Universität Bremen

Doctoral Thesis

Covariance Update in Data Assimilation for State and Parameter Estimation

Author: Yovany CORDERO HERNANDEZ
Supervisor: Prof. Dr. Angelika BUNSE-GERSTNER

A thesis submitted in fulfilment of the requirements for the degree of Dr.-Ing.
in the AG Numerik
Zentrum für Technomatematik

September 2015
Date of Defense: 15.09.2015

1\textsuperscript{st} referee: Prof. Dr. Angelika Bunse-Gerstner

2\textsuperscript{nd} referee: Prof. Dr. Nancy K. Nichols
Declaration of Authorship

Corr. § 6(5) Nr. 1-3 PromO

I herewith declare that I

- have completed this dissertation unassisted.
- did not use more than the stated sources and aid.
- have cited all references.

Gem. § 6(5) Nr. 1-3 PromO

Ich erkläre hiermit, dass ich

- die Arbeit ohne fremde Hilfe angefertigt habe.
- keine anderen, als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.
- die den benutzen Werken wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Signed:

Date:
“If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is.”

John Louis von Neumann
The goal of Data Assimilation (DA) is to estimate, with the aid of numerical methods, the true state of a dynamical system, taking into account observations (measurements), a forecast model and state, as well as statistical information of observation, model and forecast errors. DA is also crucial for improving mathematical models, which are defective or inaccurate and contain unproven data and unknown parameters respectively, for simulation and control. This research field has an extraordinary number of application areas, being the simulation and computation of climate and ocean models among the most important and studied.

In the context of data assimilation problems it is common that models depend on poorly known parameters. A well-known approach is to solve both problems, data assimilation and parameter estimation, at the same time following the so-called augmented state approach. The idea is to solve a modified DA problem, where the parameters are considered state variables and included in an augmented state vector, and the parameter evolution dynamics are incorporated into a new augmented forecast model.

Typically, parameters are not directly observed; therefore, their estimation depends on the ability of the assimilation scheme to infer parameter updates using the information obtained from the observational data. The interrelation between state variables and parameters is given by their correlations. In this work we investigate the influence of the augmented state covariance matrix on the joint state-parameter data assimilation problem. Moreover, using the augmented approach we propose a novel method based on a low-cost update of the augmented state covariance matrix. Furthermore, we find necessary and sufficient conditions for the convergence of 3D-Var methods, and in particular of our proposed strategy, when applied to linear state-parameter DA problems. The suitability of our proposed method is tested using several benchmark problems.
Zusammenfassung


Acknowledgements

This dissertation would not have been possible without the kind help of many people to whom I would like to extend my sincere thanks.

Firstly, I would like to thank my supervisor, Prof. Dr. Angelika Bunse-Gerstner, for giving me the opportunity to join her research group to conduct this investigation. Thanks for the invaluable guidance and fruitful discussions.

Special thanks to Prof. Dr. Nancy Nichols and Prof. Dr. Sarah Dance, who made possible my research stay in Reading and whose contribution was fundamental for completing an important chapter of this thesis. I would also like to thank them and Mike Green for their constant support during my visit.

I would like to show my gratitude to Dr. Dorota Kubalinska, colleague and friend, for her company and useful advice, to Dr. Lars Nerger, for his generous help and to Dr. Iwona Piotrowska-Kurczewski, for her logistic support. Thanks also to our secretary Ebba Feldmann, who always helped me with bureaucracy issues.

I am immensely grateful to my dear friend Dr. Jonathan Montalvo-Urquizo, for his generous proofreading and helpful comments on an earlier version of the manuscript. I would also like to thank my friend Dr. Adriel Domínguez García, for his support during the final stage of this investigation.

Thanks to my friend Alejandro Reyes, for the Latex tips and his friendship without limits, which made the dark days a bit brighter and gave me fuel to continue fighting.

Last but not least I would like to thank my family, who always supported me in all sort of ways and who believed in me faithfully. Thanks for walking along this way with me, despite being so far. Especially I would like to thank my wife, who has been there every day, sharing the good and bad moments and without whom everything would have been twice as hard and would have lost half of the fun.
# Contents

Declaration of Authorship ii

Abstract iv

Zusammenfassung v

Acknowledgements vi

Contents vii

List of Figures x

Acronyms xvii

List of Symbols xviii

## 1 Introduction

1.1 Goals of this investigation ........................................... 3
1.2 Contribution of the thesis ........................................... 4
1.3 Structure of the thesis .............................................. 5

## 2 Data Assimilation Overview

2.1 Problem formulation .................................................. 7
2.2 Sequential DA ...................................................... 8
    2.2.1 Optimal Interpolation and 3D-Var ............................ 9
    2.2.2 Kalman Filter .............................................. 12
    2.2.3 Extended Kalman Filter .................................... 13
    2.2.4 Ensemble Kalman Filter .................................... 14
2.3 Four dimensional variational DA ................................... 16
2.4 Summary and discussion ............................................ 18

## 3 Parameter Estimation

3.1 Parameter identifiability and sensitivity ........................... 19
3.2 Model calibration ................................................. 20
3.3 Dual state-parameter DA .......................................... 21
### Contents

3.4 Augmented state DA .................................. 23  
3.4.1 Formulation of the augmented state approach .......... 23  
3.4.2 Applications of the augmented state approach .......... 25  
3.5 Dual state-parameter vs augmented state DA .............. 25  
3.6 Summary and discussion ................................ 26

4 EKF and the Augmented State Problem 28
4.1 Forecast step .................................. 29  
4.2 Assimilation step ................................ 29  
4.3 Linear advection model. Assimilation experiments .......... 31  
4.3.1 Linear advection model ................................ 31  
4.3.2 Discrete linear advection model ....................... 32  
4.3.3 Augmented state for the advection model ............... 33  
4.3.4 Experiment set-up ................................ 34  
4.3.5 Results of standard EKF experiments ................ 35  
4.3.6 Results fixing the parameter variance ................ 39  
4.3.7 Results fixing the size of the norm of $B_{xp}$ .......... 41  
4.3.8 Results considering “model errors” .................. 43  
4.4 Summary and discussion ................................ 47

5 Methods which Combine Ideas from Different Assimilation Schemes 50
5.1 Flow-dependent $B_{xp}$ method .......................... 51  
5.1.1 Interpretation of the cross-covariance update .......... 52  
5.2 Low-rank $B_{xx}$ update. Our proposed method .......... 55  
5.2.1 Forward step covariance update ...................... 56  
5.2.2 Computational implementation ....................... 58  
5.2.3 LRSSC algorithm ................................ 59  
5.3 Summary and discussion ................................ 59

6 Convergence of 3D-Var-like Methods for the Linear Models 61
6.1 Problem statement .................................. 62  
6.2 Resulting dynamical system ................................ 62  
6.3 State error dynamical system ................................ 64  
6.4 Convergence ..................................... 65  
6.5 Full and accurate observations .......................... 66  
6.5.1 Single parameter ................................ 69  
6.5.1.1 Single parameter. FDSPC method ............... 69  
6.5.1.2 Single parameter. LRSSC method ............... 70  
6.6 Assimilation when the dimension of the observations is smaller than the number of parameters .................. 71  
6.7 Single observation and single parameter .................. 73  
6.7.1 Necessary condition for convergence .................. 73  
6.7.2 Sufficient condition for convergence .................. 77  
6.8 Precise background state estimation ...................... 79  
6.8.1 Necessary conditions for convergence ................. 79  
6.9 Numerical experiments ................................ 81  
6.9.1 1-D heat equation ................................ 81
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Sequential DA scheme.</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>KF algorithm.</td>
<td>13</td>
</tr>
<tr>
<td>2.3</td>
<td>EKF algorithm.</td>
<td>14</td>
</tr>
<tr>
<td>2.4</td>
<td>4D-Var scheme.</td>
<td>16</td>
</tr>
<tr>
<td>3.1</td>
<td>Dual state-parameter KF</td>
<td>22</td>
</tr>
<tr>
<td>4.1</td>
<td>Linear advection problem. Bed height moving at horizontal speed s.</td>
<td>32</td>
</tr>
<tr>
<td>4.2</td>
<td>Data assimilation for initial background speed 0.3 and different observation frequencies.</td>
<td>36</td>
</tr>
<tr>
<td>4.3</td>
<td>Data assimilation for different initial background speeds for observations available every 30 time steps.</td>
<td>37</td>
</tr>
<tr>
<td>4.4</td>
<td>Parameter variance and cross-covariance vector $B_{xp}$ for different initial background speeds for observations available every 30 time steps.</td>
<td>38</td>
</tr>
<tr>
<td>4.5</td>
<td>Parameter variance and norm of cross-covariance vector $B_{xp}$ for different assimilation frequencies and initial speed 0.3.</td>
<td>38</td>
</tr>
<tr>
<td>4.6</td>
<td>Data assimilation for 0.3 initial background speeds and observations available every 30 time steps with fixed parameter variance along the assimilation window.</td>
<td>39</td>
</tr>
<tr>
<td>4.7</td>
<td>Parameter variance and cross-covariance vector $B_{xp}$ for assimilations every 30 time steps and initial speed 0.3.</td>
<td>40</td>
</tr>
<tr>
<td>4.8</td>
<td>Data assimilation for 0.3 initial background speeds and observations available every 30 time steps with fixed parameter variance along the assimilation window and averaging parameter update.</td>
<td>41</td>
</tr>
<tr>
<td>4.9</td>
<td>Data assimilation for 0.3 initial background speeds and observations available every 30 time steps. After each assimilation $B_{xp}^a$ is resized such that $|B_{xp}^a|0 = c$ for a given constant $c$ and the parameter is averaged over a moving time window.</td>
<td>44</td>
</tr>
<tr>
<td>4.10</td>
<td>Parameter variance and cross-covariance vector $B_{xp}$ for assimilations every 30 time steps and initial speed 0.3. After each assimilation $B_{xp}^a$ is resized such that $|B_{xp}^a|0 = c$ for a given constant $c$ and the parameter is averaged over a time window.</td>
<td>44</td>
</tr>
<tr>
<td>4.11</td>
<td>Data assimilation for 0.3 initial background speeds and observations available every 30 time steps with fixed parameter variance along the assimilation window and averaging parameter update.</td>
<td>47</td>
</tr>
<tr>
<td>5.1</td>
<td>State-state error covariance evolution for the linear advection model, without assimilating observations and for a fixed speed value.</td>
<td>55</td>
</tr>
<tr>
<td>5.2</td>
<td>State-state error covariance evolution for the linear advection model, without assimilating observations and for an unknown speed parameter.</td>
<td>56</td>
</tr>
</tbody>
</table>
### List of Figures

5.3 Rank 1 matrices $N_{100}B_{pp}N_{100}^\top$ and $N_{200}B_{pp}N_{200}^\top$ .......................... 56

6.1 Region of admissible eigenvalues of matrix $B_{xp}^\top B_{xx}^{-1}T$ to ensure convergence. 67

6.2 Regions of admissible values of the parameter variance $\sigma_p^2$, depending on the values of $z$. .......................... 75

6.3 FDSPC method. Augmented state covariance matrix $\hat{B}$ and iteration matrix $\hat{M}$ for uncorrelated parameter errors and parameter variance $\sigma_p^2 = 0.05$. ........................................ 84

6.4 FDSPC method. Assimilated state for $\sigma_p^2 = 0.05$ ........................................ 84

6.5 Eigenvalues and unit circles. ........................................ 85

6.6 Assimilated state for $\sigma_p^2 = 25 \times 10^{-3}$ ........................................ 85

6.7 Parameter estimation along the time window for $\sigma_p^2 = 25 \times 10^{-3}$ and uncorrelated parameters. ........................................................................................................ 86

6.8 Assimilated state for $L = \frac{n}{2} \Delta z$ and $\sigma_p^2 = \sigma_p^{2*}$. .............................. 88

6.9 Parameter estimation along the time window for $L = \frac{n}{2} \Delta z$, $\sigma_p^2 = \sigma_p^{2*}$ and uncorrelated parameters. ........................................ 89

6.10 Assimilated state for $L = \frac{n}{8} \Delta z$ and $\sigma_p^2 = \sigma_p^{2*}$. .................................. 89

6.11 Parameter estimation along the time window for $L = \frac{n}{8} \Delta z$, $\sigma_p^2 = \sigma_p^{2*}$ and uncorrelated parameters. ........................................ 90

6.12 Eigenvalues and unit circles. ........................................ 90

6.13 Assimilated state for $L = \frac{n}{8} \Delta z$, $\sigma_p^2 = \sigma_p^{2*}$ and $\sigma_o^2 = 0.1$. ........... 91

6.14 Parameter estimation along the time window for $L = \frac{n}{8} \Delta z$, $\sigma_p^2 = \sigma_p^{2*}$, $\sigma_o^2 = 0.1$ and uncorrelated parameters. ........................................ 91

6.15 LRSSC. Augmented state covariance matrix $B$ and iteration matrix $\hat{M}$ for uncorrelated parameter errors and parameter variance $\sigma_p^2 = 0.05$. ........................................ 92

6.16 LRSSC. Assimilated state for $\sigma_p^2 = 25 \times 10^{-3}$ ........................................ 92

6.17 LRSSC. Parameter estimation along the time window for $\sigma_p^2 = 0.05$ and uncorrelated parameters. ........................................ 93

6.18 Covariance error augmented state matrix $B$ and iteration matrix $\hat{M}$ .......................... 94

6.19 Assimilated state for $L = \frac{n}{8} \Delta z$, $B_{pp} = B_{xx}$. ........................................ 94

6.20 Parameter estimation along the time window for $L = \frac{n}{8} \Delta z$, $B_{pp} = B_{xx}$. 95

6.21 Assimilated state for $L = n \Delta z$, $B_{pp} = B_{xx}$. ........................................ 96

6.22 Parameter estimation along the time window for $L = n \Delta z$, $B_{pp} = B_{xx}$. 96

6.23 Assimilated state for $L = n \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$. ........................ 97

6.24 Parameter estimation along the time window for $L = n \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$. ........................................ 97

6.25 Eigenvalues of $\hat{M}$ for unobserved state variables ........................................ 98

6.26 Covariance error augmented state matrix $B$ and iteration matrix $\hat{M}$ for unobserved state variables ........................................ 98

6.27 Assimilated state for unobserved states, $L = \frac{n}{8} \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$. 98

6.28 Parameter estimation along the time window for unobserved states, $L = n \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$. ........................................ 99

6.29 Assimilated state for $L = \frac{n}{8} \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$ with random initial parameter. ........................................ 100

6.30 Parameter estimation along the time window for $L = n \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$ with random initial parameter. ........................................ 100

6.31 Assimilated state for $L = \frac{n}{8} \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$. All state variables are observed except where the source is located ........................................ 101
6.32 Parameter estimation along the time window for $L = n\Delta z$, $B_{pp} = B_{xx}$ and $\sigma^2 = 0.1$. All state variables are observed except where the source is located. .......................................................... 101

7.1 Linear Advection. Assimilated parameter along the time window for different assimilation intervals and perfect observations, initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$ and $\Delta x_{obs} = 10$. .................................................. 107

7.2 Linear Advection. Norm of the assimilated state error vector for perfect observations, initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$ and $\Delta x_{obs} = 10$. .................................................. 108

7.3 Linear Advection. Background and assimilated state for $t = 0, 1, 2, 3, 4, 10$ and perfect observations, where initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$, $\Delta x_{obs} = 10$ and $\Delta t_{obs} = 25$. .................................................. 108

7.4 Linear Advection. Background and assimilated state for $t = 0, 1, 2, 3, 4, 10$ and perfect observations, where initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$, $\Delta x_{obs} = 10$ and $\Delta t_{obs} = 50$. .................................................. 109

7.5 Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window. Comparison between the FDSPC and the LRSSC methods for perfect observations, initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$, $\Delta x_{obs} = 10$ and $\Delta t_{obs} = 100$. .................. 110

7.6 Linear Advection. Norm of the assimilated state error vector for the FDSPC and the LRSSC methods for perfect observations, initial background parameter $s_b = 0.3$, $\sigma^2 = 0.01$, $\Delta x_{obs} = 10$ and $\Delta t_{obs} = 100$. .................. 110

7.7 Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window. Comparison between the FDSPC and the LRSSC methods for perfect observations, initial background parameter $s_b = 0.3$, $\sigma^2 = 0.01$, $\sigma^2 = 0.3$, $\Delta x_{obs} = 10$ and $\Delta t_{obs} = 100$. .................. 111

7.8 Linear Advection. Norm of the assimilated state error vector for the FDSPC and the LRSSC methods for perfect observations, initial background parameter $s_b = 0.3$, $\sigma^2 = 0.01$, $\sigma^2 = 0.3$, $\Delta x_{obs} = 10$ and $\Delta t_{obs} = 100$. .................. 112

7.9 Linear Advection. Assimilated parameter along the time window for different observation spacing and perfect observations, initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$ and $\Delta t_{obs} = 15$. .................. 113

7.10 Linear Advection. Norm of the assimilated state error vector for different observation spacing and perfect observations, initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$ and $\Delta t_{obs} = 15$. .................. 113

7.11 Linear Advection. Background, assimilated and true state for $t = 0, 1, 2, 3, 4, 10$ and perfect observations, where initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$, $\Delta x_{obs} = 25$ and $\Delta t_{obs} = 15$. .................. 114

7.12 Linear Advection. Background and assimilated state for $t = 0, 1, 2, 3, 4, 10$ and perfect observations, where initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$, $\Delta x_{obs} = 40$ and $\Delta t_{obs} = 15$. .................. 114

7.13 Linear Advection. Assimilated parameter along the time window. Comparison between the FDSPC (red) and the LRSSC (blue) methods for perfect observations, initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$, $\Delta x_{obs} = 50$ and $\Delta t_{obs} = 15$. .................. 115
7.14 Linear Advection. Norm of the assimilated state error vector for the FDSPC (red) and the LRSSC (blue) methods for different observation spacing and perfect observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$, $\Delta x_{\text{obs}} = 50$ and $\Delta t_{\text{obs}} = 15$. .................................................. 115
7.15 Linear Advection. Assimilated parameter along the time window for different observation variances and perfect observations, initial background parameter $s_b = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. .................................................. 116
7.16 Linear Advection. Norm of the assimilated state error vector for different observation variances and perfect observations, initial background parameter $s_b = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. .................................................. 116
7.17 Linear Advection. Assimilated parameter along the time window for different assimilation intervals and noisy observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. .................................................. 117
7.18 Linear Advection. Norm of the assimilated state error vector for perfect observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. .................................................. 118
7.19 Linear Advection. Assimilated parameter along the time window for different assimilation intervals and noisy observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. .................................................. 119
7.20 Linear Advection. Norm of the assimilated state error vector for noisy observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. .................................................. 119
7.21 Linear Advection. State assimilation at different times, for noisy observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$.120
7.22 Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window. Comparison between the FDSPC and the LRSSC methods for noisy observations, initial background parameter $s_b = 0.3$, $\sigma_o^2 = 0.01$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 50$. .......................... 120
7.23 Linear Advection. Norm of the assimilated state error vector for the FDSPC and the LRSSC methods for noisy observations, initial background parameter $s_b = 0.3$, $\sigma_o^2 = 0.01$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 50$. .......................... 121
7.24 Linear Advection. Assimilated parameter along the time window for different observation spacing and noisy observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. .................................................. 122
7.25 Linear Advection. Norm of the error in the state assimilation along the time window for different observation spacing and noisy observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta t_{\text{obs}} = 15$. .......................... 123
7.26 Linear Advection. State estimation at different times, for different observation spacing and noisy observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta t_{\text{obs}} = 15$. .......................... 123
7.27 Linear Advection. Assimilated parameter along the time window for different observation variances and noisy observations, initial background parameter $s_b = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. .......................... 124
7.28 Linear Advection. Norm of the assimilated state error vector for different observation variances and noisy observations, initial background parameter $s_b = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. .......................... 124
7.29 Linear Advection. State estimation at different times, for different observation variances and noisy observations, initial background parameter $s_b = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. .......................... 125
7.30 Linear Advection. Parameter estimation for the FDSPC and the LRSSC methods for noisy observations, initial background parameter $s_b = 0.3$, $\sigma^2_o = 1$, $\sigma^2_p = 0.5$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. .............................. 125

7.31 Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window. Comparison between the FDSPC and the LRSSC methods for noisy observations, initial background parameter $s_b = 0.3$, $\sigma^2_o = 1$, $\sigma^2_p = 0.5$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. .............................. 126

7.32 Linear Advection. Norm of the assimilated state error vector for the FDSPC and the LRSSC methods for noisy observations, initial background parameter $s_b = 0.3$, $\sigma^2_o = 1$, $\sigma^2_p = 0.5$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. .............................. 126

7.33 Advection model. RMSE for different $\alpha$’s and observation frequencies. Left: RMSE mean. Right: RMSE variance. .............................. 128

7.34 Advection model. RMSE for different $\alpha$’s and observation spacings. Left: RMSE mean. Right: RMSE variance. .............................. 128

7.35 Advection model. RMSE for different $\alpha$’s and observation variances. Left: RMSE mean. Right: RMSE variance. .............................. 129

7.36 Nonlinear oscillator. .................................................. 131

7.37 Nonlinear oscillator. Assimilated $x$ for different observation frequencies and perfect observations, $\sigma^2_o = 0.01$. .............................. 134

7.38 Nonlinear oscillator. $|x^f - x^a|$ for different observation frequencies and perfect observations, $\sigma^2_o = 0.01$. .............................. 135

7.39 Nonlinear oscillator. Assimilated $y$ for different observation frequencies and perfect observations, $\sigma^2_o = 0.01$. .............................. 135

7.40 Nonlinear oscillator. $|y^f - y^a|$ for different observation frequencies and perfect observations, $\sigma^2_o = 0.01$. .............................. 136

7.41 Nonlinear oscillator. Assimilated $m$ for different observation frequencies and perfect observations, $\sigma^2_o = 0.01$. .............................. 136

7.42 Nonlinear oscillator. Assimilated $d$ for different observation frequencies and perfect observations, $\sigma^2_o = 0.01$. .............................. 137

7.43 Nonlinear oscillator. Assimilated $x$ for different observation variances and perfect observations, $\Delta t_{\text{obs}} = 10$. .............................. 137

7.44 Nonlinear oscillator. Assimilated $y$ for different observation variances and perfect observations, $\Delta t_{\text{obs}} = 10$. .............................. 138

7.45 Nonlinear oscillator. Assimilated $m$ for different observation variances and perfect observations, $\Delta t_{\text{obs}} = 10$. .............................. 138

7.46 Nonlinear oscillator. Assimilated $d$ for different observation variances and perfect observations, $\Delta t_{\text{obs}} = 10$. .............................. 139

7.47 Nonlinear oscillator. Assimilated $x$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$. .............................. 140

7.48 Nonlinear oscillator. Assimilated $y$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$. .............................. 141

7.49 Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$. .............................. 141

7.50 Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$. .............................. 142

7.51 Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$. .............................. 142
7.52 Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$. .................................................. 143
7.53 Nonlinear oscillator. Assimilated $x$ for different observation variances and noisy observations, $\Delta t_{\text{obs}} = 10$. .................................................. 143
7.54 Nonlinear oscillator. Assimilated $y$ for different observation variances and noisy observations, $\Delta t_{\text{obs}} = 10$. .................................................. 144
7.55 Nonlinear oscillator. Assimilated $m$ for different observation variances and noisy observations, $\Delta t_{\text{obs}} = 10$. .................................................. 145
7.56 Nonlinear oscillator. Assimilated $d$ for different observation variances and noisy observations, $\Delta t_{\text{obs}} = 10$. .................................................. 145
7.57 Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$ and only $y$ is observed. ......................... 146
7.58 Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$ and only $y$ is observed. ......................... 147
7.59 Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$ and only $x$ is observed. ......................... 147
7.60 Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$ and only $x$ is observed. ......................... 148
7.61 Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$ and only $x$ is observed. The state variables are assumed to be correlated with correlation $\sigma_{xy} = 0.01$. ......................... 148
7.62 Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$ and only $x$ is observed. The state variables are assumed to be correlated with correlation $\sigma_{xy} = 0.01$. ......................... 149
7.63 Nonlinear oscillator. RMSE for different $\alpha$’s and observation variances. Left: RMSE mean. Right: RMSE variance. .................................................. 150
7.64 Nonlinear oscillator. Relative error for the assimilated parameter $m$ for different $\alpha$’s and observation variances. Left: $|\frac{m_a - m_t}{m_t}|$ mean. Right: $|\frac{m_a - m_t}{m_t}|$ variance. .................................................. 150
7.65 Nonlinear oscillator. Relative error for the assimilated parameter $m$ for different $\alpha$’s and observation variances. Left: $|\frac{d_a - d_t}{d_t}|$ mean. Right: $|\frac{d_a - d_t}{d_t}|$ variance. .................................................. 151
7.66 Lorenz 63. State assimilation for perfect observations, $\Delta t = 20$ and $\sigma_o^2 = 0.01$. .................................................. 155
7.67 Lorenz 63. Parameter assimilation for different observation frequencies and perfect observations, with $\sigma_o^2 = 0.01$. .................................................. 156
7.68 Lorenz 63. Norm of the assimilated state error vector for perfect observations and $\sigma_o^2 = 0.01$, considering different observation frequencies. .................................................. 156
7.69 Lorenz 63. Parameter assimilation for different observation variances and perfect observations assimilated every $10\Delta t$. .................................................. 157
7.70 Lorenz 63. Norm of the assimilated state error vector for different observation variances and perfect observations assimilated every $10\Delta t$. .................................................. 158
7.71 Lorenz 63. Parameter assimilation for different observation frequencies and noisy observations, with $\sigma_o^2 = 0.01$. .................................................. 159
7.72 Lorenz 63. Norm of the assimilated state error vector for noisy observations and $\sigma_o^2 = 0.01$. .................................................. 159
7.73 Lorenz 63. Parameter assimilation for different observation frequencies and noisy observations, with $\sigma_o^2 = 0.01$. .................................................. 160
7.74 Lorenz 63. Norm of the assimilated state error vector for different observation frequencies and noisy observations, with $\sigma_o^2 = 0.01$.

7.75 Lorenz 63. State assimilation for $\sigma_o^2 = 0.5$ and noisy observations assimilated every $10\Delta t$.

7.76 Lorenz 63. Parameter assimilation for different observation variances and noisy observations assimilated every $10\Delta t$.

7.77 Lorenz 63. Norm of the assimilated state error vector for different observation variances and noisy observations assimilated every $10\Delta t$.

7.78 Lorenz 63. RMSE for different $\alpha$’s and observation frequencies.

7.79 Lorenz 63. RMSE for different $\alpha$’s and observation variances.

7.80 Lorenz 63. Parameter estimation for $\alpha = 1$ and $\alpha = 0.3$, with noisy observations taken every $10\Delta t$ and $\sigma_o^2 = 0.01$.

7.81 Lorenz 63. Norm of the assimilated state error vector for $\alpha = 1$ and $\alpha = 0.3$, with noisy observations taken every $10\Delta t$ and $\sigma_o^2 = 0.01$.

C.1 Nonlinear oscillator. RMSE for different $\alpha$’s and observation variances. Left: RMSE mean. Right: RMSE variance.

C.2 Nonlinear oscillator. Relative error for the assimilated parameter $m$ for different $\alpha$’s and observation variances. Left: $|m_a - m_t|/m_t$ mean. Right: $|m_a - m_t|/m_t$ variance.

C.3 Nonlinear oscillator. Relative error for the assimilated parameter $d$ for different $\alpha$’s and observation variances. Left: $|d_a - d_t|/d_t$ mean. Right: $|d_a - d_t|/d_t$ variance.

C.4 Nonlinear oscillator. EKF assimilation of $x$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$.

C.5 Nonlinear oscillator. EKF assimilation of $y$ for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$.

C.6 Nonlinear oscillator. EKF assimilation of $m$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$.

C.7 Nonlinear oscillator. EKF assimilation of $d$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$.

C.8 Nonlinear oscillator. EKF assimilation of $x$ fixing $B_{pp}$, for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$.

C.9 Nonlinear oscillator. EKF assimilation of $y$ fixing $B_{pp}$, for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$.

C.10 Nonlinear oscillator. EKF assimilation of $m$ fixing $B_{pp}$, for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$.

C.11 Nonlinear oscillator. EKF assimilation of $d$ fixing $B_{pp}$, for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$. 
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D-Var</td>
<td>Three Dimensional Variational assimilation</td>
</tr>
<tr>
<td>4D-Var</td>
<td>Four Dimensional Variational assimilation</td>
</tr>
<tr>
<td>AOGCM</td>
<td>Atmosphere-Ocean General Circulation Model</td>
</tr>
<tr>
<td>BLUE</td>
<td>Best Linear Unbiased Estimation</td>
</tr>
<tr>
<td>DA</td>
<td>Data Assimilation</td>
</tr>
<tr>
<td>ECMWF</td>
<td>European Center Medium-Range Weather Forecast</td>
</tr>
<tr>
<td>EAKF</td>
<td>Ensemble Adjustment Kalman Filter</td>
</tr>
<tr>
<td>EKF</td>
<td>Extended Kalman Filter</td>
</tr>
<tr>
<td>EnKF</td>
<td>Ensemble Kalman Filter</td>
</tr>
<tr>
<td>EnSRF</td>
<td>Ensemble Square Root Filter</td>
</tr>
<tr>
<td>ESTKF</td>
<td>Error Subspace Transform Kalman Filter</td>
</tr>
<tr>
<td>ETKF</td>
<td>Ensemble Transform Kalman Filter</td>
</tr>
<tr>
<td>FDSPC</td>
<td>Flow-Dependent State-Parameter 3D-Var hybrid</td>
</tr>
<tr>
<td>KF</td>
<td>Kalman Filter</td>
</tr>
<tr>
<td>LEKF</td>
<td>Local Ensemble Kalman Filter</td>
</tr>
<tr>
<td>LETKF</td>
<td>Local Ensemble Transform Kalman Filter</td>
</tr>
<tr>
<td>LRSSC</td>
<td>Low-Rank State-State Covariance 3D-Var hybrid</td>
</tr>
<tr>
<td>NWP</td>
<td>Numerical Weather Prediction</td>
</tr>
<tr>
<td>OI</td>
<td>Optimal Interpolation</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>SEEK</td>
<td>Singular Evolutive Extended Kalman</td>
</tr>
<tr>
<td>SEIK</td>
<td>Singular Evolutive Interpolated Kalman</td>
</tr>
</tbody>
</table>
List of Symbols

Data assimilation

\( x \quad \text{state vector} \)
\( y \quad \text{observation vector} \)
\( f \quad \text{nonlinear forward model} \)
\( H \quad \text{nonlinear observation operator} \)
\( M \quad \text{linear forward model} \)
\( M^\top \quad \text{adjoint model} \)
\( K \quad \text{Kalman gain matrix} \)
\( B \quad \text{background error covariance} \)
\( P \quad \text{empirical background covariance} \)
\( C \quad \text{correlation matrix} \)
\( L \quad \text{correlation length scale} \)
\( n \quad \text{state dimension} \)
\( m \quad \text{observation dimension} \)
\( t \quad \text{time} \)
\( \sigma_b^2 \quad \text{background error covariance inflation factor} \)
\( \sigma_p^2 \quad \text{parameter variance} \)
\( \sigma_o^2 \quad \text{observation variance} \)
\( (\cdot)^b \quad \text{superscript referring to background} \)
\( (\cdot)^f \quad \text{superscript referring to forecast} \)
\( (\cdot)^a \quad \text{superscript referring to analysis} \)
\( (\cdot)^t \quad \text{superscript referring to the true parameter or state} \)
List of symbols

\((\cdot)_k\) \quad \text{index referring to the time step}

\(N\) \quad \text{amount of observations}

**Augmented state**

\(p\) \quad \text{parameter vector}

\(\ell\) \quad \text{amount of parameters}

\(\tilde{f}\) \quad \text{augmented state forecast model}

\(\tilde{H}\) \quad \text{augmented state observation operator}

\(H\) \quad \text{linearized augmented state observation operator}

\(M\) \quad \text{Jacobian of the model with respect to the state variables}

\(N\) \quad \text{Jacobian of the model with respect to the parameters}

\(w\) \quad \text{augmented state vector}

\(B_{xx}\) \quad \text{state-state error covariance}

\(B_{zp}\) \quad \text{state-parameter error covariance}

\(B_{pp}\) \quad \text{parameter-parameter error covariance}

\(K_x\) \quad \text{Kalman gain associated to the state update}

\(K_p\) \quad \text{Kalman gain associated to the parameter update}

**Hybrid Approach**

\(B_{xx}^b\) \quad \text{static state-state error covariance}

\(\alpha\) \quad \text{scheme parameter indicating the convex combination between covariances}

**Linear State-Parameter DA**

\(A\) \quad \text{state model matrix}

\(T\) \quad \text{parameter model matrix}

**Linear Advection**

\(x(z, t)\) \quad \text{bed height}

\(s\) \quad \text{advection speed (parameter)}

\(F\) \quad \text{forecast matrix model}

\(\Delta x_{\text{obs}}\) \quad \text{indicates the separation between observed variables}
List of symbols

\[ \mu \quad \text{upwind scheme parameter} \]

**Nonlinear Oscillator**

\[ x(t) \quad \text{mass position with respect to the equilibrium point (state variable)} \]
\[ y(t) \quad \text{mass speed (state variable)} \]
\[ x_1(z) \quad \text{true state} \]
\[ m \quad \text{square of the oscillation frequency (parameter)} \]
\[ d \quad \text{damping coefficient (parameter)} \]

**Heat Equation**

\[ \delta \quad \text{Dirac delta function} \]
\[ v(z, t) \quad \text{temperature of the rod} \]
\[ \alpha(z) \quad \text{function describing the initial temperature of the rod} \]
\[ x \quad \text{discretized temperature of the rod (state variable)} \]
\[ s \quad \text{source vector} \]
\[ p \quad \text{parameter vector} \]
\[ J \quad \text{amount of intervals in which the rod is discretized} \]
\[ \theta \quad \text{theta method parameter} \]
\[ \kappa \quad \text{diffusion constant} \]

**Lorenz 96**

\[ x \quad \text{convective intensity (state variable)} \]
\[ y \quad \text{difference in temperatures between currents (state variable)} \]
\[ z \quad \text{deviation of temperature profile from linearity (state variable)} \]
\[ \sigma \quad \text{Prandtl (parameter)} \]
\[ \rho \quad \text{normalized Rayleigh number (parameter)} \]
\[ \beta \quad \text{geometrical parameter} \]
\[ \sigma_\sigma^2 \quad \text{variance of parameter } \sigma \]
\[ \sigma_\rho^2 \quad \text{variance of parameter } \rho \]
\[ \sigma_\beta^2 \quad \text{variance of parameter } \beta \]
Discretization and Experiments

$\Delta z$  space discretization step
$\Delta t$  time discretization step
$\Delta t_{\text{obs}}$  number of time steps between observations
To Samuel
Chapter 1

Introduction

Mathematical models are used to describe the behaviour of dynamical systems, for forecasting and simulation. They represent a simplification of reality and are associated with an underlying physical problem; for example, mathematical models arising in the context of weather forecast, oil extraction, complex microelectronic systems, among others. These models, which can take the form of a differential or algebraic equation, are typically too complicated to be solved analytically. Therefore it is necessary to simplify and/or approximate them by appropriate numerical schemes, inevitably introducing errors in the mathematical model. Besides, in practical applications the information about the initial state of the system is often imprecise. These inaccuracies can drastically affect the ability of the model to correctly predict the behaviour of the system, as the errors tend to propagate over time. In order to improve the forecast, observations of the system along a time window can be used as feedback. In real applications, these observations are typically prone to errors, generated during the capture, transmission, storage or processing of the information [119].

