Parameter Estimation in Random Energy Systems using Data Assimilation

Parameterschatting in onzekere energiesysteem met behulp van data-assimilatie

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op

maandag 20 januari 2020 des middags te 12.45 uur

door

Sangeetika Ruchi

geboren op 19 augustus 1989 te Dumra, India

Promotor: Prof. dr. ir. J. E. Frank
Copromotor: Dr. S. Dubinkina

This work is part of the Industrial Partnership Program (IPP) "Computational sciences for energy research" of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organization for Scientific Research (NWO). This research program is co-financed by Shell Global Solutions International B.V. The research work was carried out at Centrum Wiskunde & Informatica (CWI), the Dutch national institute for mathematics and computer science.

Printed by Ipskamp Printing

ISBN 978-90-393-7242-5

To the Singh family

ma, papa, bhai and tari di

Thanks for being the wind beneath my wings!

539367-L-sub01-bw-Ruchi
Processed on: 18-12-2019
PDF page: 5

Acknowledgements

First and foremost, I would like to thank my supervisors Prof. Jason Frank and Dr. Svetlana Dubinkina for their constant support and encouragement throughout my PhD journey. Svetlana, from guiding me through the intricacies of the field to honing me in a good researcher, I could not have asked for a better supervisor. Thank you for always being available for meetings and discussions, for ensuring the timely progress of the thesis, for being there when I felt directionless, and mostly for showing faith in me when I had doubts on myself.

I would like to express my sincere gratitude to Dr. Marco Iglesias from the University of Nottingham, England, for inviting me to Nottingham as a guest researcher. Marco, it was a pleasure working with you on our joint paper. Our discussions and interaction with other members of the research group were extremely insightful, it proved to be helpful in timely completion of my PhD.

I would also like to thank Prof. Sebastian Reich and Dr. Jana de Wiljes from University of Potsdam, Germany, for giving me the opportunity to share my work with their data assimilation group. Jana, it has been fun working with you on our paper. Thank you for organizing the dinner and dance plans, I miss spending time with SFB1294 research group, they seem to have a perfect balance of work and play.

Nada, from taking care of the administrative hassles to even the personal ones you have been a constant source of support. You are nothing short of a wizard :). Duda, thank you for all the fun conversations and for making CWI less gloomy. I changed 5 apartments and 3 offices during my 4 years of stay, but the last ones proved to be the perfect reflection of "save the best for last". My lovely landlady Diny Ulendorp, and my "lively" office mate Ki Wai, thank you so much for looking after me and for being my foster family here in Amsterdam. Ki your passion towards your work has always motivated me and I miss our WHINE and dine sessions. Willem, you might be the best person I know in Amsterdam, sharing things with you has always been so effortless. Thank you for being there and also for translating my thesis summary :)

Paco, Prashant and Krzysztof, thank you for the squash lessons and the sometimes

deep and mostly strange evening conversations over drinks. Debarati, it has always been fun hanging out with you, your take on even the mundane of things never fails to amuse me. Thanks for all those crazy and inappropriate conversation sessions :). Beatriz and Lara, the last few months of my PhD life have been less stressful thanks to you two. I would also like to thank my friends and colleagues at CWI: Bart, Anastasia, Kristoffer, Anne, Laurent, Daan, Wouter, Yous, Fredrik, Kees, Benjamin, Jurriaan, and Hemaditya for creating a friendly atmosphere at work.

I would not have moved to Amsterdam, if not for you Vikrant, there aren't enough words to thank you. You have been a friend, a guide, a teacher and you will always be like a family to me. Vijay, thank you for showing me what resilience looks like, you always kept motivating me with your tough choice of words. Akanksha and Deepti, thank you for bearing with me all these years, listening to me yap about any and everything. Priya, thanks for keeping me sane, bearing through all those midnight calls talking about relationships, work, our families and reminiscing the good old college times. Shweta, Snigdha, Vidit da, Harish, Kritarth, Vishal, Shubham, Parimit, Himneet: just knowing you guys are out there makes me feel better.

Luis, I wish you had started your PhD journey a little sooner :) Thanks for keeping me company while I was writing my thesis and agreeing to proof-read it, over and over again. Exploring Amsterdam with you has been fun, you made the last months of my PhD probably the easiest and the best ones.

Mamma, Papa it feels strange to thank you, after-all parenthood is a thankless job :). Thank you for never telling me what I should be doing and always motivating me to do what I want to do and what I want to be, thanks for having that faith in me. Bhai and Tari di, love you for being the annoying people you are. Thank you for always having my back, for giving me the confidence that even if I fall you will be there to catch me. I am so lucky to have you all in my life.

Contents

Chapter 1

Introduction

Some of the most important physical phenomena essential for sustainable growth in recent times are described by subsurface flow processes. Perhaps the most significant example is the efficient management of earth's depleting groundwater reserves. Successful management essentially requires modelling of groundwater hydrological systems, which yields the information imperative for groundwater remediation, helps monitor level of pollutants in the subsurface and quantify thresholds for sustainable use. Another example is CO_2 storage in the subsurface repositories. Such CO_2 storage helps to reduce emission of greenhouse gases into the atmosphere but introduces a potential danger of CO_2 leakage to habitated areas or into fresh-water aquifers. Thus, it is essential to estimate the leakage rate of CO_2 which depends on accurate subsurface flow models describing the geological properties. Some other examples highlighting the importance of an accurate mathematical modelling of subsurface flow are exploring deep geothermal energy, ensuring safe storage of nuclear waste, improving technologies to remove contaminants from the subsurface, etc. In this thesis, we focus on improving predictions of petroleum production for oil/gas reservoirs. With the reduced rate of new major discoveries and severity of environment degradation, optimal utilization of current production

fields and efficient exploration measures is of utmost importance to meet the world's energy demands. All of these examples require an accurate mathematical model of subsurface flow. This is a challenging task, since geological properties of the subsurface are often uncertain.

1.1 Subsurface Reservoir Modelling

An oil/gas reservoir represents a natural accumulation of hydrocarbons within different lithological structures. The oil and gas are collected in small, connected pore spaces of the rocks and trapped within the reservoir by surrounding impermeable layers of rocks. Figure 1.1 shows a pictorial representation of a typical reservoir model. Different colors represent layers with different geophysical properties and the vertical displacements of layer are termed faults. The upward arrows represent production wells, which are used to extract oil/gas from subsurface deposits, while the downward arrows represent injection wells. These are used to place fluid deep underneath the surface into porous rock formations, either for safe storage purposes or to maintain reservoir pressure. One of the main struggles of reservoir management is to check early water breakthrough, in which the fluid injected via the injection well to maintain the pressure comes out of the production well and effects the quality of hydrocarbon recovery.

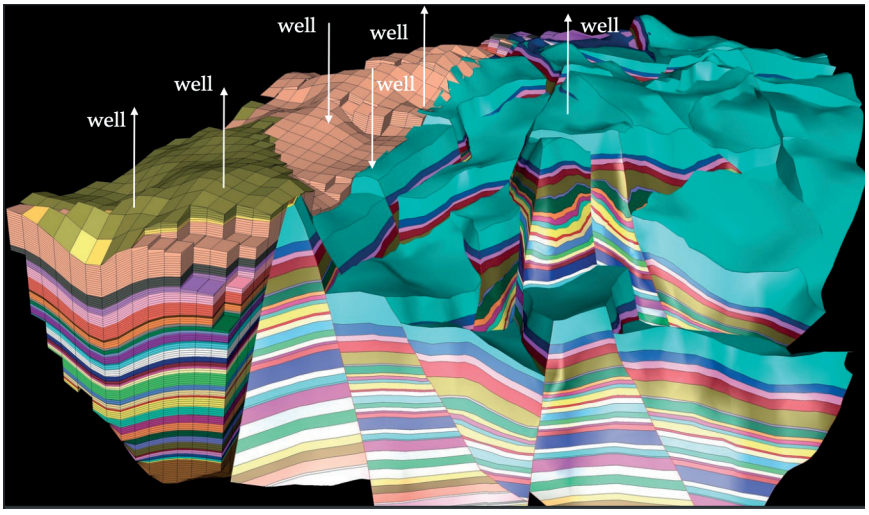

Figure 1.1: Pictorial representation of generic reservoir model.

By knowing the location of a reservoir, its structure and geological properties, we can develop a mathematical model which describes subsurface flow. Using a numerical approximation of the model, we can predict the reservoir performance. This is called solving a forward problem. However, even if the location of the reservoir is known, its structure and geological properties are often uncertain. This is because the subsurface reservoirs are buried thousands of feet below the earth surface and exhibit a highly heterogeneous structure. Usually prior information about the parameters is given, which still needs to be corrected by observations of pressure and production rates. This is called solving an inverse problem.

1.1.1 Forward Problem

If we denote the model parameter by a random variable $u \in \mathcal{U}$ then the forward model is simply presented as

$$z = \mathrm{F}(u), \tag{1.1}$$

where $z \in \mathcal{Z}$ is the model state. Here, $\mathrm{F} : \mathcal{U} \to \mathcal{Z}$ represents the set of differential equations forming the mathematical model which maps model parameter to model state, and $\mathcal{U}$ and $\mathcal{Z}$ are separable Banach spaces $(\mathcal{B}, ||\cdot||)$. The assumption of separability is necessary for a random variable to be integrable [21]. In this thesis, we consider a reservoir model defined over domain $\mathrm{D} \subset \mathbb{R}^2$ with Lipschitz boundary $\partial\mathrm{D}$. We represent the subset of strictly positive L^∞ functions on D by $\mathcal{X}$. A single phase steady state forward model is given as

$$-\nabla \cdot (k(x,y)\nabla P(x,y)) = f(x,y), \quad (x,y) \in D \tag{1.2}$$

where $\nabla = (\partial/\partial x \ \partial/\partial y)^T$, $\cdot$ denotes the dot product, $P(x,y)$ the pressure, $k(x,y)$ the permeability and $f(x,y)$ the source term. The forward problem of this second order elliptical equation is to find the solution of pressure $P(x,y)$ for given $f(x,y)$ and $k(x,y)$ under provided boundary conditions. We implement a cell-centered finite difference method to discretize the domain on a finite grid n^2, the details of which are shown later in Appendix A. Thus, the model state $P_{i,j}$, $i,j = 1,\ldots,n$ can be evaluated from equation (1.32), given in Appendix A, for a given stimulus $f_{i,j}$, $i,j = 1,\ldots,n$ and model parameters permeability $k_{i,j}$, $i,j = 1,\ldots,n$.

In 1902, J. Hadamard formulated the concept of well-posedness of problems [37].

According to Hadamard, a problem is well-posed if:

$\mathcal{C}_1$: **Existence.** There is at least one solution.

$\mathcal{C}_2$: **Uniqueness.** There is at most one solution.

$\mathcal{C}_3$: **Stability.** The solution depends continuously on parameters.

Under certain regularity conditions a forward problem is well-posed [28].

1.1.2 Inverse Problem

The inverse problem is concerned with finding u given an observation of the model state z. Let us denote an observation operator $L : \mathcal{Z} \to \mathcal{Y}$, then we can define observations as

$$y_{\text{obs}} = L(z) + \eta, \tag{1.3}$$

where $y_{\text{obs}} \in \mathcal{Y}$ is an observation, $\mathcal{Y}$ represents a subset of the separable Banach space $(\mathcal{B}, ||\cdot||)$ and $\eta \in \mathcal{Y}$ represents observational noise. The inverse problem can be formulated as finding u such that

$$y_{\text{obs}} = L(F(u)) + \eta. \tag{1.4}$$

If we define an operator $\mathcal{H}$ which maps the parameter space to the observation space, i.e. $\mathcal{H} : \mathcal{U} \to \mathcal{Y}$ by $\mathcal{H} = L \circ F$. Then, employing operator $\mathcal{H}$ we can rewrite the inverse problem as

$$y_{\text{obs}} = \mathcal{H}(u) + \eta. \tag{1.5}$$

In this thesis, u in equation (1.5) is parametrization of permeability k, and y_{obs} is observations of the pressure at a few well locations.

As for a forward problem, well-posedness of an inverse problem needs to be established. For an inverse problem the conditions $\mathcal{C}_1$ and $\mathcal{C}_2$ are identical to the ones for a forward problem. The condition $\mathcal{C}_3$, however, is

$\hat{\mathcal{C}}_3$: **Stability.** The solution depends continuously on data.

An inverse problem fails to satisfy at least one of the well-posedness conditions $\mathcal{C}_1, \mathcal{C}_2$ or $\hat{\mathcal{C}}_3$. Hence, unlike the forward problems which are mostly well-posed, the inverse problems are ill-posed. This makes the problem in hand challenging.

The formal study of ill-posed inverse problems started somewhere around the early 20th century. Early work of A. N. Tikhonov addresses the importance of inverse problems and presents regularization techniques to introduce stability in ill-posed problems which is famously known as the Tikhonov regularization [86–88]. The essence of the regularization techniques is to replace the ill-posed problem by a family of nearby well-posed problems. This is achieved by rewriting the inverse problem. A series of different regularization techniques have appeared since then, e.g. reducing the number of parameters to lessen the sensitivity towards data fluctuations, introducing constraints in order to obtain only physically viable values, etc. Over the years, geophysicists and engineers in their quest to understand the structure and internal behavior of Earth using sparse observations, contributed highly towards the development of various approaches for solving inverse problems. Around 1970, G. Backus and F. Gilbert introduced the reconstruction method, now known as Backus-Gilbert, which is frequently employed for the inversion of seismic data to achieve density profiles at the interior of the Earth. In [3, 4], these authors delve into the structure of the mathematical formulation of inverse problems which helps catapult the development of many methods of data interpretation in geosciences.

For groundwater and petroleum reservoirs in particular, numerous inverse problem approaches have been formulated. A. C. Reynolds and co-authors presented in [76] the implementation of the Gauss-Newton method for estimation of reservoir parameters. In [15, 82] authors showed the application of the neighborhood algorithm which approximates the posterior distribution by partitioning the model parameter space into regions of roughly uniform distribution. The authors of [72] implement a representer method to estimate states and parameters for a 1-D two-phase reservoir model. This approach expands all the unknown variables using a finite sum of representer functions weighted by the representer coefficients. The representer function describes the influence of the corresponding measurement on the solution and the representer coefficients determine how strong each representer should be accounted for in the final solution. In [90] researchers used the Levenberg-Marquardt method to characterize reservoir pore pressure and permeability. Another approach is presented in [50], which introduces a parameteric level-set technique to estimate subsurface properties using seismic data. They mitigate the ill-posedness of the problem by expanding the level-set function in some basis, hence reducing the number of parameters characterizing the subsurface. Thus, replacing the original least-square problem with a better behaving optimization problem. In [32, 92]

the authors reviewed methods for solving the output least-squares problem using gradient and sensitivity-based techniques. However, in this thesis we focus on the Bayesian approach to inverse problems, which has been pioneered by Andrew Stuart [21]. In the next section we describe the Bayesian approach to inverse problems in detail and discuss its role in alleviating the ill-posedness of the problem.

1.2 Bayesian Framework

The Bayesian approach determines the uncertain parameter u by evaluating the probability distribution of u given observation y_{obs}, denoted $u|y_{\text{obs}}$, and referred to as the posterior distribution. Let us consider u to be a random variable and define our prior knowledge about u in terms of a probability measure μ_0 on a measurable space $(\mathcal{U}, \mathcal{F})$, where $\mathcal{F}$ denotes the σ-algebra on $\mathcal{U}$. Assume observational noise η is an independent random variable with Gaussian distribution $\mathcal{N}(0, R)$ denoted by $\mathbb{Q}_0$. Therefore, using equation (1.5) the conditional distribution of $y_{\text{obs}}|u \sim \mathcal{N}(\mathcal{H}(u), R)$. Thus, we can define the negative log likelihood as

$$l(u; y_{\text{obs}}) = \frac{1}{2}||y_{\text{obs}} - \mathcal{H}(u)||_R^2. \tag{1.6}$$

Define ν to be a probability measure on $\mathcal{U} \times \mathcal{Y}$ as $\nu(du, dy) = \mu_0(du)\mathbb{Q}_0(dy)$. Bayes' theorem implies the existence of the conditional distribution $u|y_{\text{obs}}$.

Theorem 1. (Theorem 14 from [21]) Assume that $l : \mathcal{U} \times \mathcal{Y} \to \mathbb{R}$ is ν measurable and that, for y_{obs} $\mathbb{Q}_0$-almost surely (a.s.),

$$Z = \int_{\mathcal{U}} \exp(-l(u; y_{\text{obs}}))\mu_0(u)\, du > 0. \tag{1.7}$$

Then the conditional distribution of $u|y_{\text{obs}}$ exists and is denoted by μ. Furthermore, μ is absolutely continuous with respect to μ_0, i.e. $\mu \ll \mu_0$ and, for y_{obs} $\mathbb{Q}_0$-a.s.

$$\frac{d\mu}{d\mu_0}(u) = \frac{1}{Z}\exp(-l(u; y_{\text{obs}})). \tag{1.8}$$

To ensure the well-posedness of the posterior distribution, the key ingredient is the continuity of $l(\cdot; y_{\text{obs}})$, which in turn depends on the forward map F. From Section 1.1.1 we see that if $u \in \mathrm{L}^\infty(\mathrm{D})$ with lower bound $k_{\min} > 0$ for every $f \in H^{-1}(\mathrm{D})$, then

according to Lax-Milgram theory there exists a unique solution $P \in H_0^1(\mathrm{D})$ to expression (1.2), satisfying

$$||P||_{H_0^1} \leq \frac{1}{k_{\min}} ||f||_H^{-1} \tag{1.9}$$

where $H = (\mathrm{L}^2(\mathrm{D}), \langle \cdot, \cdot \rangle, ||\cdot||)$ is the Hilbert space and $H_0^1(\mathrm{D})$ is Sobolev space [7, 21]. This setting ensures that the forward map is well-defined, see [21]. Further, a comprehensive study of the continuity property of the posterior distribution demonstrates that for the Bayesian posterior distribution small changes in the observation y_{obs} lead to only small changes in μ, proving well-posedness [21, 81]. That is, let us denote an approximation of $\mathcal{H}(u)$ by $\mathcal{H}'(u)$, and denote the resulting modified posterior measure from μ to μ'. Then Bayesian approach ensures that under some assumptions a small change in forward error leads to a small change in inverse error. That is

$$|\mathcal{H}(u) - \mathcal{H}'(u)| = O(\delta) \implies d(\mu, \mu') = O(\delta),$$

for small enough $\delta > 0$ and some metric $d(\cdot, \cdot)$ on probability densities [80]. In this thesis, we have considered parametrization of both continuous and discontinuous permeability fields. The latter one is typical for faults or channels in the subsurface. Well-posedness of these parametrizations has been studied in [21, 46, 81]

In the next section we discuss numerical methods to approximate equation (1.8).

1.3 Data Assimilation

Data assimilation in essence minimizes the square of the mismatch between the observations and the model state to give more accurate estimations. Though data assimilation had originally been employed in the field of meteorology and oceanography for state prediction by correcting initial conditions [35], now it is one of the frequently employed approaches for parameter estimation in subsurface flow models as well [67], since it can easily be extended to estimate model parameters by implementing the method of state augmentation [48]. The idea here is to expand the state space by adding the uncertain parameters so that parameters act as the model variables in the data assimilation system and the chosen assimilation algorithm can then be applied to the augmented system in a usual way.

Considerable efforts have been devoted to develop robust data assimilation tech-

niques which could provide better assessment of state at affordable computational costs. There are mainly two classes of methods in data assimilation: the variational approach [83], which is based on minimization of an appropriate cost functional subjected to model constraints providing a single estimate; and the ensemble approach, which provides an ensemble of estimates whose variability can directly be used to evaluate the uncertainty. In this thesis, we focus on the ensemble-based approach.

Among the ensemble data assimilation methods, the ensemble Kalman filter (EnKF) is the most employed approach for parameter estimation in subsurface flow models. It was introduced by Evensen in 1994 [29] and was developed to handle non-linear models at low computational cost. EnKF updates the model states based on the assumption of Gaussian distribution, as it corrects only the mean and the variance. Initial application of EnKF in the field of groundwater hydrology is used for soil moisture estimation and is demonstrated in [75]. In petroleum engineering, [55] presented the first implementation of the ensemble Kalman filter using a dynamic two-phase model of fluid flow in a well. Since then it has been investigated by a variety of researchers in subsurface flow models, e.g. [13, 34, 67]. A detailed review of the application of ensemble Kalman filter developments in reservoir modeling can be found in [1]. As shown in these papers, the resulting updated ensemble for ensemble Kalman filters efficiently approximates the theoretical posterior distribution if the ensemble size is sufficiently large and the distributions are not far from Gaussian. However, the update equations are of the form of the Kalman filter equation that corrects only the mean and covariance matrix, which limits the performance of EnKF for models with multimodal distributions, see [25, 94].

On the other hand, particle filters [24], otherwise known as sequential Monte Carlo methods, prove to be promising for such physical systems. They are also ensemble based methods in which the probability density function is represented by a number of particles (also called samples or ensemble members) and their evolution is computed by solving the forward model for each particle. These particles are then assigned weights based on the information present in the observations of the true physical system. However, particle filtering in its original form faces the issue of ensemble collapse and also struggles to represent the actual state of the physical systems if a majority of the particles are far away from the observation. Due to these reasons particle filters were earlier assumed to be impractical for high-dimensional systems as they required a large number of particles making the process computationally very expensive. However, recent developments help to overcome these shortcomings by generating computationally efficient particle filtering

methods for high-dimensional systems as discussed in details in the review paper [89]. In this thesis, we introduce novel strategies to overcome the limitations of particle filters with the aim of approximating model parameters of different nonlinear models with strongly non-Gaussian posterior.

1.3.1 Markov chain Monte Carlo

The quantity of interest in this thesis remains the best representation of the posterior distribution of the model parameter. If we have knowledge about the target (posterior) distribution and it is rather straightforward to sample from it, then the idea of Monte Carlo simulation is simply to draw an i.i.d.set of samples $\{u_m\}_{m=1}^M$ and estimate expectations by sample averages. These M samples can be used to approximate the distribution with the following empirical point-mass function

$$\mu^M = \sum_{m=1}^{M} \delta_{u_m},$$

which converges to μ as $M \to \infty$. Here δ_u denotes the Dirac measure at u. The Monte Carlo approach is beneficial when generating random samples is easier than solving the distribution's equations to evaluate any statistical property of interest. However, it cannot be used when the distribution is too complex to generate independent samples, as often happens in high-dimensional systems. In this case, combining the Markov chain property with Monte Carlo (MCMC) has proved to be highly advantageous, see [81]. The Markov chain property ensures that every sample in a chain depends only on a previous sample. It allows the random samples to be generated in a sequential manner. The basic idea of MCMC methods is to simulate a long run of an ergodic Markov chain that is μ-invariant, and after a few iterations (burn-in) the samples of the chain can be treated as approximate samples from μ.

There are many variants of MCMC methods tuned according to the particular structure of the desired target distribution. Here, we will focus on a particular class of MCMC methods known as Metropolis-Hastings (MH) methods. An MH method first proposes a move from an arbitrary transition kernel $\mathcal{Q}$. Then it accepts or rejects the move based on the criterion which depends on the target distribution. We specifically focus on the MH approach which is designed for finite high dimensional problems stemming from the discretization of infinite dimensional systems. Otherwise, if the MH method

is tailored categorically for finite dimensional problems it might not work in the infinite dimensional limit. The reason for this is that the measures in infinite dimensions have a tendency to be mutually singular. This becomes an issue for MH methods, since their acceptance probability depends on the existence of the Radon-Nikodym derivative between two measures. Therefore, it is beneficial to design algorithms for the infinite dimensional setting, which can be expected to perform satisfactorily under refinement of finite dimensional approximations [81].

Let us assume we have a proposal kernel $\mathcal{Q}$ from which it is easy to sample. We however need to sample from a $\mathcal{P}$ kernel that is invariant under the measure μ.

Definition: Let $\mathcal{P}(u, dv)$ denote a Markov transition kernel so that $\mathcal{P}(u, \cdot)$ is a probability measure on $(\mathcal{U}, \mathcal{F})$ for each $u \in \mathcal{U}$. This transition kernel $\mathcal{P}$ will be invariant with respect to the target density μ if

$$\int_{\mathcal{U}} \mu(du)\mathcal{P}(u, \cdot) = \mu(\cdot),$$

with shorthand $\mu\mathcal{P} = \mu$ as measures on $(\mathcal{U}, \mathcal{F})$, where the integral on lefthand side is with respect to du.

A sufficient condition for μ to be an invariant measure for $\mathcal{P}$ is the detailed balance condition between $\mathcal{P}$ and μ. The Markov transition kernel $\mathcal{P}$ relates to the proposal kernel $\mathcal{Q}$ as

$$\mathcal{P}(u, dv) = \mathcal{Q}(u, dv)a(u, v) + \delta_u(dv) \int_{\mathcal{U}} (1 - a(u, w))\mathcal{Q}(u, dw), \tag{1.10}$$

where a denotes the acceptance probability and should be chosen such that $\mathcal{P}(u, dv)$ satisfies the essential condition of detailed balance with respect to μ.

Theorem 2. (Theorem 21 from [21]): We define the measures on $\mathcal{U} \times \mathcal{U}$ with proposal kernel $\mathcal{Q}$ and target density μ as

$$\begin{aligned} \nu(du, dv) &= \mathcal{Q}(u, dv)\mu(du), \\ \nu^T(du, dv) &= \mathcal{Q}(v, du)\mu(dv), \end{aligned} \tag{1.11}$$

with $\nu^T(du, dv) = \nu(dv, du)$, $\forall u, v \in \mathcal{U}$. We further assume that ν and ν^T are equivalent as measures on $\mathcal{U} \times \mathcal{U}$, and that $\nu(du, dv) = r(u, v)\nu^T(du, dv)$. Then $\mathcal{P}$ satisfies detailed

balance if and only if

$$r(u,v)a(u,v) = a(v,u),$$

thus, defining $a(u, u_{\text{prop}})$ as

$$\begin{aligned} a(u, u_{\text{prop}}) &= \min\{1, r(u, u_{\text{prop}})\} \\ &= \min\left\{1, \frac{d\nu^T}{d\nu}(u, u_{\text{prop}})\right\}, \end{aligned} \tag{1.12}$$

implies detailed balance, see [21, 85] for details.

Furthermore, we assume that the Markov chain is ergodic. Therefore after n_μ number of moves the total variation distance between $\mathcal{P}^{n_\mu}(u, \cdot)$ and μ approaches zero [84]. Hence, after a sufficiently long burn-in period we generate samples which are approximately distributed according to μ. One of the popular choices of the proposal kernel $\mathcal{Q}$ is the classical random walk algorithm in which the difference between the current and proposed states is taken to be a centered Gaussian distribution. Hence, if the prior measure μ_0 is Gaussian $\mathcal{N}(0, C)$, then according to the random walk algorithm $\mathcal{Q}(u, \cdot) = \mathcal{N}(u, \rho^2 C)$, i.e.

$$u_{\text{prop}} = u + \rho\zeta, \quad \text{where} \quad \zeta \sim \mathcal{N}(0, C). \tag{1.13}$$

However, this proposal fails to fulfill the absolute continuity condition $\nu^T \ll \nu$, and hence the MH algorithm degenerates under mesh refinement [81]. In this thesis we use the preconditioned Crank-Nicolson (pCN) algorithm of [17] which tackles this issue by slightly modifying the proposal of the random walk algorithm as

$$u_{\text{prop}} = \sqrt{1-\rho^2}u + \rho\zeta, \quad \text{where} \quad \zeta \sim \mathcal{N}(0, C). \tag{1.14}$$

If $u \sim \mathcal{N}(0, C)$, then from (1.13) $v \sim \mathcal{N}(0, (1+\rho^2)C)$, which does not preserve the underlying reference measure μ_0. If the proposal density is constructed using pCN (1.14), then the reference measure μ_0 is preserved, since $u_{\text{prop}} \sim \mathcal{N}(0, C)$. It is important to note that since the posterior has a density with respect to the infinite-dimensional prior Gaussian measure μ_0, constructing a proposal as in (1.14), which is reversible with respect to this measure, leads to the necessary condition of absolute continuity between ν^T and ν.

The parameter ρ in (1.13) and (1.14) controls the size of the move. Its value is adjusted so that the acceptance rate lies within $20 - 35\%$. It can be observed that smaller values of ρ produce proposals that do not move far from the current state and hence the acceptance is higher. This leads to a slow convergence of a chain as the chain moves slowly through the state space. On the other hand, large values of ρ lead to proposals that are unlikely to be accepted. Therefore, we need to construct moves that are large enough to probe the state space efficiently, while at the same time to keep acceptance rates high. Identifying appropriate values of ρ between these extremes is the key to making effective algorithms.

Although MCMC methods provide much more accurate solutions compared to other Bayesian data assimilation approaches, this accuracy comes at an extremely higher computational expense. Therefore, their applications are mostly limited to providing benchmarks for evaluating practically viable data assimilation algorithms.

1.3.2 Sequential Monte Carlo

Sequential Monte Carlo (SMC) is a recursive Monte Carlo method which provides a computationally cheaper technique to approximate the sequence of Bayesian posteriors. SMC methods came into existence to handle filtering in dynamical systems. However, with time they also found their ground in parameter estimation problems. The idea behind the standard sequential Monte Carlo approach uses the Bayes's theorem (Theorem 1) to transform a set of particles that approximate the prior μ_0 into a set of particles that approximate the posterior μ. This is done by means of an *importance sampling* step with μ_0 being the proposal distribution and μ the target distribution.

Importance Sampling

Our aim is to generate samples from the posterior distribution, but it is infeasible to directly sample from this distribution because it is unknown. The importance sampling approach [60] proposes that we generate samples from a prior distribution that is known.

Assume we have M particles (i.e. samples) of the model parameter $\{u_0^m\}_{m=1}^M$ drawn independently from the proposal or prior distribution μ_0, which approximate μ_0 as

$$\mu_0^M = \frac{1}{M} \sum_{m=1}^{M} \delta_{u_0^m}. \tag{1.15}$$

Then by employing the Theorem 1, available observations and log likelihood (1.6), we obtain an approximation of the posterior distribution μ as

$$\mu^M = \sum_{m=1}^{M} w^m \delta_{u_0^m}, \tag{1.16}$$

where so-called importance weights are

$$w^m = \frac{\exp(-l(u_0^m; y_{\text{obs}}))}{\sum_{s=1}^{M} \exp(-l(u_0^s; y_{\text{obs}}))}. \tag{1.17}$$

Approximation of the normalization factor Z as defined in Bayes' theorem (Theorem 1) is

$$Z \sim \sum_{m=1}^{M} \exp(-l(u_0^m; y_{\text{obs}})). \tag{1.18}$$

Thus, the importance sampling does not change the samples but changes the probability, from which they are sampled. This is the major drawback of the importance sampling. For a successful application of the importance sampling approach the posterior distribution μ and the prior distribution μ_0 should be reasonably close. If μ and μ_0 are not sufficiently similar, this algorithm performs poorly. In such situations the algorithm shows high variance in weights, i.e. a few particles have large weights, while all the other particles have weights close to zero. This is called *weight degeneracy*. Hence, the entire statistics of the system gets influenced by those few particles, which might be far away from the true state and hence provide poor approximation of the posterior distribution.

Iterative sequential Monte Carlo

An iterative approach tackles the weight degeneracy issues by constructing a smooth transition between the measures μ_0 and μ using a sequence of intermediate measures $\{\mu_t\}_{t=0}^{T}$ such that $\mu_{t=0} = \mu_0$ and $\mu_T = \mu$. To specify the iteration, we need to define the intermediate measures $\{\mu_t\}_{t=0}^{T}$ and the total number of iterations T. We first focus on defining the intermediate measures.

As these measures are mutually absolutely continuous, we can define the transition

between consecutive measures by employing (1.8)

$$\frac{d\mu_{t+1}}{d\mu_t}(u) = \frac{Z_t}{Z_{t+1}} \exp\left(-\frac{1}{T} l(u; y_{\text{obs}})\right), \tag{1.19}$$

where

$$\frac{d\mu_t}{d\mu_0}(u) = \frac{1}{Z_t} \exp\left(-\frac{t}{T} l(u; y_{\text{obs}})\right), \tag{1.20}$$

and

$$Z_t = \int \exp\left(-\frac{t}{T} l(u; y_{\text{obs}})\right) \mu_0(u) du.$$

Now, for convenience let us denote the likelihood for transition between consecutive measures by $\mathcal{L} = \exp(-l(u; y_{\text{obs}})/T)$.

