reconstruction. The second term reflects the approximation error, it depends on the approximation of  $H$  by  $R_\alpha$ . This error of the second term comes from the approximation of operator  $H$ . The second term is also called *regularization error*.

For small  $\alpha$  the regularization error will be small by Definition 2.1.23, but the first term data error will be large (compare (2.1.36)), whereas for large  $\alpha$  the stability term or data error will be small but the approximation error be large. As shown in Figure 2.1, the critical issue concerning Tikhonov and other regularisation methods is the choice of a suitable value for regularisation parameter  $\alpha$ . We have a trade-off between stability and accuracy. If  $\alpha$  is too large, we get a poor approximation of the exact solution even for exact data, and if  $\alpha$  is too small, the reconstruction becomes unstable. The optimal value of  $\alpha$  depends both on the data noise level  $\delta$  and the exact solution  $\varphi$ . For estimating a suitable value of  $\alpha$  plenty of methods are available in the literature (see for example [9, 78, 77, 100, 50, 76]). Well-known schemes are, for example, the L-curve method, generalized cross-validation and the discrepancy principle.

A regularisation scheme is called convergent, if from the convergence of the data error to zero, it follows that the regularised solution converges to the exact solution.

DEFINITION 2.1.25. We call a function  $\alpha = \alpha(\delta)$  a strategy for a regularization scheme  $R_\alpha$  if  $\alpha(\delta) \rightarrow 0$  for  $\delta \rightarrow 0$ . Such a strategy is called regular, if

$$R_{\alpha(\delta)} f^\delta \rightarrow H^{-1} f, \quad \delta \rightarrow 0 \quad (2.1.41)$$

for each  $f^\delta$  with  $\|f^\delta - f\| \leq \delta$ .

The Tikhonov regularisation  $R_\alpha = (\alpha I + H^* H)^{-1} H^* : Y \rightarrow X$  is convergent when

$$\alpha(\delta) \rightarrow 0 \quad \text{and} \quad \frac{\delta^2}{\alpha(\delta)} \rightarrow 0 \quad \text{as} \quad \delta \rightarrow 0,$$

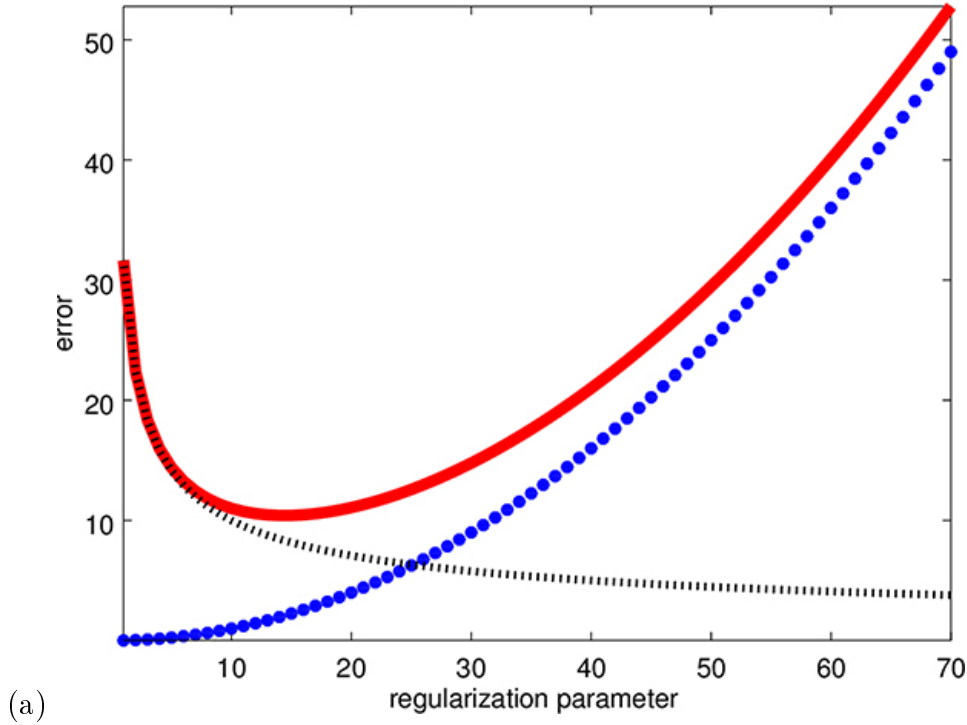


Figure 2.1: The black dot is data error, blue dotted curve is regularization error and red represent total error modeled by (2.1.40). With some regularization parameter  $\alpha > 0$  we obtain minimum error for the reconstruction, where the sum of data error contribution and regularization error has its minimum.

see [50]. So for Tikhonov regularization we can choose  $\alpha = O(\delta)$  (c.f. [9]).

There are many regularization schemes available for the solution of inverse problem. Well known regularisation schemes are *truncated singular value decomposition* (TSVD) and the *Landweber iteration* (see [50, 75, 74]). Furthermore different penalty terms can be used to include the background knowledge about the solution  $\varphi$  with different norms like the  $L_1$ -norm, see for example [9, 178, 9, 25, 39, 39]) and the article by Burger et al. ([37]).

Next, following [62, 129] we present a coherent mathematical framework for data assimilation techniques, which are used in practice, by using the result from inverse problems and regularisation theory.



### 2.1.3 Cycling, Tikhonov regularisation and 3D-VAR

The goal of data assimilation is to solve the dynamical inverse problem with measurement data  $f_1, f_2, f_3, \dots, f_k \dots$  at different time  $t_1 < t_2 < t_3 < \dots < t_k < \dots$ . For every time  $t_k$  the inverse problem is given by (2.1.3). Usually the measured data  $f_k$  has some deficiency of information due to which it is impossible to recover the state  $\varphi_k$  at time  $t_k$  completely. So it is crucial to include the dynamical evolution of the state.

Suppose we have some solution  $\varphi_k^{(a)}$  at time  $t_k$ , then the *first guess* or *background*  $\varphi_{k+1}^{(b)}$  for time  $t_{k+1}$  can be calculated by model dynamic

$$\varphi_{k+1}^{(b)} = M_k(\varphi_k^{(a)}). \quad (2.1.42)$$

Now, we assimilate the measurement  $f_{k+1}$ , i.e. calculate a *reconstruction*  $\varphi_{k+1}^{(a)}$ , which is called *the analysis* at time  $t_{k+1}$  in data assimilation. Then, by using (2.1.42) and the analysis  $\varphi_{k+1}^{(a)}$ , the background  $\varphi_{k+2}^{(b)}$  at time  $t_{k+2}$  can be calculated by replacing  $k$  to  $k+1$ . We continue this process, i.e. we again assimilate data to get to another *analysis* (or *reconstruction*) at time  $t_{k+2}$ . Then, we find the background at time  $t_{k+3}$  based on (2.1.42). This *cycling process* of calculating background and reconstruction continues marching through time. This process is known as *cycling of reconstruction and dynamics*.

DEFINITION 2.1.26 (Cycling for data assimilation). *Given an initial state  $\varphi_{(0)}$  at time  $t_0$  with observation  $f_k$ ,  $k = 0, 1, 2 \dots$  performing cycling steps:*

- (i) Propagation Step. *Calculate the background  $\varphi_{k+1}^{(b)}$  at time  $t_{k+1}$  by using system dynamic  $M_k$  by (2.1.42).*
- (ii) Analysis Step. *Calculate reconstruction or analysis  $\varphi_{k+1}^{(a)}$  based on the measurement  $f_{k+1}$  at time  $t_{k+1}$ .*

*Then by increasing index  $k$  to  $k+1$  go back to Step (i).*

The main focal point of data assimilation system is the analysis step (ii). The objective is to calculate reconstruction  $\varphi_k^{(a)}$  from data  $f_k$  using background  $\varphi_k^{(b)}$  at time step  $t_k$ . So we need to develop or use such reconstruction methods that utilizes or combines all the information optimally.

To perform the analysis we employ two approaches: ones comes from *optimization* and *optimal control theory*, and other from *stochastic and probability theory*. Here we first follow optimization approach.

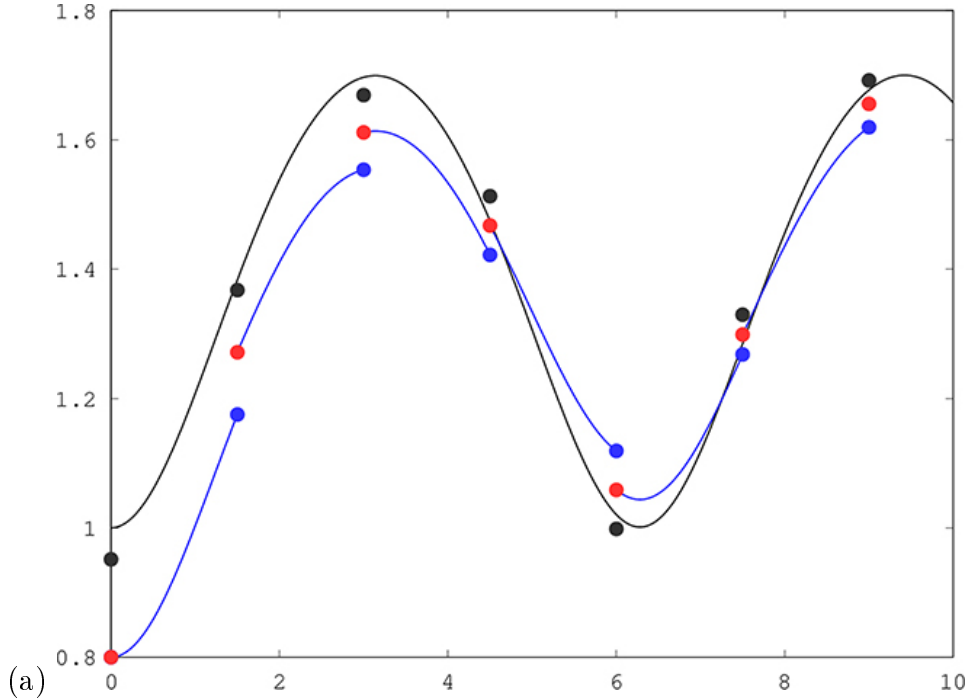


Figure 2.2: The measurements  $f$  and knowledge from previous steps, that is background or first guess  $\varphi^{(b)}$ , is used in data assimilation for generating the analysis at each time step. This resulting analysis is then used as background for next assimilation step transported by the model  $M$  of the underlying dynamical system. This is called *analysis cycle* in data assimilation. In the figure measurement  $f$  are shown by black dots, analysis in red. The black curve is the truth and the model simulations are shown as blue curves.

We can combine the given information, measurement  $f_k \in Y$  and the background  $\varphi_k^{(b)} \in X$  at step  $k$ , by minimizing the *Tikhonov functional*

$$J_k(\varphi) := \|H(\varphi) - f\|_Y^2 + \alpha \|\varphi - \varphi^{(b)}\|_X^2 \quad (2.1.43)$$

with norm  $\|\cdot\|_X$  and  $\|\cdot\|_Y$  in the state space  $X$  and the observation space  $Y$ , where  $\alpha > 0$  controls the weight between data and prior knowledge. The minimizer of (2.1.43) is called analysis denoted as  $\varphi_k^{(a)}$ . If  $\alpha$  is large then we are giving weight to the background while for small  $\alpha$  we are putting more weight on observations. Using the transformation  $\tilde{\varphi} := \varphi - \varphi^{(b)}$  the above functional becomes

$$J_k(\varphi) = \alpha \|\tilde{\varphi}_k\|_X^2 + \|(f_k - H\varphi_k^{(b)}) - H\tilde{\varphi}\|_Y^2, \quad (2.1.44)$$

which is the form of the Tikhonov functional in (2.1.33). By Theorem 2.1.22 we obtain the minimizer

$$\tilde{\varphi}_k^{(a)} = (\alpha I + H^*H)^{-1}H^*(f_k - H\varphi_k^{(b)}) \quad (2.1.45)$$

leading to minimizer of (2.1.43)

$$\varphi_k^{(a)} = \varphi_k^{(b)} + (\alpha I + H^*H)^{-1}H^*(f_k - H\varphi_k^{(b)}) \quad (2.1.46)$$

By Definition 2.1.26 the analysis (2.1.46) is called *cycled Tikhonov regularisation*.

Mostly, in data assimilation we deal with the Euclidean space  $X = \mathbb{R}^n$  and  $Y = \mathbb{R}^m$  of dimension  $n$  and  $m$ , respectively,  $m, n \in \mathbb{N}$ . The norms in the spaces  $X$  and  $Y$  are standard  $L^2$ -norms. Working with standard  $L^2$ -norms in state space  $X = \mathbb{R}^n$  leads to crucial difficulties, reflected by following example.

**Example.** Assume that  $X = \mathbb{R}^n$ ,  $Y = \mathbb{R}^1$  and  $H = (1, 0, \dots, 0)$ . This means that we measure the first variable only. The variational scheme calculates the increment

$$\delta\varphi_k = \varphi_{k+1}^{(a)} - \varphi_k^{(b)} = H^*(\alpha I + HH^*)^{-1}(f_k - H\varphi_k^{(b)}) \quad (2.1.47)$$

with  $H^* = H^T$ . This means that only the first component is updated. The other components remain unchanged. But it is highly unusual that the first variable will not influence other variables in the same or neighboring points.

- Standard  $L^2$ -norms do not take correlations into account between different variables and quantities in spatial neighborhood.
- Standard  $L^2$ -norms lead to highly unphysical and unrealistic increments.

So we define a weighted scalar product in state space  $X = \mathbb{R}^n$

$$\langle \varphi, \psi \rangle_{B^{-1}} := \varphi^T B^{-1} \psi = \langle \varphi, B^{-1} \psi \rangle \quad (2.1.48)$$

and in measurements space  $Y = \mathbb{R}^m$  by

$$\langle f, g \rangle_{R^{-1}} := f^T R^{-1} g = \langle f, g \rangle_{R^{-1}} \quad (2.1.49)$$

where  $B \in \mathbb{R}^{n \times n}$  and  $R \in \mathbb{R}^{m \times m}$  are positive definite, self-adjoint or symmetric, and invertible. These are the errors covariance matrices of respective spaces. We use the notation  $\|\cdot\|_{B^{-1}}$  and  $\|\cdot\|_{R^{-1}}$  for the weighted scalar product in the state space  $X$  and the measurement space  $Y$  respectively. The minimisation function (2.1.43) with weighted norm is represented as

$$J_k(\varphi) := \|H(\varphi) - f\|_{R^{-1}}^2 + \alpha \|\varphi - \varphi^{(b)}\|_{B^{-1}}^2. \quad (2.1.50)$$

This functional with Definition 2.1.26 of cycling is called *three-dimensional Variation data assimilation scheme (3D-VAR)* [129, 108, 46, 173, 58, 69]. It assimilates the measurement  $f_k$  sequentially at each time step  $t_k$ , that is why it is called *variational optimization in space*. If we also take time dimension into account, then it is called *four dimensional or 4D-VAR*.

We employ the notation  $H^*$  for the adjoint operator with respect to the weighted scalar product and  $H'$  for the adjoint under the standard  $\ell^2$  scalar product. The relation between them is derived as

$$\begin{aligned}
\langle f, H\varphi \rangle_{R^{-1}} &= \langle f, R^{-1}H\varphi \rangle_{L^2} \\
&= \langle H' R^{-1}f, \varphi \rangle_{L^2} \\
&= \langle H' R^{-1}f, BB^{-1}\varphi \rangle_{L^2} \\
&= \langle BH' R^{-1}f, B^{-1}\varphi \rangle_{L^2} \\
&= \langle BH' R^{-1}f, \varphi \rangle_{B^{-1}} \\
&= \langle H^*f, \varphi \rangle_{B^{-1}}.
\end{aligned} \tag{2.1.51}$$

This leads to transformation formula  $H^* = BH'R^{-1}$ . Based on the weighted scalar product (2.1.48) and (2.1.49) the minimizer (2.1.46) takes the form

$$\begin{aligned}
\varphi_k^{(a)} &= \varphi_k^{(b)} + (\alpha I + H^*H)^{-1}H^*(f_k - H\varphi_k^{(b)}) \\
&= \varphi_k^{(b)} + (\alpha I + BH'R^{-1}H)^{-1}BH'R^{-1}(f_k - H\varphi_k^{(b)})
\end{aligned} \tag{2.1.52}$$

In large data assimilated problems the dimension of the observation space  $Y$  is much smaller than the dimension of the state space  $X$ . Here  $H^*H$  is a much larger matrix than  $HH^*$ , so we transform the computation into observation space  $Y$  by using the following measurement space inversion relation

$$(\alpha I + H^*H)H^* = (\alpha H^* + H^*HH^*) = H^*(\alpha I + HH^*). \tag{2.1.53}$$

For  $\alpha > 0$  the operator  $(\alpha I + H^*H)$  is invertible in state space  $X$  and the same is true for  $(\alpha I + H^*H)$  on  $Y$ , since

$$(\alpha I + H^*H)\varphi = 0 \tag{2.1.54}$$

implies

$$\alpha \langle \varphi, \varphi \rangle + \underbrace{\langle H\varphi, H\varphi \rangle}_{\geq 0} = 0, \tag{2.1.55}$$

thus  $\varphi = 0$  and  $(\alpha I + H^*H)$  is injective and thus bijective on  $X$ . This leads to

$$(\alpha I + H^*H)^{-1}H^* = H^*(\alpha I + HH^*)^{-1}. \tag{2.1.56}$$

Using (2.1.56) the update equation (2.1.52) take the form

$$\begin{aligned}\varphi_k^{(a)} &= \varphi_k^{(b)} + H^*(\alpha I + H^*H)^{-1}(f_k - H\varphi_k^{(b)}) \\ &= \varphi_k^{(b)} + BH'R^{-1}(\alpha I + HBH'R^{-1})^{-1}(f_k - H\varphi_k^{(b)}). \\ &= \varphi_k^{(b)} + BH'(\alpha R + HBH')^{-1}(f_k - H\varphi_k^{(b)}),\end{aligned}\tag{2.1.57}$$

where the inverse  $(\alpha R + HBH')^{-1}$  is to be calculated in  $Y = \mathbb{R}^m$ . The solution is then projected into the state space by applying  $BH'$ . The solution (2.1.57) is also called *optimal interpolation* (OI) when it is applied in state space only, i.e. when the measurements are some variables of the state space. For the form (2.1.57) often the name *physical space statistical analysis system* (PSAS) is used, compare [45, 40, 65, 149].

Three dimensional variational data assimilation 3D-VAR is the application of Tikhonov regularization in each assimilation step, thus it is stable at each step.

**THEOREM 2.1.27** (Equivalence of cycled Tikhonov regularisation and 3DVar). *Three-dimensional variational data assimilation (2.1.52) or (2.1.57) is equivalent to cycled Tikhonov regularisation (2.1.46) when the norms are arising from the weighted inner products (2.1.48) and (2.1.49).*

Theorem 2.1.27 shows that 3D-VAR is merely a cycled Tikhonov regularisation in an appropriately chosen norm.

## 2.1.4 Variational Assimilation 4D-VAR

*Four dimensional variational data assimilation* (4D-VAR) technique is the temporal extension of *three dimensional variational data assimilation* (3D-VAR), which includes *model integration* as part of the observation operators. Instead of the static analysis of 3D-VAR in 4D-VAR we measure the distance between the analyzed state and distributed observations at the appropriate time. This needs the tangent linear (TL) and adjoint (AD) of the forecast model.

Thus 4D-VAR is a generalization of 3D-VAR to deal with the observations which are distributed in time. The cost functional is same but with a generalized observation operators to include the forecast model that allows comparison among observations and model state at appropriate time [69, 129, 108, 46, 133, 163]. The model is considered as a *strong constraint* in the sense that the solution has to satisfy the model equation. In this setup, 4D-VAR gives equal credibility to older as well as newer observations.

The basic aim of 4D-VAR technique is to find the analysis  $\varphi^{(a)}$  of the data assimilation problem given in Definition 2.1.3 by minimization over the *time interval*

$[t_k, t_{k+K}]$  with a window of size  $\Delta t = t_{k+K} - t_k$  to calculate the analysis state at  $t_k$ . The corresponding cost functional is given by

$$J_k(\varphi) := \alpha \|\varphi - \varphi^{(b)}\|_X^2 + \sum_{j=0}^K \|f_{k+j,k} - H(M_{k+j,k}(\varphi))\|_Y^2. \quad (2.1.58)$$

The model evolution  $M_{k+j,k}$  is provided in (2.1.2). The minimization of the functional calculates the analysis  $\varphi_k^{(a)}$  at time  $t_k$ . At other points the analysis is accumulated by forward model

$$\varphi_{k+j}^{(a)} = M_{k+j,k} \varphi_k^{(a)}, \quad j = 1, \dots, K, \quad \varphi_{k+K+1}^{(b)} = M_{k+K+1,k} \varphi_k^{(a)}. \quad (2.1.59)$$

The 4D-VAR scheme (2.1.58) and (2.1.59) compels the solution to satisfy the forward model over the interval  $[t_k, t_{k+K}]$ . The four dimensional variational data assimilation puts the same weight to all data points to accumulate its analysis.

As there are always assumptions involved in the formation of models so no model is perfect, but all are containing errors. It is a very important question for researcher how to accumulate model errors in the computations. Many strategies and techniques have been introduced to mitigate such limitation like *weak-constraint 4D-VAR*.

We can write the 4DVar functional (2.1.58) in the same form than the 3DVar functional (2.1.43) by putting all the measurement  $f_k, \dots, f_{k+K}$  into one column and disposing the sum based on the new (possibly nonlinear) operator  $\overline{H}_k$ . This leads to

$$J_k(\varphi) := \alpha \|\varphi - \varphi_k^{(b)}\|_X^2 + \|\overline{f}_k - \overline{H}_k(\varphi)\|^2 \quad (2.1.60)$$

with

$$\overline{f} := \begin{pmatrix} f_k \\ f_{k+1} \\ \vdots \\ f_{k+K} \end{pmatrix} \quad \text{and} \quad \overline{H} := \begin{pmatrix} HM_{k,k} \\ HM_{k+1,k} \\ \vdots \\ HM_{k+K,k} \end{pmatrix}. \quad (2.1.61)$$

The minimizer of (2.1.58) seeks to catch the full dynamic trajectory of the state to the given data  $f_{k+j}$ ,  $j = 0, \dots, K$ , over the time window  $[t_k, t_{k+K}]$ . As in Section 2.1.3 we can change the cost functional (2.1.58) into the form of a general Tikhonov functional (2.1.33) (see [61, 45]).

Let  $\varphi^{(a)}$  be the minimum of (2.1.58) and the background is obtain by (2.1.59). The minimization of functional (2.1.58) is called four dimensional variation data assimilation(4DVar). The repetition of minimizing (2.1.58) with the model propagation (2.1.59) is called *cycled 4DVar scheme*. It is similar to the cycled Tikhonov regularisation as describe previous Section 2.1.3.

Usually, the minimisation of functional (2.1.58) is carried out by the *gradient method*. We calculate the gradient  $\nabla_{\varphi} J(\varphi)|_{\varphi^{(\ell)}}$  at the point  $\varphi^{(\ell)}$  in state space, then update by the relation

$$\varphi^{(\ell+1)} := \varphi^{(\ell)} - h \nabla_{\varphi} J(\varphi)|_{\varphi^{(\ell)}} \quad (2.1.62)$$

with suitable *step size*  $h > 0$  and *initial condition*  $\varphi^{(0)}$  (usually  $\varphi^{(0)} := \varphi^{(b)}$ ). For simplicity we consider  $X = \mathbb{R}^n$  and  $Y = \mathbb{R}^m$  with  $\ell^2$  inner product. Let  $e_j$ ,  $j = 1, \dots, n$  are standard basis of  $\mathbb{R}^n$ . Let

$$g(\varphi) := \|f - HM\varphi\|_Y \quad (2.1.63)$$

Where  $f \in Y$  and the model operator  $M : X \rightarrow X$ . The the gradient of  $g(\varphi)$  is given by

$$\nabla_{\varphi} g(\varphi) = -2 \left( M^* H^* (f - HM\varphi) \right) \quad (2.1.64)$$

In case  $M$  is nonlinear than nonlinear version of (2.1.64) is

$$\nabla_{\varphi} g(\varphi) = -2 \left( \left( \frac{dM(\varphi)}{d\varphi} \right)^* H^* (f - HM\varphi) \right) \quad (2.1.65)$$

where  $\frac{dM(\varphi)}{d\varphi}$  is the Fréchet derivative of  $M(\varphi)$ . The derivative

$$\mathbf{M}(\varphi) := \frac{dM(\varphi)}{d\varphi} \quad (2.1.66)$$

is known as *tangent linear model* [55, 105].

In the framework of dynamic inverse problems, the dynamic model is given by a system of differential equation of the state  $\varphi(t)$

$$\dot{\varphi}(t) = F(\varphi) : t \geq 0 \quad (2.1.67)$$

with initial condition

$$\varphi(0) = \varphi_0. \quad (2.1.68)$$

where  $\varphi_0 \in X$ . The model dynamic is represented by

$$\varphi(t) = M_{t,0}(\varphi(0)) = M_{t,0}(\varphi_0) \quad (2.1.69)$$

Thus the (2.1.67) takes the form

$$F(\varphi(t)) = \frac{dM_{t,0}(\varphi_0)}{dt}, \quad t \geq 0. \quad (2.1.70)$$

The main task is the differentiation of  $\varphi(t)$  with respect to initial condition  $\varphi_0$  denoted as  $\varphi'(t) := \frac{d\varphi(t)}{d\varphi_0}$ , which is linear mapping from  $X$  into  $X$ . For  $X = \mathbb{R}^n$ ,  $\varphi(t) \in \mathbb{R}^n$  is a vector and  $\varphi'(t)$  is a  $n \times n$  matrix

$$\varphi'(t) = \begin{pmatrix} \frac{d\varphi_1}{d\varphi_{0,1}} & \cdots & \frac{d\varphi_1}{d\varphi_{0,n}} \\ \vdots & \vdots & \vdots \\ \frac{d\varphi_n}{d\varphi_{0,1}} & \cdots & \frac{d\varphi_n}{d\varphi_{0,n}} \end{pmatrix}. \quad (2.1.71)$$

We assume  $\varphi(t)$  is continuous and twice differentiable with respect to  $t$  and  $\varphi_0$ , then we can exchange the derivatives with respect to  $t$  and  $\varphi_0$ . We get

$$\frac{d\varphi'}{d\varphi_0} = \frac{d}{d\varphi_0} \frac{dM_{t,0}(\varphi_0)}{dt} = \frac{d}{dt} \frac{dM_{t,0}(\varphi_0)}{d\varphi_0} = \frac{d\varphi'(t)}{dt} \quad (2.1.72)$$

thus we have using (2.1.70)

$$\frac{d\varphi'(t)}{dt} = \frac{dF(\varphi(t))}{d\varphi_0} = F'(\varphi(t)) \frac{d\varphi(t)}{d\varphi_0} \quad (2.1.73)$$

At  $t = 0$  the initial condition takes the form

$$\varphi'(0) = \left. \frac{d\varphi}{d\varphi_0} \right|_{t=0} = I, \quad (2.1.74)$$

is a identity  $I \in X$

**THEOREM 2.1.28.** *The tangent linear model, i.e. the Fréchet derivative of  $\varphi$  with respect to the initial conditions  $\varphi_0$  of (2.1.67), is given by the solution to the system of differential equations*

$$\frac{d\varphi'(t)}{dt} = F'(\varphi(t)) \frac{d\varphi(t)}{d\varphi_0}; \quad t \geq 0 \quad (2.1.75)$$

with initial condition

$$\varphi'(0) = I \quad (2.1.76)$$

where  $F'$  is Fréchet derivative of forcing term  $F(t)$  of (2.1.67) with respect to argument  $\varphi$ .

*Proof.* See [129], Theorem 5.3.1. □

Using (2.1.69) and  $\varphi(t) = M_{t,0}(\varphi(0)) = M_{t,0}(\varphi_0)$  alone with (2.1.66) we obtain

$$\varphi'(t) = \frac{dM_{t,0}(\varphi(0))}{d\varphi_0} =: \mathbf{M}_{t,0}(\varphi_0) \quad (2.1.77)$$



for the tangent linear model. We note that the tangent linear model adjoint is square matrix of order  $n \times n$ . This matrix is large, when  $n$  is large. Thus, we need an efficient method to calculate the gradient based on it. To evaluate the adjoint in (2.1.64) we define a function  $\psi(t) \in X$  over the interval  $[t_k, t_{k+1}]$  by

$$\cdot\psi(t) = -F'(\varphi(t))^*\psi(t) \quad (2.1.78)$$

with final condition

$$\psi(t_{k+1}) = H^*(f_{k+1} - HM(\varphi_k)). \quad (2.1.79)$$

The gradient is now based on the following lemma.

LEMMA 2.1.29.  $t \in [t_k, tk + 1]$ , the inner product

$$h(t) = \langle \varphi'(\delta\varphi_0), \psi(t) \rangle \quad (2.1.80)$$

is constant over time for any  $\delta\varphi_0 \in X$ .

*Proof.* See [129], Theorem 5.3.2.  $\square$

THEOREM 2.1.30. Let  $\psi(t) \in X$  be solution of (2.1.78) and (2.1.79) on  $[t_k, tk + 1]$ . Then, the gradient (2.1.64) of (2.1.63) is given by

$$\nabla_{\varphi}g(\varphi) = -2\psi(t_k) \quad (2.1.81)$$

*Proof.* See [62], Lemma 6.1. It is shown in Theorem 2.1.30 that  $h(t)$  is constant, hence

$$\begin{aligned} (\nabla_{\varphi}g(\varphi))_j &= -2\langle \varphi'(t_{k+1})e_j, H^*(f_{k+1} - HM(\varphi_k)) \rangle \\ &= -2\langle \varphi'(t_{k+1})e_j, \psi(t_{k+1}) \rangle \\ &= -2\langle \varphi'(t_k)e_j, \psi(t_k) \rangle \\ &= -2\langle e_j, \psi(t_k) \rangle = -2\psi_j(t_k) \end{aligned} \quad (2.1.82)$$

for  $j = 1 \dots n$  by using (2.1.74) and proof is complete.  $\square$

The function defined by (2.1.78) and (2.1.79) is called *tangent linear adjoint model*, which is used for calculation of gradient of the 4D-VAR functional.

Consider  $X = \mathbb{R}^n$  and  $Y = \mathbb{R}^m$  with weighted scalar product (2.1.48) and (2.1.49). Then, the gradient of the full cost function involving model dynamic (2.1.58) is (c.f. [62])

$$\begin{aligned} \nabla_{\varphi}J_k(\varphi) &:= 2B^{-1}(\varphi - \varphi_k^{(b)}) \\ &\quad -2\sum_{j=1}^K (\mathbf{M}_{k+j,k}(\varphi))^* H^* R^{-1}(f_{k+j} - HM_{k+j,k}(\varphi)). \end{aligned} \quad (2.1.83)$$

A gradient method like (2.1.62) can be applied to minimize the functional  $J_k(\varphi)$  (2.1.58). Alternatively, we can use a Gauß-Newton method [47] to find the minimum of the functional (2.1.58). We solve  $\nabla_{\varphi} J_k(\varphi) = 0$  for finding the minimizer of the functional (2.1.58) by Newton's method, i.e.

$$\varphi^{(\ell+1)} := \varphi^{(\ell)} - \left( \nabla \nabla_{\varphi} J(\varphi)|_{\varphi^{(\ell)}} \right)^{-1} \nabla_{\varphi} J(\varphi)|_{\varphi^{(\ell)}}, \quad (2.1.84)$$

with starting guess  $\varphi^{(0)} := \varphi_k^{(b)}$ , where  $\nabla \nabla_{\varphi} J(\varphi)|_{\varphi^{(\ell)}}$  is called Hessian.

