

Jon Sætrum

Reduction of Dimensionality in Spatiotemporal Models

Doctoral thesis
for the degree of Philosophiae Doctor

Trondheim, November 2010

Norwegian University of Science and Technology
Faculty of Information Technology,
Mathematics and Electrical Engineering
Department of Mathematical Sciences

NTNU

Norwegian University of Science and Technology

Doctoral thesis
for the degree of Philosophiae Doctor

Faculty of Information Technology,
Mathematics and Electrical Engineering
Department of Mathematical Sciences

© 2010 Jon Sætrum.

ISBN 978-82-471-2363-8 (printed version)
ISBN 978-82-471-2364-5 (electronic version)
ISSN 1503-8181

Doctoral thesis at NTNU, 2010:193

Printed by NTNU-trykk

Preface

This thesis is submitted in partial fulfilment of the requirements for the degree of Philosophiae Doctor (PhD) at the Norwegian University of Science and Technology (NTNU). The research is funded by the Uncertainty in Reservoir Evaluation (URE) consortium at NTNU.

I would like to thank my supervisor, Henning Omre, for giving me a flamboyant introduction to statistics at NTNU seven years ago. His ability to not only focus on the theory, but also the interesting applications, is the major reason for my current interest in statistics. During these last three years as a PhD student, his experience and advice has been much appreciated. Furthermore, I would like to thank him for both the academic and the geographical freedom he has given me.

Secondly I would like to thank my co-supervisor Lou Durlofsky, for giving me the opportunity to visit Stanford during the academic year 08-09. The discussions I have had with Lou and his students, Jincong He in particular, have been extremely rewarding and has taught me a great deal about explaining statistical theory to non-statisticians. Our collaboration is therefore something I have enjoyed very much.

I would like to thank Inge Myrseth for our collaboration, and for making the academic conferences more fun to attend. My co-authors at the Statoil research centre in Bergen also deserve a thank you; for giving me the opportunity to work with more realistic reservoir models.

Last, but not least I would like to thank my lovely and radiant wife Hege for all her support and love. Especially I would like to thank her for taking the 15 hour flights to San Fransisco — after working 12 hour shifts offshore Norway for three weeks — during my year at Stanford.

Thesis Outline

Introduction

- Paper I: Ensemble Kalman filtering with shrinkage regression techniques**
Jon Sætrom and Henning Omre
Paper accepted for publication in *Computational Geosciences*,
DOI: 10.1007/s10596-010-9196-0.
The original publication is available at springerlink.com
- Paper II: Ensemble Kalman filtering for non-linear likelihood models using kernel-shrinkage regression techniques**
Jon Sætrom and Henning Omre
Paper submitted for publication
- Paper III: Improved uncertainty quantification in the Ensemble Kalman Filter using statistical model selection techniques**
Jon Sætrom, Joakim Hove, Jan-Arild Skjervheim, and Jon G. Vabø
Paper submitted for publication
- Paper IV: Resampling the Ensemble Kalman Filter**
Inge Myrseth, Jon Sætrom and Henning Omre
Paper submitted for publication
- Paper V: Ensemble Kalman filtering in a Bayesian regression framework**
Jon Sætrom, Inge Myrseth and Henning Omre
Paper submitted for publication
- Paper VI: Enhanced linearised reduced-order models for subsurface flow simulation**
Jincong He, Jon Sætrom and Louis J. Durlofsky
Paper submitted for publication

Contents

1	Introduction	1
1.1	Notation and Model Formulation	4
1.2	Ensemble Kalman Filter	6
1.3	Dimension Reduction Techniques	7
1.4	Overview of Thesis	8
1.4.1	Summary of Papers	9
1.4.2	Closing Remarks	14
1.5	Further Work	15
	Bibliography	17
2	Ensemble Kalman filtering with shrinkage regression techniques	27
2.1	Introduction	28
2.2	Notation and Model Formulation	29
2.3	Classical Ensemble Kalman Filter	31
2.4	Shrinkage Regression Methods	34
2.4.1	Principal Component Regression	34
2.4.2	Partial Least Squares Regression	36
2.4.3	Comments	36
2.4.4	Cross-Validation	38
2.5	Empirical Study	39
2.5.1	Results with Discussion	41
2.5.2	Linear Case	41
2.5.3	Non-Linear Case	44
2.5.4	Ensemble Size	45
2.5.5	Summary	47
2.6	Reservoir Example	48
2.6.1	Reservoir Description	48
2.6.2	Results with Discussion	51
2.7	Conclusion	60

References 2	61
APPENDIX	64
Appendix 2-A Use of Singular Value Decomposition	64
Appendix 2-B Computational Properties	65
Appendix 2-C Prior Distribution	66
Appendix 2-D Rank of the Updated Ensemble	67
3 Ensemble Kalman filtering for non-linear likelihood models using kernel-shrinkage regression techniques	71
3.1 Introduction	72
3.2 Notation and Model Formulation	74
3.3 Classical Ensemble Kalman Filter	76
3.4 Kernel-Shrinkage Regression	77
3.4.1 Kernel Methods	79
3.4.2 Comments	80
3.5 Ensemble Kalman Filtering Using Kernel-Shrinkage Regression	81
3.5.1 Kernel Ridge Regression	81
3.5.2 Kernel Principal Component Regression	82
3.5.3 Kernel Partial Least Squares Regression	83
3.5.4 Comments	83
3.5.5 Model Hyperparameter Selection	84
3.6 Empirical Study	86
3.6.1 Non-Linear Regression	87
3.6.2 Results	91
3.7 Conclusions	97
3.8 Acknowledgements	99
References 3	99
APPENDIX	103
Appendix 3-A Gaussian Random Field	103
Appendix 3-B Computational Properties	104
4 Improved uncertainty quantification in the Ensemble Kalman Filter using statistical model selection techniques	107
4.1 Introduction	108
4.2 Notation and Model Formulation	109
4.2.1 Ensemble Kalman Filter	110
4.2.2 Ensemble Kalman Filter in a Reduced Order Space	111
4.2.3 Cross-Validation	113
4.3 Case Study	114
4.3.1 Results with Discussion	115

4.4	Conclusions	125
	References 4	126
	APPENDIX	128
	Appendix 4-A Computational Properties	128
5	Resampling the Ensemble Kalman Filter	131
5.1	Introduction	132
5.2	Model Assumptions	132
5.3	The Ensemble Kalman Filter	134
	5.3.1 The conditioning step	137
5.4	Resampling	140
5.5	Resampling the EnKF	141
5.6	Empirical study	144
	5.6.1 Model description	144
	5.6.2 Results	145
	5.6.3 Gauss-linear likelihood	145
	5.6.4 Nonlinear likelihood	147
5.7	Reservoir Example	149
	5.7.1 Reservoir and model description	151
5.8	Conclusion	153
	References 5	155
6	Ensemble Kalman filtering in a Bayesian regression framework	161
6.1	Introduction	162
6.2	Notation and problem formulation	163
	6.2.1 Classical Ensemble Kalman Filter	165
	6.2.2 Properties of the Conditioned State Vector Ensemble	167
6.3	Bayesian regression	170
	6.3.1 Informative prior distribution	170
	6.3.2 Non-Informative Prior Distribution	171
	6.3.3 The Kalman Gain Posterior Distributions	172
	6.3.4 Approximate Matrix- t Distribution	173
	6.3.5 Computational Properties	174
6.4	Empirical study	175
	6.4.1 Linear Case	177
	6.4.2 Non-Linear Case	181
6.5	Conclusions	183
	References 6	185
	APPENDIX	190

Appendix 6-A	Multivariate vector and matrix probability distributions	190
Appendix 6-B	Expectation, covariance and coupling of the ensemble	191
Appendix 6-C	Underestimation prediction uncertainty	192
Appendix 6-D	Properties of some multivariate distributions	194
Appendix 6-E	Sampling from the matrix- t distribution	197
Appendix 6-F	Singular matrix variate probability distributions	199
7	Enhanced linearised reduced-order models for subsurface flow simulation	203
7.1	Introduction	205
7.2	Problem Formulation	207
7.2.1	Oil-Water Flow Equations	207
7.2.2	Solution of Discretised System using TPWL	208
7.2.3	Application Example: Reservoir Model 1	211
7.3	Local Resolution TPWL	213
7.3.1	Underfitting and Overfitting	215
7.3.2	Description of Local Resolution Scheme	216
7.3.3	Numerical Results using TPWL(LR)	218
7.4	Stability of TPWL Models	219
7.4.1	Example Showing Instability of TPWL	221
7.4.2	Stability Analysis	221
7.4.3	Stabilisation of TPWL using Optimised Basis	223
7.4.4	Stabilisation of TPWL using Modified Basis	227
7.5	Application of TPWL to Realistic Problems	231
7.5.1	Model 2: Upper Six Layers of SPE 10	231
7.5.2	Reservoir Model 3: Part of SPE 10 Upper 30 Layers	234
7.6	Use of TPWL for Production Optimisation	240
7.6.1	Direct Search Optimisation with TPWL	240
7.6.2	Optimisation Results	241
7.7	Concluding Remarks	242
References	7	244

Chapter 1

Introduction

Over the last decade, it has become increasingly important to extract useful information quickly from vast amounts of data. Examples include analysis of the human genome, development of Internet search engines, automatic recognition of handwritten postal codes, weather forecasting and petroleum reservoir evaluation. Since the introduction of Principal Component Analysis (PCA) by Pearson (1901) and Hotelling (1933), dimension reduction techniques have therefore become an integral part of contemporary statistical data analysis.

Statistical and mathematical modelling of nonlinear spatiotemporal problems in large dimensions is known to be a challenging task. High computational demands makes these problems particularly difficult to solve efficiently, which calls for approximate solutions when considering real-life applications. Reduction of the model dimensionality is therefore a natural approach to consider.

In this thesis, we apply dimension reduction techniques on high dimensional spatiotemporal models. As an inspirational source, we use two computationally demanding problems in petroleum reservoir management; petroleum reservoir characterisation and well control optimisation. Common to both problems is that a repeated number of reservoir flow simulations are needed. This involves solving a set of non-linear Partial Differential Equations (PDE) in high dimensions, which can take days or even weeks when considering traditional finite difference PDE solvers on large reservoir models.

Petroleum reservoir characterisation aims at extracting information of the earth's subsurface from seismic data, well tests, and observed production data. When considering large reservoir domains, this involves solving a spatiotemporal inverse problem in high dimensions (Omre and Tjelmeland,

1997). Bayesian inversion, where the unknown properties of the reservoir can be evaluated by generating realisations from the posterior probability density function (pdf) of interest, is therefore a natural statistical approach to consider. However, the non-linear structure of the inverse problem makes the posterior pdf analytically intractable. Moreover, high computational demands currently restrict the usefulness of asymptotically exact sampling techniques such as Markov chain Monte Carlo (MCMC) (Barndorff-Nielsen et al., 2001) or Particle Filters (PF) (Doucet et al., 2001; Künsch, 2005).

The Ensemble Kalman Filter (EnKF), introduced by Evensen (1994), is an approximate Monte Carlo method that can efficiently solve non-linear spatiotemporal inverse problems in high dimensions. A source of inspiration for the EnKF is the Kalman Filter (KF) (Kalman, 1960). The KF provides analytical expressions for the model parameters of the Gaussian posterior pdf assuming linear system dynamics with a Gaussian prior pdf and a linear Gaussian likelihood model, termed the Gauss-linear model. In the EnKF, an ensemble of realisations is used to represent the statistical properties of the posterior pdf. This is conceptually different from alternative non-linear filters, such as the extended KF (Jazwinski, 1970) and the unscented KF (Julier and Uhlmann, 1997), which uses the posterior mean and covariance to represent the statistical properties of the model. As a result, the EnKF can efficiently handle non-linear system dynamics and is reasonably robust with respect to the prior model assumptions (Evensen, 2007). Hence, the EnKF has been applied to numerous fields over the last decade, such as numerical weather prediction (Houtekamer et al., 1996; Houtekamer and Mitchell, 2001), oceanography (Keppenne and Rienecker, 2003; Leeuwenburgh et al., 2005), hydrology (Moradkhani et al., 2005), and petroleum reservoir characterisation (Nævdal et al., 2003; Aanonsen et al., 2009).

Having characterised the properties of the reservoir, the next step is often to optimise the extraction of hydrocarbons. This can be done by changing the operating conditions of the reservoir, also known as the well controls, e.g. by injecting water or gas into the reservoir at different rates or modifying the pressure in the producing wells. To accomplish this task, we need to solve a non-linear multiobjective optimisation problem. Solving this problem typically require several thousand reservoir flow simulations, meaning that approximate PDE solvers are often called for.

The Trajectory Piecewise Linearisation (TPWL) procedure, introduced by Rewiński (2003), is an approximate PDE solver which has been applied successfully to non-linear heat transfer problems (Yang and Shen, 2005), electromechanical systems (Bond and Daniel, 2007; Vasilyev et al., 2006), computational fluid dynamics (Gratton and Willcox, 2004), and reservoir

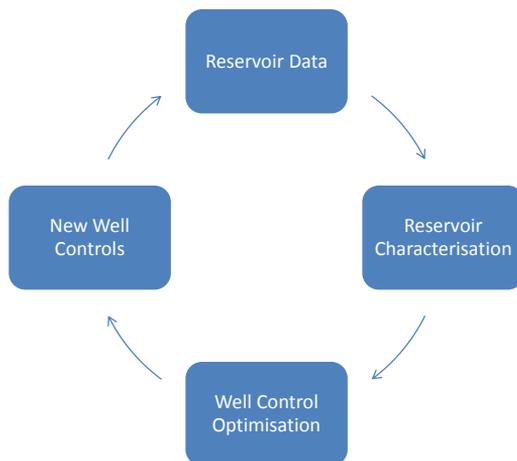


Figure 1.1: Closed-loop reservoir management.

flow simulation (Cardoso and Durlofsky, 2010). The main idea of the TPWL is that new solutions of the PDE with different well controls can be found using a Taylor expansion around previously evaluated well controls. This implies that a solution to the non-linear PDE is given by solving a set of linear equations, rather than through non-linear PDE solvers. By further applying dimension reduction techniques, this entails that we can perform reservoir flow simulation in less than a second, compared with a few hours using traditional solvers. This brings us one step closer towards automatic, or closed-loop, reservoir management, depicted in Figure 1.1. Here, the goal is to update the model automatically and in real-time as new production or seismic data arrives, to improve the future predictions.

The main focus in this thesis is solving spatiotemporal inverse problems using the EnKF, which we consider in Chapters 2 through 6. After a brief presentation of the notation and a probabilistic formulation of the spatiotemporal inverse problem, we will therefore proceed with a description of the EnKF updating scheme, presented from both a Bayesian and frequentist perspective. We will then proceed with a short description of PCA from a

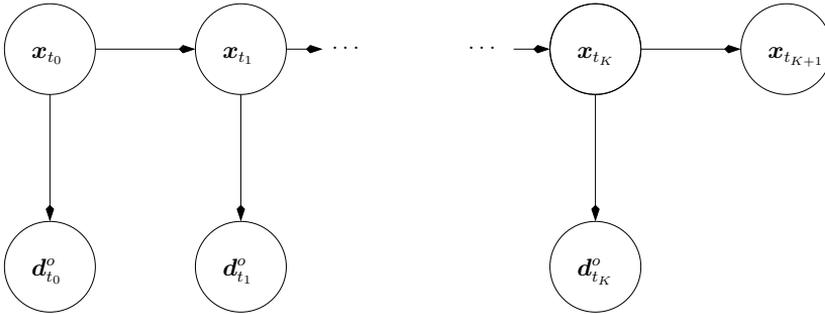


Figure 1.2: Stochastic Directed Acyclic Graph (DAG) of the spatiotemporal inverse problem.

statistical point of view, and briefly discuss a potential pitfall concerning the selection of the subspace dimension. The introductory chapter is ended with a short summary of the papers, where we emphasise new contributions in this thesis, and a discussion regarding future work.

1.1 Notation and Model Formulation

Let $\mathbf{x} \in \mathbb{R}^{n_x \times 1}$ denote that \mathbf{x} is an n_x -dimensional column vector in the real space, with its transpose denoted by \mathbf{x}^T . Further let the matrix \mathbf{A} having a rows and b columns with real entries be denoted by $\mathbf{A} \in \mathbb{R}^{a \times b}$. For notational convenience, we will use the same notation for random vectors and matrices.

A graphical description of the spatiotemporal forecast problem is given in Figure 1.2. Here $\mathbf{x}_{t_k} \in \mathbb{R}^{n_x \times 1}$ denotes the state of the unknown random vector of interest at time step k and time t_k , and correspondingly $\mathbf{d}_{t_k}^o \in \mathbb{R}^{n_d \times 1}$ denotes the vector of observed data. In a reservoir characterisation setting, \mathbf{x} can describe the unknown permeability, porosity, saturation and pressure of the reservoir, whilst \mathbf{d} can be the observed oil production rate, seismic signal, or borehole data. For simplicity, we will from now on drop the subscript t_k , and write \mathbf{x}_k , \mathbf{d}_k^o . Furthermore, \mathbf{x} and \mathbf{d} will be referred to as the state and observation vector respectively. In the predictive setting, the goal is to improve the estimates of the state vector sequentially as more data is collected, and ultimately use the updated model for prediction purposes.

Let $f(\mathbf{x}_0)$ be the prior pdf of the state vector at the initial time step. Furthermore, let $\mathbf{x}_{k+1} = \boldsymbol{\omega}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{x}_k})$, $k = 0, \dots, K$, where $\boldsymbol{\omega} : (\mathbb{R}^{n_x \times 1} \times \mathbb{R}^{n_x \times 1}) \rightarrow \mathbb{R}^{n_x \times 1}$ is a known, possibly non-linear forward function. In the

reservoir characterisation problem, $\omega(\cdot, \cdot)$ describes the fluid flow in the reservoir between consecutive time steps. Here $\epsilon_{\mathbf{x}_k}$ represents model errors in the forward model, following a known pdf $f(\epsilon_{\mathbf{x}_k})$. Note that this is in accordance with the Markov assumptions in Figure 1.2, which entails that the conditional distribution of the state vector at time step $k+1$ given the state at all previous time steps is:

$$f(\mathbf{x}_{k+1}|\mathbf{x}_{0:k}) = f(\mathbf{x}_{k+1}|\mathbf{x}_k).$$

For notational convenience, we denote the sequence $\mathbf{x}_0, \dots, \mathbf{x}_l$ by $\mathbf{x}_{0:l}$. Further note that the conditional pdf $f(\mathbf{x}_{k+1}|\mathbf{x}_k)$ is implicitly defined through the flow simulator $\omega(\cdot, \cdot)$. Hence, the prior pdf of the state vector at all time steps factorises as

$$f(\mathbf{x}_{0:(K+1)}) = f(\mathbf{x}_0) \prod_{i=0}^K f(\mathbf{x}_{i+1}|\mathbf{x}_i).$$

The observed data, \mathbf{d}_k^o , is connected to the state vector through a possibly non-linear function: $\mathbf{d}_k^o = \zeta(\mathbf{x}_k, \epsilon_{\mathbf{d}_k})$, $k = 0, \dots, K$, where $\epsilon_{\mathbf{d}_k}$ represents the likelihood model and observation errors. Thus, the likelihood function $f(\mathbf{d}_k|\mathbf{x}_k)$ is defined through $\zeta(\cdot, \cdot)$.

Having a fully specified prior and likelihood model, the spatiotemporal filter and forecast problem can be assessed by generating realisations \mathbf{x}_k^c and \mathbf{x}_{k+1}^u from the posterior pdfs $f(\mathbf{x}_k|\mathbf{d}_{0:k})$ and $f(\mathbf{x}_{k+1}|\mathbf{d}_{0:k})$ respectively for $k = 0, \dots, K$. Bayesian inversion provides a sequential solution to this problem. Using Bayes rule and the Markov properties of the prior model, we get for $k = 1, \dots, K$;

$$\begin{aligned} f(\mathbf{x}_k|\mathbf{d}_{0:k}^o) &\propto f(\mathbf{x}_k|\mathbf{d}_{0:(k-1)}^o)f(\mathbf{d}_k|\mathbf{x}_k) \\ f(\mathbf{x}_{k+1}|\mathbf{d}_{0:k}^o) &= \int f(\mathbf{x}_{k+1}|\mathbf{x}_k)f(\mathbf{x}_k|\mathbf{d}_{0:k}^o)d\mathbf{x}_k. \end{aligned} \quad (1.1)$$

In general, the conditional pdfs defined in Eq. (1.1) is only known up to an unknown normalising constant. Thus, we can apply computationally intensive techniques based on MCMC or rejection sampling to generate realisations from the posterior distributions. However, as mentioned above, these techniques can be computationally prohibitive for a reservoir characterisation problem. An approximate solution is to assume that $\mathbf{x}_k^u \sim f(\mathbf{x}_k|\mathbf{d}_{0:(k-1)}^o)$ and $\mathbf{d}_k^o \sim f(\mathbf{d}_k|\mathbf{x}_k)$ follow a probability distribution that ensures analytical tractability of $f(\mathbf{x}_k|\mathbf{d}_{0:k}^o)$. This corresponds to the model assumptions made in classical EnKF, when derived from a Bayesian perspective.

1.2 Ensemble Kalman Filter

Let $\mathbf{x}_0^{c(i)}$ be a realisation from the posterior distribution at the initial time step, $f(\mathbf{x}_0|\mathbf{d}_0^o) \propto f(\mathbf{x}_0)f(\mathbf{d}_0|\mathbf{x}_0)$. In addition, let for $k = 1, \dots, K$, $\mathbf{x}_k^{u(i)} = \boldsymbol{\omega}(\mathbf{x}_{k-1}^{c(i)}, \boldsymbol{\epsilon}_{\mathbf{x}_k}^{(i)})$ and $\mathbf{d}_k^{(i)} = \boldsymbol{\zeta}(\mathbf{x}_k^{u(i)}, \boldsymbol{\epsilon}_{\mathbf{d}_k}^{(i)})$, where we as an approximation assume that:

$$\begin{bmatrix} \mathbf{x}_k^{u(i)} \\ \mathbf{d}_k^{(i)} \end{bmatrix} \sim \text{Gauss}_{n_y} \left(\begin{bmatrix} \boldsymbol{\mu}_{\mathbf{x}_k^u} \\ \boldsymbol{\mu}_{\mathbf{d}_k} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{x}_k^u} & \boldsymbol{\Sigma}_{\mathbf{x}^u, \mathbf{d}_k} \\ \boldsymbol{\Sigma}_{\mathbf{d}, \mathbf{x}_k^u} & \boldsymbol{\Sigma}_{\mathbf{d}_k} \end{bmatrix} \right), \quad (1.2)$$

with $n_y = n_x + n_d$. Here the notation $\mathbf{y} \sim \text{Gauss}_{n_y}(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y)$ is used to denote that $\mathbf{y} \in \mathbb{R}^{n_y \times 1}$ follows the multivariate Gaussian distribution with mean $\boldsymbol{\mu}_y$ and covariance matrix $\boldsymbol{\Sigma}_y$ (Mardia et al., 1979). Because the focus from now on will be on a single time step, we will drop the subscript k from the notation.

Under the Gaussian assumption in Eq. (1.2), a realisation from the posterior pdf $f(\mathbf{x}_k|\mathbf{d}_{0:k}^o)$ can be generated by

$$\mathbf{x}^{c(i)} = \mathbf{x}^{u(i)} + \boldsymbol{\Sigma}_{\mathbf{x}^u, \mathbf{d}} \boldsymbol{\Sigma}_{\mathbf{d}}^{-1} (\mathbf{d}^o - \mathbf{d}^{(i)}). \quad (1.3)$$

This follows because the Gaussian distribution is closed under linear operations, and by computing the expected value and covariance of $\mathbf{x}^{c(i)}$ (Mardia et al., 1979).

As seen from Eq. (1.3), the covariance matrices $\boldsymbol{\Sigma}_{\mathbf{x}^u, \mathbf{d}}$ and $\boldsymbol{\Sigma}_{\mathbf{d}}$ are the only two model parameters included, forming the Kalman gain matrix, $\mathbf{K} = \boldsymbol{\Sigma}_{\mathbf{x}^u, \mathbf{d}} \boldsymbol{\Sigma}_{\mathbf{d}}^{-1} \in \mathbb{R}^{n_x \times n_d}$. In the Gauss-linear case, the Kalman gain matrix will be analytically obtainable through the Kalman filter recursions. However, in the general setting analytical tractability is lost. In the EnKF, we therefore use an ensemble of n_e realisations $\{(\mathbf{x}^{u(1)}, \mathbf{d}^{(1)}), \dots, (\mathbf{x}^{u(n_e)}, \mathbf{d}^{(n_e)})\}$ to obtain empirical estimates of the unknown covariance matrices.

For notational convenience, let $\mathbf{X} = [\mathbf{x}^{u(1)}, \dots, \mathbf{x}^{u(n_e)}] \in \mathbb{R}^{n_x \times n_e}$ and $\mathbf{D} = [\mathbf{d}^{(1)}, \dots, \mathbf{d}^{(n_e)}] \in \mathbb{R}^{n_d \times n_e}$ be the ensemble matrices for the state- and observation vector respectively. This gives a classical empirical estimate of the Kalman gain matrix:

$$\hat{\mathbf{K}} = \hat{\boldsymbol{\Sigma}}_{\mathbf{x}^u, \mathbf{d}} \hat{\boldsymbol{\Sigma}}_{\mathbf{d}}^{-1} = \mathbf{X} \mathbf{H} \mathbf{D}^T (\mathbf{D} \mathbf{H} \mathbf{D}^T)^{-1}, \quad (1.4)$$

where $\mathbf{H} = \mathbf{I} - 1/n_e \mathbf{1} \mathbf{1}^T \in \mathbb{R}^{n_e \times n_e}$ is the idempotent centring matrix. Here \mathbf{I} is the identity matrix and $\mathbf{1}$ is a column vector with all entries equal to one; both having proper dimensions. This gives the following EnKF updating scheme in a general setting:

$$\mathbf{x}^{c(i)} = \mathbf{x}^{u(i)} + \hat{\mathbf{K}} (\mathbf{d}^o - \mathbf{d}^{(i)}), \text{ for } i = 1, \dots, n_e. \quad (1.5)$$

Because we are using consistent estimators for the unknown covariance matrices, the EnKF will be consistent with the KF in the Gauss-linear model setting as the ensemble size tends to infinity (Furrer and Bengtsson, 2007).

An alternative derivation of the EnKF updating scheme is to consider the following multivariate linear regression problem:

$$[\mathbf{x}|\mathbf{d}] = \mathbf{K}\mathbf{d} + \boldsymbol{\epsilon}_r,$$

where $\mathbf{K} \in \mathbb{R}^{n_x \times n_d}$ is the unknown matrix of regression coefficients and $\boldsymbol{\epsilon}_r$ denotes the regression error term. The classical, least squares estimate of \mathbf{K} is

$$\hat{\mathbf{K}} = \arg \min_{\mathbf{K}} \text{tr} \{ (\mathbf{X}\mathbf{H} - \mathbf{K}\mathbf{D}\mathbf{H})(\mathbf{X}\mathbf{H} - \mathbf{K}\mathbf{D}\mathbf{H})^T \}, \quad (1.6)$$

where $\text{tr}\{\cdot\}$ denotes the trace operator. Solving Eq. (1.6) analytically recovers the Kalman gain estimate in Eq. (1.4) (Seber and Lee, 2003). Hence, the estimation of the unknown Kalman gain matrix can be viewed as a multivariate linear regression problem (Anderson, 2003a).

Although the EnKF has seen several successful applications, the algorithm can display poor performance resulting from the following:

- Severe violation of the Gaussian model assumptions implicitly used in the updating, such as discrete variables in the state vector.
- Highly non-linear likelihood models.
- Estimation uncertainty in the Kalman gain matrix when the ensemble size is restricted by high computational demands.
- Increasingly larger coupling of the updated ensemble members as we assimilate data sequentially.

In this thesis, we will consider remedies for the final three issues. More specifically, dimension reduction techniques known from multivariate linear regression theory (Hastie et al., 2009) are used to improve on the classical Kalman gain estimate.

1.3 Dimension Reduction Techniques

The main motivation for considering dimension reduction techniques is often to reduce the computational demands in an algorithm. Hence, a large number of dimension reduction schemes exist, which are tailored to solve a particular problem efficiently (Fodor, 2002; van der Maaten et al., 2009). In

this introductory chapter, we will only briefly present the PCA dimension reduction technique, also known as the Karhunen-Loève transform and Proper Orthogonal Decomposition. PCA is one of the most frequently used dimension reduction techniques in statistics, computer science and mathematics. This is mainly because the technique is straightforward to implement using Singular Value Decomposition (Golub and van Loan, 1996), the orthonormal properties, and that PCA is the linear dimension reduction scheme that minimises the mean squared reconstruction error (Pearson, 1901).

From a statistical perspective, the aim of PCA is to explain the covariance structure of a random variable $\mathbf{y} \in \mathbb{R}^{n_y \times 1}$ through a small number of orthogonal linear combinations termed Principal Components (PC), $z_i = \mathbf{v}_i^T \mathbf{y}$, $i = 1, \dots, p$. The PC are selected such that the variance of z_i is maximised under the orthonormality condition; $\mathbf{v}_i^T \mathbf{v}_i = 1$ for $i = 1, \dots, p$. It can be shown (Anderson, 2003b) that the i th sample PC is $z_i^{(j)} = \mathbf{v}_i^T \mathbf{y}^{(j)}$, where \mathbf{v}_i is the i th eigenvector of the empirical covariance matrix estimated based on realisations $\mathbf{y}^{(j)}$, $j = 1, \dots, n_e$. Hence, in situations where the dimension of the input vector, n_y is much larger than the subspace dimension, p , the computational demands of an algorithm can be dramatically reduced by working with the sample PC, $\mathbf{z}^{(j)} \in \mathbb{R}^{p \times 1}$ rather than $\mathbf{y}^{(j)}$.

A commonly used criterion for selection the subspace dimension, p , is based on the explained variance of the sample PC:

$$\varsigma_k = \frac{\hat{\lambda}_k}{\sum_{i=1}^r \hat{\lambda}_i}, \quad k = 1, \dots, r \quad (1.7)$$

where $\hat{\lambda}_k$ is the k th eigenvalue of the empirical covariance matrix, with a rank equal to r . That is, the subspace dimension is selected such the first p PC explains e.g. 90 percent of the total variation in the data; $p = \arg \min_k \{ \sum_{l=1}^k \varsigma_l \geq 0.90 \}$. What is often overlooked, however, is that the estimated eigenvalues can be severely biased compared with the true eigenvalues (Ledoit and Wolf, 2004). Furthermore, a model selection criterion based on ς_k does not necessarily take into account the properties that are important for the specific model, such as the predictive capabilities in a regression setting. In this thesis, the importance of proper selection of the subspace dimension and model validation is illustrated in Chapters 2, 3, 4, and 7.

1.4 Overview of Thesis

The first part of this thesis consists of Papers I through V (Chapters 2–6), where we consider the reservoir characterisation problem using the EnKF.

The focus in Papers I, II and III is shrinkage regression and model selection in an EnKF setting, whilst Papers IV and V introduces the idea of a stochastic Kalman gain matrix. Instead of using a common plug-in estimate of the Kalman gain matrix for each updated ensemble member, we can instead generate separate realisations from the matrix variate pdf of \mathbf{K} . Papers I through V are best read chronologically, although they are self-contained. Paper VI, (Chapter 7) considers the production optimisation problem using the fast TPWL reservoir flow simulator as a surrogate model.

1.4.1 Summary of Papers

A brief summary of the six scientific papers in this thesis follows, emphasising new contributions.

Paper I: Ensemble Kalman filtering with shrinkage regression techniques

Jon Sætrom and Henning Omre

In this paper, we define alternative EnKF updating schemes using the shrinkage regression techniques known as Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR). Common for both methods is that the Kalman gain matrix is estimated using multivariate linear regression where the data vector is projected into a reduced order space. The main difference between the two methods is that whilst the projection in PCR is done without considering the values of the state vector, the supervised PLSR technique uses the values of both the data and state vector.

The classical EnKF updating scheme and the two EnKF updating schemes based on PCR and PLSR are tested on two synthetic models where the former is a Gauss-linear model, whilst the latter includes a non-linear forward model. We demonstrate the importance of model validation when selecting the subspace dimension for the PCR technique, by comparing the analytical criterion based on the estimated eigenvalues and a Cross-Validation (CV) scheme. As expected, the CV scheme, which takes into account the predictive capabilities of the regression model, provides much better results. For these synthetic examples, the supervised PLSR scheme appears as slightly better compared with the PCR scheme.

Because the updated ensemble members are based on a common plug-in estimate of the unknown Kalman gain matrix, they will necessarily become increasing coupled as more data is assimilated. For this reason, we see severe problems of model overfitting and an underestimation of the prediction variance for small ensemble sizes using the classical EnKF. This is true for both

the two synthetic reservoir examples and for a synthetic reservoir model, where using the PLSR scheme led to dramatic improvements in the forecast precision, with associated prediction variance, compared with the classical EnKF scheme.

New contributions in this paper is usage of the PCR and PLSR shrinkage regression techniques to estimate the unknown Kalman gain matrix. Furthermore, we look at the importance of using a model validation scheme when selecting the subspace dimension, and suggest an automatic CV scheme to accomplish this task efficiently. Finally, we demonstrate that the collapse of the updated ensemble in the presented examples is a result of increasingly collinear ensemble members, leading to severe problems of model overfitting in the classical EnKF updating scheme for small ensemble sizes. We therefore see a significant improvement in the results when using the shrinkage based regression techniques in combination with the CV scheme.

Paper II: Ensemble Kalman filtering for non-linear likelihood models using kernel-shrinkage regression techniques

Jon Sætrum and Henning Omre

In this paper, we apply non-linear (kernel) extensions of the shrinkage regression techniques presented in Paper I, to handle highly non-linear likelihood models in an EnKF setting efficiently. Kernel methods are in most of the literature presented in mathematical terms, or simply as a black box. In this paper, however, we present kernel methods using a probabilistic discussion of a univariate non-linear regression problem with multivariate predictor variables. Kernelised version of the Ridge regression, PCR and PLSR techniques are then presented, and applied in an EnKF setting. We use a hidden Markov model with a highly non-linear likelihood model to demonstrate that the suggested updating schemes provide improved results compared with the classical EnKF updating scheme.

New contributions in this paper include the non-linear extension of the EnKF updating scheme to non-linear likelihood models using kernel-shrinkage regression techniques. Again, we demonstrate the potential problem of regression model overfitting in an EnKF setting.

Paper III: Improved uncertainty quantification in the Ensemble Kalman Filter using statistical model selection techniques

Jon Sætrom, Joakim Hove, Jan-Arild Skjervheim and Jon G. Vabø

In this paper, we focus on model selection, and the effect of model overfitting in an EnKF setting. Similar to Papers I and II, we formulate the classical EnKF updating scheme through a dimension reduction of the data ensemble. The main difference between the formulation in this paper and the formulation in Paper I, is that the likelihood model is assumed to have an additive Gaussian noise term, with zero mean and known covariance. Hence, the covariance matrix of the observation error is used directly in the estimation of the Kalman gain matrix, similar to what is done in the classical formulation of the EnKF (Evensen, 1994; Burgers et al., 1998). We proceed with a short discussion regarding selection of the subspace dimension using estimated eigenvalues, and suggest a CV scheme to prevent the potential problem of model overfitting.

New contributions in this paper include the extension of the model selection scheme based on the CV to the previously suggested EnKF algorithms based on dimension reduction. We demonstrate that the size of the subspace dimension is crucial to the predictive power of the updated model and that biased predictions occur when the subspace dimension is too large. A benchmark synthetic reservoir model developed at the Statoil research centre in Bergen, demonstrates the potential gain of using statistical model selection techniques that take into account the predictive power of the model.

Paper IV: Resampling the Ensemble Kalman Filter

Inge Myrseth, Jon Sætrom and Henning Omre

In this paper, we formulate the EnKF by generating realisations of the Kalman gain matrix from an empirically estimated sample distribution, referred to as the Resampling EnKF (ResEnKF). More specifically, non-parametric bootstrapping is used to generate individual realisations of the Kalman gain matrix in the EnKF updating scheme. This is an extension of the Hierarchical EnKF (HEnKF) (Myrseth and Omre, 2010), where the unknown covariance matrices, which defines the Kalman gain matrix, are generated separately from analytically tractable posterior distributions. A comparison between the classical EnKF, the ResEnKF, and the HEnKF updating scheme is done on a synthetic reservoir example. We observe significant improvement in the assessment of the prediction variance for both

the ResEnKF and HEnKF updating schemes compared with the classical EnKF updating scheme.

New contributions in this paper include the use of the bootstrap to generate realisations from the unknown pdf of the Kalman gain matrix. This is somewhat different from traditional applications of the bootstrap, where a bootstrap estimate of the unknown parameter is formed from the mean of several realisations from the sample distribution. Furthermore, a heuristic argument is given why the classical EnKF updating scheme introduces coupling in the updated ensemble members.

Paper V: Ensemble Kalman filtering in a Bayesian regression framework

Jon Sætrum, Inge Myrseth and Henning Omre

In this paper, we describe a hierarchical Bayesian extension of the classical EnKF algorithm, where we assign prior pdfs to the unknown model parameters. Rather than using a common plug-in estimate of the Kalman gain matrix for each updated ensemble member, we generate separate realisation from analytically matrix-variate posterior pdfs of the unknown Kalman gain matrix. Again, this can be viewed as an extension of the HEnKF scheme.

We use the connection between the EnKF updating scheme and multivariate linear regression to present theoretical results regarding the updated ensemble members when a common plug-in estimate is used for the Gauss-linear model. The results show that because the update ensemble members are coupled through the Kalman gain matrix, the predictive covariance matrix will necessarily be underestimated. These results are in accordance with previous results and empirical studies in the EnKF literature (Houtekamer and Mitchell, 1998, 1999; van Leeuwen, 1999; Sacher and Bartello, 2008; Thomas et al., 2009), although the focus in these studies is on model error and estimation error in the Kalman gain matrix.

We proceed by evaluating the posterior distribution of the Kalman gain matrix under two prior models. The first one considers a conjugate family prior (CP), whilst the non-informative Jeffreys prior (JP) (Jeffreys, 1946) is used in the second. For both cases, the posterior distribution of the Kalman gain matrix is of matrix- t type. To handle high dimensional problems efficiently, we combine the posterior distribution based on the Jeffreys prior with the shrinkage regression techniques of Paper I. This results in the definition of a singular matrix- t distribution from which we can quickly generate realisations. The synthetic models defined in Paper I, are used to evaluate the EnKF updating schemes based on Bayesian regression techniques.

Improved estimates of the prediction mean and variance is seen compared with the classical EnKF, both for the conjugate and Jeffreys prior models. When the prior hyperparameters are carefully selected, the EnKF updating scheme based on the CP distributions provides the best results. However, for a poorly selected prior distribution the CP-EnKF updating scheme obtain similar results as the classical EnKF updating scheme. We therefore conclude that the JP-EnKF updating scheme is a more robust alternative. Moreover, the JP-EnKF updating scheme requires considerably less computational effort than the CP-EnKF updating scheme.

New contributions in this paper include the theoretical results regarding the updated ensemble members in the classical setting when we treat the Kalman gain as a random matrix. These results explain that because the conditioned ensemble members are correlated, the estimated prediction variance will necessarily be underestimated. The results further explain that applying dimension reduction to the data vector, will reduce the ensemble coupling. For completeness, we extend the results in Furrer and Bengtsson (2007) to the multivariate case, where we prove that the EnKF updating scheme will underestimate the prediction variance without assigning a pdf to the unknown Kalman gain matrix.

To reduce the ensemble coupling, we use classical result known from Bayesian regression to define EnKF updating schemes where the Kalman gain matrix is generated separately for each ensemble member from a matrix- t distribution. Furthermore, we provide, to our knowledge, a new formulation of the singular matrix- t distribution, and apply this to efficiently handle models where the dimension of the state and data vector is large while the ensemble size is small. Shrinkage regression techniques are used to improve the predictive capabilities of the model.

Paper VI: Enhanced linearised reduced-order models for subsurface flow simulation

Jincong He, Jon Sætrom and Louis J. Durlofsky

In this paper, we continue the work done by Marco Cardoso (Cardoso, 2009) regarding reservoir flow simulation using the Trajectory Piecewise Linearisation (TPWL) technique. The goal of this work is to develop a fast, accurate and stable reservoir flow simulator specifically designed for well-control optimisation problems. That is, situations where multiple flow simulations are needed for a constant reservoir geology. We identify and address two problems that can degenerate the TPWL solution: Reconstruction of the reservoir state vector at important grid blocks and system stability.

We start this paper with a brief overview of the physical equations governing the oil and water flow in a petroleum reservoir, followed by a formulation of the TPWL through a first order Taylor expansion of the non-linear equations and PCA. The information from previously generated training simulations is used in both the PCA and the Taylor expansion. Next, we investigate some of the limitations of the previously defined TPWL reservoir simulator. More specifically, the limitations of using PCA as a dimension reduction technique when there is a large difference between the schedules used for training and the simulation well schedules. Because reconstruction is only necessary at a limited number of reservoir grid blocks, a localised version of the PCA is suggested, thus eliminating the reconstruction error in the TPWL scheme.

It is known that stability of the TPWL scheme is not guaranteed if the PCA dimension reduction technique is used (Bond and Daniel, 2009). The stability of the TPWL scheme is directly related to the spectral radius of the amplification matrices that are part of the TPWL representation, which we show is related to the selected subspace dimension. We therefore propose two model selection schemes, where both of them improve the stability properties of the TPWL model by optimising the subspace dimension using two different prepossessing steps. Both methods are successfully applied to various reservoir models, where the previously defined TPWL scheme was unstable.

The new contributions in this paper include the use of the localised PCA dimension reduction technique to eliminate the reconstruction error in the TPWL scheme. We use the spectral radius of the amplification matrices as an efficient tool for model selection in terms of stability. Furthermore, the TPWL scheme is combined with a generalised pattern-search optimisation procedure to solve a well-control optimisation problem efficiently.

1.4.2 Closing Remarks

The most important contribution in this thesis is the development of the theoretical results in Paper V. These results provides a formula that explain the behaviour seen in numerous empirical studies in the EnKF literature over the last decade. We can further explain that applying shrinkage regression techniques will reduce the ensemble coupling, which leads to improved EnKF results as seen in Papers I and III.

Another important contribution is the usage of kernel shrinkage regression techniques to efficiently handle non-linear likelihood models in an EnKF setting. One of the major limitations of the classical EnKF updating scheme is the implicit assumption of a linear relationship between the state and data

vector. This assumption does not hold for many spatiotemporal models in both weather forecasting and petroleum reservoir evaluation. The kernel shrinkage regression techniques can potentially handle this problem, without significantly increasing the computational demands compared with the classical EnKF updating scheme.

A third and final important contribution is the work done to further develop the extremely fast, approximate flow simulator based on the TPWL technique. By identifying that reconstruction error and stability are two major sources of error in the TPWL, we were able to address these problems efficiently. Because solving a well optimisation problem can require thousands of repeated flow simulations, developing fast approximate simulators are of obvious interest.

1.5 Further Work

In this thesis, we only consider three shrinkage regression techniques in an EnKF setting where the L_2 -norm is in focus. Alternative methods based on the L_1 -norm, such as the LASSO (Tibshirani, 1996) and Least Angle Regression (Efron et al., 2004), will therefore be interesting to consider. Furthermore, a comparative study of the shrinkage regression EnKF schemes described in this thesis and the previously suggested EnKF “shrinkage” schemes based on covariance localisation (Hamill et al., 2001) would be interesting. A more thorough investigation regarding the properties of the kernel-shrinkage regression techniques in an EnKF setting is also called for. Alternative formulations, e.g. using techniques known from Bayesian kriging (Omre, 1987), can also be looked into. We would also like to improve on the error statistics and model validation schemes considered in this thesis. Ideas based on proper scoring rules (Gneiting and Raftery, 2007), would therefore be interesting to consider.

Work currently in progress, is the combination of shrinkage regression techniques and the bootstrap, which combines Papers I, IV, and V. Preliminary results show that this approach can indeed lead to significant improvement in the quantification of the prediction uncertainty resulting from a reduction in the coupling of the updated ensemble members.

Because the TPWL scheme is currently defined as a deterministic process, it would be interesting to include a stochastic term in the formulation. With this approach, we can potentially take into account the bias introduced in the Taylor approximation, and will make it easier to quantify the uncertainty of the predicted well controls. Ideas along the lines of the probabilistic PCA (Tipping and Bishop, 1999), or the scale-corrected EnKF (Lødøen and

Omre, 2008) can be considered. Furthermore, the TPWL can potentially be used to solve both the reservoir characterisation and production optimisation problem efficiently. Preliminary work done by Jincong He (He, 2010), suggest that the TPWL can be used when both the static and dynamic properties of the reservoir model changes in the training and test simulations. However, these problems will require a larger number of training simulations compared to the production optimisation problem where the reservoir geology is constant.

A final idea that can be explored is using the TPWL solution as a proposal in a McMC based reservoir characterisation scheme. Through a careful implementation of the TPWL scheme, we can potentially generate realisations from the proposal distribution in centiseconds on modern laptop computers, rather than hours or days. Thus, we can potentially eliminate the time consuming part of the McMC scheme using the TPWL. This opens the door for solving reservoir evaluation problems where the state vector is discrete, which is currently considered the crux in petroleum reservoir characterisation. Alternatively, we can in the spirit of the linear regression view of the EnKF updating scheme, e.g construct posterior probability maps for the state vector conditioned on the data at each time step using Generalised Linear Models (McCullagh and Nelder, 1999). Multi-point geostatistical simulation tools (Strebelle, 2002; Zhang et al., 2006) can then be used to simulate a new conditioned ensemble at each time step.

Bibliography

- S. I. Aanonsen, G. Nævdal, D. S. Oliver, A. C. Reynolds, and B. Vallès. Ensemble Kalman filter in reservoir engineering - a review. *SPE Journal*, 14(3):393–412, 2009.
- J. L. Anderson. A local least squares framework for ensemble filtering. *Monthly Weather Review*, 131(4):634–642, 2003a.
- T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley, 3 edition, 2003b.
- O. E. Barndorff-Nielsen, D. R. Cox, and C. Kluppelberg. *Complex Stochastic Systems*. Chapman and Hall, 2001.
- B. N. Bond and L. Daniel. Stabilizing schemes for piecewise-linear reduced order models via projection and weighting functions. In *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pages 860–867, San Jose, California, 2007.
- B. N. Bond and L. Daniel. Stable reduced models for nonlinear descriptor systems through piecewise-linear approximation and projection. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 28(10):1467–1480, 2009.
- G. Burgers, P. J. van Leeuwen, and G. Evensen. Analysis scheme in the Ensemble Kalman Filter. *Monthly weather review*, 126(6):1719–1724, 1998.
- M. A. Cardoso. *Development and Application of Reduced-Order Modeling Procedures for Reservoir Simulation*. PhD thesis, Stanford University, 2009.
- M. A. Cardoso and L. J. Durlofsky. Linearized reduced-order models for subsurface flow simulation. *Journal of Computational Physics*, 229(3):681–700, 2010.
- A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer, 2001.
- B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression (with discussion). *Annals of Statistics*, 32(2):407–499, 2004.
- G. Evensen. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99:10143–10162, 1994.

- G. Evensen. *Data assimilation. The Ensemble Kalman Filter*. Springer, 2007.
- I. K. Fodor. A survey of dimension reduction techniques. Technical report, Lawrence Livermore National Laboratory, 2002. Technical Report UCRL-ID-148494.
- R. Furrer and T. Bengtsson. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *Journal of Multivariate Analysis*, 98(2):227–255, 2007.
- T. Gneiting and A. E. Raftery. Stricly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.
- G. Golub and C. van Loan. *Matrix Computations*. Johns Hopkins University Press, 1996.
- D. Gratton and K. Willcox. Reduced-order, trajectory piecewise-linear models for nonlinear computational fluid dynamics. In *34th AIAA Fluid Dynamics Conference and Exhibit*, pages 2004–2329, Portland, Oregon, USA, 2004.
- T. M. Hamill, J. S. Whitaker, and L. A. Snyder. Distance-dependent filtering of background error covariance estimates in an Ensemble Kalman Filter. *Monthly Weather Review*, 129(11):2776–2790, 2001.
- T. Hastie, R. Tibshirani, and J. Freidman. *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. Springer, New York, 2 edition, 2009.
- J. He. Enhanced linearized reduced-order models for subsurface flow simulation. Master’s thesis, Stanford University, 2010.
- H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 1933.
- P. L. Houtekamer and H. L. Mitchell. A sequential Ensemble Kalman Filter for atmospheric data assimilation. *Monthly Weather Review*, 129:123–137, 2001.
- P. L. Houtekamer and H. L. Mitchell. Data assimilation using an Ensemble Kalman Filter technique. *Monthly Weather Review*, 126:796–811, 1998.

- P. L. Houtekamer and H. L. Mitchell. Reply. *Monthly Weather Review*, 127(6):1378–1379, 1999.
- P. L. Houtekamer, L. Lefaiivre, J. Derome, H. Ritchie, and H. L. Mitchell. A system simulation approach to ensemble prediction. *Monthly Weather Review*, 124:1225–1242, 1996.
- A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.
- H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London (Ser. A)*, 186:453–461, 1946.
- S. J. Julier and J. K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *Proceedings of the 12th International Symposium of Aerospace/Defense Sensing, Simulation and Controls*, pages 182–193, 1997.
- R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82(Series D):35–45, 1960.
- C. L. Kepenne and M. Rienecker. Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble kalman filter. *J. Marine. Sys.*, 40-41:363–380, 2003.
- H. R. Künsch. Recursive monte carlo filters: Algorithms and theoretical analysis. *The Annals of Statistics*, 33(5):1983–2021, 2005.
- O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, 2004.
- O. Leeuwenburgh, G. Evensen, and L. Bertino. The impact of ensemble filter definition on the assimilation of temperature profiles in the tropical pacific. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3291–3300, 2005.
- O. P. Lødøen and H. Omre. Scale-corrected ensemble kalman filtering applied to production history conditioning in reservoir evaluation. *SPE Journal*, 13(2):177–194, 2008.
- K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate analysis*. Academic Press, London, 1979.

- P. McCullagh and J. Nelder. *Generalized Linear Models*. Chapman and Hall, 1999.
- H. Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser. Dual state parameter estimation of hydrological models using Ensemble Kalman Filter. *Advances in Water Resources*, 28:135–147, 2005.
- I. Myrseth and H. Omre. Hierarchical ensemble kalman filter. *SPE Journal*, 15(2):569–580, 2010.
- G. Nævdal, H. Mennseth, and E. Vefring. Near-well reservoir moitoring though Ensemble Kalman Filter., 2003. Presented at the Society of Petroleum Engineers Annual Technical Conference and Exhibition, Denver, Colerado, USA, 5-8 October 2003, SPE 84372.
- H. Omre. Bayesian kriging-merging observations and qualified guesses in kriging. *Mathematical Geology*, 19:25–39, 1987.
- H. Omre and H. Tjelmeland. *Geostatistics Wollongong 96, Vol I*, chapter Petroleum Geostatistics, pages 41–52. Kluwer Acad. Publ., 1997.
- K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.
- M. J. Rewienski. *A Trajectory Piecewise-Linear Approach to Model Order Reduction of Nonlinear Dynamical Systems*. PhD thesis, Massachusetts Institute of Technology, 2003.
- W. Sacher and P. Bartello. Sampling errors in Ensemble Kalman filtering. Part I: Theory. *Monthly Weather Review*, 136(8):3035–3049, 2008.
- G. A. F. Seber and A. J. Lee. *Linear Regression Analysis*. Wiley, 2003.
- S. Strebelle. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34(1):1–21, 2002.
- S. J. Thomas, J. P. Hacker, and J. L. Anderson. A robust formulation of the ensemble Kalman filter. *Quarterly Journal of the Royal Meteorological Society*, 135:507–521, 2009.
- R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58(1):267–288, 1996.
- M. E. Tipping and C. M. Bishop. Probabilistic Principal Component Analysis. *Journal of the Royal Statistical Society, Series B*, 61(3):611–622, 1999.

-
- L. van der Maaten, E. Postma, and J. van den Herik. Dimensionality reduction: A comparative review. Technical report, Tilburg University, 2009. Technical Report, TiCC-TR 2009-005.
- P. J. van Leeuwen. Comments on "Data assimilation using an Ensemble Kalman Filter technique". *Monthly Weather Review*, 127:1374–1377, 1999.
- D. Vasilyev, M. Rewienski, and J. White. Macromodel generation for BioMEMS components using a stabilized balanced truncation plus trajectory piecewise-linear approach. *IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems*, 25(2):285–293, 2006.
- Y. J. Yang and K. Y. Shen. Nonlinear heat-transfer macromodeling for MEMS thermal devices. *Journal of Micromechanics and Microengineering*, 15(2):408–418, 2005.
- T. Zhang, P. Switzer, and A. G. Journel. Filter-based classification of training image patterns for spatial simulations. *Mathematical Geology*, 38(1): 63–80, 2006.

Papers

Paper I

Ensemble Kalman filtering with shrinkage regression techniques

Jon Sætrum and Henning Omre

Paper accepted for publication in *Computational Geosciences*,
DOI: 10.1007/s10596-010-9196-0.

The original publication is available at springerlink.com.

Chapter 2

Ensemble Kalman filtering with shrinkage regression techniques

Abstract. The classical Ensemble Kalman Filter (EnKF) is known to underestimate the prediction uncertainty. This can potentially lead to low forecast precision and an ensemble collapsing into a single realisation. In this paper we present alternative EnKF updating schemes based on shrinkage methods known from multivariate linear regression. These methods reduce the effects caused by collinear ensemble members, and have the same computational properties as the fastest EnKF algorithms previously suggested. In addition, the importance of model selection and validation for prediction purposes is investigated, and a model selection scheme based on Cross-Validation is introduced. The classical EnKF scheme is compared with the suggested procedures on two toy examples and one synthetic reservoir case study. Significant improvements are seen, both in terms of forecast precision and prediction uncertainty estimates.

2.1 Introduction

The Ensemble Kalman Filter (EnKF) is a Bayesian data assimilation method that in recent years has become popular when considering data assimilation for nonlinear spatiotemporal models (Aanonsen et al., 2009; Evensen, 2007). The EnKF is based on the classical Kalman Filter (KF) (Kalman, 1960). Assuming a Gaussian initial prior, with a linear and Gaussian forward and likelihood model, termed the Gauss-linear model, the KF provides an analytical solution for the posterior probability distribution.

The main assumption of the EnKF is that the unconditional distribution at each timestep approximately follows a Gaussian distribution with unknown mean and covariance. In the EnKF the idea is therefore to use an ensemble of realisations to estimate these two unknown statistics. Under the Gauss-linear model, the EnKF is then consistent with the KF as the ensemble size approaches infinity (Mardia et al., 1979).

From multivariate statistics (Anderson, 2003b), we know that the classical EnKF updating scheme (Evensen, 1994) can be formulated as a multivariate regression problem (Anderson, 2003a). The least squares solution for the matrix of regression coefficients is then given as the estimated Kalman gain, expressed in terms of the estimated unknown covariance matrices. However, it is not guaranteed that this estimator is good for prediction purposes. This is especially true for regression models where the predictor variables are collinear, or when the regression model is computed based on dependent realisations (Farrer and Glauber, 1967). As shown in Myrseth et al. (2009), the traditional EnKF updating scheme under the Gaussian approximation violates the assumption of independent realisations, because they are coupled through the estimated Kalman gain matrix. Hence, we should expect that improved estimators of the regression coefficient matrix can be found in these cases.

In multivariate linear regression, there exist several different approaches to avoid overfitting a regression model with collinear predictor variables, known as shrinkage methods. Some introduce regularisation terms, while others aim to estimate the regression coefficients based on data in a reduced order space. In this paper we will consider the following methods: Principal Component Regression (PCR) (Hotelling, 1933; Jolliffe, 2002) and Partial Least Squares Regression (PLSR) (Rosipal and Krämer, 2006; Wold, 1975).

Shrinkage estimators in multivariate regression require that a prior hyperparameter is selected. For methods based on PCR or PLSR, this means to select the dimension of a reduced order subspace. The most common method used to accomplish this task is to look at the total variance explained by

the selected Principal Components (PC) (Jolliffe, 2002). However, such an approach does not take into account the predictive capabilities of the estimated regression model, and can lead to situations where we either overfit, or underfit the model to the data (Seber and Lee, 2003; Cook, 2007). The method usually applied in shrinkage regression, is therefore to base the selection of the hyperparameter on Cross-Validation (CV) (Efron, 2004) in order to avoid this potential problem.

Shrinkage estimators which are based on dimension reduction techniques are for computational reasons the natural approach to consider when assimilating high dimensional data, such as time-lapse seismic. Shrinkage type estimators for the Kalman gain have therefore already been used in an EnKF setting (Skjervheim et al., 2007). It has also been noted (Evensen, 2007, Chapter 14) that the results obtained when using shrinkage estimators are indeed dependent on the selection of the prior hyperparameter used. However, the importance of the number of components retained in the reduced order space to avoid overfitting, has seemingly been overlooked in most of the EnKF literature. Moreover, because PCR is based on an unsupervised dimension reduction technique, all components that are important for predictive purposes can potentially be discarded, unless model validation is performed (Cook, 2007; Hadi and Ling, 1998; Jolliffe, 1982, 2002).

In this paper we have formulated the EnKF using shrinkage-based regression techniques. The suggested procedures have the same computational complexity and memory requirements as the fastest implementations of the EnKF (Evensen, 2003). We have further demonstrated the approach on the following case studies: A Gauss-linear model, a non-linear forward model with a linear Gaussian likelihood, and a synthetic reservoir example.

2.2 Notation and Model Formulation

Throughout this paper the notation $\mathbf{x} \in \mathbb{R}^{n_x \times 1}$ will be used to denote that \mathbf{x} is an n_x -dimensional column vector in the real space and \mathbf{x}^T will denote its transpose. Similarly, we will write $\mathbf{A} \in \mathbb{R}^{a \times b}$ to denote that \mathbf{A} is a matrix in the real space containing a rows and b columns. Note that the same notation will be used for both scalars and random variables.

Probability density functions (pdfs) will be denoted by $f(\mathbf{x})$, and the notation $\mathbf{x} \sim f(\mathbf{x})$, implies that the random vector \mathbf{x} follows the pdf $f(\mathbf{x})$. As a special case we will let a random vector \mathbf{x} , following the Gaussian distribution with mean vector $\boldsymbol{\mu}_x$ and covariance matrix $\boldsymbol{\Sigma}_x$, be denoted by $\mathbf{x} \sim \text{Gauss}_{n_x}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$. Conditional pdfs of \mathbf{x} given \mathbf{y} will further be denoted by $f(\mathbf{x}|\mathbf{y})$.