Data assimilation (DA) is a process which combines observational data and a numerical model, taking into account all sources of errors, to estimate the most probable state of a dynamical system. This research field has an extraordinary number of application areas; being the simulation and computation of climate and ocean models one of the most important and explored.

An important aspect of any DA scheme is the precise estimation of the errors involved. Though some methods, like particle filters (see [82]), approximate the exact statistical distribution of the error, the majority of DA methods rely on a good approximation of the error covariance. The covariances of the error in the state estimation and the error associated with the observational data are used as weights, to indicate whether the state forecast should be closer to the a priori estimate (a previous estimate of the state,
coming from a former process of assimilation-forecast, but without processing the new observations) or to the observations.

In the context of DA problems, it is common that models depend on poorly known parameters. Because of the typically high sensitivity of the model with respect to the parameters, even with perfect initial data, the forecast operator can fail to accurately predict the future state of the system if the parameters are not properly estimated. In many models, parameters do not represent any specific physical property of the system, but account for characteristics of the underlying physical dynamics which are too complex to be included explicitly into the mathematical model. As an example, imagine a model describing a tsunami striking coastal communities, in which a viscosity parameter is used to represent the interaction between the water and fallen buildings. This implies that in general parameters cannot be measured or observed, which makes their calibration a much more complicated task.

Classical model calibration techniques consist in finding an optimal set of parameters which minimizes the mismatch between the system output and observational data. A serious limitation of such methods is that during the process only parameter errors are taken into account, while ignoring other sources of uncertainty. It means that a wrong estimation of the system state or simply very noisy observations can lead to completely wrong estimations of the parameter. This has motivated the developments of DA techniques to estimate the unknown parameters, which take into account all sources of error. This is the case of the augmented state approach (see [35]), which consists in considering a modified DA problem, where the parameters are added to the state vector, and the model forecast operator incorporates the equations reflecting the evolution of the parameters in time. Then, one could try to solve the augmented problem using a standard DA method. In the new problem the augmented state error covariance matrix $B$ can be written as the block matrix

$$ B = \begin{pmatrix} B_{xx} & B_{xp} \\ B_{xp}^\top & B_{pp} \end{pmatrix} \quad (1.1) $$

where the state-state error covariance $B_{xx}$ represents the covariances between errors in the estimation of state variables, the state-parameter error covariance $B_{xp}$ covariances between errors in the state and parameters estimates, and the parameter-parameter error covariance $B_{pp}$ the covariances between parameter errors.

Though in the last few years an important number of publications have reported the use of the augmented state approach for the joint state-parameter estimation (for example in [7], [62] and [139]) most of them present only empirical results for very specific models.
In this thesis we show that even for simple models, the joint state-parameter estimation via state augmentation can fail when a standard DA method is applied, mainly due to a wrong estimation of $B_{xp}$ and $B_{pp}$. Because the model parameters are never observed, the only way to estimate them is by transporting the information coming from the observed states to the parameters through the state-parameter error covariance. We also show that by modifying the way these covariances are updated the assimilation results can be drastically improved.

The update of the state-parameter error covariance is therefore a crucial aspect of the joint state-parameter DA problem. In her doctoral thesis [101] Smith investigates the performance of a Three-Dimensional Variational (3D-Var) scheme (see e.g. [125], [11] and [10]) applied to a simple model with only one parameter. This method is based on processing the observations sequentially, producing updates of the state vector. Nevertheless, the covariance matrix $B$ is kept fixed along the assimilation windows. In the aforementioned thesis it is shown that considering a time-invariant $B_{xp}$ is insufficient to correctly estimate the parameter. Instead, [101] proposes a hybrid approach, where the sequential 3D-Var ideas are combined with a flow-dependent state-parameter error covariance.

In this study we show that the results obtained by Smith can be improved if the state-state error covariance is also updated along the assimilation window. We propose a novel method which updates $B_{xx}$, as well as $B_{xp}$, at a low computational cost if the number of parameters is small. The flow-dependent $B_{xx}$ not only contributes to improve the state estimation but the parameter estimation as well. The suitability of our strategy is tested using simple nonlinear models.

The following sections present the main aims of this investigation (Section 1.1) and a summary of the main results of this study (Section 1.2). Further, Section 1.3 presents the general structure of the rest of the thesis.

### 1.1 Goals of this investigation

A main goal of this investigation is to study the joint state-parameter DA problem via state augmentation. For this purpose we focus on the classical Extended Kalman Filter (EKF) update formulas applied to the augmented state problem. Using a simple single-parameter model we intend to assess the suitability of the EKF error covariance update for the joint parameter-state data assimilation, considering different configurations of observations and initial parameter estimates.
Chapter 1. Introduction

Another important objective of this thesis is to study the hybrid approach proposed by Smith [101] and to develop a mathematical foundation for the proposed state-parameter error covariance update. Furthermore, our goal is to present a new hybrid method based on Smith’s update of the state-parameter error covariance, by incorporating a flow-dependent structure of the state-state error covariance. We intend to assess the suitability of the proposed method through numerical experiments, using different non-linear models.

It is also important to find convergence conditions for the sequential 3D-Var and the hybrid approaches presented in this investigation, applied to the linear joint parameter-state data assimilation problem. Here we refer to the convergence in the sense that the state and parameter estimates tend to be unbiased when more and more observations are assimilated. To our knowledge this issue has not been addressed in any previous investigation.

1.2 Contribution of the thesis

Among the most relevant results of this investigation we show that when using the augmented state approach for the joint state-parameter assimilation under perfect model conditions, the EKF is sometimes overconfident in the estimation of the parameter. This causes the failure of the assimilation process. Nevertheless, the results can be drastically improved by a simple modification of the state-parameter or parameter-parameter error covariance update. Besides, it is possible to obtain better results if we consider that using an incorrect value of the parameter causes an error in the forecast and the corresponding model error covariance is used to update the state error covariance during the forecast update.

It is shown in Chapter 5 that the state-parameter error covariance update of Smith’s method is closely related to the one obtained using an empirical covariance matrix approach (i.e., when a finite ensemble of states is used to approximate the covariance). Furthermore, we propose a new sequential hybrid approach for the joint state-parameter data assimilation via state augmentation. The method is based on a low computational cost update of the state-state and state-parameter error covariances. Our proposed method combines a stationary state-state error covariance with a flow-dependent term. The flow-dependency of $B_{xx}$ reflects the influence of the parameter estimation uncertainty on the statistics of the state error. Besides, $B_{xp}$ is updated similarly to its update in Smith’s method.
Chapter 1. *Introduction*

Our method proves to be effective in a range of simple nonlinear dynamical systems. By considering a flow-dependent state-state error covariance the state errors can be described more precisely, contributing to an improvement of both parameter and state estimations. The extra update of the state-state error covariance does not alter the order of the computational cost with respect to Smith’s approach.

One of the most important contributions of this investigation is the formulation of several convergence conditions for the sequential 3D-Var and in particular for both hybrid approaches, for the linear joint parameter-state DA problem. By convergence we mean that the expected value of the error in both, parameter and state estimation, converges to zero. We do not refer to the reduction of the variance of the errors.

1.3 Structure of the thesis

Chapter 2 and 3 review the most important results concerning data assimilation and parameter estimation, and discuss in detail approaches which are used in the framework of this investigation.

In chapter 4 we analyse in detail the EKF formulas applied to the augmented state problem. Special attention is given to the update of the state-parameter and parameter-parameter error covariance. We use a linear advection model to illustrate the problems that might arise when the filter is applied to the joint parameter-state assimilation and discuss possible solutions. The advantage of working with this simple model is that the parameter represents a physical characteristic of the system, therefore, the assimilation results are easier to interpret. Moreover, by comparing the forecast and the true state, it is straightforward to identify whether the value of the parameter should be incremented or decreased.

Chapter 5 introduces the idea of hybrid strategies for the general data assimilation problem. We focus on the hybrid approach proposed by Smith [99], especially designed for the state and parameter data assimilation via augmenting the state. Based on this method we propose a new hybrid strategy which, unlike the former approach, considers a flow-dependent state-state error covariance.

Convergence results for the sequential 3D-Var and both hybrid approaches, applied to the linear state and parameter data assimilation problems are obtained in Chapter 6. Experiments are conducted using a simple model to corroborate the use of the necessary and sufficient converge conditions obtained in practical situations.
To assess the performance of our hybrid approach, numerical experiments for different benchmark problems are presented in Chapter 7.

Finally, we summarize the main results of this thesis in Chapter 8, and discuss possible new lines of investigation.
Chapter 2

Data Assimilation Overview

This chapter presents the general formulation of the data assimilation (DA) problem (Section 2.1), which can be solved using methods of sequential or variational type. Section 2.2 gives a short overview of the main sequential DA methods, emphasising the schemes which are relevant to our investigation. Variational DA methods are discussed in Section 2.3. To conclude, we summarize the most important aspects of the chapter in Section 2.4.

2.1 Problem formulation

Let us consider a discrete dynamical system whose evolution in time is given by

\[ x_{k+1} = f_k(x_k) + \xi_k, \quad k = 0, 1, \ldots \]  

The state of the system at time \( t_k \) is represented by \( x_k \in \mathbb{R}^n \). The \( n \) state variables can be, for example, a set of different physical parameters at specific space points. The function \( f_k : \mathbb{R}^n \rightarrow \mathbb{R}^n \), also known as forecast model, describes the evolution of the system from time \( t_k \) to \( t_{k+1} \) and is generally nonlinear. In practice \( x_k \) is unknown and the model forecast can only be applied to an approximation of the true state, which we denote by \( x_k \). Moreover, models are just an approximation of the reality and therefore, have associated a forecast error that we represent by the stochastic variable \( \xi_k \in \mathbb{R}^n \), which we assume is unbiased and has covariance \( Q_k \). The estimate of the true state \( x_k \), given the available information up to time \( t_{k-1} \), is known as the background or forecast state and denoted by \( \hat{x}_k \) or \( x_k^f \) respectively. We assume that the error of this estimation, which we denote by \( \eta_k \), is also unbiased, and the associated background error covariance \( B_k \in \mathbb{R}^{n\times n} \) is known (or approximated).
Additionally, (noisy) observations of the system \( y \in \mathbb{R}^{m_k} \) are available at certain time points. In general it is not possible to directly observe every state variable. For example, in a weather forecast model, where the state consists of the different climatological variables at a three-dimensional grid around the earth, that would require to place satellites, buoys, etc., every 5 or 10 kilometres. The observations are related to the system state through the following equation:

\[
y_k = H_k(x^k_t) + \epsilon_k, \ k = 0, 1, \ldots, N, \tag{2.2}
\]

The observation operator \( H_k : \mathbb{R}^n \rightarrow \mathbb{R}^{m_k} \) is a mapping from the model space to the observation space. Given a state, it predicts which values are expected to be observed. The stochastic errors \( \epsilon_k \in \mathbb{R}^{m_k} \) are assumed to be unbiased, uncorrelated in time and with known covariance matrices \( R_k \in \mathbb{R}^{m_k \times m_k} \). They account not only for the noise in the observations but also for errors introduced by the observation operator (e.g., interpolation errors). In case that observations are not available at time \( t_k \) we can simply assume that \( m_k = 1, y_k = 0 \) and \( H_k = 0_{1 \times n} \), which do not produce any update of the state vector.

The goal of DA is to combine model predictions (equation (2.1)) and observations (related to the state of the system by equation (2.2)), taking into account the statistics of the errors, to determine the most probable state of the system.

DA methods can be divided into two classes: sequential and variational schemes. In the following we present in detail different methods for each approach, which are relevant to our investigation.

### 2.2 Sequential DA

Sequential methods process observations available at a certain time to correct a first guess of the state, then use the forecast model to transport the state estimation (new guess) in time until the next observations are available. Figure 2.1 represents the prototype of such a scheme.

The main advantage of these methods is that new information can be incorporated into the system as soon as it is available, which is desirable in real time applications.
2.2.1 Optimal Interpolation and 3D-Var

The 3-Dimensional Variational (3D-Var) and the Optimal Interpolation (OI) DA methods are sequential methods, based on obtaining at time $t_k$, where observations are available, a state $x^a_k$ that minimizes the cost function

$$J_k(x_k) = \left( x_k - x^b_k \right)^\top B_k^{-1} \left( x_k - x^b_k \right) + \left( y_k - H_k(x_k) \right)^\top R_k^{-1} \left( y_k - H_k(x_k) \right). \quad (2.3)$$

Here the index “$a$” stands for “analysis”. The way the minimization is performed differentiates both approaches as we explain later on. The weight matrices $B_k \in \mathbb{R}^{n \times n}$ and $R_k \in \mathbb{R}^{m_k \times m_k}$ are symmetric and positive definite. They represent the background state error and the observation error covariances respectively. It is assumed that the observation errors and the background errors are uncorrelated. Here $J_k : \mathbb{R}^n \to \mathbb{R}$ can be seen as a compromise between the background state and the observations. The quadratic form $J_b(x)$ penalizes those states faraway from the predicted one, while $J_o(x)$ does it with states that misfit the observations. The covariance matrices play an important role in the objective function. If the observation errors are much smaller than the background errors, the state resulting from the minimization process will fit the observations almost accurately. Conversely, if the background errors are much smaller, then the assimilated state will resemble the background state.

If the observation operator $H_k$ is linear it can be represented by a matrix $H_k \in \mathbb{R}^{m_k \times n}$. Then the solution of the optimization problem
\[ x_k^a = \arg \min_{x_k} J_k(x_k) \] (2.4)

can be obtained directly by the update

\[ x_k^a = x_k^b + K_k(y_k - H_k x_k^b) \] (2.5)

where \( K_k \in \mathbb{R}^{n \times m_k} \) is the so-called Kalman gain matrix

\[ K_k = B_k H_k^\top (H_k B_k H_k^\top + R_k)^{-1}. \] (2.6)

The vector \( d_k = y_k - H_k x_k^b \) is called the innovation vector. The optimal state is then the background state plus a weighted innovation. This solution is the Best Linear Unbiased Estimate (BLUE) following Bayes’ theorem (see [31]) and it is optimal in the weighted least square sense. If the observation operator is nonlinear, still a similar state update can be performed by substituting the linearization of the observation operator

\[ H_k = \frac{\partial H_k}{\partial x}(x_k^b) \] (2.7)

into (2.6) and using the nonlinear observation operator to compute the innovation vector so that

\[ d_k = y_k - H_k(x_k^b), \] (2.8)
\[ x_k^a = x_k^b + K_k d_k, \] (2.9)

though the optimality of the resulting solution is no longer guaranteed, as it depends on the non-linearity of \( H_k \). This method, consisting in computing (2.6) explicitly and substituting it into (2.9) is known as Optimal Interpolation (OI) (see [111], Section 4.2, [80], [89] and [12]), also called Statistical Interpolation, Objective Analysis or Gauss-Markov smoothing. The method was operational in numerous DA systems during the 70’s and 80’s, until the variational approaches became popular. The reason why this approach is impractical in high dimensional problems is because of the difficulties to store and apply \( B_k \), due to its size. Alternatively 3D-Var, currently operational in various DA systems (for instance, at ECMWF [92], NCEP [33], the Met. Office [73], among others), solves (2.4) with the help of a numerical optimization solver. In this thesis we use the box-constrained limited memory BFGS (L-BFGS-B) described in [26], as the test models used along this investigation have simple box constraints on the parameters.
Chapter 2. Data assimilation overview

The implementation is a MATLAB wrapper for the Fortran subroutines in [140] and considers the correction suggested in [84]. Though the models which we currently use are not high dimensional our intention is to experiment with more complex models in the future. Therefore a routine designed to solve high dimensional problems such as L-BFGS-B is chosen.

Observations are very often collected over a time window. This is for example the case in numerical weather prediction (NWP). In these situations both, 3D-Var and OI, assume that the state error does not change significantly along the time window and therefore they do not modify the background error covariance to describe the effect of the model or of new observations in the error statistics (like it occurs for instance in the Kalman Filter method). Instead, (2.4) is solved at each time step using a fixed background error covariance $B_k = B$. The new assimilated state $x^a_k$ is then transported in time using the forecast model until new observations are available. A simple pseudo-code describing the sequential 3D-Var/OI is presented in Algorithm 1.

```
Initialization $B$, $x^b$

Step 1. Assimilation Step
    IF observations not available THEN
        $x^a = x^b$
        go to Step 2
    END
    IF method==3D-Var THEN
        Obtain $x^a$ by solving (2.4) using an optimization solver
    ELSEIF method==OI THEN
        Compute $K_k$ and $x^a$ according to (2.6) and (2.5), respectively
    END

Step 2. Forecast Step
    Update $x^b$ using the forecast model
    $t=t+1$
    go to Step 1.
```

Algorithm 1: 3D-Var/OI algorithm

The set-up of a static background error covariance is itself a very complicated task and at the same time an extremely important aspect of the DA scheme. Besides accurately representing the error statistics, there are other considerations to be taken into account, for instance that the resulting matrix is well conditioned. There have been numerous studies focusing on investigating the proper generation of $B$. In [24], the generation of an ensemble-based background error covariance in a quasi-operational NWP system is discussed. The formulation and structure of the background error covariance used in the European Center Medium-Range Weather Forecast (ECMWF) are presented in [34]. Other related works are found in [13], [48], [8] and [9].
Despite its simplicity, efficiency and robustness, the fact that the background error covariance is kept invariant along the time window where observations are available (to which we will refer also as the assimilation window) can be considered as the main weakness of the method. The assumption that the errors are statistically stationary is simply not realistic in a large variety of models. In the next sections we discuss the classical Kalman Filter and Extended Kalman Filter, which consider a flow-dependent background error covariance.

2.2.2 Kalman Filter

The Kalman Filter (KF) is a linear filtering method developed by Kalman in 1960 (see [65]). It is applied to linear discrete dynamical systems and uses observations sequentially to produce a state estimate. This estimate, under certain statistical assumptions, including Gaussian distribution of the errors, is optimal, in the sense that it minimizes the error variance. The sequence of state estimates describes a Markov process, i.e.:

\[ P(x_k | x_0, x_1, \ldots, x_{k-1}) = P(x_k | x_{k-1}). \]  

(2.10)

This means basically that the history of estimations \( x_0, x_1, \ldots, x_{k-2} \) is redundant once the state \( x_{k-1} \) is computed. The information from previous assimilations is “remembered” by the filter through the background error covariance. Unlike in 3D-Var and OI, the KF updates the error covariance when the model is used to forecast the next state, which we will call a forecast step, and when the state is modified because a new observation has been assimilated, known as an assimilation step. In the following we will introduce the basic KF formulas. A deduction of such formulas and other related materials can be found in [106], [67] and [131].

Let us consider a linear data assimilation problem. Equations (2.1) and (2.2) can then be rewritten as

\[ x_{k+1} = M_k x_k + \xi_k, \]  

(2.11)

\[ y_k = H_k x_k + \epsilon_k, \quad k = 0, 1, \ldots, N, \]  

(2.12)

where the model and the observation operators are represented by the matrices \( M_k \in \mathbb{R}^{n \times n} \) and \( H_k \in \mathbb{R}^{m_k \times n} \), respectively. We assume that \( \epsilon_k \sim \mathcal{N}(0, R_k) \), where \( R_k \in \mathbb{R}^{m_k \times m_k} \) is the observation error covariance and \( \xi_k \sim \mathcal{N}(0, Q_k) \), with \( Q_k \in \mathbb{R}^{m_k \times m_k} \) being the model error covariance. Here \( \mathcal{N}(\mu, C) \) represents the Gaussian distribution with mean \( \mu \) and covariance \( C \). Other statistical assumptions are that the state errors, model errors and observation errors are mutually uncorrelated and model and observation errors are
uncorrelated in time, i.e.,

\[
\begin{align*}
\mathbb{E}[\xi_i \epsilon_j^\top] &= 0, \quad \mathbb{E}[\xi_i \eta_j^\top] = 0, \quad \mathbb{E}[\epsilon_i \eta_j^\top] = 0, \quad \forall i, j \\
\mathbb{E}[\epsilon_i^\top \epsilon_j^\top] &= 0, \quad \mathbb{E}[\xi_i^\top \xi_j^\top] = 0, \quad \forall i \neq j,
\end{align*}
\]

(2.13) (2.14)

where \(\mathbb{E}[\cdot]\) represents the expected value.

In Figure 2.2 we represent the KF method. In general, every time an observation is available the filter produces a state update, which coincides with the OI update. The difference is that the background error covariance is also updated, to reflect how the availability of new information about the system modifies the statistics of the error. The state is then forecast, using the model, until a new observation is available. At the same time the covariance is updated, to reflect how the system dynamics modify the state error.

Unlike 3D-Var and OI, each assimilated state \(x_k\) produced by the KF is the BLUE considering all observations until time \(t_k\). If observations up to time \(t_{k+p}\) are considered, then it is still possible to obtain an estimate at time \(t_k\) taking into account all the observations (including those that were processed after the KF estimate were computed) by a backwards processing of the estimates produced by the KF, from \(x_{k+p}\) to \(x_k\). This process is known as the Kalman smoother [3], [86], [71].

2.2.3 Extended Kalman Filter

If either the model or the observation operator is nonlinear, then the KF formulas can still be applied by linearising around the current state. This approach is known
as the Extended Kalman Filter (EKF). Figure 2.3 shows the idea of the EKF iteration. Nevertheless, in this case the optimality property of the solution is lost and the state and background error covariance produced can significantly differ from the optimal solution and the error statistics, respectively.

![EKF algorithm diagram](image)

**Figure 2.3: EKF algorithm.**

Despite providing the optimal solution in the weighted least square sense, the computational complexity of the KF limits the method to be used only in small to moderate dimensional models. This is also the case of the EKF. The reason is that the update of the background error covariance requires a number of model runs proportional to \( n \), which is prohibitive in applications like NWP. Notice that to compute \( M_k B_k^a M_k^\top \) the model (or tangent model \( M_k \) if the forecast operator is nonlinear) must be applied to the \( n \) columns of \( B_k^a \) and then the resulting \( n \) columns are processed by the adjoint model \( M^\top \). In such cases a single run of the forecast model (or the associated tangent and adjoint models) demands a huge computational effort. As an additional difficulty, the adjoint and linear tangent models must be developed and maintained. The importance of the KF is that it establishes the theoretical basis for the development of other approximate strategies.

### 2.2.4 Ensemble Kalman Filter

Ensemble-based methods in DA use a set of independent system simulations to approximate the error space of the state estimation. The Ensemble Kalman Filter (EnKF) is nowadays the most investigated and powerful ensemble method in the context of DA. It is a sequential method using Monte Carlo integrations [49], [66], first introduced by
Chapter 2. Data assimilation overview

Evensen in 1994 [41] (see also [60], [61], [122], [25], [44]) which, unlike other ensemble-based methods (e.g. particle filters [82], [123], [121]) assumes that all the errors are Gaussian. It is an alternative to the computationally expensive KF. The idea is to run the filter for a state ensemble \( x_{k,l}^{(l)} \), \( l = 1, 2, \ldots, N \), instead of updating a single state. Then, the state and the background error covariance, necessary for the assimilation step, can be approximated by the sample mean and covariance given by

\[
\hat{x}_k = \frac{1}{N} \sum_{l=1}^{N} x_{k,l}^{(l)}, \tag{2.15}
\]

\[
P_k = \frac{1}{N-1} \sum_{l=1}^{N} (x_{k,l}^{(l)} - \hat{x}_k)(x_{k,l}^{(l)} - \hat{x}_k)^\top, \tag{2.16}
\]

respectively. The reason why researchers have paid so much attention to these methods in the last few years is that many of the problems associated to the traditional EKF are avoided [25]. For instance, higher order statistical moments are not neglected in the error covariance evolution equation (no linearisation required). Besides, there is no need of generating and maintaining tangent and linear codes. The computational cost is also lower, as the full update of the background error covariance is avoided.

Here we explain the idea of the method as proposed by Evensen [41]. We assume that the observation operator is linear (for the nonlinear case [43] can be consulted).

Forecast:

\[
x_{k+1}^{f(l)} = f_k(x_k^{a(l)}) + \xi_k^{(l)}. \tag{2.17}
\]

Assimilation:

\[
x_k^{a(l)} = x_{k+1}^{f(l)} + K_k(y_k - H_k(x_k^{f(l)})) \tag{2.18}
\]

\[
K_k = P_k^f H_k^\top (H_k P_k^f H_k^\top + R_k)^{-1} \tag{2.19}
\]

Later it was demonstrated (see [25] and [60]) that it is necessary to perturb observations to keep the statistical consistency of the filter. Unfortunately, this introduces sampling errors. To avoid this, new methods have been developed which compute the ensemble analysis deterministically, known as square root filters. Among the many different filters we mention the Local Ensemble Kalman Filter (LEKF) [90], the Local Ensemble Transform Kalman Filter (LETKF) [107], [91] the Ensemble Square Root Filter (EnSRF) [133], the Error Subspace Transform Kalman Filter (ESTKF) [87], [88], the Singular Evolutive Interpolated Kalman (SEIK) [118], the second-order-exact EnKF, the singular second-order-exact EnKF [94], the Ensemble Transform Kalman Filter (ETKF) [17],[77], the
Ensemble Adjustment Kalman Filter (EAKF) [4] and the Singular Evolutive Extended Kalman (SEEK) filter [118], [21].

The use of a finite (for practical purposes, small) ensemble size introduces sampling errors. One of the problems is that spurious correlations appear between space-distant variables or between variables which are known to be uncorrelated. This causes a reduction of the ensemble variance and therefore an underestimation of the true one, which can lead to filter divergence [45]. To avoid this, it is necessary to apply strategies like covariance inflation [5], [6] or covariance and observation localization [58], [28], [52].

Currently there are still several unresolved issues related to the EnKF, like the proper generation of an initial ensemble, the selection of the inflation parameter and localization radius, etc. Though they are an active field of research and probably represent the future of DA, still most of the operational NWP systems are based on 3D-Var and 4D-Var approaches. This type of methods are included in this DA overview for the sake of completeness, though their use is out of the scope of this investigation.

2.3 Four dimensional variational DA

Unlike in the sequential methods, in which observations are processed as soon as they are available, four dimensional variational (4D-Var) [32], [116], [117] accumulates them along a time window in order to find the state that minimizes the weighted least square distance between the model trajectory and the observations. Therefore, all the observations in the time window are processed at the same time (see Figure 2.4).

![Figure 2.4: 4D-Var scheme.](image-url)
Chapter 2. Data assimilation overview

The function to be minimized is represented by (2.20)-(2.21) and is very similar to the 3D-Var case. The difference is that $J_o(x)$ includes all the observational data collected along the analysed time window.

$$J(x_0) = (x_0 - x_0^b)^\top B_0^{-1} (x_0 - x_0^b) + \sum_{t=0}^{N} (y_k - H_k(x_k))^\top R_k^{-1} (y_k - H_k(x_k)),$$  \hspace{1cm} (2.20)  

where $x_k$ can be seen as a nonlinear function of $x_0$, defined by the recursion:

$$x_{k+1} = f_k(x_k), \ k = 0, 1, \ldots, N - 1$$  \hspace{1cm} (2.21)  

An obvious advantage of the method is that the background error covariance in the first term of $J$ is not updated, which is the most expensive step of the KF. Moreover, under the same linearity and statistical assumptions, both methods produce the same optimal state in the following sense. The state obtained after the assimilation of the $N$-th observation by the KF is equivalent to the state obtained by transporting the solution of the variational problem forward in time using the forecast model up to time $t_N$. The optimal state trajectory is also equivalent to the sequence of states obtained by the Kalman smoother. The method has been implemented in many operational systems ([38], [39], [18]).

To minimize the cost function most of the variational methods need to compute the gradient $\nabla J$. This can be done efficiently by means of the adjoint model (see [50], [78], [126]). This code plus the Tangent Linear (TL) code, which must be obtained by modifying the routine codes computing the state forecast, are nevertheless very expensive to maintain in high dimensional models. This maintenance is a permanent task due to the frequent updates of the physical and mathematical models. Besides, the need to have a parallel implementation makes the task even more complicated (see [136]). Another disadvantage of the method is the high computational cost of the optimization scheme. This is why in most cases sub-optimal solutions are obtained after 10-100 iterations of the minimization routine. A third drawback is the possible breakdown of the tangent linear approximation due to strong instabilities of the nonlinear model, which limits the linearisation to short time intervals [124]. On the other hand, the fact that all observations have to be reprocessed once a new set of observational data is available (notice that the old and new observations should be processed at the same time to find the optimum of (2.20)) hinders the real-time response capabilities of the assimilation scheme. Furthermore, having no estimation of the assimilation error statistics is another disadvantage of the method. To overcome these difficulties, hybrid strategies have been developed, combining the 4D-Var ideas with other sequential methods (see Chapter 5).
2.4 Summary and discussion

In this chapter we presented an overview of the DA problem and important DA techniques. Special attention was given to the 3D-Var and OI, as well as to the KF and EKF, which are directly related to our investigation.

3D-Var is a robust, easy to implement and efficient sequential method, already operational in many DA systems. The fact that it uses a static background error covariance can be considered to be a limitation, as it does not capture the dynamics of the system. Conversely, in the KF, by updating the background error covariance, the assimilated information from previous observations is transported forward by the error covariance. As a result the state update does not only depend on the current observation but it is also consistent with the history of previous assimilations. Despite the optimality properties of the filter for linear forecast and observation operators under certain statistical assumptions, the cost of the covariance update is prohibitive in large dimensional models.

In the context of NWP short-time-scale errors, also called “errors of the day”, grow because of instabilities in the flow. It is known that such strong instabilities can be represented in a low dimensional subspace [93]. This suggests that it could be possible to approximate short-time-scale errors without performing the complete KF covariance update. We take this into account in Chapter 5, where ideas of 3D-Var and the EKF are combined in our sequential novel method.
Chapter 3

Parameter Estimation

A successful data assimilation (DA) depends on unbiased state prediction. Inaccurate parameter estimation is one of the main sources of model errors. This is why the correct identification of the model parameters, given a set of observations, is crucial to the success of any assimilation scheme.

Very often, parameters do not represent any physical, directly measurable characteristic of the system, but processes, which are too complex to be modelled, or not understood at all. This implies that in most of the cases it is not possible to directly observe them, and the parameter tuning must rely on evaluating the mismatch between the system response and observational data.

In Section 3.1 important theoretical aspects related to the parameter estimation problem are discussed. Section 3.2 presents an overview of model calibration techniques. In Section 3.3 we present the joint state-parameter estimation and the dual parameter-state DA. In Section 3.4 we introduce the augmented-state approach, in which the parameters are considered as additional state variables. In Section 3.5 a comparison is established between the two joint parameter-state techniques presented in this chapter.

3.1 Parameter identifiability and sensitivity

The parameter estimation problem is typically ill-posed. This means that either the uniqueness or the stability of the solution is violated. The first condition is closely related to the concept of identifiability which will be discussed in this section. The second one means that errors in the available data can produce errors in the parameter estimation of several orders of magnitude higher.
A parameter vector is identifiable if, provided a specific input-output set of data, it can be uniquely determined, supposing that the data is error-free. The concept is related to the observability, which addresses the question whether the given input-output information is enough to recover the model trajectories. For more about the relation between identifiability and observability we refer to [95]. A very complete and comprehensive overview of identifiability can be found in [63].