Usually we use an approximate Hessian instead of the full Hessian by neglecting the terms involving the gradient of tangent linear model, then leading to a quasi-Newton method. The gradient method converges linearly while the Gauß-Newton method converges superlinearly in a neighborhood of the true solution for well-posed problems. Newton and Gauß-Newton method are the same when the observation operator and dynamical model are linear. In this case the local minimizer is also global minimizer (see for example [67]).

## 2.2 Kalman Smoother and Kalman Filter

The *Kalman Filter* is a method similar to cycled Tikhonov regularisation, 3D-VAR and 4D-VAR which is used to deal with the data assimilation problems introduced in Definition 2.1.3. It is different from other methods in the sense that it not only updates the analysis at each step, it also upgrades iteratively the state space norm by using the information from former assimilation steps.

Here we describe the *Kalman filter* in a deterministic way (following [129]) and show equivalence with the *Kalman smoother* and *four dimensional variational data assimilation* introduced in Section 2.1.4 for linear operators and models. Let us study the assimilation of a measurement  $f_1 \in Y$  by the Tikhonov functional

$$J_1(\varphi) := \alpha \|\varphi - \varphi_0^{(b)}\|_{B^{-1}}^2 + \|f_1 - A(\varphi)\|_{R^{-1}}^2, \quad \varphi \in X, \quad (2.2.1)$$

We assume that  $X$  and  $Y$  are some Hilbert spaces over  $\mathbb{R}$ . The norms  $\|\cdot\|$  are weighted norms given by  $\|\varphi\|_{B^{-1}}^2 = \langle \varphi, B^{-1}\varphi \rangle$  and  $\|f\|_{R^{-1}}^2 = \langle f, R^{-1}f \rangle$ , where  $B$  and  $R$  are positive definite symmetric and invertible operators or matrices. The minimizer of functional (2.2.1) is  $\tilde{\varphi}^{(a)}$ .

Now we receive more data  $f_2 \in Y$ . Then with the measurement  $f_1$  and  $f_2$  we will proceed the same way and minimize with weighted norm we have

$$J_2(\varphi) := \alpha \|\varphi - \varphi_0^{(b)}\|_{B^{-1}}^2 + \|f_1 - A(\varphi)\|_{R^{-1}}^2 + \|f_2 - A(\varphi)\|_{R^{-1}}^2 \quad \varphi \in X, \quad (2.2.2)$$

analogous to (2.2.1). It is very flexible scheme in the sense that either assimilate  $f_1$  first then  $f_2$  or vice versa or both at a time same as above (2.2.2). When the model is not the identity, we need to incorporate the model propagation between different assimilation steps. For a linear model, this can be carried out by using

$$A_k = H_k M_{k,0} \quad (2.2.3)$$

and  $M_0 = M_{1,0}$ ,  $M_1 = M_{2,1}$ . Then the above can take the form

$$J_2(\varphi) := \alpha \|\varphi - \varphi_0^{(b)}\|_{B^{-1}}^2 + \|f_1 - HM_0(\varphi)\|_{R^{-1}}^2 + \|f_2 - M_1 M_0(\varphi)\|_{R^{-1}}^2 \quad \varphi \in X, \quad (2.2.4)$$

Let us first assimilate  $f^{(1)}$  using (2.2.1) with minimizer  $\tilde{\varphi}^{(a)}$ , then assimilate  $f^{(2)}$  in second step with changed or upgraded norm in  $X$  by including the information from first assimilation  $f^{(1)}$ . First, we assimilate  $f^{(1)}$  with minimization of the functional

$$J_1(\varphi) := (\alpha) \|\varphi - \varphi_0^{(b)}\|_{B^{-1}}^2 + \|f_1 - HM_0(\varphi)\|_{R^{-1}}^2 \quad \varphi \in X, \quad (2.2.5)$$

having minimizer  $\tilde{\varphi}^{(a)}$  and then in second step we study the assimilation of  $f_2$  by minimizing using previous information

$$J_{21}(\varphi) := \|\varphi - \tilde{\varphi}_0^{(a)}\|_{\tilde{B}^{-1}}^2 + \|f^{(2)} - HM_1 M_0(\varphi)\|_{R^{-1}}^2, \quad \varphi \in X, \quad (2.2.6)$$

with updated weighted norm having new weight  $\tilde{B}$ . The problem here is how we find the new weight  $\tilde{B}$  such that the minimizer of  $J_{21}$  is equal to the minimizer of full functional  $J_2$  in (2.2.4). This is the case if we can choose  $\tilde{B}$  such that  $J_{21} = J_2 + c$  with some constant  $c$ , where in (2.2.6)  $J_1$  is used via  $\tilde{\varphi}^{(a)}$ . The problem is solved if we can find  $\tilde{\varphi}^{(a)}$  and  $\tilde{B}$  such that – up to a constant – the first term of  $J_{21}$  is identical to  $J_1$ , that is

$$\alpha \|\varphi - \varphi_0^{(b)}\|_{B^{-1}}^2 + \|f_1 - HM_0(\varphi)\|_{R^{-1}}^2 = \|\varphi - \tilde{\varphi}^{(a)}\|_{\tilde{B}^{-1}}^2 + \tilde{c}, \quad (2.2.7)$$

where  $\tilde{c}$  is some constant depending upon  $f_1$  and  $\varphi_0$ . Now studying  $J_1$  we have

$$\begin{aligned} J_1(\varphi) &= \alpha \langle \varphi - \varphi_0^{(b)}, B^{-1}(\varphi - \varphi_0^{(b)}) \rangle + \langle f_1 - HM_0 \varphi, R^{-1}(f_1 - HM_0 \varphi) \rangle \\ &= \langle \varphi, \alpha B^{-1} \varphi \rangle - 2 \langle \varphi, \alpha B^{-1} \varphi_0^{(b)} \rangle + \langle \varphi_0^{(b)}, \alpha B^{-1} \varphi_0^{(b)} \rangle + \langle f_1, R^{-1} f_1 \rangle \\ &\quad - 2 \langle \varphi, M_0^* H^* R^{-1} f_1 \rangle + \langle \varphi, M_0^* H^* R^{-1} HM_0 \varphi \rangle \\ &= \langle \varphi, (\alpha B^{-1} + M_0^* H^* R^{-1} HM_0) \varphi \rangle - 2 \langle \varphi, (\alpha B^{-1} \varphi_0^{(b)} \\ &\quad + M_0^* H^* R^{-1} f_1) \rangle + \underbrace{\langle f_1, R^{-1} f_1 \rangle}_{const} + \underbrace{\langle \varphi_0^{(b)}, \alpha B^{-1} \varphi_0^{(b)} \rangle}_{const} \end{aligned} \quad (2.2.8)$$

Let

$$G := \alpha B^{-1} + M_0^* H^* R^{-1} HM_0 \quad (2.2.9)$$

Then (2.2.8) takes the form

$$\begin{aligned} J_1(\varphi) &= \langle \varphi, G\varphi \rangle - 2\langle \varphi, \alpha B^{-1}\varphi_0^{(b)} - M_0^*H^*R^{-1}f_1 \rangle + \underbrace{\langle f_1, R^{-1}f_1 \rangle}_{const} + \underbrace{\langle \varphi_0^{(b)}, \alpha B^{-1}\varphi_0 \rangle}_{const} \\ &= \langle \varphi, G\varphi \rangle - 2\langle \varphi, \alpha B^{-1}\varphi_0^{(b)} - M_0^*H^*R^{-1}f_1 \rangle + c \end{aligned} \quad (2.2.10)$$

with some constant  $c$ . Similarly, the first term of  $J_{21}$

$$\|\varphi - \tilde{\varphi}_0^{(a)}\|_{\tilde{B}^{-1}}^2 = \langle \varphi, \tilde{B}^{-1}\varphi \rangle - 2\langle \varphi, \tilde{B}^{-1}\tilde{\varphi}_0^{(a)} \rangle + \tilde{c} \quad (2.2.11)$$

where  $c$  and  $\tilde{c}$  do not depend on  $\varphi$ . After comparing the equations (2.2.10) and (2.2.11) we get

$$\tilde{B}^{-1} = G = (\alpha B^{-1} + M_0^*H^*R^{-1}HM_0) \quad (2.2.12)$$

and

$$\tilde{B}^{-1}\tilde{\varphi}_0^{(a)} = (\alpha B^{-1}\varphi_0^{(b)} + M_0^*H^*R^{-1}f_1). \quad (2.2.13)$$

From (2.2.13) we have

$$\begin{aligned} \tilde{\varphi}_0^{(a)} &= \tilde{B}(\alpha B^{-1}\varphi_0^{(b)} - M_0^*H^*R^{-1}f_1) \\ &= (\alpha B^{-1} + M_0^*H^*R^{-1}HM_0)^{-1}(\alpha B^{-1}\varphi_0^{(b)} + M_0^*H^*R^{-1}f_1) \\ &= (\alpha I + BM_0^*H^*R^{-1}HM_0)^{-1}(\alpha\varphi_0^{(b)} + BM_0^*H^*R^{-1}f_1) \end{aligned}$$

After some calculation using

$$\alpha I = (\alpha I + BM_0^*H^*R^{-1}HM_0) - BM_0^*H^*R^{-1}HM_0 \quad (2.2.14)$$

we get

$$\begin{aligned} \tilde{\varphi}_0^{(a)} &= \varphi_0^{(b)} + (\alpha I + BM_0^*H^*R^{-1}HM_0)^{-1}BM_0^*H^*R^{-1}(f_1 - HM_0\varphi_0^{(b)}) \\ &= \varphi_0^{(b)} + BM_0^*H^*(\alpha R + HM_0BM_0^*H^*)^{-1}(f_1 - HM_0\varphi_0^{(b)}), \end{aligned} \quad (2.2.15)$$

using inversion relation (2.1.56), which is the minimizer of  $J_1$ . It is the same as in (2.1.52) and (2.1.57). Now, the above step can incorporate iteratively new measurements  $f_1, f_2, f_3$  etc. This approach leads to the *Kalman smoother*, later on we prove that the Kalman smoother and Kalman filter are equivalent at final time. The basic difference is that the Kalman filter calculates at  $t_k$  and the Kalman smoother at  $t_0$ .

DEFINITION 2.2.1 (Kalman Smoother). *Let  $H_k : X \rightarrow Y$  and  $M_k : X \rightarrow X$ ,  $k = 0, 1, 2, \dots$  given in Definition 2.1.3 be linear and assume that measurements  $f_1, f_2, f_3, \dots$  at times  $t_1, t_2, \dots$  are given. Then, we calculate weight matrices*

$$\tilde{B}_k^{-1} = \tilde{B}_{k-1}^{-1} + M_{k,0}^* H_k^* R^{-1} H_k M_{k,0}, \quad k = 1, 2, 3, \dots \quad (2.2.16)$$

with  $\tilde{B}_0 := B$ , where  $M_{k,0}$  is defined in (2.1.2) and the analysis states  $\varphi_k^{(a)}$  at time  $t_k$  defined by

$$\tilde{\varphi}_k^{(a)} = \varphi_{k-1}^{(a)} + \tilde{B}_{k-1} M_{k,0}^* H_k^* (R + H_k M_{k,0} \tilde{B}_{k-1} M_{k,0}^* H_k^*)^{-1} (f_k - H_k M_{k,0} \varphi_{k-1}^{(a)}) \quad (2.2.17)$$

for  $k = 1, 2, \dots$ , with  $\tilde{\varphi}_0^{(a)} = \varphi_0^{(b)}$ .

Now we are ready to relate the Kalman smoother and 4DVar

THEOREM 2.2.2. *Let  $H_k : X \rightarrow Y$  and  $M_k : X \rightarrow X$ ,  $k = 0, 1, 2, \dots$  be linear operators and data  $f_1, f_2, f_3, \dots$  be given. Then, 4DVar carried out with data  $f_1, f_2, f_3, \dots, f_k$  is equivalent to the Kalman Smoother given in (2.2.1) in the sense that the minimum of the 4DVar functional (2.1.58) is given by the analysis  $\tilde{\varphi}_k^{(a)}$  defined in (2.2.17).*

*Proof.* See [62], Theorem 7.2. □

By Definition 2.2.1 the Kalman smoother works with states at  $t_0$  while the Kalman filter calculate states at time  $t_1, t_2, \dots$ . We need to propagate the states  $\tilde{\varphi}_k^{(a)}$  from time  $t_0$  to  $t_1$  by

$$\varphi_k^{(b)} = M_{k,0} \tilde{\varphi}_{k-1}^{(a)} \quad \text{and} \quad \varphi_k^{(a)} = M_{k,0} \tilde{\varphi}_k^{(a)} \quad (2.2.18)$$

for  $k = 1, 2, \dots$ , which implies

$$\varphi_k^{(b)} = M_{k-1} \varphi_{k-1}^{(a)} \quad (2.2.19)$$

evolve the state from time  $t_0$  to  $t_k$  also the weighted matrices  $\tilde{B}$  propagate from  $t_0$  to  $t_k$  by

$$B_k^{(b)} = M_{k,0} \tilde{B}_{k-1} M_{k,0}^* \quad \text{and} \quad B_k^{(a)} = M_{k,0} \tilde{B}_k M_{k,0}^* \quad (2.2.20)$$

for  $k = 1, 2, \dots$ . Propagating the analysis matrix from time  $t_{k-1}$  to  $t_k$ , the background matrix is

$$B_k^{(b)} = M_{k-1} B_{k-1}^{(a)} M_{k-1}^*. \quad (2.2.21)$$

Using the state propagation (2.2.19) and weighted matrices evolutions (2.2.21) the iterative version of (2.2.17) is given by

$$\varphi_k^{(a)} = \varphi_k^{(b)} + B_k^{(b)} H_k^* (R + H_k B_k^{(b)} H_k^*)^{-1} (f_k - H_k \varphi_k^{(b)}) \quad (2.2.22)$$

for  $k \in \mathbb{N}$ . It can also be represented in the simple form

$$\varphi_k^a = \varphi_k^{(b)} + K_k(f_k - H_k\varphi_k^{(b)}) \quad (2.2.23)$$

where  $K_k$  is

$$K_k = B_k^{(b)} H_k^* (R + H_k B_k^{(b)} H_k^*)^{-1}, \quad (2.2.24)$$

called *Kalman gain matrix*. It is basically the *Tikhonov regularisation matrix* (2.1.57). The analysis matrix  $B_k^{(a)}$  is obtained from the background matrix  $B_k^{(b)}$  by

$$(B_k^{(a)})^{-1} = (B_k^{(b)})^{-1} + H_k R^{-1} H_k \quad (2.2.25)$$

for  $k \in \mathbb{N}$ . The analysis matrix update formula is usually written as follows.

LEMMA 2.2.3. For  $k \in \mathbb{N}$  and  $B_k^{(a)}$  defined by (2.2.25) we have

$$B_k^{(a)} = (I - K_k H_k) B_k^{(b)}, \quad (2.2.26)$$

where  $K_k$  is given by (2.2.24).

*Proof.* See [62], Lemma 7.3. □

Now we are ready to define the *Kalman Filter*.

DEFINITION 2.2.4 (Kalman Filter). Starting with an initial state  $\varphi_0^{(b)}$ , and an initial weight matrix  $B_0^{(a)} := B$  for  $k \in \mathbb{N}$ , the Kalman Filter iteratively calculates an analysis  $\varphi_k^{(a)}$  at time  $t_k$  by

(i) propagate the state  $\varphi_{k-1}^{(a)}$  from  $t_{k-1}$  to  $t_k$  via (2.2.19):

$$\varphi_k^{(b)} = M_{k-1} \varphi_{k-1}^{(a)} \quad (2.2.27)$$

(ii) propagate the state  $B_{k-1}^{(a)}$  from  $t_{k-1}$  to  $t_k$  via (2.2.21):

$$B_k^{(b)} = M_{k-1} B_{k-1}^{(a)} M_{k-1}^* \quad (2.2.28)$$

(iii) calculate the Kalman gain by (2.2.24):

$$K_k = B_k^{(b)} H_k^* (R + H_k B_k^{(b)} H_k^*)^{-1} \quad (2.2.29)$$

(iv) calculate an analysis state by (2.2.22):

$$\varphi_k^{(a)} = \varphi_k^{(b)} + B_k^{(b)} H_k^* (R + H_k B_k^{(b)} H_k^*)^{-1} (f_k - H_k \varphi_k^{(b)}) \quad (2.2.30)$$

(v) calculate an analysis weight by (2.2.22):

$$B_k^{(a)} = (I - K_k H_k) B_k^{(b)} \quad (2.2.31)$$

The above mathematical setup narrates that the Kalman filter works in two stages. The first stage is called *predictor* or *forecast stage*. In this stage using some given initial condition or background state the forecast is calculated by propagating the model through time. This background state has known error covariance. Which is also evolved through time by the model to generate the forecast error covariance.

It is an important feature of the Kalman filter that these errors are also evolved through time with the model. So within the forecast stage state vectors and covariance matrices are propagated through time with the model dynamics.

The second stage is the *analysis stage*. In this stage through the ratio of errors of the background and the observations a weighting is introduced. The observations are assimilated based on this weight, which helps to adjust the forecast estimate as a best estimate to the true state. Finally, the state of the system and analysis covariance error is computed according to this weight. So in this step the state of the system and the covariance matrices are updated through new measurements and a-priori information.

Let  $M_k$  be a known matrix of linear model dynamics of the system at time  $t_k$ . The state evolution through the model dynamics is

$$\varphi_{k+1}^{true} = M_k(\varphi_k^{true}) + \eta_k \quad (2.2.32)$$

Let  $\varphi$  is a state vector containing the state parameters of the system having dimension  $n$ ,  $\varphi_k^{true}$  is the true state of the system at time  $t_k$ . Since our models are not perfect real world models. Models have errors. In (2.2.32)  $\eta_k$  is the model error at time  $t_k$  which is assumed to be unbiased, that is its expectation value is zero, written as  $\mathbb{E}(\eta_k) = 0$ . Then, the error covariance of model is

$$Cov(\eta_k) = E(\eta_k \eta_k^T).$$

It is generally unknown, but ensemble systems try to estimate it based on an ensemble of model propagations. Now, let  $y_k$  be the observation vector at time  $t_k$ , satisfy the relation

$$y_k = H(\varphi_k^{true}) + \delta_k \quad (2.2.33)$$

$H$  is the observation operator which maps the state space to the observation space. Since observations are contaminated by errors, let  $\delta_k$  be the observation error which is assumed to be unbiased, i.e.  $\mathbb{E}(\delta) = 0$ . The calculation of the covariance matrix for the observations is an important problem of current research, since it is known that measurement errors are correlated. The observation error covariance matrix is denoted as  $R = \mathbb{E}(\delta \delta^T)$ . It is assumed that the observation errors and model errors are not correlated at any time.

With the Kalman filter in both stages of forecast and analysis we try to find an *unbiased forecast* and an *unbiased analysis*, that is

$$\mathbb{E}(\varphi^{(b)} - \varphi^{(true)}) = 0 \quad \text{and} \quad \mathbb{E}(\varphi^{(a)} - \varphi^{(true)}) = 0.$$

The covariance error matrices of forecast and analysis measure the correlation between different component of the state. They are of dimension  $n \times n$  and given by

$$B^{(b)} = \mathbb{E}((\varphi^{(b)} - \varphi^{(true)})(\varphi^{(b)} - \varphi^{(true)})^T). \quad (2.2.34)$$

and

$$B^{(a)} = \mathbb{E}((\varphi^{(a)} - \varphi^{(true)})(\varphi^{(a)} - \varphi^{(true)})^T). \quad (2.2.35)$$

At the forecast stage the forecast Kalman filter equations are given by (2.2.27) and (2.2.28). These forecast equations evolve the previous model analysis state and previous analysis error covariance through time with initial state  $\varphi_0^{(a)}$  and an initial weight matrix  $B_0^{(a)} := B$ .

The update of both the analysis and the error covariance matrices at the analysis step begin with calculation of Kalman gain (2.2.29). It calculates the weighting given to the observations with respect to the ratio between forecast and observational error covariances. The Kalman gain matrix can also be found using a *minimum variance estimate*. This is obtained from a minimization of the functional (2.1.50). The analysis state and the covariance matrix are updated by (2.2.30) and (2.2.31). For Gaussian distributions and linear systems, the error covariances of the forecast and the analysis are exact (c.f. [129]).

The Kalman gain (2.2.29) calculates the weighting given to the observations with respect to the ratio between forecast and observational error covariances. The Kalman gain matrix can also be found using a *minimum variance estimate* with cost functional (2.1.50). Thus Kalman filter is optimal.

Currently the Kalman Filter is not used for numerical weather prediction because it can be used with linear models only. Clearly, numerical weather prediction models are nonlinear, which makes the Kalman filter unsuitable to this application. Furthermore due to large state space and covariance matrices it is very expensive. In numerical weather prediction the typical problem size is around  $n = 10^8$  unknowns and could easily be larger when resolution is increased. The number of measurements which are employed at each time  $t_k$  are around  $m = 10^7$ . In the Kalman Filter, this would lead to matrices  $B$  of the size  $10^8 \times 10^8$ , which has strong impact on calculation times and storage. In the assimilation cycle  $O(10^7)$  observations are assimilated. The Kalman gain involves inversion of large matrices which make Kalman filtering computationally expensive and inefficient.



There are many versions of the conventional Kalman filter. Some versions take the inverse of the covariance matrix, some take the square root of the covariance matrix, and some consider the square root of the inverse of the covariance matrices. All these variations have been developed for better performance of the Kalman filter when implemented on computers having memory storage problems and constraints on word length. Also, the conventional Kalman filter can be derived in continuous time or discrete time or in both. In the first case state covariance and state estimate are in the form of time derivatives. They are continuously integrated at each instant and do not have individual update steps, while in discrete form both state estimates and state covariances are advanced discretely and have separate update steps when ever data is processed [20, 26, 52, 96, 102, 104, 128, 166, 175].

Now we can show the equivalence of Kalman filter smoother and 4DVar as follows.

**THEOREM 2.2.5** (Equivalence of 4DVar, Kalman Filter and Kalman Smoother). *Let  $H_k : X \rightarrow Y$  for  $k \in \mathbb{N}$  and  $M_k : X \rightarrow X$  for  $k \in \mathbb{N}_0$  be linear. Let  $\varphi_k^{(a)}$  be the analysis of the Kalman Filter at time  $t_k$ ,  $\tilde{\varphi}_k^{(a)}$  the analysis of the Kalman smoother with data  $f_1, f_2, f_3, \dots, f_k$  at  $t_0$ ,  $\tilde{\varphi}_{4DVar,k}^{(a)}$  the minimizer of the 4DVar functional (2.1.4) at time  $t_0$  and define*

$$\varphi_{4DVar,k}^{(a)} = M_{k,0} \tilde{\varphi}_{4DVar,k}^{(a)}, \quad \text{for } k = 1, 2, \dots \quad (2.2.36)$$

*Then, 4DVar is equivalent to the Kalman Filter and to the Kalman Smoother in the sense that*

$$\varphi_{4DVar,k}^{(a)} = \varphi_k^{(a)} = M_{k,0} \tilde{\varphi}_{4DVar,k}^{(a)}, \quad (2.2.37)$$

*if we start the iterations with the same initial background state  $\varphi_0^{(b)}$  and the same initial weight matrix  $B_0^{(a)} := B$ .*

*Proof.* See [62], Lemma 7.5. □

## 2.3 Ensemble Methods and Ensemble Kalman Filter

Above we studied many data assimilation techniques including Tikhonov data assimilation, 3DVar, 4DVar, and the Kalman filter. Going through all these methods we find that 3DVar or Tikhonov do not carry all dynamical information we get from previous assimilation steps and they work with a fixed norm at each time step. While 4DVar includes the full trajectory over a time window, it implicitly

propagates all information, so usually this leads to better results than 3DVar. However, the Kalman filter and Bayes data assimilation are equivalent to 4DVar for linear systems. They employ updates of the covariance matrices and, thus, include all information from previous assimilation steps and propagate them through time. Generally, we expect them to lead to result comparable to 4DVAR.

For non-linear systems the Bayesian approach with non-Gaussian dynamical distribution should be better. For the Kalman filter the covariance matrices  $B_k^{(a)}$  and  $B_k^{(b)}$  are the main challenge due to their size  $n \times n$ . Their computation is infeasible for large  $n$  in terms of computation time and storage, even when super computers are employed in operational centers. There is a need of algorithms or methods which give a good approximation to the matrix  $B_k^{(b)}$  along with (2.2.30) with small computational costs. This leads to ensemble methods.

The key idea of ensemble methods is that, following the ideas of Monte-Carlo methods, we use ensembles of the states for the estimation of dynamical information. The Monte-Carlo method received a lot of attention from geophysical applications in 1990's and has been employed in operational forecast centers since 2010.

The Ensemble Kalman Filter (EnKF) was introduced by Evensen in 1994 [53]. In the article [54] the further development and application of the EnKF to weather prediction has been discussed. We can use nonlinear models in the Ensemble Kalman Filter. In case of the Kalman filter a formula is used for the  $B$  matrix update, while in the EnKF the main idea is the *ensemble* and the statistical sample of states to be used for the computation of the analysis state for each members. From the forecast ensemble then mean and covariance are calculated. They are used in the analysis step for the calculation of the Kalman gain, which is further used to assimilate observations and calculate the analysis ensemble.

The Ensemble Kalman Filter does not require a derivation of a tangent linear operator or adjoint equation and no back ground integration is needed. This simple conceptual formulation and easy implication led to strong popularity in different communities [93, 94, 136, 162]. The statistical representation of the state estimation improves the quality of predictions and calculates a good initial guess for ensemble based prediction schemes.

The ensemble approach can be discussed in many ways, but here we will approach it from an applied mathematics point of view.

**DEFINITION 2.3.1 (Ensemble).** *An ensemble of  $N$  members is any finite set of vectors  $\varphi_k^{(\ell)}$  for  $\ell = 1, \dots, N$ .*

We can propagate the ensemble through time by applying the model dynamic  $M$  or  $M_k$ , respectively. Starting with an initial ensemble  $\varphi_0^{(\ell)}$ ,  $\ell = 1, \dots, L$  this leads

to ensemble members

$$\varphi_k^{(\ell)} = M[t_{k-1}, t_k] \varphi_{k-1}^{(\ell)}, \quad k = 1, 2, 3, \dots \quad (2.3.1)$$

For an arbitrary ensemble  $\varphi^{(1)}, \dots, \varphi^{(N)}$  the mean  $\mu$  is given by

$$\mu = \frac{1}{N} \sum_{\ell=1}^N \varphi^{(\ell)} \quad (2.3.2)$$

The ensemble error covariance matrix is

$$B = \frac{1}{N-1} \sum_{\ell=1}^N (\varphi^{(\ell)} - \mu)(\varphi^{(\ell)} - \mu)^T \quad (2.3.3)$$

Here, we divide by  $(N-1)$  instead of  $(N)$ . The reason is that this leads to an unbiased estimator [23, 91, 109].  $B$  has dimension  $n \times n$ . We center the ensemble members around the ensemble mean, defining the *ensemble matrix*  $Q$  by

$$Q := \frac{1}{\sqrt{N-1}} (\varphi^{(1)} - \mu, \dots, \varphi^{(N)} - \mu) \quad (2.3.4)$$

The ensemble Kalman filter now approximates the updates in the Kalman filter in the ensemble space  $U_Q \subset X$  defined by

$$U^Q := \text{span}\{\varphi^{(1)} - \mu, \dots, \varphi^{(N)} - \mu\}. \quad (2.3.5)$$

The  $Q$  matrix (2.3.4) can be seen as some kind of square root of the covariance matrix  $B$  in (2.3.3), if a square root  $\hat{A}$  of  $A$  is given by  $A = \hat{A}\hat{A}^T$ . This definition is often used in the physical sciences. We note that  $Q$  is only a low-dimensional approximation to the full square-root of  $B$ .

With the adjoint or transpose  $Q_k^*$  of  $Q_k$  we now write

$$B_k^{(b)} := Q_k Q_k^*, \quad k \in \mathbb{N}. \quad (2.3.6)$$

This leads to an update equation

$$\varphi_k^{(a)} = \varphi_k^{(b)} + K(f - H\varphi_k^{(b)}) \quad (2.3.7)$$

with the *Kalman matrix*

$$K = Q_k^{(b)} (Q_k^{(b)})^* H_k^* (R_k + H_k Q_k^{(b)} (Q_k^{(b)})^* H_k^*)^{-1} \quad (2.3.8)$$

following (2.2.24). Clearly, by the form of  $B_k^{(b)}$  the increments  $\varphi_k^{(a)} - \mu$  are in the ensemble space  $U^Q$ .

- DEFINITION 2.3.2 (Ensemble Kalman Filter). 1. First, given an analysis ensemble  $\varphi_{(k-1)}^{(a,\xi)}$ ,  $\xi = 1, \dots, N$  at  $t_{k-1}$ , as in (2.3.6) the model dynamics  $M$  is used to propagate the ensemble members to time  $t_k$  and calculate a first guess or background ensemble  $\varphi_k^{(b,\xi)}$ ,  $\xi = 1, \dots, N$ .
2. In an analysis step at time  $t_k$ , the Ensemble Kalman Filter calculates an estimate for the matrix  $B_k^{(b)}$  by (4.1.5) based on the background ensemble  $\varphi^{(b,1)}, \dots, \varphi^{(b,N)}$ .
3. This is used to calculate an update of the ensemble mean following the Kalman filter (2.2.30). An important step is the generation of an analysis ensemble  $\varphi^{(a,1)}, \dots, \varphi^{(a,N)}$ , which fits to the analysis covariance matrix  $B_k^{(a)}$  as calculated by the Kalman filter.

The main task next is to describe how the analysis ensemble is generated. Many approaches have been suggested for this. First, we will study the family of ensembles generated by eigenvalues decomposition from self adjoint, positive definite and weighted matrix  $B := B^{(b)}$ .