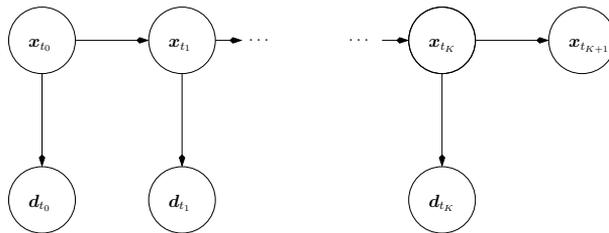


Figure 2.1: Stochastic Directed Acyclic Graph (DAG) of the model considered.

Consider the stochastic Directed Acyclic Graph (DAG) (Ripley, 1996) outlined in Figure 2.1. Here, $\mathbf{x}_{t_k} \in \mathbb{R}^{n_x \times 1}$ denotes the state of the unknown random vector of interest at timestep k and time t_k , and similarly $\mathbf{d}_{t_k} \in \mathbb{R}^{n_d \times 1}$ denotes the vector of observed data. For notational convenience we will from now on drop the subscript t_k , and simply write \mathbf{x}_k , \mathbf{d}_k . Also note that we will for simplicity refer to \mathbf{x} and \mathbf{d} as the state and observation vector respectively.

Through the Markov property of a stochastic DAG, we have conditional independence between \mathbf{x}_{k+1} and \mathbf{x}_l , $l = 0, \dots, k - 1$ given \mathbf{x}_k , meaning:

$$f(\mathbf{x}_{k+1} | \mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_0) = f(\mathbf{x}_{k+1} | \mathbf{x}_k), \quad k = 0, \dots, K.$$

In general assume that

$$\mathbf{x}_{k+1} = \boldsymbol{\omega}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{x}_k}), \quad k = 0, \dots, K, \tag{2.1}$$

where $\boldsymbol{\omega} : (\mathbb{R}^{n_x \times 1} \times \mathbb{R}^{n_x \times 1}) \rightarrow \mathbb{R}^{n_x \times 1}$ is a known, possibly non-linear forward function. Here $\boldsymbol{\epsilon}_{\mathbf{x}_k}$ represents random model errors or numerical errors in the forward model, assumed to follow a known probability distribution. Further assume that observed data \mathbf{d}_k is connected to \mathbf{x}_k by:

$$\mathbf{d}_k = \boldsymbol{\zeta}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{d}_k}), \quad k = 0, \dots, K, \tag{2.2}$$

where $\boldsymbol{\zeta} : (\mathbb{R}^{n_x \times 1} \times \mathbb{R}^{n_d \times 1}) \rightarrow \mathbb{R}^{n_d \times 1}$, is a known, possibly non-linear, function and $\boldsymbol{\epsilon}_{\mathbf{d}_k}$ represents the observation error, again assumed to follow a known pdf.

The aim in this model setting is to solve the spatiotemporal forecast problem of finding the unknown state vectors:

$$\mathbf{x}_k^c \sim f(\mathbf{x}_k | \mathbf{d}_{0:k}) \quad \text{and} \quad \mathbf{x}_{k+1}^u \sim f(\mathbf{x}_{k+1} | \mathbf{d}_{0:k}),$$

for $k = 1, \dots, K$, where we for notational convenience use the notation $\mathbf{d}_{0:l} = \{\mathbf{d}_0, \dots, \mathbf{d}_l\}$, $l > 0$. Bayesian inversion provides a sequential solution

to this problem. By defining a prior probability model for the state vector at the initial timestep, $f(\mathbf{x}_0)$, the state of the unknown vectors \mathbf{x}_k^c and \mathbf{x}_{k+1}^u , can be assessed by sampling from the respective posterior distributions. Through the use of Bayes rule and the Markov properties of a DAG, these pdfs are given as:

$$\begin{aligned} f(\mathbf{x}_k|\mathbf{d}_{0:k}) &\propto f(\mathbf{x}_k|\mathbf{d}_{0:(k-1)})f(\mathbf{d}_k|\mathbf{x}_k) \\ f(\mathbf{x}_{k+1}|\mathbf{d}_{0:k}) &= \int f(\mathbf{x}_{k+1}|\mathbf{x}_k)f(\mathbf{x}_k|\mathbf{d}_{0:k})d\mathbf{x}_k. \end{aligned} \quad (2.3)$$

Note that the conditional pdfs $f(\mathbf{x}_{k+1}|\mathbf{x}_k)$ and $f(\mathbf{d}_k|\mathbf{x}_k)$ are implicitly defined through $\boldsymbol{\omega}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{x}_k})$ and $\boldsymbol{\zeta}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{d}_k})$ defined in Eqs. (2.1) and (2.2) respectively.

Generally we only know the conditional distributions defined in Eq. (2.3) up to an unknown normalising constant. Computationally demanding techniques such as Markov chain Monte Carlo (McMC) or Rejection Sampling can therefore be used to assess the correct posterior distribution (Doucet et al., 2000). For applications such as petroleum reservoir evaluation, using these techniques are, however, computationally prohibitive because even a single evaluation of $\boldsymbol{\omega}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{x}_k})$, known as fluid flow simulation, can take several hours, or even days.

An approximative approach can be defined by assuming that \mathbf{x}_k^u and \mathbf{d}_k are from a family of probability density functions that ensures analytical tractability of $f(\mathbf{x}_k|\mathbf{d}_{0:k})$. As an example, assume that \mathbf{x}_k^u and \mathbf{d}_k are jointly Gaussian. Then the posterior distribution at timestep k will also be Gaussian. These model assumptions are equivalent to those made in the classical EnKF, which we will consider next.

2.3 Classical Ensemble Kalman Filter

In general assume that the output of the forward- and likelihood model are distributed as:

$$\begin{bmatrix} \mathbf{x}_k^u \\ \mathbf{d}_k \end{bmatrix} \sim \text{Gauss}_{n_y} \left(\begin{bmatrix} \boldsymbol{\mu}_{\mathbf{x}_k} \\ \boldsymbol{\mu}_{\mathbf{d}_k} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{x}_k} & \boldsymbol{\Sigma}_{\mathbf{x}_k, \mathbf{d}_k} \\ \boldsymbol{\Sigma}_{\mathbf{d}_k, \mathbf{x}_k} & \boldsymbol{\Sigma}_{\mathbf{d}_k} \end{bmatrix} \right), \quad (2.4)$$

where $n_y = n_x + n_d$. For notational convenience, we will from now on omit the subscript k , as the main focus is a single timestep.

Under the Gaussian assumption above, the posterior pdf $f(\mathbf{x}_k^c)$ will be Gaussian (Mardia et al., 1979) with analytically obtainable mean:

$$\boldsymbol{\mu}_{\mathbf{x}|\mathbf{d}} = \boldsymbol{\mu}_{\mathbf{x}} + \boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{d}} \boldsymbol{\Sigma}_{\mathbf{d}}^{-1} (\mathbf{d} - \boldsymbol{\mu}_{\mathbf{d}}), \quad (2.5)$$

and covariance matrix:

$$\Sigma_{\mathbf{x}|\mathbf{d}} = \Sigma_{\mathbf{x}} - \Sigma_{\mathbf{x},\mathbf{d}}\Sigma_{\mathbf{d}}^{-1}\Sigma_{\mathbf{x},\mathbf{d}}^T. \quad (2.6)$$

Let $\mathbf{y}^{(i)} = [\mathbf{x}^{u(i)T}, \mathbf{d}^{(i)T}]^T \in \mathbb{R}^{n_y \times 1}$ be a realisation from the Gaussian distribution defined in Eq. (2.4). Then,

$$\mathbf{x}^{c(i)} = \mathbf{x}^{u(i)} + \Sigma_{\mathbf{x},\mathbf{d}}\Sigma_{\mathbf{d}}^{-1}(\mathbf{d} - \mathbf{d}^{(i)}), \quad (2.7)$$

is a realisation from the Gaussian posterior distribution with mean and covariance given in Eqs. (2.5) and (2.6). This can easily be shown by recognising that the Gaussian distribution is closed under linear operations and by computing the mean and covariance of $\mathbf{x}^{c(i)}$, because the Gaussian distribution is completely specified through the first and second moments (Mardia et al., 1979).

Eq. (2.7) involves two model parameters, namely $\Sigma_{\mathbf{x},\mathbf{d}}$ and $\Sigma_{\mathbf{d}}$ forming the Kalman gain matrix,

$$\mathbf{K} = \Sigma_{\mathbf{x},\mathbf{d}}\Sigma_{\mathbf{d}}^{-1} \in \mathbb{R}^{n_x \times n_d}.$$

In a general setting such as the one considered here, these model parameters are unknown. The EnKF solution to this problem is therefore to use an ensemble of realisations to obtain empirical estimates of the unknown covariance matrices. Hence, let \mathbf{X} and \mathbf{D} be the centred state- and observation vector ensembles respectively. That is, $\mathbf{X} = [\mathbf{x}^{(1)} - \hat{\boldsymbol{\mu}}_{\mathbf{x}}, \dots, \mathbf{x}^{(n_e)} - \hat{\boldsymbol{\mu}}_{\mathbf{x}}] \in \mathbb{R}^{n_x \times n_e}$, where $\hat{\boldsymbol{\mu}}_{\mathbf{x}}$ denotes the classical estimator of the mean value with a corresponding expression for $\mathbf{D} \in \mathbb{R}^{n_d \times n_e}$. Consistent estimators for the unknown covariance matrices are then given as

$$\hat{\Sigma}_{\mathbf{d}} = \frac{1}{n_e} \mathbf{D}\mathbf{D}^T$$

and

$$\hat{\Sigma}_{\mathbf{x},\mathbf{d}} = \frac{1}{n_e} \mathbf{X}\mathbf{D}^T.$$

Replacing the unknown covariance matrices in Eq. (2.7) with any consistent empirical estimates, then ensures that $\mathbf{x}^{c(i)}$ is a sample from the Gaussian posterior distribution above as $n_e \rightarrow \infty$ (Mardia et al., 1979). Note that throughout this paper we will for notational convenience let all ensemble matrices be centred, unless otherwise stated.

Central in the EnKF updating scheme is the estimated Kalman gain matrix:

$$\hat{\mathbf{K}} = \hat{\Sigma}_{\mathbf{x},\mathbf{d}}\hat{\Sigma}_{\mathbf{d}}^{-1}. \quad (2.8)$$

From multivariate statistical theory (Mardia et al., 1979), we know that this is equivalent to the least squares estimate of the matrix of regression coefficients in a multivariate linear regression setting (Seber and Lee, 2003):

$$\hat{\mathbf{K}} = \arg \min_{\mathbf{K}} \text{tr} \{ (\mathbf{X} - \mathbf{K}\mathbf{D})(\mathbf{X} - \mathbf{K}\mathbf{D})^T \},$$

where $\text{tr}(\cdot)$ denotes the trace operator. The analytical rank of the Kalman gain is known to be equal to $\min\{n_x, n_d, n_e - 1\}$. Hence, $\hat{\mathbf{K}}$ will be rank deficient if the number of ensemble members is smaller or equal to $\min\{n_x, n_d\}$. This will have direct consequences for the computation of $\hat{\Sigma}_{\mathbf{d}}^{-1}$, because this matrix will be positive semi-definite.

Although, $\hat{\Sigma}_{\mathbf{d}}$ can be ensured to have full rank either through regularisation or Monte Carlo simulation (Myrseth et al., 2009), $\hat{\Sigma}_{\mathbf{x}, \mathbf{d}}$ may still suffer from rank deficiency. Moreover, both matrices are likely to suffer from estimation uncertainty (Furrer and Bengtsson, 2007; Houtekamer and Mitchell, 1998, 1999; van Leeuwen, 1999; Sacher and Bartello, 2008) resulting from a limited number of ensemble members. Several different approaches have been suggested in the EnKF literature in order to handle these problems such as localisation and inflation (see Myrseth and Omre (2010) or Aanonsen et al. (2009) and references therein). Most of these methods, however, focus on improving the estimates of the unknown covariance matrices, and not the Kalman gain itself. One explanation for this might be that in situations with a linear likelihood model, having additive Gaussian noise: $\mathbf{d} = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon}_{\mathbf{d}}$, the Kalman gain matrix can be written as:

$$\mathbf{K} = \Sigma_{\mathbf{x}} \mathbf{H}^T (\mathbf{H} \Sigma_{\mathbf{x}} \mathbf{H}^T + \Sigma_{\boldsymbol{\epsilon}_{\mathbf{d}}})^{-1}, \quad (2.9)$$

where $\Sigma_{\boldsymbol{\epsilon}_{\mathbf{d}}} \in \mathbb{R}^{n_d \times n_d}$ is the covariance matrix of the observations errors, assumed to be known. Hence, the covariance matrix of the state vector is the only unknown parameter that requires estimation. Note that this corresponds to the classical estimator used for linear Gaussian likelihood models in an EnKF setting (Evensen, 1994). Further note that for non-linear likelihood models with an additive noise term, the state vector can always be augmented by $\mathbf{y} = [\mathbf{x}^T, \mathbf{d}^T]^T$, so that the EnKF updating scheme can be written in a similar manner as described in Eq. (2.9). Here \mathbf{d} corresponds to the deterministic part of the likelihood model with additive noise.

In multivariate linear regression there exists several different techniques that improve the estimated matrix of regression coefficients in the presence of collinear data. These are referred to as shrinkage regression methods (Hastie et al., 2009), and we will use these approaches in an EnKF setting to obtain alternative estimates of \mathbf{K} .

2.4 Shrinkage Regression Methods

Consider the linear regression problem:

$$[\mathbf{x}|\mathbf{d}] = \mathbf{K}\mathbf{d} + \epsilon_{\mathbf{x}|\mathbf{d}},$$

where $\epsilon_{\mathbf{x}|\mathbf{d}}$ corresponds to the model error in the regression model. Here, the goal is to estimate the unknown matrix of multiple linear regression coefficients \mathbf{K} , based on a set of observed data \mathbf{X} and \mathbf{D} . An optimal estimate should then be selected such that a regression model with good prediction capabilities is provided.

Consider the well known problem of estimating a function based on the observed data. Fitting the model perfectly to the available data increases the model complexity and lowers the prediction bias with respect to the training data. However, when the model is used in a predictive setting when new data are available, this can lead to poor results with a high variability. On the other hand, underfitting the model by selecting a too simplistic model will lower the variability, but increase the bias. Hence, the optimal model should be selected such that both bias and variability are kept on a tolerable level as illustrated in Figure 2.2. As we can see from this figure, the prediction error for a test data set will tend to increase in situations where the model is overfitted to the training data.

Shrinkage regression methods focus on the model fitting problem by sacrificing the unbiasedness of the classical least squares estimator. In this paper we will consider the following methods: Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR).

2.4.1 Principal Component Regression

PCR is based on principal component analysis, which aims at explaining the structure of the data ensemble through a small number of vectors, termed Principal Components (PC):

$$\mathbf{z}_1 = (\mathbf{u}_1^T \mathbf{D})^T, \dots, \mathbf{z}_p = (\mathbf{u}_p^T \mathbf{D})^T \in \mathbb{R}^{n_e \times 1}.$$

The sample PC are selected based on the following criteria:

$$\mathbf{z}_i = (\mathbf{u}_i^T \mathbf{D})^T \leftarrow \begin{cases} \max_{\mathbf{u}_i} \{ \mathbf{u}_i^T \hat{\Sigma}_d \mathbf{u}_i \} \\ \|\mathbf{u}_i\|_2 = 1 \\ \mathbf{z}_i^T \mathbf{z}_j = 0, \text{ for all } j < i, i = 1, \dots, p, \end{cases}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. It can be shown (Anderson, 2003b) that the i th sample PC direction \mathbf{u}_i is given as the i th eigenvector of the

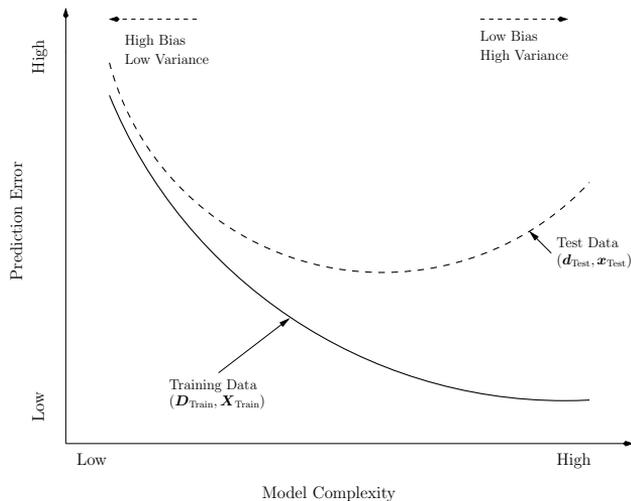


Figure 2.2: Tradeoff between Bias and Variance: Overfitting the data by increasing the model complexity tend to increase the variance in model predictions. Based on a similar figure found in Hastie et al. (2009).

covariance matrix $\hat{\Sigma}_{\mathbf{d}}$. Further, it can be shown that the variance explained by the i th PC is given by the i th eigenvalue λ_i of $\hat{\Sigma}_{\mathbf{d}}$, where $\lambda_1 > \lambda_2 > \dots > \lambda_p$. Under the assumption that the p sample PC, $\mathbf{Z} = [z_1, \dots, z_p]^T \in \mathbb{R}^{p \times n_e}$, sufficiently represents \mathbf{D} , the matrix of regression coefficients can then be estimated based on \mathbf{Z} , which gives:

$$\hat{\mathbf{K}}_{\text{PCR}} = \mathbf{X} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1}.$$

As shown in Appendix 2-A, PCR can be efficiently implemented using Singular Value Decomposition (SVD).

Because the variance of each sample PC is given as the corresponding eigenvalue of the sample covariance matrix, the criterion often used for selecting p in a PCR setting is to look at the total variance explained by the first p components given as:

$$\sum_{i=1}^p \lambda_i / \sum_{i=1}^r \lambda_i,$$

where $r = \text{rank}(\hat{\Sigma}_{\mathbf{d}})$. However, as noted above, this criterion does not take into account the predictive capabilities of the various PC, and we can end up either overfitting or underfitting the regression model.

2.4.2 Partial Least Squares Regression

Consider the DAG shown in Figure 2.3, which gives a graphical representation of the PLSR model assumptions. Similar to PCR, PLSR aims to represent \mathbf{D} in a reduced order space, before fitting the regression model. The underlying assumption is that $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_p] \in \mathbb{R}^{n_e \times p}$ and $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_p] \in \mathbb{R}^{n_e \times p}$ represent the information in \mathbf{X} and \mathbf{D} respectively. It is further assumed that $(\mathbf{w}_1, \mathbf{t}_1)$ captures more information than $(\mathbf{w}_2, \mathbf{t}_2)$, and so forth, and that the pairs $(\mathbf{w}_i, \mathbf{t}_i)$, $i = 1, \dots, p$ also explain the correlation between \mathbf{X} and \mathbf{D} (Zeng et al., 2007). Hence, while PCR only does a dimension reduction on \mathbf{D} independent of the values in \mathbf{X} , the classical PLSR algorithm selects the latent variables $\mathbf{t}_i = \mathbf{D}^T \boldsymbol{\psi}_i \in \mathbb{R}^{n_e \times 1}$ with the largest dependency on $\mathbf{w}_i = \mathbf{X}^T \mathbf{v}_i \in \mathbb{R}^{n_e \times 1}$, that is for $i = 1, \dots, p$:

$$\begin{bmatrix} \mathbf{t}_i = \mathbf{D}^T \boldsymbol{\psi}_i \\ \mathbf{w}_i = \mathbf{X}^T \mathbf{v}_i \end{bmatrix} \leftarrow \begin{cases} \max_{\boldsymbol{\psi}_i, \mathbf{v}_i} \{ \mathbf{v}_i^T \hat{\boldsymbol{\Sigma}}_{\mathbf{x}, \mathbf{d}} \boldsymbol{\psi}_i \} \\ \|\boldsymbol{\psi}_i\|_2 = 1, \|\mathbf{v}_i\|_2 = 1 \\ \mathbf{t}_i^T \mathbf{t}_j = 0, \text{ for all } j < i. \end{cases}$$

This problem can be solved sequentially using the Non-linear Iterative Partial Least Squares (NiPALS) procedure (Rosipal and Krämer, 2006), or simultaneously by computing the SVD of the estimated covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}, \mathbf{d}}$ (Barker and Rayens, 2003). The matrices of latent variables \mathbf{T} and \mathbf{W} can then be obtained using the matrices $\boldsymbol{\Psi} \in \mathbb{R}^{n_d \times p}$ and $\boldsymbol{\Upsilon} \in \mathbb{R}^{n_x \times p}$, corresponding to the first p left and right singular vectors of $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}, \mathbf{d}}$.

Similarly to PCR, we assume that the latent variables \mathbf{T} are good predictors for \mathbf{D} . In addition we assume that $\mathbf{W} = \boldsymbol{\beta} \mathbf{T} + \boldsymbol{\epsilon}_{\mathbf{T}}$ where $\boldsymbol{\beta}$ is a diagonal matrix and $\boldsymbol{\epsilon}_{\mathbf{T}}$ is a residual term. That is, there is a linear relationship between each element of \mathbf{w}_i and \mathbf{t}_i , $i = 1, \dots, p$ (Kaspar and Ray, 1993). As shown in Rosipal and Krämer (Rosipal and Krämer, 2006), the PLSR estimate of the Kalman gain \mathbf{K}_{PLSR} is then given as:

$$\hat{\mathbf{K}}_{\text{PLSR}} = \mathbf{X} \mathbf{T} (\mathbf{W}^T \mathbf{D}^T \mathbf{D} \mathbf{T})^{-1} \mathbf{W}^T \mathbf{D}^T. \quad (2.10)$$

2.4.3 Comments

The PCR and PLSR methods have the same computational complexity and memory requirements as the fastest implementation of the EnKF as explained in Appendix 2-B. Note also that the PCR and PLSR methods are in general not scale invariant. This implies that the data ensemble matrix should be standardised before the dimension reduction, when the data are collected on different scales (Mardia et al., 1979). Finally it is interesting to note that the shrinkage estimators of the Kalman gain, previously applied

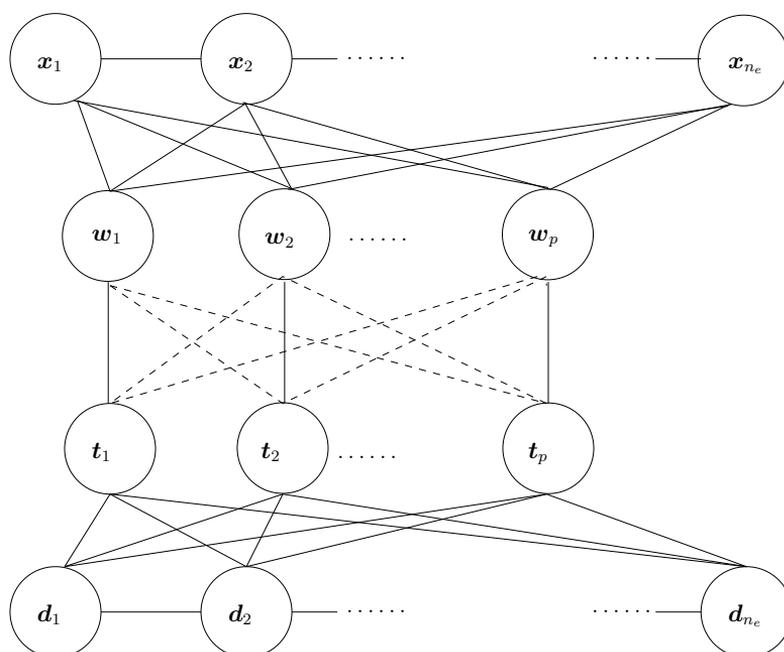


Figure 2.3: Graphical presentation of the model assumptions made in PLSR, inspired by a similar figure found in Zeng et al. (2007). Connected lines implies that there is a direct connection between the variables, while dashed lines implies an implicit connection.

in the EnKF literature (Evensen, 2007), can be thought of as hybrid methods combining Tikhonov regularisation (Tikhonov and Arsenin, 1977) and PCR. The regularisation term in these approaches is given by the covariance matrix of the observation errors, or a corresponding low rank representation.

2.4.4 Cross-Validation

The estimated matrix of regression coefficients will for all three methods be dependent on the choice of some hyperparameter θ , denoted $\mathbf{K}(\theta)$. For both PCR and PLSR, this is the dimension of the reduced order space p . As discussed above, there is a tradeoff between how well a model is fitted to the training data, and how well it is suited for prediction purposes. To determine this tradeoff, CV is often used. The idea used in CV is to randomly split the ensembles into one set used for model fitting: $\mathbf{Y}_{\text{Train}} = [\mathbf{X}_{\text{Train}}^T, \mathbf{D}_{\text{Train}}^T]^T \in \mathbb{R}^{n_y \times n_{\text{Train}}}$, and one set used for testing the prediction capabilities: $\mathbf{Y}_{\text{Test}} = [\mathbf{X}_{\text{Test}}^T, \mathbf{D}_{\text{Test}}^T]^T \in \mathbb{R}^{n_y \times n_{\text{Test}}}$. Here n_{Train} and n_{Test} are used to denote the number of members in the training and test ensembles respectively. If $\mathbf{Y}_{\text{Train}}$ consists of all realisations except $\mathbf{y}^{(i)}$, this is referred to as leave-one-out CV, and a brute implementation will increase the computational time with $\mathcal{O}(n_e^2)$ to the regression method used. Splitting the data into m sized portions randomly, and sequentially using m data points for testing and the remaining $m - 1$ part for model fitting purposes, is referred to as m -fold CV. This will increase the computational effort with $\mathcal{O}(n_e)$. Typical values used for m in m -fold CV are 5 or 10 (Hastie et al., 2009). The optimal regression model can then be selected by minimising the Predictive Error Sum of Squares (PRESS) statistic, defined as:

$$\text{PRESS}(\theta) = \sum_{i=1}^{n_{\text{Test}}} \|\mathbf{x}_{\text{Test}}^{(i)} - \hat{\mathbf{K}}(\theta)\mathbf{d}_{\text{Test}}^{(i)}\|_2^2. \quad (2.11)$$

To avoid overfitting of the regression model to the data when using the PCR and PLSR technique, a penalised version of the original PRESS statistic is given as:

$$\text{PRESS}_{\text{pen}}(p) = \sum_{i=1}^{n_{\text{Test}}} \frac{\|\mathbf{x}_{\text{Test}}^{(i)} - \hat{\mathbf{K}}(p)\mathbf{d}_{\text{Test}}^{(i)}\|_2^2}{(\min\{n_e, n_d + 1\} - p)^2}.$$

Note that $\text{PRESS}_{\text{pen}}$ is similar to the generalised PRESS statistic, which is an estimate of the leave one out CV prediction error. This follows because the effective degrees of freedom, defined as $\text{tr}(\hat{\mathbf{K}}(p)\mathbf{D})$ (Hastie et al., 2009), is equal to p for both PCR and PLSR.

Finally it should be noted that applying the CV scheme does not necessarily lead to an increase in the computational demands. This follows because CV can be equally performed in the reduced order space (Hastie and Tibshirani, 2004). A more thorough discussion regarding the computational properties of the CV scheme can be found in Appendix 2-B.

2.5 Empirical Study

We consider two test cases similar to the ones used in Myrseth and Omre (2010), where the unknown vector of interest, \mathbf{x}_k , $k = 0, \dots, 10$, with $n_x = 100$. Here $x_{j,k}$ denotes the variable of interest at timestep k and location j , with $j = 1, \dots, 100$. Observations are assumed to be made at timesteps $0, \dots, 9$, and the objective of this study is to assimilate the observed data, and predict at timestep 10.

The first test case, referred to as the linear case, is defined as a Gaussian prior at timestep zero, a linear forward function, and a linear Gaussian likelihood model:

$$\begin{aligned}\mathbf{x}_0 &\sim \text{Gauss}_{n_x}(\mathbf{0}, \Sigma_{\mathbf{x}_0}) \\ \mathbf{x}_k &= \mathbf{A}_k \mathbf{x}_{k-1} \\ \mathbf{d}_k &= \mathbf{H} \mathbf{x}_k + \boldsymbol{\epsilon}_{\mathbf{d}_k},\end{aligned}$$

where $\mathbf{0}$ is the null-vector of proper dimensions. Here $\Sigma_{\mathbf{x}_0}$ is constructed based on an exponential covariance function,

$$\text{Cov}(x_{i0}, x_{j0}) = 20 \exp \left\{ -\frac{3}{20} |i - j| \right\}, \quad (2.12)$$

the forward model is defined by the sparse matrix \mathbf{A}_k , where the elements $\mathbf{A}_{k_{l,m}}$ for $\{5(k-1) < l, m \leq 5(k+1)\}$, are displayed in Figure 2.4 with $\mathbf{A}_{k_{l,m}} = \delta_{l,m}$ otherwise, and $\boldsymbol{\epsilon}_{\mathbf{d}_k} \sim \text{Gauss}_{n_d}(\mathbf{0}, \mathbf{I})$. \mathbf{H} is a sparse matrix with elements equal to one at the grid locations displayed in Figure 2.5. Here we also see that observation $d_i = \sum_{l=-1}^1 x_{j+l}$, $i = 1, \dots, n_d$, at 13 different grid locations j .

The second test case, referred to as the non-linear case, considers the same prior and likelihood as in the linear case defined above. Here, however, the forward model is defined as the non-linear function:

$$\mathbf{x}_k = c \mathbf{A}_k (\mathbf{x}_{k-1} + \arctan(\mathbf{x}_{k-1})),$$

where $c = 0.8$ is a scaling factor which ensures an approximate alignment of the variances for the non-linear and linear case (Myrseth and Omre, 2010),

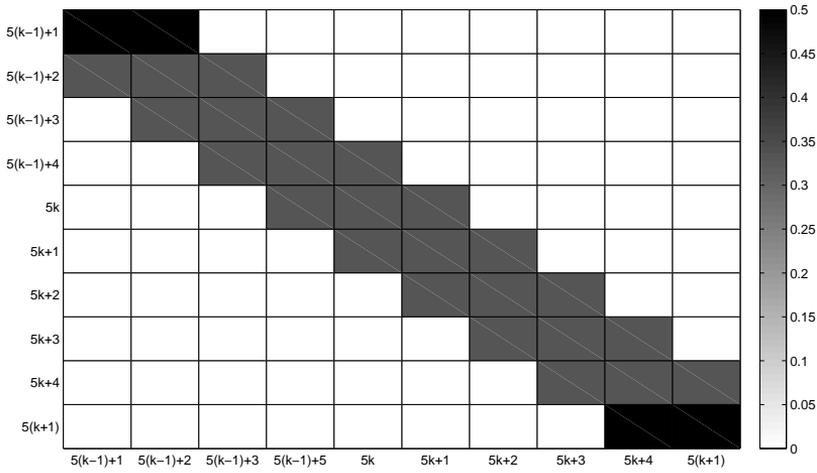


Figure 2.4: Graphical presentation of the elements of the matrix $\mathbf{A}_{k_l, m}$, for $5(k - 1) < l, m \leq 5(k + 1)$, for the forward model in the empirical case study. At all other grid locations $\mathbf{A}_{k_l, m} = \delta_{l, m}$

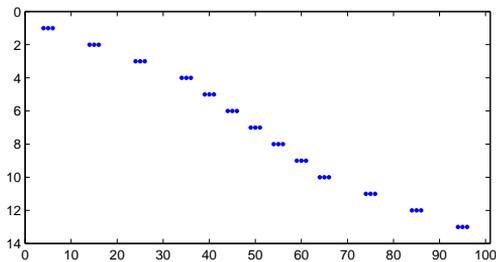


Figure 2.5: Graphical presentation of the non-zero elements of the matrix \mathbf{H} , for the likelihood model in the empirical case study.

and the functional $\arctan(\cdot)$ acts on the argument element-by-element. Note that because of the construction of \mathbf{A}_k , the state vector contains both dynamic (elements 1 through 55) and static variables (elements 56 through 100), similar to what we have in a reservoir setting.

2.5.1 Results with Discussion

We have considered the following EnKF updating schemes:

- Classical EnKF: Estimated Kalman gain matrix computed based on Eq. (2.9) using the correct Σ_{ϵ_d} .
- PCR-0.99-EnKF: p selected based on the estimated proportion of explained variance (99 percent)
- PCR-CV-EnKF: p selected based on 10-fold CV and $\text{PRESS}_{\text{pen}}$.
- PLSR-CV-EnKF: p selected based on 10-fold CV and $\text{PRESS}_{\text{pen}}$.

The same initial ensemble and random numbers were used for all four updating schemes, with two different ensemble sizes: $n_e = 100$ and $n_e = 20$. We also applied the CV criterion suggested by Hastie et al. (2009), described in Appendix 2-B, to further avoid the problem of overfitting.

2.5.2 Linear Case

For the linear case, the prediction mean, $E[\mathbf{x}_{10}^u]$, and 95% prediction interval are analytically obtainable using the KF recursions, and the results are displayed in Figure 2.6a. The results obtained when applying the four different schemes outlined above, are displayed in Figures. 2.6b through i.

As we can see from Figure 2.6b, the result obtained using the classical EnKF updating scheme is relatively reliable for $n_e = 100$. The classical EnKF solution matches the KF solution fairly well, although we do not see the same smooth behaviour in the estimated ensemble mean as seen in the KF solution. Moreover, there is a tendency of underestimating the prediction uncertainty. For $n_e = 20$, however, both the estimated mean and the prediction interval deviates dramatically from the KF solution. This is particularly true for grid nodes 1 to 35 and 65 to 100, where we have less observed data.

The estimated posterior mean obtained using the PCR-0.99-EnKF updating scheme matches the KF solution fairly well for $n_e = 100$. However, the estimated prediction interval is severely underestimated, and we are not able to capture the reference solution within the prediction interval. For

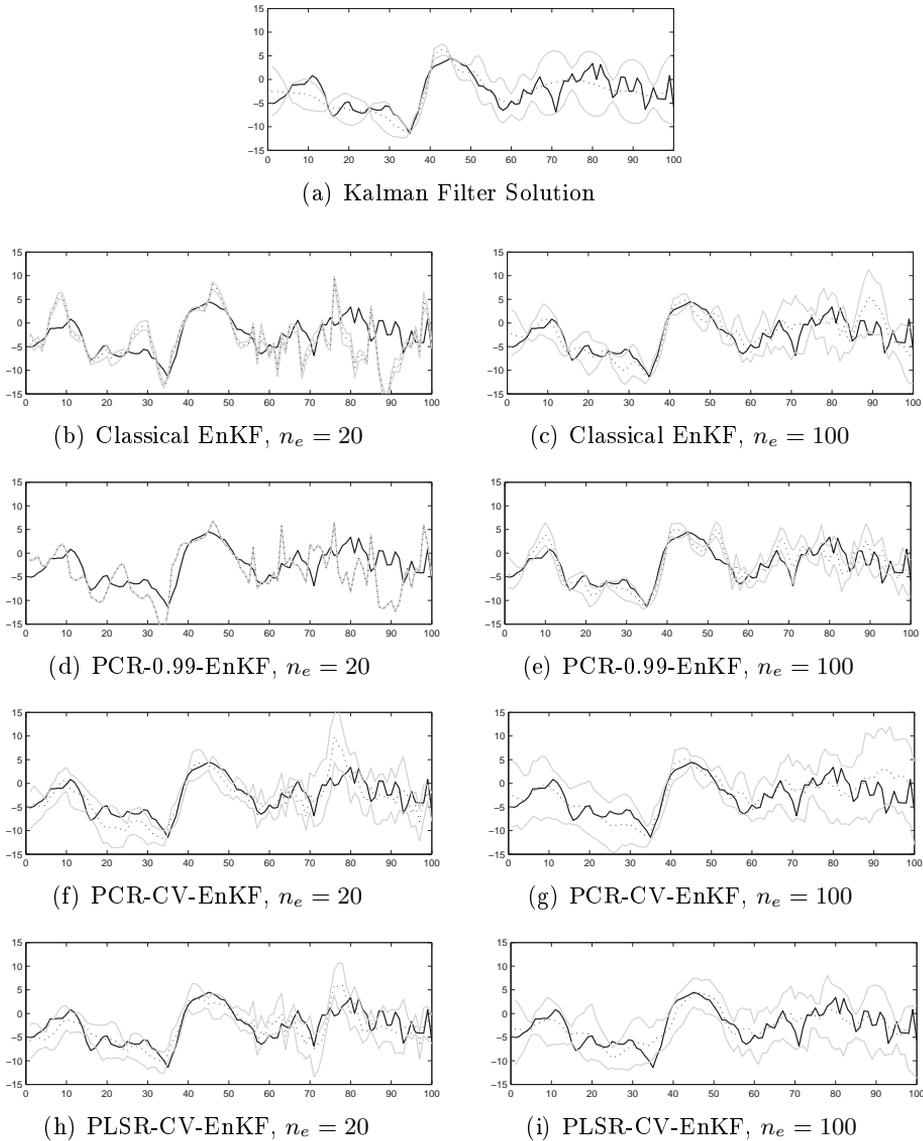


Figure 2.6: Results obtained when running four different EnKF updating schemes on the linear case with two different ensemble sizes. The figure displays the reference $\mathbf{x}_{10}^{\text{True}}$ (solid), the ensemble mean (dotted, black) and the estimated 95% confidence bounds of the prediction interval (solid, gray).

the smallest ensemble size, $n_e = 20$ the estimated posterior mean is highly variable and there is little uncertainty in the predictions as the updated ensemble has almost collapsed completely.

The obtained posterior mean using the PCR-CV-EnKF updating scheme appears to be a reliable estimate of the KF posterior mean for $n_e = 100$. Similarly we see that the estimated prediction interval matches the KF solution reasonably well between grid nodes 35 and 65. In the area where the data are observed less frequently, however, the PCR-CV-EnKF updating scheme is not able to reduce the uncertainty at the observation sites. For $n_e = 20$, the posterior mean deviates more from the KF solution and is less smooth than for $n_e = 100$. The reference solution is reasonably well captured within the prediction interval, although the prediction interval is slightly underestimated relative to the KF solution. However, the results still appear to be relatively reliable.

The PLSR-CV-EnKF scheme is able to get a good representation of the KF solution for $n_e = 100$, both in terms of the estimated posterior mean and prediction interval. By decreasing the ensemble size to $n_e = 20$, the scheme is still able to obtain reasonable results, even though the prediction uncertainty is underestimated.

High variability in the estimated posterior mean and underestimation of the prediction interval are problems occurring in all four schemes when the ensemble size is only 20. The match with the KF solution is, however, reasonably good when using the PCR-CV-EnKF and PLSR-CV-EnKF schemes. For the two other schemes, the estimated posterior mean has a high variability and we see a dramatic underestimation of the prediction uncertainty.

Increasing the ensemble size to $n_e = 100$, improves the estimates of the posterior mean and prediction intervals for all four updating schemes. However, the classical EnKF and PCR-0.99-EnKF schemes tend to underestimate the prediction uncertainty. The best overall match with the KF solution is obtained using the PLSR-CV-EnKF updating scheme.

To further quantify the performance of the four updating schemes, the algorithms are rerun 100 times using different initial ensembles. Here we consider the Root Mean Squared Error (RMSE), $\hat{\boldsymbol{\mu}}_{\mathbf{x}|\mathbf{d}} - \boldsymbol{\mu}_{\mathbf{x}|\mathbf{d}}$, where $\hat{\boldsymbol{\mu}}_{\mathbf{x}|\mathbf{d}}$ is the average of the updated ensemble after the final updating step and $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{d}}$ is the true posterior mean, and the percentage of the reference solution covered by the estimated prediction intervals. We calculate the coverage by estimating the 95%-prediction interval at each grid location and see if the reference solution is captured. The results are shown in Table 2.1. Here we also include the estimated RMSE of the posterior mean and coverage when the initial ensemble is run through the forward model without conditioning

n_e	Scheme	RMSE $\hat{\mu}_{\mathbf{x} \mathbf{d}}$	Coverage %
20	No Updating	9.96	89
20	Classical EnKF	9.96	24
20	PCR-0.99-EnKF	11.4	1
20	PCR-CV-EnKF	7.26	58
20	PLSR-CV-EnKF	6.53	59
100	No Updating	4.39	96
100	Classical EnKF	1.75	84
100	PCR-0.99-EnKF	2.42	53
100	PCR-CV-EnKF	1.74	92
100	PLSR-CV-EnKF	1.49	96

Table 2.1: Estimated Root Mean Squared Error (RMSE) of the posterior mean, $\hat{\mu}_{\mathbf{x}|\mathbf{d}}$, and coverage of the reference solution in the estimated 95% prediction intervals for the linear case based on 100 different initial ensembles.

on the observed data, referred to as the No Updating scheme. As we can see from this table, the estimated RMSE of the posterior mean decrease significantly for all four EnKF schemes when $n_e = 100$. Compared to the estimated RMSE of the initial ensemble, the PLSR-CV-EnKF scheme shows the largest improvement with a 66 percent decrease. For $n_e = 20$, however, the classical EnKF updating scheme is not able to improve the RMSE, while the PCR-0.99-EnKF scheme leads to an increase in RMSE compared to the initial ensemble. Again PLSR-CV-EnKF has the smallest RMSE with a reduction of 34 percent from the initial ensemble.

The coverage of the respective estimated 95% prediction intervals, is seen to be significantly underestimated for both the classical EnKF, and PCR-0.99-EnKF schemes. This is especially true for $n_e = 20$, where the prediction interval based on the PCR-0.99-EnKF solution only covers one percent of the reference solution. The PCR-CV-EnKF and PLSR-CV-EnKF updating schemes have similar and more reliable estimates of the prediction intervals, with the latter being slightly better than the former.

2.5.3 Non-Linear Case

For the non-linear case, analytical tractability is lost, we therefore use the results obtained with the classical EnKF with $n_e = 100\,000$, displayed in Figure 2.7a, for comparison. The four different EnKF updating schemes

n_e	Scheme	RMSE	Coverage
		$\hat{\boldsymbol{\mu}}_{\mathbf{x} \mathbf{d}}$	%
20	No Updating	10.2	89
20	Classical EnKF	8.07	21
20	PCR-0.99-EnKF	8.95	1
20	PCR-CV-EnKF	6.63	65
20	PLSR-CV-EnKF	5.88	65
100	No Updating	4.48	96
100	Classical EnKF	1.52	79
100	PCR-0.99-EnKF	2.19	56
100	PCR-CV-EnKF	1.76	92
100	PLSR-CV-EnKF	1.25	93

Table 2.2: Estimated Root Mean Squared Error (RMSE) of the posterior mean, $\hat{\boldsymbol{\mu}}_{\mathbf{x}|\mathbf{d}}$, and coverage of the reference solution in the estimated 95% prediction intervals for the non-linear case based on 100 different initial ensembles.

outlined above, provide the results shown in Figures 2.7b through i. Similar to the linear case, we see that the estimated posterior mean for the classical EnKF scheme is highly fluctuating with a severely underestimated prediction interval when $n_e = 20$. The most severe problems are in the PCR-0.99-EnKF solution, with a completely collapsed ensemble for $n_e = 20$. For both the PCR-CV-EnKF and PLSR-CV-EnKF schemes we observe relatively reasonable results when $n_e = 20$ and 100. The PCR-CV-EnKF scheme, however, tends to overestimate the prediction uncertainty at data locations for $n_e = 100$.

To quantify the performance the four schemes are rerun using 100 different initial ensembles. The results are shown in Table 2.2. Similar to the linear case, the PLSR-CV-EnKF scheme show the best performance in terms of estimated RMSE of the posterior mean and coverage for both ensemble sizes. Both the classical EnKF, and the PCR-0.99-EnKF schemes fail to cover the reference solution within their respective 95% prediction intervals for $n_e = 20$.

2.5.4 Ensemble Size

The number of ensemble members needed to achieve at least a 92% coverage of the reference solution in the respective estimated 95% prediction intervals are shown in Table 2.3. As we can see from this table, the classical

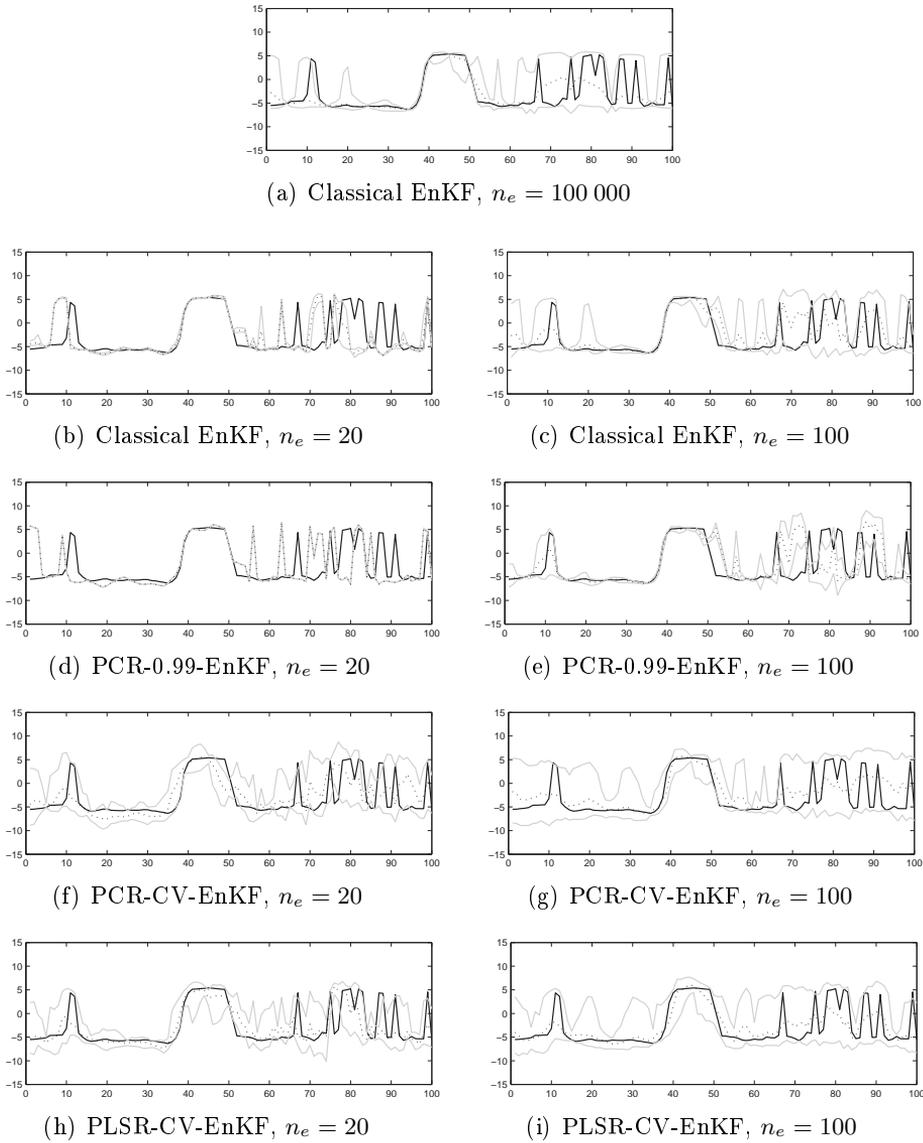


Figure 2.7: Results obtained when running four different EnKF updating schemes on the non-linear case with two different ensemble sizes. The figure displays the reference $\mathbf{x}_{10}^{\text{True}}$ (solid), the ensemble mean (dotted, black) and the estimated 95% confidence bounds of the prediction interval (solid, gray).

Scheme	Linear Case	Non-Linear Case
Classical EnKF	300	1 000
PCR-0.99-EnKF	550	1 250
PCR-CV-EnKF	100	100
PLSR-CV-EnKF	60	90

Table 2.3: Ensemble size n_e required to achieve at least a 92 percent coverage of the reference state vector within the estimated 95 percent prediction interval for the four EnKF updating schemes considered for the empirical case study. Estimates are based on 100 reruns using different initial ensembles.

EnKF updating scheme requires 5 times as many ensemble members as the PLSR-CV-EnKF scheme for the linear case, and 11 times as many ensemble members for the non-linear case.

Note also that the classical EnKF requires three times as many ensemble members in the non-linear case, compared to the linear case. This effect is believed to be caused by the non-linear forward model, because the Gaussian assumption made in Eq. 2.4 in this case is violated.

2.5.5 Summary

For both the linear and non-linear case the PLSR-CV-EnKF updating scheme gave the best representation of the reference posterior mean and prediction intervals. The PCR-CV-EnKF scheme tended to overestimate the prediction uncertainty at grid locations where data were sparsely observed for $n_e = 100$. The reason for this behaviour is that the penalised PRESS statistic ensured that only one component was selected at each updating step for both the PCR-CV-EnKF and PLSR-CV-EnKF schemes. Hence, the PCR-CV-EnKF scheme underfitted the model because some of the components important for prediction purposes were discarded. This is apparently not the case for the supervised PLSR-CV-EnKF updating scheme.

The classical EnKF suffers from estimation uncertainty, model overfitting and coupling of the ensemble members unless the ensemble size tends to infinity. For a small ensemble size this will lead to an ensemble almost collapsing, as seen in both the linear and non-linear case.

Amongst the four schemes discussed above, the PCR-0.99-EnKF updating scheme gave the least favourable representation of the prediction uncertainty for both ensemble sizes. This behaviour is expected since p was selected based on the estimated proportion of explained variance, so that

between 10 and 12 components were used at each updating step for both the linear and non-linear model. As noted in Ledoit and Wolf (2004), the estimated eigenvalues of the empirical covariance matrix is known to be severely biased unless $n_d/n_e \rightarrow 0$, and the realisations in the data ensemble are independent identically distributed. In addition, no validation of the regression model is performed to evaluate the predictive capabilities, which in this case led to severe problems of model overfitting.

2.6 Reservoir Example

We consider a small synthetic reservoir model to further evaluate the performance of the PLSR-CV-EnKF and the classical EnKF scheme. The example is similar to the reservoir model used in Hegstad and Omre (2001), although the prior models for the porosity fields, ϕ , and ln-permeability, κ , are different.

2.6.1 Reservoir Description

The reservoir grid domain is of size (10000 x 10000 x 100) ft, discretised into $n = (10 \times 10 \times 15)$ regular grid blocks, with the top of the reservoir at depth 8 325 ft shown in Figure 2.8. The reference porosity and ln-permeability fields are generated by initially sampling from the Gaussian distribution described in Appendix 2-C. The ninth vertical and horizontal cross-sections of the reference ln-permeability and porosity are shown in Figures 2.9 and 2.10. Initially the reservoir is fully saturated with oil, with pressure 5 800 psi at the equilibrium depth of 8 400 ft.

There are two horizontal production wells (P1, P2), and one vertical gas injection well (I1) at the locations displayed in Figure 2.8, where the rate of the injection is assumed to be 65 000 Mscf/D. Production data from the reference model is simulated for 4 000 days using the commercial fluid flow simulator ECLIPSETM (GeoQuest, 2009). Observations are made of the Gas/Oil Ratio (GOR) and Oil Production Rate (OPR) in the two production wells and the Bottomhole Pressure (BHP) in both the injection and production wells. Figures 2.11a through g display the observed reference production data. As we can see from these figures, the production wells switch to BHP control when the pressure reaches 4 100 psi, which happens approximately after 1 200 days of production.

The purpose of this study is to demonstrate the effect of model overfitting. To reduce the computational demands, we therefore start updating the reservoir model when we start seeing a large variability in the data ensemble,

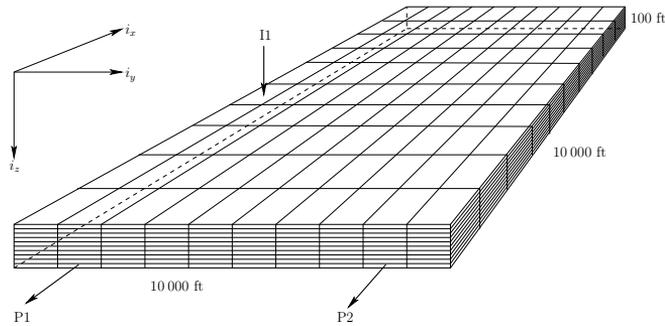


Figure 2.8: Description of the synthetic reservoir model discretised into $(10 \times 10 \times 15)$ grid blocks. The inward pointing arrow indicate the location of the injection well, while the outward pointing arrows indicate the locations of the producer wells.

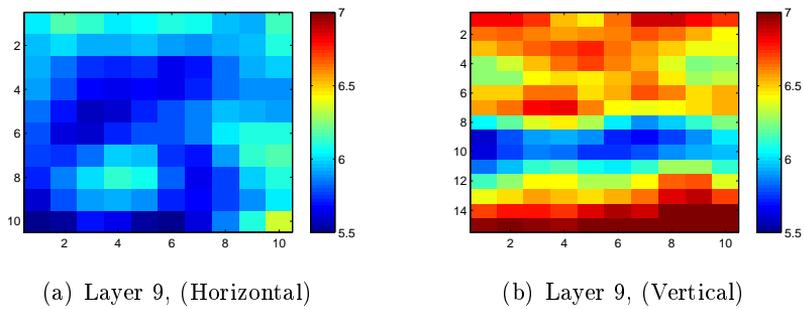


Figure 2.9: Image plot of one horizontal and vertical cross-section of the reference ln-permeability used in the synthetic reservoir example.

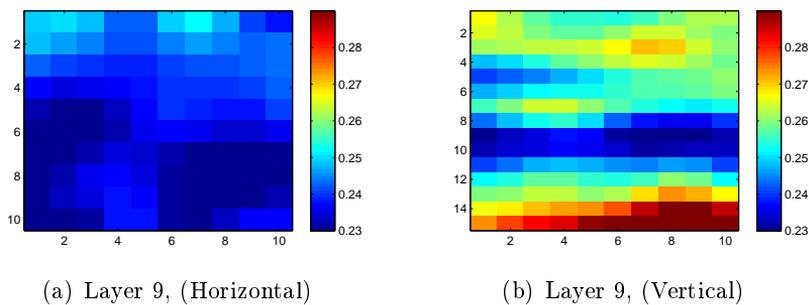


Figure 2.10: Image plot of one horizontal and vertical cross-section of the reference porosity used in the synthetic reservoir example.

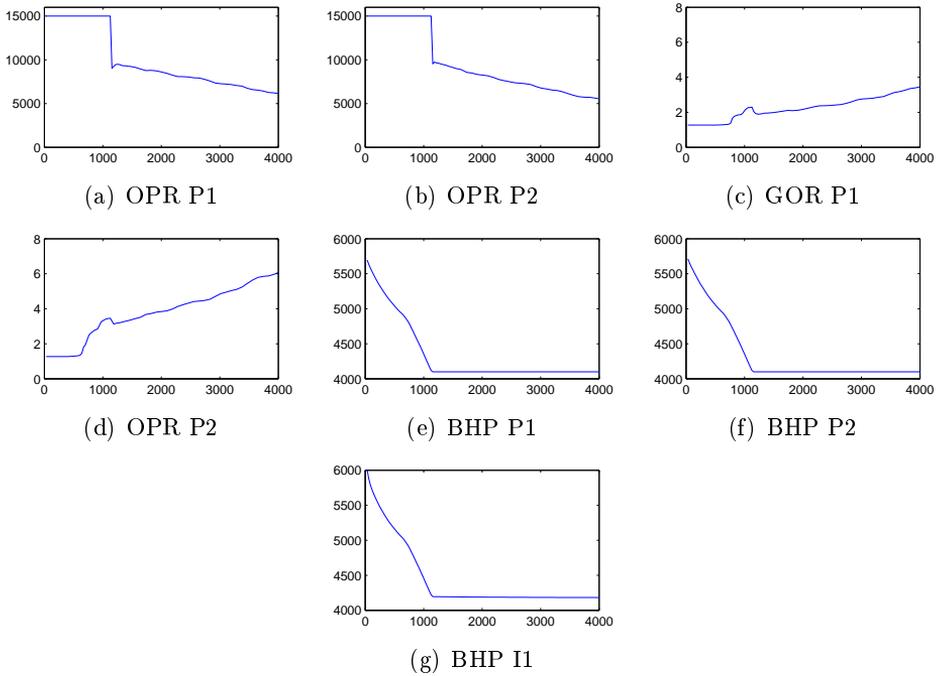


Figure 2.11: Reference production data.

after 910 days of production, as seen in Figure 2.12. We then proceed by updating the reservoir model based on the observed production data for the next 19 timesteps, following the observation schedule defined in Hegstad and Omre (2001). That is, the data is collected every 30 days for the next 540 days of production, and with one final update after 1640 days of production. Measurement errors are assumed to be additive Gaussian with a standard deviation of one percent of the observed value for the OPR and BHP, while for the GOR it is assumed to be 20 percent. In this study the reservoir state vector \mathbf{x}_k , contains κ , ϕ , logit-saturation, \mathbf{s}_k and pressure, \mathbf{p}_k .

2.6.2 Results with Discussion

The classical EnKF with four different ensemble sizes: $n_e = 20, 100, 1\,000$ and 1500, and the PLSR-CV-EnKF scheme with $n_e = 20$ are evaluated. Initially we generate 1 500 porosity and ln-permeability fields using the prior model described in Appendix 2-C. To make the results comparable, the initial ensemble members for the smaller ensemble sizes are then selected as the first 20, 100 and 1 000 members of the largest initial ensemble respectively. Note that the initial saturation and pressure are assumed to be known throughout the reservoir.

The forecasted production obtained when we restart the simulator from timestep zero and predict for 4 000 days, based on the updated κ and ϕ values are shown in Figure 2.12. As we can see from this figure, the initial ensemble fully spans the reference solution, and there is a relatively high uncertainty regarding the time of the gas breakthrough. Looking at results based on the classical EnKF with $n_e = 20$, we see that the average of the production forecast based on the updated ensemble members are missing the reference production. Moreover, reduced uncertainty in the production forecast causes the ensemble members not to span the reference production. For both the classical EnKF scheme with $n_e = 1\,500$ and the PLSR-CV-EnKF scheme with $n_e = 20$, however, the forecasts are correctly centred at the reference production and the uncertainty is considerably reduced. Note that here the CV scheme was based on minimising the PRESS statistic defined in Eq. (2.11), hence not penalising model overfitting. Nevertheless, the optimal choice of the PLSR subspace dimension was one at all updating steps.