An important aspect of the parameter estimation problem is the parameter sensitivity, which accounts for the magnitude of the model output variation as a response of a variation in the parameter. If a parameter has a very low sensitivity, then by modifying its value, little to no variation of the model output is produced. If the magnitude of this variation is smaller than the magnitude of the errors of the observational data, then it is clear that the parameter cannot be identified. Conversely, small variations of highly sensitive parameters produce a considerable variation of the model output. Obviously, the more sensitive the parameter is, the more important it is to estimate it accurately, as the estimation error has a huge impact in the system output and therefore in the forecast ability of the model. Moreover, parameter sensitivities can also be used to evaluate the parameter identifiability (see e.g., [36] and [23]). In the following sections we present different techniques used to solve the parameter estimation problem.

### 3.2 Model calibration

Model calibration is an optimization problem associated to finding an optimal set of parameters that minimizes the mismatch between a set of observations and a model. The optimization problem can be multi-objective or several mathematical criteria can be combined into a single objective function. The optimization process can be carried out either automatically or manually.

#### a) Manual calibration

Manual calibration is normally a trial-and-error process. Though supported by objective measurements, the decision over the “best” set of parameters is subjective, normally based on visual comparison of model outputs and the available data, and it depends strongly on the experience and training of the person conducting the calibration. With the increasing complexity of the models and the increasing number of model parameters, the manual calibration is a time-consuming, prone-to-error task. Therefore it is more common to be combined or supported by automatic calibration, to benefit from the strengths of both approaches [20].
b) **Automatic calibration**

Automatic calibration is part of the operation research theory. Unlike in manual calibration, the objective function is minimized using an optimization algorithm. There are two key points involved in this approach: the prescription of a suitable objective function (e.g., single or multiple criteria, minimization of the least square distances, etc.) and the selection of the optimization algorithm (e.g., global vs local methods, heuristic vs direct methods, line search or trust region methods, gradient-based or Quasi-newton methods, among others).

There are many examples in the literature for the application of automatic model calibration. In [14] a methodology for calibration of distributed models based on generalized likelihood measures is presented. Cheng [29] combines a fuzzy optimal model with a genetic algorithm to solve multi-objective rainfall-runoff model calibration. The interested reader is referred to a useful review of different calibration techniques in the context of hydrologic modelling [72].

One of the main disadvantages of model calibration is that observations have to be collected to be processed at the same time. This implies that whenever new data is available, it is necessary to reprocess the previous observations in order to use the new output information. This constitutes a limitation for being used in applications requiring real-time response. Another drawback is that it only accounts for the tuning of the model parameters, ignoring the errors of the system input and output. A significant data noise or a wrong estimation of the system state can lead to parameter estimates that cause a model bias. In the last years other strategies have been developed, which account for all sources of error, i.e., input, output and parameters, at the same time. In the following sections we introduce two of the most popular ones.

### 3.3 Dual state-parameter DA

While batch calibration accounts only for the parameter estimation, ignoring other sources of inaccuracies, DA is a technique designed mainly for estimating the system state using the model and observations. In principle, it assumes that the model parameter values are correctly set. Nevertheless, it can also be extended to the joint parameter-state estimation problem. One possibility is the dual state-parameter DA. In this approach state and parameters are updated in parallel using two different filtering processes and sharing the resulting information. For a better understanding, in Figure 3.1 a scheme of the method is represented, using as an example the KF. When an observation is available at time step $k$, only the parameter is corrected. This new estimate
is used to rerun the forecast model from time $k-1$ to $k$, obtaining a new state update. Then the observation is reused, this time to correct the state estimation.

![Diagram of Dual state-parameter KF]

**Figure 3.1:** Dual state-parameter KF. Matrices $K_p$ and $K_x$ are gain matrices obtained using the cross-covariances between state and parameter errors, and the covariances between state errors, respectively (for a formal definition of these matrices see (4.14)).

This strategy has been widely used in different kinds of models. In [113] and [114] an EnSRF is used for the simultaneous estimation of microphysical parameters and the atmospheric state using simulated radar data. A dual EnKF is successfully applied in [1] to a NWP model of intermediate complexity, the MM5 [51]. In a previous paper [2] the author had shown the applicability of a dual EnKf in a forced dissipative flow, to estimate up to 6 model parameters. Zhang shows in [138] that a parameter estimation using a dual EnKF to simulate urban land-use changes in a 13-year period (1993–2005) in Dongguan City yields better results than the, in this context, more often used logistic regression and decision-tree strategies. The applicability of the dual EnKF approach for ensemble streamflow forecasting is demonstrated in [83] using a conceptual rainfall-runoff model. Wen [132] improves the EnKF performance in the context of reservoir models by using a dual approach, where the flow equations are solved again from the previous assimilating step to the current step using the updated current permeability models. This ensures the consistency of the updated static and dynamic parameters with the flow equations. The main drawback is that it lacks of mathematical basis, not being consistent with the traditional Bayesian framework.
3.4 Augmented state DA

Another popular joint state-parameter DA approach is the augmented state. The idea, first suggested by Derber in [35], is to include state and parameters together in an augmented state vector, and to incorporate the parameters’ evolution formulas into the state forecast. The resulting DA problem is then solved using a standard DA scheme. Because we use this strategy in the context of this investigation, in the following section a formal derivation of the approach is presented.

3.4.1 Formulation of the augmented state approach

Let us consider the DA problem associated with the following dynamical system:

\[ x_{k+1} = f_k(x_k, p), \]
\[ y_k = H_k(x_k) + \epsilon_k, \quad k = 0, 1, \ldots, N, \]

with

\[ y_k \in \mathbb{R}^{m_k} \text{ (observation vector)}, \]
\[ x_k \in \mathbb{R}^{n} \text{ (state vector)}, \]
\[ \epsilon_k \in \mathbb{R}^{m_k} \text{ (observation error)} \]
\[ p \in \mathbb{R}^{\ell} \text{ (parameter vector)}, \]
\[ f_k : \mathbb{R}^{n+\ell} \to \mathbb{R}^{n} \text{ (nonlinear model)}, \]
\[ H_k : \mathbb{R}^{n} \to \mathbb{R}^{m_k} \text{ (nonlinear observation operator)}. \]

Besides, we assume that \( R_k \in \mathbb{R}^{m_k \times m_k} \) is the covariance of the observation error \( \epsilon_k \) and that these errors are unbiased. In addition to the state variables we have a set of \( \ell \) poorly known model parameters \( p \) to be estimated. Here we assume that the parameters are constants.

The idea of the augmented state approach is to consider a new state vector that contains both, parameter and state variables, modifying the forecast model to fit the new dimensions of the problem by incorporating the parameter evolution formulas. We denote the new state variable as \( w \in \mathbb{R}^{n+\ell} \):

\[ w_k = \begin{pmatrix} x_k \\ p_k \end{pmatrix}, \]

If the parameters are time-invariant, then, without any observation to be assimilated, the best estimate of the parameter at time \( t_{k+1} \) is exactly the previous estimate at time
So the new forecast function \( \tilde{f}_k : \mathbb{R}^{n+\ell} \rightarrow \mathbb{R}^{n+\ell} \) can be defined as

\[
\tilde{f}_k(w) = \left( f_k(w(1:n), w(n+1:n+\ell)), w(n+1:n+\ell) \right).
\] (3.4)

Here we use the MATLAB notation: \( v(a:b) = (v_a, v_{a+1}, \ldots, v_{b-1}, v_b) \), assuming that \( a \leq b \).

If we additionally assume that the parameters are not observed, the observation operator \( \tilde{H}_k : \mathbb{R}^{n+\ell} \rightarrow \mathbb{R}^m \) for the augmented problem results in

\[
\tilde{H}_k(w) = H_k(w(1:n)).
\] (3.5)

The resulting dynamical system for the DA problem is

\[
w_{k+1} = \tilde{f}_k(w_k)
\]
\[
y_k = \tilde{H}_k(w_k) + \epsilon_k, \quad k = 0, 1, \ldots, N.
\] (3.7)

Besides the state error covariance, extra statistical information is necessary concerning the correlation between parameters and state variables, as well as correlations between parameters. The reason is that parameters are neither modified by the forward operator nor directly observed. It means that recovering the true value of the parameters is only possible if, from the update of the state variable during the assimilation step, an update of the parameter can be inferred by using the correlations between parameters and states. The new augmented state error covariance matrix \( B \) has the following block structure:

\[
B = \begin{pmatrix} B_{xx} & B_{xp} \\ B_{xp}^\top & B_{pp} \end{pmatrix},
\] (3.8)

where \( B_{xx} \in \mathbb{R}^{n \times n} \) represents the state-state error covariance matrix, \( B_{pp} \in \mathbb{R}^{\ell \times \ell} \) the parameter-parameter error covariance and \( B_{xp} \in \mathbb{R}^{n \times \ell} \) the state-parameter cross-covariance.

Matrix \( B \) must be positive definite, as it represents the augmented state error covariance. Matrices \( B_{xx} \) and \( B_{pp} \) are symmetric positive definite matrices, as they represent the state-state and parameter-parameter error covariance. In general if \( A \) and \( D \) are symmetric positive definite matrices, and using the typical notation \( A \succ 0 \) meaning that \( A \) is a positive definite, then

\[
\begin{pmatrix} A & C \\ C^\top & D \end{pmatrix} \succ 0 \Leftrightarrow A - CD^{-1}C^\top \succ 0 \Leftrightarrow D - C^\top A^{-1}C \succ 0
\] (3.9)
Chapter 3. Parameter estimation

(see Appendix A.5 in [19]). Applied to our case:

\[ B > 0 \iff B_{pp} - B_{xp}^T B_{xx}^{-1} B_{xp} > 0. \]  

This condition can be useful to define (and if necessary rescale) the initial cross-covariance \( B_{xp} \) while keeping the augmented state error covariance \( B \) positive definite. We will be referring to it in Section 4.3.7.

### 3.4.2 Applications of the augmented state approach

The application of the augmented state ideas is well documented in the literature. Annan [7] uses this strategy with an EnKF in an intermediate complexity coupled atmospheric-ocean general circulation model (AOGCM) with pseudo-observations. In this case the model’s climatology is successfully tuned via the simultaneous estimation of 12 parameters. In [62] two parameters are tuned via an augmented state EnKF by assimilating real-data observations (wind profiler) with excellent results. Zheng [139] uses the strategy in combination with an EnKF to estimate 3 parameters in a Monte Carlo atmospheric dispersion model. Experiments were conducted using simulated and real radiation monitoring data. In both cases the method was able to reconstruct the real scene of dispersion, as well as the uncertain parameters. In her Ph.D. thesis [101] Smith used the approach in the design of a sequential strategy, which was successfully applied to morphodynamic modelling, as well as to other simple test models.

### 3.5 Dual state-parameter vs augmented state DA

Several studies have been focussed on comparing dual and state augmentation strategies. In [59] a comparison of both approaches in a real-time groundwater flow problem is established using an EnKF method. The author concludes that no significant differences were found in the quality of the analyses. Wen [132] points out, in the context of reservoir models, that the joint state-parameter estimation via augmenting the state introduces inconsistency, especially for strong heterogeneous formations. Furthermore, he shows that an approach based on a rerun of the filter at the same time step where the observations are available outperforms noniterative EnKF. Another study supporting the benefits of iterative EnKF over noniterative filters is the one conducted by Gu and Oliver in [53]. These are, nevertheless, empirical results obtained for specific models. The truth is that the dual state-parameter estimation lacks a mathematical basis, as it steps away from the traditional Bayesian framework. The augmented state approach does not propose a different assimilation strategy, but applying the already developed
DA methods to a new problem. It implies, for instance, that the state and parameter estimation is optimal according to the Best Linear Unbiased Estimate (BLUE) criterion for linear models under certain statistical assumptions, if KF or 4D-Var are applied. This is the main reason why we considered in this work the augmented state, rather than the dual approach.

With respect to the computational complexity, the augmented state approach proposes to solve a DA problem where the dimension of the state is the sum of the dimensions of the parameter and state vectors. This can considerably increase the cost of the assimilation step, if the number of parameters is relatively large. Conversely, the dual state-parameter estimation solves two different assimilation problems, of dimension $n$ and $\ell$, and therefore keeps the complexity of the optimization problem as low as possible. Nevertheless, in this investigation we are particularly interested in models where the number of parameters is relatively small. This means that the use of the augmented state approach does not increase significantly the computational complexity of the assimilation scheme.

3.6 Summary and discussion

In this chapter important concepts related to the parameter estimation problem, like ill-posedness, identifiability and parameter sensitivity, were briefly presented. We also reviewed different parameter estimation techniques.

Model calibration only considers the identification of the parameter without taking into account other sources of error (input and output errors), which can be considered a serious limitation. In this respect joint state-parameter DA techniques estimate state and parameter at the same time, considering all sources of error. This also allows to assimilate observations as soon as they are available, without having to re-process previous observations.

In particular, the augmented state approach has been widely used in various types of models with different DA techniques. The strategy is based on transforming the state-parameter estimation problem into a state estimation DA one. The state vector is augmented with the parameters and the parameter evolution over time is incorporated into the forecast model of the augmented problem. Despite the increasing number of publications related to this topic, in most of the cases only the effectiveness of a particular implementation of this strategy applied to a specific DA problem is shown. Nevertheless, little is said about the problems that arise from considering parameters as state variables, and the importance of the update of the state-parameter error covariance has not been
sufficiently investigated. Time-invariant parameters are neither directly observed nor changed by the model. It implies that their update depends exclusively on the state-parameter error covariance. It is through this covariance that the information from the observed state variables is transported to the parameters. In the next chapter we show how, even applied to a very simple model, the augmented state approach using the EKF can fail, if no special modifications of the parameter-parameter or state-parameter error covariances are considered.
Chapter 4

EKF and the Augmented State Problem

In this chapter we analyse in detail how the EKF formulas are adapted to solve the augmented state DA problem associated to (3.6)-(3.7), already introduced in Section 3.4. In general, even if a model is linear, by considering that the parameters are unknown, these may appear in the resulting augmented state model in a nonlinear way. This is the case of the test model discussed in this chapter.

For simplicity we assume that all the observation vectors have the same dimension $m$. Moreover, some important aspects of the resulting scheme are discussed with the help of a simple test model. Our motivation is to show that when the parameters are treated as state variables the parameter statistical information can be wrongly estimated. This leads to a completely wrong approximation of the parameter and therefore the failure of the assimilation scheme.

In sections 4.1 and 4.2 we introduce the EKF forecast and assimilation step formulas, respectively, for the augmented problem. Results for various numerical experiments are presented in Section 4.3, where the linear advection model is introduced. This simple model is linear if the parameter is considered to be known an fixed. Nevertheless, if the parameter is unknown the resulting augmented state problem is nonlinear, as the parameter enters the models multiplying other state variables. We show that in some cases, applying the EKF to the augmented problem leads to divergence of the filter and propose simple variations which improve the assimilation. These results are consistent with those obtained experimenting with more complex models (see Appendix C.2). Section 4.4 discusses and summarizes the results obtained.
4.1 Forecast step

In order to apply the EKF formulas to the new augmented problem a linearisation of the forecast operator for propagating the statistical information forward in time is necessary. The resulting forecast step for the augmented state is

\[
x_{k+1}^f = f_k(x_k^a, p_k^a),
\]

\[
p_{k+1}^f = p_k^a,
\]

\[
B_{k+1}^f = J_k \left( \begin{bmatrix} (B_{xx}^a)_{k}^\top & (B_{xp}^a)_{k}^\top \\ (B_{xp}^a)_{k} & (B_{pp}^a)_{k} \end{bmatrix} \right) J_k^\top,
\]

where \( J_k \) is the Jacobian of \( \tilde{f} \) at time step \( k \)

\[
J_k = \left( \begin{array}{cc} \frac{\partial f_k}{\partial x} & \frac{\partial f_k}{\partial p} \\ 0 & I_{\ell \times \ell} \end{array} \right) \bigg|_{x_k^a, p_k^a}.
\]

For ease of notation we will denote

\[
M_k := \frac{\partial f_k}{\partial x}(x_k^a, p_k^a), \quad N_k := \frac{\partial f_k}{\partial p}(x_k^a, p_k^a).
\]

Equation (4.3) can be rewritten as:

\[
\begin{bmatrix}
(B_{xx}^f)_{k+1} \\
(B_{xp}^f)_{k+1}
\end{bmatrix} =
\begin{bmatrix}
M_k & N_k \\
0 & I
\end{bmatrix}
\begin{bmatrix}
(B_{xx}^a)_{k}^\top \\
(B_{xp}^a)_{k}^\top
\end{bmatrix}
\begin{bmatrix}
M_k^\top & 0 \\
N_k^\top & I
\end{bmatrix}.
\]

After the matrix multiplications in (4.6) the update for each covariance submatrix, i.e., state-state, state-parameter and parameter-parameter error covariances, can be written as

\[
(B_{xx}^f)_{k+1} = M_k (B_{xx}^a)_{k}^\top M_k^\top + N_k (B_{xp}^a)_{k}^\top M_k^\top + N_k (B_{pp}^a)_{k}^\top N_k^\top,
\]

\[
(B_{xp}^f)_{k+1} = M_k (B_{xp}^a)_{k} + N_k (B_{pp}^a)_{k},
\]

\[
(B_{pp}^f)_{k+1} = (B_{pp}^a)_{k}.
\]

4.2 Assimilation step

During the assimilation step not only state and parameter vectors are updated but the state error covariance matrix as well. For this, it is also necessary to linearise the
observation operator $\tilde{H}_k$. If $\tilde{H}_k \in \mathbb{R}^{m_k \times (n+\ell)}$ and $H_k \in \mathbb{R}^{m_k \times n}$ denote the linearisation of $\tilde{H}_k$ and $H_k$, respectively, then

$$\tilde{H}_k = \frac{\partial \tilde{H}_k}{\partial w_k} = \begin{pmatrix} H_k & 0_{m_k \times \ell} \end{pmatrix}.\quad (4.10)$$

For ease of notation we will omit the sub-index indicating the time step. Now we use (4.10) to compute the augmented Kalman gain matrix

$$K = B_f^T \tilde{H}^\top \left( \tilde{H} B_f^T \tilde{H}^\top + R \right)^{-1}.\quad (4.11)$$

Computing the terms separately:

$$B_f^T \tilde{H}^\top = \begin{pmatrix} B_{fx}^f & B_{fp}^f \end{pmatrix} \begin{pmatrix} H^\top & 0_{\ell \times m} \end{pmatrix} = \begin{pmatrix} B_{fx}^f H^\top \cr (B_{fp}^f)^\top H^\top \end{pmatrix},\quad (4.12)$$

$$\tilde{H} B_f^T \tilde{H}^\top = \begin{pmatrix} H & 0_{m \times \ell} \end{pmatrix} \begin{pmatrix} B_{fx}^f H^\top \cr (B_{fp}^f)^\top H^\top \end{pmatrix} = H B_{fx}^f H^\top,\quad (4.13)$$

Substituting into (4.11) we obtain the Kalman gain matrix

$$K = \begin{pmatrix} K_x \\ K_p \end{pmatrix} = \begin{pmatrix} B_{fx}^f H^\top \left( H B_{fx}^f H^\top + R \right)^{-1} \\ (B_{fp}^f)^\top H^\top \left( H B_{fx}^f H^\top + R \right)^{-1} \end{pmatrix}.\quad (4.14)$$

The upper block $K_x$ of the Kalman gain matrix is exactly the gain for the state estimation problem and is used to update the state vector. The lower block, $K_p$, depends on the cross-covariance $B_{fp}^f$ and is used to update the parameter vector as shown in the following

$$w^a = \begin{pmatrix} x^a \\ p^a \end{pmatrix} = w^f + \begin{pmatrix} K_x \\ K_p \end{pmatrix} \left( y - \mathcal{H}(w^f) \right) \quad (4.15)$$

$$= \begin{pmatrix} x^f \\ p^f \end{pmatrix} + \begin{pmatrix} K_x \\ K_p \end{pmatrix} \left( y - \mathcal{H}(x^f) \right),\quad (4.16)$$

i.e.,

$$x^a = x^f + K_x \left( y - \mathcal{H}(x^f) \right),\quad (4.17)$$

$$p^a = p^f + K_p \left( y - \mathcal{H}(x^f) \right).\quad (4.18)$$
The covariance matrix is updated according to

$$
B^a = (I - K\tilde{H})B^f
$$

which results in

$$
\begin{align*}
B^a_{xx} &= B^f_{xx} - K_x H B^f_{xx}, \\
B^a_{xp} &= B^f_{xp} - K_x H B^f_{xp}, \\
B^a_{pp} &= B^f_{pp} - K_p H B^f_{xp},
\end{align*}
$$

The update of the state-state error covariance $B_{xx}$ is exactly the same as if we would consider only the state estimation data assimilation problem. The cross-covariance matrix $B^f_{xp}$ intervenes in the updates of both state-parameter and parameter-parameter error covariances, as well as in the computation of the gain matrix $K_p$, used to update the parameter vector. This matrix is, therefore, extremely important for the parameter estimation problem.

### 4.3 Linear advection model. Assimilation experiments

In this section we test the performance of the EKF, when applied to a simple model with a single unknown parameter, in order to investigate the suitability of the filter under different observation conditions.

#### 4.3.1 Linear advection model

The linear advection model describes a simple sediment transport model, in which the bed height moves forward (horizontally) with unknown constant speed (see [70], [104], [99]). The mathematical model is given by

$$
\begin{align*}
\frac{\partial x}{\partial t} + s \frac{\partial x}{\partial z} &= 0, \\
x(z,0) &= x_0(z),
\end{align*}
$$

where $x(z,t)$ is the function representing the sea bed height for the one dimensional space coordinate $z$ at time $t$ and $s$ is the unknown speed.
Figure 4.1: Linear advection problem. Bed height moving at horizontal speed $s$.

The solution for the continuous problem is very simple:

$$x(z, t) = x_0(z - st). \quad (4.26)$$

4.3.2 Discrete linear advection model

The discrete model describes the evolution of the bed height $x = (x_1, x_2, \ldots, x_n)$ for equally spaced grid points $z = (z_1, z_2, \ldots, z_n)$ from time $t$ to time $t + \Delta t$. From (4.26) $x_k(z) = x_{k-1}(z - \Delta t)$. Unless $s\Delta t$ is a multiple of the grid spacing $\Delta z$ (which depends on the parameter $s$ and therefore cannot be controlled assuming that $s$ is unknown), in general it is necessary to estimate the bed heights from space points that do not belong to the grid. We use the upwind scheme ([101], [44], [85], [69])

$$x_{k+1,j} = x_{k,j} + \frac{\Delta t}{\Delta z}(x_{k,j-1} - x_{k,j}), \quad j = 1, 2, \ldots, n \quad k = 0, 1, \ldots, N \quad (4.27)$$

where $x_{k,j} \approx x(z_j, t_k)$, $z_j = j\Delta z$ and $t_k = k\Delta t$. This is equivalent to a linear interpolation using neighbour grid points. To update the first grid point $z_1$ we impose the boundary condition

$$x(z_0, t) = x(z_n, t), \quad \forall t. \quad (4.28)$$

This condition is equivalent to assume that the bed height is a periodic function, with period equal to the length of the analysed time window. For the consistency of our discretization scheme we will assume that the horizontal displacement of the bed height during time $\Delta t$ is at most $\Delta z$ so that the bed height for a specific grid point $z_i$ in the next time step will be equal to the bed height of some point $z^* \in [z_{i-1}, z_i]$. This is equivalent to the condition

$$s \leq \frac{\Delta z}{\Delta t}, \quad (4.29)$$
which in this context is also known as the Courant–Friedrichs–Lewy stability condition. The discretized model can be rewritten as a matrix-vector product:

\[ x_{k+1} = F_k x_k, \]  

(4.30)

where

\[ F_k = F(s_k) = \begin{pmatrix} 1 - \mu s_k & 0 & \ldots & 0 & \mu s_k \\ \mu s_k & 1 - \mu s_k & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \mu s_k & 1 - \mu s_k & 0 \\ 0 & \ldots & 0 & \mu s_k & 1 - \mu s_k \end{pmatrix} \]  

(4.31)

is an \( n \times n \) matrix depending on the unknown speed \( s_k \) and \( \mu = \Delta t / \Delta z \). Notice that because of the boundary conditions in \( z_0 \), the matrix \( F_k \) is not exactly lower bi-diagonal.

An important fact is that though the continuous model describes a periodic dynamical system, the discrete version (4.30) is not periodic. The reason is that due to the linear interpolation, the norm of the state vector is decreased after each iteration. This numerical dissipation causes the data profile to diminish its height and to spread with each iteration. In fact, it can be shown that for the discrete model the steady state consists in a vector where each component is the time-invariant average of the state vector elements. In Appendix A we include a formal proof. For our experiments we will assume that the discrete model, and not the continuous one, is the true model.

### 4.3.3 Augmented state for the advection model

For a known constant speed, the model (4.30) is linear. Nevertheless, in case that the speed is an unknown parameter, the resulting augmented problem is nonlinear. Using the notation introduced in this chapter we define the augmented state model and its derivatives as follows:

\[ w_k = \begin{pmatrix} x_k \\ s_k \end{pmatrix}, \]  

(4.32)

\[ w_{k+1} = \begin{pmatrix} F((w_k)_{n+1}) & 0_{n \times 1} \\ 0_{1 \times n} & 1 \end{pmatrix} w_k = \begin{pmatrix} F(s_k) & 0_{n \times 1} \\ 0_{1 \times n} & 1 \end{pmatrix} w_k. \]  

(4.33)

To propagate the state error covariance forward we need to compute the derivatives of the model with respect to state and parameter variables. Matrix \( F_k \) does not depend on the state \( x_k \), so according to (4.5)

\[ M_k = F_k. \]  

(4.34)
Chapter 4. EKF and augmented state

The derivatives of the model with respect to the parameter are computed according to

\[
\frac{\partial f_1^k}{\partial s_k} = \frac{\partial (\mu s_k x_{k,n} + (1 - \mu s_k) x_{k,1})}{\partial s_k} = \mu (x_{k,n} - x_{k,1}) \tag{4.35}
\]

\[
\frac{\partial f_i^k}{\partial s_k} = \frac{\partial ((1 - \mu s_k) x_{k,i} + \mu s_k x_{k,i-1})}{\partial s_k} = \mu (x_{k,i-1} - x_{k,i}), \quad i = 2, 3, \ldots, n \tag{4.36}
\]

and therefore

\[
N_k = \mu \begin{pmatrix} x_{k,n} - x_{k,1} \\ x_{k,1} - x_{k,2} \\ x_{k,2} - x_{k,3} \\ \vdots \\ x_{k,n-1} - x_{k,n} \end{pmatrix} \tag{4.37}
\]

Considering that we observe the bed height at every $\Delta x_{\text{obs}}$ grid-point, starting with the first one, the observation operator at each time step $k$ is the time-invariant matrix

\[
H_k = H = \begin{pmatrix} e_1^\top \\ e_{1+\Delta x_{\text{obs}}}^\top \\ e_{1+2\Delta x_{\text{obs}}}^\top \\ \vdots \end{pmatrix}, \tag{4.38}
\]

where $e_i$ is the $i$-th canonical vector in $\mathbb{R}^n$. The matching of state and observations is reflected by

\[
y_k = \begin{pmatrix} H & 0 \end{pmatrix} w_k = H x_k. \tag{4.39}
\]

4.3.4 Experiment set-up

The EKF applied to the advection model was tested within the framework of identical twin experiments. In this framework it is considered that the true state and parameter vectors are known. Then, the observations (also called pseudo-observations) are obtained by perturbing the true state trajectory, resulting from applying the model to the true state. For the perturbations a random noise is generated with the exact statistical distribution specified by the observation error covariance $R$. The idea is that, in this context, any error in the state or parameter approximation is caused because of deficiencies associated to the assimilation scheme and not because of a misspecification of the error statistics.

Equation (4.30) represents the true model. Moreover, it is assumed that the model is perfect. This means that given the true state at an arbitrary time $t_k$ the forecast state obtained by propagating this state forward in time using the model is the true state at time $t_{k+1}$. We consider $\Delta z = 0.01$ for the space discretization of the interval $[0, 3]$. The
state variable has then dimension \( n = 301 \). The length of the time step is \( \Delta t = 0.01 \). Therefore \( \mu = 1 \) and according to (4.29) we restrict the possible values of the speed parameter to the interval \([0 \, 1]\). The true state \( x_0 \) at initial time \( t = 0 \) is obtained by evaluating a similar exponential function as the one used in [101]:

\[
    x_0(z) = \begin{cases} 
        e^{-(z-0.25)^2} / 0.01 & \text{if } z \in [0.01, 0.5] \\
        0 & \text{if not.} 
    \end{cases} 
\]  

(4.40)

Here the discontinuities of \( x_0(z) \) at the places where the definition of the function changes are not relevant, as this function is only used to be evaluated in discrete points. The background initial state \( x_b \) is generated by adding random noise to the true bed height. The true speed is set to \( s^t = 0.8 \). In our experiments \( \Delta x_{\text{obs}} = 10 \). The observation vectors \( y_k \in \mathbb{R}^{31}, k = 1, 2, \ldots \) are generated adding Gaussian noise with zero mean and variance \( \sigma_o^2 = 0.01 \) to the true state values. The observation error covariance and the observation operator are the time-invariant matrices

\[
    R = 0.01I, \\
    H = \begin{pmatrix} 
        e_1^\top \\
        e_{11}^\top \\
        e_{21}^\top \\
        \vdots \\
        e_{301}^\top 
    \end{pmatrix}. 
\]  

(4.41) (4.42)

Initial covariance estimation

The initial covariance was obtained using a Monte Carlo approach. First, 100 initial samples were generated adding Gaussian noise \( \eta_x \sim \mathcal{N}(0, 0.1I) \) to the initial state \( x_0 \) and \( \eta_p \sim \mathcal{N}(0, 0.5) \) to the speed parameter. Those initial samples were run 100 steps forward using the forecast model, generating for each of them a trace of 100 states, corresponding to \( t = 1, 2, \ldots, 100 \). Then, random samples were taken to build an empirical covariance according to (2.16).

4.3.5 Results of standard EKF experiments

Assimilation experiments were performed using a standard implementation of the EKF for different frequencies of observations and initial background parameters. In Figure 4.2 we show the results obtained for 3 different assimilations along 2000 time steps, where the state is observed every 10, 20 and 30 time steps, respectively.
We start with an initial guess of the parameter speed $s^b = 0.3$, so the movement of the wave at the beginning is considerably slower compared to the truth. When assimilations are performed every 10 time steps, the resulting assimilated state at the end of the time window is very accurate and the corresponding parameter is very close to the truth after less than 200 time steps. The data assimilation is also successful for the case when observations are taken every 20 time steps. The difference is that it takes longer for the parameter to converge to the true value, as less information is assimilated. However, when observations are assimilated every 30 time steps the results obtained are completely wrong. The recovered state is to no extent similar to the true state. If we take a look at the assimilated parameter it seems to converge, but to a much lower speed value than expected. Furthermore, it can be noticed that the size of the parameter corrections decreases along the time window.

![Figure 4.2](image)

**Figure 4.2:** Data assimilation for initial background speed 0.3 and different observation frequencies. Upper plot: true parameter (black solid line), assimilated parameter observing the state every 10 (red), 20 (green) and 30 (blue) time steps. Lower plot: true state at time step 2000 (black solid line), assimilated state at time step 2000 observing the state every 10 (red), 20 (green) and 30 (blue) time steps, noisy observations (blue circles).

New experiments are performed, this time assimilating observations every 30 time steps but with different starting background speeds. The results are shown in Figure 4.3. Notice that for the initial parameter value $s^b = 0.4$ the state assimilation at $t = 2000$ is slightly improved when compared to the one obtained with $s^b = 0.3$. This is a direct consequence of a better approximated parameter. A moderate difference between the assimilated speed and the true speed causes a critical mismatch between the assimilated state and the true state when the model is run long enough. Nevertheless if we start with $s^b = 0.5$, the assimilation scheme is capable to accurately estimate the true state and parameter. Though as observed before, the size of the parameter corrections also
decreases along the assimilation window, the initial parameter is close enough to the true value, so that it can reach the true value before the corrections become too small.

Figure 4.3: Data assimilation for different initial background speeds for observations available every 30 time steps. Upper plot: true parameter (black solid line), assimilated parameters for initial value 0.3 (red), 0.4 (green) and 0.5 (blue) of the background initial speed parameter. Lower plot: true state at time step 2000 (black solid line), assimilated state at time step 2000 for initial value 0.3 (red), 0.4 (green) and 0.5 (blue) of the background initial speed parameter, noisy observations (blue circles).