First, we note that  $B$  has a complete set of eigenvalues  $\lambda^{(1)}, \dots, \lambda^{(n)}$  with  $\lambda^{(1)} \geq \lambda^{(2)} \geq \dots \geq \lambda^{(n)}$  and orthogonal eigenvectors  $\vartheta^{(1)}, \dots, \vartheta^{(n)}$  such that

$$B\vartheta^{(\ell)} = \lambda^{(\ell)}\vartheta^{(\ell)}, \quad \ell = 1, \dots, n. \quad (2.3.9)$$

Since the covariance matrix  $B$  is positive definite and symmetric, eigenvalues are real and positive and eigenvectors are orthogonal. This leads to the diagonal matrix  $\Lambda := \{\sqrt{\lambda^{(1)}}, \dots, \sqrt{\lambda^{(n)}}\}$  and the orthogonal matrix  $W = \{\vartheta^{(1)}, \dots, \vartheta^{(n)}\}$ , such that we obtain

$$B = W\Lambda^2W^* = (W\Lambda)(W\Lambda)^*, \quad (2.3.10)$$

where we also note that  $W^* = W^{-1}$ . This correspond to the principle component analysis of quadratic form defined by

$$F(\varphi, \vartheta) := \varphi^T B \vartheta, \quad \varphi, \vartheta \in X. \quad (2.3.11)$$

$B$  defines a hyper-surface of second order with positive eigenvalues, which has level curves which form a family of ellipses with dimension  $n - 1$ . The eigenvectors  $\vartheta^{(\ell)}$ ,  $\ell = 1, \dots, n$  are the principle axes of the ellipse. According to (2.3.10) the application of  $B$  here is a combination of a projection of  $\varphi$  onto the principle axis of  $\vartheta^{(\ell)}$  of  $B$  with the multiplication by  $\lambda^{(\ell)}$ . This decomposition can be taken as a setup for the construction of a low-dimensional approximation of  $B$ .

First of all we have to define the metric for the approximation of the  $B$ -matrix and the construction of ensembles. As we have seen above the  $B$  matrices are core parts of the update of (2.2.4), (2.2.4), (2.2.30) and (2.2.31).

LEMMA 2.3.3. *We construct an ensemble of vectors by choosing the  $N - 1$  maximal eigenvalues of  $B$  and its corresponding eigenvectors  $\vartheta^{(1)}, \dots, \vartheta^{(N-1)}$ . We define*

$$Q := [\sqrt{\lambda^{(1)}}\vartheta^{(1)}, \dots, \sqrt{\lambda^{(N-1)}}\vartheta^{(N-1)}] \quad (2.3.12)$$

*Then, we have the error estimate*

$$\|B - QQ^*\|^2 = \sup_{j=N, \dots, n} |\lambda^{(j)}| = |\lambda^{(N)}| \quad (2.3.13)$$

*Proof.* See [129], Lemma 5.5.3. □

THEOREM 2.3.4. *Let the eigenvalues  $\lambda^{(1)} \geq \lambda^{(2)} \geq \dots \geq \lambda^{(N)}$  of the self-adjoint weight matrix  $B$  be ordered according to its size and let  $\varphi_{(1)}, \dots, \varphi_{(N)}$  with  $N \in \mathbb{N}$  be an arbitrary ensemble of states in  $X$ . Then, the error for the approximation of the weight matrix  $B$  by  $QQ^*$  with  $Q$  defined in (4.1.5) is estimated by*

$$\|B - QQ^*\|^2 \geq \lambda^{(N)} \quad (2.3.14)$$

*Proof.* See [129], Lemma 5.5.4. □

LEMMA 2.3.5. *Consider the approximation of  $B_k$  by an ensemble  $\varphi_k^{(1)}, \dots, \varphi_k^{(N)}$  with ensemble matrix  $Q_k$ . If the error satisfies*

$$\|B_k^{(a)} - Q_k^{(a)}(Q_k^{(a)})^*\| \leq \epsilon, \quad (2.3.15)$$

*for some  $\epsilon > 0$  then, the error estimate for the propagated ensemble at time  $t_{k+1}$  is given by*

$$\|B_{k+1}^{(b)} - Q_{k+1}^{(b)}(Q_{k+1}^{(b)})^*\| \leq \|M_k\| \|M_k^*\| \epsilon \quad (2.3.16)$$

*Proof.* See [129], Lemma 5.5.5. □

We now come to the *Square Root Filter* (SRF). The question arises how we update the ensemble matrix in ensemble Kalman filter. As we know in the case of the Kalman filter the analysis weighted matrix  $B_k^{(a)}$  is calculated via  $B_k^{(a)}$  using relation (2.2.31). In every step, an *analysis ensemble*  $Q_k^{(a)}$  needs to be constructed which keeps the update equation (2.2.31) intact, i.e. we demand

$$\begin{aligned} Q_k^{(a)}(Q_k^{(a)})^* &= (I - K_k H_k) Q_k^{(b)} (Q_k^{(b)})^*, \\ &= Q_k^{(b)} \underbrace{\left( I - (Q_k^{(b)})^* H_k^* (R_k + H_k Q_k^{(b)} (Q_k^{(b)})^* H_k^*)^{-1} H_k Q_k^{(b)} \right)}_{=: T} (Q_k^{(b)})^*. \end{aligned} \quad (2.3.17)$$

The operator in round brackets can be seen to be self-adjoint and positive, such that a square root  $S$  exists with  $T = SS^*$ . Uniqueness can be settled by writing it in the form  $W\Lambda^2W^*$  with the orthonormal matrix  $W$  which collects the normalized orthogonal eigenvectors of the matrix and where  $\Lambda^2$  is a diagonal matrix with the eigenvalues of the operator on the diagonal. The eigenvalues can be ordered according to its size and listed according to its multiplicity. Then  $\Lambda^2$  is unique, but  $W$  can have different realizations if eigenspaces of dimension larger than one exist. However, the operation of  $W\Lambda^2W^*$  on each of the eigenspaces is uniquely determined. We choose the square root given by

$$S := W\Lambda W^*, \quad (2.3.18)$$

with the diagonal matrix  $\Lambda$  constructed by taking the positive square roots of the positive diagonal entries of  $\Lambda^2$ . Then,  $S$  is well-defined, since it is the same for every orthogonal  $W$  constructed by projections onto the eigenspaces. The analysis ensemble is then updated according to

$$Q_k^{(a)} = Q_k^{(b)}S, \quad k \in \mathbb{N}, \quad (2.3.19)$$

where

$$S = \sqrt{I - (Q^{(b)})^*(H)^*(R + HQ^{(b)}(Q^{(b)})^*(H)^*)^{-1}HQ^{(b)}}. \quad (2.3.20)$$

The updated formula (2.3.19) for the ensemble is called the Ensemble Kalman Square Root Filter (EKSRF).

Filters that use randomly perturbed observation ensembles are called *stochastic filters*, while those which do not use contaminated observation ensembles are called *deterministic filters* like the square root filter (see [20, 26, 96, 102, 104, 128, 152, 166]). They do not use noisy observations, which introduce additional noise to the problem specially when the ensemble size is small [171, 73, 94]. It is efficient and accurate computationally. We finished calculating the error in the analysis  $B$ -matrix in dependence on the background  $B$ -matrix for the square root Kalman filter.

LEMMA 2.3.6. *Assume that  $\varphi_k^{(1)}, \dots, \varphi_k^{(N)}$  is an ensemble which satisfies*

$$\|B_{k+1}^{(b)} - Q_{k+1}^{(b)}(Q_k^{(b)})^*\| \leq \epsilon \quad (2.3.21)$$

*with some  $\epsilon < \|B_k^{(b)}\|$ . Then, for the analysis ensemble defined by (2.3.19) we have*

$$\|B_k^{(a)} - Q_k^{(a)}(Q_k^{(a)})^*\| \leq C\epsilon \quad (2.3.22)$$

*with some constant  $C$  not depending on  $Q_k^{(b)}$*

*Proof.* See [129], Lemma 5.5.6. □

A major advantage of the EnKF is the conservation of the state ensemble, the cost for maintaining the ensemble can be balanced by the the computational saving through the efficiency of the scheme. Using the same covariance matrix for the evolution and the Kalman gain for ensembles can help towards reducing the computational cost. The ensemble size can also be chosen small for making the method more cost effective, but we need to be careful with the selection of the ensemble size. It should be chosen in a way that it is statistically representative for the state distributions [54]. The filter is suitable for parallelization, since each ensemble member is evolved through time independently [82]. It approximates the background covariance matrix flow of the Kalman filter.

There are many problems associated with the practical use of the EnKF. The main problem is that the size of the ensemble is too small, such that it does not reflect the statistics of model. Then under-damping phenomena occur [14, 15], which lead to other problems like inbreeding, filter divergence and spurious correlations. The same problems can appear when the square root filter is used.

## Chapter 3

# Transformed and Generalized Localization for Ensemble Methods in Data Assimilation

The goal of this work is to suggest a general algorithm which allows localization in the case of *non-local observation operators*. It is not restricted by the limit of locality. To test its applicability we study the effect of the transformed localization when applied to a IASI retrieval problem [32, 148], which is known to have a strong non-locality and is of high interest to both the research community and operational centers of weather prediction.

*First*, we will carry out the *localization* in the case of *non-local operators* by using a *transformation* of the state space  $X$  and the data space  $Y$  under consideration which transforms the non-local operator into a local operator. Then, in the *transformed space* classical localization is employed to enhance the degrees of freedom of the data assimilation task. We interpret this *transformed localization* or *generalized localization* as a *projection method* in a sequence of subspaces of the state space  $X$  or observation space  $Y$ , respectively. An important complication here is the role of the covariance matrix. It needs a corresponding transformation as well, and we will discuss under which conditions the problem becomes localizable under the transformation.

*Second*, we formulate an abstract model for *localization*, which no longer takes place in the *state space*, but considers the more general algebraic model for localization, which is based on general *projection methods* in *inverse problems* or *finite elements*. We will see that classical localization for the EnKF can be understood as a special case of such a projection method, and that our transformation of the operator corresponds to choosing a particular subspace for projection.



*Third*, we study the reconstruction of an *atmospheric temperature profile* from IASI radiance measurements [41, 111, 167, 148]. In principle, the observation operator can be viewed as a discretized version of an integral equation of the form

$$f(s) = \int_G w(s, \tau) \varphi(\tau) d\tau, \quad s \in M \quad (3.0.1)$$

with  $w(s, \tau)$  being the sensitivity of the IASI measurement in channel  $s$  with respect to temperature variations at point  $\tau$  in the atmospheric column  $G$ . For IASI assimilation in each assimilation step and each atmospheric column (where data is available) we explicitly or implicitly solve the *integral equation of the first kind* (3.0.1). This is known to be an *ill-posed inverse problem*, and we show how transformed localization can help to reduce the dimensionality of the problem as it is carried out by the localization for the EnKF. Numerical examples are provided.

Before we come to our techniques in more detail, we first provide a survey about the role of localization. We will discuss the implications of *undersampling* and the role localization plays for creating a filter which does fulfill its purpose.

### 3.1 Issues Due to Small Number of Ensembles

The Ensemble Kalman filter is subjected to certain limitations. To develop its full potential we should carry out all computations with a large ensemble size. But then it's powerful computational mechanics cannot not be easily implemented because the cost is prohibitively large.

The cost depends on the ensemble size: when we reduce the number of ensembles, this reduces the cost as well. However, care must be taken reducing the ensemble size to ensure that it is not too small, since then it loses the statistical representativity and does no longer properly reflect the characteristics of the system. It should be statistically representative of the system as demanded by Kalnay [93]. We note that the choice of a statistically representative ensemble also depends on the size of state space.

A small ensemble causes problems of different nature. Obtaining the correct weight for the background error is always difficult, Reinhard Furrer and Thomas Bengtsson [63] show that this problem is exacerbated in the situation with a small number of ensemble members. They show that the filter is putting more trust to the background state than it should. This means that the measurement data is not really impact full when producing the analysis state. The filter will lead to a purposeful state estimate if the forecast state is the appropriate representation of the true state, but if our background is weighted too strongly, this is no longer the

case. The utilization of small ensembles also generates correlations between different components where there is no physical relation [69]. They are called *spurious correlations*, they cause severe problems in the naive implementation of ensemble filters.

Thus, the ensemble size is a key component for a meaningful forecast based on the ensemble Kalman filter as experimented intensely for example [82, 147]. The success of the ensemble Kalman filter will depend on the ensemble should be statistically representative but the ensemble must span the model subspace adequately as described by Oke et al. [135].

## Undersampling

We have argued that the size of the ensemble is a very critical issue for the Ensemble Kalman Filter. If the size is very large, then the computational cost and memory problems are severe. If it is very small, then the ensemble is not a proper statistical representative of state uncertainty.

The necessary size of the ensemble depends on the size of the state space. Currently, the state space size of numerical weather prediction models is at least  $10^8$ . This is very large, so we need large ensembles for it to be statistically representative.

If sample size is insufficient, then the problem is said to be under sampled. Undersampling leads to a reduced rank background error covariance matrix due to which there is a tendency that variance and covariance are under estimated. A typical consequence of filter underestimation of covariances are long range correlations [69].

## Inbreeding or Underestimation Of Covariance

The term *inbreeding* is used by Houtekamer and Mitchell [82, 146] to describe phenomena with respect to the Ensemble Kalman Filter due to undersampling. Recall that the background error covariance is evolved by model dynamics, thus known as *forecast error covariance*. If the analysis error covariance is systematically underestimated at each assimilation cycle, this situation is denoted as *inbreeding*.

The analysis error covariance  $B^{(a)}$  is always smaller than the forecast error covariance  $B^{(b)}$ . We already defined their relation by

$$B^{(a)} = (I - KH)B^{(b)} \quad (3.1.1)$$

where  $B^{(b)} = E[(\varphi^{(b)} - \varphi^t)(\varphi^{(b)} - \varphi^t)^T]$  is degree of uncertainty in the forecast state of the the system. The Kalman gain is  $K := B^{(b)}H^*(R + HB^{(b)}H^*)^{-1}$ .

Here the Kalman gain is based on the ratio of the background error covariance and the error covariance of measurements, to compute the weight that should be

given to background state and weight put on the observations. Then, by this ratio of background and observation covariance matrices the Kalman gain is used to adjust the background state estimate. Note that the adjustment of the forecast state will be incorrect if any of the observation errors and background errors are incorrectly specified.

This phenomenon is noted by many authors: Small ensembles that do not span an appropriate part of the model space give rise to inbreeding due to error in the covariance estimate (Lorenz [11]). The smaller the ensemble is, the greater the degree of undersampling that is present and the greater the chance of underestimated forecast error covariances (Ehrendorfer [48]). The Ensemble Kalman Filter updates the ensemble at analysis stage based on the observations and the Kalman gain, here with large danger of underestimation by the filter (Whitaker and Hamil [171]). In the Ensemble Kalman Filters that use contaminated observations (such as Evensen [55]) there occurs inbreeding problems due to error in the estimation of observation covariances with additional sampling error. For the Square Root EnKF we do not use noisy observation, which reduces the tendency of inbreeding. Inbreeding is a cause of many problems of filters like long range correlation and filter divergence (Hamil et al. [72, 147, 69]).

## Filter Divergence and Collapse

*Filter divergence* is a phenomenon which occurs in almost all kinds of filtering techniques. When the forecast error covariances are too large then there is uncertainty in the forecast state and undue confidence on observations. This means less weight is given to the forecast state and more to the observations.

On the other hand, if the forecast error covariance is unrealistically small, then there is undue confidence placed on the forecast state, because due to Kalman gain less weight is placed on the observations and more on forecast state. As during filtering the covariance is progressively underestimated, it places undue confidence on the forecast state estimate and less weight on observations. As a result, observation data is ignored, this is also called *filter divergence* in the sense of a *filter collapse*. If the breadth of the ensemble distribution is less in the analysis state compared to the forecast state, it implies that ensembles are converging. Inbreeding makes it difficult for a filter to adjust the erroneous forecast state to the true state and filter collapse occurs [69].

## Spurious Correlations

In the literature, the term *spurious correlation* is coined by Karl Pearson [153, 12] which is usually used for correlations of the ratio of absolute measurement rather than the a correlation of actual measurements. Spurious correlations are correlations that appear between state variables which are at significant distance and are not physically related to the true error covariance of the forecast states.

In ensemble filters, observations made at one location have influence on state variables which are physically remote from the observation by filter equations, there is a tendency of long rang correlation (Anderson [16]). Consequently physically remote observation erroneously influence the state variables. With the decrease in ensemble size and correlation between state components, the covariance error increases (Hamil et al. [72]).

In the physical world the correlations between given observation points decrease with distance. The error in the forecast error covariance matrix at grid points remote from observations is expected to be increase as compare to true correlation at that grid. Hamil et al. [72] argue that the analysis estimate of ensemble Kalman filter is less accurate as the error in the covariance estimate also known as *noise* is greater then the true correlation, known as *signal*. Correlation is inversely proportion to distance while relative error varies directly.

As correlations are expected to become smaller with distance while relative errors increase, the state variables distant from observations are expected to inherit a large *noise to signal ratio*. These are called *long range spurious correlations* and they decrease the quality of the analysis estimate. Also Hamil argues in [72] that the ratio of the covariance estimate error and the true correlation depends on ensemble size. The larger the ensemble size, the smaller the noise. Thus, spurious correlations are due to under sampling. Lorenc [11] investigates the case where we we have  $N$  ensembles generated from some probability distribution. Then, the error in the covariance estimate is proportional to  $\frac{1}{N}$  [69], the convergence order of the standard stochastic covariance estimator.

## 3.2 Remedy for undersampling

There are many methods to deal with the issue of undersampling. *Covariances inflation* and *covariance localization* [15, 72, 82, 83, 88, 136, 171]) are already mentioned in our previous Section 3.1. They are employed for the *local ensemble transform Kalman filter* [136, 162] and EnKF [82]. We will describe more details in the following parts.

### 3.2.1 Covariance Inflation

One of the common problems experienced after applying different filtering techniques is the filter divergence ([90]). As this happens when forecast error covariances become too narrow and the observations progressively lose their impact, then the analysis generated by the filter drifts away from the true state. The Kalman filter underestimates the forecast error covariance due to model errors, nonlinearities and the effects described in our previous sections. Then, for stable filtering it is necessary to *inflate* the forecast error covariance to give adequate weight to the observations. This broadens the forecast error covariance and enhances the impact of observation. If the ensemble is not inflated, then it amplifies the errors and causes filter divergence [69].

The principle of *covariance inflation* is simple: multiply the covariance matrix  $B^{(a)}$  or  $B^{(b)}$  by some inflation factor  $\beta > 0$  in each assimilation cycle. This broadens the covariance matrix and avoids filter divergence. The calculation is

$$B^{(b)} \leftarrow \beta B^{(b)},$$

where the inflation factor  $\beta$  is slightly greater than 1. Obviously, if  $\beta$  is very large it means giving more weight to observation, so  $\beta$  must be chosen with care. The specification of an optimal inflation depends on the size of the ensemble, on the model dynamics and on the type of the ensemble filter used. A popular method is choosing  $\beta$  by experimentation. Hamil et al. (2001) by experiments conclude that 1% inflation factor is optimal for a large range of numbers of ensemble members. While Whitaker and Hamil (2002) found that an inflation factor is optimal with 7% for the ensemble Kalman filter and 3% for the square root ensemble Kalman filter [72, 73, 166, 171].

Anderson and Anderson (1999) [18] call the above method *multiplicative covariance inflation*. A large inflation may disturb the physical balance [15]. Inflation does not settle the problem of spurious correlation but it overcomes the inbreeding problem. For long range correlation more sophisticated methods are needed [69].

### 3.2.2 Localization

The Ensemble Kalman Filter (EnKF) is a good alternative to the Extended Kalman filter and 4DVAR for atmospheric and oceanic assimilation, due to which we do not need to deal with tangent linear and adjoint models and we have to deal with covariance matrices of rank  $O(10^6)$  only.

The main disadvantage of the EnKF is that the ensemble size and rank of the covariance matrix are very small of  $O(10^2)$ . This means that the EnKF has state

covariance to be of excessively low rank. It leads to spurious long-range correlations and the consequence in the analysis is a bad fit to the observations.

Further, the spurious long-range correlations cause an incorrect long-range response to a single observation in both the ensemble mean and in the variance, which is distributed globally when a single observation is assimilated. To achieve feasible performance from an EnKF it is indispensable to filter out these spurious correlations; this step is known as *localization*. For localization there are different approaches like to Schur- multiply the ensemble-estimated covariance matrix by a compactly- supported distance-dependent covariance function (Houtekamer and Mitchell 2001 [83]). The function of Gaspari and Cohn (1999, equation (4.10), henceforth GC [64]) has been widely used, we explain more details below. This procedure removes all long-range correlations and increases the rank of the covariance matrices with an increase of the ability of the analysis to fit observation well.

The accuracy of background-error is a very hard issue for systems used to assimilate data. Ensemble based assimilation schemes depend on the number of ensembles of forecasts for the calculation of background error. But the size of ensembles is always smaller than the dimension of the systems state space for the real numerical predictions, so the imperfectly calculated estimate gives insufficient background error with the effect of poor performance.

A possible solution to perform a successful ensemble based assimilation when only a small sized ensemble is feasible is the use of technique called *localization*. Localization is a simple technique to localize the impact of an observation to a subset of the model state variables. The subset is defined as a close region around the observation. There are two advantages of localization. One is that it can eliminate spurious correlations of the background ensemble between distant state variables due to limited number of ensemble members. The other is that it increases the rank of the system by use of an effective number of independent ensemble members. So localization is a tool to enhance the number of degrees of freedom of the ensemble and to cease the false correlations [10, 11, 141].

Today it is well-known that ensemble filters need localization as a key tool to enhance their number of degrees of freedom and to suppress spurious correlations, compare [15, 62, 82, 83, 88, 144, 171].

Often, ensemble Kalman filters are applied to large domains  $\Omega$  (the whole globe in atmospheric models) with limited ensemble sizes ( $L = 40, \dots, 200$ ). Thus, the space spanned by the ensemble is much smaller than the state space of the underlying dynamical system and correlations between spatially well separated parameters may be large due to the representation by the stochastic process. Assuming a Gaussian distribution of errors, the expectation for the variance of the sample covariance

coefficients is

$$\mathbb{E}((QQ^T)_{ij} - B_{ij})^2 = \frac{1}{L-1} [B_{ii}B_{jj} + (B_{ij})^2]$$

[109]. The expected error does not only depend on  $B_{ij}$  but also on the diagonal elements  $B_{ii}$  and  $B_{jj}$ . The latter will always dominate the sampling error for those coefficients which actually correspond to small correlations.

### Localization in state space

Localization can be employed in both state space and observation space. In order to enhance the number of degrees of freedom of the ensemble and to suppress spurious correlations, it has been proposed to use localization on the ensemble  $B^{(b)}$  matrix, [30], [73], by replacing  $B^{(b)}$  with  $C \circ B^{(b)}$ , where  $C$  is the so called *localization matrix* and  $\circ$  is the Schur product, i.e. the element-wise product of the matrix elements  $c_{ij}$  and  $b_{ij}$  evaluated at the grid- points  $(i, j)$  of the model. This method is called *localization on B* or *covariance localization*.

A sufficient condition to ensure that  $C \circ B^{(b)}$  is a valid covariance matrix is that  $C$  itself is a valid covariance matrix (positive definite matrix). This is fulfilled if the  $c_{ij}$  are specific functions  $f(r_{ij}/l_{loc})$  of the spatial distance  $r_{ij}$  of the grid-points in the model domain, given for instance by the Gaspari & Cohn functions [64]. These functions, parameterized by the localization length scale  $l_{loc}$ , are very similar to a Gaussian with the respective standard deviation, but become exactly zero for distances larger than approximately 3 times  $l_{loc}$ . The optimal value of the localization length scale depends on actual correlation length scales and ensemble size.

In Covariance localization using  $C \circ B^{(b)}$  instead of  $B^{(b)}$  prevents application of the factorization of our upcoming equation (2.3.6) and instead of solving the low dimensional update equations (2.3.7), (2.3.20), (2.3.19) in ensemble space we have to handle a high dimensional problem in model state space. As problems of this size cannot be solved explicitly, iterative variational methods are recommended [34].

### Localization in observation space

An approach to solve the localized equations explicitly is to solve the equations in observation space (2.3.7), (2.3.8) and to apply the Schur product in observation space to  $HB^{(b)}BH^*$  and  $B^{(b)}H^*$  instead of  $B^{(b)}$ . This is equivalent for local observations (there individual observation operators  $H$  depend on only one spatial grid-point and the location of the observation is well defined) but is not strictly applicable for non-local observations. This method will be called *localization on HBH* or *observation localization*. Furthermore, in order to solve the equations efficiently the

equations have to be solved in patches [172] or sequentially [83, 171] making use of the multistep analysis method.

Another method of efficient spatial localization was proposed in [82]. Equations (2.3.7), (2.3.20) are solved individually at every model grid-point  $i$  taking into account observations at  $j$  within a distance  $r_{ij}$  from the grid-point only. The influence of the observations at the grid-point in general is weighted by applying a Schur product to the inverse of the observation error covariance matrix  $C \circ R^{-1}$ , where the localization function depends on the distance between the grid-point and the observation. This localization method will be called *localization on  $R^{-1}$*  and it is necessary for LETKF [87].

Let us now sum up the approaches in a notation suitable for application to our ensemble Kalman filter presented in the next Section 3.3.

Localization in its simplest version means that the analysis (2.2.23), see also (2.3.7), of the mean and the ensemble update (2.3.19), see also (2.3.19), is not carried out globally, but on some local set  $\Omega_x$  around a point  $x \in \mathbb{R}^3$ , where usually the set  $\Omega_x$  is a subset of a ball  $B_\rho(x)$  with center  $x$  and radius  $\rho$ . Often, weighting functions on  $B_\rho(x)$  are employed, to give observations close to  $x$  a full weight, those closer to the boundary are given less influence. This approach can be calculated in parallel on a grid of points  $x_\xi$ ,  $\xi = 1, \dots, N$ . Thus, for every set  $\Omega_{x_\xi}$ ,  $\xi = 1, \dots, N$  we obtain an analysis ensemble, which are combined into a global analysis ensemble by appropriate weighting functions. A recent deterministic error analysis for this type of localization can be found in [144], we will call it *non-coupled localization*.

A popular alternative for localization is to define a localization matrix  $C$  such that  $C(x, y) = 0$  for  $|x - y| > \rho$  with some localization radius  $\rho > 0$ . Then, we replace the covariance matrix  $B$  in the EnKF by its Schur product  $B \cdot C$ , such that the new covariance matrix is nonzero only if  $|x - y| < \rho$ . We remark that this type of localization needs to solve a full system, but now with a sparse  $B$ -matrix where spurious correlations have been removed. For a comparison of this localization with the above approach we refer to [88]. We call this localization *coupled localization*.

### 3.3 The EnKF and Localization

The Ensemble Kalman Filter (EnKF) is an ensemble data assimilation technique, which is based on a long history of works in different communities, e.g. [20, 30, 53, 55, 96, 166], see also [87] and [62]. It employs an ensemble  $\varphi^{(1)}, \dots, \varphi^{(L)}$  of states  $\varphi^{(\ell)} \in X$  on which the analysis is based. Assume that we are given data  $y_k \in Y$  at times  $t_1, t_2, t_3, \dots, t_k, \dots$ , for  $k \in \mathbb{N}$ . Here,  $Y$  is the observation space, which we



assume to be a Hilbert space, and  $X$  is the state space, which is a Hilbert space as well. The EnKF is a *cycled scheme*, which at every time  $t_k$ ,  $k = 1, 2, 3, \dots$

- propagates the ensemble from  $t_{k-1}$  to time  $t_k$  by applying the model dynamics, i.e.

$$\varphi_k^{(\ell,b)} = M(\varphi_{k-1}^{(\ell,a)}), \quad \ell = 1, \dots, L. \quad (3.3.1)$$

- At time  $t_k$  an analysis ensemble  $\varphi_k^{(\ell,a)}$  is calculated from the background or *first guess* ensemble, respectively,  $\varphi_k^{(\ell,b)}$ , and the observations  $y_k$ .

The analysis ensemble can be calculated following the approach of Evensen [53], or by a *square root filter*, compare the recent survey in [62]. Following [62], we employ the notation

$$Q_k^{(b)} := \frac{1}{\sqrt{L-1}} \left( \varphi_k^{(1,b)} - \varphi_k^{(b)}, \dots, \varphi_k^{(L,b)} - \varphi_k^{(b)} \right), \quad k \in \mathbb{N}, \quad (3.3.2)$$

where  $\varphi_k^{(b)}$  is the mean of the background ensemble. This leads to the covariance matrix  $B = Q^{(b)}(Q^{(b)})^*$  when the standard stochastic estimator is applied. Then, the *Kalman gain matrix* is given by

$$K_k = (Q_k^{(b)})(Q_k^{(b)})^* H^* (R + H(Q_k^{(b)})(Q_k^{(b)})^* H^*)^{-1}, \quad k \in \mathbb{N}. \quad (3.3.3)$$

First, the mean  $\varphi_k^{(a)}$  of the analysis ensemble is calculated using the *Kalman update equation*

$$\varphi_k^{(a)} = \varphi_k^{(b)} + K_k(y_k - H\varphi_k^{(b)}). \quad (3.3.4)$$

In the area of *inverse problems*, this is known as *Tikhonov regularization operator*. Second, we write the operation which generates the analysis ensemble in the form

$$(\varphi_k^{(1,a)}, \dots, \varphi_k^{(L,a)}) = F \left( (\varphi_k^{(1,b)}, \dots, \varphi_k^{(L,b)}) \right), \quad k \in N, \quad (3.3.5)$$

i.e.  $F$  maps the background ensemble into the analysis ensemble, it is the operator of the filtering step. As shown in [62], the square root filter can be written as

$$Q_k^{(a)} = Q_k^{(b)} \mathcal{L} \quad (3.3.6)$$

where  $\mathcal{L}$  is the square root (i.e.  $\mathcal{L}\mathcal{L}^* = A$ ) of the operators

$$A_k := I - (Q_k^{(b)})^* H_k^* \left( R + H_k Q_k^{(b)} (Q_k^{(b)})^* H_k^* \right)^{-1} H_k Q_k^{(b)}, \quad k \in \mathbb{N}. \quad (3.3.7)$$

For the original Kalman filter, an analysis covariance matrix  $B^{(a)}$  is calculated by

$$B_k^{(a)} = (I - K_k H_k) B_k^{(b)} \quad (3.3.8)$$

We note that this equation is compatible with the square root filter in the sense that

$$(Q_k^{(a)})(Q_k^{(a)})^* = (I - K_k H_k)(Q_k^{(b)})(Q_k^{(b)})^*, \quad k \in \mathbb{N}. \quad (3.3.9)$$

An important observation here is that by (3.3.3) the increment of the ensemble mean given by (3.3.4) is in the space spanned by the column of the matrix  $Q_k^{(b)}$  defined in (4.1.5). Also, we note that every ensemble member of the analysis ensemble  $Q_k^{(a)}$  defined in (3.3.6) is a linear combination of the ensemble members of  $Q_k^{(b)}$  (with coefficients depending nonlinearly on the ensemble members). Since in general, the number  $L$  of ensemble members is very small compared to the degree's of freedom of our state space  $X$ , we obtain a very crude approximation to the truth if we do not employ further measures to improve this approximation.

One key tool to achieve the above improved approximation within ensemble filters is *localization*. The basic idea of localization is to carry out the analysis given by the equations (3.3.4) and (3.3.6) not globally for all observations, but locally for observations around some point  $p \in \mathbb{R}^3$  in space. Here, we study a version of  $B$  localization to some subset  $U$ , which considers only those observations which are located in  $U$  and which restricts all functions and operators under consideration to  $U$ . Since we are in an  $L - 1$  dimensional subspace, in this case each analysis can be reduced to solving an  $(L - 1) \times (L - 1)$  dimensional linear system. This is carried out many times independently (with perfect parallelization properties), and in a final step the global analysis ensemble is constructed from the local assimilation results.