To further quantify the results obtained using the classical EnKF and PLSR-CV-EnKF updating schemes, we investigate how well the updated ensemble members span the reference ϕ and κ . The estimated posterior mean and 95 % prediction interval for κ in the 100 grid blocks of horizontal cross-section nine, are displayed Figure 2.13. Note that the results at the

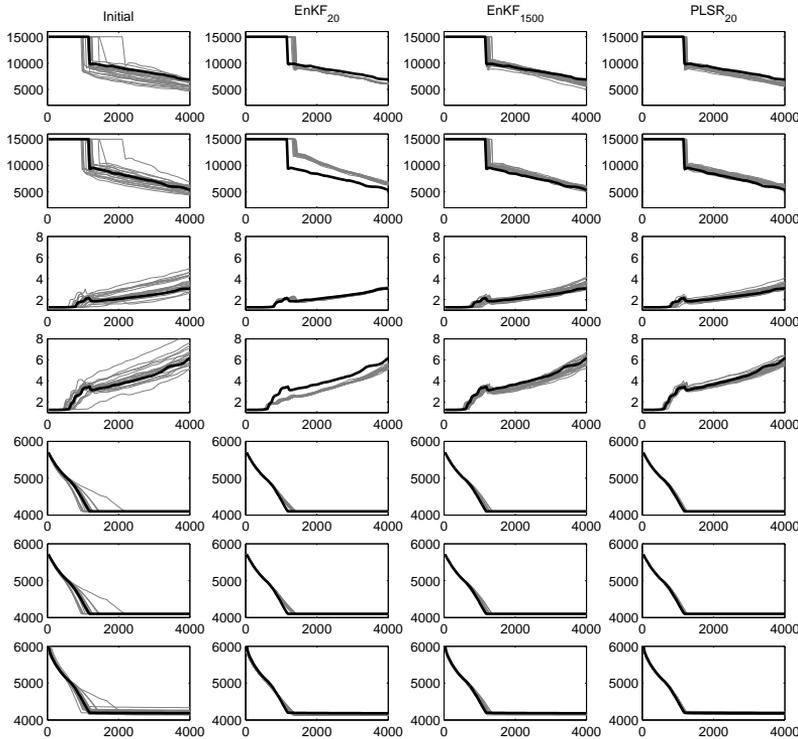


Figure 2.12: Forecasted production data based on reruns of the $n_e = 20$ first ensemble members from timestep zero compared with the reference production data (thick line). The figure shows from left to right the prediction based on the initial ensemble \mathbf{x}_0^c , the updated ensemble \mathbf{x}_{19}^c based on the classical EnKF with $n_e = 20$ (EnKF₂₀), $n_e = 1500$ (EnKF₁₅₀₀), and the updated ensemble \mathbf{x}_{19}^c based on the PLSR-CV-EnKF updating scheme with $n_e = 20$ (PLSR₂₀). The production properties considered are from top to bottom OPR P1, OPR P2, GOR P1, GOR P2, BHP P1, BHP P2 and BHP I1.

other 14 cross-sections have the same properties, with similar results for the porosity. We therefore do not present the results here.

The updated ensemble based on the classical EnKF scheme is not able to span the true ln-permeability for $n_e = 20$, and the ensemble has almost collapsed into a single realisation. Increasing the ensemble size to $n_e = 100$, does improve the uncertainty estimates. However, the ensemble mean appears to be highly variable, and we also see a tendency of overestimating κ and ϕ in most of the grid blocks. This leads to a bias in the production forecasts, not shown here, however. For the two largest ensemble sizes both the posterior mean, and prediction intervals appear to be similar. However, when rerunning the reservoir simulator from timestep zero using the updated κ and ϕ as input for $n_e = 1000$, we obtain production forecasts that deviates from the reference production curves, again not shown here. This is not the case for $n_e = 1500$, as shown in Figure 2.12. The PLSR-CV-EnKF updating scheme with $n_e = 20$, on the other hand, appears to provide a much better representation of the prediction uncertainty, although we are not able to fully cover 95% of the reference solution.

Similar to the empirical study above, the classical EnKF and PLSR-CV-EnKF updating schemes were rerun 100 times using different initial ensembles of size $n_e = 20$. The results are summarised in Table 2.4, where we estimate the scaled Residual Sum of Squares (RSS) between the ensemble members and the reference ln-permeability field, $1/n_e \sum_{i=1}^{n_e} \|\mathbf{x}^{c(i)} - \mathbf{x}^{\text{True}}\|_2^2$, and the percentage of the reference solution the estimated prediction intervals cover at each grid location. As we see from these results, the classical EnKF algorithm clearly underestimates the prediction uncertainty for the smallest ensemble size, $n_e = 20$.

Fig. 2.14 contains two realisations of ln-permeability from the initial ensemble and the corresponding realisations after the final updating step in the ninth vertical cross-section, using the classical EnKF updating scheme with four different ensemble sizes and the PLSR-CV-EnKF scheme with $n_e = 20$. We also display the estimated posterior means based on the first 20 ensemble members, which are equal at the initial timestep for all four ensemble sizes considered. The two realisations are different at the initial timestep, although the strong spatial correlation is present in both cases. From the average of the initial ensemble we see that there appears to be no particular trend in the initial model.

For the classical EnKF with $n_e = 20$, the realisations have almost collapsed and there appears to be a very high variability between neighbouring grid blocks. Moreover, the ln-permeability is well outside the range of the prior model at many of the grid locations. The realisations for $n_e = 100$

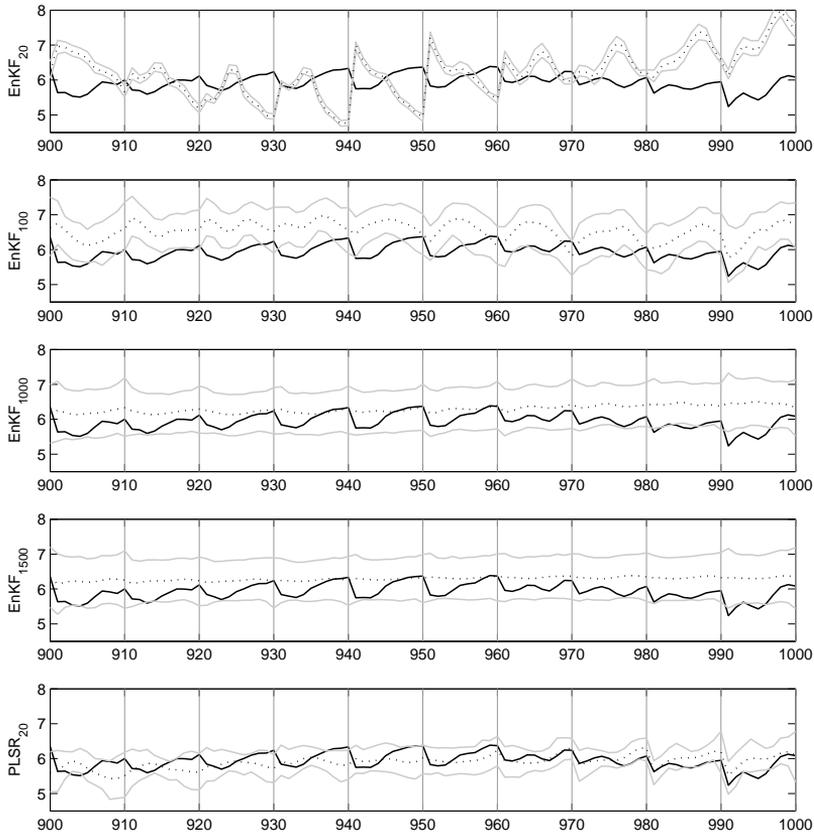


Figure 2.13: Updated ln-permeability values based on the classical EnKF updating scheme (EnKF) using four different ensemble sizes and the PLSR-CV-EnKF updating scheme (PLSR). The figure displays the reference ln-permeability (solid, black), the ensemble mean (dotted, black) and the estimated 95% confidence bounds of the prediction interval (solid, gray), for the grid blocks in the ninth horizontal cross-section, obtained based on the updated ensemble members at timestep 19. The subscript denotes the ensemble size used.

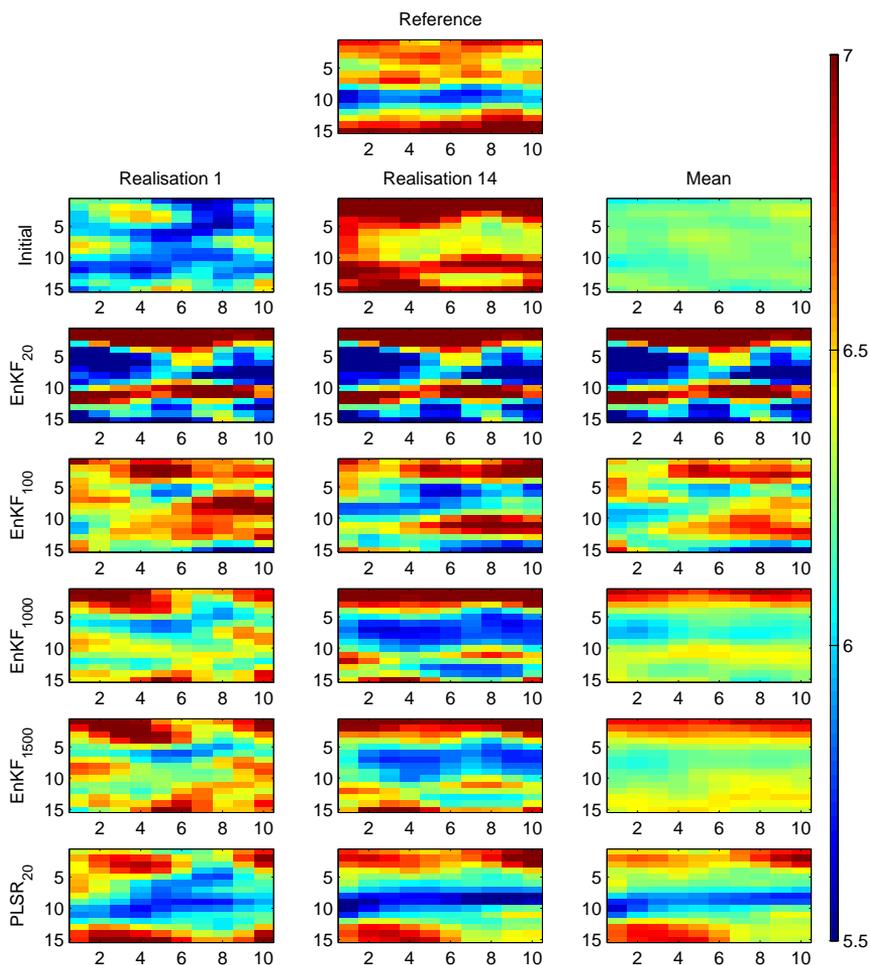


Figure 2.14: The ninth vertical cross-section for two realisations and the estimated ensemble mean based on the first 20 ensemble members for the initial ensemble, for different ensemble sizes using the classical EnKF updating scheme (EnKF) and the PLSR-CV-EnKF scheme (PLSR). The subscript denotes the ensemble size and the reference ln-permeability is shown at the top.

Scheme		Scaled RSS	Coverage %
No Updating	κ	24.1	95
Classical EnKF	κ	22.0	21
PLSR-CV-EnKF	κ	19.1	61
No Updating	ϕ	0.86	96
Classical EnKF	ϕ	0.74	21
PLSR-CV-EnKF	ϕ	0.64	62

Table 2.4: Scaled $(1/n_e)$ Residual Sum of Squares (RSS) for the the updated ensemble members $\mathbf{x}_{19}^{(i)c}$ and the reference solution, and coverage of the reference solution in the estimated 95% prediction intervals for the two static variables in the reservoir example. Estimates computed based on 100 different initial ensembles with $n_e = 20$.

appear to give a much better representation of the reference ln-permeability. However, we see that the ensemble members fail to capture the low permeable layers around horizontal cross-section 10 in the reference solution. For the two largest ensemble sizes, $n_e = 1000$ and 1500 , the spatial structure in the reference solution appears to be much better preserved in the updated realisations. We also observe the middle layer of low permeability present in the reference solution for both the updated realisations and ensemble mean. Again we see that the PLSR-CV-EnKF updating scheme with $n_e = 20$ provides updated realisation which appears to have many of the same features present in the prior model and reference solution.

Note that when the classical EnKF updating scheme with $n_e = 20$ was rerun several times, we observed that the RSS often became larger at the final updating step than for the initial ensemble. This can be explained by estimation uncertainty in the unknown covariance matrix $\Sigma_{\mathbf{x}}$ for small ensemble sizes (Evensen, 2007). Note, however, that this problem is also present in the classical EnKF updating scheme for $n_e = 100$ and $n_e = 1000$. These results suggest that errors in the EnKF solution are not only introduced by estimation uncertainty, but also by overfitting the regression model resulting from collinear ensemble members.

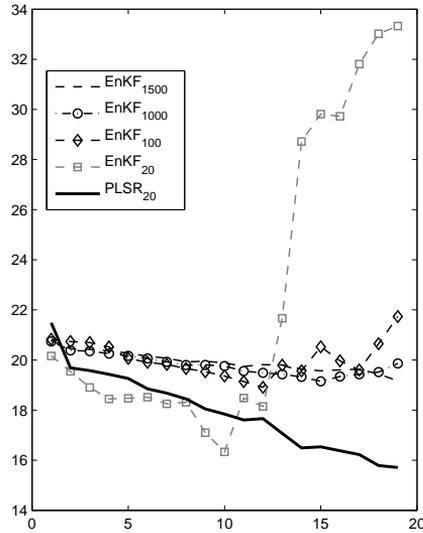
In Figures 2.15a and b, the RSS and the percentage of the reference ln-permeability located within the estimated 95% prediction interval are displayed as a function of timestep k . The RSS starts to increase after $k = 10, 12$ and 15 , for $n_e = 20, 100$ and 1000 respectively, with the most dramatic effect for the smallest ensemble size. At the same time, we see that

the coverage decrease as k increases, again with the largest effect for $n_e = 20$. For this particular reservoir model we have to increase n_e to 1 500 to have a decreasing trend in the RSS, while preserving the coverage, applying the classical EnKF updating scheme.

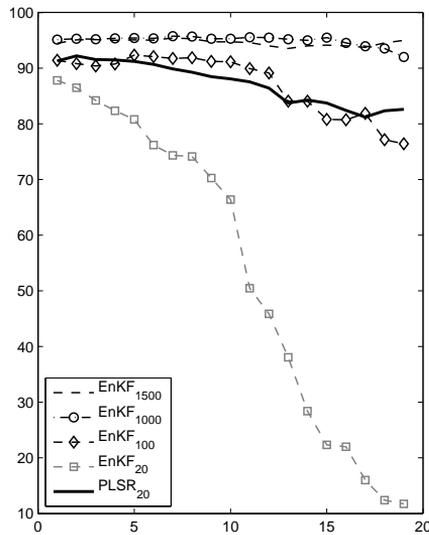
The PLSR-CV-EnKF scheme with $n_e = 20$ is able to preserve the decreasing trend in the RSS, with only a small decrease in the coverage as k increases. This appears to be caused by the reduced coupling of the updated ensemble members when using PLSR. Note, however, that both the classical EnKF and PLSR-CV-EnKF updating scheme with $n_e = 1500$ and 20 respectively, will eventually see the same increase in the RSS if data assimilation is continued into a distant future.

The increased coupling of the updated ensemble members can be quantified by the estimated rank of the updated ln-permeability ensemble, displayed in Figure 2.16. Here we have computed the loss in relative numerical rank for the updated ensemble at three different timesteps k , defined as one minus the numerical rank of the updated ln-permeability ensemble divided by the ensemble size). As seen from this figure the relative loss in rank for the updated ensemble members based on the classical EnKF updating scheme increases with timestep k . This is caused by overfitting of the regression model since the ensemble members become more collinear. We also see that this effect is more prominent for smaller ensemble sizes, leading to a larger increase in the RSS shown in Figure 2.15a. Notice that the relative rank loss is largest for $n_e = 1500$ after five updating steps for the classical EnKF updating scheme. This is expected, because there is always a probability greater than zero that some of the additional ensemble members will become collinear as we add more realisations to the ensemble. However, the large ensemble size, relative to the dimension of the data vector, reduces the effect of regression model overfitting in this case. Further note that by construction the PLSR-CV-EnKF scheme has a rank loss of one for all timesteps, as explained in Appendix 2-D, which for this case makes the approach less vulnerable to overfitting.

Finally note that if we compare the effective run-time for the reservoir example with $n_e = 20$ (rerun 100 times with different initial ensembles), the average run-time for the data assimilation step is 0.007 and 0.068 seconds the classical EnKF and PLSR-CV-EnKF updating schemes respectively. Running fluid flow simulator between two consecutive timesteps, on the other hand, has an average run-time of 25.5 and 25.0 seconds for the classical EnKF and PLSR-CV-EnKF updating schemes respectively. The reduction in the computational time in the flow simulation, which involves solving a set of non-linear Partial Differential Equations (PDE), is caused by improved



(a) Scaled RSS



(b) Coverage

Figure 2.15: Scaled $(1/n_e)$ Residual Sum of Squares (RSS) of the forecasted In-permeability ensemble members and the reference, and coverage of the reference solution in the estimated prediction intervals as a function of timesteps k . Here EnKF corresponds to the classical EnKF updating scheme, while PLSR corresponds to the PLSR-CV-EnKF scheme. The subscript denotes the ensemble size used.

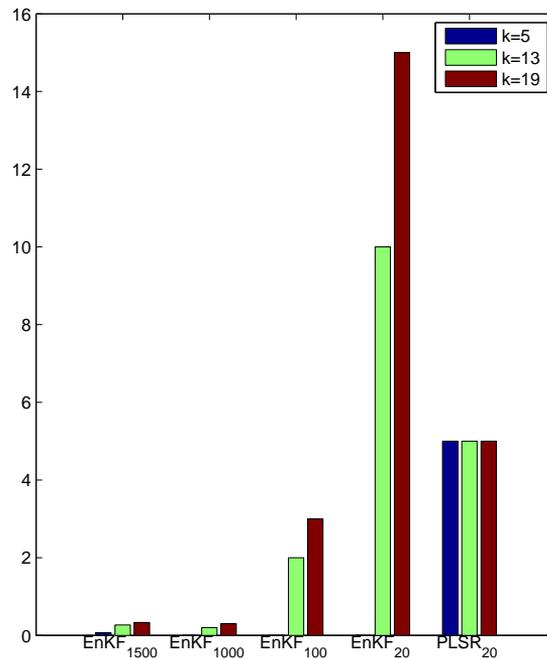


Figure 2.16: Estimated relative loss in rank (%) for the updated ln-permeability ensemble at three different timesteps, computed as one minus the numerical rank of the ensemble divided by the ensemble size. The notation EnKF corresponds to the classical EnKF updating scheme, while PLSR corresponds to the PLSR-CV-EnKF scheme. The subscript denotes the ensemble size.

convergence properties in the iterative PDE solver for the ensemble members updated using PLSR. Hence, the total run-time is smallest for the PLSR-CV-EnKF scheme, even if the data assimilation is slightly higher compared with the classical EnKF updating scheme. Note, however, that these run-times are based on a MATLAB implementation, where the classical EnKF updating scheme has been optimised (vectorised), while this is not the case for the PLSR-CV-EnKF updating scheme. The small difference in the effective run-time for the two updating schemes is therefore expected to be even smaller in reality.

2.7 Conclusion

We have formulated an EnKF updating scheme based on shrinkage regression techniques known from multivariate linear regression. In shrinkage regression, we replace the unbiased classical estimator of the Kalman gain matrix with biased alternatives, having improved predictive capabilities. Two of the techniques were considered on small linear and non-linear toy examples, namely Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR). When PCR and PLSR are combined with Cross-Validation (CV), the performance was far superior to the classical EnKF updating scheme for small ensemble sizes, with the supervised PLSR scheme providing slightly better results. However, when the subspace dimension used in PCR was selected based on the commonly applied theoretical criterion, the scheme suffered from similar problems as the classical EnKF. That is, low forecast precision and severe underestimation of the prediction uncertainty for small ensemble sizes.

We further compared the classical EnKF and PLSR updating schemes on a synthetic reservoir case study. Using the PLSR scheme we were able to obtain reasonable estimates of the prediction uncertainty for the porosity and permeability fields using only 20 ensemble members. The classical EnKF, on the other hand, required 75 times as many ensemble members to obtain similar results. This was caused by increasingly collinear ensemble members as reservoir production data was sequentially assimilated, leading to severe problems of model overfitting.

Acknowledgement

This work is funded by the Uncertainty in Reservoir Evaluation (URE) initiative at NTNU. We would also like to thank the Department of Petroleum Engineering and Applied Geophysics at NTNU for granting us access to their

computational resources and Inge Myrseth for providing the computer code for the small toy examples.

References

- S. I. Aanonsen, G. Nævdal, D. S. Oliver, A. C. Reynolds, and B. Vallès. Ensemble Kalman filter in reservoir engineering - a review. *SPE Journal*, 14(3):393–412, 2009.
- K. M. Adabir and J. R. Magnus. *Matrix Algebra*. Cambridge University Press, New York, 2005.
- J. L. Anderson. A local least squares framework for ensemble filtering. *Monthly Weather Review*, 131(4):634–642, 2003a.
- T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley, 3 edition, 2003b.
- M. Barker and W. Rayens. Partial least squares for discrimination. *Journal of Chemometrics*, 17:166–173, 2003.
- D. R. Cook. Fisher Lecture: Dimension reduction in regression. *Statistical Science*, 22(1):1–26, 2007.
- A. Doucet, S. Godsill, and C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000.
- B. Efron. The estimation of prediction error: Covariance penalties and Cross-Validation. *Journal of the American Statistical Association*, 99(467):619–632, 2004.
- G. Evensen. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99:10143–10162, 1994.
- G. Evensen. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, 2003.
- G. Evensen. *Data assimilation. The Ensemble Kalman Filter*. Springer, 2007.
- D. E. Farrer and R. R. Glauber. Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics*, 49(1): 92–107, 1967.

- R. Furrer and T. Bengtsson. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *Journal of Multivariate Analysis*, 98(2):227–255, 2007.
- GeoQuest. *ECLIPSE reference manual 2009.1*. Schlumberger, GeoQuest, 2009.
- A. S. Hadi and R. F. Ling. Some cautionary notes on the use of principal components regression. *The American Statistician*, 52(1):15–19, 1998.
- T. Hastie and R. Tibshirani. Efficient quadratic regularization for expression arrays. *Biostatistics*, 5(3):329–340, 2004.
- T. Hastie, R. Tibshirani, and J. Freidman. *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. Springer, New York, 2 edition, 2009.
- B. K. Hegstad and H. Omre. Uncertainty in production forecasts based on well observations, seismic data and production history. *Society of Petroleum Engineers Journal*, 6(4):409–425, 2001.
- I. S. Helland. Some theoretical aspects of partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 58(2):97–107, 2001.
- H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 1933.
- P. L. Houtekamer and H. L. Mitchell. Data assimilation using an Ensemble Kalman Filter technique. *Monthly Weather Review*, 126:796–811, 1998.
- P. L. Houtekamer and H. L. Mitchell. Reply. *Monthly Weather Review*, 127(6):1378–1379, 1999.
- A. Höskuldsson. PLS regression methods. *Journal of Chemometrics*, 2: 211–228, 1988.
- I. Jolliffe. A note on the use of Principal Components in regression. *Applied Statistics*, 31(3):300–303, 1982.
- I. T. Jolliffe. *Principal Component Analysis*. Springer, 2 edition, 2002.
- J. H. Kalivas. Cyclic subspace regression with analysis of the hat matrix. *Chemometrics and Intelligent Laboratory Systems*, 45(1–2):215–224, 1999.

- R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82(Series D): 35–45, 1960.
- M. H. Kaspar and W. H. Ray. Partial Least Squares modelling as successive Singular Value Decomposition. *Computers and Chemical Engineering*, 17 (10):985–989, 1993.
- O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, 2004.
- K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate analysis*. Academic Press, London, 1979.
- I. B. Myrseth and Omre. The Ensemble Kalman Filter and related filters. In L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, Y. Marzouk, L. Tenorio, B. van Bloemen Waanders, and K. Willcox, editors, *Large-scale Inverse Problems and Quantification of Uncertainty*. John Wiley and Sons, 2010.
- I. B. Myrseth, J. Sætrom, and H. Omre. Resampling the Ensemble Kalman Filter. Paper submitted for publication, 2009.
- B. D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996.
- R. Rosipal and N. Krämer. *Subspace, latent structure and feature selection techniques*, volume 2940 of *Lecture Notes in Computer Science*, chapter Overview and Recent Advances in Partial Least Squares, pages 34–51. Springer, 2006.
- S. Rännar, F. Lindgren, P. Gelandi, and S. Wold. A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. *Journal of Chemometrics*, 8:111–125, 1994.
- W. Sacher and P. Bartello. Sampling errors in Ensemble Kalman filtering. Part I: Theory. *Monthly Weather Review*, 136(8):3035–3049, 2008.
- G. A. F. Seber and A. J. Lee. *Linear Regression Analysis*. Wiley, 2003.
- J.-A. Skjervheim, G. Evensen, S. Aanonsen, B. O. Ruud, and T. A. Johansen. Incorporating 4D seismic data in reservoir simulation models using Ensemble Kalman Filter. *SPE Journal*, 12(3):282–292, 2007.

- G. Strang. *Linear Algebra and its Applications*. Thomson Learning, 1988.
- A. N. Tikhonov and V. A. Arsenin. *Solution of ill-posed problems*. Winston & Sons, Washington, 1977.
- P. J. van Leeuwen. Comments on "Data assimilation using an Ensemble Kalman Filter technique". *Monthly Weather Review*, 127:1374–1377, 1999.
- H. Wold. *Quantitative Sociology: International perspectives on mathematical and statistical model building*, chapter Path models with latent variables: The NiPALS approach, pages 307–357. Academic Press, 1975.
- X.-Q. Zeng, M.-W. Wang, and J.-Y. Nie. Text classification based on Partial Least Square analysis. In *The 22nd Annual ACM Symposium on Applied Computing, Special Track on Information Access and Retrieval*, 2007.

APPENDIX

Appendix 2-A, Use of Singular Value Decomposition

The eigenvectors of the empirically estimated covariance matrix of the data vector, $\hat{\Sigma}_{\mathbf{d}}$, are given as the column vectors $\mathbf{U} \in \mathbb{R}^{n_d \times n_d}$ obtained when performing Singular Value Decomposition (SVD) on $\frac{1}{\sqrt{n_e}} \mathbf{D} = \mathbf{U} \mathbf{S} \mathbf{V}^T$ (Strang, 1988). Moreover, the eigenvalues are given as $\lambda_i = s_{ii}^2$, where s_{ii} is the i th singular value, given as the i th diagonal element of the matrix $\mathbf{S} \in \mathbb{R}^{n_d \times r}$ where r is the rank of \mathbf{D} . Hence, the ensemble matrix \mathbf{D} can be approximated by a truncated version:

$$\mathbf{D}_p = \mathbf{U}_p \mathbf{S}_p \mathbf{V}_p^T, \tag{A-1}$$

where $\mathbf{U}_p = [\mathbf{u}_1, \dots, \mathbf{u}_p] \in \mathbb{R}^{n_d \times p}$, $\mathbf{S}_p = \text{diag}_p(\mathbf{s}_p) \in \mathbb{R}^{p \times p}$, and $\mathbf{V}_p = [\mathbf{v}_1, \dots, \mathbf{v}_p] \in \mathbb{R}^{n_e \times p}$, with \mathbf{v}_i given as the i th eigenvector of the matrix $\mathbf{G}_{\mathbf{d}} = \mathbf{D}^T \mathbf{D} \in \mathbb{R}^{n_e \times n_e}$. Here the notation $\mathbf{S}_p = \text{diag}_p(\mathbf{s}_p)$ is used to denote that the p dimensional matrix \mathbf{S}_p is a diagonal matrix with the vector \mathbf{s}_p on the main diagonal.

By use of the truncated SVD, the PCR estimate for the matrix of regression coefficient is given as:

$$\hat{\mathbf{K}}_{\text{PCR}} = \mathbf{X} \mathbf{V}_p \mathbf{S}_p^{-1} \mathbf{U}_p^T. \tag{A-2}$$

Thus, the computation of $\hat{\mathbf{K}}_{\text{PCR}}$ is efficient, both in terms of speed and memory use, compared to working with full dimensional covariance matrices.

This is especially true when n_d is larger than n_e . It should be noted that setting $p = n_e - 1$, $\hat{\mathbf{K}}_{\text{PCR}}$ minimises the mean squared error in the rank deficient case.

Appendix 2-B, Computational Properties

The classical EnKF can be modified by simply replacing the estimated Kalman gain with either the PCR or PLSR estimates. For both methods the computational complexity is $\mathcal{O}(\max\{n_d, n_x\} \cdot n_e^2)$. For the PCR estimate this corresponds to the cost of performing a SVD on \mathbf{D} and the matrix-matrix multiplication \mathbf{XV}^T , while for the PLSR estimate this is caused by the computation of the two matrices $\mathbf{G}_x = \mathbf{X}^T \mathbf{X} \in \mathbb{R}^{n_e \times n_e}$ and $\mathbf{G}_d = \mathbf{D}^T \mathbf{D} \in \mathbb{R}^{n_e \times n_e}$.

When computing the PLSR estimate of the Kalman gain, we see from Eq. (2.10) that this requires the inversion of a $p \times p$ dimensional matrix, which in the general case requires $\mathcal{O}(p^3)$ floating point operations (flops). However, it can be shown (Höskuldsson, 1988; Ränner et al., 1994), that this matrix is lower triangular, reducing the number of flops required to $\mathcal{O}(p^2)$. If both n_x and n_d are large, the classical NiPALS algorithm and computing the SVD of $\hat{\Sigma}_{x,d}$, will be computationally demanding. Note, however, that this problem can be avoided by the algorithm outlined in Ränner et al. (1994).

When m -fold CV is applied to select the optimal number of components for PCR and PLSR, the computational complexity is $\mathcal{O}(\max\{n_e \cdot n_x, (m - 1) \cdot n_d\} n_e^2)$, when the PRESS statistic is computed for all possible values for p in the truncation. Note that the computational complexity can be further reduced when $n_d > n_e - 1$. As explained in Hastie and Tibshirani (2004), the same results will be obtained if we perform the m -fold CV based on the $n_e - 1$ latent vectors, instead of the data ensemble. Hence, only a single projection into the reduced order space using the PCR or PLSR techniques will be necessary. Further note that if we replace the Euclidean norm used in the PRESS statistic in Eq. (2.11), with the (pseudo) norm:

$$\|\mathbf{a}\|_{\hat{\Sigma}_x}^2 = n_e \mathbf{a}^T \hat{\Sigma}_x \mathbf{a} = \mathbf{a}^T \mathbf{X} \mathbf{X}^T \mathbf{a},$$

we can further reduce the computational complexity to $\mathcal{O}(\max\{n_x, n_e^2, n_d\} n_e^2)$. For large dimensional state vectors, we therefore see that the additional computational demands caused by m -fold CV can be negligible as long as $n_e^2 < \max\{n_x, n_d\}$.

When overfitting is a severe problem for PCR or PLSR, the CV rule described in Hastie et al. (2009) can be applied. Rather than selecting p at the global minimum of the PRESS statistic, the optimal number of

components, p^* , should be given as the smallest number of components such that $\text{PRESS}(p^*) \leq \text{PRESS}(p^G) + \hat{\sigma}_{\text{PRESS}(p^G)}$. Here $\text{PRESS}(p^G)$ and $\hat{\sigma}_{\text{PRESS}(p^G)}$ corresponds to the estimated value and standard deviation of the PRESS statistic at the global minimum p^G respectively.

Although CV can lead to higher computational demands if an exhaustive search for the optimal number of components is carried out, it should be noted that testing for all possible combinations of p is often not required. This is especially true when using the PLSR approach because this tend to require a smaller number of components than PCR (Kalivas, 1999; Helland, 2001).

Finally note that a multivariate multiple linear regression problem can be equally treated as a sequence of multivariate linear regression problems with a univariate response variable (Hastie et al., 2009). The EnKF updating schemes based on shrinkage regression techniques is therefore straightforward to both scale and parallelise when considering high dimensional state- and data vectors with n_x and n_d in the order of $\mathcal{O}(10^6)$ - $\mathcal{O}(10^8)$.

Appendix 2-C, Prior Distribution

For the reservoir example the porosity and ln-permeability fields are described by the following prior distribution: Initially, a realisation $\mathbf{z}^{(i)} \in \mathbb{R}^{n \times 1}$ is generated from a Gaussian distribution with mean 0.25 and covariance matrix $\Sigma_{\mathbf{z}} \in \mathbb{R}^{n \times n}$, with $\Sigma_{\mathbf{z}\Delta}$ defined by an exponential covariance function:

$$c(\Delta) = \sigma_{\mathbf{z}}^2 \exp \left\{ -\sqrt{\|\Delta\|_l^2} \right\}.$$

Here $\sigma_{\mathbf{z}} = 0.02$ and

$$\|\Delta\|_l^2 = \left(\frac{\Delta_x}{l_x} \right)^2 + \left(\frac{\Delta_y}{l_y} \right)^2 + \left(\frac{\Delta_z}{l_z} \right)^2,$$

with $l_x = l_y = 10\,000$ and $l_z = 100$. We further assume that

$$[\mathbf{w}|\mathbf{z}] = 25\mathbf{z} + \epsilon_{\mathbf{w}|\mathbf{z}},$$

where $\epsilon_{\mathbf{w}|\mathbf{z}} \sim \text{Gauss}_n(\mathbf{0}, \mathbf{I})$. A convolution between the realisations and a three dimensional standard Gaussian kernel over a (5 x 5 x 5) neighbourhood is then applied to generate porosity and ln-permeability fields. This entails that $\phi^{(i)} = \mathbf{C}\mathbf{z}^{(i)}$ and $\kappa^{(i)} = \mathbf{C}\mathbf{w}^{(i)}$, where $\mathbf{C} \in \mathbb{R}^{n \times n}$ is the convolution matrix. Note that the purpose of this convolution is simply to increase the spatial continuity in the realisations. Further note that in practise the

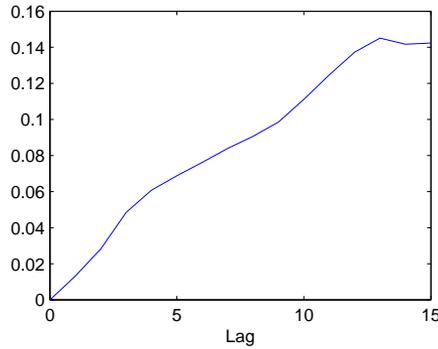


Figure 2-C-1: Estimated variogram based on realisations of the ln-permeability fields from the initial prior distribution.

MATLAB function `smooth3(·)` is used. The estimated variogram, based on realisations of the ln-permeability from the prior distribution, is shown in Figure 2-C-1. Because we assume that there is a linear relationship between the porosity and ln-permeability fields, the estimated correlation coefficient based on realisations from the prior model is approximately 0.8. We acknowledge that this linear assumption is not valid for reservoir models in general. Note, however, that the prior model ensures that there is a smooth trend in the realisations from the prior model and that they in practise are restricted to $\kappa \in [4, 8]$ and $\phi \in [0.2, 0.32]$, which might make the model violation less severe.

Appendix 2-D, Rank of the Updated Ensemble

Let $\mathbf{X}^u \in \mathbb{R}^{n_x \times n_e}$ and $\mathbf{D} \in \mathbb{R}^{n_d \times n_e}$ be two non-centred ensemble matrices. Further let $\mathbf{C} = \mathbf{I} - \frac{1}{n_e} \mathbf{1} \mathbf{1}^T$ denote the centring matrix, where $\mathbf{1}$ is a n_e -dimensional column vector with all entries equal to one. The EnKF updating scheme based on PLSR can then be written as:

$$\mathbf{X}^c = \mathbf{X}^u [\mathbf{I} + \mathbf{T} \mathbf{A}^{-1} \mathbf{W}^T \mathbf{D}^T (\mathbf{d} \mathbf{1}^T - \mathbf{D})], \quad (\text{D-1})$$

where

$$\mathbf{A} = (\mathbf{W} \mathbf{D}^T \mathbf{D} \mathbf{T}).$$

This follows from the definition of $\hat{\mathbf{K}}_{\text{PLSR}}$ and the identities:

$$\mathbf{T} = \mathbf{C} \mathbf{D}^T \mathbf{\Psi} = \mathbf{C} \mathbf{C} \mathbf{D}^T \mathbf{\Psi} = \mathbf{C} \mathbf{T} \quad (\text{D-2})$$

and $\mathbf{W} = \mathbf{C}\mathbf{W}$, where we have used that \mathbf{C} is an idempotent matrix.

For the EnKF updating scheme defined in Eq. (D-1) using p components, assuming $\text{rank}(\mathbf{X}^u) = n_e$, the following result then holds:

Result 1. *Assuming $\text{rank}(\mathbf{X}^u) = n_e$, the rank of the updated state vector ensemble for the EnKF updating scheme based on PLSR and PCR is equal to $n_e - p$.*

Proof. Consider the matrix:

$$\mathbf{H} = \mathbf{I} + \mathbf{T}\mathbf{A}^{-1}\mathbf{W}^T\mathbf{D}^T(\mathbf{d}\mathbf{1}^T - \mathbf{D}) \in \mathbb{R}^{n_e \times n_e}.$$

By use of Eq. (D-2) and the property that $\mathbf{1}^T\mathbf{C} = 0$, we obtain the identity $\mathbf{H} = \mathbf{H}^2$. Let $(\lambda_i, \mathbf{e}_i)$, $i = 1, \dots, n_e$ be an eigenvalue, -vector pair of \mathbf{H} . From the identity

$$\lambda_i \mathbf{e}_i = \mathbf{H}\mathbf{e}_i = \mathbf{H}\mathbf{H}\mathbf{e}_i = \lambda_i^2 \mathbf{e}_i,$$

we then see that all eigenvalues of the idempotent matrix \mathbf{H} are equal to zero or one. Hence, the result follows from the identity:

$$\text{rank}(\mathbf{X}^c) = \text{rank}(\mathbf{X}^u\mathbf{H}) = \text{rank}(\mathbf{H}) = \text{tr}(\mathbf{H}),$$

where second equality holds because $\text{rank}(\mathbf{X}^u) = n_e$ (Adabir and Magnus, 2005) and the third equality holds because \mathbf{H} is idempotent. The proof for the EnKF updating scheme based on PCR is along the same lines. \square

Paper II

Ensemble Kalman filtering for non-linear likelihood models using kernel-shrinkage regression techniques

Jon Sætrum and Henning Omre

Paper submitted for publication

Chapter 3

Ensemble Kalman filtering for non-linear likelihood models using kernel-shrinkage regression techniques

Abstract. One of the major limitations of the classical Ensemble Kalman Filter (EnKF) is the assumption of a linear relationship between the state vector and the observed data. Thus, the classical EnKF algorithm can suffer from poor performance when considering highly non-linear and non-Gaussian likelihood models. In this paper, we have formulated the EnKF based on kernel-shrinkage regression techniques. This approach makes it possible to handle highly non-linear likelihood models efficiently. Moreover, a solution to the pre-image problem, essential in previously suggested EnKF schemes based on kernel methods, is not required. Testing the suggested procedure on a simple, illustrative problem with a non-linear likelihood model, we were able to obtain good results when the classical EnKF failed.

3.1 Introduction

In recent years, Bayesian methods have become attractive to use when considering geophysical inverse problems (Scales and Snieder, 1997). The Ensemble Kalman Filter (EnKF) is a Bayesian method that provides a solution to highly non-linear and high dimensional spatiotemporal data assimilation problems (Aanonsen et al., 2009; Evensen, 2007). The EnKF is defined in the spirit of the classical Kalman Filter (KF) (Kalman, 1960), that provides the analytical solution of the posterior probability distribution when the prior, forward and likelihood models are Gaussian and linear, termed the Gauss-linear model.

Analytical tractability of the posterior distribution will be lost in a general model setting. Thus, we may apply techniques such as Markov chain Monte Carlo (MCMC) or rejection sampling to generate realisations from the posterior distribution of interest (Doucet et al., 2000). For high dimensional problems, however, these methods tend to be computationally prohibitive.

The EnKF approach is based on the approximation that the output of the forward and likelihood models are jointly Gaussian, with unknown mean and covariance. Using an ensemble of independent realisations to estimate the model parameters empirically, ensures that the EnKF is consistent with the KF for Gauss-linear models as the ensemble size tends to infinity (Evensen, 2007; Mardia et al., 1979).

As shown in Anderson (2003), we can equally formulate the classical EnKF updating scheme as a multivariate linear regression problem, where the Kalman gain matrix defines the unknown matrix of regression coefficients. Hence, the classical EnKF can have poor performance when considering highly non-linear likelihood models. Methods such as the Randomised Maximum Likelihood Filter (RMLF) (Oliver, 1996) can improve on the EnKF updating scheme for non-linear likelihood models. However, the RMLF algorithm requires an optimisation step making the method more computationally demanding than the traditional EnKF. This is especially true when considering high dimensional data such as time-lapse seismic. In addition, it is unclear how to use the RMLF if the error term is not additive (Myrseth and Omre, 2010).

The classical EnKF tends to underestimate the prediction uncertainty for small ensemble sizes (Furrer and Bengtsson, 2007; Houtekamer and Mitchell, 1998; Myrseth et al., 2010; Sacher and Bartello, 2008; van Leeuwen, 1999). This can potentially lead to an ensemble collapsing into one single realisation. In a recent paper, Sætrom and Omre (2010), reformulated the classical EnKF updating scheme using shrinkage regression techniques known from

multivariate statistics. It is well known from statistical literature that the unbiased classical least squares estimator for the matrix of regression coefficients is not optimal in the presence of collinear data, and can lead to severe problems of model overfitting (Hastie et al., 2009). The purpose of shrinkage regression techniques is therefore to replace the classical, unbiased estimator of the unknown matrix of regression coefficients with biased alternatives having improved predictive capabilities; e.g. using dimension reduction techniques on the predictor variables, or by regularising the estimated matrix of regression coefficients. Because the updated ensemble members will be coupled through the estimated Kalman gain matrix, collinearities between the ensemble members will eventually occur (Houtekamer and Mitchell, 1998). Hence, it is not surprising that applying shrinkage regression techniques can lead to significant improvements compared with the classical EnKF updating scheme, as illustrated in the examples considered in Sætrom and Omre (2010).

In recent years, kernel methods have become popular within the field of machine learning (Smola and Schölkopf, 2002; Taylor and Cristianini, 2004). The aim of these methods is to transform data from the original vector space into a possibly high dimensional feature space, where we assume that the underlying model assumptions, such as linear dependencies in a regression setting, are valid. Kernel methods are frequently used in non-linear Principal Component Analysis (PCA) (Sarma et al., 2008; Schölkopf et al., 1998), data mining (Huang et al., 2006), classification (Vapnik, 1998) and non-linear regression (Rosipal and Trejo, 2002). Common for these methods is that the algorithms can be reformulated through inner products in the original space. Because we can define inner products in a feature space through positive definite functions, known as kernel functions, there is no need to generate realisations in the feature space (Hofmann et al., 2008; Smola and Schölkopf, 2002).

In the current paper, we extend the EnKF updating scheme to a non-linear setting using previously defined kernel-based shrinkage regression techniques (Rosipal and Trejo, 2002; Taylor and Cristianini, 2004). We demonstrate the suggested approach on a simplistic example with a non-linear, non-Gaussian likelihood model. The procedure has the same computational complexity and memory requirements as the fastest implementations of the traditional EnKF.

We are not the first to recognise the potential of kernel methods in an EnKF setting (Caers and Park, 2008; Sarma and Chen, 2009). However, the focus of these studies is to incorporate highly non-Gaussian features of the state vector into the EnKF, which require a solution to the pre-image

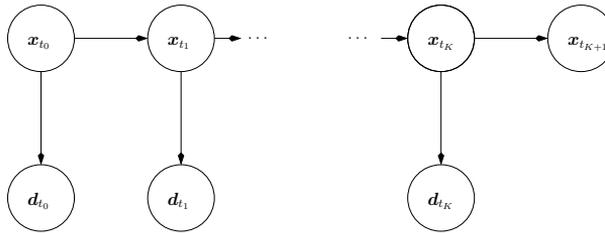


Figure 3.1: Stochastic Directed Acyclic Graph (DAG) of the model considered.

problem of mapping the state vector from the feature space back to the original space. Solving this problem using traditional approaches requires non-linear optimisation techniques (Kwok and Tsang, 2004; Mika et al., 1998; Smola and Schölkopf, 2002), which can lead to high computational demands. In the current study we use kernel methods for handling non-linearities in the likelihood model, which do not require this back-transformation.

3.2 Notation and Model Formulation

Throughout the paper, we use the notation $\mathbf{x} \in \mathbb{R}^{n_x \times 1}$ to denote that \mathbf{x} is an n_x -dimensional column vector in the real space and \mathbf{x}^T will denote its transpose. Similarly, we will write $\mathbf{X} \in \mathbb{R}^{m \times n}$ to denote that \mathbf{X} is a matrix in the real space containing m rows and n columns. Note that we will use that the same notation for both scalars and random variables. Probability density functions (pdf) will be denoted by $f(\mathbf{x})$, and the notation $\mathbf{x} \sim f(\mathbf{x})$, implies that the random vector \mathbf{x} follows the pdf $f(\mathbf{x})$. Further we will denote the conditional pdf of \mathbf{x} given \mathbf{y} by $f(\mathbf{x}|\mathbf{y})$. As a special case the notation $\mathbf{x} \sim \text{Gauss}_{n_x}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ will be used to denote that \mathbf{x} follows the n_x -dimensional multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}_x$ and covariance matrix $\boldsymbol{\Sigma}_x$.

Consider the sequence of stochastic vectors $\mathbf{x}_{t_0}, \dots, \mathbf{x}_{t_{K+1}}$; $\mathbf{x}_{t_i} \in \mathbb{R}^{n_x \times 1}$ and $\mathbf{d}_{t_0}, \dots, \mathbf{d}_{t_K}$; $\mathbf{d}_{t_i} \in \mathbb{R}^{n_d \times 1}$, outlined in Fig. 3.1. Here, \mathbf{x}_{t_k} denotes the state of the unknown random vector of interest at time step k and time t_k , and similarly \mathbf{d}_{t_k} denotes the vector of observed data. For notational convenience, we will from now on drop the subscript t_k , and simply write \mathbf{x}_k and \mathbf{d}_k . Also note that we will for simplicity refer to \mathbf{x} and \mathbf{d} as the state and observation vector respectively.

Let $f(\mathbf{x}_0)$ denote the pdf of the state vector at the initial time step. The

Markov property of the Directed Acyclic Graph (DAG) in Fig. 3.1 entails

$$f(\mathbf{x}_{k+1}|\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_0) = f(\mathbf{x}_{k+1}|\mathbf{x}_k), \quad k = 0, \dots, K.$$

We define the pdf, $f(\mathbf{x}_{k+1}|\mathbf{x}_k)$, through a known, possibly highly non-linear forward function, $\boldsymbol{\omega} : (\in \mathbb{R}^{n_x \times 1} \times \in \mathbb{R}^{n_x \times 1}) \rightarrow \in \mathbb{R}^{n_x \times 1}$, which implies

$$\mathbf{x}_{k+1} = \boldsymbol{\omega}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{x}_k}), \quad k = 0, \dots, K. \quad (3.1)$$

Here $\boldsymbol{\epsilon}_{\mathbf{x}_k} \in \mathbb{R}^{n_x \times 1}$ represents random model errors or numerical errors in the forward model, assumed to follow a known probability distribution. Thus, $f(\mathbf{x}_0)$ and $\boldsymbol{\omega}(\cdot, \cdot)$ implicitly defines the prior pdf $f(\mathbf{x}_0, \dots, \mathbf{x}_{K+1})$. We define the likelihood function, $f(\mathbf{d}_k|\mathbf{x}_k)$, through a known, non-linear function $\boldsymbol{\zeta} : (\mathbb{R}^{n_x \times 1} \times \mathbb{R}^{n_d \times 1}) \rightarrow \mathbb{R}^{n_d \times 1}$, that is,

$$\mathbf{d}_k = \boldsymbol{\zeta}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{d}_k}), \quad k = 0, \dots, K. \quad (3.2)$$

Again, $\boldsymbol{\epsilon}_{\mathbf{d}_k} \in \mathbb{R}^{n_d \times 1}$, represents the random observation error following a known pdf.

For notational convenience, we from now on let

$$\begin{aligned} \mathbf{x}_k^c &\sim f(\mathbf{x}_k|\mathbf{d}_{0:k}) \\ \mathbf{x}_k^u &\sim f(\mathbf{x}_k|\mathbf{d}_{0:k-1}), \quad k = 1, \dots, K+1, \end{aligned}$$

where $\mathbf{d}_{0:l}$ denotes the sequence $\mathbf{d}_0, \dots, \mathbf{d}_l$ for $l > 0$. Bayesian inversion provides a sequential solution to the spatiotemporal forecast problem of predicting \mathbf{x}_k^c , for $k = 1, \dots, K+1$. With such an approach, we assess the unknown vectors \mathbf{x}_k^c and \mathbf{x}_{k+1}^u by sampling from the respective posterior distributions, $f(\mathbf{x}_k|\mathbf{d}_{0:k})$ and $f(\mathbf{x}_{k+1}|\mathbf{d}_{0:k})$. Using Bayes rule and the Markov property of the DAG in Fig. 3.1, which entails $f(\mathbf{d}_k|\mathbf{x}_k, \mathbf{d}_{0:(k-1)}) = f(\mathbf{d}_k|\mathbf{x}_k)$, for $k = 1, \dots, K$, we get

$$\begin{aligned} f(\mathbf{x}_k|\mathbf{d}_{0:k}) &\propto f(\mathbf{x}_k|\mathbf{d}_{0:(k-1)})f(\mathbf{d}_k|\mathbf{x}_k) \\ f(\mathbf{x}_{k+1}|\mathbf{d}_{0:k}) &= \int f(\mathbf{x}_{k+1}|\mathbf{x}_k)f(\mathbf{x}_k|\mathbf{d}_{0:k})d\mathbf{x}_k. \end{aligned} \quad (3.3)$$

Generally, we only know the conditional distributions defined in Eq. (3.3) up to an unknown normalising constant. One possibility is to use computationally demanding techniques such as MCMC or rejection sampling to generate realisation from the correct posterior distribution (Doucet et al., 2000). However, for applications such as petroleum reservoir evaluation, these techniques are computationally prohibitive (Evensen, 2007). An approximate solution can be obtained by assuming that \mathbf{x}_k^u and \mathbf{d}_k follows a distribution that ensures analytical tractability of $f(\mathbf{x}_k|\mathbf{d}_{0:k})$, such as the Gaussian. These model assumptions are equivalent to those made in the EnKF (Evensen, 1994).

3.3 Classical Ensemble Kalman Filter

Let

$$\mathbf{x}_k^{u(i)} = \boldsymbol{\omega}(\mathbf{x}_{k-1}^{c(i)}, \boldsymbol{\epsilon}_{\mathbf{x}_{k-1}}^{(i)})$$

and

$$\mathbf{d}_k^{(i)} = \boldsymbol{\zeta}(\mathbf{x}_k^{u(i)}, \boldsymbol{\epsilon}_{\mathbf{d}_k}^{(i)}),$$

for $i = 1, \dots, n_e$, and define $\mathbf{X}_k = [\mathbf{x}_k^{u(1)}, \dots, \mathbf{x}_k^{u(n_e)}] \in \mathbb{R}^{n_x \times n_e}$ and

$\mathbf{D}_k = [\mathbf{d}_k^{(1)}, \dots, \mathbf{d}_k^{(n_e)}] \in \mathbb{R}^{n_d \times n_e}$ as the state ensemble and data ensembles matrices respectively. For notational convenience, we will from now on omit the subscript k because the focus will be on a single time step.

If we assume that the joint distribution of $(\mathbf{x}^u, \mathbf{d})$ is Gaussian, a classical updating scheme for each ensemble member would be:

$$\mathbf{x}^{c(i)} = \mathbf{x}^{u(i)} + \hat{\mathbf{K}}(\mathbf{d} - \mathbf{d}^{(i)}), \quad (3.4)$$

where

$$\hat{\mathbf{K}} = \mathbf{X} \mathbf{H} \mathbf{D}^T (\mathbf{D} \mathbf{H} \mathbf{D}^T)^{-1} \in \mathbb{R}^{n_x \times n_d}. \quad (3.5)$$

We refer to this as the classical EnKF updating scheme, where we denote the estimated Kalman gain matrix by $\hat{\mathbf{K}}$. Here

$$\mathbf{H} = \mathbf{I} - \frac{1}{n_e} \mathbf{1} \mathbf{1}^T \in \mathbb{R}^{n_e \times n_e} \quad (3.6)$$

is the idempotent centring matrix, where \mathbf{I} is the identity matrix, and $\mathbf{1}$ is a vector with each entry equal to one, both having proper dimensions. Under the Gaussian assumption stated above, $\mathbf{x}_k^{c(i)}$ will tend towards a realisation from the Gaussian posterior distribution $f(\mathbf{x}_k | \mathbf{d}_{0:k})$ as $n_e \rightarrow \infty$.

From multivariate statistical theory, we know that the estimated Kalman gain matrix is equal to the least squares estimate of the matrix of regression coefficients in a multivariate linear regression setting (Seber and Lee, 2003):

$$\hat{\mathbf{K}} = \arg \min_{\mathbf{K}} \text{tr} \{ (\mathbf{X} \mathbf{H} - \mathbf{K} \mathbf{D} \mathbf{H}) (\mathbf{X} \mathbf{H} - \mathbf{K} \mathbf{D} \mathbf{H})^T \}, \quad (3.7)$$

where $\text{tr}\{\cdot\}$ denotes the trace operator. Hence, it is not surprising that the classical EnKF can perform poorly for highly non-linear functions in the likelihood model.

Note that for situations where $n_d \geq n_e$, the matrix $\mathbf{D} \mathbf{H} \mathbf{D}^T$ in Eq. (3.5) will be singular, meaning that its inverse does not exist. To avoid this problem we can add a positive definite matrix, $(n_e - 1) \boldsymbol{\Sigma}_r$, which corresponds to

adding a regularisation term, $(n_e - 1)\text{tr}\{\mathbf{K}\boldsymbol{\Sigma}_r\mathbf{K}^T\}$, to the objective function in Eq. (3.7). This gives

$$\hat{\mathbf{K}} = \mathbf{X}\mathbf{H}\mathbf{D}^T (\mathbf{D}\mathbf{H}\mathbf{D}^T + (n_e - 1)\boldsymbol{\Sigma}_r)^{-1}, \quad (3.8)$$

which is an extension of the standard EnKF updating scheme to non-linear likelihood models, similar to the approach in Evensen (2007, Appendix A.2). Assuming the following likelihood model, $\mathbf{d} = \boldsymbol{\zeta}(\mathbf{x}^u) + \boldsymbol{\epsilon}_d$, where the independent Gaussian noise term has zero mean and covariance, $\boldsymbol{\Sigma}_r$, the Kalman gain estimate defined in Eq. (3.8) is natural to consider. If we replace the data ensemble matrix used in Eq. (3.8), with an alternative data ensemble matrix, $\tilde{\mathbf{D}} = [\boldsymbol{\zeta}(\mathbf{x}^{u(1)}), \dots, \boldsymbol{\zeta}(\mathbf{x}^{u(n_e)})]$, the estimator

$$\hat{\mathbf{K}} = \mathbf{X}\mathbf{H}\tilde{\mathbf{D}}^T (\tilde{\mathbf{D}}\mathbf{H}\tilde{\mathbf{D}}^T + (n_e - 1)\boldsymbol{\Sigma}_r)^{-1}, \quad (3.9)$$

will indeed be consistent with the estimator defined in Eq. (3.5). This follows because the data ensemble matrix can be split into two parts, $\mathbf{D} = \tilde{\mathbf{D}} + \mathbf{E}$, where $\mathbf{E} = [\boldsymbol{\epsilon}_d^{(1)}, \dots, \boldsymbol{\epsilon}_d^{(n_e)}]$ is error perturbation ensemble matrix. However, in the general model setting with $\mathbf{d} = \boldsymbol{\zeta}(\mathbf{x}, \boldsymbol{\epsilon}_d)$, the two estimators will not be consistent.

Another potential problem with the classical least squared estimate of the Kalman gain is model overfitting. This is especially true in the presence of collinear data (Farrer and Glauber, 1967). Because we couple the updated ensemble members through the estimated Kalman gain matrix, they can become increasingly collinear with time (Myrseth et al., 2010). This can potentially lead to an ensemble collapsing into a single realisation, and certainly lead to an underestimation of the prediction uncertainty (Sætrum and Omre, 2010).

Fortunately, we can handle the problems described above efficiently using kernel shrinkage regression techniques known from multivariate non-linear regression, which we will consider next.

3.4 Kernel-Shrinkage Regression

To motivate the use of kernel methods, we start this section with a probabilistic discussion of a non-linear regression problem with multivariate predictor variables:

$$x = \gamma(\mathbf{d}) + \epsilon_{x|\mathbf{d}}. \quad (3.10)$$

Here x is a univariate, centred response variable; \mathbf{d} is a centred multivariate predictor variable; $\gamma(\mathbf{d})$ is the non-linear regression function; and $\epsilon_{x|\mathbf{d}}$ is a

univariate Gaussian error term having zero mean and variance $\sigma_{x|\mathbf{d}}^2$. For notational simplicity we only consider one-dimensional response variables, because it can be shown (Hastie et al., 2009) that we under certain assumptions obtain corresponding solutions for each component in the multivariate case.

Following Williams (1998), let the non-linear function $\gamma(\cdot)$ be decomposed into L terms

$$x = \boldsymbol{\beta}^T \boldsymbol{\varphi}(\mathbf{d}) + \epsilon_{x|\mathbf{d}},$$

where $\boldsymbol{\beta} = [\beta_1, \dots, \beta_L]^T \sim \text{Gauss}_L(\mathbf{0}, \mathbf{I})$ is an unknown random vector. Here $\mathbf{0}$ is the zero vector of proper dimensions, and the L -dimensional vector $\boldsymbol{\varphi}(\mathbf{d}) = [\varphi_1(\mathbf{d}), \dots, \varphi_L(\mathbf{d})]^T$, is a collection of known link functions. Two examples of link functions are

$$\begin{aligned} \varphi_k(\mathbf{d}) &= a_{0k} + a_{1k} \mathbf{d}_k + a_{2k} \mathbf{d}_k^2 \\ \varphi_k(\mathbf{d}) &= \sigma_{\varphi_k} \exp \left\{ - \sum_{l=1}^{n_d} \frac{(d_l - \tau_{k,l})^2}{\Delta_{k,l}^2} \right\}, \end{aligned} \quad (3.11)$$

$k = 1, \dots, L$, where $a_{i,k}$, σ_{φ_k} , $\tau_{k,l}$ and $\Delta_{k,l}$ are model parameters. Under the assumption of independent observation errors for $\epsilon_{x|\mathbf{d}}$ this entails:

$$\begin{aligned} \mathbb{E}[x] &= 0 \\ \text{Cov}(x, x') &= \boldsymbol{\varphi}(\mathbf{d})^T \boldsymbol{\varphi}(\mathbf{d}') + \delta(\mathbf{d}, \mathbf{d}') \sigma_{x|\mathbf{d}}^2, \end{aligned}$$

with $\delta(\cdot, \cdot)$ being the Dirac function taking value one when the arguments are identical and zero otherwise.