Further, let us denote the μ_t-invariant Markov kernel by $\mathcal{T}_t$ which produces new proposals. An example of the Markov kernel $\mathcal{T}_t$ is an identity map. This choice results in unchanged particles as t changes and thus a classical particle filter [21]. In this thesis we use the Markov kernel of pCN-MCMC (1.14). Thus the particles change as t changes. Then we can write down the measure evolution from μ_t to μ_{t+1} as

$$\hat{\mu}_{t+1} = \mathcal{T}_t \mu_t, \tag{1.21}$$

$$\mu_{t+1} = \mathcal{L} \hat{\mu}_{t+1}. \tag{1.22}$$

Since, $\mathcal{T}_t$ preserves μ_t we can write

$$\frac{d\mu_{t+1}}{d\hat{\mu}_{t+1}}(u) = \frac{Z_t}{Z_{t+1}} \exp\left(-\frac{1}{T} l(u; y_{\text{obs}})\right). \tag{1.23}$$

Theorem 3. (Theorem 23 from [21]): We assume that the negative log likelihood $l(u; y_{\text{obs}})$ is bounded above and below for $\forall u \in \mathcal{U}$, i.e. there exists $q \in (0, 1)$ such that

$$q < \exp(-l(u; y_{\text{obs}})/T) < q^{-1}. \tag{1.24}$$

Then a distance metric $d(\mu_T^M, \mu_T)$ between measures μ_T^M and μ_T is

$$d(\mu_T^M, \mu_T) \leq \sum_{t=1}^{T} (2q^{-2})^t \frac{1}{\sqrt{M}}. \tag{1.25}$$

The above condition ensures that as $M \to \infty$, SMC converges to the true posterior

measure μ. A more detailed study of convergence of SMC methods is given by D. Crisan and A. Doucet in [18].

In iterative SMC we can predefine or tune the total number of iterations T and have a uniform step in the measure update (1.22), e.g. [27, 79]. An iterative SMC method runs as following: we approximate μ_t using M particles $\{u_t^m\}_{m=1}^M$ with importance weights $\{w_t^m\}_{m=1}^M$. The approximation μ_t^M is defined as

$$\mu_t \approx \mu_t^M = \sum_{m=1}^{M} w_t^m \delta_{u_t^m}, \tag{1.26}$$

The particles $\{u_t^m\}_{m=1}^M$ are modified into new particles $\{\hat{u}_t^m\}_{m=1}^M$ from the Markov kernel $\mathcal{T}_t$ according to (1.21). This forms an approximation of $\hat{\mu}_{t+1}$ which acts as the prior distribution in (1.23). Then an approximation of μ_{t+1} is

$$\mu_{t+1} \approx \mu_{t+1}^M = \sum_{m=1}^{M} w_{t+1}^m \delta_{\hat{u}_{t+1}^m}, \tag{1.27}$$

where

$$w_{t+1}^m = \frac{\exp(-l(\hat{u}_{t+1}^m; y_{\text{obs}})/T) w_t^m}{\sum_{m=1}^M \exp(-l(\hat{u}_{t+1}^m; y_{\text{obs}})/T) w_t^m} = \frac{\hat{w}_{t+1}^m}{\sum_{m=1}^M \hat{w}_{t+1}^m}. \tag{1.28}$$

At this stage a resampling is often required in an algorithm: when particles with low weights are abandoned and particles with high weights are dublicated. The higher the weight of a particle the more copies are generated, in order to restore the total number of particles M. There are different means of performing resampling, e.g. multimonial resampling, residual resampling, etc. However, in this thesis we advocate a resampling based on solving an optimal linear transport problem [73]. This approach to resampling maximizes correlations between particles and as a result requires much fewer samples to approximate the posterior. It has been successfully applied in dynamical systems, e.g. see [73, 74]. In this thesis we investigate its applicability for parameter estimation in elliptic inverse problems. In Section 2.2.2 we explain the method in more detail.

Adaptive sequential Monte Carlo

In this thesis, we select the sequence of measures $\{\mu_t\}_{t=0}^T$ *adaptively* based on the annealing approach of [64] as implemented in [52]. This is known as adaptive SMC. The idea of simulated annealing was first proposed in [53] exploiting the analogy be-

tween optimization and the physical process of annealing. The annealing approach presented in [64] and [24] bridges the sequence of measure using tempering parameters $0 = \phi_0 < \phi_1 ... < \phi_T \leq 1$. We evaluate these parameters based on Effective Sample Size (ESS), defined as

$$\text{ESS} = \frac{1}{\sum_{m=1}^{M} (w^m)^2}. \tag{1.29}$$

ESS takes a value between 1 and M. The parameters are found such that ESS remains above a predefined threshold limit. This allows to mitigate the weight degeneracy issue. Once the tempering parameter ϕ_t is equal to one, the iterative algorithm stops. A potential problem of the annealing approach is that the total number of iterations T is not predefined. For high-dimensional problems it can be very high due to high variance in weights. In [7], the authors present an application of an adaptive SMC method which employs tempering to deal with weight degeneracy. They show that an adaptive SMC with MCMC mutation step provides encouraging results for high-dimensional problems. We discuss adaptive SMC in more detail in Section 3.3.1.

Overview of thesis

In Chapter 2, we consider parameter estimation for a low-dimensional and a high-dimensional nonlinear problem. We present a particle filtering method which employs an optimal linear transport problem to introduce deterministic resampling. We examine the performance of the Ensemble Transform Particle Filter and Ensemble Transform Kalman Filter for estimating uncertain parameters and compare them to importance sampling.

In Chapter 3, we introduce optimal transport-based resampling in adaptive SMC. We consider two nonlinear problems, one with Gaussian posterior distribution and another with multimodal posterior distribution. We compare performance of the novel optimal transport-based adaptive SMC with ensemble Kalman inversion with mutation, while employing MCMC as a benchmark.

In Chapter 4, we consider a more realistic set-up, where uncertainty is introduced in the model but we do not account for it in data assimilation. We employ MCMC as the benchmark and compare performance of optimal transport-based adaptive SMC with regularized ensemble Kalman filter. The regularized Kalman filter is an iterative ensemble Kalman filter that implements discrepancy principle to control iterations and

prevents overfiting of observations.

The Chapters 2, 3 and 4 are based on the publications 1, 2 and 3, respectively, mentioned on the List of Publications 5. I apologize to the reader for the inconsistency in notation.

Appendix A

Cell-centered discretization

We employ a cell-centered finite difference method for discretization by implementing the grid point cluster shown below in figure 1.2, see e.g. [68].

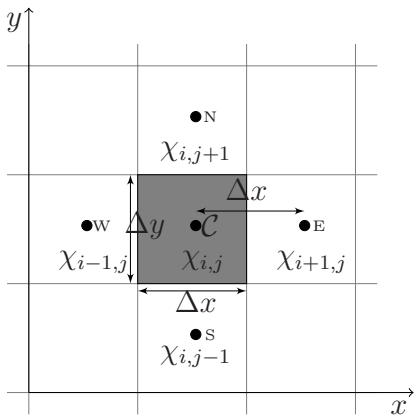

Figure 1.2: Grid point cluster

We divide the entire domain D uniformly ($\Delta x = \Delta y$) in $n \times n$ axis-parallel cells of size Δx^2. Each of these cells is denoted by $\chi_{i,j}$, $i, j = 1, \ldots, n$, and has a value of pressure $P_{i,j}$ at its center. For the shaded cell $\chi_{i,j}$ as shown in Figure 1.2, the grid point $\mathcal{C}$ is the interior point as it is surrounded by cells $\chi_{i+1,j}$, $\chi_{i-1,j}$, $\chi_{i,j+1}$ and $\chi_{i,j-1}$ in all directions. For $\mathcal{C}$ the grid point in its east side, E, has a value of pressure $P_{i+1,j}$, the grid point W in west side has a value of pressure $P_{i-1,j}$, the grid point N on the north side has a value of pressure $P_{i,j+1}$ and the grid point S in south has a value of pressure $P_{i,j-1}$. Integrating (1.2) by parts for cell $\chi_{i,j}$ and replacing the normal derivative on the edges by

$$\left(\frac{\partial P}{\partial x}\right)_{\mathcal{C},E} = \frac{P_{i+1,j} - P_{i,j}}{\Delta x}, \quad \left(\frac{\partial P}{\partial x}\right)_{\mathcal{C},W} = \frac{P_{i,j} - P_{i-1,j}}{\Delta x},$$
$$\left(\frac{\partial P}{\partial y}\right)_{\mathcal{C},N} = \frac{P_{i,j+1} - P_{i,j}}{\Delta y}, \quad \left(\frac{\partial P}{\partial y}\right)_{\mathcal{C},S} = \frac{P_{i,j} - P_{i,j-1}}{\Delta y},$$

gives the finite difference equation for the interior points as

$$\begin{aligned} &- k_{i-1/2,j}(P_{i-1,j} - P_{i,j}) - k_{i+1/2,j}(P_{i+1,j} - P_{i,j}) \\ &\qquad - k_{i,j-1/2}(P_{i,j-1} - P_{i,j}) - k_{i,j+1/2}(P_{i,j+1} - P_{i,j}) = f_{i,j}\Delta x^2. \end{aligned} \tag{1.30}$$

Here $k_{i-1/2,j} = k(x_{i-1/2}, y_j)$ is the value of permeability on the edge between grid points W and $\mathcal{C}$. If k is discontinuous along the interface we take $k_{i-1/2,j}$ as the harmonic mean of the neighboring grid points as

$$k_{i-1/2,j} = \frac{2k_{i-1,j}k_{i,j}}{k_{i-1,j} + k_{i,j}}, \tag{1.31}$$

and similarly $k_{i+1/2,j}$, $k_{i,j-1/2}$ and $k_{i,j+1/2}$ are defined. The given boundary conditions are used with the discretized equation (1.30) to derive finite difference equations for the grid points near boundaries, and finally end up with a set of linear equations of the form

$$\mathcal{K}\mathbf{P} = \mathbf{F}, \tag{1.32}$$

where $\mathcal{K}$ is a sparse, symmetric, positive definite matrix of size $n^2 \times n^2$, $\mathbf{P}$ is a vector with entries of $P_{i,j}$, $i, j = 1, \ldots, n$ and $\mathbf{F}$ is a vector with entries $f_{i,j}$ for $i, j = 1, \ldots, n$.

Chapter 2

Ensemble transform data assimilation for inverse problems

2.1 Introduction

An accurate estimation of subsurface geological properties like permeability, porosity etc. is essential for many fields specially where such predictions can have large economic or environmental impact, for instance prediction of oil or gas reservoir locations. Knowing the geological parameters a so-called forward model is solved for the model state and a prediction can be made. The subsurface reservoirs, however, are buried thousands of feet below the earth surface and exhibit a highly heterogeneous structure, which makes it difficult to obtain their geological parameters. Usually prior information about the parameters is given, which still needs to be corrected by observations of pressure and production rates. These observations are, however, known only at well locations that are often hundreds of meters apart and corrupted by errors. This gives instead of a well-posed forward problem an ill-posed inverse problem of estimating uncertain parameters,

since many possible combinations of parameters can result in equally good matches to the observations.

Different inverse problem approaches for groundwater and petroleum reservoir modelling, generally termed as history matching, have been developed over the past years, e.g. in [67] the authors implemented Markov chain Monte Carlo methods with different perturbations and tested it on a 2-D reservoir model; [90] used the Levenberg–Marquardt method to characterize reservoir pore pressure and permeability. A review of history matching developments is written in the review paper [66].

For reservoir models the term data assimilation and history matching are used interchangeably, as the goal of data assimilation is the same as that of history matching, where observations are used to improve a solution of a model. Ensemble data assimilation methods such as Ensemble Kalman filters [30] have been originally developed in meteorology and oceanography for the state estimation. Now it is one of the frequently employed approaches for parameter estimation in subsurface flow models as well [e.g. 67]. A detailed review of ensemble Kalman filter developments in reservoir engineering can be found in [1]. An ensemble Kalman filter efficiently approximates a true posterior distribution if the distribution is not far from Gaussian, as it corrects only the mean and the variance. For nonlinear models with multimodal distributions, however, an ensemble Kalman filter fails to correctly estimate the posterior, as shown in [25].

Importance Sampling (IS) is quite promising for such models as it does not have any assumptions of Gaussianity. It is also an ensemble based method in which the probability density function is represented by a number of samples. One sample corresponds to one configuration of uncertain model parameters. The forward model is solved for each sample and predicted data is computed. The weight is assigned to samples based on the observations of the true physical system and the predicted data. The drawback of IS is that it does not update the uncertain parameters but only their weight, thus a computationally unaffordable ensemble is required. In order to decrease this cost a family of particle filters [24] has been developed where IS is supplied with resampling, and each sample is called a 'particle'. Significant work on parameter estimation using particle filtering has been done in hydrology. In [63] authors used it to estimate model parameters and state posterior distributions for a rainfall-runoff model. [93] compared an ensemble Kalman filter and a particle filter with different resampling strategies for a rainfall-runoff forecast and observed that as the number of particles increases the particle filter outperforms the ensemble Kalman filter. [36] employed particle filtering to correct

the soil moisture and to estimate hydraulic parameters.

The resampling in particle filtering is, however, stochastic. The Ensemble Transform Particle Filter (ETPF) [74] is a particle filtering method that *deterministically* resamples the particles based on their weights and covariance maximization among the particles. ETPF has been used for initial condition estimations and for parameter estimations in chaotic dynamical systems with a small number of uncertain parameters (Lorenz 63 model). It has not been applied, however, in subsurface reservoir modelling for estimating a large number of uncertain parameters. In this chapter we employ it for estimating uncertain parameters in subsurface reservoir modelling. ETPF provides the equations that are solved in the space defined by the ensemble members. Therefore for comparison we employ the Ensemble Transform Kalman Filter (ETKF) [8] that also transforms the state from the model space to the ensemble space, minimises the uncertainty in the ensemble space and transforms the estimation back to the model space.

In this chapter we investigate the performance of ETPF and ETKF for parameter estimation in nonlinear problems and compare them to IS with a large ensemble. This chapter is organized as follows: in Sect. 2.2 we describe IS, ETPF, and ETKF for parameter estimation. We apply these methods in Sect. 2.3 to a one parameter nonlinear test case, where the posterior can be computed analytically, and in Sect. 2.4 to two cases of single-phase Darcy flow, where the number of parameters is 5 and 2500, respectively. In Sect. 2.5 we draw the conclusions.

2.2 Data assimilation methods

We implement an ensemble transform Kalman filter and an ensemble transform particle filter for estimating parameters of subsurface flow. Both of these methods are based on a Bayesian framework. Assume we have an ensemble of M model parameters $\{\boldsymbol{u}_m\}_{m=1}^{M}$, then according to this framework, the posterior distribution, which is the probability distribution $\pi(\boldsymbol{u}_m|\boldsymbol{y}_{\text{obs}})$ of the model parameters $\boldsymbol{u}_m$ given a set of observations $\boldsymbol{y}_{\text{obs}}$, can be estimated by the pointwise multiplication of the prior probability distribution $\pi(\boldsymbol{u}_m)$ of the model parameters $\boldsymbol{u}_m$ and the conditional probability distribution $\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{u}_m)$ of the observations given the model parameters, which is also referred as the likelihood

function,

$$\pi(\boldsymbol{u}_m|\boldsymbol{y}_{\text{obs}}) = \frac{\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{u}_m)\pi(\boldsymbol{u}_m)}{\pi(\boldsymbol{y}_{\text{obs}})}.$$

The denominator $\pi(\boldsymbol{y}_{\text{obs}})$ represents the marginal of observations and can be expressed as:

$$\pi(\boldsymbol{y}_{\text{obs}}) = \sum_{m=1}^{M} \pi(\boldsymbol{y}_{\text{obs}}, \boldsymbol{u}_m) = \sum_{m=1}^{M} \pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{u}_m)\pi(\boldsymbol{u}_m),$$

which shows that $\pi(\boldsymbol{y}_{\text{obs}})$ is just a normalisation factor.

2.2.1 Ensemble Transform Kalman Filter

Assume we have initially an ensemble of M model parameters $\{\boldsymbol{u}_m^b\}_{m=1}^M$, where b refers to a background (prior) ensemble, which are sampled from a chosen prior probability density function, then the ensemble Kalman estimate (or analysis) $\{\boldsymbol{u}_m^a\}_{m=1}^M$ is given by:

$$\boldsymbol{u}_m^a = \sum_{l=1}^{M} \left(s_{lm} + q_l - \frac{1}{M} \right) \boldsymbol{u}_l^b, \quad m = 1, \ldots, M,$$

where s_{lm} is the (l, m) entry of a matrix $\mathbf{S}$

$$\mathbf{S} = \left[\mathbf{I} + \frac{1}{M-1} (\mathbf{A}^b)^T \mathbf{R}^{-1} \mathbf{A}^b \right]^{-1/2}, \tag{2.1}$$

and q_l is the l-th entry of a column vector $\boldsymbol{q}$

$$\boldsymbol{q} = \frac{1}{M} \mathbf{1}_M - \frac{1}{M-1} \mathbf{S}^2 (\mathbf{A}^b)^T \mathbf{R}^{-1} (\bar{\boldsymbol{y}}^b - \boldsymbol{y}_{\text{obs}}).$$

Here $\mathbf{I}$ is an identity matrix of size $M \times M$, $\mathbf{1}_M$ is a vector of size M with all ones, $\bar{\boldsymbol{y}}^b$ is the mean of the predicted data defined by

$$\bar{\boldsymbol{y}}^b = \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{y}_m^b,$$

$\mathbf{A}^b$ is the background ensemble anomalies of the predicted data defined as

$$\mathbf{A}^b = \left[(\boldsymbol{y}_1^b - \bar{\boldsymbol{y}}^b) \quad (\boldsymbol{y}_2^b - \bar{\boldsymbol{y}}^b) \quad \dots \quad (\boldsymbol{y}_M^b - \bar{\boldsymbol{y}}^b)\right],$$

and $\mathbf{R}$ is the measurement error covariance. To ensure that the anomalies of analysis remain zero-centered we check whether $\mathbf{A}^a\mathbf{1}_M = \mathbf{A}^b\mathbf{S}\mathbf{1}_M = \mathbf{0}$, given $\mathbf{S}\mathbf{1}_M = \mathbf{1}_M$ and $\mathbf{A}^b\mathbf{1}_M = \mathbf{0}$. The model parameters $\boldsymbol{u}_m^b$ and the predicted data $\boldsymbol{y}_m^b$ are related by $\boldsymbol{y}_m^b = h(\boldsymbol{u}_m^b)$, where h is a nonlinear function and here we assume that the function h is known.

2.2.2 Ensemble Transform Particle Filter

In particle filtering we represent the probability distribution function using ensemble members (also called particles) as in an ensemble Kalman filter. We start by assigning prior (background) weights $\{w_m^b\}_{m=1}^M$ to M particles and then compute new (analysis) weights $\{w_m^a\}_{m=1}^M$ using the Bayes' formula and observations $\boldsymbol{y}_{\text{obs}}$

$$w_m^a = \frac{\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{u}_m^b)w_m^b}{\pi(\boldsymbol{y}_{\text{obs}})}. \tag{2.2}$$

We assume that initially all particles have equal weight, thus $w_m^b = 1/M$ for $m = 1, \dots, M$, and that the likelihood is Gaussian with error covariance matrix $\mathbf{R}$. Then from Eq. (2.2) w_m^a is given by

$$w_m^a = \frac{\exp\left[-\frac{1}{2}(\boldsymbol{y}_m^b - \boldsymbol{y}_{\text{obs}})^T\mathbf{R}^{-1}(\boldsymbol{y}_m^b - \boldsymbol{y}_{\text{obs}})\right]}{\sum_{j=1}^M \exp\left[-\frac{1}{2}(\boldsymbol{y}_j^b - \boldsymbol{y}_{\text{obs}})^T\mathbf{R}^{-1}(\boldsymbol{y}_j^b - \boldsymbol{y}_{\text{obs}})\right]}, \quad m = 1, \dots, M. \tag{2.3}$$

In Importance Sampling (IS), which will be used in this chapter as a "ground" truth, these weights define the posterior pdf. The mean parameter for IS is then

$$\bar{\boldsymbol{u}}^a = \sum_{m=1}^M \boldsymbol{u}_m^b w_m^a.$$

It is important to note that IS *does not* change the parameters $\boldsymbol{u}$, it only modifies the weight of the particles (samples). Therefore a resampling needs to be implemented for parameter estimation, which is usually stochastic. Instead particle filtering has been

modified using a deterministic coupling methodology which resulted in an ensemble transform particle filter of [74]. ETPF looks for a coupling between two discrete random variables B_1 and B_2 so as to convert the ensemble members belonging to the random variable B_2 with probability distribution $\pi(B_2 = \boldsymbol{u}_m^b) = w_m^a$ to the random variable B_1 with uniform probability distribution $\pi(B_1 = \boldsymbol{u}_m^b) = 1/M$. The coupling between these two random variables is an $M \times M$ matrix $\mathbf{T}$ whose entries should satisfy

$$t_{mj} \geq 0, \quad m, j = 1, \ldots, M, \tag{2.4}$$

$$\sum_{m=1}^{M} t_{mj} = \frac{1}{M}, \quad j = 1, \ldots, M, \tag{2.5}$$

$$\sum_{j=1}^{M} t_{mj} = w_m^a, \quad m = 1, \ldots, M. \tag{2.6}$$

An optimal coupling matrix $\mathbf{T}^*$ with elements t_{mj}^* minimizes the squared Euclidean distance

$$J(t_{mj}) = \sum_{m,j=1}^{M} t_{mj} ||\boldsymbol{u}_m^b - \boldsymbol{u}_j^b||^2 \tag{2.7}$$

and the analysis model parameters are obtained by the linear transformation

$$\boldsymbol{u}_j^a = M \sum_{m=1}^{M} t_{mj}^* \boldsymbol{u}_m^b, \quad j = 1, \ldots, M. \tag{2.8}$$

Then the mean parameter for ETPF is

$$\bar{\boldsymbol{u}}^a = \sum_{m=1}^{M} \boldsymbol{u}_m^a \frac{1}{M}.$$

We use $FastEMD$ algorithm of [69] to solve the linear transport problem and get the optimal transport matrix.

Remark: An important property of ETPF is preservation of imposed interval bounds on ensemble members. Consider an ensemble of parameters $\{\boldsymbol{u}_m^b\}_{m=1}^M$ given by

$$\boldsymbol{u}_m^b = (a_m^b \; b_m^b \; c_m^b)^T, \quad m = 1, \ldots, M,$$

where we assume all the parameters $\{a_m^b\}_{m=1}^M$, $\{b_m^b\}_{m=1}^M$ and $\{c_m^b\}_{m=1}^M$ are bounded

between 0 and 1. Therefore, the following inequalities hold:

$$\begin{aligned}
0 < a_{\min} \leq a_m^b \leq a_{\max} < 1, \quad m = 1, \ldots, M, \\
0 < b_{\min} \leq b_m^b \leq b_{\max} < 1, \quad m = 1, \ldots, M, \\
0 < c_{\min} \leq c_m^b \leq c_{\max} < 1, \quad m = 1, \ldots, M.
\end{aligned}$$

Now we assume two discrete random variables B_1 and B_2 have probability distributions given by

$$\pi(B_1 = \boldsymbol{u}_m^b) = 1/M, \quad \pi(B_2 = \boldsymbol{u}_m^b) = w_m^a,$$

with $w_m^a \geq 0$, $m = 1, \ldots, M$ and $\sum_{m=1}^{M} w_m^a = 1$. As ETPF looks for a matrix $\mathbf{T}^*$ which defines coupling between these two probability distributions, each entry of this coupling matrix satisfies the conditions given by Eq. (2.4)–(2.6). These conditions assure that each entry of the coupling matrix will be non-negative and less than 1. Since the analysis given by Eq. (2.8) is

$$\boldsymbol{u}_m^a = \begin{bmatrix} a_1^b(Mt_{1m}^*) + a_2^b(Mt_{2m}^*) + \cdots + a_M^b(Mt_{Mm}^*) \\ b_1^b(Mt_{1m}^*) + b_2^b(Mt_{2m}^*) + \cdots + b_M^b(Mt_{Mm}^*) \\ c_1^b(Mt_{1m}^*) + c_2^b(Mt_{2m}^*) + \cdots + c_M^b(Mt_{Mm}^*) \end{bmatrix}, \quad m = 1, \ldots, M,$$

these conditions lead to

$$\begin{aligned}
0 < a_{\min} \leq a_m^a \leq a_{\max} < 1, \quad m = 1, \ldots, M, \\
0 < b_{\min} \leq b_m^a \leq b_{\max} < 1, \quad m = 1, \ldots, M, \\
0 < c_{\min} \leq c_m^a \leq c_{\max} < 1, \quad m = 1, \ldots, M.
\end{aligned}$$

Thus the coupling matrix bounds the analysis ensemble members to be in the desired range. This is not observed in ETKF as the matrix $\mathbf{S}$ given by Eq. (2.1) does not impose any of the non-equality and equality constraints, so it results in values outside the bound.

2.2.3 Localization

The achievable accuracy of both ensemble Kalman filter and particle filter are limited by computational resources, i.e. memory and available processor time. The required computational resources grow dramatically with increased ensemble size, which is nec-

essary to improve accuracy. This limit of a small ensemble size introduces sampling errors. To deal with this issue localized ETKF (LETKF) was introduced in [42] and localized ETPF (LETPF) in [74]. More recent approaches to particle filter localization include [70] and [71].

We implement distance-based localization that restricts the influence of all the observations outside a certain radius around the grid point in concern. As a consequence, the actual data assimilation problem associated with each observation is of much smaller dimension. For the local update of a model parameter $\boldsymbol{u}_m(X_i)$ at a grid point X_i, we introduce a diagonal matrix $\hat{\mathbf{C}}_i \in R^{N_y \times N_y}$ in the observation space with an element

$$(\hat{\mathbf{C}}_i)_{ll} = \rho\left(\frac{||X_i - r_l||}{r_{\text{loc}}}\right), \tag{2.9}$$

where $i = 1, \ldots, n^2$, $l = 1, \ldots, N_y$, n^2 is the number of model parameters, N_y is the dimension of the observation space, r_l denotes the location of the observation, r_{loc} is a localisation radius and $\rho(\cdot)$ is a taper function, such as Gaspari-Cohn function by [33]

$$\rho(r) = \begin{cases} 1 - \frac{5}{3}r^2 + \frac{5}{8}r^3 + \frac{1}{2}r^4 - \frac{1}{4}r^5, & 0 \le r \le 1, \\ -\frac{2}{3}r^{-1} + 4 - 5r + \frac{5}{3}r^2 + \frac{5}{8}r^3 - \frac{1}{2}r^4 + \frac{1}{12}r^5, & 1 \le r \le 2, \\ 0, & 2 \le r. \end{cases}$$

Then the estimated model parameter at the location X_i is

$$\boldsymbol{u}_m^a(X_i) = \sum_{l=1}^{M}\left(s_{lm}(X_i) + q_l(X_i) - \frac{1}{M}\right)\boldsymbol{u}_l^b(X_i), \quad m = 1, \ldots, M,$$

where $s_{lm}(X_i)$ is the (l, m) entry of the localized transformation matrix $\mathbf{S}(X_i)$

$$\mathbf{S}(X_i) = \left[\mathbf{I} + \frac{1}{M-1}(\mathbf{A}^b)^T(\hat{\mathbf{C}}_i\mathbf{R}^{-1})\mathbf{A}^b\right]^{-1/2}$$

and $q_l(X_i)$ is the l-th entry of the localized column $\boldsymbol{q}(X_i)$

$$\boldsymbol{q}(X_i) = \frac{1}{M}\mathbf{1}_M - \frac{1}{M-1}\mathbf{S}(X_i)^2(\mathbf{A}^b)^T\mathbf{R}^{-1}(\bar{\boldsymbol{y}}^b - \boldsymbol{y}_{\text{obs}}).$$

Distance based localization reduces one n^2-dimensional problem into n^2-one di-

mensional problems. LETPF updates each spatially varying parameter individually. Hence, reducing the number of degrees of freedom enables application of ETPF to spatially-extended systems. LETPF modifies the likelihood and thus the weights given by Eq. (2.3) are computed locally at each grid X_i

$$w_m^a(X_i) = \frac{\exp\left[-\frac{1}{2}(\boldsymbol{y}_m^b - \boldsymbol{y}_{\text{obs}})^T(\hat{\mathbf{C}}_i\mathbf{R}^{-1})(\boldsymbol{y}_m^b - \boldsymbol{y}_{\text{obs}})\right]}{\sum_{j=1}^M \exp\left[-\frac{1}{2}(\boldsymbol{y}_j^b - \boldsymbol{y}_{\text{obs}})^T(\hat{\mathbf{C}}_i\mathbf{R}^{-1})(\boldsymbol{y}_j^b - \boldsymbol{y}_{\text{obs}})\right]}, \quad m = 1, \ldots, M, \quad (2.10)$$

where $\hat{\mathbf{C}}_i$ is the diagonal matrix given by Eq. (2.9). Then the estimated model parameter $\boldsymbol{u}_j^a(X_i)$ at the grid X_i is given by

$$\boldsymbol{u}_j^a(X_i) = M \sum_{m=1}^M t_{mj}^* \boldsymbol{u}(X_i)_m^b, \quad j = 1, \ldots, M,$$

where t_{mj}^* is an element of an optimal coupling matrix $\mathbf{T}^*$ which minimizes the squared Euclidean distance at the grid point X_i

$$J(t_{mj}) = \sum_{m,j=1}^M t_{mj}[u_m^b(X_i) - u_j^b(X_i)]^2, \quad (2.11)$$

which reduces LETPF to a univariate transport problem. It should be noted that localization can be applied only for grid-dependent parameters.

2.3 One parameter nonlinear problem

First we consider a one parameter nonlinear problem from [12]. The prior distribution is a Gaussian distribution with mean 4 and variance 1. The nonlinear observation function is

$$h(u) = \frac{7}{12}u^3 - \frac{7}{2}u^2 + 8u.$$

The system is deterministic and the true parameter $u^{\text{true}} = 6$ gives $h(u^{\text{true}}) = 48$. The observation error is drawn from a Gaussian distribution with zero mean and variance 16. Our aim here is to employ the noisy observation and estimate the parameter u using ensemble data assimilation. Such a simple one-dimensional problem allows us to understand the behavior of the methods and examine their performance in approximating the

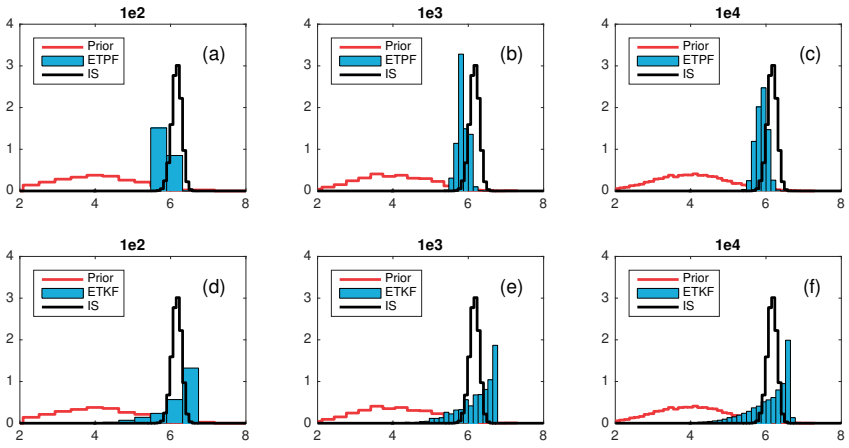

Figure 2.1: Probability density functions for the one parameter nonlinear problem. Top: ETPF, bottom: ETKF. Left: ensemble size 10^2, center: ensemble size 10^3, right: ensemble size 10^4. Prior is in red. True pdf obtained by IS with ensemble size 10^5 is in black.

posterior distribution. In Fig. 2.1 we plot the posterior probability density functions estimated by ETPF (top), ETKF (bottom) with ensemble sizes 10^2 (left), 10^3 (center), and 10^4 (right). The prior distribution is shown in red and the posterior estimated by IS with ensemble size 10^5 is shown in black. We implement IS solutions as benchmark since IS is a Monte-Carlo method, and under weak assumptions the law of large number applies. For linear models, ETKF is Monte-Carlo approximation of Kalman filter and in the limit of large ensemble size the ensemble mean and covariance of ETKF converges to exact Kalman filter [59]. However, for nonlinear models it is well-recognized that ETKF need not asymptotically converge to the true posterior. On the other hand, ETPF has been shown to converge to true posterior as $M \to \infty$ for both linear and nonlinear models [73]. We can see that ETPF provides better approximation of the reference probability density function, and the approximation improves as we increase the ensemble size. On the other hand, we observe that ETKF gives a skewed posterior. Such behavior is associated with ETKF's fundamental assumption of Gaussianity. It should be noted that ETKF is able to give a non-Gaussian (though wrong) posterior due to the nonlinearity of the map between the uncertain parameters and observations.