To be more specific, let  $p \in \mathbb{R}^3$  be a point and let  $\rho > 0$  denote our *localization radius*. We take all observations  $y_{p,\rho}$  which are located in the ball  $\mathcal{B}_\rho(p)$  of radius  $\rho$  around  $p$ . The corresponding observation operator is denoted by  $H_{p,\rho}$ .

We say that an observation  $y_j$  for  $j \in \mathbb{N}$  is *local*, if it belongs to one and only one point  $p \in \mathbb{R}^3$ . For linear observation operators  $H$  and finite dimensional spaces  $X$  and  $Y$  this means that the matrix representing  $H$  has one non-zero entry in its  $j$ -th row. Non-local measurements have more than one non-zero entry in its corresponding row of the observation operator.

Of course, for the case of discrete grids and when interpolation is employed, even observations which take place at one point in space only, lead to non-zero entries in more than one place in the observation operator  $H$ . Then, locality of an observation is linked to the particular realization of the observation operator. Later, we will only need the condition that some observation is located in a ball  $\mathcal{B}_\rho(p)$  of radius  $\rho$  around some point  $p$  in space. We call the observation to be located in  $\mathcal{B}_\rho(p)$  if the matrix  $H$  is zero outside of  $\mathcal{B}_\rho(p)$ .

### 3.4 Transformed Localization

The basic idea of a *transformed localization* is to employ a transform of both the state space  $X$  and the observation space  $Y$  such that the transformed observation operator is either local or has better localization properties.

We consider the setup of a state  $\varphi \in X$  and observations  $f \in Y$  and a linear observation operator  $H : X \rightarrow Y$ . Then, let a transformation of  $X$  given by an invertible operator  $T : X \rightarrow X$  and let a transformation of  $Y$  be given by an invertible operator  $S : Y \rightarrow Y$ . We define

$$\tilde{\varphi} := Tx, \quad \tilde{f} := Sf \quad (3.4.1)$$

and

$$\tilde{H} := SHT^{-1}. \quad (3.4.2)$$

In this case, by multiplication with  $S$  the equation

$$H\varphi = f \quad (3.4.3)$$

is transformed into

$$SH(T^{-1}T)\varphi = Sf, \quad (3.4.4)$$

which by (3.4.1) and (3.4.2) can be written as

$$\tilde{H}\tilde{\varphi} = \tilde{f}. \quad (3.4.5)$$

We call (3.4.5) the *transformed observation equation*.

In a first step, we study in what way our transforms  $S$  and  $T$  will influence the behaviour of the Ensemble Kalman Filter with square root analysis ensemble, in particular in what way they influence the *analysis mean* and the *analysis ensemble* when we do not employ localization. In this case the EnKF is robust under linear transformation. The *transform* on the ensemble matrix  $Q$  defined by (4.1.5) is given by

$$\tilde{Q} = TQ, \quad \tilde{Q}^* = Q^*T^*, \quad (3.4.6)$$

which is the matrix of the transformed ensemble members. The transformed Kalman gain matrix as given by (3.3.3) is then given by

$$\tilde{K}_k = \tilde{Q}^{(b)}(\tilde{Q}^{(b)})^*\tilde{H}^*(\tilde{R} + \tilde{H}\tilde{Q}^{(b)}(\tilde{Q}^{(b)})^*\tilde{H}^*)^{-1}, \quad k \in \mathbb{N}, \quad (3.4.7)$$

where  $\tilde{R}$  is given by

$$\tilde{R} = SRS^*. \quad (3.4.8)$$

We also note that

$$\tilde{H}^* = (T^{-1})^* H^* S^*, \quad (3.4.9)$$

such that

$$\begin{aligned} \tilde{H}\tilde{Q}^{(b)} &= SHT^{-1}TQ^{(b)} = SHQ^{(b)}, \\ (\tilde{Q}^{(b)})^*\tilde{H}^* &= (Q^{(b)})^*T^*(T^{-1})^*H^*S^* = (Q^{(b)})^*H^*S^*. \end{aligned} \quad (3.4.10)$$

LEMMA 3.4.1. *The transformed analysis increment  $\tilde{K}_k(\tilde{f}_k - \tilde{H}\tilde{\varphi}_k^{(b)})$  of the Ensemble Kalman Filter defined by (3.3.4) and (3.3.6) for the transformed ensembles (3.4.6) is given by*

$$\tilde{K}_k(\tilde{f}_k - \tilde{H}\tilde{\varphi}_k^{(b)}) = TK(f_k - H\varphi_k^{(b)}). \quad (3.4.11)$$

The analysis ensemble is given by

$$\tilde{Q}_k^{(a)} = \tilde{Q}_k^{(b)}\tilde{L} \quad (3.4.12)$$

with a matrix  $\tilde{L}$  which is the same as the update matrix  $L$  from the non-transformed case, i.e. we have

$$\tilde{L} = L. \quad (3.4.13)$$

*Proof.* We first calculate (in some places dropping the index  $k$  for brevity)

$$\begin{aligned} \tilde{K} &= \tilde{Q}^{(b)}(\tilde{Q}^{(b)})^*\tilde{H}^*(\tilde{R} + \tilde{H}\tilde{Q}^{(b)}(\tilde{Q}^{(b)})^*\tilde{H}^*)^{-1} \\ &= TQ^{(b)}(Q^{(b)})^*H^*S^*(SRS^* + SHQ^{(b)}(Q^{(b)})^*H^*S^*)^{-1} \\ &= TKS^{-1}, \end{aligned} \quad (3.4.14)$$

which leads to

$$\begin{aligned} \tilde{K}_k(\tilde{f}_k - \tilde{H}\tilde{\varphi}_k^{(b)}) &= TK_kS^{-1}(Sy_k - SHT^{-1}Tx_k^{(b)}) \\ &= TK_k(f_k - H\varphi_k^{(b)}), \end{aligned} \quad (3.4.15)$$

and we have shown (3.4.11). To investigate the analysis ensemble we calculate the transformed version of  $I - A$  with  $A$  given by (4.1.33)

$$\begin{aligned} \tilde{A} &= I - (\tilde{Q}^{(b)})^*\tilde{H}^*(\tilde{R} + \tilde{H}\tilde{Q}^{(b)}(\tilde{Q}^{(b)})^*\tilde{H}^*)^{-1}\tilde{H}\tilde{Q}^{(b)} \\ &= I - (Q^{(b)})^*H^*S^*(SRS^* + SHQ^{(b)}(Q^{(b)})^*H^*S^*)^{-1}SH^*T^{-1}TQ^{(b)} \\ &= I - (Q^{(b)})^*H^*(R + HQ^{(b)}(Q^{(b)})^*H^*Q^{(b)}) \\ &= A. \end{aligned} \quad (3.4.16)$$

Since  $\tilde{L}$  is the square root of  $\tilde{A}$ , we obtain  $\tilde{L} = L$  and the proof is complete.  $\square$   
With an analogous proof it is possible to show that the above results hold for the

Kalman Filter in general. We note that the transform of the  $B$  matrix is given by

$$\tilde{B} = TBT^* \quad (3.4.17)$$

and the transform of the  $R$  matrix by  $\tilde{R} = SRS^*$ .

**THEOREM 3.4.2.** *For the transformed Kalman Filter we have*

$$\tilde{K} = TKS^{-1}. \quad (3.4.18)$$

*For the analysis increment we obtain*

$$\begin{aligned} \tilde{\varphi}^{(a)} - \tilde{\varphi}^{(b)} &= \tilde{K}(\tilde{f} - \tilde{H}\tilde{\varphi}^{(b)}) \\ &= TKS^{-1}(\tilde{y} - \tilde{H}\tilde{x}^{(b)}) \\ &= TK(f - H\varphi^{(b)}). \end{aligned} \quad (3.4.19)$$

*Proof.* We calculate

$$\begin{aligned} \tilde{K} &= \tilde{B}\tilde{H}^*(\tilde{R} + \tilde{H}\tilde{B}\tilde{H}^*)^{-1} \\ &= TBT^*(T^{-1})^*H^*S^*\left(SRS^* + SHT^{-1}TBT^*(T^{-1})^*H^*S^*\right)^{-1} \\ &= TBH^*(R + HBH^*)^{-1}S^{-1}, \end{aligned} \quad (3.4.20)$$

where we used  $(T^{-1})^* = (T^*)^{-1}$  and  $(S^*)^{-1} = (S^*)^{-1}$ . This completes the proof.  $\square$  Finally, we note that for the *transformed* analysis matrix  $\tilde{B}^{(a)}$  in the Kalman filter we obtain

$$\begin{aligned} \tilde{B}^{(a)} &= (I - \tilde{K}\tilde{H})\tilde{B}^{(b)} \\ &= (I - TKS^{-1}SHT^{-1})TB^{(b)}T^* \\ &= T(I - KH)B^{(b)}T^* \\ &= TB^{(a)}T^*. \end{aligned} \quad (3.4.21)$$

This means that the transform and equation (3.3.4) for the covariance matrices are compatible.

### 3.4.1 Generalized localization concept

One typical understanding of localization is to solve equations on some local domain  $\mathcal{B}_\rho(p)$  around some point  $p$  and with localization radius  $\rho$ , also using only data which lives in this domain  $\mathcal{B}_\rho(p)$  as well. This means, we solve local linear systems

$$H_{p,\rho}\varphi = f_{p,\rho}, \quad (3.4.22)$$

with the operator  $H_{p,\rho}$  which is obtained from the operator  $H$  by restricting it to data defined on  $\mathcal{B}_\rho(p)$ , with  $f_{p,\rho}$  defined as the subset of data defined on  $\mathcal{B}_\rho(p)$ , and with background  $\varphi^{(b)}$ . It leads to a local analysis  $\varphi_j^{(a)}$  and to a local analysis ensemble  $\varphi_j^{(a,\ell)}$  which is valid in a neighborhood  $B_{\rho_j}(p_j)$  of  $p_j$ . We can write the *local equations* (3.4.22) in the form

$$P_{p,\rho}H\varphi = P_{p,\rho}f, \quad (3.4.23)$$

where  $P_{p,\rho}$  is a projection operator restricting to the ball defined by  $p$  and  $\rho$ . In the simplest finite dimensional case it is a matrix

$$P = \left( \begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 0 & I & 0 \\ \hline 0 & 0 & 0 \end{array} \right), \quad (3.4.24)$$

which is the identity in the region defined by  $p$  and  $\rho$  and zero in all other entries. Given a partition of unity

$$1 = \sum_{j=1}^J \chi_j, \quad (3.4.25)$$

on our domain of definition, where usually  $\chi_j$  is close to one at  $p_j$  and close to zero at all other points  $p_i$ ,  $i \neq j$ . Then, we define the global analysis by

$$\varphi^{(a)} := \sum_{j=1}^J \chi_j \cdot \varphi_j^{(a)} \quad (3.4.26)$$

based on the analysis  $\varphi_j^{(a)}$  at  $\mathcal{B}_{\rho_j}(p_j)$ . In a *generalized localization concept* we replace the simple operator  $P$  from (3.4.24) by a general family  $P_j$ ,  $j = 1, \dots, J$  of *projection operators* and solve equations

$$H_j\varphi = f_j \quad (3.4.27)$$

where

$$H_j := P_jH, \quad y_j := P_jf. \quad (3.4.28)$$

Also, let us consider appropriate subspaces  $X_j$ ,  $j = 1, \dots, J$  in  $X$ . We assume that the  $X_j$  satisfy

$$X_j \cap X_i = \{0\}, \quad i \neq j, \quad i, j = 1, \dots, J. \quad (3.4.29)$$

Now, we solve solutions to the equation

$$H_j\varphi = y_j \Leftrightarrow P_jH\varphi = P_jy \quad (3.4.30)$$

in the subspace  $X_j$  to calculate  $\varphi_j^{(a)}$ . Then, the full analysis is generated from these *partial analysis* vectors by

$$\varphi_{proj} := \sum_{j=1}^J \varphi_j^{(a)} \quad (3.4.31)$$

Equivalence of the full and the combined partial analysis is valid under particular conditions as given in the following lemma.

LEMMA 3.4.3. *Let  $H$  be a bijective operator between  $X$  and  $Y$ . Assume that  $Y_j := P_j Y$  is a sequence of projection spaces in  $Y$  and  $X_j$  is a family of subspaces  $X_j \subset X$  of  $X$ , such that*

$$Y = Y_1 \oplus \cdots \oplus Y_J, \quad X = X_1 \oplus \cdots \oplus X_J, \quad (3.4.32)$$

and

$$Y_j = H X_j. \quad (3.4.33)$$

Then, the solution  $x_{proj}$  to the projection method given by (3.4.31) using the unregularized solution to (3.4.30) in  $X_j$  is equal to the unregularized solution  $x_*$  to the equation (3.4.3).

*Proof.* By the direct sum (3.4.32) in  $Y$  we have the unique decomposition

$$f = \sum_{j=1}^J f_j, \quad (3.4.34)$$

where  $f_j = P_j f$ . We define  $\varphi_j := H^{-1} f_j$ , which is in  $X_j$  by the space condition (3.4.33). This leads to

$$\varphi_* = H^{-1} f = H^{-1} \sum_{j=1}^J f_j = \sum_{j=1}^J H^{-1} f_j = \sum_{j=1}^J \varphi_j = \varphi_{proj}, \quad (3.4.35)$$

and the proof is complete.  $\square$

*Remark.* The above reconstruction by projection looks nicely, but we have to note that calculating the decomposition (3.4.34) in general corresponds to solving a system of the type

$$f = \sum_{\xi=1}^m \tilde{f}_\xi \alpha_\xi, \quad (3.4.36)$$

where  $\tilde{f}_\xi$  are the basis vectors of the spaces  $Y_j$ . The calculation of the projection involves solving the linear system (3.4.36), i.e. the inversion of the matrix

$$A := (\tilde{f}_1, \dots, \tilde{f}_m). \quad (3.4.37)$$

In the case where  $X_j := \text{span}\{e_j\}$ ,  $j = 1, \dots, n$ , this means that  $A = H$ , and the projection is already equivalent to solving the full matrix with the observation operator, which might be an ill-posed problem, and it is not really what we want, since then the purpose of projection (as generalized localization) to simplify the problem is no longer reached. The big point here is to find the right projection space  $X_J$ . We will see that  $X_1, \dots, X_J$  can be chosen such that the spaces are orthogonal.  $\square$

In general, we are of course interested in the case of *regularized reconstructions*. This needs further conditions on the spaces and operators under consideration, since in general we will not obtain equality of the full regularized solution with the localized reconstructions. We first provide a positive results under additional assumptions and then give a counter example in the general case for dimension  $n = 2$ .

For the next arguments we will follow the notation introduced in [62], i.e. we employ the notation  $*$  for the adjoint operator or matrix with respect to the scalar product under consideration, and  $H'$  for the adjoint when the standard  $\ell^2$  scalar product is used. We also note that when  $H$  is restricted to some subspace  $X_j$  with image space  $H(X_j)$ , we use  $(H|_{X_j})^*$  to denote the adjoint with respect to the spaces  $X_j$  and  $H(X_j)$  or to  $X_j$  and  $Y_j$  if  $Y_j$  with  $H(X_j) \subset Y_j$  is given explicitly. For real valued spaces  $H'$  and  $H^T$  are identical. When we employ weighted scalar products with weights  $B^{-1}$  and  $R^{-1}$  with the covariance matrices  $B$  and  $R$  on  $X$  and  $Y$ , respectively, we obtain  $H^* = BH'R^{-1}$ . In this case, we can write the Kalman gain matrix in the form

$$A_\alpha := (\alpha I + H^*H)^{-1}H^* \quad (3.4.38)$$

with some scaling parameter  $\alpha$  where with  $\alpha = 1$  we obtain  $A_1 = K_k$ . For the following results we define

$$H_j = P_j H|_{X_j}, \quad j = 1, \dots, J, \quad (3.4.39)$$

where  $P_j$  is the projection operator defined above (3.4.27).

**THEOREM 3.4.4 (Localized Regularized Inversion).** *Under the conditions of the previous Lemma 3.4.3 we define*

$$\varphi_\alpha := A_\alpha f, \quad \varphi_{j,\alpha} := A_{\alpha,j} f_j, \quad j = 1, \dots, J, \quad (3.4.40)$$



where  $A_{j,\alpha}$  is defined by (3.4.38) with  $H$  replaced by  $H_j$ . Then we have

$$\varphi_\alpha = \varphi_{proj,\alpha} \quad \text{with} \quad \varphi_{proj,\alpha} := \sum_{j=1}^J \varphi_{j,\alpha}, \quad (3.4.41)$$

i.e. the full regularized solution  $\varphi_\alpha$  is equal to the sum of the regularized solutions under generalized localization.

*Remark.* We note that for the adjoint  $(H_j)^*$  of the operator  $H_j : X_j \rightarrow Y_j$  we have  $(H_j)^* f_j = H^* f_j$  for  $f_j \in Y_j$ .  $\square$

*Proof.* We first remark that on  $X_j$  we have

$$(\alpha I + H^* H)^{-1}|_{X_j} = (\alpha I + H_j^* H_j)^{-1}. \quad (3.4.42)$$

This yields

$$\begin{aligned} A_\alpha y_j &= (\alpha I + H^* H)^{-1} H^* f_j \\ &= (\alpha I + (H_j)^* H_j)^{-1} (H_j)^* f_j \\ &= A_{j,\alpha} f_j, \end{aligned} \quad (3.4.43)$$

for  $j = 1, \dots, J$ . With this, we calculate

$$A_\alpha f = A_\alpha \left( \sum_{j=1}^J f_j \right) = \sum_{j=1}^J A_\alpha f_j = \sum_{j=1}^J A_{j,\alpha} f_j = \sum_{j=1}^J \varphi_{j,\alpha} \quad (3.4.44)$$

and the proof for (3.4.41) is complete.  $\square$

*Example.* We now convince ourselves that in general the above equality (3.4.41) does not hold. To this end, we choose the simple setup  $X = \mathbb{R}^2$  and  $Y = \mathbb{R}^2$  and define

$$H = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} =: (\mathbf{h}_1, \mathbf{h}_2) \quad (3.4.45)$$

with vectors  $\mathbf{h}_1$  and  $\mathbf{h}_2$ . If we set  $X_1 = \text{span}\{e_1\}$  and  $X_2 = \text{span}\{e_2\}$ , we obtain

$$Y_1 = \text{span}\{\mathbf{h}_1\} = \text{span}\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}, \quad Y_2 = \text{span}\{\mathbf{h}_2\} = \text{span}\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$$

the projection operators onto  $Y_1$  and  $Y_2$  along the spaces  $Y_2$  and  $Y_1$ , respectively, is given by

$$P_1 = \frac{1}{3} \begin{pmatrix} 4 & -2 \\ 2 & -1 \end{pmatrix}, \quad P_2 = \frac{1}{3} \begin{pmatrix} -1 & 2 \\ -2 & 4 \end{pmatrix} \quad (3.4.46)$$

In this case, for  $\alpha = 0.1$  the localized solution is given by

$$\varphi_{proj,\alpha} = \begin{pmatrix} 0.65359 & -0.32680 \\ -0.32680 & 0.65359 \end{pmatrix} f, \quad (3.4.47)$$

where the full regularized inverse  $x_\alpha = A_\alpha y$  is given by

$$\varphi_\alpha = \begin{pmatrix} 0.61938 & -0.28971 \\ -0.28971 & 0.61938 \end{pmatrix} f. \quad (3.4.48)$$

The MATLAB numerical output illustrates the significant differences which appear in general.  $\square$

We next describe the role of localization for the full ensemble Kalman filter as described by (3.3.4) - (3.3.8). Let us first describe the main heuristics and then present some exact results.

- In general, if the observation operator  $H$  is non-local, then even if the  $B$  matrix is local, the term  $HBH^*$  is non-local and in (3.3.3) the inversion will need to solve a full system. In this case, localization will lead to large errors.
- However, if we transform the state space  $X$  and the observation space  $Y$  in a way such that  $\tilde{H}$  is local and  $\tilde{B}$  is local as well, then we can achieve locality of the terms  $\tilde{H}\tilde{B}\tilde{H}^*$ ,  $\tilde{B}\tilde{H}^*$  and  $\tilde{R}$ , such that  $\tilde{K}$  remains local. Then, localization applied to the transformed version of the EnKF will yield small approximation errors.

We will show in our examples that for practically important problems such as the data assimilation of infrared radiances in numerical weather prediction, the above situation can be achieved.

We will first derive an exact result for *block-local* matrices to clarify the situation further. We call a matrix  $B$  *block-local*, if it is zero outside of some blocks connecting variables located within a given number of domains. We can reorder the variables such that we obtain a block-diagonal matrix, i.e.

$$B = \left( \begin{array}{c|c|c|c} \dots & 0 & 0 & 0 \\ \hline 0 & \dots & 0 & 0 \\ \hline 0 & 0 & \dots & 0 \\ \hline 0 & 0 & 0 & \dots \end{array} \right), \quad (3.4.49)$$

where the lines indicate blocks of variables of appropriate size.

THEOREM 3.4.5. *Assume that the matrix  $B$  is block-diagonal,  $R$  is diagonal and that  $H$  is a full matrix. Then, in general, the matrix  $R + HBH^T$  is a full matrix and so is  $BH^T$ . Now, assume that we have a transformation  $T : X \rightarrow X$  and  $S : Y \rightarrow Y$  with  $T^T = T^{-1}$  and  $S^T = S^{-1}$ , such that  $\tilde{B} = TBT^T$  is block-diagonal and that  $\tilde{H} = SHT^{-1}$  is block diagonal and the location of the blocks coincide. Then, also the Kalman update matrix  $\tilde{K}$  is block-diagonal and we can fully localize the Kalman update into these blocks without loss of precision.*

*Proof.* We first remark that in the case where  $S^T = S^{-1}$  and for  $R$  diagonal with  $S^T = (s_1, \dots, s_m)$  we have  $s_j^T s_k = \delta_{jk}$  for  $j, k = 1, \dots, m$ , such that

$$\begin{aligned} SRS^T &= \begin{pmatrix} s_1^T \\ \vdots \\ s_m^T \end{pmatrix} \begin{pmatrix} r_1 & 0 & \dots & \\ 0 & r_2 & & 0 \\ \vdots & & \ddots & \\ \dots & & 0 & r_m \end{pmatrix} (s_1, \dots, s_m) \\ &= \begin{pmatrix} r_1 s_1^T s_1 & 0 & \dots & \\ 0 & r_2 s_2^T s_2 & & \dots \\ \dots & 0 & \ddots & 0 \\ \dots & \dots & 0 & r_m s_m^T s_m \end{pmatrix}, \end{aligned} \quad (3.4.50)$$

such that  $SRS^T$  remains diagonal. In the block-diagonal case, a full matrix  $H$  and the block-diagonal matrix  $\tilde{H}$  can be written in the form

$$H = \left( \begin{array}{c|c|c|c} \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{array} \right), \quad \tilde{H} = \left( \begin{array}{c|c|c|c} \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots \end{array} \right), \quad (3.4.51)$$

where the dots indicate occupied blocks and 0 stands for blocks of zeros. In this case, we readily verify the form

$$\begin{aligned} HBH^T &= \left( \begin{array}{c|c|c|c} \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{array} \right) \begin{pmatrix} \dots & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots \end{pmatrix} \begin{pmatrix} \dots & \dots \\ \dots & \dots \\ \dots & \dots \\ \dots & \dots \end{pmatrix} \\ &= \left( \begin{array}{c|c} \dots & \dots \\ \dots & \dots \end{array} \right) \end{aligned} \quad (3.4.52)$$

and

$$\begin{aligned}
 \tilde{H}\tilde{B}\tilde{H}^T &= \left( \begin{array}{c|c|c|c} \dots & \dots & 0 & 0 \\ \hline 0 & 0 & \dots & \dots \end{array} \right) \left( \begin{array}{c|c|c|c} \dots & 0 & 0 & 0 \\ \hline 0 & \dots & 0 & 0 \\ \hline 0 & 0 & \dots & 0 \\ \hline 0 & 0 & 0 & \dots \end{array} \right) \left( \begin{array}{c|c} \dots & 0 \\ \hline \dots & 0 \\ \hline 0 & \dots \\ \hline 0 & \dots \end{array} \right) \\
 &= \left( \begin{array}{c|c} \dots & 0 \\ \hline 0 & \dots \end{array} \right) \tag{3.4.53}
 \end{aligned}$$

An analogous calculation can be obtained for  $BH^T$  and  $\tilde{B}\tilde{H}^T$ , from which the final statement of the theorem is obtained.  $\square$

The above situation and result can help us to understand in what sense transformed localization helps to improve the EnKF. First, if we carry out the EnKF in an  $n$ -dimensional space with more than  $n$  ensemble members, where the ensemble  $B$  matrix has rank  $n$ , then there is no projection property of the EnKF, but it is basically solving a least squares problem in the full space  $\mathbb{R}^n$ . Still, the  $B$  matrix is an approximation, of course, but there is no further algebraic approximation by projecting on the ensemble subspace any more.

If we now start with a space whose dimension  $n$  is larger than the dimension  $L$  of the ensemble space, in addition to having the approximation of the  $B$  matrix by the ensemble estimator, we also have the projection of the EnKF analysis into the low-dimensional ensemble space. However, if we can transform our equation into the above block form with blocks of size smaller than  $L$ , we can then equivalently solve the update equation in each of the blocks, which can be fully carried out by using the ensemble. So with the transform we achieve an equivalent form which can be carried out exactly by the EnKF with projection error.

In general, we cannot expect a transformation into block form. But in important applications it is possible to define a transform  $T$  which makes the transformed covariance matrix  $\tilde{B}$  to have off-diagonal elements to be small far away from the diagonal. Then, localization in transformed space provides a good approximation to the original problem and can then be carried out efficiently by the EnKF.

To summarize, the concept of *transformed localization* is as follows:

1. Input is our measurements in a space  $Y$ , the observation operator  $H$  from the state space  $X$  into  $Y$ , and the ensemble  $Q^{(b)}$  defined in  $X$ .
2. We first calculate a transformation  $T : X \rightarrow X$  and  $S : Y \rightarrow Y$  such that  $\tilde{B} = TBT^*$  and  $\tilde{H} = SHT^{-1}$  are either diagonal or have small elements in off-diagonal matrix elements far away from the diagonal.

3. Then, we solve the localized transformed equations in each area or block given by the transformations  $T, S$ ,
4. The solution in each area or block are composed into a global analysis  $\tilde{\varphi}^{(a)}$  in transformed space.
5. The transformed analysis  $\tilde{\varphi}^{(a)}$  is mapped back into the original state space  $X$ , i.e. we calculate  $\varphi^{(a)}$ .

If we denote the localized transformed matrix by  $\tilde{B}_{k,gl}$  (where  $gl$  stands for *generalized localization*), this means we calculate

$$\tilde{\varphi}_{gl}^{(a)} := \tilde{\varphi}^{(b)} + \tilde{K}_{k,gl}(\tilde{f}_k - \tilde{H}\tilde{\varphi}_k^{(b)}), \quad (3.4.54)$$

with

$$\tilde{K}_{k,gl} := \tilde{B}_{k,gl}\tilde{H}^*(\tilde{R} + \tilde{H}\tilde{B}_{k,gl}\tilde{H}^*)^{-1} \quad (3.4.55)$$

with  $\tilde{R}$  given in (3.4.8). According to Theorem 3.4.2 and (3.4.17) it is equivalent to

$$\varphi_{gl}^{(a)} := \varphi^{(b)} + K_{k,gl}(f_k - H\varphi_k^{(b)}), \quad (3.4.56)$$

with

$$K_{k,gl} = B_{k,gl}H^*(R + HB_{k,gl}H^*)^{-1} \quad (3.4.57)$$

for  $B_{k,gl} = T^{-1}\tilde{B}_{k,gl}(T^*)^{-1}$  in the original space, transformed into each other by  $T$  and  $S$ .

### 3.4.2 Transform based on Singular Value Decomposition

Singular value decomposition SVD is a widely spread for spectral analysis of matrix and operator equations in numerical mathematics and applications. For the case of a matrix  $H \in \mathbb{R}^{m,n}$ , according to SVD we find orthonormal matrices  $V \in \mathbb{R}^{n,n}$  and  $U \in \mathbb{R}^{m,m}$  and a diagonal matrix  $S \in \mathbb{R}^{m,n}$ , such that

$$H = USV^*. \quad (3.4.58)$$

The singular value decomposition corresponds to an orthonormal basis of the state space  $X$ , given by the columns of  $V$ , i.e.

$$V = (v_1, \dots, v_n), \quad \langle v_j, v_k \rangle = \delta_{jk}, \quad j, k = 1, \dots, n, \quad (3.4.59)$$

and an orthonormal basis of the observation space  $Y$ , given by the columns of  $U$ , i.e.

$$U = (u_1, \dots, u_m), \quad \langle u_j, u_k \rangle = \delta_{jk}, \quad j, k = 1, \dots, m. \quad (3.4.60)$$

If we define

$$X_j := \text{span}\{v_j\} \quad \text{for } j = 1, \dots, n \quad (3.4.61)$$

and

$$Y_j = \text{span}\{u_j\} \quad \text{for } j = 1, \dots, m, \quad (3.4.62)$$

we obtain (3.4.33).

**COROLLARY 3.4.6.** *Singular value decomposition of the operator  $H : X \rightarrow Y$  provides a transformation of both the state space  $X$  and the observation space  $Y$  such that the resulting matrix equation (3.4.5) is diagonal and, thus, can be fully localized in the spaces  $(\tilde{X}, \tilde{Y})$ .*

Clearly, the corollary does not include the simultaneous transformation of  $B$  and  $H$ . In general, the transformed matrix  $\tilde{B} = TBT^*$  does not need to be local any more - you can easily think of covariance matrices  $B$  and transformations  $T$  where  $\tilde{B}$  is non-local. But for important applications the singular value decomposition provides the desired properties also for  $\tilde{B}$ . This leads us to our next section.

## 3.5 Transformed Localization for Atmospheric Radiance Inversion

The assimilation of atmospheric radiances is very important for numerical weather prediction, compare e.g. [154, 59, 140, 101, 110, 97, 42]. Depending on its temperature, each part of the atmosphere emits radiation in a wide range of frequencies. The radiation is attenuated and scattered from other parts of the atmosphere. Modern instruments such as IASI (c.f. [80, 111, 156]) or HIRS (c.f. [164, 161, 113, 106, 160]), ATMS (c.f. [123, 124, 66]) or AMSU-A/B (c.f. [155, 79, 21, 179, 115]) record the radiation within a satellite orbit.