In order to explain why the size of the parameter corrections decreases, we examine the size of the parameter variance and cross-covariance state-parameter vector $B_{xp}$ in Figure 4.4. After just a few assimilation steps both, $B_{pp}$ and $B_{xp}$, become very small, independently of the initial value we chose for the parameter. Similar results can be observed when assimilations are performed more frequently (Figure 4.5). A very small parameter variance indicates that the confidence in the estimation is high and therefore it should not be changed, while a small parameter-state error covariance means that the parameter and the state variables are poorly correlated and therefore little information can be inferred from changes on the state to update the parameter. If we take a look at the analysis step formulas (4.43) and (4.44) for updating the parameter and parameter-parameter error covariance

$$p^a = p^f + (B_{xp}^f)^\top H^\top (H B_{xx}^f H^\top + R)^{-1} (H(x^f) - y)$$  \hfill (4.43)

$$B_{pp}^a = B_{pp}^f - (B_{xp}^f)^\top H^\top (H B_{xx}^f H^\top + R)^{-1} H B_{xp}^f,$$  \hfill (4.44)

it is clear that if $B_{xp}$ is very small so is the parameter correction, except for very large values of the matrix inverse involved. For the sake of simplicity we have skipped the index indicating the time step. If $B_{xx}$ and $R$ are fixed, then the size of the parameter correction goes to zero when the norm of $B_{xp}$ goes to zero. On the other hand, from
If \( B_{pp} \) is small so must be \( B_{xp} \). Notice that because \((H B_{xx} H^\top + R)^{-1} \succ 0\) we have \((H B_{xp} x)^\top (H B_{xx} H^\top + R)^{-1} (H B_{xp} x) \geq 0 \ \forall x\), thus \( B_{xp} H^\top (H B_{xx} H^\top + R)^{-1} H B_{xp} \) is at least positive semi-definite. It implies that the parameter variance always decreases with each assimilation step. Because the parameter estimate is not changed by the forecast model, the variance is kept fixed during the forecast step (see (4.9)).
means that once the parameter variance is small, this value is never increased along the assimilation process. In our experiments the fact that the parameter variance becomes very small, even when the estimated value of the parameter is far from the true value, shows an overconfidence of the filter that causes the assimilation process to fail.

### 4.3.6 Results fixing the parameter variance

In order to improve the assimilation we can force the parameter variance not to decrease too much. A simple way to do this is to keep it fixed along the time window. New twin experiments were performed for an initial speed of 0.3 with observations taken every 30 time steps. For this new experiment (4.44) is substituted by $B_{pp}^a = B_{pp}^f$ in the assimilation step.

Figure 4.6 shows that a good approximation of the state at the end of the assimilation window is obtained for a fixed parameter variance of 0.5, despite the fact that the parameter value oscillates around the true parameter. Similar results were obtained when the parameter variance was set to values between 0.2 and 1, which is a huge range of values, taking into account that the parameter is restricted to the interval $[0, 1]$.

The filter detects that the wave is moving too slow and compensates it by increasing the speed, until the opposite situation arises, where the wave moves too fast, and the speed is decreased, and so on.

![Figure 4.6: Data assimilation for 0.3 initial background speeds and observations available every 30 time steps with fixed parameter variance along the assimilation window. Up: true parameter (black solid line) and assimilated parameters for initial value (blue). Down: true state at time step 2000 (black solid line) and assimilated state at time step 2000 for initial value (blue), noisy observations (blue circles).](image-url)
Chapter 4. EKF and augmented state

In Figure 4.7 we show the value of the parameter variance and the norm of $B_{xp}$. For every time step where observations are assimilated we have also represented the value that the EKF would have assigned to the parameter variance in case it would not have been kept fixed. It represents a decrease of around two orders of magnitude in most of the cases. As a side effect $\|B_{xp}\|_n$, which also decreases at every assimilation step, increases its value during the successive forecast steps. By manipulating (4.8) it is easy to see that if from time step $q$ to time step $l$ no assimilations are performed, then:

$$\left(B_{xp}\right)_l = M_{l-1} \left(B_{xp}\right)_{l-1} + N_{l-1} \left(B_{pp}\right)_{l-1}$$

$$= M_{l-1} \left(M_{l-2} \left(B_{xp}\right)_{l-2} + N_{l-2} \left(B_{pp}\right)_{l-2}\right) + N_{l-1} \left(B_{pp}\right)_{l-1}$$

$$= M_{l-2} \left(B_{xp}\right)_{l-2} + M_{l-2} N_{l-2} \left(B_{pp}\right)_{l-2} + N_{l-1} \left(B_{pp}\right)_{l-2}$$

$$\vdots$$

$$= M_{q}^{l-q} \left(B_{xp}\right)_{q} + \left(\sum_{i=0}^{l-q-1} M_{q}^{i} N_{l-1-i}\right) \left(B_{pp}\right)_{q}.$$  (4.48)

Notice that $M_k = M(s_k)$, and because the parameter does not change if new observations are not assimilated, then $M_q = M_{q+1} = \ldots = M_{l-1}$. The second term of the sum in (4.48), for a sufficiently large parameter variance $B_{pp}$, contributes to keep the norm of $B_{xp}$ large enough for the next assimilation step.

Despite the fact that the value of the assimilated parameter at a specific time step cannot be considered a good approximation of the true parameter value, its average along the assimilation window is 0.7992, very close to the true value. This suggests that
the expected value of the parameter does converge to the correct speed. Based on this empirical result and in order to have a smoother assimilation sequence, the average of the last $t_a$ parameter estimations is updated after each assimilation and used as the new parameter value for the next forward steps. Figure 4.8 shows that by averaging the parameter we achieve convergence to the true solution. Furthermore the state estimation is also considerably improved. A pseudo-code for this EKF modification can be found in Algorithm 2.

![Figure 4.8](image)

**Figure 4.8:** Data assimilation for 0.3 initial background speeds and observations available every 30 time steps with fixed parameter variance along the assimilation window and averaging parameter update. Up: true parameter (black solid line) and assimilated parameters for initial value (blue). Down: true state at time step 2000 (black solid line) and assimilated state at time step 2000 for initial value (blue), noisy observations (blue circles).

### 4.3.7 Results fixing the size of the norm of $B_{xp}$

A more direct way of avoiding too small parameter corrections is to force the cross-covariance $B_{xp}$, which is the one intervening in the parameter update, not to become too small. A simple approach is to resize the cross-covariance every time an observation is assimilated, such that $\|B_{xp}\|_2 = q$ for a given constant $q$. It is important that by manipulating a part of the covariance $B$ its positive definite property is not altered. This can be assured by taking into account the condition (3.10), which for the single parameter case is equivalent to

$$B > 0 \iff B_{pp} - B_{xp}^\top B_{xx}^{-1} B_{xp} > 0.$$  \hspace{1cm} (4.49)
Chapter 4. EKF and augmented state

Initialization $B^f$, $w^f$, $H$, $t_a$, paramAve=0, paramQueue= [], assimStep=0.

Step 1. Assimilation Step

IF observations not available THEN go to Step 3, END

Compute $K$ and $w_a$ according to (4.14) and (4.16)

$$B^{a}_{xx} = B^{f}_{xx} - K_x H B^{f}_{xx},$$
$$B^{a}_{xp} = B^{f}_{xp} - K_x H B^{f}_{xp},$$
$$B^{a}_{pp} = B^{f}_{pp}, \quad \% \text{Instead of the EKF update } B_{pp} \text{ is kept fixed}$$

assimStep++

Step 2. Update of parameter average

$p^a = w^a(\text{end} - \text{amountParam} + 1 : \text{end}),$

paramQueue.Add$(p^a))$

IF $\text{assimStep} \leq t_a$, THEN

$$\text{paramAve} = \frac{(\text{assimStep}-1)\text{paramAve}+p^a}{\text{assimStep}}$$

ELSE

$p=\text{paramQueue.Remove}(),$

\text{paramAve}+ = \frac{p^a-p}{t_a},$

$w^a(\text{end} - \text{amountParam} + 1 : \text{end}) = \text{paramAve},$

END

Step 3. Forecast Step

Update $w^f$ and $B^f$ according to (4.2), (4.1) and (4.3)

go to Step 1.

Algorithm 2: EKF modification. The parameter variance is constant along the assimilation period and the parameter estimates are averaged over a moving time window.

Let us consider the approach of updating the state-parameter cross-covariance after each assimilation step according to $B^{a}_{xp} \text{update} = \frac{q}{\| B_{xp} \|_2} B^{a}_{xp}$. For simplicity we have skipped the index indicating the time step. From (4.49), if $q \geq \sqrt{B_{pp} B_{xp}^{-1} B_{xp}}$ the positive definiteness of $B$ is violated. We aim to considerably increase the norm of $B_{xp}$, therefore this strategy is not appropriate if this lower bound is small. Another possibility is to also modify the parameter-parameter error covariance $B_{pp}$.

If we define the time dependent matrix $D$ as

$$D = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \vdots & \vdots & \vdots \\ \vdots & 0 & \ddots & 0 \\ \vdots & 1 & \ddots & 0 \\ 0 & 0 & \ldots & 0 & \frac{q}{\| B_{xp} \|_2} \end{pmatrix} \quad (4.50)$$

by adding the correction

$$B^{a}_{\text{update}} = DB^{a}_{k} D^\top \quad (4.51)$$
at the end of the assimilation step, it is guaranteed that \( B^a \succ 0 \). According to (4.51) \( B_{xx}^a \) does not change while

\[
B_{pp}^{\text{update}} = \frac{q^2}{\|B_{xp}^a\|_2^2} B_{pp}^a, \\
B_{xp}^{\text{update}} = \frac{q}{\|B_{xp}^a\|_2^2} B_{xp}^a.
\]

Algorithm 3 summarizes this assimilation step modification.

Algorithm 3: EKF analysis step modification, consisting of fixing the size of the norm of the state-parameter covariance after each assimilation step and averaging the parameter estimation over a moving time window.

We tested the effectiveness of the strategy performing experiments for different values of \( q \). For \( 20 < q < 80 \) similar results were obtained. In Figure 4.9 results for \( q = 50 \) are shown. Similar to the experiments in the previous section, the assimilated parameter values are averaged over a moving time window.

Though the quality of the assimilation is slightly worse if compared to the results obtained by fixing the parameter variance, it is much more accurate than performing a standard EKF. On the other hand, the estimated parameter at the end of the assimilation process approximates accurately the true parameter value. By inspecting Figure 4.10 it can be noticed that by fixing the size of \( B_{xp} \) after each assimilation the norm of \( B_{pp} \) is implicitly controlled.

### 4.3.8 Results considering “model errors”

The fact that the parameter variance can only decrease during the assimilation time window holds only if the model is perfect. In the presence of model errors by adding the model error covariance \( Q \) during the forecast step, the parameter variance augments. In the following we consider the wrongly estimated speed as a source of error in the model and deduce an expression for the error covariance matrix \( Q \).
Let us consider the following two different forecast models for the non augmented problem:

\begin{align*}
    x_{k+1} &= M_k(s^f)(x_k) = f(x_k), \\
    x_{k+1} &= M_k(s_k)(x_k) = f_{\text{approx}}(x_k),
\end{align*}

Figure 4.9: Data assimilation for 0.3 initial background speeds and observations available every 30 time steps. After each assimilation $B^a_{xp}$ is resized such that $\|B^a_{xp}\|_0 = c$ for a given constant $c$ and the parameter is averaged over a moving time window. Up: true parameter (black solid line) and assimilated parameters for initial value (blue). Down: true state at time step 2000 (black solid line) and assimilated state at time step 2000 for initial value (blue), noisy observations (blue circles).

Figure 4.10: Parameter variance and cross-covariance vector $B_{xp}$ for assimilations every 30 time steps and initial speed 0.3. After each assimilation $B^a_{xp}$ is resized such that $\|B^a_{xp}\|_0 = c$ for a given constant $c$ and the parameter is averaged over a time window. Up: true parameter (black solid line) and assimilated parameters for initial value (blue). Down: true state at time step 2000 (black solid line) and assimilated state at time step 2000 for initial value (blue), noisy observations (blue circles).
where equation (4.54) represents the true model, in which the true speed is used for propagating the state forward. The model \( f_{\text{approx}} \) is just an approximated or imperfect model, as it uses the best estimate of the speed at time \( k \) and not the true value to compute the next state.

If we assume that \( s_k = s^t + \Delta s_k \), we can compute an expression for the model error:

\[
e^x_k = x^t_{k+1} - x^f_{k+1} = \Delta s_k \mu \left( \begin{array}{cccccc}
1 & 0 & \ldots & 0 & -1 \\
-1 & 1 & 0 & 0 & \ldots & \ldots \\
0 & -1 & 1 & 0 & \ldots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & -1 & 1 \\
\end{array} \right) x_k
\]

\[
= \Delta s_k N_k
\]  

(4.56)

If we consider now the augmented state problem, the augmented model error is equivalent to

\[
e^w_k = \Delta s_k \left( \begin{array}{c} N_k \\ 1 \end{array} \right).
\]  

(4.57)

Though the parameter error \( \Delta s_k \) is unknown we assume it is a random variable \( \Delta s_k \sim \mathcal{N}(0, \sigma^2_k) \). Because the vector \( e^w_k \) is the multiplication of the deterministic vector \( \left( \begin{array}{c} N_k \\ 1 \end{array} \right) \) and the random scalar variable \( \Delta s_k \), it is possible to find its statistical distribution:

\[
e^w_k \sim \mathcal{N}(0, Q_k),
\]  

(4.58)

with

\[
Q_k = \left( \begin{array}{cc}
N_k \sigma^2_k & N_k \sigma^2_k \\
\sigma^2_k N_k^\top & \sigma^2_k
\end{array} \right).
\]  

(4.59)

Remark 4.3.1. In this example \( Q_k \) is exactly computed, but in general an equivalent expression can be found for the error covariance, even for multiple parameters, under the assumption that the model is almost linear with respect to the parameters in a neighbourhood of the true parameter vector. Notice that writing the Taylor expansion for a vector valued function of the model until the first order term we have:

\[
f(x_k, p^t + \Delta p) \approx f(x_k, p^t) + N_k \Delta p \Rightarrow f(x_k, p^t + \Delta p) - f(x_k, p^t) = e^x_k \approx N_k \Delta p
\]  

(4.60)

Then

\[
e^w_k \approx \left( \begin{array}{c} N_k \Delta p \\ \Delta p \end{array} \right) = \left( \begin{array}{c} N_k \\ I_{\ell \times \ell} \end{array} \right) \Delta p.
\]  

(4.61)
If we assume that $\Delta p \sim N(0, B_{pp})$, then $e^w_k$ is also Gaussian with zero mean and covariance

$$Q_k = \begin{pmatrix} N_k \\ I_{\ell \times \ell} \end{pmatrix} B_{pp} \begin{pmatrix} N_k^\top \\ I_{\ell \times \ell} \end{pmatrix} = \begin{pmatrix} N_k B_{pp} N_k^\top \\ B_{pp} N_k^\top \\ B_{pp} \end{pmatrix}$$

(4.63)

If the model is strongly nonlinear higher orders of the Taylor expansion (e.g., involving the Hessian, etc.) could be considered. Alternatively, instead of the sensitivities of the model with respect to the parameter, an ensemble of states could be used to approximate the space of the local error caused by the parameter estimation in a similar way as it is used in Section 5.1.1.

Matrix $Q_k$ is the model error covariance matrix, used for updating the augmented state error covariance $B$:

$$B_{k+1} = \begin{pmatrix} M_k & N_K \\ 0 & 1 \end{pmatrix} B_k \begin{pmatrix} M_k^\top & 0 \\ N_k^\top & 1 \end{pmatrix} + Q_k$$

(4.64)

This way we do not only guarantee that the parameter variance is increased along the assimilation window by adding $\sigma_k^2$ during the forecast step, but also the uncertainty in the parameter estimation is incorporated into the state-state error covariance.

In the experiments we assume that the variance of the parameter $\sigma^2$ decreases proportionally with each assimilation. This is based on the assumption that the parameter estimation is improved as more observations are assimilated. If instead the variance is modified according to the EKF formulas, its value is reduced too much after a few assimilations and the quality of the assimilation is affected as it was previously discussed. A pseudo-code for the forecast step modification is presented in Algorithm 4, where the reduction of $\sigma^2$ is proportional to the fraction of the assimilation window being processed.

Figure 4.11 shows that the strategy is effective for accurately approximating the parameter. The contribution of $Q$ to the state-state error covariance allows a more precise estimate of the state vector at the end of the time window.

The linear advection model, for a fixed parameter is linear, which makes possible the exact computation of the model error matrix $Q$. In practical applications, where the dependence of the model on the parameter is typically nonlinear, a linearization of the model error must be taken into account to obtain $Q$. 
% Forecast Step modification

**Step 1.** Compute vector \( v_k \) and \( \sigma^2_k \)

\[
v_k(1) = x_k(1) - x_k(n)
v_k(i) = x_k(i) - x_k(i-1), \quad i = 1, 2, \ldots, n
\]

\[
\sigma^2_k = N_{\text{obsFreq}} \cdot \text{assimStep}
\]

**Step 2.** Update \( x_f \), \( B_f^{xx} \), \( B_f^{xp} \) and \( B_{pp} \)

\[
\begin{align*}
(B^{xx})_{k+1} &= M_k (B^{a}_{xx})_k M_k^\top + N_k (B^{a}_{xp})_k M_k^\top + M_k (B^{a}_{xp})_k N_k^\top + N_k (B^{a}_{pp})_k N_k^\top + N_k \sigma^2_k N_k^\top, \\
(B^{xp})_{k+1} &= M_k (B^{a}_{xp})_k + N_k (B^{a}_{pp})_k + N_k \sigma^2_k, \\
(B^{pp})_{k+1} &= (B^{a}_{pp})_k + \sigma^2_k
\end{align*}
\]

**Algorithm 4:** EKF modification. The statistics of a model error term caused by propagating the state using an incorrectly estimated parameter are approximated and incorporated to the covariance update formulas. Besides, the parameter estimates are averaged over the assimilation time window.

![Figure 4.11: Data assimilation for 0.3 initial background speeds and observations available every 30 time steps with fixed parameter variance along the assimilation window and averaging parameter update. Up: true parameter (black solid line) and assimilated parameters for initial value (blue). Down: true state at time step 2000 (black solid line) and assimilated state at time step 2000 for initial value (blue), noisy observations (blue circles).](image)

**4.4 Summary and discussion**

In this chapter the adaptation of the EKF formulas to the augmented state DA problem was presented. In order to test the performance of EKF in this context we introduced
the linear advection model, where the advection speed is constant but unknown and should be estimated during the assimilation process.

Twin experiments were designed to test the effectiveness of the filter for estimating the state and the parameter at the same time. Good results were obtained when observations were available often enough. Nevertheless, when the size of the time interval where the state is not observed was increased, the filter diverged if the parameter estimation was not close enough to the true value after a few assimilation steps. The reason is that after a couple of assimilations the size of the parameter variances becomes very small, even if the parameter value is still incorrectly estimated. Because the parameter is unchanged by the model, its variance can only decrease along the assimilation window unless model errors are considered. The consequence is that the size of the parameter correction becomes too small, eventually converging to a wrong value.

To overcome this problem we propose a simple strategy consisting on keeping the parameter variance unchanged during the assimilation step. By doing so the parameter estimation oscillates around its true value, but the state estimation is drastically improved. In fact, the parameter mean along the assimilation window is a good approximation of the truth. It was shown that by averaging the parameter estimates over a moving time window results can be improved.

By restoring the state-parameter error covariance norm to a predetermined size after each assimilation step it was also possible to improve the state estimation, though the quality of the analysis in the experiments performed was not as good as the one obtained with the former strategy.

As a third variant we interpreted the mismatch between the state forecast obtained from using the true advection speed and the one using the estimated parameter as some sort of model error. An explicit expression for the model error covariance matrix $Q$ was obtained. The experiments show the benefits of using the error information structure coming from the parameter uncertainty to improve the quality of the analysis.

What makes the EKF solution not viable in practical applications is the high computational cost of the forward running of the model for the forecast state error covariance update, which is not avoided in any of the proposed variations. For $x \in \mathbb{R}^n$ the cost of updating the covariance in the forecast step is $O(n^3)$. In typical applications, where $n \approx 10^8$, the cost of a single update of the covariance is prohibitive. Nevertheless, in this chapter we showed that the analysis can be significantly improved if a special treatment to the covariances involving the parameters is considered. In the following chapter we will focus on less computationally expensive sequential strategies, that unlike 3D-Var, account for inexpensive updates of the state error covariance matrix. Moreover, we
will propose a novel DA method which uses ideas related with the forecast modification presented in Section 4.3.8.
Chapter 5

Methods which Combine Ideas from Different Assimilation Schemes

Sequential 3D-Var methods constitute a viable alternative to the computationally expensive EKF, as the time-consuming update of the state error covariance is avoided. The fact that this covariance is kept constant along the assimilation window makes it possible to apply the method to high-dimensional models. The drawback is that there is no flow dependency reflected in the error statistics. In other words, the influence of model dynamics is not included in the estimation of the state error.

In the last few years there have been numerous attempts to combine variational and ensemble filter methods to improve the quality of data assimilation (DA) analysis. The general idea is to take advantage of the relative computational simplicity of 3D-Var methods, while using the ensemble methods to provide a flow-dependent estimation of the state error covariance. In [57], 3D-Var is applied to different ensemble members, using independent observation perturbations. The state error covariance for each ensemble is a linear combination of a time-invariant covariance and an empirical covariance, obtained using the rest of the ensemble members. The advantage of using such combination is that the sample covariance provides flow-dependent error statistics. On the other hand, adding a time-invariant covariance can help to ameliorate some typical sampling error problems and increase the rank of the rank-deficient empirical covariance matrices. Selecting an appropriate linear combination is a crucial point here. More recently, in [128], an ETKF is incorporated in the 3D-Var scheme for the Weather Research and Forecasting Model (WRF), using extended control variables ([74]). This is applied successfully in [127] to Hurricane Track Forecasts.
A comparison between different Ensemble-3D-Var analysis schemes can be found in [129]. There it is proven that replacing the background-error covariance term in the cost function by a linear combination of the original background-error covariance with the ensemble covariance is equivalent to the analysis through augmenting the state vector with another set of control variables, preconditioned upon the square root of the ensemble covariance. In [130], a comparison between the EKTF–Optimum Interpolation hybrid and the more expensive Ensemble Square Root Filter (EnSRF) is discussed. It is shown that for large ensemble sizes the hybrid approach is nearly as accurate as the EnSRF, while proving to be more robust for smaller ensembles. This makes it even more attractive for operational purposes due to its lower computational cost. Other related strategies can be found in [24] and [40].

The development of combined strategies has not only been limited to the combination of ensemble-based methods and 3D-Var. In the literature other approaches can be found where for example, 4D-Var methods and EnKF are successfully combined (see [68], [30], [47], [102], [16] and [137]).

All these methods can be applied to the augmented state problem to estimate both, parameter and state variables. Nevertheless, we think that a method for this specific problem should take into account its special characteristics, that is, the subset of variables representing the parameters are not modified by the forecast model, they are never observed and have a big impact on the system dynamics.

Section 5.1 presents a method proposed by Smith [101] for the state and parameter DA problem. In Section 5.2 we propose a generalization of Smith’s method, considering a flow-dependent state-state error covariance matrix. A summary of the main aspects presented in this chapter can be found in Section 5.3.

### 5.1 Flow-dependent $B_{xp}$ method

One method which combines ideas from different schemes, and is especially designed to solve the state-parameter DA problem can be found in the Ph.D. thesis from Smith [101] (see also [99], [100]). In this work the author, following the augmented state approach, proposes a strategy based on 3D-Var, but considering a flow dependent $B_{xp}$, while keeping $B_{xx}$ and $B_{pp}$ fixed. The update of the state-parameter error covariance is only necessary before an assimilation takes place at time $t_k$ and is carried out according to:

$$ (B_{xp})_k = N_{k-1}B_{pp}. $$  \hspace{1cm} (5.1)
In [101] this update of the state-parameter error cross-covariance is obtained from a simplification of the corresponding EKF update (4.8). Then, at each assimilation step the optimization problem (2.3) is solved. We will refer to this strategy as the Flow-Dependent State-Parameter Covariance 3D-Var (FDSPC). A pseudo-code for this method is presented in Algorithm 5.

<table>
<thead>
<tr>
<th>Initialization</th>
<th>( B_f ), ( w_f ), ( H ), ( t_a )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1.</strong> Parameter-state cross-covariance update at ( t_k )</td>
<td>IF observations not available THEN</td>
</tr>
<tr>
<td></td>
<td>( w^a_k = w^f_k ), go to Step 3, END</td>
</tr>
<tr>
<td></td>
<td>Compute the Jacobian ( N_{k-1} ) (see 4.4), Update ( B_{xp} ) according to ( (B_{xp})<em>k = N</em>{k-1}B_{pp} ),</td>
</tr>
<tr>
<td><strong>Step 2.</strong> Assimilation Step</td>
<td>Update ( B^{-1}<em>k ) using the updated ( (B</em>{xp})_k ) in an efficient way, Minimize the augmented function (2.3) using an appropriate optimization solver</td>
</tr>
<tr>
<td><strong>Step 3.</strong> Forecast Step</td>
<td>Update ( w^f ) according to the forecast model go to Step 1.</td>
</tr>
</tbody>
</table>

**Algorithm 5:** FDSPC sequential assimilation algorithm

In this study it was observed that it is possible to obtain accurate estimates of the states and parameters when the method is applied to a range of dynamical system problems, including morphodynamic models. An obvious advantage of this method is its computational simplicity and efficiency. To incorporate some flow dependency into the 3D-Var scheme it is necessary to compute the Jacobian of the state forecast model with respect to the parameters. In many applications the number of parameters is relatively small, so this operation does not add a considerable computational load to the assimilation process. Besides, the update of \( B_{xp} \), and therefore of \( B^{-1} \), is only necessary previous to each assimilation step, rather than at each forecast step. The inverse of the augmented state error covariance can be updated in an efficient way following the same idea presented in Appendix B.

### 5.1.1 Interpretation of the cross-covariance update

The cross-covariance update (5.1) constitutes a simplification of the full forecast update (4.8), where the term \( M_k (B_{kp})^a \) has been omitted. The advantage is that it reduces the overall computational cost because no model runs are required. The simplified update depends only on the Jacobian of the model with respect to the parameters and the parameter-parameter error covariance. Related to this update we make the following
Remark 5.1.1. The special $B_{xp}$ update in (5.1) is related to the computation of an empirical flow-dependent covariance, if the model is close to linear with respect to the parameter.

Let $p \sim \mathcal{N}(p_k, B_{pp})$. The covariance $B_{pp}$ can be used to generate $r$ parameter perturbations

$$
\delta p_i = B_{pp}^{1/2} \delta_i, \; i = 1, 2, \ldots, r,
$$

with $\delta_i \sim \mathcal{N}(0, I)$. Using the model, the state $x_k$ is transported one step forward in time using each of the $r$ perturbed parameters as shown below.

\[
\begin{align*}
\delta p_1 & \rightarrow p_1^k = p_k + \delta p_1, \; x_{k+1}^1 = f_k(x_k, p_1^k) \\
\delta p_2 & \rightarrow p_2^k = p_k + \delta p_2, \; x_{k+1}^2 = f_k(x_k, p_2^k) \\
& \vdots \\
\delta p_r & \rightarrow p_r^k = p_k + \delta p_r, \; x_{k+1}^r = f_k(x_k, p_r^k) 
\end{align*}
\]

Now the mean of the augmented state samples $w_{k+1}^i = \begin{pmatrix} x_{k+1}^i \\ p_{k+1}^i \end{pmatrix}$, $i = 1, 2, \ldots, r$ is computed:

$$
\bar{w}_{k+1} = \frac{1}{r} \sum_{i=1}^{r} w_{k+1}^i = \begin{pmatrix} \bar{x}_{k+1} \\ \bar{p}_{k+1} \end{pmatrix}.
$$

(5.3)

If we subtract this mean from the samples, the result is a new ensemble with zero mean:

$$
\tilde{w}_{k+1}^i = w_{k+1}^i - \bar{w}_{k+1} = \begin{pmatrix} x_{k+1}^i - \bar{x}_{k+1} \\ p_{k+1}^i - \bar{p}_{k+1} \end{pmatrix}, \; i = 1, 2, \ldots, r.
$$

(5.4)

If we call $\tilde{W}_{k+1}$ the matrix whose columns are the centred ensemble members $\tilde{w}_{k+1}^i$, the augmented state empirical covariance at time step $k + 1$ can be estimated as:

$$
P_{k+1} = \frac{1}{r-1} \tilde{W}_{k+1} \tilde{W}_{k+1}^\top.
$$

(5.5)

Notice that matrix $P_{k+1}$ is not full rank in case that the number of ensembles is smaller than the dimension of the state vector but at least it is symmetric and positive semi-definite. We will assume for simplicity that there is only one parameter to be estimated, but the results obtained can be generalize to the multiple-parameter case. If there is only one parameter to be estimated the state-parameter error covariance matrix $B_{xp}$ is
in fact a vector, and according to the empirical covariance
\[(B_{xp})_{k+1} \approx \frac{1}{r-1} [\tilde{W}_{k+1}(1: \text{end} - 1, :)] [\tilde{W}_{k+1}(:, :)]^\top. \tag{5.6}\]

After some manipulations we can write the \(j\)-th component of \(B_{xp}\) as
\[(\{B_{xp}\}_k)_{j} \approx \frac{1}{r-1} [\tilde{W}_{k+1}(1: \text{end} - 1, j)] [\tilde{W}_{k+1}(end, :)^\top] \approx \frac{1}{r-1} \sum_{i=1}^r ((x_{i,k+1})_j - (\bar{x}_{k+1})_j) (p_{i,k+1} - \bar{p}_{k+1}) \tag{5.7}\]
\[= \frac{1}{r-1} \sum_{i=1}^r \frac{((x_{i,k+1})_j - (\bar{x}_{k+1})_j)}{(p_{i,k+1} - \bar{p}_{k+1})^2} (p_{i,k+1} - \bar{p}_{k+1})^2 \tag{5.8}\]

It is expected that if the number of samples is sufficiently large \(\bar{p}_{k+1} \approx p_k\). The reason is that
\[p_{k+1} = \frac{1}{r} \sum_{i=1}^r p_{i,k+1} \approx p_k + \frac{1}{r} \sum_{i=1}^r \delta p_i \approx p_k + \mathbb{E}(\delta p_i) = p_k. \tag{5.10}\]

On the other hand, if \(f_k\) is close to linear with respect to the parameter in some neighbourhood \(V = \{p : |p - p_k| < \epsilon\}\), and the parameter variance is small enough such that the probability that \(\delta p_i \in V\) is high, then
\[(\bar{x}_{k+1})_j = \frac{1}{r} \sum_{i=1}^r (f_k (x_k, p_{i,k}))_j \approx \left( f_k (x_k, \frac{1}{r} \sum_{i=1}^r p_{i,k}) \right)_j = (f_k (x_k, \bar{p}_{k+1}))_j \tag{5.11}\]
\[\approx (f_k (x_k, p_k))_j. \tag{5.12}\]

It means that
\[
\frac{(x_{i,k+1})_j - (\bar{x}_{k+1})_j}{(p_{i,k+1} - \bar{p}_{k+1})} \approx \frac{f_k (x_k, p_k + \delta p_i))_j - (f_k (x_k, p_k))_j}{\delta p_i} \tag{5.13}
\]
\[\approx \frac{\partial (f_k)_j}{\partial p} (x_k, p_k). \tag{5.14}\]

Substituting (5.14) into (5.9):
\[(B_{zp})_j \approx \frac{1}{r-1} \sum_{i=1}^r \frac{(x_{i}^j - \bar{x}_{j})}{(p_{i} - \bar{p})} (p_{i} - \bar{p})^2 \tag{5.15}\]
\[\approx \frac{1}{r-1} \sum_{i=1}^r \frac{\partial (f_k)_j}{\partial p} (x_k, p_k) (p_{i} - \bar{p})^2 \tag{5.16}\]
\[\approx \frac{\partial (f_k)_j}{\partial p} (x_k, p_k) \left[ \frac{1}{r-1} \sum_{i=1}^r (p_{i} - \bar{p})^2 \right] \tag{5.17}\]
Chapter 5. Methods which combine ideas from different assimilation schemes

\[
(B_{xp})_j \approx \frac{\partial (f_k)}{\partial p}(x_k, p_k)\text{Var}(p) = (N_k)_j B_{pp},
\]

which leads to the proposed update.

5.2 Low-rank $B_{xx}$ update. Our proposed method

One of the results in [101] is that in the case where observations were sparse in time, the accuracy of the parameter and state estimation drastically dropped. The problem is that when the model is run without any correction for a longer period the state-state covariance does not reflect accurately how the dynamics influence the statistics of the state error. We take as an example the linear advection model presented in Chapter 4. In Figure 5.1 the evolution of the state-state error covariance for a fixed known parameter, taking into account only the evolution of the error statistics arising from applying the forward operator.

Even after 200 time steps the structure of the covariances at the beginning of the time window and at the end are very similar. The covariance between adjacent grid points is only slightly decreased and the entries around the matrix diagonal are slightly larger. The pinch moving down the diagonal is caused by a combination of the periodic boundary conditions of the model and the fact that the initial covariance has a very small value at the position $(1, n)$.

Even after 200 time steps the structure of the covariances at the beginning of the time window and at the end are very similar. The covariance between adjacent grid points is only slightly decreased and the entries around the matrix diagonal are slightly larger. The pinch moving down the diagonal is caused by a combination of the periodic boundary conditions of the model and the fact that the initial covariance has a very small value at the position $(1, n)$.

A completely different evolution can be observed for the case where the advection speed is an unknown parameter in the model. Figure 5.2 shows how much the uncertainties of the parameter influence the covariance between state variables. To see how the parameter statistics influence the covariance between state variables equation (4.7) can be inspected.