Consider a set of centred realisations $\{(x^{(i)}, \mathbf{d}^{(i)}), i = 1, \dots, n_e\}$ and define the corresponding univariate, centred state vector $\tilde{\mathbf{x}} = [x^{(1)}, \dots, x^{(n_e)}] \in \mathbb{R}^{n_e \times 1}$ with associated centred data ensemble matrix \mathbf{D} . In addition, let \mathbf{d}^* be a new, centred data vector, with unknown centred state variable x^* . Under the Gaussian assumption above we have

$$\begin{bmatrix} x^* \\ \tilde{\mathbf{x}} \end{bmatrix} \sim \text{Gauss}_{1+n_e} \left(\begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \text{Cov}(x^*) & \text{Cov}(x^*, \tilde{\mathbf{x}}) \\ \text{Cov}(\tilde{\mathbf{x}}, x^*) & \text{Cov}(\tilde{\mathbf{x}}) \end{bmatrix} \right),$$

where $\text{Cov}(x^*) = \boldsymbol{\varphi}(\mathbf{d}^*)^T \boldsymbol{\varphi}(\mathbf{d}^*) + \sigma_{x|\mathbf{d}}^2$, $\text{Cov}(x^*, \tilde{\mathbf{x}}) = \boldsymbol{\varphi}(\mathbf{d}^*)^T \boldsymbol{\Phi}$ and $\text{Cov}(\tilde{\mathbf{x}}) = \boldsymbol{\Phi}^T \boldsymbol{\Phi} + \sigma_{x|\mathbf{d}}^2 \mathbf{I}$, with $\boldsymbol{\Phi} = [\boldsymbol{\varphi}(\mathbf{d}^{(1)}), \dots, \boldsymbol{\varphi}(\mathbf{d}^{(n_e)})] \in \mathbb{R}^{L \times n_e}$. The conditional expectation of x^* given $\tilde{\mathbf{x}}$, $\mathbb{E}[x^*|\tilde{\mathbf{x}}]$, minimises the Mean Squared Prediction Error (MSPE), $\mathbb{E}[(x^* - g(\tilde{\mathbf{x}}))^2]$, for any function $g : \mathbb{R}^{n_e} \rightarrow \mathbb{R}$ (Rao, 1973). Hence,

$$\hat{x}^* = \tilde{\mathbf{x}}^T \left(\boldsymbol{\Phi}^T \boldsymbol{\Phi} + \sigma_{x|\mathbf{d}}^2 \mathbf{I} \right)^{-1} \boldsymbol{\Phi}^T \boldsymbol{\varphi}(\mathbf{d}^*) \quad (3.12)$$

is the optimal predictor of x^* in terms of the MSPE.

An alternative formulation of the non-linear regression problem is to parametrise $\gamma(\mathbf{d})$ by a Gaussian Random Field (GRF), as defined in Appendix 3-A, with mean $E[\gamma(\mathbf{d})] = 0$, and covariance function $\text{Cov}(\gamma(\mathbf{d}), \gamma(\mathbf{d}')) = c(\mathbf{d}, \mathbf{d}')$. Using the definition of a GRF together with well-known results from multivariate Gaussian theory, the predictive mean of x^* given $\tilde{\mathbf{x}}$ is:

$$E[x^* | \tilde{\mathbf{x}}] = \tilde{\mathbf{x}}^T \left(\mathbf{C} + \sigma_x^2 \mathbf{I} \right)^{-1} \mathbf{c}^* \quad (3.13)$$

with $\mathbf{c}^* = [c(\mathbf{d}^{(1)}, \mathbf{d}^*), \dots, c(\mathbf{d}^{(n_e)}, \mathbf{d}^*)]^T \in \mathbb{R}^{n_e \times 1}$, and

$$\mathbf{C} = \begin{bmatrix} c(\mathbf{d}^{(1)}, \mathbf{d}^{(1)}) & \dots & c(\mathbf{d}^{(1)}, \mathbf{d}^{(n_e)}) \\ \vdots & \ddots & \vdots \\ c(\mathbf{d}^{(n_e)}, \mathbf{d}^{(1)}) & \dots & c(\mathbf{d}^{(n_e)}, \mathbf{d}^{(n_e)}) \end{bmatrix} \in \mathbb{R}^{n_e \times n_e}.$$

Comparing Eq. (3.12) and (3.13), we see that there is a dual formulation of the non-linear regression problem, where we can interpret inner products between vectors $\boldsymbol{\varphi}(\cdot) \in \mathbb{R}^{L \times 1}$ as covariance functions $c(\cdot, \cdot)$ of a GRF. For the link-function defined in Eq. (3.11) this can be realised by letting $L \rightarrow \infty$ and selecting $\Delta_{k,l} = \Delta$ for all k and l , which entails (Gibbs, 1997):

$$c(\mathbf{d}, \mathbf{d}') = \boldsymbol{\varphi}(\mathbf{d})^T \boldsymbol{\varphi}(\mathbf{d}') = \sigma_\varphi^2 \exp \left\{ -\frac{1}{2\Delta} \|\mathbf{d} - \mathbf{d}'\|_2^2 \right\} \quad (3.14)$$

which we recognise as the second-order exponential covariance function. The assumption that we can describe inner products between vectors $\boldsymbol{\varphi}(\mathbf{d})$ and $\boldsymbol{\varphi}(\mathbf{d}')$ through symmetric, positive definite functions, $c(\mathbf{d}, \mathbf{d}')$, is the foundation of kernel methods known from the machine learning literature (Smola and Schölkopf, 2002).

3.4.1 Kernel Methods

We now generalise the mapping $\boldsymbol{\varphi}(\cdot)$, and let $\boldsymbol{\varphi} : \mathbb{R}^{n_d} \rightarrow \mathcal{F}$, where \mathcal{F} is some unspecified inner product space (Young, 1988), which we for simplicity will refer to as the feature space. In the non-linear regression setting this entails transforming the predictor variable, \mathbf{d} , into a feature space where the linear relationship with the response variable, x , is valid, as described in Fig. 3.2.

In Eq. (3.12), we express \hat{x}^* through inner products in the feature space, which we can equally describe through values of the covariance function, $c(\cdot, \cdot)$, under the GRF assumption. In general, this will hold for any symmetric, positive definite function, $c(\cdot, \cdot)$, also known as kernel functions (Smola

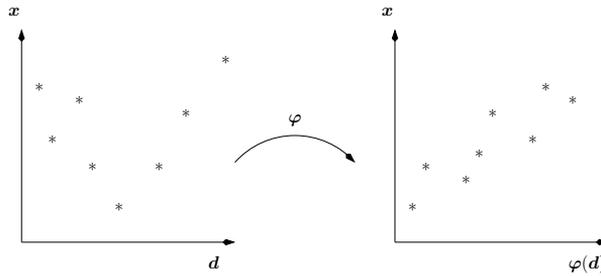


Figure 3.2: Transformation of \mathbf{d} into a feature space \mathcal{F} , where there is a linear relationship with the response \mathbf{x} .

and Schölkopf, 2002). In the machine learning literature, this is known as the "kernel-trick" (Smola and Schölkopf, 2002), which in practise means that we can reformulate any algorithm involving inner products in the input space (here \mathbb{R}^{n_d}) into a feature space \mathcal{F} , using kernel functions. The formal conditions for which the assumption, $c(\mathbf{d}, \mathbf{d}') = \langle \boldsymbol{\varphi}(\mathbf{d}), \boldsymbol{\varphi}(\mathbf{d}') \rangle$ holds are described in Hofmann et al. (2008), namely symmetry and positive definiteness. The interested reader can find further details in Smola and Schölkopf (2002) or Taylor and Cristianini (2004).

Sætrom and Omre (2010), reformulated the classical EnKF updating scheme using Principal Component Regression (PCR) (Hotelling, 1933) and Partial Least Squares Regression (PLSR) (Wold, 1975), known from shrinkage regression. Both methods can be transformed into a non-linear setting using kernelised versions; see Rosipal et al. (2001) and Rosipal and Trejo (2002) for PCR and PLSR respectively. Common to the kernel versions of the three shrinkage regression techniques is that we perform a dimension reduction in the feature space. Moreover, given a new observation vector, \mathbf{d}^* , we predict the state variable, \mathbf{x}^* , by evaluating the kernel function. Hence, the methods does not require an inverse mapping, avoiding the pre-image problem occurring when using the kernel approach on the state variables (Caers and Park, 2008; Sarma and Chen, 2009).

3.4.2 Comments

An early reference to the Bayesian formulation of the non-linear regression problem is O'Hagan (1978). However, the connection with kernel methods known from the machine learning literature was not realised until recently (Rasmussen and Williams, 2006; Williams, 1998). Moreover, readers familiar with Geostatistics will surely recognise the connection between the

probabilistic formulation above and a generalisation of kriging using non-stationary covariance models (Chiles and Delfiner, 1999; Journel and Huijbregts, 1978).

The decoupling of the state vector into univariate components, as done in the decomposition above, is theoretically only valid if the elements in the stochastic error term, $\epsilon_{\mathbf{x}|\mathbf{d}}$, are independent, which implies that the covariance $\Sigma_{\epsilon_{\mathbf{x}|\mathbf{d}}}$ is proportional to the identity matrix, \mathbf{I} . Note, however, that the methods considered in this paper will not be affected by this model assumption because we are solving the regression problem using a least squares approach. This entails that the obtained regression models are independent of the model parameter $\Sigma_{\epsilon_{\mathbf{x}|\mathbf{d}}}$ (Hastie et al., 2009).

3.5 Ensemble Kalman Filtering Using Kernel-Shrinkage Regression

We will now present the kernelised versions of the EnKF updating scheme based on three common shrinkage regression techniques; Ridge Regression (RR) (Hoerl and Kennard, 1970), PCR and PLSR. However, we will only consider kernel RR in detail. The interested reader can find a thorough description of the PCR and PLSR shrinkage regression techniques in Sætrum and Omre (2010). For simplicity, we define kernel matrices

$\mathbf{C}^* = [\mathbf{c}^*, \dots, \mathbf{c}^*] \in \mathbb{R}^{n_e \times n_e}$, with $\mathbf{c}^* = [c(\mathbf{d}^{(1)}, \mathbf{d}^*), \dots, c(\mathbf{d}^{(n_e)}, \mathbf{d}^*)]^T \in \mathbb{R}^{n_e \times 1}$, and

$$\mathbf{C} = \begin{bmatrix} c(\mathbf{d}^{(1)}, \mathbf{d}^{(1)}) & \dots & c(\mathbf{d}^{(1)}, \mathbf{d}^{(n_e)}) \\ \vdots & \ddots & \vdots \\ c(\mathbf{d}^{(n_e)}, \mathbf{d}^{(1)}) & \dots & c(\mathbf{d}^{(n_e)}, \mathbf{d}^{(n_e)}) \end{bmatrix}. \quad (3.15)$$

For all three methods, we find the estimated Kalman gain matrix by solving the following multivariate linear regression problem in a feature space, \mathcal{F} :

$$\mathbf{x} = \mathbf{K}\varphi(\mathbf{d}) + \epsilon_{\mathbf{x}|\mathbf{d}},$$

where $\epsilon_{\mathbf{x}|\mathbf{d}} \in \mathbb{R}^{n_x \times 1}$ represents the regression model error. Similar to above, $\Phi = [\varphi(\mathbf{d}^{(1)}), \dots, \varphi(\mathbf{d}^{(n_e)})]$, and we assume that all vectors and ensemble matrices are centred, unless otherwise stated.

3.5.1 Kernel Ridge Regression

Kernel Ridge Regression (RR) is a regularisation method where we select the estimated regression coefficients by minimising the mean squared error

with additional constraints, that is:

$$\hat{\mathbf{K}}_{\text{KerRR}} = \arg \min_{\mathbf{K}} \left\{ \text{tr}\{(\mathbf{X} - \mathbf{K}\Phi)(\mathbf{X} - \mathbf{K}\Phi)^T\} + \xi \text{tr}\{\mathbf{K}\mathbf{K}^T\} \right\}.$$

Solving this minimisation problem analytically gives (Seber and Lee, 2003):

$$\begin{aligned} \hat{\mathbf{K}}_{\text{KerRR}} &= \mathbf{X}\Phi^T (\Phi\Phi^T + \xi\mathbf{I})^{-1} \\ &= \mathbf{X} (\Phi^T\Phi + \xi\mathbf{I})^{-1} \Phi^T. \end{aligned}$$

Thus, we obtain kernel RR predictions for an unknown \mathbf{x}^* , with associated \mathbf{d}^* , based on

$$\mathbf{x}^* = \mathbf{X} (\mathbf{C} + \xi\mathbf{I})^{-1} \mathbf{c}^* \tag{3.16}$$

Comparing this expression with Eq. (3.13), we see that the two predictors, \mathbf{x}^* and $E[\mathbf{x}^*|\mathbf{X}]$ are identical for $\sigma_{x|\mathbf{d}}^2 = \xi$. This follows because each variable x_i^* , for $i = 1, \dots, n_x$, can be computed independently. Thus, applying kernel RR in an EnKF setting leads to the following updating scheme:

$$\mathbf{X}^c = \mathbf{X}^u + \mathbf{X}^u \mathbf{H} (\mathbf{H}\mathbf{C}\mathbf{H} + \xi\mathbf{I})^{-1} \mathbf{H}(\mathbf{C}^* - \mathbf{C}). \tag{3.17}$$

Here \mathbf{X}^c and \mathbf{X}^u are the state ensemble matrices, conditioned and unconditioned respectively and $\mathbf{H}\mathbf{C}\mathbf{H}$ is the centred kernel matrix (Schölkopf et al., 1998), with the centring matrix, \mathbf{H} , defined in Eq. (3.6).

3.5.2 Kernel Principal Component Regression

PCR is based on the assumption that most of the variability in the predictor variables can be explained through a small set of random variables, termed principal components. The estimated matrix of regression coefficients is then constructed using the principal components as predictor variables. Hence, we define a matrix of regression coefficients in a reduced order space. Kernel PCR follows directly by applying kernel PCA (Rosipal et al., 2001; Schölkopf et al., 1998) to $\varphi(\mathbf{d})$. The resulting expression for the estimated Kalman gain matrix is

$$\hat{\mathbf{K}}_{\text{KerPCR}} = \mathbf{X}\mathbf{E}_p\mathbf{\Lambda}_p^{-1}\mathbf{E}_p^T\Phi^T. \tag{3.18}$$

Here $\mathbf{E}_p \in \mathbb{R}^{n_e \times p}$ contains the p eigenvectors of the centred kernel matrix, $\mathbf{H}\mathbf{C}\mathbf{H}$, with the p largest corresponding eigenvalues given in the diagonal matrix, $\mathbf{\Lambda}_p \in \mathbb{R}^{p \times p}$. The following expression then forms an EnKF updating scheme based on kernel PCR:

$$\mathbf{X}^c = \mathbf{X}^u + \mathbf{X}^u \mathbf{E}_p \mathbf{\Lambda}_p^{-1} \mathbf{E}_p^T (\mathbf{C}^* - \mathbf{C}). \tag{3.19}$$

3.5.3 Kernel Partial Least Squares Regression

Similar to PCR, PLSR applies dimension reduction techniques to the predictor variables. The main difference between the two methods is that, PLSR uses the information available from both the response and predictor variables when projecting into the p -dimensional subspace, which entails that PLSR is based on a supervised dimension reduction technique. The kernelised version of the PLSR algorithm follows from the algorithm presented in R anner et al. (1994). This results in the following estimate of the Kalman gain matrix:

$$\hat{\mathbf{K}}_{\text{KerPLSR}} = \mathbf{X} \mathbf{T} \mathbf{A}^{-1} \mathbf{W}^T \Phi^T.$$

Here $\mathbf{A} = (\mathbf{W}^T \mathbf{C} \mathbf{T}) \in \mathbb{R}^{p \times p}$, with latent variables $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_p] \in \mathbb{R}^{n_e \times p}$ and $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_p] \in \mathbb{R}^{n_e \times p}$ given by solving sequentially for $i = 1, \dots, p$:

$$\begin{bmatrix} \mathbf{t}_i = \mathbf{H} \Phi^T \psi_i \\ \mathbf{w}_i = \mathbf{H} \mathbf{X}^T \mathbf{v}_i \end{bmatrix} \leftarrow \begin{cases} \max_{\psi_i, \mathbf{v}_i} \{ \mathbf{v}_i^T \mathbf{X} \mathbf{H} \Phi^T \psi_i \} \\ \|\psi_i\|_2 = 1, \|\mathbf{v}_i\|_2 = 1 \\ \mathbf{t}_i^T \mathbf{t}_j = 0, \text{ for all } j < i. \end{cases}$$

R anner et al. (1994) outline an efficient procedure for solving this problem when n_x and n_d are larger than n_e . This gives the following EnKF updating scheme based on kernel PLSR:

$$\mathbf{X}^c = \mathbf{X}^u + \mathbf{X}^u \mathbf{T} \mathbf{A}^{-1} \mathbf{W}^T (\mathbf{C}^* - \mathbf{C}), \quad (3.20)$$

with \mathbf{A} defined above. Note that centring of the ensemble matrices is unnecessary because of the identities $\mathbf{T} = \mathbf{H} \mathbf{T}$ and $\mathbf{W} = \mathbf{H} \mathbf{W}$ (S etrom and Omre, 2010).

3.5.4 Comments

Although the three shrinkage regression techniques presented above, and their kernelised versions, all reduce the problems caused by collinearities in the data ensemble, they accomplish this differently. Note that collinearities in the data ensemble will result in small eigenvalues in the estimated data covariance matrix, $\hat{\Sigma}_d$, or the corresponding centred kernel matrix, $\mathbf{H} \mathbf{C} \mathbf{H}$, which can result in large weights in the respective Kalman gain matrices. Whilst the RR technique reduces the weight caused by smallest eigenvalues by adding a positive constant to all eigenvalues, the PCR technique eliminates small eigenvalues, whilst retaining the p dominant. The PLSR technique works in a similar manner as the PCR, although the analysis is

slightly more complicated (see Hastie et al. (2009) for a detailed description). However, it has been noted that because the PLSR technique uses the information from both the state and data ensemble matrices in the dimension reduction, a small number of components, p , is often required in PLSR compared with PCR (Kalivas, 1999; Helland, 2001).

It should be noted that the optimality of the different shrinkage regression techniques appears to be problem dependent. That is, whilst RR leads to the smallest prediction error in one study, PCR or PLSR can be optimal in other studies. It is therefore advisable to test the predictive performance of the different methods based on the prior and likelihood model, before selecting which method to use in an EnKF setting. Finally it is important to note that the performance of all three schemes are highly dependent on the model hyperparameters used. Thus, schemes which enables an automatic selection of these model parameters are called for.

3.5.5 Model Hyperparameter Selection

The results obtained using the kernel-shrinkage regression techniques described above, will in general depend on one or more hyperparameters, θ . For kernel RR this involves selecting the size of the regularisation parameter, ξ , whilst for kernel PCR and PLSR the dimension of the reduced order space, p , is required. In addition, we need to select the kernel function, $c(\cdot, \cdot)$, which implicitly defines $\varphi(\cdot)$. This can be challenging if all symmetric, positive definite kernel functions are considered.

The usual approach for selecting $c(\cdot, \cdot)$ is to consider only certain families of kernel functions. An example is translation and rotation invariant functions $c(\mathbf{d}, \mathbf{d}') = c(\|\mathbf{d} - \mathbf{d}'\|_2^2)$, referred to as Radial Basis Function (RBF) kernels, where $\|\cdot\|_2$ is the Euclidean norm (Smola and Schölkopf, 2002). The second order exponential kernel function, defined in Eq. (3.11), is an example of such a kernel function. Another example of a kernel function is the polynomial, defined as $c(\mathbf{d}, \mathbf{d}') = (1 + \mathbf{d}^T \mathbf{d}')^\nu$, $\nu > 1$, which entails that $\varphi(\cdot)$ represents polynomials. Common for most of the kernel functions in the literature, however, is that we only need to specify one or two hyperparameters.

To avoid overfitting the regression model to the data, we can apply Cross-Validation (CV), as discussed in Sætrum and Omre (2010). CV works by sequentially splitting the state vector and data ensembles into training ensemble matrices, $(\mathbf{X}_{\text{Train}}, \mathbf{D}_{\text{Train}})$, used for model estimation, and test ensembles matrices, $(\mathbf{X}_{\text{Test}}, \mathbf{D}_{\text{Test}})$, used for model validation. Using the training ensemble matrices to estimate the Kalman gain matrix for model parameters θ , $\hat{\mathbf{K}}_{\text{Train}}(\theta)$, we can predict the state vectors based on the test

Algorithm 1: Workflow for the EnKF updating scheme based on kernel PCR.

Input: Ensemble matrices, (\mathbf{X}, \mathbf{D}) ; set of hyperparameters, $\Theta = \{\theta_1, \dots, \theta_r\}$; number CV folds, m

- 1 Randomly select indices I_1, \dots, I_m so that $\cup_{i=1}^m I_i = \{1, \dots, n_e\}$ and $\cap_{i=1}^m I_i = \emptyset$
- 2 **for** $i = 1$ to r **do**
- 3 | PRESS(i) = 0
- 4 **for** $i = 1$ to m **do**
- 5 | **for** $j = 1$ to r **do**
- 6 | | Estimate $\hat{\mathbf{K}}_{\text{KerPCR}}(\theta_j)$ using Eq. (3.18), based on centred ensemble members $(\mathbf{x}^{(k)}, \mathbf{d}^{(k)})$, for indices $k \notin I_i$
- 7 | | **for** $k \in I_i$ **do**
- 8 | | | Predict: $\hat{\mathbf{x}}^{(k)} = \hat{\mathbf{K}}_{\text{KerPCR}}(\theta_j)\varphi(\mathbf{d}^{(k)})$
- 9 | | | PRESS(j) + = $\|\hat{\mathbf{x}}^{(k)} - \mathbf{x}^{(k)}\|_2^2$
- 10 $j^* = \arg \min_j \text{PRESS}(j)$
- 11 Estimate $\hat{\mathbf{K}}_{\text{KerPCR}}(\theta_{j^*})$ using Eq. (3.18), based on full ensemble matrices \mathbf{X}, \mathbf{D}
- 12 EnKF update using Eq. (3.19)

data, $\hat{\mathbf{X}}_{\text{Test}} = \hat{\mathbf{K}}_{\text{Train}}(\theta)\mathbf{D}_{\text{Test}}$. Computing the sum of squares of the mismatch between $\hat{\mathbf{X}}_{\text{Test}}$ and \mathbf{X}_{Test} , referred to as the Predictive Error Sum of Squares (PRESS) statistic, can be used to measure of the predictive power for the chosen model. Note, however, that it is straightforward to use other objective functions than the sum of squares. We can then select the model hyperparameters by minimising the PRESS statistic for different values of θ , as discussed below. A pseudo code describing the workflow of the EnKF updating scheme based on kernel PCR, using m -fold CV to select the model parameters is given in Algorithm 1. Note that we here use the notation I_i to denote the set of indices for members of test ensemble i . A similar workflow can be used for the RR and PLSR techniques as well.

The model parameters, θ , will for most applications of kernel-shrinkage regression contain two parameters, which can be both discrete and continuous. Alternative methods for solving this problem are gradient based optimisation techniques (Nocedal and Wright, 2006), or stochastic optimisation techniques such as particle swarm optimisation (Kennedy and Eberhart, 1995) and simulated annealing (Kirkpatrick et al., 1983). In practise, how-

ever, it is often sufficient to consider a limited number discrete values for the continuous model parameters, for example $\Delta \in \{1, 2, \dots, 20\}$ for the second order exponential covariance function, rather than searching for the global optimum (Smola and Schölkopf, 2002). Hence, we do not necessarily suffer from high computational demands if we select the model parameters based on an exhaustive search rather than sophisticated optimising schemes. Also note that for applications of the EnKF where it is extremely time consuming to evaluate the forward model $\omega(\cdot)$, such as petroleum reservoir characterisation, the use of CV for model parameter selection will not increase the total computational time significantly. In addition, we emphasise that the CV scheme is straightforward to run in parallel. A discussion regarding the computational properties of the EnKF updating scheme based on kernel-shrinkage regression techniques can be found in Appendix 3-B.

Finally, because the likelihood model is known, selection of the kernel function is easier in this setting, compared to non-linear regression problems where the relation between \mathbf{d} and \mathbf{x} is unknown. An alternative approach is therefore to select the kernel function based on prior knowledge of the state vector and the likelihood model. That is, we select the hyperparameters of the kernel function based on the predictive power of the non-linear regression model created using realisations from the initial prior model.

3.6 Empirical Study

We define the state vector $\mathbf{x}_k \in \mathbb{R}^{100 \times 1}$, $k = 0, \dots, 19$ on a (10×10) regular grid domain. Here $x_{i,j,k}$ denotes the value of the state vector at location (i, j) , $i = 1, \dots, 10$, $j = 1, \dots, 10$ and time step k . We assume that the state vector is static, meaning $\mathbf{x}_{k+1} = \mathbf{x}_k$, $k = 0, \dots, 19$, with a reference model generated from the Gaussian prior model:

$$\mathbf{x}^{\text{True}} \sim \text{Gauss}_{n_x}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x).$$

Here $\boldsymbol{\mu}_x = 5 \times \mathbf{1}$ and we construct $\boldsymbol{\Sigma}_x$ based on an exponential covariance function,

$$\text{Cov}(x_{i,j,0}, x_{l,m,0}) = \exp \{ -3(\Delta)^{1.2} \}, \tag{3.21}$$

where

$$\Delta = \sqrt{\left(\frac{\Delta_x}{l_x}\right)^2 + \left(\frac{\Delta_y}{l_y}\right)^2}.$$

The range parameters l_x and l_y are one and 10 respectively.

We use the non-linear likelihood model:

$$\mathbf{d}_k = \zeta(\mathbf{x}, \boldsymbol{\epsilon}_d), \tag{3.22}$$

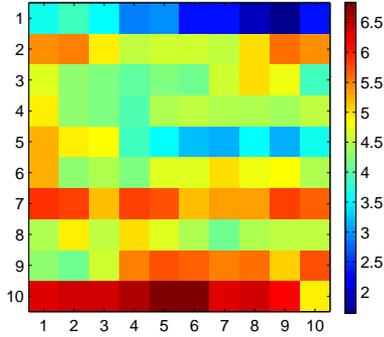


Figure 3.3: Reference state vector used in the empirical case study generated from a Gaussian prior distribution with an anisotropic covariance function.

to connect the observations \mathbf{d}_k to the state vector. Here each element $\zeta_{i,j,k} = g(x_{i,j,k}, 5) + \sigma_d \epsilon_{i,j,k}^2$, with $g(\cdot)$ defined as:

$$g(x, \beta) = \begin{cases} \frac{\sin(|x-\beta|) \cos(d)}{|x-\beta|}, & x \neq \beta \\ 1, & x = \beta. \end{cases}, \quad (3.23)$$

and $\epsilon_{i,j,k} \sim \text{Gauss}_1(0, 1)$. Hence, we have an additive error term $\epsilon_d/\sigma_d \sim \chi_1^2$, where χ_ν^2 denotes the χ^2 -distribution with ν -degrees of freedom (Casella and Berger, 2002), which makes the likelihood function both non-Gaussian and non-linear. Here $\sigma_d = 0.25$, which implies that the mean and variance of the error term are equal to 0.250 and 0.125 respectively. Fig. 3.3 shows an image plot of the reference state vector. The scatter plot of \mathbf{x}^{True} and the corresponding observed data at the initial time step, \mathbf{d}_0 , in Fig. 3.4, displays the non-linear structure of the likelihood model. At each time step $k = 0, \dots, 19$, we make new observations, \mathbf{d}_k .

3.6.1 Non-Linear Regression

To demonstrate the effect of the suggested kernel regression techniques we initially consider a univariate non-linear regression problem, with a bivariate predictor variable, based on the function defined in Eq. (3.23) as follows: Define the data ensemble matrix, \mathbf{D} , through $n_e = 91$ uniformly spaced points, $\mathbf{d} \in \mathbb{R}^{2 \times 1}$, with d_1 and d_2 having values between -4 and 4. Assume that at each of these 91 locations, we observe values of the function,

$$x = \gamma(\mathbf{d}) + \epsilon_{x|\mathbf{d}} = 1 + g(d_1, 0) + g(d_2, 0) + \epsilon_{x|\mathbf{d}},$$

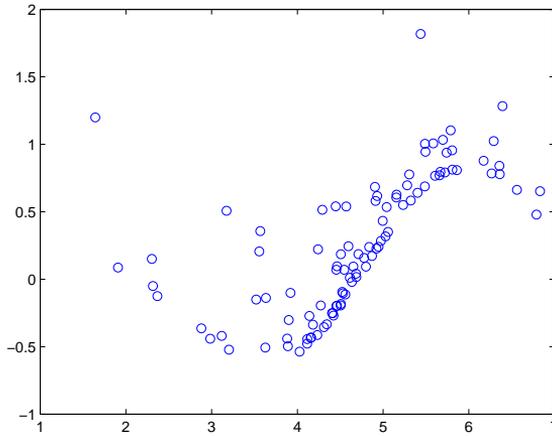


Figure 3.4: Scatter plot between the elements of \mathbf{x}^{True} and corresponding observations, \mathbf{d}^o , based on the non-linear, non-Gaussian likelihood model.

where the noise term, $\epsilon_{x|\mathbf{d}}$, is Gaussian having zero mean and standard deviation 0.5. Fig. 3.5 and 3.6 display the reference solution and the observed data.

Using the ensemble matrices \mathbf{D} and \mathbf{X} , the task is to fit a regression model based on the classical linear least squares approach and the three kernel-shrinkage regression models discussed above. Because the Gaussian kernel function, described in Eq. (3.14), is a robust choice when no prior information is available regarding the data (Smola, 1998; Smola et al., 1998), we use it to map \mathbf{d} into the feature space \mathcal{F} . To select the scaling factor in the kernel function and shrinkage factor for the three regression methods, we apply 10-fold CV. Fig. 3.7 display the results. Note that because the prior and likelihood models are fully specified, another alternative is to tailor the kernel function based on this known prior information. How this task can be accomplished in an automatic manner is, however, somewhat unclear and is a topic for further research. Hence, we apply the Gaussian kernel function for the case studies considered in this paper, even if this selection might be sub-optimal.

As expected, the plane resulting from the linear least squares method, does not provide a good prediction of the non-linear function. The three kernel methods, on the other hand, are able to capture the non-linear trend of the function, thus providing reasonable predictions. Fig. 3.8 display the absolute deviance between the reference and predicted solution for the

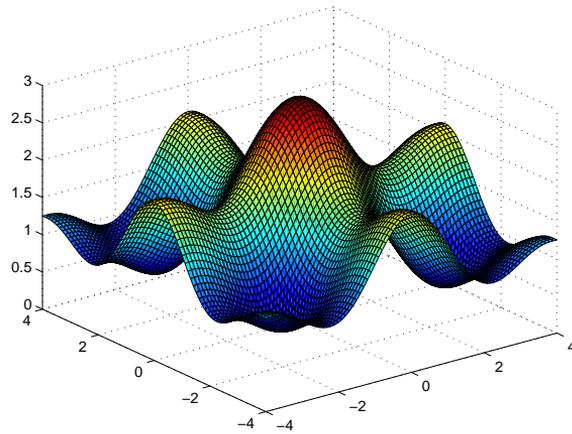


Figure 3.5: Reference solution for the non-linear regression problem considered, $x = \gamma(\mathbf{d})$, for locations, \mathbf{d} in the plane.

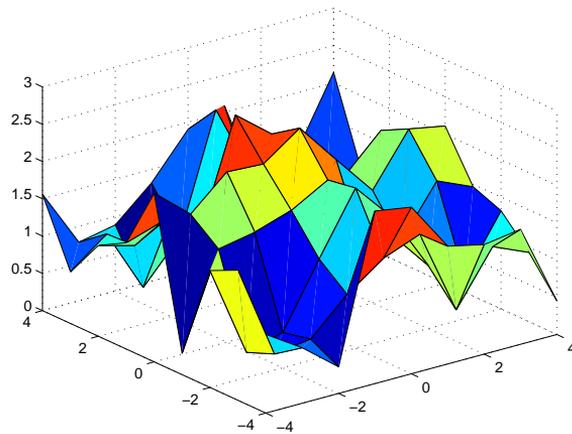


Figure 3.6: Observed data of the function $x = \gamma(\mathbf{d}) + \epsilon_{x|\mathbf{d}}$ at 81 uniformly spaced locations \mathbf{d} in the plane.

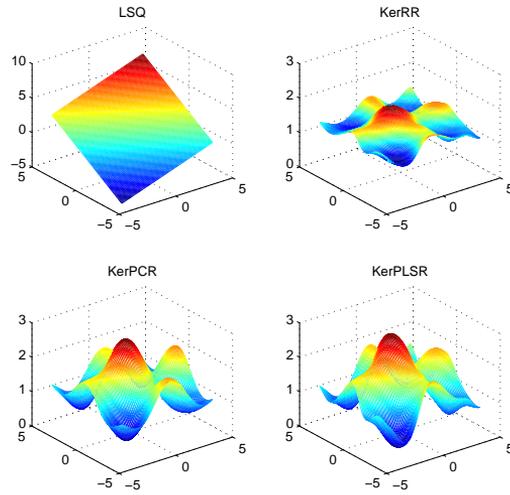


Figure 3.7: Predicted outcome based on least squares linear regression (LSQ), kernel RR (KerRR), kernel PCR (KerPCR) and kernel PLSR (KerPLSR). 10-fold CV was used to select the prior hyperparameters θ .

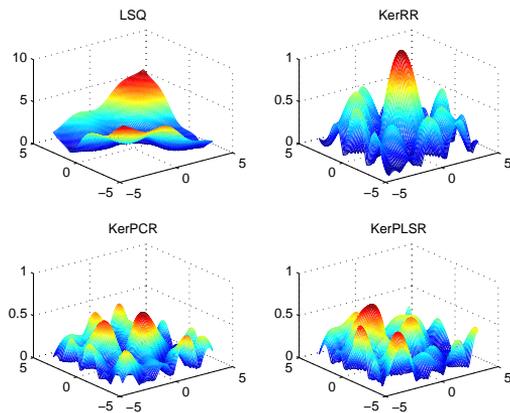


Figure 3.8: Absolute deviance between the true solution and the predicted outcome based on least squares linear regression (LSQ), kernel RR (KerRR), kernel PCR (KerPCR) and kernel PLSR (KerPLSR). 10-fold CV was used to select the prior hyperparameters θ .

four methods. Except for the larger error occurring at the origin for the kernel RR method, the deviance is within one standard deviation of the observation error at all predicted locations using the three kernel methods. The main difference between the PCR and PLSR approaches is, however, that the former requires $p = 13$ components, whilst the latter only requires $p = 1$. Note that this behaviour is in accordance with the discussion above, because PLSR uses information from both the state- and data vector in the dimension reduction.

3.6.2 Results

Based on the results seen in the non-linear regression based on the likelihood model, we proceed using the EnKF based on kernel PCR and PLSR techniques on the filtering problem with the Gaussian kernel function defined in Eq. (3.14). We use the following sampling scheme to generate the initial ensemble with $n_e = 100$:

$$\begin{aligned} \mathbf{x}^{u(1)} &= 10 \times \mathbf{1} \\ \mathbf{x}^{u(2)}, \dots, \mathbf{x}^{u(25)} &\sim \text{Gauss}_{n_x} \left(\frac{1}{1.5} \boldsymbol{\mu}_x, 2\boldsymbol{\Sigma}_x \right) \\ \mathbf{x}^{u(26)}, \dots, \mathbf{x}^{u(50)} &\sim \text{Gauss}_{n_x} \left(\frac{1}{2} \boldsymbol{\mu}_x, 2\boldsymbol{\Sigma}_x \right) \\ \mathbf{x}^{u(51)}, \dots, \mathbf{x}^{u(75)} &\sim \text{Gauss}_{n_x} (2\boldsymbol{\mu}_x, 2\boldsymbol{\Sigma}_x) \\ \mathbf{x}^{u(76)}, \dots, \mathbf{x}^{u(100)} &\sim \text{Gauss}_{n_x} (1.5\boldsymbol{\mu}_x, 2\boldsymbol{\Sigma}_x), \end{aligned}$$

with $\boldsymbol{\mu}_x$ and $\boldsymbol{\Sigma}_x$ defined above. Hence, there is a high uncertainty regarding the mean of the true underlying distribution. The purpose of this study is to assimilate observed data using the standard EnKF updating scheme for non-linear observations (Evensen, 2007, Appendix A.2) and the EnKF updating schemes based on kernel-shrinkage regression. Because $n_d = n_e$, $\hat{\boldsymbol{\Sigma}}_d$ will be singular. We therefore add a positive definite regularisation term before inverting the matrix, as explained in Section 3.3, with $\boldsymbol{\Sigma}_r = 2\sigma_d^2 \mathbf{I}$.

When selecting the prior the hyperparameters in the kernel-shrinkage regression techniques, $\boldsymbol{\theta}_k$, we consider the following two approaches:

- Automatic: $\boldsymbol{\theta} = (\sigma^2, p)$ selected based on 10-fold CV at each time step, minimising the PRESS statistic.
- Supervised: $\boldsymbol{\theta} = (\sigma^2, p)$ selected based is based on 10-fold CV at the initial time step, minimising the PRESS statistic, and remain fixed for all time steps $k = 0, \dots, 19$.

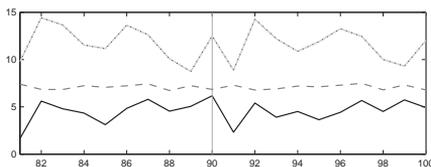
Hence, we consider five different EnKF updating schemes in total:

- Classical EnKF: The non-linear formulation of the EnKF updating scheme using the estimated Kalman gain matrix given in Eq. (3.9), assuming a Gaussian additive noise term: $\mathbf{d} = \boldsymbol{\zeta}(\mathbf{x}) + \boldsymbol{\epsilon}_d$, where $\boldsymbol{\epsilon}_d \sim \text{Gauss}_{n_d}(\mathbf{0}, 2\sigma_d^2\mathbf{I})$. This implies that the covariance of the noise term, $\boldsymbol{\Sigma}_r$, is equal to the true variance of the likelihood model (Casella and Berger, 2002), although we do not account for the non-zero mean.
- EnKF-KerPLSR1: EnKF based on the kernelised PLSR method. The dimension of the reduced order space, and the scaling parameter in the Gaussian kernel function selected based on the automatic approach
- EnKF-KerPCR1: EnKF based on the kernelised PCR method. The dimension of the reduced order space, and the scaling parameter in the Gaussian kernel function selected based on the automatic approach
- EnKF-KerPLSR2: EnKF based on the kernelised PLSR method. The dimension of the reduced order space, and the scaling parameter in the Gaussian kernel function selected based on the supervised approach.
- EnKF-KerPCR2: EnKF based on the kernelised PCR method. The dimension of the reduced order space, and the scaling parameter in the Gaussian kernel function selected based on the supervised approach

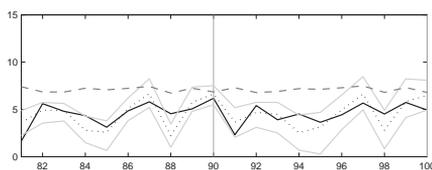
Fig. 3.9a through e, displays the results obtained in grid nodes 81 through 100 using the five different EnKF updating schemes. The results are similar for the remaining locations, and are therefore not included. Note that the initial ensembles are identical for all five schemes.

As we can see from Fig. 3.9a, the classical EnKF is not able to get a good representation of the state vector at the final updating step. The estimated posterior mean is farther from the reference state vector than at the initial time step and the ensemble has collapsed into a single realisation. Because we are using a linear updating scheme on a non-linear likelihood model, we expect that the ensemble mean is missing the reference state vector. However, it is more troubling that we are not able to obtain estimates of the prediction uncertainty. Ideally, the estimated prediction interval based on the updated ensemble, should cover 95% of the reference state vector. In this case, however, the coverage is zero percent.

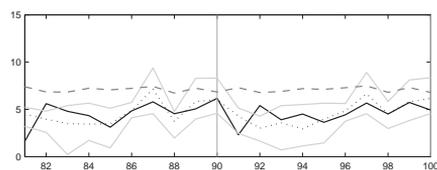
The EnKF-KerPLSR1 method is able to obtain reasonable estimates of the reference state vector, \mathbf{x}^{True} . The estimated posterior mean is centred around \mathbf{x}^{True} , and the estimated prediction interval is giving a good description of the solution uncertainty. The EnKF-KerPCR1, EnKF-KerPLSR2



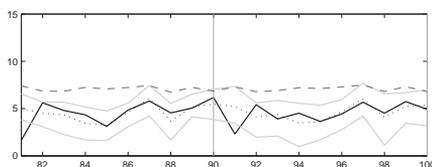
(a) Classical EnKF



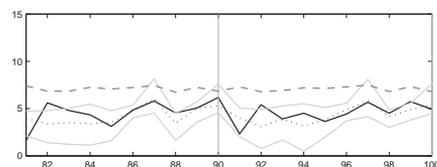
(b) EnKF-KerPLSR1



(c) EnKF-KerPCR1



(d) EnKF-KerPLSR2



(e) EnKF-KerPCR2

Figure 3.9: Results obtained when running five different EnKF updating schemes on a problem with a highly non-linear likelihood model. The figure displays, for grid nodes 81 through 100, the reference $\mathbf{x}_{10}^{\text{True}}$ (solid), the ensemble average (dotted) and the estimated 95% confidence bounds of the prediction interval (solid, light gray) at the final time step. The ensemble mean at the initial time step is shown as the dashed, dark gray line.

and EnKF-KerPCR2 updating schemes have a similar behaviour, as shown in Fig. 3.9c to e.

Fig. 3.10 contains two realisations of the state vector in the initial ensemble and the corresponding realisations after the final updating step, using the five different EnKF updating schemes. We also display the ensemble mean. At the initial time step, the realisations and the ensemble mean do not resemble the reference state vector, as expected from the prior distribution described above. Again, we notice that the realisations obtained using the Classical EnKF scheme has collapsed, and are farther from \mathbf{x}^{True} , than they were initially. The four kernel-shrinkage regression techniques, on the other hand, appear to give a good representation of the reference state vector. We especially note that the anisotropic behaviour in \mathbf{x}^{True} is captured in both the realisations and the ensemble mean for all of the four schemes based on kernel-shrinkage regression.

It is interesting to note that the results obtained using the automatic and supervised methods to select $\boldsymbol{\theta}$ produces similar results. For the supervised parameter selection approach $\boldsymbol{\theta}$ was selected as: $\boldsymbol{\theta}_{\text{PCR}} = (10, 5)$ and $\boldsymbol{\theta}_{\text{PLSR}} = (10, 1)$, whilst for the CV based selection scheme $\boldsymbol{\theta}_{\text{PCR}} \in \{[1, 7](5), [5, 30](20)\}$ and $\boldsymbol{\theta}_{\text{PLSR}} \in \{[1, 5](1), [5, 30](20)\}$. Here the notation $[1, 5](5)$ is used to denote the smallest, largest and median value of the selected parameters respectively.

To quantify the performance of the five updating schemes the model is rerun 100 times using different initial ensembles. We then compute the scaled Total Sum of Squares (TSS), given as $1/n_e \sum_{i=1}^{n_e} \|\hat{\boldsymbol{\mu}}_{\mathbf{x}^c} - \mathbf{x}^{\text{True}}\|_2^2$, and the coverage of the reference solution within the estimated 95% prediction intervals. The results are summarised in Table 3.1. Here we have included the results obtained using PLSR regression in the Euclidean space, rather than the feature space, selecting the subspace dimension p using 10-fold CV (EnKF-PLSR1). As we can see from this table, the TSS increase with 412 percent from the initial updating step when using the classical EnKF updating scheme. On average, the prediction interval only covers 6.2 percent of the reference state vector. Note, however, that the coverage for the majority of the reruns is zero percent.

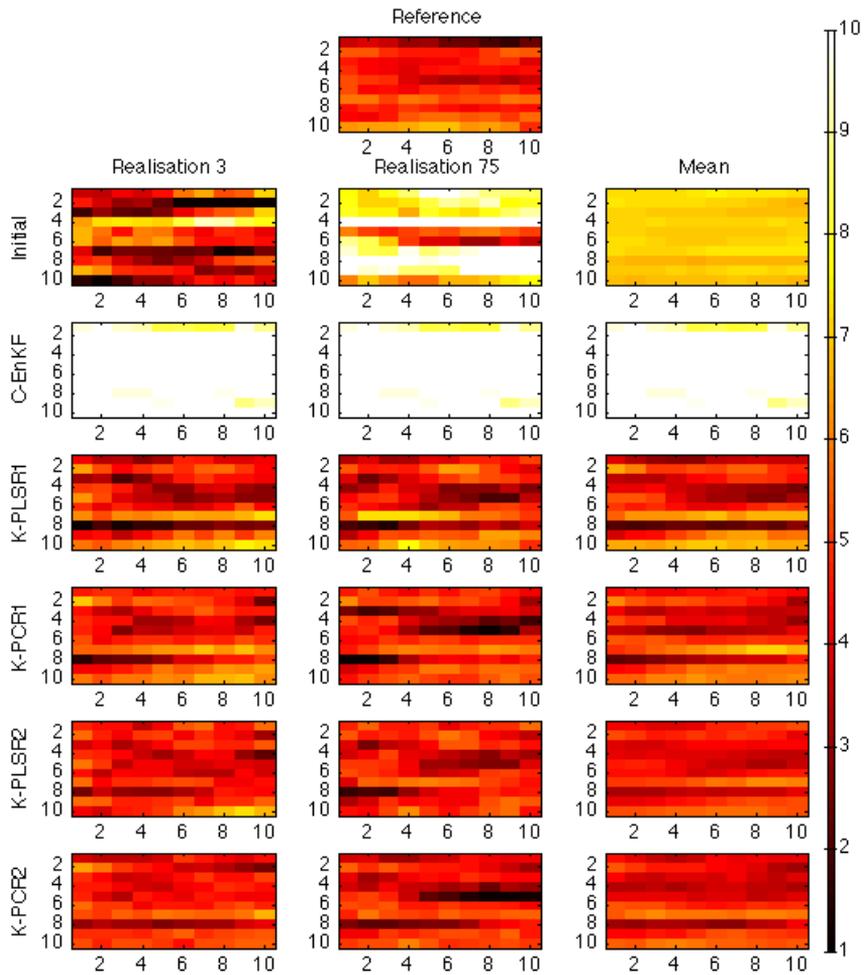


Figure 3.10: Two realisations and the estimated ensemble mean for the initial ensemble, the classical EnKF updating scheme (C-EnKF) and the four kernel-shrinkage regression techniques.

n_e	Scheme	TSS	Coverage (%)
100	No Updating	70.6	100.0
100	Classical EnKF	311.5	6.1
100	EnKF-KerPLSR1	9.3	92.2
100	EnKF-KerPCR1	9.0	96.6
100	EnKF-KerPLSR2	7.5	93.6
100	EnKF-KerPCR2	6.2	95.6
100	EnKF-PLSR1	60.5	19.3

Table 3.1: Scaled ($1/n_e$) Total Sum of Squares (TSS) of the estimated posterior mean to the reference solution, and coverage of the reference solution in the estimated 95% prediction intervals based on 100 different initial ensembles.

Using the kernelised shrinkage regression techniques reduces the TSS with between 87 and 91 percent, with the largest decrease when we use the supervised parameter selection scheme. This suggests that the PRESS statistic is not the optimal measure of the goodness-of-fit for this model. Replacing the PRESS statistic with alternative measures, is a topic for future research. The estimated mean coverage is close to the theoretical value of 95 percent for all four schemes.

Contrary to the classical EnKF updating scheme, using the EnKF-PLSR1 scheme does not lead to an increase in the TSS compared to the initial ensemble. As explained in Sætrum and Omre (2010), we expect to see this behaviour because we reduce the problem of regression model overfitting caused by collinear ensemble members, when using shrinkage regression techniques. However, the estimated prediction interval is not able to capture the reference state vector, which we expect when using a linear updating scheme on a highly non-linear likelihood model.

Fig. 3.11a and b, further illustrates the negative effect of model overfitting. Here the scaled Residual Sum of Squares (RSS) between the ensemble members and the reference solution, $1/n_e \sum_{i=1}^{n_e} \|\mathbf{x}^{c(i)} - \mathbf{x}^{\text{True}}\|_2^2$, and the coverage as a function of $k = 0, \dots, 19$ are displayed. As we can see from these figures, the RSS for the classical EnKF is increasing rapidly for the first two updating steps. At the same time, the coverage is rapidly decreasing and after two assimilation steps it is down to zero percent. A similar trend can be seen for the EnKF-PLSR1 updating scheme. However, the effect of model overfitting is not as prominent as for the classical EnKF. This illustrates the usefulness of applying shrinkage regression techniques in an EnKF updating

scheme.

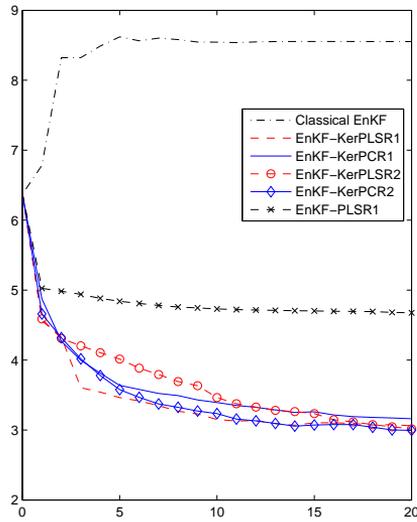
For the kernelised shrinkage regression techniques, the behaviour of the RSS and coverage is similar for all of the four updating schemes considered. Initially the RSS decrease rapidly before it stabilises at the later time steps. Similarly, the coverage is slowly converging towards the expected value of 95%, where it appears to stabilise. In addition, note that we are better able to preserve the spread of the updated ensemble members using the kernel-shrinkage regression techniques compared with the EnKF-PLSR1 updating scheme. We believe that the additional non-linearities introduced when mapping the data vector into the feature space is the cause of this behaviour. Hence, we reduce the collinearities between the updated ensemble members, which potentially can lead to model overfitting.

3.7 Conclusions

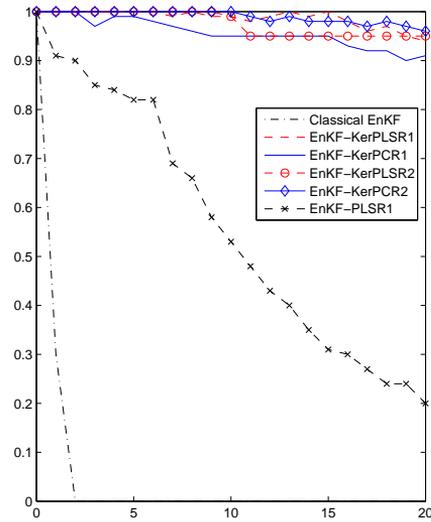
We have formulated an EnKF updating scheme based on kernel-shrinkage regression techniques to handle highly non-linear and non-Gaussian likelihood models. Contrary to previously suggested EnKF updating schemes based on kernel methods, the approach does not require solving a pre-image problem. Moreover, the computational complexity is equal to the fastest EnKF algorithms previously suggested.

We presented kernel regression as a natural extension of the classical EnKF to the non-linear case using Gaussian Random Fields (GRF). Under the assumption that the response variable is a GRF, we obtained a prediction scheme defined through the inner product between data vectors in the feature space. We can equally describe these inner products through symmetric, positive definite kernel functions in the Euclidean space, which corresponds to covariance functions for a GRF. Hence, we gave an extension of the Bayesian formulation of the classical EnKF to a non-linear setting, from which kernel Ridge Regression (RR) is a special case. In addition, we considered two additional kernel shrinkage regression techniques based on dimension reduction, namely kernel Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR).

We evaluated the performance of the three kernel-shrinkage regression techniques for a non-linear regression problem. Here we used a Cross-Validation (CV) scheme to select the prior hyperparameters of the regression models. When we applied the estimated models for prediction purposes, we obtained similar results for the kernel PCR and PLSR methods. However, the dimension of the respective reduced order models was smaller for the kernel PLSR, because this method uses both the predictor and response



(a) RSS



(b) Coverage

Figure 3.11: Scaled $(1/n_e)$ Residual Sum of Squares (RSS), for the forecasted state vector ensemble members and the reference solution, and coverage of the reference solution in the estimated 95% prediction intervals as a function of time steps k .

variables in the dimension reduction. Kernel RR gave slightly larger errors in the predictions.

EnKF updating schemes based on kernel PCR and PLSR were further tested on a hidden Markov model with a non-linear, non-Gaussian likelihood model. For comparison, we considered the standard EnKF updating scheme for non-linear likelihood models. The kernelised shrinkage regression techniques provided good estimates of unknown reference state vector, with uncertainty estimates close to the theoretical bounds. On this model, the updating scheme based on kernel PCR performed slightly better than kernel PLSR, although the subspace dimension tended to be smaller when using the supervised PLSR approach. The standard EnKF, on the other hand, completely missed the reference solution, with an ensemble collapsing after a few updating steps.

3.8 Acknowledgements

This work is funded by the Uncertainty in Reservoir Evaluation (URE) consortium at NTNU.

References

- S. I. Aanonsen, G. Nævdal, D. S. Oliver, A. C. Reynolds, and B. Vallès. Ensemble Kalman filter in reservoir engineering - a review. *SPE Journal*, 14(3):393–412, 2009.
- J. L. Anderson. A local least squares framework for ensemble filtering. *Monthly Weather Review*, 131(4):634–642, 2003.
- J. Caers and K. Park. Distance-based representation of reservoir uncertainty: the metric EnKF. In *Proceeding to the ECMOR XI, Sept 8-11, Bergen, Norway*, 2008.
- G. Casella and R. L. Berger. *Statistical Inference*. Duxbury, 2002.
- J.-P. Chiles and P. Delfiner. *Geostatistics, Modeling Spatial uncertainty*. Wiley, 1999.
- A. Doucet, S. Godsill, and C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000.

- G. Evensen. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99:10143–10162, 1994.
- G. Evensen. *Data Assimilation. The Ensemble Kalman Filter*. Springer, 2007.
- D. E. Farrer and R. R. Glauber. Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics*, 49(1): 92–107, 1967.
- R. Furrer and T. Bengtsson. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *Journal of Multivariate Analysis*, 98(2):227–255, 2007.
- M. N. Gibbs. *Bayesian Gaussian Processes for Regression and Classification*. PhD thesis, University of Cambridge, 1997.
- T. Hastie, R. Tibshirani, and J. Freidman. *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. Springer, New York, 2 edition, 2009.
- I. S. Helland. Some theoretical aspects of partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 58(2):97–107, 2001.
- A. E. Hoerl and R. W. Kennard. Ridge Regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(3):55–67, 1970.
- T. Hofmann, B. Schölkopf, and A. J. Smola. Kernel methods in machine learning. *The Annals of Statistics*, 36(3):1171–1220, 2008.
- H. Hotelling. Analysis of a complex of statistical variables into Principal Components. *Journal of Educational Psychology*, 24(6):417–441 and 498–520, 1933.
- P. L. Houtekamer and H. L. Mitchell. Data assimilation using an Ensemble Kalman Filter technique. *Monthly Weather Review*, 126:796–811, 1998.
- T. M. Huang, V. Kecman, and I. Kopriva. *Kernel Based Algorithms for Mining Huge Data Sets, Supervised, Semi-supervised, and Unsupervised Learning*. Springer, 2006.
- A. G. Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press London, 1978.

- J. H. Kalivas. Cyclic subspace regression with analysis of the hat matrix. *Chemometrics and Intelligent Laboratory Systems*, 45:215–224, 1999.
- R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82(Series D): 35–45, 1960.
- J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural Networks, Piscataway, NJ*, pages 1942–1948, 1995.
- S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.
- J. T. Kwok and I. W. Tsang. The pre-image problem in kernel methods. *IEEE Transactions on Neural Networks*, 15(6):1517–1525, 2004.
- K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979.
- S. Mika, B. Schölkopf, A. Smola, R. Müller, M. Scholz, and G. Rätsch. *Advances in neural information processing systems 11*, chapter Kernel PCA and de-noising in feature spaces. San Mateo CA: Morgan Kaufmann, 1998.
- I. B. Myrseth and H. Omre. *Large-scale Inverse Problems and Quantification of Uncertainty*, chapter The Ensemble Kalman Filter and related filters. John Wiley and Sons, 2010.
- I. B. Myrseth, J. Sætrum, and H. Omre. Resampling the Ensemble Kalman Filter. Paper submitted for publication, 2010.
- J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 2006.
- A. O’Hagan. Curve fitting and optimal design for prediction. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(1):1–42, 1978.
- D. Oliver. On conditional simulation to inaccurate data. *Journal of Mathematical Geology*, 28(6):811–817, 1996.
- C. R. Rao. *Linear Statistical Inference and Its Applications*. New York: John Wiley, 2 edition, 1973.
- C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

- R. Rosipal and L. J. Trejo. Kernel Partial Least Squares regression in reproducing kernel Hilbert space. *Journal of Machine Learning Research*, 2:97–123, 2002.
- R. Rosipal, M. Girolami, L. J. Trejo, and A. Cichocki. Kernel PCA for Feature Extraction and De-Noising in Non-Linear Regression. *Neural Computing & Applications*, 10(3):231–243, 2001.
- S. Rännner, F. Lindgren, P. Gelandi, and S. Wold. A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and Algorithm. *Journal of Chemometrics*, 8(2):111–125, 1994.
- W. Sacher and P. Bartello. Sampling errors in Ensemble Kalman filtering. Part I: Theory. *Monthly Weather Review*, 136(8):3035–3049, 2008.
- P. Sarma and W. H. Chen. Generalization of the Ensemble Kalman Filter using kernels for nongaussian Random Fields. In *Proceedings to the Reservoir Simulation Symposium, 2-4 February 2009, The Woodlands, Texas*. SPE, 2009.
- P. Sarma, L. Durlofsky, and K. Aziz. Kernel Principal Component Analysis for an efficient, differentiable parametrization of multipoint geostatistics. *Mathematical Geology*, 40(1):3–32, 2008.
- J. A. Scales and R. Snieder. To Bayes or not to Bayes? *Geophysics*, 62(4):1045–1046, 1997.
- B. Schölkopf, A. Smola, and K. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10(5):1299–1319, 1998.
- G. A. F. Seber and A. J. Lee. *Linear Regression Analysis*. Wiley, 2003.
- A. Smola. *Learning with Kernels*. PhD thesis, TU Berlin, 1998.
- A. Smola and B. Schölkopf. *Learning with Kernels*. Springer, 2002.
- A. Smola, B. Schölkopf, and K. Müller. The connection between regularization operators and support vector kernels. *Neural Networks*, 11(4):637–649, 1998.
- J. Sætrum and H. Omre. Ensemble Kalman filtering with shrinkage regression techniques, 2010. Paper accepted for publication in Computational Geosciences.

- J. S. Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
- P. J. van Leeuwen. Comments on "Data assimilation using an Ensemble Kalman Filter technique". *Monthly Weather Review*, 127:1374–1377, 1999.
- V. Vapnik. *Statistical Learning Theory*. Wiley, 1998.
- C. K. I. Williams. *Learning and inference in graphical models*, chapter Prediction with Gaussian Processes: From Linear Regression to Linear Prediction and Beyond. Kluwer, 1998.
- H. Wold. *Quantitative Sociology: International perspectives on mathematical and statistical model building*, chapter Path models with latent variables: The NiPALS approach, pages 307–357. Academic Press, 1975.
- N. Young. *An Introduction to Hilbert Space*. Cambridge University Press, 1988.