Operational centers and many academic groups have developed *radiative transfer code* to simulate the radiation and attenuation of infrared as well as microwave and near visible radiation, compare [33]. Here, we will abbreviate the operator which maps an atmospheric state  $\varphi$  onto brightness temperatures  $f$  by  $\mathcal{H}$ , i.e.  $f = \mathcal{H}(\varphi)$ . In general  $\mathcal{H}$  is non-linear, and standard codes also calculate the linearization  $H$ , i.e. the *sensitivities* of the brightness temperatures  $f$  with respect to the atmospheric variables  $\varphi$ . This leads to a linearized equation of the form

$$f = \mathcal{H}(\varphi^{(b)}) + H(\varphi^{(b)})(\varphi - \varphi^{(b)}). \quad (3.5.1)$$

The operators  $\mathcal{H}$  and  $H$  are highly non-local. We display the rows of  $H$  in an atmospheric column in Figure 3.1, where different colours are used for the different

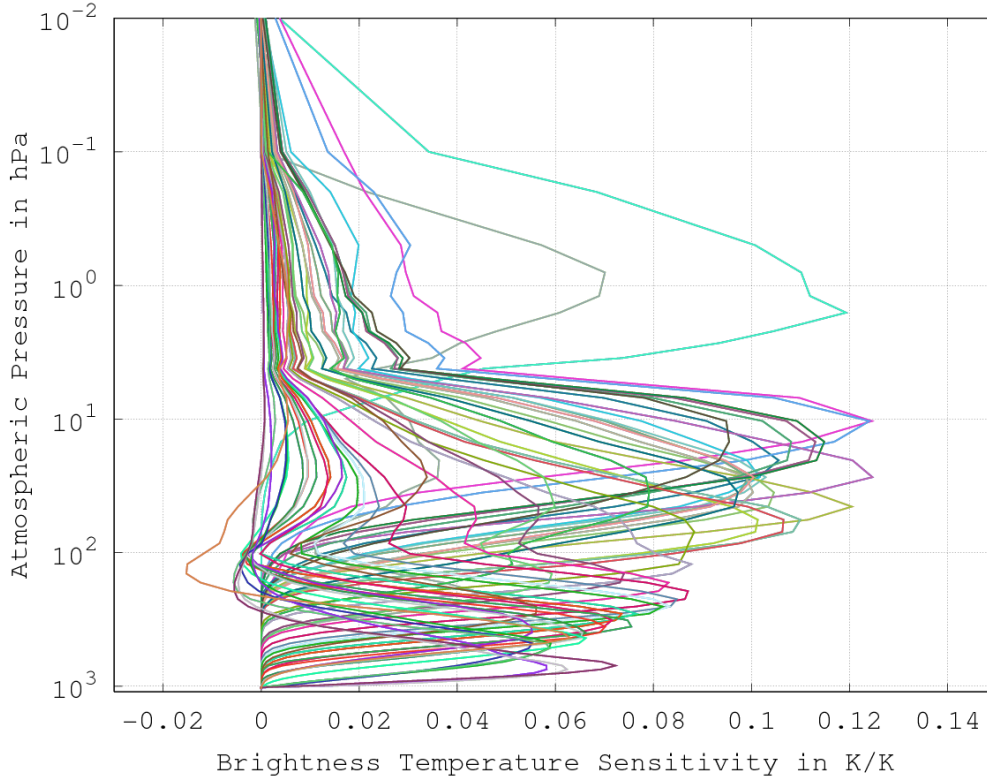


Figure 3.1: We show the sensitivity functions of the temperature sensitive IASI channels of the RTTOV operator  $H$ , i.e. the rows  $(H_{j1}, \dots, H_{jn})$  for  $j = 1, \dots, m$  are displayed.

rows of  $H$ . In the figure,  $H$  is dimensionless, since it is defined as the change of the brightness temperature in Kelvin when the temperature (in Kelvin) in a location of the atmospheric column is changed. The vertical coordinate is atmospheric pressure in hPa.

The application of the operator  $H$  can be viewed as a discretized version of an integral operator of the form

$$(Hx)(s) = \int_0^a w(s, \tau)x(\tau)d\tau, \quad s \in [s_0, s_1], \tag{3.5.2}$$

where  $w(s, \cdot)$  is one of the nonlocal functions displayed in Figure 3.1. It is well-known that such a problem is exponentially ill-posed, compare [44, 62], which is verified by the singular values which we display in Figure 3.2.

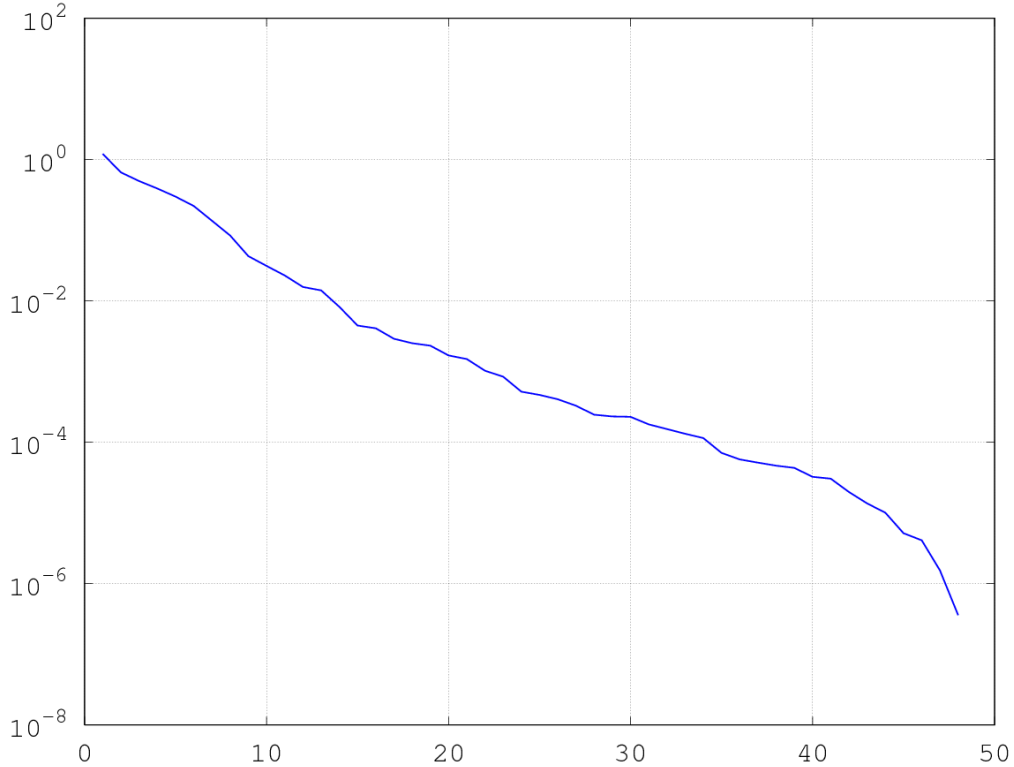


Figure 3.2: The singular values of  $H$  are displayed in a semilogarithmic scale. This shows that the singular values of the temperature sensitive IASI channels of the RTTOV operator tend to zero exponentially, and the inversion of radiances in an atmospheric column is an exponentially ill-posed inverse problem.

Assume that we have a localized background error correlation matrix  $B$ , as shown for example in Figure 3.3 (a).

1. We construct  $B$  to be either a *Gaussian* matrix, i.e. we define

$$B := \left( e^{-\sigma|\varphi_j - \varphi_k|^2} \right)_{jk=1, \dots, n}, \quad (3.5.3)$$

where  $x_j$  is the height of the level with index  $j$ , or  $B$  is the matrix as given in (3.3.3) constructed from the LETKF, compare Figure 3.4 (a). Here, we have used the experimental global LETKF as it is under development by DWD for the new ICON model. The LETKF runs with 40 ensemble members and all standard conventional and remote sensing data. We have added some Gaussian



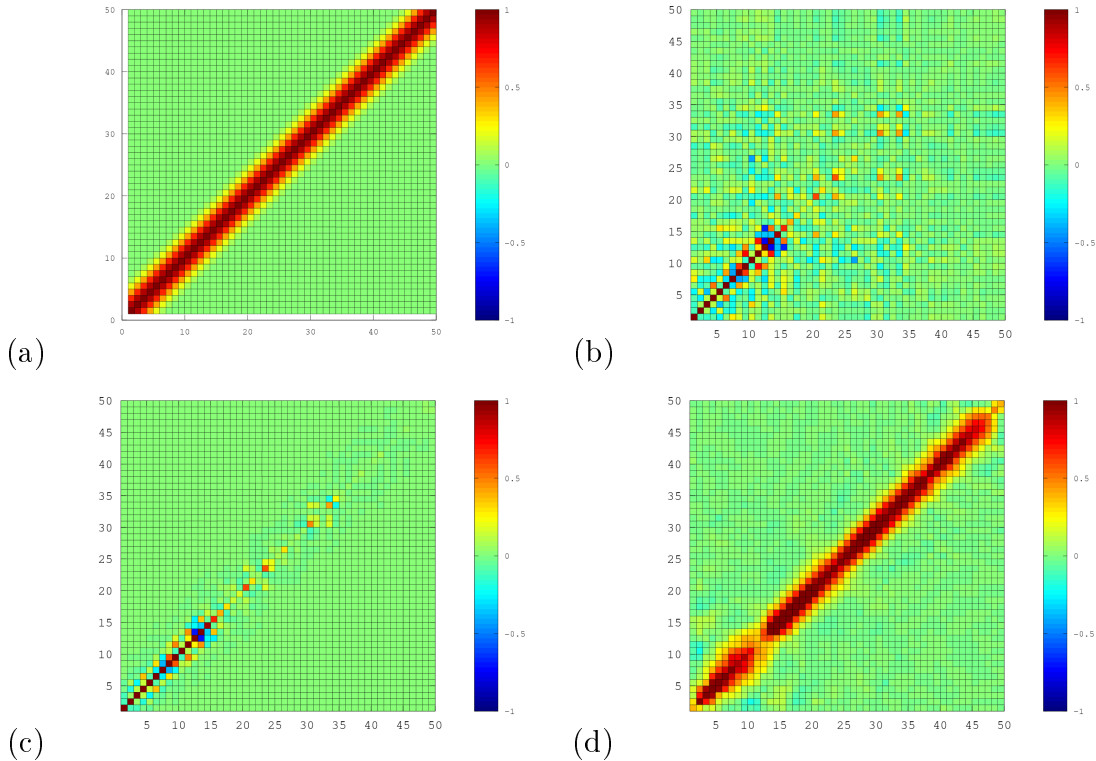


Figure 3.3: We display a local B-matrix in (a), which has been constructed according to (3.5.3) as a Gaussian matrix in the column. The transformed matrix  $\tilde{B}$  is displayed in Figure (b). We localize the matrix  $\tilde{B}$  by multiplication with the matrix shown in (a), the result is displayed in (c). Figure (d) shows the matrix  $B_{gl}$  defined as the backtransformed matrix from  $\tilde{\tilde{B}}$ .

model error term, to stabilize the system, since currently there is too much spread in the upper part of the atmosphere (low level numbers).

2. We then transform  $B$  using the transformation  $T$  which is calculated from the singular value decomposition of the RTTOV operator  $H$ . The transformed matrix  $\tilde{B} = TBT^*$  is shown in Figure 3.3 (b). It is quite remarkable that it has large values only on the diagonal and an area very close to the diagonal. We will discuss this at the end of this section. We found a similar phenomenon for the ICON-LETKF global system which is under development at DWD.
3. We have then localized the matrix  $\tilde{B}$  in the transformed space  $\tilde{Y}$  either by multiplication with a Gaussian matrix as in Figure 3.3 (a) or by setting values

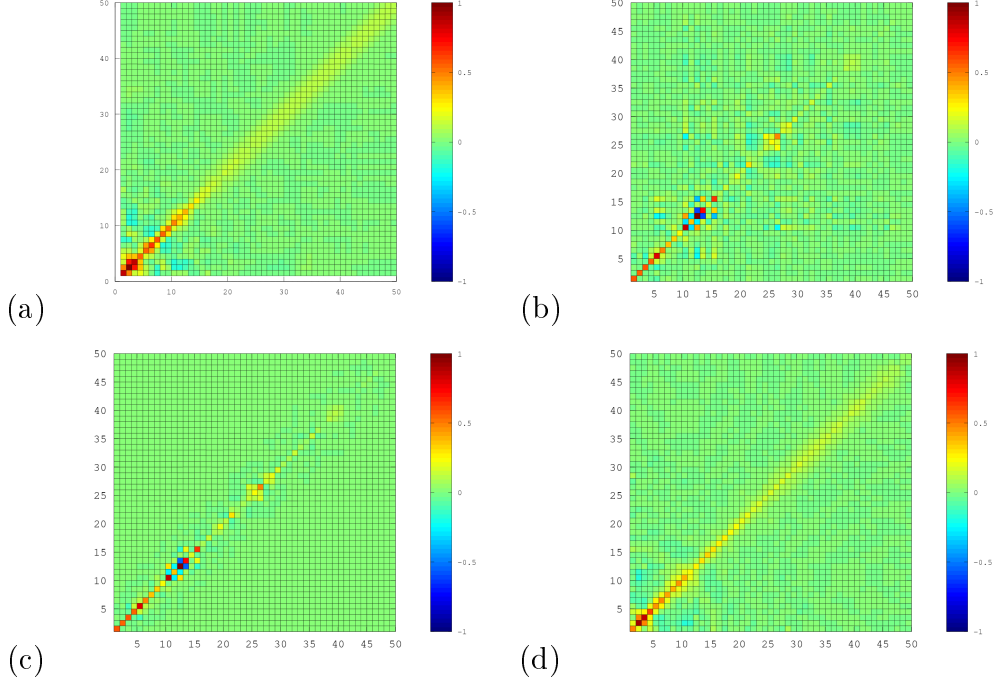


Figure 3.4: We display a  $B$ -matrix constructed from the experimental global LETKF of DWD in (a), which has been constructed by taking the standard stochastic estimator as in (3.3.9) and adding a small constant times the Gaussian covariance matrix displayed in Figure 3.3 (a). The transformed matrix  $\tilde{B}$  is displayed in Figure (b). We localize the matrix  $\tilde{B}$  by multiplication with the matrix shown in Figure 3.3 (a), the result is displayed in (c). Figure (d) shows the matrix  $B_{gl}$  defined as the backtransformed matrix from  $\tilde{B}_{gl}$ .

to zero where they were smaller than a threshold  $\epsilon > 0$ . The result  $\tilde{B}_{gl}$  for the Gaussian localization is displayed in Figure 3.3 (c).

4. Then, we transform  $\tilde{B}_{gl}$  back into the state space  $X$  by

$$B_{gl} := T^{-1} \tilde{B}_{gl} (T^*)^{-1}, \quad (3.5.4)$$

which we call the  $B$ -matrix under generalized localization.

We remark that  $B_{gl}$  is quite close to  $B$  due to the localized structure of  $\tilde{B}$  in transformed space  $\tilde{X}$ . We have carried out reconstructions using  $B$  and  $B_{gl}$  as follows.

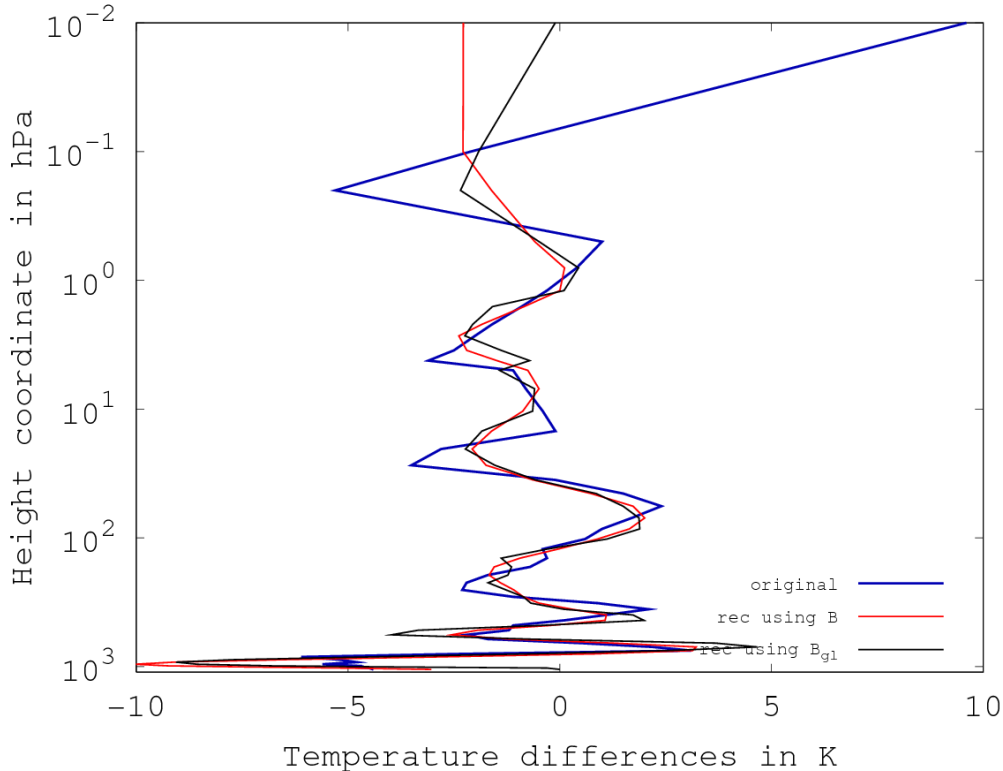


Figure 3.5: The figure shows the reconstruction of the difference between a background atmospheric temperature profile and some given temperature profile. The true difference is shown in thick blue, the reconstruction using the Gaussian  $B$ -matrix from Figure 3.3 (a) with regularization parameter  $\alpha = 0.0001$  is displayed in red, the reconstruction with  $B_{gl}$  in black.

- The blue curve shows some difference of two atmospheric profiles, some *true* profile and a background state  $\varphi^{(b)}$ .
- Then, we have calculated the corresponding brightness temperatures  $f = \mathcal{H}(\varphi)$  and  $f^{(b)} = \mathcal{H}(\varphi^{(b)})$  using RTTOV, i.e. the nonlinear operator  $\mathcal{H}$  defined above. We obtain the different  $\delta f := f - f^{(b)}$ .
- We now carry out a reconstruction to calculate  $\delta\varphi_\alpha$  by either full Tikhonov regularization (3.3.4) or by the transformed localized equations (3.4.54).

The results of our calculations are displayed in Figure 3.5, with a reasonably good approximation of the reconstruction by transformed localization (black) to the full

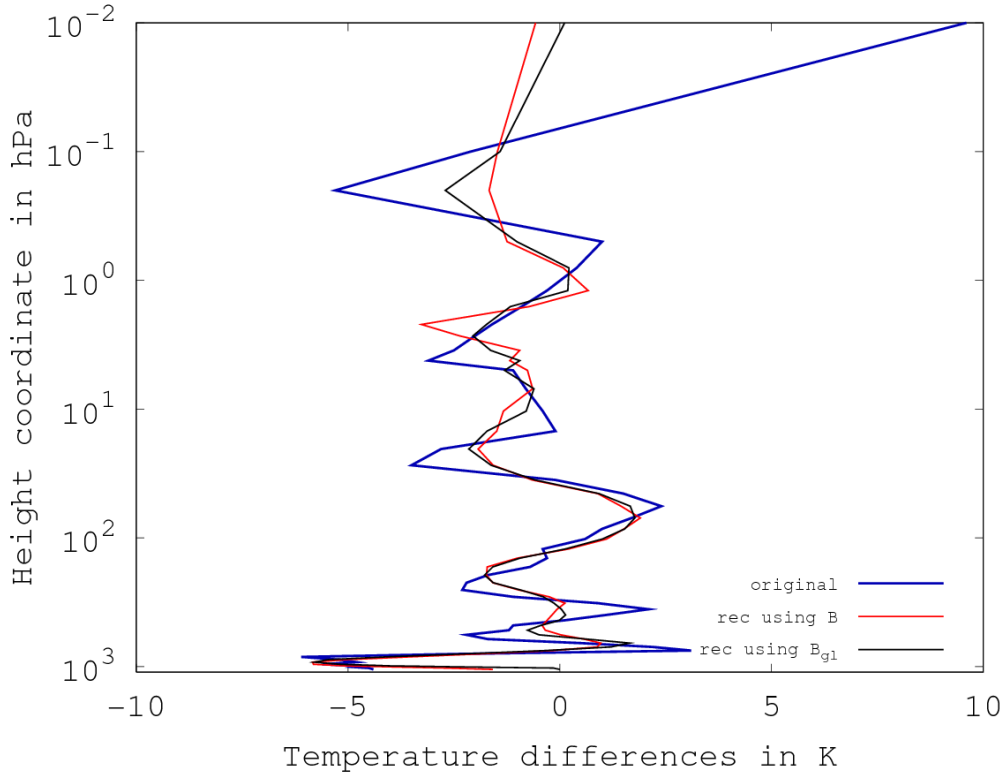


Figure 3.6: The figure shows the reconstruction of the difference between a background atmospheric temperature profile and some given temperature profile. The true difference is shown in thick blue, the reconstruction using the LETKF  $B$ -matrix displayed in Figure 3.4 with regularization parameter  $\alpha = 0.0001$  is displayed in red, the reconstruction with the corresponding  $B_{gl}$  in black.

regularized reconstruction without localization (red). We have shown that it is possible to transform the equation system into a form which is local both with respect to the covariance matrix  $\tilde{B}$  as well as the observation operator  $\tilde{H}$ .

### 3.6 Conclusions

We have developed a general concept of *transformed localization* and investigated its effect on the Ensemble Kalman Filter (EnKF) in its version described in [62]. In particular, as initial step we provide the mathematical analysis for an EnKF and its transformed version and show equivalence of both the analysis mean and the

analysis ensemble in Lemma 3.4.1 and Theorem 3.4.2.

Then, we have interpreted *localization* as a special case of a *projection method* and investigate the decomposition of a reconstruction problem into a sequence of projection subspaces in Lemma 3.4.3 and Theorem 3.4.4, where it is shown that the unlocalized projection methods in its unregularized and regularized versions provide equivalent solutions when the full space is decomposed by a sequence of subspaces in which the projected equations are solved.

We apply the method to an example where the transformed operator and the transformed covariance matrix  $\tilde{B}$  has a block structure, such that transformation effectively decomposes the problem into a sequence of subproblems with smaller dimension  $\tilde{n}$ . When this smaller dimension  $\tilde{n}$  is smaller than  $L$  with the number of linearly independent ensemble members  $L$  in the EnKF, we obtain a full and equivalent solution of the problem through localization, even if the full non-localized problem has dimension  $n \gg \tilde{n}$ . Of course, in general we cannot expect block structure by transformation, but strongly reduce the number of off-diagonal elements in the covariance matrix  $\tilde{B}$  and the transformed operator  $\tilde{H}$ , such that localization in the transformed space provides a reasonable approximation to the original operator and the EnKF achieves good analysis results in each analysis step.

In Section 3.5 we have applied the theory to a temperature retrieval problem for infrared radiance measurements from the IASI instrument. We have calculated the full regularized retrieval and the retrieval which comes from a transformed and localized version of the problem in Figures 3.5 and 3.6, where the second figure used the experimental global LETKF ensemble data assimilation system of Deutscher Wetterdienst DWD, which runs with 40 ensemble members. In this case, equivalent localization in the vertical is really needed to obtain algebraic equivalence to the full vertical reconstruction problem with 51 RTTOV-Levels in a vertical column. Theorem 3.4.5 applies, and the example shows that the the full solution with 51 degrees of freedom and approximate solution based on the localized transformed problem are quite comparable, with some stronger differences in selected locations.

Both theory and the selected example indicate that *transformed localization* is of high potential to be employed in a more realistic and cycled setup of the EnKF, which is beyond the scope of this more conceptual and analytical approach. To this end, the approach needs further development steps in subsequent work.

# Chapter 4

## On Multiscale Data Assimilation

*Multiscale methods* are very popular for modeling natural phenomena, they have attracted a lot of attention over recent years in the scientific literature, both for *direct* and *inverse* problems, compare for example [1, 2, 3, 4, 5, 6, 8, 22, 125, 142, 143] for multiscale methods in simulation and [13, 49, 51, 60, 134, 169] for their use in inverse problems.

For solving an inverse problem, the *singular value decomposition* (see [50, 68, 100, 129]) provides a traditional approach to study the different scales of the problem. Classical regularization theory can be viewed as a damping of higher modes of the inverse of the operator which maps the unknown states onto the observations.

In the area of *data assimilation*, different versions of the multi-scale approach has been tested over the past years, for example in [107, 114, 121, 122, 137, 138, 150, 151, 174, 176, 180]. For many systems in the geosciences, we have something like a natural low-scale approach by models with some course resolution, which are complemented by high-resolution models which can be run locally on nested regions. Over the past 15 years, *ensemble methods* have been highly successful in meteorological or geophysical data assimilation. Here, different scales are implicitly chosen by *localization techniques*, see for example [16, 17, 19, 29, 30, 31, 38, 89, 116, 117, 118, 119, 126, 145].

In this work, our goal is *first* a generic *theoretical understanding* of the role and effects which appear in multiscale approaches to *inverse* and *data assimilation* methods. *Second*, our goal is to analyse and test *iterative sequential multiscale*, with a numerical example from the field of integral equations as they appear in *atmospheric temperature retrieval from infrared radiances*. We will restrict our attention to *one inversion or data assimilation step*. In this case, the generic situation is described by an equation of the form

$$H\varphi = f \tag{4.0.1}$$

where  $\varphi \in X$  is in some state space  $X$  and  $f \in Y$  is the observation vector in some observation space  $Y$ . Here, we assume that  $X$  and  $Y$  are Hilbert spaces, usually for simplicity we restrict our attention to finite dimensional spaces, i.e. to the case where  $X = \mathbb{R}^n$  and  $Y = \mathbb{R}^m$  with  $n, m \in \mathbb{N}$ , where  $X$  and  $Y$  each are equipped with some norm. Most of our arguments, though, will carry over to the general infinite-dimensional case. We assume that some covariance matrix  $B \in \mathbb{R}^{n \times n}$  describes the knowledge about the relationships between the different components  $\varphi_j$ ,  $j = 1, \dots, n$  of  $\varphi \in X$ , i.e.

$$\bar{\varphi} := \mathbb{E} \{ \varphi \}, \quad B = \mathbb{E} \{ (\varphi - \bar{\varphi})(\varphi - \bar{\varphi})^T \}. \quad (4.0.2)$$

Let  $R \in \mathbb{R}^{m \times m}$  be the covariance matrix of the observation error on  $\mathbb{R}^m$ .

We will study the situation where the multiscale approach obtains its input from some given multiscale basis of  $X$ , which might be *orthogonal* or *non-orthogonal*. This includes *classical multiscale basis sets* which are obtained from generating radial basis functions, but it also includes a multiscale basis coming from two ensembles, one covering the low scale, the other the finer scale. We will study the generic approximation properties of such an approach, showing that a *sequential multiscale decomposition* leads to typical *approximation errors* when the basis is non-orthogonal. We then provide the analysis for the *iteration* of sequential multiscale, both for the case where the reconstruction is *regularized* as well as for the *unregularized* case.

For the *unregularized case* and *non-orthogonal case* we will show that in general the sequential multiscale approach cannot provide the same solution as if we would solve on all scales simultaneously. The error is given by the successive projections of the solution on the spaces perpendicular to the multiscale spaces under consideration. A generic two-dimensional example is given and visualized in Figure 4.1.

We then study *iterated sequential multiscale*, i.e. we repeat the sequential multiscale method starting with the result of each previous iteration. For the *unregularized* case we show *linear* convergence of these iterations, where the convergence speed is determined by  $\cos(\alpha)$  where  $\alpha$  is the smallest angle between the subspaces under consideration.

Inverse problems or data assimilation usually involves *regularization*, which causes further complications to the above convergence analysis. Carrying out the analysis several times – either for the sequential steps of the multiscale method or for the iteration of the sequence of multiscale steps – the role of the *regularizing* or *background term*, respectively, needs to be taken carefully into account.

We first provide a generic example of iterated regularization which shows that  $\ell$  steps of a sequential regularization with regularization parameter  $\beta$  are identical to one-step regularization with regularization parameter  $\alpha$ , when the regularization

parameters satisfy the relationship

$$\sqrt[\ell]{(1 - \alpha)} = 1 - \beta. \quad (4.0.3)$$

The general theory is worked out for the *Kalman filter*, proving that the Kalman gain matrix  $\tilde{K}$  for the iterations and the Kalman gain matrix  $K = BH'(R + HBH')^{-1}$  for the *one-step* analysis need to satisfy the equation

$$(I - \tilde{K}H)^\ell = I - KH, \quad (4.0.4)$$

where  $H$  is our *observation operator*  $H$  under consideration. The solvability of (4.0.4) is shown by studying the operator in a particular weighted Hilbert space, where  $KH$  becomes self-adjoint.

After describing our general setup in Section 4.1 we will start with the solution of the inverse problem in a subspace and its reformulation into a basis transformation of the covariance matrix and the states under consideration. This is carried out in Section 4.1.1.

The multiscale transform for the observation space is our topic in Section 4.1.2. The multiscale decomposition for the canonical transformed inverse equation based on the covariances in state and observation space is described in Section 4.1.3.

We then study the question of sequential scale splitting in Section 4.2. We first clarify that in general scale splitting does not lead to equivalence of the sequential approach to the full problem when non-orthogonal spaces are employed. For orthogonal decompositions the problem can be equivalently split into multiscale spaces. The problem of aliasing when data on different scales is employed is discussed in Section 4.2.2. Here, aliasing means that the high-frequency data may lead to large errors when the inversion in a low-frequency space is carried out.

We provide a numerical example in our Section 4.3, which demonstrates the feasibility of the multiscale inversion. We apply iterative sequential multiscale to an integral equation which is often seen as a generic example for atmospheric temperature or humidity retrieval from satellite radiances [129].

## 4.1 A Generic Multiscale Approach

The classical Tikhonov regularization for the solution of (4.0.1) is given by

$$\varphi_\alpha := (\alpha I + H^*H)^{-1}H^*f, \quad (4.1.1)$$

where  $\alpha > 0$  is known as *regularization parameter*. Here, in a functional analytic framework, we will employ the letter  $\varphi$  to denote the states, the observations will be denoted by  $f$ .



