Title	Development of efficient data assimilation methods for solute transport problems
Author(s)	Zayats, Mykhaylo
Publication Date	2018-09-18
Publisher	NUI Galway
Item record	http://hdl.handle.net/10379/10050

Development of efficient data assimilation methods for solute transport problems

by Mykhaylo Zayats

supervisor: Prof. Michael Hartnett

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy,
in the College of Engineering and Informatics, NUI Galway

March 2018

Declaration

I declare that this dissertation, in whole or in part, has not been submitted to any University as an exercise for a degree. I further declare that, except where reference is given, the work is entirely my own.

Signed:

Mykhaylo Zayats March 2018

Dedication

This thesis is dedicated to the memory of my grandmother.

Acknowledgments

I wish to thank the IBM Research and the College of Engineering and Informatics, NUI Galway for providing support for this research.

Special thanks to my supervisors: Prof. Michael Hartnett, Dr. Emanuele Ragnoli and Dr. Sergiy Zhuk for their patience, encouragements and many fruitful discussions. Thank you to our research group and to all of the staff of NUI Galway and IBM Research – Ireland for helping me during the course of this research.

Many thanks to Nadia for her love, understanding and belief in me, and to my friends and colleagues for their friendship and good company. Without them, I would not have enjoyed my time in Ireland as much as I did. Thank you to my family, and in particular, to my parents Maria and Vasyl and to my sister Tonya.

Finally, I would like to express my gratitude to all my teachers from the Ivan Franko National University of Lviv and Lviv Physics and Mathematics Lyceum. Thank you for opening the world of science to me.

Дякую Вам усім

Go raibh maith agaibh go leir

Abstract

The solution of marine contaminant transport problems is a significant research topic in civil engineering. Typically, the problem is represented as a partial differential equation described by the non-stationary, advection-diffusion operator. The underlying equation is approximated in space and time, and the state of the approximate numerical model is obtained as a solution of the corresponding linear algebraic system. However, for several reasons, numerical models do not necessarily replicate the process investigated exactly. In fact, their application and deployment generate modelling errors and discrepancies, a well-known and challenging problem to solve, which cannot be ignored in practice.

At the same time, the rapid development of measuring devices allows easier collection of data of the physical process. Typically, observations are prone to be contaminated by errors too, those generated by noise or other physical reasons. Moreover, observations are usually quite sparse in time and space. The combination and compromise between a numerical model and observations are integrated using data assimilation techniques. The development of efficient methods of data assimilation techniques in terms of estimation quality and computation speed is the main concern of this research.

The traditional algorithms of data assimilation such as minimax or Kalman filters are often used to quantify uncertainties represented by the model and observation errors. They construct an analysis state and propagate in time, taking into account model dynamics and observed information. The numerical algorithms of these filters are computationally expensive as they require multiplication and inversion of matrices of the size equal to the number of degrees of freedom of the system. Moreover, traditional filters are not scalable with respect to the number of discretisation nodes.

In this research, a combination of traditional filters with domain decomposition techniques are investigated to assess reduction of computational costs. The application of decomposition to the assimilation problem facilitates the reformulation of the global problem as a set of local subproblems coupled by continuity or transmission conditions. To solve the decomposed assimilation problem, two new approaches are considered. The first one discretises the transmission conditions directly and yields a system of differential-algebraic equations. The latter is solved by using a modified version of the minimax filter. The second approach imposes transmission conditions into the variational formulation of the local subproblems. A set of local

differential problems is solved by the iterative method of Schwarz. This approach is further extended to Kalman and ensemble filters using their equivalence with the minimax filter.

The efficiency of the proposed methods is examined using numerical experiments with different configurations including simulations with constant velocity field, periodic velocity field and velocity field generated by TELEMAC 2D for a tidal basin. The quality of estimates of the localised filters is assessed against both traditional filters and true solutions. The computational efficiency of the localised filters is evaluated, compared to the existing methods and discussed. Finally, scalability properties of the proposed algorithms are presented.

Table of contents

List of abbreviations

ADN – adaptive Dirichlet-Neumann.

ARE – algebraic Riccati equation(s).

ARN – adaptive Robin-Neumann.

BLUE – best linear unbiased estimator.

DA – data assimilation.

d-ADN – damped adaptive Dirichlet-Neumann.

DAE – differential algebraic equation(s).

d-ARN – damped adaptive Robin-Neumann.

DD – domain decomposition.

DRE – differential Riccati equation(s).

EKF – extended Kalman filter.

EnKF – ensemble Kalman filter.

EnTKF – ensemble transform Kalman filter.

EFDC – Environmental Fluid Dynamics Code.

FEM – finite element method.

FDM – finite difference method.

FP – Fokker-Planck.

LQ – linear-quadratic.

ODE – ordinary differential equation.

OI – optimal interpolation.

PDE – partial differential equation(s).

PDF – probability distribution function.

UBB – unknown but bounded.

List of figures

List of tables

1 Introduction

1.1 Mathematical Modelling

Mathematical models can be viewed in a broad sense as descriptions of physical processes using mathematical concepts and language. Consequently, the problem of modelling a physical system is to reconstruct or estimate a system state, its properties and behaviour at different instances of time and space.

There are two possible ways to describe different natural phenomena. The first is, the *a priori* approach, which constructs a mathematical description (background model) based on the laws of physics. Usually, such representation results in algebraic, differential or integral equations, the solution of which is a system state. This is a commonly used approach for applications in natural sciences such as hydrodynamics or transport of pollutants in atmospheric and marine environments.

An attractive benefit of this approach is that it is distinct from the natural environment and the problem of interest is being simulated under some ideal conditions. A major disadvantage is that a mathematical description of a process contains model errors, the unknown discrepancy between the model dynamics and actual system dynamics. Thus it can never be a perfect replica of reality. Quantifying the uncertainty due to model errors is a challenging problem; however, this is a problem that cannot be ignored in practice.

The second modelling approach, the *a posteriori* one, depends on the collection of observations of the process in question. This approach is used if a reliable a priori model of a system is difficult to construct. Also, it is often used to describe processes which are not closed systems and external factors cannot be quantified. It is popular for applications in economy, social sciences and chaotic systems.

To produce state estimates, it is possible to interpolate a collection of measured values provided by observations of the true state. Direct observations are free of modelling errors, and if they are dense enough can produce reliable results. In some applications, the state of the system can be measured directly with high accuracy; in others, direct measurement of the global system state is not feasible due to spatiotemporal gaps. Instead, the state must be inferred from available data. In order to fill observation gaps, it is necessary to rely on some background information in the form of an a priori estimate of the model state. Another limitation is that

observations are contaminated with errors. These may be random errors that tend to decrease by averaging, or systematic errors that, in contrast, do not decrease by averaging, and can be described only as components of a bounded set.

One may notice that both approaches of modelling mentioned above have the potential to complement each other. The background model thus may extend the observations, fill in the observational gaps and allows one to organise, summarise, and propagate the information from observations. On the other hand, the observations provide information about initial and boundary conditions for background model and can be used for adjusting a model to reduce modelling errors. Methods which incorporate observed information with mathematical representation of a process, by taking advantage of consistency constraints with laws of time evolution and physical properties, are referred to as *data assimilation* (DA). The development of such methods is the main focus of this thesis.

Research into the modelling of natural phenomena, up to the end of the twentieth century, may be viewed as the study of solutions of differential equations and predicting of model behaviour by deterministic solutions of these differential equations. However, model errors of mathematical descriptions of real-world processes can be eliminated only in a very small number of cases. From a theoretical point of view, the underlying physical laws are defined for an idealised case, but it is not always possible to have an accurate description of input parameters. For example, boundary or initial conditions may come from noisy measurements or in some cases may be not known. A mistreatment of model errors may result in inaccurate predictions of a system state.

It is also common that an analytical solution to the underlying partial differential equation (PDE) is not possible to find. Hence, numerical methods are required to construct approximations to the solution of the problem, thereby introducing approximation errors. The quality of an approximation depends on the number of degrees of freedom. Furthermore, the presence of errors in numerical schemes may lead to strong oscillations in a numerical solution, especially for models that are not unconditionally stable. Thus, in order to predict a state with a high accuracy for a long time window, one may need to solve a system of a very high order which can be a difficult constraint for real-time applications.

At the same time, a rapid degree of development of electronic measuring devices such as radars, buoys, etc., has begun to produce more observed data. Often, this data is sparse and may not

cover a whole domain of interest. Observed data is also prone to be contaminated by measuring errors that require special treatment. Incorporation of observations into mathematical models performed by the DA approach allows us to obtain system estimates that take advantage of both sources of information about the physical process.

DA algorithms are critically important for many real-world applications. Originally its major application was in numerical weather forecasting (Swinbank, 2010), but later DA methods were successfully adopted in hydrology (Houser et al., 2012), oceanography (Martin et al., 2015), atmospheric chemistry (Khattatov and Yudin, 2010), climate studies (Rood and Bosilovich, 2010) and others. More recently, DA methods were used in smart grids for wind and solar forecasting, and in the optimisation of power outputs of energy systems (Carpenter and Shaw, 2012). The reader is also encouraged to visit the web page http://www.ecmwf.int/ of the European Centre for Medium-Range Weather Forecasts (ECMWF) where dozens of DA research projects are listed.

Over the last few decades, many DA algorithms have been proposed starting with basic nodal averaging, up to complicated stochastic strategies based on the Fokker-Planck equations. The applicability of those methods in engineering applications may be problematic, especially for problems that are defined over a large spatial domain and require frequent analysis outputs. The main problem arises from the heavy computational demands of traditional DA approaches, and the lack of scalability properties with respect to the problem size. The size of an assimilation problem is usually defined by the size of a background model. In turn, a desirable discretisation of a model, defined over a large-scale physical domain, leads to many degrees of freedom in its discrete representation. That said, the development of efficient methods of DA is the focus of much research in this area.

An important concept, extensively used in numerical analysis for the reduction of the computational costs associated with solving PDE, is *domain decomposition* (DD). DD splits a global problem into a number of local subproblems with additional boundary conditions applied over an interface of decomposition, these are called *transmission conditions*. These conditions enforce continuity between local solutions and therefore provide equivalence with a global solution. In this thesis, DD is adapted to assimilation problems producing a family of DA methods referred to as *localised* DA (LDA) methods.

1.2 Aims and Objectives

This thesis is concerned with two-dimensional contaminant transport processes in the marine environment. The background model of the process is described by a linear advection-diffusion PDE with a non-stationary velocity flow and by a linear equation representing observations of concentration. To estimate the state of the process, online methods of DA such as minimax filter and Kalman filter are considered.

The significant disadvantage of traditional methods of DA is the requirement of large computational resources which is a strong limitation for practical use. The primary aim of this research is the development of computationally efficient methods of DA that can be applied in engineering practice. To this end, combining traditional methods of DA with DD is considered. This approach allows for localisation of computations of a global problem into local subdomains. Two different strategies of localisation are introduced here:

1) Localisation in the form of differential algebraic equation (DAE).
2) Interconnected localisation.

Algorithms of both methods are developed for the minimax filter, and the interconnected localisation is also applied for the Kalman filter and the ensemble transform Kalman filter (EnTKF). Properties of both approaches are examined using numerical experiments with idealised and realistic configurations.

The objectives of this research are the following:

1) The compilation of a literature review of the current state of the art of DA theory, focusing on the most commonly used methods and on methods that are used in marine applications.
2) The compilation of a literature review of strategies for a reduction of the computational complexity of PDE and ordinary differential equation (ODE) solvers.
3) The implementation of a FEM (finite elements method) code for a linear transport model defined by the non-stationary, advection-diffusion equation with a non-stationary velocity field.
4) The implementation and comparison of traditional methods of DA: minimax filter, Kalman filter and EnTKF.
5) Research and development of the DD numerical strategy for a FEM model of the transport equation that performs well in strongly advection-dominated flows.

6) Development of the localisation strategy of the minimax filter in the form of DAE using DD and minimax framework.
7) Development of the interconnected localised minimax filter using DD and the Schwarz iterative approach.
8) Extension of the interconnected localisation strategy to the Kalman filter and to EnTKF.
9) Computational analysis and comparison of the localised filter with traditional filters and with another filter based on DD.
10) Application of the localised filters for the numerical simulation of tracer concentration to a tidal basin.

1.3 Structure of the Thesis

This thesis is organised as follows.

Chapter 2 presents a literature review of the numerical methods of DA and DD. The review begins with the general problem statement and categorisation of assimilation strategies. Then the main categories of DA approaches, such as stochastic methods, statistic methods and optimal control methods are outlined. The advantages and disadvantages of various methods are discussed, along with relationships between them. Particular emphasis is devoted to the numerical performance of methods and variants with the lowest computational requirements. The second part of the chapter reviews the state of the art of DD methods. In relation to the transport problem, the adaptive DD methods are discussed. At the end of the chapter, the main findings of the review are summarised.

Chapter 3 elaborates on the mathematical description of the DA for the linear transport problem. The problem defined by the advection-diffusion equation, its transformation into a weak representation, discretisation in space using FEM and discretisation in time using mid-point integration rule are presented in detail. The DA methods are applied to the problem continuous in time and discrete in space, and to the problem discrete in time and space. The Kalman-Bucy and Kalman filters are derived, and an equivalence between them is established. The approximation of KF by the ensemble Kalman filter (EnKF) and EnTKF is also presented here.

Chapter 4 expands on the minimax state estimation. First, the time-space continuous formulation of the minimax filter in an operator form is explained. Then, the approximation of the continuous minimax filter using FEM is detailed, and an efficient numerical treatment of

the corresponding Riccati equation is constructed. The chapter ends by establishing a relation between the minimax filter and the Kalman-Bucy filter, and providing performance assessment and comparison of the minimax and Kalman-based filters.

Chapter 5 introduces the localisation approach by means of DD and devises the localised filter in the form of DAE. The performance of the method is demonstrated for several numerical experiments, and the results of those experiments are discussed.

Chapter 6 constructs the interconnected localised minimax filter using the DD and Schwarz iterative approach. The numerical algorithm of the method based on FEM is discussed in detail. A numerical investigation of the proposed method and its scalability properties are presented.

Chapter 7 extends the interconnected localised strategy for the KF and, after discussion on the localisation of ensembles, the interconnected localised EnTKF estimators are also devised. A performance study of the localised filters is presented first for the simulation with idealised configuration and then for the simulation of a tidal basin.

Finally, Chapter 8 contains a discussion of the main conclusions from this research and provides directions for potential future research in this area.

2 Literature Review

The main aim of this review is to present the state of the art of the DA problem, its applications to hydrodynamic systems and possible bottlenecks. A particular emphasis was on the performance of DA algorithms, their scalability for practical purposes and possible strategies for further improvements. This review is organized into two parts: the first presents the problem statement and describes current developments of the assimilation methods. The second part focuses on DD methods and their applications for DA problems.

The literature pertaining to assimilation methods is very rich. Although different methods were developed for different applications and came from different backgrounds, they can be derived in terms of several unifying frameworks. From a mathematical perspective, DA problem is naturally described by probability distributions and thus relies on a stochastic framework. It was developed initially as a mathematical theory to describe chaotic systems which later evolved into several practical methods. The stochastic framework is reviewed in more details in the first section. At the same time, in the engineering world, the first applications of data assimilation were based on simple statistical analysis of model and observation errors. Starting with very practical examples, modern Kalman theory, the Kalman framework or more generally, the statistical framework were constructed. This framework is reviewed in the second section.

Developments of those two frameworks can be seen as complementary to each other. Indeed, in the stochastic case, the most general approach is formulated from the very beginning, and all efforts are directed at an efficient solver for mathematical equations or simplification of those equations through additional constraints. In the statistical framework, several conditions about a problem structure are assumed, and the main efforts were devoted to overcome those conditions.

In the third section, the main disadvantages of both frameworks are reviewed, and the optimal control framework is presented. This framework was designed to be more robust than Kalman and practical at the same time. The equivalence between the frameworks is also provided.

Finally, the second part of the review is devoted to the state of the art of DD methods which are an essential tool in decreasing the computational efforts required to solve PDE. This part further reviews combinations of DD and DA, and describes the motivation for this research.

2.1 Problem statement

In this work, the unified notation that bridges formulations in the DA field are adopted from the publication of Ide et al. (1997). In what follows, a mathematical description of a process is called *background*, and measurements of the variables of the true process are called *observations*. Assimilation procedure is a dual process in which the background fills gaps in the observations and on the other side the observations adjust the model to be a better replication of actual physical dynamics. The result of assimilation is called an *analysis* or *forecast*, and it is assumed to be an estimate of the true state.

In the most general case, the background model is described by a PDE. That can be the Navier-Stokes equations or a shallow water model which describes hydrodynamic behaviour or a transport model defined by advection-diffusion equations (Hervouet, 2007). Later, most probably, underlying PDEs would be discretised in space and time producing a system of linear algebraic equations. The most popular methods of space discretisation include FEM (Donea and Huerta, 2003), finite difference method (FDM) (Chung, 2010), finite volumes method (FVM) (Versteeg and Malalasekera, 1995), spectral methods (SM) (Canuto et al., 1988). For time discretisation typically implicit midpoint, Crank–Nicolson or Runge-Kutta schemes (Kress, 1998) are used. In engineering practice, it is common to use sophisticated modelling toolsets based on the equations mentioned above. An example of those codes include environmental fluid dynamics code (EFDC) (Hamrick, 1992); integrated suites of solvers for free surface flow, sediments transport and wave propagation such as Telemac-Mascaret (Desombre and Lang, 2013); and coastal circulation and storm surge models (ADCIRC) (Luettich et al., 1992).

Observations of the earth systems can be made in several different ways. A traditional way is to rely on in situ measurements from ground-based stations, buoys, ships and aircraft (Thépaut and Andersson, 2010). In situ observations are generally considered to be point-wise and instantaneous. They provide relatively high spatiotemporal resolution in a small area of installation but for a global coverage an enormous amount of in situ measuring devices is required which is not possible.

Complementary to in situ are satellite observations from low earth orbit satellites (LEOs) and geostationary satellites (GEOs). For a comprehensive review of research satellite classifications and their missions, readers are referred to the review by Lahoz (2010).

The problem of assimilating observation into a mathematical model stems from the need of accurate weather analysis. Lynch (2008) published a deep insight into the history of scientific forecasting and weather prediction. The author described the main steps of developmet of DA systems and methods. The most basic algorithm of DA is a simple interpolation between the background and observation. This approach applied to weather forecasting was initially implemented by Richardson (1922). Later, it was extended to a more general algorithm called successive correction method (SCM) Cressman (1959) where the background was updated by observations weighted with some predefined functions. The SCM algorithm is attractive because of its simplicity. However, it does not properly account uncertainties in the problem which in practice lead to unsatisfactory results Lorenc (1986). For therefore, the assimilation method must be intelligent. The analysis should be conducted in an optimal way, as the best fit between the background and observations which takes into account physical behaviour and error description of the system. For instance, this may be achieved by a minimization of a penalty or a cost function.

DA methods are considered in different ways (Jazwinski, 1970):

1) When the state of the problem at time t is estimated by using data measured up to and including t, the approach is called *"filtering"*.
2) When the state of the problem at time t is estimated by using data from both past and future regarding t, the approach is called *"smoothing"*.
3) When the state of the problem at time t is estimated by using data only from the past regarding t, the approach is called *"predicting"*.

In the literature, predicting and filtering methods are often combined into sequential assimilation while smoothing methods are recalled retrospective or non-sequential (Bouttier and Courtier, 1999). This differentiation is explained by the fact that sequential algorithms can be involved in an online DA when an analysis is performed once new observations become available; in contrast, a retrospective analysis is done once all data is collected.

DA methods are also distinguished depending on a temporal structure of the background model and observations and can be categorised as follows:

1) Continuous, when both background and observations are continuous in time.
2) Continuous-discrete, when the background is continuous in time, and observations are discrete.
3) Discrete, when both background and observations are time discrete.

It should be mentioned that typically for marine application, the underlying problem is continuous and defined in terms of differential equations. Thus, depending on the nature of the observations, it is desired to devise continuous or continuous-discrete assimilation algorithms. At the same time, a continuous model of the marine process can be discretised in time so that the discrete DA can be applied. It is not always clear which type of DA algorithm to choose, and the decision depends on the particular configuration of the problem.

Another important characteristic of DA approaches is based on the underlying theoretical framework. Methods derived from different ideas are possible to categorize in the following groups:

1) Stochastic.
2) Statistical.
3) Optimal control.

In the following the main advances of each framework are reviewed.

2.2 Stochastic methods

A common way to quantify uncertainties in a system is to represent it as a random variable. This variable is called a stochastic variable, and the corresponding system is called a stochastic system. Let $\{u\}$ be a random variable with a probability distribution function (PDF) $p(u)$ which defines the a priori or background state of a physical process and let $\{y\}$ be a random variable which defines observations of the state of the same process. Both $\{u\}$ and $\{y\}$ describe one physical system but from different perspectives. Thus, it is reasonable to assume they are not independent. Furthermore, the distribution of $\{y\}$ is normally given in the form of conditional PDF $p(y|u)$ sometimes called a data distribution or likelihood function.

From a probabilistic perspective, the background describes prior information which is not complete. The presence of observations provides some evidence which can be used to update prior information. Therefore, the problem of DA is formulated as a problem of a statistical inference which aims to estimate a posteriori conditional probability $p(u|y)$ based on background distribution $p(u)$ and data distribution $p(y|u)$. Probability $p(u|y)$ can be found using the well-known Bayes formula which describes the probability law governing the process of a logical inference. The analysis provided by Bayes formula is also called Bayesian analysis. It should be mentioned that Bayesian analysis became an important branch of statistics and is

fundamental for many contemporary applications such as statistical decision, detection and estimation, pattern recognition, and machine learning. Commonly used models build on the Bayesian analysis and examples may be found in the book of Gelman et al. (2003).

A combination of background processes and observations makes the Bayesian approach an attractive tool for DA. Lorenc (1986) demonstrated an applicability of the Bayesian analysis as a general framework. Indeed, even though some methods were initially derived from other ideas they can be equivalently formulated from the Bayes formula. A survey on existing Bayesian methods in DA is presented in Wikle and Berliner (2006).

The most rigorous algorithm of DA requires direct computations of Bayes formula. To this end, it is necessary to know the PDF $p(u)$ of the background process. The evolution of the PDF is described by a deterministic Fokker-Planck (FP) equation, also known as Forward Kolmogorov's or Fokker-Planck-Kolmogorov equation (Jazwinski, 1970). FP equation is a second order parabolic partial differential equation which is defined by the corresponding drift vector and diffusion tensor. It also must satisfy suitable initial and boundary conditions.

Once the PDF $p(u)$ is found, the Bayes formula can be directly computed and the optimal estimate of the system can be found as a maximum likelihood value of conditional probability $p(u|y)$. Mathematically this approach is very general as a solution of FP equation provides complete information about the state PDF. However, it comes with severe limitations on its computational cost especially for multi-dimensional problems.

The analytical solutions of FP equations can be found only for a small number of cases: if the drift vector is linear, and diffusion tensor is constant (Risken, 1989); if separation of variables is possible and time and space components of the state vector can be split (Lo and Shizgal, 2006; Risken, 1989); in a one-dimensional case if a normalization procedure is applied (Risken, 1989); a multi-dimensional case if the drift vector and the diffusion matrix obey certain potential conditions called detailed balance conditions (Risken, 1989); if diffusion is independent and the drift vector is derived from the potential operator of a certain form (Araujo and Filho, 2012; Denisov et al., 2009; Floris, 2013; Jordan et al., 1998).

Numerical solution of multi-dimensional FP equations is also problematic especially if number of dimensions is higher than three. Pichler et al. (2011) provided a detailed comparative study of several FEM and FDM schemas applied to FP equation in four dimensions. To avoid grid discretisation, Kumar et al. (2009) studied the partition of unity finite element method (PUFEM) applied to the FP equation. Martens and von Vagner (2011) presented a Galerkin

method that uses Hermite polynomials for the expansion of weighting functions. The numerical simulations of the method were done for five- and six-dimensional systems and described in details in Martens (2012).

While meshless or global Galerkin methods simplify discretisation of the space compared to grid methods, the number of basis functions needed still increases rapidly with the dimensionality. It was concluded in Khoromskij (2012), that an efficient approach to manage high dimensionality is to represent a model operator in a tensor product structure. Sun and Kumar (2014) and Sun and Kumar (2015) developed a numerical solver for the FP equation combining tensor decomposition approach with the Chebyshev spectral differentiation. That reduced the number of degrees of freedom and maintained the accuracy of the solution as dimensionality increased. Numerical results were provided for simulation with up to fourteen -dimensional state space systems. The most promising results with operator decomposition were published in a recent work of (Cho et al., 2016). They proposed several different classes of algorithms that reduce high-dimensional equations to a problem of calculation of a sequence of low-dimensional problems. The effectiveness of the algorithms is demonstrated for numerical examples involving 40-dimensional FP equations.

The main advantage of the proposed methods of decoupling or separation of the dimensions is that they have linear rather than exponential scalability with respect to the number of dimensions. However, it should be concluded, that even for moderate size systems, the solution of the corresponding FP equation is problematic.

Another approach for computation of the Bayesian formula, referred to as particle filters, is to approximate distributions of interest with Monte Carlo sampling instead of computing them directly. That is, the distributions are approximated by a large collection of N random samples termed particles sampled directly from the state space.

Since its introduction by Gordon et al. (1993), particle filters became an active research area and have also been applied in geosciences (van Leeuwen, 2009). Several different methods have been developed depending on sampling or resampling strategies such as sequential Monte Carlo (SMC) particle filter, sequential importance sampling (SIS) filter, bootstrap filter, auxiliary particle filter and many others have been proposed. A comprehensive review of particle filters and smoothers can be found in the exposition of Chen (2003) and references therein.

However, particle filters similar to FP solvers suffer from the "curse of dimensionality" problem (Chen, 2003). There were several attempts to resolve dimensionality limitation (Brun et al., 2002; Chitchian et al., 2012; van Leeuwen, 2010), but without significant success in the general case. In fact, scalability problem still remains and, as a result, reliable results of particle filtering can be obtained only for moderately sized systems (Rebechini and van Handel, 2015).

2.3 Statistical methods

The more practical representation of the uncertainties in the system can be done through the statistical description of the corresponding errors. Typically, it is assumed that a physical process is described by the true state which is not known. However, what is known is a background model of the process and observations of the process. The difference between the background and true state is called the background error which represents "modelling error" mentioned before. In a similar fashion, the difference between the observation state and true state is referred to as the observation error. The problem of DA is to reconstruct or estimate the analysis state so that the associated analysis error is as small as possible. In the literature, these kinds of algorithms are often termed as *estimators*.

2.3.1 Optimal estimators and Kalman filters

A general insight into classical estimators can be found in Lewis et al. (2008) where the mean-square, maximum-likelihood and recursive estimators are presented. Those are formulated for a general case of a random statistical variable. Their application requires knowledge of all statistical moments of that variable which may be a hard constraint in practice.

For linearized and Gaussian models several operational estimators for the atmospheric application have been proposed in Lorenc (1981); McPherson et al. (1979); Rutherford (1972). In the review of Bouttier and Courtier (1999), typical assumptions used to simplify the statistical description of errors are discussed. The authors assume the following hypotheses about the error and model structure:

1) Errors are Gaussian so only the first two statistical moments (mean and variance) are required to obtain a complete description of the corresponding statistical variables.
2) Errors are unbiased which means that their expectation (mean) is zero.
3) Errors are uncorrelated.

4) Background model and observation operator are linear.
5) Analysis defined by correction to the background depends linearly on observations.

Using those hypotheses, a statistical variable can be described only by the first two statistical moments: mean (average) and variance. So that, together with the linearity assumption, the application of least-squares method for the minimization of the analysis error allows retrieving of the algebraic expressions for those moments. These recursive expressions are known as the best linear unbiased estimator (BLUE).

A popular method of DA called optimal interpolation (OI) can be derived from BLUE if the background covariance matrix involved in calculations of the BLUE analysis is chosen under the assumption that for each model variable, only a few observations are necessary for determining the analysis increment (Ghil and Malanotte-Rizzoli, 1991). This assumption leads to an algebraic simplification of the BLUE analysis which makes OI very attractive from the computational point of view.

However, the estimation methods described so far minimize the estimation error variance based only on the observations at a current time step. The analysis can be extended to include the observations from the past if statistical time series are considered. Those ideas were initially studied in the monograph by Wiener (1949) for the application of antiaircraft fire control system. Wiener dealt with the continuous-time system and developed the so-called Wiener filter for the steady-state case.

In 1960, Kalman analysed a discrete in time and space system in a similar fashion to the BLUE estimator. He used the same hypothesis but extended the analysis error to include data from the past. As a result, he devised the celebrated recursive Kalman filter (Kalman, 1960).

In the subsequent work, Kalman and Bucy (1961) analysed the problem entirely in the continuous domain obtaining a formulation of the filter as a solution of the Wiener-Hopf equation. The heuristic argument about equivalence between the discrete and continuous filters was provided in Kalman (1963). Later, Smith and Roberts (1978) showed that if the definition of covariance of the discrete analysis error is derived from the accepted mathematical description for the covariance of the continuous analysis error, compatibility between the discrete and continuous filters is complete. In Shald (1999) the continuous Kalman filter was obtained as a limit of the sampling interval of the discrete Kalman filter.

It should be noted, that both BLUE and Kalman filter are optimal with respect to the minimization of the analysis error and obey the similar performance. However, from the

observations incorporation standpoint, the BLUE estimator should be regarded as a crude approximation of the Kalman filter estimation for one time step (Bertino et al., 2005).

There were many attempts to develop Kalman based algorithms for various configurations. Thus, it can be reformulated not only for the filtering but also for the smoothing problems (Fletcher, 2010). For nonlinear models, an extended Kalman filter (EKF) (Lewis et al., 2008) can be adopted. EKF is not optimal in general but provides an estimate which is the best linear estimate.

Without any exaggeration, it is fair to say that the Kalman filter (and its numerous variants) has dominated the adaptive filter theory. Auger et al. (2013) have summarized the research efforts made over the past decades about the applications and the implementation of the Kalman filter algorithm in a significant number of industrial engineering and scientific fields.

The difference between stochastic and statistic methods is that FP propagates all statistical moments of corresponding PDF, while Kalman filter propagates only first two statistical moments: mean and covariance. For a general case, when no assumptions about problem structure were made, the total number of operations for one step of Kalman filter was computed by (D'Amore et al., 2012) and is of the order $O(N^3)$, where N is the size of the background model. That estimate is mainly due to the need of the background covariance matrix inversion and it demonstrates that the Kalman filter algorithm is much more efficient than the stochastic methods from the computational standpoint. However, it also shows the lack of scalability and the fact that the computational requirements grow faster than the size of the state. That imposes an unacceptable computational burden especially for large-scale systems.

For the engineering applications, many authors suggested different approximations to the original Kalman filter. The main idea of those methods is to decrease heavy performance requirements for the covariance matrix computations without a significant loss of the accuracy. Because those methods provide equivalent to the Kalman filter results only in the limit, they are considered to be sub-optimal. Nevertheless, because of the structure of the model and observation error distributions the full representation in typical oceanographic practice is of no value (Cane et al., 1996).

An excellent review of the sub-optimal schemes (SOS) and their performance for the atmospheric/oceanographic DA problems was presented in Todling and Cohn (1994). The authors distinguished six major categories of those methods:

1) Covariance modelling.

2) Dynamic simplification.
3) Reduced resolution.
4) Local approximation.
5) Limiting filtering.
6) Monte Carlo methods.

Within each category, several approximations can be used; the division into groups is not rigid and some methods may include features from different categories. Briefly, those groups are described as follows:

1) The covariance modelling category includes schemes that assume a given form for the error covariance matrix with no background dynamics taken into account. An example of such approach is the BLUE analysis and the OI analysis which were presented in the above.
2) Dynamic simplification consists of methods that use approximate but nontrivial dynamics to evolve the forecast error covariance. One of the simplest methods of this kind is a time-invariant method called the constant gain Kalman filter. It has been used for the operational storm surge forecasting in Northern Sea (Heemink et al., 1995).
3) Very similar to dynamic simplification, reduced resolution schemes apply coarser-resolution grid for the propagation of the error covariance (Fukumori and Malanotte-Rizzoli, 1995).
4) Local approximation encompasses methods that attempt to evolve the error covariance structure only for points separated by reasonably small distances.
5) Limiting filtering schemes compute an asymptotic error covariance. For instance, Verlaan and Heemink (1997) proposed to use a reduced rank approximation of the error covariance using a square root factorization. The eigenvalue decomposition of the covariance matrix allows its projection and consequently an evolution in a sub-space of dominant eigenmodes while the analysis evolves in the full space of the model. In a similar fashion, Cane et al. (1996) reduced state space of the Kalman filter by a truncation of a set of multivariate empirical orthogonal functions derived from a long model run without assimilations. Both approaches were applied in real cases and became popular in the oceanographic community.
6) Monte-Carlo methods attempt to estimate the forecast error covariance matrix by propagating the covariance of an ensemble of states. Those schemes emerged into a

popular family of methods known as the Ensemble Kalman Filters and are discussed in more detail in the following paragraph.

The applications of sub-optimal algorithms for storm surge forecasting, two and three-dimensional coastal hydrodynamics, air pollution, ocean circulation models and others were reviewed and discussed in Bertino et al. (2005).

2.3.2 Ensemble Kalman Filter

Ensemble Kalman filter (EnKF) can be viewed as a combination of the classical Kalman filtering and Monte Carlo sampling. It uses the Gaussian assumptions about the background and analysis errors and follows the time evolution of the mean and covariance. However, instead of the covariance matrix, EnKF propagates an ensemble of states which is used to approximate the covariance matrix in the subspace spanned over ensemble members. The ensemble can be reasonably small relative to the other Monte Carlo methods because it is used only to parametrize the distribution, not to sample it thoroughly. At the same time, the ensemble should be large enough to approximately span the space of possible system states at a given time, because the analysis essentially determines which linear combination of the ensemble members constructs the best estimate of the current state, given the current observations (Hunt et al., 2007).

After the first EnKF was devised by Evensen (1994) for a quasi-geostrophic model, many studies have been published to advance it. van Leeuwen and Evensen (1996) extended EnKF to the smoothing problem, Evensen (1997) demonstrated it for the strongly nonlinear problem. Bergemann and Sebastian (2012) derived the analogue of the well- known Kalman-Bucy filter for linear differential equations subject to continuous observations for the ensemble setting and nonlinear differential equations. The results of the first decade of EnKF developments and its application can be reviewed in Evensen (2003).

The advantage of EnKF algorithm is its flexibility for applications on a complex physically based simulation models. However, the computation time of a data assimilation application using EnKF may be a challenge when the simulation has high computational demand. It is practical to set the ensemble size as small as possible to keep the computation time low. But, a naive implementation of the realistic system together with the small ensemble size may cause significant sampling errors in the estimation of the error covariance matrix. If the spread of the

ensemble systematically underestimates the true error of the process EnKF may diverge (Houtekamer and Mitchell, 1998).

To overcome these problems, the ensemble covariance matrix can be empirically updated. For instance, the ensemble perturbations can be inflated to avoid underestimation of the error variance (Anderson and Anderson, 1999). The off-diagonal components of the ensemble error covariance matrix may be reduced according to the distance they are from the diagonal terms (Houtekamer and Mitchell, 1998). Many studies have shown that EnKF is usually sensitive to the choice of the covariance inflation and localization methods and parameters. Thus, it is important to optimise them (Miyoshi, 2011).

The systematic theoretical analysis of EnKF algorithm was done in Kelly et al. (2014). The authors presented the well-posedness and accuracy of the method in a discrete and continuous time in particular with the small ensemble size limit and presence of the covariance inflation.

A popular variant of EnKF is the so-called localised ensemble transform Kalman filter (LETKF) described by Hunt et al. (2007) for a discrete problem and Amezcua et al. (2014) for continuous. An example of operational implementation of LETKF for National Center for Environmental Prediction can be found in Szunyogh et al. (2007). An LETKF application for the DA problem in the Chesapeake Bay and its performance examination were described in Hoffman et al. (2012). More examples of operational EnKF implementation were described in Evensen (2009); Miyoshi (2011); Tamura et al. (2014). For the description of the algorithms of the ensemble square root filter (EnSRF), the ensemble adjustment Kalman filter (EnAKF) and the singular interpolated Kalman filter (SEIK) that are suitable for non-linear and high-dimensional systems the reader is referred to the report by van Leeuwen et al. (2012) and referenced therein.

2.3.3 Variational methods

The idea of the two basic statistical methods the BLUE estimator and Kalman filter is to minimize the mean square analysis error. The minimum is found as an analytical expression in a recursive form. It is also possible to directly minimize an error function known as a cost function by applying numerical optimisation techniques. Typically, the solution is sought iteratively by performing several evaluations of the cost function and its gradient. The DA methods which minimize the cost functional directly are called variational.

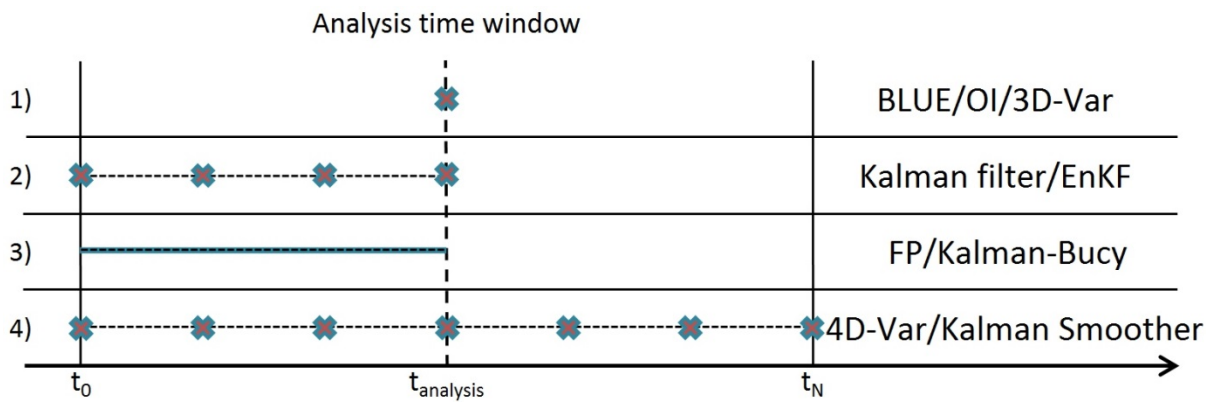

Figure 2.1 Comparison of different assimilation strategies for amount of observed data involved at analysis step.

A variational equivalent of the BLUE estimator is the three-dimensional variational analysis (3D-Var) (Courtier et al., 1998). The popularity of 3D-Var stems from its conceptual simplicity and from the ease with which complex observation operators can be used since only the problem operator and its adjoint needs to be provided. Another advantage is the ability to enforce external weak (or penalty) constraints, by putting additional terms into the cost function.

As the Kalman filter is a generalization of the BLUE estimator, in the same way the four-dimensional variational analysis (4D-Var) is a generalization of 3D-Var for observations that are distributed in time. The 4D-Var cost function is similar to the 3D-Var cost function, provided the observation operator is generalized to include a forecast model that will allow a comparison between the model state and the observations at the appropriate time. The special property of the 4D-Var analysis is that it uses all observations simultaneously, so the algorithm is regarded to be smoother (Talagrand and Courtier, 1987).

The comparison between different assimilation methods with regards to observations incorporated over time into the analysis is illustrated in Figure 2.1 where DA methods are split into four different categories:

1) The first category contains methods that incorporate observations from the current time instance such as BLUE, OI and 3D-Var.
2) Methods from the second category assimilate discrete observations from the past and up to the analysis time. Those include Kalman filter, its sub-optimal and distributed variants, EKF and ensemble filters such as EnKF and LETKF.
3) The third category defines assimilation methods which account for continuous observation from the past such as FP approach and Kalman-Bucy filter.

4) The fourth category is devoted to the methods that use discrete observations from future and past. It includes 4d-Var and Kalman smoother.

In the excellent review by Bouttier and Courtier (1999) the authors analysed 3D-Var, 4D-Var and Kalman filter estimates. When compared to the 3D-Var algorithm, 4D-Var was reported to have the following characteristics:

1) It requires perfect model assumptions, and the quality of the estimate degrades if model errors are significant.
2) Its implementation relies on the so-called adjoint operator which may be difficult to construct if the background model is complex.
3) Assimilation step is performed only when the observations from the whole reference time interval are available. In real-time systems, that will delay the availability of the analysis, whereas sequential 3D-Var models can process observations immediately after they became available (see Figure 2.1).
4) Background state is used as the initial state for the analysis, so by design 4D-Var is entirely consistent with the model equations and the four-dimensional distribution of observations until the end of assimilation interval.
5) 4D-Var is an optimal assimilation algorithm over the referenced time interval, moreover because all observations participate in the analysis it provides the smooth estimate.

Furthermore, Bouttier and Courtier (1999) defined relations between the 4D-Var and Kalman filter. The authors demonstrated that for a perfect model and the same input data, over a given interval the 4D-Var analysis is equivalent to the Kalman smoother analysis and at the end of the time interval is equal to the Kalman filter analysis at the end of the time interval. They also discussed fundamental differences between EKF and 4D-Var:

1) 4D-VAR is the preferred in realistic simulations because it is computationally less demanding than Kalman filter or EKF.
2) Because 4D-VAR is smoother, it proves more optimal than the (linear or extended) Kalman filter inside the time interval for assimilation.
3) Unlike 4D-VAR, EKF can handle model errors if they are Gaussian.
4) 4D-VAR can only be run for a finite time interval, especially if the background model is non-linear, whereas EKF can in principle be run forever.

Figure has been removed due to Copyright restrictions

Figure 2.2 Filtering versus smoothing for data assimilation. Recreated from Reichle (2008).

5) 4D-VAR itself does not provide an estimate of the analysis covariance, a specific procedure to estimate the quality of the analysis must be applied, which costs as much as running the equivalent EKF.

An illustrative example which explains the differences between filtering (3D-Var, Kalman filter) and smoothing (4D-Var) recreated from Reichle (2008) is presented in Figure 2.2.

The underlying cost function of 4D-Var analysis can be modified to encounter an imperfect model assumption by imposing additional weak constraints. Several approaches for the implementation of the weak-constraint 4D-Var were presented (Tremolet, 2006). In Courtier et al. (1994) two methods were considered for the reducing 4D-Var cost. The first method is preconditioning which speeds up minimization. And the second one is the incremental approximation of 4D-Var which is more cost-effective. Talagrand (2010) described minimization methods that can be adopted and discussed related numerical problems for practical use.

In operational weather forecasting, the variational methods rapidly gained popularity. First, 3D-Var replaced OI and later 4D-Var became the industry standard. An overview of developments in numerical weather prediction centres is presented by Rabier (2005).

Nowadays, the main competitor of the variational methods is ensemble filters. Several authors reported that ensemble data assimilation shows comparable results to the variational assimilation (Kalnay, 2010; Zupanski, 2009), and several studies were published to investigate the question. Lorenc (2003) examined the potential of various ensemble methods to replace variational in the numerical weather prediction. For the idealized Lorenc model a comparison of three methods: 4DVar using linear and adjoint models, the ensemble Kalman filter, and hybrid 4DEnVars, which is a variational method that uses a 4D ensemble covariance was done in Fairbairn et al. (2014). Overview of the operational oceanographic systems OcenView which used different variational and ensemble assimilation methods was presented in Martin et al. (2015). It is difficult to make a single choice about the best method. For different configuration they provide different results so that potential of hybrid method is high.

2.4 Optimal Control Methods

The estimation quality of the statistical methods of DA depends on the accuracy and certainty of the background and observation error description. Correct and precise definition of which, however, is generally an impossible task. For most applications, there must be concerns about the effects of parameter errors and variations in the accuracy of the estimation. The inaccuracy influence of input parameters on the estimate produced by the Kalman filter was studied, for instance, by Sage and White (1977), where the authors concluded that filtering errors may differ radically from their nominal values.

A common way to overcome Gaussian description of the uncertainties is to assume them *unknown but bounded* (UBB) and belongs to a given set of admissible errors. The latter assumption is very practical as it is more realistic and makes a data assimilation algorithm robust with respect to fluctuations in the second moments.

Figure has been removed due to Copyright restrictions

Figure 2.3 Schematic of error ellipsoid evolved in time. Recreated from Kurzhanski and Valyi (1997).

2.4.1 Minimax estimation

The main concepts and theoretical results regarding the estimation of the system with the UBB errors can be found in Milanese (1989). Several early stage algorithms of robust estimation were reviewed in Tempo and Vicino (1990).

To estimate the state of the system, one may define the worst-case mean-squared estimation error. The optimal state should minimize that error. An approach which looks for the estimate of the state as a minimum of the worst-case error is called the minimax estimator and if presented in the form of filter is referred to as the minimax filter.

A number of minimax filters were found in the literature, for instance in Krener and Kang (2011) and (Başar and Bernhard, 1991). The filter in the observer form based on the differential game theory was presented in Yaesh and Shaked (1992a) for continuous time case and in Yaesh and Shaked (1992b) for discrete time case. (Kurzhanski and Vályi, 1997) described calculus of generic bounded sets and their evolution in a dynamical system. They also presented results for a set defined as quadratic ellipsoids and devised the minimax filter for errors bounded in those ellipsoids. Implicitly-linearized minimax filter for nonlinear problems was proposed in Tchrakian et al. (2015). Zhuk (2013) devised the minimax filter for the system with background described by a DAE.

The relations between the Kalman filtering and minimax filtering or sometimes called Robust filtering have also been studied. Mintz (1972) derived the sufficient conditions for the Kalman filter to be a minimax estimator. The equivalence between the minimax filter with bounded in a quadratic ellipsoid errors and Kalman-Bucy filter was obtained in Krener (1980).

2.4.2 Solution of Riccati equation

The key ingredient and the main challenge of the continuous minimax DA approach (or equivalently Kalman-Bucy filter) is the solution of the so-called differential Riccati equation (DRE). The Riccati equation propagates error ellipsoids in time and represents an evolution of uncertainties in the system. Figure 2.3 recreated from (Kurzhanski and Vályi, 1997) provides an example of an error tubes evolution. The solution of the Riccati equation is further used for projection of the observation error in the feedback equation. RDE often arises in applications of a linear-quadratic control and is crucial for their performance.

The Riccati equation is a matrix differential equation of the second order. It is known, that the general analytical solution of this equation can be obtained if at least one particular solution is given (Reid, 1972) which, typically, is not possible in practice. That explains the necessity of numerical methods for the solution of DRE.

A straightforward numerical method is to directly discretise a time derivative and transform the equation into the non-linear Algebraic Riccati Equation (ARE). Dieci (1992) studied this approach for the symmetric, non-symmetric and stiff DRE in details. The author used backward differential formulas for the time discretisation and quasi-Newton iterations for the solution of the non-linear algebraic equation. A detailed review of iterative and Newton's methods applied to ARE with emphasis on numerical properties was published in Dieci (1991). Benner et al. (2008) reviewed several ARE solvers with a focus on large-scale equations.

The challenge in solving DRE is that for a stiff background model, associated Riccati equation may have singularities, and this is why the direct application of the standard approach may fail to integrate through those singularities (Schiff and Shnider, 1999). To overcome the above problem, the underlying DRE can be transformed applying Möbius scheme into a double size linear problem of the Hamilton structure. The authors explained why Möbius scheme passes singularities and analysed the solution with the Taylor series expansion. Frank and Zhuk (2014) extended that approach by adding reinitialisation procedure for the use of the symplectic integrators described in details in Hairer et al. (2006).

Benner and Mena (2004) analysed different numerical strategies of the solution of DRE for large-scale systems. The authors concluded that performance efficient methods for the solution of DRE remained to be a challenging problem for applications of the linear-quadratic control.

2.4.3 Summary

DA problems arise in many contemporary engineering and scientific applications. The well-known Fokker-Planck equations provide the exact description of uncertainties of Markovian processes and together with the Bayesian analysis yield general solutions of DA problems. However, this method requires extensive computational resources and efforts and becomes increasingly infeasible for the marine applications where dynamical systems are described by the high-dimensional state and should be analysed for long-term simulations. In fact, Fokker-Planck equations are mostly used only for theoretical purposes.

The statistical assimilation methods were designed to achieve better computational performance for simplified models of uncertainty propagation. They are similar to each other in that they involve restriction (or an approximation) on the underlying background model, observations and corresponding errors structure. That allows reconstructing system uncertainties using only its mean and covariance and simplifies analysis step. The robustness of assimilation approach is improved if system errors are considered to be deterministic and bounded in an ellipsoid. In this case, the assimilation problem is regarded as a minimization of the worst-case scenario of realisation of the analysis error. The solution of the minimization problem is known as the minimax filter which is a deterministic equivalent of the statistical Kalman-Bucy filter.

Nevertheless, as many authors reported, the traditional methods of DA suffer from the "curse of dimensionality" problem, are computationally expensive and not scalable. Development of performance efficient methods of DA in terms of quality estimation and computational demands is the main motivation of this research.

2.5 Domain decomposition methods

In practice, the demand for operationally efficient estimation algorithm escalates as the number of degrees of freedom of the discretised model is significantly increased (Majda et al., 2010). This is especially important for marine applications when the problem may be defined over the large domain such as a coastal zone or part of the ocean in 3D space.

In computational science a typical way for reducing computational demands is to apply some sort of problem decomposition. Decomposition methods generally refer to a splitting of a system state into smaller portions and performing computations independently on those

portions. Depending on the approach applied, this allows us to better exploit parallel nature of modern processors and reduce numerical complexity comparing to conventional algorithms.

The term “problem decomposition” can be used in context of several numerical strategies. The first strategy also known as the *data decomposition* is used for parallelisation of conventional algorithms. It is a computational approach which splits system data into chunks and assigns them to separate computational units. Then the same computations are performed on separate processors using one of the following inter processing data exchange programming paradigm: shared memory (the most common implementation is OpenMP) or distributed memory (the most common implementation is OpenMPI). Normally, data decomposition does not reduce numerical complexity of an algorithm and may even introduce additional overhead related with a data exchange, however, it improves parallelisation and scalability properties of an algorithm.

Due to a large scale nature of the marine problems, data decomposition emerged as a popular strategy. Examples of the corresponding algorithms were, for instance, presented in Drake et al. (1993) for the parallel community climate model and in (O’Donncha et al., 2014) for the parallel environmental fluid dynamics code.

The second type of problem decomposition strategies applies decomposition to mathematically reformulate a problem as a set of distributed but coupled lower order subproblems which are expected to be cheaper to compute. In the following those methods are referred to as DD methods. For a problem described by a PDE, DD produces a set of coupled local PDEs on smaller subdomains forming a partition of the original domain. A coupling between subproblems called continuity or transmission conditions can be defined based on physical properties of a process in question and is normally expressed in a form of boundary conditions over the interface of decomposition. This makes local subproblems completely independent and allows solving them separately. Similarly to data decomposition, mathematical DD allows parallel implementation, however it additionally reduces numerical cost of an algorithm comparing to a non-decomposed version.

The domain splitting is the first step of any DD approach and is usually straightforward. It can be done in two possible ways: overlapping and non-overlapping decomposition. The non-overlapping decomposition is preferred for its simplicity, but the overlapping decomposition may be required by the design of the method. In its turn, domains coupling is more difficult. There are various strategies that can be applied to different problems and particular conditions. For the non-stationary problems they can be categorized as follows:

1) Iterative methods or also called Schwarz methods. Those are derived from Schwarz's idea of DD.
2) Time fraction methods which are based on Dawson's ideas (Dawson and Dupont, 1992).
3) Methods that are based on the solution of DAE.

DD became a common tool in computational engineering for solving problems governed by PDE (Dolean et al., 2015). In this research, the application of DD to DA problems is investigated.

2.5.1 Schwarz iterative Domain Decomposition

In the literature, the first procedure of the DD for the solution of PDE was introduced by Schwarz (1870). It was a technique to prove the existence of a solution to the Laplace equation on a domain which is a combination of a rectangle and a circle. Schwarz proposed an iterative method that successively solved a problem in subdomains, going alternatively from one to the other. That approach is known as the Schwarz alternating method.

For a long time, the Schwarz method was used only as a theoretical concept. Nevertheless, a development of computational resources and the need for more efficient numerical algorithms attracted research interest to the Schwarz methodology. That interest became especially strong with the advance in parallel computing technologies.

A breakthrough in the development of DD strategies happened at the end of 1980s after the publication of Lions (1988). In that work, the author extended the alternating Schwarz algorithm to be parallel and applied it to the solution of the boundary value problem. Lions decomposed problem's domain into several overlapping subdomains. Then, local subproblems were coupled through Dirichlet transmission conditions in an iterative manner. Starting from some initial value on both subproblems, he calculated in parallel improved Dirichlet data and used it to update the adjacent subproblem. This process was repeated several times until the jump of the solution between subdomains declined. In the following works Lions (1989) continued the study of the method, introduced a convergence proof and investigated other properties of the approach.

After Lions' pioneering work many Schwarz based methods appeared for the elliptic equations. Gander (2008) provided a historical overview of the DD strategies with additive and multiplicative formulations. Toselli and Widlund (2005) presented Dirichlet-Dirichlet,

Dirichlet-Neumann and Neumann-Neumann variants of the Schwarz approach together with various preconditioners for the resulting decomposed system.

At the same time, Lions (1990) also showed that the convergence of the classical Schwarz method depends on the size of an overlap. Increasing overlapping area for a better convergence may be a severe limitation on the applicability of the method. To avoid this restriction, it was suggested to use a convex combination of the Dirichlet and Neumann conditions for passing data through the common interface. Lions discovered that for that combination the convergence does not depend on the overlap anymore. Additionally, in Gander et al. (2001) it was shown that the Dirichlet conditions at the artificial interfaces inhibit the information exchange between subdomains and therefore slow down the convergence of the algorithm. In order to improve the performance of the classical Schwarz method, they also replaced the Dirichlet conditions by the Robin conditions. This algorithm was called optimised Schwarz method, since it contains scalar parameters in the transmission conditions which should be optimised to get a better convergence.

Unlike the classical Schwarz methods, the optimised ones are possible for both overlapping and non-overlapping decomposition. Even more, usually it requires fewer iterations, and it was demonstrated that for some problems using a particular form of the Robin conditions the method converges after a finite number of iterations. For instance, Gander et al. (2003) demonstrated the convergence of the Laplace and Helmholtz equations in two iterations.

For the non-stationary problems, Schwarz waveform relaxation was applied to the parallel in time solution (Gander et al., 1999). Later, different convergence studies were presented: for linear advection-diffusion-reaction problems (Gander and Halpern, 2007), for its semi-linear counterpart with overlapping decomposition (Descombes et al., 2011) and non-overlapping decomposition (Caetano et al., 2010). The decomposition of the both space and time domains was suggested and analysed by Giladi and Keller (2002) and Gander and Stuart (1998).

Although many techniques have been developed to study the convergence of the classical and optimised Schwarz methods, the problem with nonlinear equations for multi-dimensional space and general shape multi-subdomains is still open (Tran, 2011).

2.5.2 Adaptive Domain Decomposition

The particular interest of this research is a solution of the transport phenomena in the marine environment described by the linear advection-diffusion equation. Thus, a proper choice of the decomposition strategy is essential for the performance of the method.

Depending on the nature of the problem, the Schwarz methods applied to advection-diffusion problems can be grouped into two categories (Quarteroni and Valli, 1999). The Dirichlet-Neumann techniques provide a good approximation when the diffusive part is dominant, but may produce discrepancies if the advective part becomes more relevant. The natural interface conditions techniques provide a good approximation when the advection part is dominant, enforcing boundary conditions over the interface taking into account the hyperbolic nature of the problem, but may produce discrepancies when the diffusive part of the problem is dominant. Based on the physical properties of the underlying transport problem it is reasonable to consider the second group of methods which is known in the literature as the adaptive Dirichlet-Neumann DD.

One of the first works that can be viewed in the framework of adaptive methods was presented by Gastaldi and Gastaldi (1993). The authors described DD method applied to the two-dimensional non-stationary advection equation which is of the hyperbolic type. Decomposition of the domain was introduced with respect to the inflow/outflow areas of problem's velocity field over the interface. Following that, the family of coupled problems is reformulated in terms of functional equations as a boundary mapping of Dirichlet inflow data into outflow portion which essentially leads to the Steklov-Poincare operator. Those equations were suggested to be solved in an iterative way. The authors further demonstrated the contraction property of the Steklov-Poincare operator that proved the convergence of the iterative algorithm for both continuous and discrete problems. They also showed the equivalence of the iterative algorithm and iteration-by-subdomain strategy (Schwarz approach). In the following paper (Gastaldi, 1994), numerical experiments demonstrated the efficiency of the proposed method. It should be noted, at that stage, the authors did not provide a specific name for their approach.

The terminology of the adaptive methods was introduced in Carlenzoli and Quarteroni (1995). It was a parallel study to the previous one in which, however, the authors aimed to build a decomposition strategy for the advection-diffusion problem with the advection dominant part. In that work, the one-dimensional stationary advection-diffusion problem was analysed by partitioning in two non-overlapping subdomains. The study of the convergence of the classical Dirichlet-Neumann and Neumann-Neumann methods revealed that the convergence rate

deteriorates for the advection dominated problem, i.e., problem with the high Peclet number. Moreover, the convergence was shown to depend on the partition point and value of the Peclet number. The authors concluded that the boundary conditions over the interface introduced by the classical scheme were inconsistent with the hyperbolic limit of the problem.

To improve convergence, they described a decomposition which took into account the direction of the characteristic lines. To this end, the interface of the decomposition was split into inflow and outflow zones with the relaxed Dirichlet or Robin conditions on the inflow and relaxed Neumann conditions on the outflow. The derived strategies were called adaptive Dirichlet-Neumann (ADN) and adaptive Robin-Neumann (ARN) decomposition respectively. Again, the set of coupled local problems was solved in an iterative way.

The proofs of the convergence of both strategies were done for a simple geometry problem. The authors also estimated the admissible intervals for the relaxation parameters and concluded that ARN is more stable than ADN for a problem with small diffusion, but performs poorly when diffusion is significant, whereas the AND keeps its properties of convergence.

The numerical aspects of ADN and ARN were further investigated in Carlenzoli and Quarteroni (1995) where those methods were applied to the incompressible Navier-Stokes equations with high Peclet number. The authors also discussed common numerical problems such as the so-called cross point situation when one point of the grid belongs to more than two subdomains; when the advection vector field forms different angles of incidence with the interface; the choice of relaxation parameter.

Trotta (1996) extended adaptive methods by introducing the damped version of the ADN and ARN algorithms. He investigated the influence of damping on convergence and stability. However, most of the explanations were heuristic and based on the results of numerical simulations without proper mathematical justifications.

Study of the parallel version of the adaptive methods with applications to non-stationary problems can be found in Ciccoli (1996). In the paper, the author presented the multiplicative and additive versions of the algorithms. He provided the continuous analysis of the damped methods and demonstrated that the damping introduces discontinuity between subdomains. The level of the discontinuity was further shown to be acceptable with the approximation error for a small diffusion coefficient. At the same time, the author stated that the damping improved the convergence rate which was illustrated by numerical experiments. Additionally, the

sensitivity of the convergence on the mesh size of the decomposition and diffusion coefficient was considered as well.

Rigorous theoretical study of the adaptive decomposition strategies was published by Gastaldi et al. (1996) and Gastaldi et al. (1998). Using the methodology of the Fourier transform the authors studied different configurations of algorithms for one and two-dimensional problems. The main conclusion of the works was that for a finite-dimensional problem (discrete case) both ADN and ARN enjoy good convergence rates for a reasonable choice of the initial guess. The convergence was additionally improved for small diffusion if damped versions (d-ADN, d-ARN) were applied.

2.5.3 Time fraction Domain Decomposition

The iterative DD methods can be directly applied to the elliptic problems. The dynamic problems additionally require implicit semi-discretisation in time, yielding an elliptic problem at each time step. Then, the Schwarz methods are applied to reconstruct boundary data on the interface. However, the Schwarz methods work in space coordinates only and ignore the dynamic nature of the problem.

Dawson and Dupont (1992) proposed the algorithm that takes the advantage of the information gained in the previous time steps. This method is known as the explicit/implicit domain decomposition (EIDD). Based on an implicit discretisation in time, it makes use of the results of the previous time steps to predict the values on the interface at the current time step. Such an approach reduces the computation and communication costs because it is non-iterative. Instead, it raises the question of the accuracy and stability as it uses lagged results from the previous time steps, i.e., an explicit treatment, in contrast to spatial iterations that predict the interface values.

The analysis revealed that EIDD methods are only conditionally stable while Schwarz methods could maintain the good stability condition of implicit temporal discretisation schemes. In order to improve the stability, different stabilization techniques can be used. Detailed review and discussion of the stabilization mechanisms were done in Zhuang and Sun (2002). As a further improvement, Liao et al. (2009) presented a new class of corrected explicit-implicit DD (CEIDD) methods.

There are other time fraction methods not of EIDD type that experience good stability conditions. For instance, explicit predictor implicit corrector (EPIC) methods (Jun and Mai,

2006), hybrid predictor-corrector methods (Lapin and Pieska, 2002) or methods proposed by Dryja and Tu (2007).

It should be noted though, that time fraction methods are much less studied and applied in practice than the Schwarz one. Moreover, their convergence and stability properties depend on the size of the time discretisation increment. That means that benefits of introducing DD may be outweighed by the need of the decreasing of the time step or loss of stability. At the same time, the Schwarz methods are stable and for the optimised or adaptive variants, the convergence depends on the physical properties of the model and is independent of time-space discretisation.

2.5.4 DAE based Domain Decomposition

For the elliptic problems an alternative approach to Schwarz is the so-called Tearing and Interconnecting (TI) methods such as FETI, or FETI-DP (Farhat et al., 1999). The main idea of the TI methods is to represent the interface continuity of the solution between subdomains (transmission conditions) through an additional algebraic constraint; later, Lagrange multipliers enforce the algebraic constraint.

There were a few attempts to adapt the same approach for the dynamic problems. However, in this case, the algebraic constraint should be incorporated into the differential equation that leads to a system of DAE. DAE problems can be solved only in particular cases and do not have a general attractive solver (Hairer and Wanner, 1996).

Nakshatrala et al. (2008) manipulated the Schur complement of the algebraic operator to solve DAE. Zheng et al. (2008) proposed to use the explicit Runge-Kutta-Chebyshev (RKC) projection method (Zheng and Petzold, 2006) for the solution of the governing DAE system derived for Navier-Stokes problem. The authors analysed the stability of the method and showed that it performs best for moderately stiff diffusion-dominated problems. For the convection-dominated system, it was suggested to use a semi-Lagrangian formulation with the RKC method to form a finite element semi-Lagrangian explicit RKC method.

2.6 Combination of domain decomposition and data assimilation

Several approaches of problem decomposition have been used in a development of efficient DA solvers. Due to the large scale nature of marine problems, data decomposition emerged as

a popular strategy exploiting parallel computational resources advantageous. A parallel version of the Kalman filter was described in Lyster et al. (1997) for the problem of tracer transport in the atmosphere. The authors suggested two algorithms of data decomposition. The first algorithm, referred to as the operator decomposition, partitions state vector and columns of covariance matrix and distributes them among processors. The second algorithm distributes a whole column of covariance matrix among processors and acts as a model matrix on a vector, but not a matrix. This approach is called *covariance decomposition*.

Parallelization of ensemble filters using operator decomposition was, for instance presented in Keppenne (2000) where a parallel variant of multivariate ensemble Kalman filter (MvEnKF) was described for a spectral shallow water model. Parallel sequential ensemble Kalman filter for atmospheric data assimilation was developed in Houtekamer and Mitchell (2001).