2.4 Single-phase Darcy flow

We consider a steady-state single-phase Darcy flow model defined over an aquifer of two-dimensional physical domain $D = [0, 1] \times [0, 1]$, which is given by

$$-\nabla \cdot (k(x, y)\nabla P(x, y)) = f(x, y), \quad (x, y) \in D$$
$$P(x, y) = 0, \quad (x, y) \in \partial D$$

where $\nabla = (\partial/\partial x \ \partial/\partial y)^T$, $\cdot$ denotes the dot product, $P(x, y)$ the pressure, $k(x, y)$ the permeability, and ∂D the boundary of domain D. Here $f(x, y)$ is the source term, which we assume is known and is given by by $2\pi^2 \cos(\pi x)\cos(\pi y)$. The forward problem of this second order elliptical equation is to find the solution of pressure $P(x, y)$ for given $f(x, y)$ and $k(x, y)$. We, however, are interested in finding permeability $k(x, y)$ given noisy observations of pressure at a few locations.

We perform numerical experiments with synthetic observations, where instead of a measuring device a model is used to obtain observations. We implement a cell-centered finite difference method to discretize the domain D into $n \times n$ grid cells X_i of size Δx^2 and solve the forward model with the true parameters. Then the synthetic observations are obtained by

$$\boldsymbol{y}_{\text{obs}} = \mathbf{L}(\mathbf{P}) + \eta,$$

with an element of $\mathbf{L}(\mathbf{P})$ being a linear functional of pressure, namely

$$L_l(\mathbf{P}) = \frac{1}{2\pi\sigma^2} \sum_{i=1}^{n^2} \exp\left(-\frac{||X_i - r_l||^2}{2\sigma^2}\right) P_i \Delta x^2, \quad l \in 1, \ldots, N_y$$

where $n = 50$, $\sigma = 0.01$, r_l denotes the location of the observation and $N_y = 16$, which is the number of observations. The observation locations are spread uniformly across the domain D and η denotes the observation noise drawn from a normal distribution with zero mean and standard deviation of 0.09. This form of the observation functional and parametrization of the uncertain parameters given below guarantee the continuity of the forward map from the uncertain parameters to the observations and thus the existence of the posterior distribution as shown by [46].

2.4.1 Five parameter nonlinear problem

For our first numerical experiment with Darcy flow, we consider a low-dimensional problem where the permeability field is defined by mere 5 parameters similarly to [46]. We assume that the entire domain $D = [0, 1] \times [0, 1]$ is divided into two subdomains D_1 and D_2 as shown in Fig. 2.2. Each subdomain of D represents a layer and is assumed to have a permeability function $k(\mathbf{X})$, where an element of $\mathbf{X}$ is defined by X_i for $i = 1, \ldots, n^2$.

Parameters a and b denote the thickness of the bottom layer on the left and the right side, respectively, which correspondingly defines the slope of the interface. A parameter c defines a vertical fault. The vertical fault displaces the layer up or down depending on $c < 0$ or $c > 0$, respectively, and its location is assumed to be fixed at $x = 0.5$.

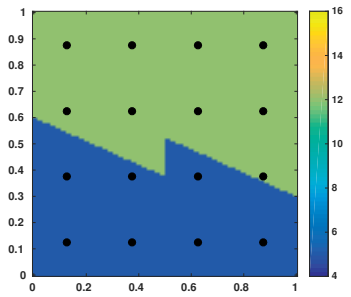

Figure 2.2: True permeability of the 5 parameter nonlinear problem with dots representing the observation locations.

Further, for this test case we assume piecewise constant permeability within each of the subdomains, hence $k(\mathbf{X})$ is given by

$$k(\mathbf{X}) = k_1 \delta_{D_1}(\mathbf{X}) + k_2 \delta_{D_2}(\mathbf{X}),$$

where k_1 and k_2 represent permeability of the subdomain D_1 and D_2, respectively, and δ is characteristic function. Then the parameters defining the permeability field for this configuration are

$$\boldsymbol{u} = (a \ b \ c \ \log(k_1) \ \log(k_2))^T.$$

We assume that the true parameters are $a^{\text{true}} = 0.6$, $b^{\text{true}} = 0.3$, $c^{\text{true}} = -0.15$, $k_1^{\text{true}} = 12$ and $k_2^{\text{true}} = 5$. These parameters are used to create synthetic observations. Figure 2.2 shows the true permeability with dots representing the observation locations. Next, we

assume that the five uncertain parameters are drawn from a uniform distribution over a specified interval, namely $a, b \sim \mathcal{U}[0,1]$, $c \sim \mathcal{U}[-0.5, 0.5]$, $k_1 \sim \mathcal{U}[10,15]$ and $k_2 \sim \mathcal{U}[4,7]$.

As it was pointed out in Sect. 2.2.2, ETPF updates the parameters within the original range of an initial ensemble, while ETKF does not. Therefore a change of variables has to be performed for ETKF so that the updated parameters are physically viable. In order to be consistent we perform the change of variables for ETPF as well. As the domain D is $[0,1] \times [0,1]$, the parameters a and b should lie within the interval $[0,1]$. To enforce this constraint we substitute a according to

$$a' = \log\left(\frac{a}{1-a}\right), \quad a' \in R$$

and similarly b is substituted by b'. Thus the uncertain parameters are now $\boldsymbol{u}' = (a'\ b'\ c\ \log(k_1)\ \log(k_2))^T$.

We first analyse the performance of ETPF and ETKF in estimation of the posterior distribution of parameters using IS as the benchmark. In Fig. 2.3 we plot probability density functions for parameters a (a)–(d), c (e)–(h) and $\log(k_2)$ (i)–(l), as the parameters b and $\log(k_1)$ show similar results. The posterior distribution obtained by IS with ensemble size 10^6 is plotted as a black line and the true value of parameters is plotted as a black line with crosses. The posterior of ETPF is shown at the top and the posterior of ETKF at the bottom. ETPF and ETKF used 10^3 (odd columns) and 10^4 (even columns) ensemble members. In order to perform an objective comparison between the distribution we compute the Kullback-Leibler divergence of a posterior π obtained by either ETPF or ETKF and the posterior π^{IS} obtained by IS

$$\mathrm{D}_{\mathrm{KL}}(\pi^{\mathrm{IS}} \parallel \pi) = \sum_{i=1}^{N_{\mathrm{b}}} \pi^{\mathrm{IS}}(u_i) \log \frac{\pi^{\mathrm{IS}}(u_i)}{\pi(u_i)} (u_i - u_{i-1}), \tag{2.12}$$

where $N_{\mathrm{b}} = 20$ is the number of bins. The Kullback-Leibler divergence for parameters a, c and $\log(k_2)$ is displayed in the titles of Fig. 2.3, where we observe that ETKF outperforms ETPF.

In order to check the sensitivity of the results to the initial parameter ensemble we perform $R = 10$ simulations based on a random draw of an initial ensemble from the same prior distributions. We conduct the numerical experiments varying from 10 to 10^3, with an increment of 50 in the interval $50 - 10^3$. In Fig. 2.4 we plot the true parameters,

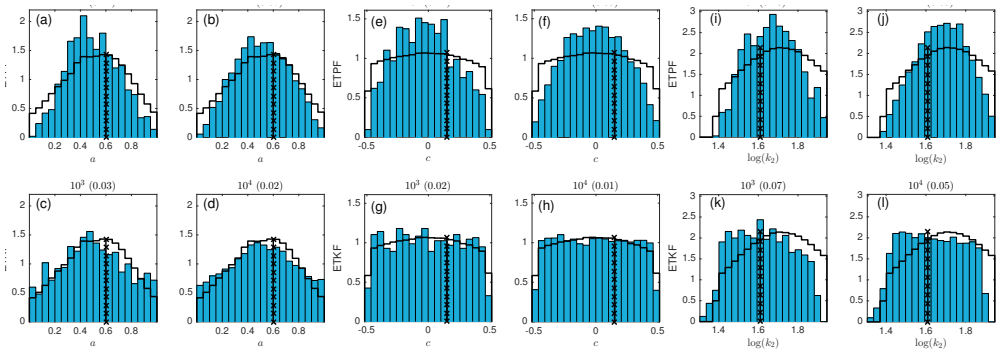

Figure 2.3: Probability density functions for the parameters a (a)–(d), c (e)–(h), and $\log(k_2)$ (i)–(l). The posterior obtained by IS with ensemble size 10^6 is plotted as a black line and the true values of parameters are plotted as black crosses. The posterior of ETPF is shown at the top and the posterior of ETKF at the bottom. ETPF and ETKF used 10^3 (odd columns) and 10^4 (even columns) ensemble members.

the mean estimated by IS, the mean $\bar{\bar{\boldsymbol{u}}}^a$ and the spread $\bar{\bar{\boldsymbol{u}}}^a \pm \bar{\boldsymbol{u}}^a_{\text{std}}$ of estimated parameters averaged over 10 simulations

$$\bar{\bar{u}}^a_i = \frac{1}{R}\sum_{r=1}^{R} \bar{u}^{a,r}_i, \quad \bar{u}^a_{\text{std}} = \frac{1}{R}\sum_{r=1}^{R}\sqrt{\frac{1}{M-1}\sum_{m=1}^{M}(u^{a,r}_{i,m} - \bar{u}^{a,r}_i)^2},$$
$$\text{where } \bar{u}^{a,r}_i = \frac{1}{M}\sum_{m=1}^{M} u^{a,r}_{i,m},\ r = 1,\dots,R, \quad (2.13)$$

M is ensemble size, $i = 1,\dots,5$ is parameter index, and the superscript a is for analysis. We observe that all the methods including IS have a bias in the estimations of geometrical parameters, which is due to a small number of observations. ETPF and ETKF perform comparably in terms of mean estimation, though some are better estimated by ETKF and other are better estimated by ETPF. Comparing the error in pressure of the mean parameters we observe that the methods are equivalent (thus not shown), which is a manifestation of the ill-posedness of the problem. In Fig. 2.4 we see that the spread from ETPF is smaller than from ETKF for each parameter. Both methods are slightly underdispersive as the spread to error ratio is below 1. For ensemble size 10^3 the the spread to error ratio for ETKF is (0.95 0.88 0.88 0.97 0.98) and for ETPF is

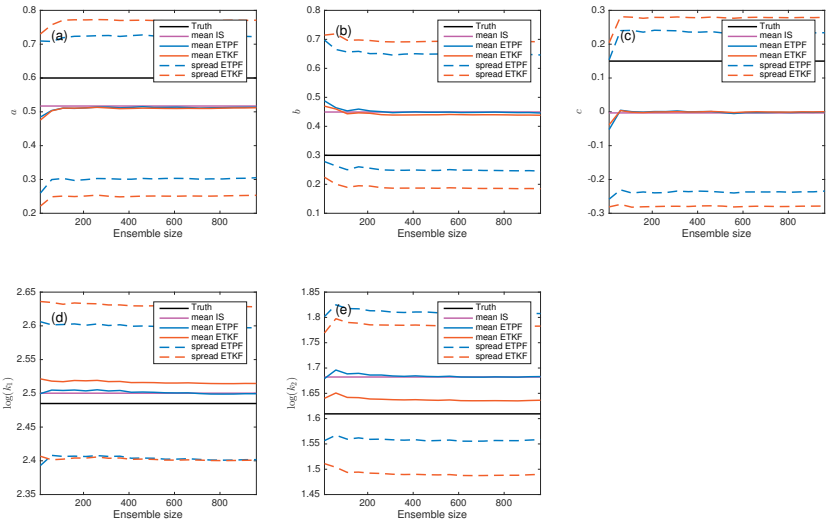

Figure 2.4: $\bar{\bar{\boldsymbol{u}}}^a$ and $\bar{\bar{\boldsymbol{u}}}^a \pm \bar{\boldsymbol{u}}^a_{\text{std}}$ w.r.t ensemble size: (a) for the parameter a, (b) for b, (c) for c, (d) for $\log(k_1)$, (e) for $\log(k_2)$. ETPF is shown in blue, ETKF in red, the true parameters are in black and the mean of IS in magenta.

(0.92 0.81 0.84 0.99 0.86) for (a b c $\log(k_1)$ $\log(k_2)$). Thus ETKF gives better ratio for all the parameters but $\log(k_1)$.

We compute an average of the relative error over all parameters

$$\text{RE}^{a,r} = \frac{1}{5}\sum_{i=1}^{5} \frac{|\bar{u}_i^{a,r} - u_i^{\text{true}}|}{|u_i^{\text{true}}|}, \ r = 1, \ldots, R,$$

and the data misfit

$$\text{misfit}^{a,r} = (\bar{\boldsymbol{y}}^{a,r} - \boldsymbol{y}_{\text{obs}})^T R^{-1} (\bar{\boldsymbol{y}}^{a,r} - \boldsymbol{y}_{\text{obs}}), \ r = 1, \ldots, R \tag{2.14}$$

after data assimilation. The same metrics are computed before data assimilation and denoted by a superscript b. In Fig. 2.5(a)–(b) we plot ($\text{misfit}^{a,r} - \text{misfit}^{b,r}$) and ($\text{RE}^{a,r} - \text{RE}^{b,r}$), respectively, for each simulation r as a function of ensemble size. ETPF is shown in blue and ETKF in red. The black line is at the zero level. Positive values of the differences mean an increase of either data mismatch or relative error after data assimilation. We observe a data misfit decrease for both ETPF and ETKF except at an ensemble size 10. RE does not always decrease for ETPF: for some simulations ETPF is

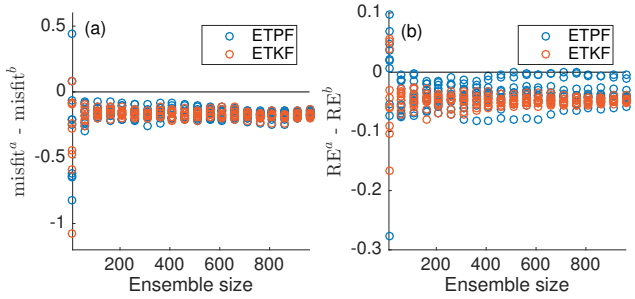

Figure 2.5: misfit$^{a,r}-$misfit^{b,r} (a) and RE$^{a,r}-$RE^{b,r} (b) w.r.t ensemble size. ETPF is shown in blue, ETKF in red and the zero level in black. One circle is for one simulation.

at the zero level or slightly above it, while for ETKF the sole exception is at an ensemble size 10.

2.4.2 High-dimensional nonlinear problem

Next, we consider a high-dimensional problem where the dimension of the uncertain parameter is $n^2 = 2500$. The domain D is now not divided into subdomains. However, unlike in the previous test case, here we implement a spatially varying permeability field. We assume the log permeability is generated by a random draw from a Gaussian distribution $\mathcal{N}(\log(\mathbf{5}), \mathbf{C})$. Here $\mathbf{5}$ is an n^2 vector with all elements 5. $\mathbf{C}$ is assumed to be an exponential correlation elements

$$C_{i,j} = \exp(-3(|h_{i,j}|/v)),\ i, j = 1, \ldots, n^2.$$

Here $h_{i,j}$ is the distance between two spatial locations and v is the correlation range which is taken to be 0.5. For the log permeability we use Karhunen-Loeve expansions of the form

$$\log(k_j) = \log(5) + \sum_{i=1}^{n^2} \sqrt{\lambda_i}\nu_{i,j}\mathcal{Z}_i, \quad \text{for} \quad j = 1, \ldots, n^2, \tag{2.15}$$

where λ and ν are eigenvalues and eigenfunctions of $\mathbf{C}$, respectively, and the vector $\mathcal{Z}$ of dimension n^2 is drawn iid from a Gaussian distribution with zero mean and variance one. Making sure that the eigenvalues are sorted in descending order $\mathcal{Z}_i \sim \mathcal{N}(0, 1)$ produces

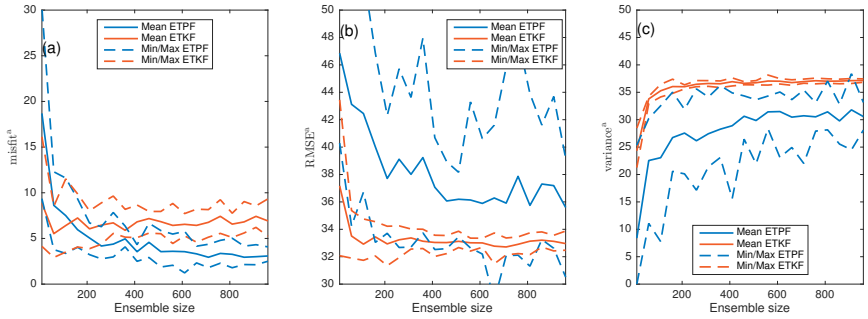

Figure 2.6: Mean, minimum and maximum over 10 simulations after data assimilation for the data misfit (a), RMSE (b), and variance (c). ETPF is shown in blue and ETKF in red.

$\log(\mathbf{k}) \sim \mathcal{N}(\log(\mathbf{5}), \mathbf{C})$. The uncertain parameter is thus $\boldsymbol{u} = \mathcal{Z}$ with the dimension $n^2 = 2500$.

We perform $R = 10$ different simulations based on a random draw of an initial ensemble from the prior distribution. We conduct the numerical experiments for ensemble sizes varying from 10 to 10^3, with an increment of 50 in the interval $50 - 10^3$. We compute the root mean square error (RMSE) of the log permeability field

$$\text{RMSE}^{r,a} = \sqrt{\left(\log(\overline{\mathbf{k}}^{a,r}) - \log(\mathbf{k}^{\text{true}})\right)^T \left(\log(\overline{\mathbf{k}}^{a,r}) - \log(\mathbf{k}^{\text{true}})\right)}, \quad r = 1, \ldots, R,$$

and variance

$$\text{variance}^{r,a} = \frac{1}{M-1} \sum_{m=1}^{M} \left((\log(\mathbf{k}_m^{a,r}) - \log(\overline{\mathbf{k}}^{a,r})\right)^T \left(\log(\mathbf{k}_m^{a,r}) - \log(\overline{\mathbf{k}}^{a,r})\right), \quad r = 1, \ldots, R.$$

We also compute the data misfit for each simulation after data assimilation by Eq. (2.14). In Fig. 2.6 we plot mean, minimum and maximum over 10 simulations after data assimilation for the data misfit (left), RMSE (center), and variance (right). ETPF is shown in blue and ETKF in red. We observe that ETPF is underdispersive compared to ETKF as particle filters are highly degenerative compared to Kalman filters. The misfit given by ETPF is smaller than the one given by ETKF for almost all simulations at ensemble sizes greater than 150. The RMSE on the contrary is larger. In Fig. 2.7(a)–(b) we plot

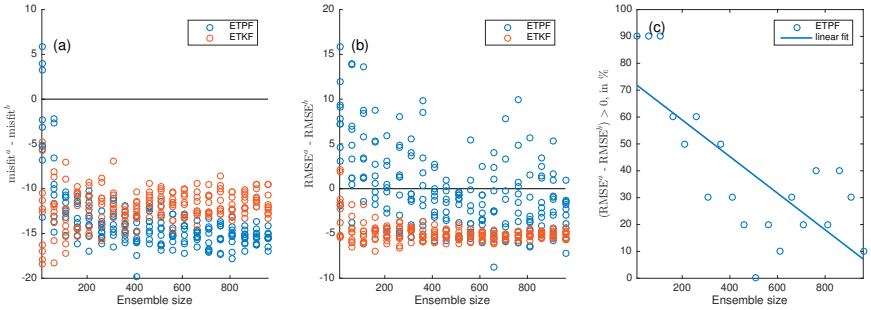

Figure 2.7: misfit^{a,r} − misfit^{b,r} (a) and RMSE^{a,r} − RMSE^{b,r} (b) w.r.t ensemble size. ETPF is shown in blue, ETKF in red and zero level in black. One circle is for one simulation. For ETPF % of simulations that result in (RMSE^a − RMSE^b) > 0 and a linear fit as a function of ensemble size are shown in (c).

(misfit^{a,r} − misfit^{b,r}) and (RMSE^{a,r} − RMSE^{b,r}), respectively, as a function of ensemble size for a simulation $r = 1, \ldots, 10$. The superscript b is for the metrics before data assimilation and the superscript a is for the metrics after data assimilation. ETKF always provides a decrease in both the data misfit and RMSE except at ensemble size 10. ETPF gives a decrease in the data misfit though an increase in RMSE, which indicates that ETPF overfits the data. However, as the ensemble size increases this happens less often as can be seen in Fig. 2.7(c), where we plot for ETPF a percentage of simulations that result in (RMSE^a − RMSE^b) > 0 and a linear fit as a function of ensemble size.

In Fig. 2.8 we plot log permeability fields. In Fig. 2.8(a) the true permeability is shown with dots representing the observation locations, and in Fig. 2.8(d) the mean permeability field obtained by IS with ensemble size 10^5. The RMSE provided by IS is 32.62. In Fig. 2.8(b), (e) and Fig. 2.8(c), (f) we display mean permeability fields obtained with ensemble size 10^3 by ETPF and ETKF, respectively. In Fig. 2.8(b–c) we plot the mean log permeabilities for the smallest RMSE over simulations, which is 30.51 for ETPF and 32.48 for ETKF. In Fig. 2.8(d–e) we plot the mean log permeabilities for the largest RMSE over simulations, which is 39.2 for ETPF and 33.87 for ETKF. We observe that ETKF as well as IS provide smooth mean permeability fields that have smaller absolute values than the true permeability. ETPF gives higher variations of the mean permeability field and is in an excellent agreement with the true permeability for a good initial ensemble shown in Fig. 2.8(b). This means that ETPF sensitivity to the initial sample is due to sampling error and that the spatial variability of ETPF

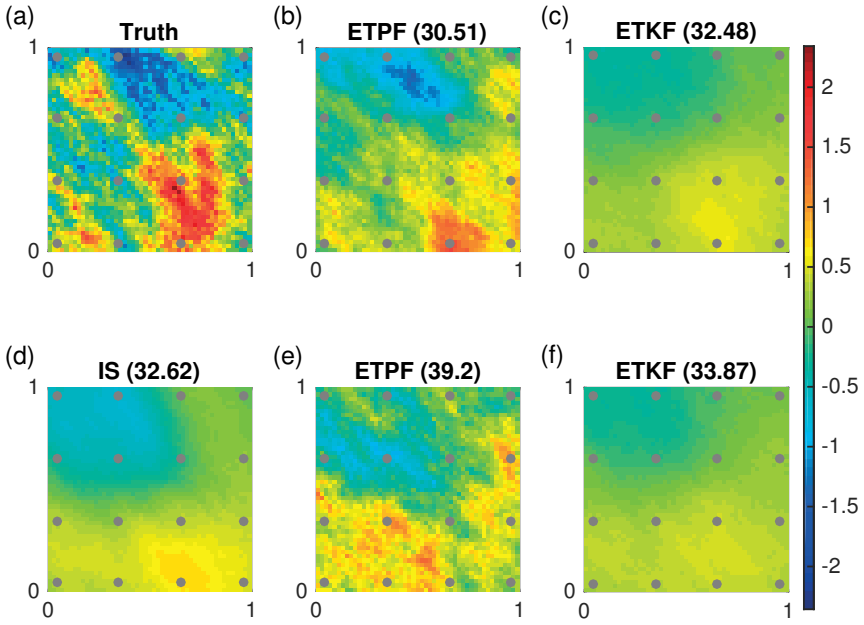

Figure 2.8: Log permeability field with dots representing the observation locations. Truth is shown in (a) and mean obtained by IS with ensemble size 10^5 in (d). Mean obtained with ensemble size 10^3 by ETPF shown in (b), (e) and by ETKF in (c) ,(f), where (b–c) are at the smallest RMSE and (e–f) are at the largest RMSE over simulations. The corresponding RMSE is given in brackets.

is a result of sampling error. It should be noted that IS with ensemble size 10^3 and this good initial ensemble gives the RMSE 30.51 and the same mean log permeability field as ETPF shown in Fig. 2.8(b). However, IS does not change the parameters, only their weights, while ETPF does change the parameters. Therefore ETPF shares an advantage of IS, namely representing the correct posterior, but does not suffer from its lack of resampling. In Fig. 2.9 we plot the variance of the permeability fields obtained with ensemble size 10^5 by IS (d), with ensemble size 10^3 by ETPF (b), (e) and ETKF (c) ,(f). Fig. 2.9(b–c) is for the smallest RMSE and Fig. 2.9(e–f) is for the largest RMSE. ETKF provides smoother spatial variability of log-permeability than ETPF, complying with the smooth mean permeability field obtained by ETKF.

In Fig. 2.10 we show squared error $(\overline{\mathcal{Z}}^a - \mathcal{Z}^{\text{true}})^2$ in blue for ETPF and in red for ETKF for three leading modes $\mathcal{Z}_1$ (a), $\mathcal{Z}_2$ (b), and $\mathcal{Z}_3$ (c), where the solid line indicates

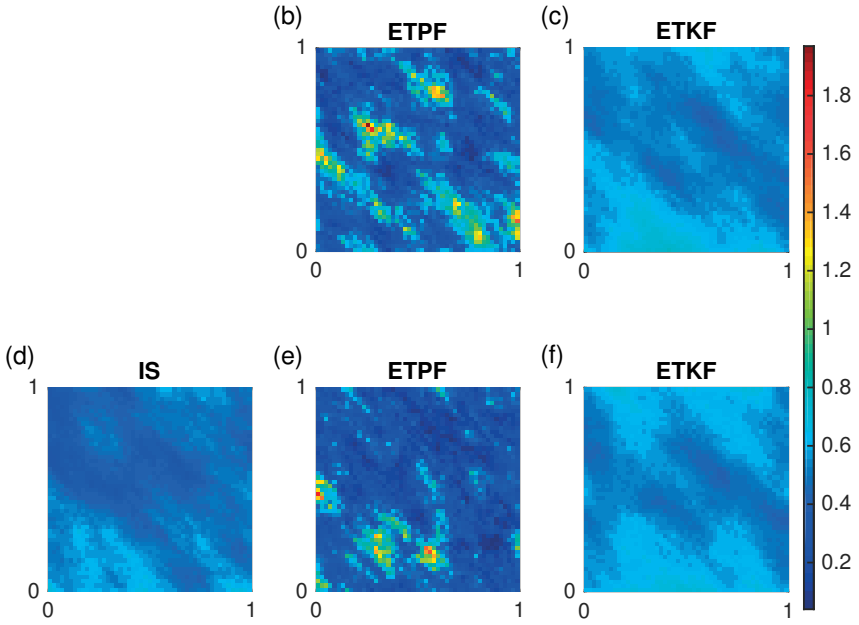

Figure 2.9: Variance of log permeability fields: obtained with ensemble size 10^5 by IS (d), with ensemble size 10^3 by ETPF (b), (e), and ETKF (c), (f). Variance at the smallest RMSE (b–c) and at the largest RMSE (e–f) over simulations.

the median and shaded area is for 25 and 75 percentile over 10 simulations. We observe that in terms of the estimation of the three leading modes ETPF attains lower squared error as the ensemble size increases, with the lowest achieved error among 10 simulations of ETPF outperforms that of ETKF even at smaller ensemble sizes. In Fig. 2.11 we plot the posterior of $\mathcal{Z}_1$ (left), $\mathcal{Z}_2$ (center), and $\mathcal{Z}_3$ (right) obtained by IS with ensemble size 10^6 and by ETPF (top) and ETKF (bottom) with ensemble size 10^4. The posterior of these modes is roughly approximated by ETPF as shown in Fig. 2.11 (a)–(c). ETKF provides a skewed posterior of the modes shown in Fig. 2.11 (d)–(f), which was also observed in the one parameter nonlinear problem, see Fig. 2.1(f). In order to perform an objective comparison between the probabilities we compute the Kullback-Leibler divergence of a posterior π obtained by either ETPF or ETKF and the posterior π^{IS} obtained by IS according to Eq. (2.12). ETPF gives the Kullback-Leibler divergence 0.21, 0.42, and 0.6, while ETKF 0.16, 0.07, and 0.5 for the modes $\mathcal{Z}_1$, $\mathcal{Z}_2$, and $\mathcal{Z}_3$, respectively. Thus ETKF gives a better approximation of the true pdf. Since the first

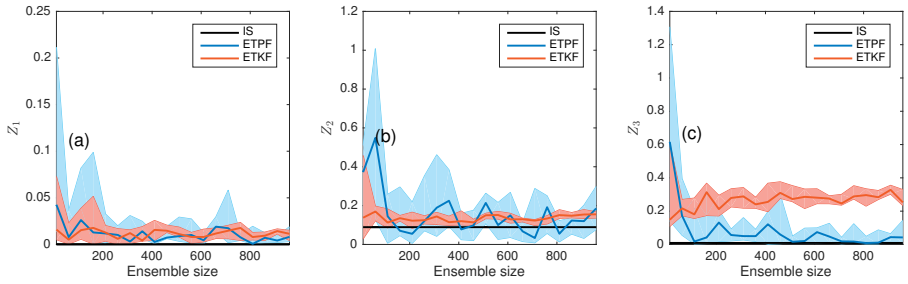

Figure 2.10: Squared error between the true and the mean estimated modes for $\mathcal{Z}_1$ (a), $\mathcal{Z}_2$ (b), and $\mathcal{Z}_3$ (c) w.r.t ensemble size. ETPF is shown in blue and ETKF in red with solid lines for median and shaded area for 25 and 75 percentile over 10 simulations. IS with ensemble size 10^5 is in black.

modes are well estimated by ETPF and last modes are not (not shown), we use only three leading modes in the Karhunen-Loeve expansion given by Eq. (2.15) when computing the estimated log permeability, keeping the number of uncertain parameters the same, namely 2500. In Fig. 2.12(a) we observe that ETPF gives lower RMSE than ETKF for large ensemble sizes independent of an initial sample. While at small ensemble sizes, the lowest attained RMSE among the 10 simulations of ETPF outperforms that of ETKF. Moreover, ETPF is not overfitting the data anymore since RMSE always decreases after data assimilation except at small ensemble sizes shown in Fig. 2.12(b). In Fig. 2.13 we show the mean fields for the best and worst initial samples of 10^4 size. ETPF gives RMSE at the best sample 31.1 and the worst sample 32.98. By comparing it to 30.51 and 39.2 obtained using the full Karhunen-Loeve expansions, we observe that the maximum RMSE over simulations decreased substantially, while the minimum RMSE only slightly increased. ETKF gives RMSE at the best sample 32.27 and the worst sample 33.23. (Compare to 32.48 and 33.9 using the full Karhunen-Loeve expansions). Thus ETKF slightly decreases both maximum and minimum RMSE over simulations. ETPF is more affected by sampling noise at small scales, so using a truncated representation of the fields significantly improves the results for ETPF. ETKF is filtering out the small scales that are not observed and thus is less affected by the truncation.

Next we apply LETPF and LETKF. The optimal localization radius between 0.2 and 1.2 was obtained in terms of the smallest RMSE and shown in Table 2.1. It should be noted that smaller localization radius for LETPF than for LETKF was also observed

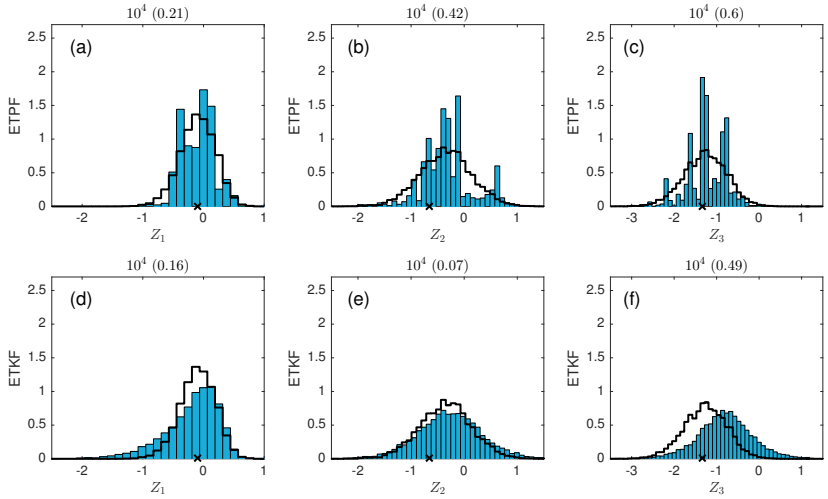

Figure 2.11: The posterior probability density function of parameters $\mathcal{Z}_1$ (left), $\mathcal{Z}_2$ (center), and $\mathcal{Z}_3$ (right). The posterior obtained by IS with ensemble size 10^6 is plotted as a black line and the true parameter as a black cross. The posterior of ETPF is shown at the top and the posterior of ETKF at the bottom. Both ETPF and ETKF used 10^4 ensemble members. The Kullback-Leibler divergence is in brackets.

by [14] and is probably related to more noisy approximation of the posterior by LETPF than by LETKF. In Fig. 2.14 we plot misfit, RMSE and variance.