Either starting with Tikhonov regularization in  $X$  equipped with a norm weighted by  $B^{-1}$  or using a Bayesian approach, it is well known (see e.g. [37], [62] or [129]), that the regularized solution of (4.0.1) can be obtained by a minimization of

$$J(\varphi) := \|\varphi - \varphi^{(b)}\|_{B^{-1}}^2 + \|f - H\varphi\|_{R^{-1}}^2, \quad (4.1.2)$$

where the minimizer is given by

$$\varphi^{(a)} := \varphi^{(b)} + BH^T(R + HBH^T)^{-1}(f - H\varphi^{(b)}). \quad (4.1.3)$$

For ensemble methods, often the covariance matrix  $B$  is given by the standard covariance estimator

$$B = QQ^T \quad (4.1.4)$$

with  $Q$  defined by

$$Q = \frac{1}{\sqrt{L-1}}(\varphi^{(1)} - \mu, \dots, \varphi^{(L)} - \mu), \quad (4.1.5)$$

where  $\varphi^{(\xi)}$  denotes an ensemble of states under consideration and  $\mu$  is its mean

$$\mu := \frac{1}{L} \sum_{\xi=1}^L \varphi^{(\xi)}. \quad (4.1.6)$$

A key idea of a *multi-scale inversion method* or *data assimilation method* is to carry out an inversion on a lower scale first and then move to the next finer scale. Let

$$U = \tilde{U}_1 \oplus \tilde{U}_2 \oplus \tilde{U}_3 \oplus \dots \oplus \tilde{U}_M \quad (4.1.7)$$

be a multi-scale decomposition of  $U$  into a hierarchy of spaces defined by

$$U_j := \bigoplus_{\xi=1}^j \tilde{U}_\xi, \quad j = 1, 2, \dots, M. \quad (4.1.8)$$

We assume that the spaces are mutual linearly independent. Instead of solving the full minimization problem minimizing

$$J(\varphi) := \|f - H\varphi^{(b)} - H\varphi\|_{R^{-1}}^2, \quad \varphi \in U, \quad (4.1.9)$$

we solve problems with  $\tilde{\varphi} \in \tilde{U}_\xi$  successively. In each step  $\xi = 1, 2, 3, \dots, M$  we solve the equation

$$H\varphi = f^{(\xi)}, \quad \varphi \in \tilde{U}_\xi, \quad (4.1.10)$$

where we start with  $f^{(1)} := f - H\varphi^{(b)}$ . Starting with  $\varphi^{(0)} = \varphi^{(b)}$  for  $\xi = 1, 2, 3, \dots, M$  we denote the minimizer of

$$J_\xi(\varphi) := \alpha \|\varphi\|_{B^{-1}}^2 + \|H\varphi - f^{(\xi)}\|_{R^{-1}}^2, \quad \varphi \in \tilde{U}_\xi, \quad (4.1.11)$$

by  $\tilde{\varphi}^{(\xi)}$ , where

$$f^{(\xi)} := f^{(\xi-1)} - H\tilde{\varphi}^{(\xi-1)} \quad (4.1.12)$$

and  $\alpha > 0$  is our regularization parameter. We also allow  $\alpha = 0$ , in this case the problem is called *unregularized*. We carry out the loop of (4.1.11) - (4.1.12) for  $\xi = 1, \dots, M$ . This leads to successive elements  $\tilde{\varphi}^{(1)}, \tilde{\varphi}^{(2)}, \dots, \tilde{\varphi}^{(M)}$  which are in the spaces  $\tilde{U}_1, \dots, \tilde{U}_M$ . The corresponding solutions up to level  $j$  are given by

$$\varphi^{(j)} := \varphi^{(b)} + \sum_{\xi=1}^j \tilde{\varphi}^{(\xi)}, \quad j = 1, 2, 3, \dots, M, \quad (4.1.13)$$

which is an element of  $U_j$  defined in (4.1.8).

### 4.1.1 A Multiscale Step in State Space

Here, we study one reconstruction step of (4.1.11) - (4.1.12). Let the space under consideration be denoted by  $U \subset X$ . We assume that a set of linearly independent *ensemble members* or a *multiscale basis*  $\varphi^{(1)}, \dots, \varphi^{(L)}$  of  $U$  is given. We drop unnecessary indices for simplicity.

**THEOREM 4.1.1.** *Let  $\Phi = (\varphi^{(1)}, \dots, \varphi^{(L)})$  be a basis of  $U \subset \mathbb{R}^n$  for  $L \leq n$ , i.e. it is a matrix with columns being linearly independent vectors spanning  $U$ . Then, the minimization of  $J(\varphi)$  given in (4.1.2) on the affine subspace  $\varphi^{(b)} + U$  defined by*

$$\varphi = \varphi^{(b)} + \sum_{\ell=1}^L \alpha_\ell \varphi^{(\ell)} \quad (4.1.14)$$

for  $\alpha = (\alpha_1, \dots, \alpha_L)$ , is carried out by

$$\alpha^{(a)} = \check{B}\check{H}^T(R + \check{H}\check{B}\check{H}^T)^{-1}\check{f}, \quad (4.1.15)$$

or alternatively by

$$\alpha^{(a)} = (\alpha I + \check{B}\check{H}^T R^{-1} \check{H})^{-1} \check{B}\check{H}^T R^{-1} \check{f}, \quad (4.1.16)$$

and

$$\varphi^{(a)} := \varphi^{(b)} + \Phi \alpha^{(a)}, \quad (4.1.17)$$

using the abbreviations

$$\check{f} = f - H\varphi^{(b)}, \quad \check{B}^{-1} = \Phi^T B^{-1} \Phi \quad \text{and} \quad \check{H} = H\Phi. \quad (4.1.18)$$

*Remark.* We note that when the dimension of  $U$  is  $L$  and the dimension of  $Y$  is  $m$ , the matrix  $\check{B}^{-1}$  is an  $L \times L$  matrix and  $\check{H}$  is an  $m \times L$  matrix. Thus, the inversion of (4.1.16) is taking place in the  $L$ -dimensional space only, where we might have  $L \ll m$ . In the case where  $L > m$  we will prefer to use (4.1.15).

*Proof.* With  $\varphi = \varphi^{(b)} + \Phi\alpha$  for  $\alpha \in \mathbb{R}^L$  we rewrite (4.1.2) into

$$J(\alpha) := \|\Phi\alpha\|_{B^{-1}}^2 + \|(f - H\varphi^{(b)}) - H\Phi\alpha\|_{R^{-1}}^2, \quad \alpha \in \mathbb{R}^L. \quad (4.1.19)$$

We remark that if  $B$  is invertible and  $\Phi$  is of maximal rank, then  $\Phi^T B^{-1} \Phi$  is invertible in  $\mathbb{R}^{L \times L}$ , since

$$\Phi^T B^{-1} \Phi \alpha = 0 \Rightarrow \langle \Phi \alpha, B^{-1} \Phi \alpha \rangle = 0 \Rightarrow \Phi \alpha = 0 \Rightarrow \alpha = 0, \quad (4.1.20)$$

such that  $\Phi^T B^{-1} \Phi$  is injective and thus surjective on  $\mathbb{R}^L$ . Then, with the definitions of  $\check{H}$ ,  $\check{B}$  and  $\check{f}$  from above we transform (4.1.19) into

$$J(\alpha) := \|\alpha\|_{\check{B}^{-1}}^2 + \|\check{f} - \check{H}\alpha\|_{R^{-1}}^2, \quad \alpha \in \mathbb{R}^L. \quad (4.1.21)$$

Minimization of (4.1.21) with respect to  $\alpha$  leads to (4.1.15), such that  $\alpha^{(a)}$  is the minimizer of (4.1.19) and (4.1.17) is a minimizer of (4.1.2) over  $\varphi^{(b)} + U$ . For the different forms of the update operator we refer to the recent book [129], Chapter 5.  $\square$

In the case where  $U$  is the whole space and  $\Phi$  is linearly independent, we obtain

$$\check{B} = \Phi^{-1} B (\Phi^T)^{-1} \quad (4.1.22)$$

with the inverse  $\Phi^{-1}$  of  $\Phi$  and the inverse  $(\Phi^T)^{-1}$  of the transpose  $n \times n$ -matrix  $\Phi$ . In the case where  $\Phi$  is an orthonormal basis of  $X$ , the transform of the  $B$  matrix is obtained by taking the inverse of  $\check{B}^{-1}$ , which according to (4.1.18) leads to

$$\begin{aligned} \check{B} &= \Phi^{-1} B (\Phi^T)^{-1} \\ &= \Phi^T B \Phi. \end{aligned} \quad (4.1.23)$$

In the orthonormal case and when  $B$  is given by the covariance estimator (4.1.4), we obtain

$$\check{B} = \Phi^T B \Phi = \Phi^T Q Q^T \Phi = (\Phi^T Q) (\Phi^T Q)^T, \quad (4.1.24)$$

which is the estimated covariance for the *multi-scale coefficients*. If we are in the non-orthonormal case, we obtain

$$\check{B} = \Phi^{-1} B (\Phi^T)^{-1} = (\Phi^{-1} Q) (\Phi^{-1} Q)^T \quad (4.1.25)$$

based on  $(\Phi^T)^{-1} = (\Phi^{-1})^T$ , which holds for invertible matrices on  $\mathbb{R}^n$ .

We need to extend our transformation to the case where  $\Phi$  is a matrix of vectors of dimension  $n \times L$  and where  $B \in \mathbb{R}^{n \times n}$  is rank deficient. Here, we can replace  $B^{-1}$  by its *pseudo-inverse*

$$B^\dagger := \lim_{\alpha \rightarrow 0} (\alpha I + B^T B)^{-1} B^T, \quad (4.1.26)$$

compare [129], Section 3.2. If  $\text{span}\{\Phi\} \cap N(B) = \{0\}$  the matrix  $\Phi^T B^\dagger \Phi$  is invertible and (4.1.18) with  $B^{-1}$  replaced by  $B^\dagger$  is well-defined and we obtain

$$\check{B} = \left( \Phi^T B^\dagger \Phi \right)^{-1} = \Phi^\dagger (B^\dagger)^\dagger (\Phi^T)^\dagger = \Phi^\dagger B (\Phi^T)^\dagger. \quad (4.1.27)$$

If we search for the solution in ensemble space, i.e. if  $U$  is the space defined by the columns of  $Q$  defined in (4.1.5), then (4.1.25) leads to

$$\check{B} = Q^\dagger Q Q^T (Q^T)^\dagger = I. \quad (4.1.28)$$

## 4.1.2 A Multiscale Step in Observation Space

Next, let us study the sequential multiscale approach based on a multiscale decomposition of the observation space. Let  $\psi_1, \dots, \psi_m$  be a *basis* in observation space  $Y$ , i.e. we have  $\psi_j$  are linearly independent for  $j, \xi = 1, \dots, m$  and span  $Y$ , and define

$$\Psi := (\psi_1, \dots, \psi_m). \quad (4.1.29)$$

Then, the coefficients  $\gamma$  of a vector  $f \in Y$  with respect to  $\psi_1, \dots, \psi_m$  is given by  $\gamma = \Psi^{-1} f$ . We can now transform the equation (4.1.3) with respect to the basis  $\psi_1, \dots, \psi_m$ . The solution of (4.1.3) is obtained by first solving

$$(R + HBH^T)z = f - H\varphi^{(b)}, \quad (4.1.30)$$

and then calculating the *increments*

$$\varphi^{(a)} - \varphi^{(b)} = BH^T z \quad (4.1.31)$$

in the second step, where the solution  $z$  in observation space is mapped back into state space  $X$ . Equation (4.1.30) is multiplied by  $\Psi$  from the left and we insert  $I = \Psi\Psi^{-1}$  on the left of the variable  $z$  to obtain

$$\left( \Psi^{-1}(R + HBH^T)\Psi \right) \Psi^{-1} z = \Psi^{-1}(f - H\varphi^{(b)}) \quad (4.1.32)$$

which with

$$\begin{aligned} A &:= \Psi^{-1}(R + HBH^T)\Psi \\ &= R + \Psi^{-1}HBH^T\Psi, \end{aligned} \quad (4.1.33)$$

(where we used  $R = rI$  and  $\Psi^{-1}\Psi = I$ ) and  $\gamma = \Psi^{-1}z$  and  $\gamma^o := \Psi^{-1}(f - H\varphi^b)$  takes the form

$$A\gamma = \gamma^o. \quad (4.1.34)$$

Here  $\gamma^o$  are the coordinates of the observation increment  $f - H\varphi^b$  with respect to the basis  $\psi_1, \dots, \psi_m$  and  $\gamma$  denotes the coefficient vector of  $z = \Psi\gamma$  with respect to  $\psi_1, \dots, \psi_m$ . The full inversion now takes the form

$$\begin{aligned} \varphi^{(a)} &= \varphi^{(b)} + BH^T\Psi A^{-1}\Psi^{-1}(f - H\varphi^{(b)}) \\ &= \varphi^{(b)} + B\tilde{H}^T A^{-1}\gamma^o, \end{aligned} \quad (4.1.35)$$

where

$$\tilde{H} := \Psi^T H \quad (4.1.36)$$

is the operator  $H$  when evaluated in coordinates of  $\psi_1, \dots, \psi_m$ .

Now consider the case where  $\psi_1, \dots, \psi_m$  is a basis of a subspace  $W$  of  $Y$ . We replace  $\Psi^{-1}$  by the *pseudo-inverse*  $\Psi^\dagger$  and obtain a projection of  $f - H\varphi^{(b)}$  onto  $W$  by

$$\gamma^o = \Psi^\dagger(f - H\varphi^{(b)}). \quad (4.1.37)$$

Equation (4.1.30) is transformed into a projected version of this equation with respect to the coefficients of  $\Psi$  in  $W$ . The solution is mapped back into the state space by  $BH^T\psi$ .

We first note that if  $\Psi$  diagonalizes the self-adjoint matrix  $HBH^T$ , then the calculation step (4.1.32) completely decouples and becomes the solution of a diagonal matrix equation (4.1.34), which is very fast.

If  $\Psi$  is a *multiscale basis* with scales ordered from low scales to high scales, we can construct low-scale approximations to the solution of the data assimilation problem by solving the problem in a subspace  $W$  of the first  $L$  coefficients  $\gamma_L^o$  of  $\gamma^o$  to obtain low-scale approximation to  $A$  by calculating

$$A_L := rI_L + \Psi_L^\dagger HBH^T \Psi_L \quad (4.1.38)$$

with the first  $L$  columns of  $\Psi$  collected into the matrix  $\Psi_L$ . Then, we can continue with a sequential method as described in (4.1.11)-(4.1.12).

Further *multiscale approximations* to the solution of (4.1.32) or (4.1.34), respectively, are given by any approximation  $\tilde{A}$  to  $A$  in (4.1.33). This could be setting small elements of the transformed matrix

$$\tilde{B} := \Psi^{-1}HBH^T\Psi \quad (4.1.39)$$

to zero or higher-mode off-diagonal elements of  $\tilde{B}$ . Approximating  $\tilde{B}$  by a simpler matrix in coefficient space can lead to a high speed-up of the calculation once the

transformation is carried. Further, it can lead to a reduction of noise when the  $B$ -matrix is estimated. In the case where  $B = QQ^T$ , we obtain

$$\tilde{B} = \psi^{-1}HQQ^TH^T\psi = (\psi^{-1}HQ)(\psi^THQ)^T, \quad (4.1.40)$$

In the case of an orthonormal matrix  $\Psi$  we have  $\Psi^{-1} = \Psi^T$ , i.e. we can calculate the covariance in transformed observation space based on the coefficients  $\gamma_Q := \psi^T Q_Y$  of the *ensemble of observation equivalents*  $Q_Y := HQ$ . For general non-orthonormal multiscale decompositions this is not the case.

### 4.1.3 The Role of SVD for the Transformed Equation

For completeness here we also review another well-known standard type of a multi-scale decomposition of a transformed problem based on the *standard transformation* of the inversion step. By using

$$\tilde{H} = R^{-\frac{1}{2}}HB^{\frac{1}{2}}, \quad \tilde{f} = R^{-\frac{1}{2}}f \quad (4.1.41)$$

we can transform the increment

$$\delta\varphi = \varphi^{(a)} - \varphi^{(b)} = BH^T(R + HBH^T)^{-1}(y - H\varphi^{(b)}) \quad (4.1.42)$$

into

$$\delta\varphi = B^{\frac{1}{2}}\tilde{H}^T(I + \tilde{H}\tilde{H}^T)^{-1}(\tilde{f} - R^{-\frac{1}{2}}H\varphi^{(b)}). \quad (4.1.43)$$

Let  $\tilde{H} = V_1SV_2^T$  with orthonormal matrices  $V_1$  and  $V_2$  be the singular value decomposition of  $\tilde{H}$ . The diagonal matrix  $S$  contains the singular values  $\mu_n$  of  $\tilde{H}$ . Then, we obtain

$$\begin{aligned} \delta\varphi &= B^{\frac{1}{2}}V_2SV_1^T(I + V_1S^2V_1^T)^{-1}(\tilde{f} - R^{-\frac{1}{2}}H\varphi^{(b)}) \\ &= B^{\frac{1}{2}}V_2S(I + S^2)^{-1}V_1^T(\tilde{f} - R^{-\frac{1}{2}}H\varphi^{(b)}) \end{aligned} \quad (4.1.44)$$

In the language of inverse problems, the term  $S(I + S^2)^{-1}$  is nothing else than the spectral version of the Tikhonov regularization as presented in [129], Theorem 3.1.8, which with the singular values  $\mu_n$  is a multiplication operator with spectral factors given by  $\mu_n/(\alpha + \mu_n^2)$ , here with  $\alpha = 1$ . We obtain a special *multiscale decomposition* of our analysis step by the basis given by  $V_1$ , where then the scales are split and after solving the problem for each scale separately we get back to a full solution by the basis given by  $V_2$ .

In general, the singular value decomposition of  $\tilde{H}$  is not feasible for large problems. Also, taking the square root of  $B$  is a tough problem for very high-dimensional

problems as they naturally appear in the geosciences. If  $B$  is approximated by an ensemble approach as in (4.1.4), we can apply *localization techniques* to obtain an efficient localized calculation of the square root of  $B$ . There is a wide literature on localization today, see for example some recent work [129, 145, 173]. Here, we will restrict our attention to the analysis for a generic setup where some non-orthogonal basis is considered as generic input.

## 4.2 Non-Equivalence and Iterations of the Sequential Multiscale Approach

In this part we first study the non-equivalence of sequential multiscale to the full inversion step, for both the unregularized and the regularized case. We show that an iteration of the sequential steps will converge under appropriate assumptions. We will interpret our analysis as a type of *aliasing* in the framework of well-posed or ill-posed inversion.

### 4.2.1 On the Equivalence and Non-Equivalence of Sequential Multiscale

In general, the sequential multiscale approach does not lead to the correct solution of the full inverse problem under consideration.

**THEOREM 4.2.1.** *Consider the sequential multiscale method defined in (4.1.11)-(4.1.12). If the spaces  $\tilde{U}_1, \dots, \tilde{U}_M$  are not orthogonal to each other with respect to  $\langle \cdot, \cdot \rangle_{H^*R^{-1}H}$ , then in general the sequential solution  $\varphi^{(M)}$  defined in (4.1.13) will not coincide with the solution to the full inverse problem given by (4.1.1).*

*Proof.* Let us study the generic case  $M = 2$  and first consider the unregularized solution and an injective operator  $H$  first, such that

$$\langle \varphi, \psi \rangle_{H^*R^{-1}H} := \langle H\varphi, R^{-1}H\psi \rangle, \quad \varphi, \psi \in X \quad (4.2.1)$$

defines a scalar product on the state space  $X$ . For the case  $M = 2$  we assume that the subspaces  $\tilde{U}_1$  and  $\tilde{U}_2$  satisfy  $X = \tilde{U}_1 \oplus \tilde{U}_2$ . In this case, if  $\tilde{U}_1 \perp \tilde{U}_2$ , then  $\tilde{U}_1 = \tilde{U}_2^\perp$  and  $\tilde{U}_2 = \tilde{U}_1^\perp$ . Let  $P_1$  be the orthogonal projector of  $X$  onto  $\tilde{U}_1$  and  $P_2$  the corresponding projector onto  $\tilde{U}_2$ . Then,  $I - P_2$  is the orthogonal projector onto  $\tilde{U}_2^\perp = \tilde{U}_1$  and  $I - P_1$  is the orthogonal projector onto  $\tilde{U}_1^\perp = \tilde{U}_2$ .

LEMMA 4.2.2. *If  $\tilde{U}_1$  and  $\tilde{U}_2$  are not perpendicular with respect to  $\langle \cdot, \cdot \rangle_{H^*R^{-1}H}$ , then the projector  $(I - P_2)(I - P_1)$  is a non-zero mapping, i.e. there is  $z \in X$  such that*

$$(I - P_2)(I - P_1)z \neq 0. \quad (4.2.2)$$

*Proof.* We remark that if  $\tilde{U}_1$  is not perpendicular to  $\tilde{U}_2$ , then we have  $\tilde{U}_1^\perp$  is not perpendicular to  $\tilde{U}_2^\perp$  just since  $\tilde{U}_1 = \tilde{U}_2^\perp$  and  $\tilde{U}_2 = \tilde{U}_1^\perp$ . This means that there are elements in  $\tilde{U}_1^\perp$  whose projection onto  $\tilde{U}_2^\perp$  is non-zero, which proves (4.2.2).  $\square$

We now study the difference between the full unregularized solution of the problem (4.1.2) and the sequential process given by (4.1.11)-(4.1.12). We note that if the unregularized sequential solutions are not identical to the full unregularized solution, this cannot be the case for the regularized solutions, since for large covariance matrices  $B$  the regularization term is very small and the regularized solution will be close to the unregularized solutions.

For the case of true data we have  $f = H\varphi^*$  we write the first step as

$$\tilde{\varphi}^{(1)} = \arg \min_{\varphi \in \tilde{U}_1} \|\varphi - \varphi^*\|_{H^*R^{-1}H}^2, \quad (4.2.3)$$

the second step as

$$\tilde{\varphi}^{(2)} = \arg \min_{\varphi \in \tilde{U}_2} \|\varphi - (\varphi^* - \tilde{\varphi}^{(1)})\|_{H^*R^{-1}H}^2 \quad (4.2.4)$$

and the full solution as

$$\varphi^{(tot)} = \arg \min_{\varphi \in X} \|\varphi - \varphi^*\|_{H^*R^{-1}H}^2 \quad (4.2.5)$$

The solutions can be understood as orthogonal projections, in particular we have

$$\tilde{\varphi}^{(1)} = P_1\varphi^*, \quad (4.2.6)$$

$$\begin{aligned} \tilde{\varphi}^{(2)} &= P_2(\varphi^* - \tilde{\varphi}^{(1)}) \\ &= P_2\varphi^* - P_2P_1\varphi^*. \end{aligned} \quad (4.2.7)$$

This leads to

$$\varphi^{(2)} = (P_1 + P_2 - P_2P_1)\varphi^*. \quad (4.2.8)$$

The difference between the full solution and the sequential solution in the case of injective  $H$  is given by

$$\begin{aligned} \varphi^{(tot)} - \varphi^{(2)} &= \varphi^* - (P_1 + P_2 - P_2P_1)\varphi^* \\ &= \left\{ I - (P_1 + P_2 - P_2P_1) \right\} \varphi^*. \end{aligned} \quad (4.2.9)$$



Finally, we note that

$$I = P_1 + (I - P_1) \quad (4.2.10)$$

$$= P_1 + [P_2 + (I - P_2)](I - P_1) \quad (4.2.11)$$

$$= P_1 + P_2 - P_2P_1 + (I - P_2)(I - P_1), \quad (4.2.12)$$

such that

$$\varphi^{(tot)} - \varphi^{(2)} = (I - P_2)(I - P_1)\varphi^*. \quad (4.2.13)$$

By Lemma 4.2.2 the right-hand side of the above term is non-zero in general, and thus we have shown that the sequential method does not coincide with the full minimization procedure under the above conditions. We note that equation (4.2.13) can also be used to estimate the error between the two methods.

In the case where  $H$  is not injective, the minimizer with minimal norm is the pseudo-inverse with respect to the scalar product (4.2.1). In this case the above arguments will still work when we replace  $\tilde{U}_\xi$  by  $\tilde{U}_\xi \cap N(H^*R^{-1}H)^\perp$ , on which the operator is injective and all arguments apply.

Finally, the case *with regularization* has the same property that the sequential solution and the full solution will not coincide in general. This is clear since for small regularization parameters the regularized solution is close to the unregularized solution (if it exists), and thus they must be different in the two cases under consideration.  $\square$

We demonstrate the situation by a two-dimensional unregularized example to visualize the situation. For  $M = 2$  and  $\varphi^{(b)} = 0$  we show that in general the successive solution is not the same as a solution which is calculated in one big step. Assume that  $R = I$  and  $H = I$ . In this case the solution of (4.1.10) in  $U_1$  corresponds to an orthogonal projection  $P^{(1)}f$  of  $f$  onto  $U_1$ . Then,  $f^{(2)} = f^{(1)} - P^{(1)}f$  is perpendicular to  $U^{(1)}$ . In the second step, the solution of (4.1.10) corresponds to an orthogonal projection of  $f^{(2)}$  onto  $U^{(2)}$ . We visualize this situation in Figure 4.1. In particular, the result for the full minimization would be the measured true solution  $\varphi = f$ . But in the successive case, we end up with the gray solution  $\varphi^{(2)}$  in the second step.

The above result shows that in general we need to be careful with the successive solution of the minimization problem (4.1.10). We will see next that when we guarantee orthogonality of the subspaces, we are fine.

We now come to a generic result which shows the validity and convergence of the multi-scale minimization problem.

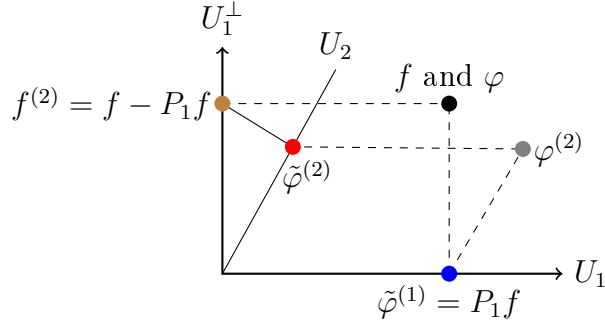


Figure 4.1: The successive solution of the equations leads first to the point  $P_1 f$  shown in blue. The reduced data for step two  $f^{(2)}$  is shown as a brown dot. This is projected onto  $U_2$  in step two, see the red dot visualizing  $\tilde{\varphi}^{(2)}$ , leading to the solution  $\varphi^{(2)}$  shown by the gray point. However, the true solution  $\varphi$  would be the point  $f$  itself in this example. Thus, we see that in the general case the successive solution of the minimization problem cannot lead to correct results.

**THEOREM 4.2.3.** *Assume that  $H$  is linear and injective on  $U$  and assume that the spaces  $\tilde{U}_1, \dots, \tilde{U}_M$  are mutual orthogonal with respect to the scalar product defined by*

$$\langle \varphi, \psi \rangle_{H^T R^{-1} H} := \langle \varphi, H^T R^{-1} H \psi \rangle, \quad \varphi, \psi \in X. \quad (4.2.14)$$

*Then, for the unregularized approach the full minimizer  $\varphi$  of (4.1.9) and the iterative minimizer  $\varphi^{(M)}$  coincide.*

*Proof.* We have worked out the error estimate for the unregularized approach for the case  $M = 2$  in the equations (4.2.1) to (4.2.13). When we have orthogonality of  $\tilde{U}_1$  and  $\tilde{U}_2$ , we obtain  $(I - P_2)(I - P_1) = 0$ , such that the difference between the full solution and the successive solution vanishes. This applies inductively to all subspaces  $\tilde{U}_1, \tilde{U}_2, \dots$ , such that we obtain full equivalence.  $\square$

We also note at this point that the regularized successive solutions will not be identical even if we have orthogonality in the spaces. Differences come from the role of the regularizing term, since in each regularization step the regularized solution is between the background and the unregularized solution, but does not use the same norm but a weight which involves the covariance matrix  $B$ .

## 4.2.2 General Form of Aliasing Problems

One important phenomenon of signal processing is *aliasing*. It means that if you sample a signal with high frequency components and you try to reconstruct its low frequency parts from a low number of sampling points, the high frequency parts of your signal will be transformed into spurious lower frequency signal. It is possible to avoid aliasing either by using more sampling points even when you reconstruct the low modes only, or you need to run a *filter* on your signal before sampling and then only sample the filtered signal and transmit it.

The aliasing problem also appears when you try to solve multiscale inverse problems by using a reduced number of measurements coming from a signal which contains higher modes. In general, the aliasing problem is a null-space phenomenon on the signal or state reconstruction operator.

Let  $H : X \rightarrow Y$  be our observation operator in the finite dimensional case, i.e.  $X = \mathbb{R}^n$ ,  $Y = \mathbb{R}^m$ , and let  $\hat{H} : X \rightarrow \hat{Y}$  be a selection of the rows of  $H$ , i.e.  $\hat{Y} = \mathbb{R}^{\hat{m}}$  with  $\hat{m} < m$ . Let us assume that  $m = n$  and  $H$  is injective. Then,  $\hat{H}$  will not be injective, and there is some nullspace  $U := N(\hat{H})$  of  $\hat{H}$  in  $X$ . The solution of the reduced data equation

$$\hat{H}\varphi = \hat{f}, \quad \varphi \in V, \quad (4.2.15)$$

in some reduced subspace  $V \subset X$  does not need to coincide with the orthogonal or any other particular projection  $Px_0$  of the true solution  $x_0$  of

$$H\varphi = f, \quad \varphi \in X \quad (4.2.16)$$

onto the space  $V$ . If the nullspace  $U$  of  $\hat{H}$  does not coincide with  $V^\perp$  or the space along which  $P$  projects, there are many solutions  $\varphi_0 + \Delta\varphi$  of (4.2.15) which are projected onto different solutions  $\varphi \in V$  by  $P$ . If, however, we have  $U = V^\perp$  or more general  $U$  is the space along which  $P$  projects, then the projection of any solution  $\varphi_0 + \Delta\varphi$  with  $\Delta\varphi \in U$  will lead to the projection  $P\varphi = P\varphi_0$  in  $V$ , i.e. we obtain the correct projection in the low-dimensional subspace  $V$ .

In the framework of (4.2.15) filtering of the data corresponds to the transformation of  $\hat{f}$  into data which are in the image space of  $V$  under  $\hat{H}$ , i.e. the filter needs to construct  $\hat{f}^{filter} := \hat{H}(P\varphi_0)$  as data for which the correct state estimate  $P\varphi_0 \in V$  is obtained when a solution to (4.2.15) in  $V$  is calculated. As argued in Lemma 4.2.2, the use of the data  $\hat{f}$  can lead to arbitrarily large errors when solving the unregularized reduced equation.

## 4.2.3 Convergence for Iterated Sequential Multiscale

We now define an *iterated* approach to *sequential multiscale*. The basic idea is to carry out the multiscale approach several times to improve on the solution in the

case where orthogonality of the spaces is not given.

DEFINITION 4.2.4 (Iterative sequential multiscale). *We carry out the sequential multiscale approach iteratively by the following steps.*

1. *We start with a first application of the sequential multiscale as given by equations (4.1.11) - (4.1.12). For  $\ell = 1$  we define  $\varphi^{(\xi, \ell)} := \varphi^{(\xi)}$  and  $f^{(\xi, \ell)} := f^{(\xi)}$  for  $\xi = 1, \dots, M$ .*
2. *In each iteration step for  $\ell = 1, 2, 3, \dots$  we define new data  $f^{(1, \ell+1)}$  by*

$$f^{(1, \ell+1)} := f - H\varphi^{(M, \ell)} \quad (4.2.17)$$

and

$$\varphi^{(b, \ell+1)} = \varphi^{(M, \ell)}. \quad (4.2.18)$$

3. *Then, the sequential multiscale method is iterated until some convergence criterion is satisfied or stopping rule fires.*

The goal of the following paragraphs is to carry out the convergence analysis of the iterative sequential multiscale algorithm for the *unregularized* case and to show convergence when true data  $f = H\varphi^*$  with an injective operator  $H$  are given. The regularized case will be discussed in the following Section 4.2.4

LEMMA 4.2.5. *The error vector for the  $M$ -th iteration of the sequential multiscale is given by*

$$\varphi^* - \varphi^{(M)} = (I - P_M)(I - P_{M-1}) \dots (I - P_1)\varphi^*. \quad (4.2.19)$$

*Proof.* We carry out a proof by induction. Assume that we have already shown

$$\varphi^* - \varphi^{(M-1)} = (I - P_{M-1}) \dots (I - P_1)\varphi^*. \quad (4.2.20)$$

Then, we obtain

$$\begin{aligned} \varphi^* - \varphi^{(M)} &= \varphi^* - (\varphi^{(M-1)} + \tilde{\varphi}^{(M)}) \\ &= \varphi^* - \varphi^{(M-1)} - \tilde{\varphi}^{(M)} \\ &= \varphi^* - \varphi^{(M-1)} - P_M(\varphi^* - \varphi^{(M-1)}) \\ &= (I - P_M)(\varphi^* - \varphi^{(M-1)}) \\ &= (I - P_M)(I - P_{M-1}) \dots (I - P_1)\varphi^* \end{aligned} \quad (4.2.21)$$

and the proof is complete.  $\square$

We can now estimate the error for the iterated sequential multiscale as follows.