Development of decomposition methods for DA problems produced another type of DD method called *discrete decomposition* that are commonly applied to filters. It is similar to data decomposition in that it is applied on a discrete level by splitting system vectors and matrices and similar to continuous DD in that it reformulates conventional filters as a set of coupled distributed lower order filters which are cheaper to compute and maintain similar level of accuracy. Discrete and domain decomposition of DA problem additionally requires restriction of error functions and corresponding error descriptions.

Early evidence of discrete DD approach can be found in van der Vooren (1990) in which the authors applied decomposition with a multi-level algorithm for the coordination of analysis steps between local Kalman filters while the propagation step was not modified. The approach was tested for a one-dimensional shallow water model. It was reported that the proposed filter was computed faster than conventional Kalman filters with a similar level of accuracy. In the literature those methods were referred to as distributed Kalman filters. Several variants of decomposition or distribution of Kalman filters have been developed. Those approaches are built on a discrete level and ignore the continuous nature of the background problem defined as a PDE. Also, they implement algebraic transformations for problem distributions which require a specific problem structure and may suffer from scalability issues (Wang et al., 2015). For a comprehensive literature review of those methods, readers are referred to the bibliographic survey by (Mahmoud and Khalid, 2013).

Discrete decomposition became a common tool also for ensemble filters. To achieve better computational efficiency of ensemble filters compared to Kalman filters, the size of an

ensemble is normally taken to be small. Since an ensemble is used to approximate correlations between state components, small amount of ensemble members may produce difficulties in estimation of correlations between distant observations points. As a consequence, the analyses may diverge from the actual system state (Houtekamer and Mitchell, 1998). On the other hand, an ensemble defines a subspace which approximates system uncertainties. If the size of an ensemble is small, the forecast errors may grow in directions not accounted by the subspace (Hunt et al., 2007).

In the literature (Brusdal et al., 2003; Ott et al., 2002; Penduff et al., 2002), it is suggested that the problem may be overcome by the analysis localisation procedure which is analogue to discrete decomposition methods. The idea of localization is to split the system into smaller subsystems and compute analyses on each subsystem locally. This allows an approximation of the global uncertainty space in a piecewise manner by a set of low-order ensembles.

Ott et al. (2004) presented and studied an algorithm of local ensemble Kalman filter (LEKF) for atmospheric data assimilation. The authors suggested that for each analysis point only observations inside a circle with a certain radius, referred to as an influence radius, are included in the analysis, assuming observations from outside of the circle have no impact on analysis. It was concluded that since global analysis is replaced by a number of smaller size analyses, LEKF demands lower computational cost compared to the full size EnKF at the same time providing similar accuracy estimates.

Hunt et al. (2007) presented a local version of the ensemble transform Kalman filter (LETKF) by splitting observations into separate batches that have uncorrelated errors. Since the localization may produce discontinuities between local analyses, the authors suggested reducing the impact of observations close to the boundary of the local domains by artificially increasing its error variance. An implementation of LETKF for a Global Forecast System is described in details in Szunyogh et al. (2007). An alternative approach that address the problem of smoothness is introduced in Janjic et al. (2011) where, for each local domain, Schur product localization of analysis error covariance matrix is used.

An attractive property of the domain localisation of ensemble filters is that local analysis is performed independently and thus can be easily computed in parallel. Development of the parallel data assimilation framework PDAF that supports combination of several versions of local ensemble filters with existing numerical models is done in Nerger and Hiller (2013). Nino et al. (2016) presented a parallel implementation of LEKF based on Cholesky decomposition

for the Atmospheric General Circulation Model. Nevertheless, it should be stressed that the approaches presented in Hunt et al. (2007), Ott et al. (2004) and others are concerned only with the analysis step and ignore the background propagation step which results in lack of scalability of those algorithms.

There are examples of application of DD to optimal control problems. Benamou (1999) presented an approach for the solution of the quadratic control system governed by the wave equation. The authors applied DD to the optimality system composed of direct and adjoint state equations and recast it as a set of local optimality systems which are further suggested to solve using synthesis method. Another example is given in Heinkenschloss and Herty (2007) where LQ optimal control problem is considered for a parabolic PDE. In this case, DD is applied to the optimality system using Schur complement formulation obtained after applying suitable reordering of rows and columns of a space discretised problem. In both cases it was reported that the application of DD resulted in significant computational improvements. These examples inspired further development of DD methodology for combining it with the minimax filter that is constructed using LQ control theory.

Combining DD and filtering is less investigated then other decomposition strategies. Fujimoto and Kawahara (2001) presented an algorithm of domain decomposition based on Schwarz alternating approach for a shallow water model and Kalman filter. The authors reported good convergence using splitting in two subdomains while usage of higher number of subdomains resulted in divergence of the algorithm. In a recent publication by Battistelli et al. (2015), the DD has been applied for a consensus Kalman filter and investigated for a process described by a diffusion equation. In contrast to distributed Kalman filters DD applies decomposition to a continuous space domain. The resulting localised problem is resolved using an approach based on splitting the time intervals into substeps and propagating both state estimate and state covariance over those substeps. The stability and convergence properties of the consensus Kalman filter are analysed in detail in Battistelli et al., (2016). It should be noted that this algorithm demonstrates significant computational reduction comparing to global Kalman filter as well as good scalability properties. A version of this approach adapted to the contaminant transport problem will be later used in the thesis for a comparison with the methods proposed herein.

There are methods of problem decomposition that have been used in combination with offline assimilation methods. Several strategies of data decomposition for an implementation of a parallel 4d-var algorithm are described in Rantakokko (1997). In (D'Amore et al., 2013) and

(D'Amore et al., 2014) the authors constructed localised version of 3D-Var method by applying state decomposition on a discrete level. To this end, a cost function was split between subdomain and a coupling between subdomains required by a decomposition strategy was enforced through an additional constraint in each local cost function. The assimilation method obtained was applied to Mediterranean Forecasting System and significant improvements of computational time comparing to the conventional 3D-Var were reported.

2.7 Discussion

The development of performance efficient methods of DA in terms of quality estimation and computational demands is an active research area. Modern methods used in the operational DA represent a trade-off between computational cost and efficiency. To reduce the computational requirements, especially for large scale problems such as those occurring in marine modeling, many strategies may be applied. For instance, instead of the traditional Kalman filter, it is common in engineering practice to use various sub-optimal schemes of the Kalman filter. The computational cost of variational DA methods such as 3D-Var or 4D-Var and computational cost of ensemble filters may be controlled by limiting the amount of minimization iterations performed or by restricting an ensemble size correspondingly.

More recently, several problem decomposition approaches were developed to overcome computational burden of conventional DA methods. Data decomposition applied to conventional filters allows performing computations in parallel more efficiently but it does not decrease numerical complexity of the underlying method.

Discrete decomposition approaches such as distributed Kalman filters or localised ensemble Kalman filters depend on a particular structure of underlying system matrices and in a general case may lose their advantageous. Discrete decomposition is also limited to grid methods such as FEM of FD where system vector components are geometrically associated to a point. If, for instance, discrete system is obtained using approximation by functions with global support as in case of spectral methods, discrete decomposition is not possible. Finally, those methods are specifically designed for decomposition of an analysis equation and cannot be applied to a background propagation, therefore additional decomposition of background model is essential. Although many efforts have been dedicated to overcoming computational burden of DA problems, the problem remains open.

At the same time, DD technique emerged as a popular tool for computational price reduction. It was successfully applied to the solution of many PDE reducing numerical cost of an algorithm and providing nearly linear scalability of performance. In the literature, application of DD to the solution of DA problems in particular for transport problems is not developed enough. This provides the second motivation of this research: to construct efficient and scalable DA strategies for contaminant transport problems based on combinations of conventional filters with DD.

3 Data Assimilation methods

This chapter presents a framework that is used as a foundation for a development of DA methods combined with DD. While the intention of this development is to make the framework as general as possible, it will be adopted in particular for settings that described physical processes of pollutant transport in the marine environment.

The main components of this framework include linear PDEs that describe an underlying processes and traditional Kalman and minimax filters. Those filters are demonstrated for a problem continuous in time and space problem which permits a combination with DD that could be further introduced on a continuous level. A transformation or discretisation of continuous filters into discrete filters is provided is this chapter as well and is used later in the thesis for a development of several localised filters. Space discretisation is performed using Galerkin approximations. The later makes the framework suitable for various numerical methods, in particular, FEM with bilinear basis function is used in this work.

3.1 Abstract assimilation problem

From a philosophical point of view, a DA problem can be considered as a differential game involving two opposing parties. The first party is nature, which drives a physical process referred to as the true process, and state, which is denoted by $u^t(t,x)$. The second party is an engineer who aims to construct an analysis state $u^a(t,x)$ that would be as close to the true state as possible.

In order to compute the best estimate, an engineer should rely on some prior information regarding the problem and nature behaviour. Typically this information includes:

1) The time interval $[t_0; t_1]$ over which the estimation should be performed.
2) The domain Ω which is a predefined area of the space where the physical process is happening.
3) The background model that is described by the dynamic system (parabolic or hyperbolic) of PDE:

$$\frac{\partial u(t,x)}{\partial t} = L\big(u(t,x)\big) + e_m(L,t,x) \tag{3.1}$$

where $u(t,x)$ describes the state of the system at time $t \in [t_0; t_1]$ and space $x \in \Omega$; L is the known model operator, $e_m(L,t,x)$ is the model error presented as an additive term and may depend on the model operator. It should be mentioned that if $e_m(L,t,x) = 0$ then the underlying equation is deterministic and its solution is referred to as a background and denoted by $u^b(t,x)$.

4) The initial and boundary conditions that are applied at the initial moment t_0 and over the boundaries of the domain Ω.

5) Observations: the engineer can empirically measure data related to the state of the system. The observations are typically presented in the following way

$$y(t,x) = Hu(t,x) + e_o(t,x) \tag{3.2}$$

where $y(t,x)$ is the observation variable, $e_o(t,x)$ is the error of the observation process and H is a linear observation operator defined by the given kernel function $h(z)$ as follows

$$Hu(t,x) = \int_\Omega h(y-x)u(t,y)dy \tag{3.3}$$

The main challenge for an engineer is to find an optimal combination of a model and observations taking into account model and observation errors controlled by nature which are not deterministic. The policy of nature is not known but an engineer may collect statistical information about those errors such as average, variance and statistical moments. In the limit of a large number of the realisations, it is expected that statistics would converge to the actual probability distributions that depend only on the physical process responsible for the errors.

There are several strategies of the formal description of the model and observation errors. The stochastic approach which is known to provide the most general approach for uncertainty quantification is addressed in the next section.

Reliable methods of data assimilation are devised if several conditions of problem specifications are met. The first group of those conditions is related to the model configuration; the analysis is greatly simplified if the model is assumed to be linear. Second, the continuous

problem may be discretised first, and then an assimilation algorithm is applied to its discrete counterpart.

The second group of conditions is related to the description of the errors. The Gaussian function is an example of a structure-specific distribution that is widely used in practice. If a PDF is Gaussian, its shape is defined only by the first two statistical moments: mean and variance. The following section presents data assimilation methods that assume the Gaussian structure of the underlying process and optimise its mean and variance. The analysis is presented for both discrete and continuous cases.

3.1.1 Problems continuous in time and space

This research concerns the physical process of transport of a pollutant in the marine environment. Let the problem be defined on a spatiotemporal domain $([t_0; t_1] \times \Omega)$ where $\Omega \subset \mathbb{R}^2$ is a bounded domain with Lipschitz boundary. The differential operator L that describes the transport phenomena is the advection-diffusion operator defined as (Donea and Huerta, 2003)

$$\begin{gathered} L: H^1([t_0; t_1] \times \Omega) \to L^2([t_0; t_1] \times \Omega) \\ u \mapsto Lu = \epsilon \Delta u - \nabla \cdot (\mu u) \end{gathered} \tag{3.4}$$

where $\epsilon \geq 0$ is a constant diffusion coefficient and vector $\mu = [\mu_1, \mu_2]^T$ is a vector of velocity field. Hence, the background PDE (3.1) is rewritten in the form of the linear dynamic equation

$$\begin{cases} \dfrac{\partial u(t,x)}{\partial t} = \epsilon \Delta u(t,x) - \nabla \cdot \big(\mu(t,x) u(t,x)\big) + e_m(t,x) \\ u(0,x) = u^0(x) + e_0(x) \\ u(t,x) = 0, \qquad x \in \partial\Omega \end{cases} \tag{3.5}$$

where $u_0(x)$ is the initial condition reflecting the initial concentration. The homogeneous boundary condition for u reflects the following assumption: the domain Ω is chosen to be large enough in order to guarantee that no concentration is entering the domain.

Observations of the state are also described by the following linear equation

$$y(t,x) = \int_\Omega h(z - x) u(t,y) dz + e_o(t,x) \tag{3.6}$$

where $h(z)$ is a given function, which in this work is chosen as Dirac's delta function leading to the so-called direct observations equation.

For the description of uncertainties in the system (3.5)-(3.6), Kalman (1960) suggested to accept the following assumptions:

1) The model, observation and initial errors $e_m(t,x)$, $e_o(t,x)$, $e_0(x)$ are considered to be normal. That is, their probability distributions are described by the multidimensional Gaussian function:

$$p(e(t,x)) = \frac{1}{\sqrt{2\pi Q(t,t,x,x)}} \exp\left(-\frac{\left(e(t,x) - \overline{e}(t,x)\right)^2}{2Q(t,t,x,x)}\right) \tag{3.7}$$

where $\overline{e}(t,x)$ is mean or average or expected value of a stochastic variable $e(t,x)$ and its covariance is given by

$$\begin{aligned} &\mathrm{cov}\big(e(t_1,x_1), e(t_2,x_2)\big) \\ &\qquad = \overline{\big(e(t_1,x_1) - \overline{e}(t_1,x_1)\big)\big(e(t_2,x_2) - \overline{e}(t_2,x_2)\big)} \\ &\qquad = Q(t_1,t_2,x_1,x_2)\delta(t_1 - t_2) \end{aligned} \tag{3.8}$$

where $Q(t_1,t_2,x_1,x_2)$ is a positive continuous function and $\delta(\cdot)$ is Dirac's delta function.

In a shorter form the fact that variable e is Gaussian distributed is denoted as follows:

$$p(e) = N(\mu, Q) \text{ or } e \sim N(\mu, Q) \tag{3.9}$$

2) The model, observation and initial errors $e_m(t,x)$, $e_o(t,x)$, $e_0(x)$ are unbiased, i.e., have zero expectations and their covariance are defined by positive functions $Q(t,x)$, $R(t,x)$, $Q_0(x)$ respectively. Hence

$$\begin{aligned} e_m(t,x) &\sim N\big(0, Q(t,t,x,x)\big); \\ e_o(t,x) &\sim N\big(0, R(t,t,x,x)\big); \\ e_0(x) &\sim N(0, Q_0(x,x)); \end{aligned} \tag{3.10}$$

3) $e_m(t,x)$, $e_o(t,x)$, $e_0(x)$ are considered to be mutually uncorrelated:

$$\mathrm{cov}(e_m, e_o) = 0; \mathrm{cov}(e_m, e_0) = 0;\ \mathrm{cov}(e_o, e_0) = 0 \tag{3.11}$$

If a stochastic variable satisfies the above conditions, it is called white Gaussian noise. The benefit of imposing those conditions to reconstruct PDFs of the model or observation errors

that are white Gaussian noise is that one needs to reconstruct only first two statistical moments of the underlying distribution: mean and covariance.

The problem defined by (3.5)-(3.6) together with the error description (3.10) is continuous in both time and space. In the following, the corresponding problem of its state estimation is referred as a time-space continuous DA problem.

3.1.2 Variational formulation of the problem

Considering the difficulty of finding an analytical solution for the PDE, the problem (3.5)-(3.6) is reformulated to be feasible for numerical approximations (Evans, 1998). A common approach is to look for the weak solution of the underlying differential equation (3.5) following the Galerkin method. For this reason, the fixed-in-time test function $v(x)$ is introduced, chosen from the Sobolev space $H^1(\Omega)$. The PDE (3.5) is then multiplied by that test function and the problem becomes reformulated as a variational problem:

$$\begin{cases} \left(v, \dfrac{\partial u}{\partial t}\right) = (v, \epsilon \Delta u) - \big(v, \nabla \cdot (\mu u)\big) + (v, e_m) \\ \big(v, u(0,x)\big) = \big(v, u^0(x)\big) + \big(v, e_0(x)\big) \end{cases} \tag{3.12}$$

where $(\cdot,\cdot)$ defines the scalar product in $L_2(\Omega)$:

$$(v, u) = \int_\Omega vu dx \tag{3.13}$$

The solution of the problem (3.12) is called a weak solution. Due to the properties of the linear advection-diffusion operator, the weak solution of (3.12) satisfies the original continuous problem (3.6) in a weak sense (Evans, 1998).

The variational formulation (3.12) is further transformed by the application of the divergence theorem (Evans, 1998) to the integrals of (3.12). In that case, the diffusive term is transformed as

$$\begin{aligned} (v, \epsilon \Delta u) &= \epsilon \int_\Omega v(x) \Delta u(t,x) dx \\ &= -\int_\Omega \epsilon \nabla^{\mathrm{T}} v(t,x) \nabla u(t,x) dx + \int_{\partial\Omega} \epsilon v(x) \frac{\partial u(t,x)}{\partial n} dx \\ &= -\epsilon(\nabla v, \nabla u) + \epsilon \left(v, \frac{\partial u}{\partial n}\right)_{\partial\Omega} \end{aligned} \tag{3.14}$$

where $n(x)$ is the outward normal vector to the surface at point x.

Moreover, applying the divergence theorem to the advective term yields

$$
\begin{aligned}
(v, \mu \cdot \nabla u) &= \int_{\Omega} v(x) \Big(\nabla \cdot \big(\mu(t,x) u(t,x) \big) \Big) dx \\
&= -\int_{\Omega} \big(\nabla v(t,x) \big) \cdot \mu(t,x) u(t,x) dx \\
&\qquad + \int_{\partial\Omega} \big(\mu(t,x) \cdot n(x) \big) v(x) u(t,x) dx \\
&= -(\mu \cdot \nabla v, u) + \big((\mu \cdot n) v, u \big)_{\partial\Omega}
\end{aligned}
\tag{3.15}
$$

Substituting (3.14) and (3.15) into the equation (3.12) returns

$$
\begin{cases}
\left(v, \dfrac{\partial u}{\partial t} \right) = -\epsilon (\nabla v, \nabla u) + (\mu \cdot \nabla v, u) \\
\qquad\qquad + \epsilon \left(v, \dfrac{\partial u}{\partial n} \right)_{\partial\Omega} - \big((\mu \cdot n) v, u \big)_{\partial\Omega} + (v, e_m) \\
\big(v, u(0,x) \big) = \big(v, u^0(x) \big) + \big(v, e_0(x) \big)
\end{cases}
\tag{3.16}
$$

The usage of the divergence theorem permits the incorporation of the boundary conditions into the variational formulation (3.16) through the boundary integrals. Those integrals require boundary information about the function $u(t,x)$, known as Dirichlet data, and information about the normal derivative of $u(t,x)$, known as Neumann data. If the Dirichlet and Neumann data is imposed through the boundary integrals, it is said that the boundary conditions are satisfied in a weak sense.

The variational formulation of the observation equation is straightforward and is written in the following way

$$
(v, y) = (v, Hu) + (v, e_o) \tag{3.17}
$$

3.1.3 Problems continuous in time and discrete in space

The discrete projection of the continuous weak solution is obtained as a truncation of the corresponding infinite expansion of the function. Let $\phi_i(x) \in H^1(\Omega), i = 1, \dots, \infty$ be a complete system of a basis function in the space $H^1(\Omega)$. Then there exists a set of coefficients $\mathbf{u}_i(t)$, such that

$$u(t,x) = \sum_{i=1}^{\infty} \mathbf{u}_i(t)\phi_i(x) \tag{3.18}$$

By taking $u^N(t,x)$ to be the sum of only the first N terms in (3.18), one gets a finite dimensional approximation of $u(t,x)$ that has the desired property of separated time and space variables. This approach is known as the semi-discrete Galerkin approximation.

Different choices of the basis functions define different methods with different approximation properties. In this work, a Finite Element Method (FEM) with bilinear basis functions is considered. The domain Ω is approximated by rectangular finite elements and the vertices $x_i, i = 1, \dots, N_{nd}$ of the elements form the FEM grid. At each node of the grid, the basis function $\phi_i(x)$ is defined as a tensor product of one-dimensional piecewise linear functions that are called hat functions or Currant functions, so that the following interpolation condition holds:

$$\phi_j(x_i) = \delta_{ij} \; i,j = 1, \dots, N_{nd} \tag{3.19}$$

The FEM approximation of the continuous function $u(t,x)$ is then written as follows

$$u(t,x) \approx u^N(t,x) = \sum_{i=1}^{N_{nd}} \mathbf{u}_i(t)\phi_i(x) = \boldsymbol{\phi}^T(x)\mathbf{u}(t) \tag{3.20}$$

where $\boldsymbol{\phi}(x) = \left[\phi_1(x), \dots, \phi_{N_{nd}}(x)\right]^T$ is a vector of FEM basis functions and $\mathbf{u}(t) = \left[\mathbf{u}_1(t), \dots, \mathbf{u}_{N_{nd}}(t)\right]^T$ is a vector of unknown coefficients of the expansion.

The stochastic function $e(t,x)$ is approximated similarly to (3.20)

$$e(t,x) \approx e^N(t,x) = \sum_{i=1}^{N_{nd}} \mathbf{e}_i(t)\phi_i(x) = \boldsymbol{\phi}^T(x)\mathbf{e}(t) \tag{3.21}$$

where $\mathbf{e}(t) = \left[\mathbf{e}_1(t), \dots, \mathbf{e}_{N_{nd}}(t)\right]^T$ is a vector of expansion coefficients and because of the property of the FEM basis function (3.19) it follows that

$$\mathbf{e}_i(t) = e(t, x_i) \; i = 1, \dots, N_{nd} \tag{3.22}$$

Inserting the given FEM approximations (3.20) and (3.21) into the variational representation (3.12) and taking N test functions $v(x) = \phi_i(x), i = 1, \dots, N$ denoted in vector form $\boldsymbol{\phi}(x)$, the

space-discrete background model for the vector of coefficients $\mathbf{u}(t)$ is obtained in the form of ordinary differential equations (ODE):

$$\begin{aligned}\int_{\Omega} \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\frac{d\mathbf{u}(t)}{dt}dx \\ &= -\int_{\Omega} \epsilon\nabla^T\boldsymbol{\phi}(x)\nabla\boldsymbol{\phi}^T(x)\mathbf{u}(t)dx \\ &+ \int_{\Omega} \nabla^T\boldsymbol{\phi}(x)\mu(t,x)\boldsymbol{\phi}^T(x)\mathbf{u}(t)dx \\ &+ \int_{\partial\Omega} \epsilon\boldsymbol{\phi}(x)\frac{d\boldsymbol{\phi}^T(x)}{dn(x)}\mathbf{u}(t)dx \\ &- \int_{\partial\Omega} \mu^T(t,x)n(x)\boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\mathbf{u}(t)dx \\ &+ \int_{\Omega} \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\mathbf{e}_m(t)dx\end{aligned} \tag{3.23}$$

with the initial conditions

$$\int_{\Omega} \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\boldsymbol{u}(0)dx = \int_{\Omega} \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\boldsymbol{u}^0dx + \int_{\Omega} \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\boldsymbol{e}_0dx \tag{3.24}$$

The boundary integrals in (3.23) are split into zones $\partial\Omega_N$ and $\partial\Omega_D$ where proper boundary conditions are defined. Thus, the boundary integrals computed using known boundary data form a vector $\tilde{\mathbf{f}}(t)$

$$\begin{aligned}\tilde{\mathbf{f}}(t) = \int_{\partial\Omega_N} \epsilon\boldsymbol{\phi}(x)\frac{d\boldsymbol{\phi}^T(x)}{dn(x)}\mathbf{u}_N(t)dx \\ - \int_{\partial\Omega_D} \mu^T(t,x)n(x)\boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\mathbf{u}_D(t)dx\end{aligned} \tag{3.25}$$

If the boundary data is not prescribed, corresponding integrals contribute to the matrix $\tilde{\mathbf{S}}_{\partial\Omega}(t)$

$$\tilde{\mathbf{S}}_{\partial\Omega}(t) = \int_{\partial\Omega_D} \epsilon\boldsymbol{\phi}(x)\frac{d\boldsymbol{\phi}^T(x)}{dn(x)}dx - \int_{\partial\Omega_N} \mu^T(t,x)n(x)\boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)dx \tag{3.26}$$

Introducing the matrix notations

$$\mathbf{M} = \int_{\Omega} \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)dx \tag{3.27}$$

$$\tilde{\mathbf{S}}_{\Omega}(t) = -\int_{\Omega} \epsilon\nabla^T\boldsymbol{\phi}(x)\nabla\boldsymbol{\phi}^T(x)dx + \int_{\Omega} \nabla^T\boldsymbol{\phi}(x)\mu(t,x)\boldsymbol{\phi}^T(x)dx \tag{3.28}$$

And denoting $\tilde{\mathbf{S}}(t) = \tilde{\mathbf{S}}_{\Omega}(t) + \tilde{\mathbf{S}}_{\partial\Omega}(t)$ the ODE system (3.23)-(3.24) can be rewritten as

$$\begin{cases} \mathbf{M}\dfrac{d\mathbf{u}(t)}{dt} = \tilde{\mathbf{S}}(t)\mathbf{u}(t) + \tilde{\mathbf{f}}(t) + \mathbf{M}\mathbf{e}_m(t) \\ \mathbf{M}\mathbf{u}(0) = \mathbf{M}\mathbf{u}^0 + \mathbf{M}\mathbf{e}_0 \end{cases} \tag{3.29}$$

The set of functions $\phi_i(x)$ is linearly independent, meaning that the mass matrix $\mathbf{M}$ is invertible. This quality allows for the stiffness matrix and source vector to be denoted as $\mathbf{S}(t) = \mathbf{M}^{-1}\tilde{\mathbf{S}}(t)$ and $\mathbf{f}(t) = \mathbf{M}^{-1}\tilde{\mathbf{f}}(t)$. In the new notations, the system (3.29) becomes

$$\begin{cases} \dfrac{d\mathbf{u}(t)}{dt} = \mathbf{S}(t)\mathbf{u}(t) + \mathbf{f}(t) + \mathbf{e}_m(t) \\ \mathbf{u}(0) = \mathbf{u}^0 + \mathbf{e}_0 \end{cases} \tag{3.30}$$

Here $\mathbf{S}$ is the stiffness matrix that corresponds to the model operator L and vector $\mathbf{f}(\mathrm{t})$ reflects the fact that the boundary conditions are imposed in a weak sense.

Discretisation of the observation equation (3.6) is done in a similar fashion

$$\begin{aligned} &\int_{\Omega} \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\mathbf{y}(t)dx \\ &\qquad = \int_{\Omega} \boldsymbol{\phi}(x) \int_{\Omega} \boldsymbol{\phi}^T(x)\tilde{\mathbf{H}}\boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z)\mathbf{u}(t)dz\,dx \\ &\qquad + \int_{\Omega} \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\mathbf{e}_o(t)dx \end{aligned} \tag{3.31}$$

where matrix $\tilde{\mathbf{H}}$ is a discretised observation kernel function

$$\tilde{\mathbf{H}} = \{h(x_i - z_j)\}_{i,j=1}^{N_{nd}} \tag{3.32}$$

Accounting for the notation of the mass matrix and denoting the discrete observation operator as $\mathbf{H} = \tilde{\mathbf{H}}\mathbf{M}$ yields the observation equation in the discrete form

$$\mathbf{y}(t) = \mathbf{H}\mathbf{u}(t) + \mathbf{e}_o(t) \tag{3.33}$$

The statistics of the space discretised errors $\mathbf{e}_m(t)$, $\mathbf{e}_o(t)$ and $\mathbf{e}_0$ are retrieved from the expression (3.22). Indeed, normality of the continuous model error $e_m(t,x)$ and (3.22) imply normal distribution of the components $\mathbf{e}_{m,i}(t)$ and, as a result, normal distribution of the vector of the model error $\mathbf{e}_m(t)$ itself. Moreover, this also implies a zero average of the error vector $\bar{\mathbf{e}}_m(t) = 0$.

The covariance matrix of the model error is computed from the definition of the covariance and the relation (3.22)

$$\begin{aligned}\mathbf{Q}(t) &= \overline{\mathbf{e}_m(t)\mathbf{e}_m^T(t)} = \overline{\left(\mathbf{e}_{m,i}(t)\mathbf{e}_{m,j}(t)\right)}_{ij} \\ &= \overline{\left(e_m(t,x_i)e_m(t,x_j)\right)}_{ij} = \left(Q(t,t,x_i,x_j)\right)_{ij}\end{aligned} \tag{3.34}$$

The PDF for $\mathbf{e}_m(t)$ that is equivalent to the continuous one (3.7) is written as follows

$$\begin{aligned}p(\mathbf{e}_m(t)) &= \frac{1}{\sqrt{(2\pi)^k|\mathbf{Q}(t)|}} \\ &\times \exp\left(-\frac{1}{2}\left(\mathbf{e}_m(t)-\overline{\mathbf{e}}_m(t)\right)^T\mathbf{Q}^{-1}(t)(\mathbf{e}_m(t)-\overline{\mathbf{e}}_m(t))\right)\end{aligned} \tag{3.35}$$

Using the same arguments as above, the space discrete observation and initial errors $\mathbf{e}_o(t)$ and $\mathbf{e}_0$ are normally distributed with a zero average. The covariance of the discrete observation and initial errors are obtained as

$$\mathbf{R}(t) = \left(R(t,t,x_i,x_j)\right)_{ij} \text{ and } \mathbf{Q}^0 = \left(Q_0(x_i,x_j)\right)_{ij} \tag{3.36}$$

The latter also demonstrates that if the continuous error satisfies the Gaussian white noise assumptions, then its space-discretised counterpart satisfies those assumptions as well.

Finally, the time continuous and space discrete DA problem is formulated as a state estimation problem of the system described by the linear ODE

$$\begin{cases}\dfrac{d\mathbf{u}(t)}{dt} = \mathbf{S}(t)\mathbf{u}(t) + \mathbf{f}(t) + \mathbf{e}_m(t) \\ \mathbf{u}(0) = \mathbf{u}^0 + \mathbf{e}_0\end{cases} \tag{3.37}$$

with available observations in a linear form

$$\mathbf{y}(t) = \mathbf{H}\mathbf{u}(t) + \mathbf{e}_o(t) \tag{3.38}$$

The errors are described by the following statistics:

$$\mathbf{e}_m(t) \sim N(0,\mathbf{Q}(t));\ \mathbf{e}_o \sim N(0,\mathbf{R}(t));\ \mathbf{e}_0 \sim N(0,\mathbf{Q}^0); \tag{3.39}$$

3.1.4 Problems discrete in time and space

In order to solve the differential equation (3.37), implicit midpoint rule is adopted (Iserles, 1996). This is done by integrating (3.37) over the time interval $[t_0, t]$, so that

$$\mathbf{u}(t) - \mathbf{u}(t_0) = \int_{t_0}^{t} \mathbf{S}(t)\mathbf{u}(t)\, dt + \int_{t_0}^{t} \mathbf{f}(t)\, dt + \int_{t_0}^{t} \mathbf{e}_m(t)\, dt \tag{3.40}$$

The computation of the integral terms in (3.40) is done by the midpoint quadrature

$$\int_a^b g(t)\, dt \approx (b-a) g\left(\frac{a+b}{2}\right) \tag{3.41}$$

And $\mathbf{u}\left(\frac{a+b}{2}\right)$ is approximated as

$$\mathbf{u}\left(\frac{a+b}{2}\right) \approx \frac{\mathbf{u}(a) + \mathbf{u}(b)}{2} \tag{3.42}$$

Applying (3.41) and (3.42) to (3.40) yields the discrete in time and space equation for $\mathbf{u}(t)$

$$\begin{aligned} \mathbf{u}(t) - \mathbf{u}(t_0) = (t-t_0)\mathbf{S}\left(\frac{t+t_0}{2}\right)\frac{\mathbf{u}(t_0) + \mathbf{u}(t)}{2} \\ + (t-t_0)\mathbf{f}\left(\frac{t+t_0}{2}\right) + \int_{t_0}^{t} \mathbf{e}_m(\tau)\, d\tau \end{aligned} \tag{3.43}$$

Rearranging the terms in the last formula gives

$$\begin{aligned} \left(\mathbf{I} - \frac{t-t_0}{2}\mathbf{S}\left(\frac{t+t_0}{2}\right)\right)\mathbf{u}(t) \\ = -\left(\mathbf{I} - \frac{t-t_0}{2}\mathbf{S}\left(\frac{t+t_0}{2}\right)\right)\mathbf{u}(t_0) + 2\mathbf{u}(t_0) \\ + (t-t_0)\mathbf{f}\left(\frac{t+t_0}{2}\right) + \int_{t_0}^{t} \mathbf{e}_m(\tau)\, d\tau \end{aligned} \tag{3.44}$$

Moving the multiplier from left side to right side, the recurrent expression for the system state is obtained

$$\begin{aligned} \mathbf{u}(t) = 2\left(\left[\mathbf{I} - \frac{t-t_0}{2}\mathbf{S}\left(\frac{t+t_0}{2}\right)\right]^{-1} - \mathbf{I}\right)\mathbf{u}(t_0) \\ + (t-t_0)\left[\mathbf{I} - \frac{t-t_0}{2}\mathbf{S}\left(\frac{t+t_0}{2}\right)\right]^{-1}\mathbf{f}\left(\frac{t+t_0}{2}\right) \\ + \left[\mathbf{I} - \frac{t-t_0}{2}\mathbf{S}\left(\frac{t+t_0}{2}\right)\right]^{-1}\int_{t_0}^{t} \mathbf{e}_m(\tau)\, d\tau \end{aligned} \tag{3.45}$$

The time variable is now uniformly discretised with the interval T such that $t_0 = nT$; $t = (n+1)T$; $T = t - t_0$. The time discretisation permits the derivation of the following notations:

for the state vector

$$\mathbf{u}_0 = \mathbf{u}(0);\ \mathbf{u}_n = \mathbf{u}(nT);\ \mathbf{u}_{n+1} = \mathbf{u}\big((n+1)T\big) \tag{3.46}$$

for the model matrix

$$\mathbf{S}_n = 2\left(\left[\mathbf{I} - \frac{T}{2}\mathbf{S}\left(nT + \frac{T}{2}\right)\right]^{-1} - \mathbf{I}\right) \tag{3.47}$$

for the source vector

$$\mathbf{f}_n = T\left(\mathbf{I} - \frac{T}{2}\mathbf{S}\left(nT + \frac{T}{2}\right)\right)^{-1} \mathbf{f}\left(nT + \frac{T}{2}\right) \tag{3.48}$$

and for the model error vector

$$\mathbf{e}_{m,n} = \left(\mathbf{I} - \frac{T}{2}\mathbf{S}\left(nT + \frac{T}{2}\right)\right)^{-1} \int_0^T \mathbf{e}_m(nT + \tau)\, d\tau \tag{3.49}$$

Using (3.46)-(3.49), the time and space discretised background model (3.45) is written in the form of the algebraic recurrence relation

$$\begin{cases} \mathbf{u}_{n+1} = \mathbf{S}_n\mathbf{u}_n + \mathbf{f}_n + \mathbf{e}_{m,n} \\ \mathbf{u}_0 = \mathbf{u}^0 + \mathbf{e}_0 \end{cases} \tag{3.50}$$

While there are many possible choices for the numerical approximation of definite integrals of (3.40), the midpoint rule (3.41) was chosen for this research as its application for the integration of deterministic ODE systems is known to have good stability properties even for stiff systems and provides a second order of accuracy (Levy, 2010).

In a similar fashion, the observations equation (3.38) is integrated in time, but, for consistency of the time discretisation, applying the left rectangles rule for computation of the time integrals

$$T\mathbf{y}(nT) = T\mathbf{H}\mathbf{u}(nT) + \int_0^T \mathbf{e}_o(nT + t)dt \tag{3.51}$$

With respect to the notations (3.46), the discrete in space and time observations equation is written

$$\mathbf{y}_n = \mathbf{H}\mathbf{u}_n + \mathbf{e}_{o,n} \tag{3.52}$$

where

$$\mathbf{e}_{o,n} = \int_0^T \mathbf{e}_o(nT + t)dt \tag{3.53}$$

The next step is to analyse the error statistics. From the relation (3.49) and the properties of normal distribution, it is evident that $\mathbf{e}_{m,n}$, $\mathbf{e}_{o,n}$ and $\mathbf{e}_0$ are normally distributed with zero mean. The model error covariance $\mathbf{Q}_n^D$ is found from its definition

$$\begin{aligned} \mathbf{Q}_n^D &= \overline{\mathbf{e}_{m,n}\mathbf{e}^T{}_{m,n}} \\ &= \left(\mathbf{I} - \frac{T}{2}\mathbf{S}(nT)\right)^{-1} \int_0^T \int_0^T \overline{\mathbf{e}_m(nT+\tau)\mathbf{e}^T{}_m(nT+\sigma)}\, d\tau\, d\sigma \\ &\quad \times \left(\mathbf{I} - \frac{T}{2}\mathbf{S}(nT)\right)^{-1^T} \end{aligned} \tag{3.54}$$

Since

$$\overline{\mathbf{e}_m(nT+\tau)\mathbf{e}_m^T(nt+\sigma)} = \mathbf{Q}(nT+\tau)\delta(\tau-\sigma) \tag{3.55}$$

then

$$\int_0^T \int_0^T \mathbf{Q}(nT+\tau)\delta(\tau-\sigma)d\tau\, d\sigma = \int_0^T \mathbf{Q}(nT+\sigma)\, d\sigma \tag{3.56}$$

Again, applying the midpoint rule (3.41) for the latter integral

$$\int_0^T \mathbf{Q}(nT+\sigma)\, d\sigma = T\mathbf{Q}\left(nT + \frac{T}{2}\right) \tag{3.57}$$

and inserting it into the expression for the model error covariance (3.54) one gets

$$\begin{aligned} \mathbf{Q}_n^D &= \left(\mathbf{I} - \frac{T}{2}\mathbf{S}(nT)\right)^{-1} T\mathbf{Q}\left(nT + \frac{T}{2}\right) \\ &\quad \times \left(\mathbf{I} - \frac{T}{2}\mathbf{S}(nT)\right)^{-1^T} \end{aligned} \tag{3.58}$$

That formulation is further simplified by denoting

$$\widetilde{\mathbf{Q}}_{n+0.5} = \mathbf{Q}\left(nT + \frac{T}{2}\right) \tag{3.59}$$

Since $\left\|\frac{T}{2}\mathbf{S}(nT)\right\| \leq 1$ for a sufficiently small T, covariance matrix (3.58) is rewritten using the matrix Neumann series

$$(\mathbf{I} - \mathbf{X})^{-1} = \mathbf{I} + \mathbf{X} + \mathbf{X}^2 + \cdots, \qquad \text{provided } \|\mathbf{X}\| \leq 1 \tag{3.60}$$

in the following way

$$\begin{aligned}\mathbf{Q}_n^D = \Bigg(\mathbf{I} + \frac{T}{2}\mathbf{S}(nT) + \frac{T^2}{4}\mathbf{S}^2(nT) \\ + \cdots\Bigg) T\widetilde{\mathbf{Q}}_{n+0.5}\left(\mathbf{I} + \frac{T}{2}\mathbf{S}(nT) + \frac{T^2}{4}\mathbf{S}^2(nT) + \cdots\right)^T\end{aligned} \tag{3.61}$$

$$\begin{aligned}\mathbf{Q}_n^D = T\widetilde{\mathbf{Q}}_{n+0.5} + \frac{T^2}{2}\widetilde{\mathbf{Q}}_{n+0.5}\mathbf{S}^T(nT) + \frac{T^2}{2}\mathbf{S}(nT)\widetilde{\mathbf{Q}}_{n+0.5} \\ + \frac{T^3}{4}\mathbf{S}(nT)\widetilde{\mathbf{Q}}_{n+0.5}\mathbf{S}^T(nT) + \cdots\end{aligned} \tag{3.62}$$

The approximation of the background model (3.50) is of the second order accuracy, which implies it is enough to provide the second order of accuracy for the time discretisation of the covariance matrix. This means that all terms in (3.61) that contain T to the power of two or higher can be neglected. From this argument, it follows that

$$\mathbf{Q}_n^D = T\widetilde{\mathbf{Q}}_{n+0.5} = T\mathbf{Q}\left(nT + \frac{T}{2}\right) \tag{3.63}$$

The temporal discretisation of the model error covariance (3.63) should be regarded as an approximation of a continuous function by a piecewise constant function.

From the definition of the observations error covariance, it can be shown that

$$\begin{aligned}\mathbf{R}_n^D &= \overline{\mathbf{e}_{o,n}, \mathbf{e}_{o,n}^T} \\ &= \frac{1}{T^2}\int_0^T \int_0^T \overline{\mathbf{e}_o(nT + t), \mathbf{e}_o^T(nT + \sigma)}\, dt\, d\sigma\end{aligned} \tag{3.64}$$

Using the fact that the time continuous observations error covariance is given as

$$\overline{\mathbf{e}_o(t), \mathbf{e}_o^T(\sigma)} = \mathbf{R}(t)\delta(t - \tau) \tag{3.65}$$

and applying the left rectangle integration rule, it is found that

$$\mathbf{R}_n^D = \frac{1}{T^2}\int_0^T\int_0^T \mathbf{R}(nT+t)\delta(t-\tau)dt\,d\sigma$$
$$= \frac{1}{T^2}\int_0^T \mathbf{R}(nT+\tau)d\sigma \approx \frac{1}{T}\mathbf{R}\left(nT+\frac{T}{2}\right) \tag{3.66}$$

In the end, a DA problem is formulated as a state estimation problem of the recurrent algebraic equation, also known as a difference equation

$$\begin{cases}\mathbf{u}_{n+1} = \mathbf{S}_n\mathbf{u}_n + \mathbf{f}_n + \mathbf{e}_{m,n} \\ \mathbf{u}_0 = \mathbf{u}^0 + \mathbf{e}_0\end{cases} \tag{3.67}$$

with observations in the algebraic form

$$\mathbf{y}_n = \mathbf{H}\mathbf{u}_n + \mathbf{e}_{o,n} \tag{3.68}$$

and the error statics

$$\mathbf{e}_{m,n}(t) \sim N(0,\mathbf{Q}_n);\ \mathbf{e}_{o,n} \sim N(0,\mathbf{R}_n);\ \mathbf{e}_0 \sim N(0,\mathbf{Q}^0); \tag{3.69}$$

that can be equivalently formulated with regard to the continuous error description

$$\mathbf{e}_{m,n}(t) \sim N\left(0, T\mathbf{Q}\left(nT+\frac{T}{2}\right)\right);\ \mathbf{e}_o \sim N\left(0, \frac{1}{T}\mathbf{R}\left(nT+\frac{T}{2}\right)\right);$$
$$\mathbf{e}_0 \sim N(0,\mathbf{Q}^0); \tag{3.70}$$

3.2 Kalman based filters

From the error description (3.69) and the equation of the background model evolution (3.67), the stochastic background state transition is written in the form of the Gaussian conditional probability

$$p(\mathbf{u}_n|\mathbf{u}_{n-1}) \sim N(\mathbf{S}_{n-1}\mathbf{u}_{n-1} + \mathbf{f}_n, \mathbf{Q}_n) \tag{3.71}$$

Considering the fact that the model error $\mathbf{e}_{m,n}$ is white noise, the stochastic processes $\{\mathbf{u}_n\}$ defined by (3.67) is Markovian and has the initial distribution

$$p(\mathbf{u}_0) \sim N(\mathbf{u}^0, \mathbf{Q}^0) \tag{3.72}$$

Similarly, the observations equation (3.68) is written as the conditional probability

$$p(\boldsymbol{y}_n|\mathbf{u}_n) \sim N(\mathbf{H}\mathbf{u}_n, \mathbf{R}_0) \tag{3.73}$$

The aim is to infer the probability distribution of the observations (3.73) into the background probability (3.71) and get the analysis probability $p(\mathbf{u}_n|\mathbf{Y}_n)$, where $\mathbf{Y}_n = [\mathbf{y}_1, \dots, \mathbf{y}_n]$. This can be done using Bayes' formula, written for a time discrete system

$$p(\mathbf{u}_n|\mathbf{Y}_n) = \frac{p(\mathbf{y}_n|\mathbf{u}_n)p(\mathbf{u}_n|\mathbf{Y}_{n-1})}{\int p(\mathbf{y}_n|\mathbf{u}_n)p(\mathbf{u}_n|\mathbf{Y}_{n-1})d\mathbf{u}_n} \tag{3.74}$$

Using the transition probability (3.71) and a marginal rule:

$$p(\mathbf{u}_n|\mathbf{Y}_{n-1}) = \int p(\mathbf{u}_n|\mathbf{u}_{n-1})p(\mathbf{u}_{n-1}|\mathbf{Y}_{n-1})d\mathbf{u}_n \tag{3.75}$$

Results in

$$p(\mathbf{u}_n|\mathbf{Y}_n) = \frac{p(\mathbf{y}_n|\mathbf{u}_n)\int p(\mathbf{u}_n|\mathbf{u}_{n-1})p(\mathbf{u}_{n-1}|\mathbf{Y}_{n-1})d\mathbf{u}_{n-1}}{\int p(\mathbf{y}_n|\mathbf{u}_n)\left(\int p(\mathbf{u}_n|\mathbf{u}_{n-1})p(\mathbf{u}_{n-1}|\mathbf{Y}_{n-1})d\mathbf{u}_{n-1}\right)d\mathbf{u}_n} \tag{3.76}$$

From (3.76) and properties of the Gaussian functions, it is concluded that

$$p(\boldsymbol{u}_n|\mathbf{Y}_n) \sim N(\boldsymbol{\mu}, \mathbf{P}) \tag{3.77}$$

Where $\boldsymbol{\mu}$ and $\mathbf{P}$ are the mean and covariance of $\{\mathbf{u}_n\}$, which may be derived from (3.76). The following section presents an approach for computation of $\boldsymbol{\mu}$ and $\mathbf{P}$ based on algebraic analysis of the error variances that is known as the Kalman filter.

3.2.1 Discrete Kalman filter

The idea of the Kalman filter is to recursively evolve the mean and covariance of the analysis state from a time step n to a time step $n+1$ starting with the initial mean $\mathbf{u}^0$ and the initial covariance $\mathbf{Q}^0$. At each time step, the analysis is constructed such that it incorporates all available observations from the past while maintaining a minimal variance.

Let $\mathbf{u}_n^t$ be the true state of the system presented by the model with observations discretised in time and space (3.67)-(3.69). Assume that at the time step n, analysis mean $\overline{\mathbf{u}}_n^a$ and covariance of the analysis error, given by

$$\mathbf{P}_n^a = \overline{(\mathbf{u}_n^a - \mathbf{u}_n^t)(\mathbf{u}_n^a - \mathbf{u}_n^t)^T} \tag{3.78}$$

are known. The first step of the Kalman algorithm is to perform the background propagation of the analysis

$$\mathbf{u}_{n+1}^{b} = \mathbf{S}_n \mathbf{u}_n^a + \mathbf{f}_n + \mathbf{e}_{m,n} \tag{3.79}$$

Taking the mean of the left and right sides together with the assumption of unbiased errors yields the background mean

$$\overline{\mathbf{u}}_{n+1}^{b} = \mathbf{S}_n \overline{\mathbf{u}}_n^a + \mathbf{f}_n \tag{3.80}$$

To compute the background error covariance

$$\mathbf{P}_{n+1}^{b} = \overline{\left(\mathbf{u}_{n+1}^{b} - \overline{\mathbf{u}}_{n+1}^{t}\right)\left(\mathbf{u}_{n+1}^{b} - \overline{\mathbf{u}}_{n+1}^{t}\right)^T} \tag{3.81}$$

expression (3.79) is inserted into the last formula

$$\mathbf{P}_{n+1}^{b} = \overline{\left(\mathbf{S}_n \mathbf{u}_n^a + \mathbf{e}_{m,n} - \mathbf{S}_n \overline{\mathbf{u}}_n^t\right)\left(\mathbf{S}_n \mathbf{u}_n^a + \mathbf{e}_{m,n} - \mathbf{S}_n \overline{\mathbf{u}}_{n+1}^t\right)^T} \tag{3.82}$$

Multiplying the terms in brackets and regrouping them returns

$$\begin{aligned} \mathbf{P}_{n+1}^{b} = {} & \boldsymbol{S}_n \overline{(\mathbf{u}_n^a - \mathbf{u}_n^t)(\mathbf{u}_n^a - \mathbf{u}_n^t)^T} \boldsymbol{S}_n^T + \overline{\mathbf{e}_{m,n}(\mathbf{u}_n^a - \mathbf{u}_n^t)^T} \boldsymbol{S}_n^T \\ & + \boldsymbol{S}_n \overline{(\mathbf{u}_n^a - \mathbf{u}_n^t)\mathbf{e}_{m,n}^T} + \overline{\mathbf{e}_{m,n}\mathbf{e}_{m,n}^T} \end{aligned} \tag{3.83}$$

Due to the fact that the error and state variables are mutually uncorrelated, their cross products vanish so that

$$\mathbf{P}_{n+1}^{b} = \boldsymbol{S}_n \overline{(\mathbf{u}_n^a - \mathbf{u}_n^t)(\mathbf{u}_n^a - \mathbf{u}_n^t)^T} \boldsymbol{S}_n^T + \overline{\mathbf{e}_{m,n}\mathbf{e}_{m,n}^T} \tag{3.84}$$

Taking into account the description of the model error (3.69) and analysis error (3.78) the background covariance evolution in time step $n+1$ is written in the form of Lyapunov equation

$$\mathbf{P}_{n+1}^{b} = \mathbf{S}_n \mathbf{P}_n^a \mathbf{S}_n{}^T + \mathbf{Q}_n \tag{3.85}$$

with a given initial condition $\mathbf{P}_n^a = \mathbf{Q}^0$.

It should be further noted that the superscript b in $\mathbf{u}_{n+1}^{b}$ and $\mathbf{P}_{n+1}^{b}$ denotes the fact that those are the values before an observation is made at time t_{n+1}, while $\mathbf{u}_{n+1}^{a}$ and $\mathbf{P}_{n+1}^{a}$ denote values after observations at time instance t_{n+1} are incorporated into the analysis. The second step of

the algorithm is to update the background mean and covariance by taking into account information from the observations.

Let $\mathbf{d}_{n+1}$ define the difference between the actual observations of the state and the expected background state

$$\mathbf{d}_{n+1} = \mathbf{y}_{n+1} - \mathbf{H}\mathbf{u}_{n+1}^{b} \tag{3.86}$$

The analysis state is found as the correction of the background state by the vector $\mathbf{d}_{n+1}$ that is weighted by a matrix parameter $\mathbf{K}$:

$$\mathbf{u}_{n+1}^{a} = \mathbf{u}_{n+1}^{b} + \mathbf{K}\mathbf{d}_{n+1} \tag{3.87}$$

In order to get the optimal choice for $\mathbf{K}$, the covariance of the analysis error is computed

$$\mathbf{P}_{n+1}^{a} = \overline{(\mathbf{u}_{n+1}^{a} - \mathbf{u}_{n+1}^{t})(\mathbf{u}_{n+1}^{a} - \mathbf{u}_{n+1}^{t})^{T}} \tag{3.88}$$

Taking into account the expression for the analysis state (3.87) and the observation difference (3.86) the analysis error is rewritten as

$$\mathbf{u}_{n+1}^{a} - \mathbf{u}_{n+1}^{t} = \mathbf{u}_{n+1}^{b} + \mathbf{K}\left(\mathbf{y}_{n} - \mathbf{H}\mathbf{u}_{n+1}^{b}\right) - \mathbf{u}_{n+1}^{t} \tag{3.89}$$

Inserting the observation equation (3.68) and regrouping the terms, one gets

$$\mathbf{u}_{n+1}^{a} - \mathbf{u}_{n+1}^{t} = -\mathbf{u}_{n+1}^{t} + \mathbf{u}_{n+1}^{b} + \mathbf{K}\mathbf{H}\mathbf{u}_{n}^{t} - \mathbf{K}\mathbf{H}\mathbf{u}_{n+1}^{b} + \mathbf{K}\mathbf{e}_{o,n+1} \tag{3.90}$$

$$\mathbf{u}_{n+1}^{a} - \mathbf{u}_{n+1}^{t} = -(\mathbf{I} - \mathbf{K}\mathbf{H})\left(\mathbf{u}_{n+1}^{t} - \mathbf{u}_{n+1}^{b}\right) + \mathbf{K}\mathbf{e}_{o,n+1} \tag{3.91}$$

The analysis error covariance is then computed from the last equality

$$\begin{aligned}\mathbf{P}_{n+1}^{a} = (\mathbf{I} - \mathbf{K}\mathbf{H})\overline{\left(\mathbf{u}_{n+1}^{t} - \mathbf{u}_{n+1}^{b}\right)\left(\mathbf{u}_{n+1}^{t} - \mathbf{u}_{n+1}^{b}\right)^{T}}(\mathbf{I} - \mathbf{K}\mathbf{H})^{T} \\ + \mathbf{K}\overline{\mathbf{e}_{o,n+1}\mathbf{e}_{o,n+1}^{T}}\mathbf{K}^{T}\end{aligned} \tag{3.92}$$

The covariance $\mathbf{P}_{n+1}^{a}$ is further transformed using (3.81) and (3.69)

$$\mathbf{P}_{n+1}^{a} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_{n+1}^{b}(\mathbf{I} - \mathbf{K}\mathbf{H})^{\mathrm{T}} + \mathbf{R}_{n+1} \tag{3.93}$$

By rearranging terms, it is shown that

$$\mathbf{P}_{n+1}^{a} = \mathbf{P}_{n+1}^{b} - \mathbf{KHP}_{n+1}^{b} - \mathbf{P}_{n+1}^{b}(\mathbf{KH})^{T} + \mathbf{K}\left(\mathbf{HP}_{n+1}^{b}\mathbf{H}^{T} + \mathbf{R}_{n+1}\right)\mathbf{K}^{T} \tag{3.94}$$

Since it is desired that the analysis state is as close as possible to the true state, the value of the analysis error covariance is optimised by the proper choice of parameter **K**. To this end, the optimal choice of **K** is found as a solution of the optimality equation

$$\frac{\delta \mathbf{P}_{n+1}^{a}}{\delta \mathbf{K}} = 0 \tag{3.95}$$

or

$$\frac{\delta\left(\mathbf{P}_{n+1}^{b} - \mathbf{KHP}_{n+1}^{b} - \mathbf{P}_{n+1}^{b}(\mathbf{KH})^{T} + \mathbf{K}\left(\mathbf{HP}_{n+1}^{b}\mathbf{H}^{T} + \mathbf{R}_{n+1}\right)\mathbf{K}^{T}\right)}{\delta \mathbf{K}} = 0 \tag{3.96}$$

It is easy to check, that the following expression of **K**

$$\mathbf{K} = \mathbf{P}_{n+1}^{b}\mathbf{H}^{T}\left(\mathbf{HP}_{n+1}^{b}\mathbf{H}^{T} + \mathbf{R}_{n+1}\right)^{-1} \tag{3.97}$$

verifies the optimality equation (3.96).

The optimum value for **K** in terms of minimization of the analysis error covariance given by (3.97) is known as the Kalman gain. Once it is computed, the analysis mean is found from (3.87) as

$$\overline{\mathbf{u}}_{n+1}^{a} = \overline{\mathbf{u}}_{n+1}^{b} + \mathbf{K}(\mathbf{y}_{n+1} - \mathbf{H}\overline{\mathbf{u}}_{n+1}^{b}) \tag{3.98}$$

Finally, substituting Kalman gain (3.97) into (3.93) produces the expression for the optimal analysis error covariance

$$\begin{aligned} \mathbf{P}_{n+1}^{a} &= \mathbf{P}_{n+1}^{b} - \mathbf{KHP}_{n+1}^{b} - \mathbf{P}_{n+1}^{b}\mathbf{H}^{T}\mathbf{K} + \mathbf{P}_{n+1}^{b}\mathbf{H}^{T}\mathbf{K} \\ &= (\mathbf{I} - \mathbf{KH})\mathbf{P}_{n+1}^{b} \end{aligned} \tag{3.99}$$

3.2.2 The Kalman filter algorithm

The algorithm of the Kalman filter for the solution of the problem (3.67)-(3.69) starts with the initial conditions for the covariance and mean

$$\mathbf{P}_{0}^{a} = \mathbf{Q}^{0};\ \mathbf{u}_{0} = \mathbf{u}^{0} \tag{3.100}$$

The background update (prediction) of the error covariance and estimate are found as

$$\mathbf{P}_{n+1}^{b} = \mathbf{S}_n \mathbf{P}_n^a \mathbf{S}_n^T + \mathbf{Q}_n \tag{3.101}$$

$$\mathbf{u}_{n+1}^{b} = \mathbf{S}_n \mathbf{u}_n^a + \mathbf{f}_n \tag{3.102}$$

The observation update (correction) of the error covariance and estimate are therefore computed from

$$\mathbf{K}_{n+1} = \mathbf{P}_{n+1}^{b} \mathbf{H}^T \left(\mathbf{H}\mathbf{P}_{n+1}^{b} \mathbf{H}^T + \mathbf{R}_{n+1}\right)^{-1} \tag{3.103}$$

$$\mathbf{P}_{n+1}^{a} = (\mathbf{I} - \mathbf{K}_{n+1}\mathbf{H})\mathbf{P}_{n+1}^{b} \tag{3.104}$$

$$\mathbf{u}_{n+1}^{a} = \mathbf{u}_{n+1}^{b} + \mathbf{K}_{n+1}\left(\mathbf{y}_{n+1} - \mathbf{H}\mathbf{u}_{n+1}^{b}\right) \tag{3.105}$$

If at a time point t_k there are no observations available, only the background update step is performed.

In practical implementation of the filter, equation (3.104) may be used in the equivalent stabilized form:

$$\mathbf{P}_{n+1}^{a} = (\mathbf{I} - \mathbf{K}_{n+1}\mathbf{H})\mathbf{P}_{n+1}^{b}(\mathbf{I} - \mathbf{K}_{n+1}\mathbf{H})^T + \mathbf{K}_{n+1}\mathbf{R}_{n+1}\mathbf{K}_{n+1}^T \tag{3.106}$$

The advantage of this symmetric representation is that it guarantees positive semidefiniteness of $\mathbf{P}_{n+1}^{a}$ in the presence of the round off errors that usually accrue in computer simulations (Lewis et al., 2008).

The background and observation updates of the estimate can be combined into one equation

$$\mathbf{u}_{n+1}^{a} = (\mathbf{S}_n - \mathbf{K}_{n+1}\mathbf{H})\mathbf{u}_n^a + \mathbf{f}_n + \mathbf{K}_{n+1}\mathbf{y}_{n+1} \tag{3.107}$$

Another common equivalent formulation of the observation update step of the filter is defined in (Lewis et al., 2008) as

$$\mathbf{P}_{n+1}^{a} = \left[\left(\mathbf{P}_{n+1}^{b}\right)^{-1} + \mathbf{H}^T \mathbf{R}_{n+1}^{-1} \mathbf{H}\right]^{-1} \tag{3.108}$$

$$\mathbf{u}_{n+1}^{a} = \mathbf{u}_{n+1}^{b} + \mathbf{f}_n + \mathbf{P}_{n+1}^{a} \mathbf{H}^T \mathbf{R}_{n+1}^{-1}\left(\mathbf{y}_{n+1} - \mathbf{H}\mathbf{u}_{n+1}^{b}\right) \tag{3.109}$$

The computation of the analysis covariance by (3.108) requires fewer operations of the matrix multiplication then (3.103)-(3.104), although, it contains the additional operation of the matrix inversion.

3.2.3 Continuous Kalman filter

The time continuous and space-discrete problem (3.37)-(3.39) is discretised in time uniformly with a time interval of length T, and the time and space discrete formulation is derived in (3.67)-(3.70). The solution of the latter problem was deducted based on the estimated state variance optimisation and was written in a recurrence form known as the Kalman filter. That raises the question whether or not Kalman filtering can be reformulated in the time continuous manner.

The algebraic structure of the fully discrete system indeed simplifies the analysis in contrast to a differential one. For this reason, the ideas of the discrete Kalman filter cannot be directly applied to the continuous case. On the other hand, time discretisation of the continuous problem was done in terms of the approximation of the corresponding time integrals (see section 3.1.4). Thus, as the discretisation interval T goes to zero, the discrete model converges to the continuous one. Following on from that argument, the continuous Kalman filter may be derived from the time limit of the discrete counterpart.

In order to do that, the equations of the error covariance (3.103)-(3.104) to be manipulated are written, accounting for the error description (3.70), as

$$\mathbf{K}_{n+1} = \mathbf{P}_{n+1}^{b}\mathbf{H}^{T}\left(\mathbf{H}\mathbf{P}_{n+1}^{b}\mathbf{H}^{T} + \frac{\mathbf{R}(nT + T/2)}{T}\right)^{-1} \tag{3.110}$$

$$\mathbf{P}_{n+1}^{a} = (\mathbf{I} - \mathbf{K}_{n+1}\mathbf{H})\mathbf{P}_{n+1}^{b} \tag{3.111}$$

Dividing both parts by the size of the time interval T yields

$$\frac{1}{T}\mathbf{K}_{n+1} = \mathbf{P}_{n+1}^{b}\mathbf{H}^{T}\left(\mathbf{H}\mathbf{P}_{n+1}^{b}\mathbf{H}^{T}T + \mathbf{R}(nT + T/2)\right)^{-1} \tag{3.112}$$

The behaviour of the discrete Kalman gain $\mathbf{K}_{n+1}$ as T tends to zero is now examined:

$$\lim_{T\to\infty}\frac{1}{T}\mathbf{K}_{n+1} = \mathbf{P}_{n+1}^{b}\mathbf{H}^{T}\mathbf{R}^{-1}(nT + T/2) \tag{3.113}$$

This implies that

$$\lim_{T\to\infty} \mathbf{K}_{n+1} = 0 \tag{3.114}$$

which suggests that the discrete Kalman gain tends to zero when the sampling period becomes small. It is important to handle this fact properly when designing discrete Kalman filters for continuous-time systems. From (3.111), it also means that

$$\mathbf{P}_{n+1}^{a} = \mathbf{P}_{n+1}^{b} \tag{3.115}$$

as T goes to zero.