At small ensemble sizes both LETKF and LETPF give smaller misfit, smaller RMSE but larger variance than ETKF and ETPF. For large ensembles LETKF performs worse than ETKF, which is due to the imposed range on localization radius, meaning that 1.2 is not optimal. Comparing the performance of LETPF to (L)ETKF we observe that at small ensemble sizes LETKF still outperforms ETPF but at large ensemble sizes LETPF performs now comparably to ETKF. Moreover, LETPF overfits the data less often than ETPF. That is, for ensemble size 10 LETPF overfits 40% of simulations and ETPF overfits 90% of simulations (not shown). For ensemble size greater than 150 we do not observe overfiting in LETPF, while ETPF still overfits though to much lesser extent (not shown).

In Fig. 2.15–2.16 we plot mean and variance of the log permeability field at ensemble size 10^3 for ETPF (b), (e) and ETKF (c), (f) with localization at the smallest RMSE (b), (c) and largest RMSE (e), (f) over simulations, which are 32.29 and 34.08 for ETPF

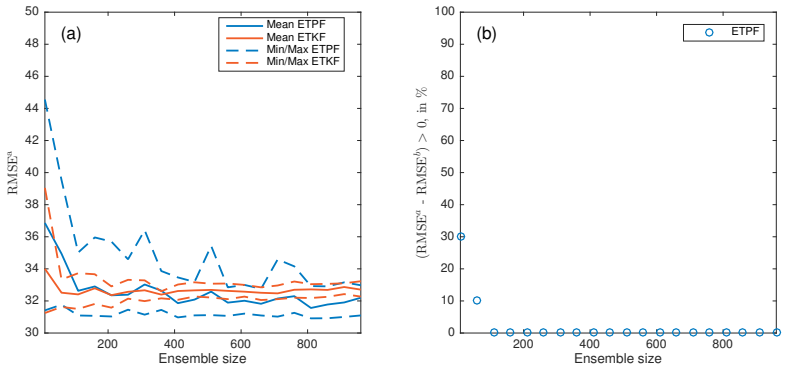

Figure 2.12: Using only three leading modes in the KL expansion. Panel (a): RMSE after data assimilation w.r.t ensemble size with mean, minimum and maximum over 10 simulations for ETPF shown in blue and ETKF in red. Panel (b): % of simulations that result in $(\text{RMSE}^a - \text{RMSE}^b) > 0$ for ETPF.

Table 2.1: Optimal localization radius for LETPF and LETKF at different ensemble sizes M.

M	10	110	210	...	910
LETPF	0.2	0.6	0.6	...	0.6
LETKF	0.2	1.2	1.2	...	1.2

and 32.92 and 34.09 for ETKF, respectively. We observe that localization decreases the sampling noise, and the spatial variability of the mean field obtained by ETPF at ensemble size 10^3 resembles IS at ensemble size 10^5. The variance obtained by ETPF with localization shown in Fig. 2.16(b), (e) has also improved.

The posterior estimation of the leading mode $\mathcal{Z}_1$, however, degraded, while of $\mathcal{Z}_2$ and $\mathcal{Z}_3$ improved. The Kullback-Leibler divergence for the leading mode is 0.73 (compare to 0.21 without localization), and for second and third is 0.2 and 0.18, correspondingly (compare to 0.42 and 0.6 without localization). Variance of the posteriors is larger when localization is applied for both methods. The localized weights given by Eq. (2.10) vary less than the non-localized weights given by Eq. (2.3). Therefore the localized pdf is less noisy than the non-localized. Moreover unlike ETKF, LETPF does not converge

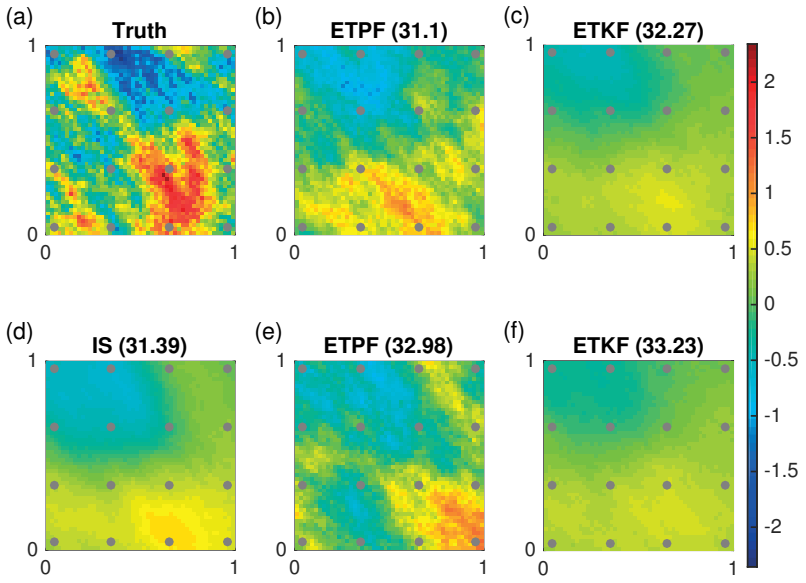

Figure 2.13: Same as figure 2.8, but using only three leading modes in the KL expansion.

to ETPF as the localization radius goes to infinity due to the transport problem being univariate for LETPF and multivariate for ETPF.

2.5 Conclusions

MCMC methods remain the most reliable methods for estimating the posterior distributions of uncertain model parameters and states. However, they also remain computationally expensive. Ensemble Kalman filters (ETKF) provide computationally affordable approximations but rely on the assumptions of Gaussian probabilities. For nonlinear models even if the prior is Gaussian the posterior is not Gaussian anymore. Particle filtering on the other hand does not have such an assumption but requires a resampling step, which is usually stochastic. ETPF is a particle filtering method that deterministically resamples the particles based on their importance weights and covariance maximization among the particles.

ETPF certainly outperforms ETKF for a one parameter nonlinear test case by giving

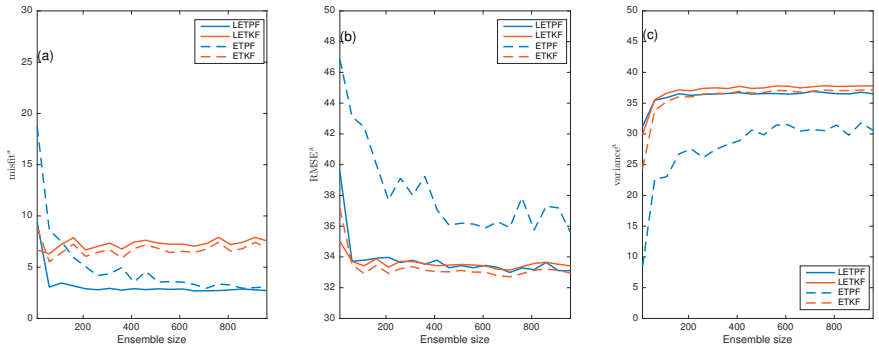

Figure 2.14: Mean over 10 simulations after data assimilation for the data misfit (a), RMSE (b), and variance (c). LETPF is shown in solid blue and LETKF in solid red. ETPF is shown in dashed blue and ETKF in dashed red.

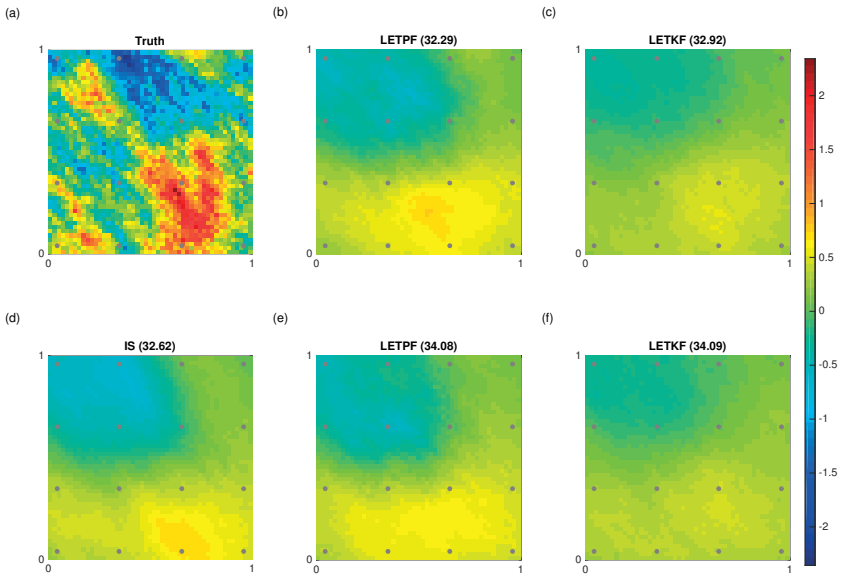

Figure 2.15: Same as figure 2.8, but with localization.

a better posterior estimation. For the five parameter test case, the mean estimations obtained by ETPF are not consistently better than the ones obtained by ETKF and the spread is smaller. The Kullback-Leibler divergence from ETKF is smaller than that from ETPF for all the parameters. When the number of uncertain parameters is large

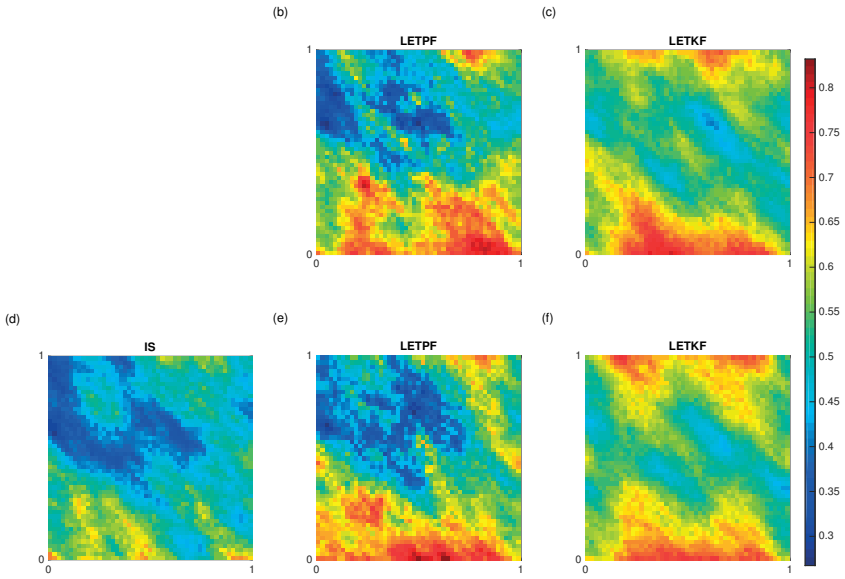

Figure 2.16: Same as figure 2.9, but with localization.

(2500) a decrease in number of degrees of freedom is essential. This is attained by localization. At large ensemble sizes LETPF performs as well as LETKF, while at small ensemble sizes LETKF still outperforms LETPF. Even though LETPF overfits the data less often than ETPF, localization destroys the property of ETPF to retain the imposed bounds. This deteriorates a posterior estimation of the leading mode. Another plausible drawback of localization is an assumption of observations being local, which might not be the case for inverse modelling. An alternative approach to improve ETPF performance is instead of applying localization to use only leading modes in the approximation of log permeability, as they are better estimated by the method. However, one needs to know at which mode to truncate and this is highly dependent on the covariance matrix of log permeability.

To conclude, we believe ETPF is promising for inverse modelling. However, it struggles for problems with prior substantially different from posterior and numerous accurate observations. In Chapter 3 we introduce an annealing extension of ETPF which is formulated to overcome these limitations.

Chapter 3

Transform-based filtering for Bayesian inverse problems

3.1 Introduction

We consider the inverse problem of inferring unknown parameters in models described by partial differential equations (PDEs), given incomplete noisy data /observations of the model outputs. We adopt the Bayesian approach where the unknowns are random functions with a prescribed prior measure that encompasses our prior statistical knowledge of the unknown. The solution to the Bayesian inversion problem is the posterior, i.e. the conditional distribution of the unknown parameters given the observed data. We can use the posterior to compute estimates of the unknown together with the degree of confidence in those estimates. We are interested in problems where the parameter-to-output map from the underlying PDE model is nonlinear. These are particularly challenging problems since the resulting posterior cannot be obtained analytically even when the prior and the noise distributions are assumed Gaussian. Hence, sampling methods are

required to approximate (expectations under) the posterior which, in turn, is defined on a very high dimensional space after discretisation of the PDEs that define the forward problem.

Markov chain Monte Carlo (MCMC) is the method of choice to sample the Bayesian posterior [51]. In particular, there is a class of MCMC methods constructed in functional settings with mesh-invariant properties suitable for PDE-constrained identification problems [17]. However, the most standard version of these methods often exhibit excessively long correlations (e.g. up to 10^4 [46, 52]), a situation particularly exacerbated with highly-peaked (possibly multimodal) posteriors such as those arising when observational noise is small. Very long MCMC chains (e.g. over 10^7 steps) are thus required (i) to ensure that MCMC fully explores the posterior measure thus capturing possibly multiple modes and (ii) to produce sufficient independent samples to compute accurate posterior statistics. Since every step of MCMC involves at least one PDE solve, these methods become impractical for costly large-scale simulations. While more efficient MCMC can be used to approximate the posterior [9, 61], their proposals often required high-order derivatives of the likelihood which are not available in many applications where the simulator is accessible only in a black-box fashion.

Sequential Monte Carlo (SMC) samplers [23] offer a different sampling approach for approximating the Bayesian posterior. In the context of large-scale Bayesian inversion, adaptive SMC methods construct particle approximations of a sequence of intermediate measures that interpolate (e.g. via tempering) between the prior and the posterior. Particle positions and their weights are adapted on-the-fly to enable a controlled transition between those intermediate measures, thus facilitating to gradually move from a simple prior to a possibly complex posterior. The transition between two intermediate measures involves an importance resampling (IR) step by which the particles are weighted according to the tempered likelihood and then resampled according to those weights. This step is then followed by mutation of particles induced by sampling from a kernel with the IR measure as its invariant measure; this is typically conducted via running MCMC chains with the aforementioned target measure.

Adaptive SMC samplers for solving Bayesian inverse problems have been proposed in [52] and applied for the identification of the initial condition in the Navier-Stokes equations. This work showed that SMC can produce accurate approximations of the Bayesian posterior at a computational cost an order of magnitude smaller than those obtained via state-of-the-art MCMC. The same adaptive SMC sampler was used in [44]

to infer permeability in a moving boundary problem arising in porous media flow. A theoretical framework for adaptive SMC framework was developed in [7] and tested numerically by inferring hydraulic conductivity in a groundwater flow model.

Despite the computational advantages of using SMC samplers, their computational cost still poses severe limitations for its application to practical large-scale inverse problems. The cost of a single iteration (IR+mutation) within SMC is proportional to $J \times N_\mu$ where J is the number of particles and N_μ is the number of mutation MCMC moves. Therefore, each iteration could involve, say, over 10^4 PDE solves even for relatively small J and N_μ (e.g. $J = 10^3$ and $N_\mu = 10$). Hence, if the posterior is complex hence requiring several intermediate measures, the cost of SMC is prohibited unless high performance (HPC) resources are available to scale the cost of SMC with respect to J. While parallelisation is indeed one of the main advantages of SMC, the availability of HPC with $10^4 - 10^5$ processors for typical engineering and geophysical (practical) applications is the exception rather than norm. It is worth mentioning that reducing the cost of SMC via using small number of samples and/or reducing the number of mutation steps can be substantially detrimental to the accuracy of the particle approximation provided by SMC; see for example the work of [44] where SMC with limited number of particles $(10^2 - 10^3)$ results in very poor approximations of the Bayesian posterior. Recent work aimed at reducing the computational cost of SMC samplers includes the development of multilevel versions [6, 49].

In this chapter we investigate the feasibility of an alternative, potentially more computationally affordable approach to approximate the Bayesian posterior within the adaptive tempering SMC setting for Bayesian PDE-constrained inverse problems [7, 52]. The proposed approach consists of replacing the resampling step in SMC with a deterministic linear transformation that maps the system of particles that approximate two consecutive measures. At each iteration step within SMC, the transformation is obtained via solving an optimal transportation problem which, in turn, defines a deterministic coupling between two discrete random variables with realisations defined by the particles and with probabilities determined by their corresponding weights. Replacing resampling by an optimal transformation within Bayesian algorithms was proposed in [73] where it was shown that the linear transport map leads to samples that converge to the posterior measures in large ensemble limit. In the context of data assimilation of partially observed dynamic systems, the idea of replacing IR by optimal transport maps is at the core of the so-called ensemble Transform Particle filter (ETPF) [73, 74]. The novelty

of our approach lies in transferring the application of optimal transport to compute the transition between measures in the tempering scheme within SMC.

Numerous work on data assimilation has shown that, when relatively small numbers of particles are used, ETPF provides more accurate state estimations compared to standard IR-based particle filters due to the sampling errors introduced by resampling. While methods such as ensemble Kalman filter (EnKF) can work well for small ensemble sizes compared to IR-based methods, they rely on Gaussian approximations which are often a severe limitation when the underlying distribution is, for example, multimodal. In contrast, the optimal transport within ETPF does not rely on Gaussian approximations and has been shown to be first order consistent for the mean, and to converge to the posterior measure in the large-ensemble size limit [73]. Here we investigate whether those well known advantages of ETPF can be exploited within the setting of adaptive SMC for Bayesian inversion. As a proof-of-concept we apply the proposed algorithm to a Bayesian elliptic inverse problem arising in subsurface flow. The goal is to infer hydraulic conductivity from pressure measurements. We consider two parameterisations of the conductivity field aimed at assessing the method under two levels of complexity. In the first one we assume that the log-conductivity is a smooth function characterised by a Gaussian random field under the prior. The second parameterisation consists of a channelised permeability that is described by a set of geometric parameters together with two random fields in the regions inside and outside the channel. While the first parameterisation yields posteriors which are relatively well approximated by Gaussians, the second parameterisation can result in multimodal distributions which are more difficult to capture with Gaussian approximations.

We compared the performance of the proposed technique against a fully resolved posterior computed by the preconditioned Crank-Nicolson (pcn)-MCMC with sufficient steps to ensure that a chain is properly converged. We then compare the proposed technique against monomial based SMC as well as an ensemble Kalman inversion (EKI) technique that arises naturally from the adaptive SMC setting. This EKI methodology has been proposed in [11] as an alternative of [45]. Here this approach is modified to incorporate a mutation with the invariant measure.

3.2 Forward and Inverse Problem

We consider Bayesian inversion, which demands formulation of both a forward problem and an inverse problem. The forward problem consists of finding pressure from hydraulic conductivity. The "inverse" problem consists of two parts. The first part is parametrization of hydraulic conductivity by a random variable. The second part is employment of the Bayes rule to obtain the posterior distribution of the random variable from a given prior and a likelihood. The likelihood involves forward problem evaluation. Thus the Bayesian inversion employs the forward problem within the inverse problem.

3.2.1 Forward Model

The forward problem consist of the identification of the hydraulic conductivity, $\kappa(x)$, of a two-dimensional confined aquifer for which the physical domain is $D = [0, 6] \times [0, 6]$. Assuming that the flow within the aquifer is single-phase steady-state Darcy flow, the piezometric head $h(x)$, is given by the solution of [5]

$$-\nabla \cdot \kappa \nabla h \quad = f \qquad \text{in} \quad D \tag{3.1}$$

where f represents a known recharge term. We use the benchmark from [10, 39, 45] where f has the following form

$$f(x_1, x_2) = \begin{cases} 0 & \text{if} \quad 0 < x_2 \leq 4, \\ 137 & \text{if} \quad 4 < x_2 < 5, \\ 274 & \text{if} \quad 5 \leq x_2 < 6. \end{cases} \tag{3.2}$$

and where the boundary conditions are given by

$$h(x_1, 0) = 100, \quad \frac{\partial h}{\partial x}(6, x_2) = 0, \quad -\kappa \frac{\partial h}{\partial x}(0, x_2) = 500, \quad \frac{\partial h}{\partial y}(x_1, 6) = 0. \tag{3.3}$$

We wish to infer $\kappa \in X := \{L^\infty(D; \mathbb{R}) | \operatorname{ess\,inf}_{x \in D} \mathrm{f}(x) > 0\}$ from point observations of h collected at M locations denoted by $\{x_i\}_{i=1}^M \subseteq D$. To this end, we consider smoothed point observations defined by

$$\ell_j(h) = \int_D \frac{1}{2\pi\epsilon^2} e^{-\frac{1}{2\epsilon^2}(x - x_i)^2} h(x) \, \mathrm{d}x$$

where $\epsilon > 0$. Let us define the forward map $G : X \to \mathbb{R}^M$ by

$$G(\kappa) = (\ell_1(h), \ldots, \ell_M(h)). \tag{3.4}$$

which maps permeability into predictions of hydraulic head at measurement locations. Assume that we have noisy measurements of $\{\ell_j(h)\}_{j=1}^M$ of the form

$$y_j = l_j(h) + \eta_j, \qquad j = 1, \ldots, M$$

where η_j represents measurement noise. Our aim is to reconstruct $\kappa \in X$ given $y = (y_1, \ldots, y_M) \in \mathbb{R}^M$.

Parameterisation of permeability

We consider the following two parameterisations of the permeability function $\kappa(x)$ that we wish to identify from observations of the Darcy flow model (3.1)-(3.3).

P1: For the first model the parameter that we consider is simply the natural logarithm of κ, i.e. $u(x) = \log \kappa(x)$.

P2: The second model consist of parameterisation of a piecewise continuous permeability of the form

$$\kappa(x) = \exp(u_1(x))\chi_{D_c}(x) + \exp(u_2(x))\chi_{D\setminus D_c}(x)$$

where $\kappa_1 = \exp(u_1(x))$ and $\kappa_2 = \exp(u_2(x))$ are continuous permeabilities inside and outside a sinusoidal channel with channel domain denoted by D_c. The geometry of the channel is parameterized by five parameters $\{d_i\}_{i=1}^5$ as described in Figure 3.1. The lower boundary of the channel is given by

$$x_2 = d_1 \sin(d_2 x_1/6) + \tan(d_3)x_1 + d_4$$

where we use the notation $x = (x_1, x_2) \in D$ in terms of the horizontal and vertical components. The upper boundary of the channel is given by $x_2 + d_5$. For this permeability model the parameters of interest are comprised in

$$u = (d_1, \cdots, d_5, u_1, u_2)$$

where we assume that each d_i is restricted to an interval $A_i \equiv [d_i^-, d_i^+]$.

We define the following parameter space

$$U = \begin{cases} L^\infty(D;\mathbb{R}) & \text{for P1,} \\ \prod_{i=1}^5 A_i \times L^\infty(D;\mathbb{R}^2) & \text{for P2,} \end{cases}$$

with metric

$$|u|_U = \begin{cases} ||u||_\infty & \text{for P1,} \\ \sum_{i=1}^5 |d_i| + ||u_1||_\infty + ||u_2||_\infty & \text{for P2,} \end{cases}$$

where $\sum_{i=1}^5 |d_i|$, $||u_1||_\infty$ and $||u_2||_\infty$ are of comparable magnitudes.

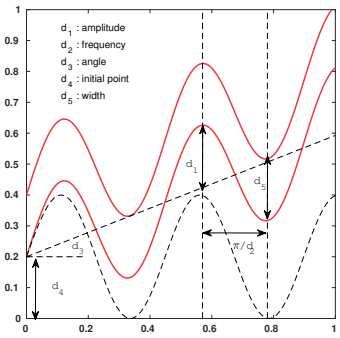

Figure 3.1: Geometrical configuration of channel flow.

The parameterizations described earlier define an abstract map $F : U \to X$ from the space of parameters to the space of admissible permeabilities, via

$$F(u) = \kappa. \tag{3.5}$$

We define the parameter-to-observations map $\mathcal{G} : U \to \mathbb{R}^M$ by $\mathcal{G} = G \circ F$ and reformulate the inverse problem (3.4) in terms of finding the parameter $u \in U$, given $y \in \mathbb{R}^M$ that satisfies

$$y = \mathcal{G}(u) + \eta \tag{3.6}$$

for $\eta = (\eta_1, \ldots, \eta_M) \in \mathbb{R}^M$. The continuity of the parameter-to-observations map $\mathcal{G}$ for this, and more general cases, has been established in [46, 81].

3.2.2 The Bayesian Inverse Problem

In order to address the inverse problem formulated via (3.6) we adopt the Bayesian framework [81] where η is a random vector and u is a random function. We put a prior, $\mu_0(u)$, on the unknown u, and define the random variable $y|u$ under the standard assumption that $\eta \sim N(0, \sigma^2 I)$ independent of u. The solution to the inverse problem in the Bayesian setting is the posterior measure on $u|y$. In the following sections we introduce the prior and likelihood which by the infinite-dimensional framework of [81] ensure that the posterior measure exists and is continuous with respect to appropriate metrics.

The Prior

For P1 we consider a Gaussian prior $\mu_0 = N(m, C)$ with mean m and covariance C. We define C via a correlation function given by the Wittle-Matern correlation function defined by [62]:

$$c(x, y) = \sigma_0^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{|x-y|}{\ell} \right)^{\nu} K_\nu \left(\frac{|x-y|}{\ell} \right), \tag{3.7}$$

where Γ is the gamma function, ℓ is the characteristic length scale, σ_0^2 is an amplitude scale and K_ν is the modified Bessel function of the second kind of order ν. The parameter ν controls the regularity of the samples.

For P2 we assume independence between geometric parameters and log- permeabilities and thus consider a prior of the form

$$\mu_0(du) = \Pi_{i=1}^5 \pi_0^{A_i}(d_i) \otimes N(m_1, C_1) N(m_2, C_2) \tag{3.8}$$

where $\pi_0^A(x)$ is the uniform density defined by

$$\pi_0^A(x) = \begin{cases} \frac{1}{|A|} & x \in A, \\ 0 & x \notin A. \end{cases} \tag{3.9}$$

In expression (3.8) $N(m_1, C_1)$ and $N(m_2, C_2)$ are two Gaussians such as those described earlier in terms of the correlation function from (3.7).

The likelihood

We assume the unknown u is independent of the observational noise $\eta \sim N(0, \sigma^2)$. We note that $y|u \sim N(\mathcal{G}(u), \sigma^2 I)$, hence the likelihood is given by

$$l(u, y) \propto \exp(-\Phi(u, y)), \tag{3.10}$$

where $\Phi(u, y)$ is the data misfit defined by

$$\Phi(u, y) = \frac{1}{2\sigma^2}||y - \mathcal{G}(u)||^2. \tag{3.11}$$

The Posterior

The selection of prior measures from subsection 3.2.2 satisfies $\mu_0(U) = 1$; i.e. samples from μ_0 are in U almost surely [46, 81]. This property, together with the continuity of the forward map defined in subsection 3.2.1, can be used in the Bayesian framework of [46, 81] to conclude that (i) the posterior measure $\mu(u)$ on $u|y$ exists and is absolutely continuous with respect to the prior μ_0; and (ii) μ_0 and has a density with respect to μ_0 given by the following Bayes' rule

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0} = \frac{1}{Z} l(u, y), \tag{3.12}$$

where

$$Z = \int_U l(u, y)\mu_0(\mathrm{d}u). \tag{3.13}$$

3.3 Sequential Monte Carlo for Bayesian inversion

Since we consider a highly nonlinear model, an iterative approach to Bayesian inversion is essential. In the framework of SMC it is performed by tempering (or annealing), when the prior measure is bridged to the posterior measure not at once but through tempered measures. It should be noted that the number of tempered measures is not predefined, which could be a potential computational burden. In order to avoid filter degeneracy both resampling and mutation (or jittering) has to be performed. In the "classical" approach we perform monomial resampling, which we propose to replace by resampling

based on optimal transport.

3.3.1 Adaptive SMC

The SMC approach to Bayesian inversion involves bridging the prior μ_0 and the posterior μ via a sequence of intermediate artificial measures $\{\mu_n\}_{n=0}^N$, with $\mu_N = \mu$, defined by

$$\frac{\mathrm{d}\mu_n}{\mathrm{d}\mu_0}(u) \propto l_n(u,y) \equiv l(u,y)^{\phi_n} \tag{3.14}$$

where $\{\phi_n\}_{n=0}^N$ is a set of tempering parameters that satisfy $0 = \phi_0 < \phi_1 < \cdots < \phi_N = 1$. Expression (3.14) formally implies

$$\frac{\mathrm{d}\mu_n}{\mathrm{d}\mu_{n-1}}(u) = \frac{1}{Z_n} l(u,y)^{(\phi_n - \phi_{n-1})} \tag{3.15}$$

where

$$Z_n \equiv \int_X l(u,y)^{(\phi_n - \phi_{n-1})} \mu_{n-1}(\mathrm{d}u) \tag{3.16}$$

Let us then assume that at the iteration level $n-1$, the tempering parameter ϕ_{n-1} has been specified, and that a set of particles $\{u_{n-1}^{(j)}\}_{j=1}^J$ provides the following approximation (with equal weights) of the intermediate measure μ_{n-1}:

$$\mu_{n-1}^J(u) \equiv \frac{1}{J} \sum_{j=1}^J \delta_{u_{n-1}^{(j)}}(u), \tag{3.17}$$

which is similar or equal to $\mu_{n-1}(u)$, i.e. $\frac{1}{J}\sum_{j=1}^J \delta_{u_{n-1}^{(j)}}(u) \simeq \mu_{n-1}(u)$. Then from (3.16) it follows that

$$Z_n \simeq \sum_{j=1}^J l(u_{n-1}^{(j)}, y)^{(\phi_n - \phi_{n-1})} \tag{3.18}$$

and thus, for any measurable f, we have that

$$\mathbb{E}^{\mu_n}(f(u)) \equiv \int_X f(u)\mu_n(du) = \frac{1}{Z_n} \int_X f(u) l(u,y)^{(\phi_n - \phi_{n-1})} \mu_{n-1}(\mathrm{d}u)$$

$$\simeq \Big[\sum_{j=1}^{J} l(u_{n-1}^{(j)}, y)^{(\phi_n - \phi_{n-1})}\Big]^{-1} \sum_{j=1}^{J} l(u_{n-1}^{(j)}, y)^{(\phi_n - \phi_{n-1})} f(u_{n-1}^{(j)}),$$

$$= \sum_{j=1}^{J} W_n^{(j)} f(u^{(j)}). \tag{3.19}$$

The importance weights for the approximation of μ_n are given by

$$W_n^{(j)} = \mathcal{W}_{n-1}^{(j)}[\phi_n] \equiv \frac{l(u_{n-1}^{(j)}, y)^{\phi_n - \phi_{n-1}}}{\sum_{s=1}^{J} l(u_{n-1}^{(s)}, y)^{\phi_n - \phi_{n-1}}}, \tag{3.20}$$

where ϕ_n is to be chosen adaptively, as described below. From (3.19) we see that the importance (normalized) weights $W_n^{(j)}$ assigned to each particle $u_{n-1}^{(j)}$ define the following empirical (particle) approximation of μ_n:

$$\mu_n^J(u) \equiv \sum_{j=1}^{J} W_n^{(j)} \delta_{u_{n-1}^{(j)}}(u). \tag{3.21}$$

Selection-Resampling Step

From the previous subsection it follows that adaptive SMC requires us to select the tempering parameters ϕ_n so that the two consecutive measures μ_{n-1} and μ_n are sufficiently close for the IS approximating to be accurate. To this end, a common procedure [47] involves imposing a threshold on the effective sample size (ESS) defined by

$$\mathrm{ESS}_n(\phi) \equiv \left[\sum_{j=1}^{J} (\mathcal{W}_{n-1}^{(j)}[\phi])^2\right]^{-1}, \tag{3.22}$$

which, in turn, provides a measure of the quality of the population. In other words, ϕ_n is defined by the solution to

$$\mathrm{ESS}_n(\phi) = J_{\mathrm{thresh}}, \tag{3.23}$$

for a user-defined parameter J_{thresh} on the ESS. A bisection algorithm on the interval $(\phi_{n-1}, 1]$ can be used to solve (3.23) [45]. If $\mathrm{ESS}_n(1) > J_{\mathrm{thresh}}$, then we can simply set $\phi_n = 1$ as no further tempering is thus required.

Once the tempering parameter ϕ_n has been computed via (3.23), normalised weights

(3.20) can be computed. Since some of these can be very low, resampling with replacement according to these weights is then required to discard particles associated with those low weights. After resampling, a new set of equally-weighted particles denoted by $\hat{u}_n^{(j)}$ $(j = 1, \ldots, J)$ provide a particle approximation of the measure μ_n.