THEOREM 4.2.6. *The error vector for the iterated sequential multiscale in the unregularized case can be calculated by*

$$\varphi^* - \varphi^{(M,\ell)} = \left[ (I - P_M)(I - P_{M-1}) \dots (I - P_1) \right]^\ell \varphi^* \quad (4.2.22)$$

for  $\ell = 1, 2, 3, \dots$

*Proof.* For  $\ell = 1$  the result has been shown in Lemma 4.2.5. If we now carry out the iterative procedure, by definition (4.2.17) we start with the difference between the true data and the last outcome of the sequential multiscale, which is

$$\begin{aligned} f^{(1,\ell+1)} &= f - H\varphi^{(M,\ell)} \\ &= H\left(\varphi^* - \varphi^{(M,\ell)}\right). \end{aligned} \quad (4.2.23)$$

This means that by applying Lemma 4.2.5 inductively, in each application of the multiscale method we increase  $\ell$  by one in the error estimate (4.2.22).  $\square$

In each step of the method, the error is reduced by a factor  $\cos(\alpha)$ , where  $\alpha$  is the angle between the subspaces under consideration. This is linear convergence for the iterative sequential multiscale method, as a consequence of its understanding as a sequential projection method. We summarize this immediate consequence in the following corollary.

COROLLARY 4.2.7 (Convergence Iterated Sequential Multiscale). *The sequential multiscale given by (4.2.17)-(4.2.18) in the unregularized case shows linear convergence with factor  $c = \cos(\alpha)$  where  $\alpha$  is a lower estimate for the angles between the subspaces  $\tilde{U}_\xi$ , i.e. the error  $e_{\xi,\ell} = \|\varphi^{\xi,\ell} - \varphi^*\|_{H^*R^{-1}H}$  in the  $\ell$ -th iteration and after step  $\xi$  of this iteration is estimated by*

$$e_{\xi,\ell} \leq \left( \cos(\alpha) \right)^{\xi+M(\ell-1)} \|\varphi^{(b)} - \varphi^*\|_{H^*R^{-1}H} \quad (4.2.24)$$

#### 4.2.4 Regularized Sequential Multiscale

We have seen that we obtain convergence in the unregularized case when we iterate the sequential multiscale, with a linear error estimate which is determined by the angle between the multi-scale subspaces.

The *regularized* case is much more involved, since the background term of the regularizer needs to be taken care of in a special way. To see the challenge let us first study the generic but simplified situation where we start some iteration with

$$\gamma_0 := \varphi^{(b)} \quad (4.2.25)$$

and then iterate by

$$\gamma_\ell = \gamma_{\ell-1} + \beta(\varphi^* - \gamma_{\ell-1}) \quad (4.2.26)$$

for  $\ell = 1, 2, 3, \dots$  with some parameter  $\beta \in (0, 1)$ . This corresponds to the increment (4.1.3) where perfect data  $\varphi^*$  are given and where we have an observation operator  $H = I$  and covariance matrices  $B = \frac{\beta}{1-\beta}I$  and  $R = I$ . It is the case where we do not have different multi-scale spaces, but carry out an iteration in one single space only. In this case, the iterations can be calculated exactly as follows.

LEMMA 4.2.8. *The iteration (4.2.25)-(4.2.26) leads to*

$$\gamma_\ell = \gamma_0 + \left\{ \sum_{\xi=1}^{\ell} \binom{\ell}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \quad (4.2.27)$$

$$= \gamma_0 + \left\{ 1 - (1 - \beta)^\ell \right\} (\varphi^* - \gamma_0) \quad (4.2.28)$$

$$= \varphi^* - (1 - \beta)^\ell (\varphi^* - \gamma_0) \quad (4.2.29)$$

for  $\ell = 1, 2, 3, \dots$

*Proof.* We use induction for the proof. For  $\ell = 1$  it is

$$\begin{aligned} \gamma_1 &= \gamma_0 + \left\{ \sum_{\xi=1}^1 \binom{1}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \\ &= \gamma_0 + \beta(\varphi^* - \gamma_0), \end{aligned} \quad (4.2.30)$$

which coincides with our claim. Suppose the formula is true for  $\ell = n$

$$\gamma_n = \gamma_0 + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \quad (4.2.31)$$

Then we obtain for  $\ell = n + 1$

$$\begin{aligned} \gamma_{n+1} &= \gamma_n + \beta(\varphi^* - \gamma_n) \\ &= \gamma_0 + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \\ &\quad + \beta \left( \varphi^* - \left( \gamma_0 + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \right) \right) \\ &= \gamma_0 + \beta(\varphi^* - \gamma_0) + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \end{aligned}$$

$$\begin{aligned}
& -\beta \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \\
= & \gamma_0 + \beta(\varphi^* - \gamma_0) + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \\
& + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+2} \beta^{\xi+1} \right\} (\varphi^* - \gamma_0) \\
= & \gamma_0 + \beta(\varphi^* - \gamma_0) + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \\
& + \left\{ \sum_{\xi=2}^{n+1} \binom{n}{\xi-1} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \\
= & \gamma_0 + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) + \beta(\varphi^* - \gamma_0) \\
& + \left\{ \sum_{\xi=2}^n \binom{n}{\xi-1} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) + \binom{n}{n} (-1)^{n+2} \beta^{n+1} (\varphi^* - \gamma_0) \\
= & \gamma_0 + \left\{ \sum_{\xi=1}^n \binom{n}{\xi} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) \\
& + \left\{ \sum_{\xi=1}^n \binom{n}{\xi-1} (-1)^{\xi+1} \beta^\xi \right\} (\varphi^* - \gamma_0) + \binom{n}{n} (-1)^{n+2} \beta^{n+1} (\varphi^* - \gamma_0) \\
= & \gamma_0 + \sum_{\xi=1}^n \left\{ \binom{n}{\xi} + \binom{n}{\xi-1} \right\} (-1)^{\xi+1} \beta^\xi (\varphi^* - \gamma_0) + \binom{n}{n} (-1)^{n+2} \beta^{n+1} (\varphi^* - \gamma_0) \\
= & \gamma_0 + \sum_{\xi=1}^n \left\{ \binom{n+1}{\xi} \right\} (-1)^{\xi+1} \beta^\xi (\varphi^* - \gamma_0) + \binom{n+1}{n+1} (-1)^{n+2} \beta^{n+1} (\varphi^* - \gamma_0) \\
= & \gamma_0 + \sum_{\xi=1}^{n+1} \left\{ \binom{n+1}{\xi} \right\} (-1)^{\xi+1} \beta^\xi (\varphi^* - \gamma_0), \tag{4.2.32}
\end{aligned}$$

such that the formula is true for  $\ell = n + 1$ . Hence, it is true for all  $n \in \mathbb{N}$ . By using the binomial expansion we have

$$(1 - X)^\ell = \left\{ \sum_{\xi=0}^{\ell} \binom{\ell}{\xi} (-1)^\xi (X)^\xi \right\}$$

$$(1 - X)^\ell = 1 + \left\{ \sum_{\xi=1}^{\ell} \binom{\ell}{\xi} (-1)^\xi (X)^\xi \right\} \quad (4.2.33)$$

So the equation (4.2.31) becomes

$$\begin{aligned} \gamma_\ell &= \gamma_0 + \left\{ 1 - (1 - \beta)^\ell \right\} (\varphi^* - \gamma_0) \\ &= \varphi^* - (1 - \beta)^\ell (\varphi^* - \gamma_0) \end{aligned} \quad (4.2.34)$$

for  $\ell = 1, 2, 3, \dots$  and the proof is complete.  $\square$

Let us compare the  $\ell$ -th iteration in the form (4.2.28) with the one-step regularized increment given by

$$\varphi^{(a)} = \gamma_0 + \alpha(\varphi^* - \gamma_0) \quad (4.2.35)$$

with replacing  $\beta$  by the regularization parameter  $\alpha$ . To achieve equivalence of a  $\ell$ -step method with the one-step method, we need to choose

$$\alpha = 1 - (1 - \beta)^\ell \Leftrightarrow \sqrt[\ell]{1 - \alpha} = 1 - \beta. \quad (4.2.36)$$

In this simplified case, when we use iterations for the regularized case, to achieve an appropriate balance between the background and the data, we need to adapt the relative weight using the  $\ell$ -th square root of the regularization parameter  $1 - \alpha$ .

To complete our analysis, we now study the general but still complete-scale case, where the iteration is given by

$$\gamma_0 = \varphi^{(b)} \quad (4.2.37)$$

$$\gamma_\ell = \gamma_{\ell-1} + \tilde{K}(H\varphi^* - H\gamma_{\ell-1}) \quad (4.2.38)$$

with some Kalman gain operator  $\tilde{K}$  which we need to determine appropriately. With the analysis given in the proof of Lemma 4.2.8 we now obtain the following result.

**THEOREM 4.2.9.** *The iterations (4.2.37)-(4.2.38) lead to*

$$\gamma_\ell = \gamma_0 + \left( 1 - (1 - \tilde{K}H)^\ell \right) (\varphi^* - \gamma_0) \quad (4.2.39)$$

for  $\ell = 1, 2, 3, \dots$



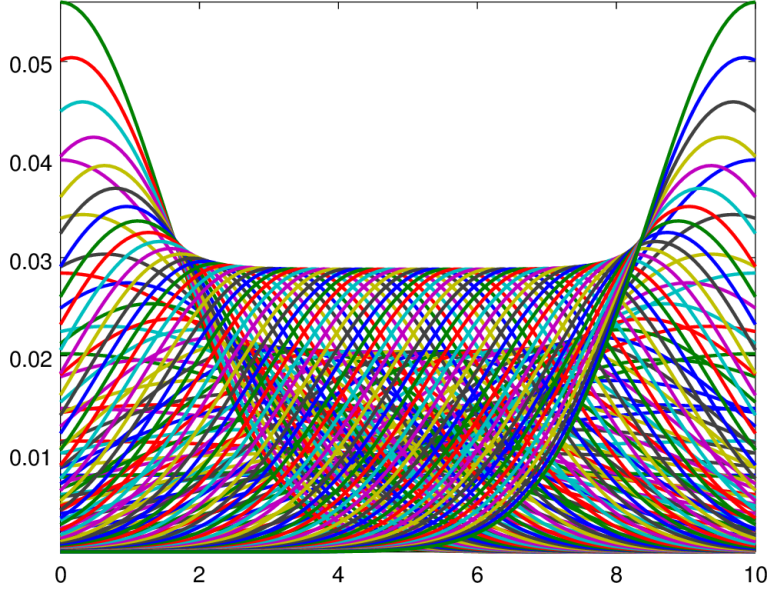


Figure 4.2: We show the basis functions used for the demonstration of the iterative multiscale method on the interval  $[a, b] = [0, 10]$ . The functions are generated by scaling and translating a Gaussian basis wavelet. All functions are scaled to have the  $L^2$ -norm  $\|\varphi\|_{L^2([a,b])} = 1$ , leading to some amplification at the edges. In this image, we display  $n = 2^7 = 128$  basis functions.

*Proof.* We basically use a modification of the proof of Lemma 4.2.8, which was the simplified case in which an observation operator  $H = I$  and covariance matrices  $B = \frac{\beta}{1-\beta}I$  and  $R = I$  were studied. Here, we need to be careful with the operator nature of  $\tilde{K}H$ . We have

$$(I - \tilde{K}H)^\ell = \left\{ \sum_{r=1}^{\ell} \binom{\ell}{r} (-1)^r (\tilde{K}H)^r \right\} \quad (4.2.40)$$

since there is no difficulty with the commutative law of powers of an operator  $A = \tilde{K}H$ . So, induction is carried out following the steps of Lemma 4.2.8 to obtain

$$\gamma_\ell = \gamma_0 + \sum_{\xi=1}^{n+1} \left\{ \binom{n+1}{\xi} (-1)^{\xi+1} (\tilde{K}H)^\xi \right\} (\varphi^* - \gamma_0) \quad (4.2.41)$$

Now again, by using the binomial theorem representation, we derive

$$\gamma_\ell = \gamma_0 + \left\{ I - (I - \tilde{K}H)^\ell \right\} (\varphi^* - \gamma_0) \quad (4.2.42)$$

$$= \varphi^* - (1 - \tilde{K}H)^\ell(\varphi^* - \gamma_0) \quad (4.2.43)$$

and the proof is complete.  $\square$

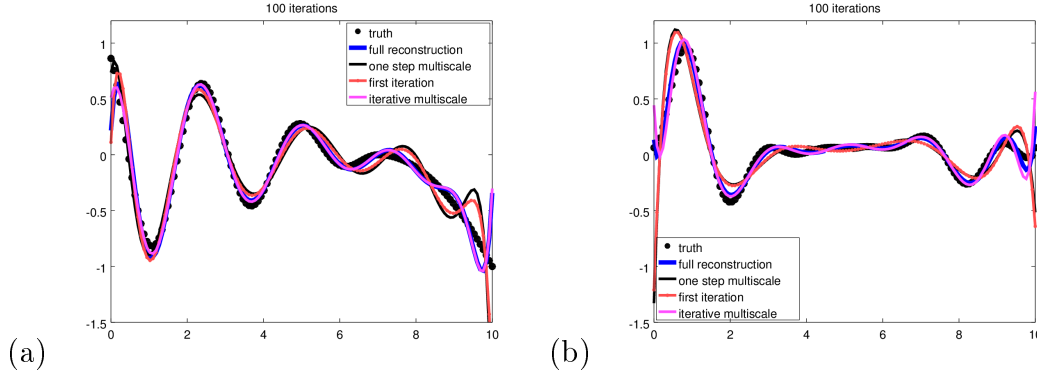


Figure 4.3: Two reconstruction results of the *multiscale method* and the *iterative multiscale method* after  $\ell = 100$  iterations is displayed here. The blue curve shows the regularized reconstruction when the full basis is employed. The classical multiscale method leads to the black curve. The red curve shows the first step of the iteration, which is similar to the black one, but with a modified regularization parameter for the two steps. Then, iteration leads to the convergence as shown by the magenta line, where we stopped after  $\ell$  iterations, we chose  $\alpha = 10^{-9}$ . For a convergence study see Figure 4.4.

To achieve equivalence of an  $\ell$ -step reconstruction with the one-step reconstruction

$$\varphi^{(a)} = \varphi^{(b)} + KH(\varphi^* - \varphi^{(b)}), \quad (4.2.44)$$

we need to determine  $\tilde{K}$  such that

$$(I - \tilde{K}H)^\ell = I - KH \quad (4.2.45)$$

as the counterpart of the postulate (4.2.36). It corresponds to taking the  $\ell$ -th root of the operator  $I - KH = I - BH^*(R + HBH^*)^{-1}H$ , with  $K$  defined by the second term of (4.1.3). For symmetric positive operators, the square root is well-defined. Here, we obtain this property by using the weighted scalar products

$$\langle \varphi, \tilde{\varphi} \rangle_{B^{-1}} := \langle \varphi, B^{-1}\tilde{\varphi} \rangle_{L^2} \quad (4.2.46)$$

and

$$\langle \psi, \tilde{\psi} \rangle_{R^{-1}} := \langle \psi, R^{-1}\tilde{\psi} \rangle_{L^2} \quad (4.2.47)$$

for functions  $\varphi, \tilde{\varphi} \in X$  and  $\psi, \tilde{\psi} \in Y$ . Using equation (5.2.11) as in (5.2.14) of [129] with  $H^{*,w}$  being the adjoint with respect to the scalar products (4.2.46) and (4.2.47), we see that we have  $BH^*R^{-1} = H^{*,w}$  and thus

$$\begin{aligned} I - KH &= I - BH^*(R + HBH^*)^{-1}H \\ &= I - H^{*,w}(I + HH^{*,w})^{-1}H. \end{aligned} \quad (4.2.48)$$

Clearly, we have

$$\left( H^{*,w}(I + HH^{*,w})^{-1}H \right)^{*,w} = H^{*,w}(I + HH^{*,w})^{-1}H, \quad (4.2.49)$$

i.e. we have a self-adjoint positive operator. Thus, the  $l$ -th square root of  $I - KH$  exists and can be calculated using the singular value decomposition of the operator in the space with weighted scalar product. We now collect the above results into the following theorem.

**COROLLARY 4.2.10.** *Assume that for  $\ell \in \mathbb{N}$  we determine  $\tilde{K}$  such that (4.2.45) is satisfied. Then, after  $\ell$  steps the iteration (4.2.37) - (4.2.38) is identical to the regularized solution (4.2.44).*

For practical applications of the *sequential multiscale method* taking the  $\ell$ -th root of the operator  $K$  is not feasible. Usually, the reason why we consider multiscale is the complexity of the task under consideration. We suggest to employ the solution algorithms as they are available on each of the scales under consideration, adapt the regularization parameter according to (4.2.36) and then iterate the multiscale algorithms. Numerical tests confirm the feasibility and convergence of this approach in our final section.

### 4.3 Numerical Examples

This last section serves to test the above method of *iterative sequential multiscale* for solving integral equations of the form

$$f(x) = \int_a^b k(x, y)\varphi(y) ds(y), \quad x \in [a, b] \quad (4.3.1)$$

where  $b \geq a$  are real numbers,  $k$  is a continuous kernel

$$k : [a, b] \times [a, b] \rightarrow \mathbb{R} \quad (4.3.2)$$

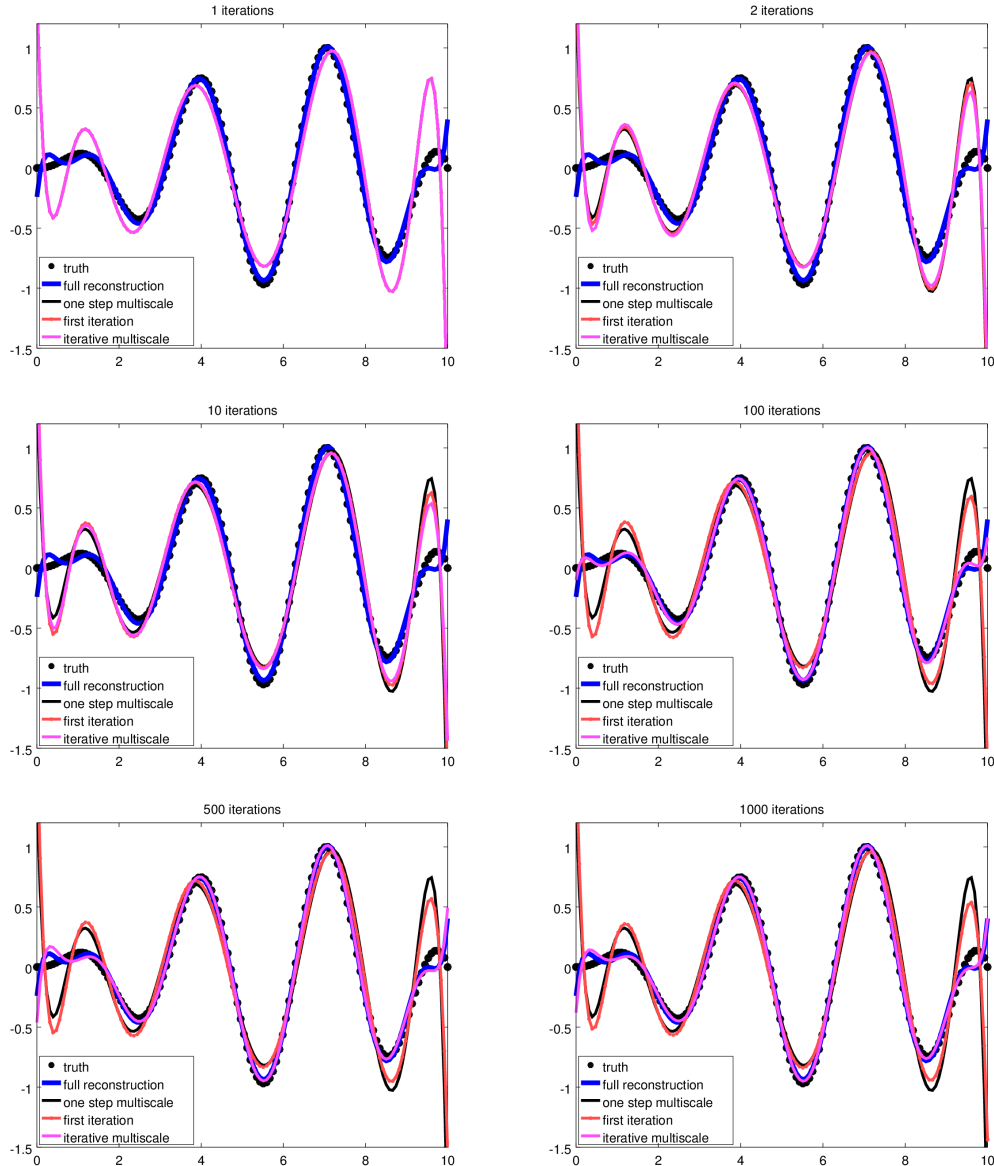


Figure 4.4: The results of the *multiscale method* and the *iterative multiscale method* after  $\ell = 1, 2, 10, 100, 500, 1000$  iterations is displayed here. The blue curve shows the regularized reconstruction of the truth in a space with  $\ell = 128$  degrees of freedom, when the basis with  $\ell = 32$  elements is employed. The classical multiscale method with dimension  $\ell = 16$  of the course space and  $\ell = 16$  for the fine space leads to the black curve. The red curve shows the first step of the iteration, which is similar to the black one, but with a modified regularization parameter for the two steps. The iterative multiscale solution converges to the blue solution, i.e. to the solution when all its basis functions are used in one step.

and we work in  $L^2(a, b)$ . For numerical testing we choose the example of a Gaussian kernel

$$k(x, y) := e^{-\sigma|x-y|^2}, \quad x, y \in [a, b], \quad (4.3.3)$$

with  $\sigma = 3$ , which serves as a generic example of atmospheric temperature or humidity retrieval from satellite radiances, compare [129, 126]. Here, we choose some example densities

$$\varphi_1(y) := \sin(c_6 \cdot (y - y_1)) \cdot |y - y_1| \cdot (y < y_1) + c_7 \cdot (y - y_1)^2 \cdot (y > y_1) \quad (4.3.4)$$

$$\varphi_2(y) := c_5 \cdot ((y - y_0)^{c_3} \cdot \sin(c_4 * y) \cdot \sin(y) \cdot (b - y) \cdot (b/2)^{-1/2} + 1) \quad (4.3.5)$$

$$\varphi_3(y) := y^{c_1} \cdot \sin(y) \cdot \cos(y) \cdot (b - y)^{c_2} \quad (4.3.6)$$

for  $y \in [a, b]$  with  $a = 0$ ,  $b = 10$ ,  $c_1 = 1.5$ ,  $c_2 = 0.8$ ,  $c_3 = 1.9$ ,  $c_4 = 2.1$ ,  $c_5 = 2$ ,  $c_6 = 2.4$ ,  $c_7 = -0.8$ ,  $y_0 = 5$  and  $y_1 = 7$ . The functions are not in the span of the basis functions. Examples with  $\varphi_1$  and  $\varphi_2$  are shown in Figure 4.3 (a) and (b), the reconstruction of  $\varphi_3$  is displayed in Figure 4.4. We calculate the simulated measured data  $f(x)$  by numerical quadrature with trapezoidal or rectangular rule. The observation operator  $H$  is given by the integral (4.3.1), i.e. we have

$$(H\varphi)(x) := \int_a^b k(x, y)\varphi(y) ds(y), \quad x \in [a, b] \quad (4.3.7)$$

with kernel  $k$  given by (4.3.2) and (4.3.3). For the numerical simulation of  $H$  we have used a collocation scheme with a regular grid on  $[a, b]$  with  $n$  points, defined by

$$x_k := a + k \cdot \frac{b - a}{n - 1}, \quad k = 0, \dots, n - 1 \quad (4.3.8)$$

Since we want to investigate convergence, we have used *true data* which were generated by numerical collocation. We remark that the problem is *exponentially ill-posed* such that even when the dimension  $n$  of our space is just modestly high an algorithm without regularization does not lead to any reasonable solution.

There are many possibilities to define and test a multiscale basis on the interval  $[a, b]$ . Here, we carry out tests with a *Gaussian type radial basis function*, i.e. for  $N \in \mathbb{N}$  and  $n = 2^N$  we define

$$\varphi_{\ell,k}(x) := c_{\ell,k} e^{-\rho_\ell |x - x_{\ell,k}|^2}, \quad x \in [a, b] \quad (4.3.9)$$

for  $\ell = 1, \dots, N$  and  $k = 1, \dots, 2^{\ell-1}$  where

$$x_{\ell,k} := a + \frac{k}{2^{\ell-1}}(b - a), \quad \rho_\ell = 2^{\ell-1}\rho_0 \quad (4.3.10)$$

and  $\varphi_{0,0} \equiv c_{0,0}$ . The normalized basis functions (4.3.9) are displayed in Figure 4.2.

For multiscale splitting we chose the lowest  $L$  modes for the space  $U_1 = \tilde{U}_1$  and defined  $\tilde{U}_2$  to be the space spanned by the basis functions  $L+1, \dots, \tilde{n}$ , where  $\tilde{n}$  is the maximal number of basis functions under consideration. For the iterations we show two cases with  $\ell = 100$  in Figure 4.3. Figure 4.3 displays the results for  $n = 128$ ,  $L = 16$ ,  $\tilde{n} = 32$  and a regularization parameter  $\alpha = 10^{-9}$ . We show the truth in black dots, the full reconstruction without any multiscale splitting in blue, the one-step regularized multiscale in black, the first iteration of the iterative regularized multiscale method in red and the result of the iterative multiscale method in pink. As expected, the one-step multiscale has much larger errors than the full solution, arising from the non-orthogonality of the multi-scale spaces. The convergence is studied in a third example in Figure 4.4, where we display the result of iterations with  $\ell = 1, 2, 10, 100, 500, 1000$ . The iterative multiscale solution (pink line) confirms the convergence to the solution in the full subspace (blue line) according to Theorem 4.2.9 and Corollary 4.2.10.

## 4.4 Conclusions

The broad idea of a multiscale approach is to decompose your problem into different scales or levels and to use these decompositions either for constructing appropriate approximations or to solve smaller problems on each of these levels, leading to increased stability or increased efficiency.

The idea of sequential multiscale as studied in our Section 4.1 is to first solve the problem in a large-scale subspace and then successively move to finer scale spaces. The equivalence or non-equivalence of this approach to solving the full-scale problem at once has been studied in Section 4.2.1. When the spaces under consideration are orthogonal, we show equivalence for the unregularized problem. For the regularized problems equivalence can only be achieved when the regularization parameters are modified appropriately.

In general the sequential multiscale approach is not equivalent to a full solution, but we have shown in Theorem 4.2.6 and Corollary 4.2.7 that under appropriate assumptions we obtain convergence of an *iterative* sequential multiscale version of the method.

For the *regularized case* we have developed a strategy to appropriately adapt the regularization when an iterative approach is taken. We have first studied a simplified situation in Lemma 4.2.8, for which the calculation of the modified regularization parameter could be carried out explicitly. Then, the *general operator-based case* is treated, with a core equivalence result in Theorem 4.2.9.

Finally, we demonstrate the validity of the iterative sequential multiscale approach by testing the method on integral equations as they appears for atmospheric temperature retrieval from infrared satellite radiances and for image denoising, here for a simplified case in one dimension. The examples show the practical validity of the results and confirm the non-equivalence and convergence theory.

The analysis and methods of this work provide basic insight into the convergence behavior of multiscale approaches to inversion and data assimilation. Here, the static case has been treated, where iteration takes place at one given point in time with fixed measurements. For data assimilation, where iteration is carried out naturally by cycled systems (compare [129]), we are interested in the extension to the case where a dynamical model is applied between different iteration steps. This is of high interest to many researchers and practitioners, the use of multiscale approaches in a cycled environment will be part of our future work.

# CURRICULUM VITAE/LEBENS LAUF

## Personal Data

Name: Aamir Nadeem  
Date of Birth : 1st day of May, 1977  
Place of Birth: Toba Tek Singh, Pakistan  
Citizenship: Pakistani  
Marital Status: Married  
Parents: Mohammad Boota  
Najabaat Begum  
Spouse Name: Ayesha Aamir

## Secondary Education

1991-1993: Govt. Abbasi High School, Rawalpindi  
1993-1995: Government Asghar Mall College Rawalpindi

## Higher Education

1995-1997: B.Sc. in Mathematics and Physics  
from University of the Punjab, Lahore  
1999-2001: M.Sc. in Mathematics  
from University of the Punjab, Lahore  
2006-2008: M.Phil in Mathematics  
from Quaid-i-Azam University, Islamabad

## Occupational Career

2001-2005: Lecturer, Department of Mathematics,  
Pace College Rawalpindi  
2005-2010 : Lecturer, Department of Mathematics  
Federal Urdu University Of Arts, Science and Technology, Islamabad  
2010 tillday: Assistant professor, Department of Mathematics  
Federal Urdu University Of Arts, Science and Technology, Islamabad



# Bibliography

- [1] In Timothy J. Barth, Tony Chan, and Robert Haimes, editors, *Multiscale and Multiresolution Methods: Theory and Applications*, volume 20 of *Lecture Notes in Computational Science and Engineering*, pages X, 394. Springer-Verlag Berlin Heidelberg, 2002.
- [2] In Horst J. Neugebauer and Clemens Simmer, editors, *Dynamics of Multiscale Earth Systems*, volume 97 of *Lecture Notes in Earth Sciences*, pages XIV, 362. Springer-Verlag Berlin Heidelberg, 2003.
- [3] In Alexander Mielke, editor, *Analysis, Modeling and Simulation of Multiscale Problems*, pages XVIII,697. Springer Berlin Heidelberg, 2006.
- [4] In William W. Hager, Shu-Jen Huang, Panos M. Pardalos, and Oleg A. Prokopyev, editors, *Multiscale Optimization Methods and Applications*, volume 82 of *Nonconvex Optimization and Its Applications*, pages XVII, 407. Springer US, 2006.
- [5] In Ronald DeVore and Angela Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation: Dedicated to Wolfgang Dahmen on the Occasion of his 60th Birthday*, pages XXIV,660. Springer Berlin Heidelberg, 2009.
- [6] In Björn Engquist, Olof Runborg, and Per Lötsted, editors, *Multiscale Modeling and Simulation in Science*, volume 66 of *Lecture Notes in Computational Science and Engineering*, pages XIV,320. Springer Berlin Heidelberg, 2009.
- [7] *Data Assimilation: Making Sense of Observations*. Springer-Verlag Berlin Heidelberg, 2010.
- [8] In Björn Engquist, Olof Runborg, and Yen-Hsi R Tsai, editors, *Numerical Analysis of Multiscale Computations: Proceedings of a Winter Workshop at the Banff International Research Station 2009*, volume 82 of *Lecture Notes in Computational Science and Engineering*, pages X,430. Springer-Verlag Berlin Heidelberg, 2012.