Inserting the expression (3.111) for $\mathbf{P}_n^a$ into the background covariance (3.101) produces

$$\mathbf{P}_{n+1}^{b} = \mathbf{S}_n(\mathbf{I} - \mathbf{K}_n\mathbf{H})\mathbf{P}_n^b\mathbf{S}_n^T + \mathbf{Q}(nT + T/2)T \tag{3.116}$$

$$\mathbf{P}_{n+1}^{b} = \mathbf{S}_n\mathbf{P}_n^b\mathbf{S}_n^T - \mathbf{S}_n\mathbf{K}_n\mathbf{H}\mathbf{P}_n^b\mathbf{S}_n^T + \mathbf{Q}(nT + T/2)T \tag{3.117}$$

The first term of the right-hand side is transformed as follows

$$\begin{aligned} \mathbf{S}_n\mathbf{P}_n^b\mathbf{S}_n^T &= (\mathbf{S}_n - \mathbf{I} + \mathbf{I})\mathbf{P}_n^b(\mathbf{S}_n^T - \mathbf{I} + \mathbf{I}) = \\ &= \Big((\mathbf{S}_n - \mathbf{I})\mathbf{P}_n^b + \mathbf{P}_n^b\Big)(\mathbf{S}_n^T - \mathbf{I} + \mathbf{I}) \\ &= (\mathbf{S}_n - \mathbf{I})\mathbf{P}_n^b(\mathbf{S}_n^T - \mathbf{I}) + (\mathbf{S}_n - \mathbf{I})\mathbf{P}_n^b + \mathbf{P}_n^b(\mathbf{S}_n^T - \mathbf{I}) + \mathbf{P}_n^b \end{aligned} \tag{3.118}$$

Substituting it back into (3.117) and moving $\mathbf{P}_n^b$ to the left side yields

$$\begin{aligned} \mathbf{P}_{n+1}^{b} - \mathbf{P}_n^b &= (\mathbf{S}_n - \mathbf{I})\mathbf{P}_n^b(\mathbf{S}_n^T - \mathbf{I}) + (\mathbf{S}_n - \mathbf{I})\mathbf{P}_n^b \\ &+ \mathbf{P}_n^b(\mathbf{S}_n^T - \mathbf{I}) - \mathbf{S}_n\mathbf{K}_n\mathbf{H}\mathbf{P}_n^b\mathbf{S}_n^T + \mathbf{Q}(nT + T/2)T \end{aligned} \tag{3.119}$$

The last equality is divided by T as $T \to 0$

$$\begin{aligned} \lim_{T\to 0}\frac{\mathbf{P}_{n+1}^{b} - \mathbf{P}_n^b}{T} = \lim_{T\to 0}\frac{1}{T}\Big((\mathbf{S}_n - \mathbf{I})\mathbf{P}_n^b(\mathbf{S}_n^T - \mathbf{I}) + (\mathbf{S}_n - \mathbf{I})\mathbf{P}_n^b \\ + \mathbf{P}_n^b(\mathbf{S}_n^T - \mathbf{I}) - \mathbf{S}_n\mathbf{K}_n\mathbf{H}\mathbf{P}_n^b\mathbf{S}_n^T + \mathbf{Q}(nT + T/2)T\Big) \end{aligned} \tag{3.120}$$

The limit on the left-hand side defines a time derivative of the matrix $\mathbf{P}(t)$. To compute the limits on the right-hand side, the discrete model matrix $\mathbf{S}_n$ given by (3.47) is expanded using Neumann series (3.60) so that

$$\mathbf{S}_n - \mathbf{I} = 2\left(\mathbf{I} + \frac{T}{2}\mathbf{S}(nT) + \frac{T^2}{4}\ \mathbf{S}^2(nT) + \cdots\right) - \mathbf{I} - \mathbf{I}$$
$$= T\mathbf{S}(nT) + \frac{T^2}{2}\ \mathbf{S}^2(nT) + \cdots \tag{3.121}$$

From the expression (3.121) the following limits are computed

$$\lim_{T\to 0}\frac{(\mathbf{S}_n - \mathbf{I})}{T} = \lim_{T\to 0}\left(\mathbf{S}(nT) + \frac{T}{2}\mathbf{S}^2(nT) + \cdots\right) = \mathbf{S}(t) \tag{3.122}$$

$$\lim_{T\to 0}\frac{(\mathbf{S}_n - \mathbf{I})\mathbf{P}_n^b(\mathbf{S}_n^T - \mathbf{I})}{T}$$
$$= \lim_{T\to 0}\left(\mathbf{S}(nT) + \frac{T}{2}\mathbf{S}^2(nT) + \cdots\right)\mathbf{P}_n^b\left(T\mathbf{S}(nT)\right. \tag{3.123}$$
$$\left. + \frac{T^2}{2}\ \mathbf{S}^2(nT) + \cdots\right) = \mathbf{S}(t)\mathbf{P}_n^b 0 = 0$$

Assuming that $t = nT$ and accounting for (3.113) and (3.122)-(3.123) in (3.120) produces the continuous equation for the estimation error covariance

$$\dot{\mathbf{P}}(t) = \mathbf{S}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{S}^T(t) + \mathbf{Q}(t) - \mathbf{P}(t)\mathbf{H}\mathbf{R}^{-1}\mathbf{H}^T\mathbf{P}(t) \tag{3.124}$$

Here $\mathbf{P}(t)$ denotes error covariance of the time continuous estimate. Due to (3.115) the continuous covariance is not separated into the background and analysis parts as in the discrete case.

In a similar fashion, the discrete equation for the estimate is written as

$$\frac{\mathbf{u}_{n+1} - \mathbf{u}_n}{T} = \frac{\mathbf{S}_n - \mathbf{I}}{T}\mathbf{u}_n + \frac{\mathbf{f}_n}{T} + \frac{\mathbf{K}_{n+1}}{T}(\mathbf{y}_n - \mathbf{H}\mathbf{u}_n) \tag{3.125}$$

Taking the limit of T as it approaches zero and using limits (3.113), (3.121) and (3.122) returns the ODE referred to as the feedback equation for the time continuous state estimate

$$\begin{cases} \dfrac{d\mathbf{u}(t)}{dt} = \mathbf{S}(t)\mathbf{u}(t) + \mathbf{f}(t) + \mathbf{P}(t)\mathbf{H}\mathbf{R}^{-1}(t)\big(\mathbf{y}(t) - \mathbf{H}\mathbf{u}(t)\big) \\ \mathbf{u}(0) = \mathbf{u}^0 + \mathbf{e}_0 \end{cases} \tag{3.126}$$

where the covariance $\mathbf{P}(t)$ is the solution of the matrix differential Riccati equation

$$\begin{cases} \dfrac{d\mathbf{P}(t)}{dt} = \mathbf{S}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{S}^T(t) + \mathbf{Q}(t) - \mathbf{P}(t)\mathbf{H}\mathbf{R}^{-1}\mathbf{H}^T\mathbf{P}(t) \\ \mathbf{P}(0) = \mathbf{Q}_0 \end{cases} \tag{3.127}$$

The set of the equations (3.126)-(3.127) defining the time continuous Kalman filter are also called the Kalman-Bucy filter.

To solve the Kalman-Bucy filter, it again is discretised in time. The discretised Riccati equation is in the form similar to (3.119), but without additional term that according to (3.123) vanishes as T goes to zero. On a discrete level, however, and correspondingly in a practical implementation, they result in different estimates from the Kalman and the Kalman-Bucy filters.

Another conceptual difference between discrete and continuous filters applied to the time continuous DA problem is that in the case of Kalman filter the problem is discretised first and then optimised, while in Kalman-Bucy case, the problem is optimised first and then the continuous optimal estimate is discretised to be computed. Consequently, it is expected that the Kalman-Bucy filter produces better quality estimates than its discrete counterpart.

Finally, for the linear system (3.37), the Riccati equation (3.127) is the simplified FP equation, i.e., FP equation without a quadratic term and reflects the Bayesian assimilation of the observed data (Jazwinski, 1970).

3.3 Ensemble Kalman filters

The Kalman-Bucy and Kalman filters provide a significant computational advantage compared to the FP method. Conversely, the FP analysis is a more general approach in contrast to the Kalman filters, which have several design limitations. First of all, they are defined for linear problems. In the case of a nonlinear problem, the algorithm is instead applied to the linearized version of the background equation, an approach known as the EKF (Lewis et al., 2008). It performs well for moderately nonlinear systems, but because a nonlinear transition model does not maintain the Gaussian structure of the background distribution (3.77), the resulting estimate loses its optimality.

A second disadvantage is the sequential nature of the Kalman algorithms. For both the Kalman-Bucy and the Kalman filters, the estimate is obtained after a fixed amount of operations are performed. In a practical implementations of DA for marine applications where computational demands are usually very high, it is desired to have a more robust algorithm that can balance the computational speed and the quality of the estimate in an optimal way.

Over the last few years, ensemble Kalman Filters (EnKF) (Evensen, 2009) have become popular methods in the DA community. The classical Kalman filter computes Bayes' inference (3.76) of the background state and observations in an algebraic way resulting in (3.103)-(3.105). In contrast, the EnKF provides a Monte Carlo implementation of the Bayesian update (3.76).

EnKF represents the distribution of the system state using an ensemble

$$\mathbf{U}_n = [\mathbf{u}_n^1, \mathbf{u}_n^2, \dots, \mathbf{u}_n^K] \tag{3.128}$$

which is comprised of model states sampled around the average, and replaces the covariance matrix by the sample covariance computed from the ensemble. The main advantage of the EnKF algorithm is that the uncertainties of the system are quantified in the space spanned over ensemble members and advancing the PDF in time is achieved by advancing each member of the ensemble.

In more detail, the algorithm starts with the selection of the initial ensemble

$$\mathbf{U}_0^a = \mathbf{U}_0 = [\mathbf{u}^{0,1} + \mathbf{e}_0^1, \mathbf{u}^{0,2} + \mathbf{e}_0^2, \dots, \mathbf{u}^{0,K} + \mathbf{e}_0^K] \tag{3.129}$$

where $\mathbf{e}_0^i$ is a random realization of the initial condition error.

The background model then updates the ensemble and its average

$$\mathbf{U}_{n+1}^b = \mathbf{S}_n \mathbf{U}_n^a + \mathbf{f}_n \tag{3.130}$$

$$\overline{\mathbf{U}}_{n+1}^b = \frac{1}{K} \sum_{s=1}^{K} \mathbf{u}_{n+1}^{b,s} \tag{3.131}$$

The observation equation updates observation ensemble and observation average

$$\mathbf{Y}_{n+1}^b = \mathbf{H}\mathbf{U}_{n+1}^b \tag{3.132}$$

$$\overline{\mathbf{Y}}_{n+1}^b = \frac{1}{K} \sum_{s=1}^{K} \mathbf{H}\mathbf{u}_{n+1}^{b,s} \tag{3.133}$$

Next, the perturbation matrices of the ensembles are computed

$$\widetilde{\mathbf{U}}_{n+1} = \mathbf{U}_{n+1} - \overline{\mathbf{U}}_{n+1}^b \mathbf{1}_{1\times K} \tag{3.134}$$

$$\widetilde{\mathbf{Y}}_{n+1} = \mathbf{Y}_{n+1} - \overline{\mathbf{Y}}_{n+1}^{b}\mathbf{1}_{1\times K} \tag{3.135}$$

The simplest formulation of EnKF is obtained if the covariance matrix is computed as

$$\mathbf{C}_{n+1} = \frac{\widetilde{\mathbf{U}}_{n+1}\widetilde{\mathbf{U}}_{n+1}^{T}}{K-1} \tag{3.136}$$

Based on that, the ensemble propagation is achieved by applying Kalman update (3.105) for each member of the ensemble with the ensemble covariance matrix

$$\begin{aligned}\mathbf{U}_{n+1}^{a} = \mathbf{U}_{n+1}^{b} + \mathbf{C}_{n+1}\mathbf{H}_{n+1}^{T}(\mathbf{H}_{n+1}\mathbf{C}_{n+1}\mathbf{H}_{n+1}^{T} + \mathbf{R})^{-1}\big(\mathbf{D}_{n+1} \\ - \mathbf{H}_{n+1}\mathbf{U}_{n+1}^{b}\big)\end{aligned} \tag{3.137}$$

where $\mathbf{D} = \mathbf{y}_{n+1}\mathbf{1}_{1\times K}$ is the matrix of the actual observations.

From the central limit theorem, it is known that the ensemble covariance (3.136) converges to the actual error covariance as fast as $1/\sqrt{K}$ (Kelly et al., 2014). That means the EnKF estimate is equivalent to the Kalman filter estimate in the limit of the ensemble size. However, the finite size ensemble tends to underestimate the covariance of the ensemble comparing to the covariance computed by the standard Kalman Filter. As a result, the influence of the observations on the state estimate decreases and for a long integration window may lead to the desynchronization between the estimate and the true state.

An ensemble method that deals with the covariance inflation problem and has attracted a substantial amount of attention is ensemble transform Kalman Filter (EnTKF) (Hunt et al., 2007). In this case, covariance is computed similarly as in the square-root filter

$$\mathbf{P}_{n+1}^{a} = \left[(K-1)\mathbf{I}\frac{1}{\rho} + \widetilde{\mathbf{Y}}_{n+1}^{T}\mathbf{R}_{n+1}^{-1}\widetilde{\mathbf{Y}}_{n+1}\right]^{-1} \tag{3.138}$$

$$\mathbf{W} = [(K-1)\mathbf{P}_{n+1}^{a}]^{\frac{1}{2}} \tag{3.139}$$

where ρ denotes an inflation parameter chosen empirically and the $1/2$ power in the second expression denotes the square root of a symmetric matrix.

Finally, the ensemble propagation equation becomes

$$\begin{aligned}\mathbf{U}_{n+1}^{a} = \overline{\mathbf{U}}_{n+1}^{b}\mathbf{1}_{1\times K} \\ + \widetilde{\mathbf{U}}_{n+1}\left(\mathbf{W} + \mathbf{P}_{n+1}^{a}\widetilde{\mathbf{Y}}_{n+1}^{T}\mathbf{R}_{n+1}^{-1}\left(\mathbf{y}^{o} - \overline{\mathbf{y}}_{n+1}^{b}\right)\mathbf{1}_{1\times K}\right)\end{aligned} \tag{3.140}$$

The main idea of the ensemble Kalman filtering is to quantify the uncertainty in the space derived by ensemble members. If the size K of the ensemble is much smaller then M the size of the model space, the calculations of the covariance matrix and state estimate are performed with significant computational advantage. On the other hand, approximation of the model space by the much smaller ensemble space may lead to the loss of accuracy. During the analysis step, an optimal estimate is constructed based on the available observations as the linear combination of the ensemble members, so that, the ensemble space should be large enough to approximately span all possible states of the system.

In order to improve the robustness of the EnTKF, a localisation procedure is applied. The idea of the localisation is to split the global space of uncertainties into a number of local spaces and conduct analysis for each of them separately. In doing this, the ensemble members in different regions would have to span only over the space of local states and, as a result, the global state space is approximated in a piecewise manner. In other words, localisation allows for the reduction of the size of the ensemble even further, while maintaining the same level of the accuracy.

Hunt et al. (2007) presented a localisation procedure for batch processing of observations. The authors suggest partitioning of available observations into separate batches while adhering to the following restrictions:

1) All of the observations in a given batch must be used in the same subset of the local analyses.
2) Observations in different batches must have uncorrelated errors so that each batch corresponds to a block in a block-diagonal decomposition of the observation error covariance matrix. For efficiency reasons, one should make the batches as large as possible while still meeting the restriction of the first condition.

Another benefit of localisation is that computations of each local covariance are independent of each other and thus can be organised in parallel.

3.4 Summary

This chapter presents in detail a basic framework for solving contaminant transport problem based on Galerkin FEM approximations that would be used later for combining DA with DD. In more details three of the most common forms of the contaminant transport problem are

derived. Beginning with the continuous formulation described by the PDE, two main approximations were explained: the spatial discretisation performed by the FEM method and the temporal discretisation obtained by the midpoint integration rule.

Based on the background model representation, several conventional filters are introduced: for a time continuous problem, the Kalman-Bucy filter is devised and for a discrete problem, algorithms of the Kalman and EnKF filters were obtained. The equivalence between different filters is presented as well.

4 Development of a minimax filter

The statistical methods of DA rely on exact descriptions of model uncertainties. Those methods require the Gaussian white noise assumption about the error structure and perform optimisation of the parameters of the corresponding Gaussian distribution. The quality of the estimation produced by those filters is sensitive to the accuracy of the definition of the input parameters, and if the probability law is arbitrary or unknown little can be said about the system state estimate (Jazwinski, 1970).

The accuracy of the error description can be relaxed if those errors are considered to be unknown but bounded. The latter attribute is a natural assumption as it reflects the fact that the energy of the system is limited. Once the set of all possible error realisations is known, the state estimate can be found as the realisation that minimizes the error of the worst-case scenario. This approach of DA is known as *minimax filtering*.

In this chapter, a minimax filter is developed in a feedback form following an approach presented by Zhuk (2013) with two extensions. The first extension is that model errors are assumed to have linear constraints defined by a L^∞-type norm, while typically those errors are considered to have quadratic constraints governed by a L^2-type norm. The L^∞ constraints are easier to define then the L^2 constraints and thus more desired in practice. The second extension is that the filter is defined for a continuous DA problem in an operator form and later discretised using Galerkin and FEM approximations. This will be used in the following chapters for a derivation of a localised version of a minimax filter. Another benefit of this extension is that it follows the paradigm "optimise and discretise" in contrast to "discretise and optimise".

4.1 Filter derivation

The background model for the continuous DA problem for the concentration transport process is defined on the spatiotemporal domain ($[t_0; t_1] \times \Omega$) and is given by the advection-diffusion equation with constant diffusion coefficient ϵ and velocity flow field $\mu = [\mu_1, \mu_2]^T$

$$\begin{cases} \dfrac{\partial u(t,x)}{\partial t} = \epsilon \Delta u(t,x) - \nabla \cdot \big(\mu(t,x) u(t,x)\big) + e_m(t,x) \\ u(0,x) = u^0(x) + e_0(x) \\ u(t,x) = 0, \qquad x \in \partial\Omega \end{cases} \tag{4.1}$$

The observation equation is written as

$$y(t,x) = Hu(t,x) + e_o(t,x) \tag{4.2}$$

where H is a linear operator

$$Hu(t,x) = \int_{\Omega} h(z-x)u(t,z)dz \tag{4.3}$$

In the following, DA problem (4.1)-(4.3) is solved using deterministic treatment of errors.

4.1.1 Representation of errors

The fact that the model error $e_m(t,x)$ and initial condition error $e_0(x)$ are uncertain but bounded can be reflected by assuming that they are taken from the given convex bounded set

$$\mathcal{L}_m^{\infty} = \{e_0(x), e_m(t,x): |e_0(x)| \leq \tilde{q}_0(x), |e_m(x)| \leq \tilde{q}(t,x)\} \tag{4.4}$$

where $\tilde{q}_0(x)$ and $\tilde{q}(t,x)$ are given weighting functions such that

$$0 < \underline{q} < \tilde{q}(t,x) < \overline{q} < +\infty \tag{4.5}$$

$$0 < \underline{q}_0 < \tilde{q}_0(x) < \overline{q}_0 < +\infty \tag{4.6}$$

and $\underline{q}$, $\overline{q}$, $\underline{q}_0$, $\overline{q}_0$ are known real numbers.

The functions $q(t,x)$ and $q_0(x)$ should be considered as design parameters that quantify a level of confidence in the initial condition and state equation. The function $\tilde{q}_0(x)$ specifies zones of Ω where knowledge of the initial condition $u^0(x)$ is either more precise or less. The time-varying function $\tilde{q}(t,x)$ defines zones of Ω where the background equation (4.1) holds almost exactly if $|e_m| \approx 0$ in that zone or only up to a significant error if $|e_m| \gg 0$. Statistically this corresponds to the uniform distribution of the elements from the set $\mathcal{L}_m^{\infty}$, i.e., any combination of $(e_0; e) \in \mathcal{L}_m^{\infty}$ has equal probability to appear in (4.1).

It is also assumed that the observation error $e_o(t,x)$ is white noise with bounded second moments where $e_o(t,x)$ is an element of the following convex bounded set

$$\mathcal{L}_o^{\infty} = \left\{e_o(t,x): \overline{e_o(t,x)} = 0, \overline{e_o^2(t,x)\tilde{r}(t,x)} \leq 1\right\} \tag{4.7}$$

and $\tilde{r}(t,x)$ is a given function such that

$$0 < \underline{r} < \tilde{r}(t,x) < \overline{r} < +\infty \tag{4.8}$$

for known $\underline{r}$ and $\overline{r}$.

In a similar fashion, the function $\tilde{r}(t,x)$ introduces asymptotical flexibility into the description of the observation uncertainties. Note that equation (4.7) does not imply any restrictions on the structure of the distribution of $e_o(t,x)$.

The bounded set (4.4) is further transformed into a set of quadratic constraints and it is straightforward to verify that

$$e_0^2(x)\frac{1}{\tilde{q}_0^2(x)} \leq 1;\ e_m^2(x)\frac{1}{\tilde{q}^2(t,x)} \leq 1 \tag{4.9}$$

Integrating the first expression over time interval $[t_0; t_1]$ and the second expression over spatio-temporal domain $([t_0; t_1] \times \Omega)$ yields

$$\int_\Omega \frac{e_0^2(x)}{\tilde{q}_0^2(x)} dx \ \leq A(\Omega) \tag{4.10}$$

$$\int_{t_0}^{t_1} \int_\Omega \frac{e_m^2(t,x)}{\tilde{q}^2(t,x)} dx\, dt \leq (t_1 - t_0)A(\Omega) \tag{4.11}$$

where $A(\Omega) = \int_\Omega d\Omega$ is the area of the spatial domain Ω.

Denoting $T = t_1 - t_0$ and combining the last two inequalities returns

$$\int_\Omega \frac{e_0^2(x)}{\tilde{q}_0^2(x)} dx + \int_{t_0}^{t_1} \int_\Omega \frac{e_m^2(t,x)}{\tilde{q}^2(t,x)} dx\, dt \leq (T+1)A(\Omega) \tag{4.12}$$

With the introduction of the new weighting functions

$$q_0(x) = \frac{1}{\gamma_T \tilde{q}_0^2(x)}, \qquad q(x) = \frac{1}{\gamma_T \tilde{q}^2(x)} \tag{4.13}$$

where $\gamma_T = (T+1)A(\Omega)$, the set (4.4) is rewritten in the form of quadratic ellipsoid:

$$\begin{aligned} \mathcal{L}_m^2 = \Big\{ e_0(x), e_m(t,x) : \int_\Omega e_0^2(x) q_0(x)\, dx \\ + \int_{t_0}^{t_1} \int_\Omega e_m^2(t,x) q(t,x) dx\, dt \leq 1 \Big\} \end{aligned} \tag{4.14}$$

Applying a similar transformation to the observation set (4.7), one gets a quadratic ellipsoid for the observation error:

$$\mathcal{L}_o^2 = \left\{ e_o(t,x) : \overline{e_o(t,x)} = 0, \int_{t_0}^{t_1} \int_{\Omega} \overline{e_o^2(t,x) r(t,x)} dx\, dt \leq 1 \right\} \tag{4.15}$$

where the new weighting function is defined as

$$r(t,x) = \frac{\tilde{r}(t,x)}{\gamma_T} \tag{4.16}$$

The bounded set $\mathcal{L}_m^\infty$ is derived by the L^∞-type norm, and the set $\mathcal{L}_m^2$ is derived by the L^2-type norm. The latter fact explains the usage of the superscripts ∞ and 2 in the notations of the error ellipsoids. It also gives an interesting geometrical interpretation: due to the nature of L^∞-type norm and L^2-type norm, the shape of the $\mathcal{L}_m^\infty$ set is rectangular and the shape of the $\mathcal{L}_m^2$ is ellipsoidal. The level set of L^∞-type norm in the 2-dimensional Euclidean space:

$$L = \left\{ l = (l_1, l_2)^T : \max_i |l_i| \leq 1 \right\} \tag{4.17}$$

is a rectangle, and the level set of L^2-type norm is a circle (ellipsoid), i.e.,

$$L = \left\{ l = (l_1, l_2)^T : \sum_{i=1}^{2} l_i^2 \leq 1 \right\} \tag{4.18}$$

Following this logic, the sets $\mathcal{L}_m^\infty$ and $\mathcal{L}_o^\infty$ would be referred as the L^∞ model and observation rectangles and the sets $\mathcal{L}_m^2$ and $\mathcal{L}_o^2$ as the L^2 model and observation ellipsoids. Approximating $\mathcal{L}_m^\infty$ and $\mathcal{L}_o^\infty$ by $\mathcal{L}_m^2$ and $\mathcal{L}_o^2$ should be regarded as approximating a rectangle by an ellipsoid that contains it and the discrepancy introduced by that approximation is reflected by γ_T.

A hybrid treatment of the system uncertainties implies that the model error is considered to be deterministic while the observation error is stochastic and the constraints are imposed on its statistical moments. Additionally, it is possible to consider the observation error to be a deterministic and unknown but bounded variable. In that case, the $\mathcal{L}_o^\infty$ and $\mathcal{L}_o^2$ sets would be written in an identical way to the $\mathcal{L}_m^\infty$ and $\mathcal{L}_m^2$ sets, but without the term corresponding to the initial error. If both model and observation error are deterministic, they are bounded either by a general rectangle ($\mathcal{E}_m^\infty$) or an ellipsoid ($\mathcal{E}_m^2$):

$$\begin{aligned} \mathcal{E}_m^\infty = \Big\{ & e_0(x), e_m(t,x), e_m(t,x) : |e_0(x)| \leq \tilde{q}_0(x), \\ & |e_m(x)| \leq \tilde{q}(t,x), |e_o(t,x)| \leq \frac{1}{\tilde{r}(t,x)} \Big\} \end{aligned} \tag{4.19}$$

$$\mathcal{E}_m^2 = \left\{ e_0(x), e_m(t,x), e_o(t,x) : \int_\Omega e_0^2(x) q_0(x)\, dx \right.$$
$$+ \int_{t_0}^{t_1} \int_\Omega e_m^2(t,x) q(t,x) dx\, dt \tag{4.20}$$
$$\left. + \int_{t_0}^{t_1} \int_\Omega e_o^2(t,x) r(t,x) dx\, dt \leq 1 \right\}$$

In this case, observation and model errors do not possess independence of each other. For instance, the increase or decrease of the model error parameters also allows for, mathematically, the increase or decrease of the observation errors. Thus, in this research the hybrid description (4.14)-(4.15) is taken.

The main advantage of the representation of the uncertainties in the form of (4.4)-(4.7), (4.14)-(4.15) or (4.19)-(4.20) comparing to the entirely probabilistic or Gaussian description is that it does not require precise knowledge about stochastic variables, but only defines its possible boundaries in an asymptotic way which in practice is more reliable.

4.1.2 Optimal control problem

Since the bounded set $\mathcal{L}_m^\infty$ is deterministic, by applying the background propagation model (4.1) to all possible combinations $(e_0, e_m) \in \mathcal{L}_m^\infty$ one obtains a set of trajectories or solutions $\mathcal{R}$. It follows that the background maps the sets $\mathcal{L}_m^\infty$ onto $\mathcal{R}$. The observation equation (4.2) then selects all of the trajectories from $\mathcal{R}$ that are compatible with the observed data $y(t)$ up to the observation error described as an element $e_o(t) \in \mathcal{L}_o^\infty$. A section of $\mathcal{R}$, taken at time instant t_1, $\mathcal{R}(t_1)$ is called a reachability set. $\mathcal{R}(t_1)$ is comprised of all functions v such that $u(t_1, x) = v(x)$ where $u(t,x)$ satisfies (4.1) for some $(e_0, e_m) \in \mathcal{L}_m^\infty$, and u verifies (4.2) for the given observed data $y(t)$, $t \in (t_0, t_1)$ and some $e_o(t,x) \in \mathcal{L}_o^\infty$. The reachability set $\mathcal{R}(t_1)$ contains full information about all of the possible states of the model (4.1) which are compatible with the uncertainties treatment $\mathcal{L}_m^\infty, \mathcal{L}_o^\infty$ and data $y(t)$, $t \in (0, t_1)$.

The question is then how to select an analysis state $u^a(t_1)$ that should be considered to be the best with respect to some criteria estimate of the system (4.1)-(4.2) from the set of admissible states $\mathcal{R}(t_1)$. One way of doing this is to construct the minimax centre of the reachability set $\mathcal{R}(t_1)$. The minimax centre of $\mathcal{R}(t_1)$ is called a component of $\mathcal{R}(t_1)$ which minimizes the

largest distance to other components of $\mathcal{R}(t_1)$. In other words $u^a(t_1)$ should satisfy the following

$$\sup_{u(t_1)\in\mathcal{R}(t_1)} \ell\big(u^a(t_1) - u(t_1)\big) \leq \sup_{\tilde{u}(t_1),u(t_1)\in\mathcal{R}(t_1)} \ell\big(\tilde{u}(t_1) - u(t_1)\big) \tag{4.21}$$

where ℓ denotes the distance between elements in $\mathcal{R}(t_1)$ and is given in the form of the linear functional

$$\ell(u) = \int_\Omega l(x)u(t_1,x)dx \tag{4.22}$$

The estimation error represented by the left-hand side of the inequality (4.21) is given by the diameter of the reachability set $\mathcal{R}(t_1)$.

Mathematically, the estimation problem is described as the optimal control problem to find $u^a(t_1)$, which minimizes the convex cost function (4.21) for a given realization of observations (4.2) and differential constraints (4.1).

In the study by Zhuk (2009) it has been shown that to find the solution of this control problem one needs to solve Euler-Lagrange equations, which, in particular, implies that, to compute $u^a(t_2)$ for $t_2 > t_1$, one needs to solve Euler-Lagrange equation for $t \in (t_0, t_2)$ as $u^a(t_2)$ cannot be expressed as a function of $u^a(t_1)$ and observations $y(t)$, $t \in (t_1, t_2]$. In other words, the estimate $u^a(t)$ does not possess the Markovian property. The reason for this is as follows: $\mathcal{L}_m^\infty$ is a level set of the $L^\infty(\Omega)$-norm, and the dual norm of the latter is given by $L^1(\Omega)$-norm. Hence, $\mathcal{L}_m^\infty$ does not coincide with its dual set. This is obvious in the case of two-dimensional Euclidean space where $\mathcal{L}_m^\infty$ corresponds to a rectangle and its dual will be a rhombus. On the other hand, L^2-norm coincides with its dual norm and this property of the norm is necessary and sufficient to get a Markovian estimate (see for instance (Zhuk, 2009)).

The lack of the Markovian property of the estimate for L^∞ uncertainties does not allow for the construction of a filter in a feedback form. At the same time, recalculation of the Euler-Lagrange equations over the whole interval $[t_0, t]$ when each new observations $y(t)$ became available introduces significant redundancy in contrast to the desired feedback representation.

The L^2 type description of the uncertainties allows the construction of a suboptimal Markovian estimate, provided $(e_0, e_m) \in \mathcal{L}_m^2$ and $e_o \in \mathcal{L}_o^2$. In that case, the estimate $u^a(t)$ of $u(t)$ with minimal mean-squared estimation error

$$\sup_{(e_0,e_m)\in\mathcal{L}_m^2, e_o\in\mathcal{L}_o^2} \ell\big(u^a(t_1)-u(t_1)\big) \leq \sup_{(e_0,e_m)\in\mathcal{L}_m^2, e_o\in\mathcal{L}_o^2} \ell\big(v(y)-u(t_1)\big) \tag{4.23}$$

where $v(y)$ is the functional of observations, admits the following representation

$$\ell(u^a) = \int_{\Omega_T} r(t,x)(Hp)(t,x)y(t,x)dxdt \tag{4.24}$$

provided p and z solve the following Hamiltonian system of equations:

$$\begin{cases} \dfrac{\partial w}{\partial t} = -L^* w + H^* rHp \quad \text{in } [t_0,t_1]\times\Omega \\ w(t_1,x) = l(x) \quad \text{on } \Omega \\ w(t,x) = 0 \quad \text{on } [t_0,t_1]\times\partial\Omega \\ \dfrac{\partial p}{\partial t} = Lp + qw \quad \text{in } [t_0,t_1]\times\Omega \\ p(t_0,x) = q_0(x)w(0,x) \quad \text{on } \Omega \\ p(t,x) = 0 \quad \text{on } [t_0,t_1]\times\partial\Omega \end{cases} \tag{4.25}$$

where L^* and H^* are the adjoint of L and H.

The relation between w and p is expressed by the linear mapping $P: L^2(\Omega) \to L^2(\Omega)$ such that

$$p(t,x) = (Pw)(t,x) = \int_\Omega k(t,x,z)w(z)dz \tag{4.26}$$

Operator P represents a bounded integral operator and $k(t,x,z)$ is its kernel function continuous in time.

Inserting (4.26) into the equation for p in (4.25) produces

$$\left(\frac{\partial}{\partial t}(Pw)\right)(t,x) = (LPw)(t,x) + qw \tag{4.27}$$

and respectively

$$\left(\left(\frac{\partial}{\partial t}P\right)w\right)(t,x) + \left(P\left(\frac{\partial}{\partial t}w\right)\right)(t,x) = (LPw)(t,x) + qw \tag{4.28}$$

The term $\frac{\partial}{\partial t}w$ is obtained from the first equation of the Hamiltonian system so that

$$\left(\left(\frac{\partial}{\partial t}P\right)w\right)(t,x) + \left((-PL^* + PH^*rHP)w\right)(t,x) = (LPw)(t,x) + qw(t,x) \tag{4.29}$$

$$\left(\frac{\partial}{\partial t}P\right)w(t,x) = (LP)w(t,x) + (PL^*)w(t,x) + qw(t,x) - (PH^*rHP)w(t,x) \tag{4.30}$$

Equation (4.30) is known as the operator Riccati equation for the operator P, which is referred to in the following as the Riccati operator. Equation (4.30) can be further represented for the kernel function $k(t,x,z)$ as follows

$$\begin{cases} \dot{k} = L_x k + L_z k + q^{-1}(t,x)\delta(z-x) - P\left(H^*r(Hk)\right) \\ k(0,x,z) = q_0^{-1}(x)\delta(z-x) \text{ on } \Omega \end{cases} \tag{4.31}$$

where $L_x k$ denotes the result of an application of L to k with respect to the variable x and δ is the Dirak delta function. To distinguish equation (4.31) from (4.30), (4.31) is referred to as the kernel Riccati equation.

Accounting for (4.26) and (4.30), the estimate or analysis $u^a(t)$ is represented as a filter

$$\begin{cases} \dot{u} = Lu + PH^*r(y - Hu) \text{ in } [t_0, t_1] \times \Omega \\ u = 0 \text{ on } [t_0, t_1] \times \partial\Omega \\ u(0) = 0 \text{ on } \Omega \end{cases} \tag{4.32}$$

For any linear continuous functional $l: L^2(\Omega) \to R$ the estimate $u^a(t)$ computed from (4.32) satisfies

$$\begin{aligned} &\sup_{(e_0,e_m)\in\mathcal{L}_m^2, e_o\in\mathcal{L}_o^2} \left(\ell\left(u^a(t_1)\right) - \ell(u(t_1))\right)^2 \\ &\leq \sup_{(e_0,e_m)\in\mathcal{L}_m^2, e_o\in\mathcal{L}_o^2} \left(v(y) - \ell\left(u^a(t_1)\right)\right)^2 \\ &= \int_\Omega l(x)(Pl)(t,x)dx \end{aligned} \tag{4.33}$$

for any continuous function $v: L^2(\Omega) \to R$. In other words, for any $l \in L^2(\Omega)$ the worst-case mean-squared error of $\ell(u^a)$ (the left-hand side of (4.33)) is minimal in the class of all continuous functionals of measurement $v(y)$. Equations (4.31)-(4.32) represent a so-called minimax filter and its solution $u^a(t_1)$ will be referred to as a minimax estimate.

4.1.3 Discretisation of a continuous minimax filter

Continuing the solution method introduced in section 4.1.2, the error ellipsoids (4.14)-(4.15) and minimax filter equations (4.31)-(4.32) are discretised in space using a FEM approximation approach. Assuming that a set of nodes x_i, $i = 1, \dots, N_{nd}$ comprises a FEM discretisation grid, a vector of the nodal basis function

$$\boldsymbol{\phi}(x) = [\phi_i(x);\ i = 1, \dots, N_{nd}]^T \tag{4.34}$$

is defined at each node x_i, such that

$$\phi_i(x_j) = \delta_{ij} \tag{4.35}$$

This definition is used to approximate a continuous function $u(t, x)$:

$$\begin{aligned} u(t,x) &\approx \sum_{i=1}^{N_{nd}} u(t, x_i)\, \phi_i(x) = \sum_{i=1}^{N_{nd}} \mathbf{u}_i(t)\phi_i(x) \\ &= \boldsymbol{\phi}^T(x)\mathbf{u}(t) = \mathbf{u}^T(t)\boldsymbol{\phi}(x) \end{aligned} \tag{4.36}$$

If the function $u(t, x)$ is known, owing to the interpolation property of FEM basis functions (4.35), coefficients of the expansion $\mathbf{u}_i(t)$ are easily computed using $\mathbf{u}_i(t) = u(t, x_i)$. A two variable function $k(t, x, z)$ is approximated in both spatial directions x and z and is written as

$$\begin{aligned} k(t,x,z) &\approx \sum_{i=1}^{N_{nd}} \sum_{j=1}^{N_{nd}} k(t,x,z)\phi_i(x)\phi_j(z) \\ &= \sum_{i=1}^{N_{nd}} \sum_{j=1}^{N_{nd}} \mathbf{K}_{ij}(t)\phi_i(x)\phi_j(z) = \boldsymbol{\phi}^T(x)\mathbf{K}(t)\boldsymbol{\phi}(z) \end{aligned} \tag{4.37}$$

where K is kernel matrix. A particular case of the discretisation is considered for Dirac delta function $\delta(x)$. Defining $f(x, z) = g(x, z)\delta(x - z)$, applying the definition of the Dirac delta function and FEM approximation (4.36) gives

$$\int_\Omega g(x,z)\delta(x-z)dx = g(z,z) \approx \boldsymbol{\phi}^T(z)\mathbf{G}(t)\boldsymbol{\phi}(z) \tag{4.38}$$

where $\mathbf{G}(t) = \{g(z_i, z_j)\}_{i,j=1}^{N_{nd}}$. On the other hand, recalling that $\mathbf{M} = \int_\Omega \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)dx$ is a mass matrix:

$$\int_{\Omega} g(x,z)\delta(x-z)dx \approx \int_{\Omega} \boldsymbol{\phi}^T(z)\mathbf{G}(t)\boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\mathbf{D}\boldsymbol{\phi}(z)dx$$
$$= \boldsymbol{\phi}^T(z)\mathbf{GMD}(t)\boldsymbol{\phi}(z) \tag{4.39}$$

Although the entries $\mathbf{D}_{ij} = \boldsymbol{\delta}(x_i - z_j)$ of the matrix $\mathbf{D}$ cannot be defined in the classical meaning, combining (4.38) and (4.39) results in

$$\mathbf{MD} = \mathbf{I}, \mathbf{D} = \mathbf{M}^{-1} \tag{4.40}$$

which means that the Dirac delta function in a FEM discretised space is equivalent to the inverted mass matrix.

Using the expansion (4.36), the model error is discretised as

$$e_m(t,x) \approx \boldsymbol{\phi}^T(x)\mathbf{e}_m(t) \tag{4.41}$$

Here each component of the model error vector $\mathbf{e}_m(t) = [\mathbf{e}_{m,i}(t) = e_m(t,x_i)]$ describes uncertainties of the associated component of the discrete analysis vector $\mathbf{u}_m^a(t)$. Taking that discretisation into account, the integral of the weighted quadratic model error is approximated by

$$\int_{\Omega} e_m^2(t,x)q(t,x)dx = \int_{\Omega} e_m(t,x)\int_{\Omega} q(t,z)\delta(x-z)e_m(t,z)\,dx$$
$$\approx \int_{\Omega} \mathbf{e}_m^T(t)\boldsymbol{\phi}(x)\int_{\Omega} \boldsymbol{\phi}^T(x)\,\widetilde{\mathbf{Q}}(t)\mathbf{M}^{-1}\boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z)\mathbf{e}_m(t)dx \tag{4.42}$$

where

$$\widetilde{\mathbf{Q}}(t) = \mathrm{diag}\{q(t,z_i)\}_{i=1}^{N_{nd}} \tag{4.43}$$

Substituting the definition of the mass matrix into (4.42) generates an approximation of the error integral as

$$\int_{\Omega} e^2(t,x)q(t,x)dx \approx \mathbf{e}_m^T(t)\mathbf{M}\widetilde{\mathbf{Q}}(t)\mathbf{e}_m(t) \tag{4.44}$$

Applying the same approximation as in (4.41)-(4.44) to the initial error and observation error, the ellipsoids $\mathcal{L}_m^2$ and $\mathcal{L}_o^2$ discretised by the FEM basis functions are obtained

$$\mathbb{L}_m^2 = \left\{\mathbf{e}_0, \mathbf{e}_m(t): \mathbf{e}_0^T\mathbf{M}\widetilde{\mathbf{Q}}_0\mathbf{e}_0 + \int_{t_0}^{t_1} \mathbf{e}_m^T(t)\mathbf{M}\widetilde{\mathbf{Q}}(t)\mathbf{e}_m(t)dt \leq 1\right\} \tag{4.45}$$

$$\mathbb{L}_o^2 = \left\{ \mathbf{e}_o(t) : \int_{t_0}^{t_1} \overline{\mathbf{e}_o^T(t)\mathbf{M}\widetilde{\mathbf{R}}(t)\mathbf{e}_o(t)} dt \leq 1 \right\} \tag{4.46}$$

where matrices $\widetilde{\mathbf{Q}}_0$ and $\widetilde{\mathbf{R}}(t)$ are defined as

$$\widetilde{\mathbf{Q}}_0(t) = \mathrm{diag}\{q_0(z_i)\}_{i=1}^N, \qquad \widetilde{\mathbf{R}}(t) = \mathrm{diag}\{r(t, z_i)\}_{i=1}^N \tag{4.47}$$

The structure of the feedback equation (4.32) is similar to the structure of the background equation (4.1) except that it is deterministic and does not contain any stochastic terms. The notation of the stiffness matrices $\mathbf{S}$, $\tilde{\mathbf{S}}$ and observation matrices $\mathbf{H}$, $\widetilde{\mathbf{H}}$ that are used here are equivalent to the notation derived in Section 3.2.3, so that only the discretisation of the variational equivalent of the source term $s = PH^* r^{-1}(y - Hu)$ is presented in detail.

Since $e_o = y - Hu$, using the definition of the observation operator (4.3) and definition of the Riccati operator (4.26), the source term is written as

$$s(t,x) = PH^* r^{-1} e_o = \int_\Omega k(t,x,z) \int_\Omega h(z-\zeta)\, r(\zeta) e_o(t,\zeta) dz \tag{4.48}$$

Applying the FEM expansions (4.36)-(4.37) in the (4.48), one obtains the following approximation

$$\begin{aligned} s(t,x) &\approx \boldsymbol{\phi}^T(x)\mathbf{s}(t) \\ &= \int_\Omega \boldsymbol{\phi}^T(x)\mathbf{K}(t)\boldsymbol{\phi}(z) \int_\Omega \boldsymbol{\phi}^T(z)\widetilde{\mathbf{H}}^T\boldsymbol{\phi}(\zeta)\boldsymbol{\phi}^T(\zeta)\widetilde{\mathbf{R}}(t)\mathbf{e}_o(t) d\zeta\, dz \end{aligned} \tag{4.49}$$

Reordering the integration and using the definition of the mass matrix results in

$$\begin{aligned} &\boldsymbol{\phi}^T(x)\mathbf{s}(t) \\ &= \boldsymbol{\phi}^T(x)\mathbf{K}(t) \int_\Omega \boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z) dz\, \widetilde{\mathbf{H}}^T \int_\Omega \boldsymbol{\phi}(\zeta)\boldsymbol{\phi}^T(\zeta) d\zeta\, \widetilde{\mathbf{R}}(t)\mathbf{e}_o(t) \\ &= \boldsymbol{\phi}^T(x)\mathbf{K}(t)\mathbf{M}\widetilde{\mathbf{H}}\mathbf{M}\widetilde{\mathbf{R}}(t)\mathbf{e}_o(t) \end{aligned} \tag{4.50}$$

The discretised variational equivalent of the source term is computed if the discrete observation equation (see Section 3.2.3) is substituted into (4.50)

$$\left(\boldsymbol{\phi}(x), \boldsymbol{\phi}^T(x)\mathbf{s}(t)\right) = \mathbf{M}\mathbf{K}(t)(\widetilde{\mathbf{H}}\mathbf{M})^T\mathbf{M}\widetilde{\mathbf{R}}(t)(\mathbf{y}(t) - \widetilde{\mathbf{H}}\mathbf{M}\mathbf{u}(t)) \tag{4.51}$$

In order to construct a FEM discretisation of the operator Riccati equation (4.31), it should be reformulated in a variational way. This is done by multiplying equation (4.31) on the left and

on the right by the test functions $v_x(x), v_z(z) \in L^2(\Omega)$. Thus, the weak solution of the Riccati equation is computed as a solution of the variational equivalent

$$\begin{aligned}\left(\left(v_x, \frac{\partial k}{\partial t}\right), v_z\right) &= \left((v_x, L_x k), v_z\right) + \left((v_x, L_z k), v_z\right) \\ &+ \left(\left(v_x, q^{-1}\delta(z-x)\right), v_z\right) \\ &- \left(\left(v_x, P\left(H^* r\delta(z-x)(Hk)\right)\right), v_z\right)\end{aligned} \tag{4.52}$$

where the operator L_x/L_z is an advection-diffusion operator defined in x/z direction

$$L_x k(t,x,z) = \sum_{i=1}^{2}\left[\epsilon \frac{\partial^2 k(t,x,z)}{\partial {z_i}^2} - \mu_i \frac{\partial k(t,x,z)}{\partial z_i}\right] \tag{4.53}$$

or represented in vector form

$$L_x k(t,x,z) = \epsilon \Delta_x k(t,x,z) - \nabla_x \cdot \left(\mu(t,x) k(t,x,z)\right) \tag{4.54}$$

The variational equation (4.52) is also supplemented with the initial conditions

$$\left((v_x, k(0,x,z)), v_z\right) = \left(\left(v_x, q_0^{-1}(x)\delta(z-x)\right), v_z\right) \tag{4.55}$$

Taking the test functions $v_x(x)$ and $v_z(z)$ as vector FEM basis functions $\boldsymbol{\phi}(x)$ and $\boldsymbol{\phi}(z)$ and inserting the approximation of the kernel (4.37) into the variational formulation (4.52) yields the matrix differential Riccati equation, presented for convenience as

$$D = A_x + A_z + B + C \tag{4.56}$$

where components of the equation are discretised as follows:

The term D is denoted as

$$\begin{aligned} D &= \int_\Omega \int_\Omega \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x) \frac{\partial \mathbf{K}(t)}{\partial t} \boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z) dx\, dz \\ &= \int_\Omega \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x) dx \frac{\partial \mathbf{K}(t)}{\partial t} \int_\Omega \boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z) dx\, dz \end{aligned} \tag{4.57}$$

Recalling notation of the mass matrix, one obtains

$$D = \mathbf{M}\, \dot{\mathbf{K}}(t) \mathbf{M}^T \tag{4.58}$$

The term A_x is given by

$$A_x = \int_\Omega \int_\Omega \boldsymbol{\phi}(x)\{\epsilon \Delta_x \boldsymbol{\phi}^T(x)\mathbf{K}(t)\boldsymbol{\phi}(z) - \nabla_x^T\big(\mu(t,x)\boldsymbol{\phi}^T(x)\big)\mathbf{K}(t)\boldsymbol{\phi}(z)\}\boldsymbol{\phi}^T(z)\, dxdz \tag{4.59}$$

and separating variables in (4.59) results in

$$A_x = \int_\Omega \epsilon \boldsymbol{\phi}(x)\Delta_x \boldsymbol{\phi}^T(x)dx\ \mathbf{K}(t)\int_\Omega \boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z)dz - \int_\Omega \boldsymbol{\phi}(x)\nabla_x^T\big(\mu(t,x)\boldsymbol{\phi}^T(x)\big)dx\,\mathbf{K}(t)\int_\Omega \boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z)dz \tag{4.60}$$

The diffusion and advection terms are then transformed according to the divergence theorem (Evans, 1998)

$$\int_\Omega \epsilon \boldsymbol{\phi}(x)\Delta_x \boldsymbol{\phi}^T(x)dx = -\int_{\boldsymbol{\Omega}} \epsilon \nabla^T \boldsymbol{\phi}(x)\nabla \boldsymbol{\phi}^T(x)dx + \int_{\partial\Omega} \epsilon \frac{\partial \boldsymbol{\phi}(x)}{n(x)}\boldsymbol{\phi}^T(x)dx \tag{4.61}$$

$$\int_\Omega \boldsymbol{\phi}(x)\nabla_x^T\big(\mu(t,x)\boldsymbol{\phi}^T(x)\big)dx = -\int_\Omega \nabla_x^T \boldsymbol{\phi}(x)\mu(t,x)\boldsymbol{\phi}^T(x)d\mathrm{x} + \int_{\partial\Omega} \mu^T(t,x)n(x)\boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)dx \tag{4.62}$$

Because the boundary conditions in the background model (4.1) are defined exactly, it is consistent to assume homogeneous boundary conditions for the kernel function $k(t,x,y)$. After (4.61)-(4.62) are inserted into (4.60), the boundary integrals vanish and A_x becomes

$$A_x = -\left(\int_{\boldsymbol{\Omega}} \epsilon \nabla^T \boldsymbol{\phi}(x)\nabla \boldsymbol{\phi}^T(x)dx - \int_\Omega \nabla_x^T \boldsymbol{\phi}(x)\mu(t,x)\boldsymbol{\phi}^T(x)d\mathrm{x}\right)\mathbf{K}(t)\int_\Omega \boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z)dz \tag{4.63}$$

Introducing the mass and stiffness matrix, A_x can be written

$$A_x = \tilde{\mathbf{S}}(t)\mathbf{K}(t)\mathbf{M}^T(t) \tag{4.64}$$

The term A_z is computed in a similar way as A_x and is given by

$$A_z = \mathbf{MK}(t)\tilde{\mathbf{S}}^T(t) \tag{4.65}$$

Discretisation of the free term B is done using the representation of the discrete Dirac delta function (4.40)

$$\begin{aligned} B &= \int_\Omega \int_\Omega \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\widetilde{\mathbf{Q}}^{-1}(t)\mathbf{M}^{-1}\boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z)\, dxdz \\ &= \int_\Omega \boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)dx\, \widetilde{\mathbf{Q}}^{-1}(t)\mathbf{M}^{-1} \int_\Omega \boldsymbol{\phi}(z)\boldsymbol{\phi}^T(z)dz \\ &= \mathbf{M}\widetilde{\mathbf{Q}}^{-1}(t) \end{aligned} \tag{4.66}$$

The quadratic term C is defined as

$$C = \int_\Omega \int_\Omega \boldsymbol{\phi}(x) f(x,z)\boldsymbol{\phi}(z)\, dxdz \tag{4.67}$$

where $f(x,z)$ is written in an expanded form

$$\begin{aligned} f(x,y) &= P\big(H^* r(Hk)\big) \\ &= \int_\Omega k(t,x,\alpha)(H^* rHk)(t,\alpha,z)dz \\ &= \int_\Omega k(t,x,\alpha)H^* r(t,\alpha) \int_\Omega h(\xi-\alpha)k(t,\xi,z)d\xi\, d\alpha \\ &= \int_\Omega k(t,x,\alpha) \int_\Omega h(\alpha-\zeta) \\ &\qquad \times r(t,\zeta) \int_\Omega h(\xi-\zeta)k(t,\xi,z)d\xi\, d\xi\, d\alpha \end{aligned} \tag{4.68}$$

To compute $f(x,y)$, the FEM expansion is applied to (4.68)

$$\begin{aligned} f(x,y) \approx \boldsymbol{\phi}(x)\mathbf{K}(t) \int_\Omega \boldsymbol{\phi}(\alpha)\boldsymbol{\phi}^T(\alpha)d\alpha\, \widetilde{\mathbf{H}} \int_\Omega \boldsymbol{\phi}(\zeta)\boldsymbol{\phi}^T(\zeta)d\zeta\, \widetilde{\mathbf{R}}(t)\widetilde{\mathbf{H}} \\ \times \int_\Omega \boldsymbol{\phi}(\xi)\boldsymbol{\phi}^T(\xi)d\xi\, \mathbf{K}(t)\boldsymbol{\phi}(z) \end{aligned} \tag{4.69}$$

and presented in a matrix form as

$$f(x,y) \approx \boldsymbol{\phi}(x)\mathbf{K}(t)(\widetilde{\mathbf{H}}\mathbf{M})^T\mathbf{M}\widetilde{\mathbf{R}}(t)\widetilde{\mathbf{H}}\mathbf{MK}(t)\boldsymbol{\phi}(z) \tag{4.70}$$

Inserting (4.70) into (4.67) allows the determination of the quadratic term C

$$C = \mathbf{M}\mathbf{K}(t)(\widetilde{\mathbf{H}}\mathbf{M})^T\mathbf{M}\widetilde{\mathbf{R}}(t)\widetilde{\mathbf{H}}\mathbf{M}\mathbf{K}(t)\mathbf{M} \tag{4.71}$$

Finally, the discretised Riccati matrix differential equation is obtained by substituting discretised terms (4.58), (4.64), (4.65), (4.66) and (4.71) into (4.56)

$$\begin{aligned}\mathbf{M}\dot{\mathbf{K}}(t)\mathbf{M} = \tilde{\mathbf{S}}(t)\mathbf{K}(t)\mathbf{M} + \mathbf{M}\mathbf{K}(t)\tilde{\mathbf{S}}^T(t) + \mathbf{M}\widetilde{\mathbf{Q}}^{-1}(t)\\ - \mathbf{M}\mathbf{K}(t)(\widetilde{\mathbf{H}}\mathbf{M})^T\mathbf{M}\widetilde{\mathbf{R}}(t)\widetilde{\mathbf{H}}\mathbf{M}\mathbf{K}(t)\mathbf{M}\end{aligned} \tag{4.72}$$

The last equation is further transformed by multiplying on the left and on the right by $(\mathbf{M}^T)^{-1}$ and $\mathbf{M}^{-1}$

$$\begin{aligned}\dot{\mathbf{K}} = \mathbf{M}^{-1}\tilde{\mathbf{S}}(t)\mathbf{K} + \mathbf{K}\left(\mathbf{M}^{-1}\tilde{\mathbf{S}}(t)\right)^T + \widetilde{\mathbf{Q}}^{-1}(t)\mathbf{M}^{-1}\\ - \mathbf{K}(t)(\widetilde{\mathbf{H}}\mathbf{M})^T\mathbf{M}\widetilde{\mathbf{R}}(t)\widetilde{\mathbf{H}}\mathbf{M}\mathbf{K}(t)\end{aligned} \tag{4.73}$$

Discretisation of the initial conditions (4.55) is done in a similar fashion to the discretisation of the free term (4.66) and produces

$$\mathbf{M}\dot{\mathbf{K}}(0)\mathbf{M} = \mathbf{M}\widetilde{\mathbf{Q}}_0^{-1} \tag{4.74}$$

$$\dot{\mathbf{K}}(0) = \widetilde{\mathbf{Q}}_0^{-1}\mathbf{M}^{-1} \tag{4.75}$$

As a result, the discrete kernel Riccati equation is given by equations (4.73) and (4.75).

4.1.4 Algorithm of time continuous minimax filter

Denoting $\mathbf{S}(t) = \mathbf{M}^{-1}\tilde{\mathbf{S}}(t)$, $\mathbf{H} = \widetilde{\mathbf{H}}\mathbf{M}$, $\mathbf{Q}_0 = \mathbf{M}\widetilde{\mathbf{Q}}_0$, $\mathbf{Q}(t) = \mathbf{M}\widetilde{\mathbf{Q}}(t)$ and $\mathbf{R}(t) = \mathbf{M}\widetilde{\mathbf{R}}(t)$, the algorithm of the space discrete minimax filter for the model, initial conditions and observation errors bounded by the ellipsoidal sets is written as follows:

$$\mathbb{L}_m^2 = \left\{\mathbf{e}_0, \mathbf{e}_m(t): \mathbf{e}_0^T\mathbf{Q}_0\mathbf{e}_0 + \int_{t_0}^{t_1}\mathbf{e}_m^T(t)\mathbf{Q}(t)\mathbf{e}_m(t)dt \leq 1\right\} \tag{4.76}$$

$$\mathbb{L}_o^2 = \left\{\mathbf{e}_o(t): \int_{t_0}^{t_1}\overline{\mathbf{e}_o^T(t)\mathbf{R}(t)\mathbf{e}_o(t)}dt \leq 1\right\} \tag{4.77}$$

The solution of the FEM discretised DA problem (4.1)-(4.2) is found as the solution of the feedback equation

$$\begin{cases} \dfrac{d\mathbf{u}(t)}{dt} = \mathbf{S}(t)\mathbf{u}(t) + \mathbf{f}(t) + \mathbf{K}(t)\mathbf{H}^T\mathbf{R}(t)(\mathbf{y}(t) - \mathbf{H}\mathbf{u}(t)) \\ \mathbf{u}(0) = \mathbf{u}^0 \end{cases} \tag{4.78}$$

where $\mathbf{K}(t)$ is the solution of the Riccati equation

$$\begin{cases} \dfrac{d\mathbf{K}(t)}{dt} = \mathbf{S}(t)\mathbf{K} + \mathbf{K}\mathbf{S}^T(t) + \mathbf{Q}^{-1}(t) - \mathbf{K}(t)\mathbf{H}^T\mathbf{R}(t)\mathbf{H}\mathbf{K}(t) \\ \mathbf{K}(t_0) = \mathbf{Q}_0^{-1} \end{cases} \tag{4.79}$$

The Riccati equation is a quadratic equation that balances model errors and observation errors based on the dynamics of the system. It has the following logical interpretation, the increase in the level of the uncertainty represented by the Riccati matrix is driven by three factors: (i) model dynamics, (ii) model error and (iii) observation error. The model dynamics does not change the amount of the uncertainty, but instead redistributes it over the domain. The model errors represented by the positive linear term in the Riccati equation increases the amount of the uncertainty, while the observation errors represented as the inverted multiplier of the quadratic term decreases the uncertainty. If the level of the observation discrepancies is low, then the quadratic term is high and neutralizes the model error fast, which leads to a fast convergence of the estimate. For the large observation error and correspondingly small quadratic term the model error may prevail, thereby increasing the matrix **K**. However, because the rate of growth of the quadratic term is much faster than the rate of growth of the linear term, matrix **K** will grow only to some certain level and then will get balanced.

The error of the minimax analysis is given by the inequality (4.33) written in a discrete case as

$$\sup_{(\mathbf{e}_0,\mathbf{e}_m)\in\mathbb{L}_m^2,\mathbf{e}_o\in\mathbb{L}_o^2} \left(\boldsymbol{l}\mathbf{M}\left(\mathbf{u}^a(t_1) - \mathbf{u}(t_1)\right)\right)^2 \leq \boldsymbol{l}\mathbf{M}\mathbf{K}(t_1)\mathbf{M}\boldsymbol{l} \tag{4.80}$$

for an vector $\boldsymbol{l}$ which is a FEM projection of an arbitrary element l.

One may notice the similar expressions of the time continuous minimax filter (4.78)-(4.79) and Kalman-Bucy filter. Both are found as a solution of the Riccati equation with a proper choice of the weighting matrices.

From the design point of view, the choice of the weighting matrices is done based on information about the system being modelled. Particularly, the error description meaning plays the same role as the covariances of the Kalman-Bucy filter. If the weights are the inverses of the covariances, then the Kalman-Bucy and minimax estimates are identical.

The heuristic argument of the filter equivalence is the following. The minimax estimate is the centre of the reachability set which has the property to provide the minimal distance to other elements in the set. The Kalman-Bucy filter is based on the minimization of the analysis variance while variance is a measure of a distance between elements of a statistical space. Since the reachability set of the minimax filter and the analysis set of the Kalman-Bucy filter are driven by the same process and they begin from the same centres, the propagation of those sets should coincide.

4.1.5 Solution of the Riccati equation

The solution of the feedback equation (3.30) is straightforward and obtained by the application of the mid-point time integration method. The solution of the Riccati equation (4.79) is more difficult. First of all, the equation is quadratic; this implies that a non-linear equation should be solved if implicit methods are used for time integration. Moreover, one of the properties of the Riccati equation is the existence of singularities that are dependent on the initial conditions. Direct integration by traditional time integration methods such as Runge-Kutta methods fail to compute those singularities (Schiff and Shnider, 1999).

A common way to overcome the singularity burden is by applying the Möbius Transformation. That transformation maps the original nonlinear problem into the double sized linear Hamiltonian problem that can be effectively solved by symplectic methods. For some matrix $\mathbf{A}(t)$ and invertible matrix $\mathbf{B}(t)$, substitute

$$\mathbf{K}(t) = \mathbf{A}(t)\mathbf{B}^{-1}(t) \tag{4.81}$$

into equation (4.79), then

$$\begin{aligned}\frac{d\mathbf{A}(t)}{dt}\mathbf{B}^{-1}(t) - \mathbf{A}(t)\mathbf{B}^{-1}(t)\frac{d\mathbf{B}(t)}{dt}\mathbf{B}^{-1}(t) \\ = \mathbf{S}(t)\mathbf{A}(t)\mathbf{B}^{-1}(t) + \mathbf{A}(t)\mathbf{B}^{-1}(t)\mathbf{S}^T(t) + \mathbf{Q}^{-1}(t) \\ - \mathbf{A}(t)\mathbf{B}^{-1}(t)\mathbf{H}^T\mathbf{R}(t)\mathbf{H}\mathbf{A}(t)\mathbf{B}^{-1}(t)\end{aligned} \tag{4.82}$$

Multiplying (4.82) by $\mathbf{B}(t)$ and regrouping the terms with $\mathbf{A}(t)\mathbf{B}^{-1}(t)$ multiplier produces

$$\begin{aligned}\frac{d\mathbf{A}(t)}{dt} - \mathbf{A}(t)\mathbf{B}^{-1}(t)\frac{d\mathbf{B}(t)}{dt} \\ = \mathbf{S}(t)\mathbf{A}(t) + \mathbf{Q}^{-1}(t)\mathbf{B}(t) \\ + \mathbf{A}(t)\mathbf{B}^{-1}(t)\big(\mathbf{S}^T(t)\mathbf{B}(t) - \mathbf{H}^T\mathbf{R}(t)\mathbf{H}\mathbf{A}(t)\big)\end{aligned} \tag{4.83}$$

It is easy to show that if $\mathbf{A}(t)$ and $\mathbf{B}(t)$ solve the equations

$$\begin{cases} \dfrac{d\mathbf{A}(t)}{dt} = \mathbf{S}(t)\mathbf{A}(t) + \mathbf{Q}^{-1}(t)\mathbf{B}(t) \\ \dfrac{d\mathbf{B}(t)}{dt} = \mathbf{S}^T(t)\mathbf{B}(t) - \mathbf{H}^T\mathbf{R}(t)\mathbf{H}\mathbf{A}(t) \end{cases} \tag{4.84}$$

then they verify (4.83).

If the system in (4.83) is written in the form

$$\frac{d}{dt}\begin{pmatrix}\mathbf{A}(t)\\ \mathbf{B}(t)\end{pmatrix} = \begin{pmatrix}\mathbf{S}(t) & \mathbf{Q}^{-1}(t)\\ \mathbf{H}^T\mathbf{R}(t)\mathbf{H} & \mathbf{S}^T(t)\end{pmatrix}\begin{pmatrix}\mathbf{A}(t)\\ \mathbf{B}(t)\end{pmatrix} \tag{4.85}$$

it is known as a Hamiltonian system, which has the following initial conditions

$$\begin{pmatrix}\mathbf{A}(t_0)\\ \mathbf{B}(t_0)\end{pmatrix} = \begin{pmatrix}\mathbf{Q}_0^{-1}\\ \mathbf{I}\end{pmatrix} \tag{4.86}$$

To solve the linear problem in (4.85)-(4.86), suppose the time interval $[t_0; t_1]$ is uniformly discretised into L subintervals of the length $\Delta t = (t_1 - t_0)/L$. Discretising using the mid-point rule, the Hamiltonian system is presented at time step $n+1$ as

$$\begin{pmatrix}\mathbf{A}_{n+1}\\ \mathbf{B}_{n+1}\end{pmatrix} = \begin{bmatrix}\mathbf{I} - \dfrac{\Delta t}{2}\mathbf{S}_{n+0.5} & \dfrac{\Delta t}{2}\mathbf{Q}_{n+0.5}^{-1}\\ \dfrac{\Delta t}{2}\mathbf{F}_{n+0.5} & \mathbf{I} - \dfrac{\Delta t}{2}\mathbf{S}_{n+0.5}^T\end{bmatrix}^{-1}\begin{pmatrix}\mathbf{A}_n\\ \mathbf{B}_n\end{pmatrix} - \begin{pmatrix}\mathbf{A}_n\\ \mathbf{B}_n\end{pmatrix} \tag{4.87}$$

where $\mathbf{S}_{n+0.5} = \mathbf{S}(t_0 + (n+0.5)\Delta t)$, $\mathbf{Q}_{n+0.5} = \mathbf{Q}(t_0 + (n+0.5)\Delta t)$ and $\mathbf{F}_{n+0.5} = \mathbf{H}^T\mathbf{R}(t_0 + (n+0.5)\Delta t)\mathbf{H}$.