APPENDIX

Appendix 3-A, Gaussian Random Field

The collection of random variables $\{r(\mathbf{x}_1), \dots, r(\mathbf{x}_{n_e})\}$ is a Gaussian Random Field (GRF) if any subset of the random variables $\{r(\mathbf{x}_i)\}$ has a joint Gaussian distribution. A GRF is completely specified through a mean and covariance function, denoted $m(\mathbf{x})$, and $c(\mathbf{x}, \mathbf{x}')$ respectively, for any vectors \mathbf{x} and $\mathbf{x}' \in \mathbb{R}^{n_x \times 1}$. Here $m(\mathbf{x}) = \mathbb{E}[r(\mathbf{x})]$ and

$$c(\mathbf{x}, \mathbf{x}') = \mathbb{E} \left[(r(\mathbf{x}) - m(\mathbf{x})) (r(\mathbf{x}') - m(\mathbf{x}'))^T \right]$$

Whilst $m : \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ can be any function, the following criteria must be satisfied for the covariance function $c : (\mathbb{R}^{n_x \times 1} \times \mathbb{R}^{n_x \times 1}) \rightarrow \mathbb{R}$:

1. Symmetric: $c(\mathbf{x}, \mathbf{x}') = c(\mathbf{x}', \mathbf{x})$.
2. Positive semidefinite:

$$\int c(\mathbf{x}, \mathbf{x}') f(\mathbf{x}) f(\mathbf{x}') d\mu(\mathbf{x}) d\mu(\mathbf{x}') \geq 0, \quad (\text{A-1})$$

for all \mathbf{x} and $\mathbf{x}' \in \mathbb{R}^{n_x \times 1}$ and $f \in L_2(\mathbb{R}^{n_x}, \mu)$.

Within the class of valid covariance functions are the stationary and isotropic:

$$c(\mathbf{x}, \mathbf{x}') = c(\|\mathbf{x} - \mathbf{x}'\|_2),$$

where $\|\cdot\|_2$ is the Euclidean norm.

Appendix 3-B, Computational Properties

If the model parameters, θ are specified, the computational complexity of the EnKF updating schemes based on kernel RR, kernel PCR and kernel PLSR is $\mathcal{O}(\max\{n_d, n_x, n_e\}n_e^2)$. Hence, we have the same computational complexity and memory requirements as the fastest EnKF algorithms previously suggested when $n_e < \max\{n_x, n_d\}$. Further note that we can write the updating scheme based on kernel RR, similarly to kernel PCR, using the n_e estimated eigenvectors and values of the centred kernel matrix. This follows because \mathbf{HCH} by construction is symmetric, positive semidefinite.

When the model parameters, θ are selected based on a CV scheme, as outlined in Algorithm 1, the computational complexity of the respective kernelised EnKF updating schemes is $\mathcal{O}(r \max\{n_x, n_d, n_e\}n_e^2)$. In the example considered in this paper, $r = n_\sigma n_e$, where n_σ denotes the number of discrete values of σ we consider in the Gaussian kernel function, which in this study is $n_e/4$. Hence, in terms of the computational complexity, it is preferable to pre-select the model parameters because the computational demands decrease by a factor n_e^2 . However, in the general case we expect that using pre-selected model parameters will be a less robust choice, compared with the alternative of selecting the model parameters at each updating step using CV.

Paper III

**Improved uncertainty quantification in the
Ensemble Kalman Filter using statistical model
selection techniques**

Jon Sætrum, Joakim Hove, Jan-Arild Skjervheim and Jon G. Vabø

Paper submitted for publication

Chapter 4

Improved uncertainty quantification in the Ensemble Kalman Filter using statistical model selection techniques

Abstract. The Ensemble Kalman Filter (EnKF) is a sequential Monte Carlo method for solving non-linear spatiotemporal inverse problems in high dimensions, such as petroleum reservoir evaluation. Although the EnKF has seen successful applications in numerous areas, the classical EnKF algorithm can severely underestimate the prediction uncertainty. This can lead to biased production forecasts and an ensemble collapsing into a single realisation.

In this paper we combine a previously suggested EnKF scheme based on dimension reduction in the data space, with an automatic Cross-Validation (CV) scheme to select the subspace dimension. The properties of both the dimension reduction and the CV scheme, are well known in the statistical literature. In an EnKF setting, the former can reduce the effects caused by collinear ensemble members, whilst the latter can guard against model overfitting by evaluating the predictive capabilities of the EnKF scheme. The model selection criterion traditionally used for determining the subspace dimension, on the other hand, does not take the predictive power of the EnKF scheme into account, and can potentially lead to severe problems of model overfitting. A reservoir case study is used to demonstrate that the CV scheme can substantially improve the reservoir predictions with associated uncertainty estimates.

Is not included due to copyright

Paper IV

Resampling the Ensemble Kalman Filter

Inge Myrseth, Jon Sætrom and Henning Omre

Paper submitted for publication

Chapter 5

Resampling the Ensemble Kalman Filter

Abstract. Ensemble Kalman filters (EnKF) based on a small ensemble tend to provide collapse of the ensemble over time. It is shown that this collapse is caused by positive coupling of the ensemble members due to use of one common estimate of the Kalman gain for the update of all ensemble members at each time step. This coupling can be avoided by resampling the Kalman gain from its sampling distribution in the conditioning step. In the analytically tractable Gauss-linear model finite sample distributions for all covariance matrix estimates involved in the Kalman gain estimate are known and hence Kalman gain resampling can be done. For the general nonlinear case we introduce the resampling ensemble Kalman filter (ResEnKF) algorithm. The resampling strategy in the algorithm is based on bootstrapping of the ensemble and Monte Carlo simulation of the likelihood model. An empirical study demonstrates that ResEnKF provides more reliable prediction intervals than traditional EnKF, on the cost of somewhat less accuracy in the point predictions. In a synthetic reservoir study, it is shown the the hierarchical ensemble Kalman filter (HEnKF) provides more reliable predictions and prediction intervals than both ResEnKF and traditional EnKF. The HEnKF requires additional modelling, however.

Is not included due to copyright

Paper V

Ensemble Kalman filtering in a Bayesian regression framework

Jon Sætrom, Inge Myrseth and Henning Omre

Paper submitted for publication

Chapter 6

Ensemble Kalman filtering in a Bayesian regression framework

Abstract. Bayesian methods have in recent years become popular when considering problems in geosciences, such as sequential data assimilation in high dimensions. Ensemble based Monte Carlo methods, such as the Ensemble Kalman Filter (EnKF), are attractive to use because they are easy to implement and computationally fast. However, high computational demands will often restrict the ensemble size. Problems resulting from estimation uncertainty and dependencies between the ensemble members can therefore occur. As a result, the traditional EnKF updating schemes can lead to unreliable predictions with a severe underestimation of the prediction interval. In this paper, we present analytical expressions for the prediction variance and the downward bias of the prediction intervals. Furthermore, we introduce alternative EnKF updating schemes based on Bayesian regression techniques. The main idea is to replace the traditional plug-in estimate of the Kalman gain matrix with individual realisations from a matrix-variate distribution for each updated ensemble member. We evaluate the performance of the suggested schemes through simulation on synthetic case studies. The results reveal that we can dramatically improve the accuracy of the forecast and predictions intervals, especially for small ensemble sizes.

6.1 Introduction

The Ensemble Kalman Filter (EnKF) is a Monte Carlo method for solving non-linear spatiotemporal inverse problems in high dimensions (Evensen, 2007). Applications include numerical weather prediction (Houtekamer et al., 1996; Houtekamer and Mitchell, 2001), oceanography (Keppenne and Rienecker, 2003; Leeuwenburgh et al., 2005), hydrology (Moradkhani et al., 2005) and petroleum reservoir characterisation (Aanonsen et al., 2009). The EnKF is based on the traditional Kalman Filter (KF) (Kalman, 1960), which provides an analytical solution for the posterior probability density function (pdf) of interest, assuming linear system dynamics and linear, Gaussian assumptions, termed the Gauss-linear model.

Analytical tractability of the posterior pdf will be lost for non-linear system dynamics and non-Gaussian distributions. Techniques such as Markov chain Monte Carlo or importance sampling can therefore be used to generate realisations correctly from the posterior pdf (Doucet et al., 2000; Liu, 2001; West and Harrison, 1999). High dimensional problems will, however, restrict the tractability of these methods because of high computational demands. Moreover, approximate solutions such as the extended KF (Gelb, 1974; Jazwinski, 1970), where the forward model is linearised, can lead to an unstable solution, in addition to potentially high computational demands (Evensen, 1992).

The EnKF solution to the spatiotemporal forecast problem is based on a Monte Carlo approach with sequential forecasting and assimilation of available observations. The updating step is based on a linearisation where the required unknown covariances are assessed from the Monte Carlo ensemble. From these estimates we can estimate the weights in the linearisation, referred to as the Kalman gain matrix. This entails that for the Gauss-linear model, the EnKF will be consistent with the KF as the ensemble size tends to infinity (Mardia et al., 1979). When the ensemble size is finite, however, problems resulting from estimation uncertainty and ensemble collinearities are known to occur (Furrer and Bengtsson, 2007; Houtekamer and Mitchell, 1998; Sacher and Bartello, 2008). As a consequence, the updated ensemble members will fail to correctly represent the statistical properties of the posterior distribution.

From classical multivariate statistics (Anderson, 2003b), it is well known that estimation of the unknown Kalman gain matrix can be equally formulated as a multivariate linear regression problem (Anderson, 2003a; Sætrum and Omre, 2010). Hence, the EnKF updating scheme can be reformulated based on known regression techniques aiming at improving both the accuracy

and variance estimates of the forecasts. Examples are shrinkage regression techniques for collinear data, where the unbiased least squares estimator is replaced by biased alternatives having improved predictive capabilities (Farrer and Glauber, 1967; Hastie et al., 2009). Indeed, such an approach can lead to considerable improvements compared to the classical EnKF algorithm for small ensemble sizes (Sætrum and Omre, 2010). However, the ensemble members will, similar to the classical EnKF, be coupled over time because the same Kalman gain estimate is used to update every ensemble member.

In the current paper we present theoretical results for the bias and covariances in the forecast based on the classical updating scheme, taking into account the uncertainty of the unknown Kalman gain matrix. Furthermore, we have formulated an alternative EnKF updating scheme based on Bayesian regression techniques. In this scheme, each ensemble member is updated based on a Kalman gain matrix independently generated from a matrix variate distribution, rather than using one common plug-in estimate. We consider both conjugate and non-informative prior distributions, and an approximate dimension reducing scheme for high dimensional models is suggested. Synthetic examples inspired by petroleum reservoir evaluation problems are used to empirically evaluate the performance of the suggested procedures.

6.2 Notation and problem formulation

Throughout this paper the notation $\mathbf{x} \in \mathbb{R}^{n_x \times 1}$ will be used to denote that \mathbf{x} is an n_x -dimensional column vector in the real space and \mathbf{x}^T will denote its transpose. Similarly, a matrix \mathbf{A} in the real space containing a rows and b columns will be denoted by $\mathbf{A} \in \mathbb{R}^{a \times b}$. For simplicity, the same notation will be used for random vectors and matrices.

Consider the stochastic Directed Acyclic Graph (DAG) outlined in Figure 6.1. Here $\mathbf{x}_{t_k} \in \mathbb{R}^{n_x \times 1}$ denotes the state of the unknown random vector of interest at time step k and time t_k , and similarly $\mathbf{d}_{t_k}^o \in \mathbb{R}^{n_d \times 1}$ denotes the vector of observed data. For notational convenience, we will from now on drop the subscript t_k , and write \mathbf{x}_k , \mathbf{d}_k^o . Moreover, \mathbf{x} and \mathbf{d} will be referred to as the state and observation vector respectively.

Let $f(\mathbf{x}_0)$ denote the prior pdf of the state vector at the initial time step. Through the Markov properties of a stochastic DAG, we have conditional independence between \mathbf{x}_{k+1} and \mathbf{x}_l , $l = 0, \dots, k-1$ given \mathbf{x}_k , which implies

$$f(\mathbf{x}_{k+1} | \mathbf{x}_k, \dots, \mathbf{x}_0) = f(\mathbf{x}_{k+1} | \mathbf{x}_k).$$

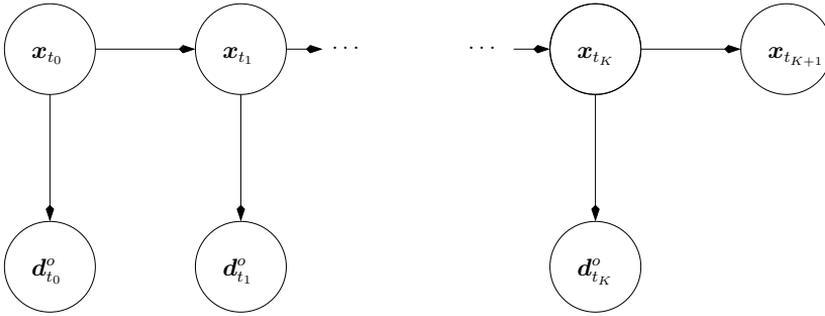


Figure 6.1: Stochastic Directed Acyclic Graph (DAG) of the model considered.

In general assume that

$$\mathbf{x}_{k+1} = \boldsymbol{\omega}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{x}_k}), \quad k = 0, \dots, K, \quad (6.1)$$

where $\boldsymbol{\omega} : (\mathbb{R}^{n_x \times 1} \times \mathbb{R}^{n_x \times 1}) \rightarrow \mathbb{R}^{n_x \times 1}$ is a known, possibly non-linear forward function. Here $\boldsymbol{\epsilon}_{\mathbf{x}_k}$ represents random model errors and/or numerical errors in the forward model, assumed to follow a known probability distribution. This implies that we implicitly get a fully specified prior model of the unknown state vector through $f(\mathbf{x}_0)$ and $\boldsymbol{\omega}(\cdot, \cdot)$. The function:

$$\mathbf{d}_k^o = \boldsymbol{\zeta}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{d}_k}), \quad k = 0, \dots, K, \quad (6.2)$$

connecting the observed data \mathbf{d}_k^o to \mathbf{x}_k , where $\boldsymbol{\zeta} : (\mathbb{R}^{n_x \times 1} \times \mathbb{R}^{n_d \times 1}) \rightarrow \mathbb{R}^{n_d \times 1}$, is a known, possibly non-linear, function and $\boldsymbol{\epsilon}_{\mathbf{d}_k}$ represents the likelihood model and observation errors, again assumed to follow a known pdf.

For notational convenience, let $\mathbf{x}_k^c \sim f(\mathbf{x}_k | \mathbf{d}_{0:k}^o)$ and $\mathbf{x}_{k+1}^u \sim f(\mathbf{x}_{k+1} | \mathbf{d}_{0:k}^o)$, for $k = 1, \dots, K$, where we use $\mathbf{d}_{0:l}^o$ to denote the sequence $\mathbf{d}_0^o, \dots, \mathbf{d}_l^o$. The objective in this model setting is to solve the spatiotemporal forecast problem of predicting \mathbf{x}_k given observations $\mathbf{d}_{0:(k-1)}^o$, for $k = 1, \dots, K + 1$. Bayesian inversion provides a sequential solution to this problem. Through Bayes rule and the Markov properties of the prior model, we have:

$$\begin{aligned} f(\mathbf{x}_k | \mathbf{d}_{0:k}^o) &\propto f(\mathbf{x}_k | \mathbf{d}_{0:(k-1)}^o) f(\mathbf{d}_k | \mathbf{x}_k) \\ f(\mathbf{x}_{k+1} | \mathbf{d}_{0:k}^o) &= \int f(\mathbf{x}_{k+1} | \mathbf{x}_k) f(\mathbf{x}_k | \mathbf{d}_{0:k}^o) d\mathbf{x}_k. \end{aligned} \quad (6.3)$$

Note that the conditional pdf $f(\mathbf{x}_{k+1} | \mathbf{x}_k)$ and $f(\mathbf{d}_k | \mathbf{x}_k)$ are implicitly defined through $\boldsymbol{\omega}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{x}_k})$ and $\boldsymbol{\zeta}(\mathbf{x}_k, \boldsymbol{\epsilon}_{\mathbf{d}_k})$ defined in Eq. (6.1) and (6.2) respectively.

In the general case, we only know the conditional distributions defined in Eq. (6.3) up to an unknown normalising constant. Computationally demanding techniques, such as Markov chain Monte Carlo (MCMC) or rejection sampling, can then be used to generate realisations correctly from the posterior distribution (Doucet et al., 2000). However, for spatiotemporal problems in high dimensions, such as petroleum reservoir evaluation and weather forecasting, these techniques are computationally prohibitive. This follows because even a single evaluation of $\omega(\mathbf{x}_k, \epsilon_{\mathbf{x}_k})$, which involves solving non-linear partial differential equations in dimensions of order $10^6 - 10^9$, can take several hours or even days. An approximate solution can be defined by assuming that \mathbf{x}_k^u and \mathbf{d}_k^o follow a distribution that ensures analytical tractability of $f(\mathbf{x}_k | \mathbf{d}_{0:k}^o)$, for example the multivariate Gaussian distribution. These model assumptions are equivalent to those made in the classical EnKF (Evensen, 2007), which we will consider next.

6.2.1 Classical Ensemble Kalman Filter

Let $\mathbf{x}_0^{c(i)}$ be a realisation from the unspecified conditional distribution at the initial time step, $f(\mathbf{x}_0 | \mathbf{d}_0^o)$. Further let $\mathbf{x}_k^{u(i)} = \omega(\mathbf{x}_{k-1}^{c(i)}, \epsilon_{\mathbf{x}_k}^{(i)})$ and $\mathbf{d}_k^{(i)} = \zeta(\mathbf{x}_k^{u(i)}, \epsilon_{\mathbf{d}_k}^{(i)})$, for $k > 0$, where we at each time step, as an approximation, assume:

$$\begin{bmatrix} \mathbf{x}_k^{u(i)} \\ \mathbf{d}_k^{(i)} \end{bmatrix} \sim \text{Gauss}_{n_y} \left(\begin{bmatrix} \boldsymbol{\mu}_{\mathbf{x}_k^u} \\ \boldsymbol{\mu}_{\mathbf{d}_k} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{x}_k^u} & \boldsymbol{\Sigma}_{\mathbf{x}_k^u, \mathbf{d}_k} \\ \boldsymbol{\Sigma}_{\mathbf{d}_k, \mathbf{x}_k^u} & \boldsymbol{\Sigma}_{\mathbf{d}_k} \end{bmatrix} \right), \quad (6.4)$$

with $n_y = n_x + n_d$. Here the notation $\mathbf{y} \sim \text{Gauss}_{n_y}(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y)$ is used to denote that $\mathbf{y} \in \mathbb{R}^{n_y \times 1}$ follows the multivariate Gaussian distribution with mean $\boldsymbol{\mu}_y$ and covariance matrix $\boldsymbol{\Sigma}_y$ as defined in Appendix 6-A. For notational convenience, we will from now on omit the subscript k , because the focus is on updating at a single time step.

Under the Gaussian assumption in Eq. (6.4), the posterior pdf $f(\mathbf{x}_k | \mathbf{d}_{0:k}^o)$ is Gaussian with analytically obtainable mean:

$$\boldsymbol{\mu}_{\mathbf{x}^u | \mathbf{d}} = \boldsymbol{\mu}_{\mathbf{x}^u} + \boldsymbol{\Sigma}_{\mathbf{x}^u, \mathbf{d}} \boldsymbol{\Sigma}_{\mathbf{d}}^{-1} (\mathbf{d}^o - \boldsymbol{\mu}_{\mathbf{d}}), \quad (6.5)$$

and covariance matrix:

$$\boldsymbol{\Sigma}_{\mathbf{x}^u | \mathbf{d}} = \boldsymbol{\Sigma}_{\mathbf{x}^u} - \boldsymbol{\Sigma}_{\mathbf{x}^u, \mathbf{d}} \boldsymbol{\Sigma}_{\mathbf{d}}^{-1} \boldsymbol{\Sigma}_{\mathbf{d}, \mathbf{x}^u}, \quad (6.6)$$

where \mathbf{d}^o is the vector of observed data. Furthermore, it is straightforward to show that

$$\mathbf{x}^{c(i)} = \mathbf{x}^{u(i)} + \boldsymbol{\Sigma}_{\mathbf{x}^u, \mathbf{d}} \boldsymbol{\Sigma}_{\mathbf{d}}^{-1} (\mathbf{d}^o - \mathbf{d}^{(i)}), \quad (6.7)$$

is one realisation from the Gaussian posterior distribution with mean and covariance given in Eq. (6.5) and (6.6) (Evensen, 2007).

Eq. (6.7) involves two model parameters, namely $\Sigma_{\mathbf{x}^u, \mathbf{d}}$ and $\Sigma_{\mathbf{d}}$ forming the Kalman gain matrix,

$$\mathbf{K} = \Sigma_{\mathbf{x}^u, \mathbf{d}} \Sigma_{\mathbf{d}}^{-1} \in \mathbb{R}^{n_x \times n_d}. \quad (6.8)$$

For the Gauss-linear model, \mathbf{K} will be analytically tractable and given by the Kalman recursions (Furrer and Bengtsson, 2007; Kalman, 1960). Analytical tractability is, however, lost in the general model setting we consider here. The EnKF solution to this problem is to use an ensemble of n_e realisations $\{(\mathbf{x}^{u(1)}, \mathbf{d}^{(1)}), \dots, (\mathbf{x}^{u(n_e)}, \mathbf{d}^{(n_e)})\}$, from which we can obtain empirical estimate of the unknown covariance matrices.

Let $\mathbf{X} = [\mathbf{x}^{u(1)}, \dots, \mathbf{x}^{u(n_e)}] \in \mathbb{R}^{n_x \times n_e}$ and $\mathbf{D} = [\mathbf{d}^{(1)}, \dots, \mathbf{d}^{(n_e)}] \in \mathbb{R}^{n_d \times n_e}$ denote the ensemble matrices for the state- and observation vector respectively. Thus, the expression

$$\hat{\mathbf{K}} = \hat{\Sigma}_{\mathbf{x}^u, \mathbf{d}} \hat{\Sigma}_{\mathbf{d}}^{-1} = \mathbf{X} \mathbf{H}_{n_e} \mathbf{D}^T (\mathbf{D} \mathbf{H}_{n_e} \mathbf{D}^T)^{-1}, \quad (6.9)$$

defines the classical estimate of the Kalman gain matrix in a general setting (Anderson, 2003b), where $\mathbf{H}_{n_e} = \mathbf{I}_{n_e} - 1/n_e \mathbf{1}_{n_e, 1} \mathbf{1}_{n_e, 1}^T \in \mathbb{R}^{n_e \times n_e}$ is the idempotent centring matrix. Here \mathbf{I}_{n_e} is the n_e -dimensional identity matrix and $\mathbf{1}_{n_e, 1}$ is a n_e -dimensional vector where all the entries are equal to one.

Replacing the unknown Kalman gain matrix in Eq. (6.7), with the empirical estimate in Eq. (6.9), thus defines the standard EnKF updating scheme in a general form:

$$\mathbf{x}^{c(i)} = \mathbf{x}^{u(i)} + \hat{\mathbf{K}}(\mathbf{d}^o - \mathbf{d}^{(i)}), \text{ for } i = 1, \dots, n_e. \quad (6.10)$$

We will from now on refer to the replacement of the unknown Kalman gain matrix with the empirical estimate in Eq. (6.9) as the standard EnKF. Note that we here motivate the EnKF updating scheme by a Gaussian prior model assumption. However, because the conditioned state vector, $\mathbf{x}^{c(i)}$, at each updating step is given as a linear combination of $\mathbf{x}^{u(i)}$, \mathbf{d}^o and $\mathbf{d}^{(i)}$, with the model parameters only involved in the Kalman gain matrix, we can still capture non-Gaussian properties present in the general posterior distribution at the initial time step, $f(\mathbf{x}_0 | \mathbf{d}_0^o)$, at the later time steps (Evensen, 2007).

Under the Gauss-linear model, we have convergence in distribution of the EnKF updating scheme based on $\hat{\mathbf{K}}$ towards the classical KF scheme as $n_e \rightarrow \infty$ (Furrer and Bengtsson, 2007). However, for finite ensemble sizes the standard EnKF updating scheme will produce a conditional ensemble which underrepresents the variability of \mathbf{x}^c (Furrer and Bengtsson, 2007;

van Leeuwen, 1999; Sacher and Bartello, 2008). In the following section we will use classical multivariate statistics to establish theoretical results regarding properties of the mean and covariance of the conditioned state vector ensemble.

6.2.2 Properties of the Conditioned State Vector Ensemble

Following Myrseth et al. (2009), we assume that the estimated Kalman gain matrix, $\hat{\mathbf{K}}$, follows a matrix variate distribution, thus accounting for both the model and parameter uncertainty. Consider the standard EnKF updating scheme defined in Eq. (6.10). Using known results for the conditional expectation and covariance we get:

$$\begin{aligned} \mathbb{E}[\mathbf{x}^{c(i)}] &= \boldsymbol{\mu}_{\mathbf{x}^u|\mathbf{d}} + (\boldsymbol{\Gamma}_{\hat{\mathbf{K}}} - \mathbf{K})(\mathbf{d}^o - \boldsymbol{\mu}_d) \\ \text{Cov}(\mathbf{x}^{c(i)}) &= \boldsymbol{\Sigma}_{\mathbf{x}^u|\mathbf{d}} + (\boldsymbol{\Gamma}_{\hat{\mathbf{K}}} - \mathbf{K})\boldsymbol{\Sigma}_d(\boldsymbol{\Gamma}_{\hat{\mathbf{K}}} - \mathbf{K})^T \\ &\quad + \mathbb{E}\left[(\hat{\mathbf{K}} - \boldsymbol{\Gamma}_{\hat{\mathbf{K}}})(\boldsymbol{\Sigma}_d + \boldsymbol{\Delta}_{\mathbf{d}^o})(\hat{\mathbf{K}} - \boldsymbol{\Gamma}_{\hat{\mathbf{K}}})^T\right], \end{aligned} \quad (6.11)$$

for $i = 1, \dots, n_e$. Furthermore, the covariance between ensemble members $\mathbf{x}^{c(i)}$ and $\mathbf{x}^{c(j)}$ for $i, j = 1, \dots, n_e$, $i \neq j$ is

$$\text{Cov}(\mathbf{x}^{c(i)}, \mathbf{x}^{c(j)}) = \mathbb{E}\left[(\hat{\mathbf{K}} - \boldsymbol{\Gamma}_{\hat{\mathbf{K}}})\boldsymbol{\Delta}_{\mathbf{d}^o}(\hat{\mathbf{K}} - \boldsymbol{\Gamma}_{\hat{\mathbf{K}}})^T\right]. \quad (6.12)$$

Here $\mathbb{E}[\hat{\mathbf{K}}] = \boldsymbol{\Gamma}_{\hat{\mathbf{K}}}$, \mathbf{K} is the true Kalman gain matrix, defined in Eq. (6.8), and $\boldsymbol{\Delta}_{\mathbf{d}^o} = (\mathbf{d}^o - \boldsymbol{\mu}_d)(\mathbf{d}^o - \boldsymbol{\mu}_d)^T \in \mathbb{R}^{n_d \times n_d}$ is a matrix with rank equal to one. The details are given in Appendix 6-B.

Under the Gaussian assumption in Eq. (6.4), with $n_e > n_x + n_d$, the finite sampling distribution of the estimated Kalman gain in Eq. (6.9) is analytically tractable, and given as (Kabe, 1968; Kaufman, 1969; Kshirsagar, 1961; Wegge, 1971):

$$\hat{\mathbf{K}} \sim \text{MatrixT}_{n_x, n_d}(\mathbf{K}, \boldsymbol{\Sigma}_d^{-1}, \boldsymbol{\Sigma}_{\mathbf{x}^u|\mathbf{d}}, n_e - n_d), \quad (6.13)$$

with $\boldsymbol{\Sigma}_{\mathbf{x}^u|\mathbf{d}}$ defined in Eq. (6.6), and the matrix- t distribution defined in Appendix 6-A. Consider the Kalman gain estimate in Eq. (6.9) as one realisation from the matrix- t distribution in Eq. (6.13). Gupta and Nagar (2000, Theorem 4.3.2) then yields:

$$\begin{aligned} \mathbb{E}[\mathbf{x}^{c(i)}] &= \boldsymbol{\mu}_{\mathbf{x}^c} = \boldsymbol{\mu}_{\mathbf{x}^u|\mathbf{d}} \\ \text{Cov}(\mathbf{x}^{c(i)}) &= \boldsymbol{\Sigma}_{\mathbf{x}^c} = \boldsymbol{\Sigma}_{\mathbf{x}^u|\mathbf{d}} \left(1 + \frac{n_d}{n_e - n_d - 2} + \frac{\text{tr}\{\boldsymbol{\Sigma}_d^{-1} \boldsymbol{\Delta}_{\mathbf{d}^o}\}}{n_e - n_d - 2} \right), \end{aligned} \quad (6.14)$$

for $i = 1, \dots, n_e$, and further for $i, j = 1, \dots, n_e$ $i \neq j$

$$\text{Cov}(\mathbf{x}^{(c(i))}, \mathbf{x}^{(c(j))}) = \frac{\text{tr}\{\Sigma_{\mathbf{d}}^{-1} \Delta_{\mathbf{d}^o}\}}{n_e - n_d - 2} \Sigma_{\mathbf{x}^u | \mathbf{d}}. \quad (6.15)$$

Notice that the posterior mean, $\boldsymbol{\mu}_{\mathbf{x}^c}$, coincides with the conditional mean $\boldsymbol{\mu}_{\mathbf{x}^u | \mathbf{d}}$, whilst the posterior covariance, $\Sigma_{\mathbf{x}^c}$, is larger than $\Sigma_{\mathbf{x}^u | \mathbf{d}}$, thus taking into account the additional uncertainty when estimating the unknown model parameters. However, as we increase the ensemble size, the uncertainty in the Kalman gain estimate is reduced and $\Sigma_{\mathbf{x}^c}$ tends to $\Sigma_{\mathbf{x}^u | \mathbf{d}}$. The problematic result appears in Eq. (6.15), which demonstrates that the updated ensemble members are coupled.

Assume that the single joint realisation $\{\mathbf{x}^{u(i)}, \mathbf{d}^{(i)}\}$ is independent of $\hat{\mathbf{K}}$. Although this assumption is strictly not valid for the standard EnKF for small ensemble sizes, we can accomplish this using two separate ensembles; one used for estimation and one used for prediction, similar to the double EnKF introduced in Houtekamer and Mitchell (1998).

Proceeding with the EnKF updating scheme in Eq. (6.10), we obtain an updated ensemble $\{\mathbf{x}^{c(1)}, \dots, \mathbf{x}^{c(n_e)}\}$. Let the posterior mean and covariance be estimated from the updated ensemble by $\hat{\boldsymbol{\mu}}_{\mathbf{x}^c} = \frac{1}{n_e} \sum_{i=1}^{n_e} \mathbf{x}^{c(i)}$ and $\hat{\Sigma}_{\mathbf{x}^c} = 1/(n_e - 1) \sum_{i=1}^{n_e} (\mathbf{x}^{c(i)} - \hat{\boldsymbol{\mu}}_{\mathbf{x}^c})(\mathbf{x}^{c(i)} - \hat{\boldsymbol{\mu}}_{\mathbf{x}^c})^T$. The expectation and covariance of the posterior empirical mean become

$$\begin{aligned} \text{E}[\hat{\boldsymbol{\mu}}_{\mathbf{x}^c}] &= \boldsymbol{\mu}_{\mathbf{x}^c} \\ \text{Cov}(\hat{\boldsymbol{\mu}}_{\mathbf{x}^c}) &= \Sigma_{\mathbf{x}^u | \mathbf{d}} \left(\frac{1}{n_e} + \frac{n_d}{n_e(n_e - n_d - 2)} + \frac{\text{tr}\{\Sigma_{\mathbf{d}}^{-1} \Delta_{\mathbf{d}^o}\}}{n_e - n_d - 2} \right). \end{aligned} \quad (6.16)$$

Hence, the posterior empirical mean will be an unbiased estimator for the prediction mean, $\boldsymbol{\mu}_{\mathbf{x}^c}$. Note, however, that for small values of $n_e - n_d$ we expect to see a relatively large variability in the posterior empirical mean, resulting from the last term in Eq. (6.16), which increases with the dimension of the data vector, n_d .

The expected value of the estimated posterior covariance matrix becomes

$$\text{E}[\hat{\Sigma}_{\mathbf{x}^c}] = \Sigma_{\mathbf{x}^u | \mathbf{d}} \left(1 + \frac{n_d}{n_e - n_d - 2} \right),$$

and hence

$$\text{E}[\hat{\Sigma}_{\mathbf{x}^c} - \Sigma_{\mathbf{x}^c}] = -\frac{\text{tr}\{\Sigma_{\mathbf{d}}^{-1} \Delta_{\mathbf{d}^o}\}}{n_e - n_d - 2} \Sigma_{\mathbf{x}^u | \mathbf{d}}. \quad (6.17)$$

That is, if we use a common Kalman gain matrix in the EnKF updating scheme, the ensemble members will be positively correlated, which leads to an underestimation of the prediction covariance matrix.

Note that if we instead use independent realisations of the Kalman gain matrices generated from the matrix- t distribution in Eq. (6.13), the ensemble members $\{\mathbf{x}^{c(1)}, \dots, \mathbf{x}^{c(n_e)}\}$, will remain independent and consequently $\mathbb{E} \left[\hat{\Sigma}_{\mathbf{x}^c} \right] = \Sigma_{\mathbf{x}^c}$.

Consider a special case where the likelihood model is Gauss-linear, $\mathbf{d} = \mathbf{B}\mathbf{x} + \boldsymbol{\epsilon}_d$, with $\mathbf{B} \in \mathbb{R}^{n_d \times n_x}$ known and $\boldsymbol{\epsilon}_d \sim \text{Gauss}_{n_d}(\mathbf{0}_{n_d,1}, \Sigma_{\boldsymbol{\epsilon}_d})$ independent of \mathbf{x} . The estimated Kalman gain matrix can then be given as:

$$\hat{\mathbf{K}}_{\text{GL}} = \hat{\Sigma}_{\mathbf{x}^u} \mathbf{B}^T \left(\mathbf{B} \hat{\Sigma}_{\mathbf{x}^u} \mathbf{B}^T + \Sigma_{\boldsymbol{\epsilon}_d} \right)^{-1}, \quad (6.18)$$

which corresponds to the original formulation in Evensen (1994). Here $\hat{\Sigma}_{\mathbf{x}^u}$ is the unbiased estimate of the covariance matrix $\Sigma_{\mathbf{x}^u}$ and $\mathbf{0}_{a,b}$ is used to denote a $a \times b$ -dimensional matrix containing only zeros. Because the mean and covariance of $\hat{\mathbf{K}}_{\text{GL}}$ is analytically intractable we, similar to Furrer and Bengtsson (2007), further assume $\mathbf{B}^T \mathbf{B} = \mathbf{I}_{n_x}$ and $\boldsymbol{\epsilon}_d \sim \text{Gauss}_{n_d}(\mathbf{0}_{n_d,1}, \sigma^2 \mathbf{I}_{n_d})$. Consider the estimated posterior covariance matrix based on the standard EnKF, $\hat{\Sigma}_{\mathbf{x}^u|\mathbf{d}} = \hat{\Sigma}_{\mathbf{x}^u} - \hat{\mathbf{K}}_{\text{GL}} \left(\mathbf{B} \hat{\Sigma}_{\mathbf{x}^u} \mathbf{B}^T + \Sigma_{\boldsymbol{\epsilon}_d} \right) \hat{\mathbf{K}}_{\text{GL}}^T$, with $\hat{\mathbf{K}}_{\text{GL}}$ defined in Eq. (6.18), then Result 2 in Appendix 6-C gives a bound for the negative bias:

$$\text{tr} \left\{ \mathbb{E} \left[\hat{\Sigma}_{\mathbf{x}^u|\mathbf{d}} - \Sigma_{\mathbf{x}^u|\mathbf{d}} \right] \right\} < -\frac{2}{n_e - 1} \frac{\sigma^4}{(\lambda_1 + \sigma^2)^3} \sum_{i=1}^{n_d} \lambda_i^2 + \mathcal{O}(n_e^{-2}), \quad (6.19)$$

where $\lambda_1 > \dots > \lambda_{n_d}$ are the eigenvalues of $\mathbf{B} \Sigma_{\mathbf{x}^u} \mathbf{B}^T$. As explained in Furrer and Bengtsson (2007), we can generalise the result through a transformation of the variables \mathbf{x} and \mathbf{d} .

It is worth noting that the predictions of the EnKF scheme are characterised by the parameters $(\boldsymbol{\mu}_{\mathbf{x}^c}, \Sigma_{\mathbf{x}^c})$ while $\Sigma_{\mathbf{x}^u|\mathbf{d}}$ is merely a model parameter. Consequently the former parameters and the result in Eq. (6.17) are of primary interest in a predictive setting as defined here. Finally note that it is straightforward to verify both the bias under the finite sampling distribution in Eq. (6.17), and the lower bound in Eq. (6.19) for the standard EnKF through exhaustive simulation.

The results above illustrates a critical weakness with the standard EnKF scheme: Using the same estimated Kalman gain matrix to update all ensemble members, leads to correlated ensemble members after one updating step. Consequently, by using the classical covariance matrix estimates based on the updated ensemble, we underestimate the posterior covariance matrix. More troubling, however, is that we amplify the ensemble coupling as we perform sequential data assimilation. Hence, it is not surprising that the

updated ensemble can ultimately collapse into one single realisation when the standard EnKF updating scheme is applied.

Myrseth and Omre (2010a), suggested a hierarchical Bayesian approach, termed the Hierarchical EnKF (HEnKF), as an alternative to using empirical estimates of the covariance matrix. The HEnKF replace the estimated covariance matrices with realisations from the analytically tractable posterior distribution, assuming conjugate prior distributions for the mean and covariance in Eq. (6.4). Hence, the coupling between the updated ensemble will be reduced, leading to an improved representation of the prediction uncertainty. In the current paper, we have extended the HEnKF to a more general setting, using results known from Bayesian regression (Box and Tiao, 1992; Press, 1982). This entails sampling the Kalman gain matrix directly for each ensemble member through analytically tractable matrix variate pdfs.

6.3 Bayesian regression

We will now consider the posterior distribution of the Kalman gain matrix, \mathbf{K} , through a hierarchical Bayesian approach using both a natural conjugate prior, similar to Myrseth and Omre (2010a), and a non-informative Jeffreys' prior (Press, 1982).

6.3.1 Informative prior distribution

Consider the prior model assumption in Eq. (6.4), and let in addition $[\boldsymbol{\mu}_y | \boldsymbol{\Sigma}_y] \sim \text{Gauss}_{n_y}(\boldsymbol{\eta}_y, \xi^{-1} \boldsymbol{\Sigma}_y)$ and $\boldsymbol{\Sigma}_y \sim W_{n_y}^{-1}(\boldsymbol{\Psi}_y^{-1}, \nu)$, where $\mathbf{y} = [\mathbf{x}^{uT}, \mathbf{d}^T]^T \in \mathbb{R}^{n_y \times 1}$. Here, the inverse Wishart distribution is defined in Appendix 6-A, and $\boldsymbol{\eta}_y$, $\boldsymbol{\Psi}_y$, ξ and ν are known prior hyperparameters. Let $\mathbf{Y} = [\mathbf{X}; \mathbf{D}] \in \mathbb{R}^{n_y \times n_e}$ denote the ensemble matrix for the combined state- and observation vector, \mathbf{y} , then (Anderson, 2003b):

$$[\boldsymbol{\mu}_y | \boldsymbol{\Sigma}_y, \mathbf{Y}] \sim \text{Gauss}_{n_y} \left(\boldsymbol{\eta}_y^c, \frac{1}{\xi^c} \boldsymbol{\Sigma}_y \right)$$

$$[\boldsymbol{\Sigma}_y | \mathbf{Y}] \sim W_{n_y}^{-1}(\boldsymbol{\Psi}_y^{c-1}, \nu^c),$$

with

$$\boldsymbol{\eta}_y^c = \begin{bmatrix} \boldsymbol{\eta}_{x^u}^c \\ \boldsymbol{\eta}_d^c \end{bmatrix} = \frac{n_e \hat{\boldsymbol{\mu}}_y + \xi \boldsymbol{\eta}_y}{\xi^c}$$

$$\boldsymbol{\Psi}_y^c = \begin{bmatrix} \boldsymbol{\Psi}_{x^u}^c & \boldsymbol{\Psi}_{x^u, d}^c \\ \boldsymbol{\Psi}_{d, x^u}^c & \boldsymbol{\Psi}_d^c \end{bmatrix} = \boldsymbol{\Psi}_y + (n_e - 1) \hat{\boldsymbol{\Sigma}}_y + \frac{n_e \xi}{\xi^c} (\hat{\boldsymbol{\mu}}_y - \boldsymbol{\eta}_y)(\hat{\boldsymbol{\mu}}_y - \boldsymbol{\eta}_y)^T, \tag{6.20}$$

$\nu^c = \nu + n_e$ and $\xi^c = \xi + n_e$. Here, $\hat{\boldsymbol{\mu}}_{\mathbf{y}}$ is the average and $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}$ is the traditional unbiased estimate of the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}}$ of the ensemble \mathbf{Y} . Hence, by Result 5, Appendix 6-D, the Kalman gain matrix given the joint ensemble is distributed as

$$[\mathbf{K}|\mathbf{Y}] \sim \text{MatrixT}_{n_x, n_d} \left(\boldsymbol{\Gamma}_{\mathbf{K}}^c, (\boldsymbol{\Psi}_{\mathbf{d}}^c)^{-1}, \boldsymbol{\Psi}_{\mathbf{x}^u|\mathbf{d}}^c, \nu^c - n_x + 1 \right), \quad (6.21)$$

where

$$\boldsymbol{\Gamma}_{\mathbf{K}}^c = \boldsymbol{\Psi}_{\mathbf{x}^u, \mathbf{d}}^c \boldsymbol{\Psi}_{\mathbf{d}}^{c-1} \quad (6.22)$$

$$\boldsymbol{\Psi}_{\mathbf{x}^u|\mathbf{d}}^c = \boldsymbol{\Psi}_{\mathbf{x}^u}^c - \boldsymbol{\Gamma}_{\mathbf{K}}^c \boldsymbol{\Sigma}_{\mathbf{d}}^c \boldsymbol{\Gamma}_{\mathbf{K}}^{cT}. \quad (6.23)$$

Replacing the estimated Kalman gain matrix, $\hat{\mathbf{K}}$ in Eq. (6.10) with independent realisations from the matrix- t distribution in Eq. (6.21) thus defines the EnKF updating scheme.

6.3.2 Non-Informative Prior Distribution

In the hierarchical Bayesian approach outlined above, we need to select prior hyperparameters for both the unknown mean and covariance. When considering high dimensional spatiotemporal inverse problems, this task can be far from trivial. A non-informative prior distribution can be used to avoid this problem.

Consider the multivariate linear regression problem:

$$\mathbf{X} = \mathbf{K}\mathbf{D} + \mathbf{R}, \quad (6.24)$$

where we for notational convenience assume that $\mathbf{X} \in \mathbb{R}^{n_x \times n_e}$ and $\mathbf{D} \in \mathbb{R}^{n_d \times n_e}$ are centred ensemble matrices, $\mathbf{K} \in \mathbb{R}^{n_x \times n_d}$ is the matrix of regression coefficients, or the Kalman gain, and $\mathbf{R} \in \mathbb{R}^{n_x \times n_e}$ is the matrix of regression model error independent of \mathbf{K} and \mathbf{D} . Under the EnKF model assumptions in Eq. (6.4):

$$\mathbf{D} \sim \text{MatrixG}_{n_d, n_e}(\mathbf{0}_{n_d, n_e}, \mathbf{I}_{n_e}, \boldsymbol{\Sigma}_{\mathbf{d}}).$$

Further let the regression error term be

$$\mathbf{R} \sim \text{MatrixG}_{n_x, n_e}(\mathbf{0}_{n_x, n_e}, \mathbf{I}_{n_e}, \boldsymbol{\Sigma}_{\mathbf{r}}),$$

where the matrix-variate Gaussian distribution is defined in Appendix 6-A. As shown in Mardia et al. (1979), the conditional likelihood of \mathbf{X} given \mathbf{D} can be written as:

$$f(\mathbf{X}|\mathbf{D}, \mathbf{K}, \boldsymbol{\Sigma}_{\mathbf{r}}) \propto |\boldsymbol{\Sigma}_{\mathbf{r}}|^{-n_e/2} \exp \left\{ -\frac{1}{2} \text{tr} (\boldsymbol{\Psi}_{\mathbf{K}} \boldsymbol{\Sigma}_{\mathbf{r}}^{-1}) \right\}, \quad (6.25)$$

where

$$\Psi_{\mathbf{K}} = n_e \left[(\mathbf{K} - \hat{\mathbf{K}}) \hat{\Sigma}_{\mathbf{d}} (\mathbf{K} - \hat{\mathbf{K}})^T + \hat{\Sigma}_{\mathbf{r}} \right]. \quad (6.26)$$

Here $\hat{\mathbf{K}}$, given in Eq. (6.9), and $\hat{\Sigma}_{\mathbf{r}} = 1/n_e (\mathbf{X} - \hat{\mathbf{K}}\mathbf{D})(\mathbf{X} - \hat{\mathbf{K}}\mathbf{D})^T$ are the maximum likelihood estimates of the Kalman gain and residual covariance matrix respectively, while $\hat{\Sigma}_{\mathbf{d}} = 1/n_e \mathbf{D}\mathbf{D}^T$.

We will now consider the posterior distribution of \mathbf{K} for the non-informative Jeffreys' prior (Jeffreys, 1946; Press, 1982):

$$f(\mathbf{K}, \Sigma_{\mathbf{r}}) \propto |\Sigma_{\mathbf{r}}|^{-(n_x+1)/2}. \quad (6.27)$$

By Bayes' rule using the pdfs in Eq. (6.25) and (6.27) we get (Press, 1982):

$$[\mathbf{K}|\mathbf{Y}] \sim \text{MatrixT}_{n_x, n_d} \left(\hat{\mathbf{K}}, \hat{\Sigma}_{\mathbf{d}}^{-1}, \hat{\Sigma}_{\mathbf{r}}, n_e - (n_x + n_d - 1) \right) \quad (6.28)$$

An EnKF updating scheme is then defined by replacing the Kalman gain matrix in Eq. (6.10) with independent realisations from the matrix- t distribution in Eq. (6.28).

6.3.3 The Kalman Gain Posterior Distributions

Similar to the approach in Section 6.3.2, we can derive the posterior distribution based on a fully specified conjugate prior model for \mathbf{K} and $\Sigma_{\mathbf{r}}$. That is (Press, 1982):

$$f(\mathbf{K}, \Sigma_{\mathbf{r}}) = f(\mathbf{K}|\Sigma_{\mathbf{r}})f(\Sigma_{\mathbf{r}}),$$

with $[\mathbf{K}|\Sigma_{\mathbf{r}}] \sim \text{MatrixG}_{n_x, n_d}(\Gamma_{\mathbf{K}}, \Theta, \Sigma_{\mathbf{r}})$, and $\Sigma_{\mathbf{r}} \sim W_{n_x}^{-1}(\Psi_{\mathbf{r}}, \nu)$. Such an approach does, however, make it less clear how the unknown hyperparameters should be specified. In addition, the approach outlined in Section 6.3.1 makes the posterior distribution consistent with the posterior distribution using a non-informative prior when $n_e \rightarrow \infty$. Moreover, the presented approach increases the degrees of freedom in the prior covariance matrix of \mathbf{K} from $\frac{1}{2}(n_x(n_x + 1) + n_d(n_d + 1))$ to $\frac{1}{2}(n_x + n_d)(n_x + n_d + 1)$ when directly defining a prior distribution for the Kalman gain. This is due to the structure of the Kronecker product (Drèze and Richard, 1983; Press, 1982).

It should be noted that choosing conjugate or Jeffreys' prior distributions in Bayesian linear regression is not without issues (Broemeling, 1985; Brown, 1993; Dawid, 1988; Press, 1982; Robert, 2001). In this study, however, we have adopted these priors mainly because of their analytical properties. Of course we might instead consider approaches based on generalised natural conjugate priors (Press, 1982), or selecting non-conjugate prior distributions (Brown et al., 1999).

Because the same prior distributions are used in the HEnKF (Myrseth and Omre, 2010a) and the approach outlined in Section 6.3.1, we should expect that there is a close connection between the Bayesian regression approach based on the informative prior and the HEnKF. Indeed, there are many similarities between the two approaches because the covariance hyperparameter matrix, $\Psi_{\mathbf{y}}^c$, is used in both posterior distributions. The major difference, however, lies in the sampling of the respective Kalman gain matrices. For the HEnKF we start by sampling $\Sigma_{\mathbf{y}}^{(i)} \sim W_{n_y}^{-1}(\Psi_{\mathbf{y}}^c, \nu^c)$, thus generating a realisation of the Kalman gain as $\mathbf{K}^{(i)} = \Sigma_{\mathbf{x}, \mathbf{d}}^{(i)} \Sigma_{\mathbf{d}}^{(i)-1}$. For $\mathbf{K} \sim \text{MatrixT}_{n_x, n_d} \left(\Gamma_{\mathbf{K}}^c, (\Psi_{\mathbf{d}}^c)^{-1}, \Psi_{\mathbf{x}|\mathbf{d}}^c, \nu^c - n_x + 1 \right)$, however, $\mathbf{K}^{(i)} = \Gamma_{\mathbf{K}}^c + \Gamma_{\mathbf{K}}^{(i)}$, where $\Gamma_{\mathbf{K}}^c$ is the posterior mean $E[\mathbf{K}^{(i)}]$, and $\Gamma_{\mathbf{K}}^{(i)} \in \mathbb{R}^{n_x \times n_d}$ is a random matrix, given as a product of the square root of a matrix following the inverted Wishart distribution and a matrix following the matrix Gaussian distribution as explained in Result 6, Appendix 6-E. Hence, the latter approach will have improved computational properties compared with the HEnKF in the general case, because we do not need to invert the matrix $\Sigma_{\mathbf{d}}^{(i)}$ for each realisation. Note, however, that the updated realisations based on the HEnKF updating scheme are expected to be less coupled. This follows because we avoid using the constant matrix $\Gamma_{\mathbf{K}}^c$ for all updated ensemble members, which for the Bayesian approaches considered here involves the empirically estimated covariance matrix $\hat{\Sigma}_{\mathbf{y}}$.

6.3.4 Approximate Matrix- t Distribution

When we consider a non-informative prior distribution, the posterior distribution of \mathbf{K} , defined in Eq. (6.28), will not exist for $n_e < n_x + n_d$. Moreover, if we consider high dimensional reservoir models, generating realisations from the matrix- t distribution can be highly computationally demanding. We will therefore consider an approximate version of the singular matrix- t distribution, inspired by Principal Component Analysis (PCA) (Mardia et al., 1979) and shrinkage methods in a multivariate regression setting (Brown, 1993; Hastie et al., 2009). This is motivated by the property that the matrix- t distribution is closed under linear transformations $\mathbf{U} = \mathbf{A}\mathbf{T}\mathbf{B}$, for constant matrices \mathbf{A} and \mathbf{B} of proper dimensions (Gupta and Nagar, 2000).

As shown in Díaz-García and Gutiérrez-Jáimez (2006), for the case with the Jeffreys' prior distribution with $\text{rank}(\hat{\Sigma}_{\mathbf{d}}) = \text{rank}(\hat{\Sigma}_{\mathbf{r}}) = n_e - 1$, $[\mathbf{K}|\mathbf{Y}] \sim \text{MatrixT}_{n_x, n_d}^{n_e - 1, n_e - 1} \left(\hat{\mathbf{K}}, \hat{\Sigma}_{\mathbf{d}}^-, \hat{\Sigma}_{\mathbf{r}}, 1 \right)$, where the singular matrix- t distribution is defined in Appendix 6-F and $(\hat{\Sigma}_{\mathbf{d}})^-$ is the Moore-Penrose inverse (Strang, 1988). To ensure that the degrees of freedom is larger than one, we there-

fore propose to approximate the singular matrix- t variate distribution by selecting the dimension of the respective reduced order subspaces p and q , smaller than the rank of $\hat{\Sigma}_d$ and $\hat{\Sigma}_r$. This corresponds to PCA, where the assumption is that a small set of variables, termed principal components, explain most of the variability in the data (Mardia et al., 1979). To improve the predictive capabilities of the regression model (Sætrum and Omre, 2010), we can replace the MLE estimate of the Kalman gain matrix in Eq. (6.18) with e.g. the Principal Component Regression (PCR) estimate: $\hat{K}_{\text{PCR}} = \mathbf{X} \mathbf{V}_D^q \mathbf{S}_D^{q-1} \mathbf{U}_D^{qT}$. Here $\mathbf{D} \approx \mathbf{U}_D^q \mathbf{S}_D^q \mathbf{V}_D^{qT}$ is the truncated Singular Value Decomposition (SVD) of \mathbf{D} , keeping only the first q components corresponding to the q largest singular values (Golub and van Loan, 1996). This implies that for the approximate matrix- t distribution based on PCR,

$$\mathbf{K} \sim \text{MatrixT}_{n_x, n_d}^{p, q} \left(\hat{K}_{\text{PCR}}, \left(\hat{\Sigma}_d^q \right)^-, \hat{\Sigma}_r^p, n_e - (q - p - 1) \right), \quad (6.29)$$

where $\hat{\Sigma}_d^q$ and $\hat{\Sigma}_r^p$ denotes the low rank approximation of the two matrices based on a truncated eigenvalue decomposition keeping $q \leq \min\{n_d, n_e - 1\}$ and $p \leq \min\{n_x, n_e - 1\}$ non-zero eigenvalues respectively. Alternative estimators for the Kalman gain matrix based on dimension reduction techniques such as Partial Least Squares Regression (PLSR) (R anner et al., 1994; Wold, 1975), can be used as well. Finally note that we use Cross-Validation (CV) to select the subspace dimensions p and q (Seber and Lee, 2003; S etrom and Omre, 2010).

6.3.5 Computational Properties

As shown in Appendix 6-E, we can generate a realisation from the matrix- t distribution using $n_x^2 n_d / 2 + n_d^2 n_x / 2 + (2/3 + 1/3) n_x^3 + \mathcal{O}(n_x \max\{n_x, n_d\})$ floating point operations (flops), when the Cholesky factorisation of the respective matrices are given. Here the first two terms are the number of flops required for multiplying an upper diagonal matrix with a full matrix of dimension $n_x \times n_d$, and $(2/3 + 1/3) n_x^3$ are the leading terms for computing the matrix inverse based on back substitution and multiplication of two upper triangular matrices (Golub and van Loan, 1996). For a singular matrix- t distributed random matrix, the computational demands are $\mathcal{O}(n_x n_d \max\{p, q\})$, given the SVD of the respective matrices and the estimated Kalman gain matrix.

When considering the suggested procedure in an EnKF setting, however, only $n_d^2 + 2n_x n_d + 2n_x^2 + \mathcal{O}(\max\{n_x, n_d\})$ flops are required to update

one realisation in the fully specified prior setting. For the singular matrix- t distributed Kalman gain, the computational demand is $\mathcal{O}(n_e \max\{n_x, n_d\})$ flops, which is the same as for the classical EnKF scheme presented in Section 6.2.1 for the rank deficient case (Evensen, 2003). The memory requirement for the informative prior case is $\mathcal{O}(\max\{n_x^2, n_d^2\})$, which implies that approximate techniques such as localisation (Evensen, 2007) must be used when n_x or n_d is large. For the singular matrix- t distribution defined through the non-informative Jeffreys' prior, the memory requirement is $\mathcal{O}(n_e \max\{n_x, n_d\})$, which corresponds to the memory requirement of the classical EnKF.

6.4 Empirical study

To evaluate the performance of the EnKF updating schemes based on the Bayesian regression techniques presented in Sections 6.3.1 and 6.3.2, we consider two synthetic case studies, similar to the ones used in Myrseth and Omre (2010b). Here the unknown state vector of interest $\mathbf{x}_k \in \mathbb{R}^{100 \times 1}$, is defined for $k = 0, \dots, 10$, where $x_{j,k}$ denotes the variable of interest at time step k and location $j \in \{1, \dots, 100\}$. The purpose of this study is to assimilate observed data, \mathbf{d}_k^o made at time steps $0, \dots, 9$, and predict at time step 10.

The first test case, referred to as the linear case, considers a Gaussian prior at the initial time step, a linear forward function and a linear Gaussian likelihood model:

$$\begin{aligned}\mathbf{x}_0 &\sim \text{Gauss}_{n_x}(\mathbf{0}_{n_x,1}, \boldsymbol{\Sigma}_{\mathbf{x}_0}) \\ \mathbf{x}_k &= \mathbf{A}_k \mathbf{x}_{k-1} \\ \mathbf{d}_k &= \mathbf{B} \mathbf{x}_k + \boldsymbol{\epsilon}_{\mathbf{d}_k}.\end{aligned}$$

Here an exponential covariance function defines $\boldsymbol{\Sigma}_{\mathbf{x}_0}$, and \mathbf{A}_k defines a linear smoothing envelope moving from left to right in time over the grid domain. The error term of the likelihood model is Gaussian, $\boldsymbol{\epsilon}_{\mathbf{d}_k} \sim \text{Gauss}_{n_d}(\mathbf{0}_{n_d,1}, \mathbf{I}_{n_d})$ and $\mathbf{B} \in \mathbb{R}^{n_d \times n_x}$ is defined as a sparse matrix, so that $d_i = \sum_{l=-1}^1 x_{j+l}$, $i = 1, \dots, n_d$, at $n_d = 13$ different grid locations $j \in \{1, \dots, 100\}$. Further details can be found in Sætrum and Omre (2010).

The second test case, referred to as the non-linear case, has the same prior and likelihood model as the linear case above, with a forward model defined through the non-linear function:

$$\mathbf{x}_k = c \mathbf{A}_k (\mathbf{x}_{k-1} + \arctan(\mathbf{x}_{k-1})).$$

Here $c = 0.8$ is a scaling factor which ensures approximate alignment of the variances for the non-linear and linear case, and the functional $\arctan(\cdot)$ acts on the argument element by element (Myrseth and Omre, 2010b).

Figure 6.2, displays the reference state vectors, $\mathbf{x}_9^{\text{True}}$ and $\mathbf{x}_{10}^{\text{True}}$, for the two test cases. Notice the smoother behaviour in the left part of the grid nodes owing to the construction of \mathbf{A}_k . Furthermore, the state vector contains both dynamic and static variables, where the dynamic variables are located in grid nodes 1 – 55 and the static variables are at nodes 56-100.