- [9] R Acar and C R Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse problems*, 10(6):1217, 1994.
- [10] A.C.Lorenc. A global three-dimensional multivariate statistical interpolation scheme. *Monthly Weather Review*, 109(4):701–721, 1981.
- [11] A.C.Lorenc. The potential of the ensemble kalman filter for nwp - a comparison with 4d-var. *Quarterly Journal of the Royal Meteorological Society*, 129(595):3183–3203, 2003.
- [12] John Aldrich. Correlations genuine and spurious in pearson and yule. *Statistical science*, pages 364–376, 1995.
- [13] Habib Ammari and Hyeonbae Kang. *Inverse Problems, Multi-scale Analysis and Effective Medium Theory: Workshop in Seoul, Inverse Problems, Multi-scale Analysis, and Homogenization, June 22-24, 2005, Seoul National University, Seoul, Korea*, volume 408. American Mathematical Soc., 2006.
- [14] Brain D O Anderson and John B Moore. *Optimal Filtering*. Dover Books on Electrical Engineering Series. Dover Publications, Incorporated, 2012.
- [15] Jeffery L Anderson. An ensemble adjustment Kalman filter for data assimilation. *Monthly weather review*, 129(12), 2001.
- [16] Jeffery L Anderson. Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D:Nonlinear Phenomena*, 230:99–111, 2007.
- [17] Jeffrey L Anderson. An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus A*, 59(2):210–224, 2007.
- [18] Jeffrey L Anderson and Stephen L Anderson. A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127(12):2741–2758, 1999.
- [19] J.L. Anderson. Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus A*, 61(1):72–83, 2009.
- [20] A Andrews. A square root formulation of the Kalman covariance equations. *AIAA Journal*, 6(6):1165–1166., June 1968.

- [21] Hartmut H Aumann, Moustafa T Chahine, Catherine Gautier, Mitchell D Goldberg, Eugenia Kalnay, Larry M McMillin, Hank Revercomb, Philip W Rosenkranz, Fellow IEEE, William L Smith, David H Staelin, Fellow IEEE, L Larrabee Strow, and Joel Susskind. AIRS/AMSU/HSB on the aqua mission: design, science objectives, data products, and processing systems. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(2):253–264, Feb 2003.
- [22] Jacek Banasiak, Vincenzo Capasso, Mirosław Lachowicz, and Jacek Miekisz. In Vincenzo Capasso and Mirosław Lachowicz, editors, *Multiscale Problems in the Life Sciences: From Microscopic to Macroscopic*, volume 1940 of *C.I.M.E. Foundation Subseries*, pages XII, 330. Springer-Verlag Berlin Heidelberg, 2008.
- [23] R.J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*. Manchester Physics Series. Wiley, 1989.
- [24] S. Barnett and R.G. Cameron. *Introduction to Mathematical Control Theory*. Oxford applied mathematics and computing science series. Clarendon Press, 1985.
- [25] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1):183–202, 2009.
- [26] J F Bellantoni and K W Dodge. A square root formulation of the Kalman-schmidt filter. *AIAA Journal*, 5(7):1309–1314., June 1967.
- [27] T. Bick, C. Simmer, S. Trömel, K. Wapler, H.J. Hendricks Franssen, K. Stephan, U. Blahak, C. Schraff, H. Reich, Y. Zeng, and R. Potthast. Assimilation of 3d radar reflectivities with an ensemble Kalman filter on the convective scale. *Quarterly Journal of the Royal Meteorological Society*, 142(696):1490–1504, 2016.
- [28] Craig H Bishop, Brian J Etherton, and Sharanya J Majumdar. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly weather review*, 129(3):420–436, 2001.
- [29] Craig H Bishop and Daniel Hodyss. Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 133(629):2029–2044, 2007.
- [30] Craig H Bishop and Daniel Hodyss. Ensemble covariances adaptively localized with eco-rap. Part 1: tests on simple error models. *Tellus A*, 61(1):84–96, 2009.

- [31] Craig H Bishop and Daniel Hodyss. Ensemble covariances adaptively localized with eco-rap. Part 2: a strategy for the atmosphere. *Tellus A*, 61(1):97–111, 2009.
- [32] Denis Blumstein, Gilles Chalon, Thierry Carlier, Christian Buil, Philippe Hebert, Thierry Maciaszek, G Ponce, Thierry Phulpin, Bernard Tournier, Denis Simeoni, Patrick Astruc, Alain Clauss, Gohkahn Kayal, and R Jegou. IASI instrument: technical overview and measured performances. volume 5543, pages 196–207, 2004.
- [33] Niels Bormann, Alan Geer, Tomas Wilhelmsson, ECMWF Research Department, and European Centre for Medium Range Weather Forecasts. *Operational Implementation of RTTOV-10 in the IFS*. ECMWF technical memorandum. European Centre for Medium-Range Weather Forecasts, 2011.
- [34] Mark Buehner and Martin Charron. Spectral and spatial localization of background-error correlations for data assimilation. *Quarterly journal of the Royal Meteorological Society*, 133(624):615–630, 2007.
- [35] Mark Buehner, P L Houtekamer, Cecilien Charette, Herschel L Mitchell, and Bin He. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiment. *Monthly Weather Review*, 138(5):1550–1566, 2010.
- [36] Mark Buehner, P L Houtekamer, Cecilien Charette, Herschel L Mitchell, and Bin He. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observation. *Monthly Weather Review*, 138(5):1567–1586, 2010.
- [37] Martin Burger, Hendrik Dirks, and Jahn Mueller. Inverse problems in imaging. In Mike Cullen, Melina A Freitag, Stefan Kindermann, and Robert Scheichl, editors, *Large Scale Inverse Problems*, volume 13 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter, Berlin, 2013.
- [38] William F Campbell, Craig H Bishop, and Daniel Hodyss. Vertical covariance localization for satellite radiances in ensemble Kalman filters. *Monthly Weather Review*, 138(1):282–290, 2010.
- [39] Emmanuel J Candes, Michael B Wakin, and Stephen P Boyd. Enhancing sparsity by reweighted l1 minimization. *Journal of Fourier analysis and applications*, 14(5-6):877–905, 2008.

- [40] Stephen E Cohn, Arlindo Da Silva, Jing Guoand, Meta Sienkiewicz, and David Lamich. Assessing the effects of data selection with the dao physical-space statistical analysis system. *Monthly Weather Review*, 126(11):2913–2926, 1998.
- [41] A D Collard. Selection of IASI channels for use in numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 133(629):1977–1991, 2007.
- [42] A D Collard and A P McNally. The assimilation of infrared atmospheric sounding interferometer radiances at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 135(641):1044–1058, 2009.
- [43] David Colton and Rainer Kress. *Integral equation methods in scattering theory*. Pure and applied mathematics. Wiley, 1983.
- [44] David Colton and Rainer Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Applied Mathematical Sciences. Springer, 2012.
- [45] Philippe Courtier. Dual formulation of four-dimensional variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, 123(544):2449–2461, 1997.
- [46] Roger Daley. *Atmospheric Data Analysis*. Cambridge Atmospheric and Space Science Series. Cambridge University Press, 1993.
- [47] J. Dennis and R. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics, 1996.
- [48] Martin Ehrendorfer. A review of issues in ensemble-based Kalman filtering. *Meteorologische Zeitschrift*, 16(6):795–818, 2007.
- [49] E.Miller and A. Willsky. A multiscale approach to solving one dimensional inverse problems. In *Time-Frequency and Time-Scale Analysis, 1992., Proceedings of the IEEE-SP International Symposium*, pages 129–132, Oct 1992.
- [50] Heinz W Engl, Martin Hankle, and Andreas Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications. Springer Netherlands, 2000.
- [51] Bjorn Engquist and Christina Frederick. Nonuniform sampling and multiscale computation. *Multiscale Modeling & Simulation*, 12(4):1890–1901, 2014.

- [52] Geir Evensen. Sampling strategies and square root analysis schemes for the enkf. *Ocean Dynamics*, 54(6):539–560.
- [53] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.
- [54] Geir Evensen. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367, 2003.
- [55] Geir Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Earth and Environmental Science. Springer, 2009.
- [56] Geir Evensen, Gerrit Burgers, and Peter Jan van Leeuwen. Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126(6):1719–1724, 1998.
- [57] Geir Evensen and Peter Jan van Leeuwen. An ensemble Kalman smoother for nonlinear dynamics. *Monthly Weather Review*, 128:1852–1867, 2000.
- [58] D Fairbairn, S R Pring, A C Lorenzb, and I Roulstone. A comparison of 4dvar with ensemble data assimilation methods. *Quarterly Journal of the Royal Meteorological Society*, 140(678):281–294, 2014.
- [59] Nadia Fourrié and Jean-Noël Thépaut. Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 129(592):2425–2439, 2003.
- [60] Christina Frederick and Bjorn Engquist. Numerical methods for multiscale inverse problems. *arXiv preprint arXiv:1401.2431*, 2014.
- [61] M. A. Freitag, N. K. Nichols, and C. J. Budd. Resolution of sharp fronts in the presence of model error in variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 139(672):742–757, 2013.
- [62] Melina Freitag and Roland Potthast. Synergy of inverse problems and data assimilation techniques. In Mike Cullen, Melina A Freitag, Stefan Kindermann, and Robert Scheichl, editors, *Large Scale Inverse Problems*, volume 13 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter, Berlin, 2013.

- [63] Reinhard Furrer and Thomas Bengtsson. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *Journal of Multivariate Analysis*, 98(2):227 – 255, 2007.
- [64] Gregory Gaspari and Stephen E. Cohn. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554):723–757, 1999.
- [65] Michael Ghil and Paola Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. *Adv. Geophys*, 33:141–266, 1991.
- [66] Mitchell D Goldberg and Dr Fuzhong Weng. Advanced technology microwave sounder. In John J Qu, Wei Gao, Menas Kafatos, Robert E Murphy, and Vincent V Salomonson, editors, *Earth Science Satellite Remote Sensing*, pages 243–253. Springer Berlin Heidelberg, 2006.
- [67] S. Gratton, A. S. Lawless, and N. K. Nichols. Approximate gauss-newton methods for nonlinear least squares problems. *SIAM Journal on Optimization*, 18(1):106–132, 2007.
- [68] Charles W Groetsch. *Inverse problems in the mathematical sciences*. Theory and Practice of Applied Geophysics Series. Vieweg, 1993.
- [69] Rodolfo Guzzi. *Data Assimilation: Mathematical Concepts and Instructive Examples*. Springer, 2016.
- [70] Jacques Hadamard. *Lectures on Cauchy’s Problem in Linear Partial Differential Equations*. New Haven, CT: Yale University Press, 1923.
- [71] G J Haltiner and R T Williams. *Numerical prediction and dynamic meteorology*. Wiley, 1980.
- [72] Thomas M Hamill, Steven L Mullen, Chris Snyder, David P Baumhefner, and Zoltan Toth. Ensemble forecasting in the short to medium range: Report from a workshop. *Bulletin of the American Meteorological Society*, 81(11):2653–2664, 2000.
- [73] Thomas M. Hamill, Jeffrey S. Whitaker, and Chris Snyder. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, 129(11), 2001.

- [74] Martin Hanke, Andreas Neubauer, and Otmar Scherzer. A convergence analysis of the landweber iteration for nonlinear ill-posed problems. *Numerische Mathematik*, 72(1):21–37, 1995.
- [75] Per Christian Hansen. The truncatedsvd as a method for regularization. *BIT Numerical Mathematics*, 27(4):534–553.
- [76] Per Christian Hansen. Analysis of discrete ill-posed problems by means of the l-curve. *SIAM review*, 34(4):561–580, 1992.
- [77] Per Christian Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Mathematical Modeling and Computation. Society for Industrial and Applied Mathematics, 1998.
- [78] Per Christian Hansen, James G Nagy, and Dianne P O’leary. *Deblurring Images: Matrices, Spectra, and Filtering (Fundamentals of Algorithms 3) (Fundamentals of Algorithms)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2006.
- [79] Timothy J Hewison and Roger Saunders. Measurements of the AMSU-B antenna pattern. *Geoscience and Remote Sensing, IEEE Transactions on*, 34(2):405–412, Mar 1996.
- [80] F Hilton, N C Atkinson, S J English, and J R Eyre. Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments. *Quarterly Journal of the Royal Meteorological Society*, 135(639):495–505, 2009.
- [81] Bernd Hofmann. *Regularization for applied inverse and ill-posed problems: a numerical approach*. Teubner-Texte zur Mathematik. B.G. Teubner, 1986.
- [82] P L Houtekamer and Herschel L Mitchell. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126(3), 1998.
- [83] P L Houtekamer and Herschel L Mitchell. A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129(1):123–137, 2001.
- [84] P L Houtekamer and Herschel L Mitchell. Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3269–3289, 2005.
- [85] P L Houtekamer, Herschel L Mitchell, and Xingxiu Deng. Model error representation in an operational ensemble Kalman filter. *Monthly Weather Review*, 137(7), 2009.



- [86] P L Houtekamer, Herschel L Mitchell, Gérard Pellerin, Mark Buehner, Martin Charron, Lubos Spacek, and Bjarne Hansen. Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Monthly Weather Review*, 133(3), 2005.
- [87] Brian R Hunt, Eric J Kostelich, and Istvan Szunyogh. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230(1-2):112–126, 2007.
- [88] Tijana Janjic, Lars Nerger, Alberta Albertella, Jens Schröter, and Sergey Skachko. On domain localization in ensemble-based Kalman filter algorithms. *Monthly Weather Review*, 139(7):2046–2060, 2011.
- [89] Tijana Janjic, Lars Nerger, Alberta Albertella, Jens Schröter, and Sergey Skachko. On domain localization in ensemble-based Kalman filter algorithms. *Monthly Weather Review*, 139(7):2046–2060, 2011.
- [90] A.H. Jazwinski. *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engineering. Elsevier Science, 1970.
- [91] Jari Kaipio and Erkki Somersalo. *Statistical and computational inverse problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2005.
- [92] R E Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, March 1960.
- [93] Eugenia Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 2003.
- [94] Eugenia Kalnay, Brian Hunt, Edward Ott, and Istvan Szunyogh. *Ensemble forecasting and data assimilation: two problems with the same solution*. Cambridge University Press, 2006.
- [95] B. Kaltenbacher, A Neubauer, and O Scherzer. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*. Radon Series on Computational and Applied Mathematics. De Gruyter, 2008.
- [96] Paul G Kaminski, Arthur E Bryson, JR, and Stanley F Schmidt. Discrete square root filtering:a survey of current techniques. *IEEE Transactions on Automatic Control*, AC-16(6):727–736, Mar 1971.

- [97] Masahiro Kazumori, Quanhua Liu, Russ Treadon, and John C Derber. Impact study of AMSR-E radiances in the NCEP global data assimilation system. *Monthly Weather Review*, 136(2):541–559, 2008.
- [98] Joseph B Keller. Inverse problems. *The American Mathematical Monthly*, 83(2):107–118, 1976.
- [99] Paul Kirchgessner, Lars Nerger, and Angelika Bunse-Gerstner. On the choice of an optimal localization radius in ensemble Kalman filter methods. *Monthly Weather Review*, (6):2165–2175, 2014.
- [100] Andreas Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems*. Applied Mathematics Science. Springer-Verlag New York, 2011.
- [101] Dieter Klaes, Jörg Ackermann, Rainer Schraidt, Tim Patterson, Peter Schlüssel, Pepe Phillips, Arlindo Arriaga, and Jochen Grandell. The ATOVS and AVHRR product processing facility of EPS. *Advances in Space Research*, 36(5):996–1002, 2005. Atmospheric Remote Sensing: Earths Surface, Troposphere, Stratosphere and Mesosphere-I.
- [102] M Koda and H Koyama. Square root filtering formulation for satellite autonomous navigation. In *A bridge between control science and technology.*, volume 5, pages 2951–2955. Oxford and New York, Pergamon Press, 1985.
- [103] Rainer Kress. *Linear Integral Equations*, volume 82 of *Applied Mathematical Sciences*. Springer New York, 1999.
- [104] Evan Kwiatkowski and Jan Mandel. Convergence of the square root ensemble Kalman filter in the large ensemble limit. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):1–17, 2015.
- [105] François-Xavier Le Dimet and Olivier Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A: Dynamic Meteorology and Oceanography*, 38(2):97–110, 1986.
- [106] Hai-Tien Lee, Arnold Gruber, Robert G Ellingson, and Istvan Laszlo. Development of the HIRS outgoing longwave radiation climate dataset. *Journal of Atmospheric and Oceanic Technology*, 24(12):2029–2047, 2007.
- [107] Yoonsang Lee and Andrew J. Majda. Multiscale methods for data assimilation in turbulent systems. *Multiscale Modeling & Simulation*, 13(2):691–713, 2015.

- [108] J M Lewis, S Lakshmivarahan, and S Dhall. *Dynamic Data Assimilation: A Least Squares Approach*, volume 13 of *Dynamic data assimilation: a least squares approach*. Cambridge University Press, 2006.
- [109] Stephane Mallat, George Papanicolaou, and Zhifeng Zhang. Adaptive covariance estimation of locally stationary processes. *The Annals of Statistics*, 26(1):1–47, 1998.
- [110] J Le Marshall, J Jung, T Zapotocny, J Derber, R Treadon, S Lord, M Goldberg, and W Wolf. The application of airs radiances in numerical weather prediction. *Australian Meteorological Magazine*, 55(3):213–217, 2006.
- [111] P Martinet, L Lavanant, N Fourri, F Rabiera, and A Gambacortac. Evaluation of a revised IASI channel selection for cloudy retrievals with a focus on the mediterranean basin. *Quarterly Journal of the Royal Meteorological Society*, 2013.
- [112] Boris A Marx and Roland Potthast. On instabilities in data assimilation algorithms. *GEM - International Journal on Geomathematics*, 3(2):253–278, 2012.
- [113] C Mätzler, Institution of Engineering, and Technology. *Thermal Microwave Radiation: Applications for Remote Sensing*. IET electromagnetic waves series. Institution of Engineering and Technology, 2006.
- [114] Zhiyong Meng and Fuqing Zhang. Limited-area ensemble-based data assimilation. *Monthly Weather Review*, 139(7):2025–2045, 2011.
- [115] Jarno Mielikainen, Bormin Huang, Hung-Lung Allen Huang, and Roger Saunders. Development of GPU-based RTTOV-7 IASI and AMSU-A forward models. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 2773–2776, July 2011.
- [116] T. Miyoshi. The gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform kalman filter. *Monthly Weather Review*, 139(5):1519–1535, 2011.
- [117] T. Miyoshi and K. Kondo. A multi-scale localization approach to an ensemble kalman filter. *SOLA*, 9:170–173, 2013.
- [118] T. Miyoshi and Y. Sato. Assimilating satellite radiances with a local ensemble transform kalman filter (LETKF) applied to the jma global model (GSM). *SOLA*, 3:37–40, 2007.

- [119] T. Miyoshi, S. Yamane, and T. Enomoto. Localizing the error covariance by physical distances within a local ensemble transform kalman filter(LETKF). *SOLA*, 3:89–92, 2007.
- [120] Takemasa Miyoshi, Yoshiaki Sato, and Takashi Kadowaki. Ensemble Kalman filter and 4d-var intercomparison with the japanese operational global analysis and prediction system. *Monthly Weather Review*, 138(7), 2010.
- [121] Nicola Montaldo and John D Albertson. Multi-scale assimilation of surface soil moisture data for robust root zone moisture predictions. *Advances in water resources*, 26(1):33–44, 2003.
- [122] Carsten Montzka, Valentijn R.N.Pauwels, Harrie-Jan Hendricks Franssen, Xu-jun Han, and Harry Vereecken. Multivariate and multiscale data assimilation in terrestrial systems: A review. *Sensors*, 12(12):16291–16333, 2012.
- [123] Christina Muth, Paul S Lee, James C Shiue, and W Allan Webb. Advanced technology microwave sounder on NPOESS and NPP. In *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International*, volume 4, pages 2454–2458 vol.4, Sept 2004.
- [124] Christina Muth, W A Webb, W Atwood, and P Lee. Advanced technology microwave sounder on the national polar-orbiting operational environmental satellite system. In *Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International*, volume 1, page 4pp, July 2005.
- [125] Siegfried MÃ¼ller. *Adaptive Multiscale Schemes for Conservation Laws*, volume 27 of *Lecture Notes in Computational Science and Engineering*. Springer Berlin Heidelberg, 2003.
- [126] Aamir Nadeem and Roland Potthast. Transformed and generalized localization for ensemble methods in data assimilation. *Mathematical Methods in the Applied Sciences*, 39(4):619–634, 2016.
- [127] Aamir Nadeem, Roland Potthast, and Andreas Rhodin. On sequential multi-scale inversion and data assimilation. revised.
- [128] S Nakamori. Square-root algorithms of rls wiener filter and fixed-point smoother in linear discrete stochastic systems. *Applied Mathematics and Computation*, 203(1):186–193, 2008.

- [129] Gen Nakamura and Roland Potthast. *Inverse Modeling*. 2053-2563. IOP Publishing, 2015.
- [130] Lars Nerger, Tijana Janjic, Jens Schröter, and Wolfgang Hiller. A regulated localization scheme for ensemble-based Kalman filters. *Quarterly Journal of the Royal Meteorological Society*, 138(664):802–812, 2012.
- [131] Lars Nerger, Svenja Schulte, and Angelika Bunse-Gerstner. On the influence of model nonlinearity and localization on ensemble Kalman smoothing. *Quarterly Journal of the Royal Meteorological Society*, 140:2249–2259, 2014.
- [132] Arnold Neumaier. Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM Rev.*, 40(3):636–666, Sep 1998.
- [133] Nancy Kay Nichols. Mathematical concepts of data assimilation. In *Data assimilation*, pages 13–39. Springer, 2010.
- [134] James Nolen, Grigorios A. Pavliotis, and Andrew M. Stuart. Multiscale modelling and inverse problems. In Ivan G. Graham, Thomas Y. Hou, Omar Lakkis, and Robert Scheichl, editors, *Numerical Analysis of Multiscale Problems*, volume 83 of *Lecture Notes in Computational Science and Engineering*, pages 1–34. Springer Berlin Heidelberg, 2012.
- [135] Peter R Oke, Pavel Sakov, and Stuart P Corney. Impacts of localisation in the enkf and enoi: experiments with a small model. *Ocean Dynamics*, 57(1):32–45, 2007.
- [136] Edward Ott, Brain R Hunt, Istvan Szunzogh, Aleksey V Zimin, Eric J Kostelich, Matteo Corazza, Eugenia Kalnay, D J Patil, and James A Yorke. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A*, 56(5):415–428, 2004.
- [137] Ming Pan and Eric F Wood. A multiscale ensemble filtering system for hydrologic data assimilation. part ii: Application to land surface modeling with satellite rainfall forcing. *Journal of Hydrometeorology*, 10(6):1493–1506, 2009.
- [138] Ming Pan, Eric F. Wood, Dennis B. McLaughlin, Dara Entekhabi, and Lifeng Luo. A multiscale ensemble filtering system for hydrologic data assimilation. part i: Implementation and synthetic experiment. *Journal of Hydrometeorology*, 10(3):794–806, 2009.

- [139] Seon Ki Park and Liang Park, editors. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*. Springer-Verlag Berlin Heidelberg, 2013.
- [140] C L Parkinson. Aqua: an earth-observing satellite mission to examine water and other climate variables. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(2):173–183, Feb 2003.
- [141] David F. Parrish and John C. Derber. The national meteorological center’s spectral statistical- interpolation analysis system. *Monthly Weather Review*, 120(8):1747–1763, 1992.
- [142] G A Pavliotis and Andrew Stuart. *Multiscale Methods: Averaging and Homogenization*, volume 53 of *Texts in Applied Mathematics*. Springer-Verlag New York, 2008.
- [143] Clemens Pechstein. *Finite and Boundary Element Tearing and Interconnecting Solvers for Multiscale Problems*, volume 90 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag Berlin Heidelberg, 2013.
- [144] A Perianez, H Reich, and R Potthast. Finite-element type error analysis for ensemble methods in data assimilation. Preprint, 2014.
- [145] A Perianez, H Reich, and R Potthast. Optimal localization for ensemble Kalman filter systems. *Journal of the Meteorological Society of Japan*, 92(6):585–597, 2014.
- [146] Ruth E Petrie. Localization in the ensemble Kalman filter. Master’s thesis, University of reading, United Kingdom, 2008.
- [147] Ruth E Petrie and Sarah L Dance. Ensemble-based data assimilation and the localisation problem. *Weather*, 65(3):65–69, 2010.
- [148] Von John J Qu, Wei Gao, Menas Kafatos, Robert E Murphy, and Vincent V Salomonson. *Earth Science Satellite Remote Sensing: Vol.1: Science and Instruments*. Earth Science Satellite Remote Sensing. Springer, 2007.
- [149] Florence Rabier, Philippe Courtier, Jean Pailleux, Olivier Talagrand, and Drasko Vasiljevic. A comparison between four-dimensional variational assimilation and simplified sequential assimilation relying on three-dimensional variational analysis. *Quarterly Journal of the Royal Meteorological Society*, 119(512):845–880, 1993.

- [150] Santha R. Akella et al. Reservoir multiscale data assimilation using the ensemble Kalman filter. *Applied Mathematics*, 2(02):165–180, 2011.
- [151] David R. Stauffer and Nelson L. Seaman. Multiscale four-dimensional data assimilation. *Journal of Applied Meteorology*, 33(3):416–434, 1994.
- [152] Pavel Sakov and Peter Oke. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus A*, 60(2), 2008.
- [153] N. J. Salkind. *Encyclopedia of Measurement and Statistics*. SAGE Publications, 2006.
- [154] R. Saunderson, M. Matricardi, and P. Brunel. An improved fast radiative transfer model for assimilation of satellite radiance observations. *Quarterly Journal of the Royal Meteorological Society*, 125(556):1407–1425, 1999.
- [155] Roger Saunders, Timothy J. Hewison, Stephen J. Stringer, and Nigel C. Atkinson. The radiometric characterization of AMSU-B. *Microwave Theory and Techniques, IEEE Transactions on*, 43(4):760–771, Apr 1995.
- [156] Roger W. Saunders, Thomas A. Blackmore, Brett Candy, Peter N. Francis, and Tim J. Hewison. Monitoring satellite radiance biases using NWP models. *Geoscience and Remote Sensing, IEEE Transactions on*, 51(3):1124–1138, March 2013.
- [157] A. Schomburg, C. Schraff, and R. Potthast. A concept for the assimilation of satellite cloud information in an ensemble Kalman filter: single-observation experiments. *Quarterly Journal of the Royal Meteorological Society*, 141(688):893–908, 2015.
- [158] C. Schraff, H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Perianez, and R. Potthast. Kilometre-scale ensemble data assimilation for the cosmo model (kenda). *Quarterly Journal of the Royal Meteorological Society*, 142(696):1453–1472, 2016.
- [159] T. Schuster, B. Kaltenbacher, B. Hofmann, and K.S. Kazimierski. *Regularization Methods in Banach Spaces*. Radon series on computational and applied mathematics. De Gruyter, 2012.
- [160] Lei Shi and John J. Bates. Three decades of intersatellite-calibrated high-resolution infrared radiation sounder upper tropospheric water vapor. *Journal of Geophysical Research: Atmospheres*, 116(D4), 2011.

- [161] Domenico Solimini. *Microwave Radiometry and Remote Sensing of The Environment*. Taylor & Francis, 1995.
- [162] Istvan Szunyogh, Eric J Kostelich, Gyorgyi Gyarmati, Eugenia Kalnay, Brian R Hunt, Edward Ott, Elizabeth Satterfield, and James A Yorke. A local ensemble transform Kalman filter data assimilation system for the ncep global model. *Tellus A*, 60(1):113–130, 2008.
- [163] Monique Tanguay, Luc Fillion, Ervig Lapalme, and Manon LajoieManon. Four-dimensional variational data assimilation for the canadian regional deterministic prediction system. *Monthly Weather Review*, 140(5):1517–1538, 2012.
- [164] Owen E Thompson. HIRS-AMTS satellite sounding system test-theoretical and empirical vertical resolving power. *Journal of Applied Meteorology*, 21(10):1550–1561, 1982.
- [165] AndreĀ Nikolaevich Tikhonov and VasiliĀ İi, Ai,ĵkovlevich Arsenin. *Solutions of ill-posed problems*. Scripta series in mathematics. Winston, 1977.
- [166] Michael K Tippett, Jeffrey L Anderson, Craig H Bishop, Thomas M Hamill, and Jeffrey S Whitaker. Ensemble square root filters. *Monthly Weather Review*, 131(7):1485–1490, 2003.
- [167] L Ventress and A Dudhia. Improving the selection of IASI channels for use in numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 2013.
- [168] Curtis R. Vogel. *Computational Methods for Inverse Problems*. Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics, 2002.
- [169] Jiang Wan and Nicholas Zabaras. A bayesian approach to multiscale inverse problems using the sequential monte carlo method. *Inverse Problems*, 27(10):105004, 2011.
- [170] Yanfei Wang, Anatoly G. Yagola, and Changchun Yang. *Optimization and Regularization for Computational Inverse Problems and Applications*. Springer-Verlag Berlin Heidelberg, 2011.
- [171] Jeffrey S. Whitaker and Thomas M. Hamill. Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130(7), 2002.



- [172] Jeffrey S Whitaker, Thomas M. Hamill, Xue Wei, Yucheng Song, and Zoltan Toth. Ensemble data assimilation with the NCEP global forecast system. *Monthly Weather Review*, 136(2), 2008.
- [173] Fatih Yaman, Valery G Yakhno, and Roland Potthast. A survey on inverse problems for applied sciences. *Mathematical Problems in Engineering*, 2013:19, 2013.
- [174] Y.Xie, S.Koch, J.McGinley, S.Albers, P.E.Bieringer, M.Wolfson, and M.Chan. A space-time multiscale analysis system: A sequential variational analysis approach. *Monthly Weather Review*, 139(4):1224–1240, 2011.
- [175] Yong Zhou, YuFeng Zhang, and JuZhong Zhang. A new adaptive square-root unscented Kalman filter for nonlinear system. *Applied Mechanics and Materials*, 300-301:623–626, 2013.
- [176] Yuhua Zhou, Dennis McLaughlin, Dara Entekhabi, and Gene-Hua Crystal Ng. An ensemble multiscale filter for large nonlinear data assimilation problems. *Monthly Weather Review*, 136(2):678–698, 2008.
- [177] Jiang Zhu, Fei Zheng, and Xichen Li. A new localization implementation scheme for ensemble data assimilation of non-local observations. *Tellus A*, 63(2), 2011.
- [178] Xiao Xiang Zhu and Richard Bamler. Tomographic sar inversion by-norm regularization - the compressive sensing approach. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(10):3839–3846, 2010.
- [179] Cheng-Zhi Zou and Wenhui Wang. Intersatellite calibration of AMSU-A observations for weather and climate applications. *Journal of Geophysical Research: Atmospheres*, 116(D23), 2011.
- [180] Yu Zou and Roger Ghanem. A multiscale data assimilation with the ensemble Kalman filter. *Multiscale Modeling & Simulation*, 3(1):131–150, 2005.