Once, the $\mathbf{A}_{n+1}$ and $\mathbf{B}_{n+1}$ are found, the Riccati matrix is constructed according to (4.81) as

$$\mathbf{K}_{n+1} = \mathbf{A}_{\mathrm{n}+1}\mathbf{B}_{n+1}^{-1} \tag{4.88}$$

The stability properties of the recurrent expression (4.87) are further improved if the re-initialisation mechanism is employed (Frank and Zhuk, 2014). That requires the reassignment of matrices $\mathbf{A}_{\mathrm{n}+1}$ and $\mathbf{B}_{\mathrm{n}+1}$ after the Riccati matrix computation step (4.88) is performed

$$\begin{cases}\mathbf{A}_{\mathrm{n}+1} \leftarrow \mathbf{A}_{\mathrm{n}+1}\mathbf{B}_{n+1}^{-1} = \mathbf{K}_{n+1}\\ \mathbf{B}_{n+1} \leftarrow \mathbf{B}_{n+1}\mathbf{B}_{n+1}^{-1} = \mathbf{I}\end{cases} \tag{4.89}$$

While this procedure does not change the Riccati matrix $\mathbf{K}_{n+1}$ mathematically, it has the effect of improving its numerical performance. Substituting (4.89) into (4.87), the final expression for the solution of the Hamiltonian system is obtained

$$\begin{pmatrix}\mathbf{A}_{n+1}\\ \mathbf{B}_{n+1}\end{pmatrix} = \begin{bmatrix}\mathbf{I}-\frac{\Delta t}{2}\mathbf{S}_{n+0.5} & \frac{\Delta t}{2}\mathbf{Q}_{n+0.5}^{-1}\\ \frac{\Delta t}{2}\mathbf{F}_{n+0.5} & \mathbf{I}-\frac{\Delta t}{2}\mathbf{S}_{n+0.5}^{T}\end{bmatrix}^{-1}\begin{pmatrix}\mathbf{K}_n\\ \mathbf{I}\end{pmatrix}-\begin{pmatrix}\mathbf{K}_n\\ \mathbf{I}\end{pmatrix} \tag{4.90}$$

4.2 Comparison of conventional filters

In this section the minimax filter is compared versus conventional filters such as Kalman filter and EnTKF with different ensemble size. The comparison is done from two perspectives: estimation quality and numerical complexity.

4.2.1 Estimation quality

The quality of estimates of the filters is illustrated here with a set of numerical examples with an idealised configuration that serves illustrative purposes. In this experiment, a two-dimensional transport problem is described by the advection-diffusion equations (4.1) on a rectangular domain of the size $[0,1] \times [0,4]$ metres. That domain is further discretised by 900 bilinear finite elements (60 elements over the *x*-axis and 15 elements over the *y*-axis) which produce the FEM grid comprised of 976 nodes (61 nodes over the *x*-axis and 16 nodes over the *y*-axis). The underlying flow field is defined by the constant vector-function $\mu = [0.2;\ 0]\ m/s$ that translates initial concentration of a contaminant spill over the x-axis and a constant diffusion coefficient $\epsilon = 10^{-5}\ m^2/s$. The time step is taken to be $0.1\ s$ and the length of the simulation is set to be 200 time steps allowing the concentration to completely transition from the right to the left of the domain.

The resulting FEM model is quite imprecise in that it diverges from the analytical solution which is available in this case. This has been constructed intentionally in order to illustrate that the filters can improve the quality of the numerical solutions by using observed data and without knowing the initial conditions.

In order to introduce the analytical solution and observations of that problem, define the following two-dimensional Gaussian function:

$$u_a(t,x,y) = \frac{1}{\sigma^2 2\pi}\ e^{-\frac{1}{2}\left(\frac{x-x_0-m_x}{\sigma}\right)^2} e^{-\frac{1}{2}\left(\frac{y-y_0-m_y}{\sigma}\right)^2} \tag{4.91}$$

where σ, m_x and m_y are the diffusion and advection of the initial concentration $u_a(0, x, y)$; x_0 and y_0 define its center. Let $u_a(t, x, y)$ be the function as in (4.91) with parameters

$$\sigma = 0.1 + 2t\epsilon = 0.1 + 2t10^{-5} \tag{4.92}$$

$$m_x = t\,\mu_x = 0.2t, \qquad m_y = t\,\mu_y = 0 \tag{4.93}$$

$$x_0 = 0.5, \qquad y_0 = 0.5 \tag{4.94}$$

It is not difficult to check that the function $u_a(t, x, y)$ satisfies the original advection-diffusion equation (4.1) with the idealised flow field μ. In what follows, this serves as a ground-truth, and, in particular, the observations are sampled by restricting $u_a(t, x, y)$ onto the nodes of the FEM grid at each time step. The observations are corrupted by the observation noise. To guarantee boundedness of the observation noise, it is assumed to be uniformly distributed within the interval $[-1; 1]$, so the largest value of its component is approximately 10% of the contaminant concentration largest value.

According to (4.13) and (4.16), the error ellipsoids are defined by the functions q_0, q and r that are chosen to be constant in time and space and by the ellipsoidal approximation error factor $\gamma_T = (20 + 1) \cdot 4 = 84$. Hence,

$$\widetilde{\mathbf{Q}} = \frac{2}{\gamma_T}\mathbf{I}, \qquad \widetilde{\mathbf{Q}}_0 = \frac{0.1}{\gamma_T}\mathbf{I}, \qquad \widetilde{\mathbf{R}} = \frac{3}{\gamma_T}\mathbf{I} \tag{4.95}$$

and correspondingly

$$\mathbf{Q}(t) = \mathbf{M}\widetilde{\mathbf{Q}}, \qquad \mathbf{Q}_0 = \mathbf{M}\widetilde{\mathbf{Q}}_0, \qquad \mathbf{R} = \mathbf{M}\widetilde{\mathbf{R}} \tag{4.96}$$

These choices reflect the moderate level of trust in the FEM model and the absence of the initial conditions. The weighting matrix $\mathbf{R}$ is taken as the reciprocal of the variance of the $[-1; 1]$-uniformly distributed random variable. It should be noted that the described model of the observations noise is a robust version of the conventional statistical noise description. In what follows, the described experiment configuration is referred to as the test experiment configuration.

The estimates generated by the minimax filters $u_{minimax}$ are compared against other solutions by applying the following error metrics:

- Spatial norm:

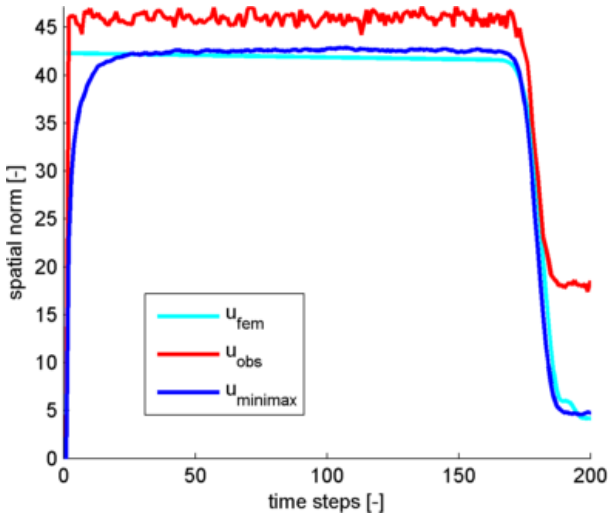

Figure 4.1 The spatial norm of the FEM solution, observations and minimax estimate plotted over time.

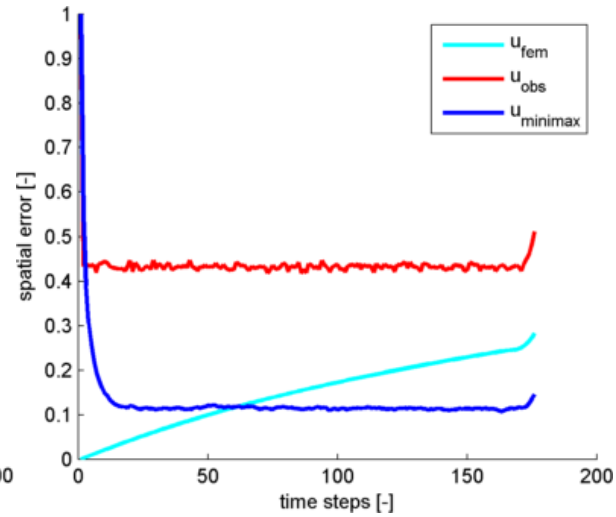

Figure 4.2 The spatial error of the FEM solution, observations and minimax estimate plotted over time.

$$n_s(u)(t) = \|u(t)\| \tag{4.97}$$

- Spatial error:

$$e_s(u)(t) = \frac{\|u(t) - u_a(t)\|}{\|u_a(t)\|} \tag{4.98}$$

- Estimation error:

$$e_e(u) = \frac{\sum_{k=0}^{L}\|u(t_0 + k\Delta t) - u_a(t_0 + k\Delta t)\|}{\sum_{k=0}^{L}\|u_a(t_0 + k\Delta t)\|} \tag{4.99}$$

where $\|\cdot\|$ denotes Euclidean norm in a discrete space, i.e., $\|\mathbf{u}\| = \sqrt{\mathbf{u}_1^2 + \cdots + \mathbf{u}_{N_{nd}}^2}$ and $u_a(t)$ is known analytical solution.

The effect of the assimilation of observations into the model produced by the minimax filter is demonstrated in Figure 4.1 and Figure 4.2. They plot the spatial norm and spatial error of the FEM solution $\mathbf{u}_{\text{fem}}$ with known initial conditions, which shows the quality of the model and as Figure 4.2 suggests the spatial error generated by the FEM increases in time. The spatial error of the observations plotted in Figure 4.2 does not increase in time, however it is relatively high (approximately 43%) due to the presence of the noise. At the same time, the minimax filter does not know the initial conditions, but it takes the advantage of the available error description to combine the model and observations. As a result, it constructs an estimate, spatial norm of which is close to the spatial norm of $\mathbf{u}_{\text{fem}}$ and according to Figure 4.2 the minimax

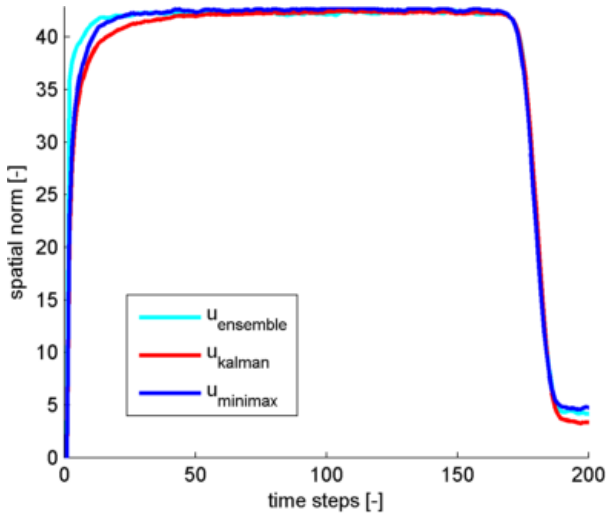

Figure 4.3 The spatial norm of the EnTKF with ensemble 1000, Kalman and minimax estimates plotted over time.

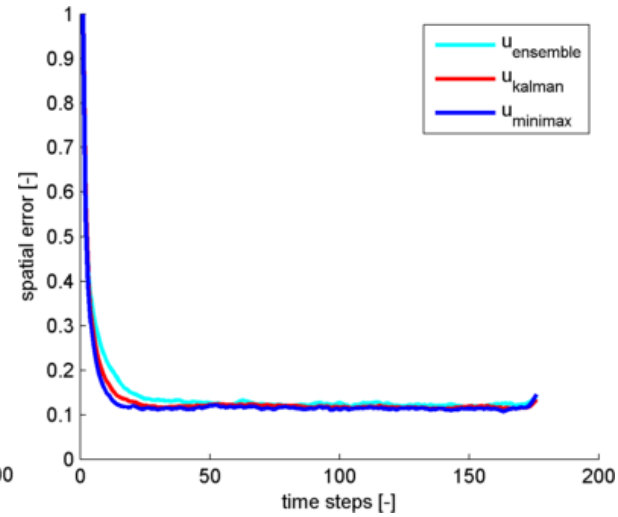

Figure 4.4 The spatial error of the EnTKF with ensemble 1000, Kalman and minimax estimates plotted over time.

spatial error is much smaller than the spatial error of the observations and is not increasing in time as the spatial error of the model. Improvements in terms of the error reduction provided by the minimax filter are further confirmed by the lower level of the estimation error: $e_e(\mathbf{u}_{\text{fem}}) = 16.3\%$, $e_e(\mathbf{u}_{\text{obs}}) = 43.6\%$, $e_e(\mathbf{u}_{\text{minimax}}) = 15\%$.

Figure 4.3 and Figure 4.4 depict comparison between estimates obtained from minimax filter $\mathbf{u}_{\text{minimax}}$, Kalman filter $\mathbf{u}_{\text{kalman}}$ and EnTKF $\mathbf{u}_{\text{ensemble}}$ with the ensemble size equal to 1000. As it can be seen from those images all three filters provide similar results and estimation error is slightly in favour of the minimax filter $e_e(\mathbf{u}_{\text{minimax}}) = 15\%$, the next one is the Kalman filter $e_e(\mathbf{u}_{\text{kalman}}) = 15.6\%$ and then EnTKF $e_e(\mathbf{u}_{\text{ensemble}}) = 16.5\%$. Since the minimax filter is equivalent to the Kalman-Bucy filter with respect to the parameters interpretation, no comparison between the two is presented.

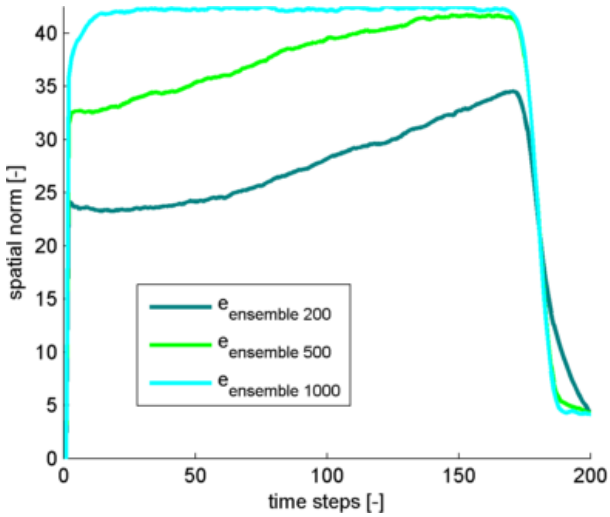

Figure 4.5 The spatial norm of the EnTKF estimates generated for different ensemble size: 200, 500, 1000 and plotted over time.

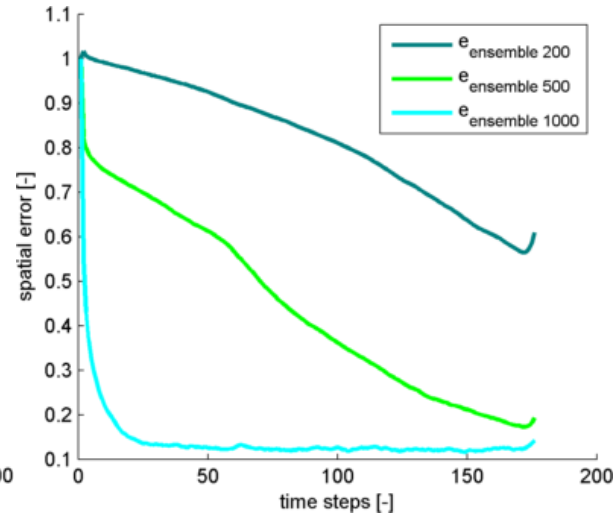

Figure 4.6 The spatial error of the EnTKF estimates generated for different ensemble size: 200, 500, 1000 and plotted over time.

The quality of estimate of EnTKF is also demonstrated for three different configurations of the ensemble size: 200, 500 and 1000. Figure 4.5 and Figure 4.6 plot spatial norm and spatial error of those filters. As expected, the increase of the ensemble size improves the quality of the estimate. It should be also noted, that the EnTKF estimates converge to the Kalman filter estimates not only with the increase of the ensemble size but also with time. The time convergence is explained by the fact that the velocity flow is taken to be constant for this particular experiment which leads to a time-constant model matrix that together with a constant error matrices provide a stationary solution for the covariance equation. Regardless of the ensemble size, the ensemble filters approach to that stationary solution but with the speed dependent on the ensemble size. In the case of a non-stationary velocity field, the EnTKF estimate may not converge to the Kalman estimate with time, which is a significant disadvantage.

4.2.2 Computational costs

In general, the filtering algorithms at each time instance can be split into two steps. During the first step, the model matrices such as stiffness matrix, mass matrix, source vector, etc. are computed using FEM. This is the common part of all filtering algorithms used in this experiment. Moreover, usually the first step is much faster comparing to the second step. The second step performs manipulations with the model matrices and is particular for each

algorithm. In the following analysis, only the cost of the second step is calculated to compare different filtering approaches.

The most time-consuming activities of the filtering algorithms are the matrix multiplication and matrix inversion. The basic algorithm of the matrix multiplication requires $\mathcal{O}({N_{nd}}^3)$ operations where N_{nd} is the number of the nodes of the FEM grid. Another common alternative is the more advanced Strassen's algorithm. That algorithm takes the advantage of the multiplication of matrices of the size 2×2 which requires only 7 multiplications in contrast to the usual 8. Applying this recursively gives an algorithm with the numerical cost $\mathcal{O}\left({N_{nd}}^{\log_2 7}\right) \approx \mathcal{O}\left({N_{nd}}^{2.807}\right)$. Even though Strassen's algorithm works faster, it is not necessary the best choice in practice. For instance, in this work for matrix operations the numerical library called OpenBLAS is employed. It contains a set of matrix routines that are implemented to support multiprocessor architecture of the modern computers. Because of the complexity of Strassen's algorithm, it uses basic matrix multiplication that is parallelised more efficiently. For this reason, the cost of the filter algorithms is computed for the basic matrix multiplication. Since the matrix inversion depends on the matrix multiplication itself, its cost is $\mathcal{O}({N_{nd}}^3)$ as well. It should be noted that matrix summation, or matrix by vector multiplication require $\mathcal{O}({N_{nd}}^2)$ operations and vector operations usually require $\mathcal{O}(N_{nd})$ operations which allows omitting them in the present analysis.

Let us denote by $\mathcal{C}_{MN} = \mathcal{O}(M^2 N)$ a unit of numerical complexity representing the amount of operations required for multiplication of matrices of the size $M \times N$. Then, inversion of dense matrix of the size $N \times N$ has complexity $\mathcal{C}_{NN}$. An application of FEM approximations to the underlying PDE generates sparse stiffness and mass matrices which allows resolving free run models, i.e., models with known initial conditions and without observations filtering, using $\mathcal{O}({N_{nd}}^2)$ operations. In contrast resolution of filters requires solving covariance matrix equations which leads to inversions and multiplications of dense matrices.

Complexity of filters can be expressed in terms of the introduced complexity units. To compute the estimate by the minimax filter, one needs to solve the feedback and Riccati equations (4.78)-(4.79). The feedback equation solved by the midpoint integration rule is not equivalent to the free run of the model due to a presence of a Riccati matrix in the innovation term. In fact, it requires $2\mathcal{C}_{N_{nd}N_{nd}}$ operations for the matrix multiplication and matrix inversion. The solution of the Riccati equation given by the expression (4.87) requires $8\mathcal{C}_{N_{nd}N_{nd}}$ operations for the doubled size matrix inversion of the midpoint method, $2\mathcal{C}_{N_{nd}N_{nd}}$ operations for the

multiplication by the matrix $\mathbf{K}_k$ and $2\mathcal{C}_{N_{nd}N_{nd}}$ for the matrix inversion and multiplication of the Möbius transformation (4.88). The total cost of the Riccati equation computation is $12\mathcal{C}_{N_{nd}N_{nd}}$ and the total cost of the minimax estimation is $14\mathcal{C}_{N_{nd}N_{nd}}$.

The total number of operations for the Kalman filter is $7\mathcal{C}_{N_{nd}N_{nd}}$, which includes 1 matrix multiplication during background estimate propagation, 2 units of complexity for 2 operations of the matrix multiplication during the background covariance update and 4 units of complexity for 2 operations of the matrix multiplication and 2 operations of the matrix inversion during the observation covariance update.

The complexity of the EnTKF filter depends on the size of the ensemble K. It consists of three components: the background propagation needs $\mathcal{C}_{N_{nd}N_{nd}}$ operations for the matrix multiplication and $\mathcal{C}_{N_{nd}K}$ for propagating each ensemble member; the covariance propagation that requires $\mathcal{C}_{N_{nd}K} + \mathcal{C}_{KN_{nd}}$ operations for the matrix multiplication and $2\mathcal{C}_{KK}$ operations for the matrix inversion and the eigen decomposition of the covariance matrix; and the ensemble propagation that is performed in $\mathcal{C}_{N_{nd}K} + 2\mathcal{C}_{KN_{nd}} + \mathcal{C}_{KK}$ operations involved in matrix multiplication. In total, the cost of EnTKF algorithm is $\mathcal{C}_{N_{nd}N_{nd}} + 3\mathcal{C}_{N_{nd}K} + 3\mathcal{C}_{KN_{nd}} + 3\mathcal{C}_{KK}$. It should be noted, that the EnTKF cost is the cheapest option, provided ensemble size $K \ll N_{nd}$, however, if the size of the ensemble K reaches or exceeds N_{nd} the EnTKF become slower than Kalman or minimax filters.

Table 4.1 Numerical and practical performance of various filters. Results of the CPU time measurements are averaged over 4 runs.

Minimax	Kalman	EnTKF		
		$K = 100$	$K = 200$	$K = 500$
$14\mathcal{C}_{N_{nd}N_{nd}}$	$7\mathcal{C}_{N_{nd}N_{nd}}$	$\mathcal{C}_{N_{nd}N_{nd}} + 3\mathcal{C}_{N_{nd}K} + 3\mathcal{C}_{KN_{nd}} + 3\mathcal{C}_{KK}$		
$244\ s$	$157\ s$	$71.1\ s$	$83\ s$	$136\ s$

Numerical computational costs of different filters built on top of the FEM model are summarised in Table 4.1. These results are further confirmed by the CPU execution time (total time across all cores) averaged over 4 runs of the corresponding methods for a test experiment configuration which are also presented in Table 4.1. The simulations were performed using Intel Core i7 processor containing 4 cores with 2.2 GHz clock speed each and 16 GB RAM. For a matrix data structures and manipulation Armadillo library (Sanderson and Curtin, 2016)

was used. For general matrix operations, it was linked to the OpenBLAS library which implements shared memory parallelisation (Xianyi et al., 2016) and for solving sparse linear systems it was linked to the SuperLU library (Li et al., 1999).

4.3 Discussion

The attractive benefit of the minimax filter is the flexibility of treatment of the uncertainties in the system. In contrast to the statistical methods, a minimax framework does not require an exact description but only needs a set of admissible realisations of those uncertainties. Moreover, it is shown that mathematically minimax is equivalent to the Kalman-Bucy filter.

From a computational point of view, the estimation quality of the minimax algorithm is similar to the estimate provided by the Kalman algorithm, while the computational cost is at least two times higher. It should be noted, however, that the cost of both methods is significantly higher than the free run model. Another problem is that the computational demands of those algorithm grow faster then the size of the underlying system, i.e., the minimax and Kalman filters are not scalable.

EnTKF allows for a reduction in the cost of estimation if the ensemble size is smaller than the size of the system, but this reduction comes with a decline in the quality of the estimate. Also, the ensemble size defines the size of the subspace in which uncertainties of the system are approximated. Thus, the choice of the ensemble size depends on the system size implying that the EnTKF algorithms are also not scalable.

To conclude, EnTKF is more flexible in terms of balancing computational time and estimation error comparing to the Kalman filter and minimax filter. Kalman filter performs faster than the minimax filter resulting in a similar estimation error. At the same time, due to its deterministic nature, the minimax filter provides a more robust error treatment, a desired property in engineering practice. The biggest disadvantage of the conventional filters is the lack of the scalability with respect to a number of finite elements.

5 Minimax filter in the form of DAE

The applicability of the traditional filters is problematic if the number of the degrees of freedom of the discretised model is large. According to the literature (Todling and Cohn, 1994), this problem may be tackled by introducing various sub-optimal approximations for the underlying uncertainty model that require less computational effort but reduce the quality of the estimate (see Section 2.3.1). More recently, with differing levels of success, there were several attempts to apply decomposition to the traditional filters (see Section 2.6).

In this chapter, a new approach that aims to localise computations on smaller subdomains of the original domain is presented. More specifically, the original problem for a large domain is decomposed into subproblems assigned to related subdomains fulfilled by transmission conditions. The set of subproblems is then discretised and recast as a DAE. To solve that DAE, the minimax filter is reformulated to incorporate algebraic constraints that are produced by the discretisation of the transmission conditions into the model. This approach is similar to discrete decomposition and particularly to distributed filters in that it acts on a discrete level. At the same, the resulting minimax DAE filter is equivalent to the traditional minimax filter while distributed filters are only approximations of traditional filters. Another distinction of the proposed method is a treatment of transmission conditions. Since transmission conditions are normally enforced through boundary conditions (Quarteroni and Valli, 1999) their design depends on physical properties of an underlying problem. While here, transmission conditions are presented as additional algebraic constraints of an optimisation problem which makes the approach more robust with respect to a problem nature. For instance, minimax DAE filter can be applied for both advection dominated and diffusion dominated transport problems.

The main results contained in this chapter are published in a paper by the author as Ragnoli et al. (2014). However, there are some differences between the contents of this chapter and that paper: i) sections 5.1.1 - 5.2.2 of this chapter elaborate on the development of the DAE minimax filter in more detail than the corresponding sections of that paper; ii) for consistency with other chapters of the thesis, Experiment 1 from that paper was replaced in this chapter by an experiment with the test experiment configuration; iii) Section 5.3.3 herein contains analysis of the computational performance of the algorithm. It should also be noted that although that paper was co-authored with others, the mathematical formulation of the algorithm, code development and numerical simulations were performed by the author of this thesis.

5.1.1 Global and Localised Problems

The definition of the approach starts with the background model for the concentration transport process, then the DA problem is formulated on a continuous level in terms of a minimax framework. For a given spatiotemporal domain $([t_0; t_1] \times \Omega)$, the process is described by the differential equation

$$\begin{cases} \dfrac{\partial u(t,x)}{\partial t} = Lu(t,x) + e_m(t,x) \\ u(0,x) = u^0(x) + e_0(x) \\ u(t,x) = 0, \qquad x \in \partial\Omega \end{cases} \tag{5.1}$$

where operator L is an advection-diffusion operator with a constant diffusion coefficient ϵ and velocity field μ

$$\begin{gathered} L: H^1([t_0; t_1] \times \Omega) \to L^2([t_0; t_1] \times \Omega) \\ u \mapsto Lu = \epsilon\Delta u - \nabla \cdot (\mu u) \end{gathered} \tag{5.2}$$

The observation equation is written as

$$y(t,x) = Hu(t,x) + e_o(t,x) \tag{5.3}$$

where H is linear operator

$$Hu(t,x) = \int_\Omega h(z-x)u(t,z)dz \tag{5.4}$$

The model error $e_m(t,x)$, the initial condition error $e_0(x)$ and observation error $e_o(t,x)$ are uncertain but bounded.

It is assumed that the deterministic problem from (4.1) is well-posed and has a unique solution u_G for any initial condition u^0. In the following the domain Ω, the problem (5.1)-(5.2) and the observation equation (5.3)-(5.4) are referred to as the global domain, global problem and global observation equation correspondingly.

To construct a numerical solution for the global problem, DD is applied. Assume a geometrical decomposition of the global domain Ω into N non-overlapping subdomains $\Omega_1, \dots, \Omega_N$. The common boundary between two adjacent subdomains, Ω_i and Ω_j, is denoted by $\Gamma_{i,j} = \partial\Omega_i \cap \partial\Omega_j$ and the set of common boundaries in the ith subdomain is designated by $\Gamma_i = \cup_j \Gamma_{ij}$, where

j goes over the adjacent subdomain of Ω_i, and the union of all possible $\Gamma_{i,j}$ is denoted by $\Gamma = \cup_{ij} \Gamma_{ij}$ and is referred to as the interface of the decomposition.

A function $f(x)$ defined on the global domain Ω can be decomposed using its representation by the characteristic functions of the decomposition χ_i, $i = 1, \dots, N$ as follows:

$$f(x) = \sum_{i=1}^{N} f_i(x)\chi_i(x) \tag{5.5}$$

where

$$\chi_i(x) = \begin{cases} 1 & x \in \Omega_i \\ 0.5 & x \in \Gamma_{ij} \\ 0 & otherwise \end{cases} \tag{5.6}$$

and $f_i(x)$ is the restrictions of $f(x)$ into local subdomains Ω_i such that $f_i(x) = f(x)$ for $x \in \Omega_i$, $i = 1, \dots, N$. In such a way, decomposition of the global functions $u(t,x)$, $u^0(x)$, $e_m(t,x)$ and $e_0(x)$ is defined by introducing a set of local functions $u_i(t,x)$, $u_i^0(x)$, $e_{m,i}(t,x)$ and $e_{0,i}(x)$.

The local solutions $u_i(t,x)$ for $i = 1, \dots, N$ are found from the local subproblems

$$\begin{cases} \dfrac{\partial u_i}{\partial t} = L_i u_i + e_{m,i}, & \text{in } \Omega_i \\ u_i(t,x) = 0, & \text{on } \partial\Omega \cap \partial\Omega_\mathrm{i} \\ u_i(0,x) = u_i^0(x) + e_{0,i}(x) \end{cases} \tag{5.7}$$

where the operator L_i denotes the restriction of the operator L on the subdomain Ω_i with the appropriate boundary condition provided $\partial\Omega_i \cap \partial\Omega \neq \emptyset$.

It can be seen from the formulation of (5.7) that the local problems and their corresponding solutions are independent of each other. This is only possible if the concentration spill does not cross the interface of the decomposition, which implies that the velocity flow is orthogonal to the interface. Otherwise, the subproblems are coupled and they should exchange information with each order. The exchange of information is organised through an additional equation known as the transmission conditions

$$B(u_1, \dots, u_N) = 0 \tag{5.8}$$

which synchronises computations of the local subproblems and forces continuity between local solutions (Toselli and Widlund, 2005). Equation (5.8) is valid for deterministic problems and

since the local problems (5.7) contain uncertainties, it is expected that equation (5.8) will hold approximately up to some error $e_a(t,x)$, which is known as the algebraic error that is defined on the interface Γ and is expected to be very small. Equation (5.8) can be reformulated as

$$B(u_1, \dots, u_N) = e_a \tag{5.9}$$

Since the only requirement that is imposed on B is the consistency of a generated solution, there are different possible choices of the operator B. One of the most common choices, which is used in this work, is comprised of the local trace operators

$$B_{i,j}: \Omega_i \mapsto \Gamma_{i,j}, \qquad i = 1, \dots, M, j = 1, \dots, N \tag{5.10}$$

that assigns to $u_i \in H^1(\Omega_\mathrm{i})$ its value over the boundary $\Gamma_{i,j}$, provided $\Gamma_{i,j} \neq \emptyset$ and M is the number of boundary segments. If $\Gamma_{i,j} = \emptyset$ for some j, the trace is taken to be empty, i.e., $B_{i,j} = 0$. The operator B then verifies if the traces from the adjacent subdomains are equal with respect to the algebraic error

$$B_{i,j}u_i(t,x) + B_{j,i}u_j(t,x) = e_a(t,x) \tag{5.11}$$

or

$$u_i(t,x) - u_j(t,x) = e_a(x) \text{ for } x \in \Gamma_{i,j} \tag{5.12}$$

for all segments of the interface Γ.

Taking into account (5.7) and (5.11), the global problem is reconstructed via a set of local problems on each subdomain

$$\frac{\partial}{\partial t}\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} L_1 & 0 & \cdots & 0 \\ 0 & L_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & L_N \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix} + \begin{pmatrix} e_{m,1} \\ e_{m,2} \\ \vdots \\ e_{m,N} \end{pmatrix} \tag{5.13}$$

with transmission conditions that enforces the continuity of the solution across the interface, which are also known as the Dirichlet transmission conditions

$$Bu = \begin{pmatrix} B_{1,1} & \cdots & B_{1,N} \\ \vdots & \ddots & \vdots \\ B_{M,1} & \cdots & B_{M,N} \end{pmatrix}\begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} = e_a \tag{5.14}$$

with initial conditions

$$\begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} u_1^0 \\ \vdots \\ u_1^0 \end{pmatrix} + \begin{pmatrix} e_{0,1} \\ \vdots \\ e_{0,N} \end{pmatrix} \tag{5.15}$$

and with boundary conditions

$$\begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \text{on } \partial\Omega \cap \partial\Omega_1 \\ \vdots \\ \text{on } \partial\Omega \cap \partial\Omega_\mathrm{N} \end{matrix} \tag{5.16}$$

Similarly, the observation equation is formulated on the local subdomains using (5.5) as

$$\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} H_1 u_1 \\ \vdots \\ H_N u_N \end{pmatrix} + \begin{pmatrix} e_{o,1} \\ \vdots \\ e_{o,N} \end{pmatrix} \tag{5.17}$$

where local observation operator H_i is the following restriction of the global observation operator H

$$H_i u_i(t,x) = \int_\Omega h_i(z-x) u_i(t,z) dz \tag{5.18}$$

Problem (5.13)-(5.18) is hereby referred to as the localised problem. The solution of that problem is called the localised solution, which is comprised of the local solutions u_i by using (5.5). The equivalence between the global and localised problems in particular the existence and uniqueness of the solution of the localised problem is described by the following theorem:

Theorem 1. If u_G is a solution of the global problem then it is a solution of the localised problem. Moreover, if u_G is unique, it is a unique solution for both the global and decomposed problem.

Proof: Let u_{Gi} denote the restriction of u_G to the subdomain Ω_i. It is obvious that $Bu_G = 0$. Thus, it remains to be proven that u_{Gi} is a solution of the i-th local problem. Since $u_{Gi}\,|_{\Gamma_i} = u_G\,|_{\Gamma_i}$, where Γ_i is the interface of subdomain Ω_i (the common boundaries with other subdomains), it follows that u_{Gi} solves the i-th local problem. The uniqueness is an obvious consequence.■

It should be noted that the Dirichlet transmission conditions (5.12) is the most basic way of data transmission. While mathematically it provides continuity of the solution across subdomains, it does not guarantee continuity of the solution's derivatives which may be a desired property in numerical simulations. To this end, it is possible to use operators that transmit more data such as Neumann or Robin type operators which, furthermore, can be used

in combination with overlapping decomposition. There are only two requirements regarding a the design of transmission operator: (i) it should provide equivalence between localised and global problems; (ii) it should be time independent. Otherwise, regardless of the choice of transmission operator, its discretisation produces a matrix similar to (5.14) that will be used as an optimisation constraint for a development of minimax DAE filter.

5.1.2 Finite Elements approximation

The discretisation of the local problem as well as of the global problem, is obtained here through the FEM approximation. Each local subdomain Ω_i is approximated by polygonal finite elements and the i-th local solution u_i is approximated by piecewise linear basis functions defined in the nodes of the corresponding element. Therefore, the i-th local solution in a discrete space is represented by the vector of coefficients of the FEM expansion:

$$\mathbf{u}_i(t) = \left[\mathbf{u}_{i,1}(t), \dots, \mathbf{u}_{i,N_{nd}^i}(t)\right]^T \tag{5.19}$$

where N_{nd}^i is the number of nodes in the i-th FEM grid.

To find the vector of coefficients $\mathbf{u}_i(t)$, the FEM representation of u_i is substituted into the weak formulation of the local problem (5.7) that also incorporates boundary conditions (5.16) (through the boundary integral, see Section 3.2.2) and generates a reduced model for the coefficients:

$$\begin{cases} \dfrac{d\mathbf{u}_i}{dt} = \mathbf{S}_i(t)\mathbf{u}_i(t) + \mathbf{e}_{m,i}(t) \\ \mathbf{u}_i(0) = \mathbf{u}_i^0 + \mathbf{e}_{0,i} \end{cases} \tag{5.20}$$

where $\mathbf{S}_i(t)$ is the stiffness matrix corresponding to the discretisation of the local advection-diffusion operator (see Section 3.2.3) and $\mathbf{u}_i^0$, $\mathbf{e}_m(t)$, $\mathbf{e}_o$ are vectors of projection coefficients of the initial conditions, model error and initial error.

The local observation equation is given by

$$\mathbf{y}_i(t) = \mathbf{H}_i\mathbf{u}_i(t) + \mathbf{e}_{o,i}(t) \tag{5.21}$$

where $\mathbf{H}_i$ is discretised by the FEM local observation operator (5.18) (see Section 3.2.3). $\mathbf{y}_i$ and $\mathbf{e}_{i,o}$ are vectors of projection coefficients of the observations and observation error functions respectively.

The discretised version of the trace operator (5.10) is given by the matrix $\mathbf{B}_{ij}$ of size $N_{nd}^{i} \times N_{nd}^{\Gamma_{ij}}$, where $N_{nd}^{\Gamma_{ij}}$ is the amount of FEM nodes on the common boundary Γ_{ij}. The elements of the matrix $\mathbf{B}_{ij}$ are taken from the set $\{-1,0,1\}$ so the product $\mathbf{B}_{ij}\mathbf{u}_i$ returns a vector of components of $\mathbf{u}_i$ that are defined in the nodes that belong to Γ_{ij} and the following equality holds

$$\mathbf{B}_{i,j}\mathbf{u}_i(t) + \mathbf{B}_{j,i}\mathbf{u}_j(t) = \mathbf{e}_a|_{\Gamma_{i,j}}(t) \tag{5.22}$$

In other words, matrices $\mathbf{B}_{ij}$ and $\mathbf{B}_{ji}$ are taken such that solutions on adjacent sub-domains coincide at the nodes of the common part of the FEM grid up to a given error:

$$\mathbf{u}_{i,k}(t) - \mathbf{u}_{j,m}(t) = \mathbf{e}_{a,n}(t) \tag{5.23}$$

Here, subscript k in the index of a node in subdomain Ω_i which coincides with m-th node in the domain Ω_j and subscript n is the index of the same node in the numbering of interface nodes.

Comprising the set of the stiffness matrices of the N local problems (5.20) discretised by FEM results in the localised stiffness matrix

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{S}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{S}_N \end{pmatrix} \tag{5.24}$$

Using the set of vectors $\mathbf{u}_i$, N solutions of the local problems (5.20) are comprised into the vector of the localised solution

$$\mathbf{u} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N]^T \tag{5.25}$$

From the set of vectors $\mathbf{u}_i^0$, N initial conditions of the local problems (5.20) are comprised into a vector of the localised initial condition

$$\mathbf{u^0} = [\mathbf{u}_1^0, \mathbf{u}_2^0, \dots, \mathbf{u}_N^0]^T \tag{5.26}$$

and the set of vectors $\mathbf{e}_{i,m}$ and $\mathbf{e}_{i,0}$ are condensed into a vector of the localised model error and the localised initial error

$$\mathbf{e}_m = \left[\mathbf{e}_{1,m}, \mathbf{e}_{2,m}, \dots, \mathbf{e}_{N,m}\right]^T \tag{5.27}$$

$$\mathbf{e}_0 = \left[\mathbf{e}_{1,0}, \mathbf{e}_{2,0}, \dots, \mathbf{e}_{N,0}\right]^T \tag{5.28}$$

Finally, the discrete localised background model is obtained from

$$\begin{cases} \dfrac{d\mathbf{u}(t)}{dt} = \mathbf{S}(t)\mathbf{u}(t) + \mathbf{e}_m(t) \\ \mathbf{u}(0) = \mathbf{u}^0 + \mathbf{e}_0 \end{cases} \tag{5.29}$$

with the transmission conditions given in the form of an algebraic constraint

$$\mathbf{Bu} = \begin{pmatrix} \mathbf{B}_{1,1} & \cdots & \mathbf{B}_{1,N} \\ \vdots & \ddots & \vdots \\ \mathbf{B}_{M,1} & \cdots & \mathbf{B}_{M,N} \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_N \end{pmatrix} = \mathbf{e}_a \tag{5.30}$$

In a similar fashion, the localised observation equation is written

$$\mathbf{y}(t) = \mathbf{Hu}(t) + \mathbf{e}_o(t) \tag{5.31}$$

where $\mathbf{H}$ is localised observations operator and $\mathbf{y}(t)$ and $\mathbf{e}_o(t)$ are localised observations and observation error vectors such that

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{H}_N \end{pmatrix} \tag{5.32}$$

$$\mathbf{y} = [\boldsymbol{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]^T \tag{5.33}$$

$$\mathbf{e}_o = \left[\mathbf{e}_{o,1}, \mathbf{e}_{o,2}, \dots, \mathbf{e}_{o,N}\right]^T \tag{5.34}$$

The problem (5.29)-(5.31) is referred to as the discrete localised problem. By Theorem 1, the unique solvability of the discretised global problem implies the unique solvability of the corresponding discrete localised problem represented by (5.29)-(5.31).

5.2 Minimax filter for localised problem

From (5.20) it is necessary that the local solution $\mathbf{u}_i$ depends on the local model error $\mathbf{e}_{m,i}$. At the same time, the algebraic error $\mathbf{e}_{\boldsymbol{a}}$ measures the difference between coefficients of $\mathbf{u}_i$ in the shared nodes on common boundaries between subdomains. This implies that the realisation of the model error influences the realisation of the algebraic error meaning that the errors $\mathbf{e}_m$ and $\mathbf{e}_{\boldsymbol{a}}$ are correlated.

Let's define

$$\mathbf{e} = \begin{pmatrix} \mathbf{e}_m \\ \mathbf{e}_a \end{pmatrix}, \mathbf{F} = \begin{pmatrix} \mathbf{I} \\ \mathbf{0} \end{pmatrix}, \mathbf{C} = \begin{pmatrix} \mathbf{S} \\ \mathbf{B} \end{pmatrix} \tag{5.35}$$

where

$$\begin{gathered} \mathbf{e}_m \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd}}), \qquad \mathbf{e}_a \in \mathbb{C}\left([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd}^{\Gamma}}\right) \\ \mathbf{e} \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd}+N_{nd}^{\Gamma}}) \end{gathered} \tag{5.36}$$

$$\begin{gathered} \mathbf{S} \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd}\times N_{nd}}), \qquad \mathbf{B} \in \mathbb{C}\left([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd}^{\Gamma}\times N_{nd}}\right) \\ \mathbf{F}, \mathbf{C} \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{(N_{nd}+N_{nd}^{\Gamma})\times(N_{nd}+N_{nd}^{\Gamma})}) \end{gathered} \tag{5.37}$$

and N_{nd} denotes the sum of the FEM nodes from each subdomain and N_{nd}^{Γ} is the amount of FEM nodes on the interface of decomposition.

The background model (5.29)-(5.30) is then rewritten in the form of a DAE

$$\begin{cases} \dfrac{d}{dt}(\mathbf{Fu}) = \mathbf{Cu} + \mathbf{e} \\ \mathbf{Fu}(t_0) = \mathbf{u}_0 + \mathbf{e}_0 \end{cases} \tag{5.38}$$

and the error of the initial conditions, $\mathbf{e}_0$, and the general error, **e,** are assumed to be inside the following ellipsoid

$$\mathbb{L}_m = \left\{ \mathbf{e}: \mathbf{e}_0^T \mathbf{Q}_0^T \mathbf{e}_0 + \int_{t_0}^{t_1} \mathbf{e}^T(t)\mathbf{Q}(t)\mathbf{e}(t)dt \leq 1 \right\} \tag{5.39}$$

with weighting matrices $\mathbf{Q}_0$ and $\mathbf{Q}$ symmetric and positive-definite matrices

$$\mathbf{Q}_0 = \mathbf{Q}_0^T > 0; \ \mathbf{Q}(t) = \mathbf{Q}^T(t) > 0 \tag{5.40}$$

such that

$$\mathbf{Q}_0^{-1}, \mathbf{Q}_0 \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd}\times N_{nd}}) \tag{5.41}$$

$$\mathbf{Q}^{-1}, \mathbf{Q} \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{(N_{nd}+N_{nd}^{\Gamma})\times(N_{nd}+N_{nd}^{\Gamma})}) \tag{5.42}$$

For the observation equation

$$\mathbf{y}(t) = \mathbf{Hu}(t) + \mathbf{e}_o(t) \tag{5.43}$$

the localised observation error $\mathbf{e}_o$ is the realization of a random process such that

$$\mathbb{L}_o = \left\{ \overline{\mathbf{e}_o(t)} = 0, \int_{t_0}^{t_1} \mathbf{e}_o(t)^T \mathbf{R}(t) \mathbf{e}_o(t) dt \leq 1 \right\} \tag{5.44}$$

where weighting matrix $\mathbf{R}(t)$ is taken to be

$$\mathbf{R}(t) = \mathbf{R}^T(t) > 0, \qquad \mathbf{R}^{-1}, \mathbf{R} \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd} \times N_{nd}}) \tag{5.45}$$

Solving the discrete localised problem (5.29)-(5.31) is equivalent to finding an estimate of the solution to the system of DAE (5.38) and (5.43) with the system uncertainties described by the quadratic ellipsoidal sets (5.39) and (5.44). Since (5.38) also contains algebraic equations (the discretised transmission conditions), the minimax filter cannot be applied directly. Instead, the minimax framework is adopted for the estimation of the system (5.38) and (5.43).

5.2.1 Derivation of the filter

From the Kalman duality principle, introduced in (Zhuk, 2013), for linear functionals ℓ and v defined as

$$\ell(\mathbf{u}) = l^T \mathbf{F} \mathbf{u}(t_1) = l_s^T \mathbf{u}(t_1), \qquad l \in \mathbb{R}^{N_{nd} + N_{nd}^{\Gamma}}, l_s \in \mathbb{R}^{N_{nd}} \tag{5.46}$$

$$v(\mathbf{y}) = \int_{t_0}^{t_1} \mathbf{v}(t) \mathbf{y}(t) dt\,, \;\; \mathbf{v}(t) \in \mathbb{R}^{N_{nd}} \tag{5.47}$$

the estimate $\mathbf{u}^a(t)$ of $\mathbf{u}(t)$ with the minimal mean-squared estimation error

$$\sup_{(\mathbf{e}_0, \mathbf{e}) \in \mathbb{L}_m, \mathbf{e}_o \in \mathbb{L}_o} \ell(\mathbf{u}^a(t_1) - \mathbf{u}(t_1)) \leq \sup_{(\mathbf{e}_0, \mathbf{e}) \in \mathbb{L}_m, \mathbf{e}_o \in \mathbb{L}_o} \ell(v(\mathbf{y}) - \mathbf{u}(t_1)) \tag{5.48}$$

admits the following representation

$$\ell(\mathbf{u}^a) = \int_{\Omega_T} \mathbf{R} \mathbf{H}(t) \mathbf{p}(t) \mathbf{y}(t) dt \tag{5.49}$$

provided $\mathbf{p}$ and $\mathbf{z}$ such that

$$\mathbf{p} \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd}}), \qquad \mathbf{z} \in \mathbb{C}\left([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd} + N_{nd}^{\Gamma}}\right) \tag{5.50}$$

solve the following dual and direct equations

$$\begin{cases} \frac{d}{dt}(\mathbf{F}^T\mathbf{z}) = -\mathbf{C}^T\mathbf{z} + \mathbf{H}^T\mathbf{R}\mathbf{H}\mathbf{p} \\ \mathbf{F}^T\mathbf{z}(t_1) = \mathbf{F}^T l \end{cases} \tag{5.51}$$

$$\begin{cases} \frac{d}{dt}(\mathbf{F}\mathbf{p}) = \mathbf{C}\mathbf{p} + \mathbf{Q}^{-1}\mathbf{z} \\ \mathbf{F}\mathbf{p}(t_0) = \mathbf{Q}_0^{-1}\mathbf{z}(t_0) \end{cases} \tag{5.52}$$

The vector $\mathbf{z}$ and the matrix $\mathbf{Q}^{-1}$ are split according to the blocks $\mathbf{S}$ and $\mathbf{B}$ of matrix $\mathbf{C}$

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}_s \\ \mathbf{z}_b \end{pmatrix}, \mathbf{Q}^{-1} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{Q}_2^T & \mathbf{Q}_4 \end{pmatrix} \tag{5.53}$$

such that

$$\mathbf{Q}_1 \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd} \times N_{nd}}), \qquad \mathbf{Q}_2 \in \mathbb{C}\left([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd} \times N_{nd}^{\Gamma}}\right) \tag{5.54}$$

$$\mathbf{Q}_4 \in \mathbb{C}([\mathrm{t}_0, \mathrm{t}_1], \mathbb{R}^{N_{nd}^{\Gamma} \times N_{nd}^{\Gamma}})$$

Let introduce the following definitions

$$\mathbf{M} = \mathbf{S} - \mathbf{Q}_2\mathbf{Q}_4^{-1}\mathbf{B} \tag{5.55}$$

$$\mathbf{N} = \mathbf{B}^T\mathbf{Q}_4^{-1}\mathbf{B} + \mathbf{H}^T\mathbf{R}\mathbf{H} \tag{5.56}$$

$$\mathbf{L} = \mathbf{Q}_1 - \mathbf{Q}_2\mathbf{Q}_4^{-1}\mathbf{Q}_2^T \tag{5.57}$$

where $\mathbf{L}$ is the Schur complement of the matrix $\mathbf{Q}_4$ (Toselli and Widlund, 2005) and $\mathbf{Q}_4^{-1} = (\mathbf{Q}_4^{-1})^T$ is invertible since $\mathbf{Q}$ is positive definite matrix. Using, (5.55)-(5.57) the following Lemma can be proved.

Lemma 1: The problem in (5.51)-(5.52) is equivalent to the following ODE boundary-value problem:

$$\begin{cases} \dot{\mathbf{z}}_s = -\mathbf{M}^T\mathbf{z}_s + \mathbf{N}\mathbf{p}, & \mathbf{z}_s(t_1) = l_s \\ \dot{\mathbf{p}} = \mathbf{M}\mathbf{p} + \mathbf{L}\mathbf{z}_s, & \mathbf{p}(t_0) = \mathbf{Q}_0^{-1}\mathbf{z}_s(t_0) \end{cases} \tag{5.58}$$

Proof: Considering the splitting of vector $\mathbf{z}$ and matrix $\mathbf{Q}^{-1}$ defined by (5.53), the system of equations (5.51) and (5.52) are then rewritten as follows

$$\begin{cases} \dot{\mathbf{z}}_s = -\mathbf{S}^T\mathbf{z}_s - \mathbf{B}^T\mathbf{z}_b + \mathbf{H}^T\mathbf{R}\mathbf{H}\mathbf{p} \\ 0 = \mathbf{B}\mathbf{p} + \mathbf{Q}_2^T\mathbf{z}_s + \mathbf{Q}_4\mathbf{z}_b \\ \mathbf{z}_s(t_1) = l_s \end{cases} \tag{5.59}$$

$$\begin{cases}\dot{\mathbf{p}} = \mathbf{Sp} + \mathbf{Q}_1\mathbf{z}_s + \mathbf{Q}_2\mathbf{z}_b \\ \mathbf{p}(t_0) = \mathbf{Q}_0^{-1}\mathbf{z}_s(t_0)\end{cases} \tag{5.60}$$

From the algebraic part of system (5.59), $\mathbf{z}_b$ is derived as

$$\mathbf{z}_b = -\mathbf{Q}_4^{-1}(\mathbf{Bp} + \mathbf{Q}_2^T\mathbf{z}_s) \tag{5.61}$$

Inserting the expression of $\mathbf{z}_b$ (5.61) into (5.59) and (5.60) and regrouping the terms, the flowing equations for $\mathbf{z}_s$ and p are obtained

$$\begin{cases}\dot{\mathbf{z}_s} = (\mathbf{B}^T\mathbf{Q}_4^{-1}\mathbf{Q}_2^T - \mathbf{S}^T)\mathbf{z}_s - (\mathbf{B}^T\mathbf{Q}_4^{-1}\mathbf{B} + \mathbf{H}^T\mathbf{RH})\,\mathbf{p} \\ \dot{\mathbf{p}} = (\mathbf{Q}_1 - \mathbf{Q}_2\mathbf{Q}_4^{-1}\mathbf{Q}_2^T)\mathbf{z}_s + (\mathbf{S} - \mathbf{Q}_2\mathbf{Q}_4^{-1}\mathbf{B})\mathbf{p}\end{cases} \tag{5.62}$$

Taking into account the notation (5.55)-(5.57) in the formulation of (5.62) completes the proof.

■

The solution of the system (5.58) is obtained using the standard argument from LQ control (Reid, 1972). Let matrix **K** defines the relationship between $\mathbf{z}_s$ and **p**

$$\mathbf{p} = \mathbf{Kz}_s \tag{5.63}$$

Inserting (5.63) into the second equation of (5.58) and differentiating it results in

$$\dot{\mathbf{K}}\mathbf{z}_s = \mathbf{MKz}_s - \mathbf{K}\dot{\mathbf{z}_s} + \mathbf{Lz}_s \tag{5.64}$$

and inserting the first equation of (5.58) into (5.63) leads to the following equation

$$\dot{\mathbf{K}}\mathbf{z}_s = \mathbf{MKz}_s + \mathbf{KM}^T\mathbf{z}_s - \mathbf{KNKz}_s + \mathbf{Lz}_s \tag{5.65}$$

with the initial condition

$$\mathbf{K}(t_0)\mathbf{z}_s(t_0) = \mathbf{Q}_0^{-1}\mathbf{z}_s(t_0) \tag{5.66}$$

Matrix $\mathbf{K}(t)$, therefore, needs to satisfy the following matrix differential Riccati equation

$$\begin{cases}\dot{\mathbf{K}} = \mathbf{MK} + \mathbf{KM}^T - \mathbf{KNK} + \mathbf{L} \\ \mathbf{K}(t_0) = \mathbf{Q}_0^{-1}\end{cases} \tag{5.67}$$

Now assuming that $\mathbf{K}(t)$ solves (5.67) and using (5.63) it is easy to see that $\mathbf{z}_s$ solves the following equation

$$\begin{cases}\dot{\mathbf{z}_s} = -\mathbf{M}^T\mathbf{z}_s + \mathbf{NKz}_s \\ \mathbf{z}_s(t_1) = l_s\end{cases} \tag{5.68}$$

and so

$$\ell(\mathbf{u}^a) = \int_{t_0}^{t_1} \mathbf{RH}(t)\mathbf{Kz}_s\mathbf{y}(t)dt \tag{5.69}$$

Since (5.69) holds true for any l, it follows that $\mathbf{u}^a$ is the minimax estimate of the state vector of the DAE (5.38) and can be found from the feedback equation defined as follows (Zhuk, 2013)

$$\begin{cases} \dot{\mathbf{u}} = (\mathbf{M} - \mathbf{KN})\mathbf{u} + \mathbf{KH}^T\mathbf{Ry}(t) \\ \mathbf{u}(t_0) = \mathbf{u}^0 \end{cases} \tag{5.70}$$

The set of equations (5.67)-(5.70) defines a minimax filter for the DAE (5.38)-(5.43).

5.2.2 The algorithms with activation and deactivation

The proposed minimax filter (5.38)-(5.43) can be used in two different settings:

A. For a known initial condition with no observations. In this case the observation matrix $\mathbf{H}$ is set to be zero, meaning that the second term in the (5.56) of the Riccati equation (5.67) and the term $\mathbf{K}^T\mathbf{H}^T\mathbf{Ry}(t)$ in the feedback equation (5.70) are both zero. The inverse of $\mathbf{Q}_0$ in (5.67) controls the norm of the error related with the initial conditions. If the initial conditions are known exactly, matrix $\mathbf{Q}_0$ approaches to infinity, so its inverse is the zero matrix. In this setting the algorithm can be considered as the DD solver in which exchange of information between the adjacent subdomains is carried out by the minimax filter.
B. For an unknown initial condition, but with available noisy observations. The observation matrix $\mathbf{H}$ becomes non trivial and the level of noise is described by the matrix $\mathbf{R}$. Hence, the feedback equation (5.70) has a nonzero term $\mathbf{K}^T\mathbf{H}^T\mathbf{Ry}(t)$ with zero initial conditions.

Furthermore, application of DD to the transport equation allows for more robust organisation of computations then in the traditional filter. Assuming the location of concentration spill is detected, the actual calculation takes place only over those subdomains where the concentration is above zero. In particular, the subdomain gets activated/deactivate once the concentration reaches/leaves the boundaries of that subdomain. Depending on a geometrical decomposition, this enables the exclusion of non-active subdomains from the computations and may decrease the size of both the Riccati equation (5.67) and the feedback equation (5.70).

```
Require:
  T                    // number of time steps
  globalProblem        // description of the global physical problem
  subProblems = DecomposeProblem(globalProblem)
  for t = 1 to T do
    actProblems = GetActiveSubproblem(subProblems)
    for subproblem in actProblems do
      DiscretiseSubproblemByFem(subproblem, t)        // as in (5.20)-(5.21)
      ConstructDAE(subproblem, daeProblem, t)         // produce (5.29)-(5.31)
      If HasObservations(subproblem) then
        InitObservations(subproblem, t)
      endif
    endfor
      ComputeFilterMatrices(daeProblem, t)
      SolveRiccatiEquation(daeProblem, t)             // solve (5.67)
      SolveFeedbackEquation(daeProblem, t)            // solve (5.70)
  Endfor
```

Figure 5.1 Algorithm of the DAE minimax filter.

The algorithms of the minimax filter for DAE problem with activation/deactivation are described in Figure 5.1.

5.3 Numerical experiments

The feasibility of solving the localised problem (5.13)-(5.18) via (5.67)-(5.70) is addressed here by performing a set of numerical experiments that benchmark solutions of the localised problem versus solutions of the global problem. The set of numerical experiments is conducted using FEM approximations for both the global problem and the localised problem.

To investigate properties of the filter, two experiments have been performed. Experiment 1 represents the test experiment configuration that employs a trivial underlying flow. Simulations of Experiment 1 are carried out in both settings A and B. Experiment 2 is carried out for Setting B and employs a complex underlying flow that is generated using the Environmental Fluid

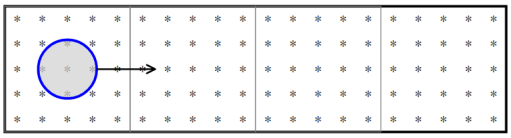

Figure 5.2. Configuration of Experiment 1.

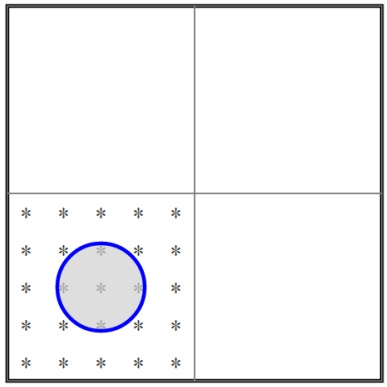

Figure 5.3. Configuration of Experiment 2.

Dynamics Code (EFDC), a complex hydrodynamic model that is used in marine and coastal applications (Hamrick, 1992).

In both experiments, the matrix $\mathbf{Q}$ that defines the control ellipsoid (5.39) is constructed to minimise the algebraic error $\mathbf{e}_a$ and to constrain the model error $\mathbf{e}_m$. In order to do so, $\mathbf{Q}$ is assumed to be block-decomposed due to the differential and algebraic parts of equation (5.38).

Hence, it takes the form

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{Q}_4 \end{pmatrix} \tag{5.71}$$

where

$$\mathbf{Q}_1 = \frac{c_m}{\gamma_T}\mathbf{M}, \qquad \mathbf{Q}_4 = c_a\mathbf{I} \tag{5.72}$$

and $\gamma_T = 84$, $c_m = 2$. Algebraic constant $c_a > 0$ is taken to be 10, 100 or 1000 reflecting different levels of confidence in the continuity equation (5.30).

It is assumed that the initial conditions are non zero only in the first subdomain and are generated by a standard linear spline (4.91). Moreover, observations are present only in the first subdomain and only the first subdomain is activated at the beginning of both experiments.

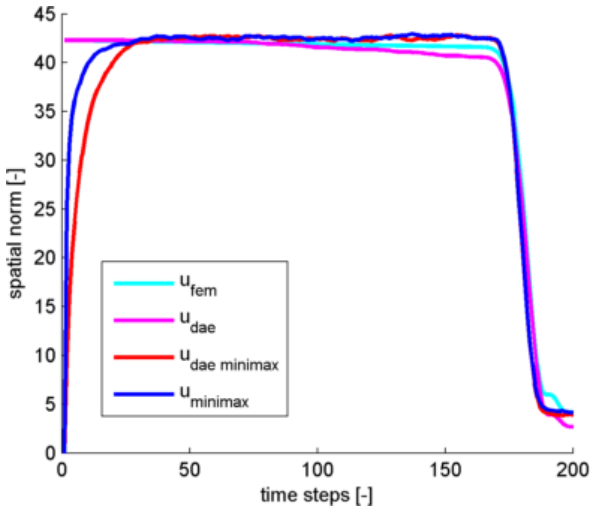

Figure 5.4 The spatial norm of the FEM solution, DAE estimate, DAE minimax estimates and minimax estimate plotted over time for the problem from Experiment 1.

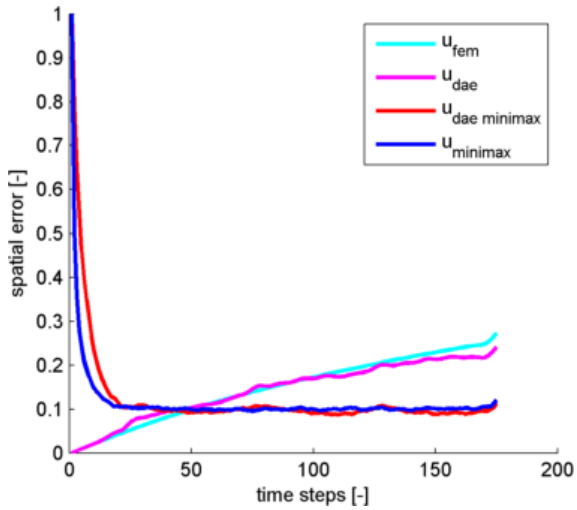

Figure 5.5 The spatial error of the FEM solution, DAE estimate, DAE minimax estimate and minimax estimates plotted over time for the problem from Experiment 1.

5.3.1 Experiment 1

In this experiment a set of numerical simulations is performed adopting the test experiment configuration (see Section 4.2.1) for a DA problem (4.1). The global domain is further decomposed into four equal size subdomains $[0,1] \times [0,1]$ metres along horizontal axis as in Figure 5.2. The underlying problem is defined by the constant velocity flow $\mu = [0.2;\ 0]^T\ m/s$ and the constant diffusion coefficient $\epsilon = 10^{-5}\ m^2/s$. The timestep is taken to be 0.1 s and the length of the simulation is set to be 200 time steps. It is also assumed that the analytical solution of this problem is given by the Gaussian function $u_a(t, x, y)$ defined in (4.91) with the parameters (4.92)-(4.94).