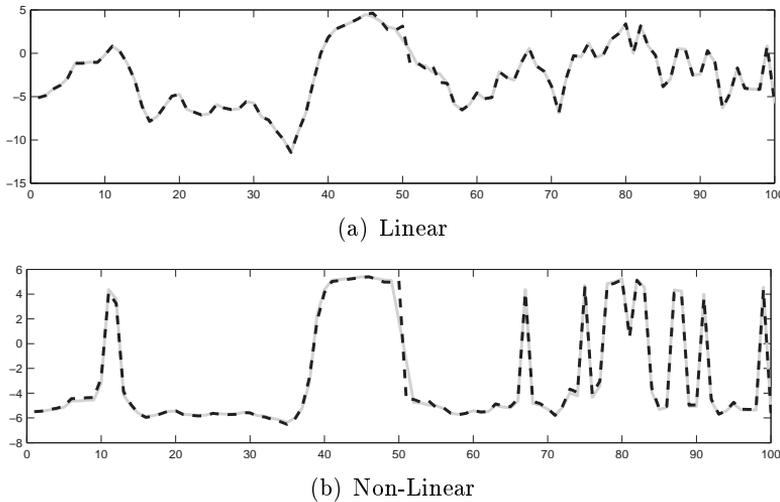


Figure 6.2: Reference $\mathbf{x}_9^{\text{True}}$ (dashed, black) and $\mathbf{x}_{10}^{\text{True}}$ (solid, grey) for the linear and non-linear models considered.

For both the linear and non-linear case the following EnKF updating schemes are considered:

- Classical EnKF: Estimated Kalman gain matrix computed based on Eq. (6.18), using the correct Σ_{ϵ_d} .
- CP-EnKF: Kalman gain matrix generated independently for each ensemble member based on the matrix- t distribution in Eq. (6.21), obtained using a Conjugate Prior (CP) distribution on the mean and covariance as described in Section 6.3.1.
- JP-PLSR-EnKF: Kalman gain matrix generated independently for each ensemble member using the approximate matrix- t distribution based on the non-informative Jeffreys' prior, described in Section 6.3.4.

In addition, $\hat{\mathbf{K}}$ is replaced with a shrinkage regression estimate based on PLSR, where the dimension of the respective subspaces p and q are selected based on 10-fold CV, minimising the Predictive Error Sum of Squares (PRESS) (Hastie et al., 2009).

We consider two different ensemble sizes, $n_e = 20$ and $n_e = 100$. In addition, the initial ensembles are identical for all three updating schemes. For the CP-EnKF updating scheme we select the prior model hyperparameters $\boldsymbol{\eta}_x = \boldsymbol{\mu}_x$, $\xi = 0.0001$, $\boldsymbol{\Psi}_x = 10 \times \boldsymbol{\Sigma}_x$ and $\nu = n_x + n_d + 3$, with $\boldsymbol{\Psi}_{x,d} = \boldsymbol{\Psi}_x \mathbf{B}^T$ and $\boldsymbol{\Psi}_d = 10 \times \boldsymbol{\Sigma}_d$. Thus, preserving the structure of the reference model in $\boldsymbol{\eta}_y$ and $\boldsymbol{\Psi}_y$, although the uncertainty regarding the size of $\boldsymbol{\eta}_y$ and $\boldsymbol{\Psi}_y$ is high.

6.4.1 Linear Case

Because we consider a Gauss-linear model, the prediction mean $E[\mathbf{x}_{10}^u]$, and 95% prediction interval are analytically obtainable using the KF recursions, see Figure 6.3a. Notice that the reference state vector is centred around the prediction mean, with reduced uncertainty in the 95% prediction intervals in the neighbourhood of the observation sites. The results for the three different EnKF updating schemes are shown in Figures 6.3b through g.

As seen from Figure 6.3b, the Classical EnKF is able to produce reliable results for $n_e = 100$. Both the estimated posterior mean and prediction interval matches the analytical KF solution fairly well, albeit the ensemble average is not as smooth as the theoretical one. Moreover, there is a tendency to underestimate the prediction variance, as expected from Eq. (6.19). For $n_e = 20$, the effects caused by increased estimation uncertainty and coupling of the ensemble members are more prominent. This results in an estimated posterior mean and prediction interval deviating dramatically from the KF solution, as expected from the discussion in Section 6.2.2.

The CP-EnKF updating scheme provides reasonable estimates of the posterior mean and prediction interval for $n_e = 100$, as seen in Figure 6.3. However, the bounds of the prediction interval is slightly increased at certain grid locations compared with the KF solution based on $\boldsymbol{\Sigma}_{\mathbf{x}^u|\mathbf{d}}$. Because we are sampling the Kalman gain matrix, rather than using a common plug-in estimate, we should expect to see this behaviour for $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}^c}$, in light of the theoretical results in Section 6.2.2. For $n_e = 20$ the width of the estimated prediction interval is reduced, and appears more narrow than the KF prediction interval. Because the matrix- t distribution for the Kalman gain matrix in Eq. (6.21) requires the empirical covariance matrix as a model parameter, the ensemble members will be somewhat coupled. However, this coupling is reduced compared with the Classical EnKF updating scheme, leading to

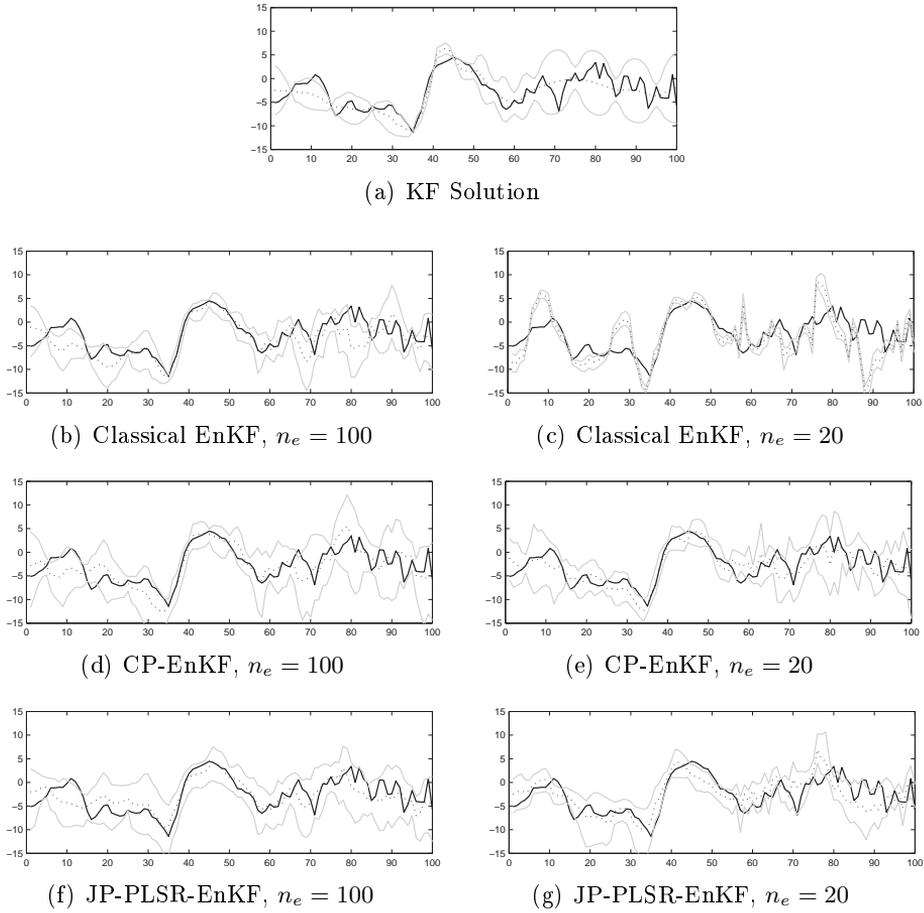


Figure 6.3: Results obtained when running four different EnKF updating schemes on the linear case with two different ensemble sizes. The figure displays the reference $\mathbf{x}_{10}^{\text{True}}$ (solid), the ensemble mean (dotted, black) and the estimated 95% confidence bounds of the prediction interval (solid, grey).

improved estimates of both the prediction mean and interval.

Both the estimated posterior mean and prediction interval resembles the analytical KF solution for the JP-PLSR-EnKF updating scheme with $n_e = 100$, as seen in Figure 6.3f. The smooth behaviour of the posterior mean, $\boldsymbol{\mu}_{\mathbf{x}^c}$ is present, with reduced uncertainty in the prediction variance around the observation sites. For $n_e = 20$, however, the ensemble average appears to be less smooth, and the prediction intervals are reduced. The results are comparable to the CP-EnKF results, and appears to be much better than the ones obtained with the Classical EnKF scheme.

To further quantify the performance of the three EnKF updating schemes, the case study is rerun 100 times with different initial ensembles. We consider the Root Mean Squared Error (RMSE) of the ensemble average to the analytical posterior mean and the percentage of the reference solution covered by the estimated prediction interval. Table 6.1 displays the results, together with the estimated RMSE and coverage computed when the initial ensemble is run through the forward model without conditioning on the observed data, referred to as the No Updating scheme. Ideally, the RMSE should be small and the Coverage should be close to 95 %. The former represents accuracy while the latter is a measure for the reliability of the empirically estimated prediction intervals. For $n_e = 100$, the RMSE decreases significantly for all three updating schemes, relative to the RMSE of the No Updating scheme. The CP-EnKF updating scheme shows the largest improvement with a 62% decrease in the RMSE, although it is not found to be significantly better than the Classical EnKF updating scheme with a p -value of 0.92. The hypothesis test is based on the traditional two sample t -statistic with unequal variance (Casella and Berger, 2002). The RMSE of the JP-PLSR-EnKF updating scheme is slightly larger compared with the RMSE of the other two methods.

As expected from the discussion in Section 6.2.2, sampling the Kalman gain matrix, as done in the CP-EnKF and JP-PLSR-EnKF updating schemes, will increase the width of the prediction intervals, whilst it will be narrower for the Classical EnKF updating scheme. For this case with $n_e = 100$ we find the coverage the CP-EnKF updating scheme and the No Updating scheme to be almost equal, and the value is close to 95% as it should be. Testing for equivalence between the coverage of the CP-EnKF and JP-PLSR-EnKF updating schemes, leads to a p -value of 0.007, and hence should be rejected at a 99% level of confidence.

For $n_e = 20$, the CP-EnKF updating scheme again leads to the largest improvement in the RMSE relative to the No Updating scheme with a 63 percent decrease. The JP-PLSR-EnKF updating scheme shows the second

n_e	Scheme	RMSE	Coverage (%)
100	No Updating	4.60 ± 0.14	95.1 ± 2.2
100	Classical EnKF	1.77 ± 0.30	81.1 ± 5.8
100	CP-EnKF	1.74 ± 0.28	95.3 ± 3.2
100	JP-PLSR-EnKF	1.91 ± 0.25	88.4 ± 3.8
20	No Updating	4.68 ± 0.27	82.3 ± 4.2
20	Classical EnKF	3.89 ± 1.05	23.1 ± 6.3
20	CP-EnKF	1.73 ± 0.29	85.0 ± 4.3
20	JP-PLSR-EnKF	3.02 ± 0.52	63.8 ± 9.4

Table 6.1: Estimated Root Mean Squared Error (RMSE) of the posterior mean and coverage of the reference solution in the estimated 95% prediction intervals, plus-minus one standard deviation, for the linear case based on 100 different initial ensembles.

largest improvement with a 35 percent decrease, whilst the Classical EnKF updating scheme has a 17 percent decrease in the RMSE. We find all three updating schemes to give significantly different RMSE values at a 99.999% level of confidence.

The Classical EnKF updating scheme has a coverage of only 23%, which demonstrates a dramatic underestimation of the prediction intervals caused by the coupling of the ensemble members. Furthermore, we observe that the coverage of the CP-EnKF updating scheme is found to be significantly larger than the coverage of the Classical EnKF and JP-EnKF updating schemes. However, the coverage of the No Updating scheme is not found to be significantly different from the CP-EnKF updating scheme with a p -value of 0.37. It is worth noticing that the coverage for both schemes is significantly smaller than the ideal value of 95%. This illustrates the importance of carefully selecting the initial ensemble for small ensemble sizes, n_e . The methods described in Evensen (2007) and Oliver and Chen (2009) outline remedies for this problem. However, we do not pursue these approaches in this paper.

We obtain the best results for the linear problem using the CP-EnKF updating scheme, but the results appear to rely heavily on the selected prior hyperparameters. If the model is rerun with a poorly selected prior model, the CP-EnKF updating scheme provides terrible results, with a large increase in the RMSE and large decrease in the coverage. This is particularly true for $n_e = 20$, although we do not show these results here. For high dimensional

reservoir evaluation problems, the task of selecting the prior hyperparameters is far from trivial. In addition, because of high computational and memory demands, the CP-EnKF updating scheme will require additional approximations, such as enforcing sparsity in the model parameters through localisation (Evensen, 2007). This suggests that the JP-PLSR-EnKF updating scheme is the most robust alternative, although the CP-EnKF scheme with a carefully chosen prior model performed best for the empirical evaluation considered here.

6.4.2 Non-Linear Case

For the non-linear case, analytical tractability of the posterior distribution is lost. Hence, the results obtained with the Classical EnKF with $n_e = 100000$, displayed in Figure 6.4a, are used for comparison. Again, we consider the three updating schemes outlined above, with results shown in Figure 6.4b through g.

Similar to the linear case, the Classical EnKF produces fairly good results for $n_e = 100$, whilst the results are considerably poorer for $n_e = 20$, with a highly fluctuating posterior mean and an ensemble almost collapsing into a single realisation. For both the CP-EnKF and JP-PLSR-EnKF updating schemes we observe reasonably good estimates of the posterior mean and prediction interval, with similar characteristics as seen in the linear case.

To quantify the performance, the three schemes are rerun 100 times using different initial ensembles. The results are summarised in Table 6.2 in a similar format as Table 6.1. For $n_e = 100$, all three updating schemes shows a significant improvement in the RMSE compared with the initial ensemble. The Classical EnKF scheme produce the best results with a 68 percent decrease. However, we do not find the RMSE to be significantly better than the RMSE of the other two updating schemes at a 99% level of confidence. The CP- and JP-PLSR-EnKF schemes have a coverage significantly lower than the No Updating scheme and 95% with p -values of 0.03 and 0.02 respectively. In addition, the Classical EnKF scheme has a coverage which is significantly smaller than all the other schemes.

For $n_e = 20$, the CP-EnKF scheme has by far the largest decrease in RMSE compared to the No Updating scheme, with a 62% reduction. The JP-PLSR-EnKF scheme shows a 41% reduction, whilst the Classical EnKF scheme has a 30% reduction. We find all RMSE values to be significantly different at a 99% level of confidence. The coverage of the CP-EnKF scheme is smaller than the coverage of the No Updating scheme, but not significantly so. However, both schemes have a coverage which is significantly smaller than the ideal value of 95%. The JP-PLSR-EnKF scheme displays a cover-

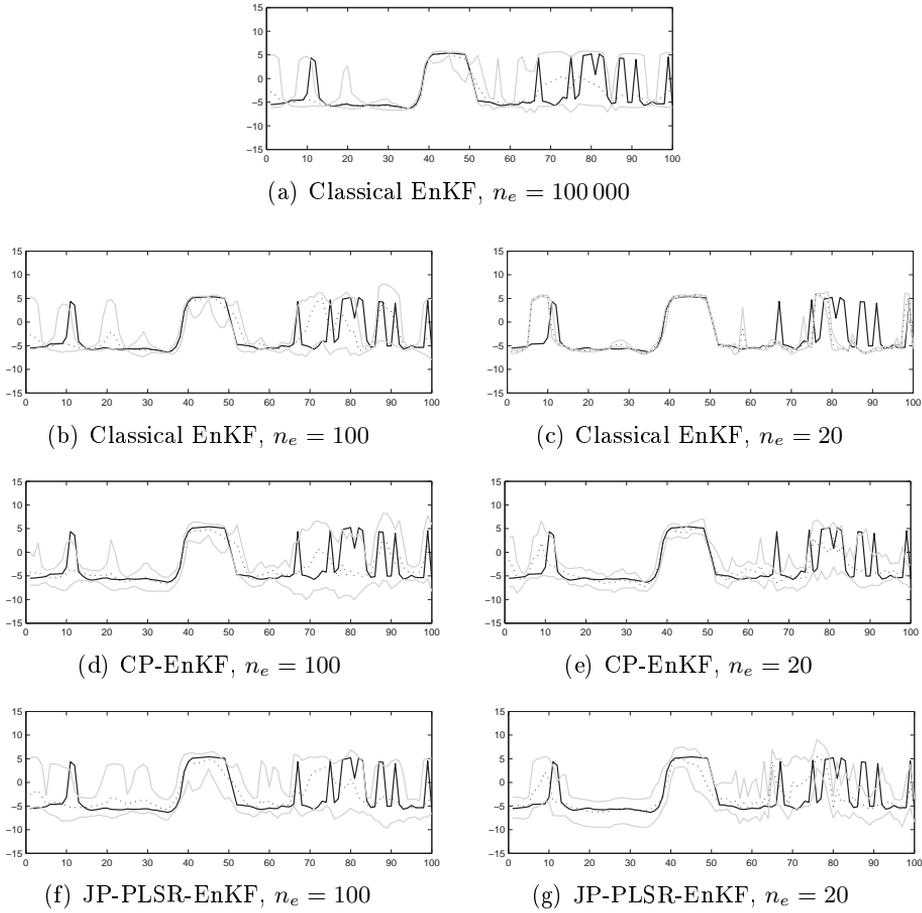


Figure 6.4: Results obtained when running four different EnKF updating schemes on the linear case with two different ensemble sizes. The figure displays the reference $\mathbf{x}_{10}^{\text{True}}$ (solid), the ensemble mean (dotted, black) and the estimated 95% confidence bounds of the prediction interval (solid, grey).

n_e	Scheme	RMSE	Coverage (%)
100	No Updating	4.45 ± 0.15	95.0 ± 1.9
100	Classical EnKF	1.42 ± 0.28	81.8 ± 5.1
100	CP-EnKF	1.50 ± 0.24	90.2 ± 2.7
100	JP-PLSR-EnKF	1.57 ± 0.21	90.0 ± 2.9
20	No Updating	4.57 ± 0.32	83.2 ± 4.5
20	Classical EnKF	3.21 ± 0.59	28.3 ± 7.4
20	CP-EnKF	1.73 ± 0.32	80.1 ± 4.0
20	JP-PLSR-EnKF	2.72 ± 0.46	70.4 ± 8.7

Table 6.2: Estimated Root Mean Squared Error (RMSE) of the posterior mean and coverage of the reference solution in the estimated 95% prediction intervals, plus-minus one standard deviation, for the non-linear case based on 100 different initial ensembles.

age of 70%. As expected, the Classical EnKF scheme has a dramatically low coverage of 28%. The poor performance of the Classical EnKF scheme is caused by the coupling of the ensemble members, which increase the variance of the ensemble average and make the empirical variance biased downward compared with prediction variance. Note that small ensemble sizes is a relative statement, it is actually the size of the difference between the ensemble size and number of observations, $n_e - n_d$, that is important. Again note that the apparent success of the CP-EnKF updating scheme is highly dependent on the prior hyperparameters selected, similar to what was seen in the linear case. This leave the JP-PLSR-EnKF updating scheme as the most promising one.

6.5 Conclusions

In this paper we have formulated alternative EnKF updating schemes based on classical results known from Bayesian regression. Rather than using a common estimate of the Kalman gain matrix in the update, each ensemble member is updated using a unique Kalman gain matrix generated from an analytically tractable matrix variate probability distribution. Furthermore, we have developed theoretical results valid for the EnKF updating scheme on a Gauss-linear model, where we prove that the updated ensemble members will be positively correlated. This will necessarily lead to an underestimation of the empirical prediction variance when the conditioned ensemble is used

for estimation. Explicit expressions for the bias in finite sample cases have been developed. Moreover, the results explain why we tend to see a large variability in the posterior empirical mean for small ensemble sizes. These results are troubling, because the ensemble coupling will amplify during data assimilation, and potentially lead to an ensemble collapsing into a single realisation. The suggested EnKF updating schemes based on Bayesian regression reduces the coupling between the updated ensemble members, which leads to improved predictions and associated prediction intervals.

Two alternative EnFK updating schemes are defined and empirically tested. The first one is based on a conjugate prior distribution on the unknown mean and covariance matrix (CP-EnKF). The second one uses a non-informative Jeffreys' prior distribution on a multivariate regression model combined with Partial Least Squares Regression (JP-PLSR-EnKF), leading to a singular matrix- t distributed Kalman gain matrix defined in a reduced order space. The main objective of this additional dimension reduction is to improve the Kalman gain matrix estimate and to handle large dimensional problems with small ensemble sizes. Cross-validation is used to select the subspace dimension by evaluating the predictive strength of the regression model.

The two suggested schemes were tested on two synthetic models, inspired by reservoir evaluation problems, and compared with the results obtained using the classical EnKF updating scheme. With a small ensemble size of 20, both the CP-EnKF and JP-PLSR-EnKF updating schemes performed significantly better than the classical EnKF updating scheme. The coupling of the ensemble members in the classical EnKF scheme caused the predictions and associated prediction intervals to be highly unreliable. In accordance with the theoretical results presented in this paper, increasing the ensemble size to 100 improved the results for classical EnKF updating scheme and made it comparable to the results obtain with the alternative EnKF scheme based on Bayesian regression.

The CP-EnKF scheme appeared as the most reliable EnKF approach in the empirical study, closely followed by the JP-PLSR-EnKF scheme. The classical EnKF scheme produced unreliable predictions and associated prediction intervals for small ensemble sizes relative to the number of observations. The CP-EnKF scheme is, contrary to the JP-PLSR-EnKF scheme, computationally demanding to perform and highly dependent on the specification of a representative prior model. Consequently, we recommend using the JP-PLSR-EnKF scheme for high dimensional spatiotemporal problems with a large number of observations.

Acknowledgement

This work is funded by the Uncertainty in Reservoir Evaluation (URE) initiative at NTNU.

References

- S. I. Aanonsen, G. Nævdal, D. S. Oliver, A. C. Reynolds, and B. Vallès. Ensemble Kalman filter in reservoir engineering - a review. *SPE Journal*, 14(3):393–412, 2009.
- J. L. Anderson. A local least squares framework for ensemble filtering. *Monthly Weather Review*, 131(4):634–642, 2003a.
- T. W. Anderson. *An introduction to multivariate statistical analysis*. Wiley, 3 edition, 2003b.
- T. W. Anderson. Asymptotic theory for principal component analysis. *The Annals of Mathematical Statistics*, 34(1):122–148, 1963.
- T. Bodnar and Y. Okhrin. Properties of the singular, inverse and generalized inverse partitioned Wishart distributions. *Journal of Multivariate Analysis*, 99(10):2389–2405, 2008.
- G. E. P. Box and G. C. Tiao. *Bayesian inference in statistical analysis*. Wiley, 1992.
- L. D. Broemeling. *Bayesian Analysis of Linear Models*. Marcel Dekker, Inc, 1985.
- P. J. Brown. *Measurement, Regression, and Calibration*. Oxford Science Publications, 1993.
- P. J. Brown, T. Fearn, and M. Vannucci. The choice of variables in multivariate regression: A non-conjugate Bayesian decision theory approach. *Biometrika*, 86(3):635–648, 1999.
- G. Casella and R. L. Berger. *Statistical Inference*. Duxbury, 2002.
- W. F. Caselton, L. Kan, and J. V. Zidek. Quality data networks that minimize entropy. In *Statistics in the Environmental and Earth Sciences*, pages 10–38. Wiley, 1992.
- A. P. Dawid. The infinite regress and its conjugate analysis. In *Bayesian Statistics 3*, pages 95–110. Oxford University Press, 1988.

- W. L. Deemer and I. Olkin. The Jacobians of certain matrix transformations useful in multivariate analysis. *Biometrika*, 38:345–367, 1951.
- A. Doucet, S. Godsill, and C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10:197–208, 2000.
- J. H. Drèze and J.-F. Richard. Bayesian analysis of simultaneous equation systems. In *Handbook of Econometrics*, pages 517–598. Elsevier, 1983.
- J. Díaz-García and R. Gutiérrez-Jáimez. Distribution of the generalised inverse of a random matrix and its applications. *Journal of Statistical Planning and Inference*, 136:183–192, 2006.
- J. A. Díaz-García. A note about measures and jacobians of singular random matrices. *Journal of Multivariate Analysis*, 98:960–969, 2007.
- J. A. Díaz-García and R. Gutiérrez-Jáimez. Singular matrix and matrix variate t distributions. *Journal of Statistical Planning and Inference*, 139(7):2382–2387, 2009.
- J. A. Díaz-García, R. Gutiérrez-Jáimez, and K. V. Mardia. Wishart and pseudo-Wishart distributions and some applications to shape theory. *Journal of Multivariate Analysis*, 63(1):73–87, 1997.
- G. Evensen. Using the extended Kalman filter with a multilayer quasi-geostrophic ocean mode. *J. Geophys. Res.*, 97(C11):17905–17924, 1992.
- G. Evensen. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99:10143–10162, 1994.
- G. Evensen. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, 2003.
- G. Evensen. *Data assimilation the Ensemble Kalman Filter*. Springer, 2007.
- D. E. Farrer and R. R. Glauber. Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics*, 49(1): 92–107, 1967.
- R. Furrer and T. Bengtsson. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *Journal of Multivariate Analysis*, 98(2):227–255, 2007.

- A. Gelb. *Applied Optimal Estimation*. MIT Press Cambridge, 1974.
- M. A. Girshick. On the sampling theory of roots of determinantal equations. *The Annals of Mathematical Statistics*, 10(3):203–224, 1939.
- G. Golub and C. van Loan. *Matrix Computations*. Johns Hopkins University Press, 1996.
- A. K. Gupta and D. K. Nagar. *Matrix Variate Distributions*. Chapman & Hall, 2000.
- T. Hastie, R. Tibshirani, and J. Freidman. *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. Springer, New York, 2 edition, 2009.
- P. L. Houtekamer and H. L. Mitchell. A sequential Ensemble Kalman Filter for atmospheric data assimilation. *Monthly Weather Review*, 129:123–137, 2001.
- P. L. Houtekamer and H. L. Mitchell. Data assimilation using an Ensemble Kalman Filter technique. *Monthly Weather Review*, 126:796–811, 1998.
- P. L. Houtekamer, L. Lefaiivre, J. Derome, H. Ritchie, and H. L. Mitchell. A system simulation approach to ensemble prediction. *Monthly Weather Review*, 124:1225–1242, 1996.
- A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.
- H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London (Ser. A)*, 186:453–461, 1946.
- M. C. Jones. Generating inverse wishart matrices. *Communications in Statistics - Simulation and Computation*, 14:511–514, 1985.
- D. G. Kabe. On the distribution of the regression coefficient matrix of a normal distribution. *The Australian Journal of Statistics*, 10:21–23, 1968.
- R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82(Series D): 35–45, 1960.
- G. M. Kaufman. Conditional prediction and unbiasedness in structural equations. *Econometrica*, 37(1):44–49, 1969.

- C. L. Keppenne and M. Rienecker. Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble kalman filter. *J. Marine. Sys.*, 40-41:363–380, 2003.
- A. M. Kshirsagar. Some extensions of the multivariate t -distribution and the multivariate generalization of the distribution of the regression coefficient. *Proceedings of the Cambridge Philosophical Society*, 57:80–85, 1961.
- O. Leeuwenburgh, G. Evensen, and L. Bertino. The impact of ensemble filter definition on the assimilation of temperature profiles in the tropical pacific. *Quarterly Journal of the Royal Meteorological Society*, 131(613): 3291–3300, 2005.
- J. S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer, New York, 2001.
- K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979.
- H. Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser. Dual state parameter estimation of hydrological models using Ensemble Kalman Filter. *Advances in Water Resources*, 28:135–147, 2005.
- I. Myrseth and H. Omre. Hierarchical ensemble kalman filter. *SPE Journal*, 15(2):569–580, 2010a.
- I. Myrseth, J. Sætrum, and H. Omre. Resampling the Ensemble Kalman Filter. Paper submitted for publication, 2009.
- I. B. Myrseth and Omre. The Ensemble Kalman Filter and related filters. In L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, Y. Marzouk, L. Tenorio, B. van Bloemen Waanders, and K. Willcox, editors, *Large-scale Inverse Problems and Quantification of Uncertainty*. John Wiley and Sons, 2010b.
- D. S. Oliver and Y. Chen. Improved initial sampling for the Ensemble Kalman Filter. *Computational Geosciences*, 13(1):13–27, 2009.
- J. S. Press. *Applied multivariate analysis: using Bayesian and frequentist methods of inference*. Robert E. Krieger Publishing Company, 2 edition, 1982.
- C. P. Robert. *The bayesian choice*. Springer, 2 edition, 2001.
- S. M. Ross. *Probability Models*. Academic Press, 2003.

- S. Ranner, F. Lindgren, P. Gelandi, and S. Wold. A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. *Journal of Chemometrics*, 8:111–125, 1994.
- W. Sacher and P. Bartello. Sampling errors in Ensemble Kalman filtering. Part I: Theory. *Monthly Weather Review*, pages 3035–3049, 2008.
- S. R. Searle. *Matrix Algebra Useful for Statistics*. Wiley, 1982.
- G. A. F. Seber and A. J. Lee. *Linear Regression Analysis*. Wiley, 2003.
- W. B. Smith and R. R. Hocking. Algorithm AS 53: Wishart variate generator. *Journal of the Royal Statistical Society. Series C, Applied Statistics*, 21(3):341–345, 1972.
- G. Strang. *Linear Algebra And Its Applications*. Thomson Learning, 1988.
- J. Setrom and H. Omre. Ensemble Kalman filtering with shrinkage regression techniques. *Computational Geosciences*, 2010. In press, DOI: 10.1007/s10596-010-9196-0.
- P. J. van Leeuwen. Comments on "Data assimilation using an Ensemble Kalman Filter technique. *Monthly Weather Review*, 127:1374–1377, 1999.
- L. L. Wegge. The finite sampling distribution of least squares estimators with stochastic regressors. *Econometrica*, 39(2):241–251, 1971.
- M. West and J. Harrison. *Bayesian Forecasting and Dynamic Models*. Springer, 2 edition, 1999.
- H. Wold. Path models with latent variables: The NiPALS approach. In *Quantitative Sociology: International perspectives on mathematical and statistical model building*, pages 307–357. Academic Press, 1975.

APPENDIX

Appendix 6-A, Multivariate vector and matrix probability distributions

Definition 1 (Gaussian distribution). Consider a random vector $\mathbf{y} \in \mathbb{R}^{k \times 1}$. Then

$$\mathbf{y} \sim \text{Gauss}_k(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y),$$

means that \mathbf{y} has a *multivariate Gaussian* probability density function (pdf)

$$f(\mathbf{y}) = \frac{1}{(2\pi)^{k/2} |\boldsymbol{\Sigma}_y|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu}_y)^T \boldsymbol{\Sigma}_y^{-1} (\mathbf{y} - \boldsymbol{\mu}_y) \right\}. \quad (\text{A-1})$$

The mean and covariance of \mathbf{y} are $E[\mathbf{y}] = \boldsymbol{\mu}_y$ and $\text{Cov}(\mathbf{y}) = \boldsymbol{\Sigma}_y$ respectively.

Definition 2 (Matrix Gaussian distribution). Consider the random matrix $\mathbf{X} \in \mathbb{R}^{a \times b}$. Then

$$\mathbf{X} \sim \text{MatrixG}_{a,b}(\mathbf{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma}),$$

where $\mathbf{M} \in \mathbb{R}^{a \times b}$, $\boldsymbol{\Omega} \in \mathbb{R}^{b \times b}$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{a \times a}$, means that \mathbf{X} follows the multivariate *matrix Gaussian* distribution, with corresponding pdf (Gupta and Nagar, 2000)

$$f(\mathbf{X}) = \frac{|\boldsymbol{\Sigma}|^{-b/2} |\boldsymbol{\Omega}|^{-a/2}}{(2\pi)^{ab/2}} \exp \left\{ -\frac{1}{2} \text{tr} (\boldsymbol{\Omega}^{-1} (\mathbf{X} - \mathbf{M})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \mathbf{M})) \right\}, \quad (\text{A-2})$$

where $\text{tr}(\cdot)$ denotes the trace operator. The mean of \mathbf{X} is $E[\mathbf{X}] = \mathbf{M}$ and the covariance matrix of the stacked matrix, $\text{Vec}(\mathbf{X})$, is $\text{Cov}(\text{Vec}(\mathbf{X})) = (\boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})$, where (\otimes) is the Kronecker product (Gupta and Nagar, 2000).

Definition 3 (Wishart distribution). Let

$$\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\nu)} \stackrel{\text{i.i.d.}}{\sim} \text{Gauss}_p(\mathbf{0}_{p,1}, \boldsymbol{\Sigma})$$

with $\nu > p$. Then

$$\mathbf{M} = \sum_{i=1}^{\nu} \mathbf{x}_i \mathbf{x}_i^T,$$

is said to have the *Wishart* distribution, $\mathbf{M} \sim W_p(\boldsymbol{\Sigma}, \nu)$, with pdf:

$$f(\mathbf{M}) = \frac{|\mathbf{M}|^{(\nu-p-1)/2} \exp \left\{ -\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{M}) \right\}}{2^{\nu p/2} |\boldsymbol{\Sigma}|^{\nu/2} \Gamma_p(\nu)}.$$

Here

$$\Gamma_p(\nu) = \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma((\nu + 1 - i)/2), \quad (\text{A-3})$$

with

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx. \quad (\text{A-4})$$

The mean of \mathbf{M} is $\mathbb{E}[\mathbf{M}] = \nu \boldsymbol{\Sigma}$ (Mardia et al., 1979).

Definition 4 (Inverse Wishart distribution). Let $\mathbf{M} \sim W_k(\boldsymbol{\Sigma}, \nu)$, then $\mathbf{Z} = \mathbf{M}^{-1}$ is said to have the *inverse Wishart* distribution, $\mathbf{Z} \sim W_k^{-1}(\boldsymbol{\Sigma}, \nu)$, with pdf (Mardia et al., 1979):

$$f(\mathbf{Z}) = \frac{|\mathbf{Z}|^{-(\nu+k+1)/2} \exp\left\{-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{Z}^{-1})\right\}}{2^{\nu k/2} |\boldsymbol{\Sigma}|^{\nu/2} \Gamma_k(\nu)}. \quad (\text{A-5})$$

The mean of \mathbf{Z} is $\mathbb{E}[\mathbf{Z}] = \boldsymbol{\Sigma}^{-1}/(\nu - k - 1)$ (Gupta and Nagar, 2000).

Definition 5 (Matrix- t distribution). Consider the random matrix $\mathbf{T} \in \mathbb{R}^{a \times b}$. Then

$$\mathbf{T} \sim \text{MatrixT}_{a,b}(\mathbf{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma}, \nu),$$

where $\mathbf{M} \in \mathbb{R}^{a \times b}$, $\boldsymbol{\Omega} \in \mathbb{R}^{b \times b}$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{a \times a}$, means that \mathbf{T} follows the multivariate *matrix- t* distribution, with corresponding pdf (Gupta and Nagar, 2000)

$$f(\mathbf{T}) = C_t |\mathbf{I}_b + \boldsymbol{\Omega}^{-1}(\mathbf{T} - \mathbf{M})^T \boldsymbol{\Sigma}^{-1}(\mathbf{T} - \mathbf{M})|^{-(\nu+a+b-1)/2}, \quad (\text{A-6})$$

where

$$C_t = \frac{\Gamma_a((\nu + a + b - 1))}{\pi^{ab/2} \Gamma_a((\nu + a - 1))} |\boldsymbol{\Sigma}|^{-b/2} |\boldsymbol{\Omega}|^{-a/2}.$$

The mean of \mathbf{T} is $\mathbb{E}[\mathbf{T}] = \mathbf{M}$ and the covariance matrix of the stacked matrix, $\text{Vec}(\mathbf{T})$, is $\text{Cov}(\text{Vec}(\mathbf{T})) = (\boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})/(\nu - 2)$ (Gupta and Nagar, 2000).

Appendix 6-B, Expectation, covariance and coupling of the ensemble

The EnKF updating scheme is defined as: $\mathbf{x}^{c(i)} = \mathbf{x}^{u(i)} + \hat{\mathbf{K}}(\mathbf{d}^o - \mathbf{d}^{(i)})$. Using Results 3 and 4 in Appendix 5-D, the posterior mean and covariance of $\mathbf{x}^{c(i)}$ are then given as:

$$\begin{aligned} \mathbb{E}[\mathbf{x}^c] &= \mathbb{E}\left[\mathbb{E}[\mathbf{x}^c | \hat{\mathbf{K}}]\right] \\ &= \boldsymbol{\mu}_{\mathbf{x}|\mathbf{d}} + (\boldsymbol{\Gamma}_{\mathbf{K}} - \mathbf{K})(\mathbf{d}^o - \boldsymbol{\mu}_{\mathbf{d}}) \end{aligned} \quad (\text{B-1})$$

and

$$\begin{aligned} \text{Cov}(\mathbf{x}^c) &= \text{Cov}\left(\mathbb{E}[\mathbf{x}^c | \hat{\mathbf{K}}]\right) + \mathbb{E}\left[\text{Cov}(\mathbf{x}^c | \hat{\mathbf{K}})\right] \\ &= \Sigma_{\mathbf{x}|d} + (\Gamma_{\mathbf{K}} - \mathbf{K})\Sigma_d(\Gamma_{\mathbf{K}} - \mathbf{K})^T \\ &\quad + \mathbb{E}\left[(\hat{\mathbf{K}} - \Gamma_{\mathbf{K}})(\Sigma_d + \Delta_{d^o})(\hat{\mathbf{K}} - \Gamma_{\mathbf{K}})^T\right]. \end{aligned} \quad (\text{B-2})$$

Here $\mu_{\mathbf{x}|d} = \mu_{\mathbf{x}} + \mathbf{K}(d^o - \mu_d)$, $\mathbb{E}[\hat{\mathbf{K}}] = \Gamma_{\mathbf{K}}$, $\Sigma_{\mathbf{x}|d} = \Sigma_{\mathbf{x}} - \mathbf{K}\Sigma_d\mathbf{K}^T$ and $\Delta_{d^o} = (d^o - \mu_d)(d^o - \mu_d)^T$, where $\mathbf{K} = \Sigma_{\mathbf{x},d}\Sigma_d^{-1}$ is the true Kalman gain matrix. The covariance between two ensemble members $\mathbf{x}^{c(i)}$ and $\mathbf{x}^{c(j)}$ is:

$$\begin{aligned} \text{Cov}(\mathbf{x}^{c(i)}, \mathbf{x}^{c(j)}) &= \mathbb{E}\left[\mathbb{E}\left[(\mathbf{x}^{c(i)} - \mathbb{E}[\mathbf{x}^{c(i)}])(\mathbf{x}^{c(j)} - \mathbb{E}[\mathbf{x}^{c(j)}])\right] \middle| \hat{\mathbf{K}}\right] \\ &= \mathbb{E}\left[(\hat{\mathbf{K}} - \Gamma_{\mathbf{K}})\Delta_{d^o}(\hat{\mathbf{K}} - \Gamma_{\mathbf{K}})^T\right]. \end{aligned} \quad (\text{B-3})$$

Appendix 6-C, Underestimation prediction uncertainty

Result 2. Let $\mathbf{X} \in \mathbb{R}^{n_x \times n_e}$ be a centred ensemble of realisations from a Gaussian distribution and $\hat{\Sigma}_{\mathbf{x}} = \frac{1}{n_e - 1} \mathbf{X} \mathbf{X}^T$ be the unbiased estimate of the covariance matrix of \mathbf{x} . Assuming a Gauss-linear likelihood model, $\mathbf{d} = \mathbf{B}\mathbf{x} + \epsilon_d$, where $\mathbf{B}^T \mathbf{B} = \mathbf{I}_{n_x}$ and $\epsilon_d \sim \text{Gauss}_{n_d}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_d})$, and $\mathbf{B}\Sigma_{\mathbf{x}}\mathbf{B}^T$ has n_d distinct eigenvalues, $\lambda_1 > \dots > \lambda_{n_d} > 0$, then:

$$\text{tr}\left\{\mathbb{E}\left[\hat{\Sigma}_{\mathbf{x}|d} - \Sigma_{\mathbf{x}|d}\right]\right\} < -\frac{2}{n_e - 1} \frac{\sigma^4}{(\lambda_1 + \sigma^2)^3} \sum_{i=1}^{n_d} \lambda_i^2 + \mathcal{O}(n_e^{-2}),$$

where

$$\hat{\Sigma}_{\mathbf{x}|d} = \hat{\Sigma}_{\mathbf{x}} - \hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T \left(\mathbf{B}\hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T + \sigma^2 \mathbf{I}_{n_d}\right)^{-1} \mathbf{B}\hat{\Sigma}_{\mathbf{x}}.$$

Proof. Using well known matrix algebra (Furrer and Bengtsson, 2007; Searle, 1982):

$$\begin{aligned} \text{tr}\left\{\hat{\Sigma}_{\mathbf{x}|d} - \Sigma_{\mathbf{x}|d}\right\} &= \text{tr}\left\{\left(\mathbf{B}\hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T + \sigma^2 \mathbf{I}_{n_d}\right)^{-1}\right\} \\ &\quad - \text{tr}\left\{\left(\mathbf{B}\Sigma_{\mathbf{x}}\mathbf{B}^T + \sigma^2 \mathbf{I}_{n_d}\right)^{-1}\right\}. \end{aligned}$$

Because $\mathbf{B}\hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T$ is a symmetric, positive definite matrix, we can write:

$$\mathbf{B}\hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T = \mathbf{U}\hat{\Lambda}\mathbf{U}^T,$$

where $\mathbf{U} \in \mathbb{R}^{n_d \times n_d}$ contains the n_d eigenvector of $\mathbf{B}\hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T$ and $\hat{\Lambda} \in \mathbb{R}^{n_d \times n_d}$ is a diagonal matrix containing the n_d corresponding eigenvalues, $\hat{\lambda}_1, \dots, \hat{\lambda}_{n_d}$. Because $\mathbf{B}\hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T$ share the same eigenvectors as $\left(\mathbf{B}\hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T + \sigma^2\mathbf{I}_{n_d}\right)^{-1}$, we have:

$$\text{tr} \left\{ \left(\mathbf{B}\hat{\Sigma}_{\mathbf{x}}\mathbf{B}^T + \sigma^2\mathbf{I}_{n_d} \right)^{-1} \right\} = \sum_{i=1}^{n_d} \frac{1}{\hat{\lambda}_i + \sigma^2}.$$

Proceeding with a Taylor expansion of $1/(\hat{\lambda}_i + \sigma^2)$ around the i th eigenvalue of $\mathbf{B}\Sigma_{\mathbf{x}}\mathbf{B}^T$ yields:

$$\begin{aligned} \text{tr} \left\{ \hat{\Sigma}_{\mathbf{x}|d} - \Sigma_{\mathbf{x}|d} \right\} &= \sum_{i=1}^{n_d} \frac{\sigma^4}{(\lambda_i + \sigma^2)^2} \left(\hat{\lambda}_i - \lambda_i \right) \\ &\quad - \sum_{i=1}^{n_d} \frac{\sigma^4}{(\lambda_i + \sigma^2)^3} \left(\hat{\lambda}_i - \lambda_i \right)^2 \\ &\quad + \sum_{i=1}^{n_d} \frac{\sigma^4}{(\lambda_i + \sigma^2)^4} \left(\hat{\lambda}_i - \lambda_i \right)^3 \\ &\quad - \mathcal{O}(\max_i (\hat{\lambda}_i - \lambda_i)^4) \\ &< \sum_{i=1}^{n_d} \frac{\sigma^4}{(\lambda_{n_d} + \sigma^2)^2} \left(\hat{\lambda}_i - \lambda_i \right) \\ &\quad - \sum_{i=1}^{n_d} \frac{\sigma^4}{(\lambda_1 + \sigma^2)^3} \left(\hat{\lambda}_i - \lambda_i \right)^2 \\ &\quad + \sum_{i=1}^{n_d} \frac{\sigma^4}{(\lambda_{n_d} + \sigma^2)^4} \left(\hat{\lambda}_i - \lambda_i \right)^3 \\ &\quad - \mathcal{O}(\max_i (\hat{\lambda}_i - \lambda_i)^4). \end{aligned}$$

The empirical covariance matrix, $\hat{\Sigma}_{\mathbf{x}}$, is an unbiased estimator of $\Sigma_{\mathbf{x}}$, thus

$$\mathbb{E} \left[\sum_{i=1}^{n_d} (\hat{\lambda}_i - \lambda_i) \right] = 0.$$

The result then follows from Girshick (1939):

$$\mathbb{E} \left[(\hat{\lambda}_i - \lambda_i)^2 \right] = \frac{2\lambda_i^2}{n_e - 1} + \mathcal{O}(n_e^{-2})$$

and

$$\mathbb{E} \left[(\hat{\lambda}_i - \lambda_i)^j \right] = \mathcal{O}(n_e^{-2}), \quad j \geq 3.$$

□

Remark. This result is similar to Corollary 2 in Furrer and Bengtsson (2007). However, a lower bound for the bias was not given for the multivariate case, in addition to some unclear elements in the proof (R. Furrer, personal communication, 2010). Finally note that the assumption of n_d distinct eigenvalues can be relaxed using the results in Anderson (1963). However, these results are only valid for sufficiently large ensemble sizes, n_e , contrary to the proof above which holds for any n_e .

Appendix 6-D, Properties of some multivariate distributions

Result 3.

$$\mathbb{E}[\mathbf{X}] = \mathbb{E}[\mathbb{E}[\mathbf{X}|\mathbf{Y}]]. \quad (\text{D-1})$$

Proof. For the univariate case, this is shown in Ross (2003, Chapter 3). From the definition of the expectation of a vector, or matrix (Gupta and Nagar, 2000), the result follows. □

Result 4.

$$\text{Cov}(\mathbf{X}) = \mathbb{E}[\text{Cov}(\mathbf{X}|\mathbf{Y})] + \text{Cov}(\mathbb{E}[\mathbf{X}|\mathbf{Y}]). \quad (\text{D-2})$$

Proof. The result follows using the definition of the covariance (Gupta and Nagar, 2000) of a vector or matrix together with Result 3. □

Result 5. Consider the random matrix $\Sigma \sim W_n^{-1}(\Psi^{-1}, \nu)$, where Σ and Ψ are decomposed into submatrices:

$$\Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_2 \end{bmatrix}, \quad \Psi = \begin{bmatrix} \Psi_1 & \Psi_{12} \\ \Psi_{12}^T & \Psi_2 \end{bmatrix}.$$

Further define

$$\begin{aligned} \Sigma_{1|*} &= \Sigma_1 - \Sigma_{12}\Sigma_2^{-1}\Sigma_{12}^T \in \mathbb{R}^{n_1 \times n_1} \\ \mathbf{T} &= \Sigma_{12}\Sigma_2^{-1} \in \mathbb{R}^{n_1 \times n_2} \\ \Sigma_* &= \Sigma_2 \in \mathbb{R}^{n_2 \times n_2}. \end{aligned}$$

Then

$$\begin{aligned}\boldsymbol{\Sigma}_{1|*} &\sim W_{n_1}^{-1}(\boldsymbol{\Psi}_{1|*}^{-1}, \nu) \\ \boldsymbol{T} &\sim \text{Matrix}\Gamma_{n_1, n_2}(\boldsymbol{\Gamma}, (\boldsymbol{\Psi}_*)^{-1}, \boldsymbol{\Sigma}_{1|*}, \nu - n_1 + 1) \\ \boldsymbol{\Sigma}_* &\sim W_{n_2}^{-1}(\boldsymbol{\Psi}_*^{-1}, \nu - n_1),\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{\Psi}_{1|*} &= \boldsymbol{\Psi}_1 - \boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_2^{-1}\boldsymbol{\Psi}_{12}^T \\ \boldsymbol{\Gamma} &= \boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_2^{-1} \\ \boldsymbol{\Psi}_* &= \boldsymbol{\Psi}_2.\end{aligned}$$

Proof. Because the block matrix $\boldsymbol{\Sigma}$ only depends on the three matrices $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_{12}$ and $\boldsymbol{\Sigma}_2$, we can write

$$f(\boldsymbol{\Sigma}) = f(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_{12}, \boldsymbol{\Sigma}_2).$$

As shown in Caselton et al. (1992), we can express $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$ as:

$$\begin{aligned}\boldsymbol{\Sigma} &= \begin{bmatrix} \boldsymbol{I}_{n_1} & \boldsymbol{T} \\ \mathbf{0}_{n_2, n_1} & \boldsymbol{I}_{n_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{1|*} & \mathbf{0}_{n_1, n_2} \\ \mathbf{0}_{n_2, n_1} & \boldsymbol{\Sigma}_* \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{n_1} & \boldsymbol{T} \\ \mathbf{0}_{n_2, n_1} & \boldsymbol{I}_{n_2} \end{bmatrix}^T \\ &= \begin{bmatrix} \boldsymbol{\Sigma}_{1|*} + \boldsymbol{T}\boldsymbol{\Sigma}_*\boldsymbol{T}^T & \boldsymbol{T}\boldsymbol{\Sigma}_* \\ \boldsymbol{\Sigma}_*\boldsymbol{T}^T & \boldsymbol{\Sigma}_* \end{bmatrix} \\ \boldsymbol{\Psi} &= \begin{bmatrix} \boldsymbol{I}_{n_1} & \boldsymbol{\Gamma} \\ \mathbf{0}_{n_2, n_1} & \boldsymbol{I}_{n_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Psi}_{1|*} & \mathbf{0}_{n_1, n_2} \\ \mathbf{0}_{n_2, n_1} & \boldsymbol{\Psi}_* \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{n_1} & \boldsymbol{\Gamma} \\ \mathbf{0}_{n_2, n_1} & \boldsymbol{I}_{n_2} \end{bmatrix}^T,\end{aligned}$$

where the different matrices are defined above. Now, by general properties of the determinant and the trace operator (Mardia et al., 1979, Appendix A),

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}_{1|*}^{-1} & -\boldsymbol{\Sigma}_{1|*}^{-1}\boldsymbol{T} \\ -\boldsymbol{T}^T\boldsymbol{\Sigma}_{1|*}^{-1} & \boldsymbol{\Sigma}_*^{-1} + \boldsymbol{T}^T\boldsymbol{\Sigma}_{1|*}^{-1}\boldsymbol{T} \end{bmatrix},$$

and

$$|\boldsymbol{\Sigma}| = |\boldsymbol{\Sigma}_*| \cdot |\boldsymbol{\Sigma}_{1|*}|.$$

Thus,

$$\begin{aligned}
 f(\boldsymbol{\Sigma}_1(\cdot), \boldsymbol{\Sigma}_{12}(\cdot), \boldsymbol{\Sigma}_2(\cdot)) &\propto |\boldsymbol{\Sigma}|^{-(\nu+n+1)/2} \exp\left\{-\frac{1}{2}\text{tr}(\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1})\right\} \\
 &\propto |\boldsymbol{\Sigma}_*|^{-(\nu+n+1)/2} \times |\boldsymbol{\Sigma}_{1|*}|^{-(\nu+n+1)/2} \times \\
 &\quad \exp\left\{-\frac{1}{2}\text{tr}\left(\boldsymbol{\Psi}_{1|*}\boldsymbol{\Sigma}_{1|*}^{-1} + \boldsymbol{\Psi}_*\boldsymbol{\Sigma}_*^{-1}\right)\right\} \times \\
 &\quad \exp\left\{-\frac{1}{2}\text{tr}\left(\boldsymbol{\Psi}_*(\mathbf{T} - \boldsymbol{\Gamma})^T\boldsymbol{\Sigma}_{1|*}^{-1}(\mathbf{T} - \boldsymbol{\Gamma})\right)\right\}.
 \end{aligned}$$

Then,

$$f(\boldsymbol{\Sigma}_{1|*}, \mathbf{T}, \boldsymbol{\Sigma}_*) = f(\boldsymbol{\Sigma}_1(\cdot), \boldsymbol{\Sigma}_{12}(\cdot), \boldsymbol{\Sigma}_2(\cdot)) |J|,$$

where by Deemer and Olkin (1951) and Mardia et al. (1979, Appendix B)

$$\begin{aligned}
 J &= \begin{vmatrix} \frac{\partial \boldsymbol{\Sigma}_1(\cdot)}{\partial \boldsymbol{\Sigma}_{1|*}} & \frac{\partial \boldsymbol{\Sigma}_1(\cdot)}{\partial \mathbf{T}} & \frac{\partial \boldsymbol{\Sigma}_1(\cdot)}{\partial \boldsymbol{\Sigma}_*} \\ \frac{\partial \boldsymbol{\Sigma}_{12}(\cdot)}{\partial \boldsymbol{\Sigma}_{1|*}} & \frac{\partial \boldsymbol{\Sigma}_{12}(\cdot)}{\partial \mathbf{T}} & \frac{\partial \boldsymbol{\Sigma}_{12}(\cdot)}{\partial \boldsymbol{\Sigma}_*} \\ \frac{\partial \boldsymbol{\Sigma}_2(\cdot)}{\partial \boldsymbol{\Sigma}_{1|*}} & \frac{\partial \boldsymbol{\Sigma}_2(\cdot)}{\partial \mathbf{T}} & \frac{\partial \boldsymbol{\Sigma}_2(\cdot)}{\partial \boldsymbol{\Sigma}_*} \end{vmatrix} \\
 &= \begin{vmatrix} \mathbf{I}_{n_1^2} & (*) & (**) \\ \mathbf{0}_{n_1 n_2, n_1^2} & \text{diag}_{n_1}(\boldsymbol{\Sigma}_*) & (***) \\ \mathbf{0}_{n_2^2, n_1^2} & \mathbf{0}_{n_2^2, n_1 n_2} & \mathbf{I}_{n_2^2} \end{vmatrix} \\
 &= |\mathbf{I}_{n_1^2}| |\text{diag}_{n_1}(\boldsymbol{\Sigma}_*)| |\mathbf{I}_{n_2^2}| \\
 &= |\boldsymbol{\Sigma}_*|^{n_1}.
 \end{aligned}$$

Here $\text{diag}_{n_1}(\mathbf{W})$ is a diagonal block matrix, where the n_1 diagonal elements are the matrices \mathbf{W} . Hence,

$$\begin{aligned}
 f(\boldsymbol{\Sigma}_{1|*}, \mathbf{T}, \boldsymbol{\Sigma}_*) &\propto |\boldsymbol{\Sigma}_*|^{-(\nu-2n_1+n+1)/2} \times |\boldsymbol{\Sigma}_{1|*}|^{-(\nu+n+1)/2} \\
 &\quad \exp\left\{-\frac{1}{2}\text{tr}\left(\boldsymbol{\Psi}_{1|*}\boldsymbol{\Sigma}_{1|*}^{-1} + \boldsymbol{\Psi}_*\boldsymbol{\Sigma}_*^{-1} + \boldsymbol{\Psi}_*(\mathbf{T} - \boldsymbol{\Gamma})^T\boldsymbol{\Sigma}_{1|*}^{-1}(\mathbf{T} - \boldsymbol{\Gamma})\right)\right\}.
 \end{aligned} \tag{D-3}$$

From this last expression we see that

$$f(\boldsymbol{\Sigma}_*) \propto |\boldsymbol{\Sigma}_*|^{-(\nu-n_1+n_2+1)/2} \times \exp\left\{-\frac{1}{2}\text{tr}(\boldsymbol{\Psi}_*\boldsymbol{\Sigma}_*^{-1})\right\},$$

which we recognise as the kernel of the inverted Wishart pdf, thus

$$\boldsymbol{\Sigma}_* \sim W_g^{-1}(\boldsymbol{\Psi}_*^{-1}, \nu - n_1).$$

Using Eq. (D-3) we see that

$$f(\boldsymbol{\Sigma}_{1|*}, \mathbf{T}) \propto |\boldsymbol{\Sigma}_{1|*}|^{-(\nu+n+1)/2} \times \exp \left\{ -\frac{1}{2} \text{tr} \left([\boldsymbol{\Psi}_{1|*} + (\mathbf{T} - \boldsymbol{\Gamma}) \boldsymbol{\Psi}_* (\mathbf{T} - \boldsymbol{\Gamma})^T] \boldsymbol{\Sigma}_{1|*}^{-1} \right) \right\}.$$

Thus, by using the property of the sum of determinants (Mardia et al., 1979),

$$\begin{aligned} f(\mathbf{T}) &\propto \int |\boldsymbol{\Sigma}_{1|*}|^{-(\nu+n+1)/2} \\ &\quad \exp \left\{ -\frac{1}{2} \text{tr} \left([\boldsymbol{\Psi}_{1|*} + (\mathbf{T} - \boldsymbol{\Gamma}) \boldsymbol{\Psi}_* (\mathbf{T} - \boldsymbol{\Gamma})^T] \boldsymbol{\Sigma}_{1|*}^{-1} \right) \right\} d\boldsymbol{\Sigma}_{1|*} \\ &\propto |\boldsymbol{\Psi}_{1|*} + (\mathbf{T} - \boldsymbol{\Gamma}) \boldsymbol{\Psi}_* (\mathbf{T} - \boldsymbol{\Gamma})^T|^{-(\nu+n_2)/2} \\ &\propto \left| \mathbf{I}_{n_2} + \boldsymbol{\Psi}_* (\mathbf{T} - \boldsymbol{\Gamma})^T \boldsymbol{\Psi}_{1|*}^{-1} (\mathbf{T} - \boldsymbol{\Gamma}) \right|^{-(\tilde{\nu}+n_1+n_2-1)/2}, \end{aligned}$$

where $\tilde{\nu} = \nu - n_1 + 1$, which entails

$$\mathbf{T} \sim \text{MatrixT}_{n_1, n_2}(\boldsymbol{\Gamma}, \boldsymbol{\Psi}_*^{-1}, \boldsymbol{\Psi}_{1|*}, \nu - n_1 + 1).$$

Similarly,

$$\begin{aligned} f(\boldsymbol{\Sigma}_{1|*}) &\propto |\boldsymbol{\Sigma}_{1|*}|^{-(\nu+n+1)/2} \exp \left\{ -\frac{1}{2} \text{tr} \left(\boldsymbol{\Psi}_{1|*} \boldsymbol{\Sigma}_{1|*}^{-1} \right) \right\} \\ &\quad \times \int \exp \left\{ -\frac{1}{2} \text{tr} \left((\mathbf{T} - \boldsymbol{\Gamma}) \boldsymbol{\Psi}_* (\mathbf{T} - \boldsymbol{\Gamma})^T \boldsymbol{\Sigma}_{1|*}^{-1} \right) \right\} d\mathbf{T}. \\ &\propto |\boldsymbol{\Sigma}_{1|*}|^{-(\nu+n+1)/2} \exp \left\{ -\frac{1}{2} \text{tr} \left(\boldsymbol{\Psi}_{1|*} \boldsymbol{\Sigma}_{1|*}^{-1} \right) \right\} |\boldsymbol{\Sigma}_{1|*}|^{n_2/2} \\ &\propto |\boldsymbol{\Sigma}_{1|*}|^{-(\nu+n_1+1)/2} \exp \left\{ -\frac{1}{2} \text{tr} \left(\boldsymbol{\Psi}_{1|*} \boldsymbol{\Sigma}_{1|*}^{-1} \right) \right\}, \end{aligned}$$

where we have used $\mathbf{T}|\boldsymbol{\Sigma}_{1|*} \sim \text{MatrixG}_{n_1, n_2}(\boldsymbol{\Gamma}, \boldsymbol{\Psi}_*^{-1}, \boldsymbol{\Sigma}_{1|*})$, which completes the proof. \square

Appendix 6-E, Sampling from the matrix- t distribution

Result 6 (Realisation from the Matrix t -distribution). *Let*

$$\boldsymbol{\Sigma} \sim W_u^{-1}(\boldsymbol{\Psi}^{-1}, \nu + u - 1)$$

independent of $\mathbf{X} \sim \text{MatrixG}_{u, g}(\mathbf{0}_{u, g}, \boldsymbol{\Omega}, \mathbf{I}_u)$. *Then*

$$\mathbf{T} = \boldsymbol{\Sigma}^{1/2} \mathbf{X} + \boldsymbol{\Gamma} \sim \text{MatrixT}_{u, g}(\boldsymbol{\Gamma}, \boldsymbol{\Omega}, \boldsymbol{\Psi}, \nu). \quad (\text{E-1})$$

Proof. By Definitions 4 and 2 in Appendix 6-A

$$f(\boldsymbol{\Sigma}, \mathbf{X}) \propto |\boldsymbol{\Sigma}|^{-(\nu+u-1+u+1)/2} \exp \left\{ -\frac{1}{2} \text{tr} (\boldsymbol{\Sigma}^{-1} \boldsymbol{\Psi} + \boldsymbol{\Omega}^{-1} \mathbf{X}^T \mathbf{X}) \right\},$$

Following Gupta and Nagar (2000), let $\mathbf{T} = \mathbf{g}(\mathbf{X}, \boldsymbol{\Sigma}) = \boldsymbol{\Sigma}^{1/2} \mathbf{X} + \boldsymbol{\Gamma}$. A transformation of the variables gives,

$$f(\boldsymbol{\Sigma}, \mathbf{T}) = f(\boldsymbol{\Sigma}, \mathbf{g}^{-1}(\mathbf{T}, \boldsymbol{\Sigma})) |\mathbf{J}_{\mathbf{x}\mathbf{t}}|,$$

where by Deemer and Olkin (1951),

$$\begin{aligned} \mathbf{J}_{\mathbf{x}\mathbf{t}} &= \frac{\partial \boldsymbol{\Sigma}^{-1/2}(\mathbf{T} - \boldsymbol{\Gamma})}{\partial \mathbf{T}} \\ &= |\boldsymbol{\Sigma}|^{-g/2}. \end{aligned}$$

Hence,

$$\begin{aligned} f(\mathbf{T}) &\propto \int |\boldsymbol{\Sigma}|^{-(\nu+u-1+u+1+g)/2} \\ &\quad \exp \left\{ -\frac{1}{2} \text{tr} ([\boldsymbol{\Psi} + (\mathbf{T} - \boldsymbol{\Gamma})\boldsymbol{\Omega}^{-1}(\mathbf{T} - \boldsymbol{\Gamma})^T] \boldsymbol{\Sigma}^{-1}) \right\} d\boldsymbol{\Sigma} \\ &\propto |\boldsymbol{\Psi} + (\mathbf{T} - \boldsymbol{\Gamma})\boldsymbol{\Omega}^{-1}(\mathbf{T} - \boldsymbol{\Gamma})^T|^{-(\nu+u+g-1)/2}, \end{aligned}$$

and the result follows by the general properties of the determinant and trace operators (Mardia et al., 1979). \square

Result 7 (Fast sampling from the matrix- t distribution). *Let $\mathbf{U}_{\boldsymbol{\Omega}} \in \mathbb{R}^{b \times b}$ and $\mathbf{U}_{\boldsymbol{\Sigma}} \in \mathbb{R}^{a \times a}$ be the upper triangular Cholesky factors of the positive definite matrices $\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}$ respectively. Further let*

$$\mathbf{A} = \begin{bmatrix} u_{1,1} & z_{1,2} & z_{1,3} & \cdots & z_{1,a-1} & z_{1,a} \\ 0 & u_{2,2} & z_{2,3} & \cdots & z_{2,a-1} & z_{2,a} \\ 0 & 0 & u_{3,3} & \cdots & z_{2,a-1} & z_{2,a} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & u_{a-1,a-1} & z_{a-1,a} \\ 0 & 0 & \cdots & 0 & 0 & u_{a,a} \end{bmatrix} \in \mathbb{R}^{a \times a}, \quad (\text{E-2})$$

where $u_{i,i} = \sqrt{v_i}$, with $v_i \stackrel{\text{iid}}{\sim} \chi^2(\nu + a - i)$, and $z_{i,j} \stackrel{\text{iid}}{\sim} \text{Gauss}_1(0, 1)$, and let $\mathbf{W} \in \mathbb{R}^{a \times b}$, have elements $w_{i,j} \stackrel{\text{iid}}{\sim} \text{Gauss}_1(0, 1)$. Then

$$\mathbf{T} = \mathbf{U}_{\boldsymbol{\Sigma}}^T \mathbf{A}^{-1} \mathbf{W} \mathbf{U}_{\boldsymbol{\Omega}} + \mathbf{B} \sim \text{MatrixT}_{a,b}(\mathbf{B}, \boldsymbol{\Omega}, \boldsymbol{\Sigma}, \nu). \quad (\text{E-3})$$

Proof. By Smith and Hocking (1972) and Jones (1985), $\mathbf{A}^{-1}\mathbf{A}^{T^{-1}} \sim W_a^{-1}(\mathbf{I}_a, \nu + u - 1)$. Thus, by Result 6,

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{W} \sim \text{MatrixT}_{a,b}(\mathbf{0}_{a,b}, \mathbf{I}_b, \mathbf{I}_a, \nu).$$

The result then follows by Theorem 4.3.8 in Gupta and Nagar (2000). \square

Appendix 6-F, Singular matrix variate probability distributions

Definition 6 (Singular matrix Gaussian distribution). Let $\Sigma \in \mathbb{R}^{a \times a}$ and $\Omega \in \mathbb{R}^{b \times b}$ be positive semidefinite matrices with rank p and q respectively. Then \mathbf{W} follows the *singular matrix Gaussian* distribution, $\mathbf{W} \sim \text{MatrixG}_{a,b}^{p,q}(\mathbf{M}, \Omega, \Sigma)$, if

$$\mathbf{W} = \mathbf{U}_\Sigma \mathbf{Z} \mathbf{U}_\Omega^T + \mathbf{M}, \quad (\text{F-1})$$

where $\mathbf{Z} \sim \text{MatrixG}_{p,q}(\mathbf{0}_{p,q}, \Lambda_\Omega, \Lambda_\Sigma)$, and

$$\begin{aligned} \Sigma &= \mathbf{U}_\Sigma \Lambda_\Sigma \mathbf{U}_\Sigma^T \\ \Omega &= \mathbf{U}_\Omega \Lambda_\Omega \mathbf{U}_\Omega^T. \end{aligned}$$

Here $\mathbf{U}_\Sigma \in \mathbb{R}^{a \times p}$ is the matrix containing the p first eigenvectors of Σ corresponding to the p non-zero eigenvalues given as the p elements of the diagonal matrix $\Lambda_\Sigma \in \mathbb{R}^{p \times p}$, with a similar notation for $\mathbf{U}_\Omega \in \mathbb{R}^{b \times q}$ and $\Lambda_\Omega \in \mathbb{R}^{q \times q}$. The pdf can be found in Díaz-García et al. (1997).