Mutation Phase

In order to add diversity to the resampled particles $\hat{u}_n^{(j)}$ computed in the selection-resampling step, a mutation step is included in most SMC methodologies. This mutation consists of sampling from a Markov kernel $\mathcal{K}_n$ with invariant distribution μ_n. This can be achieved by running N_μ steps of an MCMC algorithm that has target distribution equal to μ_n. An example of MCMC suitable for the parameterisation P1 of section 3.2.1 is the preconditioned Crank-Nicolson (pcn)-MCMC [17] displayed in Algorithm 1. This algorithm samples from the target μ_n with reference measure $\mu_0 = N(m, C)$; we recall these two measures are related by (3.12). The resulting particles denoted by $\{u_n^{(j)}\}_{j=1}^J$ $(u_n^{(j)} \sim \mathcal{K}_n(\hat{u}_n^{(j)}, \cdot))$ provide a particle approximation of μ_n in the form

$$\mu_n^J \equiv \frac{1}{J} \sum_{j=1}^{J} \delta_{u_n^{(j)}} \tag{3.24}$$

Convergence of (3.24) to μ_n in the large ensemble size limit can be found in [7]. The complete adaptive SMC sampler is displayed in Algorithm 2.

3.3.2 Optimal Transport within SMC

In this section we assume that $X = \mathbb{R}^K$. We denote U_{n-1} a discrete random variable with realisations $\{u_{n-1}^{(j)}\}_{j=1}^J$ and probabilities $\{W_n^{(j)}\}_{j=1}^J$. We denote U_n the random variable with samples $\{\hat{u}_{n-1}^{(j)}\}_{j=1}^J$ with equal weights. The aim is to replace the resampling step in the method above with resampling that maximizes the covariance between U_{n-1} and U_n. Such a resampling is performed by finding a coupling (joint distribution) between the posterior defined by the weights $\{W_n^{(j)}\}_{j=1}^J$ and the uniform probability density such that it maximizes the covariance between U_{n-1} and U_n. Let us assume that the two consecutive measures μ_{n-1} and μ_n are defined on a measurable space $(\Omega, \mathcal{F})$. A coupling of μ_{n-1} and μ_n consists of a pair (U_{n-1}, U_n) of random variables such that μ_{n-1} is the law of $U_{n-1} : \Omega \to \mathcal{U}$ and μ_n is the law of $U_n : \Omega \to \mathcal{U}$. A coupling is called deterministic if

Algorithm 1 pcn-MCMC to generate samples from a μ_n-invariant Markov kernel with $\mu_0 = N(m, C)$

Select $\beta \in (0, 1)$ and an integer N_μ.
for $j = 1, \ldots, J$ **do**
Initialize $\nu^{(j)}(0) = \hat{u}_n^{(j)}$
while $\alpha \leq N_\mu$ **do**
(1) **pcN proposal**. Propose u_{prop} from

$$u_{\text{prop}} = \sqrt{1-\beta^2}\nu^{(j)}(\alpha) + (1-\sqrt{1-\beta^2})m + \beta\xi, \qquad \text{with } \xi \sim N(0, C)$$

(2) Set $\nu^{(j)}(\alpha+1) = u_{\text{prop}}$ with probability $a(\nu^{(j)}(\alpha), u)$ and $\nu^{(j)}(\alpha+1) = \nu^{(j)}(\alpha)$ with probability $1 - a(\nu^{(j)}(\alpha), u)$, where

$$a(u, v) = \min\left\{1, \frac{l(u, y)^{\phi_n}}{l(v, y)^{\phi_n}}\right\}, \text{ with } l \text{ defined in (3.10)}$$

(3) $\alpha \leftarrow \alpha + 1$
end while
end for

there exists a measurable function $\Psi : \mathcal{U} \to \mathcal{U}$ such that $U_n = \Psi(U_{n-1})$. This measurable function Ψ is called transport map and leads to dependent random variables. Unlike couplings, deterministic couplings do not always exist. On the other hand there may be infinitely many deterministic couplings. An example of a deterministic coupling is an optimal coupling. An optimal coupling is a solution of the Monge-Kantorovitch miminization problem

$$\inf \int_{\mathcal{U}\times\mathcal{U}} c(u_{n-1}, \hat{u}_{n-1}) \mathrm{d}\ell(u_{n-1}, \hat{u}_{n-1}),$$

where $u_{n-1} \in \mathcal{U}$ and $\hat{u}_{n-1} \in \mathcal{U}$. The infimum is taken over all joint probability measures ℓ on the product space $\mathcal{U} \times \mathcal{U}$ (also termed as transport plan) with marginals μ_{n-1} and μ_n and $c(u_{n-1}, \hat{u}_{n-1})$ is a cost function on $\mathcal{U} \times \mathcal{U}$. The joint measures achieving the infimum are called optimal transport plans. The optimal coupling is unique if the measure μ_{n-1} possesses some regularity properties and the cost function $c(u_{n-1}, \hat{u}_{n-1})$ is convex [91]. It appeared that such a coupling simultaneously minimizes the expectation between $||u_{n-1} - \hat{u}_n||^2$ and is defined as the solution of the Monge-Kantorovitch problem

Algorithm 2 SMC algorithm for High-Dimensional Inverse Problems

Let $\{u_0^{(j)}\}_{j=1}^J \sim \mu_0$ be the initial ensemble of J particles.
Define the tunable parameters J_{thresh} and N_μ.
Set $n = 0$ and $\phi_0 = 0$
while $\phi_n < 1$ **do**
$n \to n + 1$
Compute the likelihood (3.10) $l(u_{n-1}^{(j)}, y)$ (for $j = 1, \ldots, J$)
Compute the tempering parameter ϕ_n:
if $\min_{\phi \in (\phi_{n-1}, 1)} \text{ESS}_n(\phi) > J_{\text{thresh}}$ **then**
set $\phi_n = 1$.
else
compute ϕ_n such that $\text{ESS}_n(\phi) \approx J_{\text{thresh}}$
using a bisection algorithm on $(\phi_{n-1}, 1]$.
end if
Computing weights from expression (3.20) $W_n^{(j)} \equiv \mathcal{W}_{n-1}^{(j)}[\phi_n]$
Resample. Let $(p^{(1)}, \ldots, p^{(J)}) \in \mathcal{R}(W_n^{(1)}, \ldots, W_n^{(J)})$, where $\mathcal{R}$ denotes multinomial resampling with replacement.
Set $\hat{u}_n^{(j)} \equiv u_{n-1}^{(p^{(j)})}$ and $W_n^{(j)} = \frac{1}{J}$
Mutation. Sample $u_n^{(j)} \sim \mathcal{K}_n(\hat{u}_n^{(j)}, \cdot)$ via Algorithm 1.
end while
Approximate μ_n by $\mu_n^J \equiv \frac{1}{J} \sum_{j=1}^J \delta_{u_{n,r}^{(j)}}$

with cost function $c(u_{n-1}, \hat{u}_n) = ||u_{n-1} - \hat{u}_n||^2$. Here, $\hat{u}_n$ represents a new set of particles after transformation.

For the coupling of the posterior defined by weights $\{W_n^{(j)}\}_{j=1}^J$ and the uniform probability density, the above described coupling results in a $J \times J$ matrix T^* with non-negative entries T_{ij}^* that satisfy

$$\sum_{i=1}^J T_{ij}^* = \frac{1}{J}, \qquad \sum_{j=1}^J T_{ij}^* = W_i, \tag{3.25}$$

and minimizes

$$\sum_{i,j=1}^J T_{ij} ||u_{n-1}^{(i)} - u_{n-1}^{(j)}||^2 \tag{3.26}$$

for $T_{ij} = T_{ij}^*$. This is a *linear* transport problem of finding J^2 unknowns. Then the

linear transformation gives new samples according to

$$\hat{u}_n^{(j)} := \sum_{i=1}^{J} P_{ij} u_{n-1}^{(j)} \qquad \text{for } j = 1, \dots, J, \tag{3.27}$$

where $P_{ij} = JT_{ij}^*$.

The deterministic optimal transformation (3.27) converges weakly to the solution of the underlying continuous Monge-Kantorovitch problem as $J \to \infty$ [73]. ETPF is first order consistent, since

$$\overline{\hat{u}}_n = \frac{1}{J}\sum_{j=1}^{J} \hat{u}_n^{(j)} = \frac{1}{J}\sum_{j=1}^{J}\sum_{i=1}^{J} P_{ij} u_{n-1}^{(j)} = \sum_{j=1}^{J}\sum_{i=1}^{J} T_{ij}^* u_{n-1}^{(j)} = \sum_{j=1}^{J} W_n^{(j)} u_{n-1}^{(j)}.$$

There also exists a second-order accurate ETPF [22], which however does not satisfy $T_{ij}^* \geq 0$. The main difference between resampling based on optimal transport and monomial resampling is that the former is optimal in the sense of the Monge-Kantorovitch problem, while the latter is non-optimal in that sense.

The computational complexity of finding the minimizer of (3.27) is in general $\mathrm{O}(J^3 \ln J)$, which has been reduced to $\mathrm{O}(J^2 \ln J)$ in [69]. The wall clock time at $J = 100$ is 0.3 seconds for SMC with optimal resampling, while 0.03 seconds for both SMC with monomial resampling and EKI. It can be further improved by employing fast iterative methods for finding approximate minimizers using the Sinkhorn distance [19], which was implemented in [22] for the second-order accurate ETPF. The algorithm of earth's moving distances of [69] is available as both *MATLAB* and *Python* codes and is used here. The complete adaptive optimal transport based SMC sampler is displayed in Algorithm 3.

3.3.3 Gaussian Approximation of SMC via ensemble Kalman inversion

A natural approximation that arises from the adaptive SMC framework described in subsection 3.3.1 involves ensemble Kalman inversion (EKI) [44]. More specifically, let us assume that at the $n-1$ iteration level, we approximate μ_{n-1} with a Gaussian $\hat{\mu}_{n-1} = N(m_{n-1}, C_{n-1})$ where the mean m_{n-1} and covariance C_{n-1} are the empirial mean and covariance of the particles (assumed with equal weights) at the current iteration level.

Algorithm 3 Optimal transport based SMC algorithm for High-Dimensional Inverse Problems

Let $\{u_0^{(j)}\}_{j=1}^J \sim \mu_0$ be the initial ensemble of J particles.
Define the tunable parameters J_{thresh} and N_μ.
Set $n = 0$ and $\phi_0 = 0$
while $\phi_n < 1$ **do**
$n \to n + 1$
Compute the likelihood (3.10) $l(u_{n-1}^{(j)}, y)$ (for $j = 1, \ldots, J$)
Compute the tempering parameter ϕ_n:
if $\min_{\phi \in (\phi_{n-1}, 1)} \text{ESS}_n(\phi) > J_{\text{thresh}}$ **then**
set $\phi_n = 1$.
else
compute ϕ_n such that $\text{ESS}_n(\phi) \approx J_{\text{thresh}}$
using a bisection algorithm on $(\phi_{n-1}, 1]$.
end if
Computing weights from expression (3.20) $W_n^{(j)} \equiv \mathcal{W}_{n-1}^{(j)}[\phi_n]$
Resample based on optimal transport. Compute $\mathcal{D}_{ij} = ||u_{n-1}^{(i)} - u_{n-1}^{(j)}||^2$ (for $i, j = 1, \ldots, J$). Supply $\{\mathcal{D}_{ij}\}_{i,j=1}^J$ and $\{W_n^{(j)}\}_{j=1}^J$ to the earth's moving distances algorithm of Pele & Werman. The output is the coupling $\{T_{ij}^*\}_{i,j=1}^J$.
Compute new samples $\hat{u}_n^{(j)}$ (3.27) and set $W_n^{(j)} = \frac{1}{J}$.
Mutation. Sample $u_n^{(j)} \sim \mathcal{K}_n(\hat{u}_n^{(j)}, \cdot)$ via Algorithm 1.
end while
Approximate μ_n by $\mu_n^J \equiv \frac{1}{J} \sum_{j=1}^J \delta_{u_{n,r}^{(j)}}$

That is,

$$m_{n-1} \equiv \frac{1}{J} \sum_{j=1}^J u_{n-1}^{(j)}, \quad C_{n-1} \equiv \frac{1}{J-1} \sum_{j=1}^J (u_{n-1}^{(j)} - m_{n-1}) \otimes (u_{n-1}^{(j)} - m_{n-1}) \tag{3.28}$$

If we now linearise the forward map around m_{n-1} and replace Frechet derivatives of the forward map with covariances/cross-covariances as in [45], it can be shown that the application to Bayes' rule yields an approximate posterior $\hat{\mu}_n = N(m_n, C_n)$ with mean and covariance given by

$$m_n = m_{n-1} + C_{n-1}^{u\mathcal{G}} (C_{n-1}^{\mathcal{G}\mathcal{G}} + \alpha_n \Gamma)^{-1} (y - \overline{\mathcal{G}}_{n-1}), \tag{3.29}$$

$$C_n = C_{n-1} - C_{n-1}^{u\mathcal{G}} (C_{n-1}^{\mathcal{G}\mathcal{G}} + \alpha_n \Gamma)^{-1} C_{n-1}^{\mathcal{G}u}, \tag{3.30}$$

where Γ is the observation error covariance.

$$C_{n-1}^{u\mathcal{G}} \equiv \frac{1}{J-1}\sum_{j=1}^{J}(u_{n-1}^{(j)} - m_{n-1}) \otimes (\mathcal{G}(u_{n-1}^{(j)}) - \overline{\mathcal{G}}_{n-1}), \tag{3.31}$$

$$C_{n-1}^{\mathcal{G}\mathcal{G}} \equiv \frac{1}{J-1}\sum_{j=1}^{J}(\mathcal{G}(u_{n-1}^{(j)}) - \overline{\mathcal{G}}_{n-1}) \otimes (\mathcal{G}(u_{n-1}^{(j)}) - \overline{\mathcal{G}}_{n-1}), \tag{3.32}$$

with $\overline{\mathcal{G}}_{n-1} \equiv \frac{1}{J}\sum_{j=1}^{J}\mathcal{G}(u_{n-1}^{(j)})$ and where

$$\alpha_n = \frac{1}{\phi_n - \phi_{n-1}}. \tag{3.33}$$

Since we are interested in a particle approximation of $\hat{\mu}_n = N(m_n, C_n)$, we can use the following expression

$$\hat{u}_n^{(j)} = u_{n-1}^{(j)} + C_{n-1}^{u\mathcal{G}}(C_{n-1}^{\mathcal{G}\mathcal{G}} + \alpha_n\Gamma)^{-1}(y_n^{(j)} - \mathcal{G}_n(u_{n-1}^{(j)})), \tag{3.34}$$

where

$$y_n^{(j)} \equiv y + \eta_n^{(j)}, \qquad \eta_n^{(j)} \sim N(0, \alpha_n\Gamma). \tag{3.35}$$

Standard Kalman filter arguments [54] can be used to show that the particle approximation provided by (3.34)-(3.35) converges to $\hat{\mu}_n$ as $J \to \infty$. We note in passing that, within the adaptive SMC framework used here, the regularisation/inflation parameter α_n in formulas (3.33) is computed based on the ESS criteria discussed in subsection 3.3.1.

It is important to emphasize that, in general, the approximate Gaussian measure $\hat{\mu}_n$ coincides with μ_n only when the forward map is linear and the prior μ_0 is Gaussian. The approximation provided by EKI will deteriorate when we depart from Gaussian-linear assumptions. Therefore, we propose to conduct MCMC mutations to each of the particles in (3.34) with the aim of improving the approximation of each posterior measure μ_n. The complete EKI-based algorithm is displayed in Algorithm 4. We recognise that this is only an ad-hoc approach for which exact sampling of the posterior (as $J \to \infty$) is not ensured. A more rigorous (i.e. fully-Bayesian approach) that we leave for future work is to use EKI in the proposal design for the importance sampling step within SMC;

this is done for data assimilation settings in [16].

Algorithm 4 EKI approximation to SMC

Let $\{u_0^{(j)}\}_{j=1}^J \sim \mu_0$ be the initial ensemble of J particles.
Define the tunable parameters J_{thresh} and N_μ.
Set $n = 0$ and $\phi_0 = 0$
while $\phi_n < 1$ **do**
  $n \to n + 1$
  Compute the likelihood (3.10) $l(u_{n-1}^{(j)}, y)$ (for $j = 1, \ldots, J$)
  Compute the tempering parameter ϕ_n:
  if $\min_{\phi \in (\phi_{n-1}, 1)} \text{ESS}_n(\phi) > J_{\text{thresh}}$ **then**
    set $\phi_n = 1$.
  else
    compute ϕ_n such that $\text{ESS}_n(\phi) \approx J_{\text{thresh}}$
    using a bisection algorithm on $(\phi_{n-1}, 1]$.
  end if
  Generate particles $\{\hat{u}_n^{(j)}\}_{j=1}^J$ according to (3.34).
  Mutation. Sample $u_n^{(j)} \sim \mathcal{K}_n(\hat{u}_n^{(j)}, \cdot)$ via Algorithm 1.
end while
Approximate μ_n by $\mu_n^J \equiv \frac{1}{J} \sum_{j=1}^J \delta_{u_{n,r}^{(j)}}$

3.4 Numerical experiments

In this section we perform numerical experiments to infer P1 and P2 parameters. We compare optimal transport based SMC to both monomial based SMC and EKI, which we denote *optimal*, *monomial*, and *Kalman*, respectively. We analyze method performance with respect to a pcn-MCMC solution, which we denote as *reference*. We combine 50 independent chains each of length 10^6 and 10^5 burn-in period and thinning 10^3.

Observations of pressure were obtained from the true permeability with observation noise from normal distribution with zero mean and standard deviation of 2% of L^2-norm of the true pressure. We should note that both the true random variable and an initial ensemble of parameterized permeability are drawn from the same prior distribution, as

the prior includes knowledge about geological properties. However, the true solution is computed on a fine grid and an initial guess on a coarse grid, which is half the resolution of the fine grid. The uncertain parameter for P1 inference has the dimension of the coarse grid, i.e. $4900 = 70^2$. For P2 inference, we first define both inside and outside the channel log-permeability over the entire domain and then by employing the geometrical parameters generate the channel configurations. Since, we generate the log-permeability based on the correlation function defined by equation (3.7), this approach helps us avoid any additional error due to loss of regularity due to sudden jump in log-permeability over adjacent grids. Thus, the uncertain parameter for P2 inference has the dimension twice that of the coarse grid plus the dimension of the geometrical parameters, i.e. $5005 = 50^2 + 50^2 + 5$.

For log-permeability parameters, the prior is a normal distribution with mean 5 for P1, and for P2 with mean 15 outside the channel and 100 inside the channel. For geometrical parameters, the prior is uniform: $d_1 \sim U[0.05\times 6,\ 0.35\times 6]$, $d_2 \sim U[\pi/2,\ 6\pi]$, $d_3 \sim U[-\pi/2,\ \pi/2]$, $d_4 \sim U[0,\ 6]$, $d_5 \sim U[0.02\times 6,\ 0.7\times 6]$. For tempering we choose the effective ensemble size threshold $J_{\text{thresh}} = J/3$ and for mutations the length of Markov chain $N_\mu = 10$ to save computational costs. For P2, we use the Metropolis-within-Gibbs methodology of [46] to separate geometrical parameters and log-permeability parameters within the mutation step, since it allows to better exploit the structure of the prior. The proposal design for the geometric parameters within the Metropolis-within-Gibbs consist of local moves within the intervals of the prior with a step size that we tune to achieve acceptance rates between 20% and 30%. Geometrical parameters that fall outside those intervals are projected back via a projection that preserves reversibility of the proposal with respect to the prior [46]. We perform numerical experiments with different ensemble sizes of 100, 500, and 1000. We perform 10 simulations with different realizations of the initial ensemble to check the robustness of results.

For log-permeability, we compute the L^2 norm of the error in the mean with respect to the reference

$$\text{Error} = ||\bar{u} - \bar{u}^{\text{ref}}||, \quad \text{where} \quad \bar{u} = \frac{1}{J}\sum_{j=1}^{J} u^{(j)}.$$

We investigate the performance of the proposed approach to approximate the marginal posterior, $p(d_i)$, of each geometric parameter d_i $(i = 1, \ldots, 5)$ defined in parameterisation P2. To this end, we compute the Kullback-Leibler divergence with respect to the

reference/true posterior marginal (denoted by $p^{\text{ref}}(d_i)$) computed via MCMC:

$$\mathrm{D}_{\mathrm{KL}}(p^{\text{ref}} \parallel p) = \sum_{j=1}^{J_{\mathrm{b}}} p^{\text{ref}}(d_i^j) \log \frac{p^{\text{ref}}(d_i^j)}{p(d_i^j)}, \tag{3.36}$$

where $J_{\mathrm{b}} = J/10$ is a chosen number of bins and $p(d_i^j)$ is approximated by the weights. The results (median, 25 and 75 percentiles) that we report below for both the error in the mean and the KL divergence are computed over 10 experiments corresponding to independent choices of the prior ensemble.

3.4.1 Numerical inference for P1

For P1, we perform a numerical experiment using 36 uniformly distributed observations. In Figure 3.2, we plot error in the mean log-permeability with respect to reference. We observe that while optimal transport based SMC outperforms monomial based SMC for all ensemble sizes, EKI outperforms both SMC methods. This is due to the nature of the P1 parametrization and only two degrees of freedom (mean and variance) of EKI. In Figure 3.3, we plot mean log-permeability for a simulation with smallest error at

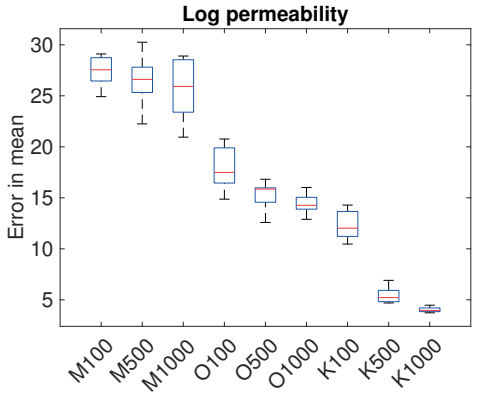

Figure 3.2: Box plot of the error in the mean log-permeability for P1 inference. Central mark is the median, edges of the box are the 25th and 75th percentiles, whiskers extend to the most extreme datapoints over 10 independent simulations. On x-axis numbers stand for ensemble sizes, M stands for monomial based SMC, O for optimal transport based SMC, and K for EKI.

ensemble size 100 and reference mean log-permeability. We see that monomial based

SMC gives a less smooth estimation compared to optimal transport based SMC, EKI, and reference, which leads to larger error.

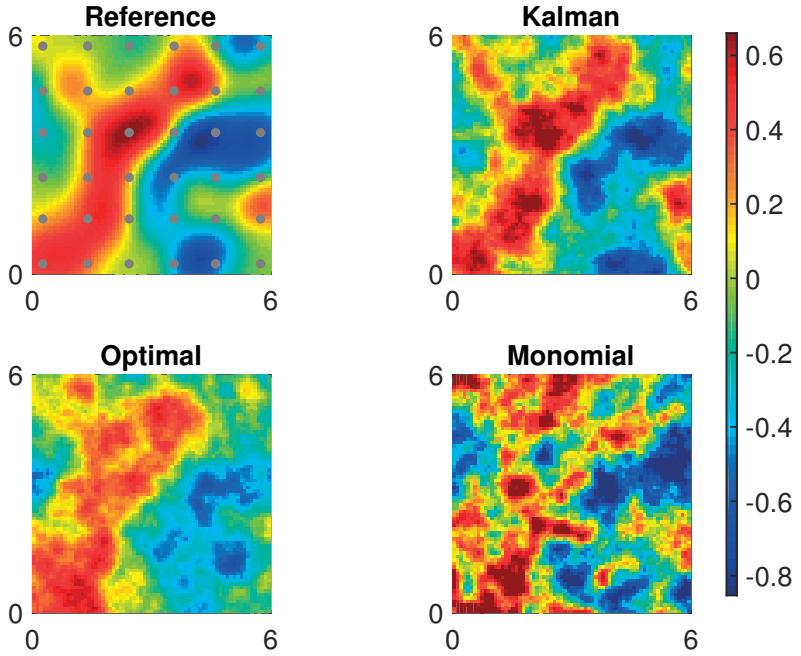

Figure 3.3: Mean log-permeability for P1 inference for the lowest error at ensemble size 100. Observation locations are shown in circles.

For ensemble sizes considered here, the number of tempering steps on average is 15 for optimal transport based SMC, and 17 for both monomial based SMC and EKI. Thus in terms of computational cost optimal transport based SMC is equivalent to monomial based SMC, since the computational complexity of the forward model is higher than $O(J \ln J)$.

3.4.2 Numerical inference for P2

For P2, we perform a numerical experiment using 9 uniformly distributed observations. For ensemble size considered here, the number of tempering steps on average is 8 for EKI, and 7 for both optimal transport based SMC and monomial based SMC. In Figure 3.4, we plot error in the mean log-permeability with respect to reference for permeability

outside channel on the left and for permeability inside channel on the right. We observe that while optimal transport based SMC still outperforms monomial based SMC for all ensemble sizes, it is now comparable to EKI. This is due to a small number of observations. In Figures 3.5–3.6, we plot mean log-permeability for a simulation with

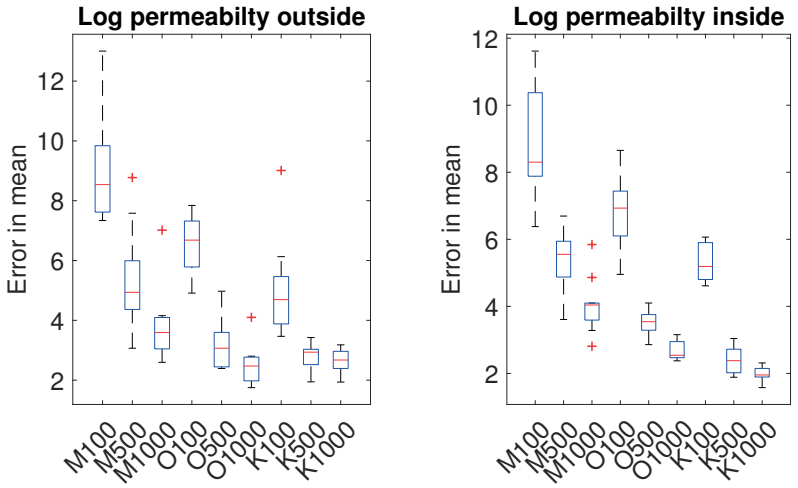

Figure 3.4: Box plot of the error in the mean log-permeability for P2 inference. Central mark is the median, edges of the box are the 25th and 75th percentiles, whiskers extend to the most extreme datapoints, and crosses are outliers over 10 independent simulations. On the left: outside channel, on the right: inside channel. On x-axis numbers stand for ensemble sizes, M stands for monomial based SMC, O for optimal transport based SMC, and K for EKI.

smallest error at ensemble size 100 and reference mean log-permeability for permeability outside channel and for permeability inside channel, respectively. We see that monomial based SMC gives a less smooth estimation compared to optimal transport based SMC, EKI, and reference, which leads to larger error.

In Figure 3.7, we show posterior estimations of geometrical parameters. We see that all the parameters except amplitude and width exhibit strongly non-Gaussian behaviour. In Figure 3.9, we show a trace plot of the geometrical parameter frequency using a chain of MCMC to ensure proper mixing of chain. We can observe from this figure that the two modes are being sampled within each chain and that the chain is properly mixed. In Figure 3.8, we plot KL divergence for geometrical parameters. We observe that EKI performs better than optimal transport based SMC for amplitude and width, while

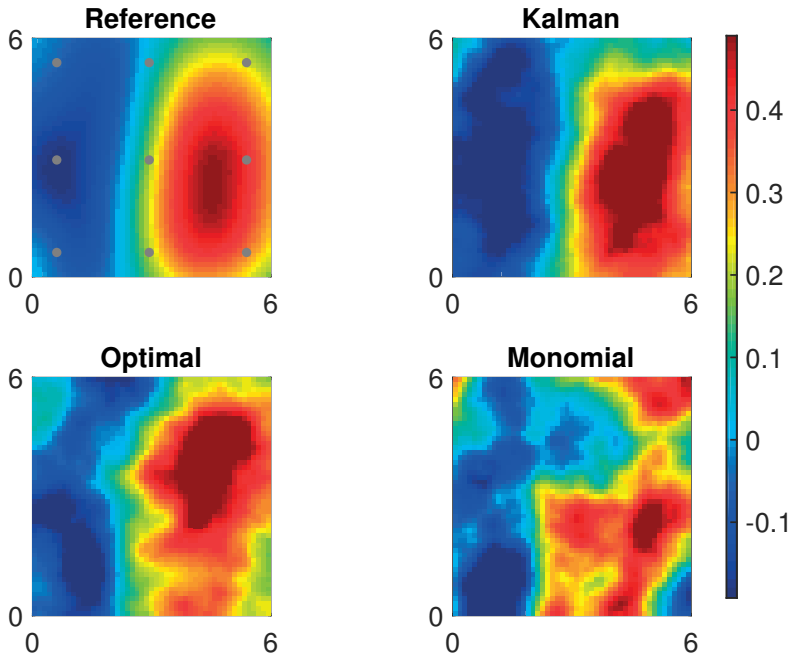

Figure 3.5: Mean log-permeability outside channel for P2 inference for the lowest error at ensemble size 100. Observation locations are shown in circles.

worse for other parameters. We should note that the two different modes of frequency shown in Figure 3.7 provide two significantly different channel configurations, thus it is important to correctly estimate the pdf. Monomial based SMC performs comparably to optimal transport based SMC though not consistently better or worse. We should recall, however, that optimal transport based SMC outperforms monomial based SMC for log-permeability both inside and outside channel. In Figure 3.10, we show mean field of permeability over the channelized domain for the lowest error at ensemble size 1000.

3.5 Conclusions

Accurate estimation of the posterior distribution of uncertain model parameters of strongly nonlinear problems remains a challenging problem. Parameters are high dimensional, they are not observed, and they do not have a dynamical equation. Moreover,

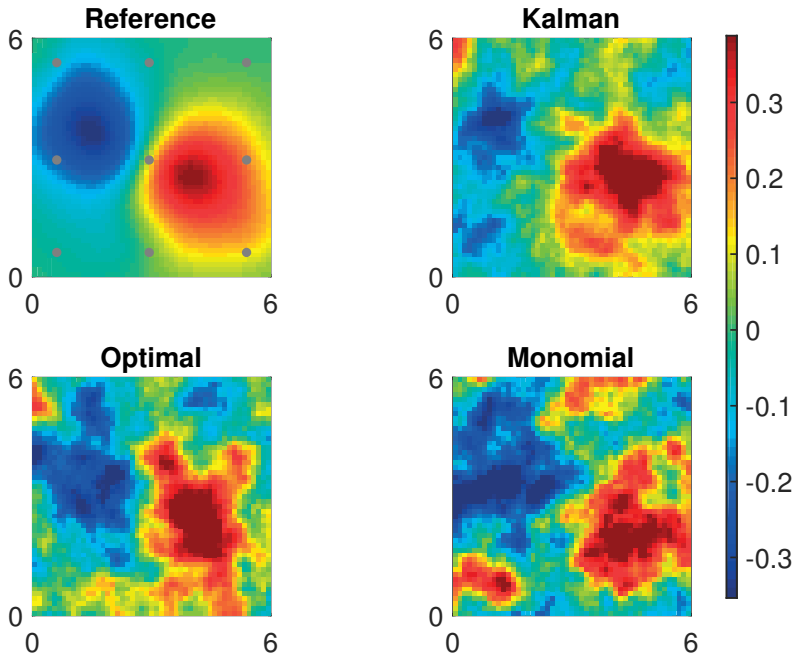

Figure 3.6: Mean log-permeability inside channel for P2 inference for the lowest error at ensemble size 100. Observation locations are shown in circles.

due to nonlinearity of models even a Gaussian prior of parameters might result in non-Gaussian posterior. Since MCMC is computationally unfeasible for high-dimensional problems, adaptive SMC is an alternative to estimate posterior distributions in the Bayesian framework. However, adaptive SMC still requires large ensembles.

In order to reduce computational cost, we proposed to introduce optimal transport based resampling from [73] to adaptive SMC. Optimal transport based resampling creates new samples by maximizing variance between prior and posterior. It has been already shown for state estimation and parameter estimation with low dimension, that a particle filter with optimal transport based resampling outperforms a particle filter with monomial based resampling. As it was aimed to estimate time-evolving model states of chaotic systems, simple inflation was sufficient to mutate particles.

Here we have adopted optimal transportation to elliptic Bayesian inverse problems. We have shown that optimal transport-based SMC has a high potential for Bayesian

inversion of high-dimensional parameters. The parameterisation of the channelised permeability was particularly useful since it involves geometric parameters with marginal posteriors that display non-Gaussian features (e.g. bimodality in the frequency parameter; see Figure 3.7) which are often difficult to characterise via EKI. Indeed, for this case the proposed approach provides more accurate approximations to the marginal posteriors (quantified via KL divergence) than those approximated with EKI. Compared to the standard monomial-based SMC we did not observe substantial differences in the level of approximation of the aforementioned marginals. However, the proposed transport-based SMC outperforms the monomial-based version in approximating the high-dimensional (marginal) posteriors of the two spatially-variable log-permeability fields that we infer in the present setting (measured in terms of the error in the mean and variance).