Setting A is used to demonstrate correctness of the proposed DD approach with Dirichlet transmission conditions. For this setting, the initial conditions are known and given by $u_a(0, x, y)$. The solutions of the DAE filter (5.67), (5.70) $\mathbf{u}_{\text{dae}}$ are compared to the solution $\mathbf{u}_{\text{fem}}$ of the free run of the global problem discretised by 900 bilinear finite elements. Each local problem is discretised with 900 finite elements and with $c_a = 1000$. Results are first shown in Figure 5.4, in which the spatial norm of the DAE and FEM solutions $\mathbf{u}_{\text{dae}}$ and $\mathbf{u}_{\text{fem}}$ are plotted over time. Figure 5.5 shows the similarity of the spatial errors of $\mathbf{u}_{\text{dae}}$ and $\mathbf{u}_{\text{fem}}$.

The minimax estimation error at each time step is controlled by the Riccati matrix via

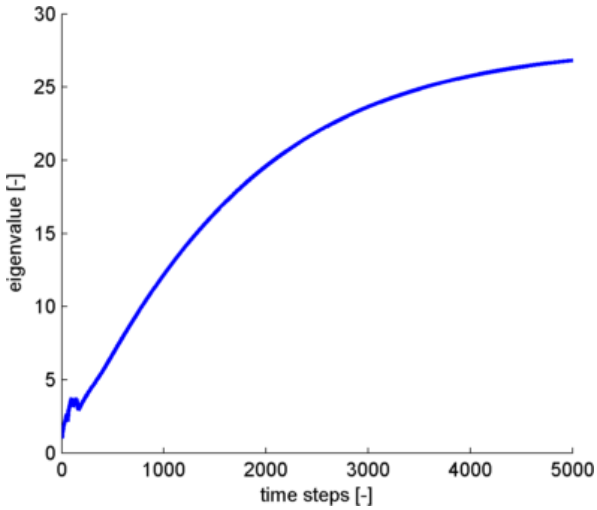

Figure 5.6 The largest eigenvalue of the Riccati matrix of the problem from Experiment 1A plotted over time.

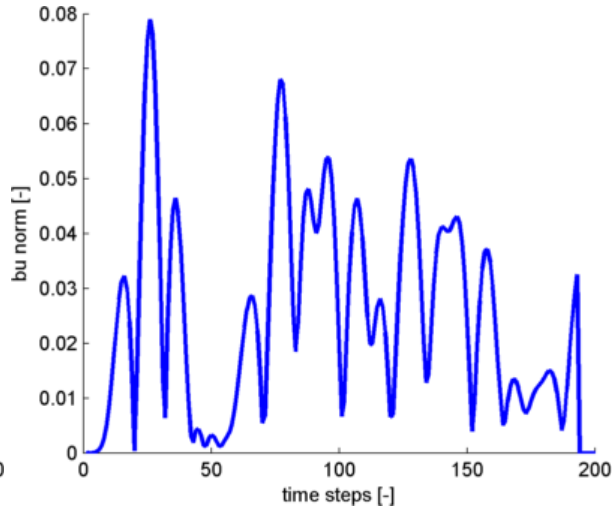

Figure 5.7 The spatial norm of the algebraic constraint of the problem from Experiment 1A plotted over time.

$$l^T\left(\mathbf{u}_{\text{true}}(t) - \mathbf{u}_{\text{minimax}}(t)\right) \leq [l^T\mathbf{K}(t)l]^{\frac{1}{2}} \tag{5.73}$$

where l is an arbitrary vector. Selecting $\mathrm{l} = l_\lambda$ to be an eigenvector that corresponds to the largest eigenvalue λ of the matrix $\mathbf{K}(t)$ results in the following error estimate:

$$l_\lambda{}^T\left(\mathbf{u}_{\text{true}}(t) - \mathbf{u}_{\text{minimax}}(t)\right) \leq \sqrt{\lambda} \tag{5.74}$$

Figure 5.6 tracks the maximum eigenvalue of the Riccati matrix that bounds the minimax estimation error in (5.74). Since the velocity field in this experiment is constant in time, the maximum Riccati eigenvalue is computed for 5000 time steps to demonstrate its convergence to some value and as a result to prove that estimation error is bounded in time.

In Figure 5.7 the norm of the algebraic constraint $\mathbf{Bu}_{\text{dae}}$ is tracked over time. As it can be seen in the graph, this norm is decreasing over time proving convergence property of the DAE minimax filter. This is explained by the fact that, Riccati equation incorporates algebraic constraint into the model and over time the influence of the algebraic constraint on an estimate increases.

The same experiment was made for other configurations as reported in Table 5.1 where the estimation error and the norm of algebraic constraint are presented. From the Table 5.1 it can be observed that the increase in the algebraic constraint c_a corresponds to the decrease in the norm of $\mathbf{Bu}_{\text{dae}}$ and to the decrease of the estimation error e_e.

Table 5.1 Comparison of metrics for different values of algebraic constant c_a

c_a	$\|\mathbf{B}\mathbf{u}_{\text{dae}}\|$	e_e
1	26.6153	26.78 %
10	4.2693	16.23 %
100	1.3715	16.18 %
1000	0.4352	16.17 %

For Setting B, the underlying flow field and domain splitting are the same as above. The initial conditions are not known, but the global domain is completely observed. Observations are generated numerically by adding to $u_a(t, x, y)$ noise sampled from the uniformly distributed interval $[-1, 1]$. The spatial norm and spatial error of the estimate $\mathbf{u}_{\text{dae minimax}}$ obtained from the DAE minimax filter are provided in Figure 5.4 and Figure 5.5 correspondingly which also show the speed of convergence of $\mathbf{u}_{\text{dae minimax}}$ to the traditional minimax estimate $\mathbf{u}_{\text{minimax}}$ and analytical solution. Similarity of the quality of the DAE minimax estimate $\mathbf{u}_{\text{minimax}}$ to the estimation quality of traditional minimax filter is further confirmed be the following estimation errors: $e_e(\mathbf{u}_{\text{dae minimax}}) = 16.8\ \%$ and $e_e(\mathbf{u}_{\text{minimax}}) = 15\ \%$.

It should be noted that, while in Experiment 1A there were no observations, the algebraic part of DAE (5.52) could be considered as the observation equation. Since the pair $(\mathbf{S}, \mathbf{B})$ is not observable (there can be different initial conditions for the corresponding FEM model, leading to solutions that are in the null-space of $\mathbf{B}$), it follows that the algebraic part of DAE does not lead to a decrease in the eigenvalues of the Riccati matrix representing the level of the uncertainty in the system. On the other hand, by adding actual observations a decrease of the maximal eigenvalue of the Riccati matrix is observed.

5.3.2 Experiment 2

For the second test, the underlying flow field was generated by EFDC. The global domain is decomposed into four subdomains as in Figure 5.3 discretised by 1225 finite elements each. Due to the complex behaviour of the underlying flow, the analytical solution for this experiment is not known. So that, the initial conditions for Setting A are defined by the linear spline with parameters as in (4.94) and FEM solution of the corresponding free-run model $\mathbf{u}_{fem}$ is considered as a ground truth, i.e., $\mathbf{u}_{\text{true}} = \mathbf{u}_{\text{fem}}$. For Setting B, observations are generated

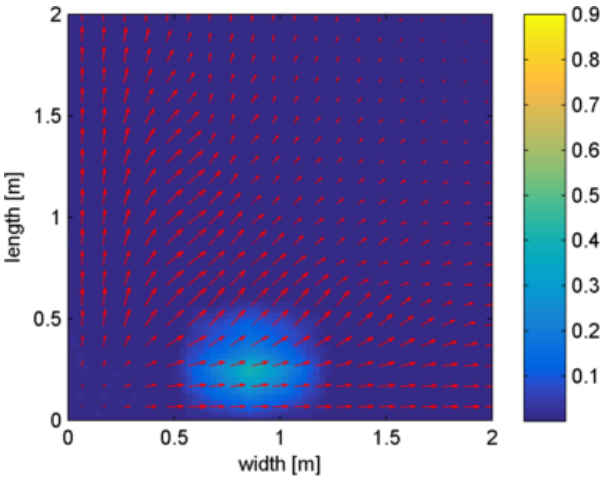

Figure 5.8 EFDC generated velocity field and DAE minimax estimate plotted at time step 80.

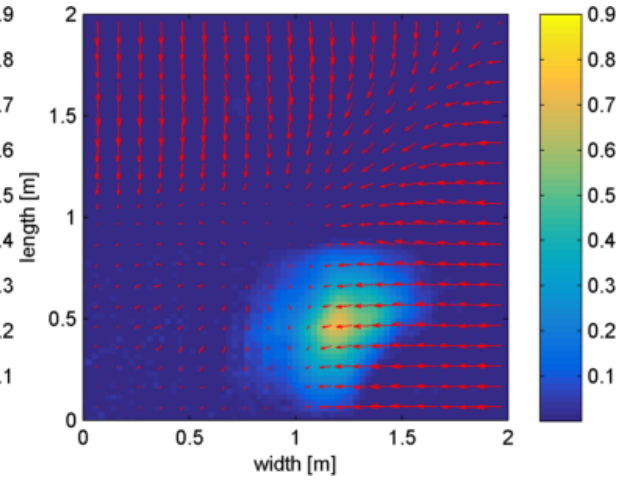

Figure 5.9 EFDC generated velocity field and DAE minimax estimate plotted at time step 430.

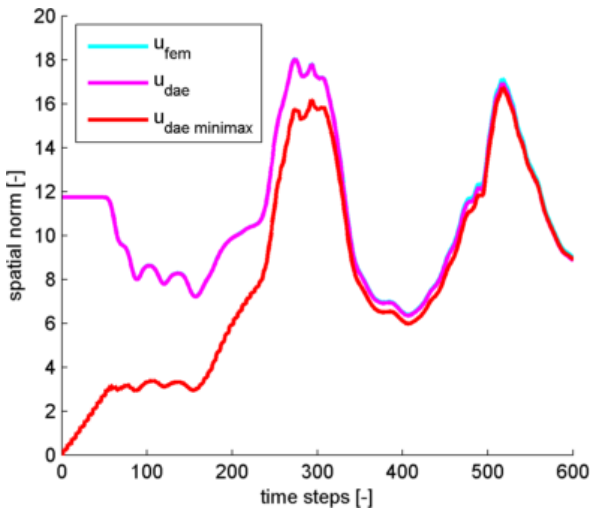

Figure 5.10 The spatial norm of the FEM solution, DAE estimate and DAE minimax estimates plotted over time for the problem from Experiment 2.

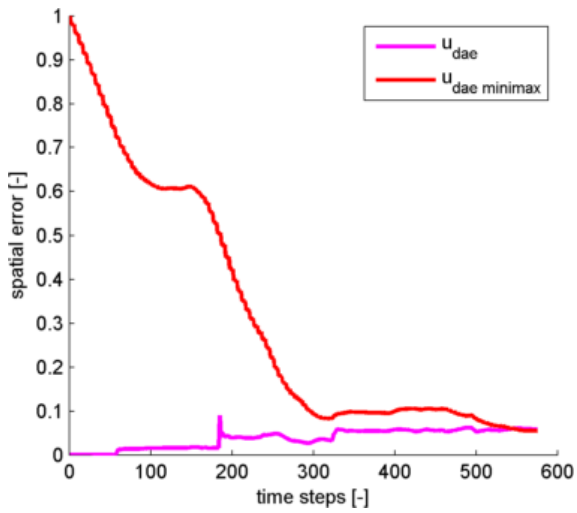

Figure 5.11 The spatial error of the DAE estimate and DAE minimax estimate plotted over time for the problem from Experiment 2.

by restricting $\mathbf{u}_{\text{fem}}$ on the first subdomain (see Figure 5.3), sampling it every five time-steps and adding to it noise sampled from the uniformly distributed interval $[-0.1,0.1]$ so that its largest value is approximately 10% of contaminant concentration largest value.

In comparison to Experiment 1, the underlying flow field is not stationary and changes its direction over time and space. That makes the concentrations to cross the boundaries between the subdomains several times. This is indeed shown in Figure 5.8 and Figure 5.9 where velocity flow field is depicted at two distinct instances of time: (i) at time step 80 in Figure 5.8 and (ii) at time step 420 in Figure 5.9.

Results of the simulations are shown in Figure 5.10 and Figure 5.11 where spatial norms and spatial errors of DAE estimate for Setting A $\mathbf{u}_{\text{dae}}$ and DAE minimax estimate for Setting B $\mathbf{u}_{\text{dae minimax}}$ are plotted over time. As it can be seen, even for a complicated velocity field

minimax filter for DD provides similar results to the FEM solution. Comparing to Experiment 1, this experiment employs a denser FEM grind which results in a better convergence of $\mathbf{u}_{\text{dae}}$ to $\mathbf{u}_{\text{fem}}$. This is justified by the estimation error $e_e(\mathbf{u}_{\text{dae}}) = 4.27\%$ provided $\mathbf{u}_{\text{true}} = \mathbf{u}_{\text{fem}}$.

Figure 5.8 and Figure 5.9 also demonstrate how DAE minimax filter recreates the true state for Setting B from noisy observations. The convergence of the DAE minimax estimate for Setting B $\mathbf{u}_{\text{dae minimax}}$ to the FEM solution that is used for observation sampling is further shown in Figure 5.10 and Figure 5.11 after approximately 300 time steps.

For the simple fluid flow taken in the Experiment 1 the corresponding stiffness matrix is stationary and stable in a Lyapunov sense. In the present experiment, the fluid flow is non-stationary and exhibits a complex behaviour that together with the FEM approximation error amounts to the non-stationary stiffness matrix with positive eigenvalues. This artificial instability is resolved by the stabilizing effect brought by the minimax filter that injects the stabilizing feed-back into the state equation.

5.3.3 Computational performance

Computational performance of the DAE minimax filter is different to the computational performance of the traditional minimax filter in two ways. The application of DD provides computational benefits since it reformulates underlying mass and stiffness matrices into a block structured matrices. This reduces the number of operations involved in their multiplications or inversions and as a result, speeds up computations of the feedback equation. On the other hand, the application of DD also provides computational overhead. The decomposition interface nodes are included several times in the state vector of the DAE in adjacent local subproblems. As a result, the size of the discrete systems (5.67), (5.70) of the DAE minimax filter is increased compared to the size of the system of the traditional minimax filter; this produces the computational overhead.

Assuming the total number of FEM nodes is N_{nd}^{fem}, the complexity of the traditional minimax filter is given as $14\mathcal{C}_{N_{nd}^{fem} N_{nd}^{fem}}$ (see Section 4.2.2). Let N_{nd}^{Γ} describes the number of FEM nodes over the interface of decomposition, then the size of the DAE minimax filter is $N_{nd} = N_{nd}^{fem} + N_{nd}^{\Gamma}$ and its computational complexity is $13\mathcal{C}_{N_{nd} N_{nd}}$ or $13\mathcal{C}_{N_{nd}^{fem}+N_{nd}^{\Gamma}, N_{nd}^{fem}+N_{nd}^{\Gamma}}$. These estimates suggest that the optimal number of subdomains should be chosen such that it balances the benefits and overhead introduced by the DAE minimax filter.

The theoretical considerations are further tested on two numerical simulations. Simulation 1 is performed using global domain of the size $[0,6] \times [0,1]$ metres and is decomposed in three different ways: with 2, 3 and 6 subdomains. Simulation 2 is performed using global domain of the size $[0,8] \times [0,1]$ metres and is also decomposed in three different ways: with 2, 4 and 8 subdomains. The remaining configurations are similar to those as in the Experiment 1.B.

Table 5.2 Comparison of CPU time measurements for experiments with different number of subdomains

Simulation 1			Simulation 2		
Traditional minimax	DAE minimax		Traditional minimax	DAE minimax	
Time	number of subdomains	Time	Time	number of subdomains	Time
$789\ s$	2	$761\ s$	$1910\ s$	2	$1811\ s$
	3	$786\ s$		4	$1857\ s$
	6	$873\ s$		8	$2064\ s$

Measurements of execution CPU time of those simulations are presented in Table 5.2. Those results demonstrate that indeed for a small number of subdomains in the DAE minimax filter, the benefits related to a block structure of mass and stiffness matrices are more significant than the overhead driven by the increase of the number of interface nodes N_{nd}^{Γ} and the corresponding increase of state vector. In this case, the DAE minimax filter with 2 and 3 subdomains in Simulation 1 and the DAE minimax filter with 2 and 4 subdomains in Simulation 2 are computed faster than the corresponding traditional minimax filters. If the number of subdomain is increasing, the associated overhead becomes more significant than the associated benefits. In Table 5.2 this is shown for the DAE minimax filter with 6 subdomains in Simulation 1 and for the DAE minimax filter with 8 subdomains in Simulation 2.

5.4 Discussion and conclusions

This chapter presented an approach for DA problems that combines both DD and minimax filters. The original problem was reformulated into DAE and solved using the Kalman duality principle and the minimax framework.

The proposed approach follows the standard idea of decomposition, but, enforces transmission conditions on a discrete level as an optimisation constraint in an algebraic form. This allows for independent choice of boundary conditions between subdomains so that the approach can be easily applied for complicated physical processes including heterogeneous problems.

Of note is that the application of DD did not reduce the complexity of the presented method in comparison to the traditional minimax filter as expected. While the localised stiffness and observation matrices (5.24) and (5.32) possess a block-matrix structure (due to DD), the presence of the algebraic constraint inside terms (5.55) and (5.56), and correspondingly inside the Riccati equation (5.67), blends the matrix components which leads to a general form of the Riccati matrix. As a result, the full size Riccati and feedback equations (5.67) and (5.70) are solved.

The decomposition influences the computational performance of the DAE minimax algorithm in two ways: on the one hand, it decreases the computational cost associated with the computation of local matrices, but on the other it increases the computational cost associated with the solution of the Riccati equation. The numerical experiments demonstrate that for a small number of subdomains, the execution CPU time of the presented approach is smaller than for the traditional minimax approach. When the number of subdomains exceeds a certain threshold, the execution CPU time of the DAE minimax filter is higher than for the traditional minimax filter. As Table 5.2 suggests, for Simulation 1 and Simulation 2 the optimal number of subdomains is 2. However, it is possible that for simulations with larger numbers of FEM nodes, the optimal number of subdomains would be higher.

For the simulations used in this chapter, the algorithm of the DAE minimax filter is implemented in a sequential way and performed using a single processor machine with 4 cores. An additional advantage of the presented algorithm in contrast to the traditional algorithm is that computations of local matrices can be naturally organised in parallel on a multi-processor computer. For instance, each subdomain can be assigned to a separate computation node. This will not change the execution CPU time but will decrease the execution wall-clock time compared to the non-parallel implementation.

The application of DD for transport phenomena has a potential for the activation or deactivation of subproblems that are, or are not, involved in the computations for the current time step. This may reduce the computational cost of the method. Activation/deactivation may be implemented as the addition/elimination of the corresponding blocks in the matrices of the Riccati equation

(5.67) which destroys the optimality properties of the minimax framework and breaks the equivalence with the traditional minimax filter. While numerical simulations with subdomain activation show good results for free run model, for DAE minimax subproblem deactivation leads to the deterioration of the estimates. More stable deactivation procedure requires additional investigation.

6 Interconnected localised minimax filter

To address the problem of the high computational demands of the traditional minimax filter in the previous chapter, the approach which combines DD with the minimax filtering was developed in the form of DAE. In that case, the transmission of data between subdomains was represented by a set of trace operators, discretised and enforced through the additional algebraic constraint which produced the DAE.

In this chapter, another approach of DD is used in combination with the minimax filter. In this case, the transmission conditions are enforced as additional boundary conditions over common boundaries between subdomains. Comparing to the DAE problem, the localised problem is continuous and can be solved using the Schwarz iterative approach. As a result, the estimation problem is transformed from one general filter into a number of small local filters that are interconnected.

The choice of the Schwarz DD techniques that is suitable for the solution of advection-diffusion problems depends on the nature of the problem, in particular, on the configuration of boundary conditions. It is possible to distinguish two groups of Schwarz DD methods (Quarteroni and Valli, 1999; Toselli and Widlund, 2005): (i) the Dirichlet-Neumann techniques provide a good approximation when the diffusive part is dominant, but may produce discrepancies when the advective part becomes more relevant; (ii) the natural boundary conditions techniques provide a good approximation when the advection part is dominant, enforcing boundary conditions over the interface taking into account the hyperbolic nature of the problem, but may produce discrepancies when the diffusive part of the problem is dominant. Since this work deals with advection-dominated flows, it is reasonable to adopt the second approach, which is also known in the literature as adaptive Dirichlet-Neumann (ADN) DD (Gastaldi et al., 1998).

Some of the results of this chapter were published by the author in Ragnoli et al. (2015), however, this chapter contains several significant extensions relative to that paper: i) in this chapter problem errors are defined in terms of $\mathcal{L}^\infty$ ellipsoids in contrast to $\mathcal{L}^2$ ellipsoids used in that paper (see Section 6.1.1); ii) the localised filter is defined for a continuous in time and space problem (see Section 6.1.1); iii) the FEM approximation of the localised filter are presented in a much more detailed way (see Section 6.1.2); iv) variants of the filter with the reinitialisation procedure and pseudo-observations are introduced in Section 6.1.4 and Section

6.1.5; v) an additional set of numerical simulation with non-stationary velocity flow is added (see Section 6.2.2); vi) last but not least, this chapter provides an investigation of the computational performance of the localised minimax filter in Section 6.2.3. The main contribution of the author of this thesis to that paper are the mathematical formulation of the localised filters, code development and numerical simulations.

6.1 Global and Localised Problems

The global problem is the continuous DA problem defined on a spatiotemporal domain $([t_0; t_1] \times \Omega)$ with a background model

$$\begin{cases} \dfrac{\partial u(t,x)}{\partial t} = Lu(t,x) + e_m(t,x) \\ u(0,x) = u^0(x) + e_0(x) \\ u(t,x) = 0, \qquad x \in \partial\Omega \end{cases} \tag{6.1}$$

where

$$\begin{aligned} &L: H^1([t_0; t_1] \times \Omega) \to L^2([t_0; t_1] \times \Omega) \\ &u \mapsto Lu = \epsilon\Delta u - \nabla \cdot (\mu u), \epsilon = const \end{aligned} \tag{6.2}$$

The observation equation is written as

$$y(t,x) = Hu(t,x) + e_o(t,x) \tag{6.3}$$

where H is the linear operator

$$Hu(t,x) = \int_\Omega h(z-x)u(t,z)dz \tag{6.4}$$

The model error $e_m(t,x)$ and the initial condition error $e_0(x)$ are uncertain but bounded inside the L^∞-ellipsoid

$$\mathcal{L}_m^\infty = \{e_0(x), e_m(t,x): |e_0(x)| \le \tilde{q}_0(x), |e_m(x)| \le \tilde{q}(t,x)\} \tag{6.5}$$

The observation error $e_o(t,x)$ is assumed to be a white noise with bounded second moments taken as an elements of the following convex bounded set

$$\mathcal{L}_o^\infty = \left\{e_o(t,x): \overline{e_o(t,x)} = 0, \overline{e_o^2(t,x)\tilde{r}(t,x)} \le 1\right\} \tag{6.6}$$

The weighting functions $\tilde{q}_0(x)$, $\tilde{q}(t,x)$ and $\tilde{r}(t,x)$ are positive and bounded.

In the following, the continuous global problem (4.1)-(4.7) is localised by splitting it into subproblems using the ADN DD approach. It should be noted that not only the background and observation equations but also the error ellipsoids are restricted onto local subdomain. Since L^{∞}-ellipsoid imposes pointwise restrictions on its elements, it is in a good agreement with the geometrical decomposition of the domain of interest. After that, the minimax filters are formulated for each subproblem independently and exchange information through the inherited mechanism from ADN method. Finally, the set of local filters is computed using iterative methods.

6.1.1 Interconnected localised minimax filters

Let the domain Ω be divided into N non-overlapping subdomains $\Omega_1, \dots, \Omega_N$. Besides the common boundary between subdomains $\Gamma_{i,j}$ and decomposition interface Γ (see Section 5.1.1), the inflow and outflow parts of the $\Gamma_{i,j}$ and $\partial\Omega_i$ are defined as follows

$$\Gamma_{i,j}^{in} = \{x \in \Gamma_{i,j} : \mu(x) \cdot n(x) < 0\} \tag{6.7}$$

$$\Gamma_{i,j}^{out} = \{x \in \Gamma_{i,j} : \mu(x) \cdot n(x) > 0\} \tag{6.8}$$

$$\partial\Omega_i^{in} = \{x \in \partial\Omega_i : \mu(x) \cdot n(x) < 0\} \tag{6.9}$$

$$\partial\Omega_i^{out} = \{x \in \partial\Omega_i : \mu(x) \cdot n(x) > 0\} \tag{6.10}$$

$$\Gamma_i^{in} = \{\Gamma_{i,j}^{in} : \Gamma_{i,j}^{in} \neq \emptyset\} \tag{6.11}$$

$$\Gamma_i^{out} = \{\Gamma_{i,j}^{out} : \Gamma_{i,j}^{out} \neq \emptyset\} \tag{6.12}$$

where $n(x)$ is the outward normal vector at the point x.

The continuous background model (4.1) is equivalent to a set of local backgrounds

$$\begin{cases} \dfrac{\partial u_i}{\partial t} = L_i u_i + e_i \\ u_i(t,x) = 0, \qquad \text{on } \partial\Omega \cap \partial\Omega_i \\ u_i(t,x) = u_j(t,x), \qquad \text{on } \Gamma_{i,j}^{in} \in \Gamma_i^{in} \\ \dfrac{\partial u_i(t,x)}{\partial n} = \dfrac{\partial u_j(t,x)}{\partial n}, \qquad \text{on } \Gamma_{i,j}^{out} \in \Gamma_i^{out} \\ u_i(0,x) = u_{0,i}(x) + e_{0,i}(x) \end{cases} \tag{6.13}$$

where the local operator L_i is the restriction of the original operator L on Ω_i, and e_i, $u_{0,i}$ and $e_{0,i}$ are the restrictions of e, u^0 and e_0 onto Ω_i (see Section 5.1.1). The localised model and initial conditions errors $e_{0,i}$ and e_i belong to

$$\begin{aligned} \mathcal{L}_{m,i}^{\infty} = \{ & e_{0,i}(x), e_{m,i}(t,x) : |e_{0,i}(x)| \le \tilde{q}_{0,i}(x), \\ & |e_{m,i}(t,x)| \le \tilde{q}_i(t,x) \text{ on } \Omega_i \} \end{aligned} \tag{6.14}$$

the restriction of $\mathcal{L}_m^{\infty}$ onto Ω_i. The restriction of the observation equation is obvious and defined as

$$\begin{aligned} & y_i(t,x) = H_i u_i(t,x) + e_{o,i}(t,x), \\ & H_i u_i(t,x) = \int_{\Omega_i} h(x-y) u_i(x,t) dx \end{aligned} \tag{6.15}$$

where the local observation error $e_{o,i}$ belongs to the local L^{∞} observations ellipsoid

$$\mathcal{L}_{o,i}^{\infty} = \left\{ e_{o,i}(x), \overline{e_{o,i}^2(t,x)\tilde{r}_i(t,x)} \le 1 \text{ on } \Omega_i \right\} \tag{6.16}$$

The functions $\tilde{q}_{0,i}$, $\tilde{q}_i$ and $\tilde{r}_i$ are corresponding restrictions of $\tilde{q}_0$, $\tilde{q}$ and $\tilde{r}$.

Equations (6.13)-(6.16) define the i-th local problem and a set of N local problems (6.13)-(6.16) define the localised problem. The choice of the boundary conditions on the interface boundaries Γ_i^{in} and Γ_i^{out} guarantees the continuity of the solution of the localised problem across the interface Γ. Boundary conditions on external boundaries $\partial\Omega \cap \partial\Omega_i$ are inherited from the global problem which implies the equivalence between the localised and global solutions.

The localised problem described above is an application of the DD technique known as the ADN method (Gastaldi et al., 1998). Since advection-dominated flows are considered, a further modification of the formulation (6.13) is possible. For the pure advection problems, the outflow boundary conditions on $\Gamma_{i,j}^{out}$ are not required as it follows from the physical properties of the flow $\mu(t,x)$. This suggests the hyperbolic nature of the background model can be incorporated

into (6.13) by imposing the homogeneous Neumann condition, which leads to a damped ADN (d-ADN) decomposition. The latter is known to work well for advection-dominated problems (Ciccoli, 1996). The actual computational scheme is then carried out by solving for u_i over Ω_i and iterating until convergence, a so called iterative Schwarz approach (Quarteroni and Valli, 1999). More specifically, it starts with a set of initial solutions $\{u_i^0\}$, and compute $\{u_i^{k+1}\}$ from $\{u_i^k\}$, $n \geq 0$ by solving numerically the following problem:

$$\begin{cases} \dfrac{\partial u_i^{k+1}}{\partial t} = L_i u_i^{k+1} + e_{m,i} \\ u_i^{k+1}(t,x) = 0, \qquad \text{on } \partial\Omega^{in} \cap \partial\Omega_i^{in} \\ u_i^{k+1}(t,x) = u_j^k(t,x), \qquad \text{on } \Gamma_{i,j}^{in} \in \Gamma_i^{in} \\ \dfrac{\partial u_i^{k+1}(t,x)}{\partial n} = 0, \qquad \text{on } \Gamma_{i,j}^{out} \in \Gamma_i^{out} \\ u_i^{k+1}(0,x) = u_{0,i}(x) + e_{0,i}(x) \end{cases} \tag{6.17}$$

The heuristic purpose of the Schwarz iterations defined in (6.17) is to enforce the continuity of the solution of the problem (6.17) along the interfaces. Once this is achieved, the iterative process can be stopped. While this work does not study the rate of the convergence of the iterative Schwarz d-ADN method, it should be noted that if the direction of the flow is constant, only one iteration of the Schwarz method is required. In the general case, it can be shown that the sequence $\{u_i^k\}$ converges weakly in $H^1(\Omega_1) \times \ldots \times H^1(\Omega_N)$ to the unique solution of the localised problem, provided the latter exists (Gastaldi et al., 1998).

The minimax estimate (analysis) $u_i^{a,k+1}$ of u_i^{k+1}, the solution of the k-th Schwarz iteration for i-th local problem (sometimes referred to as the $(k+1,i)$-filter), is introduced given observations y_i, data from the previous Schwarz iteration u_i^k, and assuming that $e_{0,i}, e_{m,i} \in \mathcal{L}_{m,i}^2$, and $e_{o,i} \in \mathcal{L}_{o,i}^2$.

To construct minimax filters for each subproblem, the local error ellipsoids $\mathcal{L}_{m,i}^\infty$ and $\mathcal{L}_{o,i}^\infty$ are approximated by L^2-ellipsoids (see Section 4.1.1). As a result, denoting $T = t_1 - t_0$, the model error and initial conditions error are bounded in

$$\mathcal{L}_{m,i}^{2} = \left\{ e_{0,i}(x), e_{m,i}(t,x) : \int_{\Omega_i} e_{0,i}^{2}(x)\tilde{q}_{0,i}^{-2}(x)d\Omega_i \right.$$
$$+ \int_{[t_0;t_1]\times\Omega_i} e_{m,i}^{2}(t,x)\tilde{q}_i^{-2}(t,x)d\Omega_i dt \tag{6.18}$$
$$\left. \leq (T+1)A(\Omega_i) \right\}$$

which obviously contains $\mathcal{L}_{m,i}^{\infty}$. It needs to be stressed that the union of the "small" ellipsoids $\mathcal{L}_{m,i}^{2}$, approximating $\mathcal{L}_{m,i}^{\infty}$ is contained in the large ellipsoid $\mathcal{L}_{m}^{2}$ approximating the entire $\mathcal{L}_{m}^{\infty}$.

Similarly, $\mathcal{L}_{o}^{\infty}$ is approximated by

$$\mathcal{L}_{o,i}^{\infty} = \left\{ e_o : \int_{[t_0;t_1]\times\Omega_i} \overline{e_{o,i}^{2}(t,x)\tilde{r}_i(t,x)dxdt} \leq (1+T)A(\Omega_i) \right\} \tag{6.19}$$

Ellipsoids (6.18)-(6.19) can be equivalently rewritten as

$$\mathcal{L}_{m,i}^{2} = \left\{ e_{0,i}(x), e_{m,i}(t,x) : \int_{\Omega_i} e_{0,i}^{2}(x)q_{0,i}(x)dx \right.$$
$$\left. + \int_{\Omega_i\times(0,T)} e_{m,i}^{2}(t,x)q_{m,i}(t,x)dxdt \leq 1 \right\} \tag{6.20}$$

$$\mathcal{L}_{o,i}^{\infty} = \left\{ e_{o,i} : \int_{\Omega_i\times(0,T)} \overline{e_{o,i}^{2}(t,x)r_i(t,x)dxdt} \leq 1 \right\} \tag{6.21}$$

where

$$q_{0,i} = \frac{1}{\gamma_{T,i}\tilde{q}_{0,i}^{2}}, \qquad q_{m,i} = \frac{1}{\gamma_{T,i}\tilde{q}_{m,i}^{2}}, \qquad r_i = \frac{\tilde{r}_i}{\gamma_{T,i}} \tag{6.22}$$

$$\gamma_{T,i} = (1+T)A(\Omega_i) \tag{6.23}$$

and factor $\gamma_{T,i}$ reflects the error of approximating $\mathcal{L}_{m,i}^{\infty}$ and $\mathcal{L}_{o,i}^{\infty}$ by $\mathcal{L}_{m,i}^{2}$ and $\mathcal{L}_{o,i}^{2}$.

Each local solution u_i^{k+1} of the i-th local problem (6.13), (6.15) with the errors described by (6.19) and (6.20) is the sum of a "mean" local solution a_i^{n+1} and noisy part b_i, i.e.,

$$u_i^{k+1} = a_i^{k+1} + b_i \tag{6.24}$$

provided a_i^{k+1} solves

$$\begin{cases} \dfrac{\partial a_i^{k+1}}{\partial t} = L_i a_i^{k+1} \\ a_i^{k+1}(t,x) = 0, \qquad \text{on } \partial\Omega^{in} \cap \partial\Omega_i^{in} \\ a_i^{k+1}(t,x) = a_j^k(t,x) + b_j(t,x), \qquad \text{on } \Gamma_{i,j}^{in} \in \Gamma_i^{in} \\ \dfrac{\partial a_i^{k+1}(t,x)}{\partial n} = 0, \qquad \text{on } \Gamma_{i,j}^{out} \in \Gamma_i^{out} \\ a_i^{k+1}(0,x) = u_{0,i}(x) \end{cases} \tag{6.25}$$

and b_i solves

$$\begin{cases} \dfrac{\partial b_i}{\partial t} = L_i b_i + e_i \\ b_i(t,x) = 0, \qquad \text{on } \partial\Omega^{in} \cap \partial\Omega_i^{in} \\ b_i(t,x) = 0, \qquad \text{on } \Gamma_{i,j}^{in} \in \Gamma_i^{in} \\ \dfrac{\partial b_i(t,x)}{\partial n} = 0, \qquad \text{on } \Gamma_{i,j}^{out} \in \Gamma_i^{out} \\ b_i(0,x) = e_{0,i}(x) \end{cases} \tag{6.26}$$

From (6.25) it is concluded that a_i^{k+1} depends linearly on b_j and a_j^k, where Ω_j is adjacent to Ω_i. Hence the minimax estimate of a_i^{k+1} is given by $a_i^{a,k+1}$, the solution of (6.25) which corresponds to $a_i^{k+1} = u_j^{a,k}$ on $\Gamma_{i,j}^{in} \in \Gamma_i^{in}$, where $u_j^{a,k}$ denotes the (k,j)-filter obtained on the k-th iteration of the Schwarz iterative procedure. Since $u_i^{k+1} = a_i^{k+1} + b_i$, it follows that

$$y_i = H_i u_i^{k+1} + e_{o,i} = H_i a_i^{k+1} + H_i b_i + e_{o,i} \tag{6.27}$$

so the noisy part b_i^{k+1} can be estimated from the shifted local measurements $\tilde{y}_i = y_i - H_i a_i^{a,k+1}$. Although the noisy part b_i^{k+1} is independent of the corresponding noisy parts b_j^{k+1}, its minimax estimate $b_j^{a,k+1}$ does depend on observations $\tilde{y}_i$ which, in turn, depend on $a_j^{a,k}$, so $b_j^{a,k+1}$ changes over the course of the Schwarz iterative procedure. That said, the minimax estimate $u_i^{a,k+1}$ can be computed as the sum of $a_i^{a,k+1}$ and $b_i^{a,k+1}$ as

$$\ell_i\left(u_i^{a,k+1}\right) = \ell_i\left(a_i^{a,k+1}\right) + \ell_i\left(b_i^{a,k+1}\right) \tag{6.28}$$

where, analogously to Section 4.1.2, the minimax estimate $b_i^{a,k+1}$ is represented as follows

$$\ell_i\left(b_i^{a,k+1}\right) = \int_{\Omega_{T,i}} r_i(t,x)(H_i p_i)(t,x)\tilde{y}_i(t,x)dxdt \tag{6.29}$$

provided w_i and p_i solve the following Hamiltonian system of equations:

$$
\begin{cases}
\dfrac{\partial w_i}{\partial t} = -L_i^* w_i + H_i^* r_i(t,x) H_i p_i \\
w_i(T,x) = l_i(x), \qquad \text{on } \Omega_i \\
w_i(t,x) = 0, \qquad \text{on } \partial\Omega^{in} \cap \partial\Omega_i^{in} \\
w_i(t,x) = 0, \qquad \text{on } \Gamma_{i,j}^{in} \in \Gamma_i^{in} \\
\dfrac{\partial w_i(t,x)}{\partial n} = 0, \qquad \text{on } \Gamma_{i,j}^{out} \in \Gamma_i^{out} \\
\dfrac{\partial p_i}{\partial t} = -L_i p_i + q_i w_i \\
p_i(0,x) = q_{0,i}(x) w_i(0,x), \qquad \text{on } \Omega_i \\
p_i(t,x) = 0, \qquad \text{on } \partial\Omega^{in} \cap \partial\Omega_i^{in} \\
p_i(t,x) = 0, \qquad \text{on } \Gamma_{i,j}^{in} \in \Gamma_i^{in} \\
\dfrac{\partial p_i(t,x)}{\partial n} = 0, \qquad \text{on } \Gamma_{i,j}^{out} \in \Gamma_i^{out}
\end{cases}
\tag{6.30}
$$

where l_i stands for the restriction of l onto Ω_i.

The relation between w_i and p_i is expressed by the local linear mapping $P_i: L^2(\Omega_i) \to L^2(\Omega_i)$ such that

$$
p_i(t,x) = (P_i w_i)(t,x) = \int_\Omega k_i(t,x,z) w_i(z) d\gamma \tag{6.31}
$$

Substituting (4.26) into the Hamiltonian system (6.30) the local kernel function $k_i(t,x,z)$ referred to as the local Riccati kernel is found from the following local operator Riccati equation (see Section 4.1.2):

$$
\begin{cases}
\dot{k}_i = L_{x,i} k_i + L_{z,i} k_i + q_i^{-1}(t,x)\delta(z-x) - P_i\big(H_i^* r_i (H_i k_i)\big) \\
k_i(0,x,z) = q_{0,i}^{-1}(x)\delta(z-x) \text{ on } \Omega
\end{cases}
\tag{6.32}
$$

The minimax estimate $b_i^{a,k+1}$ is then found from ODE:

$$
\begin{cases}
\dot{b}_i = L_i b_i + P_i H_i^* r_i(\tilde{y}_i - H_i b_i) \\
b_i(t,x) = 0 \quad \text{on } [t_0, t_1] \times \partial\Omega \\
b_i(0,x) = 0 \quad \text{on } \Omega
\end{cases}
\tag{6.33}
$$

Since the estimate $u_i^{a,k+1}$ is found as the sum of estimates $a_i^{a,k+1}$ and $b_i^{a,k+1}$, the equations (6.25) (with modified boundary condition) and (4.32) can be added resulting in a feedback equation for the estimate $u_i^{a,k+1}$

$$\begin{cases} \dot{u}_i^{k+1} = L_i u_i^{k+1} + P_i H_i^* r_i\left(y_i - H_i u_i^{k+1}\right) \\ u_i^{k+1}(t,x) = 0 \quad \text{on } \partial\Omega^{in} \cap \partial\Omega_i^{in} \\ u_i^{k+1}(t,x) = u_j^k, \quad \text{on } \Gamma_{i,j}^{in} \in \Gamma_i^{in} \\ \dfrac{\partial u_i^{k+1}(t,x)}{\partial n} = 0, \quad \text{on } \Gamma_{i,j}^{out} \in \Gamma_i^{out} \\ u_i^{k+1}(0) = 0 \quad \text{on } \Omega \end{cases} \tag{6.34}$$

with the local worst-case mean-squared estimation error given by

$$\begin{aligned} &\sup_{(e_{0,i},e_{m,i})\in\mathcal{L}_{m,i}^2, e_{o,i}\in\mathcal{L}_{o,i}^2} \ell_i\left(u_i^{a,k+1}(t_1)\right) - \ell_i\left(u_i(t_1)\right)^2 \\ &= \int_\Omega l_i(x)(P_i l_i)(t_1,x)dx \end{aligned} \tag{6.35}$$

The set of equations (6.34), (4.26) and (4.31) defines localised minimax filter. In order to solve localised minimax filter, the Schwarz iterative procedure should be employed.

6.1.2 Finite Element Approximation for the $(k+1, i)$ - filter

The discretisation of the localised problem and localised minimax filter is obtained similarly to the global problem (see section 3.2.3) and traditional minimax filter (see section 4.1.3). However, since the d-ADN method requires comprehensive treatment of the boundary conditions, the FEM discretisation of the i-th subproblem on k-th Schwarz iteration is developed in details.

The FEM discretisation of the local problem consists of:

(i) reformulating the problem (6.17) in the weak form,

(ii) applying the Galerkin projection method to construct

$$\mathbf{u}_i^{k+1} = \left(\mathbf{u}_{i,1}^{k+1}(t), \dots, \mathbf{u}_{i,N_{nd}^i}^{k+1}(t)\right)^T \tag{6.36}$$

the FEM approximation of the local solution u_i^{k+1}

$$u_i^{k+1} \approx \sum_{s=1}^{N_{nd}^i} \mathbf{u}_{i,s}^{k+1}(t)\, \phi_s(x) \tag{6.37}$$

in the so-called FEM space, spanned over basis function $\phi_s(x)$, $s = 1, \dots, N_{nd}^i$. Note that $x = (x^1, x^2)$ is a point in a two-dimensional space, and $\phi_s(x)$ is a two dimensional function.

A weak formulation of the problem (6.17) is obtained by multiplying (6.17) by a test function $v \in H_0^1(\Omega_i)$; integrating it over the spatial domain Ω_i and applying the divergence theorem in order to enforce the boundary conditions (in the weak sense). As a result, the solution of the problem (6.17) is found as a solution of the following i-th local weak problem:

$$\int_{\Omega_i} \dot{u}_i^{k+1} v dx = \tag{6.38}$$

$$-\int_{\Omega_i} \epsilon \left(\frac{\partial u_i^{k+1}}{\partial x^1} \frac{\partial v}{\partial x^1} + \frac{\partial u_i^{k+1}}{\partial x^2} \frac{\partial v}{\partial x^2} \right) dx \tag{6.39}$$

$$+\int_{\Omega_i} u_i^{k+1} \left(\mu_1 \frac{\partial v}{\partial x^1} + \mu_2 \frac{\partial v}{\partial x^2} \right) dx \tag{6.40}$$

$$+\int_{\partial\Omega \cap \partial\Omega_i} \epsilon \frac{\partial u_i^{k+1}}{\partial n} v dx \tag{6.41}$$

$$+\int_{\Gamma_{i,j}^{in} \in \Gamma_i^{in}} \epsilon \frac{\partial u_i^{k+1}}{\partial n} v dx \tag{6.42}$$

$$+\int_{\Gamma_{i,j}^{out} \in \Gamma_i^{out}} \epsilon \frac{\partial u_i^{k+1}}{\partial n} v dx \tag{6.43}$$

$$-\int_{\partial\Omega \cap \partial\Omega_i} (\mu_1 n_1 + \mu_2 n_2) u_i^{k+1} v dx \tag{6.44}$$

$$-\int_{\Gamma_{i,j}^{in} \in \Gamma_i^{in}} (\mu_1 n_1 + \mu_2 n_2) u_j^{k} v dx \tag{6.45}$$

$$-\int_{\Gamma_{i,j}^{out} \in \Gamma_i^{out}} (\mu_1 n_1 + \mu_2 n_2) u_i^{k+1} v dx \tag{6.46}$$

where $d\Omega$ and $d\Gamma$ denote the differentials for the integrals over the subdomains and parts of their boundaries. In d-ADN the decomposition integral (6.43) vanishes.

The FEM discretisation of (6.17) proceeds by means of polygonal finite elements Λ_m, $m = 1,..,N_{el}^i$, i.e., the domain Ω_i is divided into a finite number of polygons Λ_m with vertices x_r, $r = 1..N_{nd}^i$, $\Omega \approx \cup \Lambda_m$. The vertices x_r form the FEM grid, and at each node x_r of this grid, the corresponding piece-wise linear basis function ϕ_s ("hat function") is defined that satisfies

$$\phi_s(x_r) = \delta_{sr} \tag{6.47}$$

where δ_{sr} is the Kronecker delta. The vertices of the FEM grid that are involved in the computation of the boundary integrals (6.41)-(6.46) are grouped in the following inflow/outflow subsets of indices

$$D_{in/out} = \{r: x_r \in \Gamma_{i,j}^{in/out}\} \tag{6.48}$$

$$N_{in} = \{r: \exists m \text{ that } \partial\Lambda_m \cap \Gamma_{i,j}^{in} \neq \emptyset \text{ and } x_r \in \Lambda_m\} \tag{6.49}$$

$$I = \{r: x_r \notin D_{in/out} \cup N_{in}\} \tag{6.50}$$

To find the coefficients $\mathbf{u}_{is}^{k+1}$ the representation (6.37) is substituted into (6.38)-(6.46) which leads to the FEM model for the coefficients:

$$\begin{cases} \mathbf{M}_i \dfrac{d\mathbf{u}_i^{k+1}(t)}{dt} = \tilde{\mathbf{S}}_i(t)\mathbf{u}_i^{k+1}(t) + \tilde{\mathbf{f}}_i\left(t; \mathbf{u}_j^k\right) + \mathbf{M}_i\mathbf{e}_i \\ \mathbf{u}_i^{k+1}(0) = \mathbf{u}_i^0 + \mathbf{e}_{0,i} \end{cases} \tag{6.51}$$

where $\mathbf{u}_i^0$ is the FEM approximation of the restriction of u^0 onto Ω_i, $\mathbf{e}_{m,i}$ and $\mathbf{e}_{0,i}$ are the vectors of coefficients of the spatial FEM discretisation of the model and initial errors,

$$\mathbf{M}_i = \left\{ \int_{\Omega_i} \phi_s \phi_r dx \right\}_{s,r=1}^{N_{nd}^i} \tag{6.52}$$

is the local mass matrix, $\mathbf{S}_i(t)$ is the local stiffness matrix defined by

$$\tilde{\mathbf{S}}_i(t) = \mathbf{S}_i^{\Omega}(t) + \mathbf{S}_i^{N_{in}}(t) + \mathbf{S}_i^{D_{out}}(t) \tag{6.53}$$

where

$$\mathbf{S}_i^{\Omega}(t) = \begin{pmatrix} \mathbf{S}_{D_{in}D_{in}}^{\Omega} & \mathbf{S}_{D_{out}D_{in}}^{\Omega} & \mathbf{S}_{N_{in}/\Gamma D_{in}}^{\Omega} & \mathbf{S}_{ID_{in}}^{\Omega} \\ \mathbf{S}_{D_{in}D_{out}}^{\Omega} & \mathbf{S}_{D_{out}D_{out}}^{\Omega} & \mathbf{S}_{N_{in}/\Gamma D_{out}}^{\Omega} & \mathbf{S}_{ID_{out}}^{\Omega} \\ \mathbf{S}_{D_{in}N_{in}/\Gamma}^{\Omega} & \mathbf{S}_{D_{out}N_{in}/\Gamma}^{\Omega} & \mathbf{S}_{N_{in}/\Gamma N_{in}/\Gamma}^{\Omega} & \mathbf{S}_{IN_{in}/\Gamma}^{\Omega} \\ \mathbf{S}_{D_{in}I}^{\Omega} & \mathbf{S}_{D_{out}I}^{\Omega} & \mathbf{S}_{N_{in}/\Gamma I}^{\Omega} & \mathbf{S}_{II}^{\Omega} \end{pmatrix} \tag{6.54}$$

and

$$\mathbf{S}_{X,Y}^{\Omega} = \left\{ -\int_{\Omega_i} \epsilon \left(\frac{\partial \phi_s}{\partial x^1} \frac{\partial \phi_r}{\partial x^1} + \frac{\partial \phi_s}{\partial x^2} \frac{\partial \phi_r}{\partial x^2} \right) dx \right.$$
$$+ \int_{\Omega_i} \phi_s \left(\mu_1 \frac{\partial \phi_r}{\partial x^1} + \mu_2 \frac{\partial \phi_r}{\partial x^2} \right) dx \tag{6.55}$$
$$\left. + \int_{\partial\Omega \cap \partial\Omega_i} \epsilon \frac{\partial \phi_s}{\partial n} \phi_r dx \right\}_{s \in X, r \in Y}$$

with X and Y corresponding to the subsets of the indices of the basis functions, e.g., $X = D_{in}$ and $Y = D_{out}$. In fact, $\mathbf{S}_i^{\Omega}$ absorbs the integrals (6.39), (6.40) and (6.41). $\mathbf{S}_i^{N_{in}}(t)$ is defined as follows

$$\mathbf{S}_i^{N_{in}}(t) = \begin{pmatrix} \mathbf{S}_{D_{in}D_{in}}^{N_{in}} & 0 & \mathbf{S}_{N_{in}/\Gamma D_{in}}^{N_{in}} & 0 \\ 0 & 0 & 0 & 0 \\ \mathbf{S}_{D_{in}N_{in}/\Gamma}^{N_{in}} & 0 & \mathbf{S}_{N_{in}/\Gamma N_{in}/\Gamma}^{N_{in}} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.56}$$

with

$$\mathbf{S}_{X,Y}^{N_{in}} = \left\{ \int_{\Gamma_{i,j}^{in} \in \Gamma_i^{in}} \epsilon \frac{\partial \phi_s}{\partial n} \phi_r dx \right\}_{s \in X, r \in Y} \tag{6.57}$$

Clearly, $\mathbf{S}_i^{N_{in}}(t)$ absorbs (6.42). (6.46) is absorbed by $\mathbf{S}_i^{D_{out}}(t)$ given by

$$\mathbf{S}_i^{D_{out}}(t) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\mathbf{S}_{D_{out},D_{out}}^{D_{out}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.58}$$

where

$$\mathbf{S}_{D_{out},D_{out}}^{D_{out}} = \left\{ \int_{\Gamma_{i,j}^{out} \in \Gamma_i^{out}} (\mu_1 n_1 + \mu_2 n_2) \phi_s \phi_r dx \right\}_{s \in D_{out}, r \in D_{out}} \tag{6.59}$$

The local source vector absorbs the integrals (6.45) and (6.44) (the latter equals 0 as the global problem has homogeneous Dirichlet boundary condition). It is defined by

$$\tilde{\mathbf{f}}_i(t; \mathbf{u}_j^n) = \left[\mathbf{S}_{D_{in},D_{in}}^{D_{in}}(t) \mathbf{u}_j^{n,D_{out}}, 0_{D_{out}}, 0_{N_{in}/\Gamma}, 0_I \right]^T = \mathbf{S}_i^{D_{in}}(t) \mathbf{u}_j^n \tag{6.60}$$

where $\mathbf{S}_i^{D_{in}}$ is defined by substituting "out" by "in" in the definition of $\mathbf{S}_i^{D_{out}}$. $\mathbf{u}_j^{n,D_{out}}$ denotes the sub-vectors of $\mathbf{u}_j^n$ with components $\mathbf{u}_{js}^n$ such that $s \in D_{in}^i \cap D_{out}^j$ (here D_{in}^i denotes D_{in} of Ω_i). $\mathbf{u}_i^{D_{out}}$ is defined analogously.

The block-structure of the stiffness matrix suggests the following vector splittings

$$\mathbf{u}_i^{n+1} = \left[\mathbf{u}_i^{D_{in}}, \mathbf{u}_i^{D_{out}}, \mathbf{u}_i^{N_{in}/\Gamma}, \mathbf{u}_i^{I}\right]^T, \qquad \mathbf{u}_i^{N_{in}} = \left[\mathbf{u}_i^{D_{in}}, \mathbf{u}_i^{N_{in}/\Gamma}\right]^T \tag{6.61}$$

Discretisation of the local observation equation (6.15) and local error ellipsoids (6.20)-(6.21) is analogous to their global counterparts (see Section 3.2.3 and Section 4.1.3). Introducing the following notations

$$\widetilde{\mathbf{H}}_i = \{h_i(x_s - z_r)\}_{s,r=1}^{N_{nd}^i}, \qquad \mathbf{H}_i = \widetilde{\mathbf{H}}_i \mathbf{M}_i \tag{6.62}$$

$$\widetilde{\mathbf{Q}}_{0,i}(t) = diag\{q_{0,i}(x_s)\}_{s=1}^{N_{nd}^i}, \qquad \mathbf{Q}_{0,i} = \mathbf{M}_i \widetilde{\mathbf{Q}}_{0,i} \tag{6.63}$$

$$\widetilde{\mathbf{Q}}_i(t) = diag\{q_i(x_s)\}_{s=1}^{N_{nd}^i}, \qquad \mathbf{Q}_i = \mathbf{M}_i \widetilde{\mathbf{Q}}_i \tag{6.64}$$

$$\widetilde{\mathbf{R}}_i(t) = diag\{r_i(x_s)\}_{s=1}^{N_{nd}^i}, \qquad \mathbf{R}_i = \mathbf{M}_i \widetilde{\mathbf{R}}_i \tag{6.65}$$

$$\mathbf{S}_i = \mathbf{M}_i^{-1} \widetilde{\mathbf{S}}_i, \qquad \mathbf{f}_i = \mathbf{M}_i^{-1} \widetilde{\mathbf{f}}_i \tag{6.66}$$

the FEM discretised local observation equation is written as

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{u}_i + \mathbf{e}_{o,i} \tag{6.67}$$

where $\mathbf{y}_i = \left(y_i(x_1, t), \dots, y_i\left(x_{N_{nd}^i}\right)\right)^T$, and the discretised local model and observation errors L^2-ellipsoids are written as

$$\mathbb{L}_{m,i}^2 = \left\{\mathbf{e}_{0,i}, \mathbf{e}_{m,i}(t) \colon \mathbf{e}_{0,i}^T \mathbf{Q}_{0,i} \mathbf{e}_{0,i} + \int_{t_0}^{t_1} \mathbf{e}_{m,i}^T(t) \mathbf{Q}_i(t) \mathbf{e}_{m,i}(t) dt \le 1\right\} \tag{6.68}$$

$$\mathbb{L}_{o,i}^2 = \left\{\mathbf{e}_{o,i}(t) \colon \int_{t_0}^{t_1} \overline{\mathbf{e}_{o,i}^T(t) \mathbf{R}_i(t) \mathbf{e}_{o,i}(t)} dt \le 1\right\} \tag{6.69}$$

Finally, the discrete local minimax estimate $\mathbf{u}_i^{a,k+1}$ of $\mathbf{u}_i^{k+1}$, the solution of the problem (6.51), (6.67)-(6.69) is found as the solution of the discrete local feedback equation

$$\begin{cases} \dfrac{d\mathbf{u}_i^{k+1}}{dt} = \mathbf{S}_i\mathbf{u}_i^{k+1}(t) + \mathbf{f}_i\left(t;\mathbf{u}_j^k\right) + \mathbf{K}_i\mathbf{H}_i^T\mathbf{R}_i\left(\mathbf{y}_i - \mathbf{H}_i\mathbf{u}_i^{k+1}\right) \\ \mathbf{u}_i^{k+1}(0) = \mathbf{u}_i^0 \end{cases} \tag{6.70}$$

where $\mathbf{K}_i$ is the solution of the discrete local Riccati equation

$$\begin{cases} \dfrac{d\mathbf{K}_i}{dt} = \mathbf{S}_i\mathbf{K}_i + \mathbf{K}_i\mathbf{S}_i^T + \mathbf{Q}_i^{-1} - \mathbf{K}_i\mathbf{H}_i^T\mathbf{R}_i\mathbf{H}_i\mathbf{K}_i \\ \mathbf{K}_i(t_0) = \mathbf{Q}_{0,i}^{-1} \end{cases} \tag{6.71}$$

The equations (6.70)-(6.71) define $(k+1,i)$-minimax filter in a discrete form and represent the FEM approximation of the $(k+1,i)$-filter. A set of N local filters (6.70)-(6.71) are referred to as the *discrete localised minimax filter*.

The feedback equation (6.70) has two "correctors": the first one $\mathbf{K}_i\mathbf{H}_i^T\mathbf{R}_i\left(\mathbf{y}_i - \mathbf{H}_i\mathbf{u}_i^{k+1}\right)$ steers the $(k+1,i)$-filter towards the observed data and requires the solution of the corresponding Riccati equation (6.71); and the second one, $\mathbf{f}_i\left(t;\mathbf{u}_j^k\right)$ enforces the continuity across the interfaces between subdomains and requires data from adjacent subproblem. The solution the system of N equations (6.70) is obtained using Schwarz iterative approach.

6.1.3 Error estimates of the localised filter

Discretisation of the integral error estimate (6.35) is straightforward and is written as

$$\begin{aligned} \sup_{(\mathbf{e}_{0,i},\mathbf{e}_{m,i})\in\mathbb{L}_{m,i}^2,\mathbf{e}_{o,i}\in\mathbb{L}_{o,i}^2} & \left(\boldsymbol{l}_i^T\mathbf{M}_i(\mathbf{u}_i^{a,k+1}(t_1) - \mathbf{u}_i(t_1))\right)^2 \\ & \leq \boldsymbol{l}_i^T\mathbf{M}_i\mathbf{K}_i(t_1)\mathbf{M}_i\boldsymbol{l}_i \end{aligned} \tag{6.72}$$

where $\boldsymbol{l}_i$ is FEM discretisation of arbitrary function $l_i(x)$. Denoting the analysis error as

$$\mathbf{e}_i = \mathbf{u}_i^{a,k+1}(t_1) - \mathbf{u}_i(t_1) \tag{6.73}$$

the integral error estimate of the local minimax filter is written

$$\boldsymbol{l}_i^T\mathbf{M}_i\mathbf{e}_i \leq (\boldsymbol{l}_i^T\mathbf{M}_i\mathbf{K}_i(t_1)\mathbf{M}_i\boldsymbol{l}_i)^{\frac{1}{2}} \tag{6.74}$$

Here (6.74) is called an integral error estimate as the right-hand side of (6.72) is an approximation of the integral

$$\int_\Omega l_i(x)(P_il_i)(t,x)dx \approx \boldsymbol{l}_i^T\mathbf{M}_i\mathbf{K}_i(t_1)\mathbf{M}_i\boldsymbol{l}_i \tag{6.75}$$

and does not depend on the FEM discretisation up to the error level of the FEM approximation.

A straightforward way to obtain a pointwise estimate is to take a vector $\boldsymbol{l}_i$ as $\boldsymbol{l}_i = \mathbf{M}_i^{-1}\boldsymbol{l}_i^s$ where $\boldsymbol{l}_i^s = (0, \dots, 1, \dots, 0)$ is a vector with only one non-zero component corresponding to the point x_s of the FEM grid. Then

$$\boldsymbol{l}_i^{s^T}\mathbf{M}_i^{-1}\mathbf{M}_i\mathbf{e}_i \leq \boldsymbol{l}_i^{s^T}\mathbf{M}_i^{-1}\mathbf{M}_i\mathbf{K}_i\mathbf{M}_i\mathbf{M}_i^{-1}\boldsymbol{l}_i^s \tag{6.76}$$

$$\boldsymbol{l}_i^{s^T}\mathbf{e}_i \leq \boldsymbol{l}_i^{s^T}\mathbf{K}_i\boldsymbol{l}_i^s \tag{6.77}$$

As a result, a pointwise estimate of s-th error component is obtained

$$\mathbf{e}_{i,s} \leq \mathbf{K}_{i,ss} \tag{6.78}$$

At the same time, from the initial condition of the Riccati equation (6.71) and the expression (6.63) it follows that

$$\mathbf{K}_i(t_0) = \widetilde{\mathbf{Q}}_{0,i}^{-1}\mathbf{M}_i^{-1} \tag{6.79}$$

The computation of components of the mass matrix $\mathbf{M}_i$ depends on the size of the finite elements. Since the size of subdomain is fixed, the increase in the number of elements (or equivalently number of FEM nodes N_{nd}^i) per subdomain reduces the size of those finite elements. That leads to the decay of values of the mass matrix components and, consequently, to the increase of the components of the inverted mass matrix $\mathbf{M}_i^{-1}$ as N_{nd}^i increases. At the same time, the diagonal matrix $\widetilde{\mathbf{Q}}_{0,i}^{-1}$ represents values of a function computed at FEM nodes and is constant with respect to N_{nd}^i due to the properties of the FEM. That said, from (6.79) it is concluded that components of the matrix $\mathbf{K}_i(t_0)$ increase as N_{nd}^i increases. According to (6.64), the free term $\mathbf{Q}_i^{-1}$ of the Riccati equation (6.71) has the same influence on $\mathbf{K}_i(t)$ and in particular on $\mathbf{K}_i(t_1)$. That mean, the estimate (6.78) is meaningless since

$$\mathbf{e}_{i,s} \leq \mathbf{K}_{i,ss} \to \infty, \quad \text{provided } N_{nd}^i \to \infty \tag{6.80}$$

To overcome this, one should use a different error estimate, namely

$$\boldsymbol{l}_i^{s^T}\mathbf{e}_i \leq \left(\boldsymbol{l}_i^{s^T}\mathbf{K}_i(t_1)\mathbf{M}_i\boldsymbol{l}_i^s\right)^{\frac{1}{2}} \tag{6.81}$$

The rationale behind this is that $\mathbf{K}_i(t_1)\mathbf{M}_i$ is a discrete version of the operator in (4.26). Since that operator represents a mapping between two regular functions its components are bounded

for any N_{nd}^{i}. Even though (6.81) is not derived here directly, as the proposed framework is optimal for the integral estimates like (6.74), the validity of (6.81) is confirmed by the numerical experiments.