After 100 time steps there are a few entries of the matrix, whose absolute value are much larger (by several orders of magnitude) than the rest. They correspond to state
variables representing the bed height at grid points around the forecast maximum. This structure is similar to the one of $N_{100}^\top B_{pp} N_{100}$, that is, the rank 1 matrix obtained using the first order derivative of the model with respect to the parameter and the estimated variance of the parameter (Figure 5.3). This matrix is exactly the upper block matrix of the model error covariance in (4.60). Here we assume that the variance of the parameter is $B_{pp} = 0.1$, but a similar structure of the covariance is obtained even if the variance is set to a much smaller value.

\[ \begin{bmatrix} a \end{bmatrix} \]

\[ \begin{bmatrix} b \end{bmatrix} \]

\[ \begin{bmatrix} c \end{bmatrix} \]

Figure 5.3: Rank 1 matrices $N_{100} B_{pp} N_{100}^\top$ and $N_{200} B_{pp} N_{200}^\top$

### 5.2.1 Forward step covariance update

We propose an update of $B_{xx}$ during the forecast, which keeps the computational efficiency of the FDSPC, but at the same time uses the derivatives of the model with respect to the parameter to obtain a flow-dependent state-state error covariance. Instead of the EKF state-state error covariance update (4.7)

\[
\begin{bmatrix} B_{xx} \end{bmatrix}_{k+1} = M_k (B_{xx}^a)_k M_k^\top + N_k (B_{xp}^a)_k M_k^\top + M_k (B_{xp}^a)_k N_k^\top + N_k (B_{pp}^a)_k N_k^\top ,
\]

we use a much simpler one:

\[
\begin{bmatrix} B_{xx} \end{bmatrix}_{k+1} = \alpha_1(k)B_{xx}^k + \alpha_2(k) \left( N_k B_{pp} N_k^\top \right) ,
\]
which combines a time-invariant background state-state covariance $B_{xx}^b$ with a flow-dependent low rank matrix. The static covariance term represents the error in longer-time-scales and contributes to keep the rank of the state-state error covariance full, while the flow dependent (and typically low rank) matrix accounts for the errors in shorter-time-scales, caused by applying the forecast model using a wrong parameter value.

To complete the forecast step the update of the state-parameter error covariance is performed according to (5.1), while the parameter-parameter error covariance is invariant over the time window. Moreover, all the matrix updates must be performed only before an observation is assimilated.

This update uses the Remark 4.3.1 in Section 4.3.8. First we consider a standard 3D-Var method, but assuming that the parameter vector is a known constant. Then we compute a model error matrix $Q_k$, which reflects the effect of the parameter estimation error in the model bias. This matrix, for the multiple parameter case, is obtained in (4.63). Then the resulting augmented state error covariance update is given by

$$
\begin{pmatrix}
B_{xx}^f_{k+1} \\
B_{xp}^f_{k+1}
\end{pmatrix}
= \begin{pmatrix}
B_{xx}^b & 0_{n \times \ell} \\
0_{\ell \times n} & 0_{\ell \times \ell}
\end{pmatrix}
+ \begin{pmatrix}
N_k B_{pp} N_k^\top & N_k B_{pp} \\
B_{pp} N_k^\top & B_{pp}
\end{pmatrix}
= \begin{pmatrix}
B_{xx}^b + N_k B_{pp} N_k^\top & N_k B_{pp} \\
B_{pp} N_k^\top & B_{pp}
\end{pmatrix}
$$

The functions $\alpha_1$ and $\alpha_2$ in the update of $B_{xx}$ are included to provide the flexibility of specifying how much importance is given to the long-time-scale and short-time-scale errors. In this work we will only consider time-invariant convex combinations, which means:

$$
\begin{align*}
\alpha_1(t) &= \alpha, \\
\alpha_2(t) &= 1 - \alpha,
\end{align*}
$$

for a given $\alpha > 0$, though other time-dependent approaches should be studied in further investigations. Substituting into (5.20):

$$
\left( B_{xx}^f \right)_{k+1} = \alpha B_{xx}^b + (1 - \alpha) \left( N_k B_{pp} N_k^\top \right).
$$

For a small number of parameters the update of $B_{xx}$ is the sum of two matrices: a full-rank time-invariant and low-rank flow-dependent one. Because of the low-rank term we name our approach Low-Rank State-State Covariance Update (LRSSC) method.
5.2.2 Computational implementation

The matrix $B^{-1}$ in (2.3) can be computed according to the following lemma:

**Remark 5.2.1.** The inverse of the augmented state error covariance matrix can be updated following the four-step computation:

1. 
   
   \[
   B^{-1}_{xx} = \frac{B^{-1}_{xx}}{\alpha_1} - \frac{B^{-1}_{xx}}{\alpha_1} N \left( \frac{B^{-1}_{pp}}{\alpha_2} + \frac{N^\top B^{-1}_{xx}N}{\alpha_1} \right)^{-1} \frac{N^\top B^{-1}_{xx}}{\alpha_1} \]

2. 
   
   $B_{xp} = NB_{pp}$

3. 
   
   $S = \left( B_{pp} - B_{xp}^\top B^{-1}_{xx} B_{xp} \right)^{-1}$

4. 
   
   \[
   B^{-1} = \begin{pmatrix} B_{xx} & B_{xp} \\ B_{xp}^\top & B_{pp} \end{pmatrix}^{-1} = \left( \begin{array}{cc} B^{-1}_{xx} + B^{-1}_{xx} B_{xp} S B_{xp}^\top B^{-1}_{xx} & -B^{-1}_{xx} B_{xp} S \\ -S B_{xp}^\top B^{-1}_{xx} & S \end{array} \right). \]

(5.26)

For ease of notation we skipped the index indicating the time step. Equation (5.26) is a direct result of applying the Sherman-Morrison-Woodbury matrix identity [135] to the proposed $B_{xx}$ update:

\[
B^{-1}_{xx} = \left( \alpha_1 B^{-1}_{xx} + \alpha_2 NB_{pp}N^\top \right)^{-1}
\]

(5.30)

Once $B_{xx}$ is updated, $B^{-1}$ can be computed according to the block matrix inverse identity presented in [56]:

\[
\begin{pmatrix} A & U \\ V & C \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}USV A^{-1} & -A^{-1}US \\ -SVA^{-1} & S \end{pmatrix},
\]

(5.31)

where $S := (C - VA^{-1}U)^{-1}$ is the inverse of the Schur complement of $A$ in the block matrix.

To evaluate the cost function (2.3), we do not need to compute the matrix $B^{-1}$ explicitly. In fact, in numerous real applications it is not possible to store $B$ or $B^{-1}$, because of the high dimensionality of the model. Instead, a computational code to compute the
product $B^{-1}\delta z$ is available, for a given vector $\delta z = \begin{pmatrix} \delta x \\ \delta p \end{pmatrix}$. As shown in Appendix B, this can be done at a computational cost of $O(\ell r)$, where $r$ is the cost of evaluating $(B_{xx}^b)^{-1}\delta x$ and $\ell$ is the number of unknown parameters. This does not alter the order of computing the simpler covariance update proposed by Smith. In a huge number of applications $\ell$ is relatively small, so both strategies are very efficient.

### 5.2.3 LRSSC algorithm

We can summarize our strategy in the following pseudo-code:

**Initialization** $B^f$, $w^f$, $H$, $t_a$

**Step 1.** Parameter-state cross-covariance update at $t_k$

- IF observations not available THEN
  - $w^a_k = w^f_k$,
  - go to Step 3,
- END

- Compute the Jacobian $N_{k-1}$ (4.4),
- Update $B_{xp}$ according to $(B_{xp})_k = N_{k-1}B_{pp}$,
- Update $B_{xx}$ according to $(B_{xx})_k = \alpha B_{xx}^b + (1 - \alpha)N_{k-1}B_{pp}N_{k-1}^\top$.

**Step 2.** Assimilation Step

- Find the state that minimizes (2.3) using some optimization solver.
- To evaluate $B_k^{-1}w_k$ follow the idea presented in Appendix B.

**Step 3.** Forecast Step

- Update $w^f_k$ according to the forecast model,
- go to Step 1.

![Algorithm 6: Sequential LRSSC assimilation algorithm.](image)

### 5.3 Summary and discussion

In this chapter we presented a brief review of several strategies which combine ideas from different DA schemes. We focused on the FDSPC approach, which is designed to solve the joint state-parameter DA problem via state augmentation. The strategy is based on updating only the state-parameter error covariance, using the derivatives of the model with respect to the parameter. We showed that there is a relation between this cross-covariance update and the computation of an empirical covariance. Though in good results were obtained when the method was applied to several test problems, the fact that the state-state error covariance is kept fixed along the assimilation window limits the accuracy of the state estimation when observations are inaccurate or sparse in time.
We proposed an approach, where the state-parameter error covariance is updated according to the FDSPC formulation. Additionally, the state-state error covariance is also updated, using a linear combination of a static background error covariance and a generally low-rank matrix (assuming that the number of parameters is small). We expect to improve the assimilation results by keeping a flow-dependency in the state-state error covariance.

One of the crucial aspects of our proposal is the selection of the weighted functions $\alpha_1$ and $\alpha_2$. In this work we analyse only the case where both functions are constant. Furthermore, we assume that the linear combination is convex, which means that $\alpha_1 = \alpha$, $\alpha_2 = 1 - \alpha$, with $0 \leq \alpha \leq 1$. If $\alpha_1 \approx 1$ then the FDSPC and the LRSSC are very similar (the FDSPC being a particular case of our more general formulation for $\alpha = 1$).

In the next chapter we study the convergence of both strategies, and in general of 3D-Var methods, when applied to the linear state-parameter DA problem.
Chapter 6

Convergence of 3D-Var-like Methods for the Linear Models

Though in practical applications dynamical models are generally nonlinear, the study of the linear ones is of crucial importance, as many of the data assimilation (DA) methods are based on solving this simple case. In this chapter we investigate the convergence of 3D-Var like methods and in particular of the Flow-Dependent State-Parameter Covariance 3D-Var (FDSPC) proposed by Smith in [101] and our proposed Low-Rank State-State error Covariance update (LRSSC) method for models which are linear with respect to both parameter and state vectors.

In Section 6.1 we state the linear dynamical system for which the state and parameter DA is to be performed, and present the associated augmented state problem. Section 6.2 shows how the forecast and assimilation steps can be condensed in a new time-invariant dynamical system that describes the evolution of the state and parameter estimation from one assimilation to the next. In Section 6.3 the dynamical system for the state error is deduced, and its convergence is analysed in Section 6.4. Special convergence conditions can be found in section 6.5, for the case where the full state is observed and observations are very accurate. Section 6.6 studies the case where less observations than parameters are available. Special conditions for convergence when the state estimation is very accurate are found in Section 6.8. Assimilation experiments are conducted for the linear heat equation in Section 6.9, in order to investigate how the theoretical results apply, even if some conditions are relaxed. Finally we summarize and discuss the principal results of this chapter in Section 6.10.
6.1 Problem statement

We consider the joint state-parameter DA problem associated to the dynamical system

$$x_k = Ax_{k-1} + Tp,$$  \hspace{0.5cm} (6.1)  $$y_k = Hx_k + \eta_k,$$  \hspace{0.5cm} (6.2)

where $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^m$, $p \in \mathbb{R}^\ell$ and $\eta_k \sim \mathcal{N}(0, R)$. Matrices $A \in \mathbb{R}^{n \times n}$, $T \in \mathbb{R}^{n \times \ell}$ and $H \in \mathbb{R}^{m \times n}$ are time-invariant. The forecast model is linear with respect to both parameters and state variables. Besides, we assume that the model is perfect.

Following the augmented state approach, (6.1)-(6.2) is transformed into

$$w_{k+1} = \begin{pmatrix} x_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} Ax_k + Tp_k \\ 0 \end{pmatrix} w_k,$$ \hspace{0.5cm} (6.3)  $$y_k = \begin{pmatrix} H \\ 0 \end{pmatrix} w_k + \eta_k.$$ \hspace{0.5cm} (6.4)

6.2 Resulting dynamical system

We will assume that observations are available at every time step. Then a 3D-Var assimilation of a sequence of observations for the linear problem is equivalent to applying at each time step the Kalman Filter assimilation step update where $B_{xx}$, $B_{xp}$ and $B_{pp}$ are fixed in time:

$$x_k^a = x_k^f + K_x(y_k - Hx_k^f)$$ \hspace{0.5cm} (6.5)  $$= Ax_k^a - 1 + Tp_k^a + K_x(y_k - H(Ax_k^a - 1 + Tp_k^a))$$ \hspace{0.5cm} (6.6)  $$= (I - K_x H) Ax_k^a - 1 + (I - K_x H) Tp_k^a - 1 + K_x y_k$$ \hspace{0.5cm} (6.7)  $$= \begin{pmatrix} (I - K_x H) \end{pmatrix} Ax_k^a - 1 + \begin{pmatrix} (I - K_x H) \end{pmatrix} Tp_k^a - 1 + \begin{pmatrix} K_x \end{pmatrix} y_k.$$ \hspace{0.5cm} (6.8)  $$p_k^a = p_k^f + K_p(y_k - Hx_k^f)$$ \hspace{0.5cm} (6.9)  $$= p_k^{a-1} + K_p(y_k - H(Ax_k^{a-1} + Tp_k^{a-1}))$$ \hspace{0.5cm} (6.10)  $$= p_k^{a-1} + K_p(y_k - K_p H Ax_k^{a-1} - K_p H T p_k^{a-1})$$ \hspace{0.5cm} (6.11)  $$= -K_p H Ax_k^{a-1} + (I - K_p H) T p_k^{a-1} + K_p y_k.$$ \hspace{0.5cm} (6.12)

We have separated the Kalman gain matrix as the two block-row matrix

$$K = \begin{pmatrix} K_x \\ K_p \end{pmatrix} = \begin{pmatrix} B_{xx} H^T (H B_{xx} H^T + R)^{-1} \\ B_{xp} H^T (H B_{xx} H^T + R)^{-1} \end{pmatrix}.$$ \hspace{0.5cm} (6.13)
The evolution of the augmented state vector after a sequence of assimilation-forecast cycle can be described by the dynamical system

\[
\begin{pmatrix}
x^a_k \\
p^a_k
\end{pmatrix} = \tilde{M} \begin{pmatrix}
x^a_{k-1} \\
p^a_{k-1}
\end{pmatrix} + \begin{pmatrix}
K_x \\
K_p
\end{pmatrix} y_k,
\]

(6.14)

where

\[
\tilde{M} := \begin{pmatrix}
(I - K_x H) A & (I - K_x H) T \\
-K_p H A & I - K_p H T
\end{pmatrix}.
\]  
(6.15)

The assumption that observations are available at every time step is not very restrictive. In fact, if instead we have observations available only every \(s\) time steps, matrices \(A\) and \(T\) in (6.15) are substituted by \(A' = A^s\) and \(T' = \left(\sum_{i=0}^{s-1} A^s\right) T\), respectively, as we show in the following.

The evolution of the state and parameter vectors from time \(t\) (where the state was observed) to time \(t + s\) (where the next observation is available) can be obtained as:

\[
\begin{pmatrix}
x^f_{k+s} \\
p^f_{k+s}
\end{pmatrix} = \begin{pmatrix}
A & T \\
0 & I
\end{pmatrix}^s \begin{pmatrix}
x^a_k \\
p^a_k
\end{pmatrix}.
\]  
(6.16)

Now notice that

\[
\begin{pmatrix}
A & T \\
0 & I
\end{pmatrix}^2 = \begin{pmatrix}
A & T \\
0 & I
\end{pmatrix} \begin{pmatrix}
A & T \\
0 & I
\end{pmatrix} = \begin{pmatrix}
A^2 & AT + T \\
0 & I
\end{pmatrix} = \begin{pmatrix}
A^2 & (A + I)T \\
0 & I
\end{pmatrix},
\]

(6.17)

\[
\begin{pmatrix}
A & T \\
0 & I
\end{pmatrix}^3 = \begin{pmatrix}
A^2 & (A + I)T \\
0 & I
\end{pmatrix} \begin{pmatrix}
A & T \\
0 & I
\end{pmatrix} = \begin{pmatrix}
A^3 & (A^2 + A + I)T \\
0 & I
\end{pmatrix},
\]

(6.18)

\[
\vdots \quad \text{(direct proof using induction)}
\]

\[
\begin{pmatrix}
A & T \\
0 & I
\end{pmatrix}^s = \begin{pmatrix}
A^s & \left(\sum_{i=0}^{s-1} A^s\right) T \\
0 & I
\end{pmatrix}.
\]  
(6.19)

The dynamical system matrix \(\tilde{M}\) is constant along the assimilation window, which means that the system is time invariant. In the FDSPC at every assimilation step the corresponding covariance matrix is

\[
B_k = \begin{pmatrix}
B_{xx} & N_k B_{pp} \\
B_{pp} N_k^\top & B_{pp}
\end{pmatrix},
\]  
(6.20)

with \(N = \frac{\partial f}{\partial p}(x^a, p^a)\) and \(f\) the forecast model. Here the problem is linear, which implies \(N_k = T\). The matrix \(B_k\) is constant along the assimilation process, and therefore \(\tilde{M}\) is
time-invariant as well. The resulting gain matrix is

\[ K = \begin{pmatrix} B_{xx} H^\top (HB_{xx} H^\top + R)^{-1} \\ B_{pp} T^\top H^\top (HB_{xx} H^\top + R)^{-1} \end{pmatrix}. \]  

(6.21)

If our proposed LRSSC method is applied, using (5.25) it can be directly checked that the strategy is equivalent to 3D-Var with gain

\[ K = \begin{pmatrix} (\alpha B_{xx} + (1 - \alpha)TB_{pp} T^\top) H^\top (\alpha B_{xx} + (1 - \alpha)TB_{pp} T^\top) H^\top + R \end{pmatrix}^{-1} \]

(6.22)

6.3 State error dynamical system

If we denote the true state at time step \( k \) as \( w^t_k \) then

\[ e^w_{k+1} = w^a_{k+1} - w^t_{k+1} \]

(6.23)

\[ = \tilde{M} w^a_k + (Ky_{k+1} - w^t_{k+1}). \]

(6.24)

The second addend in (6.24) is equivalent to

\[ Ky_{k+1} = -\tilde{M} w^a_k - \tilde{M} w^t_k + K\eta_{k+1}. \]

(6.25)

Substituting (6.29) in (6.24):

\[ e^w_{k+1} = \tilde{M} w^a_k + K\eta_{k+1} = \tilde{M} e^w_k + K\eta_{k+1}. \]

(6.30)

Equation (6.30) represents the dynamical system for the augmented state error.
6.4 Convergence

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues, real or complex, of the square matrix $A \in \mathbb{R}^{n \times n}$ and $\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}$ the spectral radius of $A$. Then the convergence of the 3D-Var method applied to the DA problem (6.1)-(6.2) is related to the spectral radius of $\tilde{M}$ as enunciated in the following theorem:

**Theorem 6.4.1.** The expected value of the error dynamical system in (6.30) converges to zero if and only if $\rho(\tilde{M}) < 1$.

**Proof.** Let's write the recursion (6.30) explicitly:

\[
\begin{align*}
    e_1^w &= \tilde{M}w_0 + K\eta_1 \\
    e_2^w &= \tilde{M}^2w_0 + \tilde{M}K\eta_1 + K\eta_2 \\
    & \vdots \\
    e_k^w &= \tilde{M}^k w_0 + \sum_{i=0}^{k-1} \tilde{M}^i \eta_{k-i}. 
\end{align*}
\]

Now computing the expected value of the error for any random variable with expected value $\bar{\eta}$ (our case is even easier because we assume that $\bar{\eta} = 0$):

\[
\begin{align*}
    \mathbb{E}(e_k^w) &= \mathbb{E}(\tilde{M}^k w_0) + \mathbb{E}\left(\sum_{i=0}^{k-1} \tilde{M}^i \bar{\eta}_{k-i}\right) \\
    &= \tilde{M}^k w_0 + \sum_{i=0}^{k-1} \tilde{M}^i \mathbb{E}(\eta_{k-i}) \\
    &= \tilde{M}^k w_0 + \left(\sum_{i=0}^{k-1} \tilde{M}^i\right) \bar{\eta}. 
\end{align*}
\]

It is clear that if $\rho(\tilde{M}) > 1$, then $\tilde{M}^k w_0$ diverges, while if $\rho(\tilde{M}) < 1$, then $(I - \tilde{M})^{-1}$ exists (all its eigenvalues are positive) and (6.37) is equivalent to

\[
\begin{align*}
    \mathbb{E}(e_k^w) &= \tilde{M}^k w_0 + (I - \tilde{M}) (I - \tilde{M})^{-1} \bar{\eta}.
\end{align*}
\]

In the limit

\[
\lim_{k \to +\infty} \mathbb{E}(e_k^w) = (I - \tilde{M})^{-1} \bar{\eta},
\]

and because $\bar{\eta} = 0$ the error converges to 0 and therefore, the expected value of the assimilated state and parameter variables converge to their true values.
Notice that the fact that the expected value of the estimation error converges to zero does not mean that the estimates of the state and parameter vectors converge to their true values, but that their expected values do. Therefore, in this chapter, when we refer to the convergence of a certain method, we mean that the obtained estimates are unbiased.

### 6.5 Full and accurate observations

As a special case we analyse the convergence, when the whole state is observed ($H = I$) and observations are very accurate ($R \approx 0$). Substituting $R = 0$ we can analyse the convergence for the limit case, when observations are infinitely accurate. It is clear that the assimilated state after each iteration is in this case $x_k^a = x_k^l = y_k$. The convergence depends only on whether the parameter (which is not observed) can be recovered throughout the assimilation process or not.

Substituting $H$ and $R$ in (6.13):

$$
\begin{pmatrix}
K_x \\
K_p
\end{pmatrix} = \begin{pmatrix}
B_{xx} I (IB_{xx} I + 0)^{-1} \\
B_{xp} I (IB_{xx} I + 0)^{-1}
\end{pmatrix} = \begin{pmatrix}
I \\
B_{xp} B_{xx}^{-1}
\end{pmatrix}.
$$

(6.40)

The matrix $\tilde{M}$ is then:

$$
\tilde{M} = \begin{pmatrix}
0 & 0 \\
-B_{xp} B_{xx}^{-1} A & I - B_{xp} B_{xx}^{-1} T
\end{pmatrix}.
$$

(6.41)

Computing the characteristic polynomial of $\tilde{M}$:

$$
\det(\tilde{M} - \lambda I) = \det \begin{pmatrix}
-\lambda I & 0 \\
-B_{xp} B_{xx}^{-1} A & (1 - \lambda) I - B_{xp} B_{xx}^{-1} T
\end{pmatrix} = -\lambda^n \det((1 - \lambda) I - B_{xp} B_{xx}^{-1} T) = 0.
$$

(6.42)

The method converges if and only if all solutions of

$$
\det((1 - \lambda) I - B_{xp} B_{xx}^{-1} T) = 0,
$$

(6.44)

which are the eigenvalues of $\tilde{M}$, have absolute value smaller than one. Equation (6.44) is equivalent to the condition that $\lambda' = 1 - \lambda$ is an eigenvalue of $B_{xp} B_{xx}^{-1} T$. The convergence is then attained if and only if

$$
|1 - \lambda (B_{xp} B_{xx}^{-1} T)| < 1.
$$

(6.45)
Let $z \in \mathbb{C}$, $z = a + bi$.

$$|1 - z| < 1 \iff |1 - a - bi| < 1 \iff (1 - a)^2 + b^2 < 1$$

$$a^2 + b^2 - 2a < 0 \iff |z|^2 < 2 \text{Re}(z)$$

From (6.47) it follows that the method converges if and only if for all eigenvalues $\lambda_i$ of matrix $B_{xx}^{-1}T$ it holds that $|\lambda_i|^2 < 2 \text{Re}(\lambda_i)$. This is also equivalent to the statement that all the eigenvalues must be in the interior of the circle with centre in $(1,0)$ and radius 1 represented in Figure 6.1.

If specifically the FDSPC or the LRSSC are used ($B_{xp} = TB_{pp}$), using the factorization $B_{pp} = B_{pp}^{\frac{1}{2}}B_{pp}^{\frac{1}{2}}$, an affine transformation can be found for which the following holds:

$$\lambda(B_{xx}^{-1}TB_{xx}T) = \lambda(B_{pp}^{-\frac{1}{2}}TB_{pp}^{-\frac{1}{2}}TB_{pp}^{-\frac{1}{2}})$$

The matrix in (6.51) is symmetric, moreover, positive definite if $T$ is a full column rank matrix and therefore the eigenvalues are real and positive, which implies that the condition (6.45) is equivalent to:

$$\lambda(B_{pp}^{-\frac{1}{2}}TB_{xx}^{-1}TB_{pp}^{\frac{1}{2}}) < 2$$

The convergence speed is proportional to the proximity of the eigenvalues $\lambda_i$ to 1. It is clear that if all the eigenvalues of $\hat{M}$ are close to zero the sequence $\{\hat{M}^n\}$ converges to
Chapter 6. Convergence

the zero matrix very fast. In fact we can prove the following theorem:

**Theorem 6.5.1.** If the matrix $T$ in (6.1) has full column rank, then if all the state variables are directly observed and $R = 0$, there exists a cross-covariance matrix $B_{xp}$ such that the 3D-Var converges in 2 iterations.

**Proof.** We will show that if

$$B_{xp}^T B_{xx}^{-1} T = I,$$  \hspace{1cm} (6.53)

then after two iterations the true parameter is assimilated. The evolution of the state for the k-th iteration, with $k > 1$, can be written as follows:

$$
\begin{pmatrix}
  x_a^k \\
p_a^k
\end{pmatrix}
= \tilde{M} 
\begin{pmatrix}
  x_a^{k-1} \\
p_a^{k-1}
\end{pmatrix}
+ \begin{pmatrix}
  K_x \\
K_p
\end{pmatrix} y_k
\hspace{1cm} (6.54)

= \begin{pmatrix}
0 & 0 \\
-B_{xp}^T B_{xx}^{-1} A & 0
\end{pmatrix} 
\begin{pmatrix}
x_a^{k-1} \\
p_a^{k-1}
\end{pmatrix}
+ \begin{pmatrix}
I \\
B_{xp}^T B_{xx}^{-1}
\end{pmatrix} y_k
\hspace{1cm} (6.55)

= \begin{pmatrix}
y_k \\
B_{xp} B_{xx}^{-1} (y_k - Ax^a_{k-1})
\end{pmatrix}.
\hspace{1cm} (6.56)

As expected $x^a_k = y_k = x^l_k$, which does not depend on the iteration $k$ being analysed. This means that for $k > 1$ then also $x^a_{k-1} = x^l_{k-1}$. Rewriting the parameter update:

$$p_k = B_{xp}^T B_{xx}^{-1} (y_k - Ax^a_{k-1})$$  \hspace{1cm} (6.57)

$$= B_{xp}^T B_{xx}^{-1} (x^l_k - Ax^l_{k-1} - Tp^l + Tp^l)$$  \hspace{1cm} (6.58)

$$= B_{xp}^T B_{xx}^{-1} (x^l_k - x^l_{k-1} + Tp^l)$$  \hspace{1cm} (6.59)

$$= B_{xp}^T B_{xx}^{-1} Tp^l$$  \hspace{1cm} (6.60)

$$= p^l.$$  \hspace{1cm} (6.61)

Now it only remains to prove that such a matrix $B_{xp}$ that fulfils (6.53) exists. If we can find a matrix $X \in \mathbb{R}^{\ell \times n}$, such that $(XB_{xx}^{-1} T)^{-1}$ exists, then

$$B_{xp} = X^\top (XB_{xx}^{-1} T)^{-1} \Rightarrow B_{xp}^T B_{xx}^{-1} T = (XB_{xx}^{-1} T)^{-1} X B_{xx}^{-1} T = I.$$  \hspace{1cm} (6.62)

Let $X = Y T^\top$, with $Y \in \mathbb{R}^{\ell \times \ell}$ invertible. Matrix $B_{xx}^{-1}$ is symmetric and positive definite, so we can decompose it as $B_{xx}^{-1} = S^\top S$, with $S \in \mathbb{R}^{n \times n}$ invertible (see e.g. Cholesky factorization [115]). Then:

$$XB_{xx}^{-1} T = YT^\top S^\top ST = Y(ST)^\top ST.$$  \hspace{1cm} (6.63)
Matrices $S$ and $T$ are of full column rank, and the number of rows is larger or equal than the number of columns in both cases. Then $ST$ is the product of two injective mappings, which is also injective and therefore the resulting matrix has also full column rank. To show that $A^\top A$ is invertible for any full column rank matrix $A$ we suppose that there exists a vector $x \neq 0$ such that $A^\top Ax = 0$. Then:

$$A^\top Ax = 0 \Rightarrow x^\top A^\top Ax = 0 \Rightarrow \|Ax\|_2 = 0 \Rightarrow Ax = 0,$$

(6.64)

which is a contradiction because $A$ is a full column rank matrix. Then matrix $(ST)^\top ST$ is invertible and the product of two invertible matrices is also invertible.

As a result, if $B_{xp} = TY^\top (YT^\top B_{xx}^{-1}T)^{-T}$, then 3D-Var recovers the parameter after two assimilations, or in only one assimilation, supposing that the available state estimation is the true state.

In general it is assumed that the error statistics are given, but very often there is little or no information about the parameter error. This analysis provides a guide not only for choosing an appropriate parameter-parameter covariance in the presence of full and accurate observations, but to predict the behaviour of the assimilation method according to the statistical information available.

### 6.5.1 Single parameter

Here we analyse the scenario of a single parameter. In this case $T \in \mathbb{R}^n$ and $B_{pp} \in \mathbb{R}$ is the variance of the parameter, which we will rename as $\sigma_p^2$. Matrix $B_{xp}^\top B_{xx}^{-1}T \in \mathbb{R}$ and the convergence condition (6.45) is reduced to:

$$0 < B_{xp}^\top B_{xx}^{-1}T < 2.$$

(6.65)

#### 6.5.1.1 Single parameter. FDSPC method

If the FDSPC is applied ($B_{xp} = \sigma_p^2 T$), then (6.65) is transformed to:

$$0 < \sigma_p^2 T^\top B_{xx}^{-1}T < 2.$$

(6.66)

The parameter variance $\sigma_p^2$ is positive and $B_{xx}^{-1} > 0$, which implies $T^\top B_{xx}^{-1}T > 0$, so the left inequality always holds.

If we consider $B_{xx} = \sigma_b^2 C$, where $C$ is a matrix, describing the structure of correlations between state variables (which is common for many 1D models), the condition in (6.66)
can be rewritten as:
\[
\frac{\sigma_p^2}{\sigma_0^2} < \frac{2}{T^\top C^{-1} T}.
\] (6.67)

The larger the variance of the state vector, the less precision on the estimation of the parameter is required to ensure convergence. On the other hand, if the state variance is very small, so must be the parameter variance. Though the correlation structure of the errors is normally given, a variance parameters \(\sigma_b^2\) could be used to regulate the size of the covariance. This inequality should be taken into account in such cases.

As we proved before, if \(B_{xp} = T Y^\top (Y T^\top B_{xx}^{-1} T)^{-1}\), the method converges after 2 iterations. This is the case if we take \(\sigma_p^2 = \frac{1}{T^\top B_{xx}^{-1} T}\) (Y=1).

### 6.5.1.2 Single parameter. LRSSC method

If we consider the LRSSC method, instead of (6.66) we have
\[
0 < \sigma_p^2 T^\top \left(\alpha B_{xx} + (1 - \alpha)\sigma_p^2 T T^\top \right)^{-1} T < 2,
\] (6.68)
with \(0 \leq \alpha \leq 1\). For the special case where \(\alpha = 1\) the LRSSC and the FDSPC are equivalent. To analyse the case \(0 < \alpha < 1\) we use the Sherman-Morrison-Woodbury formula (see [56]):
\[
\left(\alpha B_{xx} + (1 - \alpha)\sigma_p^2 T T^\top \right)^{-1} = \left\{\begin{array}{l}
\alpha B_{xx}^{-1} - \frac{\frac{\sigma_p^2}{\alpha^2} B_{xx}^{-1} T}{1 + \frac{\sigma_p^2}{\alpha^2} (1 - \alpha) T^\top B_{xx}^{-1} T} \\
\frac{\sigma_p^2 (1 - \alpha) B_{xx}^{-1} T T^\top B_{xx}^{-1}}{\alpha (\alpha + \frac{\sigma_p^2}{\alpha^2} (1 - \alpha) T^\top B_{xx}^{-1} T)} \\
\frac{\sigma_p^2 (1 - \alpha) B_{xx}^{-1} T T^\top B_{xx}^{-1}}{\alpha (\alpha + \frac{\sigma_p^2}{\alpha^2} (1 - \alpha) T^\top B_{xx}^{-1} T)}
\end{array}\right. \] (6.69)

Now substituting in (6.68):
\[
\sigma_p^2 T \left(\alpha B_{xx}^{-1} + (1 - \alpha)\sigma_p^2 T T^\top \right)^{-1} T = \sigma_p^2 T \left(\frac{B_{xx}^{-1} - (1 - \alpha)\sigma_p^2 T T^\top B_{xx}^{-1}}{\alpha (\alpha + \frac{\sigma_p^2}{\alpha^2} (1 - \alpha) T^\top B_{xx}^{-1} T)} \right) T
\] (6.70)
\[
= \frac{1}{\alpha} \left(\frac{\sigma_p^2 T B_{xx}^{-1} T - (1 - \alpha) (\sigma_p^2 T B_{xx}^{-1} T)^2}{\alpha (1 - \alpha)\sigma_p^2 T B_{xx}^{-1} T}\right)
\] (6.71)
\[
= \frac{1}{\alpha} \left(\frac{\alpha \sigma_p^2 T B_{xx}^{-1} T}{\alpha + (1 - \alpha)\sigma_p^2 T B_{xx}^{-1} T}\right)
\] (6.72)
\[ \sigma_p^2 T^\top \left( \alpha B_{xx} + (1 - \alpha)\sigma_p^2 TT^\top \right)^{-1} T = \frac{\sigma_p^2 T^\top B_{xx}^{-1} T}{\alpha + (1 - \alpha)\sigma_p^2 T^\top B_{xx}^{-1} T}. \] (6.76)

Let us denote \( z = \sigma_p^2 T^\top B_{xx}^{-1} T \), then:

\[ 0 < \frac{z}{\alpha + (1 - \alpha)z} < 2. \] (6.77)

The left inequality always holds, as \( z \) and \( \alpha \) are positive quantities, so

\[ \frac{z}{\alpha + (1 - \alpha)z} < 2 \Leftrightarrow (1 + 2\alpha)z < 2(1 + 2\alpha). \] (6.78)

If \( \alpha \leq \frac{1}{2} \) the left hand side is negative and therefore the inequality holds. For \( \frac{1}{2} < \alpha < 1 \) the convergence is attained when

\[ z < \frac{2\alpha}{1 + 2\alpha} \Leftrightarrow z < \frac{2 - 2(1 - \alpha)}{1 - 2(1 - \alpha)}. \] (6.79)

The right hand side of the inequality is a monotone decreasing rational function of \( \alpha \) with image in \((2, +\infty)\) for \( \frac{1}{2} < \alpha < 1 \), which means that the upper boundary of \( z \) to ensure convergence is less restrictive, compared to Smith’s method (for which it was \( z < 2 \)). This allows us to have more freedom to choose the variance parameters \( \sigma_p^2 \) and \( \sigma_b^2 \). Furthermore our strategy always converges when in the convex combination \( \alpha B_{xx} + (1 - \alpha)T \sigma_p^2 T^\top \) more weight is given to the parameter derivative information than to the constant state-state covariance. This is something to consider in practical applications, when the variance of the observations is small.