Definition 7 (Singular Wishart distribution). Let $\mathbf{Z} \sim \text{MatrixG}_{u,\nu}^{p,\nu}(\mathbf{0}_{u,\nu}, \mathbf{I}_\nu, \Sigma)$ with $\nu < u$, then

$$\mathbf{M} = \mathbf{Z} \mathbf{Z}^T \quad (\text{F-2})$$

follows the *singular Wishart* distribution, $\mathbf{M} \sim W_u^r(\Sigma, \nu)$, with $r = \min\{p, \nu\}$. The pdf can be found in Díaz-García et al. (1997).

Definition 8 (Singular inverse Wishart distribution). Assume $\mathbf{V} \sim W_u^p(\Sigma, \nu)$, and let $\mathbf{U}_\mathbf{V} \in \mathbb{R}^{u \times p}$ be the matrix containing the p left singular vectors of \mathbf{V} corresponding to the p non-zero singular values in the diagonal matrix $\mathbf{S}_\mathbf{V} \in \mathbb{R}^{p \times p}$. Then

$$\mathbf{W} = \mathbf{U}_\mathbf{V} \mathbf{S}_\mathbf{V}^{-2} \mathbf{U}_\mathbf{V}^T \quad (\text{F-3})$$

follows the *singular inverse Wishart* distribution, $\mathbf{W} \sim W_u^{-p}(\Sigma, \nu)$, where the pdf can be found in Bodnar and Okhrin (2008).

Definition 9 (Singular matrix- t distribution). Let $\mathbf{U}_\Sigma \in \mathbb{R}^{a \times p}$ be the matrix containing the p first eigenvectors of $\Sigma \in \mathbb{R}^{a \times a}$ corresponding to the p non-zeros eigenvalues given as the p elements of the diagonal matrix $\Lambda_\Sigma \in \mathbb{R}^{p \times p}$, with a similar notation for the matrix $\Omega \in \mathbb{R}^{b \times b}$, $\mathbf{U}_\Omega \in \mathbb{R}^{b \times q}$ and $\Lambda_\Omega \in \mathbb{R}^{q \times q}$. Assume $\mathbf{Z} \sim \text{MatrixG}_{p,q}^{p,q}(\mathbf{0}_{p,q}, \Lambda_\Omega, \mathbf{I}_p)$ independent of $\mathbf{V} \sim W_p^{-1}(\Lambda_\Sigma^{-1}, \nu + p - 1)$, then

$$\mathbf{T} = \mathbf{U}_\Sigma \mathbf{W}^{1/2} \mathbf{Z} \mathbf{U}_\Omega^T + \mathbf{M} \quad (\text{F-4})$$

follows the *singular matrix- t* distribution, $\mathbf{T} \sim \text{MatrixT}_{a,b}^{p,q}(\mathbf{M}, \Omega, \Sigma, \nu)$.

Remark. This definition is different from the one used in Díaz-García and Gutiérrez-Jáimez (2009), where they define the singular matrix- t distribution through $\mathbf{Y} = (\mathbf{W}^{1/2})^{-1} \mathbf{X} + \mathbf{M}$, where $\mathbf{W} \sim W_a^{\tilde{p}}(\Sigma, \nu)$ independent of $\mathbf{X} \sim \text{MatrixG}_{a,b}^{a,q}(\mathbf{0}_{a,b}, \Omega, \mathbf{I}_p)$, $\tilde{p} = \min\{p, \nu\}$ and the superscript $(-)$ denotes the Moore-Penrose inverse of a matrix (Strang, 1988).

Result 8 (Probability density function of the singular matrix- t distribution). *The pdf of $\mathbf{T} \sim \text{MatrixT}_{a,b}^{p,q}(\mathbf{M}, \Omega, \Sigma, \nu)$ is*

$$f(\mathbf{T}) = c(p, q, \nu) \frac{|\mathbf{I}_b + \Omega^{-1}(\mathbf{T} - \mathbf{M})^T \Sigma^{-1}(\mathbf{T} - \mathbf{M})|^{\nu+p+q-1/2}}{|\Lambda_\Sigma|^{q/2} |\Lambda_\Omega|^{p/2}}, \quad (\text{F-5})$$

where

$$c(p, q, \nu) = \frac{\Gamma_p(\nu + p + q - 1)}{\pi^{pq/2} \Gamma_p(\nu + p - 1)},$$

with $\Gamma(\cdot)$ defined in Eq. (A-3) and the superscript $(-)$ denotes the Moore-Penrose inverse of a matrix.

Proof. Using Eq. (F-4) and the results in Díaz-García (2007),

$$\mathbf{T} | \mathbf{W} \sim \text{MatrixG}_{a,b}^{p,q}(\mathbf{M}, \Omega, \mathbf{U}_\Sigma \mathbf{W} \mathbf{U}_\Sigma^T).$$

Hence, the joint pdf of \mathbf{T} and \mathbf{W} is:

$$f(\mathbf{T}, \mathbf{W}) = c_1 |\mathbf{W}|^{-(\nu+2p+q)/2} \exp \left\{ -\frac{1}{2} \text{tr}(\Theta^{-1} \mathbf{W}^{-1}) \right\}, \quad (\text{F-6})$$

where

$$c_1 = \frac{(2\pi)^{-pq/2} |\Lambda_\Sigma|^{(\nu+p-1)/2}}{|\Lambda_\Omega|^{p/2} 2^{(\nu+p-1)p/2} \Gamma_p(\nu + p - 1)}$$

and

$$\Theta^{-1} = \Lambda_\Sigma + \mathbf{U}_\Sigma^T (\mathbf{T} - \mathbf{M}) \Omega^{-1} (\mathbf{T} - \mathbf{M})^T \mathbf{U}_\Sigma.$$

The result follows by integrating Eq. (F-6) with respect to \mathbf{W} and using general properties of the determinant (Mardia et al., 1979). \square

Paper VI

**Enhanced linearised reduced-order models for
subsurface flow simulation**

Jincong He, Jon Sætrum and Louis J. Durlofsky

Paper submitted for publication

Chapter 7

Enhanced linearised reduced-order models for subsurface flow simulation

Abstract. Trajectory piecewise linearisation (TPWL) represents a promising approach for constructing reduced-order models. Using TPWL, new solutions are represented in terms of expansions around previously simulated (and saved) solutions. High degrees of efficiency are achieved when the representation is projected into a low-dimensional space using a basis constructed by proper orthogonal decomposition of snapshots generated in a training run. In recent work, a TPWL procedure applicable for two-phase subsurface flow problems was presented. The method was shown to perform well for cases with no density differences between phases, though accuracy and robustness were found to degrade when there were substantial differences in phase densities. In this work, these limitations are shown to be related to model accuracy at key locations and model stability. Enhancements addressing both of these issues are introduced. A new TPWL procedure, referred to as local resolution TPWL, enables key grid blocks (such as those containing injection or production wells) to be represented at full resolution; i.e., these blocks are not projected into the low-dimensional space. This leads to high accuracy at selected locations, and will be shown to improve the accuracy of important simulation quantities such as injection and production rates. Next, two techniques for enhancing the stability of the TPWL model are presented. The first approach involves a basis optimisation procedure in which the number of columns in the basis matrix is determined to minimise the spectral radius of an appropriately defined amplification matrix. The second procedure incorporates a basis matrix constructed using snapshots from a

simulation with equal phase densities. Both approaches are compatible with the local resolution procedure. Results for a series of test cases demonstrate the accuracy and stability provided by the new treatments. Finally, the TPWL model is used as a surrogate in a direct search optimisation algorithm, and comparison with results using the full-order model demonstrate the efficacy of the enhanced TPWL procedures for this application.

7.1 Introduction

Optimisation and uncertainty quantification are essential components in many model-based design procedures. The associated computations, which typically require large numbers of simulations, can be extremely time-consuming if highly resolved models are used. This may be the case even if parallel processing is applied, as multiobjective optimisation and optimisation under uncertainty, in which simulations are performed over a large number of models, can quickly occupy multiple cores.

The use of reduced-order models provides a means for accelerating these simulations. Our interest here is in subsurface flow modelling, which includes simulation of oil reservoirs, aquifers, and carbon sequestration operations. Several reduced-order modelling procedures based on proper orthogonal decomposition have been previously applied within this context; see, e.g., (Vermeulen et al., 2004; van Doren et al., 2006; Heijn et al., 2004; Cardoso et al., 2009). For the nonlinear problems associated with oil reservoir simulation, the speedups achieved by these procedures were, however, relatively modest, at most about a factor of 10.

Trajectory piecewise linearisation (TPWL) is a promising approach for model-order reduction that can potentially provide much larger speedups. TPWL represents new solutions of the governing equations in terms of linear expansions around previously simulated (saved) states. This requires that we perform one or more training simulations, from which the states and converged Jacobian matrices at each time step are saved. High degrees of efficiency are achieved by projecting the saved states and matrices into a low-dimensional subspace. This projection can be accomplished in different ways. In the implementation discussed in this paper, the projection matrix is constructed by proper orthogonal decomposition (POD) of the saved states. The TPWL approach was first introduced in Rewiński (2003) and Rewiński and White (2003) for the modelling of nonlinear circuits and micromachined devices. Since then it has been applied in a number of disciplines including computational fluid dynamics (Gratton and Willcox, 2004), nonlinear heat-transfer modelling (Yang and Shen, 2005) and electromechanical systems (Vasilyev et al., 2006; Bond and Daniel, 2007b).

In recent work Cardoso and Durlofsky (2010b) applied trajectory piecewise linearisation (TPWL) procedures for oil reservoir simulation. Systems involving two fluid components and two phases – oil and water – were considered. For test runs involving equal density fluids, TPWL results were shown to be in close agreement with reference (full-order) simulations for control schedules that were within the general range of those used in the

training runs. Runtime speedups of a factor of 200-1000 were observed for the examples considered. For cases in which the fluid densities differed considerably, however, instabilities were observed in some runs. This can lead to inaccuracy or, in some instances, to the blowup of the TPWL solution. This is of concern as oil and water phases often display density differences in practical cases.

Our goal in this work is to enhance the TPWL procedure presented in Cardoso and Durlofsky (2010b) to address the limitations noted above. We proceed in two important directions. First, we introduce a localisation treatment in which key grid blocks, such as those containing injection or production wells, are represented at full resolution; i.e., the states in these blocks are not projected into the low-dimensional subspace. A missing point estimation procedure (Astrid and Verhoeven, 2006) is used to determine which grid blocks (in addition to well blocks) to represent explicitly. This localisation will be shown to improve the accuracy of the overall TPWL representation and to have relatively little impact on run times (assuming the number of resolved blocks is not too large). The second enhancement is the use of stabilisation procedures. Two such approaches are investigated. In one approach, a stabilised basis is determined in a preprocessing step in which various combinations of basis vectors are considered, with the goal of minimising the spectral radii of the amplification matrices that appear in the TPWL representation. In the other stabilisation approach, the POD basis is determined using simulations involving equal density fluids. Accuracy is restored through use of the localisation treatment. Both procedures will be shown to greatly improve TPWL stability and performance.

The stability of TPWL models has been studied previously and our development here builds on earlier work. Rather than use the traditional Galerkin projection, Bond and Daniel (Bond and Daniel, 2007a, 2008, 2009) computed a left projection matrix that stabilises the system. This new left projection matrix is based on Lyapunov stability theory. However, if only stability is considered, the accuracy of the reduced-order model can degrade considerably. Thus, in Bond and Daniel (2008), a stabilisation procedure was proposed where the left projection matrix is constrained by Lyapunov theory to guarantee stability and the difference between the basis matrix and this new left projection matrix is optimised to recover accuracy. Bui-Thanh et al. (2004, 2007) formulated a goal-oriented, model-constrained optimisation problem to determine the optimal basis under Galerkin projection. This approach was also shown to improve stability. This method differs from that of Bond and Daniel (2008) in that it retains the Galerkin projection and optimises both the basis matrix and the left projection matrix. However,

when both accuracy and stability are considered, both of these approaches can lead to high computational demands.

This paper proceeds as follows. In Section 7.2, the equations for the subsurface flow of oil and water are presented, followed by a brief description of the finite volume approach used for their solution. The POD-based TPWL representation is then described. In Section 7.3 we present the local resolution approach in which key grid blocks are resolved explicitly. The enhanced accuracy provided by this treatment is illustrated with an example. The two approaches for stabilising the TPWL representation are presented in Section 7.4. The impact of the local resolution and stabilisation procedures is demonstrated through two examples, both of which contain $O(10^5)$ grid blocks, in Section 7.5. Next, in Section 7.6, the enhanced TPWL method is combined with a generalised pattern search (GPS) optimisation technique and applied to a production optimisation problem. Finally, in Section 7.7, we present conclusions and suggestions for future work.

7.2 Problem Formulation

The governing equations for oil-water flow and the basic TPWL formulation were presented in detail in Cardoso and Durlofsky (2010b). For the sake of completeness, an abbreviated description is also included here.

7.2.1 Oil-Water Flow Equations

The equations governing the two-phase flow of oil and water in porous formations are derived by combining expressions for mass conservation with Darcy's law. Using the subscript j to designate component/phase ($j=o$ for oil and w for water), these equations can be written as:

$$\frac{\partial}{\partial t} (\phi \rho_j S_j) - \nabla \cdot [\rho_j \lambda_j \mathbf{k} (\nabla p_j - \rho_j g \nabla D)] + q_j^w = 0, \quad (7.1)$$

where \mathbf{k} is the (diagonal) absolute permeability tensor, $\lambda_j = k_{rj}/\mu_j$ is the phase mobility, with k_{rj} the relative permeability to phase j and μ_j the phase viscosity, p_j is phase pressure, ρ_j is the phase density, g is gravitational acceleration, D is depth, t is time, ϕ is porosity, S_j is saturation and q_j^w is the source/sink term. Eq. 7.1 is written slightly differently here than in Cardoso and Durlofsky (2010b). Specifically, the source term q_j^w here differs by a constant factor of ρ_j^0 (where ρ_j^0 is the reference density of phase j) from that in Cardoso and Durlofsky (2010b), and the definition of λ_j differs by a factor of ρ_j/ρ_j^0 . The general oil-water model is completed by enforcing

the saturation constraint ($S_o + S_w = 1$) and by specifying a capillary pressure relationship $p_c(S_w) = p_o - p_w$. Eq. 7.1 is nonlinear since functions of unknowns (e.g., $k_{rj}(S_j)$) multiply unknowns.

We take p_o and S_w to be primary variables (p_w and S_o can be immediately computed once these are known). Eq. 7.1 is solved numerically using a fully-implicit finite volume procedure. Discretised forms for all terms are discussed in Cardoso and Durlofsky (2010b) and Aziz and Settari (1986). Basically, the flow from block to block is given by the interface transmissibility multiplied by the difference in block pressures, the accumulation term is handled using a first-order implicit (backward Euler) method, and the source term is treated using a well index representation, in which well rates are expressed in terms of an appropriately defined transmissibility times the pressure difference between the well block and the well. Using these representations, and defining $\mathbf{x} = [p_o, S_w]$ as the state vector and \mathbf{u} as the well controls (in this case the wells are controlled by specifying bottom hole pressure or BHP), the discrete system for the fully-implicit formulation can be written as:

$$\mathbf{g}(\mathbf{x}^{n+1}, \mathbf{x}^n, \mathbf{u}^{n+1}) = \mathbf{A}(\mathbf{x}^{n+1}, \mathbf{x}^n) + \mathbf{F}(\mathbf{x}^{n+1}) + \mathbf{Q}(\mathbf{x}^{n+1}, \mathbf{u}^{n+1}). \quad (7.2)$$

Here \mathbf{g} is the residual we seek to drive to zero, n and $n + 1$ designate time level, and \mathbf{A} , \mathbf{F} and \mathbf{Q} are the discretised accumulation, flux and source/sink terms, respectively.

Typically, a full-order reservoir simulator is used to solve Eq. 7.2. Newton's method, with the Jacobian matrix given by $\partial \mathbf{g} / \partial \mathbf{x}$, is applied. This is computationally expensive because Eq. 7.2 can be highly nonlinear and practical models may contain on the order of 10^5 or 10^6 grid blocks. We now describe the application of the TPWL approach for the efficient solution of Eq. 7.2.

7.2.2 Solution of Discretised System using TPWL

The idea of trajectory piecewise linearisation is to linearise the residual equation around states saved from previous (training) simulations. Here, at any given time, we linearise around a single point on the training trajectory. Methods that involve weighted linearisations around multiple points also exist (e.g., Rewiński and White (2003)), though optimal weights can be difficult to determine and improper weighting can lead to stability problems.

Given the current state \mathbf{x}^n , we designate the closest saved state encountered during the training run as \mathbf{x}^i . To determine \mathbf{x}^{n+1} , we linearise Eq. 7.2

around the state $(\mathbf{x}^{i+1}, \mathbf{x}^i, \mathbf{u}^{i+1})$. This gives

$$\mathbf{g}^{n+1} = \mathbf{g}^{i+1} + \frac{\partial \mathbf{g}^{i+1}}{\partial \mathbf{x}^{i+1}} (\mathbf{x}^{n+1} - \mathbf{x}^{i+1}) + \frac{\partial \mathbf{g}^{i+1}}{\partial \mathbf{x}^i} (\mathbf{x}^n - \mathbf{x}^i) + \frac{\partial \mathbf{g}^{i+1}}{\partial \mathbf{u}^{i+1}} (\mathbf{u}^{n+1} - \mathbf{u}^{i+1}), \quad (7.3)$$

where $\mathbf{g}^{n+1} = \mathbf{g}(\mathbf{x}^{n+1}, \mathbf{x}^n, \mathbf{u}^{n+1})$ and $\mathbf{g}^{i+1} = \mathbf{g}(\mathbf{x}^{i+1}, \mathbf{x}^i, \mathbf{u}^{i+1})$. Here $\mathbf{g}^{i+1} = 0$ because it is the residual of the training simulation. Defining the Jacobian matrix as the derivative of the residual with respect to the state,

$$\mathbf{J}^{i+1} = \frac{\partial \mathbf{g}^{i+1}}{\partial \mathbf{x}^{i+1}}, \quad (7.4)$$

and using the fact that, upon solution $\mathbf{g}^{n+1} = 0$, Eq. 7.3 can be expressed as

$$\mathbf{J}^{i+1} (\mathbf{x}^{n+1} - \mathbf{x}^{i+1}) = - \left[\frac{\partial \mathbf{A}^{i+1}}{\partial \mathbf{x}^i} (\mathbf{x}^n - \mathbf{x}^i) + \frac{\partial \mathbf{Q}^{i+1}}{\partial \mathbf{u}^{i+1}} (\mathbf{u}^{n+1} - \mathbf{u}^{i+1}) \right]. \quad (7.5)$$

Given \mathbf{x}^n and saved information, Eq. 7.5 allows us to linearly compute \mathbf{x}^{n+1} ; i.e., no iteration is required. This equation is, however, still in a high-dimensional space.

In order to reduce the dimension of Eq. 7.5, we employ linear order reduction. This entails representing the state \mathbf{x} in terms of a reduced state \mathbf{z} and a basis matrix Φ using

$$\mathbf{x} \approx \Phi \mathbf{z}. \quad (7.6)$$

There are many ways to construct Φ including optimal Hankel model (Adamjan et al., 1971; Bettayeb et al., 1980), balanced truncation (Moore, 1981), Krylov subspace methods (Feldmann and Freund, 1995), and proper orthogonal decomposition (POD) (Lumley, 1967). Within the context of TPWL, previous researchers have used Krylov subspace methods (e.g., (Rewienski, 2003; Vasilyev et al., 2006; Yang and Shen, 2005)) and POD (e.g., (Gratton and Willcox, 2004)). The POD method was applied successfully for our problem in Cardoso and Durlofsky (2010b) and will be used again here.

In POD, the high-dimensional space is represented by a set of orthogonal basis vectors, which are the singular vectors of a snapshot matrix \mathbf{X} . Each column of \mathbf{X} is simply the pressure or saturation state at a particular time step saved from a training simulation. POD is optimal in the sense that it minimises the mean squared reconstruction error for the snapshots (Pearson, 1901). Therefore, it is reasonable to assume that POD can represent the states of other (test) simulations if these states are somewhat similar to those encountered during training runs. In this work, consistent with van Doren

et al. (2006), Cardoso and Durlofsky (2010b) and Cardoso and Durlofsky (2010a), POD is used to reduce pressure and saturation separately. As discussed in Cardoso and Durlofsky (2010b), to generate the pressure basis matrix we use oil potential φ_o ($\varphi_o = p_o/\rho_o - gD$) rather than pressure. This was found to improve TPWL stability in Cardoso and Durlofsky (2010b).

If we denote the number of simulation grid blocks as N_c and the number of reduced pressure and saturation basis vectors as l_p and l_s respectively, the dimension of the problem can be reduced from $2N_c$ to $l_p + l_s$. This is accomplished by expressing \mathbf{x} in Eq. 7.5 using Eq. 7.6 and by premultiplying both sides of Eq. 7.5 by Φ^T (also called Galerkin projection (Berkooz and Titi, 1993)), which gives

$$\mathbf{J}_r^{i+1} (\mathbf{z}^{n+1} - \mathbf{z}^{i+1}) = - \left[\left(\frac{\partial \mathbf{A}_r^{i+1}}{\partial \mathbf{x}^i} \right)_r (\mathbf{z}^n - \mathbf{z}^i) + \left(\frac{\partial \mathbf{Q}_r^{i+1}}{\partial \mathbf{u}^{i+1}} \right)_r (\mathbf{u}^{n+1} - \mathbf{u}^{i+1}) \right], \quad (7.7)$$

where

$$\begin{aligned} \mathbf{J}_r^{i+1} &= \Phi^T \mathbf{J}^{i+1} \Phi, \\ \left(\frac{\partial \mathbf{A}_r^{i+1}}{\partial \mathbf{x}^i} \right)_r &= \Phi^T \left(\frac{\partial \mathbf{A}^{i+1}}{\partial \mathbf{x}^i} \right) \Phi, \\ \left(\frac{\partial \mathbf{Q}_r^{i+1}}{\partial \mathbf{u}^{i+1}} \right)_r &= \Phi^T \left(\frac{\partial \mathbf{Q}^{i+1}}{\partial \mathbf{u}^{i+1}} \right). \end{aligned} \quad (7.8)$$

Rearranging Eq. 7.7 we have

$$\mathbf{z}^{n+1} = \mathbf{z}^{i+1} - (\mathbf{J}_r^{i+1})^{-1} \left[\left(\frac{\partial \mathbf{A}_r^{i+1}}{\partial \mathbf{x}^i} \right)_r (\mathbf{z}^n - \mathbf{z}^i) + \left(\frac{\partial \mathbf{Q}_r^{i+1}}{\partial \mathbf{u}^{i+1}} \right)_r (\mathbf{u}^{n+1} - \mathbf{u}^{i+1}) \right]. \quad (7.9)$$

Eq. 7.9 can be solved efficiently for two reasons. First, as a result of linearisation, the new reduced state \mathbf{z}^{n+1} can be calculated directly from \mathbf{z}^n without any iteration. Second, all of the terms in Eq. 7.9 are in low-dimensional space, which means that the matrix operations are fast to compute. Specifically, the evaluation of Eq. 7.9 basically involves only two matrix-vector products and four vector additions in the reduced space, which can theoretically be performed in a fraction of a second for typical reduced-space dimension (~ 100 -500) and number of time steps (~ 300). Therefore, most of the total TPWL computation time is spent on training simulations and projections, which only need to be done once during preprocessing.

This completes our description of the basic TPWL formulation. See Cardoso and Durlofsky (2010b) for additional details and algorithms and a discussion of some implementation issues. In the following section we will

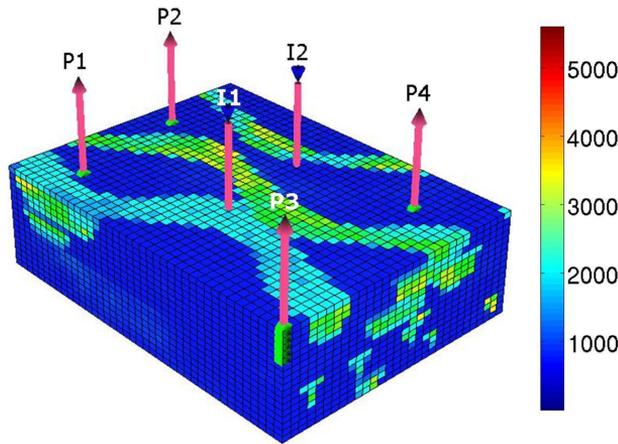


Figure 7.1: Model 1 with four production wells and two injection wells. Permeability in the x -direction (in mD) is shown.

present an application example of the basic TPWL method and discuss some of its limitations, which we will address in subsequent sections.

7.2.3 Application Example: Reservoir Model 1

We now apply the basic TPWL procedure described in Section 7.2.2 to a reservoir simulation model. The simulation model, shown in Figure 7.1, is part of the so-called Stanford VI geological model developed by Castro (2007). The model represents a fluvial system with high-permeability channels embedded in a low-permeability background region. The portion of the model considered here contains 20,400 grid blocks (with $n_x = 30$, $n_y = 40$, $n_z = 17$, where n_x , n_y and n_z indicate the number of blocks in each coordinate direction). The dimension of the full-order problem is 40,800 (pressure and saturation unknowns in each grid block). In this model the fluid and rock compressibility and the capillary pressure between the two phases are neglected. We specify oil and water viscosities as $\mu_o = 3$ cp, $\mu_w = 0.5$ cp. The fluids are here specified to have equal densities ($\rho_o = \rho_w = 55$ lb/ft³). The relative permeability relationships are given by

$$k_{ro}(S_w) = k_{ro}^0 \left(\frac{1 - S_w - S_{or}}{1 - S_{wr} - S_{or}} \right)^a, \quad k_{rw}(S_w) = k_{rw}^0 \left(\frac{S_w - S_{wr}}{1 - S_{wr} - S_{or}} \right)^b. \quad (7.10)$$

We set $k_{ro}^0 = k_{rw}^0 = 1$, $S_{wr} = S_{or} = 0.2$ and $a = b = 2$, as was used in Cardoso and Durlofsky (2010b).

To systematically investigate the performance of TPWL, we define test

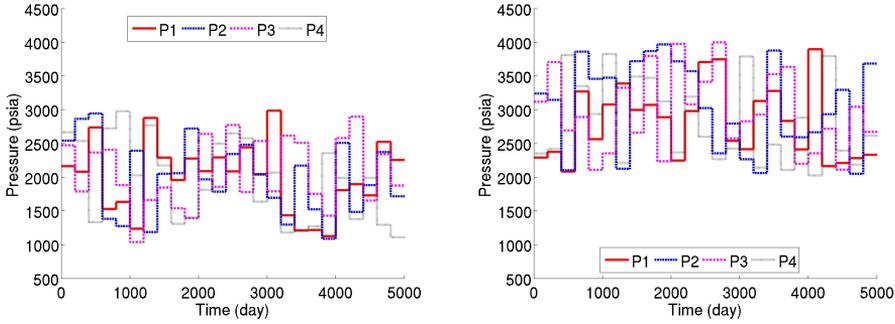


Figure 7.2: Training (left) and target (right) producer BHP schedules for Model 1.

cases based on the level of perturbation from the training run. Shown on the left in Figure 7.2 is the bottom hole pressure (BHP) control schedule for the production wells in the training simulation, which we designate as $\mathbf{u}_{training}$. This schedule is generated randomly, with BHPs between 1000 psi and 3000 psi, and is updated every 200 days. On the right is a different (target) schedule also generated randomly with the same update frequency as the training schedule, though this schedule varies between 2000 psi and 4000 psi. Test cases are specified as a weighted combination of these two schedules; i.e.,

$$\mathbf{u}_{test} = (1 - \alpha)\mathbf{u}_{training} + \alpha\mathbf{u}_{target}, \quad (7.11)$$

where α represents the “distance” of the test controls from the training controls. When α is near zero, test cases are close to the training run and TPWL would be expected to provide accurate results. As α increases toward 1, test cases are further from the training run and larger errors are expected. For this example, the BHPs of the two injection wells are held constant at 6000 psi throughout the simulations.

In this work, errors are quantified in terms of the mismatch of the production rates (for both oil and water) and water injection rates between the full-order solution (Q_{full}) and TPWL simulations (Q_{tpwl}). For example, for oil production rate, the error for the j th well (E_o^j) is calculated as:

$$E_o^j = \frac{\int_0^T |Q_{o,full}^j - Q_{o,tpwl}^j| dt}{\int_0^T Q_{o,full}^j}, \quad (7.12)$$

where subscript o designates oil and T is the total simulation time. The overall average error of the oil production rates, designated E_o , is computing

	$\alpha = 0$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
TPWL $_{40,60}^{E_o}$	0.0039	0.0047	0.0071	0.0108	0.0144	0.0193
TPWL $_{40,60}^{E_w}$	0.0024	0.0110	0.0246	0.0441	0.0733	0.1266
TPWL $_{40,60}^{E_i}$	0.0121	0.0156	0.0209	0.0272	0.0354	0.0456

Table 7.1: Relative error in TPWL solutions for various test schedules for Model 1 for basic TPWL with $l_p = 40$ and $l_s = 60$.

by averaging E_o^j over all wells:

$$E_o = \frac{1}{n_{pw}} \sum_{j=1}^{n_{pw}} E_o^j, \quad (7.13)$$

where n_{pw} is the total number of production wells. Similar expressions are used to compute average water production error and water injection error (E_w and E_i , respectively).

We now compare full-order simulation results, generated using Stanford's General Purpose Research Simulator (GPRS) (Cao, 2002; Jiang, 2007), with TPWL results. For the TPWL model we use $l_p = 40$ and $l_s = 60$. Results for $\alpha = 1.0$ for oil and water production rates are shown in Figure 7.3, while results for water injection rates appear in Figure 7.4. The errors in the TPWL solution for other levels of perturbation are shown in Table 7.2.3. Discrepancies are evident between the full-order and TPWL results, especially for water production and water injection rates. In Section 7.3 these discrepancies are discussed in detail and a local resolution method is proposed to improve the accuracy. However, it should be noted here that this inaccuracy results from the use of $\alpha = 1.0$, which means that the test case is quite different from the training run. As shown in Table 7.2.3, the use of smaller values of α leads to TPWL results that are much more accurate.

7.3 Local Resolution TPWL

As shown in Section 7.2.3, the basic TPWL method can display inaccuracy for cases with large perturbations. In this chapter we will first analyse the problem of underfitting and overfitting. We then propose the local resolution method to enhance the accuracy of the basic TPWL.

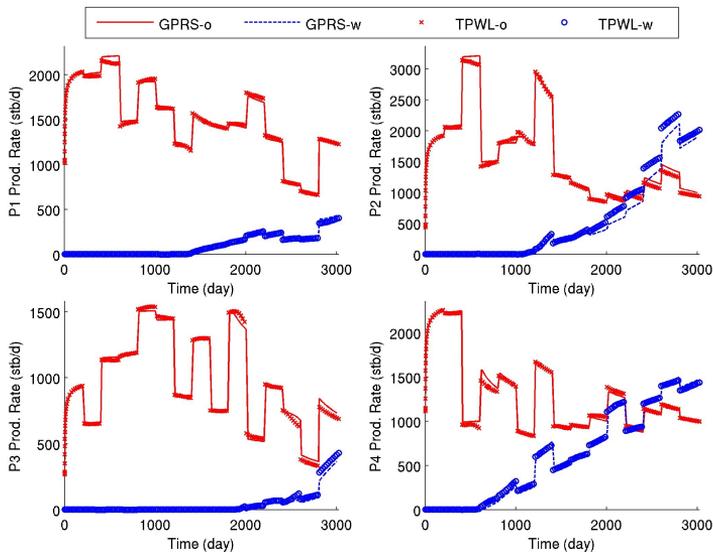


Figure 7.3: Model 1 production rates for $\alpha = 1.0$ using basic TPWL ($l_p = 40$, $l_s = 60$).

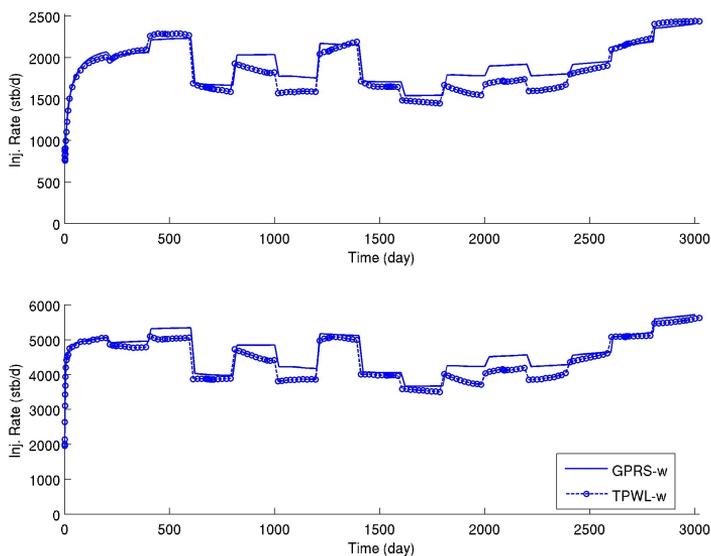


Figure 7.4: Model 1 injection rates for $\alpha = 1.0$ using basic TPWL ($l_p = 40$, $l_s = 60$).

7.3.1 Underfitting and Overfitting

Since Φ is orthonormal, we have $\mathbf{z} = \Phi^T \mathbf{x}$. POD will by construction minimise the mean squared reconstruction error of the training snapshots, $(1/S)\sum_{i=1}^S |\Phi(\Phi^T \mathbf{x}^i) - \mathbf{x}^i|^2$, where S is the number of snapshots (Pearson, 1901). Thus, if we consider TPWL solutions based on the full-order model (Eq. 7.5) and the reduced-order model (Eq. 7.9) for $\mathbf{u}^n = \mathbf{u}^i$, and if we take the number of basis vectors stored in Φ equal to the number of snapshots (and if the snapshots are based on pressure rather than potential), both approaches will reproduce the training states exactly (for Eq. 7.9 we need to apply $\mathbf{x} = \Phi \mathbf{z}$). However, this will not be the case for states from a new set of well controls $\mathbf{u}^n \neq \mathbf{u}^i$.

The POD dimension reduction technique is based on the assumption that most of the variability in the snapshot matrices can be represented using a limited number of orthonormal basis vectors $\Phi = [\phi_1, \dots, \phi_l]$ (Mardia et al., 1979). We therefore expect that selecting too few basis vectors will potentially lead to the problem of model underfitting, which can result in large errors in the TPWL solution, both for $\mathbf{u}^n = \mathbf{u}^i$ and $\mathbf{u}^n \neq \mathbf{u}^i$. On the other hand, including too many basis vectors will potentially cause model overfitting (Hastie et al., 2009), leading to large errors in the TPWL solution for $\mathbf{u}^n \neq \mathbf{u}^i$. This can occur because in the basis matrix obtained from POD, the basis vectors that correspond to smaller eigenvalues are more subject to noise in the training snapshots. Thus, including them in Φ can lead to a deterioration of the representation of \mathbf{x} stored in \mathbf{z} .

These effects can be observed through an assessment of the sensitivity of the TPWL results to different numbers of basis vectors in Φ . Errors in injection rate (E_i) for various α are presented in Table 7.3.1. The smallest error at each value of α is shown in bold. The other errors (E_o and E_w) display generally similar behaviour and are not shown. It is apparent from the table that selecting a small number of basis vectors ($l_p = 5$, $l_s = 5$) leads to large errors in the TPWL solution, particularly as α increases. Interestingly, however, error does not necessarily decrease monotonically as we include more basis vectors in Φ . Consistent with this, the use of the largest l_p and l_s does not provide the most accurate results for water injection for any value of α .

This is in part due to error introduced through overfitting. It is additionally because the POD reduction scheme focuses on the global reconstruction error, not the reconstruction error at the well blocks. Note also that the reason we do not observe a monotonic decrease in the error for the training schedule ($\alpha = 0$) is because, as noted earlier, we use oil potential rather than pressure snapshots to compute the pressure basis matrix. This results in im-

	$\alpha = 0$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
TPWL $_{5,5}^{E_i}$	0.0513	0.1245	0.1521	0.2078	0.2602	0.2921
TPWL $_{20,30}^{E_i}$	0.0298	0.0355	0.0423	0.0494	0.0564	0.0664
TPWL $_{40,60}^{E_i}$	0.0121	0.0156	0.0209	0.0272	0.0354	0.0456
TPWL $_{70,90}^{E_i}$	0.0183	0.0177	0.0187	0.0215	0.0281	0.0381
TPWL $_{120,120}^{E_i}$	0.0326	0.0332	0.0351	0.0367	0.0383	0.0416

Table 7.2: Relative error in TPWL solutions for various test schedules for Model 1 with equal density. The notation TPWL $_{l_p, l_s}$ in this and subsequent tables denotes the numbers of pressure and saturation basis vectors retained in Φ . The smallest error at each value of α is shown in bold.

proved accuracy in the TPWL solution for test schedules that differ from the training schedule (Cardoso and Durlofsky, 2010b), although it leads to an increase in the reconstruction error for the training schedule.

7.3.2 Description of Local Resolution Scheme

For our applications, we are particularly interested in the pressure and saturation at well locations because they directly affect injection and production rates, which are the key quantities needed for production optimisation. To compute these quantities, we construct $\mathbf{x}_w^{n+1} = \Phi_w \mathbf{z}^{n+1}$, where Φ_w includes only the rows of Φ corresponding to the grid blocks containing wells (subscript w here denotes well). Although we are especially interested in maintaining accuracy in Φ_w , the POD dimension reduction technique, by construction, minimises the *global* reconstruction error of the training snapshots – it does not preferentially weight information at well locations. We can therefore expect reconstruction of the saturation and pressure at well locations to be suboptimal.

This effect is illustrated in Figure 7.5, where we plot the maximum reconstruction error (relative to the full-order solution) of the saturation and pressure at the injection and production wells, as a function of l , for $\alpha = 1.0$. From this figure we see that reconstruction of saturation at the production wells is problematic, and that increasing the number of basis vectors included in Φ can lead to a clear increase in the reconstruction error.

It is thus evident that reconstruction error at well locations can have a large impact on the accuracy of TPWL results for production and injection rates. To eliminate reconstruction error at the well locations, we therefore propose a TPWL procedure in which selected key grid blocks are repre-

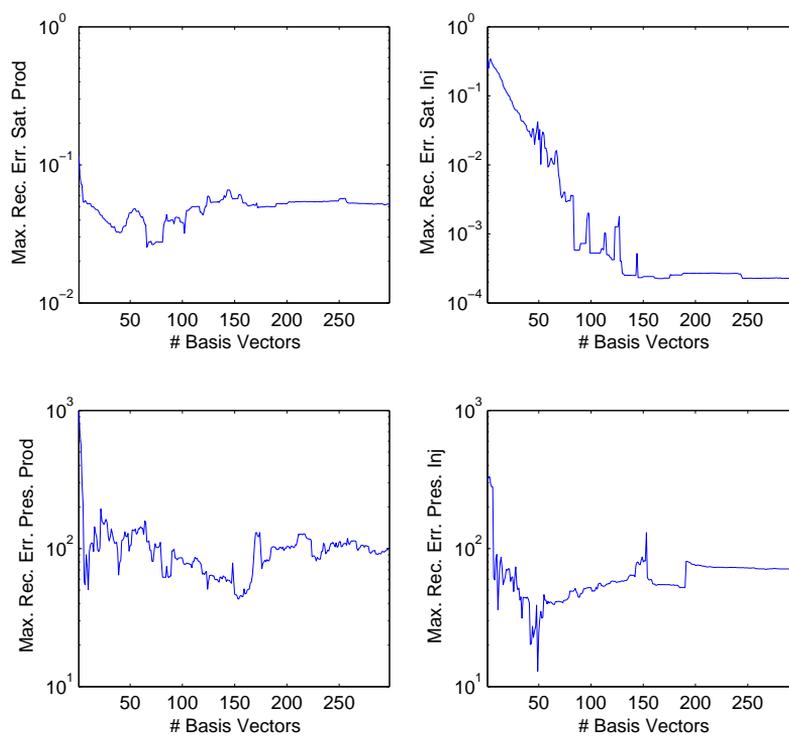


Figure 7.5: Log plot of the maximum reconstruction error for saturation and pressure at production and injection wells for Model 1 with $\alpha = 1.0$.

sented using the full-order (linearised) model. We let \mathbf{x}_{LR} designate the full-order states for n_{LR} selected grid blocks and \mathbf{x}_G the full-order states for the remaining grid blocks. Then, in place of Eq. 7.6, we write:

$$\begin{bmatrix} \mathbf{x}_{LR} \\ \mathbf{x}_G \end{bmatrix} \approx \begin{bmatrix} \Phi_{LR} & 0 \\ 0 & \Phi_G \end{bmatrix} \begin{bmatrix} \mathbf{z}_{LR} \\ \mathbf{z}_G \end{bmatrix}, \quad (7.14)$$

where Φ_{LR} is taken to be the identity matrix (thus $\mathbf{z}_{LR} = \mathbf{x}_{LR}$). This means that the grid blocks associated with \mathbf{x}_{LR} are not subject to reduction, so high local resolution is maintained. Note that, although the local and global grid blocks are decoupled in Eq. 7.14, they do couple in the TPWL scheme defined in Eq. 7.9 through $(\mathbf{J}_r^i)^{-1}$.

In our implementation, the n_{LR} locally-resolved blocks include the well blocks and possibly additional blocks that are important for the flow solution. To determine these additional blocks, we apply the missing point estimation (MPE) procedure suggested by Astrid and Verhoeven (2006). In this approach we retain blocks that have the largest impact on the condition number of $\Phi^T \Phi$, which are blocks that strongly affect the flow solution. A computationally efficient algorithm for MPE is described in Cardoso et al. (2009). In our implementation, locally resolved blocks are determined separately for the saturation and pressure portions of Φ . The final set of n_{LR} locally-resolved blocks is the union of these two sets of blocks.

The local resolution method provides flexibility for improving the accuracy of the TPWL representation. As is evident in Figure 7.5 and Table 7.3.1, this cannot necessarily be accomplished by adding more basis vectors. Even if it could, the maximum number of basis vectors is limited by the number of snapshots. The local resolution method does not have these limitations, and the TPWL model thus defined approaches the full-order model as n_{LR} is increased. Thus, the local resolution method enables us to achieve a balance between accuracy and efficiency.

7.3.3 Numerical Results using TPWL(LR)

We now apply the local resolution TPWL scheme, designated TPWL(LR), to the case considered earlier (Model 1 with $\alpha = 1.0$). For these runs we use $l_p = 40$, $l_s = 60$ and $n_{LR} = 26$ (local resolution only at well blocks), which corresponds to a total of 152 unknowns. Results are shown in Figures 7.6 and 7.7. Comparing these results with those from the basic TPWL scheme using $l_p = 40$, $l_s = 60$ (Figures 7.3 and 7.4), we see that by eliminating the reconstruction error at wells, there is a large improvement in the TPWL results, particularly for the water injection rates.

Results for flow rate errors for a range of α are shown in Table 7.3.3. Errors are shown for TPWL(LR) (with $l_p = 40$, $l_s = 60$, $n_{LR} = 26$) as well as for the basic method for two different sets of (l_p, l_s) . The smallest errors for each value of α are shown in bold. TPWL(LR) is the most accurate in all cases, and for some quantities a large reduction in error is observed. These results clearly demonstrate the efficacy of the use of local resolution within a TPWL model.

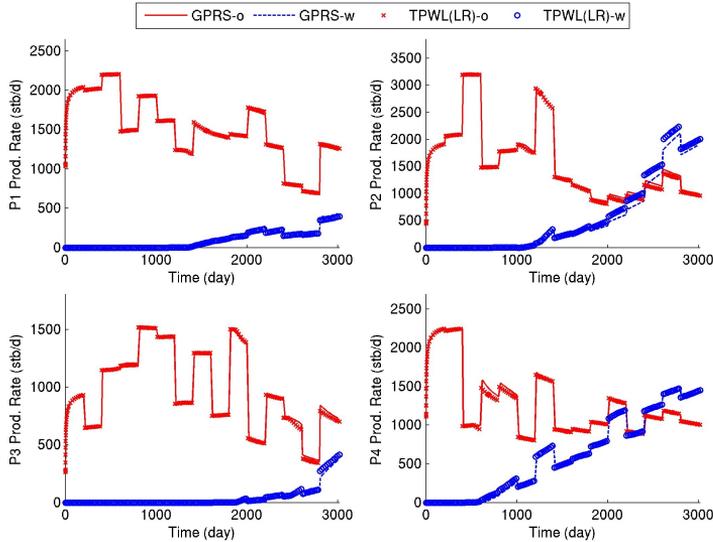


Figure 7.6: Model 1 production rates for $\alpha = 1.0$ using local resolution TPWL ($l_p = 40$, $l_s = 60$, $n_{LR} = 26$).

7.4 Stability of TPWL Models

As demonstrated through extensive examples in Cardoso and Durlofsky (2010b) and Cardoso and Durlofsky (2010a), the TPWL procedure can provide reasonable accuracy and robustness for cases with equal phase densities. It was, however, also reported in Cardoso and Durlofsky (2010b) that the method can become unstable when large density differences between the two phases exist. In this section we will first show an example of this instability. We will then discuss stability criteria and present two methods for stabilising TPWL.

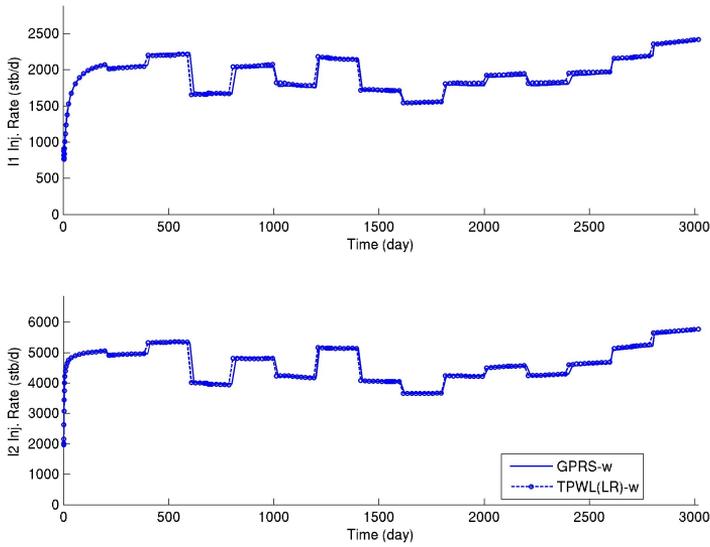


Figure 7.7: Model 1 injection rates for $\alpha = 1.0$ using local resolution TPWL ($l_p = 40, l_s = 60, n_{LR} = 26$).

	$\alpha = 0$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
TPWL $^{E_o}_{40,60}$	0.0039	0.0047	0.0071	0.0108	0.0144	0.0193
TPWL $^{E_o}_{70,90}$	0.0099	0.0113	0.0154	0.0212	0.0281	0.0357
TPWL $^{E_o}_{LR}$	0.0002	0.0020	0.0041	0.0065	0.0089	0.0121
TPWL $^{E_w}_{40,60}$	0.0024	0.0110	0.0246	0.0441	0.0733	0.1266
TPWL $^{E_w}_{70,90}$	0.0128	0.0229	0.0529	0.1106	0.2135	0.4126
TPWL $^{E_w}_{LR}$	0.0002	0.0090	0.0193	0.0350	0.0547	0.1032
TPWL $^{E_i}_{40,60}$	0.0121	0.0156	0.0209	0.0272	0.0354	0.0456
TPWL $^{E_i}_{70,90}$	0.0183	0.0177	0.0187	0.0215	0.0281	0.0381
TPWL $^{E_i}_{LR}$	0.0003	0.0013	0.0025	0.0037	0.0049	0.0067

Table 7.3: Relative error in TPWL solutions for various test schedules for Model 1. Basic TPWL and local resolution TPWL with $l_p = 40, l_s = 60, n_{LR} = 26$ are compared. The smallest error at each value of α is shown in bold.

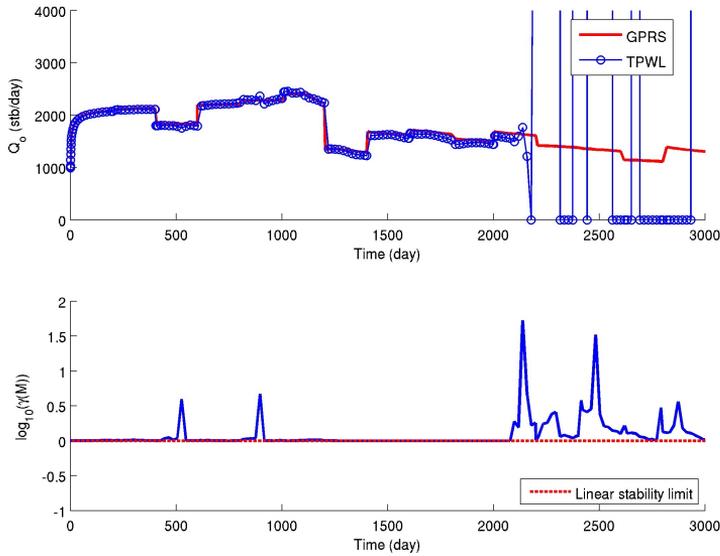


Figure 7.8: Upper: production rate for well P1 using basic TPWL with different phase densities (Model 1, $\alpha = 0.3$, $l_p = 70$, $l_s = 100$). Lower: log of spectral radius of amplification matrix.

7.4.1 Example Showing Instability of TPWL

We consider a model that is the same as that used in Section 7.3, except that now the densities for the two phases are different (here we set $\rho_o = 45$ lb/ft³ and $\rho_w = 55$ lb/ft³). Figure 7.8 (upper) displays the oil production rate for well P1 for $\alpha = 0.3$. Small spikes in the solution can be observed at around 500 days and 900 days. Shortly after 2000 days the solution becomes completely unstable and blows up. Results for all other wells display similar behaviours and are not shown here. Clearly, the method requires improvement if it is to be applied to problems of this type. We now consider the stability of the TPWL model.

7.4.2 Stability Analysis

We first consider the linearised full-order system, Eq. 7.5. This equation can be viewed as a piecewise linear discrete-time system,

$$\mathbf{x}^{n+1} = \mathbf{M}^{i+1} \mathbf{x}^n + \mathbf{b}^{i+1}, \quad (7.15)$$

where \mathbf{b} is a vector that involves the source term and \mathbf{M} is the amplification matrix given by

$$\mathbf{M}^{i+1} = -(\mathbf{J}^{i+1})^{-1} \frac{\partial \mathbf{A}^{i+1}}{\partial \mathbf{x}^i}. \quad (7.16)$$

For a constant \mathbf{M} , the system defined in Eq. 7.15 is stable if and only if the spectral radius of \mathbf{M} ($\gamma(\mathbf{M})$) is less than or equal to 1 (Gelfand, 1941). Here, stability means that the error in \mathbf{z}^n will not be amplified in \mathbf{z}^{n+1} .