6.1.4 Localised minimax filter with reinitialisation

The minimax estimate $\mathbf{u}_i^{a,k+1}$ computed from (6.70)-(6.71) depends on the size of the time interval T and the size of the subdomain $A(\Omega_i)$ that, according to (6.23), are included in a constant $\gamma_{T,i}$. Through (6.63)-(6.65), $\gamma_{T,i}$ is used in the computations of the ellipsoid matrices $\mathbf{Q}_{0,i}$, $\mathbf{Q}_i$, and $\mathbf{R}_i$ and reflects the quality of the approximation of L^∞-ellipsoids $\mathcal{L}_{m,i}^\infty$ and $\mathcal{L}_{o,i}^\infty$ by the L^2-ellipsoids $\mathcal{L}_{m,i}^2$ and $\mathcal{L}_{o,i}^2$. To examine the influence of $\gamma_{T,i}$ assume the uniform rescaling of the discrete error ellipsoids $\mathbb{L}_{m,i}^2$ and $\mathbb{L}_{o,i}^2$ by a positive constant α. The local minimax filter is then written as

$$\begin{cases} \dfrac{d\mathbf{u}_i^{k+1}}{dt} = \mathbf{S}_i\mathbf{u}_i^{k+1}(t) + \mathbf{f}_i\left(t;\mathbf{u}_j^k\right) + \alpha\mathbf{K}_i\mathbf{H}_i^T\mathbf{R}_i\left(\mathbf{y}_i - \mathbf{H}_i\mathbf{u}_i^{k+1}\right) \\ \mathbf{u}_i^{k+1}(0) = \mathbf{u}_i^0 \end{cases} \tag{6.82}$$

$$\begin{cases} \dfrac{d\mathbf{K}_i}{dt} = \mathbf{S}_i\mathbf{K}_i + \mathbf{K}_i\mathbf{S}_i^T + \dfrac{1}{\alpha}\mathbf{Q}_i^{-1} - \alpha\mathbf{K}_i\mathbf{H}_i^T\mathbf{R}_i\mathbf{H}_i\mathbf{K}_i \\ \mathbf{K}_i(t_0) = \dfrac{1}{\alpha}\mathbf{Q}_{0,i}^{-1} \end{cases} \tag{6.83}$$

Introducing substitution $\alpha\mathbf{K}_i = \widetilde{\mathbf{K}}_i$, (6.82)-(6.83) becomes

$$\begin{cases} \dfrac{d\mathbf{u}_i^{k+1}}{dt} = \mathbf{S}_i\mathbf{u}_i^{k+1}(t) + \mathbf{f}_i\left(t;\mathbf{u}_j^k\right) + \widetilde{\mathbf{K}}_i\mathbf{H}_i^T\mathbf{R}_i\left(\mathbf{y}_i - \mathbf{H}_i\mathbf{u}_i^{k+1}\right) \\ \mathbf{u}_i^{k+1}(0) = \mathbf{u}_i^0 \end{cases} \tag{6.84}$$

$$\begin{cases} \dfrac{d\widetilde{\mathbf{K}}_i}{dt} = \mathbf{S}_i\widetilde{\mathbf{K}}_i + \widetilde{\mathbf{K}}_i\mathbf{S}_i^T + \mathbf{Q}_i^{-1} - \widetilde{\mathbf{K}}_i\mathbf{H}_i^T\mathbf{R}_i\mathbf{H}_i\widetilde{\mathbf{K}}_i \\ \widetilde{\mathbf{K}}_i(t_0) = \mathbf{Q}_{0,i}^{-1} \end{cases} \tag{6.85}$$

(6.84)-(6.85) demonstrates that the estimate $\mathbf{u}_i^{a,k+1}$ is invariant to the suggested rescaling, but that the error estimate (6.74) is not. For large values of $\gamma_{T,i}$ which is possible when $T \gg 1$ or/and $A(\Omega_i) \gg 1$ the worst-case estimation error may be quite large. As it follows from (6.74), the larger Riccati matrix $\mathbf{K}_i$ corresponds to larger estimation error; on the other hand, small $\gamma_{T,i}$ neutralize the impact of the source term in the Riccati equation (6.71) and amplifies the contribution of the quadratic term. Hence, it is desired to keep the factor $\gamma_{T,i}$ small.

It should be stressed that for a global minimax filter, the constant γ_T depends on the size of the problem domain Ω which is fixed; however, for a localised minimax filter, due to decomposition of Ω, $\gamma_{T,i}$ can be decreased by taking small local subdomains Ω_i. In addition, thanks to the Markovian property of $\mathbf{u}_i^{a,k+1}$, the size of the estimation horizon T can be taken as small as required. To this end, a procedure of time decomposition referred to as the reinitialisation procedure may be applied as follows:

1) For the time interval $[t_0, t_1]$, chose any $T = \varepsilon > 0$, $t_0^* = t_0$ and $t_1^* = t_0 + \varepsilon$.
2) Compute $\mathbf{u}_i^{a,k+1}$ over (t_0^*, t_1^*) by using (6.70)-(6.71) where $\gamma_{T,i} = (\varepsilon + 1)A(\Omega_i)$.
3) Reinitialise the filter by taking $t_0^* = t_1^*$, $t_1^* = t_0 + \varepsilon$ and compute $\mathbf{u}_i^{a,k+1}$ over the new interval (t_0^*, t_1^*) using (6.70)-(6.71) starting the Riccati equation from $(1 + \varepsilon)\mathbf{K}_i(t_0^*)$ where $\mathbf{K}_i(t_0^*)$ is the endpoint Riccati matrix from the previous time interval.
4) Repeat step 3) until t_1^* reaches t_1.

A constant $\gamma_{T,i}$ that represents the quality of the approximation of $\mathcal{L}_{m,i}^{\infty}$ and $\mathcal{L}_{o,i}^{\infty}$ ellipsoids by $\mathcal{L}_{m,i}^{2}$ and $\mathcal{L}_{o,i}^{2}$ ellipsoids, even though does not influence the estimate, is crucial for the estimation error that may explode for a large time window or large problem domain. It turns out that the proposed reinitialisation procedure together with the problem decomposition facilitates the mitigation of the error to approximate $\mathcal{L}_{m,i}^{\infty}$ and $\mathcal{L}_{o,i}^{\infty}$ by $\mathcal{L}_{m,i}^{2}$ and $\mathcal{L}_{o,i}^{2}$ and significantly reduce the estimation error of the localised minimax filter.

6.1.5 Algorithm for the localised minimax filter with pseudo-observations

The interconnections between the discrete local filters $\mathbf{u}_i^{a,k+1}$ are implemented by means of the source terms $\mathbf{f}_i\left(t; \mathbf{u}_j^{a,k}\right)$. As a result, the information from the interface (1D set in this case) is spread around in the domain and affects the nodes of the local estimate $\mathbf{u}_i^{a,k+1}$ which are not necessarily close to the mentioned interface. This, in turn, allows for pushing the information brought by observations $\mathbf{y}_i$ on the domain Ω_i to the internal FEM nodes of the adjacent domains.

```
Require:
  T                    // number of time steps
  globalProblem        // description of the global physical problem
  errorLevel           // acceptable level of the Schwarz iteration error
  GetInterfaceError()  // computes the difference between estimates on the interface
                       node obtained from adjacent subdomains
subProblems = DecomposeProblem(globalProblem)
for t = 1 to T do
  actProblems = GetActiveSubproblem(subProblems)
  for subproblem in actProblems do
    DiscretiseSubproblemByFem(subproblem, t)          // as in (6.51)
    UpdateBoundaryData(subproblem, actProblems, t)    // using (6.60)
    If HasObservations(subproblem) then
      InitObservations(subproblem, t)
    else
      InitPseudoObservations(subproblem, t)
    endif
    ReinitRiccatiEquation(subproblem, t)
    SolveRiccatiEquation(subproblem, t)               // as in (6.71)
    SolveFeedbackEquation(subproblem, t)              // as in (6.70)
  endfor

  error = GetInterfaceError(subproblem, t)
  while error > errorLevel do
    for subproblem in actProblems do
      UpdateBoundaryData(subproblem, actProblems, t)  // using (6.60)
      SolveFeedbackEquation(subproblem, t)            // as in (6.70)
    endfor
    error = GetInterfaceError(subproblems, t)
  endwhile
endfor
```

Figure 6.1 Algorithm for the localised minimax filter method.

On the other hand, the impact of observations on a local estimate depends on the structure of the local observation matrix $\mathbf{H}_i$. If the global observations $\mathbf{y}(t,x)$ are localised at a specific region of the global domain Ω (e.g., an observation kernel function $h(t,x)$ has compact support within a subdomain of Ω), it is possible that $h(t,x)$ vanishes over a number of subdomains Ω_i which results in the zero local observation matrix $\mathbf{H}_i = 0$. This, in turn, may impact the uncertainty propagation associated with the local filters. As it follows from (6.70), the so-called innovation term $\mathbf{K}_i\mathbf{H}_i^T\mathbf{R}_i\left(\mathbf{y}_i - \mathbf{H}_i\mathbf{u}_i^{k+1}\right)$ disappears if $\mathbf{H}_i = 0$, so the impact of model errors from Ω_i is, in fact, neglected as the proposed procedure cannot communicate the corresponding information to the Riccati matrices on the adjacent subdomains. In this case, the local estimation error represented by means of the discrete Riccati operator $\mathbf{K}_i$ may be underestimated.

A possible solution used in this work is to introduce "pseudo" observations. The Dirichlet data that comes from the adjacent subdomains can be treated as "pseudo" observations, and the local minimax filter is written for those observations. In this way, the impact of the model errors from adjacent domains can impact the estimate $\mathbf{u}_i^{a,k+1}$.

Finally, it should be mentioned, that the independence of the local feedback and Riccati equations allow for the activation/deactivation of the corresponding subdomains (see Section 5.2.2). The algorithm of the localised minimax filter with reinitialisation procedure, pseudo-observations and activation/deactivation is summarised by algorithm presented in Figure 6.1.

6.2 Numerical Experiments

The efficacy of the interconnected minimax filters is illustrated here with a set of numerical examples. First, time discretised representation of (6.70) is constructed using mid-point integration rule

$$\begin{aligned}\mathbf{u}_{i,n+1}^{k+1} = -\mathbf{u}_{i,n}^{k+1} &\\ &+ \left(\mathbf{I} - \frac{\Delta t}{2}\mathbf{S}_{i,n+0.5} + \frac{\Delta t}{2}\mathbf{G}\mathbf{H}_i\right)^{-1} \\ &\times \left[2\mathbf{u}_{i,n}^{k+1} + \mathbf{f}_{i,n+0.5}\left(\mathbf{u}_{j,n+0.5}^{k}\right) + \Delta t\mathbf{G}\mathbf{y}_{i,n+0.5}\right]\end{aligned} \tag{6.86}$$

$$\mathbf{u}_{i,0}^{k+1} = \mathbf{u}_{0,i}$$

where

$$\mathbf{G} = \frac{1}{2}\left(\mathbf{K}_{i,n} + \mathbf{K}_{i,n+1}\right)\mathbf{H}_i^T\mathbf{R}_{i,n+0.5} \tag{6.87}$$

The subscript n denotes the index of the points of the uniform time discretisation with the step Δt. The subscript $n + 0.5$ means that the corresponding matrix or vector is evaluated in the middle of the time interval $[t_n, t_{n+1}]$.

For the discretisation of the matrix Riccati Equation (6.71), a non-standard numerical integration technique is applied (see Section 4.1.5). First, the Möbius transformation maps the equation onto its Hamiltonian representation that is further solved by the symplectic midpoint method with reinitialisation at each time step. The discrete in time system of linear Hamiltonian equations that corresponds to (6.71) is written:

$$\begin{pmatrix}\mathbf{U}_{n+1}\\ \mathbf{V}_{n+1}\end{pmatrix} = 2\begin{pmatrix} I - \frac{\Delta t}{2}\mathbf{S}_{i,n+0.5} & \frac{\Delta t}{2}\mathbf{Q}_i^{-1} \\ \frac{\Delta t}{2}\mathbf{H}_i^T\mathbf{R}_{i,n+0.5}\mathbf{H}_i & I - \frac{\Delta t}{2}\mathbf{S}_{i,n+0.5}^T \end{pmatrix}^{-1} \begin{pmatrix}\mathbf{K}_{i,n}\\ \mathrm{I}\end{pmatrix} - \begin{pmatrix}\mathbf{K}_{i,n}\\ \mathrm{I}\end{pmatrix} \tag{6.88}$$

and the i-th local Riccati matrix is then found as

$$\begin{aligned} \mathbf{K}_{i,n+1} &= \mathbf{U}_{n+1}\mathbf{V}_{n+1}^{-1}, \quad n > 0 \\ \mathbf{K}_{i,0} &= \mathbf{Q}_{0,i}^{-1} \end{aligned} \tag{6.89}$$

To obtain the estimate of a solution of the DA problem (4.1)-(4.7) in two spatial dimensions and with the linear advection-dominated background, the fully discrete interconnected localised minimax filter (6.86)-(6.89) is iterated according to the Figure 6.1 in a set of two idealised experiments: one with a stationary flow field and another one with a non-stationary periodic flow field. In both experiments, the localised filters are compared against the ground-truth analytical solution and also compared to the global (non-localised) minimax filter, i.e., the standard minimax filter which approximates $\mathcal{L}_m^\infty$ and $\mathcal{L}_o^\infty$ by $\mathcal{L}_m^2$ and $\mathcal{L}_o^2$, and does not use DD and reinitialization. This latter comparison illustrates the following points:

1) L^2 non-decomposed filter does overestimate uncertainty which makes it problematic in practical use.
2) Interconnected localised minimax filters provide quite accurate uncertainty estimates in the considered examples
3) Significant reduction of the computational cost in the case of the localised filters.
4) Good scalability properties of the localised minimax filter.

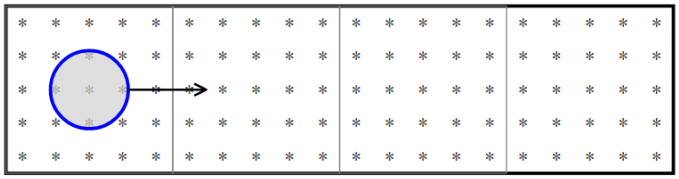

Figure 6.2 Configuration of the DD, initial conditions and observations for the Experiment 1.

6.2.1 Experiment 1

In this experiment a set of numerical simulations is performed adopting the test experiment configuration (see Section 4.2.1) for a DA problem. This problem is solved using the free-run FEM model, global minimax filter and localised minimax filter. To construct a localised filter, DD is applied by decomposing the global domain into 4 subdomains over the x-axis (see Figure 6.2) of the equal size $[0,1] \times [0,1]$ metres and discretised by 225 finite elements each. The underlying flow field is defined by the constant vector-function $\mu = [0.2;\ 0]^T\ m/s$ and the constant diffusion coefficient $\epsilon = 10^{-5}\ m^2/s$. The timestep is taken to be $0.1\ s$ and the length of the simulation is set to be 200 time steps allowing the concentration to completely transition from the right to the left of the domain.

Observations of the concentration in this experiment are produced using the Gaussian function $u_a(t,x,y)$ defined in (4.91) with the parameters (4.92)-(4.94). Since the function $u_a(t,x,y)$ satisfies the original advection-diffusion equation (4.1) with the idealised stationary flow field μ defined as above, it is considered to be a ground truth. The observations of the concentration are sampled by restricting $u_a(t,x,y)$ onto the nodes in subdomains Ω_i, $i \in I_{obs} = \{1, \dots, 4\}$ as demonstrated on Figure 6.2. The observations are corrupted by the observation noise with values uniformly distributed within the interval $[-1; 1]$.

To describe uncertainties in the system, the ellipsoids are taken as defined by the functions q_0, q and r, constant in time and space. Ellipsoid matrices of the local filter at i-th subdomain are chosen as follows:

$$\widetilde{\mathbf{Q}}_i = \frac{2}{\gamma_{T,i}}\mathbf{I}, \qquad \widetilde{\mathbf{Q}}_{0,i} = \frac{0.1}{\gamma_{T,i}}\mathbf{I}, \qquad \widetilde{\mathbf{R}}_i = \frac{3}{\gamma_{T,i}}\mathbf{I} \tag{6.90}$$

where $\gamma_{T,i} = (1 + \Delta t)A(\Omega_i) = (1 + 0.1)1 = 1.1$ reflects the quality of the approximation of L^∞-ellipsoids by L^2-ellipsoids using DD and reinitialisation and

$$\mathbf{Q}_i = \mathbf{M}_i\widetilde{\mathbf{Q}}_i, \qquad \mathbf{Q}_{0,i} = \mathbf{M}_i\widetilde{\mathbf{Q}}_{0,i}, \qquad \mathbf{R}_i = \mathbf{M}_i\widetilde{\mathbf{R}}_i \tag{6.91}$$

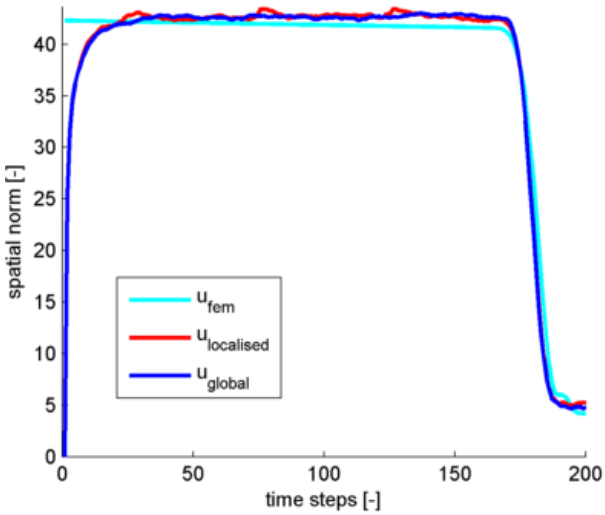

Figure 6.3 The spatial norm of the FEM solution, localised minimax estimate and global minimax estimate plotted over time.

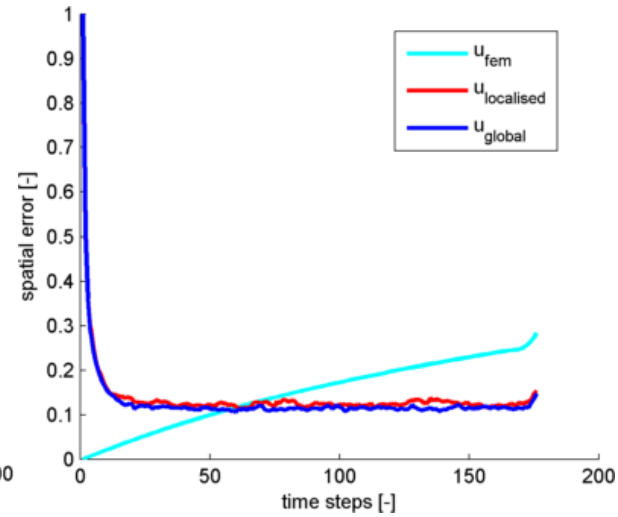

Figure 6.4 The spatial error of the FEM solution, localised minimax estimate and global minimax estimate plotted over time.

The model error matrix $\widetilde{\mathbf{Q}}_i$ is chosen to reflect the level of trust in the FEM model, in a similar fashion, $\widetilde{\mathbf{Q}}_{0,i}$ account for the absence of initial conditions. The weighting matrix $\widetilde{\mathbf{R}}_i$ represents the variance of the $[-1; 1]$-uniformly distributed random variable.

The performance of the localised estimate $\mathbf{u}_{\text{localised}}$ is compared against the estimate $\mathbf{u}_{\text{global}}$ of the so-called global filter which has been obtained by approximating the original L^∞-ellipsoids by the L^2-ellipsoids but without decomposition and reinitialisation. To compute the global filter equations (6.70)-(6.71) are used with $\Omega_i = \Omega$ and the ellipsoids' matrices $\mathbf{Q}$, $\mathbf{Q}_0$ and $\mathbf{R}$ defined as in (6.91)-(6.90), but to maintain the consistency between descriptions of the global and local model errors, and of the observation errors, the parameter $\gamma_{T,i}$ is replaced by its global equivalent

$$\gamma_T = (T+1)A(\Omega) = (T+1)NA(\Omega_{\text{i}}) = 88 \tag{6.92}$$

where factor $N = 4$ reflects the fact that $A(\Omega) = 4A(\Omega_i)$ and the size of time integration horizon is set to be $T = 20$.

In Figure 6.3 and Figure 6.4 the spatial norm and the spatial error of the localised filters are compared against the global filter estimate and non-decomposed (mono-domain) free-run FEM solution $\mathbf{u}_{\text{fem}}$ of the problem with the exact initial condition $u_a(0, x, y)$. Figure 6.3 shows that the spatial norm of the ground-truth is estimated correctly by $\mathbf{u}_{\text{fem}}$. The localised filter estimate $\mathbf{u}_{\text{localised}}$ tend to estimate the norm correctly as well, furthermore it converges to the estimate of the global filter. Figure 6.4 further shows that, as it was expected, the $\mathbf{u}_{\text{localised}}$ quickly

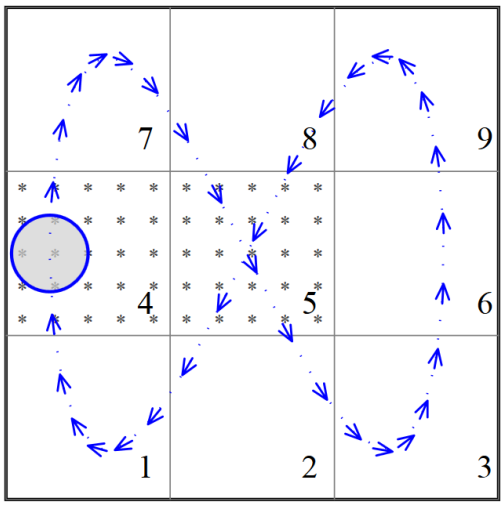

Figure 6.5 Configuration of the Experiment 2.

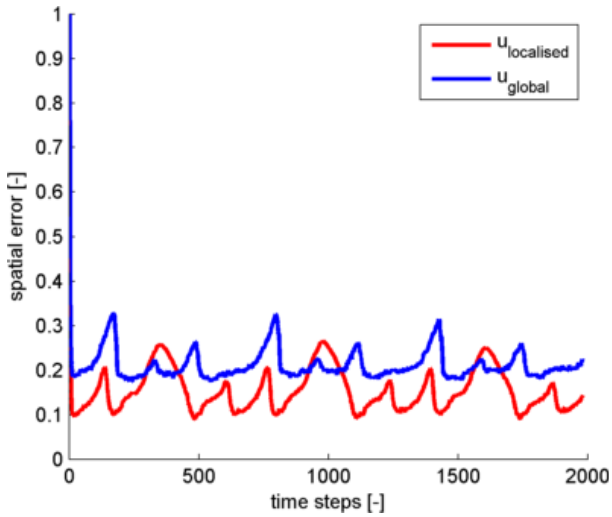

Figure 6.6 The spatial error of the localised minimax filter estimate and the global minimax filter estimate plotted over time.

reconstruct the initially unknown solution by accounting available observations and produces very similar estimate to $\mathbf{u}_{\text{global}}$. The respective estimation errors are $e_e(\mathbf{u}_{\text{fem}}) = 16.3\%$, $e_e(\mathbf{u}_{\text{localised}}) = 15.6\%$ and $e_e(\mathbf{u}_{\text{global}}) = 15\%$.

6.2.2 Experiment 2

In this experiment a two-dimensional rectangular domain $[0,3] \times [0,3]$ metres has been discretised by 2025 bilinear finite elements. DD is applied by decomposing the domain into 9 equal size subdomains $[0,1] \times [0,1]$ m each over the x and the y-axis (see Figure 6.5) and discretised by 225 finite elements. The underlying flow field μ is defined by the time dependent periodic functions:

$$\mu_x(t,x,y) = \sin(\pi - t/10) * 0.12 \tag{6.93}$$

$$\mu_y(t,x,y) = \sin(\pi/2 - t/5) * 0.24 \tag{6.94}$$

The size of the time step is taken to be $\Delta t = 0.1$ and the length of the simulation is set to be 2000 time steps allowing for the spill to make three loops following the trajectory depicted in Figure 6.5 (one loop requires 630 time steps).

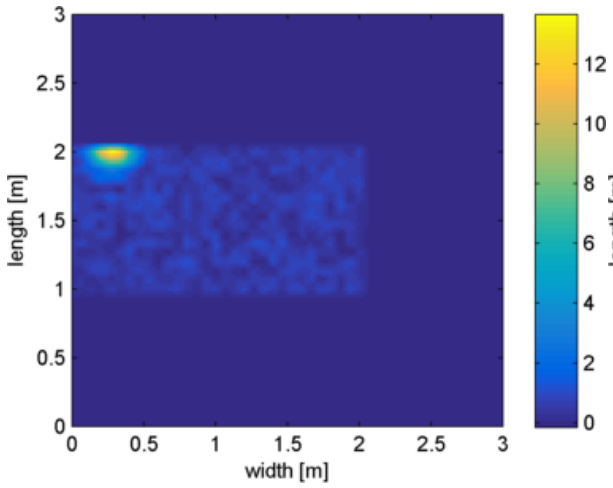

Figure 6.7 Observations at the time step 25. Relative error 84.4%.

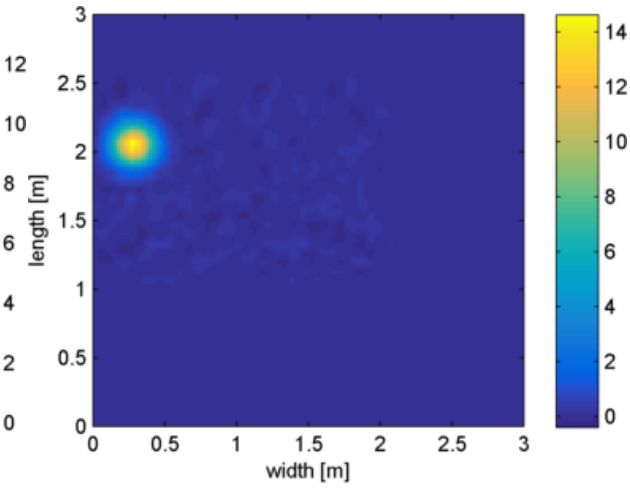

Figure 6.8 Estimate by localised filters at the time step 25. Relative error 10.8%.

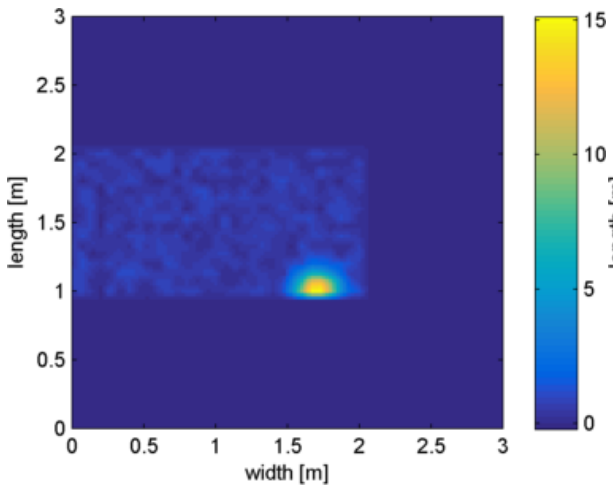

Figure 6.9 Observations at the time step 180. Relative error 68.3%.

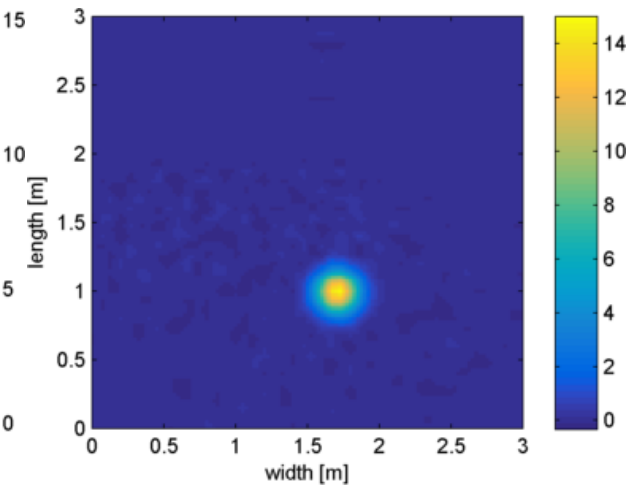

Figure 6.10 Estimate by localised filters at the time step 180. Relative error 10.3%.

As in the first experiment, the analytical solution $u_a(t,x,y)$ is defined in the form of the Gaussian function (4.91) with the following parameters:

$$\sigma = 0.1 + 0.01t, \qquad x_0 = 0.25;\ y_0 = 1.5 \tag{6.95}$$

$$m_x = t\,\mu_x = \sin\left(\pi - \frac{t}{10}\right) * 0.12\,t \tag{6.96}$$

$$m_y = t\,\mu_y = \sin\left(\frac{\pi}{2} - \frac{t}{5}\right) * 0.24\,t \tag{6.97}$$

The observations are generated by restricting the function $u_a(t,x,y)$ onto the nodes in subdomains $\Omega_i, i \in I_{obs} = \{3,4\}$. This is achieved by setting $h(z-x) = \delta(z-x)$ for $y \in \Omega_i$, and 0 for $z \notin \Omega_i$, i $\in I_{obs}$. As a result, the observation matrix $\mathbf{H}_i$ is a diagonal matrix with diagonal components equal to 1 for the observed nodes and 0 otherwise. The observation noise

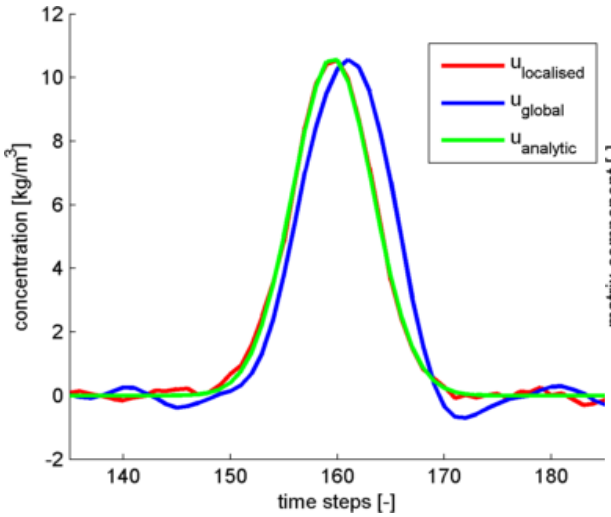

Figure 6.11 The estimate of the localised filter, global filter and analytical solution computed at the point x=1.4, y=1.4 plotted over time steps [135,185].

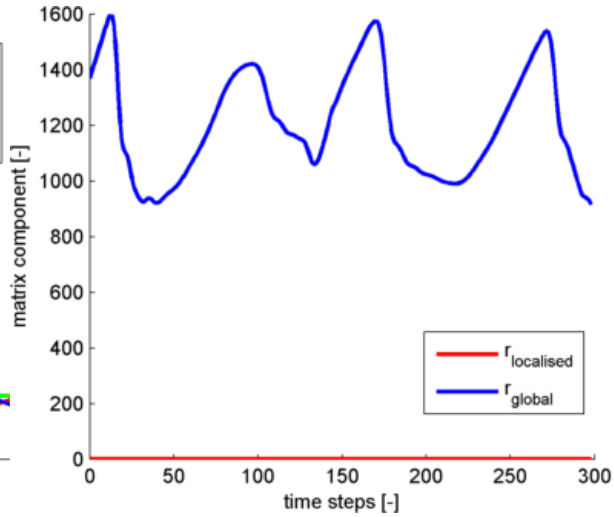

Figure 6.12 Components of the localised and global Riccati operator corresponding to the point x=1.4, y=1.4 plotted over time steps [0,300].

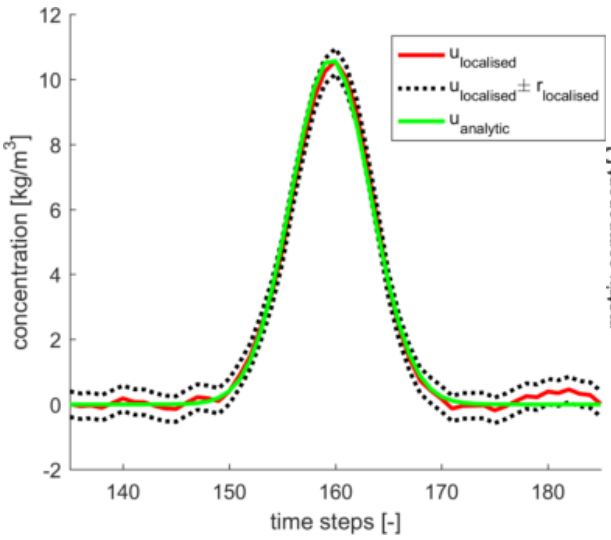

Figure 6.13 The estimate, ellipsoid of the estimate and analytical solution computed at the point x=1.4, y=1.4 plotted over time steps [135,185].

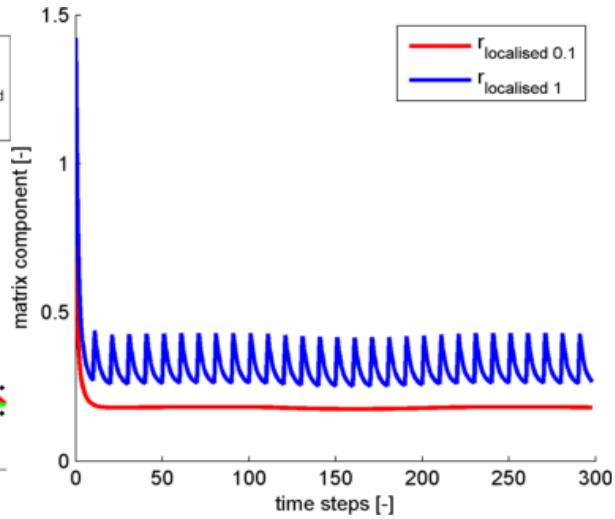

Figure 6.14 Components of the Riccati operator corresponding to the point x=1.4, y=1.4 computed by localised filter with reinitialisation intervals 0.1 and 1 plotted over time steps [0,300].

is taken to be uniformly distributed over the interval $[-0.5; 0.5]$. The sensors' locations are shown in Figure 6.5.

The parameters of the local filter at i-th subdomain are chosen as follows:

$$\widetilde{\mathbf{Q}}_i = \frac{5}{\gamma_{T,i}}, \qquad \widetilde{\mathbf{Q}}_{0,i} = \frac{1.2}{\gamma_{T,i}}, \qquad \widetilde{\mathbf{R}}_i = \frac{12}{\gamma_{T,i}} \tag{6.98}$$

and $\gamma_{T,i} = 1.1$. The choice of matrices (6.98) describes a moderate level of trust in the FEM model over the subdomain Ω_i, low confidence in the initial condition for the filter and a high trust to the observations.

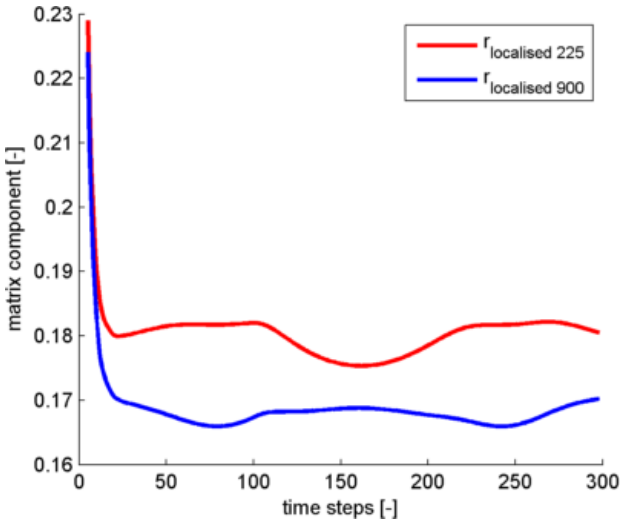

Figure 6.15 Components of the Riccati operator corresponding to the point x=1.4, y=1.4 computed by localised filter with 225 and 900 FEM elements per subdomain plotted over time steps [5,300].

Figure 6.13 shows the estimated value at the spatial point $x = 1.4$, $y = 1.4$ and demonstrates that the ground-truth is contained inside the ellipsoid. Examples of the observed fields are shown in Figure 6.7 and Figure 6.9, and the corresponding estimates generated by the localised filters are shown in Figure 6.8 and Figure 6.10. These figures show that even though the spill is not fully observed by the sensors, the local filters manage to reconstruct it with a reasonable precision level.

Similarly to Experiment 1, the localised filter estimate $\mathbf{u}_{\text{localised}}$ is compared versus the global filter estimate $\mathbf{u}_{\text{global}}$, configured using (6.98) and $\gamma_T = (T+1)A(\Omega) = 1809$. Figure 6.6 presents the spatial errors of the localised filters and the global filter. As one would expect, because of the nonstationary (in time) periodic behaviour of the underlying velocity field μ, there are intervals where the errors are decreasing when the spill is entering the observable area and increasing when spill is leaving the observable area. At the same time, it is concluded, that in general both errors are not increasing over time and obey periodic behaviour. The respective estimation errors are in favour of the localised filter: $e_e(\mathbf{u}_{\text{localised}}) = 16\%$ and $e_e(\mathbf{u}_{\text{global}}) = 19\%$.

Figure 6.11-Figure 6.12 also suggests that even though the estimates are close to each other, the global filter overestimates the uncertainty in the system. Indeed, the uncertainty overestimation is demonstrated in Figure 6.12 where diagonal components of the Riccati operator $\mathbf{K}_i\mathbf{M}_i$ and $\mathbf{KM}$ are plotted. Those components are computed at the spatial point $x = 1.4, y = 1.4$ and represent the uncertainty estimate provided by each of the filters via (6.81). It can be seen that the localised filter's ellipsoid is much tighter than that of the global filter.

A comparison analysis of the impact of the reinitialisation procedure on the estimation error is in Figure 6.14. The components of $\mathbf{K}_i\mathbf{M}_i$ are obtained from the localised filter with the reinitialisation interval ε equal to the time step of numerical integration $\Delta t = 0.1$ are compared against the same components of $\mathbf{K}_i\mathbf{M}_i$, corresponding to the reinitialization interval of length $\varepsilon = 1$. It is shown that the decrease of the reinitialisation interval leads to the decrease of the Riccati matrix components which, in turn, reduces the estimation error.

Finally, components of $\mathbf{K}_i\mathbf{M}_i$ corresponding to the point $x = 1.4, y = 1.4$ computed with different FEM resolutions (225 elements and 900 elements per subdomain) are depicted in Figure 6.15. As expected, the increase of FEM degrees of freedom, does not increase the components of $\mathbf{K}_i\mathbf{M}_i$ and the corresponding pointwise estimation error.

6.2.3 Computational Performance

The computational performance of the global minimax filter for a system with N_{nd} degrees of freedom is given by

$$c_g(N_{nd}) = 14\mathcal{C}_{N_{nd}N_{nd}} \tag{6.99}$$

where a unit of complexity $\mathcal{C}_{N_{nd}N_{nd}}$ represents the amount of operations required for dense matrix multiplication or inversion and is given as $\mathcal{O}(N_{nd}^3)$ (see Section 4.3.2). Assume that the global domain is decomposed into N subdomains, each of them containing N_{nd}^i FEM nodes. Using the rough approximation of N_{nd} by the $N \cdot N_{nd}^i$ a unit of complexity $\mathcal{C}_{N_{nd}N_{nd}}$ can be expressed in terms of N_{nd}^i as

$$\mathcal{C}_{N_{nd}N_{nd}} = N^3\mathcal{C}_{N_{nd}^iN_{nd}^i} \tag{6.100}$$

so that

$$c_g(N_{nd}) = 14N^3\mathcal{C}_{N_{nd}^iN_{nd}^i} \tag{6.101}$$

Complexity of i-th local minimax filter is $14\mathcal{C}_{N_{nd}^iN_{nd}^i}$. However, to estimate the computational performance of the localised minimax filter, Schwarz iterations should be taken into account.

Assume p Schwarz iterations were performed, which implies that the equation (6.86) was computed p times. Since the velocity flow depends on time but does not depend on a Schwarz

iteration, the inflow/outflow zones of the subdomain boundaries are fixed for all iterations. It means that in (6.86) only the source vector $\mathbf{f}_{i,k+0.5}\left(\mathbf{u}_{j,k+0.5}^{n}\right)$ is updated during each iteration, while much more expensive computations of the local model matrix $\left(\mathbf{I}-\frac{\Delta t}{2}\mathbf{S}_{i,n+0.5}+\frac{\Delta t}{2}\mathbf{GH}_i\right)^{-1}$ should be performed only once at each time step for each subproblem. Therefore, only $2N\mathcal{C}_{N_{nd}^i N_{nd}^i}$ operations are required for the equation (6.86)-(6.87), and $12N\mathcal{C}_{N_{nd}^i N_{nd}^i}$ for the equation (6.88)-(6.89). The total amount of arithmetic operations for the algorithm of the localised minimax filter for one time step is estimated as

$$c_l(N) = 14N\mathcal{C}_{N_{nd}^i N_{nd}^i} \tag{6.102}$$

From the estimates (6.101) and (6.102) it is concluded that the localised minimax filter provides significant complexity reduction comparing to the traditional minimax filter.

A further way of detailing the estimation of the total number of operations (6.102) is to count the number of nodes on the interface as $N_{nd}^{\Gamma} = kN$, where k is the average number of the interface nodes per subdomain and N_{nd}^{Γ} is the total number of the interface nodes. Since NN_{nd}^i includes interface nodes twice (from adjacent subdomains) it is equal to $N_{nd} + N_{nd}^{\Gamma}$, and N_{nd}^i can be expressed as

$$N_{nd}^i = \frac{N_{nd} + N_{nd}^{\Gamma}}{N} = \frac{N_{nd} + kN}{N} \tag{6.103}$$

Using (6.103), (6.102) is updated to

$$c_l(N) = \frac{14}{N^2}\mathcal{C}_{N_{nd}+kN, N_{nd}+kN}\mathcal{O}((N_{nd} + kN)^3) \tag{6.104}$$

If only one domain is used for the localised filter, i.e., $N = 1$ and $k = 0$, (6.104) implies that $c_l(N)$ becomes equal to c_g. On the other hand, when $N \to \infty$, then $c_l(N) \to \infty$. Using (6.104) and taking $c_l'(N) = 0$

$$\frac{14}{N^4}(3k(N_{nd} + kN)^2N^2 - (N_{nd} + kN)^3 2N) = 0 \tag{6.105}$$

$$3kN - 2(N_{nd} + kN) = 0 \tag{6.106}$$

an optimal choice of N is found as

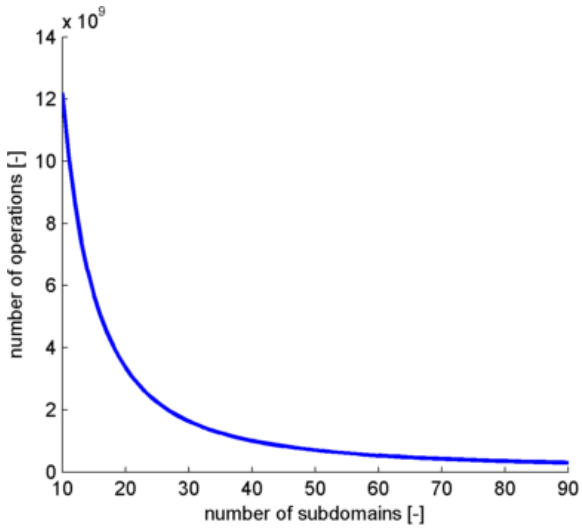

Figure 6.16 Theoretical complexity of the localised filter plotted over a number of fixed-size subdomains.

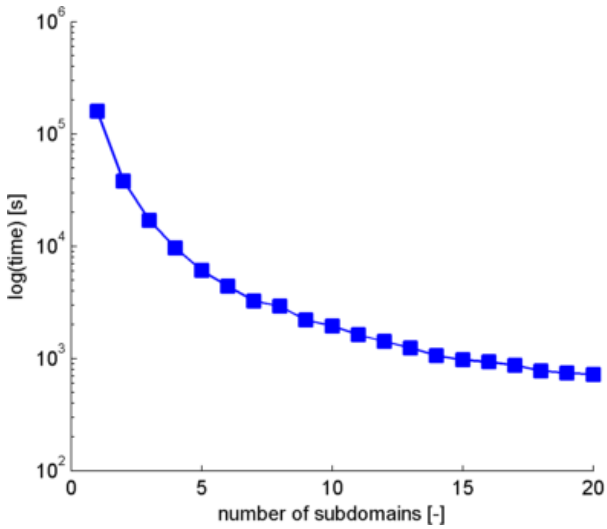

Figure 6.17 Logarithm of computational CPU time of the localised filter measured for the problem decomposed with a varying number of equal-size subdomains.

$$N = \frac{2N_{nd}}{k} \tag{6.107}$$

From Figure 6.16 where the graph of (6.104) is plotted, it can be inferred that an optimal N is achieved at the expense of having relatively small subdomains (in terms of number of FEM elements contained). That is further proved by the expression for the optimal N given in (6.107). This implies that, for a fixed number of FEM elements, the localised minimax filtering strategy has an immediate impact in diminishing the computational costs till the size of the subdomains becomes relatively small. When that happens is that the effect of DD levels off and, if the number of subdomain goes beyond the certain threshold, the computational performance does not improve further.

This is demonstrated for an experiment with a global domain of the size $[0,1] \times [0,20]$ metres and discretised by 4500 bilinear finite elements. Velocity flow, initial conditions and uncertainty ellipsoids in this experiment are configured identically as in the Experiment 1. Figure 6.17, plots computational time of localised minimax filters obtained by applying DD with various number of subdomain from the interval [1,20]. As expected, the smallest computational time achieved in the neighbourhood of $N = 19$ subdomains, without any further significant benefit when N increases.

Another important benefit of the localised strategy with DD is that it is nearly linearly scalable with regards to the number of finite elements. The increase of finite elements in a global filter leads to the cubical increase of the complexity as suggested by (6.99). While in the case of the

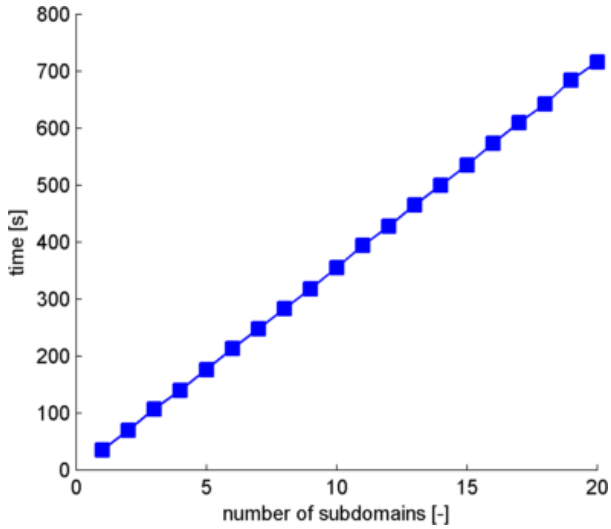

Figure 6.18 Computational CPU time of the localised filter measured for the problems with a different number of fixed-size subdomains.

localised filter, additional finite elements can be arranged into additional subdomains the size of which is fixed. As the complexity estimation (6.102) states, that would produce a linear increase of the computational costs. This is indeed shown in Figure 6.18, where the computational CPU time for problems with a different amount of finite elements and a fixed number of finite elements per subdomains is plotted over time.

6.3 Discussion and conclusions

In this chapter, the problem of efficient coupling of the contaminant transport model in the advection-dominated flows with measurements is investigated. The traditional minimax filter, despite its benefits, has several significant disadvantages for a practical usage. First, it is an L^2-type filter. It has been designed for L^2-type model errors and measurement noises, and, hence, it does not apply directly to the case of L^∞-noises that are more desired in practice. A straightforward way to apply the traditional minimax filter in this case is to approximate a L^∞-ellipsoid by a L^2-ellipsoid; this is very much like approximating a rectangle by the minimal ellipsoid which contains it. This approximation is quite crude, especially if the size of the computational domain or/and the estimation horizon are large. Second, the minimax filter is very demanding computationally. The need of the effective solvers for the Riccati equations requires approximately 12 times more arithmetic operations for the minimax filter then for a background model solver in a general case of Galerkin approximation and even more if FEM approximations are used. Moreover, minimax filter is not scalable, and it costs escalates if the

amount of degrees of freedom increases. This problem is especially important for large scale systems like those that occur in marine modelling.

To resolve the aforementioned issues, a new DA algorithm in the form of the localised interconnected minimax filters is proposed. Specifically, the computational domain of the problem is geometrically decomposed into smaller subdomains. Then the underlying PDE and observation equation are restricted to the introduced subdomains by means of a suitable DD technique to maintain the continuity of the solutions across the interfaces between the subdomains.

The ellipsoidal approximation issue is resolved as a consequence of the spatial and temporal decomposition. The spatial decomposition results in a splitting of a L^{∞}-ellipsoid (large rectangle) into a number of small rectangles and then each small rectangle is approximated by a small ellipsoid (L^2-ellipsoid) that contains corresponding rectangle. The error of approximating L^{∞}-ellipsoid by L^2-ellipsoid on a small sub-domain can be made quite small so the union of small L^2-ellipsoids which contains a large rectangle is smaller than a large L^2-ellipsoid. The temporal decomposition, known as the reinitialisation procedure splits filtering time interval into subintervals of much smaller size and reinitialise error propagation for each subinterval. In fact, the spatial and temporal decomposition significantly mitigate an ellipsoidal approximation error.

Another consequence of the spatial decomposition is a reduction of the computational costs of the localised filter, since computing N estimates on small subdomains is cheaper than computing estimate on a large domain as required by the traditional minimax filter.

The efficiency of the approach is demonstrated on a set of numerical examples. These experiments are characterised by idealised simulations of concentration being transported either by a constant flow field or a non-stationary periodic flow field. The benchmark for estimation is given by a correspondent known analytical solution and traditional minimax filter estimate. Even though the localised filter is not equivalent to the global filter, the results of the numerical simulations demonstrate good convergence between the localised and global filters.

Finally, the proposed algorithm of the interconnected localised minimax filter is Markovian (the current estimate depends on the previous one and the current observation), computationally efficient and scalable. It delivers both integral and pointwise estimates which converge to the corresponding continuous quantities over each local subdomain.

7 Interconnected localisation of Kalman filters

Chapter 5 and Chapter 6 presented two different DA approaches based on combining minimax filtering and DD. In both cases, the original continuous DA problem is transformed into a localised DA problem, which is resolved using different methods resulting in two distinct localised filters that are based on the minimax framework. To distinguish between the filters, the filter that is devised in Chapter 5 is called DAE minimax filter, and the filter constructed in Chapter 6 is called the interconnected minimax filter.

The DAE minimax filter is constructed on a discrete level using results from LQ control theory and a deterministic nature of DA problems defined in terms of minimax framework. This makes an extension of the minimax DAE filter for Kalman filters problematic. The interconnected minimax filter uses Schwarz iterative DD which is applied to the continuous problem. The minimax filter is applied only to the local problems, making the approach robust with respect to the estimator being used. Thus, the interconnected localised filtering approach is further extended in this chapter. Using the equivalence between the minimax filter and the Kalman-Bucy filter (see Section 4.1.4) and then the equivalences between the Kalman-Bucy filter and Kalman filter (see Section 3.2.3), the interconnected localised Kalman filter is derived. Using the approximation of Kalman filter by ensemble filters, in particular EnTKF (see Section 3.3), the two algorithms of the localised interconnected EnTKF are derived.

7.1 Decomposed Kalman

The equivalence of the minimax filter and the Kalman-Bucy filter (see Section 4.2) implies the equivalence between the corresponding estimation problems with deterministic error represented by quadratic ellipsoids, and statistical error represented using the Gaussian distribution. Since the interconnected localised minimax filter is comprised of minimax filters defined on local subdomains, the continuous local problems with local errors defined in local ellipsoids (see Section 6.1.1) are equivalently represented by the local problems with statistically defined local errors. Applying a FEM approximation to that continuous DA problem (see Section 3.2.3) and discretising it in time (see Section 3.2.4), one obtains a localised problem discrete in time and space for the state estimate $\mathbf{u}_{i,n}^{k}$ on the k-th Schwarz iteration as follows

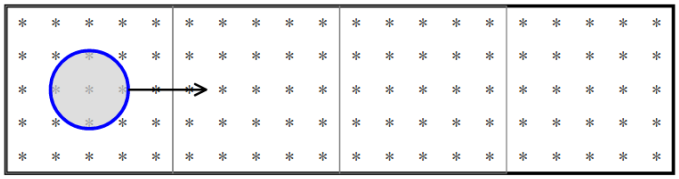

Figure 7.1 Configuration of numerical experiment with four completely observed subdomains.

$$\begin{cases} \mathbf{u}_{i,n+1}^{k+1} = \mathbf{S}_{i,n}\mathbf{u}_{i,n}^{k+1} + \mathbf{f}_{i,n}(\mathbf{u}_{j,n}^{k}) + \mathbf{e}_{i,n}^{m} \\ \mathbf{u}_{0,i}^{k+1} = \mathbf{u}_{i}^{0} + \mathbf{e}_{i}^{0} \end{cases} \tag{7.1}$$

$$\mathbf{y}_{i,n} = \mathbf{H}_i\mathbf{u}_{i,n}^{k+1} + \mathbf{e}_{i,n}^{o} \tag{7.2}$$

$$\mathbf{e}_{0,i} \sim N(0, \mathbf{Q}_i^0);\ \mathbf{e}_{m,i,n} \sim N\big(0, \mathbf{Q}_{i,n}\big);\ \mathbf{e}_{o,i} \sim N\big(0, \mathbf{R}_{i,n}\big); \tag{7.3}$$

where the subscript i is an index of the subproblem and the subscript n is the time step index. Matrix $\mathbf{S}_{i,n}$ is the model matrix, $\mathbf{H}_i$ is the observation matrix, $\mathbf{f}_{i,n}$ is the source vector, which depends on the estimate from the adjacent subproblem, and $\mathbf{u}_{0,i}$ is vector of initial conditions that are defined on i-th subproblem at time step n. Equations (7.3) reflect the fact that the local errors $\mathbf{e}_{i,n}^{o}$, $\mathbf{e}_{i,n}^{m}$ and $\mathbf{e}_{i,n}^{o}$ are normally distributed statistical variables with zero mean and corresponding covariance matrices $\mathbf{Q}_i^0$, $\mathbf{Q}_{i,n}$ and $\mathbf{R}_{i,n}$.

The local estimate $\mathbf{u}_{i,n}^{a,k}$ of $\mathbf{u}_{i,n}^{k}$ on i-th subdomain is found as a solution of the following local Kalman filter with the background update

$$\mathbf{P}_{i,n+1}^{b} = \mathbf{S}_{i,n}\mathbf{P}_{i,n}^{a}{\mathbf{S}_{i,n}}^{T} + \mathbf{Q}_{i,n} \tag{7.4}$$

$$\mathbf{u}_{i,n+1}^{b,k+1} = \mathbf{S}_{i,n}\mathbf{u}_{i,n}^{a} + \mathbf{f}_{i,n}(\mathbf{u}_{j,n}^{k}) \tag{7.5}$$

the observation update

$$\mathbf{P}_{i,n+1}^{a} = \left[\left(\mathbf{P}_{i,n+1}^{b}\right)^{-1} + \mathbf{H}_i^T\mathbf{R}_{n+1}^{-1}\mathbf{H}_i\right]^{-1} \tag{7.6}$$

$$\mathbf{u}_{i,n+1}^{a} = \mathbf{u}_{i,n+1}^{b} + \mathbf{P}_{i,n+1}^{a}\mathbf{H}_i^T\mathbf{R}_{i,n+1}^{-1}\left(\mathbf{y}_{i,n+1} - \mathbf{H}_i\mathbf{u}_{i,n+1}^{b,k+1}\right) \tag{7.7}$$

and the initial conditions

$$\mathbf{P}_{i,0}^{a} = \mathbf{Q}_i^0;\ \mathbf{u}_{i,0}^{a} = \mathbf{u}^0 \tag{7.8}$$

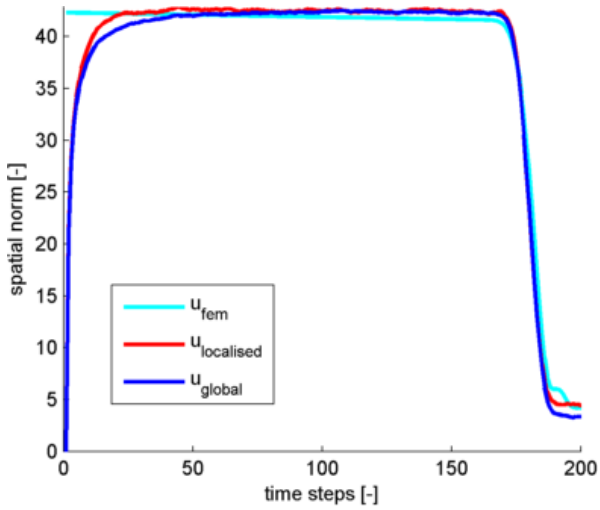

Figure 7.2 The spatial norm of the FEM solution, localised Kalman filter estimate and global Kalman filter estimate plotted over time.

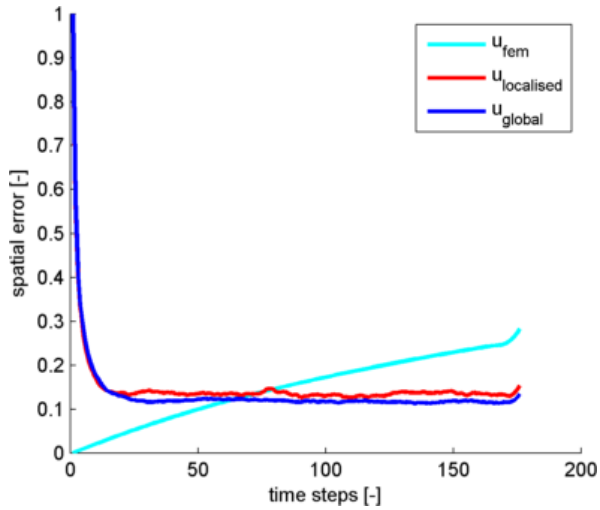

Figure 7.3 The spatial error of the FEM solution, localised Kalman filter estimate and global Kalman filter estimate plotted over time.

The set of N local Kalman filter (7.4)-(7.8) is referred to in the following as the *interconnected localised Kalman filter*. It should be stressed that in the limit as the time discretisation interval approaches zero, the localised Kalman filter is equivalent to the localised Kalman-Bucy filter.

The quality of the estimate $\mathbf{u}_{i,n}^{a}$ computed by the localised Kalman filter is examined here by performing an experiment that was designed following the test experiment configuration. The experiment adopts constant flow field $\mu = [0.2; 0]^T\ m/s$ and the global domain of the size $[0,1] \times [0,4]$ metres decomposed into four equal-size subdomains as in Figure 7.1 covered by noisy observations (for a detailed configuration of the experiment, see Section 4.3.1). The results of the simulations are depicted in Figure 7.2 and Figure 7.3 where spatial norms of the FEM solution $n_s(\mathbf{u}_{\text{fem}})$, localised Kalman filter $n_s(\mathbf{u}_{\text{localised}})$ and and global Kalman filter $n_s(\mathbf{u}_{\text{global}})$ and corresponding spatial errors $e_s(\mathbf{u}_{\text{fem}})$, $e_s(\mathbf{u}_{\text{localised}})$ and $e_s(\mathbf{u}_{\text{global}})$ are compared. It can be seen from those images that the localised Kalman filter provides a very similar quality of the estimate to the traditional Kalman filter. The same is concluded from the estimation errors of both filters:

$$e_e(\mathbf{u}_{\text{localised}}) = 16.5\% \text{ and } e_e(\mathbf{u}_{\text{global}}) = 15.6\% \tag{7.9}$$

The localised Kalman filter also inherits the same scalability properties as the localised minimax filter.

7.2 Localised ensemble filters

The interconnected localised ensemble filter is obtained by replacing the local Kalman estimator (7.4)-(7.8) by the local ensemble estimator, in particular by EnTKF. To this end, one needs to define $\mathbf{U}_{i,n}$ which are the restrictions of the ensemble of system states $\mathbf{U}_n$ on the i-th subdomain, and $\mathbf{U}_{0,i}$, the local initial ensemble. Time propagation of $\mathbf{U}_{i,n}$ is then split in two steps:

i) Computation of the background $\mathbf{U}_{i,n}^b$ and

ii) Computation of the analysis or estimate $\mathbf{U}_{i,n}^a$ of $\mathbf{U}_{i,n}$.

Let $\mathbf{F}_{i,n}$ denote the source ensemble that is derived from the source vector-function $\mathbf{f}_{i,n}(\mathbf{u}_{j,n}^k)$ of the background model (7.1). The local background ensemble $\mathbf{U}_{i,n}^a$ is propagated to the next time step $n+1$ using the background model as follows

$$\mathbf{U}_{i,n+1}^b = \mathbf{S}_{i,n}\mathbf{U}_{i,n}^a + \mathbf{F}_{i,n} \tag{7.10}$$

Then, the local observation ensemble $\mathbf{Y}_{i,n+1}^b$ and the local perturbation matrices $\widetilde{\mathbf{U}}_{i,n+1}$ and $\widetilde{\mathbf{Y}}_{i,n+1}$ are propagated using $\mathbf{U}_{n+1}^b$

$$\mathbf{Y}_{i,n+1}^b = \mathbf{H}_i\mathbf{U}_{i,n+1}^b \tag{7.11}$$

$$\widetilde{\mathbf{U}}_{i,n+1} = \mathbf{U}_{i,n+1}^b - \overline{\mathbf{U}}_{i,n+1}^b\mathbf{1}_{1\times K} \tag{7.12}$$

$$\widetilde{\mathbf{Y}}_{i,n+1} = \mathbf{Y}_{i,n+1}^b - \overline{\mathbf{Y}}_{i,n+1}^b\mathbf{1}_{1\times K} \tag{7.13}$$

The local covariance matrix of the local EnTKF is given by

$$\mathbf{P}_{i,n+1}^a = C_i(\mathbf{Y}_{i,n+1}) \tag{7.14}$$

where the matrix function C_i is defined by

$$C_i(\mathbf{Y}_{i,n+1}) = \left[(K-1)\mathbf{I}\frac{1}{\rho} + \widetilde{\mathbf{Y}}_{i,n+1}^T\mathbf{R}_{i,n+1}^{-1}\widetilde{\mathbf{Y}}_{i,n+1}\right]^{-1} \tag{7.15}$$

The local analysis ensemble $\mathbf{U}_{i,n+1}^a$ is given by the function A_i

$$\mathbf{U}_{i,n+1}^a = A_i(\mathbf{U}_{i,n+1}^b, \mathbf{P}_{i,n+1}^a) \tag{7.16}$$

which is defined as

$$A_i = \overline{\mathbf{U}}^b_{i,n+1}\mathbf{1}_{1\times K} + \widetilde{\mathbf{U}}_{i,n+1}\Big(\mathbf{W}_i + \mathbf{P}^a_{i,n+1}\widetilde{\mathbf{Y}}^T_{i,n+1}\mathbf{R}^{-1}_{i,n+1}\Big(\mathbf{y}^o_i - \overline{\mathbf{Y}}^b_{i,n+1}\Big)\mathbf{1}_{1\times K}\Big) \tag{7.17}$$

where

$$\mathbf{W}_i = \left[(K-1)\mathbf{P}^a_{i,n+1}\right]^{\frac{1}{2}} \tag{7.18}$$

To complete the algorithm of the localised EnTKF (7.10)-(7.18), a proper choice of $\mathbf{F}_{i,n}$ is needed.

The source ensemble $\mathbf{F}_{i,n}$ should accommodate $\mathbf{u}_{j,n}$, data from the j-th adjacent subproblem represented by the vector of discretised boundary conditions $\mathbf{f}_{i,n}(\mathbf{u}_{j,n})$. Since $\mathbf{u}_{j,n}$ is not known, a straight-forward way is to compute it as a solution of the deterministic counterpart of the background model (7.1) and then use the solution $\mathbf{u}_{j,n}$ for the propagation of each ensemble member. In this case, the algorithm of the localised EnTKF is as follows

$$\mathbf{u}^{b,k+1}_{i,n+1} = \mathbf{S}_{i,n}\mathbf{u}^a_{i,n} + \mathbf{f}_{i,n}\big(\mathbf{u}^{b,k}_{j,n+1}\big) \tag{7.19}$$

$$\mathbf{U}^b_{i,n+1} = \mathbf{S}_{i,n}\mathbf{U}^a_{i,n} + \mathbf{f}_{i,n}\big(\mathbf{u}^b_{j,n+1}\big)\mathbf{1}_{1\times K} \tag{7.20}$$

$$\mathbf{P}^a_{i,n+1} = C_i\big(\mathbf{Y}_{i,n+1}\big) \tag{7.21}$$

$$\mathbf{U}^a_{i,n+1} = A_i\big(\mathbf{U}^b_{i,n+1}, \mathbf{P}^a_{i,n+1}\big) \tag{7.22}$$

where equation (7.19), for the background state $\mathbf{u}^b_{i,n+1}$, is solved using the Schwarz iterative approach and the background ensemble is propagated using $\mathbf{u}^b_{i,n+1}$ via (7.20).