### 6.6 Assimilation when the dimension of the observations is smaller than the number of parameters

In real applications, it is not possible to observe all the state variables. In fact, very often we have \( m \ll n \). On the other hand, one of the applications of parameter estimation in DA is to estimate model errors (for example, in the 1D heat equation presented at the end of this chapter). In those cases the number of parameters can be equal to the number of state variables. In this section we address the convergence of 3D-Var for the case \( m < l \).

**Theorem 6.6.1.** Let \( m \) be the dimension of the observations \( \ell \) the dimension of the parameter vector and \( n \) the dimension of the state vector in the linear parameter and state assimilation problem (6.3)-(6.4). If \( \ell > m, \ell = m + r \), then 1 is an eigenvalue of \( \bar{M} \) of multiplicity at least \( r \).
Proof. The iteration matrix \( \tilde{M} \) defined in (6.15) can be rewritten as:

\[
\tilde{M} = M - \begin{pmatrix} B_{xx}S \\ B_{xp}^T S \end{pmatrix} \begin{pmatrix} A & T \end{pmatrix},
\]

(6.80)

where

\[
S := H^T (HB_{xx}H^T + R)^{-1} H
\]

(6.81)

and \( M \) was already defined in (6.3). Besides, we define

\[
G := \begin{pmatrix} B_{xx}S \\ B_{xp}^T S \end{pmatrix} \begin{pmatrix} A & T \end{pmatrix},
\]

(6.82)

as we will be using it later. In the following we analyse the rank of the matrix \( \tilde{M} - I \) to investigate whether 1 is an eigenvalue of \( \tilde{M} \).

\[
\tilde{M} - I = M - I - \begin{pmatrix} B_{xx}S \\ B_{xp}^T S \end{pmatrix} \begin{pmatrix} A & T \end{pmatrix}
\]

(6.83)

\[
= \begin{pmatrix} A - I & T \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} B_{xx}S \begin{pmatrix} A & T \end{pmatrix} \\ B_{xp}^T S \begin{pmatrix} A & T \end{pmatrix} \end{pmatrix}
\]

(6.84)

Analysing the rank of \( B_{xp}^T S \begin{pmatrix} A & T \end{pmatrix} \):

\[
\text{rank}(B_{xp}^T S \begin{pmatrix} A & T \end{pmatrix}) \leq \min(\text{rank}(B_{xp}), \text{rank}(S), \text{rank}(A \ T)) \leq \min(\ell, n, \text{rank}(S)).
\]

(6.85)

(6.86)

Besides,

\[
\text{rank}(S) = \text{rank}(H^T (HB_{xx}H^T + R)^{-1} H) \leq \min(\text{rank}(H), \text{rank}(HB_{xx}H^T + R)) \leq \min(n, m) \leq m,
\]

(6.87)

(6.88)

(6.89)

so

\[
\text{rank}(B_{xp}^T S \begin{pmatrix} A & T \end{pmatrix}) \leq \min(\ell, n, m) \leq m.
\]

(6.90)

Because the last \( \ell = m + r \) rows of \( \tilde{M} - I \) are exactly the rows of \( B_{xp}^T S \begin{pmatrix} A & T \end{pmatrix} \), there are at least \( r \) rows that are linearly dependent. This means that \( \text{dim ker}(\tilde{M} - I) \geq r \) and therefore 1 is an eigenvalue of multiplicity at least \( r \).

\[
\square
\]

Corollary 6.6.2. Let \( m \) be the dimension of the observations, \( \ell \) the dimension of the parameter vector and \( n \) the dimension of the state in (6.3)-(6.4). If \( \ell > m, \ell = m + r, \)
then 3D-Var does not converge.

Proof. Direct consequence of Theorem 6.6.1

3D-Var-like methods have, therefore, limitations when an estimation of the model state variables and the model bias via state-augmentation is desired, at least for the linear case. In Section 6.9.6 we show experiments which confirm this affirmation.

6.7 Single observation and single parameter

Now we investigate the convergence when only a minimum of information is available. Let us consider that only the j-th state variable is observed. If \( \ell > 1 \), we already proved that the method does not converge. We suppose then that there is only one parameter to be estimated. The linear operator \( H \) is just the canonical vector \( e_j^\top \) and the variance of the observation error will be denoted by \( \sigma_o^2 \). Substituting our special \( H \) and \( R \) into (6.81), we have:

\[
S = e_j(e_j^\top B_{xx} e_j + \sigma_o^2)^{-1} e_j^\top = \frac{1}{\sigma_x^2 + \sigma_o^2} e_j e_j^\top.
\]  

(6.91)

Here, \( \sigma_x^2 \) is the variance of the observed state. This is a square matrix, where the only entry different from 0 is \( S_{jj} = \frac{1}{\sigma_x^2 + \sigma_o^2} \). The parameter-parameter error covariance is a positive real number and we rename it as \( \sigma_p^2 \). If we denote the j-th row of a matrix \( X \) as \( X_j \) (\( T_j \) is a scalar) we can express the matrix \( G \) defined in (6.82) as:

\[
G = \frac{1}{\sigma_x^2 + \sigma_o^2} \begin{pmatrix} B_{xx} e_j e_j^\top \\ B_{xp} e_j e_j^\top \end{pmatrix} \begin{pmatrix} A \\ T_j \end{pmatrix}
\]  

(6.92)

\[
= \frac{1}{\sigma_x^2 + \sigma_o^2} \begin{pmatrix} B_{xx} e_j \\ B_{xp} e_j \end{pmatrix} e_j^\top \begin{pmatrix} A \\ T_j \end{pmatrix}
\]  

(6.93)

\[
= \frac{1}{\sigma_x^2 + \sigma_o^2} \begin{pmatrix} B_{xx} e_j \\ B_{xp} e_j \end{pmatrix} \begin{pmatrix} e_j^\top A \\ e_j^\top T \end{pmatrix} = \frac{1}{\sigma_x^2 + \sigma_o^2} \begin{pmatrix} (B_{xx})_{jj} \\ B_{xp} \end{pmatrix} \begin{pmatrix} A_j \\ T_j \end{pmatrix}.
\]  

(6.94)

For ease of notation, in the rest of this section, we will denote the state-state covariance \( B_{xx} \) simply as \( B \). It is important not to confuse it with the full augmented state covariance.

6.7.1 Necessary condition for convergence

**Theorem 6.7.1.** If the 3D-Var, applied to the linear joint state-parameter DA problem (6.1)-(6.2) when only the j-th state variable with variance \( \sigma_x^2 \) is observed, converges, then
it holds:
\[
\left| \text{tr}(A) + 1 - \frac{1}{\sigma_x^2 + \sigma_o^2} \left( A_j B_j^T + T_j B_{xpj} \right) \right| < n + 1.
\] (6.95)

**Proof.** Let \( \lambda_i, i = 1, \ldots, n + 1 \) be the eigenvalues of \( \tilde{M} \) and \( \lambda^* = \arg \max |\lambda_i| \)
\[
\frac{|\text{tr}(\tilde{M})|}{n + 1} = \frac{\sum_{i=1}^{n+1} |\lambda_i|}{n + 1} \leq \frac{(n + 1)|\lambda^*|}{n + 1} \leq |\lambda^*| \] (6.96)

The condition \(|\lambda^*| < 1\), necessary for the convergence, is only satisfied when
\[
\frac{|\text{tr}(\tilde{M})|}{n + 1} < 1. \] (6.97)

Substituting
\[
\text{tr}(\tilde{M}) = \text{tr}(M) - \text{tr}(G) = \text{tr}(A) + 1 - \frac{1}{\sigma_x^2 + \sigma_o^2} \left( \text{tr}(B_j^T A_j) + \text{tr}(B_{xpj} T_j) \right) \] (6.98)
\[
= \text{tr}(A) + 1 - \frac{1}{\sigma_x^2 + \sigma_o^2} \left( A_j B_j^T + B_{xpj} T_j \right) \] (6.99)
\[
= \text{tr}(A) + 1 - \frac{1}{\sigma_x^2 + \sigma_o^2} \left( A_j B_j^T + B_{xpj} T_j \right) \] (6.100)

into (6.97):
\[
\frac{|\text{tr}(\tilde{M})|}{n + 1} \leq 1 \iff |\text{tr}(A) + 1 - \frac{1}{\sigma_x^2 + \sigma_o^2} \left( A_j B_j^T + B_{xpj} T_j \right) | \leq n + 1 \] (6.101)

is a necessary condition for convergence.

If the variances of the observed variable and the observations are very small, the only possibility to achieve convergence is that \( B_{xpj} T_j \approx -A_j B_j^T \). On the other hand, if the parameter does not play a role in the model update of the observed variable (\( T_j = 0 \)), then it is possible to find a necessary lower boundary for the sum \( \sigma_x^2 + \sigma_o^2 \) for the method to converge. In the following we explore this theorem in more detail for both, FDSPC and LRSSC methods. Because of the special structure of the parameter-state cross-covariance matrix \( B_{xp} = \sigma_p^2 T \), conditions on \( B_{xp} \) can be translated into conditions on the parameter variance \( \sigma_p^2 \).

**FDSPC method**

If the FDSPC is considered, after substituting \( B_{xp} = \sigma_p^2 T \) in (6.101) the following corollary can be directly obtained:

**Corollary 6.7.2.** If the FDSPC method, applied to the linear joint state-parameter DA problem (6.1)-(6.2) when only the \( j \)-th state variable with variance \( \sigma_x^2 \) is observed,
converges, then it holds:

$$\left| \text{tr}(A) + 1 - \frac{1}{\sigma^2 + \sigma_0^2} \left( A_j B_j^\top + \sigma^2 T_j^2 \right) \right| < n + 1. \tag{6.102}$$

**Proof.** Notice that $B_{xp} T_j = (\sigma_p^2 T_j) T_j = \sigma_p^2 T_j^2 T_j = \sigma_p^2 T_j^2$. The corollary is a direct result of substituting this product into (6.101).

After some manipulation (6.102) is equivalent to

$$\frac{(\text{tr}(A) - n)}{T_j^2} \left( \sigma_p^2 + \sigma_0^2 \right) - \frac{A_j B_j^\top}{T_j^2} < \sigma_p^2 < \frac{(\text{tr}(A) + n + 2)}{T_j^2} \left( \sigma_p^2 + \sigma_0^2 \right) - \frac{A_j B_j^\top}{T_j^2}. \tag{6.103}$$

The inequalities in (6.103) show how the admissible boundaries of $\sigma_p^2$ for convergence change for different values of $\sigma_p^2 + \sigma_0^2$. Let us denote $z = \sigma_p^2 + \sigma_0^2$. As a function of $z$ these boundaries are straight lines intersecting at $(0, -\frac{A_j B_j^\top}{T_j^2})$. A necessary condition for convergence is that $(z, \sigma_p^2)$ lie in the intersection of the region between the lines and the upper-right quadrant of the Cartesian Plane. Figure 6.2 and Table 6.1 explain all the possible cases.

![Figure 6.2](image)

**Figure 6.2:** Regions of admissible values of the parameter variance $\sigma_p^2$, depending on the values of $z$.

<table>
<thead>
<tr>
<th>$A_j B_j^\top$</th>
<th>$\text{tr}(A) &lt; -n - 2$</th>
<th>$-n - 2 \leq \text{tr}(A) &lt; n$</th>
<th>$\text{tr}(A) \geq n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt; 0$</td>
<td>a</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td>$&gt; 0$</td>
<td>d</td>
<td>e</td>
<td>f</td>
</tr>
</tbody>
</table>

**Table 6.1:** Conditions for cases shown in Figure 6.2.
(a) If \( z \geq \frac{A_jB_j^T}{\text{tr}(A)+n+2} \), the method does not converge. Otherwise \( \sigma_p^2 \) is upper bounded, with the boundary inversely proportional to the size of \( z \). If \( z < \frac{A_jB_j^T}{\text{tr}(A)+n} \), it is also lower bounded, with the boundary inversely proportional to the size of \( z \) as well.

(b) The value of \( \sigma_p^2 \) is upper bounded, with this boundary proportional to the size of \( z \). In the special case where \( \text{tr}(A) = -n-2 \) then \(-A_jB_j^T\) is an upper boundary for \( \sigma_p^2 \). For \( 0 < z \leq \frac{A_jB_j^T}{\text{tr}(A)+n} \) the variance \( \sigma_p^2 \) is lower bounded with the boundary inversely proportional to the size of \( z \).

(c) \( \sigma_p^2 \) is upper bounded, with the boundary proportional to the size of \( z \). If \( \text{tr}(A) = n \) then \(-A_jB_j^T\) is a lower boundary for \( \sigma_p^2 \), otherwise the lower boundary is proportional to the size of \( z \).

(d) The method does not converge.

(e) To achieve convergence it is necessary that \( z > \frac{A_jB_j^T}{\text{tr}(A)+n+2} \). \( \sigma_p^2 \) is upper bounded, with this boundary proportional to the size of \( z \). In the special case where \( \text{tr}(A) = -n-2 \) the method does not converge.

(f) To achieve convergence it is necessary that \( z > \frac{A_jB_j^T}{\text{tr}(A)+n+2} \). \( \sigma_p^2 \) is upper bounded, with both boundaries proportional to the size of \( z \). If \( \text{tr}(A) > n \) and \( z > \frac{A_jB_j^T}{\text{tr}(A)-n} \), \( \sigma_p^2 \) is lower bounded, with the boundary proportional to the size of \( z \).

The special case when \( A_j \perp B_j \) can be analysed as the case \( A_jB_j^T < 0, A_jB_j^T \approx 0 \).

If \( \sigma_x^2 \) and \( \sigma_o^2 \) are very small there is only a very small interval of admissible values of \( \sigma_p \), if at all, for which the necessary condition for convergence is fulfilled. This interval length goes to zero as the sum of the variances goes to zero. This means that the more precise our estimation of the state and the observations, the more precise the estimation of our parameter must be.

**Our approach**

An equivalent result for our LRSSC is:

**Corollary 6.7.3.** If the LRSSC method, applied to the linear joint state-parameter DA problem (6.1)-(6.2) when only the \( j \)-th state variable with variance \( \sigma_x^2 \) is observed, then it holds:

\[
\left| \text{tr}(A) + 1 - \frac{\alpha A_jB_j^T + (1-\alpha)\sigma_x^2T_jA_jT + \sigma_p^2T_j^2}{\alpha \sigma_x^2 + (1-\alpha)\sigma_p^2T_j^2} \right| < n + 1.
\] (6.104)
Proof. Now the state-state covariance matrix is $B' = \alpha B + (1 - \alpha)\sigma_p^2 TT^\top$. On the other hand, to compute the variance $\sigma_x^2$ of the observed variable:

\[
\sigma_x^2 = e_j^\top \left( \alpha B + (1 - \alpha)\sigma_p^2 TT^\top \right) e_j = \alpha e_j^\top B_{xx} e_j + (1 - \alpha)\sigma_p^2 T_j^2.
\]

Computing $B'_j$:

\[
B'_j = e_j^\top \left( \alpha B + (1 - \alpha)\sigma_p^2 TT^\top \right) = \alpha B_j^\top + (1 - \alpha)\sigma_p^2 T_j T^\top.
\]

The rest of the proof is a direct substitution into (6.95).

6.7.2 Sufficient condition for convergence

Necessary conditions are a useful tool to discard possible values of the parameter variance $\sigma_p^2$, for which the method diverges. In the following we present a sufficient condition for convergence.

**Theorem 6.7.4.** If for the linear joint state-parameter DA problem (6.1)-(6.2), when only the $j$-th state variable is observed, it holds:

\[
\frac{q}{\sqrt{n}} + \frac{\sqrt{\|A\|_F^2 + \|T\|_F^2 + l + \frac{\|A_j\|^2 + T_j^2}{\sigma_x^2 + \sigma_o^2} + \frac{2(B_j A^\top T_j + B_{xp} T_j)}{\sigma_x^2 + \sigma_o^2}}}{n} < 1,
\]

where

\[
q = \text{tr}(A) + 1 - \frac{1}{\sigma_x^2 + \sigma_o^2} (A_j B_j^\top + B_{xp} T_j),
\]

then 3D-Var applied to the problem converges.

**Proof.** We will use the following theorem found in [134]:

**Theorem 6.7.5.** Let $A$ be an $n \times n$ complex matrix, and let

\[
m = \frac{\text{tr}(A)}{n}, \quad s_a^2 = \frac{\text{tr}(A^* A)}{n} - |m|^2.
\]

Then

\[
|m| \leq |\lambda_1| \leq |m| + s_a \sqrt{n - 1},
\]
where $\lambda_1 = \max \{|\lambda_i|\}$.

Substituting $A$ by $\tilde{M}$ into the right inequality in (6.111):

$$|\lambda^*| < \frac{|\text{tr}(\tilde{M})|}{n + l} + \sqrt{n + l - 1}S_a, \quad S_a^2 = \frac{\text{tr}(\tilde{M}^\top \tilde{M})}{n + l} - \left(\frac{\text{tr}(\tilde{M})}{n + l}\right)^2$$

(6.112)

A sufficient condition then to ensure convergence is:

$$\frac{|\text{tr}(\tilde{M})|}{n + l} + \sqrt{n + l - 1}S_a < 1$$

(6.113)

Throughout the proof we will use some properties of the trace of a matrix: Let $c$ be a scalar, $x$ and $y$ two vectors of dimension $n$ and $A$, $B$ are matrices of dimension $n \times n$ then, the following four properties hold:

$$\text{tr}(xy^\top) = \sum_{i=1}^{n} x_i y_i = y^\top x$$

(6.114)

$$\text{tr}(A + B) = \sum_{i=1}^{n} (A_{ii} + B_{ii}) = \sum_{i=1}^{n} A_{ii} + \sum_{i=1}^{n} B_{ii} = \text{tr}(A) + \text{tr}(B)$$

(6.115)

$$\text{tr}(cA) = \sum_{i=1}^{n} cA_{ii} = c \sum_{i=1}^{n} A_{ii} = c \text{tr}(A)$$

(6.116)

$$\text{tr}(A^\top A) = \text{tr}(AA^\top) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij}^2 = \|A\|_F^2$$

(6.117)

Computing the trace of $\tilde{M}^\top \tilde{M}$:

$$\tilde{M}^\top \tilde{M} = \left(M - \frac{1}{\sigma_x^2 + \sigma_o^2} uw^\top\right)^\top \left(M - \frac{1}{\sigma_x^2 + \sigma_o^2} uw^\top\right)$$

(6.118)

$$= M^\top M + \frac{1}{(\sigma_x^2 + \sigma_o^2)^2} uw^\top uw^\top - \frac{2}{\sigma_x^2 + \sigma_o^2} (wu^\top M + (wu^\top M)^\top)$$

(6.119)

$$M^\top M = \begin{pmatrix} A^\top & 0 \\ T^\top & I \end{pmatrix} \begin{pmatrix} A & T^\top \\ 0 & I \end{pmatrix} = \begin{pmatrix} A^\top A & A^\top T \\ T^\top A & T^\top T + I \end{pmatrix}$$

(6.120)

$$\text{tr}(\tilde{M}^\top \tilde{M}) = \text{tr}(M^\top M) + \frac{\text{tr}(wu^\top uw^\top)}{(\sigma_x^2 + \sigma_o^2)^2} - \frac{2\text{tr}(wu^\top M)}{\sigma_x^2 + \sigma_o^2}$$

(6.121)

$$= \text{tr}(A^\top A) + \text{tr}(T^\top T) + l + \frac{\|u\|_2^2 \text{tr}(wu^\top)}{(\sigma_x^2 + \sigma_o^2)^2} - \frac{2\text{tr}(wu^\top M)}{\sigma_x^2 + \sigma_o^2}$$

(6.122)

$$= \|A\|_F^2 + \|T\|_F^2 + l + \frac{\|u\|_2^2 \|w\|_2^2}{(\sigma_x^2 + \sigma_o^2)^2} - \frac{2u^\top M w}{\sigma_x^2 + \sigma_o^2}$$

(6.123)
Substituting vectors $u$ and $w$:

$$
\text{tr}(\tilde{M}^\top \tilde{M}) = \|A\|_F^2 + \|T\|_F^2 + l + \frac{(||A_j||_F^2 + T_j^2)(||B_j||_F^2 + B_{xpj}^2)}{(\sigma_x^2 + \sigma_o^2)^2} - \frac{2(A_j B_T^j + T_j B_j T + B_{xpj} T_j)}{\sigma_x^2 + \sigma_o^2} \sigma_x^2 + \sigma_o^2 (6.124)
$$

The rest of the proof is a straightforward substitution of $\text{tr}(\tilde{M})$ and $\text{tr}(\tilde{M}^\top \tilde{M})$ into (6.113).

This condition seems to be very restrictive, especially when $1 - \frac{|q|}{n+1} \approx 0$, but on the other hand, very easy to evaluate. It is possible to adapt this sufficient condition to any of the two approaches studied with the proper substitutions.

### 6.8 Precise background state estimation

In this section we consider the case where the background state estimation is very accurate, i.e., $B_{xx} \approx 0$. In the following we analyse the limit situation in which $B_{xx} = 0$. In this case the assimilation of observations will not alter the state forecast. The question is whether the parameters, which do change with each assimilation depending on the observations, converge to the truth. Because the eigenvalues of $\tilde{M}$ can be considered as a continuous function of the covariance inflation factor $\sigma_b^2$ in $B_{xx} = \sigma_b^2 C$, we expect that the conditions found here can be also useful if the condition $\sigma_b^2 = 0$ is relaxed to $\sigma_b^2 \approx 0$.

According to (6.13) and (6.15):

$$
K = \begin{pmatrix}
0 H^\top (H0H^\top + R)^{-1} \\
B_{xp} H^\top (H0H^\top + R)^{-1}
\end{pmatrix} = \begin{pmatrix}
0 \\
B_{xp} H^\top R^{-1}
\end{pmatrix}, \quad (6.125)
$$

$$
\tilde{M} = \begin{pmatrix}
A & T \\
-B_{xp} H^\top R^{-1} A & I - B_{xp} H^\top R^{-1} H T
\end{pmatrix}. \quad (6.126)
$$

### 6.8.1 Necessary conditions for convergence

**Theorem 6.8.1.** Let $(\lambda, v)$ be a pair of eigenvalue and eigenvector of $A$, and let $|\lambda| \geq 1$. If 3D-Var applied to the linear joint state-parameter DA problem (6.1)-(6.2) with $B_{xx} = 0$ converges then $v \notin \ker(B_{xp} H^\top R^{-1} H)$. 

Proof. Let us suppose that \( v \in \ker(B_{xp}^T H^T R^{-1} H) \). We have then

\[
\begin{align*}
Av &= \lambda v, \quad \text{(6.127)} \\
B_{xp}^T H^T R^{-1} Hv &= 0. \quad \text{(6.128)}
\end{align*}
\]

We will use now (6.127) to expand (6.128):

\[
\begin{align*}
-B_{xp}^T H^T R^{-1} H \lambda v &= 0 \quad \text{(6.129)} \\
-B_{xp}^T H^T R^{-1} H Av &= 0 \quad \text{(6.130)} \\
-B_{xp}^T H^T R^{-1} H Av + (I - B_{xp}^T H^T R^{-1} HT)0 &= \lambda 0 \quad \text{(6.131)} \\
Av + T0 &= \lambda v. \quad \text{(6.132)}
\end{align*}
\]

(6.131) and (6.132) are equivalent to

\[
\tilde{M}\begin{pmatrix} v \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} v \\ 0 \end{pmatrix}. \quad \text{(6.133)}
\]

As \( v \neq 0 \), then \( \lambda \) is also an eigenvalue of \( \tilde{M} \), and because \( |\lambda| \geq 1 \) the method diverges. \( \square \)

**Theorem 6.8.2.** If 3D-Var applied to the linear joint state-parameter DA problem (6.1)-(6.2) with \( B_{xx} = 0 \) converges, then \( \forall v \in \ker(B_{xp}^T H^T R^{-1} H) , v \neq 0 \) it holds \((A - I)v \notin \text{Im} T\).

Proof. Let's suppose that \( \exists v_x, v_p, v_x \in \mathbb{R}^n, v_x \neq 0 \) and \( v_p \in \mathbb{R}^l \) such that:

\[
\begin{align*}
B_{xp}^T H^T R^{-1} H(-v_x) &= 0, \quad \text{(6.134)} \\
Tv_p &= (A - I)(-v_x) \iff Av_x + T v_p = v_x. \quad \text{(6.135)}
\end{align*}
\]

Manipulating the system (6.134)-(6.135):

\[
\begin{align*}
B_{xp}^T H^T R^{-1} H(v_x) &= 0 \quad \text{(6.136)} \\
v_p + B_{xp}^T H^T R^{-1} H v_x &= v_p \quad \text{(6.137)} \\
B_{xp}^T H^T R^{-1} H v_x &= B_{xp}^T H^T R^{-1} H(A v_x + T v_p) \quad \text{(6.138)}
\end{align*}
\]

and after subtracting (6.137) and (6.138):

\[
v_p = -B_{xp}^T H^T R^{-1} H Av_x + (I - B_{xp}^T H^T R^{-1} HT)v_p. \quad \text{(6.139)}
\]
(6.135) and (6.139) are equivalent to
\[ \tilde{M} \begin{pmatrix} v_x \\ v_p \end{pmatrix} = \begin{pmatrix} v_x \\ v_p \end{pmatrix}, \]
(6.140)
which means that 1 is an eigenvalue of \( \tilde{M} \) and therefore there is no convergence.

Both necessary conditions involve the nullspace of the matrix \( B_{zp}^T H^T R^{-1} H \). For the specific case of the FDSPC and the LRSSC methods \( B_{zp}^T H^T R^{-1} H = B_{pp}^T H^T R^{-1} H \). Because \( B_{pp} \) is invertible the nullspace of the aforementioned matrix is equivalent to the nullspace of the matrix \( T^T H^T R^{-1} H \). If additionally we suppose that \( H \) is a full rank matrix, square matrix (i.e., all the states are observed) then the nullspace of \( T^T \) plays an important role in the theorems.

### 6.9 Numerical experiments.

In this section we test the theoretical results previously shown with the aid of the linear 1-D heat equation model. In most of the experiments we apply the simpler FDSPC. In cases where it is relevant, we also show the results when the LRSSC is applied.

#### 6.9.1 1-D heat equation

The heat equation models the distribution of heat on a certain space domain over time. The 1-D case considers a one dimensional space domain (for example, a metal rod). We assume that our space domain is \( z \in [0, 1] \) and the time domain \( t \in [0, T] \). If we add a heat source of intensity \( \frac{10}{3} \) at \( z = \frac{1}{4} \), the mathematical model describing the change of temperature is
\[ v_t = \kappa v_{zz} + \frac{10}{3} \delta(z - \frac{1}{4}), \]
(6.141)
with initial conditions
\[ v(z, 0) = \alpha(z) \]
(6.142)
and boundary conditions
\[ v(0, t) = v(1, t) = 0. \]
(6.143)
Here \( \delta \) is the Dirac delta function, \( v(z, t) \) represents the temperature of the rod at time \( t \) and position \( z \), and \( \kappa \) is the diffusion constant. For the discretization of (6.141) we replace the time derivative by a forward finite difference scheme and the second order
space derivative by a central differencing scheme, obtaining:

\[ x_j^{k+1} - x_j^k = \kappa \frac{\Delta t}{\Delta z^2} (x_{j-1}^k - 2x_j^k + x_{j+1}^k) + s_j \Delta t \]  \hspace{1cm} (6.144)

with initial conditions

\[ x_0^0 = \alpha (j \Delta z) \]  \hspace{1cm} (6.145)

and boundary conditions

\[ x_0^k = x_J^k = 0. \]  \hspace{1cm} (6.146)

Here \( x_j^k \approx v(j \Delta z, k \Delta t) \) for \( j = 0, 1, \ldots, J \) and \( k = 0, 1, \ldots, N \) and \( \Delta z = \frac{1}{J}, \Delta t = \frac{T}{N}, \) with \( s_j \) being the discretization

\[ s_j = \begin{cases} 
\frac{10}{3\Delta z} & \text{if } j = \frac{J}{4} \\
0 & \text{otherwise}
\end{cases} \]  \hspace{1cm} (6.147)

Here we assume that \( J \) is divisible by 4. The discrete model (6.144) can be rewritten as

\[ x_k = Ax_{k-1} + Ts, \]  \hspace{1cm} (6.148)

where \( T = I_{(J-1)\times(J-1)}, x_k = (x_1^k, x_2^k, \ldots, x_{J-1}^k) \) and \( A \in \mathbb{R}^{(J-1)\times(J-1)} \) is defined as:

\[ A = \begin{pmatrix}
1 - 2\mu & \mu \\
\mu & 1 - 2\mu & \mu \\
& \ddots & \ddots & \ddots \\
& & \mu & 1 - 2\mu & \mu \\
& & & \mu & 1 - 2\mu
\end{pmatrix}, \]  \hspace{1cm} (6.149)

where \( \mu = \kappa \frac{\Delta t}{\Delta z^2} \). It is not difficult to prove that the system is stable (see [27]) if \( 0 \leq \mu \leq 0.5 \).

### 6.9.2 Experiment setup

We performed twin experiments using the perfect model (6.148) to generate the observations. For the forecast model we assume we don’t have any information about the heat source, so a parameter vector \( p \) is introduced

\[ x_k = Ax_{k-1} + p, \]  \hspace{1cm} (6.150)

to account for the constant bias error of the imperfect model \( x_k = Ax_{k-1} \). The goal is to obtain a good estimation of both the state and the parameter.
The space domain $[0, 1]$ is divided into 16 equidistant intervals, which implies $J = 16$ (the dimension of the state is $n = 15$) and $\Delta z = \frac{1}{16}$. The assimilation time window is $[0, \frac{1}{2}]$, subdivided into $N = 50$ time steps ($\Delta t = \frac{1}{100}$). Observations are available at every time step.

We consider that the observations are uncorrelated and share the same variance $\sigma_o^2$, so $R_k = \sigma_o^2 I$, $i = 0, 1, \ldots$. The state-state error covariance has the form

$$B_{xx} = \sigma_b^2 C,$$  

(6.151)

where $\sigma_b^2$ is the background error variance and $C$ is the Markov Correlation Matrix, defined using the correlation function

$$c_{ij} = e^{-\frac{|i-j|\Delta z}{L}},$$  

(6.152)

where $L$ is the background correlation length scale. It is assumed that the correlation between two state variables depends only on the spatial distance between them and that it decreases with the separation. In our experiments $\sigma_b^2 = 0.1$, $L = \frac{3}{4}\Delta z = \frac{15}{64}$. The diffusion constant is $\kappa = 0.1$. These values make $\mu = 0.256$. For both the true and background initial state, the temperature is homogeneous, set to 1 along the rod.

In the experiment, two different parameter-parameter error covariances $B_{pp}$ were considered. In one case it is assumed that the parameter errors are uncorrelated and with similar variance $\sigma_p^2$. The covariance is then $B_{pp} = \sigma_p^2 I$. For the second case we assume that the errors are correlated and that the correlation is similar to the one between the corresponding state errors, so $B_{pp} = \sigma_p^2 C$. Finally, because in this case $T = I$ the state-parameter error covariance is $B_{xp} = N_k B_{pp} = TB_{pp} = B_{pp}$.