Moreover, optimal transport based SMC still underestimates variance (not shown), which could be improved by considering second order consistent optimal transport resampling instead of first order. However, second order consistent optimal transport resampling does not necessarily provide non-negative transformations. Finally, optimal transport resampling does not need to be restricted to finite dimensions, at least theoretically [14], with the challenge of finding such a minimizer computationally.

So far in this thesis, we have considered permeability being the only source of model error. However, there exists more than one source of model error, e.g. model reduction, discretization error, etc. In Chapter 4, we consider the Darcy flow problem where both permeability and boundary condition have uncertainties. We compare tempered adaptive SMC to regularized ensemble Kalman filter.

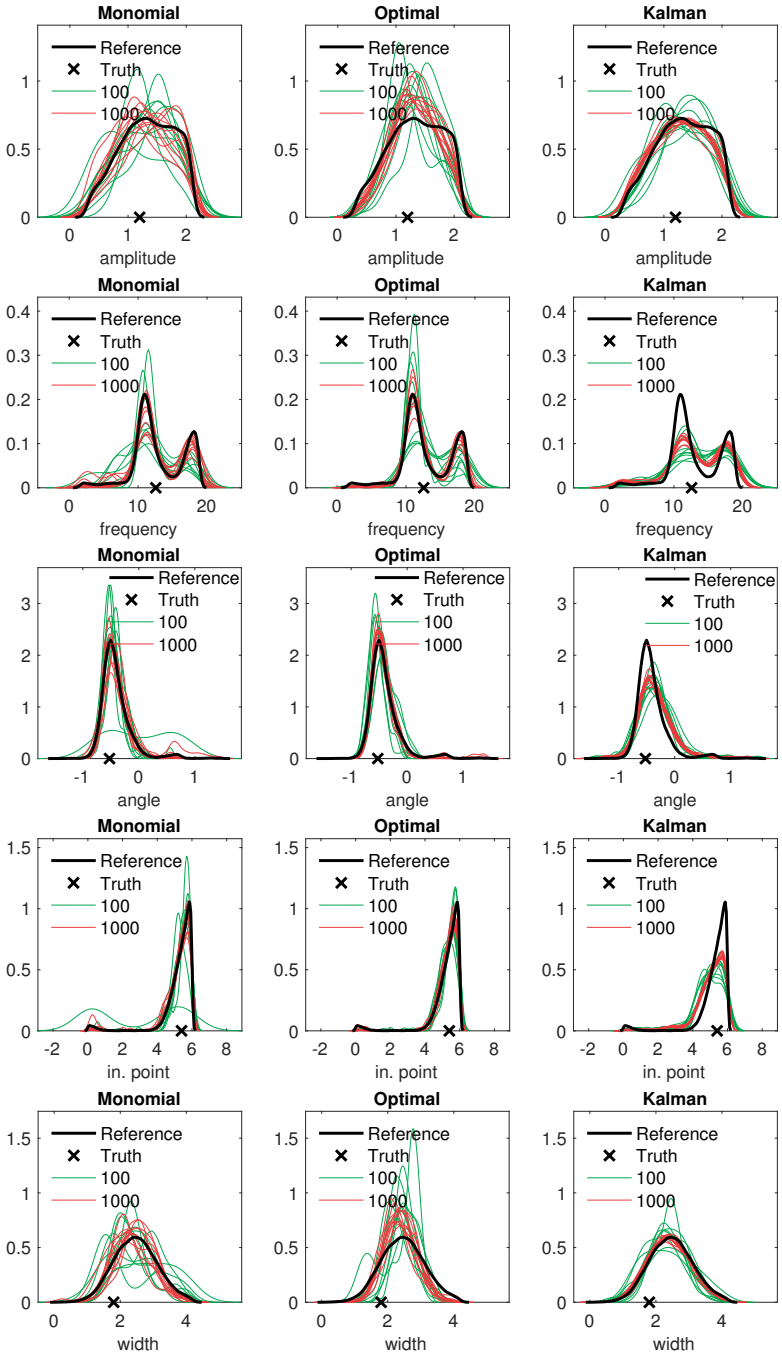

Figure 3.7: Posterior of geometrical parameters for P2 inference. In black is reference, in green 10 simulations of ensemble size 100, in red 10 simulations of ensemble size 1000. The true parameters are shown as black cross.

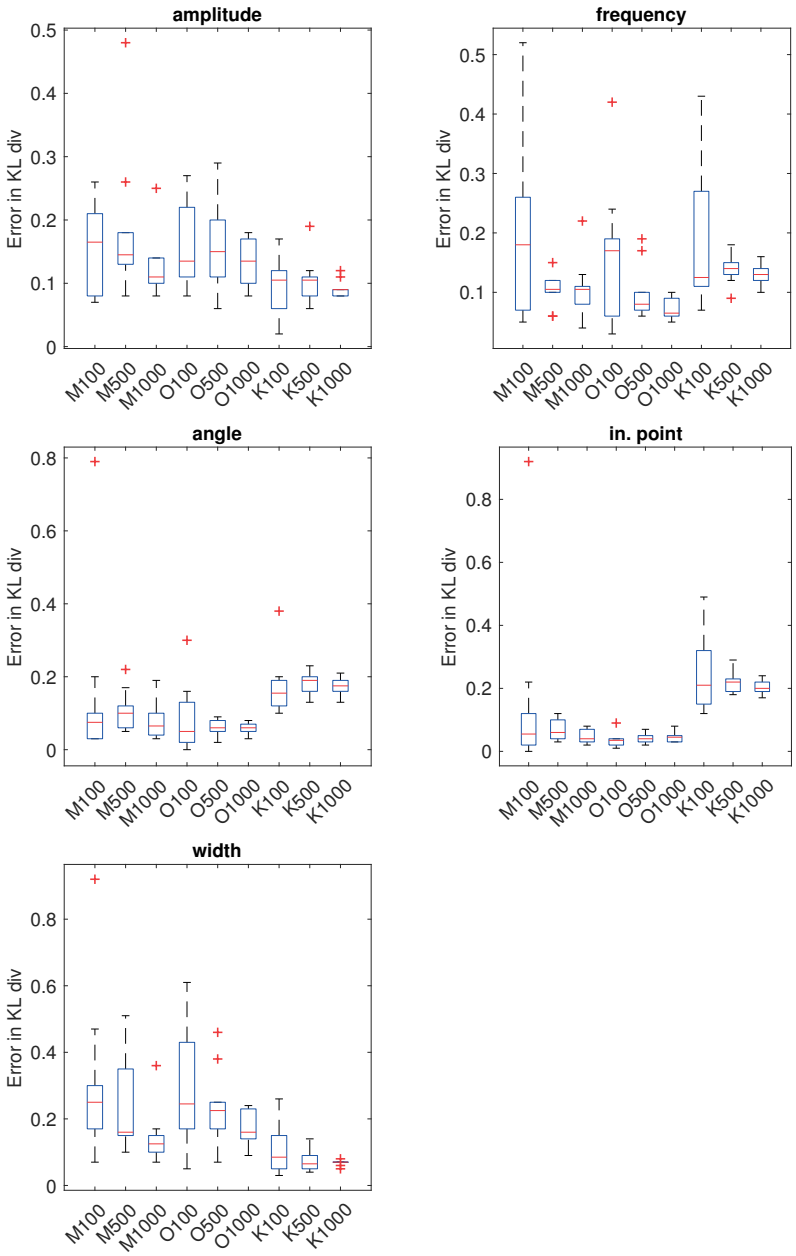

Figure 3.8: Box plot of KL divergence for geometrical parameters for P2 inference. Central mark is the median, edges of the box are the 25th and 75th percentiles, whiskers extend to the most extreme datapoints, and crosses are outliers. On x-axis numbers stand for ensemble sizes, M stands for monomial based SMC, O for optimal transport based SMC, and K for EKI.

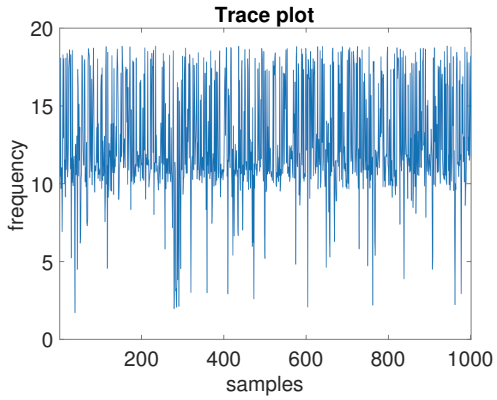

Figure 3.9: Trace plot of frequency from a pcn-MCMC chain.

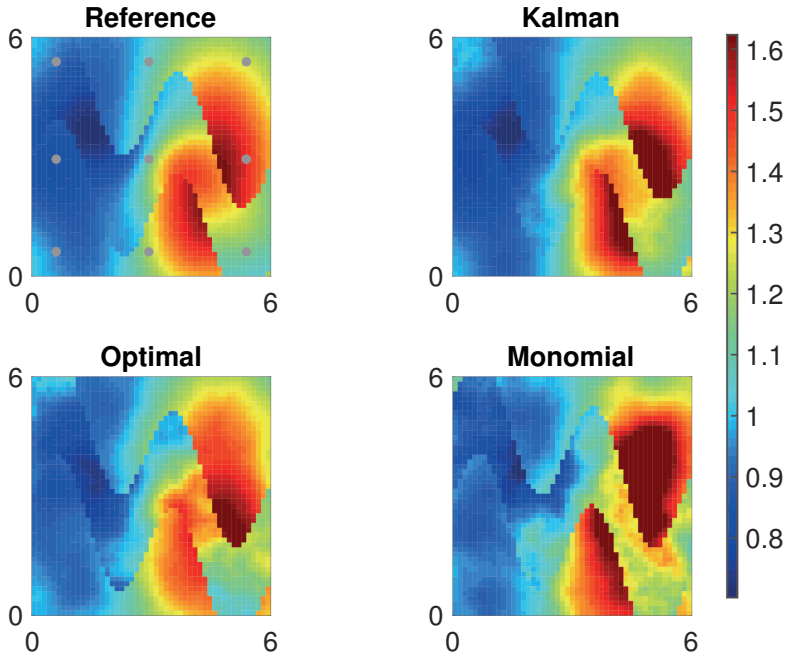

Figure 3.10: Mean permeability for P2 inference for the lowest error at ensemble size 1000. Observation locations are shown in circles.

Chapter 4

Tempered Particle filter for Elliptical Inverse problem with Uncertain Boundary Condition

4.1 Introduction

Ensemble-based data assimilation deals with estimation of uncertain parameters and states of a model constrained by available observations using an ensemble. It is widely employed in many fields, for example meteorology [40] and reservoir engineering [27]. In meteorology one is interested in estimation of uncertain initial conditions of a high-dimensional chaotic system, in reservoir engineering—of estimating high-dimensional uncertain parameters, of permeability for example, of a deterministic non-chaotic system.

However, sources of model error are not limited to random coefficients of a PDE. In inverse problems, another source of model error is model reduction for example, where a complex model is replaced by a simple one. It has been acknowledged, i.e [51], that

accounting for model error in data assimilation greatly improves parameter-state estimation. Recent advances in accounting for model error in ensemble-based data assimilation include: extension of iterative ensemble Kalman filter to include additive model error [79], addition of model error in the randomized maximum likelihood to correctly sample the posterior without marginalization [65], non-additive though Gaussian model error update for Bayesian inversion [20], and adaptation of machine learning techniques for data assimilation [56]. By additive model error we mean that if G is an erroneous approximation of the true model G^{true}, then $G(\cdot) = G^{\text{true}}(\cdot) + q$ where q is model error.

However, most of these works have considered either additive model error or Gaussian model error, with the sole exception of [65] where Gaussian anamorphosis was used. However, for high-dimensional problems finding a transformation to multivariate Gaussian probability is computationally challenging. Though the assumption of additive Gaussian model error simplifies an optimization problem, model error is not limited to being additive nor Gaussian. Therefore, it is essential for a data assimilation method to account for model error in a most general way. A straightforward example of such a data assimilation method is Markov Chain Monte Carlo (MCMC). However, for high-dimensional systems it is impractical.

An alternative to MCMC is particle filtering [24]. Particle filtering is based on proposing an ensemble from a prior that is not necessarily close to the target posterior and to correct for this mismatch by computing importance weights. The importance weights are defined as a function of ensemble estimations and available observations. The ensemble is then resampled according to the estimated posterior. Particle filtering in its original form worked only for low-dimensional problems. However, due to recent advances of employing localization [70, 71] it has been proven competitive for high-dimensional problems as well.

There are different approaches to resampling in particle filtering, but most of them are stochastic. An ensemble transform particle filter [74] employs *deterministic* resampling, which reduces sampling error and thus needs a smaller ensemble than a typical particle filter. It also has a localized version. In Chapter 2, we have employed the method to solve an inverse problem of uncertain permeability. We have shown that though localization makes the ensemble transform particle filter deteriorate a posterior estimation of the leading modes, it makes the method applicable to high-dimensional problems. In Chapter 3, instead of localization we have implemented tempering to the ensemble transform particle filter (TETPF). We have shown that iterations based on

temperatures [52, 64] handle notably strongly nonlinear cases and that TETPF is able to predict multimodal distributions for high-dimensional problems.

In this chapter, we consider a steady-state single-phase Darcy flow model. This groundwater model was first used as benchmark for inverse modelling in [10]. It has been used as a test model for the identification of parameters, for example with iterative regularization methods [39], an ensemble Kalman approach [45], and in our previous work with particle filtering [77, 78]. In this chapter, in addition to uncertain log permeability defined as a Gaussian process, we assume an error in boundary conditions that is non-Gaussian distributed. We note that error in boundary conditions of the groundwater model gives a nonlinear response, thus model error is non-additive. We compare TETPF to a regularized ensemble Kalman filter [43].

The regularized ensemble Kalman filter is a robust ensemble-based data assimilation method, which assumes Gaussian probabilities. It solves an optimization problem for the mean and approximates variance with an ensemble. It has been employed for history matching applications for example in [57, 58]. It has been shown that the ensemble Kalman filter is able to estimate skewed probabilities with frequent observations [26, 31]. However, it fails to estimate multimodal probabilities, e.g. [78].

Our goal is to investigate whether pressure estimation is sensitive to uncertain boundary conditions, and to compare T(L)ETPF to R(L)EnKF (with and without localization). We assume that we have information regarding the sources of errors. That is, we assume we have precise knowledge that the uncertainties arises from boundary conditions and permeabilties. We further assume that we have information about the statistical properties of errors and the prior probability density function. This is an idealized setting, that however allows to check sensitivity of a data assimilation method to uncertainty in boundary conditions, which is also relevant for practical applications.

4.1.1 Bayesian Inference

Both T(L)ETPF and R(L)EnKF are based on Bayesian inference. Assume $\boldsymbol{u} \in \tilde{\mathcal{U}}$ and $\boldsymbol{q} \in \tilde{\mathcal{Q}}$ are two independent random variables. Later on we will have different assumptions on $\boldsymbol{u}$ and $\boldsymbol{q}$. We denote by $g\colon \tilde{\mathcal{U}} \times \tilde{\mathcal{Q}} \to \mathcal{Y}$ the nonlinear forward operator that arises from a model under consideration. In other words, g maps the space $\tilde{\mathcal{U}} \times \tilde{\mathcal{Q}}$ of uncertain quantities to the observation space $\mathcal{Y}$ defined in terms of observable quantities,

which are related to the solution of the model as

$$y = g(u, q).$$

Assume that $\boldsymbol{y}_{\text{obs}} \in \mathcal{Y}$ is an observation of y. Then according to the Bayes' formula

$$\pi(\boldsymbol{u}, \boldsymbol{q}|\boldsymbol{y}_{\text{obs}}) \propto \pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{u}, \boldsymbol{q})\pi(\boldsymbol{u})\pi(\boldsymbol{q}) \tag{4.1}$$

up to a constant of normalization, where π is a probability density function.

For any smooth function $f : \tilde{\mathcal{U}} \to \tilde{\mathcal{U}}$, its expectation is defined as

$$\overline{f(\boldsymbol{u})} = \int d\boldsymbol{u} f(\boldsymbol{u}) \int d\boldsymbol{q} \pi(\boldsymbol{u}, \boldsymbol{q}|\boldsymbol{y}_{\text{obs}}). \tag{4.2}$$

It is common, e.g. [31], to express the joint probability density function as

$$\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{u}, \boldsymbol{q}) = \int \pi(\boldsymbol{y}_{\text{obs}}, \boldsymbol{y}|\boldsymbol{u}, \boldsymbol{q}) d\boldsymbol{y} = \int \pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{y})\pi(\boldsymbol{y}|\boldsymbol{u}, \boldsymbol{q}) d\boldsymbol{y} \tag{4.3}$$

and $\pi(\boldsymbol{y}|\boldsymbol{u}, \boldsymbol{q}) = \delta(\boldsymbol{y} - g(\boldsymbol{u}, \boldsymbol{q}))$, where the transition density is the Dirac delta function.

4.2 Tempered Ensemble Transform Particle Filter

The goal of the Bayesian approach is to compute the posterior given by Eq. (4.1)–(4.3). Sequential Monte Carlo (SMC) is an approximation of the Bayesian posterior. An SMC method creates a finite sample from a prior, that is easy to sample from, and corrects for the differences between the prior and the posterior by computing so-called importance weights. Finally, a resampling is performed according to those weights in order to create a new sample.

4.2.1 Importance weights

We consider discrete random variables and define $\mathcal{U} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_M\} \subset \tilde{\mathcal{U}}$, $\boldsymbol{u}_i \in \mathbb{R}^n$, and $\mathcal{Q} = \{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_M\} \subset \tilde{\mathcal{Q}}$, $\boldsymbol{q}_i \in \mathbb{R}^m$. The model has unknown quantities $\boldsymbol{u}^{\text{true}} \in \mathbb{R}^n$ and $\boldsymbol{q}^{\text{true}} \in \mathbb{R}^m$ that we wish to estimate from noisy observations $\boldsymbol{y}_{\text{obs}} \in \mathbb{R}^\kappa$, where $\kappa < n$ and $\kappa < m$.

$$\boldsymbol{y}_{\text{obs}} := g(\boldsymbol{u}^{\text{true}}, \boldsymbol{q}^{\text{true}}) + \eta,$$

where $\eta \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R})$ with $\boldsymbol{R}$ being a known covariance matrix of the observation noise. The conditional probability density function is given by

$$\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{y}) \propto \exp\left[-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{y}_{\text{obs}})'\boldsymbol{R}^{-1}(\boldsymbol{y}-\boldsymbol{y}_{\text{obs}})\right].$$

Thus a discrete approximation to Eq. (4.3) is

$$\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{u},\boldsymbol{q}) \propto \exp\left[-\frac{1}{2}\left(g(\boldsymbol{u},\boldsymbol{q})-\boldsymbol{y}_{\text{obs}}\right)'\boldsymbol{R}^{-1}\left(g(\boldsymbol{u},\boldsymbol{q})-\boldsymbol{y}_{\text{obs}}\right)\right],$$

where $'$ denotes the transpose. We assume the priors $\pi(\boldsymbol{u})$ and $\pi(\boldsymbol{q})$ are uniform, then denoting $\boldsymbol{v} = [\boldsymbol{u}\ \boldsymbol{q}]'$ the expectation of a function f of $\boldsymbol{v}$ is

$$\overline{f(\boldsymbol{v})} \approx \sum_{i=1}^{M} f(\boldsymbol{v}_i) w_i.$$

Here the importance weights are

$$w_i = \frac{h(\boldsymbol{v}_i)}{\sum_{j=1}^{M} h(\boldsymbol{v}_j)}, \quad \text{where} \quad h(\boldsymbol{v}) = \exp\left[-\frac{1}{2}\left(g(\boldsymbol{u},\boldsymbol{q})-\boldsymbol{y}_{\text{obs}}\right)'\boldsymbol{R}^{-1}\left(g(\boldsymbol{u},\boldsymbol{q})-\boldsymbol{y}_{\text{obs}}\right)\right]. \tag{4.4}$$

4.2.2 Tempering

An SMC method suffers when the likelihood $h(\boldsymbol{v})$ Eq. (4.4) is peaked, which could be due to very accurate data, amount of data or when the prior poorly approximates the posterior. A tempering iterative approach tackles this problem by introducing temperatures $0 = \phi_0 < \cdots < \phi_T = 1$ and corresponding bridging likelihoods $h(\boldsymbol{v})^{(\phi_t - \phi_{t-1})}$ for $t = 1, \ldots, T$. A tempering parameter ϕ_t is typically chosen based on effective ensemble size

$$\text{ESS}_t(\phi_t) := \left[\sum_{i=1}^{M} w_i^2(\phi_t)\right]^{-1}, \tag{4.5}$$

such that ESS does not drop below a certain threshold $1 < M_{\text{thresh}} < M$. A bisection algorithm on the interval $(\phi_{t-1}, 1]$ can be used to solve (4.5) [45]. If $\text{ESS}_t(1) > M_{\text{thresh}}$, then we can simply set $\phi_t = 1$ as no further tempering is thus required.

4.2.3 Deterministic resampling

In order to avoid filter degeneracy, each tempering iteration t needs to be supplied with resampling. Resampling is typically performed by a stochastic approach, which introduces an additional error. In TETPF a tempering iteration t is accompanied by a *deterministic* resampling based on optimal transportation. This resampling transforms particles with weights defined in terms of bridging likelihood

$$w_i^{(t)} = \frac{h\left(\boldsymbol{v}_i^{(t)}\right)^{(\phi_t - \phi_{t-1})}}{\sum_{j=1}^{M} h\left(\boldsymbol{v}_j^{(t)}\right)^{(\phi_t - \phi_{t-1})}},$$

where $h(\boldsymbol{v})$ is from 4.4, to particles with uniform weights $1/M$ by maximizing the correlation between the particles. Thus, the optimal transport $\boldsymbol{S}$ is an $M \times M$ matrix with s_{ij} that satisfy

$$s_{ij} \geq 0, \qquad \sum_{i=1}^{M} s_{ij} = \frac{1}{M}, \qquad \sum_{j=1}^{M} s_{ij} = \frac{h\left(\boldsymbol{v}_i^{(t)}\right)^{(\phi_t - \phi_{t-1})}}{\sum_{j=1}^{M} h\left(\boldsymbol{v}_j^{(t)}\right)^{(\phi_t - \phi_{t-1})}}, \tag{4.6}$$

and minimizes the cost function

$$\sum_{i,j=1}^{M} s_{ij} \left\| \boldsymbol{v}_i^{(t)} - \boldsymbol{v}_j^{(t)} \right\|^2. \tag{4.7}$$

This gives rise to a resampling with replacement and a stochastic transport matrix $\boldsymbol{S}$. In order to have a deterministic optimal transformation the following proposal is adopted

$$\tilde{\boldsymbol{v}}_j = M \sum_{i=1}^{M} \boldsymbol{v}_i^{(t)} \tilde{s}_{ij} \quad \text{for} \quad j = 1, \ldots, M, \tag{4.8}$$

where $\tilde{s}_{ij}$ is a solution to the optimization problem Eq. (4.6)–(4.7). To solve the linear transport problem for multivariate variables Eq. (4.6)–(4.7), we use *FastEMD* algorithm of [69]. Its computational complexity is of order $M^2 \ln M$, and the algorithm is available as a *MATLAB* and a *Python* subroutine.

Localization

The Ensemble Transform Particle Filter, as any particle filter, does not have assumptions about the posterior. Therefore it still demands a large ensemble. For high-dimensional problems this is computationally unfeasible. Hence one has to decrease the number of degrees of freedom, i.e. by distance-based localization of [74, 77] abbreviated here LETPF.

Assume we have a numerical grid of $N \times N$ size with grid cells denoted by X^l for $l = 1, \ldots, N^2$. Assume that the uncertain parameter $\boldsymbol{u}$ is not grid-based. We assume, however, that there exists a matrix $\mathcal{A}$ such that $\log(\boldsymbol{k}) = \mathcal{A}\boldsymbol{u}$ is grid-based, thus $\log(k^l) = \log[k(X^l)]$. Then for the local update of an uncertain parameter $\log(k^l)$ we introduce a diagonal matrix $\hat{\mathbf{C}}^l \in \mathbb{R}^{\kappa\times\kappa}$ in the observation space with an element

$$(\hat{C}^l)_{\ell,\ell} = \rho\left(\frac{||X^l - r^\ell||}{r^{\text{loc}}}\right) \quad \text{for} \quad \ell = 1, \ldots \kappa. \tag{4.9}$$

Here r^ℓ denotes the location of the observation, r^{loc} is a localization radius and $\rho(\cdot)$ is a taper function, such as Gaspari-Cohn function by [33]

$$\rho(r) = \begin{cases} 1 - \frac{5}{3}r^2 + \frac{5}{8}r^3 + \frac{1}{2}r^4 - \frac{1}{4}r^5, & 0 \le r \le 1, \\ -\frac{2}{3}r^{-1} + 4 - 5r + \frac{5}{3}r^2 + \frac{5}{8}r^3 - \frac{1}{2}r^4 + \frac{1}{12}r^5, & 1 \le r \le 2, \\ 0, & 2 \le r. \end{cases} \tag{4.10}$$

The localization radius r^{loc} is typically tuned by a trial-and-error approach in terms of estimation error.

LETPF modifies the likelihood Eq. (4.4) as following

$$h^l(\boldsymbol{v}) = \exp\left[-\frac{1}{2}\left(g(\boldsymbol{u},\boldsymbol{q}) - \boldsymbol{y}_{\text{obs}}\right)'\left(\hat{\mathbf{C}}^l\mathbf{R}^{-1}\right)\left(g(\boldsymbol{u},\boldsymbol{q}) - \boldsymbol{y}_{\text{obs}}\right)\right], \tag{4.11}$$

where $\hat{\mathbf{C}}^l$ is the diagonal matrix given by Eq. (4.9). Then the optimal transport $\boldsymbol{S}^l$ is an $M \times M$ matrix with entries s_{ij}^l that satisfy

$$s_{ij}^l \ge 0, \qquad \sum_{i=1}^{M} s_{ij}^l = \frac{1}{M}, \qquad \sum_{j=1}^{M} s_{ij}^l = \frac{h^l\left(\boldsymbol{v}_i^{(t)}\right)^{(\phi_t - \phi_{t-1})}}{\sum_{j=1}^{M} h^l\left(\boldsymbol{v}_j^{(t)}\right)^{(\phi_t - \phi_{t-1})}}, \tag{4.12}$$

and minimizes the cost function

$$\sum_{i,j=1}^{M} s_{ij}^{l} \left[\log\left(k_i^{l,(t)}\right) - \log\left(k_j^{l,(t)}\right) \right]^2. \tag{4.13}$$

The estimated parameter $\log(\tilde{k}^l)$ is given by

$$\log(\tilde{k}_j^l) = M \sum_{i=1}^{M} \tilde{s}_{ij}^{l} \log\left(k_i^{l,(t)}\right) \quad \text{for} \quad j = 1, \ldots, M, \tag{4.14}$$

where $\tilde{s}_{ij}^{l}$ is a solution to the optimization problem Eq. (4.12)–(4.13). We note that localization reduces LETPF to a univariate transport problem. Solving the optimal transport problem for univariate variables Eq. (4.12)–(4.13) is computationally less expensive than for multivariate variables Eq. (4.6)–(4.7), since the marginal computational cost is in sorting M numbers. We use an algorithm described in [74] to solve the univariate linear transport problem Eq. (4.12)–(4.13). LETPF, however, loses direct dependence on the parameters at other grid cells. Update of the uncertain grid-based parameters $\log(k^l)$ could be performed in parallel for each $l = 1, \ldots, N^2$. Then the estimated model parameter is $\tilde{\boldsymbol{u}} = \mathcal{A}^{-1} \log(\tilde{\boldsymbol{k}})$.

To estimate $\boldsymbol{q}$, we solve the optimal transport problem with s_{ij}^{G} that satisfy

$$s_{ij}^{\mathrm{G}} \geq 0, \qquad \sum_{i=1}^{M} s_{ij}^{\mathrm{G}} = \frac{1}{M}, \qquad \sum_{j=1}^{M} s_{ij}^{\mathrm{G}} = \frac{h\left(\left[\tilde{\boldsymbol{u}}_i \; \boldsymbol{q}_i^{(t)}\right]'\right)^{(\phi_t - \phi_{t-1})}}{\sum_{j=1}^{M} h\left(\left[\tilde{\boldsymbol{u}}_j \; \boldsymbol{q}_j^{(t)}\right]'\right)^{(\phi_t - \phi_{t-1})}}, \tag{4.15}$$

and minimize the cost function

$$\sum_{i,j=1}^{M} s_{ij}^{\mathrm{G}} \left(\|\tilde{\boldsymbol{u}}_i - \tilde{\boldsymbol{u}}_j\|^2 + \left\| \boldsymbol{q}_i^{(t)} - \boldsymbol{q}_j^{(t)} \right\|^2 \right). \tag{4.16}$$

The estimated parameter $\tilde{\boldsymbol{q}}$ is given by

$$\tilde{\boldsymbol{q}}_j = M \sum_{i=1}^{M} \tilde{s}_{ij}^{\mathrm{G}} \boldsymbol{q}_i^{(t)}, \quad j = 1, \ldots, M, \tag{4.17}$$

where $\tilde{s}_{ij}^{\mathrm{G}}$ is a solution to the optimization problem Eq. (4.15)– (4.16). Finally, we set

$\tilde{\boldsymbol{v}} = [\tilde{\boldsymbol{u}}\ \tilde{\boldsymbol{q}}]'$.

4.2.4 Mutation

The advantage of deterministic resampling is that it reduces sampling noise. The disadvantage of deterministic resampling is that for a deterministic and non-chaotic system the filter collapse is unavoidable unless particle mutation is introduced. The mutation is performed over an index $1 < \tau < \tau_{\max}$ with prescribed $\tau_{\max}$. At the first inner iteration $\tau = 1$ we assign $\boldsymbol{v} = \tilde{\boldsymbol{v}}$.

We denote by v_i^ℓ a component of $\boldsymbol{v}_i$, where $1 \leq \ell \leq n+m$. If v_i^ℓ has a Gaussian prior, then we use the preconditioned Crank-Nicolson pcn-MCMC method from [17]

$$v_i^{\ell,\text{prop}} = \sqrt{1-\beta^2} v_i^\ell + \beta \xi_i \quad \text{for} \quad i = 1, \ldots, M, \tag{4.18}$$

where $\{\xi_i\}_{i=1}^M$ is from normal distribution. The parameter $\beta \in [0,\ 1)$ is a free parameter. For uniform prior $U[a,\ b]$, we use random walk

$$v_i^{\ell,\text{prop}} = v_i^\ell + \xi_i \quad \text{for} \quad i = 1, \ldots, M, \tag{4.19}$$

where $\xi_i \sim U[a-b,\ b-a]$, and we project $v_i^{\ell,\text{prop}}$ to $[a,\ b]$ when necessary. The proposal Eq. (4.18)– (4.19) is accepted

$$\boldsymbol{v} = \boldsymbol{v}^{\text{prop}} \quad \text{with the probability} \quad \min\left\{1, \frac{h(\boldsymbol{v}^{\text{prop}})^{\phi_{t+1}}}{h(\boldsymbol{v})^{\phi_{t+1}}}\right\}, \tag{4.20}$$

and the inner iteration τ is increased by one. We choose β based on acceptance rate being between 20% and 30% by the last iteration T. The mutation Eq. (4.18)–(4.20) is repeated until $\tau = \tau_{\max}$, then we assign $\boldsymbol{v}^{(t+1)} = \boldsymbol{v}$.

After that, the next tempering iteration proceeds by computing the weights Eq. (4.4), computing the new temperature ϕ based on Eq. (4.5) ESS $\geq M_{\text{thresh}}$, performing deterministic resampling either by Eq. (4.6)–(4.8) for the non-localized method or by Eq. (4.12)–(4.17) for the localized method, and concluding by mutation Eq. (4.18)–(4.20) for $\tau_{\max}$ iterations. The algorithms stops when the final temperature ϕ reaches one. Recall that T is the total number of tempering iterations that corresponds to ϕ reaching one. Thus T is not predefined, which can lead to computationally unfeasible iteration times. TETPF demands $TM(\tau_{\max}+1)$ evaluations of the model g, and TLETPF

demands $TM(\tau_{\max}+2)$ evaluations of the model g.