The disadvantage of the filter (7.19)-(7.22) is that the term $\mathbf{f}_{i,n}\big(\mathbf{u}^b_{j,n+1}\big)$ in (7.20) ensures continuity only between deterministic local solutions, even though, each ensemble member of $\mathbf{U}^b_{i,n+1}$ also contains noise. As a result, ensemble propagation (7.19)- (7.20) produces jumps between the adjacent local ensembles that may lead to the underestimation of the uncertainty.

The above issue can be overcome if instead of the background state $\mathbf{u}^b_{i,n+1}$, each member of the ensemble $\mathbf{U}^b_{i,n+1}$ is propagated by taking data from an adjacent ensemble. The resulting localised EnTKF is then formulated as follows

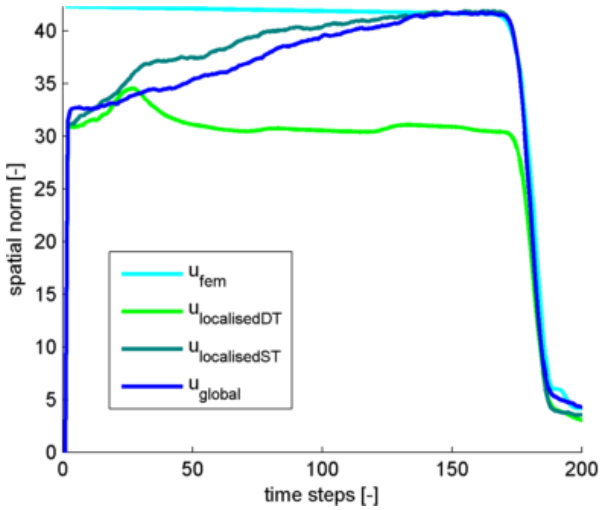

Figure 7.4 The spatial norm of the FEM solution, localised EnTKF DT and EnTKF ST estimates with 100 ensemble members and global EnKTF estimate with 500 ensemble members plotted over time.

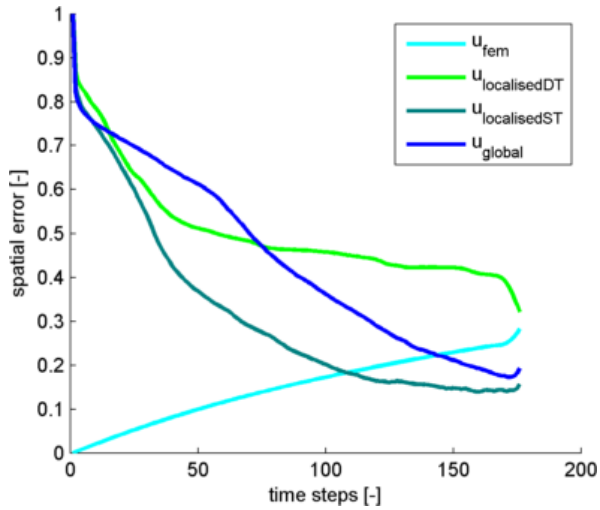

Figure 7.5 The spatial error of the FEM solution, localised EnTKF DT and EnTKF ST estimates with 100 ensemble members and global EnKTF estimate with 500 ensemble members plotted over time.

$$\mathbf{U}_{i,n+1}^{b,k+1} = \mathbf{S}_{i,n}\mathbf{U}_{i,n}^{a} + \mathbf{f}_{i,n}\left(\mathbf{U}_{j,n+1}^{b,k}\right) \tag{7.23}$$

$$\mathbf{P}_{i,n+1}^{a} = C_i\left(\mathbf{Y}_{i,n+1}\right) \tag{7.24}$$

$$\mathbf{U}_{i,n+1}^{a} = A_i\left(\mathbf{U}_{i,n+1}^{b}, \mathbf{P}_{i,n+1}^{a}\right) \tag{7.25}$$

The algorithm of the localised EnTKF (7.23)-(7.25) ensures stochastic transmission conditions between the local ensembles. However, this algorithm requires an iterative solution of the background equation (7.23) for each member of the ensemble, which makes it more computationally demanding than (7.19)-(7.22). From here, the algorithm (7.19)-(7.22) is referred to as the interconnected localised EnTKF with the deterministic transmission (DT) and the algorithm (7.23)-(7.25) is now called the interconnected localised EnTKF with the stochastic transmission (ST).

A set of numerical experiments was performed using the test experiment configuration to compare the quality of the estimates between the two algorithms of the localised EnTKF filters: EnTKF DT and EnTKF ST. The configuration of the experiments uses the constant flow field $\mu = [0.2; 0]^T$ and the global domain of size $[0,1] \times [0,4]$ metres decomposed into four equal size completely observed subdomains and initial spatial distribution of a pollutant tracer as in Figure 7.1 (for a detailed description of the test experiment configuration see Section 4.3.1).

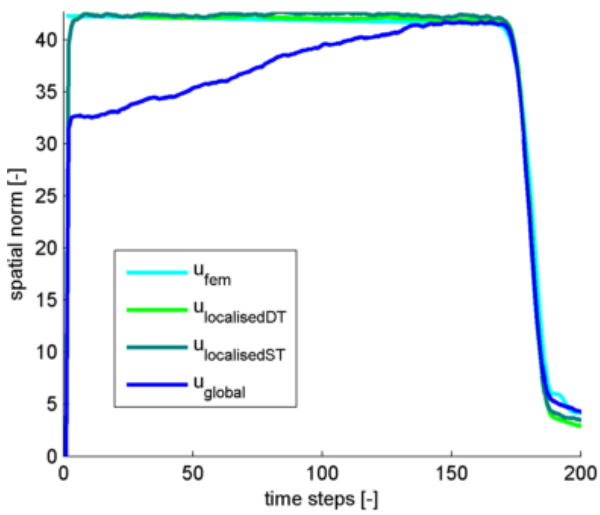

Figure 7.6 The spatial norm of the FEM solution, localised EnTKF DT, localised EnTKF ST and global EnKTF estimates with 500 ensemble members plotted over time.

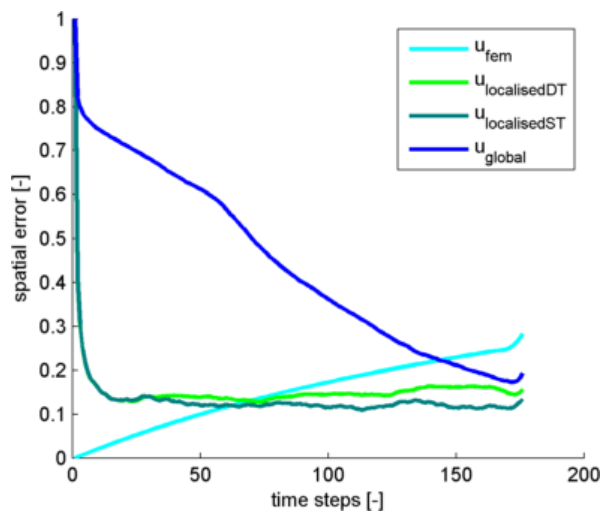

Figure 7.7 The spatial error of the FEM solution, localised EnTKF DT, localised EnTKF ST and global EnKTF estimates with 500 ensemble members plotted over time.

The results of the simulations are depicted in Figure 7.4-Figure 7.7, where the spatial norm and the spatial error of the localised EnTKF DT estimate $\mathbf{u}_{\text{localisedDT}}$ and localised EnTKF ST estimate $\mathbf{u}_{\text{localisedST}}$ are demonstrated for ensembles containing 100 and 500 members. Figure 7.4 and Figure 7.5 show that if the ensemble size is 100, the localised EnTKF ST provides better quality estimates than the localised EnTKF DT. The better estimates from the EnTKF ST is also confirmed by the corresponding estimation errors: $e_e(\mathbf{u}_{\text{localisedDT}}) = 47.8\%$ and $e_e(\mathbf{u}_{\text{localisedST}}) = 37.2\%$. Such large errors are explained by a small number of ensemble members for the given integration window. Initially both estimates are very similar, but after approximately 30 time steps, they begin to diverge. This occurs after 30 time steps, because, at this point, the concentration plume reaches the second subdomain and starts to spill into it. The spillage implies that the source term (7.20) starts to play an important role in the communication between subdomains. Since the localised EnTKF DT underestimates the uncertainty over the interface of decomposition, using EnTKF DT results in a reduction of the quality of the estimate. However, if the number of ensemble members is increasing, the effect of the uncertainty underestimation is negligible. This is illustrated in Figure 7.6 and Figure 7.7, where estimates were obtained for the ensemble with 500 members. These estimates demonstrate much better convergence and an improvements of the estimation errors $e_e(\mathbf{u}_{\text{localisedDT}}) = 16.8\%$ and $e_e(\mathbf{u}_{\text{localisedST}}) = 15.1\%$, which are in a much narrower range.

Establishing the relation between the localised and the traditional EnTKF is more difficult than in the case of the Kalman filter or minimax filter. The purpose of the ensemble filtering is to quantify the uncertainty within the space that is derived by ensemble members. If the localised

version of the EnTKF is used, a space of uncertainties is approximated in a piecewise fashion where local ensembles approximate the corresponding part of uncertainties and communicate with each other through appropriate boundary conditions, which are represented by the source vector. Instead of the global ensemble space of the size K, this approximation of uncertainties allows for the use of local ensembles that cover a smaller subdomain with a smaller amount of members, or $K_i \leq K$.

Figure 7.4 and Figure 7.5 also present a comparison between the traditional EnTKF $\mathbf{u}_{\text{global}}$ with the ensemble size 500 and the localised EnTKF $\mathbf{u}_{\text{localisedDT}}$ and $\mathbf{u}_{\text{localisedST}}$ using four local ensembles of size 125. The results demonstrate that a piecewise approximation of the system of uncertainties provides better results than a traditional approximation. The same is inferred by the estimation errors: $e_e(\mathbf{u}_{\text{localisedDT}}) = 47.8\%$, $e_e(\mathbf{u}_{\text{localisedST}}) = 37.2\%$ and $e_e(\mathbf{u}_{\text{global}}) = 48.4\%$. A further increase of the ensemble size of the localised filters to 500 members results in significantly better estimates from the localised EnTKF, as it can be seen from Figure 7.6 and Figure 7.7, than in the estimation obtained from the traditional EnTKF with an ensemble size of 500.

Finally, it should be noted that the idea of performing the analysis in smaller domains is not new. (Hunt et al., 2007) suggest to propagate the global ensemble, but to partition available observations into separate batches that have uncorrelated errors, perform a local analysis and update ensemble locally as well:

$$\mathbf{U}_{n+1}^{b} = \mathbf{S}_n \mathbf{U}_n^{a} + \mathbf{f}_n \tag{7.26}$$

$$\mathbf{P}_{i,n+1}^{a} = C_i\left(\mathbf{Y}_{i,n+1}\right) \tag{7.27}$$

$$\mathbf{U}_{i,n+1}^{a} = A_i\left(\mathbf{U}_{i,n+1}^{b}, \mathbf{P}_{i,n+1}^{a}\right) \tag{7.28}$$

This approach does not allow for the splitting of the batch of observation with correlated errors because of the absence of the transmission mechanism between local ensembles. In contrast, the localised EnTKF DT or localised EnTKF ST inherit subdomain communication mechanisms from d-ADN. Another advantage of the localisation method presented here is that the background ensemble propagation is also localised in contrast to (7.26) meaning that the algorithm can be completely parallelised.

7.3 Comparison of localised filters

To compute the computational costs of the interconnected localised Kalman and EnTKF filters, assume that in total, N_{nd} FEM nodes are split into N subdomains with N_{nd}^i nodes each, and p iterations of the Schwarz approach are needed to solve the background problem. One time step of the localised Kalman filter (7.4)-(7.8) requires $7N\mathcal{C}_{N_{nd}^i N_{nd}^i}$ arithmetic operations, where $N\mathcal{C}_{N_{nd}^i N_{nd}^i}$ operations are used for the estimate computation (7.5), (7.7) and $6N\mathcal{C}_{N_{nd}^i N_{nd}^i}$ operations are needed for the covariance analysis (7.4), (7.6).

The algorithm of the interconnected localised EnTKF DT requires $N\mathcal{C}_{N_{nd}^i N_{nd}^i}$ operations for the solution of the background problem (7.19), $N\mathcal{C}_{N_{nd}^i K}$ operations are used for (7.20), $N\left(\mathcal{C}_{N_{nd}^i K} + \mathcal{C}_{K N_{nd}^i} + 2\mathcal{C}_{KK}\right)$ operations are used for the covariance propagation (7.21) and $N\left(\mathcal{C}_{N_{nd}^i K} + 2\mathcal{C}_{K N_{nd}^i} + \mathcal{C}_{KK}\right)$ operations are used for the ensemble update (7.22). For the localised EnTKF ST, the background propagation is performed in $N\mathcal{C}_{N_{nd}^i N_{nd}^i}$ for the model matrix computation and $pN\mathcal{C}_{N_{nd}^i K}$ for matrix multiplication. The total complexity of the localised filters presented in this research is given in the following table:

Table 7.1 Computational complexity of traditional and localised filter.

	Minimax filter	Kalman filter
Global	$14\mathcal{C}_{N_{nd}N_{nd}}$	$7\mathcal{C}_{N_{nd}N_{nd}}$
Localised	$14N\mathcal{C}_{N_{nd}^i N_{nd}^i}$	$7N\mathcal{C}_{N_{nd}^i N_{nd}^i}$
	EnTKF	
Global	$\mathcal{C}_{N_{nd}N_{nd}} + 3\mathcal{C}_{N_{nd}K} + 3\mathcal{C}_{KN_{nd}} + 3\mathcal{C}_{KK}$	
Localised DT	$N\left(\mathcal{C}_{N_{nd}^i N_{nd}^i} + 3\mathcal{C}_{N_{nd}^i K} + 3\mathcal{C}_{K N_{nd}^i} + 3\mathcal{C}_{KK}\right)$	
Localised ST	$N\left(\mathcal{C}_{N_{nd}^i N_{nd}^i} + (2+p)\mathcal{C}_{N_{nd}^i K} + 3\mathcal{C}_{K N_{nd}^i} + 3\mathcal{C}_{KK}\right)$	

The theoretical results in Table 7.1 are justified by the results from numerical experiments conducted using the test experiment configuration (see Section 4.3.1) with a decomposition of the global domain into four equal size completely observed subdomains and initial spatial

distribution of a pollutant tracer as in Figure 7.1. Consolidated results of those experiments are provided in Table 7.2 and Table 7.3 where the execution time and the estimation errors are presented for the following simulations: (i) global minimax, Kalman and EnTKF filters (Section 4.2.1); (ii) localised minimax filter (Section 6.2.1); (iii) localised Kalman filter (Section 7.1), and (iv) localised ensemble filters (Section 7.2).

Table 7.2 Execution CPU time and estimation errors of the global and localised minimax and Kalman filters. Results of the time measurements are averaged over 4 runs.

	Minimax filter		Kalman filter	
	Execution time	Estimation error	Execution time	Estimation error
Global	$244\ s$	15.0 %	$157\ s$	15.6 %
Localised	$39.1\ s$	15.6 %	$18.6\ s$	16.5 %

Table 7.3 Execution CPU time and estimation errors of the global and localised EnTKF filters with 100 and 500 ensemble members. Results of the time measurements are averaged over 4 runs.

	EnTKF, $\mathrm{K} = 100$		EnTKF, $\mathrm{K} = 500$	
	Execution time	Estimation error	Execution time	Estimation error
Global	$71.1\ s$	89.7 %	$136\ s$	48.4 %
Localised DT	$11.7\ s$	47.8 %	$98\ s$	16.8 %
Localised ST	$20.8\ s$	37.2 %	$143\ s$	15.1 %

As expected, the results in Table 7.2 demonstrate that the localised filters based on d-ADN DD provide significant computational advantages comparing to the traditional global algorithms while maintaining the same quality of estimation. The results given in Table 7.3 also confirm theoretical complexities of localised EnTKF filters. These results further demonstrate that the complexity of the localised ensemble filters is in a balance between number of subdomains N and the size of an ensemble K.

The difference between ensemble filters and minimax or Kalman filters for decomposition is the method of approximation of the uncertainties. In minimax or Kalman filter, the uncertainty space represented by the Riccati matrix (or covariance matrix for Kalman filter) is strictly

defined over the degrees of freedom that are associated with the spatial domain. The geometrical decomposition of the global spatial domain produces smaller spatial subdomains, which result in smaller covariance matrices. In the case of the localised EnTKF, the ensemble of states approximates the uncertainty space. So that, the application of decomposition to the spatial domain splits the global ensemble into local fractions, but the size of each fraction remains the same. For this reason, as demonstrated in Table 7.3, the localised EnTKF DT and ST algorithms are cheaper than EnTKF for a small ensemble size since the background propagation dominates the computational price. For larger ensembles, the analysis step becomes more demanding, and the total execution time is higher for localised EnTKF than for the global EnTKF because for localised filter analysis is performed on each subdomain (see Table 7.3). However, since the spatial subdomains of the localised problems are smaller than the original domain, the local ensemble size can be decreased, as it should represent a smaller size uncertainty space. The level of the decrease is not obvious as it is not directly associated with the decomposition applied, but, for instance, the localised EnTKF DT and EnTKF ST using 100 ensemble members provide a smaller estimation error then EnTKF using 500 ensemble members (see Figure 7.4 and Figure 7.5). It should also be noted that, although the localised EnTKF ST provides a better quality of the estimate than the localised EnTKF DT, it is also more computationally expensive. Thus, it may be reasonable to use the localised EnTKF DT with a higher amount of ensemble members than the localised EnTKF ST.

Computational performance of the localised filters is further demonstrated using the following numerical experiment. Consider a global domain of the size $[0,1] \times [0, N]$ discretised into $225N$ FEM elements where N is an integer variable such that $N \in \{1,2, \dots ,20\}$. The experiment adopts the constant flow field $\mu = [0.2; 0]^T m/s$, the constant diffusion coefficient $\epsilon = 10^{-5}\ m^2/s$ and is performed for 2000 time steps of $0.1\ s$ each. The global domain is decomposed into N equal-size subdomain over the horizontal axis and discretised by 225 FEM elements. The initial conditions of the problem are not known, but each subdomain is observed by noisy observations. To solve this DA problem, the following localised filters are used: minimax localised filter, Kalman localised filter, EnKTF DT with 100 ensemble members and EnTKF ST with 100 ensemble members.

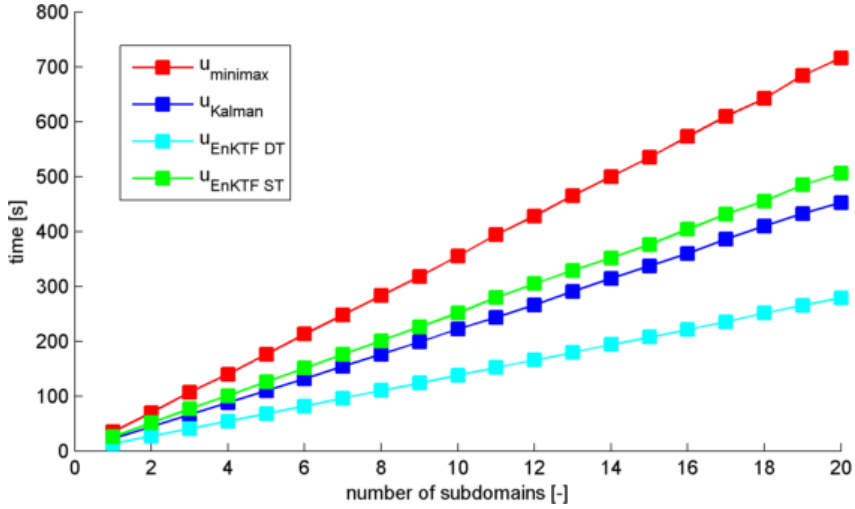

Figure 7.8 Computational time of the minimax, Kalman, EnKTF DT with 100 ensemble members and EnTKF with 100 ensemble members localised filters measured for the problems with a different number of fixed-size subdomains.

Figure 7.8 plots the execution CPU time of the corresponding filters for different values of N. The figure demonstrates that the linear increase of the number of subdomains N results in the linear increase of the execution time. This result demonstrates that as suggested by the computational complexities given in Table 7.1 the localised filters are nearly scalable with respect to an increasing number of finite elements organised in additional subdomains. Theoretical complexity estimates given in Table 7.1 are further confirmed by the slopes of lines from Figure 7.8. Those slopes represent an increase of computational time for each additional subdomain and have the following values: $s(\mathbf{u}_{\text{minimax}}) = 35.3$, $s(\mathbf{u}_{\text{kalman}}) = 21.9$, $s(\mathbf{u}_{\text{entkfDT}}) = 13.4$ and $s(\mathbf{u}_{\text{entkfST}}) = 25.6$. As expected, the minimax filter produces the steepest slope while EnTKF DT provides the flattest slope.

The interconnected localisation approach provides potential for even greater reduction of execution time. This indeed can be achieved since computations on each subdomain are independent of each other and can be organised in parallel on a supercomputer, i.e., each subdomain is assigned to a separate computational node and resolved in parallel. The near linear scalability of the algorithms suggests that parallel scalability should be achieved as well.

7.4 Comparison with consensus Kalman filter

In this section, interconnected localised filters and in particular interconnected localised Kalman filter are compared to the consensus Kalman filter that was recently proposed in Battistelli et al. (2015) and later analysed in more details in Battistelli et al. (2016). The

consensus Kalman filter also uses DD for reducing computational costs and recasts the localised problem using approach similar to Schwarz iterative approach. To be specific, it constructs "consensus" between adjacent subdomains by exchanging local state and covariance data across the decomposition interface.

It is not straightforward to apply the algorithm of the consensus Kalman filter presented in Battistelli et al., (2015) and Battistelli et al., (2016) to the contaminant transport problem, and significant modification are required. The main modifications are the following:

1) Instead of the diffusion problem operator, the advection-diffusion operator is considered. This also requires an adjustment of the FEM approximation schema.
2) To satisfy the hyperbolic nature of the contaminant transport problem, the original Robin transmission condition is replaced by the d-ADN transmission condition.
3) Boundary conditions are incorporated through boundary integrals and corresponding source terms.
4) Additionally, the mid-point rule is used for a time discretisation instead of the backward Euler rule. This is done to allow for a consistent comparison with the localised filters.

The adapted version of the consensus Kalman filter algorithm for the contaminant transport problem is defined for a time interval $[t_n, t_{n+1}]$ and a decomposition of a global domain into N subdomains in the following steps:

i) initialisation:

$$\mathbf{u}_{i,n}^{b} = \mathbf{u}_{i,n}^{a}, \qquad \mathbf{P}_{i,n}^{b} = \mathbf{P}_{i,n}^{a} \tag{7.29}$$

ii) consensus construction: split time interval $[t_n, t_{n+1}]$ into L subintervals $[t_n + l\delta, t_n + (l+1)\delta]$, where $\delta = (t_{n+1} - t_n)/L$ and $l = [0, \dots, L-1]$. For each subinterval propagate background state and covariance for each subdomain $i = 1, \dots, N$ by exchanging boundary data

$$\mathbf{u}_{i,n+(l+1)\delta}^{b} = \mathbf{S}_{i,n+l\delta}\mathbf{u}_{i,n+l\delta}^{a} + \mathbf{S}_{i,n+l\delta}^{D_{in}}\mathbf{u}_{j,n+l\delta}^{b} \tag{7.30}$$

$$\begin{aligned}\mathbf{P}_{i,n+(l+1)\delta}^{b} = {} & \mathbf{S}_{i,n+l\delta}\mathbf{P}_{i,n+l\delta}^{b}\mathbf{S}_{i,n+l\delta}{}^{T} + \mathbf{Q}_{i,n+l\delta} \\ & + \mathbf{S}_{i,n+l\delta}^{D_{in}}\mathbf{P}_{j,n+l\delta}^{b}\left(\mathbf{S}_{i,n+l\delta}^{D_{in}}\right)^{T}\end{aligned} \tag{7.31}$$

Here, the sub-index $n + l\delta$ denotes that the corresponding matrix or vector are computed at a time instance $t_n + l\delta$. While matrices $\mathbf{S}_{i,n+l\delta}$ and $\mathbf{Q}_{i,n+l\delta}$ are equivalent to those as in the

interconnected localised Kalman filter, the matrix $\mathbf{S}_{i,n+l\delta}^{D_{in}}$ represents discretisation of a boundary integral for a Dirichlet condition over the inflow zone (see Section 6.1.2). The matrix $\mathbf{P}_{j,n+l\delta}^{b}$ represents the covariance of boundary nodes taken from adjacent subdomains. It is computed in a manner similar to the vector of boundary values $\mathbf{u}_{j,n+l\delta}^{b}$.

iii) the observation update:

$$\mathbf{P}_{i,n+1}^{a} = \left[\left(\mathbf{P}_{i,n+1}^{b}\right)^{-1} + \mathbf{H}_i^T \mathbf{R}_{n+1}^{-1} \mathbf{H}_i\right]^{-1} \tag{7.32}$$

$$\mathbf{u}_{i,n+1}^{a} = \mathbf{u}_{i,n+1}^{b} + \mathbf{P}_{i,n+1}^{a} \mathbf{H}_i^T \mathbf{R}_{i,n+1}^{-1} \left(\mathbf{y}_{i,n+1} - \mathbf{H}_i \mathbf{u}_{i,n+1}^{b,k+1}\right) \tag{7.33}$$

The initial conditions of the algorithm are given as

$$\mathbf{P}_{i,0}^{a} = \mathbf{Q}_i^0;\ \mathbf{u}_{i,0}^{a} = \mathbf{u}^0 \tag{7.34}$$

Comparing the consensus and interconnected localised Kalman filter algorithms one may notice that they share an identical observation update step but differ in the background propagation step. Specifically, the two main differences are the following:

1) To coordinate local solutions across the decomposition interface, the consensus Kalman filter performs iterations in the temporal dimension by splitting each time step into a fixed number of substeps. In contrast, the localised filter performs iterations in the Schwarz dimension. It should be noted that this difference results in different convergence properties of those filters. For example, for a constant velocity field, the Schwarz approach converges in a fixed number of iterations while consensus approach converges only if the size of substep approaches zero.
2) In addition to the exchange of local state estimates between adjacent subdomains, the consensus Kalman filter also exchanges local covariances through the decomposition interface. This is done by introducing an extra term in the background covariance propagation equation (7.31) that incorporates boundary covariances obtained from adjacent subdomains.

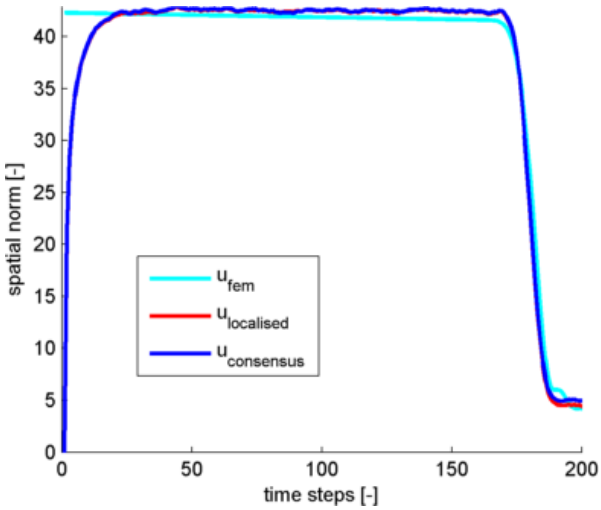

Figure 7.9 The spatial norm of the FEM solution, interconnected localised Kalman filter and consensus Kalman filter estimates plotted over time.

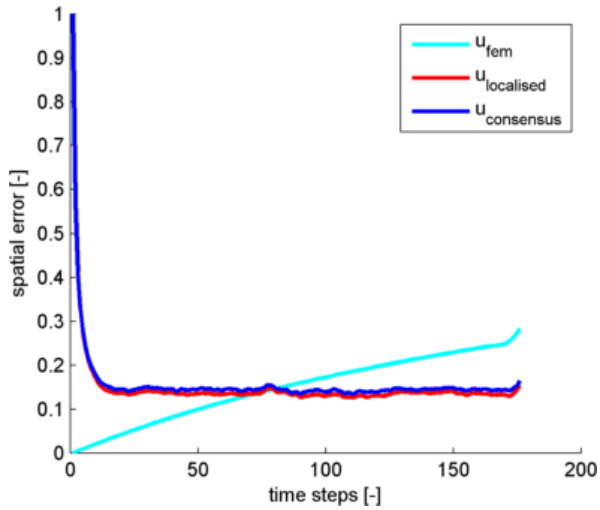

Figure 7.10 The spatial error of the FEM solution, interconnected localised Kalman filter and consensus Kalman filter estimates plotted over time.

To compare quality of estimates of the interconnected localised Kalman filter $\mathbf{u}_{\text{localised}}$ and consensus Kalman filter $\mathbf{u}_{\text{consensus}}$ a numerical simulation with the test experiment configuration is performed. For consistency, the interconnected localised Kalman filter was executed using 1 Schwarz iteration and the consensus Kalman filter was done with 1 consensus iteration. The results of this simulation are demonstrated in Figure 7.9 and Figure 7.10 where the spatial norms and spatial errors of the corresponding filters are presented. It can be seen that both estimates are very similar. This is further confirmed by the estimations errors: $e_e(\mathbf{u}_{\text{localised}}) = 16.5\%$ and $e_e(\mathbf{u}_{\text{consensus}}) = 17.3\%$. The slightly higher estimation error of the consensus filter is due to its incorporation of transmission conditions from the previous iteration, i.e., from the previous time substep. Increasing number of consensus iterations would increase number of time points where computations are performed and thus would improve quality of estimates. However, the question of convergence of the consensus filter and localised filter requires additional investigations.

To compare the computational efficiency of the filters, assume that only one iteration of both filters is performed. In this case, the consensus filter performs two extra matrix multiplications in (7.31) associated with the exchange of covariances between subdomains comparing to the interconnected localised filter. For instance, the execution time of the experiment considered here is $18.6\ s.$ for the localised filter and $19.7\ s.$ for the consensus filter. Each extra iteration of the consensus filter is done at a different time instance and requires recomputing FEM matrices, background state and covariance propagation (7.30)-(7.31). Iterations of the localised filter are done in the Schwarz dimension using the same matrices computed during the first

iteration. That means each additional iteration of the consensus filter requires an extra matrix inversion and two matrix multiplications comparing to the localised filter. For instance, execution time of the interconnected localised filter with 2 Schwarz iterations is $20.4\ s$. while execution time of the consensus Kalman filter with 2 consensus iterations is $29.2\ s$. That means that from the computational point of view, the consensus Kalman filter is more expensive then the interconnected localised Kalman filter.

The results of this comparison can be extrapolated to other localised filters using results from Table 7.2 and Table 7.3. It is concluded that the consensus Kalman filter is faster than the interconnected localised minimax filter for a small number of iterations but provides less accurate results. The accuracy of the consensus filter may be improved by increasing number of consensus iterations however that may be computationally more demanding that the interconnected localised minimax filter. Interconnected localised EnTKF are cheaper than the consensus filter for small number of ensemble members but are less accurate.

7.5 Tidal basin numerical experiment

To assess the performance of the interconnected localised filter, numerical experiments for a DA problem of a tracer transport with a realistic velocity flow field are presented in this section. For this purpose, a simulation of a tidal basin is considered. A tidal basin is a physical model designed to generate tidally induced water circulation and tracer transport that can represent physical features of a natural prototype (Olbert et al., 2013). It is comprised of two main parts: 1) the weir that is used to generate tides, and 2) the working area. An advantage of tidal basin is that it can be constructed in laboratory conditions and experiments may be performed in a fully controlled environment where domain setup and forcing are user-defined. Therefore, a tidal basin facilitates a coherent comparison between laboratory and numerical outputs. This experiment deals only with the numerical simulation of the tidal basin while laboratory simulations are left for future work.

A transport of tracer in a marine environment is typically modelled by an advection-diffusion equation with turbulent diffusion. In this experiment, DA is used to predict tracer concentration using linear advection-diffusion PDE without turbulent diffusivity as a background model and incorporating observations into that background model. To solve a DA problem, two interconnected localised filters are selected: (i) the localised minimax filter and (ii) the localised EnTKF DT filter. These choices are explained by the fact that the localised minimax filter is

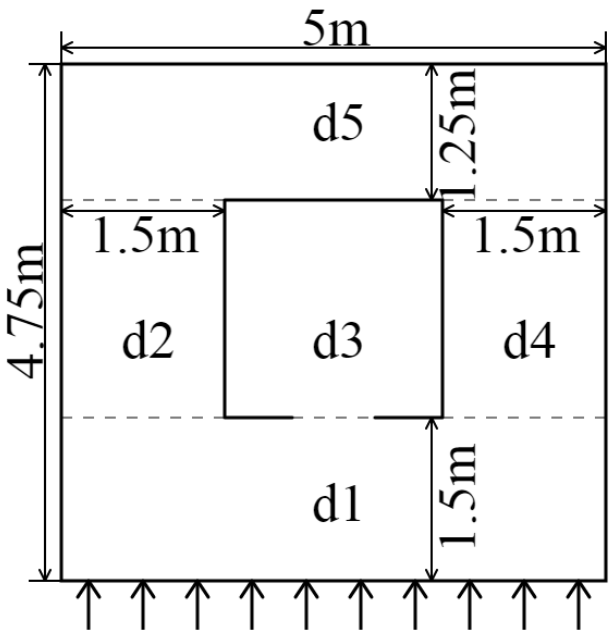

Figure 7.11 Schematic illustration and dimensions of the working area in the tidal basin.

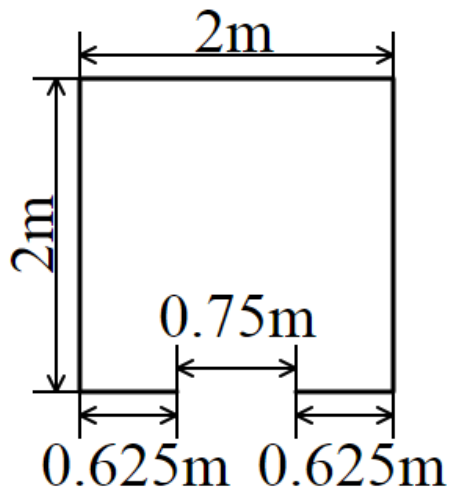

Figure 7.12 Schematic illustration and dimensions of the harbour of the tidal basin.

the most general in terms of error treatment while the localised EnTKF DT has the lowest computational requirements provided a low level of the ensemble size.

The velocity flow field and transport of concentration are predicted using two-dimensional depth-integrated numerical model TELEMAC 2D (Desombre and Lang, 2013). It is applied to a physical model of a single entrance symmetrical square harbour placed in the working area of the tidal basin. The schematics and dimensions of the working area of the tidal basin are presented in Figure 7.11, and the dimensions of the harbour are presented in Figure 7.12. The initial water elevation $h_0(x)$ in the working area is $0.3m$. The water movement is induced by the weir placed on the south edge of working area as in Figure 7.11 producing tides. In TELEMAC 2D that is represented by the forcing boundary conditions applied to the bottom boundary of the working area (see Figure 7.11) and given in the form of harmonic function:

$$h(t,x) = h_0(x) + h_A \sin(\frac{2\pi}{T}) \tag{7.35}$$

where the amplitude $h_A = 0.05m$ and the cycle $T = 2s$.

The TELEMAC solver is configured to discretise the basin by 3200 linear finite elements in space and 1000 time steps of the length $0.01s$ each. In addition, it is also assumed that a tracer is released in the harbour with the initial concentration $1\ kg/m^3$ inside the harbour and 0 otherwise. TELEMAC is set up to generate tracer concentration $\mathbf{u}_{\text{telemac}}$ using turbulent diffusion with diffusion coefficient $0.01\ m^2/s$.

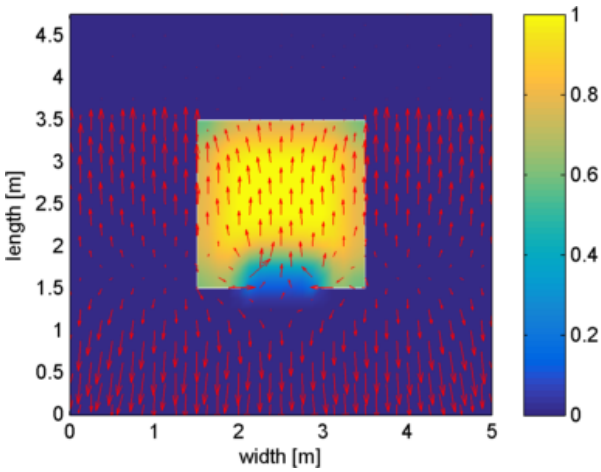

Figure 7.13 Tracer concentration predicted by TELEMAC at time step 200 when the first tide is entering the harbour.

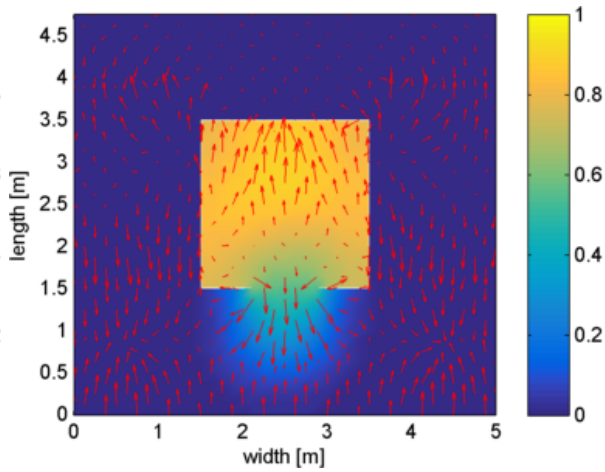

Figure 7.14 Tracer concentration predicted by TELEMAC at time step 850 when the tide is leaving the harbour.

Results of TELEMAC simulations are demonstrated in Figure 7.13 and Figure 7.14 where concentration of the tracer $\mathbf{u}_{\text{telemac}}$ is depicted. Figure 7.13 shows $\mathbf{u}_{\text{telemac}}$ at the time step 200 when the first tide reached the harbour and fresh water is entering the gate, leading to a decrease of the tracer concentration around the gate. Figure 7.14 demonstrates $\mathbf{u}_{\text{telemac}}$ at the time step 850 after the tide that was reflected from the top wall of the harbour is leaving the gate propagating tracer concentration outside.

The DA problem considered in this paragraph is to model a tracer evolution without a turbulent diffusion by using the linear advection-diffusion equation and noisy observations sampled from $\mathbf{u}_{\text{telemac}}$. The underlying flow field in this case is taken from the TELEMAC simulation and $\mathbf{u}_{\text{telemac}}$ is considered to be a ground-truth solution of this problem.

To model tracer concentration, TELEMAC uses a non-conservative advection-diffusion equation (Hervouet, 2007). To replicate the TELEMAC model, the background equation describing tracer concentration u of the DA problem is taken in the following form:

$$\frac{\partial u}{\partial t} = \epsilon \Delta u - (\nabla \cdot \mu) u \tag{7.36}$$

where $\epsilon \geq 0$ is a constant diffusion coefficient and vector $\mu = [\mu_1, \mu_2]^T$ is a vector of TELEMAC generated velocity field. Since the advection term of equation (7.36) is different from the advection term in the conservative advection-diffusion equation (3.4) its treatment requires additional explanation.

A weak representation of the advection term $(\nabla \cdot \mu) u$ is obtained using a test function v as:

$$(\mu \cdot \nabla u, v) = \int_{\Omega} \mu \cdot v \nabla u dx$$

$$= -\int_{\Omega} (\nabla \cdot \mu v) u dx + \int_{\partial\Omega} (\mu \cdot n) u v dx \tag{7.37}$$

$$= -\int_{\Omega} (\nabla \cdot \mu) u v dx - \int_{\Omega} \mu \cdot u \nabla v dx + \int_{\partial\Omega} (\mu \cdot n) u v dx$$

Imposing inflow Dirichlet boundary data $u_{|\,\partial\Omega}$ onto the boundary integral, (7.37) becomes

$$\begin{aligned}(\mu \cdot \nabla u, v) = &-\int_{\Omega} (\nabla \cdot \mu) u v dx - \int_{\Omega} \mu \cdot u \nabla v dx \\ &+ \int_{\partial\Omega_{in}} (\mu \cdot n) u_{|\,\partial\Omega} v dx + \int_{\partial\Omega_{out}} (\mu \cdot n) u v dx\end{aligned} \tag{7.38}$$

The last equation requires computation of the term $\int_{\Omega} (\nabla \cdot \mu) u v dx$, which is problematic because of the derivative of the numerically generated velocity field. However, $\int_{\Omega} (\nabla \cdot \mu) u v dx$ can be expressed from (7.37) as

$$\begin{aligned}&\int_{\Omega} (\nabla \cdot \mu) u v dx \\ &= -\int_{\Omega} \mu \cdot v \nabla u dx - \int_{\Omega} \mu \cdot u \nabla v dx + \int_{\partial\Omega} (\mu \cdot n) u v dx\end{aligned} \tag{7.39}$$

Inserting (7.39) into (7.37) results in

$$\begin{aligned}(\mu \cdot \nabla u, v) = &\int_{\Omega} \mu \cdot v \nabla u dx - \int_{\partial\Omega_{in}} (\mu \cdot n) u v dx \\ &+ \int_{\partial\Omega_{in}} (\mu \cdot n) u_{|\,\partial\Omega} v dx\end{aligned} \tag{7.40}$$

The advection term $\int_{\Omega} \mu \cdot v \nabla u dx$ is symmetric to the term $\int_{\Omega} \mu \cdot u \nabla v dx$ that represents advection of the conservative equation (3.15) and thus is computed similarly. The time derivative term, diffusion term and boundary terms of the weak formulation of equation (7.36) are computed identically to the corresponding terms in (3.4) and (3.16) (see Section 3.1.3 and Section 6.1.2).

To solve this DA problem by the interconnected localised filters, the tidal basin is decomposed into five subdomains as it can be seen in Figure 7.11. Each subdomain is discretised in space by bilinear finite elements as follows: domain $d1$ by 480 elements, domains $d2$ and $d4$ by 192

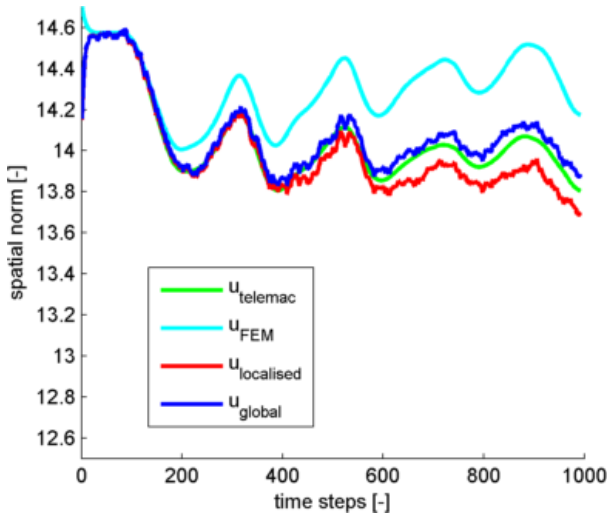

Figure 7.15 The spatial norm of the TELEMAC solution, FEM solution, localised minimax estimate and global minimax estimate plotted over time.

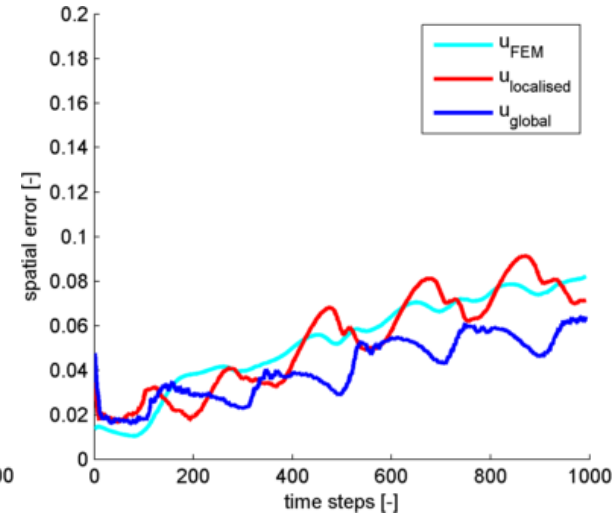

Figure 7.16 The spatial error of the FEM solution, localised minimax estimate and global minimax estimate plotted over time.

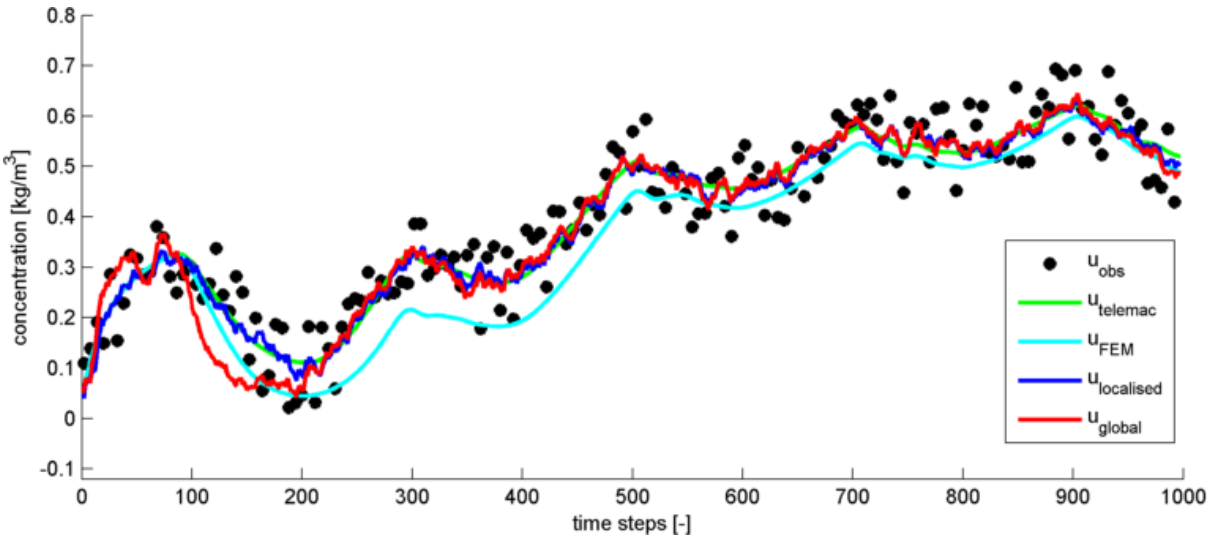

Figure 7.17 The observation, TELEMAC solution, FEM solution, localised minimax estimate and global minimax estimate taken in the point x=2.5, y=1.5 and plotted over time.

elements each, domain $d3$ (the harbour) by 256 elements and domain $d5$ by 400 elements. The observations $\mathbf{u}_{\text{obs}}$ are generated by restricting $\mathbf{u}_{\text{telemac}}$ to the third subdomain $d3$ and adding uniformly distributed random noise selected from the interval $[-0.1; 0.1]$.

The parameters of the localised minimax filter are configured as follows

$$\widetilde{\mathbf{Q}}_i = \frac{1}{\gamma_{T,i}}, \qquad \widetilde{\mathbf{Q}}_{0,i} = \frac{10}{\gamma_{T,i}}, \qquad \widetilde{\mathbf{R}}_i = \frac{300}{\gamma_{T,i}} \tag{7.41}$$

where $\gamma_{T,i} = (1 + \Delta t)A(\Omega_i)$, $\widetilde{\mathbf{Q}}_i$ reflects a moderate level of trust in the model, $\widetilde{\mathbf{Q}}_{0,i}$ describes unknown initial conditions and $\widetilde{\mathbf{R}}_i$ is the reciprocal of the variance of the uniformly distributed on the interval $[-0.1; 0.1]$ variable.

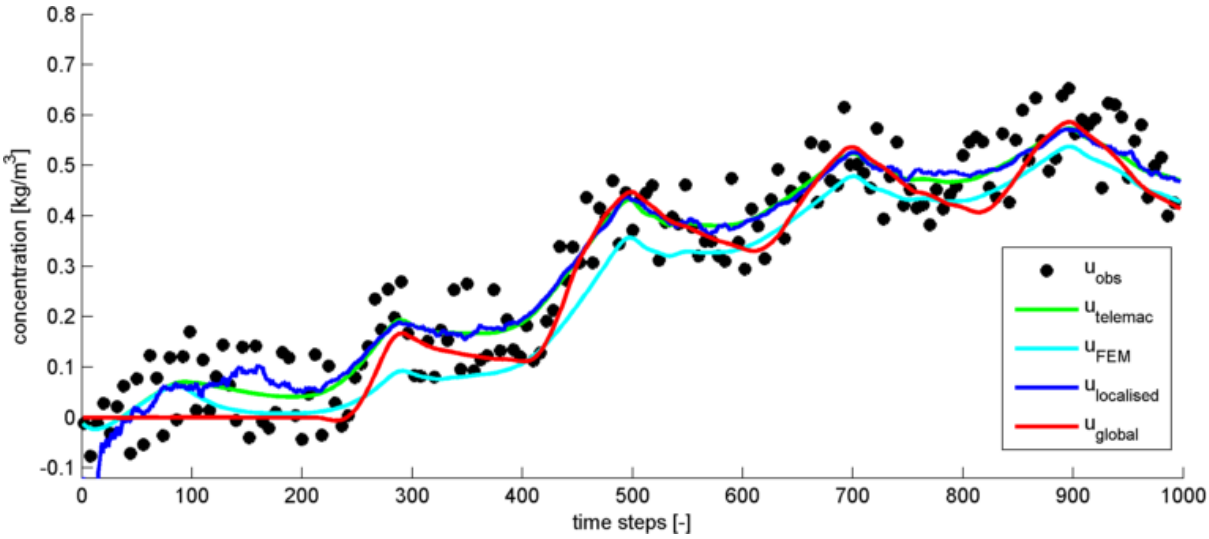

Figure 7.18 The observation, TELEMAC solution, FEM solution, localised minimax estimate and global minimax estimate taken in the point x=2.5, y=1.375 and plotted over time.

The results of estimation of the solution of the DA problem are shown in Figure 7.15 and Figure 7.16 where the spatial norm and the spatial error of the estimate $\mathbf{u}_{\text{localised}}$ obtained from the localised minimax filter is compared versus TELEMAC solution $\mathbf{u}_{\text{telemac}}$, free run FEM model with known initial condition $\mathbf{u}_{\text{fem}}$ and the traditional minimax estimate $\mathbf{u}_{\text{global}}$. It can be seen that initially, when the tracer is in the observed area, the corresponding spatial errors are low. But when part of the tracer leaves the harbour into the unobserved area, the spatial errors start to grow. In general, spatial errors of the localised and global filters are growing with intervals of errors decrease corresponding to the time when a tide is leaving the working area providing better updates from the observed data. It should be noted that the spatial errors of the localised filter exhibits more complex behaviour due to a DD and errors associated with its convergence.

A comparison of the estimation errors of the FEM solution $\mathbf{u}_{\text{fem}}$ and d-ADN FEM solution $\mathbf{u}_{\text{dADN}}$, $e_e(\mathbf{u}_{\text{fem}}) = 5.8\%$ and $e_e(\mathbf{u}_{\text{dADN}}) = 9.6\%$ suggests the significance of the discrepancies introduced by the decomposition method. At the same time, Figure 7.15 and Figure 7.16 demonstrate that global and localised estimates are similar. This is further confirmed by the corresponding estimation errors $e_e(\mathbf{u}_{\text{localised}}) = 6.8\%$ and $e_e(\mathbf{u}_{\text{global}}) = 5.7\%$. Finally, the localised minimax code executed in $432s$ while the traditional filter code ran for $7156s$.

The results of simulations are further presented for a point taken in the middle of the harbour entrance ($x = 2.5m$, $y = 1.5m$) in Figure 7.17. The figure compares localised and global estimates with FEM and TELEMAC solutions and also with numerically generated noisy observations. Initially, the concentration is increasing due to diffusion and then it increases or

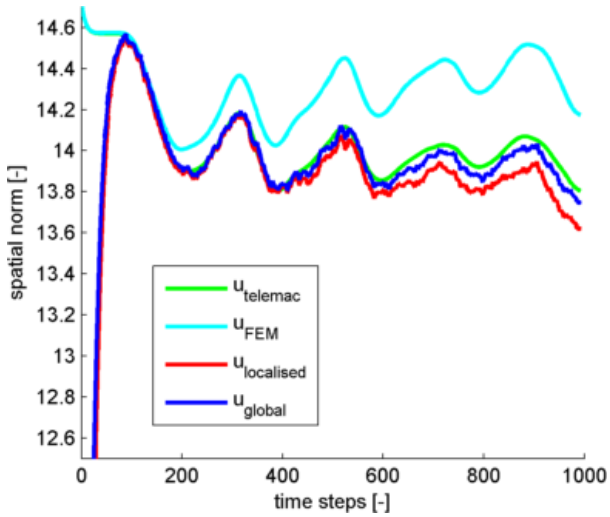

Figure 7.19 The spatial norm of the TELEMAC solution, FEM solution, localised EnTKF DT estimate and global EnTKF estimate plotted over time.

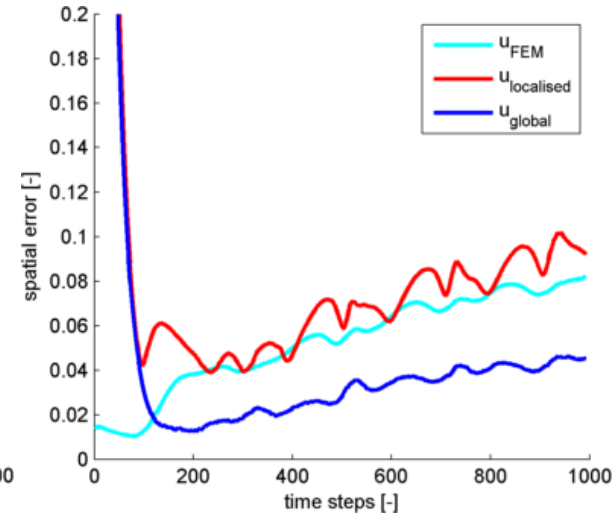

Figure 7.20 The spatial error of the FEM solution, localised EnTKF DT estimate and global EnTKF estimate plotted over time.

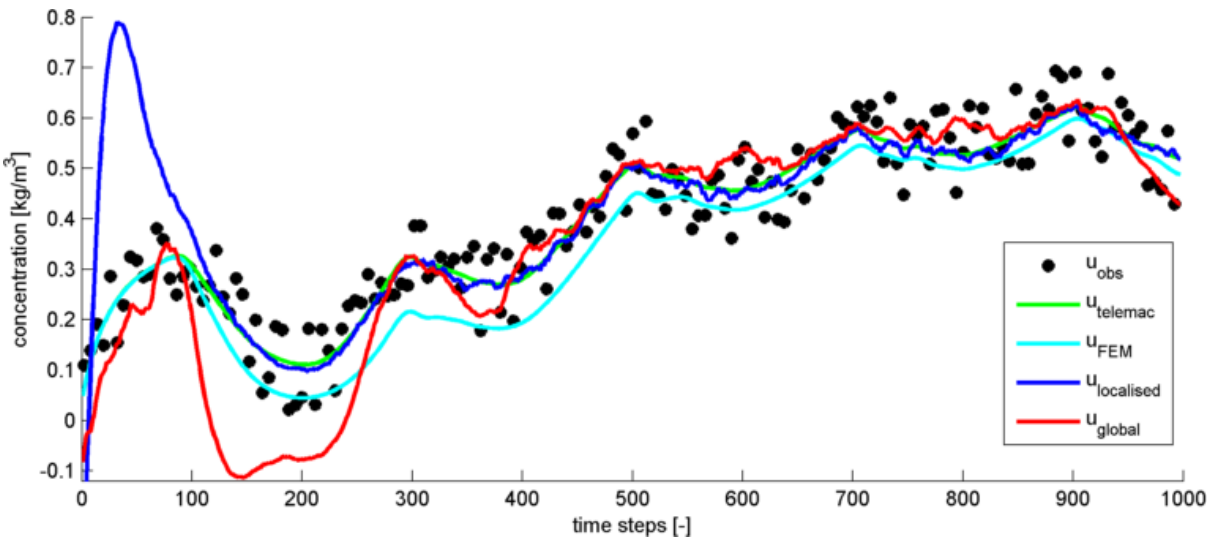

Figure 7.21 The observation, TELEMAC solution, FEM solution, localised EnTKF DT estimate and global EnTKF estimate taken in the point x=2.5, y=1.5 and plotted over time.

decreases depending if the tide is entering or leaving the harbour while the general trend is the increase of tracer concentration. It can also be seen that the localised minimax estimate is accurate and converges to the global minimax estimate.

Another estimate is plotted in Figure 7.18 for the point taken immediately outside of the harbour entrance ($x = 2.5m$, $y = 1.375m$) which is in the domain $d1$ and is not observed. In this case, the local subproblem does not contain any observations, but, as shown in Figure 7.18, communication between subdomains results in improvement of the estimate when the tide leaves the harbour and propagates estimated data from the harbour. When the concentration in

this point is influenced by the tide that comes from the not observed e.g. not filtered area and moves into the harbour, the quality of estimate decreases.

The tidal basin simulation was also conducted using the localised EnTKF DT with the ensemble size 100. A comparison of the spatial norm and spatial error of the localised EnTKF DT $\mathbf{u}_{\text{localised}}$ with the traditional EnTKF estimate $\mathbf{u}_{\text{global}}$ as well as FEM and TELEMAC solutions $\mathbf{u}_{\text{fem}}$ and $\mathbf{u}_{\text{telemac}}$ is presented in Figure 7.19 and Figure 7.20. The estimates of the tracer concentration in the point taken in the middle of the harbour entrance are shown in Figure 7.21. As it is expected, the convergence of the localised EnTKF DT is slower than of the convergence of the localised minimax filter which explains high level of discrepancies at the initial stage of estimation (see Figure 7.21). Also the estimation quality is lower as well, the corresponding estimation error are $e_e(\mathbf{u}_{\text{localised}}) = 20.1\%$ and $e_e(\mathbf{u}_{\text{global}}) = 17.8\%$. At the same time, the localised EnTKF DT code took only $153s$ to execute which is the computationally most efficient estimator presented here. For comparison, global EnTKF code war running for $2038s$. Figure 7.20 and the estimation norm of the traditional EnTKF also shows that the quality of the estimate of EnTKF is better that for a traditional minimax filter. This happens because the EnTKF method is constructed for a background model without model errors while the minimax filter considers modelling errors.

7.6 Discussion and conclusions

In this chapter, the interconnected localisation strategy was further extended for Kalman filter and EnTKF. Using the equivalence properties between minimax and Kalman filters, the interconnected localised Kalman filter is constructed in a straightforward way. Localisation of EnTKF is more complicated than localisation of Kalman filter due to its mechanism of the uncertainty propagation. Depending on the treatment of the transmission conditions, two different algorithms were suggested: (i) the localised EnTKF DT with the deterministic transmission and (ii) the localised EnTKF ST with the stochastic transmission. It was explained and demonstrated that the localised EnTKF ST provides better quality estimates, but the localised EnTKF DT is less computationally demanding for the same ensemble size.

A comparison of the theoretical computational complexity of the various interconnected localised filters was also presented. That comparison demonstrates significant computational benefits of the localised methods and their nearly linear scalability properties with respect to a

number of degrees of freedom. The theoretical comparison is further justified by the numerical experiments with idealised configuration. It is concluded that the localised minimax filter is faster than the traditional filters, but slower than other localised filters. The fastest computations can be achieved if the localised EnTKF DT with a small ensemble size is used. Furthermore, the benefits of the localised filters are explained and demonstrated with a numerical simulation that compares it with the so-called consensus Kalman filter (Battistelli et al., 2016, 2015) which also uses DD for computational costs reduction.

Finally, the performance of the interconnected localised minimax filter and interconnected EnTKF DT was assessed for the tidal basin simulation with complex velocity flows. The numerical experiments demonstrate the computational efficacy of the proposed approaches.

8 Summary and Conclusions

8.1 Summary and discussion

Assimilation of available observations into a model described by the PDE is an important engineering problem. While a mathematical solution of the problem is known and can be obtained from the solution of the corresponding FP equation, its application in practice is often limited due to high computational requirements. It is common to use linear assimilation methods such as the Kalman filter or its more robust deterministic equivalent: the minimax filter. A disadvantage of those methods is that they are not scalable, and computational costs escalate for large-scale problems that typically occur in marine modelling applications. As demonstrated in the literature review, over the last few decades, there were many attempts to construct more efficient assimilation methods. This problem, however, still remains active.

This thesis is mainly concerned with the construction of performance efficient DA methods. To this end, it is suggested to localise computations by coupling traditional DA methods with DD. The application of DD techniques for the solution of PDE has several advantages: the computational cost can be optimally managed, different geometric forms of subdomains can be exploited, various patterns of the physical solution can be captured using different numerical schemes in corresponding subdomains and the resulting subproblems can be solved independently. Here, the applicability of DD is extended to the assimilation problem.

In this thesis, the process of contaminant transport is considered. Starting with the time-space continuous formulation of the problem given by the linear advection-diffusion equation, it is reformulated in a weak sense. The weak problem is solved using Galerkin approximations. Specifically, a FEM approach with bilinear basis functions is employed. Discretisation in time is performed using the mid-point rule which is chosen due to its stability properties. According to those representations, the corresponding DA problems are equivalently introduced together with their optimal estimation methods. For a problem continuous in time and discrete in space, the Kalman-Bucy filter is devised. For a problem discrete in time and space, an algorithm of the Kalman filter is obtained which is further approximated by the EnKF.

The usage of Kalman filters relies on the exact description of the model uncertainties which are assumed to be Gaussian white noise. The quality of the estimation produced by those filters is sensitive to the accuracy of the definition of the input parameters. At the same time, an

accurate statistical description of the model and observation errors often is not available in practical situations. To increase the robustness of an estimation method, instead of a statistical description of errors, it is assumed that modeling errors are unknown, but bounded. The latter attribute is a natural assumption as it reflects the fact that the energy of the system is limited. Once the set of all possible error realisations is known, the state estimate is found using a minimax filter which finds the realisation that minimizes the error of the worst-case scenario.

In this work, a minimax filter is developed in a feedback form as an extension of the approach presented by Zhuk (2013). The first extension is that model errors are assumed to have linear constraints defined by L^{∞}-type norm. Measurement errors are assumed to be a realization of a random process, so the seconds moments of those errors are assumed to be unknown but pointwise bounded so that, in fact, it belongs to the L^{∞} ellipsoid produced by the nonnegative definite operator. The L^{∞} ellispsoids are approximated by the L^2 ellipsoids which is geometrically interpreted as approximating a rectangle by a circle that contains it. For L^2 ellipsoids, the estimation problem is given as an LQ control problem and its solution is provided by the minimax filter which represents the solution of the corresponding optimality system. The second extension is that the filter is defined for a continuous DA problem in an operator form and later discretised using Galerkin and FEM approximations following the "optimise and discretise" paradigm. An application of FEM produces discrete feedback and matrix DRE. Benner and Mena (2004) analysed several of the most common approaches for DRE solution. For the problem considered in this research, symplectic mid-point schema (Hairer et al., 2006) with Möbius transformation (Benner and Mena, 2004) and reinitialisation (Frank and Zhuk, 2014) was chosen because of its capability to integrate through singularities of DRE. Following the introduction of the traditional filters, the mathematical equivalence between the Kalman-Bucy filter and the minimax filter is explained suggesting a chain of equivalences between the minimax filter, Kalman filter and EnTKF.