As can be seen from Eq. 7.16, the amplification matrix \mathbf{M}^{i+1} of Eq. 7.15 is the same as that for the training simulation. Therefore, the full-order linearised system (Eq. 7.15) displays the same numerical stability properties as the original system. This is not the case, however, for the reduced-order model. In this case, the counterpart to Eq. 7.9 is

$$\mathbf{z}^{n+1} = \mathbf{M}_r^{i+1} \mathbf{z}^n + \mathbf{b}_r^{i+1}, \quad (7.17)$$

where the amplification matrix is given by

$$\mathbf{M}_r^{i+1} = -(\mathbf{\Phi}^T \mathbf{J}^{i+1} \mathbf{\Phi})^{-1} \mathbf{\Phi}^T \left(\frac{\partial \mathbf{A}^{i+1}}{\partial \mathbf{x}^i} \right) \mathbf{\Phi}. \quad (7.18)$$

For general matrices \mathbf{J}^{i+1} and $\partial \mathbf{A}^{i+1} / \partial \mathbf{x}^i$, $\gamma(\mathbf{M}_r^{i+1})$ can be greater than 1 even when $\gamma(\mathbf{M}^{i+1}) \leq 1$. In general, only special choices of $\mathbf{\Phi}$ can maintain the stability of the system. We note that the spectral radius of \mathbf{M}_r^{i+1} affects the stability of the linearised system at the time step when it is used. An isolated unstable \mathbf{M}_r^{i+1} will amplify the error at a specific time step and may create a spike in the solution. The solution may still be able to recover if the perturbation is not too large and if subsequent \mathbf{M}_r^{i+1} are stable. However, if we have several consecutive time steps with unstable \mathbf{M}_r^{i+1} , the error will amplify and the solution may blow up. Therefore, to ensure that the error does not accumulate over time, it is necessary to require the linearised system to be stable for all time steps.

Figure 7.8 (lower) presents the log of the spectral radius of the amplification matrix at each time step. Instability occurs for $\log \gamma > 0$. It is apparent that the spikes and eventual blowup of the solution correspond to values of $\log \gamma$ that are much larger than zero and thus unstable. The isolated unstable points at around 500 days and 900 days produce small spikes in the solution, while the consecutive unstable points starting at around 2200 days lead to solution blowup. Thus it is clear that the use of the TPWL method with a POD basis and Galerkin projection can result in instability for this problem.

The loss of stability for reduced-order models, especially those based on POD and Galerkin projection, has been studied previously and several

methods to enhance stability have been proposed. Two basic types of stabilisation procedures have been considered. The idea of the first set of methods is to compute a left projection matrix that stabilises the system, rather than use Φ^T as in Galerkin projection. In Bond and Daniel (2007a), Bond and Daniel (2008) and Bond and Daniel (2009), this new left projection matrix was based on Lyapunov stability theory. However, if only stability is considered, the accuracy of the reduced-order model can degrade considerably. If accuracy is also taken into account, a matrix optimisation has to be solved to obtain an optimal left projection matrix, which is computationally expensive for large systems. The other group of methods focuses on finding a basis other than POD that can guarantee stability under Galerkin projection. In Bui-Thanh et al. (2007), a goal-oriented, model-constrained optimisation problem was formulated to determine the optimal Φ . However, the procedure involves calculation in the full-order space. Even when the basis matrix is parametrised by the snapshots, as in Bui-Thanh et al. (2007), the optimisation can still be computationally demanding.

Neither of the approaches described above maintains both POD and Galerkin projection while stabilising the result. We will present two relatively efficient ways to stabilise TPWL in the following sections. Both apply a POD basis and Galerkin projection.

7.4.3 Stabilisation of TPWL using Optimised Basis

The first stabilisation method is based on the observation that the stability of the reduced-order model is sensitive to l_p and l_s , which define the number of columns in Φ . The dependence of \mathbf{M}_r on Φ is evident in Eq. 7.18.

Figure 7.9 depicts $\log(\gamma(\mathbf{M}_r))$ for different values of l_p and l_s for two particular saved points, $i = 70$ and $i = 100$, for the problem with different phase densities defined in Section 7.4.1. In the figure, the lower bound of the colour bar is the linear stability limit. Therefore the dark blue regions indicate combinations of l_p and l_s that give a stable linearised reduced system while other colours represent different levels of instability. It can be seen that the relation between stability and the number of basis vectors is somewhat random and that the use of more basis vectors does not necessarily lead to improved stability. This demonstrates that the widely used energy criterion (see, e.g., Cardoso and Durlofsky (2010b)) for selecting the number of basis vectors based on singular values may lead to stability problems. However, it can also be seen that for many of the (l_p, l_s) combinations, the spectral radius is less than 1. This means that if these combinations are used to generate the basis, the resulting reduced system will be stable.

The idea of basis optimisation is to define a range of l_p and l_s and to

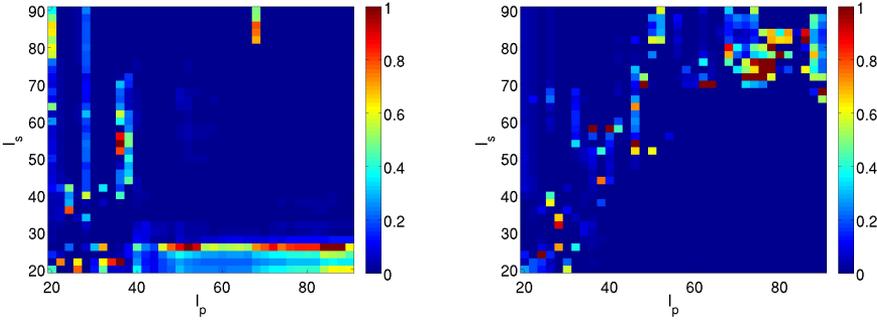


Figure 7.9: $\log_{10} \mathbf{M}_r^i$ for $i = 70$ and $i = 100$.

determine values (designated l_p^i and l_s^i) that minimise $\gamma(\mathbf{M}_r)$; i.e.,

$$(l_p^i, l_s^i) = \arg \min_{l_p^i, l_s^i} \gamma(\mathbf{M}_r^i). \quad (7.19)$$

This is accomplished using an exhaustive search over the allowable range of l_p and l_s (with prescribed increments in l_p and l_s). If we select different l_p and l_s for different time steps, which may be necessary in some cases, the reduced space changes in time. It is therefore necessary to map \mathbf{z} in one reduced space to \mathbf{z} in another reduced space. This is accomplished using

$$\mathbf{z}_\beta = \Phi_\beta^T \Phi_\alpha \mathbf{z}_\alpha, \quad (7.20)$$

where subscripts α and β indicate the two reduced spaces. We will refer to this procedure as basis switching.

The optimised basis (OB) procedure is summarised in Algorithm 7. First, a range for l_p and l_s is specified. Then the spectral radius is calculated for selected combinations of l_p and l_s . The combination that gives the smallest γ is selected. This method takes into account both accuracy and stability. Accuracy is controlled approximately by the search range, and stability is improved by choosing the optimal (l_p, l_s) combinations.

The optimisation procedure is reasonably efficient because \mathbf{J}_r^i and $(\partial \mathbf{A}^{i+1} / \partial \mathbf{x}^i)_r$ only need to be constructed once, for $l_{p,\max}$ and $l_{s,\max}$. Denoting these matrices as $\mathbf{J}_{r,\max}^i$ and $(\partial \mathbf{A}^{i+1} / \partial \mathbf{x}^i)_{r,\max}$, \mathbf{J}_r^i and $(\partial \mathbf{A}^{i+1} / \partial \mathbf{x}^i)_r$ for any other (l_p^i, l_s^i) combination are just submatrices of $\mathbf{J}_{r,\max}^i$ and $(\partial \mathbf{A}^{i+1} / \partial \mathbf{x}^i)_{r,\max}$ and can be extracted directly. Therefore the matrix operations inside the optimisation loop are all in reduced space. Furthermore, the optimal (l_p^i, l_s^i) combinations only need to be determined once during

Algorithm 7: Selecting the Number of Basis Vectors to Achieve Stability.

Input: Training data obtained using GPRS. Search region for $l_p^i \in [l_{p,\min}, l_{p,\max}]$ and $l_s^i \in [l_{s,\min}, l_{s,\max}]$

- 1 $[\Phi, \Lambda] = \text{SVD}(\mathbf{X})$
- 2 Keep first $l_{p,\max} + l_{s,\max}$ components
- 3 Compute $\mathbf{J}_{r,\max}^i$ and $(\partial \mathbf{A}^{i+1} / \partial \mathbf{x}^i)_{r,\max}$ based on $l_{p,\max}$ and $l_{s,\max}$ for all time steps i
- 4 **for each time step do**
- 5 **for** $l_p^i = l_{p,\min}$ **to** $l_{p,\max}$ **do**
- 6 **for** $l_s^i = l_{s,\min}$ **to** $l_{s,\max}$ **do**
- 7 Extract columns $1, \dots, l_p, 1, \dots, l_s$, of $\mathbf{J}_{r,\max}^i$ and $(\partial \mathbf{A}^{i+1} / \partial \mathbf{x}^i)_{r,\max}$
- 8 Construct amplification matrix \mathbf{M}_r^i
- 9 Calculate $\gamma(\mathbf{M}_r^i)$
- 10 Select the optimal values l_p^i and l_s^i that minimise $\gamma(\mathbf{M}_r^i)$
- 11 Construct the basis Φ^i

the preprocessing and will not add to the runtime of TPWL. Thus this optimisation does not overly affect the efficiency of the TPWL model.

This method differs from the two types of methods developed previously in that it maintains the advantages of POD and the Galerkin projection. Specifically, POD provides optimal accuracy in representing the snapshots and the Galerkin projection is straightforward and efficient. It is, however, important to note that the algorithm does not guarantee stability. Nevertheless, it does lead to considerable improvements in stability and enables the solution of challenging problems with large differences in density between phases.

Figure 7.10 displays the results when applying this technique to the problem described in Section 7.4.1. In this case a single set of optimised l_p and l_s was used ($l_p = 45, l_s = 60$); i.e., basis switching was not required. It is evident from Figure 7.10 that, after basis optimisation, $\gamma(\mathbf{M}_r)$ is close to 1, which means that TPWL is stabilised. The oil production for well P1 is seen to be in reasonable agreement with the reference full-order results, with the solution blowup after 2000 days eliminated (compare with Figure 7.8). We note that the use of local resolution will act to further improve the accuracy of the TPWL solution.

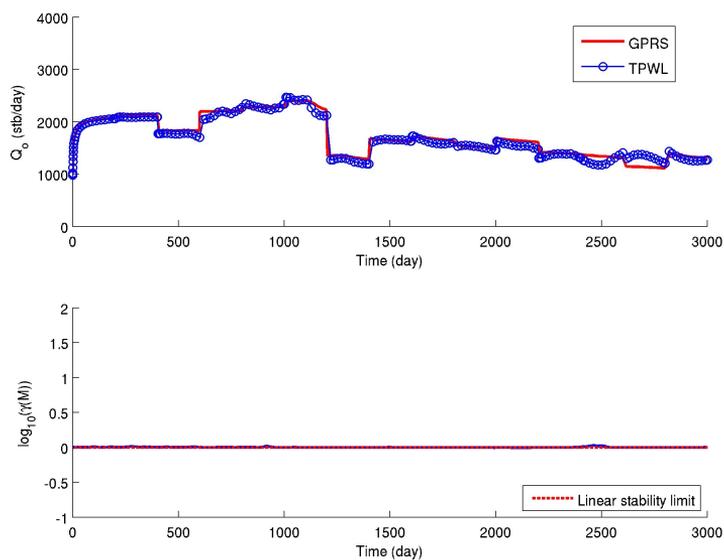


Figure 7.10: Upper: production rate for well P1 using TPWL with optimised basis for case with different phase densities (Model 1, $\alpha = 0.3$, $l_p = 45$, $l_s = 60$). Lower: log of spectral radius of amplification matrix.

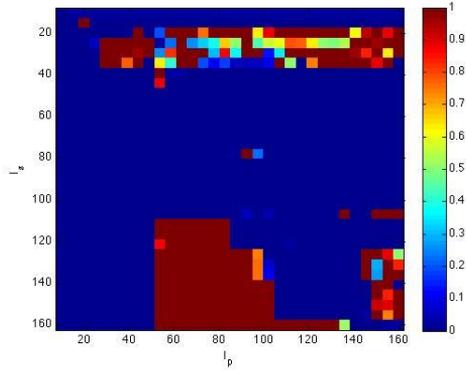
It is of interest to note that spikes in $\gamma(\mathbf{M}_r)$ usually correspond to spikes in the condition number of \mathbf{J}_r^{-1} (or \mathbf{J}_r). This is potentially of concern since \mathbf{J}_r^{-1} appears not only in \mathbf{M}_r but also in \mathbf{b}_r in Eq. 7.17. Thus, even when the system is stable, spikes in the condition number of \mathbf{J}_r^{-1} can still cause inaccuracy in the solution. In such cases, we may need to determine optimum (l_p, l_s) such that both $\gamma(\mathbf{M}_r)$ and the condition number of \mathbf{J}_r^{-1} are minimised to assure solution stability and accuracy. This will be addressed in future work.

7.4.4 Stabilisation of TPWL using Modified Basis

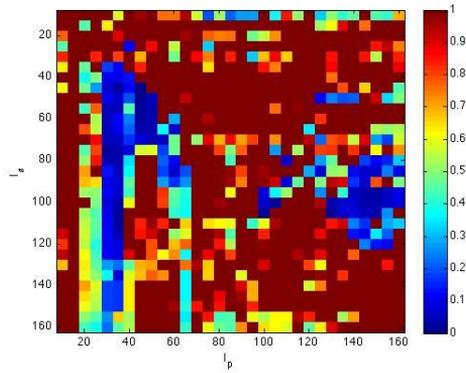
We now introduce a stabilisation procedure in which Φ is derived from a problem with better stability characteristics than the target problem with $\Delta\rho \neq 0$. This approach can be motivated with reference to stability maps. Figure 7.11(a) shows the maximum spectral radius for all saved points as a function of l_p and l_s for Model 1 with $\rho_o = \rho_w$. Here we see that the TPWL scheme is unstable only for a limited number of combinations of l_p and l_s . However, the stability map for the same reservoir model but with $\Delta\rho = 10$, shown in Figure 7.11(b), indicates only isolated regions where the TPWL model is stable. This motivates the construction of a model reduction scheme where the basis matrix is formed using snapshots generated from a reservoir model with $\Delta\rho = 0$. The states and Jacobian matrices used in the actual TPWL solution (Eq. 7.9) are still from the specific problem of interest (with $\Delta\rho \neq 0$). We will refer to this procedure as the equal density projection (EDP) scheme. A stability analysis of the EDP scheme for different values of l_p and l_s reveals that the regions where the TPWL solution is stable, seen in Figure 7.11(c), correspond to the stability regions in Figure 7.11(a). Thus the approach inherits the stability properties of the $\Delta\rho = 0$ TPWL model.

Because the EDP basis matrix is constructed using snapshots which are different from the states in the actual solution, the global reconstruction error for \mathbf{z}^i will necessarily increase dramatically compared with the approaches considered above. However, combining the EDP scheme with local resolution TPWL effectively eliminates the reconstruction error at key locations such as well blocks.

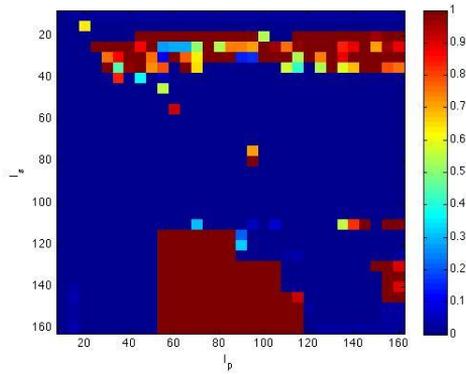
Figure 7.12 shows the oil and water production rates for Model 1 with $\Delta\rho = 10$ and $\alpha = 0.3$. Here we apply both EDP and LR and set $l_p = 70$, $l_s = 100$ and $n_{LR} = 27$ in the TPWL(EDP+LR) solution. Note that the use of $l_p = 70$, $l_s = 100$ in the basic TPWL method leads to instability (see Figure 7.8). As is evident in Figure 7.12, the TPWL(EDP+LR) scheme, by contrast, is able to provide a stable and reasonably accurate solution relative to the full-order results. This is further illustrated in the results for water



(a) $\log_{10} \max\{\gamma(M^i)\}$, $\Delta\rho = 0$, POD



(b) $\log_{10} \max\{\gamma(M^i)\}$, $\Delta\rho = 10$, POD



(c) $\log_{10} \max\{\gamma(M^i)\}$, $\Delta\rho = 10$, EDP

Figure 7.11: Maximum value for the spectral radius of the amplification matrix, M_r^i , as a function of number of basis vectors retained for Model 1 using (a) $\Delta\rho = 0$ with standard POD, (b) $\Delta\rho = 10$ with standard POD, and (c) $\Delta\rho = 10$ with EDP scheme.

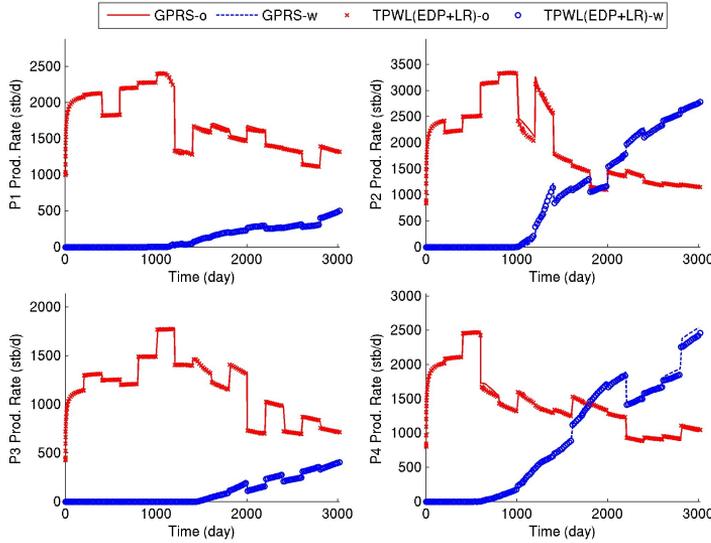


Figure 7.12: Model 1 production rates for $\alpha = 0.3$ using TPWL(EDP+LR) ($l_p = 70$, $l_s = 100$, $n_{LR} = 27$).

injection rates shown in Figure 7.13.

The reason why the basis matrix generated from snapshots of the corresponding equal density case (Φ_{ED}) leads to better stability than the basis matrix generated from snapshots of the actual (different density) case appears to be related to the magnitudes of the components appearing in the two basis matrices. Referring to the basis from the simulation with different densities as Φ_{DD} , we observe that Φ_{DD} tends to have more isolated extreme values than Φ_{ED} . These appear to derive from the fact that changes between snapshots in simulations with different densities are more localised than in cases with equal densities. More specifically, in runs with $\Delta\rho \neq 0$, substantial changes in the solution from one time step to the next may occur in only a few grid blocks, with little change in the global solution. This behaviour, which is less prominent in the equal density case, results in extreme values in Φ_{DD} . This in turn can lead to high condition numbers in \mathbf{J}_r and solution instability. By contrast, Φ_{ED} contains more evenly distributed values and as a result provides better model stability.

This argument is also consistent with results from the random projection (RP) scheme (Vempala, 2004). In the RP scheme, the pressure and saturation snapshots are replaced with random vectors generated independently

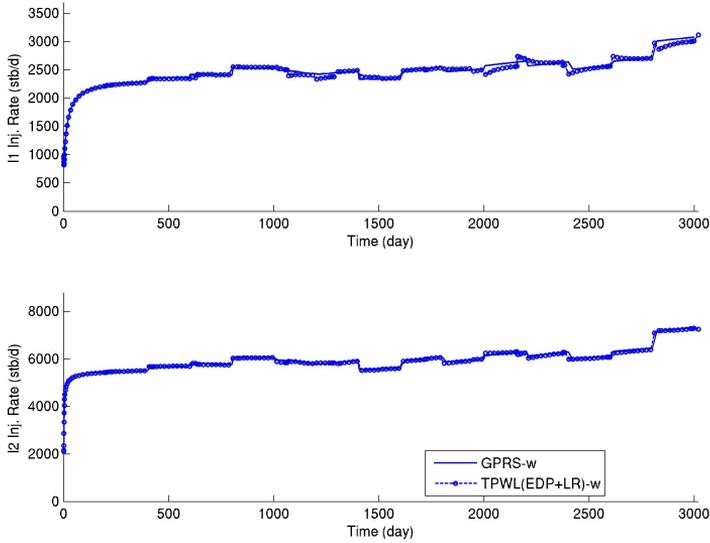


Figure 7.13: Model 1 injection rates for $\alpha = 0.3$ using TPWL(EDP+LR) ($l_p = 70$, $l_s = 100$, $n_{LR} = 27$).

from a Gaussian probability distribution with zero mean and unit variance. The variance between snapshots is thus evenly distributed spatially. Orthonormalisation of these matrices provides the basis matrix. Interestingly, this approach leads to a stable TPWL scheme for any combination of l_p and l_s . In numerous tests, we achieved stable TPWL results for all of the reservoir models considered in this paper, including cases where the basic TPWL method exhibited instability. The accuracy of the RP scheme is, however, quite poor as it does not use any information from actual snapshots. The EDP scheme can be viewed as an enhanced or supervised version of the RP scheme in the sense that, by using a particular set of snapshots, it achieves stability at the cost of accuracy. As noted above, accuracy is recovered through use of local resolution. We note finally that the accuracy of the RP scheme can also be improved using local resolution, but n_{LR} needs to be very large before adequate levels of accuracy are consistently achieved.

The two stabilisation schemes proposed in this section have somewhat different features. The optimised basis method requires only one full-order simulation, though it does require basis optimisation computations. In many cases it performs reasonably well without the use of local resolution. However, it has limited flexibility in the choice of l_p and l_s as they are determined

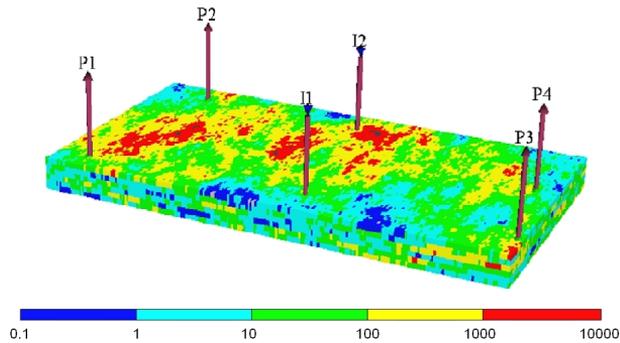


Figure 7.14: Upper six layers of the SPE 10 reservoir model (79,200 grid blocks) with four producers and two injectors. Permeability in x -direction (in mD) is shown.

based on stability properties. The EDP method, by contrast, requires two full-order simulations and it must be combined with local resolution to provide reasonable accuracy. It has more flexibility, however, in the choice of l_p and l_s . Thus both methods have advantages and limitations, and further application and development of both approaches appears to be warranted.

7.5 Application of TPWL to Realistic Problems

In this section we apply the new TPWL procedures to two realistic reservoir models. The models contain $O(10^5)$ grid blocks and the oil and water phases are of different densities. For these examples we combine the stabilisation methods with the local resolution procedure, and present results for a range of perturbations α . The basic TPWL procedure has difficulty providing stable solutions for these cases.

7.5.1 Model 2: Upper Six Layers of SPE 10

The geological model used here is shown in Figure 7.14. The model, referred to as Model 2, comprises the upper six layers of the so-called SPE 10 geological model, developed in Christie and Blunt (2001). The model contains 79,200 grid blocks (with $n_x = 60$, $n_y = 220$, $n_z = 6$). This model was also studied in Cardoso and Durlofsky (2010b), where it was applied for an example with equal phase densities, for which accurate TPWL results were reported. The problems that can arise using the basic TPWL procedure for cases with unequal densities were also illustrated using this model in Cardoso and Durlofsky (2010b). Model 2 includes four producers, which

are perforated in the upper two layers, and two injectors, perforated in the lower two layers. The fluid and rock-fluid properties are the same as were used in Model 1 (see Section 7.2.3) except that here we set the density for water to be $\rho_w=60 \text{ lb/ft}^3$, which gives a larger density difference.

The performance of the TPWL procedure for this problem is again studied using perturbation tests. The training schedule and target BHP schedules for this case are shown in Figure 7.5.1. For producers, the training BHP schedules vary randomly between 1000 psi and 3000 psi and are changed every 200 days. The target BHP schedules vary randomly between 1000 psi and 4000 psi; they also change every 200 days. Clear differences are evident between the two schedules. For the injectors, the training BHP schedules are constant at 8000 psi while the target BHP schedules vary randomly between 7000 psi and 9000 psi. Again, input BHP schedules for test cases are generated as weighted combinations of the training and target schedules, as defined in Eq. 7.11. The training simulation was run for 5000 days and produced 311 pressure and saturation snapshots.

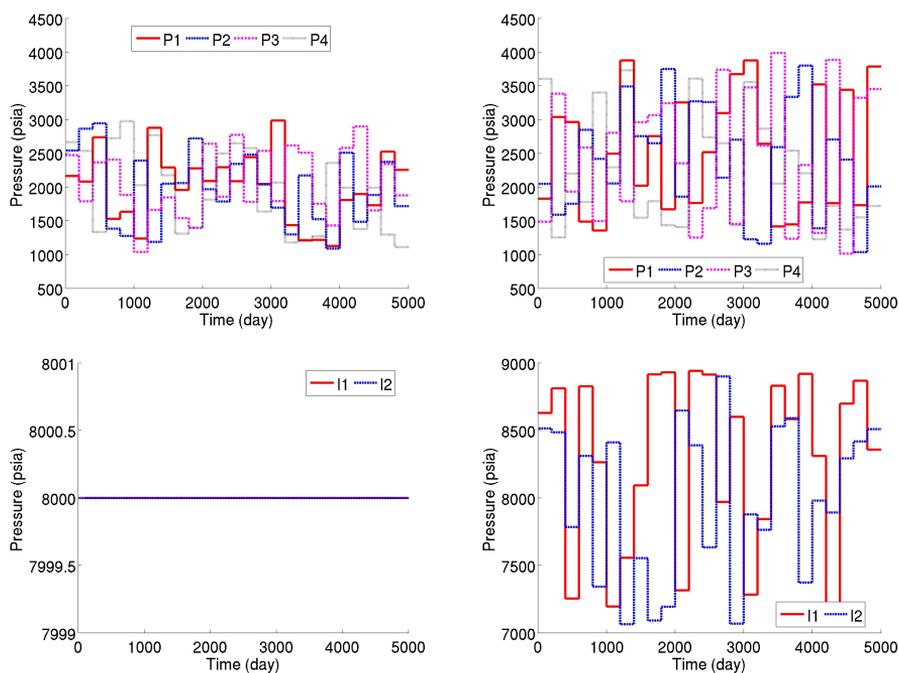


Figure 7.15: Training (left) and target (right) BHP schedules for producers and injectors for Model 2.

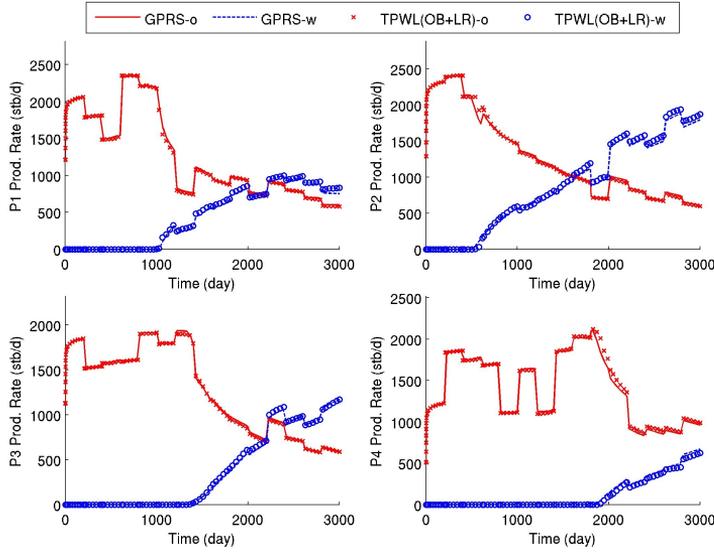


Figure 7.16: Model 2 production rates for $\alpha = 0.5$ using TPWL with optimised basis and local resolution ($l_p = 90$, $l_s = 85$, $n_{LR} = 154$).

We apply both the optimised basis method and the EDP procedure to generate stable TPWL models. For the optimised basis method, we considered l_p in the range $[40, 90]$ and l_s in the range $[60, 90]$. The search increment for both was 5. A basis with $l_p = 90$ and $l_s = 85$ was found to be stable for the entire simulation, so no basis switching was applied in this case. A total of 154 blocks were locally resolved. These include the 12 well blocks and 144 additional blocks determined by the MPE procedure.

Test cases were run for values of α from 0 to 1. Results for oil and water production rates, for $\alpha = 0.5$, are shown in Figure 7.16. Water injection rates are presented in Figure 7.17. It is clear that the TPWL model performs reasonably well for this case. Although slight mismatches are observed for some quantities, the general level of accuracy of the TPWL solution is quite acceptable and the method is clearly stable.

We now present results for this case using the EDP method for stabilisation. This requires that the training simulations are run twice – once using the actual densities and once using equal densities. The equal density run is used only to construct Φ ; the saved states and matrices are from the run using the actual densities. To allow direct comparison with TPWL(OB), we use $l_p = 90$, $l_s = 85$ and $n_{LR} = 154$.

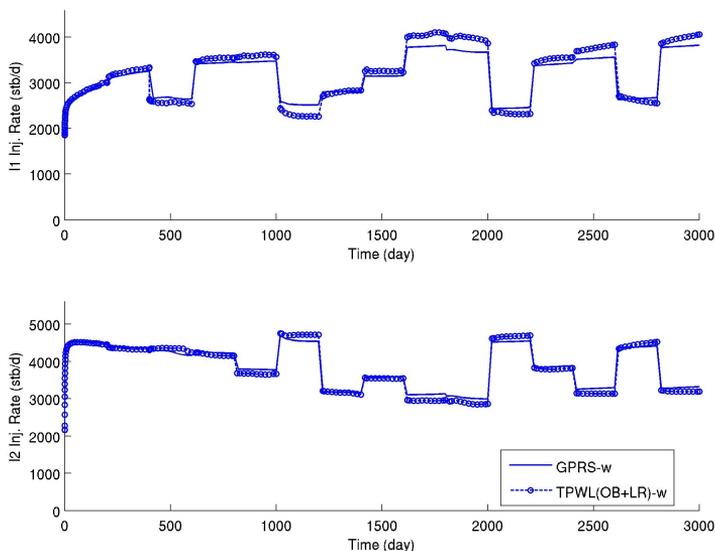


Figure 7.17: Model 2 injection rates for $\alpha = 0.5$ using TPWL with optimised basis and local resolution ($l_p = 90$, $l_s = 85$, $n_{LR} = 154$).

Figures 7.18 and 7.19 display results for oil and water production rates and water injection rates for $\alpha = 0.5$. We again observe stability and a reasonable degree of accuracy in the TPWL results. In fact, these results are quite comparable to those using the optimised basis procedure, shown in Figures 7.16 and 7.17.

Table 7.5.1 presents errors for oil and water production rates and water injection rates for the two sets of TPWL solutions at five values of α . Errors for both methods increase consistently with α , as would be expected. The EDP TPWL method is slightly more accurate than TPWL with the optimised basis, though the magnitudes of the errors are quite comparable.

7.5.2 Reservoir Model 3: Part of SPE 10 Upper 30 Layers

We now consider a more complex case. This model, also extracted from the geological description in Christie and Blunt (2001), contains 108,000 grid blocks ($n_x = 60$, $n_y = 60$, $n_z = 30$). The permeability distribution is shown in Figure 7.20. The model is referred to as Model 3 and includes four producers perforated in the upper 12 layers and two injectors perforated in the lower 12 layers. There are thus a total of 72 well blocks. To avoid large discrepancies in the well rates, well indices were set to a specified value. The

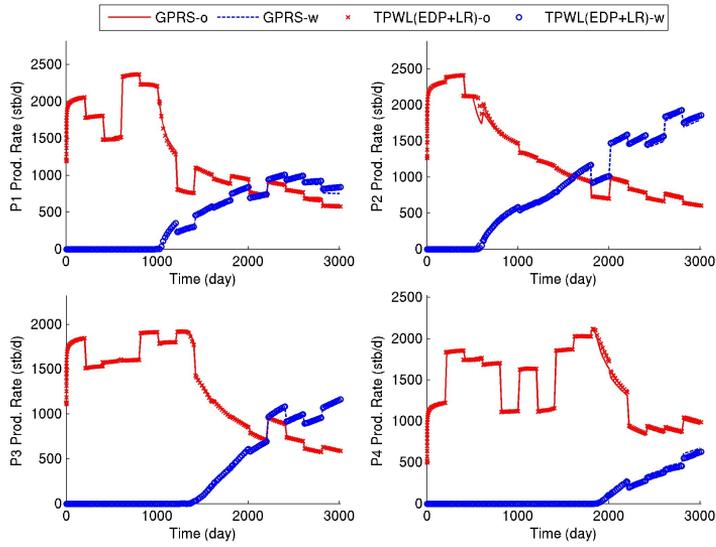


Figure 7.18: Model 2 production rates for $\alpha = 0.5$ using TPWL with EDP basis and local resolution ($l_p = 90$, $l_s = 85$, $n_{LR} = 154$).

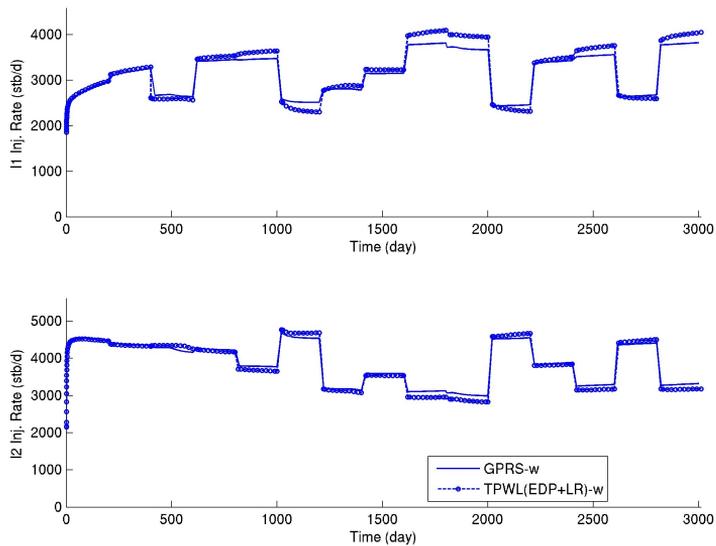


Figure 7.19: Model 2 injection rates for $\alpha = 0.5$ using TPWL with EDP basis and local resolution ($l_p = 90$, $l_s = 85$, $n_{LR} = 154$).

Method/ α	0.1	0.3	0.5	0.7	0.9
TPWL(OB+LR), E_o	0.0029	0.0088	0.0142	0.0203	0.0263
TPWL(OB+LR), E_w	0.0116	0.0351	0.0582	0.0819	0.1063
TPWL(OB+LR), E_i	0.0065	0.0194	0.0321	0.0449	0.0585
TPWL(EDP+LR), E_o	0.0022	0.0067	0.0116	0.0166	0.0220
TPWL(EDP+LR), E_w	0.0097	0.0293	0.0495	0.0702	0.0912
TPWL(EDP+LR), E_i	0.0059	0.0174	0.0292	0.0409	0.0531

Table 7.4: Errors for TPWL(OB+LR) and TPWL(EDP+LR) for Model 2.

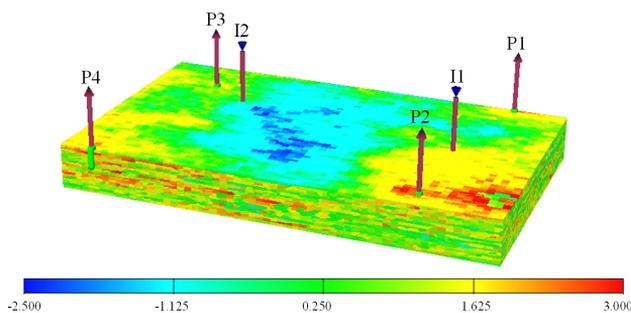


Figure 7.20: Portion of the SPE 10 reservoir model (108,000 grid blocks) with four producers and two injectors. Log of permeability in x -direction is shown.

fluid and rock-fluid properties are the same as were used in Model 2; the density difference between phases is again 15 lb/ft³. This case is challenging not only because of its size but also because it has more and thicker layers, which means that density-driven gravitational effects can be large.

The training and target BHP schedules are the same as were used for Model 2, as shown in Figure 15. The training simulation was run for 5000 days. A total of 314 pressure and saturation snapshots were saved.

We again present results using both stabilisation methods. For the optimised basis method, l_p and l_s were both evaluated over the range [60, 120] with increments of 5. The optimal combination, $l_p = 90$ and $l_s = 90$, was found to be stable for the entire run, so basis switching was not required. Local resolution is applied only to the 72 well blocks. Therefore, the total dimension of the reduced space is 324, compared with 204,000 for the full-order model.

Test cases were again run for α between 0 and 1. Figures 7.21 and 7.22 display results for production and injection rates for $\alpha = 0.5$. The results

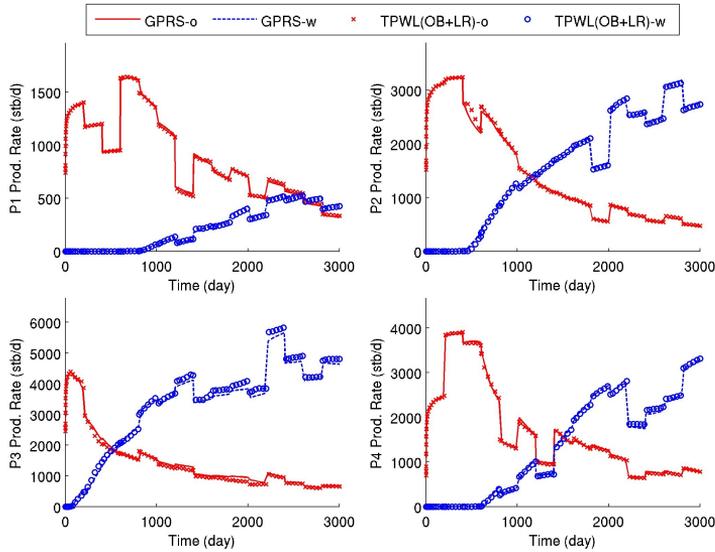


Figure 7.21: Model 3 production rates for $\alpha = 0.5$ using TPWL with optimised basis and local resolution ($l_p = 90$, $l_s = 90$, $n_{LR} = 72$).

are stable and generally accurate and are of about the same quality as those presented for Model 2. This is encouraging, as Model 3 represents a more challenging test case.

We also generated a TPWL model for this case using the EDP approach, again with $l_p = 90$, $l_s = 90$ and $n_{LR} = 72$. As indicated above, constructing this TPWL model requires that two training runs be performed. Results for production and injection rates for $\alpha = 0.5$ are shown in Figures 7.23 and 7.24. These results are again comparable to those using the optimised basis method.

Errors for both sets of runs for a range of α are presented in Table 7.5.2. As was observed for Model 2, error increases with increasing α and the EDP method is slightly more accurate than the optimised basis method, though both provide comparable levels of accuracy.

The runtimes for the full-order models for the two examples considered in this section were 30 minutes and 50 minutes respectively on an Opteron dual-core CPU. The TPWL models, by contrast, required at most a few seconds. Thus runtime speedups of $O(10^3)$ were achieved. We reiterate, however, that TPWL requires a training run and additional overhead comparable to about another full-order simulation.

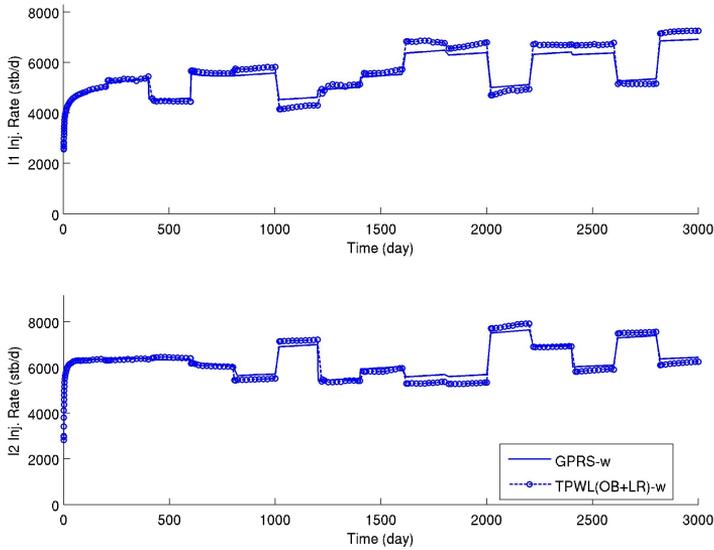


Figure 7.22: Model 3 injection rates for $\alpha = 0.5$ using TPWL with optimised basis and local resolution ($l_p = 90$, $l_s = 90$, $n_{LR} = 72$).

Taken in total, the results presented in this section demonstrate that our enhanced TPWL procedures are able to provide stable results of reasonable accuracy for challenging reservoir simulation problems. This suggests that these approaches may indeed be applicable in computational optimisation or uncertainty quantification procedures. In the next section, we apply the enhanced TPWL method to a production optimisation problem.

Method/ α	0.1	0.3	0.5	0.7	0.9
TPWL(OB+LR), E_o	0.0044	0.0132	0.0222	0.0308	0.0398
TPWL(OB+LR), E_w	0.0063	0.0181	0.0298	0.0412	0.0526
TPWL(OB+LR), E_i	0.0064	0.0192	0.0326	0.0457	0.0595
TPWL(EDP+LR), E_o	0.0033	0.0098	0.0164	0.0229	0.0295
TPWL(EDP+LR), E_w	0.0043	0.0128	0.0217	0.0308	0.0403
TPWL(EDP+LR), E_i	0.0057	0.0173	0.0291	0.0401	0.0530

Table 7.5: Errors for TPWL(OB+LR) and TPWL(EDP+LR) for Model 3.

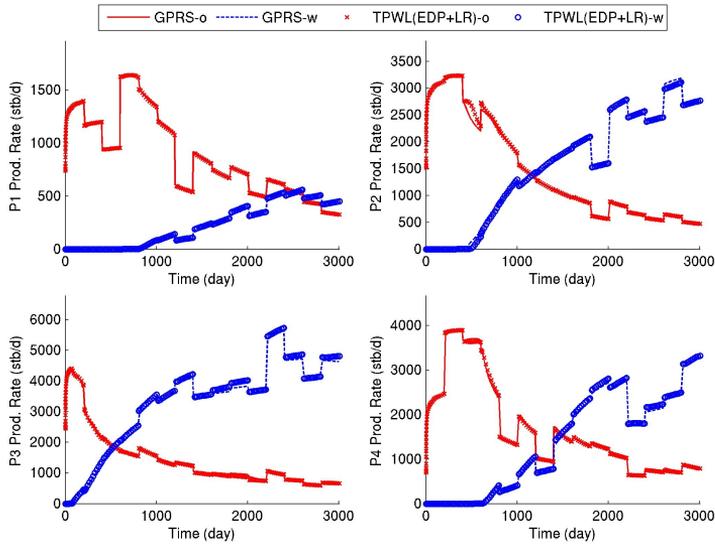


Figure 7.23: Model 3 production rates for $\alpha = 0.5$ using TPWL with EDP basis and local resolution ($l_p = 90, l_s = 90, n_{LR} = 72$).

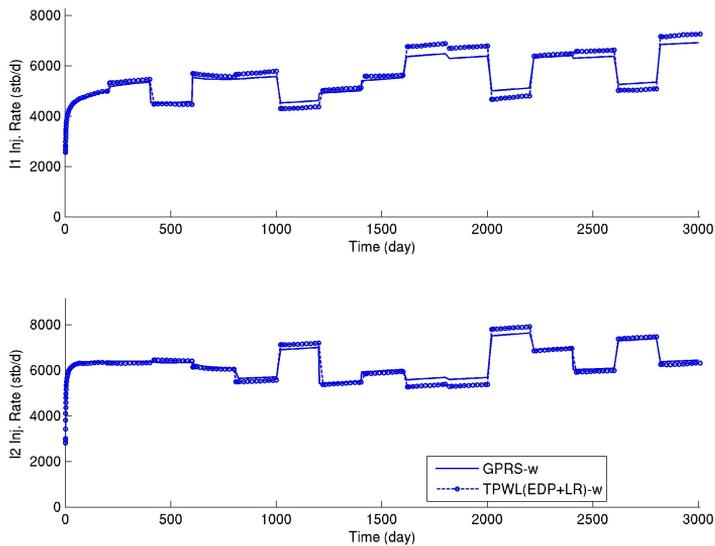


Figure 7.24: Model 3 injection rates for $\alpha = 0.5$ using TPWL with EDP basis and local resolution ($l_p = 90, l_s = 90, n_{LR} = 72$).

7.6 Use of TPWL for Production Optimisation

In this section we use TPWL as a surrogate within a generalised pattern search optimisation algorithm. The optimisation targets the maximisation of net present value for oil production under water injection. We note that TPWL was used previously as a surrogate in gradient-based optimisations (with gradients computed numerically) in Cardoso and Durlofsky (2010b), and Cardoso and Durlofsky (2010a). That work did not consider systematic retraining of the TPWL model, which we do incorporate here.

7.6.1 Direct Search Optimisation with TPWL

Surrogate modelling is widely used for simulation-based optimisation when the full-order (high-fidelity) model is computationally expensive to evaluate. A surrogate model should be computationally inexpensive and at least locally accurate. TPWL should be well suited for use as a surrogate as it appears to provide a sufficiently accurate approximation of the true solution within a reasonably sized neighbourhood around the training case.

The direct search method used here is generalised pattern search (GPS). GPS computes a sequence of points that approach an optimal point. The algorithm applies polling, which entails the evaluation of solutions defined by a stencil (aligned with the coordinate directions) in the search space. The central point of the stencil is the current (best) solution. If an improvement in the cost function is found, the stencil is shifted such that it is centred on the improved point. If an improved solution is not found, the stencil size is decreased. See Kolda et al. (2003) for more detail on GPS and Echeverría Ciaurri et al. (2010) for application of GPS and related procedures to oil production optimisation problems.

Our approach for incorporating TPWL into GPS is depicted in Figure 7.25. We start by performing a training simulation with well BHPs defined by the initial guess. The states and Jacobian matrices are saved and the stabilised TPWL model is constructed using the basis optimisation procedure described earlier. Then, the GPS optimisation is started using the TPWL surrogate for function evaluations. After a specified number of function evaluations are performed, GPS is paused and a training simulation is run at the current best point (the specified number of function evaluations can vary during the course of the optimisation). TPWL is then retrained at this point and GPS is resumed. It occasionally happens that, upon retraining, the objective function of the current point, evaluated using the full-order model, is suboptimal relative to that of the previous full-order solution. This inconsistency can occur when the TPWL solution loses accuracy because it

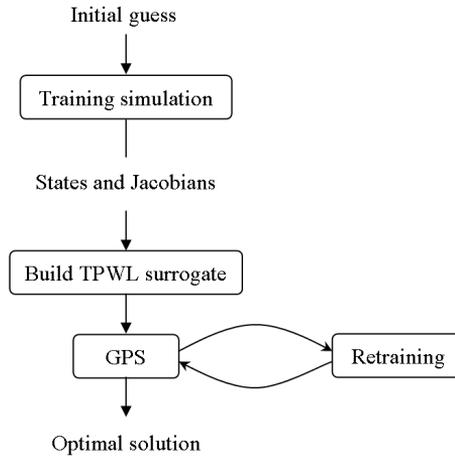


Figure 7.25: Flowchart for generalised pattern search with TPWL.

is too far from the most recent training case. When this problem is detected, we restart the search from the previous retraining point and reduce the number of function evaluations until the next retraining. The size of the GPS mesh may also be reduced. We note that it should be possible to incorporate more sophisticated criteria, possibly based on mass balance errors in the TPWL model (which are straightforward to compute), for retraining.

7.6.2 Optimisation Results

In this example we optimise using a small model to enable comparison of results with the surrogate (TPWL) procedure to those using the full-order model. The reservoir model for this case comprises the first four layers of Model 1 (Model 1 is shown in Figure 7.1) and contains 4800 grid blocks. The fluid and rock-fluid properties and well locations are the same as for Model 1 in Section 7.4.1 ($\rho_o \neq \rho_w$ for this case). In this example, we optimise the production well BHPs to maximise net present value (NPV) over five years (1800 days) of production. The BHP of each well is changed every 200 days. Thus there are nine control variables for each producer, giving a total of 36 control variables. Injection well BHPs are set to 6000 psi for the entire simulation. The oil price is specified to be \$80/bbl while the cost of water produced and injected are \$36/bbl and \$18/bbl, respectively. Water prices are set to be artificially high to limit the use of water. The bounds for the production well BHPs are 1000 psi and 3000 psi. Initially, the BHPs for the four production wells are set to 1500 psi for the entire production period.

The TPWL model here used basis optimisation with basis switching.

Method	Init. NPV (\$10 ⁶)	Final NPV(\$10 ⁶)	# of full sim.
Full-order GPS	49.9	170.1	2500
TPWL guided GPS	49.9	169.0	15

Table 7.6: Optimisation results using TPWL surrogate and full-order models.

Typical TPWL parameter values were $l_p = 65$, $l_s = 75$ and $n_{LR} = 50$. The evolution of NPV with the number of simulations is shown in Figure 7.26 and summarised in Table 7.6. In the figure, the red curve represents the optimisation results using the full-order simulation model while the blue curve presents results using the TPWL model. The circles indicate points where the TPWL surrogate model was retrained. It is evident that, using only 15 full-order training simulations, the TPWL guided optimisation provides essentially the same result as was achieved using the full-order simulations. We note, however, that the detailed well BHP controls do show some differences between the two optimisations. It is frequently observed in oil production optimisations that, even when the resulting objective functions are essentially identical, the well controls determined by different optimisation procedures can be somewhat different (see, e.g., Echeverría Ciaurri et al. (2010)). This presumably results from the structure of the objective function in the high-dimensional search space.

The TPWL overhead in this case required the equivalent of about another 10 training simulations. Thus the equivalent number of full-order runs associated with TPWL model construction was about 25, which is a factor of 100 less than that required for the optimisation based on full-order simulations. Because our TPWL implementation is currently in Matlab, we expect that TPWL model construction could be accelerated considerably through a careful C++ implementation.

Another example, involving a larger reservoir model, the determination of optimal BHPs for both injection and production wells, and nonlinear constraints in the optimisation, is presented in He (2010). This example, along with those presented in Cardoso and Durlofsky (2010b) and Cardoso and Durlofsky (2010a), further demonstrate the applicability of TPWL for use in oil production optimisations.

7.7 Concluding Remarks

In this work, we developed several new techniques for use in trajectory piecewise linearisation (TPWL) procedures. Although our target application area

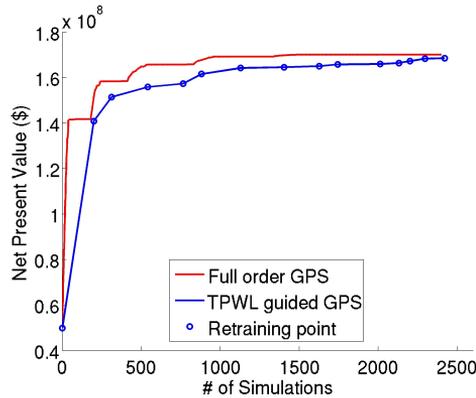


Figure 7.26: NPV evolution during optimisation.

is subsurface flow – specifically oil reservoir simulation – the methods presented here should also find application in other disciplines. The specific methods developed in this paper include a local resolution treatment, in which particular grid blocks are not projected into the low-dimensional space but are instead represented in the original full-order space, and two stabilisation procedures. One of the stabilisation techniques (referred to as OB) entails the determination of the optimum number of columns in each portion of the basis matrix while the other approach (EDP) applies a basis derived from a simulation with phases of equal density. The use of snapshots from equal-density simulations is shown to lead to basis matrices with better stability properties. The EDP approach, when used by itself, results in a degradation of accuracy in the TPWL model. Accuracy in key quantities such as well rates is recovered, however, by using the EDP method in combination with the local resolution treatment.

Numerical simulations were performed using two highly heterogeneous reservoir models, both containing $O(10^5)$ grid blocks. These results demonstrated the accuracy and robustness of the new TPWL procedures. The EDP approach was observed to be slightly more accurate than the OB method, though the two techniques exhibited levels of accuracy that were generally comparable. The EDP technique is more flexible in terms of the size of the basis matrix, though it requires an additional full-order training run and must be used with local resolution. Since neither stabilisation approach is superior in all regards, we believe further development and testing of both methods are warranted.

The TPWL model was also used as a surrogate within a generalised pattern search (GPS) optimisation procedure. Retraining was applied to

update the TPWL model during the course of the GPS run. The objective function computed using the surrogate model was very close to that using the full-order model. A reduction of about a factor of 100 in the number of full-order simulations required for the optimisation was achieved through use of the TPWL surrogate. Future work should be directed toward formalising the determination of when retraining is required in optimisation applications. It may also be useful to view the TPWL surrogate within the context of space-mapping procedures, which are widely used in surrogate-based optimisation.

Acknowledgements

We are grateful to the industry sponsors of the Stanford Reservoir Simulation Research (SUPRI-B) and Smart Fields Consortia for partial funding of this work. We also thank Marco Cardoso for providing the original TPWL code and Obi Isebor for discussions regarding the use of generalised pattern search.

References

- V. Adamjan, D. Arov, and M. Krein. Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. *Math. USSR Sbornik*, 15:31–73, 1971.
- P. Astrid and A. Verhoeven. Application of least squares MPE technique in the reduced order modeling of electrical circuits. In *17th International Symposium on Mathematical Theory of Networks and Systems*, pages 1980–1986, Kyoto, Japan, July 2006.
- K. Aziz and A. Settari. *Fundamentals of Reservoir Simulation*. Elsevier Applied Science Publishers, 1986.
- G. Berkooz and E. Z. Titi. Galerkin projections and the proper orthogonal decomposition for equivariant equations. *Physics Letter A*, 174(1-2):94–102, 1993.
- M. Bettayeb, L. Silverman, and M. Safonov. Optimal approximation of continuous-time system. In *Proceedings of the IEEE Conference on Decision and Control*, volume 1, Albuquerque, NM, 1980.
- B. N. Bond and L. Daniel. Stabilizing schemes for piecewise-linear reduced order models via projection and weighting functions. In *Proceedings of the*

- IEEE/ACM International Conference on Computer-Aided Design*, pages 860–867, San Jose, California, 2007a.
- B. N. Bond and L. Daniel. A piecewise-linear moment-matching approach to parameterized model-order reduction for highly nonlinear systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 26(12):2116–2129, 2007b.
- B. N. Bond and L. Daniel. Guaranteed stable projection-based model reduction for indefinite and unstable linear systems. In *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pages 728–735, San Jose, California, 2008.
- B. N. Bond and L. Daniel. Stable reduced models for nonlinear descriptor systems through piecewise-linear approximation and projection. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 28(10):1467–1480, 2009.
- T. Bui-Thanh, M. Damodaran, and K. Willcox. Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. *AIAA Journal*, 42(8):1505–1516, 2004.
- T. Bui-Thanh, K. Willcox, O. Ghattas, and B. van Bloemen Waanders. Goal-oriented, model-constrained optimization for reduction of large-scale systems. *Journal of Computational Physics*, 224(2):880–896, 2007.
- H. Cao. *Development of Techniques for General Purpose Simulators*. PhD thesis, Stanford University, 2002.
- M. A. Cardoso and L. J. Durlofsky. Use of reduced-order modeling procedures for production optimization. *SPE Journal*, 2010a. In Press.
- M. A. Cardoso and L. J. Durlofsky. Linearized reduced-order models for subsurface flow simulation. *Journal of Computational Physics*, 229(3):681–700, 2010b.
- M. A. Cardoso, L. J. Durlofsky, and P. Sarma. Development and application of reduced-order modeling procedures for subsurface flow simulation. *International Journal for Numerical Methods in Engineering*, 77(9):1322 – 1350, 2009.
- S. A. Castro. *A Probabilistic Approach to Jointly Integrate 3D/4D Seismic Production Data and Geological Information for Building Reservoir Models*. PhD thesis, Stanford University, 2007.

- M. A. Christie and M. J. Blunt. Tenth SPE comparative solution project: A comparison of upscaling techniques. *SPE Reservoir Evaluation & Engineering*, 4(4):308 – 317, 2001.
- D. Echeverría Ciaurri, O. J. Isebor, and L. J. Durlofsky. Application of derivative-free methodologies to generally constrained oil production optimization problems. In *International Conference on Computational Science (ICCS)*, Amsterdam, The Netherlands, 2010.
- P. Feldmann and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14:639–649, 1995.
- I. Gelfand. Zur theorie der caractere der abelschen topologischen gruppen. *Mat.Sb.*, 9(51):49–50, 1941.
- D. Gratton and K. Willcox. Reduced-order, trajectory piecewise-linear models for nonlinear computational fluid dynamics. In *34th AIAA Fluid Dynamics Conference and Exhibit*, pages 2004–2329, Portland, Oregon, USA, 2004.
- T. Hastie, R. Tibshirani, and J. Freidman. *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. Springer, New York, 2 edition, 2009.
- J. He. Enhanced linearized reduced-order models for subsurface flow simulation. Master’s thesis, Stanford University, 2010.
- T. Heijn, R. Markovinović, and J. D. Jansen. Generation of low-order reservoir models using system-theoretical concepts. *SPE Journal*, 9(2):202–218, 2004.
- Y. Jiang. *A Flexible Computational Framework for Efficient Integrated Simulation of Advanced Wells and Unstructured Reservoir Models*. PhD thesis, Stanford University, 2007.
- T. G. Kolda, R. M. Lewis, and V. Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Review*, 45(3):385–482, 2003.
- J. L. Lumley. Atmospheric turbulence and radio wave propagation. *Journal of Computational Chemistry*, 23(13):1236–1243, 1967.
- K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979.

- B. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26:17–31, 1981.
- K. Pearson. On lines and planes of closest fit to points in space. *Philosophical Magazine*, 2:559–572, 1901.
- M. Rewienski and J. White. A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 22(2):155–170, 2003.
- M. J. Rewienski. *A Trajectory Piecewise-Linear Approach to Model Order Reduction of Nonlinear Dynamical Systems*. PhD thesis, Massachusetts Institute of Technology, 2003.
- J. F. M. van Doren, R. Markovinović, and J. D. Jansen. Reduced-order optimal control of water flooding using proper orthogonal decomposition. *Computational Geosciences*, 10:137–158, 2006.
- D. Vasilyev, M. Rewienski, and J. White. Macromodel generation for BioMEMS components using a stabilized balanced truncation plus trajectory piecewise-linear approach. *IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems*, 25(2):285–293, 2006.
- S. S. Vempala. *The Random Projection Method*. American Mathematical Society, 2004.
- P. T. M. Vermeulen, A. W. Heemink, and C. B. M. T. Stroet. Reduced models for linear groundwater flow models using empirical orthogonal functions. *Advances in Water Resources*, 27:57–69, 2004.
- Y. J. Yang and K. Y. Shen. Nonlinear heat-transfer macromodeling for MEMS thermal devices. *Journal of Micromechanics and Microengineering*, 15(2):408–418, 2005.