4.3 Regularized Ensemble Kalman Filter

REnKF is based on the Ensemble Kalman Filter with perturbed observations

$$\boldsymbol{y}_i^{\eta} = \boldsymbol{y}_{\text{obs}} + \boldsymbol{\eta}_i \quad \text{for} \quad i = 1, \ldots, M,$$

where $\eta_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$ with $\boldsymbol{R}$ being a known covariance matrix of the observation noise. We define an M-dimensional vector with all elements equal to 1 as $\boldsymbol{1}_M$. REnKF solves the following set of equations for $t = 0, \ldots, T-1$ with $\boldsymbol{v}^{(0)}$ being an initial ensemble

$$\begin{aligned}
\boldsymbol{B}^{\text{gg}} &= \frac{1}{M-1}\left(g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)}) - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\boldsymbol{1}_M'\right)\left(g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)}) - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\boldsymbol{1}_M'\right)', \\
\boldsymbol{B}^{\text{vg}} &= \frac{1}{M-1}\left(\boldsymbol{v}^{(t)} - \overline{\boldsymbol{v}^{(t)}}\boldsymbol{1}_M'\right)\left(g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)}) - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\boldsymbol{1}_M'\right)', \\
\boldsymbol{v}_i^{(t+1)} &= \boldsymbol{v}_i^{(t)} + \boldsymbol{B}^{\text{vg}}\left(\boldsymbol{B}^{\text{gg}} + \mu^{(t)}\boldsymbol{R}\right)^{-1}\left(\boldsymbol{y}_i^{\eta} - g(\boldsymbol{u}_i^{(t)}, \boldsymbol{q}_i^{(t)})\right) \quad \text{for} \quad i = 1, \ldots, M. \quad (4.21)
\end{aligned}$$

The regularized parameter $\mu^{(t)}$ is chosen such that

$$\mu^{(t)}\left\| \boldsymbol{R}^{1/2}\left(\boldsymbol{B}^{\text{gg}} + \mu^{(t)}\boldsymbol{R}\right)^{-1}\left(\boldsymbol{y}_{\text{obs}} - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\right)\right\| \geq \Omega\left\|\boldsymbol{R}^{-1/2}\left(\boldsymbol{y}_{\text{obs}} - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\right)\right\| \quad (4.22)$$

for predefined $\Omega \in (0,1)$. This is achieved by the bisection method $\mu^{\tau+1} = 2^{\tau}\mu^0$ for $\tau = 0, \ldots, \tau_{\max}$ and an initial guess μ^0. We assign $\mu^{(t)} = \mu^{\tau_{\max}}$, where $\tau_{\max}$ is the first integer for which Eq. (4.22) holds.

Finally, REnKF is stopped based on discrepancy principle, namely when

$$\left\|\boldsymbol{R}^{-1/2}\left(\boldsymbol{y}_{\text{obs}} - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\right)\right\| \leq 1/\Omega\left\|\boldsymbol{R}^{-1/2}\eta\right\| \quad (4.23)$$

with η being the observation noise. Thus the total number of iterations T is not predefined, as in T(L)ETPF. A rule of thumb is to choose $\Omega \in (0.5, 1)$, and we choose $\Omega = 0.7$ for all the numerical experiments. REnKF demands $TM + 1$ evaluations of the model g.

4.3.1 Localization

In the Ensemble Kalman filter, covariance-based localization of the Kalman gain [38, 41] can be applied in order to remove spurious correlations due to a small ensemble size. We assume again having a numerical grid of $N \times N$ size with grid cells denoted by X^l for $l = 1, \ldots, N^2$. Assume that the uncertain parameter $\boldsymbol{u}$ is not grid-based. We assume, however, that there exists a matrix $\mathcal{A}$ such that $\log(\boldsymbol{k}) = \mathcal{A}\boldsymbol{u}$ is grid-based, thus $\log(k^l) = \log[k(X^l)]$. Then Eq. (4.21) for a localized EnKF, denoted here LEnKF is rewritten as

$$\log\left(\boldsymbol{k}_i^{(t+1)}\right) = \log\left(\boldsymbol{k}_i^{(t)}\right) + \hat{\mathcal{C}} \circ \boldsymbol{B}^{\mathrm{kg}} \left(\boldsymbol{B}^{\mathrm{gg}} + \mu^{(t)}\boldsymbol{R}\right)^{-1} \left(\boldsymbol{y}_i^{\eta} - g(\boldsymbol{u}_i^{(t)}, \boldsymbol{q}_i^{(t)})\right),$$
$$\boldsymbol{q}_i^{(t+1)} = \boldsymbol{q}_i^{(t)} + \boldsymbol{B}^{\mathrm{qg}} \left(\boldsymbol{B}^{\mathrm{gg}} + \mu^{(t)}\boldsymbol{R}\right)^{-1} \left(\boldsymbol{y}_i^{\eta} - g(\boldsymbol{u}_i^{(t)}, \boldsymbol{q}_i^{(t)})\right) \quad \text{for} \quad i = 1, \ldots, M.$$

Here $\circ$ denotes the element-wise product and $\hat{\mathcal{C}}$ is a distance-based correlation matrix, an element of which is

$$\hat{\mathcal{C}}_{l,\ell} = \rho\left(\frac{||X^l - r^{\ell}||}{r^{\mathrm{loc}}}\right) \quad \text{for} \quad l = 1, \ldots, N^2 \quad \text{and} \quad \ell = 1, \ldots, \kappa, \tag{4.24}$$

where r^{ℓ} denotes the location of the observation, r^{loc} is a localization radius and ρ is given by Eq. (4.10). The covariance matrices $\boldsymbol{B}^{\mathrm{gg}}$, $\boldsymbol{B}^{\mathrm{kg}}$ and $\boldsymbol{B}^{qg}$ are

$$\boldsymbol{B}^{\mathrm{gg}} = \frac{1}{M-1}\left(g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)}) - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\mathbf{1}'_M\right)\left(g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)}) - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\mathbf{1}'_M\right)',$$
$$\boldsymbol{B}^{\mathrm{kg}} = \frac{1}{M-1}\left(\log(\boldsymbol{k}^{(t)}) - \overline{\log(\boldsymbol{k}^{(t)})}\mathbf{1}'_M\right)\left(g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)}) - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\mathbf{1}'_M\right)',$$
$$\boldsymbol{B}^{\mathrm{qg}} = \frac{1}{M-1}\left(\boldsymbol{q}^{(t)} - \overline{\boldsymbol{q}^{(t)}}\mathbf{1}'_M\right)\left(g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)}) - \overline{g(\boldsymbol{u}^{(t)}, \boldsymbol{q}^{(t)})}\mathbf{1}'_M\right)'.$$

RLEnKF also demands $TM + 1$ evaluations of the model g, as REnKF.

4.4 Numerical experiment

We consider a test case of estimating uncertain Gaussian permeability. We consider a steady-state single-phase Darcy flow model defined over an aquifer of a two-dimensional

physical domain $D = [0, 6] \times [0, 6]$, which is given by

$$-\nabla \cdot [k(x, y)\nabla P(x, y)] = \mathcal{F}(x, y) \quad \text{for} \quad (x, y) \in D,$$

where $\nabla = (\partial/\partial x \ \partial/\partial y)'$, $\cdot$ is the dot product, $P(x, y)$ is the pressure, $k(x, y)$ is the permeability, and the source term is

$$\mathcal{F}(x, y) = \begin{cases} 0 & \text{for} \quad 0 \leq y \leq 4, \\ 137 & \text{for} \quad 4 < y \leq 5, \\ 274 & \text{for} \quad 5 < y \leq 6. \end{cases}$$

The boundary conditions are a combination of Dirichlet and Neumann boundary conditions

$$P(x, 0) = 100, \quad \frac{\partial P}{\partial x}(6, y) = 0, \quad -k(0, y)\frac{\partial P}{\partial x}(0, y) = 500(1 + q), \quad \frac{\partial P}{\partial y}(x, 6) = 0,$$

where q denotes error in boundary conditions.

We implement a cell-centred finite difference method to discretize the domain D into $N \times N$ grid cells X^l of size Δx^2. We solve the true forward model G on a fine grid $N = N_{\mathrm{f}} = 140$ for the true solution. Then the synthetic observations are obtained by

$$\boldsymbol{y}_{\mathrm{obs}} = \boldsymbol{L}(\boldsymbol{P}^{\mathrm{true}}) + \boldsymbol{\eta}.$$

An element of $\boldsymbol{L}(\boldsymbol{P}^{\mathrm{true}})$ is a linear functional of pressure, namely

$$L^{\ell}(\boldsymbol{P}^{\mathrm{true}}) = \frac{1}{2\pi\sigma^2} \sum_{l=1}^{N_{\mathrm{f}}^2} \exp\left(-\frac{||X^l - r^{\ell}||^2}{2\sigma^2}\right) P^{\mathrm{true},l} \Delta x_{\mathrm{f}}^2 \quad \text{for} \quad \ell = 1, \ldots, \kappa,$$

where r^{ℓ} denotes the location of the observation, $\kappa = 36$ the number of observations, $\Delta x_{\mathrm{f}} = 6/N_{\mathrm{f}}$, and $\sigma = 0.01$. The observation locations are displayed in Fig. 4.4 as circles.

Observation noise is denoted by $\boldsymbol{\eta}$ and it is drawn from $\mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$. Observation error covariance $\boldsymbol{R}$ is known, and we choose it such that the norm of the noise is 1% of the norm of the data. Such a small noise makes the data assimilation problem hard to solve, since the likelihood is very peaked and a non-iterative data assimilation approach fails.

Both the true permeability and an initial ensemble are drawn from the same prior

distribution as the prior includes knowledge about geological properties. We assume log permeability is generated by a random draw from a Gaussian distribution $\mathcal{N}(\log(\mathbf{5}), \boldsymbol{C})$. Here $\mathbf{5}$ is an N^2 vector with all elements being 5 and $\boldsymbol{C}$ is the Whittle-Matern correlation, an element of which is given by

$$C^{l\ell} = \frac{1}{\Gamma(1)} \frac{d^{l\ell}}{\delta} \Upsilon_1 \left(\frac{d^{l\ell}}{\delta} \right) \quad \text{for} \quad l, \ell = 1, \ldots, N^2.$$

Here $d^{l\ell}$ is the distance between two spatial locations, $\delta = 0.5$ is the correlation length, Γ is the gamma function, and Υ_1 is the modified Bessel function of the second kind of order one. We denote by λ and γ eigenvalues and eigenfunctions of $\boldsymbol{C}$, respectively. Then, following a Karhunen-Loeve expansion, log permeability is

$$\log(k^l) = \log(5) + \sum_{\ell=1}^{N^2} \sqrt{\lambda^\ell} \gamma^{\ell l} u^\ell \quad \text{for} \quad l = 1, \ldots, N^2,$$

where u^ℓ is i.i.d. from a normal distribution for $\ell = 1, \ldots, N^2$.

Therefore the initial parameter u is drawn from $\mathcal{N}(0, 1)$, while the initial boundary condition error q is drawn from a uniform distribution $U[0\ 0.5]$. We then solve the incorrect forward model g on a coarse grid $N = N_\mathrm{c} = 70$. The uncertain parameter $\boldsymbol{u}$ has the dimension $n = 4900$, which makes the dimension of $\boldsymbol{v}$ $n + m = 4901$. We perform 40 numerical experiments with both T(L)ETPF and R(L)EnKF to check initial sample sensitivity. We conduct numerical experiments with ensemble sizes 100 and 1000. We compare the methods to a pcn-MCMC method without model error in boundary conditions. An MCMC experiment was conducted using 100 chains with the lengths 10^6, burn-in period 10^5, and thinning period 10^3 each. For T(L)ETPF we choose $\tau_{\max} = 20$ and $\beta = 0.045$ for mutation, since this choice gives good mixing with acceptance rate at the final tempering iteration around 20%. We set the threshold for ESS to be $M_{\text{thresh}} = M/3$.

We define the root mean square error (RMSE) of a mean field $\overline{\Xi} = 1/M \sum_{i=1}^{M} \Xi_i$ as

$$\mathrm{RMSE}(\Xi) = \sqrt{\left(\overline{\Xi} - \Xi^{\mathrm{MCMC}}\right)^T \left(\overline{\Xi} - \Xi^{\mathrm{MCMC}}\right)}$$

for either log permeability $\Xi = \log(\boldsymbol{k})$ or pressure $\Xi = \boldsymbol{P}$. To choose a favouring localization radius, we perform a numerical experiment with r^{loc} ranging from one to

six with an increment of one. Then we define the favouring localization radius as a localization radius that gives the smallest RMSE in terms of mean log permeability for that numerical experiment. For TLETPF the favouring localization radius is $r^{\text{loc}} = 1$ for both ensemble sizes 100 and 1000. For RLEnKF the favouring localization radius is $r^{\text{loc}} = 3$ for both ensemble sizes 100 and 1000.

4.4.1 Data assimilation without localization

In this section we compare REnKF to TETPF, thus without localization. In Fig. 4.1, we plot RMSE of mean log permeability on the left and of mean pressure on the right at different ensemble sizes for both TETPF and REnKF. When the uncertainty is only in permeability, REnKF outperforms TETPF for estimation of $\log(k)$, and as a consequence, for estimation of pressure, as can be seen in Fig. 4.1(a–b) and as has been reported in the literature, e.g. [77, 78]. When uncertainty is in both permeability and boundary conditions, we investigate the method's performance for two numerical set-ups. The first numerical set-up is when we account for model error in boundary conditions. The second numerical set-up is when we do not account for model error in boundary conditions. In both set-ups REnKF outperforms TETPF as seen in Fig. 4.1(c–f).

Let us now examine these numerical experiments in more detail. First, we compare numerical experiments when we account for error in boundary conditions to numerical experiments with no error in boundary conditions. Comparing Fig. 4.1(d) to Fig. 4.1(b), we observe that the pressure estimation does not change. At the same time comparing Fig. 4.1(c) to Fig. 4.1(a), we observe that the log permeability estimation becomes worse when model error is present, as to be expected.

Next, we compare numerical experiments when we do not account for error in boundary conditions to numerical experiments when we do account for error in boundary conditions. Comparing Fig. 4.1(e) to Fig. 4.1(c), we observe that for REnKF at ensemble size 1000 and TETPF at both ensemble sizes log permeability is better estimated when a method does not account for error in boundary conditions. Pressure is also better estimated by both methods when error in boundary conditions is not accounted for, as seen from comparing Fig. 4.1(f) to Fig. 4.1(d).

In order to further investigate this result, we compare REnKF and TETPF to an estimation from a pcn-MCMC with error in boundary conditions. In Fig. 4.2 we plot the posterior approximations of q. We recall that the prior for q is $U[0\ 0.5]$. We observe

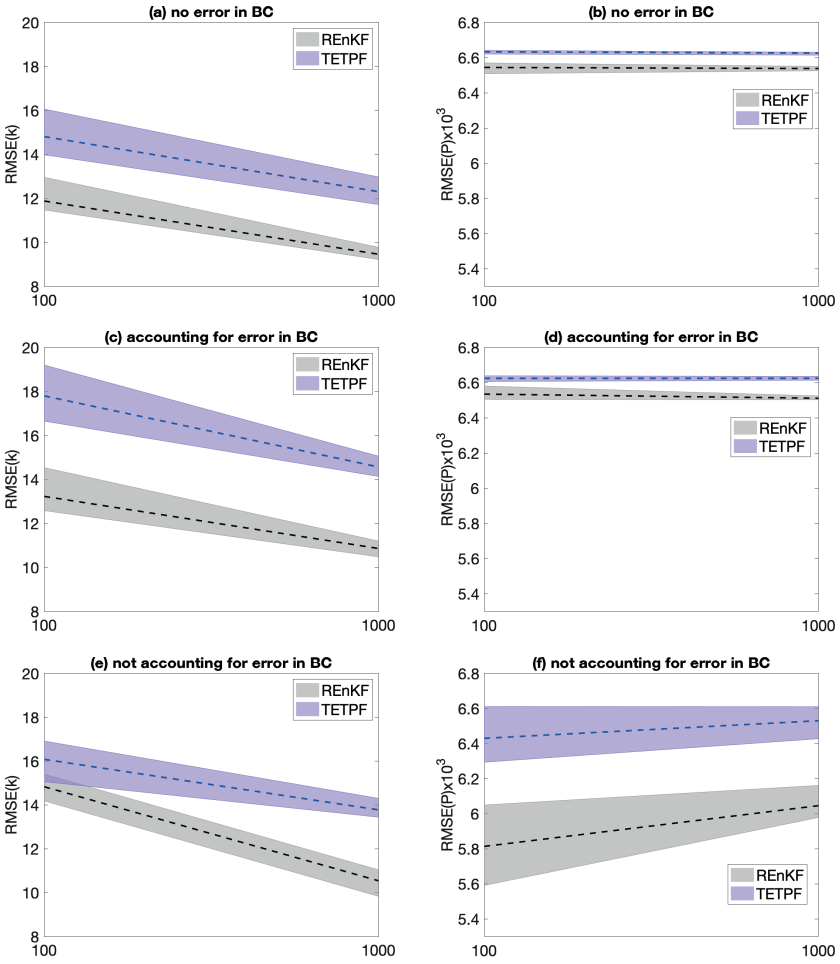

Figure 4.1: RMSE for $\log(\boldsymbol{k})$ is on the left. RMSE $\times 10^3$ for pressure is on the right. A dashed line indicates the median, a shaded area indicates the $25-75$ percentile range over 40 simulations. Ensemble size is on x-axis. REnKF is shown in grey. TETPF is shown in blue. (a)–(b) no model error, (c)–(d) accounting for model error, and (e)–(f) not accounting for model error.

that MCMC gives a skewed posterior. TETPF has better resemblance to the posterior obtained by MCMC, while REnKF gives negative values of q (negative values of q for TETPF and MCMC are only due to kernel representation for plotting). We however observed that pressure is better estimated by both methods when error in boundary conditions is not accounted for. Thus by allowing for error in boundary conditions and

not constraining it by data, we obtain a better pressure estimation.

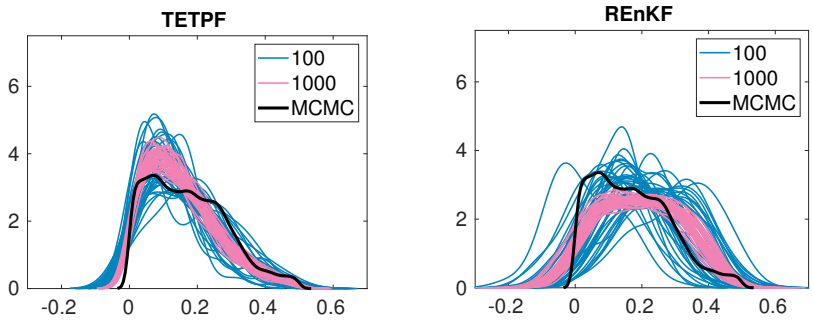

Figure 4.2: Posterior of error in boundary conditions q. On the left is TETPF and on the right is REnKF. Results for the ensemble size 100 are shown in blue and for 1000 in pink, where one line is for one simulation out of 40. MCMC is shown in black.

4.4.2 Data assimilation with localization

Next we compare RLEnKF to TLETPF, thus with localization. In Fig. 4.3, we plot RMSE of mean log permeability on the left and of mean pressure on the right at different ensemble sizes for both TLETPF and RLEnKF. When uncertainty is only in permeability, TLETPF outperforms RLEnKF for estimation of $\log(k)$ and as a consequence for estimation of pressure, as can be seen in Fig. 4.3(a–b). This shows that localization drastically improves TETPF, such that it outperforms RLEnKF even for normally distributed $\log(k)$.

Next, we compare numerical experiments when we account for error in boundary conditions to numerical experiments with no error in boundary conditions. Comparing Fig. 4.3(d) to Fig. 4.3(b), we observe that pressure estimation does not change. At the same time comparing Fig. 4.3(c) to Fig. 4.3(a), we observe that the log permeability estimation becomes worse when model error is present, as is to be expected. Here TLETPF again outperforms RLEnKF but only at a large ensemble size 1000. For comparison, we show results of MCMC, when error in boundary conditions is accounted for. In Fig. 4.4 we plot mean of log permeability at the top and variance of log permeability at the bottom for MCMC, TLETPF, and RLEnKF at ensemble size 1000. We observe that both methods give a reasonably good approximation of the MCMC mean log permeability.

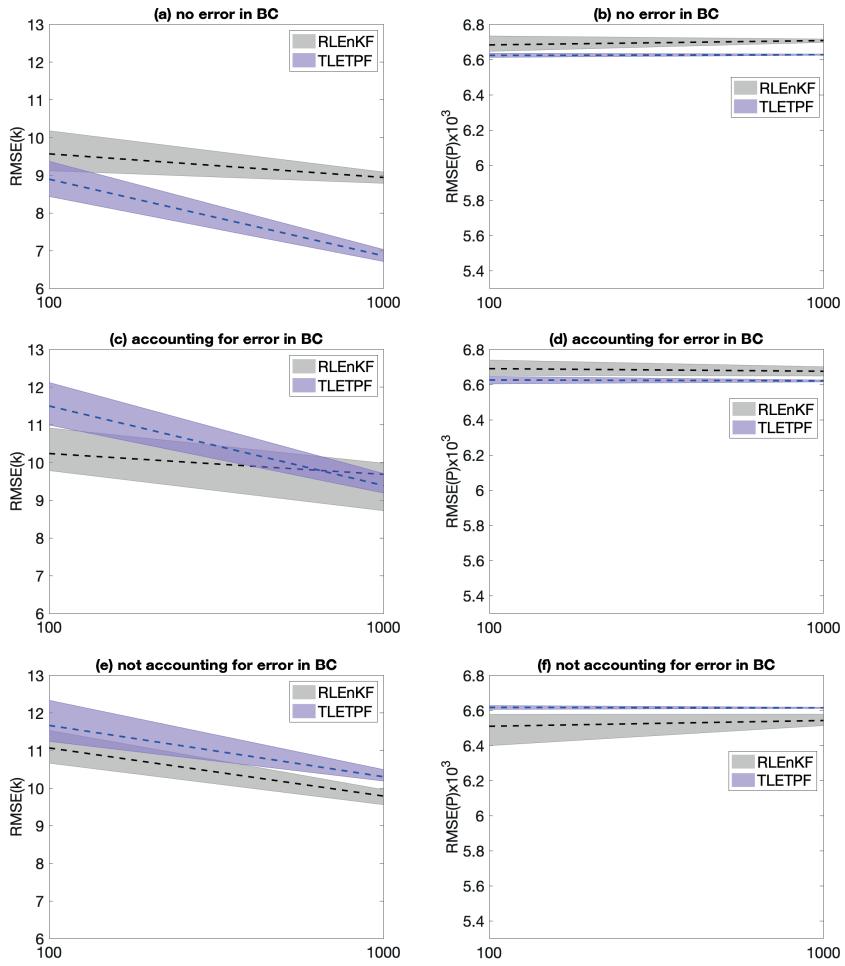

Figure 4.3: RMSE for $\log(\boldsymbol{k})$ is on the left. RMSE $\times 10^3$ for pressure is on the right. A dashed line indicates the median, a shaded area indicates the $25-75$ percentile range over 40 simulations. Ensemble size is on x-axis. RLEnKF is shown in grey. TLETPF is shown in blue. (a)–(b) no model error, (c)–(d) accounting for model error, and (e)–(f) not accounting for model error.

The variance, however, is underestimated.

Next, we compare numerical experiments when we do not account for error in boundary conditions to numerical experiments when we do account for error in boundary conditions. Comparing Fig. 4.3(e) to Fig. 4.3(c), we observe that for RLEnKF and TLETPF log permeability is better estimated when a method accounts for error in boundary con-

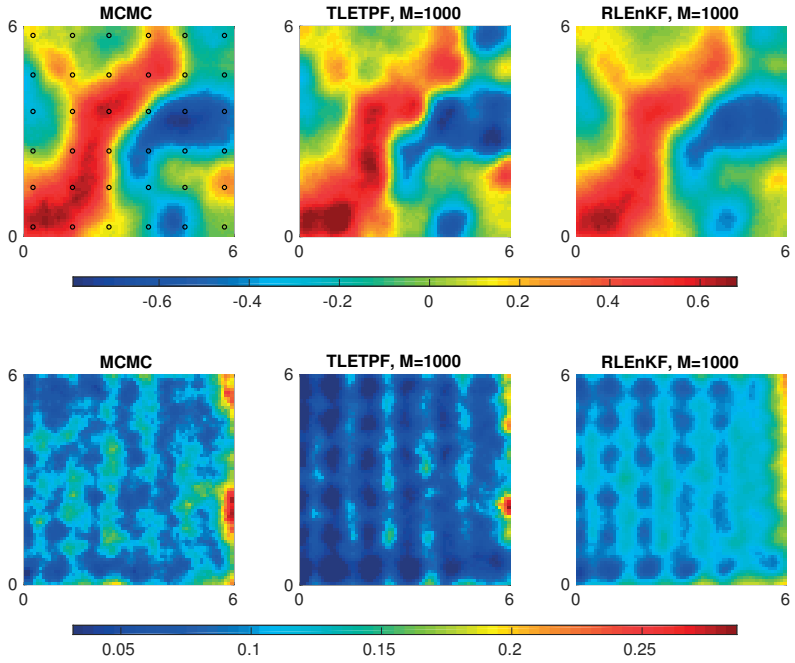

Figure 4.4: Mean $\log(\boldsymbol{k})$ at the top and variance of $\log(\boldsymbol{k})$ at the bottom for a numerical experiment in which error in boundary conditions is accounted for. MCMC is on the left, with circles for the observation locations. A simulation with ensemble size 1000 is in the middle for TLETPF and on the right for RLEnKF.

ditions. This is opposite to results of REnKF and TETPF. When it comes to pressure estimation, TLETPF gives equivalent performance for all three test cases. This means that TLETPF is not sensitive to error in boundary conditions in terms of pressure estimation. RLEnKF, however, better estimates pressure when error in boundary conditions is not accounted for, as seen from comparing Fig. 4.3(f) to Fig. 4.3(d). The same was observed in experiments with REnKF. When comparing RLEnKF and TLETPF in terms of the posterior approximations of q shown in Fig. 4.5, we observe that none of the methods give estimations close to MCMC.

R(L)EnKF at large ensemble size

We further investigate R(L)EnKF (with and without localization) performance in the case when error in boundary conditions is not accounted for. We increase ensemble

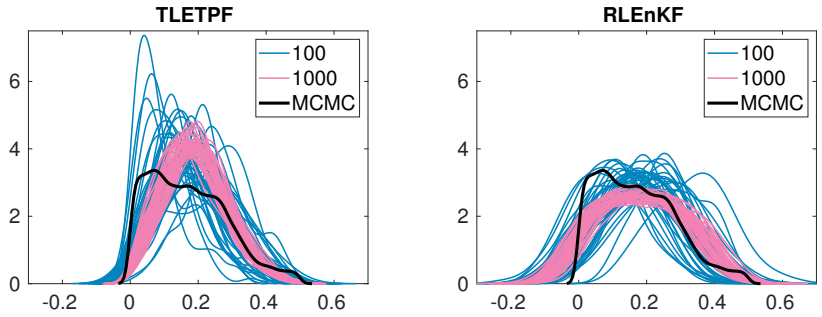

Figure 4.5: Posterior of error in boundary conditions q. On the left is TLETPF and on the right is RLEnKF. Results for ensemble size 100 are shown in blue and for 1000 in pink, where one line is for one simulation out of 40. MCMC is shown in black.

size to $M = 6000$ for the localized version and to $M = 7700$ for the non-localized version. We have tested localization radius between 1 and 8 with an increment of 1, and $r^{\text{loc}} = 6$ gave the smallest RMSE. We compute RMSE of mean log permeability and mean pressure and display them in Fig. 4.6 in grey for R(L)EnKF. We observe that as ensemble size increases pressure becomes better estimated. For 1000 and 6000 members, the improvement in P can be explained by the improvement in $\log(k)$, as seen on the left of Fig. 4.6. Comparing RLEnKF with 6000 members to REnKF with 7700 members, we observe that better pressure estimation can only be explained by better estimation of q. In Fig. 4.6, we show TLETPF with ensemble size 1000 in pink. We observe that R(L)EnKF outperforms TLETPF greatly. For TLETPF we were not able to increase the ensemble size beyond 1000 due to high computational costs.

4.4.3 Computational costs

In Tab. 4.1, we show the number of iterations that a method takes on average. We should note that for R(L)EnKF we put a limit of 20 iterations. When performing numerical experiments where no model error is present or when model error is accounted for, we observed that R(L)EnKF attains the stopping criterion (4.23) in less than 20 iterations for $\Omega = 0.7$. When performing numerical experiments when model error is not accounted for, we observed that in the majority of experiments after 20 iterations the data misfit starts to increase and the stopping criterion is not attained. Decreasing Ω to values

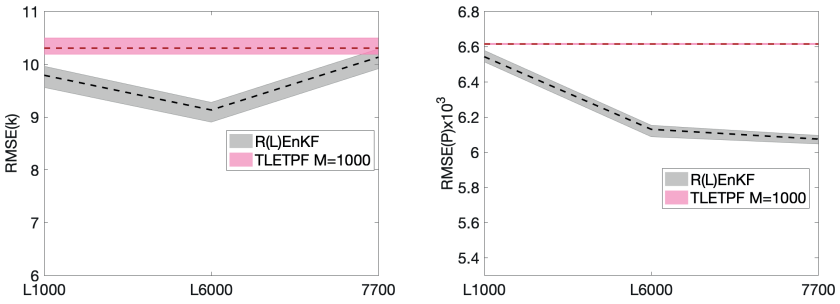

Figure 4.6: RMSE for $\log(\boldsymbol{k})$ is on the left, and RMSE $\times 10^3$ for pressure is on the right for the test case where error in boundary conditions is not accounted for. A dashed line indicates the median, a shaded area indicates the $25-75$ percentile range over 40 simulations. On the x-axis numbers stand for ensemble sizes and L stands for a localized method. R(L)EnKF is shown in grey. TLETPF for ensemble size 1000 is shown in pink.

less than 0.5 solves the issue of attaining the stopping criterion but the data misfit is too high for sensible estimations. By the trial-and-error approach, we discovered that it is preferable to keep $\Omega = 0.7$ and put a limit of 20 total iterations. For T(L)ETPF we do not put any limit on total number of iterations. The wall clock of one iteration of TETPF is 70 seconds, of TLETPF is 78 seconds, of REnKF is 9 seconds, and of RLEnKF is 10 seconds.

Table 4.1: Mean over 40 simulations of number of iterations for T(L)ETPF and R(L)EnKF at ensemble size 1000.

	TETPF	TLETPF	REnKF	RLEnKF
no ME	13	28	8	11
accounting for ME	14	35	11	15
not accounting for ME	19	30	19	19

4.5 Conclusions

We have considered a 2D steady-state Darcy flow with uncertain permeability. Observations of pressure at 36 locations were used to infer permeability on a grid of 70×70. The

corresponding inverse problem was solved using two methods: regularized (localized) ensemble Kalman filter—R(L)EnKF—and tempered (localized) ensemble transform particle filter—T(L)ETPF. The ensemble Kalman filter updates uncertain parameters based on an assumption of Gaussian probabilities. Therefore mean and variance need to be estimated. The ensemble Kalman filter makes further assumption that variance can be approximated by an ensemble. Due to small ensemble size, variables exhibit spurious correlations. Spurious correlations can be removed be employing localization—where the covariance matrix in the Kalman gain is multiplied by a distance-based matrix, for example. Moreover, it has been recognized that the ensemble Kalman filter overfits observations, and thus requires regularization. Therefore in our numerical experiments we have used the regularized (localized) ensemble Kalman filter.

The ensemble transform particle filter makes an assumption that probability can be approximated by an empirical measure. It employs optimal transport to transform a prior measure into a posterior. It correctly estimates mean but not variance. The ensemble transform particle filter is a computationally demanding algorithm, since its complexity is of order $M^2 \ln M$ where M is the ensemble size. The ensemble transform particle filter suffers from spurious correlations as does the ensemble Kalman filter, and thus requires localization. There exists only one type of localization for the ensemble transform particle filter, which requires solving a univariate optimal transport problem at every grid, although that can be done in parallel. It has been also recognized that when observations are very accurate or numerous, the ensemble transform particle filter needs tempering—in which the prior is transformed into the posterior over several iterations. Therefore in our numerical experiments we have used a tempered (localized) ensemble transform particle filter.