For the implementation of the code of traditional filters, matrix operations are performed using data decomposition approach provided by the OpenBLAS library which implements shared memory parallelisation. Nevertheless, performed numerical simulations demonstrate that the traditional filters require significant computational cost and are not scalable as suggested by the theoretical complexity estimates.

To construct more efficient methods of DA, two strategies of coupling filters and DD are presented in this research. The first strategy follows the standard idea of decomposition applied to the underlying background problem which produces a set of additional equations enforcing

continuity of the solution across the interface of decomposition known as transmission conditions. Thus, the original PDE is reformulated as a DAE and is discretised in space using FEM approximations, although any other spatial discretisation techniques may be used. Since obtaining solutions to DAE systems in a general case is very problematic (Hairer and Wanner, 1996), it is suggested to consider discrete transmission conditions as an additional set of observations with a very small error. The state of a DAE is then estimated using minimax estimation approach where transmission conditions are treated as an algebraic optimisation constraint. The latter is accomplished by applying methods of LQ control theory, in specific the generalised Kalman duality principle proposed in Zhuk (2013). As a result, the DAE minimax filter is devised.

The results of numerical simulations demonstrate that the DAE minimax filter provides similar quality of estimates to the traditional minimax filter. However, it should be explained that from a computational point of view the approach does not provide significant benefits. While the localised stiffness and observation matrices of the DAE minimax filter possess a block matrix structure (due to DD), the presence of the algebraic constraint in the Riccati equation blends the matrix components and leads to a general form of the Riccati matrix. Thus, the full-size Riccati and feedback equations should be solved which require the same amount of operations as solving traditional minimax filters.

The computational costs of the DAE minimax filter may be decreased if the structure and patterns of the transmission matrix and Riccati equation are studied. The transmission matrix has non-zero components only for boundary nodes, thus if FEM nodes are renumbered that may lead to the Riccati equation which obeys a certain block structure that may be cheaper to solve than the original one.

An attractive property of the localisation of the minimax filter by means of DAE is that the approach does not depend on physical properties of the modelled process over the decomposition interface. Since the discrete transmission conditions are treated as an algebraic constraint of an optimisation problem it can be chosen independently from the problem operator provided its continuous equivalent satisfies continuity of a decomposed problem. Although it should be noted that the choice of transmission conditions may effect estimation quality of the method. For example, the choice of the Dirichlet transmission conditions that are used in this work does not depend on flow patterns of the underlying transport process and can be used for both advection and diffusion dominated flows. Also, it may be replaced by the Robin transmission conditions in a straightforward way, which in the case of FEM

approximation would guarantee not only continuity of the solution but also continuity of its first derivative. Additionally, matrix patterns produced by different transmission conditions may have an impact on the computational performance of the algorithm.

The DAE minimax filter can be solely used for the solution of the decomposed problem. In a general case, it is more expensive than classical decomposition solvers such as the Schwarz iterative approach, due to the need for the Riccati equation computation. However, if the stiffness matrix of the underlying problem is stationary, the solution of the Riccati equation approaches to the solution of the stationary Riccati equation also known as the algebraic Riccati equation (equation without time-derivative term). The resulting method for the solution of the decomposed problem would require only one solution of the algebraic Riccati equation which is cheaper than the Schwarz approach. That may have a potential application for the solution of the problems, where the model operator is independent of time, and the system dynamics is driven by the boundary conditions or source terms that do not impact the stiffness matrix.

The idea to apply LQ control to recast local subdomain that is used by the DAE minimax filter is similar to the virtual controls approach. Several variants of this approach are described in Discacciati et al. (2012) where they are used to solve heterogeneous problems, for instance, problem of advection/advection-diffusion coupling, by combining DD and control theory. In that regard, the DAE minimax filter can be seen as a combination of the virtual controls approach and minimax framework. Application of DA to heterogeneous problems is another potential application for the DAE minimax filter since the traditional filters cannot be applied to heterogeneous problems directly.

Another resemblance to the DAE minimax filter approach can be founded in (D'Amore et al., 2013) and (D'Amore et al., 2014) where DD-oceanVar approach is designed. DD-oceanVar also uses DD with the transmission conditions enforced as an additional constraint of an optimisation problem, i.e., an extra term of a cost function. Although the authors reported improvement in a computational time of the DD-oceanVar approach in comparison to no DD method, it should be stressed that in their work the 3d-Var estimator is used. Since 3d-Var analysis takes into account observations only from a current time step its estimation quality is lower than the one that is obtained using minimax/Kalman filters (Bouttier and Courtier, 1999).

The second strategy of coupling filters and DD presented in this thesis is the interconnected localisation of the minimax filters. This strategy employs DD to the assimilation problem with pointwise bounded errors controlled by L^∞ ellipsoids. Usage of L^∞ ellipsoids allows for

decomposing not only of operators of the background problem but also error ellipsoids. Thus, the global problem is reformulated as a set of independent local subproblems which are interconnected by transmission conditions enforced through local boundary conditions. After the localisation of the problem is done, the minimax filter is applied to each subproblem. To resolve this set of local filters, each of them is discretised in space and time using FEM and mid-point methods and computed using the Schwarz iterative approach.

In contrast to the DAE minimax approach, in this case, the choice of the transmission conditions should satisfy physical properties of the underlying problem, in particular, configuration of boundary conditions. For this need, an adaptive method called d-ADN is used since it satisfies the hyperbolic limit of the advection-dominated problem and assigns Dirichlet or Neumann conditions according to inflow/outflow zones of the corresponding local boundary (Gastaldi et al., 1998).

Numerical experiments performed in this work demonstrate that the interconnected localised and global versions of the minimax filter provide very similar quality of estimates. Those experiments further reveal a crucial role of the decomposition method on the estimation quality of the interconnected localised minimax filter. It is known that the d-AND method performs well for advection dominated flows which are characterised by a high value of the Péclet number. However, the Péclet number is averaged over a spatial domain measure and may not reflect the full complexity of a flow field well. For instance, during the initial period of the numerical tidal basin simulation, the flow in the harbour is produced by diffusion only while the tide, generated on the south boundary, is propagating into the working area and generates a high Péclet number. This inconsistency results in discrepancies caused by the incorrect treatment of the boundary conditions in the d-AND method between the harbour and working area. It is known that the damped adaptive Robin-Neumann (d-ARN) DD method (Gastaldi et al., 1998) is less sensitive to the Péclet number, but usually requires more Schwarz iterations for convergence. Other factors that influence the convergence of DD methods are damping (Ciccoli, 1996) and relaxation (Marini and Quarteroni, 1989) of boundary conditions over the interface of decomposition. An optimal choice of parameters of those factors and their influence on the convergence of the localised minimax filters requires additional investigation. It should be also explained that in the case when the flow is advection dominated in some areas or instances of time and is diffusion dominated in another areas or instances of time it is possible to construct hybrid transmission conditions that are changing depending on the Péclet number. For decomposition interface segments with a high value of the Péclet number, the

adaptive boundary conditions are used such as d-ADN or ARN, but if the Péclet number over a segment is low, then Dirichlet or Robin conditions should be applied.

The interconnection mechanism of the localised minimax filter provides communication between local estimates using boundary conditions of local subproblems. Those boundary conditions are imposed on the weak representation through boundary integrals which further produce source terms of feedback equations. Those source terms provide spreading of uncertainties between subproblems. However, since the local uncertainty equations (represented by the Riccati equation or error covariance propagation) are independent, the connection between local uncertainties through source terms is weak. To improve the uncertainty propagation, a strategy with pseudo filters is suggested. This strategy employs additional local filters with pseudo-observations defined over decomposition interface boundaries on subdomains without actual observations. Parameters of those pseudo filters are configured to represent uncertainty transmission between adjacent subproblems.

The approximation of L^∞ ellipsoids by L^2 ellipsoids that is used for the minimax filter derivation can be geometrically illustrated as approximating a circle by the smallest rectangle which contains it. The quality of the ellipsoidal approximation depends on a domain size and length of the estimation window T. For the global minimax filter this approximation can be rather crude especially for a large domain size and long integration window. The localised minimax filter constructed here mitigates ellipsoidal approximation error in two ways. First, because of decomposition, ellipsoidal approximation is performed on smaller subdomains. So that, the "small" ellipsoids L_i^∞ (circles) approximate the "small" ellipsoids L_i^2 (rectangles) and the union of the ellipsoids L_i^2 is contained in the large ellipsoid L^2 approximating the entire L^∞. Increasing the amount of subdomains that decompose the global domain, results in reduction of the ellipsoidal approximation error. In addition, the reinitialisation procedure is introduced for mitigating the impact of the estimation window T on ellipsoidal approximation error. This is achieved thanks to the Markovian property of the minimax estimate that allows the size of the estimation horizon to be taken as small as the size of the time integration step. As demonstrated in the numerical experiments, the proposed reinitialisation procedure together with the error decomposition facilitate the mitigation of the ellipsoidal approximation error and significantly reduce the error estimates of the localised minimax filter comparing to the global minimax filter.

The computational demands of the interconnected localised minimax filter and global minimax filter are compared by estimating a number of operations required for a one timestep integration

of the corresponding filter. It is shown that the localisation significantly reduces computational requirement in contrast to a global filter. Those estimates further suggest that the localised algorithm possesses scalability properties with regard to the number of finite elements. Theoretical considerations are indeed confirmed by the numerical experiments. The main computational reduction introduced by the localisation approach is related to the treatment of the Riccati equation. While the global minimax filter is solving the global Riccati equation, the interconnected localised filter is solving a set of smaller size Riccati equations. It should be noted that the same solver is used for both global and local DREs, specifically the symplectic mid-point rule with Möbius transformation and reinitialisation.

The resolution of local optimality systems using local DRE as used in this work is not the only possible option. For an application in optimal control, Benamou (1999) applied DD to the optimality system which resulted in a set of local optimality systems. But instead of using the Riccati equation, those local systems were reformulated using synthesis techniques and feedback law computations. Similarly, Heinkenschloss and Herty (2007) performed reordering of rows and columns to apply Schur complement formulation. The drawback of both approaches is a presence of the final conditions which are inherited from a backward equation of the optimality system. The latter makes those approaches unattractive for online DA problems since the corresponding systems should be recomputed every time an estimation window is increased.

The interconnected localisation strategy is further extended for Kalman filter and EnTKF. Using the equivalence properties between the minimax filter and the Kalman filter obtained here for the mid-point discretisation of the DA problem, the interconnected localised Kalman filter is constructed in a straightforward way. As expected, computational demands of the interconnected localised Kalman filter are decreased by an order compared to the global Kalman filter and the localised filter possess a scalability property. Numerical simulations of the interconnected localised Kalman filter were successfully performed for a decomposition with up to twenty subdomains in contrast to the algorithm combining DD and Kalman filter presented by Fujimoto and Kawahara (2001) where convergence was reported for a decomposition into two subdomains only.

Although the Kalman filter is defined on a spatially-temporally discrete domain, a continuous level information regarding the underlying PDE is used to construct the interconnected localised version of the Kalman filter. It means that the current approach cannot be applied to a DA problem defined on a discrete level. At the same time, distributed Kalman filters are using

discrete decomposition in combination with the traditional Kalman filter. Various methods have been proposed in the field of distributed Kalman filtering, but many of them still suffer from scalability issues or require a specific problem structure (Wang et al., 2015). Recently, Battistelli et al. (2015) presented an approach which uses continuous decomposition of the underlying PDE with FEM and discrete level transmission conditions enforced using an iterative approach that splits time steps into substeps and computes a consensus at each of them. The interconnected localised Kalman filter is compared to this method known as the consensus Kalman filter using the test experiment configuration. It is shown that for a small number of iterations, the localised approach provides slightly lower estimation error and is computationally cheaper. Furthermore, it is demonstrated that each additional iteration of the consensus Kalman filter is significantly more expensive that an additional iteration of the interconnected localised Kalman filter since it requires additional matrix operations on local subdomains to construct consensus. The results of this comparison are further extrapolated to other interconnected localised methods.

Localisation of ensemble filters can be achieved by associating a local region of a certain influence radius with each grid point (Ott et al., 2004) or by splitting observations into separate batches that have uncorrelated errors (Hunt et al., 2007). By contrast, in this work localisation is done by decomposing the original DA problem into a set of local subproblems in the same way as it was performed for the minimax and Kalman filters and applying ensemble filter on each subproblem following the equivalence between ensemble filters and the Kalman filter. In particular, this localisation is obtained for EnTKF filter and it can be easily extended for other variants of ensemble filters. Depending on the treatment of the transmission conditions, two different algorithms were proposed: (i) the localised EnTKF DT with deterministic transmission, and (ii) the localised EnTKF ST with stochastic transmission. It is explained and demonstrated that the localised EnTKF ST provides better quality estimates, while the localised EnTKF DT limits the amount of transmitted data which reduces the quality of estimates on the one side but decreases computational demands on the other.

Establishing the relationship between the localised and the traditional EnTKF is more difficult than in the case of the Kalman filter or minimax filter. The computational complexity of EnTKF filter considered here depends not only on the size of a state vector but also on the number of ensemble members. The idea of ensemble filtering is to quantify the uncertainty within the space that is derived by ensemble members. If the interconnected localised version of the EnTKF is used, a space of uncertainties is approximated in a piecewise manner where local

ensembles approximate the corresponding part of global uncertainties and communicate with each other through appropriate boundary conditions. Instead of the global ensemble space of size K, this approximation of uncertainties permits using local ensembles that cover a smaller subdomain with a smaller number of members $K_i < K$. This raises a question for further investigation: how much more can we decrease the amount of local ensemble members for a given increase in the number of subdomains?

A comparison of the theoretical computational complexity of the various interconnected localised and traditional filters is also presented. That comparison demonstrates significant computational benefits of the localised methods and their nearly linear scalability properties with respect to a number of degrees of freedom. The theoretical comparison is further justified by the numerical experiments with idealised configuration. In practice, EnTKF and ensemble filters, in general, are preferred to the minimax/Kalman filters due to lower computational demands which are easy to control by changing ensemble size. However, it is concluded that the localised minimax filter is faster than the traditional filters including EnTKF with relatively small number of ensembles, but slower than other localised filters. Quickest computations can be achieved if the localised EnTKF DT with a small ensemble size is used. Finally, the performance of the interconnected localised minimax filter and interconnected localised EnTKF DT filter was assessed for the numerical simulation with complex velocity flows of a tidal basin. The numerical experiments demonstrate the computational efficiency of the proposed approaches.

In this research, the Schwarz iterative algorithm, that is at the core of the devised localised filters is implemented in an alternating manner. It should be noted that it can be easily extended to a parallel version. Indeed, computations on each subdomain are independent of each other and for implementation on a supercomputer can be executed inside different computational nodes in parallel. The near linear scalability of the algorithms suggests that parallel scalability should be achieved as well. This makes the approach of interconnected localisation an attractive tool for engineering applications.

For numerical simulations, a 2D FEM model was developed for the solution of the linear non-stationary advection-diffusion PDE. The solution of the PDE is obtained using Galerkin approximations that adopt basis functions defined in an abstract space. In this work, basis functions are constructed using bilinear finite elements, however other types of FEM basis functions are also possible as well as other Galerkin approximation such as Spectral Methods with higher order non-local polynomial basis functions. Moreover, the proposed localised

filters can be extended to 3D advection-diffusion problems. Indeed, the divergence theorem that provides means for a cross domain communication can be also applied in 3D problems. Thus, the numerical schema can be easily extended to the 3D equation with the appropriate choice of 3D basis functions. This would result in a modified approach for local mass matrix, stiffness matrix and boundary vector computation while other parts of the localised filters algorithms would remain the same.

Since the localised filters follow standard practices of the Galerkin method, the existing codes can be harnessed if they provide mass and stiffness matrices and implement boundary conditions through the force terms. The localised filters can be also built on top of existing codes if they implement the Schwarz iterative approach for the solution of the background problem. Although in this case, the complexity estimates related to the solution of the feedback equation may change.

To summarise, the major novel elements of this research are (i) development of the DAE localised minimax filter with transmission conditions given in the form of an algebraic constraint; (ii) development of the minimax filter variant for L^{∞} errors' ellipsoids; (iii) development of the interconnected localised minimax filter based on a damped version of ADN DD and Schwarz approach; and (iv) using cross filters equivalences, development of the interconnected localised Kalman filter and two variants of the interconnected localised EnTK filters with deterministic and stochastic transmission conditions.

8.2 Conclusions

The main conclusions of this research relate to the interconnected localisation methods of DA. The three main conclusions are:

- The interconnected localisation strategy that applies DD to the time-space continuous assimilation problem with non-Gaussian error description was introduced here in order to decrease computational demands of the assimilation problem. The application of the interconnected localisation strategy to the traditional minimax filter resulted in the interconnected localised minimax filter that has much lower computational demands, and yet a similar quality of estimation compared with the original minimax filter. For instance, a tidal basin simulation took $7156s$ to run using a global minimax filter; the

same scenario only took $432s$ to run when an interconnected localised minimax filter was implemented.

- Using the equivalence between the minimax filter and the Kalman filter obtained here for the mid-point time approximations, the interconnected localised Kalman filter is introduced. Using the equivalence between the Kalman filter and EnTKF, two different algorithms of the interconnected localised EnTKF filter are devised: (i) the localised EnTKF DT with deterministic transmission and (ii) the localised EnTKF ST with stochastic transmission. Performance of the new filters was studied on a set of numerical experiments and significant computational advantages were reported. Again comparing simulations of a tidal basin, a model using global EnTKF took $2038s$ to run, whereas a model incorporating an interconnected localised EnTKF DT run the same scenario in only $153s$.
- The interconnected localised minimax, Kalman and EnTKF filters are nearly linearly scalable due to the inherent nature of the DD approach. Scalability is indeed achieved if the additional degrees of freedom are organised into a separate fixed-size subdomain. Since the subdomains are independent, and continuity between them is enforced by the Schwarz iterative approach, the increase in the number of the fixed-size subdomain results in a linear increase in computational costs.

Several other significant conclusions are also drawn from this research:

- The robustness of the minimax filter is improved by assuming pointwise bounded errors controlled by L^∞ ellipsoids instead of L^2 ellipsoids. It was demonstrated that the ellipsoidal approximation errors generated are mitigated by the localisation strategy and time reinitialisation procedure proposed here.
- The localisation of the minimax filter can be achieved in the form of DAE. In this case, the transmission conditions are discretised by a FEM and written as an additional algebraic constraint enforced by the minimax filter. The localised stiffness and observation matrices of the DAE minimax filter possess block matrix structure (due to DD), but the presence of the algebraic constraint in the Riccati equation blends the matrix components and leads to a general form of the Riccati matrix. Thus, the full-size Riccati and feedback equations should be solved. From a computational point of view, the approach has an equivalent performance to the traditional minimax filter.
- The DAE minimax filter does not depend on physical properties of the modelled process over the interface of decomposition and thus can be easily applied to

heterogeneous problems. The approach can also incorporate constraints that are given in an algebraic form.

- A 2D FEM model was developed for the solution of the linear non-stationary transport phenomena described by the advection-diffusion PDE and observations contaminated with noise. The model supports a number of assimilation methods (traditional and localised) which were tested for different configurations such as constant and periodic velocity flows or tidal basin simulations. The structure of the implemented code is presented in Appendix A.
- The approaches presented herein are robust with respect to the spatial and temporal discretisation methods. Since the FEM approximation is used here, discretisation of local filters by any other Galerkin approximations may be obtained in a straightforward way. Discretisation by FDM is more complicated as it uses a different mechanism of boundary condition implementation.

8.3 Recommendations for future work

Ideas for further improvements and future extensions of the results undertaken in this thesis are suggested:

- Conduct performance comparisons of the interconnected localised filter with different adaptive transmission conditions such as: ADN, adaptive Robin-Neumann (ARN), adaptive conditions with/without dumping and relaxation.
- Expand the interconnected localisation idea to other DD approaches. For instance, time fraction methods such as explicit predictor implicit corrector methods (Jun and Mai, 2006) or corrected explicit-implicit DD methods (Liao et al., 2009) could be considered.
- Investigate possible computational advantages from leveraging sparse structure of the underlying FEM matrices and/or using specialised scalable numerical linear algebra techniques such as multigrid methods.
- Implement a parallel version of the interconnected localised filters on HPC cluster or supercomputer. Analyse computational properties and scalability of the parallel version of the localised filters.
- Investigate optimal amount of ensemble members of interconnected localised EnTKF for a given number of subdomains.

- Extend the interconnected localised DA to non-linear problems. For example, EnTKF is naturally designed for non-linear problems. Also, a variant of Kalman filter called Extended Kalman filter can be used with non-linear problems as well.
- Apply the interconnected localisation approach to other assimilation methods; for instance, using equivalence properties between 4d-Var and Kalman filter.
- Investigate structure and patterns of the transmission matrix and corresponding Riccati equation of the DAE minimax filter for a potential computational cost reduction.
- Investigate performance of the DAE minimax filter based on a stationary solution of the corresponding Riccati equation applied to resolving a decomposed problem where the model operator is independent of time, and the system dynamics is driven by the boundary conditions and/or source term that does not influence the stiffness matrix.
- Although NUI Galway has a tidal basin facility, it lacks concentration measuring devices, so only numerical simulations of the tidal basin were conducted in this research. However, the images of concentration can be recorded with a camera. The numerical simulations can incorporate those images using a non-direct observation operator, i.e., where observations provide only qualitative information such as whether the concentration at a current point is zero or not. This approach is known in the literature as image assimilation (Titaud et al., 2008).

Appendix A. Main data types of the FEM model

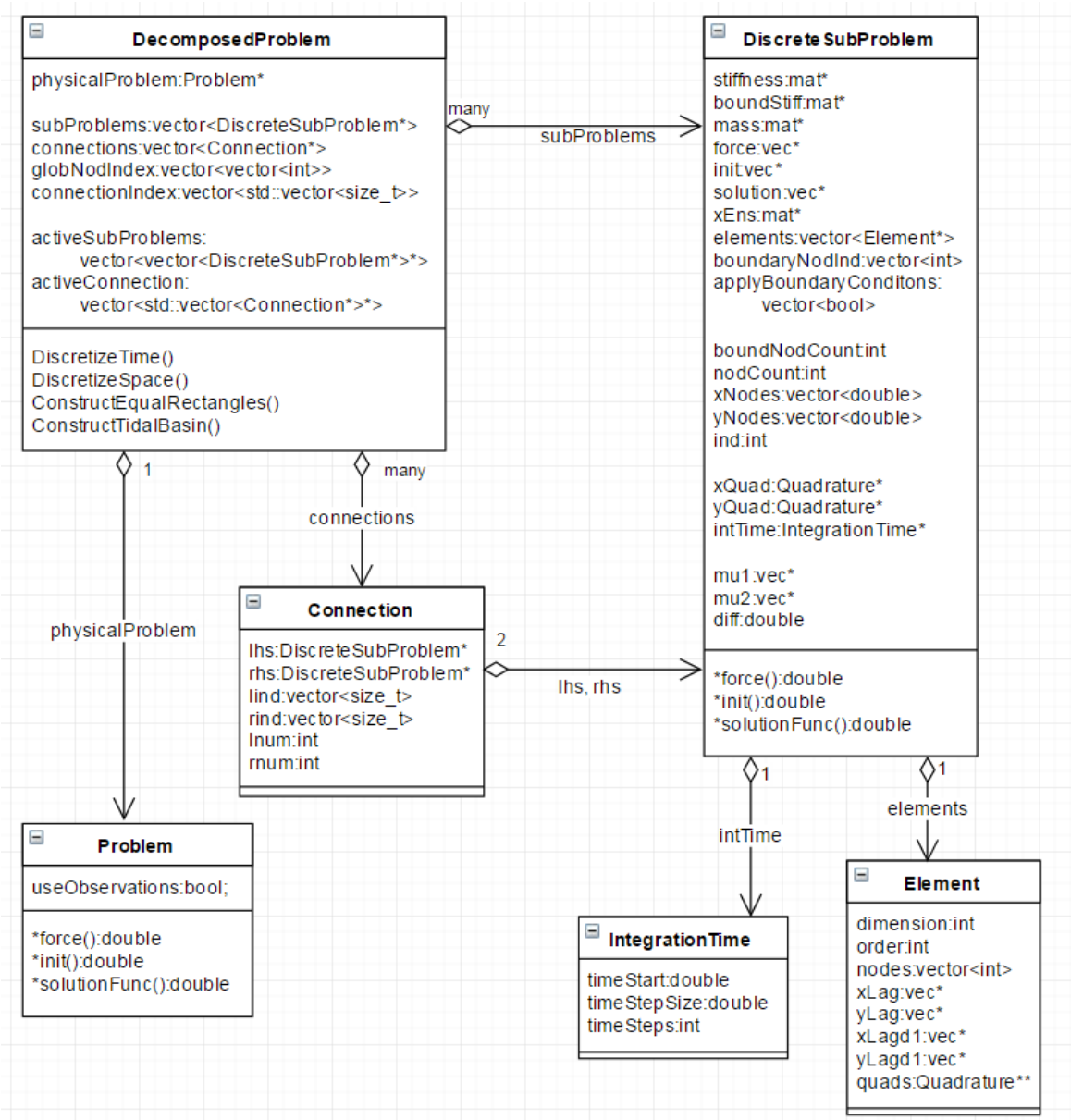

Figure A.0.1 FEM model data types.

Figure A.0.1 presents the main data types and relations between them that are used in the FEM code. The object **DiscreteSubProblem** contains information about a FEM discretisation of a local subdomain including list of FEM grid nodes (*xNodes* and *yNodes*), list of boundary nodes (*boundaryNodInd*), list of finite elements (*elements*), mass matrix (*mass*), stiffness matrix (*stiffness*), boundary stiffness matrix (*boundStiff*), vector of initial conditions (*init*), list of solutions at each time step (*solution*), integration time object (*intTime*) vectors of a velocity flow (*mu1* and *mu2*), diffusion coefficient (*diff*) and others.

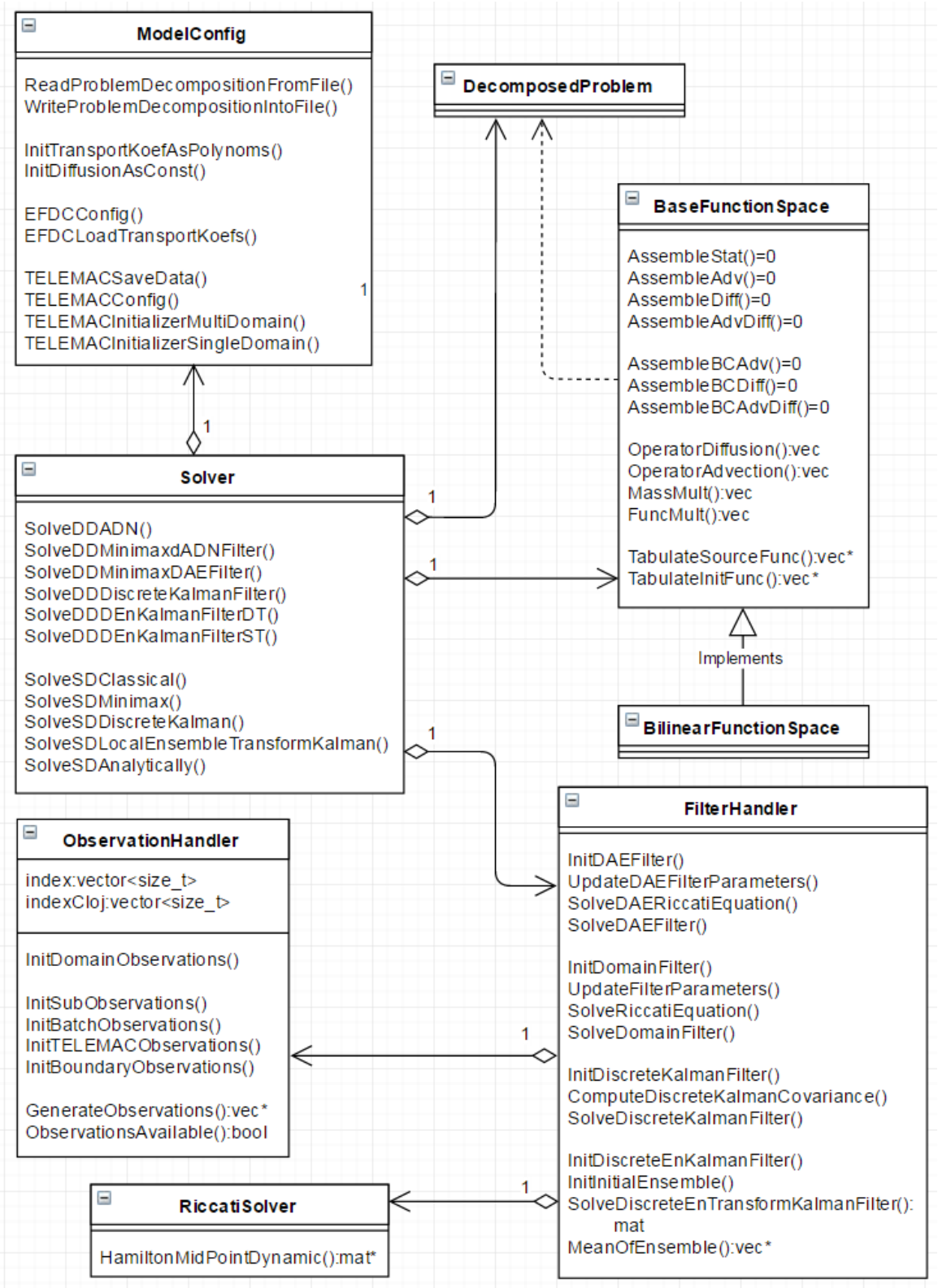

Figure A.0.2 Main functional components of the FEM model.

The main data type is object **DecomposedProblem**. It contains a list of FEM discretised local problems (*subProblems*), connections between them represented by a list of **Connection** objects (*connections*) and indices that are mapping nodes from a local numbering to a global

numbering (*globNodIndex* and *connectionIndex*). Each **Connection** object describes a discretised physical connection between local subproblems and has two references to **DiscreteSubProblem** objects (*lhs* and *rhs*) and two lists of nodes for each subdomain that are taken from the common boundary (*lind* and *rind*). The **DecomposedProblem** object also contains the **Problem** object (*physicalProbelm*) with configurations of physical problem and is used for initialisation (*ConstructEqualRectangles()* and *ConstructTidalBasin()*) and discretisation (*DiscretizeTime()* and *DiscretizeSpace()*) of a model.

The main functional objects are depicted in Figure A.0.2. For the configuration of the **DecomposedProblem**, the **ModelConfig** is used. This object contains methods for reading and writing problem configuration from a file (*ReadProblemDecompositionFromFile()* and *WriteProblemDecompositionIntoFile()*), reading EFDC and TELEMAC data (*EFDCLoadTransportKoefs()* and *TELEMACInitializerMultiDomain*()). The implementation of FEM was done in **BaseFunctionalSpace** and **BilinearFunctionalSpace** objects where methods for computation and assembling of FEM matrices are placed. The **FilterHandler** object, together with the **ObservationHadnler** object provide an implementation of filtering algorithms. The core component of the presented model is the **Solver** object equipped with various subroutines that use other objects for initialisation, discretisation and computation of assimilation problems using both traditional (*SolveSD…()*) and localised (*SolveDD…()*) filters.

References

Amezcua, J., Ide, K., Kalnay, E., Reich, S., 2014. Ensemble transform Kalman–Bucy filters. Quarterly Journal of the Royal Meteorological Society 140, 995–1004.

Anderson, J., Anderson, S., 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Monthly Weather Review 127, 2741–2758.

Araujo, M., Filho, E., 2012. A General Solution of Fokker-Planck Equation. Journal of Statistical Physics 146, 610–619.

Auger, F., Hilairet, M., Guerrero, J., Monmasson, É., Orlowska-Kowalska, T., Katsura, S., 2013. Industrial applications of the Kalman filter: a review. IEEE Transactions on Industrial Electronics 60, 5458–5471.

Başar, T., Bernhard, P., 1991. H-infinity optimal control and related minimax design problems: a dynamic game approach, System & Control: Foundations & Applications. Birkhauser, Boston Basel Berlin.

Battistelli, G., Chisci, L., Forti, N., Pelosi, G., Selleri, S., 2015. Distributed Finite Element Kalman Filter, in: Proceedings of the European Control Conference (ECC).

Battistelli, G., Chisci, L., Forti, N., Selleri, S., Pelosi, G., 2016. Decentralized consensus finite-element Kalman filter for field estimation. arXiv:1604.02392 [cs.SY].

Benamou, J.D., 1999. Domain Decomposition, Optimal Control of Systems Governed by Partial Differential Equations, and Synthesis of Feedback Laws. Journal of Optimization Theory and Applications 102, 15–36.

Benner, P., Li, J.-R., Penzl, T., 2008. Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems. Numerical Linear Alfebra with Applications 15, 755–777.

Benner, P., Mena, H., 2004. BDF methods for large-scale differential Riccati equations. Proceedings of Mathematical Theory of Network and Systems, MTNS 2004 10.

Bergemann, K., Sebastian, R., 2012. An ensemble Kalman-Bucy filter for continuous data assimilation. Meteorologische Zeitschrift 21, 213–219.

Bertino, L., Evensen, G., Wackernagel, H., 2005. Sequential data assimilation techniques in oceanography. International Statistical Review 71, 223–241.

Bouttier, F., Courtier, P., 1999. Data assimilation concepts and methods (Meteorological Training Course Lecture Series).

Brun, O., Teuliere, V., Garcia, J.-M., 2002. Parallel Particle Filtering. Journal of Parallel and Distributed Computing 62, 1186–1202.

Brusdal, K., Brankart, J., Halberstadt, G., Evensen, G., Brasseur, P., van Leeuwen, P., Dombrowsky, E., Verron, J., 2003. A demonstration of ensemble-based assimilation methods with a layered OGCM from the perspective of operational ocean forecasting systems. Journal of Marine Systems 40, 253–289.

Caetano, F., Gander, M., Halpern, L., Szeftel, J., 2010. Schwarz waveform relaxation algorithms for semilinear reaction-diffusion equations. Networks and Heterogeneous Media 5, 487–505.

Cane, M., Kaplan, A., Miller, R., Benyang, T., Hackert, E., Busalacchi, A., 1996. Mapping tropical Pacific sea level: Data assimilation via a reduced state space Kalman filter. Journal of Geophysical Research 101, 22599–22617.

Canuto, C., Hussaini, Y., Quarteroni, A., Zang, T., 1988. Spectral Methods in Fluid Dynamics, Springer Series in Computational Physics. Springer-Verlag.

Carlenzoli, C., Quarteroni, A., 1995. Adaptive Domain Decomposition methods for advection-diffusion problems., in: Modeling, Mesh Generation, and Adaptive Numerical Methods for Partial Differential Equation, IMA Volumes in Mathematics and Its Applications. pp. 165–199.

Carpenter, R., Shaw, B., 2012. Weather Information for Smart Grid, in: Proceedings of the 3rd Conf. Weather, Climate, New Energy Economy. Presented at the 3rd Conf. Weather, Climate, New Energy Economy, New Orleans.

Chen, Z., 2003. Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond. Statistics 182, 1–69.

Chitchian, M., van Amesfoort, A., Simonetto, A., Keviczky, T., Sips, H., 2012. Particle Filters on Multi-Core Processors (No. Report number PDS-2012-001), Parallel and Distributed Systems Report Series. Delft University of Technology.

Cho, H., Venturi, D., Karniadakis, G., 2016. Numerical Methods for Highly-Dimensional Probability Density Function Equations. Journal of Computational Physics 305, 817–837.

Chung, T.J., 2010. Computational Fluid Dynamics, Second Edition. ed. Cambridge University Press.

Ciccoli, M., 1996. Adaptive Domain Decomposition Algorithms and Finite Volume/Finite Element Approximation for Advection-Diffusion Equations. Journal of Scientific Computing 11, 299–341.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, J., Hamrud, M., Hollinsworth, A., Fabier, F., Fisher, M., 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. Quarterly Journal of the Royal Meteorological Society 124, 1783–1807.

Courtier, P., Thépaut, J.-N., Hollingsworth, A., 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. Quarterly Journal of the Royal Meteorological Society 120, 1367–1387.

Cressman, G., 1959. An Operational Objective Analysis System. Monthly Weather Review 87, 367–374.

D'Amore, L., Arcucci, R., Carracciuolo, L., Murli, A., 2014. A Scalable Approach for Variational Data Assimilation. Journal of Scientific Computing 61, 239–257.

D'Amore, L., Arcucci, R., Carracciuolo, L., Murli, A., 2013. DD-OceanVar: a Domain Decomposition fully parallel Data Assimilation software for the Mediterranean Forecasting System. Procedia Computer Science 18, 1235 – 1244.

D'Amore, L., Arcucci, R., Murli, A., 2012. Numerical Methods for Data Assimilation: Kalman Filter. CMCC Research Paper 165.

Dawson, C., Dupont, T., 1992. Explicit/implicit conservative Galerkin domain decomposition procedures for parabolic problems. Mathematics of computation 58, 21–34.

Denisov, S., Horsthemke, W., Hanggi, P., 2009. Generalized Fokker-Planck Equation: Derivation and Exact Solutions. The European Physical Journal B 68, 567–575.

Descombes, S., Dolean, V., Gander, M., 2011. Schwarz Waveform Relaxation Methods for Systems of Semi-Linear Reaction-Diffusion Equations, in: Domain Decomposition Methods in Science and Engineering XIX, Lecture Notes in Computational Science and Engineering. Springer, pp. 423–430.

Desombre, J., Lang, P., 2013. Telemac Modelling System: Telemac-2D Software.

Dieci, L., 1992. Numerical integration of the differential Riccati equation and some related issues. SIAM Journal on Numerical Analysis 29, 781–815.

Dieci, L., 1991. Some numerical considerations and Newton's method revisited for solving algebraic Riccati equations. IEEE Transactions on Automatic Control 36, 608–616.

Discacciati, M., Gervasio, P., Quarteroni, A., 2012. Heterogeneous Mathematical Models in Fluid Dynamics and Associated Solution Algorithms, in: Multiscale and Adaptivity: Modeling, Numerics and Applications. Springer-Verlag Berlin Heidelberg.

Dolean, V., Jolivet, P., Nataf, F., 2015. An Introduction to Domain Decomposition Methods: algorithms, theory and parallel implementation. Master., France.

Donea, J., Huerta, A., 2003. Finite Element Methods for Flow Problems. Wiley.

Drake, J., Flanery, R., Walker, D., Worley, P., Foster, I., Michalakes, J., Stevens, R., Hack, J., Williamson, D., 1993. The message passing version of the parallel community climate model. Parallel Supercomputing in Atmospheric Science 500–513.

Dryja, M., Tu, X., 2007. A domain decomposition discretization of parabolic problems. Numerische Mathematik 625–640.

Evans, L., 1998. Partial Differential Equations, Graduate Studies in Mathematics. American Mathematical Society.

Evensen, G., 2009. Data Assimilation: the Ensemble Kalman Filter. Springer-Verlag.

Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics 53, 343–367.

Evensen, G., 1997. Advanced data assimilation for strongly nonlinear dynamics. Monthly Weather Review 125, 1342–1354.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research 99, 10143–10162.

Fairbairn, D., Pring, S.R., Lorenc, A.C., Roulstone, I., 2014. A comparison of 4DVar with ensemble data assimilation methods. Quarterly Journal of the Royal Meteorological Society 140, 281–294.

Farhat, C., Macedo, A., Tezaur, R., 1999. FETI-H: a scalable domain decomposition method for high frequency exterior Helmholtz problems. Proceedings of the Eleventh International Conference on Domain Decomposition Methods.

Fletcher, T., 2010. The Kalman Filter Explained.

Floris, C., 2013. Numeric Solution of the Fokker-Planck-Kolmogorov Equation. Engineering 5, 975–988.

Frank, J., Zhuk, S., 2014. Symplectic Mobius integrators for LQ optimal control problems, in: Proceedings of the IEEE Conference on Decision and Control. Presented at the IEEE Conference on Decision and Control.

Fujimoto, M., Kawahara, M., 2001. Domain Decomposition for Kalman Filter Method and Its Application to Tidal Flow at Onjuku Coast. Proceedings of 12th International Conference on Domain Decomposition Methods.

Fukumori, I., Malanotte-Rizzoli, P., 1995. An approximate Kalman filter for ocean data assimilation: An example with an idealized Gulf Stream model. Journal of Geophysical Research: Oceans 100, 6777–6793.

Gander, M., 2008. Schwarz methods over the course of time. Electronic Transactions on Numerical Analysis 31, 228–255.

Gander, M., Halpern, L., 2007. Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems. SIAM Journal on Numerical Analysis 45, 666–697.

Gander, M., Halpern, L., Nataf, F., 2003. Optimal Schwarz waveform relaxation for the one dimensional wave equation. SIAM Journal of Numerical Analysis 41, 1643–1681.

Gander, M., Halpern, L., Nataf, F., 2001. Optimized Schwarz methods, in: Twelfth International Conference of Domain Decomposition Methods. pp. 15–27.

Gander, M., Halpern, L., Nataf, F., 1999. Optimal convergence for overlapping and non-overlapping Schwarz waveform relaxation, in: Eleventh International Conference of Domain Decomposition Methods. pp. 27–36.

Gander, M., Stuart, A., 1998. Space-Time Continuous Analysis of Waveform Relaxation for the Heat Equation. SIAM Journal of Scientific Computing 19, 2014–2031.

Gastaldi, F., Gastaldi, L., 1993. On a Domain Decomposition for the Transport Equation: Theory and Finite Element Approximation. IMA Journal of Numerical Analysis 14, 111–135.

Gastaldi, F., Gastaldi, L., Quarteroni, A., 1998. ADN and ARN Domain Decomposition Methods for Advection-Diffusion Equations, in: Proceedings of the Ninth International Conference on Domain Decomposition Methods. Presented at the Ninth International Conference on Domain Decomposition Methods, pp. 334–341.

Gastaldi, F., Gastaldi, L., Quarteroni, A., 1996. Adaptive Domain Decomposition methods for advection dominated equations. East-West J. Numer. Math. 4, 165–206.

Gastaldi, L., 1994. A Domain Decomposition for the Transport Equations. Contemporary Mathematics 157, 97–102.

Gelman, A., Carlin, J., Stern, H., Rubin, D., 2003. Bayesian Data Assimilation, Second Edition. ed, Texts in Statistical Sciences. Chapman and Hall/CRC.

Ghil, M., Malanotte-Rizzoli, P., 1991. Data assimilation in meteorology and oceanography. Advances in Geophysics 33, 141–266.

Giladi, E., Keller, H., 2002. Space-time domain decomposition for parabolic equation. Numerische Mathematik 93, 279-313 pp.

Gordon, N.J., Salmond, D.J., Smith, A.F.M., 1993. Novel Approach to Nonlinear/non-Gaussian Bayesian State Estimation. IEE Proceedings-F 140.

Hairer, E., Lubich, C., Wanner, G., 2006. Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations, Second Edition. ed. Springer-Verlag Berlin Heidelberg.

Hairer, E., Wanner, G., 1996. Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems, Second Revised Edition. ed, Springer Series in Computational Mathematics. Springer.

Hamrick, J., 1992. A Three-Dimensional Environment Fluid Dynamics Computer Code: Theoretical and Computational Aspects, Special Report in Applied Marine Science and Ocean Engineering. Virginia Institute of Marine Science,College of William & Mary.

Heemink, A.W., Bolding, K., Verlaan, M., 1995. Storm surge forecasting using Kalman filtering (a review). Delft University of Technology, Faculty of Technical Mathematics and Informatics.

Heinkenschloss, M., Herty, M., 2007. A spatial domain decomposition method for parabolic optimal control problems. Journal of Computational and Applied Mathematics 88 – 111.

Hervouet, J.-M., 2007. Hydrodynamics of Free Surface Flows. Wiley.

Hoffman, M., Miyoshi, T., Haine, T., Ide, K., Brown, C., Murtugudde, R., 2012. An advanced data assimilation system for the Chesapeake Bay: Performance evaluation. Journal of Atmospheric and Oceanic Technology 29, 1542–1557.

Houser, P., De Lannoy, G., Walker, J., 2012. Hydrologic Data Assimilation, in: Approaches to Managing Disaster - Assessing Hazards, Emergencies and Disaster Impacts. Editor: Tiefenbacher J. InTech.

Houtekamer, P., Mitchell, H., 2001. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. Monthly Weather Review 129, 123–137.

Houtekamer, P., Mitchell, H., 1998. Data assimilation using an ensemble Kalman filter technique. Monthly Weather Review 126, 796–811.

Hunt, B., Kostelich, E., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. Physica D: Nonlinear Phenomena 230, 112–126.

Ide, K., Courtier, P., Ghil, M., Lorenc, A., 1997. Unified Notation For Data Assimilation: Operational, Sequential and Variational. J. Met. Soc. Japan 75, 181–189.

Iserles, A., 1996. A First Course in the Numerical Analysis of Differential Equations. Cambridge University Press.

Janjic, T., Nerger, L., Albertella, A., Schroter, J., Skachko, S., 2011. On Domain Localization in Ensemble-Based Kalman Filter Algorithms. Monthly Weather Review 139, 2046–2060.

Jazwinski, A., 1970. Stochastic Processes and Filtering Theory, Mathematics in Science and Engineering. Academic Press, New York and London.

Jordan, R., Kinderlehrer, D., Otto, F., 1998. The Variational Formulation Of The Fokker-Planck Equation. SIAM J. of Mathematical Analysis 29, 1–17.

Jun, Y., Mai, T.-Z., 2006. IPIC domain decomposition algorithm for parabolic problems. Applied Mathematics and Computation 177, 352–364.

Kalman, R.E., 1963. New methods in Wiener filtering theory, in: Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability, Edited by JL Bogdanoff and F. Kozin. John Wiley & Sons, New York.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82, 35–45.

Kalman, R.E., Bucy, R.S., 1961. New results in linear filtering and prediction theory. Journal of Basic Engineering 83, 95–108.

Kalnay, E., 2010. Ensemble Kalman Filter: current status and potential, in: Data Assimilation: Making Sense of Observations. Editors: Lahoz W, Swinbank R, Khattatov B. Springer-Verlag Berlin Heidelberg, pp. 69–92.

Kelly, D.T.B., Law, K.J.H., Stuart, A.M., 2014. Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time. Nonlinearity 27, 2579–2601.

Keppenne, C., 2000. Data Assimilation into a Primitive-Equation Model with a Parallel Ensemble Kalman Filter. Monthly Weather Review 128, 1971–1981.

Khattatov, B., Yudin, V., 2010. Representation and Modelling of Uncertainties in Chemistry and Transport Models, in: Data Assimilation: Making Sense of Observations. Editors: Lahoz W, Swinbank R, Khattatov B. Springer Berlin Heidelberg.

Khoromskij, B., 2012. Tensors-Structured Numerical Methods In Scientific Computing: Survey On Recent Advances. Chemometrics and Intelligent Laboratory Systems 110, 1–19.

Krener, A., 1980. Kalman-Bucy and Minimax filtering. IEEE Transactions on Automatic Control AC-25, 291–292.

Krener, A., Kang, W., 2011. Linear time invariant minimax filtering. Annual Reviews in Control 35, 166–171.

Kress, R., 1998. Numerical Analysis, Graduate Text in Mathematics. Springer.

Kumar, M., Chakravorty, S., Singla, P., Junkins, J., 2009. The Partition of Unity Finite Element Approach With hp-Refinement for the Stationary Fokker-Planck Equation. Journal of Sound and Vibration 327, 144–162.

Kurzhanski, A., Vályi, I., 1997. Ellipsoidal calculus for estimation and control, Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston.

Lahoz, W., 2010. Research Satellites, in: Data Assimilation: Making Sense of Observations. Editors: Lahoz W, Swinbank R, Khattatov B. Springer Berlin Heidelberg, pp. 301–323.

Lapin, A., Pieska, J., 2002. On the parallel Domain Decomposition algorithms for time-dependent problems. Lobachevskii Journal of Mathematics 10, 27–44.

Levy, D., 2010. Introduction to Numerical Analysis. University of Maryland.

Lewis, F., Xie, L., Popa, D., 2008. F., L Xie, D Popa. Optimal and Robust Estimation With an Introduction to Stochastic Control Theory. Second Edition. CRC Press.

Li, X., Demmel, J., Gilbert, J., Grigori, L., Shao, M., Yamazaki, I., 1999. SuperLU Users' Guide. Lawrence Berkeley National Laboratory.

Liao, H., Shi, H., Sun, Z., 2009. Corrected explicit-implicit domain decomposition algorithms for two-dimensional semilinear parabolic equations. Science in China Series A: Mathematics 52, 2362–2388.

Lions, P.-L., 1990. On the Schwarz Alternating Method. II. Stochastic interpretation and order properties., in: Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, R. Glowinski, J. Periaux, and O. Widlund (Eds.). SIAM, Philadelphia, PA, pp. 47–70.

Lions, P.-L., 1989. On the Schwarz Alternating Method. III. A variant for nonoverlapping subdomains., in: Second International Symposium on Domain Decomposition Methods, T. Chan, R. Glowinski, J. Periaux, and O. Widlund (Eds.). SIAM, Philadelphia, PA, pp. 222–223.

Lions, P.-L., 1988. On the Schwarz Alternating Method. I., in: First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Periaux (Eds.). SIAM, Philadelphia, PA, pp. 1–42.

Lo, J., Shizgal, B., 2006. Spectral Convergence Of The Quadrature Discretization Method In The Solution Of The Schrödinger And Fokker-Planck Equations: Comparison With Sinc Methods. The Journal of Chemical Physics 125.

Lorenc, A., 2003. The potential of the ensemble Kalman filter for NWP — a comparison with 4D-Var. Quarterly Journal of the Royal Meteorological Society 129, 3183–3203.

Lorenc, A., 1986. Analysis Methods for Numerical Weather Prediction. Quarterly Journal of the Royal Meteorological Society 112, 1177–1194.

Lorenc, A., 1981. A global three-dimensional multivariate statistical interpolation scheme. Monthly Weather Review 109, 701–721.

Luettich, R.A., Westerink, J.J., Scheffner, N.W., 1992. ADCIRC: an advanced three-dimensional circulation model for shelves coasts and estuaries, report 1: theory and methodology of ADCIRC-2DDI and ADCIRC-3DL (Dredging Research Program Technical Report DRP-92-6, U.S. Army Engineers Waterways Experiment Station, Vicksburg, MS).

Lynch, P., 2008. The Origins of Computer Weather Prediction and Climate Modeling. Journal of Computational Physics 227, 3431–3444.

Lyster, P., Cohn, S., Menard, R., Chang, L.-P., Lin, S.-J., Olsen, R., 1997. Parallel Implementation of a Kalman Filter for Constituent Data Assimilation. Monthly Weather Review 125, 1674–1686.

Mahmoud, M., Khalid, H., 2013. Distributed Kalman filtering: A bibliographic review. IET Control Theory and Applications 7, 483–501.

Majda, A., Harlim, J., Gershgorin, B., 2010. Mathematical strategies for filtering turbulent dynamical systems. Dynamical Systems 27, 441–486.

Marini, L., Quarteroni, A., 1989. A Relaxation Procedure for Domain Decomposition Methods Using Finite Elements. Numerical Mathematics 55, 575–598.

Martens, W., 2012. A Semi-Analytical Method of Solving the Fokker-Planck-Equation for High-Dimensional Nonlinear Mechanical Systems. Proceedings in Applied Mathematics and Mechanics 243–244.

Martens, W., von Vagner, U., 2011. Advances in solving high dimensional Fokker-Planck equations, in: Proceedings of ENOC 2011. Presented at the ENOC 2011, Rome, Italy.

Martin, M.J., Balmaseda, M., Bertino, L., Brasseur, P., Brassington, G., Cummings, J., Fujii, Y., Lea, D.J., Lellouche, J.-M., Mogensen, K., Oke, P.R., Smith, G.C., Testut, C.-E., Waagbø, G.A., Waters, J., Weaver, A.T., 2015. Status and future of data assimilation in operational oceanography. Journal of Operational Oceanography 8, 28–48.

McPherson, R.D., Bergman, K.H., Kistler, R.E., Rasch, G.E., Gordon, D.S., 1979. The NMC operational global data assimilation system. Monthly Weather Review 107, 1445–1461.

Milanese, M., 1989. Estimation and prediction in the presence of unknown but bounded uncertainty: a survey, in: Robustness in Identification and Control. Editors: Milanese M. Tempo R. and Vicino A. Plenum Press, New York and London, pp. 3–24.

Mintz, M., 1972. A Kalman Filter as a Minimax estimator. Journal of Optimization and Applications 9.

Miyoshi, T., 2011. The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. Monthly Weather Review 139, 1519–1535.

Nakshatrala, K.B., Hjelmstad, K.D., Tortorelli, D.A., 2008. A FETI-based domain decomposition technique for time-dependent first-order systems based on a DAE approach. International Journal For Numerical Methods In Engineering 75, 1385–1415.

Nerger, L., Hiller, W., 2013. Software for ensemble-based data assimilation system — Implementation strategies and scalability. Computers & Geosciences 55, 110–118.

Nino, E., Sandu, A., Deng, X., 2016. A Parallel Implementation of the Ensemble Kalman Filter Based on Modified Cholesky Decomposition. arXiv:1606.00807 [cs.NA].

O'Donncha, F., Ragnoli, E., Suits, F., 2014. Parallelisation study of a three-dimensional environmental flow model. Computers & Geosciences 64, 96–103.

Olbert, A.I., Nash, S., Ragnoli, E., Hartnett, M., 2013. Turbulence modelling using 3DVAR data assimilation in laboratory conditions, in: Proceedings of Oceans. Presented at the Oceans, San-Diego.

Ott, E., Hunt, B., Szunyogh, I., Corazza, M., Kalnay, E., Patil, D., Yorke, J., Zimin, A., Kostelich, E., 2002. Exploiting Local Low Dimensionality of the Atmospheric Dynamics for Efficient Ensemble Kalman Filtering. arXiv:physics/0203058 [physics.ao-ph].

Ott, E., Hunt, B., Szunyogh, I., Zimin, A., Kostelich, E., Corazza, M., Kalnay, E., Patil, D., Yorke, J., 2004. A local ensemble Kalman filter for atmospheric data assimilation. Tellus 56A, 415–428.

Penduff, T., Brasseur, P., Testut, C.-E., Barnier, B., Verron, J., 2002. A four-year eddy-permitting assimilation of sea-surface temperature and altimetric data in the South Atlantic Ocean. Journal of Marine Research 805–833.

Pichler, L., Masud, A., Bergman, L., 2011. Numerical Solution of The Fokker-Planck Equation By Finite Difference And Finite Element Methods - A Comparative Study, in: 3rd ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, M. Papadrakakis, M. Fragiadakis, V. Plevris (Eds.). Corfu, Greece.

Quarteroni, A., Valli, A., 1999. Domain Decomposition Methods for Partial Differential Equations, Numerical Mathematics and Scientific Computation. Clarendon Press, Oxford.

Rabier, F., 2005. Overview of global data assimilation developments in numerical weather-prediction centres. Quarterly Journal of the Royal Meteorological Society 131, 3215–3233.

Ragnoli, E., Zhuk, S., Zayats, M., Hartnett, M., 2015. Localised filters for Linear Advection-Diffusion Equations, in: Proceedings of the IEEE Conference on Decision and Control (CDC). pp. 449–454.

Ragnoli, E., Zhuk, S., Zayats, M., Hartnett, M., 2014. Domain Decomposition for a Linear Advection-Diffusion Equation by means of minimax filtering, in: Proceedings of the European Control Conference (ECC). pp. 2733–2738.

Rantakokko, J., 1997. Strategies for parallel variational data assimilation. Parallel Computing 23, 2017–2039.

Rebechini, P., van Handel, R., 2015. Can Local Particle Filters Beat The Curse Of Dimensionality? The Annals of Applied Probability 25, 2809–2866.

Reichle, R., 2008. Data assimilation methods in the Earth sciences. Advances in Water Resources 31, 1411–1418.

Reid, W., 1972. Riccati Differential Equations, Mathematics in Science and Engineering. Academic Press, New York and London.

Richardson, L., 1922. Weather Prediction by Numerical Process. Cambridge University Press.

Risken, H., 1989. The Fokker-Planck Equations Methods of Solution and Applications, Second edition. ed, Springer Series in Synergetics. Springer-Verlag.

Rood, R., Bosilovich, M., 2010. Reanalysis: Data Assimilation for Scientific Investigation of Climate, in: Data Assimilation: Making Sense of Observations. Editors: Lahoz W, Swinbank R, Khattatov B. Springer Berlin Heidelberg.

Rutherford, I., 1972. Data assimilation by statistical interpolation of forecast error fields. Journal of the Atmospheric Sciences 29, 809–815.

Sage, A., White, C., 1977. Optimum Systems Control, Second edition. ed. Prentice-Hall, Englewood Cliffs, New Jersey.

Sanderson, C., Curtin, R., 2016. Armadillo: a template-based C++ library for linear algebra. Journal of Open Source Software 1.

Schiff, J., Shnider, S., 1999. A Natural Approach to the Numerical Integration of Riccati Differential Equations. SIAM J. Numer. Anal. 36, 1392–1413.

Schwarz, H., 1870. Uber einen Grenzübergang durch alternierendes Verfahren. Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich 272–286.

Shald, S., 1999. The continuous Kalman filter as the limit of the discrete Kalman filter. Stochastic analysis and applications 17, 841–856.

Smith, M.W.A., Roberts, A.P., 1978. An exact equivalence between the discrete-and continuous-time formulations of the Kalman filter. Mathematics and Computers in Simulation XX 2, 102–109.

Sun, Y., Kumar, M., 2015. A numerical Solver for High Dimensional Transient Fokker-Planck Equation in Modeling Polymeric Fluids. Journal of Computational Physics 289, 149–168.

Sun, Y., Kumar, M., 2014. Numerical Solution of High Dimensional Stationary Fokker-Planck Equations via Tensor Decomposition and Chebyshev Spectral Differentiation. Computers and Mathematics with Applications 67, 1960–1977.

Swinbank, R., 2010. Numerical Weather Prediction, in: Data Assimilation: Making Sense of Observations. Editors: Lahoz W, Swinbank R, Khattatov B. Springer Berlin Heidelberg.

Szunyogh, I., Satterfield, E., Aravequia, J., Fertig, E., Gyarmati, G., Kalnay, E., Hunt, B., Kostelich, E., Kuhl, D., Ott, E., Yorke, J., 2007. The local ensemble transform Kalman Filter and its implementation on the NCEP global model at the University of Maryland. Proceedings of the ECMWF Workshop on Flow dependent aspects of data assimilation.

Talagrand, O., 2010. Variational assimilation, in: Data Assimilation: Making Sense of Observations. Editors: Lahoz W, Swinbank R, Khattatov B. Springer-Verlag Berlin Heidelberg, pp. 41–67.

Talagrand, O., Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. Quarterly Journal of the Royal Meteorological Society 113, 1311–1328.

Tamura, H., Bacopoulos, P., Wang, D., Hagen, S., Kubatko, E., 2014. State estimation of tidal hydrodynamics using ensemble Kalman filter. Advances in Water Resources 63, 45–56.

Tchrakian, T., Zhuk, S., Nogueira, A., 2015. Robust Traffic Density Estimation Using Discontinuous Galerkin Formulation of a Macroscopic Model. IEEE 18th International Conference on Intelligent Transportation Systems (ITSC).

Tempo, R., Vicino, A., 1990. Optimal algorithms for system identification: a review of some recent results. Mathematics and Computers in Simulation 32 585–595.

Thépaut, J.-N., Andersson, E., 2010. The Global Observing System, in: Data Assimilation: Making Sense of Observations by Lahoz W, Swinbank R, Khattatov B. Springer Berlin Heidelberg, pp. 263–281.

Titaud, O., Vidard, A., Souopgui, I., Francois-Xavier, L.D., 2008. Assimilation of Image Sequences in Numerical Models. [Research Report] RR-6701.

Todling, R., Cohn, S., 1994. Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. Monthly Weather Review 122, 2530–2557.

Toselli, A., Widlund, O., 2005. Domain Decomposition methods - Algorithms and Theory. Springer.

Tran, M.-B., 2011. Convergence Properties of Overlapping Schwarz Domain Decomposition Algorithms. arXiv:1104.4294 [math.NA].

Tremolet, Y., 2006. Accounting for an imperfect model in 4D-Var. Quarterly Journal of the Royal Meteorological Society 132, 2483–2504.

Trotta, R.L., 1996. Multidomain Finite Elements for Advection-Diffusion Equations. Applied Numerical Mathematics 21, 91–118.

van der Vooren, M., 1990. Domain Decomposition for Data Assimilation in Flow Forecasting Models, in: Realization and Modelling in System Theory, Progress in Systems and Control Theory. pp. 585–593.

van Leeuwen, P.J., 2010. Nonlinear Data Assimilation in Geosciences: an Extremely Efficient Particle Filter. Quarterly Journal of the Royal Meteorological Society 136, 1991–1999.

van Leeuwen, P.J., 2009. Particle Filtering in Geophysical Systems. Monthly Weather Review 137, 4089–4114.

van Leeuwen, P.J., Evensen, G., 1996. Data assimilation and inverse methods in terms of a probabilistic formulation. Monthly Weather Review 124, 2898–2913.

van Leeuwen, P.J., Vetra-Carvalho, S., Nerger, L., Heemink, A., van Velzen, N., Verlaan, M., Altaf, M.U., Beckers, J.-M., Barth, A., Brasseur, P., Brankart, J.-M., de Mey, P., Bertino, L., 2012. Non-linear data assimilation methods V0 report, SANGOMA project report.

Verlaan, M., Heemink, A.W., 1997. Tidal flow forecasting using reduced rank square root filters. Stochastic Hydrology and Hydraulics 11, 349–368.

Versteeg, H.K., Malalasekera, W., 1995. An Introduction to Computational Fluid Dynamics. The Finite Volume Method. Longman Scientific & Technical.

Wang, Y.-S., You, S., Matni, N., 2015. Localized Distributed Kalman Filters for Large-Scale Systems. 5th IFAC Workshop on Distributed Estimation and Control in Networked Systems.

Wiener, N., 1949. Extrapolation, interpolation, and smoothing of stationary time series. MIT press, Cambridge, MA.

Wikle, C., Berliner, M., 2006. A Bayesian Tutorial for Data Assimilation. Physica D.

Xianyi, Z., Qian, W., Saar, W., 2016. OpenBLAS: An optimized BLAS library. http://www.openblas.net/.

Yaesh, I., Shaked, U., 1992a. Game theory approach to optimal linear state estimation and its relation to the minimum H-infimum-norm estimation. IEEE Transactions on Automatic Control 37.

Yaesh, I., Shaked, U., 1992b. Game theory approach to state estimation of linear discrete-time processes and its relation to H-infimum-optimal estimation. International Journal of Control 55, 1443–1452.

Zheng, Z., Petzold, L., 2006. Runge–Kutta–Chebyshev projection method. Journal of Computational Physics 976–991.

Zheng, Z., Simeon, B., Petzold, L., 2008. A stabilized explicit Lagrange multiplier based domain decomposition method for parabolic problems. jcp 5272–5285.

Zhuang, Y., Sun, X.-H., 2002. Stabilized explicit-implicit domain decomposition methods for the numerical solution of parabolic equations. SIAM Journal of Scientific Computing 24, 335–358.

Zhuk, S., 2013. Kalman Duality Principle for a Class of Ill-Posed Minimax Control Problems with Linear Differential-Algebraic Constraints. Appl Math Optim 289–309.

Zhuk, S., 2009. State Estimation for a Dynamical System Described by a Linear Equation With Unknown Parameters. Ukrainian Mathematical Journal 61, 214–235.

Zupanski, M., 2009. Theoretical and practical issues of ensemble data assimilation in weather and climate, in: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, Editors: Park S.K., Xu L. Springer-Verlag Berlin Heidelberg.