In the first group of experiments all the state variables are directly observed at every time step. The observation operator is therefore $H = I$. The goal is to check how the conditions found in section 6.5 apply to the case where $R$ is not zero, but $R = \sigma_o^2 I$, with $\sigma_o^2$ considerably small. Because the eigenvalues of $\tilde{M}$ depend continuously on the value of $\sigma_o^2$ we expect that this conditions are also useful for small observation variances.

### 6.9.3 FDSPC for uncorrelated parameter errors

We assume that $B_{pp} = \sigma_p^2 I = 0.05 \times I$, so there is no correlation between parameter errors. The variance of observation errors is $\sigma_o^2 = 1e - 06$. The structure of the matrices $B$ and $\tilde{M}$ are shown in Figure 6.3.
Chapter 6. Convergence

After assimilating 40 observations (see Figure 6.4) the state obtained approximates the true state extremely well. Nevertheless, if we let the model run freely for another 40 time steps without any assimilation, the results are completely wrong. This is due to the fact that the parameter vector, which accounts for the bias in the model, is wrongly estimated. The peaks over the zero temperature line reveal that in the corresponding position of the assimilated parameter, the entry is positive (heat source), while those below zero corresponds to negative entries (cold source). The true parameter vector is zero everywhere, except at the fourth position.

From Section 6.5 we know that for $\sigma_{o}^{2}$ small enough, the method converges if and only if all the eigenvalues of matrix $B_{pp} T^\top B_{xx}^{-1} T$ are inside the unit circle with center in $(1, 0)$. In this case

$$B_{pp} T^\top B_{xx}^{-1} T = \frac{\sigma_{p}^{2}}{\sigma_{b}^{2}} C^{-1}. \quad (6.153)$$
For ease of notation we denote
\[ D := \frac{\sigma_p^2}{\sigma_b^2} C^{-1}. \] (6.154)

Figure 6.5 shows the eigenvalues of \( D \) and \( \tilde{M} \).

Some eigenvalues of \( \tilde{M} \) lie outside the unit circle with center in \((0, 0)\), which is equivalent to the statement that some eigenvalues of \( D \) lie outside the unit circle with center in \((1, 0)\). Therefore the method does not converge. If \( \sigma_p^2 = 25 \times 10^{-3} \), that is, half of its previous size, the eigenvalues of \( D \) are also halved. From 6.5 it is clear that they all lie inside the corresponding unit circle. Figure 6.6 shows the results after 40 assimilation steps and subsequent 40 forecast steps without assimilation.

**Figure 6.5:** Eigenvalues and unit circles.

**Figure 6.6:** Assimilated state for \( \sigma_p^2 = 25 \times 10^{-3} \). Left upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state(red dashed line) after \( N = 40 \) time steps. Left lower: true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. Right: state errors after \( \frac{N}{2} \) assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).
This time the state estimation, even after performing 40 forecast steps without assimilation, is accurate. The reason is that the parameters (model bias) are also well approximated. Figure 6.7 shows the parameter estimation along the assimilation time window.

All the entries of the parameter vector are close to zero, except the fourth one, whose value approximates accurately the intensity of the heat source.

**Convergence speed and condition number of $C$**

From Figure 6.7 it can be noticed that, though the parameter vector converges to the truth, the convergence speed is not particularly high. In Section 6.5 we showed that

$$\text{eig}(\tilde{M}) = 1 - \text{eig}(D).$$

(6.155)

The convergence speed is inversely proportional to the absolute value of the eigenvalues of $\tilde{M}$. If they are all close to 0, the convergence will be fast, while if any of the eigenvalues is near the unit circle with center in $(0,0)$, the convergence speed will be low. This is equivalent to the statement that, if all the eigenvalues of $D$ are close to 1, the convergence will be fast, while if any of the eigenvalues is near the unit circle with center in $(1,0)$, the convergence speed will be low.

**Remark 6.9.1.** When the parameter vector represents a model bias and therefore $T = I$, then in case that the parameters are uncorrelated with similar variance the convergence
speed of FDSPC is inversely proportional to the condition number of state-state covariance matrix.

This can be directly checked. Because $C$ is symmetric:

$$
\text{cond}(D) = \text{cond} \left( \frac{\sigma_p^2}{\sigma_b^2} C^{-1} \right) = \text{cond}(C^{-1}) = \text{cond}(C) = \frac{|\lambda_{\text{max}}|}{|\lambda_{\text{min}}|},
$$

(6.156)

where $\lambda_{\text{max}}$ and $\lambda_{\text{min}}$ are the eigenvalues of $D$ with largest and smallest absolute value, respectively. Because any complex number inside the unit circle with center in $(1, 0)$ has absolute value smaller than 2, in case of convergence

$$
|\lambda_{\text{min}}| = \frac{|\lambda_{\text{max}}|}{\text{cond}(C)} < \frac{2}{\text{cond}(C)},
$$

(6.157)

must be fulfilled. If the condition number of $C$ is large, it means that there will be at least one eigenvalue of $D$ very close to zero and therefore to the unit circle with center in $(1, 0)$, causing a slow convergence speed. On the other hand, if the condition of $C$ is small, it is always possible to find some value of $\sigma_p^2$ for which both extreme eigenvalues are equally distant to $(1, 0)$ and close enough to ensure fast convergence. For example, if the eigenvalues are all real, as in this case, it would be interesting to know for which value of $\sigma_p^2$ the equality

$$
\lambda_{\text{max}} - 1 = 1 - \lambda_{\text{min}}
$$

(6.158)

holds.

$$
\begin{align*}
\lambda_{\text{max}} - 1 &= 1 - \lambda_{\text{min}} \quad \text{(6.159)} \\
\lambda_{\text{max}} + \lambda_{\text{min}} &= 2 \quad \text{(6.160)} \\
\text{cond}(C)\lambda_{\text{min}} + \lambda_{\text{min}} &= 2 \quad \text{(6.161)} \\
\lambda_{\text{min}} &= \frac{2}{\text{cond}(C) + 1}. \quad \text{(6.162)}
\end{align*}
$$

If $\text{cond}(C) \approx 1$, then $\lambda_{\text{min}} \approx \lambda_{\text{max}} \approx 1$. Moreover

$$
\lambda_{\text{min}}(D) = \lambda_{\text{min}} \left( \frac{\sigma_p^2}{\sigma_b^2} C^{-1} \right)
$$

(6.163)

$$
= \frac{\sigma_p^2}{\sigma_b^2} \lambda_{\text{min}}(C^{-1})
$$

(6.164)

$$
= \frac{\sigma_p^2}{\sigma_b^2 \lambda_{\text{max}}(C)}.
$$

(6.165)
so

\[
\frac{2}{\text{cond}(C) + 1} = \frac{\sigma_p^2}{\sigma_b^2 \lambda_{\text{max}}(C)}
\]

(6.166)

\[
\sigma_p^2 = \frac{2\sigma_b^2 \lambda_{\text{max}}(C)}{\text{cond}(C) + 1}
\]

(6.167)

\[
= \frac{2\sigma_b^2 \lambda_{\text{max}}(C) \lambda_{\text{min}}(C)}{\lambda_{\text{max}}(C) + \lambda_{\text{min}}(C)}.
\]

(6.168)

Equation (6.168) represents the optimal parameter variance concerning convergence speed. To illustrate how the condition number of \( C \) affects the convergence speed, two experiments are performed. In one the correlation length scale \( L \) is increased, which increases the condition number of \( C \). In the second one \( L \) is decreased, causing the opposite effect. In both cases we will use the optimal value found in (6.168), which will be denoted as \( \sigma_p^{2*} \).

Figure 6.8: Assimilated state for \( L = \frac{n}{2} \Delta z \) and \( \sigma_p^2 = \sigma_p^{2*} \). (a) upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state (red dashed line) after \( N = 40 \) time steps. (a) lower: the true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. (b) state errors after \( N/2 \) assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).

Figure 6.8 shows the results after doubling the correlation length scale. The condition number of \( C \) is increased from its previous value of 43.8 to 128.5, affecting the convergence speed. It takes longer for the parameter to approach the true value and with only 40 assimilation steps, the estimation obtained is obviously worse than the one using \( L = \frac{n}{4} \Delta z \).

Setting \( L = \frac{n}{8} \Delta z \) causes the opposite effect: the condition number of \( C \) is decreased to 13.3 and this speeds up the convergence of the method as it is shown in Figure 6.10 and 6.11. The state error at the end of the time window is two orders of magnitude smaller than with a doubled correlation scale length. The error after the last assimilation is comparable to the one after 40 time steps without new observations to assimilate.
Convergence

Figure 6.9: Parameter estimation along the time window for $L = \frac{n}{s} \Delta z$, $\sigma^2_p = \sigma^2_{p*}$ and uncorrelated parameters. Each plot accounts for an entry of the parameter vector. True parameter (black) and estimated parameter (red).

Figure 6.10: Assimilated state for $L = \frac{n}{s} \Delta z$ and $\sigma^2_p = \sigma^2_{p*}$. Left upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state (red dashed line) after $N = 40$ time steps. Left lower: the true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. Right: state errors after $N$ assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).

Convergence for larger $\sigma^2_o$

A last experiment was performed setting $\sigma^2_o = 0.1$ and $L = \frac{n}{s} \Delta z$ for $\sigma^2_p = \sigma^2_{p*}$, to investigate whether the criteria to choose the parameter error variance is still valid when the observation error variance is not very small. Looking at the eigenvalues of $\tilde{M}$ and $D$ in Figure 6.12 it can be seen that by making the eigenvalues of $D$ lie inside the corresponding circle, the condition $\rho(\tilde{M}) < 1$ is fulfilled.
The state and parameter approximation obtained are shown in Figure 6.13. The parameters are recovered to an acceptable level of accuracy and the magnitude of the state errors is similar to the magnitude of the observation errors. Though these are just empirical results, they suggest that choosing the parameter error variance $\sigma_p^2$ according to (6.168) could be also suitable for the case where the observations are noisy.

6.9.4 LRSSC for uncorrelated parameter errors

Now we apply the LRSSC method, choosing $\alpha = 0.5$. We consider that the variance for each parameter is $\sigma^2 = 0.05$, exactly as in the first experiment where the FDSPC failed.
Figure 6.13: Assimilated state for $L = \frac{n}{8} \Delta z$, $\sigma_p^2 = \sigma_p^{2*}$ and $\sigma_o^2 = 0.1$. Left upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state (red dashed line) after $N = 40$ time steps. Left lower: the true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. Right: state errors after $\frac{N}{2}$ assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).

Figure 6.14: Parameter estimation along the time window for $L = \frac{n}{8} \Delta z$, $\sigma_p^2 = \sigma_p^{2*}$, $\sigma_o^2 = 0.1$ and uncorrelated parameters. Each plot accounts for an entry of the parameter vector. True parameter (black) and estimated parameter (red).

to recover the true parameter vector. Figure 6.15 shows the structure of the iteration matrix $\tilde{M}$ and the covariance $B$. In this case the eigenvalues of the former are smaller than 1, so the method converges as shown in Figures 6.16 and 6.17.

For this specific case we get:
Figure 6.15: LRSSC. Augmented state covariance matrix $B$ and iteration matrix $\tilde{M}$ for uncorrelated parameter errors and parameter variance $\sigma_p^2 = 0.05$.

Figure 6.16: LRSSC. Assimilated state for $\sigma_p^2 = 25 \times 10^{-3}$. Left upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state (red dashed line) after $N = 40$ time steps. Left lower: the true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. Right: state errors after $\frac{N}{2}$ assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).

\[
|\lambda(B_{pp}T^T(\alpha B_{xx} + (1 - \alpha)TB_{pp} T^T)^{-1}T)| < 2
\]

\[
\Downarrow
\]

\[
\sigma_p^2 |\lambda((\alpha B_{xx} + (1 - \alpha)\sigma_p^2 I)^{-1})| = \frac{\sigma_p^2}{|\lambda(\alpha B_{xx} + (1 - \alpha)\sigma_p^2 I)|} < 2.
\]

This is equivalent to the following condition for each eigenvalue $\lambda$ of $B_{xx}$:

\[
\frac{\sigma_p^2}{|\alpha \lambda + (1 - \alpha)\sigma_p^2|} < 2.
\]
If we choose $\alpha < \frac{1}{2}$, and taking into account that $\lambda > 0$ because $B_{xx} > 0$, then

$$\frac{\sigma_p^2}{|\alpha \lambda + (1 - \alpha)\sigma_p^2|} = \frac{\sigma_p^2}{\alpha \lambda + (1 - \alpha)\sigma_p^2} < \frac{\sigma_p^2}{(1 - \alpha)\sigma_p^2} = \frac{1}{1 - \alpha} < 2,$$

(6.172)

and therefore, the method converges. For any other value of $\alpha$ we observe that

$$\min(1, \frac{\sigma_p^2}{\lambda}) < \frac{\sigma_p^2}{\alpha \lambda + (1 - \alpha)\sigma_p^2} < \max(1, \frac{\sigma_p^2}{\lambda}).$$

(6.173)

If the FDSPC converges, it means that

$$\max_{\lambda} \frac{\sigma_p^2}{\lambda} < 2,$$

(6.174)

and therefore the LRSSC also converges. This result is similar to the one found for a single parameter model in section 6.5.1.2.

### 6.9.5 Correlated parameter errors

Now we consider the case where the parameter errors are correlated. We assume that the covariance between parameter errors is proportional to the covariance between the correspondent state errors. It means that $B_{pp} = \sigma_p^2 C = cB_{xx}$, with $c = \frac{\sigma_p^2}{\sigma_b^2}$. In our experiments we set $c = 1$. The reason for this choice is the result from Theorem 6.5.1, which states that in the presence of perfect observations and a fully observed state, the
method needs only two iterations to converge if

\[ B_{pp} = (T^\top B_{xx}^{-1}T)^{-1}. \] (6.175)

Because here \( T = I \), equation (6.175) is equivalent to \( B_{pp} = B_{xx} \). For this parameter-parameter covariance matrix the state-state covariance matrix \( B'_{xx} \) for the LRSSC would be:

\[ B'_{xx} = \alpha B_{xx} + (1 - \alpha)TB_{pp}T^\top = \alpha B_{xx} + (1 - \alpha)B_{xx} = B_{xx}. \] (6.176)

Therefore, the LRSSC and the FDSPC methods are equivalent in this case. Though here observations are not perfect their variance is small, so we expect to have good results.

The structure of the resulting \( B \) and \( \tilde{M} \) for \( \sigma_o^2 = 10^{-6} \) is shown in Figure 6.18. It can be seen that the eigenvalues of \( \tilde{M} \) are very close to zero, which predicts a very fast convergence.

![Figure 6.18: Covariance error augmented state matrix \( B \) and iteration matrix \( \tilde{M} \)](image)

![Figure 6.19: Assimilated state for \( L = \frac{n}{2} \Delta z \), \( B_{pp} = B_{xx} \). Left upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state (red dashed line) after \( N = 40 \) time steps. Left lower: the true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. Right: state errors after \( \frac{N}{2} \) assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).](image)
Remarkably, the state error is even smaller, after letting the assimilated state be transported forward by the model, without any assimilation being performed (see Figure 6.19). This is due to a very good estimation of the parameters, resulting in a model with a very small bias and which, therefore, correctly represents the dynamics of the system. The forecast model works as some sort of smoother to the assimilated solution, driving it to the correct state, because the model bias is very precise and the state variables are not being pulled by observations out of their trajectory. In Figure 6.20 it can be seen that the parameter vector is very accurately estimated, even after just a few assimilation steps.

One advantage of taking $B_{pp} = B_{xx}$ is that, unlike for $B_{pp} = \sigma_p^2 I$, the convergence speed is not drastically affected when $\text{cond}(C) \gg 1$ for small $\sigma_o^2$. To corroborate it we run the assimilation process again, this time for $L = n \Delta z = 15 \frac{16}{16}$. For such a big correlation length scale the condition number of $C$ is 329.5 (instead of 13.3, like in the previous experiment). The results in Figures 6.21 and 6.22 prove that the estimates obtained are still very accurate, despite the bad conditioning of $C$.

**Convergence for larger $\sigma_o^2$**

As for the uncorrelated case we investigate what happens if the observation error variance is not so small. In our experiment $\sigma_o^2 = 0.1$, $L = n \Delta z$. The results for the state and parameter estimation are shown in Figures 6.23 and 6.24, respectively.
The errors augment as expected, because the observations are much more imprecise, but the errors for the state estimations and the observation errors are still of the same order. The results suggest that even for larger values of the observation error variance, good results can be obtained by setting $B_{pp} = B_{xx}$.

### 6.9.6 Unobserved states

If some of the state variables are not observed, then $m < l$ and according to the results of Section 6.6 a 3D-Var-like method in general does not converge. In the experiments
we observe every second variable, starting with position 2, that is: $x_2, x_4,$ etc. The observations are independent, with variance $\sigma_o^2 = 0.1$ and the correlation length scale is $L = \frac{n}{8} \Delta z$. Figure 6.25 shows that 1 is an eigenvalue of $\tilde{M}$ as expected. The structure of $\tilde{M}$ and $B$ is shown in Figure 6.26. Figure 6.27 shows the assimilated state at the end of the assimilation window and after 40 a posteriori runs of the model. The assimilated values of the parameter along the time window are shown in Figure 6.28.

Though the state assimilation is quite accurate, the parameter vector is not very accurately estimated. After further experiments we concluded that there are two factors
Chapter 6. Convergence

Figure 6.25: Eigenvalues of $\tilde{M}$ for unobserved state variables

\begin{figure}[ht]
\centering
\includegraphics[width=\textwidth]{figure625}
\caption{Eigenvalues of $\tilde{M}$ for unobserved state variables}
\end{figure}

$B$ $\tilde{M}$

Figure 6.26: Covariance error augmented state matrix $B$ and iteration matrix $\tilde{M}$ for unobserved state variables

\begin{figure}[ht]
\centering
\includegraphics[width=\textwidth]{figure626}
\caption{Covariance error augmented state matrix $B$ and iteration matrix $\tilde{M}$ for unobserved state variables}
\end{figure}

Figure 6.27: Assimilated state for $L = n\Delta z$, $B_{pp} = B_{xx}$ and $\sigma^2_o = 0.1$. Left upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state (red dashed line) after $N = 40$ time steps. Left lower: the true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. Right: state errors after $N^2$ assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).

that contribute to the accuracy of the state estimation in this experiment. The first one is that the initial background parameter vector is the zero vector, which differs from the
true parameter in only one entry (the position where the heat source is located). In general it is not possible to start with such a good first guess of the parameter vector. The second one is that the state variable, corresponding to the temperature of the position where the source is located, is observed. We suppose that the location of the source is unknown, so observing the specific variable corresponding to the source position is not always possible.

To support this hypothesis, two last experiment results are presented. In the first one the same experiment set-up is used, with the only change being that the initial parameter is generated randomly ($p \sim \mathcal{N}(0, B_{pp})$). In the second one we observe all the state variables except the one corresponding to the temperature in the location of the source.

When the initial parameter is generated randomly, the size of the state error at the end of the assimilation window as well as after 40 a posteriori model running is doubled, as shown in Figure 6.29. From Figure 6.30 it is obvious that the quality of the parameter estimation is very poor. On the other hand, though we augment the number of observations, if the state describing the temperature in the portion of the rod where the source is located is not observed, the corresponding parameter is poorly estimated (see Figure 6.32), resulting in a strong underestimation of the temperature of the rod at this point, as shown in Figure 6.31.
Figure 6.29: Assimilated state for $L = \frac{n}{2} \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$ with random initial parameter. Left upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state (red dashed line) after $N = 40$ time steps. Left lower: the true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. Right: state errors after $\frac{N}{2}$ assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).

Figure 6.30: Parameter estimation along the time window for $L = n \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$ with random initial parameter. Each plot accounts for an entry of the parameter vector. True parameter (black) and estimated parameter (red).
Figure 6.31: Assimilated state for $L = \frac{n}{2} \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$. All state variables are observed except where the source is located. Left upper: observations (magenta circles), the true state (solid black line), forecast state without assimilation of observations (dotted blue line) and assimilated state (red dashed line) after $N = 40$ time steps. Left lower: the true state, forecast state without assimilation of observations and assimilated state after running the forecast model for another 40 time steps without assimilating observations. Right: state errors after $\frac{N}{2}$ assimilation steps (solid black line), state errors after running the forecast model for another 40 time steps without assimilating observations (dashed red line).

Figure 6.32: Parameter estimation along the time window for $L = n \Delta z$, $B_{pp} = B_{xx}$ and $\sigma_o^2 = 0.1$. All state variables are observed except where the source is located. For each component $p_i$, $i = 1, 2, \ldots, 15$ of the parameter vector it is represented the true (black) and estimate (red) value.
6.10 Summary and discussion

In this chapter we stated and proved several convergence conditions for the 3D-Var method applied to a simple linear state and parameter DA problem. These conditions explain basically how the definition of the state-parameter covariance affects the convergence of the method. For this specific problem both FDSPC and LRSSC constitute a particular case of 3D-Var as indicated at the end of Section 6.2, and special conditions for them were stated where appropriate.

If the full state is observed and the observation errors are very small, it was shown that the method’s convergence depends on the eigenvalues of \( B_{xp}^T B_{xx}^{-1} T \) being inside the unit circle with centre in \((1,0)\). In Section 6.5.1.2 we proved that for the single parameter case the convergence condition for the LRSSC is less restrictive than for the FDSPC method. This was also proven in Section 6.9.4 for the case where the parameter vector represents a model bias \((T = I)\) and the parameters are uncorrelated with similar variance. Moreover, if the parameter errors are uncorrelated and with similar variance, the convergence speed of the both methods is inversely proportional to the condition number of the state-state correlation matrix. However, in the presence of correlated parameters the convergence speed is not drastically affected if we augment the condition number of the state-state covariance by changing the correlation length-scale, as shown in Section 6.9.5. These theoretical results were tested using the simple 1-D heat equation model. The experimental results suggest that the conditions found can be empirically used even if observations are noisy.

In Section 6.6 it was proven that convergence is only attained if the dimension of the parameter vector is smaller than the number of observations. This is not a huge limitation in typical joint state-parameter estimation problems, in which generally the number of parameters is considerably smaller than the size of the observation vector. Nevertheless, if the augmented state approach is used to estimate model errors [141], the parameter, which represents the model bias, is generally defined in the same space as the model state variable. In this case, and under linearity conditions, 3D-Var-like methods, including the two approaches studied, are not a proper choice unless the full state is observed.

Special necessary and sufficient conditions for convergence were found for the single parameter case where only one state variable is observed. For the FDSPC it was shown that if the variances of the observation and of the observed state variable are small, then the interval of possible values of the parameter variance to attain convergence is also small; in other words, the degree of certainty in the parameter estimation must be very high if both the observed variable estimate and the observation are very precise.
Even in case that the state estimation is very precise, recovering the true parameter is only possible if certain conditions involving the kernel of $B_{xp}^T H^T R^{-1} H$ are fulfilled.

In the next chapter we test the suitability of the LRSSC by experimenting with nonlinear test models and the results are compared with those obtained by applying the FDSPC.
Chapter 7

Nonlinear Experiments

In Chapter 5 we presented two strategies that combine the 3D-Var method ideas with a low computational cost update of the error covariance matrix, to keep a flow dependency on the error statistics. In the FDSPC proposed by Smith [101], only the state-parameter error covariance is updated, to avoid processing the high dimensional state-state matrix. The author shows that this type of update is enough to achieve a high accuracy estimation for both, parameter and state variables, in a variety of data assimilation (DA) problems. Our hybrid method, the LRSSC, differs from the FDSPC in that the state-state error covariance matrix $B_{xx}$ is also updated, using the derivatives of the model with respect to the parameters and the parameter-parameter error covariance.

We proved in the last chapter that for the linear state-parameter DA problem several convergence conditions are less restrictive for the LRSSC when compared with those obtained for the FDSPC (see sections 6.5.1.2 and 6.9.4). We expect that considering a flow-dependent $B_{xx}$ can also improve the assimilation results when the method is applied to nonlinear models.

In this chapter we investigate the effectiveness of our hybrid strategy by conducting twin assimilation experiments using different test problems, and comparing our results those obtained using the FDSPC.

In Section 7.1 we start with the linear advection model previously presented in this work. Section 7.2 discusses a two-parameter nonlinear damped oscillating system and in Section 7.3 the Lorenz 63 model is presented. In each section we introduce the model and its discretization, give details about the experiment set-up and finally show the results that are relevant to this investigation. Section 7.4 summarizes our observations and conclusions.
Chapter 7. Nonlinear experiments

7.1 Advection model

We first start with the linear advection model introduced in Section 4.3.1, in which the speed of the bed height is the unknown parameter to be estimated. Basically we will use the same upwind scheme to obtain the discrete dynamical system (4.31).

7.1.1 Error covariances.

The state-state error covariance matrix $B_{xx}^b$ is generated according to

$$
(B_{xx}^b) = \sigma_b^2 C,
$$

(7.1)

where $C = c_{ij}$ is a matrix describing the correlation between the entries $i$ and $j$ of the state vector. In the experiments we used $\sigma_b^2 = 0.05$ and $C$ is the Markov matrix [98]

$$
c_{ij} = e^{-\frac{r_{ij}}{L}}, \quad i, j = 1, \ldots, n,
$$

(7.2)

also used by Smith for the same test problem. Here $r_{ij}$ is the distance between gridpoints $i$ and $j$, i.e., $r_{ij} = |i-j|\Delta z$, and $L$ is a correlation length scale. The larger $L$, the stronger is the correlation between state variables corresponding to distant gridpoints. On the other hand, if $L$ is small, only the correlation between neighbour gridpoints plays a significant role in the assimilation process.

As there is only one parameter to be estimated (the advection speed), the covariance $B_{pp}$ is a scalar and it is renamed as $\sigma_p^2$. In the experiments we set $\sigma_p^2 = 0.1$. The state-parameter error covariance is updated previous to an assimilation step using (4.37):

$$
(B_{xp})_k = N_k B_{pp} = \sigma_p^2 \begin{pmatrix}
  x_{k,1} - x_{k,n} \\
  x_{k,2} - x_{k,1} \\
  x_{k,3} - x_{k,2} \\
  \vdots \\
  x_{k,n} - x_{k,n-1}
\end{pmatrix},
$$

(7.3)
For $B_{xx}$, the resulting update according to (5.25) is

$$(B_{xx})_k = \alpha_1 B_{xx}^b + \alpha_2 \sigma_p^2 \begin{pmatrix}
    x_{k,1} - x_{k,n} \\
    x_{k,2} - x_{k,1} \\
    x_{k,3} - x_{k,2} \\
    \vdots \\
    x_{k,n} - x_{k,n-1}
\end{pmatrix}
\begin{pmatrix}
    x_{k,1} - x_{k,n}, x_{k,1} - x_{k,n}, \ldots, x_{k,1} - x_{k,n}.
\end{pmatrix}$$

(7.4)

The observation errors are considered to be uncorrelated and with equal variance, i.e., $R_k = \sigma_o^2 I$.

### 7.1.2 Experimental set-up

The LRSSC was tested within the framework of identical twin experiments. We assume that the discretized model (4.30) is the true model and that it is perfect. We consider $\Delta z = 0.01$ and $\Delta t = 0.01$. As in section 4.3.4 we restrict the possible speed values to the interval $[0, 1]$ to keep the consistency of the discretization. This way also stability is guaranteed.

In our experiments we observed that the accuracy of the state and parameter estimation depends on several factors, such as the observation spacing ($\Delta x_{\text{obs}}$), observation frequency ($\Delta t_{\text{obs}}$), observation error variance ($\sigma_o^2$), the selection of $\alpha$, among others. Here we show assimilation results which are representative for a larger set of experiments conducted in the framework of this investigation. For the conducted experiments $\alpha$ is set to 0.3, unless stated otherwise. This is by no means the optimal value of the scheme parameter. Our goal is to assess the performance of the method when an $\alpha$ different from 1 is considered. Later in this section we discuss how the selection of $\alpha$ has an impact on the assimilation results.

The true state is obtained by evaluating the function

$$x_t(z) = \begin{cases} 
    2e^{-\frac{(z-0.35)^2}{0.01}} & \text{if } z \in [0.01, 0.5] \\
    0 & \text{if not}
\end{cases}$$

(7.5)

at the 301 gridpoints. The background initial state $x_b$ is generated by perturbing a slightly displaced true initial state. This is done by evaluating the function

$$g(z) = x_t(z + 0.10) + \beta(0.05C),$$

(7.6)
at the gridpoints, where $\beta(V)$ denotes a Gaussian noise with zero mean and covariance $V$. The true speed is $s_t = 0.8$ and the initial background speed is slower: $s_b = 0.53$. The parameter variance is set to $\sigma_p^2 = 0.1$. Every $\Delta x_{\text{obs}}$ state variable is observed, starting with the first one. The observation errors are uncorrelated and unbiased with variance $\sigma_o^2$ and are available every $\Delta t_{\text{obs}}$ time steps. The correlation length scale for the background state-state error covariance is twice the space between observations, i.e., $L = 2\Delta z \Delta x_{\text{obs}}$.

7.1.3 Experimental results

Perfect observations

In this section we analyse the results obtained considering perfect observations. This means that observations are generated using the trajectory of the true state given by the forecast model along the assimilation window, without adding any additional noise.

![Figure 7.1: Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window for different assimilation intervals and perfect observations, initial background parameter $s_b = 0.53$, $\sigma_p^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. Assimilations for $\Delta t_{\text{obs}} = 5$ (blue), $\Delta t_{\text{obs}} = 15$ (green), $\Delta t_{\text{obs}} = 25$ (red), $\Delta t_{\text{obs}} = 50$ (yellow) and $\Delta t_{\text{obs}} = 100$ (magenta). The black line corresponds to the parameter true value.](image)

We first investigate the influence of the observation frequency on the quality of the analysis. In Figure 7.1 the sequence of parameter assimilations for $\Delta x_{\text{obs}} = 10$, $\sigma_o^2 = 0.01$ and different observation frequencies between $5\Delta t$ and $100\Delta t$ time steps are shown. In
Figure 7.2: Linear Advection. Norm of the assimilated state error vector for perfect observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. Assimilations for $\Delta t_{\text{obs}} = 5$ (blue), $\Delta t_{\text{obs}} = 15$ (green), $\Delta t_{\text{obs}} = 25$ (red), $\Delta t_{\text{obs}} = 50$ (yellow) and $\Delta t_{\text{obs}} = 100$ (magenta).

Figure 7.3: Linear Advection. Background and assimilated state for $t = 0, 1, 2, 3, 4, 10$ and perfect observations, where initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 25$. In the graphic: forecast state (blue dashed line), assimilated state (red), true state (black) and observations (blue circles).

All cases the value of the parameter is recovered to a high level of accuracy. The more frequent the observations are assimilated, the faster is the convergence of the assimilated parameter to its true value. In general the parameter oscillates around the true value,
with the amplitude of the oscillations getting smaller while more observations are assimilated until convergence to the true value is achieved. Especially if the observations are sparse in time, the scheme requires a larger number of time steps until the parameter estimate starts oscillating around the true value.

The norm of the difference between the true state and the assimilated state for every time step is presented in Figure 7.2. The assimilated state converges faster to the truth when observations are taken more frequently. The lack of precision on the parameter estimate can be identified from the peaks of the norm trajectory in the plot. If the speed is not accurately estimated, the error norm is anyway reduced with each assimilation step because the observations pull the state variables towards their true values. In the following forecast steps, nevertheless, a mismatch between the true and assimilated parameter causes a fast increase of the error norm.

Comparing Figures 7.3 and 7.4 it becomes clear that doubling the observation frequency affects the assimilation results. In Figure 7.4, at time \( t = 4 \), the difference between the forecast and the assimilated bed height peaks is considerable (which occurs when the model runs without any speed correction for a long period). As a result there is a broad interval where the forecast and the true state differ extremely from each other. The resulting assimilated state has as a consequence a wider and lower wave peak. Still at time \( t = 10 \) the problem persists. After time \( t = 20 \) a very precise estimate of the state is obtained for all the analysed cases.
Figure 7.5: Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window. Comparison between the FDSPC (red) and the LRSSC (blue) methods for perfect observations, initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 100$. True parameter (black).

When we compared the performance of the FDSPC and the LRSSC, no significant differences could be detected for a small $\Delta t_{\text{obs}}$. For large $\Delta t_{\text{obs}}$, nevertheless, our approach proved to converge faster.

In Figures 7.5 and 7.6 the results for $\Delta t_{\text{obs}} = 100$ are shown. The FDSPC takes longer to accurately approximate the parameter and this is reflected in a larger state error.

Figure 7.6: Linear Advection. Norm of the assimilated state error vector for the FDSPC (red) and the LRSSC (blue) methods for perfect observations, initial background parameter $s_b = 0.53$, $\sigma^2 = 0.01$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 100$. 
norm along the time window. For time $t = 20$ both methods have recovered state and parameter to a high level of accuracy. When the initial background speed was set to 0.3, both methods failed to recover parameter and state precisely. In this case the small parameter variance $\sigma_p^2 = 0.1$ does not reflect the real statistics of the parameter along the time window.

In Figures 7.7 and 7.8 results are shown for $\sigma_p^2 = 0.3$. Again both methods recover accurately parameter and state, but the FDSPC takes much longer to converge.