We have performed experiments with ensemble sizes 100 and 1000 for both R(L)EnKF and T(L)ETPF. We have shown that localization improves estimations obtained by REnKF and TETPF in all considered numerical experiments with 2D steady-state Darcy flow. First, we have considered uncertainty only in permeability. For ensemble sizes 100 and 1000, we have shown that TLETPF outperforms both REnKF and RLEnKF. Next, we have considered uncertain permeability and uncertain boundary conditions. We have shown that TLETPF outperforms both REnKF and RLEnKF but only at a large ensemble size of 1000. Finally, we have considered uncertainty in both permeability and boundary conditions, but we did not account for error in boundary conditions in the data assimilation. In this set-up RLEnKF outperforms both TETPF and TLETPF. More-

over, T(L)ETPF is computationally more demanding than R(L)EnKF and requires more iterations than R(L)EnKF.

Chapter 5

Discussion and Outlook

This thesis investigates the ensemble data assimilation approach to accurately estimate the probability of the petrophysical parameters defining the reservoir models. Among the frequently employed ensemble data assimilation methods in reservoir modeling problems, variations of ensemble Kalman filter have proved to be highly efficient. However, the ensemble Kalman filter approach corrects for the mean and the variance of the parameters and thus makes the assumption of Gaussian probability, which is not always the case in practical scenarios. Such a fundamental assumption presents a severe drawback when addressing systems with non-Gaussian posteriors, e.g. reservoir models with channel flow configurations (Section 3.4.2). In this thesis we explore another class of ensemble data assimilation methods: particle filters. Particle filters have the advantage of not making any assumption of Gaussianity regarding the posterior. They are capable of handling non-linear systems and can provide more complete information rather than the first two moments. For a long time particle filters have been overlooked as they are known to suffer from the curse of dimensionality. The curse of dimensionality makes it highly improbable that the samples will reach the high-probability areas of posterior

distribution. In such situations the samples show high variance in weights (a posterior approximation), i.e. a few samples have large weights, while all the other samples have weights close to zero. This is termed as weight degeneracy, that is measured in terms of effective ensemble size (ESS) and in scenarios of extreme weight degenracy ESS = 1. Thus the posterior is approximated by a few (or even one) particles that might be far away from the state. Therefore, particle filters appeared impractical even for systems with dimensions 10, because the ESS quickly reduces to one, unless a very large ensemble size is used [89]. However, with ongoing developments in particle filtering, they are emerging as a promising data assimilation method. Examples of such developments are: resampling (when samples with low weights are abandoned and samples with high weight are duplicated to attenuate weight degeneracy), mutation (when Markov kernels are used to modify the samples in order to maintain diversity in sample population) and localization (when grid-dependent parameters are updated by taking into account only nearby observations to reduce dimension of problem associated with each observation). This thesis in particular focuses on a deterministic resampling based on optimal transportation, which is coined in [74] as ensemble transform particle filter (ETPF). Unlike the standard resampling approaches, ETPF modifies the particles such that the covariance between them is maximized.

To investigate the behavior of ETPF for parameter estimation problems, we undertake a steady-state single-phase Darcy flow model with permeability as an uncertain parameter. Such a simple test case allows for an in-depth examination of the method. Since the parameter-to-output map from the underlying model is nonlinear, the resulting posterior is non-Gaussian even if the prior and the noise distribution are Gaussian. The results obtained in this thesis show that for a low-dimensional problem ETPF provides a good approximation of the posterior. For a high-dimensional problem, ETPF requires updating uncertain grid-dependent parameters locally due to otherwise occurring weight degeneracy. A distance-based localization is performed such that each grid-dependent parameter is updated one at a time by taking into account only nearby observations. Localization improves the mean and variance approximation, though worsens the posterior distribution approximation.

When the prior differs substantially from the posterior or numerous accurate observations are available, then the issue of weight degeneracy in particle filtering becomes more prominent. To handle such scenarios an iterative approach of particle filtering based on tempering is advised. An adaptive sequential Monte Carlo (SMC) method is

an iterative particle filtering approach which provides a particle approximation of intermediate measures between the prior and the posterior, instead of the posterior measure directly. As the iteration proceeds, the intermediate measure gets further away from the prior but closer to the posterior. The intermediate measures are chosen based on ESS being more than some predefined value, which handles the weight degeneracy issue. We demonstrated that the deterministic resampling approach based on optimal transport in adaptive SMC is promising for inverse modelling. The numerical experiments show that the proposed approach estimates well both Gaussian and non-Gaussian distributions at small ensemble size of 100. Adaptive SMC with ensemble Kalman filter does not converge to the true non-Gaussian posterior, while with monomial resampling requires much more samples than 100 to estimate well the true Gaussian posterior. However, more theoretical studies have to be performed before the optimal transport-based adaptive SMC approach is ready for realistic applications, due to time-dependency of an underlying flow model and the multiphase nature of a flow, for example.

Moreover, the ensemble transform particle filter we have used throughout this thesis is first-order accurate. Therefore, it underestimates the variance for finite ensemble sizes. In 2017, [2] introduced the second-order accurate ensemble transform particle filter which correctly estimates variance. This approach modifies the optimal transportation map by adding a correction term based on the solution of a Riccati equation. However, the second-order accurate ETPF loses the desirable property of non-negative transformations. The numerical experiments, which are not presented in this thesis, demonstrated that the additional computational burden of solving the Riccati equation is not justifiable when compared with the gain in variance.

Last but not least, the ensemble transform particle filter is computationally demanding. Sinkhorn approximation of the ETPF [19] alleviates the computational expense but decreases the ensemble spread. Therefore, employing Sinkhorn approximation to the second-order accurate ensemble transform particle filter both decreases computational costs and improves variance estimation [2]. Though, it is important to note that Sinkhorn approximation modifies the original cost function by introducing a regularization term and each different regularization parameter results in a different minimizer of the cost function [2]. Thus, the resulting transportation map is not optimal. The performance of this approximation requires tuning of the regularization parameter, which impacts the quality of the transport map and hence the updated samples. Therefore, the second-order accurate ETPF has a high potential for inverse problems in reservoir

modeling but requires further investigation for its application for inverse problems in reservoir modelling.

Bibliography

[1] Aanonsen, S. I., Nævdal, G., Oliver, D. S., Reynolds, A. C., Vallès, B., et al. (2009). The ensemble Kalman filter in reservoir engineering–a review. *SPE Journal*, 14(03):393–412.

[2] Acevedo, W., de Wiljes, J., and Reich, S. (2017). Second-order accurate ensemble transform particle filters. *SIAM Journal on Scientific Computing*, 39(5):A1834–A1850.

[3] Backus, G. and Gilbert, F. (1968). The resolving power of gross earth data. *Geophysical Journal International*, 16(2):169–205.

[4] Backus, G. and Gilbert, F. (1970). Uniqueness in the inversion of inaccurate gross earth data. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 266(1173):123–192.

[5] Bear, J. (1972). *Dynamics of Fluids in Porous Media*. Dover Pulications, New York.

[6] Beskos, A., Jasra, A., Law, K., Marzouk, Y., and Zhou, Y. (2018). Multilevel sequential Monte Carlo with dimension-independent likelihood-informed proposals. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):762–786.

[7] Beskos, A., Jasra, A., Muzaffer, E., and Stuart, A. (2015). Sequential Monte Carlo methods for Bayesian elliptic inverse problems. *Statistics and Computing*, 25(4):727–737.

[8] Bishop, C. H., Etherton, B. J., and Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform Kalman filter. part I: Theoretical aspects. *Monthly Weather Review*, 129(3):420–436.

[9] Bui-Thanh, T. and Girolami, M. (2014). Solving large-scale PDE-constrained Bayesian inverse problems with Riemann manifold Hamiltonian Monte Carlo. *Inverse Problems*, 30(11):114014.

[10] Carrera, J. and Neuman, S. P. (1986). Estimation of aquifer parameters under transient and steady state conditions: III. Application to synthetic and field data. *Water Resour. Res.*, 22.

[11] Chada, N. K., Iglesias, M. A., Roininen, L., and Stuart, A. M. (2018). Parameterizations for ensemble Kalman inversion. *Inverse Problems*, 34(5):055009.

[12] Chen, Y. and Oliver, D. S. (2013). Levenberg–Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computational Geosciences*, 17(4):689–703.

[13] Chen, Y. and Zhang, D. (2006). Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Advances in Water Resources*, 29(8):1107–1122.

[14] Cheng, Y. and Reich, S. (2015). Data assimilation: a dynamical system perspective. *Frontiers in Applied Dynamical Systems: Reviews and Tutorials*, 2:75–118.

[15] Christie, M., Demyanov, V., and Erbas, D. (2006). Uncertainty quantification for porous media flows. *Journal of Computational Physics*, 217(1):143–158.

[16] Chustagulprom, N., Reich, S., and Reinhardt, M. (2016). A hybrid ensemble transform particle filter for nonlinear and spatially extended dynamical systems. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):592–608.

[17] Cotter, S. L., Roberts, G. O., Stuart, A. M., and White, D. (2013). MCMC methods for functions: modifying old algorithms to make them faster. *Statistical Science*, 28:424–446.

[18] Crisan, D. and Doucet, A. (2000). Convergence of sequential Monte Carlo methods. *Signal Processing Group, Department of Engineering, University of Cambridge, Technical Report CUEDIF-INFENGrrR38*, 1.

[19] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300. Lake Tahoe, Nevada.

[20] Daniela, C., Dunlop, M., Somersalo, E., and Stuart, A. (2018). Iterative updating of model error for Bayesian inversion. *Inverse Problems*, 34(2):025008.

[21] Dashti, M. and Stuart, A. M. (2016). The Bayesian approach to inverse problems. *Handbook of Uncertainty Quantification*, pages 1–118.

[22] de Wiljes, J., Reich, S., and Stannat, W. (2018). Long-time stability and accuracy of the ensemble Kalman-Bucy filter for fully observed processes and small measurement noise. *SIAM Journal on Applied Dynamical Systems*, 17:1152–1181.

[23] Del Moral, P., Doucet, A., and Jasra, A. (2006). Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436.

[24] Doucet, A., de Freitas, N., and Gordon, N. (2001). *Sequential Monte-Carlo Methods in Practice*. Springer-Verlag, New York.

[25] Dovera, L. and Della R., E. (2011). Multimodal ensemble Kalman filtering using Gaussian mixture models. *Computational Geosciences*, 15(2):307–323.

[26] Dubinkina, S. (2018). Relevance of conservative numerical schemes for an ensemble Kalman filter. *Quarterly Journal of the Royal Meteorological Society*, 144(711):468–477.

[27] Emerick, A. A. and Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55:3–15.

[28] Evans, L. C. (2010). *Partial differential equations.* American Mathematical Society, Providence, R.I.

[29] Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162.

[30] Evensen, G. (2009). *Data assimilation: the ensemble Kalman filter*. Springer Science & Business Media.

[31] Evensen, G. (2018). Analysis of iterative ensemble smoothers for solving inverse problems. *Computational Geosciences*, 22(3):885–908.

[32] Ewing, R. E., Pilant, M. S., Wade, J. G., and Watson, A. T. (1994). Estimating parameters in scientific computation: A survey of experience from oil and groundwater modeling. *IEEE Computational Science & Engineering*, 1(3):19–31.

[33] Gaspari, G. and Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554):723–757.

[34] Geir, N., Johnsen, L. M., Aanonsen, S. I., Vefring, E. H., et al. (2003). Reservoir monitoring and continuous model updating using ensemble Kalman filter. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

[35] Ghil, M. and Malanotte-Rizzoli, P. (1991). Data assimilation in meteorology and oceanography. *Advances in geophysics*, 33:141–266.

[36] Guingla, P., Antonio, D., De Keyser, R., De Lannoy, G., Giustarini, L., Matgen, P., and Pauwels, V. (2012). The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter. *Hydrology and Earth System Sciences*, 16(2):375–390.

[37] Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, pages 49–52.

[38] Hamill, T. M. and Whitaker, J. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, 129:2776–2790.

[39] Hanke, M. (1997). A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, 13:79–95.

[40] Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H., and Mitchell, H. L. (1996). A system simulation approach to ensemble prediction. *Monthly Weather Review*, 124:1225–1242.

[41] Houtekamer, P. L. and Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129(1):123–137.

[42] Hunt, B., Kostelich, E., and I., S. (2007). Efficient data assimilation for spatialtemporal chaos: A local ensemble transform Kalman filter. *Physica D*, 230:112–137.

[43] Iglesias, M. (2015). Iterative regularization for ensemble data assimilation in reservoir models. *Computational Geosciences*, 19(1):177–212.

[44] Iglesias, M., Park, M., and Tretyakov, M. (2018). Bayesian inversion in resin transfer molding. *Inverse Problems*, 34(10):105002.

[45] Iglesias, M. A. (2016). A regularizing iterative ensemble Kalman method for PDE-constrained inverse problems. *Inverse Problems*, 32:025002.

[46] Iglesias, M. A., Lin, K., and Stuart, A. M. (2014). Well-posed Bayesian geometric inverse problems arising in subsurface flow. *Inverse Problems*, 30(11):114001.

[47] Jasra, A., Stephens, D., Doucet, A., and Tsagaris, T. (2011). Inference for Levy-driven stochastic volatility models via adaptive sequential Monte Carlo. *Scandinavian Journal of Statistics*, 38(1):1–22.

[48] Jazwinski, A. H. (2007). *Stochastic processes and filtering theory*. Courier Corporation.

[49] Jonas, L., Papaioannou, I., and Ullmann, E. (2018). Multilevel sequential Monte Carlo for Bayesian inverse problems. *Journal of Computational Physics*, 368:154–178.

[50] Kadu, A., van Leeuwen, T., and Mulder, W. A. (2016). Salt reconstruction in full-waveform inversion with a parametric level-set method. *IEEE Transactions on Computational Imaging*, 3(2):305–315.

[51] Kaipio, J. and Somersalo, E. (2004). *Statistical and Computational Inverse Problems*, volume 160. Springer Science & Business Media.

[52] Kantas, N., Beskos, A., and Jasra, A. (2014). Sequential Monte Carlo for inverse problems: a case study for the Navier Stokes equation. *SIAM/ASA Journal Uncertainty Quantification*, 2:464–489.

[53] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *science*, 220(4598):671–680.

[54] Lehtinen, M., Paivarinta, L., and Somersalo, E. (1989). Linear inverse problems for generalised random variables. *Inverse Problems*, 5(4):599.

[55] Lorentzen, R. J., Fjelde, K. K., Frøyen, J., Lage, A., Nævdal, G., Vefring, E. H., et al. (2001). Underbalanced drilling: Real time data interpretation and decision support. In *SPE/IADC drilling conference*. Society of Petroleum Engineers.

[56] Luo, X. (2019). Ensemble-based kernel learning for a class of data assimilation problems with imperfect forward simulators. *PLOS ONE*, 14(7):1–40.

[57] Luo, X., Stordal, A. S., Lorentzen, R. J., and Naevdal, G. (2015). Iterative ensemble smoother as an approximate solution to a regularized minimum-average-cost problem: Theory and applications. *Society of Petroleum Engineers Journal*, 20(5):962–982.

[58] Ma, X. and Bi, L. (2019). A robust adaptive iterative ensemble smoother scheme for practical history matching applications. *Computational Geosciences*, 23(3):415–442.

[59] Mandel, J., Cobb, L., and Beezley, J. D. (2011). On the convergence of the ensemble Kalman filter. *Applications of Mathematics*, 56(6):533–541.

[60] Marshall, A. W. (1956). The use of multi-stage sampling schemes in Monte Carlo computations. In *Symposium on Monte Carlo Methods*, pages 123–140.

[61] Martin, J., Wilcox, L., Burstedde, C., and Ghattas, O. (2012). A stochastic newton MCMC method for large-scale statistical inverse problems with application to seismic inversion. *SIAM Journal on Scientific Computing*, 34(3):A1460–A1487.

[62] Matérn, B. (1986). *Spatial Variation*. Lecture Notes in Statistics, No. 36, Springer.

[63] Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R. (2005). Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in water resources*, 28(2):135–147.

[64] Neal, R. (2001). Annealed importance sampling. *Statistics and Computing*, 11(2):125–139.

[65] Oliver, D. (2017). Metropolized randomized maximum likelihood for improved sampling from multimodal distributions. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):259–277.

[66] Oliver, D. S. and Chen, Y. (2011). Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1):185–221.

[67] Oliver, D. S., Reynolds, A. C., and Liu, N. (2008). *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press.

[68] Patankar, S. (1980). *Numerical heat transfer and fluid flow*. CRC press.

[69] Pele, O. and Werman, M. (2009). Fast and robust earth mover's distances. In *Computer vision, 2009 IEEE 12th international conference on*, pages 460–467. IEEE.

[70] Penny, S. G. and Miyoshi, T. (2016). A local particle filter for high-dimensional geophysical systems. *Nonlinear Processes in Geophysics.*, 23:391–405.

[71] Poterjoy, J. (2016). A localized particle filter for high-dimensional nonlinear systems. *Monthly Weather Review*, 144(1):59–76.

[72] Przybysz-Jarnut, J. K., Hanea, R. G., Jansen, J.-D., and Heemink, A. W. (2007). Application of the representer method for parameter estimation in numerical reservoir models. *Computational Geosciences*, 11(1):73–85.

[73] Reich, S. (2013). A nonparametric ensemble transform method for Bayesian inference. *SIAM Journal on Scientific Computing*, 35(4):A2013–A2024.

[74] Reich, S. and Cotter, C. (2015). *Probabilistic forecasting and Bayesian data assimilation*. Cambridge University Press.

[75] Reichle, R. H., McLaughlin, D. B., and Entekhabi, D. (2002). Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, 130(1):103–114.

[76] Reynolds, A. C., He, N., Chu, L., Oliver, D. S., et al. (1996). Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data. *SPE Journal*, 1(04):413–426.

[77] Ruchi, S. and Dubinkina, S. (2018). Application of ensemble transform data assimilation methods for parameter estimation in reservoir modeling. *Nonlin. Processes Geophys.*, 25:731–746.

[78] Ruchi, S., Dubinkina, S., and Iglesias, M. A. (2019). Transform-based particle filtering for elliptic bayesian inverse problems. *Inverse Problems*, 35(11):115005.

[79] Sakov, P., Haussaire, J. M., and Bocquet, M. (2018). An iterative ensemble Kalman filter in the presence of additive model error. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1297–1309.

[80] Stuart, A. and Taeb, A. (2018). Data assimilation and inverse problems. *arXiv*.

[81] Stuart, A. M. (2010). Inverse problems: a Bayesian perspective. In *Acta Numerica*, volume 19, pages 451 – 559.

[82] Subbey, S., Christie, M., and Sambridge, M. (2004). Prediction under uncertainty in reservoir modeling. *Journal of Petroleum Science and Engineering*, 44(1):143–153.

[83] Talagrand, O. and Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478):1311–1328.

[84] Tierney, L. (1994). Markov chains for exploring posterior distributions. *the Annals of Statistics*, pages 1701–1728.

[85] Tierney, L. et al. (1998). A note on Metropolis-Hastings kernels for general state spaces. *The Annals of Applied Probability*, 8(1):1–9.

[86] Tikhonov, A. N. (1943). On the stability of inverse problems. In *Dokl. Akad. Nauk SSSR*, volume 39, pages 195–198.

[87] Tikhonov, A. N. (1963). Solution of incorrectly formulated problems and the regularization method. *Soviet Math.*, 4:1035–1038.

[88] Tikhonov, A. N. and Arsenin, V. Y. (1977). Solutions of ill-posed problems. 1977. *WH Winston, Washington, DC*, 330.

[89] van Leeuwen, P. J. (2010). Nonlinear data assimilation in geosciences: an extremely efficient particle filter. *Quarterly Journal of the Royal Meteorological Society*, 136(653):1991–1999.

[90] Vefring, E. H., Nygaard, G. H., Lorentzen, R. J., Naevdal, G., Fjelde, K. K., et al. (2006). Reservoir characterization during underbalanced drilling (ubd): methodology and active tests. *SPE Journal*, 11(02):181–192.

[91] Villani, C. (2008). *Optimal transport: old and new*, volume 338. Springer Science & Business Media.

[92] Watson, A., Wade, J., and Ewing, R. (1994). Parameter and system identification for fluid flow in underground reservoirs. In *Proceedings of the Conference Inverse Problems and Optimal Design in Industry*, pages 81–108. Springer.

[93] Weerts, A. H. and El Serafy, G. Y. (2006). Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. *Water Resources Research*, 42(9).

[94] Zafari, M., Reynolds, A. C., et al. (2005). Assessing the uncertainty in reservoir description and performance predictions with the ensemble Kalman filter. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Summary

Efficient utilization of oil/gas reservoirs is of paramount importance to meet increasing global energy demand. The predictions of reservoir models have large economic and environmental impact, and hence, it becomes imperative to develop reliable mathematical models to thoroughly describe their behavior. In this thesis, we investigate the ensemble transform particle filter approach to accurately estimate the petrophysical parameters of reservoir models. For this purpose, we undertake a steady-state single-phase Darcy flow model with permeability as an uncertain parameter. In particular, we focus on problems where the posterior is non-Gaussian even if the prior and the noise distributions are Gaussian.

The standard method of choice to tackle such problems is Markov chain Monte Carlo (MCMC) sampling approach, which is known to be computationally expensive. In this thesis we explore data assimilation alternatives that are computationally more affordable. In Chapter 2, we investigate the performance of an Ensemble Transform Particle Filter (ETPF) for inverse problems. The ETPF is developed on the backbone of a Bayesian approach of sequential Monte Carlo with the framework of a linear transport problem. We examine the performance of the ETPF and an Ensemble Kalman filter (EnKF) for estimating uncertain parameters in low and high-dimensional systems and compare them to importance sampling. The numerical experiments show that while the ETPF outperforms the EnKF for a low-dimensional system, it struggles for a high-dimensional system

and requires localization to decrease the degrees of freedom. The localized version of ETPF updates uncertain grid-dependent parameters locally due to otherwise occurring weight degeneracy. Although localization helps in improving the estimation of mean and variance, the Kullback-Leibler divergence shows that it deteriorates the approximation of the posterior distribution. Moreover, the numerical experiments highlight that the ETPF suffers from sampling error and the issue of weight degeneracy becomes prominent when handling numerous accurate observations. Therefore, in Chapter 3, we introduce an optimal transport-based resampling in an adaptive sequential Monte Carlo (SMC) to tackle these shortcomings of ETPF. We study the behavior of the novel approach in an elliptic inverse problem of inferring hydraulic conductivity from pressure measurements, where we employ MCMC as a reference solution. Numerical experiments emphasize that the optimal transport-based adaptive SMC is promising for high-dimensional systems with non-Gaussian distributed parameters. Moreover, the results show that the optimal transport-based adaptive SMC requires smaller ensemble sizes compared to a monomial-based adaptive SMC to achieve a good estimation of the true non-Gaussian posterior. Moreover, the results highlight that an adaptive SMC with ensemble Kalman filter does not converge to the true non-Gaussian posterior. To further investigate the performance of the proposed method in a more realistic set-up, in Chapter 4 we consider a steady-state single-phase Darcy flow model, where permeability and boundary conditions are uncertain. In this chapter we compare the optimal transport-based adaptive SMC to a regularized ensemble Kalman filter (REnKF), which is an alternative to the adaptive SMC with ensemble Kalman filter. The REnKF is an iterative ensemble Kalman method that also introduces intermediate measures, which however are chosen not based on effective sample size (ESS) but on a discrepancy principle. The discrepancy principle ensures that an ensemble approximation does not overfit observations, since it leads to increase in error approximation otherwise. Additionally, we implement a distance-based localization within the tempering approach for the optimal transport-based adaptive SMC and a covariance localization for REnKF. We implement MCMC as a reference solution for all the test cases. The numerical experiments demonstrate that localization improves estimations obtained by the REnKF and the optimal transport-based adaptive SMC in all considered scenarios. For the case when uncertainty is only in permeability, the localised optimal transport-based adaptive SMC outperforms both localised and non-localized REnKF for both considered ensemble size 100 and 1000. When uncertainty is in both permeability and boundary conditions, the results show that the localized opti-

mal transport-based adaptive SMC outperforms the localized and non-localized REnKF but only at a large ensemble size of 1000. Furthermore, when uncertainty is both in permeability and in boundary conditions but error in boundary conditions is not accounted for in the data assimilation, the localized REnKF outperforms both the localized and non-localized optimal transport-based adaptive SMC. Moreover, the optimal transport-based adaptive SMC is computationally more demanding than the REnKF and requires more iterations than the REnKF.

Samenvatting

Efficiënt gebruik van olie-/gasreservoirs is van het grootste belang om aan de toenemende wereldwijde vraag naar energie te voldoen. De voorspellingen van reservoirmodellen hebben grote economische en ecologische gevolgen, en daarom wordt het noodzakelijk betrouwbare wiskundige modellen te ontwikkelen om hun gedrag grondig te beschrijven. In dit proefschrift onderzoeken we het Ensemble Transform Particle Filter om de petrofysische parameters van reservoirmodellen nauwkeurig te schatten. Daarvoor gebruiken we een steady-state één-fase Darcy-stroommodel met permeabiliteit als een onzekere parameter. We richten ons in het bijzonder op problemen waarbij de a-posteriori-verdeling niet-Gaussiaans is, zelfs als de a-priori-verdeling en de ruisverdelingen Gaussiaans zijn.

De standaardkeuze om dergelijke problemen aan te pakken is Markov Chain Monte Carlo (MCMC) bemonstering, waarvan bekend is dat deze computationeel zwaar is. In dit proefschrift onderzoeken we alternatieven voor data-assimilatie die computationeel efficiënter zijn.

In Hoofdstuk 2 onderzoeken we de prestaties van een Ensemble Transform Particle Filter (ETPF) voor inverse problemen. Het ETPF is ontwikkeld met een Bayesiaanse aanpak van Sequentiële Monte Carlo binnen het kader van een lineair transportprobleem. We onderzoeken de prestaties van het ETPF en een Ensemble-Kalman-Filter (EnKF) voor het schatten van onzekere parameters in laag- en hoogdimensionale systemen en vergelijken ze met Importance Sampling. De numerieke experimenten tonen

aan dat hoewel het ETPF beter presteert dan het EnKF voor een laag-dimensionaal systeem, het worstelt met een hoog-dimensionaal systeem en dat lokalisatie vereist is om het aantal vrijheidsgraden te verminderen. De gelokaliseerde versie van het ETPF werkt onzekere grid-afhankelijke parameters lokaal bij vanwege een anders optredende gewichtsdegeneratie. Hoewel lokalisatie helpt bij het verbeteren van de schatting van het gemiddelde en de variantie, blijkt uit de Kullback-Leibler-divergentie dat het de benadering van de a-posteriori-verdeling verslechtert. Bovendien laten de numerieke experimenten zien dat het ETPF lijdt aan een bemonsteringsfout en dat het probleem van gewichtsdegeneratie een prominente rol speelt bij het omgaan met grote aantallen nauwkeurige waarnemingen.

Daarom introduceren we in Hoofdstuk 3 een herbemonstering gebaseerd op optimaal transport in een adaptieve Sequentiële Monte Carlo (SMC) om deze tekortkomingen van ETPF aan te pakken. We bestuderen het gedrag van de nieuwe aanpak in een elliptisch invers probleem van het afleiden van hydraulische geleidbaarheid uit drukmetingen, waarbij we MCMC als referentieoplossing gebruiken. Numerieke experimenten tonen dat de optimaal transport adaptieve SMC veelbelovend is voor hoog-dimensionale systemen met niet-Gaussiaans verdeelde parameters. Bovendien laten de resultaten zien dat de optimaal transport adaptieve SMC een kleinere ensemblegrootte nodig heeft in vergelijking met een monomiale adaptieve SMC om een goede schatting van de ware niet-Gaussiaanse a-posteriori-verdeling te verkrijgen. Bovendien tonen de resultaten dat een adaptieve SMC met Kalman-filter niet convergeert naar de ware niet-Gaussiaanse a-posteriori-verdeling.

Om de prestaties van de voorgestelde methode in een meer realistische context verder te onderzoeken, beschouwen we in Hoofdstuk 4 een steady-state één-fase Darcy-stroommodel, waar permeabiliteit en randvoorwaarden onzeker zijn. In dit hoofdstuk vergelijken we de optimaal transport adaptieve SMC met een geregulariseerd Ensemble-Kalman-Filter (REnKF), dat een alternatief is voor de adaptieve SMC met Ensemble-Kalman-Filter. De REnKF is een iteratieve ensemble-Kalman-methode die ook tussentijdse maten introduceert, die echter niet worden gekozen op basis van effectieve steekproefgrootte (ESS) maar op basis van een discrepantieprincipe. Het discrepantieprincipe zorgt ervoor dat een ensemblebenadering de waarnemingen niet overfit, omdat het anders tot een toename van de foutbenadering leidt. Daarnaast implementeren we een op afstand gebaseerde lokalisatie binnen de temperingsaanpak voor de optimaal transport adaptieve SMC en een covariantielokalisatie voor REnKF.

We implementeren MCMC als een referentieoplossing voor alle testgevallen. De numerieke experimenten tonen aan dat lokalisatie schattingen verbetert verkregen door de REnKF en de optimaal transport adaptieve SMC in alle beschouwde scenario's. Voor het geval waarin alleen de permeabiliteit onzeker is, presteert de gelokaliseerde optimaal transport adaptieve SMC beter dan zowel gelokaliseerde als niet-gelokaliseerde REnKF voor de beide beschouwde ensemble-groottes 100 en 1000. Wanneer zowel de permeabiliteit als de randvoorwaarden onzeker zijn, tonen de resultaten dat de gelokaliseerde optimaal transport adaptieve SMC de gelokaliseerde en niet-gelokaliseerde REnKF overtreft, maar alleen bij een groot ensemble van 1000. Bovendien, wanneer zowel permeabiliteit als randvoorwaarden onzeker zijn, maar de data-assimilatie geen rekening houdt met de fout in de randvoorwaarden, overtreft de gelokaliseerde REnKF zowel de gelokaliseerde als de niet-gelokaliseerde optimaal transport adaptieve SMC. Bovendien is de optimaal transport adaptieve SMC computationeel zwaarder dan de REnKF, en heeft deze meer iteraties nodig.

List of Publications

1. *Application of ensemble transform data assimilation methods for parameter estimation in reservoir modeling.* **S. Ruchi** and S. Dubinkina. Nonlinear Processes in Geophysics, 25 (2018), pp. 731-746, https://doi.org/10.5194/npg-25-731-2018.

2. *Transform-based particle filtering for elliptic Bayesian inverse problems* **S. Ruchi**, S. Dubinkina and M. A. Iglesias. Inverse Probelms, (2019), https:// doi.org/10.1088 /1361-6420/ab30f3.

3. *Comparison of regularized ensemble Kalman filter and tempered ensemble transform particle filter for an elliptic inverse problem with uncertain boundary conditions.* S. Dubinkina and **S. Ruchi**. Computational Geosciences, Springer, (2019), https://doi.org/10.1007/s10596-019-09904-w. (accepted)

4. *Sinkhorn-approximated transform data assimilation method for elliptic Bayesian inverse problems.* **S. Ruchi**, S. Dubinkina and J. de Wiljes. (in preparation)

Curriculum Vitae

Sangeetika Ruchi was born in August 1989 in Dumra, India. She studied Mechanical Engineering at G.B. Pant University, Pantnagar, India, and completed her Bachelor's thesis on *"Experimental study of heat transfer in a rectangular duct with and without roughness"* under the supervision of Prof. Dr. Lokesh Varshney. After obtaining a Bachelor of Technology degree in 2011, she worked as a teaching assistant for one year at Amrapali Institute, India. She taught the Bachelor course of *mechanical vibration* and *fluid machinery* and conducted the lab classes of *internal combustion engines*. She moved to U.S.A. in 2012 to pursue her Master's in Mechanical Engineering at University of Cincinnati, Ohio, under the University Graduate Scholarship covering her tuition fees. During her Master's she followed various mechanical and aerospace engineering courses. She completed her Master's thesis on *"Computational modeling of laser treatment of port-wine stains based on reduced scattering method"* under the guidance of Prof. Dr. Milind Jog. She received a Master of Science degree in 2015. After defending her Master's thesis she worked as a volunteer at a government-aid primary school in Haldwani, India for three months. She assisted in setting up computer labs at the school and taught computer science. In July 2015, Sangeetika started her PhD research in Applied Mathematics at Centrum Wiskunde & Informatica, Amsterdam, the

Netherlands, under the supervision of Dr. Svetlana Dubinkina and Prof. Dr. Jason Frank. During her PhD she spent three months at University of Nottingham under the supervision of Dr. Marco Iglesias, which led to a joint paper.

After completing the research for her PhD thesis she spent three months at University of Potsdam, Germany, in the group of Prof. Dr. Sebastian Reich under the *SFB1294-Data Assimilation* fellowship for PhD researchers. Collaboration with an active member of the group Dr. Jana de Wiljes has led to a joint paper tentatively titled *"Sinkhorn-approximated transform data assimilation method for elliptic Bayesian inverse problems"*. After the completion of her research fellowship at University of Potsdam, Sangeetika took a position of Design Engineer at the Metrology department of ASML in Eindhoven, the Netherlands.