Figure 7.9 shows the sequence of assimilated parameters for different observation densities ($\Delta x_{\text{obs}} = 5, 15, 25, 50, 100$), this time setting the parameter variance back to $\sigma_p^2 = 0.1$. The convergence speed also depends on the spatial distribution of observations (see Figure 7.10). Even for $\Delta x_{\text{obs}} = 50$ the scheme recovers parameter and state to a high level of accuracy, but fails for $\Delta x_{\text{obs}} = 100$. This is expected if it is taken into account that in this case only 3 out of 301 state variables are observed. Comparing Figures 7.11 and 7.12 it can be seen how the quality of the assimilated state is worsened when the number of state variables observed is reduced.
the parameter estimation is completely wrong and so is the assimilated state (Figure 7.14).

A similar situation was observed when experiments were carried out for different observation variances. Even if the observations are perfect, the fact that the observation error covariances are different from zero means that the assimilation scheme is not sure about the accuracy of observations and this will therefore affect the assimilation results. For larger variances, for example $\sigma^2_o = 1$, the scheme does not trust the observations enough and it takes longer to recover the parameter and state as shown in Figures 7.15 and 7.16. For a very small variance ($\sigma^2_o = 0.001$ or smaller) the error oscillates strongly until the scheme is stabilized. A possible explanation is that in general the assimilated state takes a value between the observation and the prediction. If the observation variance is very small the observed states are updated practically with the value of the observations. This corresponds to the maximal distance possible between the assimilated and the forecast state. Therefore the size of the parameter correction is also maximized, as the scheme tries to match the forecast and assimilated states by augmenting or decreasing the speed drastically. Now, if the speed was overestimated, because the state is already corrected, in the next assimilation step the peak of the forecast state is in front of the peak of the true state and therefore the new speed will be underestimated, and vice versa. This can cause, at the beginning of the time window, that the parameter oscillates from its maximal value 1 to its minimal value 0, as shown in Figure 7.15.

Figure 7.8: Linear Advection. Norm of the assimilated state error vector for the FD-SPC (red) and the LRSSC (blue) methods for perfect observations, initial background parameter $s_b = 0.3$, $\sigma^2_o = 0.01$, $\sigma^2_p = 0.3$, $\Delta x_{obs} = 10$ and $\Delta t_{obs} = 100$. 
Figure 7.9: Linear Advection. Assimilated parameter along the time window for different observation spacing and perfect observations, initial background parameter $s_b = 0.53$, $\sigma_b^2 = 0.01$ and $\Delta t_{\text{obs}} = 15$. Assimilations for $\Delta x_{\text{obs}} = 5$ (blue), $\Delta x_{\text{obs}} = 15$ (green), $\Delta x_{\text{obs}} = 25$ (red), $\Delta x_{\text{obs}} = 50$ (yellow) and $\Delta x_{\text{obs}} = 100$ (magenta). The black line represents the parameter true value.

Figure 7.10: Linear Advection. Norm of the assimilated state error vector for perfect observations, initial background parameter $s_b = 0.53$, $\sigma_b^2 = 0.01$ and $\Delta t_{\text{obs}} = 15$. Assimilations for $\Delta x_{\text{obs}} = 5$ (blue), $\Delta x_{\text{obs}} = 15$ (green), $\Delta x_{\text{obs}} = 25$ (red), $\Delta x_{\text{obs}} = 50$ (yellow) and $\Delta x_{\text{obs}} = 100$ (magenta).
Chapter 7. Nonlinear experiments

Figure 7.11: Linear Advection. Background, assimilated and true state for $t = 0, 1, 2, 3, 4, 10$ and perfect observations, where initial background parameters $s_b = 0.53$, $\sigma^2_o = 0.01$, $\Delta x_{\text{obs}} = 25$ and $\Delta t_{\text{obs}} = 15$. In the graphic: forecast state (blue dashed line), assimilated state (red), true state (black) and observations (blue circles).

Figure 7.12: Linear Advection. Background, assimilated and true state for $t = 0, 1, 2, 3, 4, 10$ and perfect observations, where initial background parameters $s_b = 0.53$, $\sigma^2_o = 0.01$, $\Delta x_{\text{obs}} = 40$ and $\Delta t_{\text{obs}} = 15$. In the graphic: forecast state (blue dashed line), assimilated state (red), true state (black) and observations (blue circles).
Figure 7.13: Linear Advection. Assimilated parameter along the time window. Comparison between the FDSPC (red) and the LRSSC (blue) methods for perfect observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$, $\Delta x_{\text{obs}} = 50$ and $\Delta t_{\text{obs}} = 15$.

Figure 7.14: Linear Advection. Norm of the assimilated state error vector for the FDSPC (red) and the LRSSC (blue) methods for different observation spacing and perfect observations, initial background parameter $s_b = 0.53$, $\sigma_o^2 = 0.01$, $\Delta x_{\text{obs}} = 50$ and $\Delta t_{\text{obs}} = 15$. 
Chapter 7. Nonlinear experiments

Figure 7.15: Linear Advection. Assimilated parameter along the time window for different observation variances and perfect observations, initial background parameter $s_0 = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.1$ (red) and $\sigma_o^2 = 1$ (magenta). The black line represents the parameter true value.

Figure 7.16: Linear Advection. Norm of the assimilated state error vector for different observation variances and perfect observations, initial background parameter $s_0 = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.1$ (red) and $\sigma_o^2 = 1$ (magenta).
Noisy observations

A similar set of experiments was conducted considering this time noisy observations. As explained in Section 4.3.4 the noise added to the perfect observations is generated using the exact statistical distribution of the observation errors given by the observation error covariance $R = \sigma_o^2 I$. For these experiments $\sigma_p^2 = 0.1$.

We repeated the experiments varying the observation frequency from $5\Delta t$ to $100\Delta t$. If we analyse Figure 7.17, it is very difficult to find any significant difference in the behaviour of the assimilated parameter for different frequencies after a certain time interval, which we call tune-up phase. During this early stage of the assimilation process the assimilated parameter gets closer to the true value, but experiencing drastic changes. In general, the parameter oscillates around the true value, no matter for how long the assimilations are performed. The tune-up phase, in which the parameter oscillations are wider, lasts longer for larger $\Delta t_{\text{obs}}$. However, after the model stabilizes there are no obvious disadvantages compared to less frequent observations. In fact, when observations are taken every $5\Delta t$, the trajectory of the assimilated parameter is more chaotic, especially in the time interval [14, 20].

![Figure 7.17: Linear Advection. Assimilated parameter along the time window for different assimilation intervals and perfect observations, initial background parameter $s_0 = 0.53$, $\sigma_o^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. Assimilations for $\Delta t_{\text{obs}} = 5$ (blue), $\Delta t_{\text{obs}} = 15$ (green), $\Delta t_{\text{obs}} = 25$ (red), $\Delta t_{\text{obs}} = 50$ (yellow) and $\Delta t_{\text{obs}} = 100$ (magenta). The black line represents the parameter true value.](image)

From Figure 7.18 it can be observed that the error norm oscillates in a similar range for all the observation frequencies. This was already pointed out by Smith in [101]. In this investigation the author mentioned that averaging the assimilated parameter
along a moving time window helped to make the analysis smoother. We made a similar
observation in this study when applying the EKF modifications in Section 4.3.

In Figure 7.19 results are shown for the case where the parameter is averaged after 10
seconds of tuning along a 10 seconds moving time window. The sequence of assimilated
parameters is smoothed in all cases. After 20 seconds the true value of the parameter
is recovered to a high level of accuracy. For observations taken every $100\Delta t$, recovering
the true parameter takes much longer. In this case averaging the parameter slows down
the process. Even if the new assimilated value is close to the truth, averaging it with
the former estimations, which are very inaccurate, pulls it away from the true value. If
a longer tuning time window is chosen the true parameter is recovered faster.

In Figure 7.20 it can be observed how the imprecise parameter estimation for $\Delta x_{\text{obs}} = 100$
causes a slow stabilization of the error norm. Once the error is stabilized the size of the
norm is equivalent to the case where no parameter average was considered. Figure 7.21
shows the state estimation at different time steps. In general, the scheme is very good
in estimating the peak of the wave, but struggles to accurately estimate the flat regions.
Notice, nevertheless, that the absolute error for the estimation of each state variable is
in general smaller or very close to 0.1, which is the standard deviation of observations.
Therefore, the results are as good as it can be expected, as they are consistent with the
variability of the observational data.

![Figure 7.18: Linear Advection. Norm of the assimilated state error vector for noisy
observations, initial background parameter $s_b = 0.53$, $\sigma^2_o = 0.01$ and $\Delta x_{\text{obs}} = 10$.
Assimilations for $\Delta t_{\text{obs}} = 5$ (blue), $\Delta t_{\text{obs}} = 15$ (green), $\Delta t_{\text{obs}} = 25$ (red), $\Delta t_{\text{obs}} = 50$
(yellow) and $\Delta t_{\text{obs}} = 100$ (magenta).](image)
Figure 7.19: Linear Advection. Assimilated parameter along the time window for different assimilation intervals and perfect observations, initial background parameter $s_b = 0.53$, $\sigma_b^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. Assimilations for $\Delta t_{\text{obs}} = 5$ (blue), $\Delta t_{\text{obs}} = 15$ (green), $\Delta t_{\text{obs}} = 25$ (red), $\Delta t_{\text{obs}} = 50$ (yellow) and $\Delta t_{\text{obs}} = 100$ (magenta). The horizontal black solid line represents the parameter true value. The vertical black dashed line indicates the time from which the parameter estimates are averaged.

Figure 7.20: Linear Advection. Norm of the assimilated state error vector for noisy observations, initial background parameter $s_b = 0.53$, $\sigma_b^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. Assimilations for $\Delta t_{\text{obs}} = 5$ (blue), $\Delta t_{\text{obs}} = 15$ (green), $\Delta t_{\text{obs}} = 25$ (red), $\Delta t_{\text{obs}} = 50$ (yellow) and $\Delta t_{\text{obs}} = 100$ (magenta). For this experiment the parameter is averaged after 5 seconds.
Chapter 7. Nonlinear experiments

Figure 7.21: Linear Advection. State assimilation at different times, for noisy observations, initial background parameter $s_0 = 0.53$, $\sigma^2 = 0.01$ and $\Delta x_{\text{obs}} = 10$. Assimilations for $\Delta t_{\text{obs}} = 5$ (blue), $\Delta t_{\text{obs}} = 15$ (green), $\Delta t_{\text{obs}} = 25$ (red), $\Delta t_{\text{obs}} = 50$ (yellow) and $\Delta t_{\text{obs}} = 100$ (magenta). True state (black).

Figure 7.22: Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window. Comparison between the FDSPC and the LRSSC methods for noisy observations, initial background parameter $s_0 = 0.3$, $\sigma^2 = 0.01$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 50$. Upper plot: true parameter (black), assimilated parameter for the FDSPC (red) and for the LRSSC (blue). Lower plot: observations (blue circles) true state (black), FDSPC assimilated state (red) and LRSSC assimilated state (blue).
The two methods can be compared from Figures 7.22 and 7.23, where the initial background parameter is set to $s^b = 0.3$ (as no significant differences were detected for $s^b = 0.53$). The quality of the parameter assimilation with our method is slightly better. A more significant difference was found in the quality of the state assimilation. Our method manages to adjust the peak of the wave accurately, while if a static state-state error covariance is used there are noticeable differences between the true and the assimilated state. Our state estimation is also better in the flat areas. As a measure of the quality of the assimilation we use the Root Mean Square Error (RMSE) applied to the estimates along the assimilation window.

The standard RMSE is a frequently used measure of the difference between values predicted by the model and those that are actually observed. We use a similar formula to measure the quality of the analysis, but here depending on the assimilated and true states:

$$RMSE = \sqrt{\frac{1}{N-t_i+1} \sum_{k=t_i}^{N} \| x^a_k - \hat{x}^a_k \|^2_2},$$

where $t_i$ is the time step where the scheme is considered to be tuned up and the parameter starts to be averaged and $N$ the number of time steps. For the FDSPC the RMSE = 1.2701, while for the LRSSC the RMSE = 0.9138. This represents an error reduction of 28%. To get an idea about the quality of this approximation it should be considered that if the error is distributed uniformly into all the state components, a RMSE of 1 would correspond to a deviation of 0.0576 from the truth, which is considerably smaller than the observation’s standard deviation, which is 0.1.

![Figure 7.23: Linear Advection. Norm of the assimilated state error vector for the FDSPC (red) and the LRSSC (blue) methods for noisy observations, initial background parameter $s_b = 0.3$, $\sigma^2 = 0.01$, $\Delta x_{obs} = 10$ and $\Delta t_{obs} = 50$.](image)
When experiments were performed for different observation spacings (experiments were conducted for $s^b = 0.53$), the results for both parameter and state estimations were equivalent for $\Delta x_{\text{obs}} = 5, 15, 25$. Only for $\Delta x_{\text{obs}} = 50$ the assimilated state and parameter differed completely from the truth, even if longer assimilation windows were considered (see Figures 7.24 and 7.25). In Figure 7.26 we show the assimilated state for different time steps. It is very difficult to find any significant difference in the quality of the assimilation for all the observation spacings smaller than 50.

![Figure 7.24: Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window for different observation spacing and noisy observations, initial background parameter $s_0 = 0.53$, $\sigma^2_o = 0.01$ and $\Delta t_{\text{obs}} = 15$, where the parameter estimates are averaged over a moving time window. Assimilations for $\Delta x_{\text{obs}} = 5$ (blue), $\Delta x_{\text{obs}} = 15$ (green), $\Delta x_{\text{obs}} = 25$ (red) and $\Delta x_{\text{obs}} = 50$ (yellow). The horizontal black solid line represents the parameter true value. The vertical black dashed line indicates the time from which the parameter is averaged.](image)

Experiments with different observation variances were conducted. For large variances, the tune-up time of 5 seconds proved to be insufficient in most of the experiment runs. In Figure 7.27 we show the assimilated parameter for a tune-up phase of 10 seconds. For $\sigma^2_o = 0.001, 0.01, 0.1$ the parameter is very accurately recovered. Even for the extreme case $\sigma^2_o = 1$, which means that the variance of the observations is larger than half of the maximum bed height, the level of accuracy of the estimation is high. Nevertheless, the precision of the state assimilation is significantly affected by a larger observation variance (see Figure 7.29 ). Figure 7.28 shows how the state error levels consistently decrease when smaller variances are considered.
Figure 7.25: Linear Advection. Norm of the error in the state assimilation along the time window for different observation spacing and noisy observations, initial background parameter $s_b = 0.53$, $\sigma^2_o = 0.01$ and $\Delta x_{\text{obs}} = 10$. Assimilations for $\Delta x_{\text{obs}} = 5$ (blue), $\Delta x_{\text{obs}} = 15$ (green), $\Delta x_{\text{obs}} = 25$ (red) and $\Delta x_{\text{obs}} = 50$ (yellow).

Figure 7.26: Linear Advection. State estimation at different times, for different observation spacing and noisy observations, initial background parameter $s_b = 0.53$, $\sigma^2_o = 0.01$ and $\Delta t_{\text{obs}} = 15$. Assimilations for $\Delta x_{\text{obs}} = 5$ (blue), $\Delta x_{\text{obs}} = 15$ (green), $\Delta x_{\text{obs}} = 25$ (red) and $\Delta x_{\text{obs}} = 50$ (yellow). True state (black).
Chapter 7. Nonlinear experiments

Figure 7.27: Linear Advection. Assimilated parameter along the time window for different observation variances and noisy observations, initial background parameter $s_b = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.1$ (red) and $\sigma_o^2 = 1$ (magenta). The horizontal black solid line represents the parameter true value. The vertical black dashed line indicates the time from which the parameter estimates are averaged.

Figure 7.28: Linear Advection. Norm of the assimilated state error vector for different observation variances and noisy observations, initial background parameter $s_b = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.1$ (red) and $\sigma_o^2 = 1$ (magenta).
Figure 7.29: Linear Advection. State estimation at different times, for different observation variances and noisy observations, initial background parameter $s_0 = 0.53$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. Assimilations for $\sigma^2_0 = 0.001$ (blue), $\sigma^2_0 = 0.01$ (green), $\sigma^2_0 = 0.1$ (red) and $\sigma^2_0 = 1$ (magenta). True state (black).

Figure 7.30: Linear Advection. Parameter estimation for the FDSPC (red) and the LRSSC (blue) methods for noisy observations, initial background parameter $s_0 = 0.3$, $\sigma^2_p = 0.5$, $\sigma^2_o = 1$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. The horizontal black solid line represents the parameter true value. The vertical black dashed line indicates the time from which the parameter estimates are averaged.
Figure 7.31: Linear Advection. Assimilated parameter along the time window and assimilated state at the end of the time window. Comparison between the FDSPC and the LRSSC methods for noisy observations, initial background parameter $s_b = 0.3$, $\sigma_2^2 = 1$, $\sigma_p^2 = 0.5$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. True state (black), FDSPC assimilated state (red) and LRSSC assimilated state (blue).

Figure 7.32: Linear Advection. Norm of the assimilated state error vector for the FDSPC (red) and the LRSSC (blue) methods for noisy observations, initial background parameter $s_b = 0.3$, $\sigma_p^2 = 0.5$, $\sigma_2^2 = 1$, $\Delta x_{\text{obs}} = 10$ and $\Delta t_{\text{obs}} = 15$. 
When we compared the performance of both hybrid approaches, we did not find significant differences in the accuracy of the assimilated parameters. In Figure 7.30 we show the parameter estimates for very noisy observations \((\sigma_o^2 = 1, \text{which represents approximately one half of the maximum bed height at time } t = 0 \text{ and is slightly larger than the maximum bed height at time } t = 12)\) and a very inaccurate initial parameter \((s_b = 0.3)\). In this case it was necessary to set the parameter variance to 0.5 to reflect the uncertainty of the parameter estimation. In both cases the parameter was surprisingly precisely recovered. Nevertheless, there were significant differences with respect to the state assimilation. During the tune-up phase, the FDSPC estimation is even better (see \(t = 4.2 \text{ or } t = 6.3 \) in Figure 7.31), but once the model is stabilized the precision of the LRSSC is noticeably higher, as shown in 7.32.

**Selection of \(\alpha\)**

In order to investigate how the selection of \(\alpha\) in the update of the state-state error covariance \(B_{xx}^{f} = \alpha B_{xx}^{p} + (1 - \alpha)N_k B_{pp} N_k^\top\) introduced in (5.25) influences the quality of the state assimilation, experiments were performed combining different \(\alpha\)’s and observation frequencies, spacings and variances. For each combination 20 independent experiments were conducted. For each category the mean and variance of the RMSE (as defined in (7.7)) over all experiments were computed.

In Figure 7.33 results for different \(\alpha\)’s and \(\Delta t_{\text{obs}}\)’s are shown. For these experiments we set \(\sigma_o^2 = 0.01\) and \(\Delta x_{\text{obs}} = 10\). Moving along the columns in Figure 7.33 it can be noticed that for a fixed \(\alpha\) the RMSE slightly varies, but there is no evident relation between the observation frequency and the size of the RMSE. This was already observed in previous experiments. There is nevertheless an obvious relation between the size of \(\alpha\) and the quality of the assimilation. In this case, the smaller \(\alpha\), the smaller the RMSE we obtain. The state-state error covariance approximates a rank 1 matrix when \(\alpha\) goes to zero. This means that the information coming from the parameter derivatives is crucial in the estimation of the state error. If the full EKF update is performed, the state-state error covariance is very similar to the rank 1 matrix \(N_k B_{pp} N_k^\top\) (see Figures 5.3 and 5.2). The fact that the RMSE variance is small indicates that the mean is representative of what occurs in individual experiments.

Results of experiments for different \(\Delta x_{\text{obs}}\)’s are shown in 7.34. They were conducted observing the state every 15\(\Delta t\). If \(\alpha = 0.05\) is considered, the scheme fails to approximate the state within any reasonable boundaries for \(\Delta x_{\text{obs}} = 40, 45\). Even for \(\Delta x_{\text{obs}} = 15, 50\) the error variance is extremely high, which indicates that for some experiments the error was considerably higher compared to the mean. As expected, for a fixed \(\alpha\) the error is in general smaller for smaller values of \(\Delta x_{\text{obs}}\). Nevertheless it can be noticed that
the RMSE for $\Delta x_{\text{obs}} = 45$ is much larger than, for example, for $\Delta x_{\text{obs}} = 50$. One possible explanation is that for this specific distribution of observations in combination with the observation frequency, the peak of the wave is not often observed, affecting the quality of the assimilation. Here again, more precise results are obtained if smaller $\alpha$’s are considered. The same was observed when experiments for different observation variances were conducted (see Figure 7.35). The improvement of the state assimilation quality, when applying smaller $\alpha$’s, is more noticeable for large observation variances.

**Figure 7.33:** Advection model. RMSE for different $\alpha$’s and observation frequencies. Left: RMSE mean. Right: RMSE variance.

**Figure 7.34:** Advection model. RMSE for different $\alpha$’s and observation spacings. Left: RMSE mean. Right: RMSE variance.

### 7.1.4 Summary

The LRSSC performed very well for most of the experiments conducted. Only when observations were very infrequent, very few or very noisy the scheme failed to recover state and parameter to an acceptable level of accuracy. The divergence of the filter in such cases can be associated with the observability of the state variables and the
parameter. The observability depends on several factors, including the distribution and frequency of observations. For example, let us assume that the observation frequency is such that the bed height is completing a whole cycle in the space domain. Then, it is impossible to detect whether the speed of the bed height is the true speed or simply the bed height is not moving at all and the speed is zero. Similarly, if the combination of the observation distribution and observation frequency is such that the peak of the bed height is never observed, and instead only observations of flat areas are available, then the speed cannot be recovered. These are just extreme cases, but they help to understand the observability underlying problem.

When the variance of the observations was large the assimilation scheme takes longer to recover the true parameter and state, even if observations are perfect. The fact that the scheme does not trust observations enough slows down the convergence speed of the scheme. When observations were generated according to the observation variance the error in the state estimate was lower bounded in proportion to the observation variance, which means that the quality of the assimilation cannot be improved if the quality of the observational data is not improved as well.

We noticed, specially when observations were noisy, that in general the parameter estimates oscillate around the true value, even if the state is very well approximated. We showed that averaging the parameter over a moving time window contributes to smooth the parameter estimates and in most of the cases the parameter approximation is very close to the truth at the end of the time window. Nevertheless, it was also shown that the averaging can slow down the process of recovering the parameter. This can cause the assimilation process to fail, even in cases where without applying the averaging strategy the state estimate was good. We discuss this issue in more detail in Section 7.4.

Figure 7.35: Advection model. RMSE for different $\alpha$’s and observation variances. Left: RMSE mean. Right: RMSE variance.
In general, the smaller $\alpha$ is chosen, the better results are obtained. Nevertheless, this behaviour is specific for this model, where the evolution of $B_{xx}$, according to the EKF update for the augmented problem, is closely related to the rank 1 matrix $N_k B_{pp} N_k^\top$ as can be seen by comparing figures 5.2 and 5.3. When experimenting with other models a different behaviour was observed, as shown in upcoming sections.

\section*{7.2 Nonlinear oscillator}

Our second model is a damped, nonlinear oscillator with two parameters, also used by Smith in [99]. In general, a damped, periodically forced nonlinear oscillator is governed by the second order differential equation

$$\ddot{x} + d\dot{x} + f(x) = e(t), \quad (7.8)$$

where $d$ is a damping coefficient, $x = x(t)$ a function of time, $f(x)$ a nonlinear restoring term and $e(t)$ a periodic function of $t$ and $\dot{x}, \ddot{x}$ denote the first and second time derivatives. This equation was first introduced by Duffing [37] and is therefore also known as the Duffing oscillator. It has been studied by many researchers and is used for describing a variety of situations. Interested readers can consult [22], [54], [105] and [110], among others. More about the steady states and regular and chaotic phenomena associated with this type of equations can be found in [120]. In particular we are interested in the case: $f(x) = mx + x^3$, $e(t) = 0$, for which (7.8) takes the form of:

$$\ddot{x} + d\dot{x} + mx + x^3 = 0. \quad (7.9)$$

Equation (7.9) can be used to describe the displacement of a single mass attached to a spring with nonlinear elasticity and linear damping. In this case $m, d > 0$, where $m$ is the square of the oscillation frequency and the nonlinear restoring force $-(mx + x^3)$ is equivalent to the force exerted by the spring under a displacement $x$ of the mass with respect to the equilibrium point. Equation (7.9) can be rewritten as a first order system of differential equations:

$$\begin{align*}
\dot{x} &= y \quad (7.10) \\
\dot{y} &= -dy - mx - x^3. \quad (7.11)
\end{align*}$$
In (7.10)-(7.11) \( x \) represents the relative position of the mass with respect to the equilibrium point and \( y \) its speed. For this system it is known that the only stable equilibrium is \((x, y) = (0, 0)\), which is obvious, considering which physical dynamical system is represented.

The motivation for experimenting with this simple example is to test how the method performs when applied to a nonlinear model with more than one unknown parameter.

### 7.2.1 Discretization

To solve (7.10)-(7.11) numerically we use a second order Runge-Kutta scheme. The resulting discrete system is

\[
x_{k+1} = \left(1 - m \frac{\Delta t^2}{2} - \frac{\Delta t^2}{2} x_k^2\right) x_k + \left(\Delta t - d \frac{\Delta t^2}{2}\right) y_k = f_1(x_k, y_k, m, d) \tag{7.12}
\]

\[
y_{k+1} = \left(-m \Delta t + dm \frac{\Delta t^2}{2} + \left(d \frac{\Delta t^2}{2} - \frac{\Delta t}{2}\right) x_k^2\right) x_k + \left(1 - d \Delta t - m \frac{\Delta t^2}{2} + d^2 \frac{\Delta t^2}{2}\right) y_k
- \frac{\Delta t}{2} (x_k + \Delta t y_k) = f_2(x_k, y_k, m, d). \tag{7.13}
\]

Equations (7.10) and (7.11) represent the nonlinear model \( f : \mathbb{R}^2 \to \mathbb{R}^2 \) considering \( m \) and \( d \) as fixed parameters. For solving the parameter estimation problem

\[
p_k = \begin{pmatrix} m_k \\ d_k \end{pmatrix} \tag{7.14}
\]

is added to the original state vector, to obtain the augmented-state

\[
w_k = \begin{pmatrix} x_k \\ y_k \\ m_k \\ d_k \end{pmatrix}. \tag{7.15}
\]
The two equations
\[ m_{k+1} = m_k, \quad (7.16) \]
\[ d_{k+1} = d_k, \quad (7.17) \]

stating that the parameters are constant, added to (7.12)-(7.13) constitute the augmented-state forecast model \( \tilde{f} : \mathbb{R}^4 \to \mathbb{R}^4 \).

### 7.2.2 State-parameter and state-state error covariances.

To obtain the state-parameter error covariance \( B_{xp} \) it is necessary to compute the derivatives of \( f \) with respect to both parameters:

\[
N_k = \left| \begin{array}{cc}
\frac{\partial f_1}{\partial m} & \frac{\partial f_1}{\partial d} \\
\frac{\partial f_2}{\partial m} & \frac{\partial f_2}{\partial d}
\end{array} \right|_{x_k^n, y_k^n, m_k^n, d_k^n}, \quad (7.18)
\]

where
\[
\frac{\partial f_1}{\delta m} = -\frac{\Delta t^2}{2} x_k, \quad (7.19)
\]
\[
\frac{\partial f_1}{\delta d} = -\frac{\Delta t^2}{2} y_k, \quad (7.20)
\]
\[
\frac{\partial f_2}{\delta m} = \left(-\Delta t + d \frac{\Delta t^2}{2} \right) - \frac{\Delta t^2}{2} y_k, \quad (7.21)
\]
\[
\frac{\partial f_2}{\delta d} = m \frac{\Delta t^2}{2} x_k + \frac{\Delta t^2}{2} x_k^3 - (\Delta t - d \Delta t^2) y_k. \quad (7.22)
\]

Even for such a simple problem, to have an a priori estimation of the parameter errors can be challenging. In the experiments we assume that the parameters are uncorrelated, meaning that \( B_{pp} \) is the diagonal matrix

\[
B_{pp} = \begin{pmatrix}
\sigma_m^2 & 0 \\ 0 & \sigma_d^2
\end{pmatrix}, \quad (7.23)
\]

where \( \sigma_m^2, \sigma_d^2 \) are the variances of the parameters \( m \) and \( d \) respectively. The state-parameter and state-state error covariance updates are

\[
(B_{xp})_k = N_k B_{pp} = \left| \begin{array}{cc}
\sigma_m^2 \frac{\partial f_1}{\partial m} & \sigma_d^2 \frac{\partial f_1}{\partial d} \\
\sigma_m^2 \frac{\partial f_2}{\partial m} & \sigma_d^2 \frac{\partial f_2}{\partial d}
\end{array} \right|_{x_k^n, y_k^n, m_k^n, d_k^n}, \quad (7.24)
\]
\[
(B_{xx})_k = \alpha B_{xx}^b + (1 - \alpha)(B_{xp})_k N_k^\top \quad (7.25)
\]
\[(B_{xx})_k = \alpha B_{xx}^b + (1 - \alpha) \begin{pmatrix} \sigma^2_m \left( \frac{\partial f_1}{\partial m} \right)^2 + \sigma^2_d \left( \frac{\partial f_1}{\partial d} \right)^2 & \sigma^2_m \frac{\partial f_1}{\partial m} \frac{\partial f_2}{\partial m} + \sigma^2_d \frac{\partial f_1}{\partial d} \frac{\partial f_2}{\partial d} \\ \sigma^2_m \frac{\partial f_1}{\partial m} \frac{\partial f_2}{\partial m} + \sigma^2_d \frac{\partial f_1}{\partial d} \frac{\partial f_2}{\partial d} & \sigma^2_m \left( \frac{\partial f_2}{\partial m} \right)^2 + \sigma^2_d \left( \frac{\partial f_2}{\partial d} \right)^2 \end{pmatrix} \right|_{x_k^a, y_k^a, m_k^a, d_k^a} \tag{7.26}
\]

### 7.2.3 Experimental set-up

Twin experiments are conducted considering that the discretized augmented model $\tilde{f}$ is perfect. In the experiments we use a similar set-up as the one used by Smith in [101] in order to make the comparison easier. The discretization step is $\Delta t = 0.1$. The initial true displacement is $x^0 = 2$ and the initial true speed $y^0 = 0$. The true parameters are $d^t = 0.05$ and $m^t = 1$. We assume that the parameters are uncorrelated and their respective variances are $\sigma^2_d = 0.01$ and $\sigma^2_m = 0.1$. We generate pseudo-observations for both displacement and speed, in regular time intervals, using the true state trajectory and adding random noise of zero mean and variance $\sigma^2_o = 0.01$, so

\[R = \begin{pmatrix} \sigma^2_o & 0 \\ 0 & \sigma^2_o \end{pmatrix} \tag{7.27}\]

and

\[\tilde{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{7.28}\]

The initial background displacement and speed are generated adding random noise with variance $\sigma^2_b = 0.01$ to the true values. The initial state errors are assumed to be uncorrelated, therefore

\[B_{xx}^b = \begin{pmatrix} \sigma^2_b & 0 \\ 0 & \sigma^2_b \end{pmatrix}. \tag{7.29}\]

### 7.2.4 Experimental results

We conducted a large number of experiments, varying the initial value of the parameters and the observation frequency and variance. We observed that if the damping parameter $d$ is relatively large ($d^t > 0.2$), it is difficult to obtain a quality assimilation, because the system is damped very fast to the steady state, where the parameters are no longer identifiable. We present the results obtained for the initial estimated parameters $d_0 = 0.08$, $m_0 = 0.6$ and $\alpha = 0.8$.

**Perfect observations**

In Figures 7.37 and 7.39 the assimilated values of the state variables $x$ and $y$, respectively, are shown for observation frequencies between $5\Delta t$ and $50\Delta t$. For frequencies up
to $25\Delta t$ the scheme is able to recover both variables to a high level of accuracy. For these frequencies it is very difficult to distinguish the true values of the state variables from their estimations by looking at the plot. To make the comparison easier the absolute values of the estimation errors for both state variables are shown in Figures 7.38 and 7.40. Larger frequencies require a longer assimilation time window (for example for $\Delta t_{\text{obs}} = 25$) to recover the state and parameters. If $\Delta t_{\text{obs}} = 50$ the quality of the estimation is very poor. During the assimilation step the state is pulled towards the observations but the oscillations’ amplitude remains small. This is because the parameter estimation is also deficient, as shown in Figures 7.41 and 7.42. In these figures a dependency between convergence speed and observation frequency can be observed. At the end of the assimilation window the estimation of $m$ is very accurate for all assimilation frequencies, except for $50\Delta$. The estimation of $d$ is also precise, though for $\Delta t_{\text{obs}} = 25$ a small difference between the assimilated and true value can be appreciated.

Another set of experiments was conducted to investigate how the observation variance affects the assimilation. Though observations are perfect, to which extent we trust in them has an impact on the assimilation quality. In Figures 7.43 and 7.44 the assimilated values of the state variables $x$ and $y$ are shown for observations taken every $10\Delta t$ and observation variances between 0.001 and 0.5. For variances between 0.001 and 0.1 the state estimation is very accurate after $t = 70$ and it is not possible to distinguish the difference between the true and the estimated value.
Figure 7.38: Nonlinear oscillator. $|x^t - x^a|$ for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta).

Figure 7.39: Nonlinear oscillator. Assimilated $y$ for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue, almost not distinguishable), $5\Delta t$ (green, almost not distinguishable), $10\Delta t$ (red, almost not distinguishable), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). Observations for $\Delta t_{\text{obs}}=50$ (blue circles). $y^t$ (black).
Figure 7.40: Nonlinear oscillator. $|y^t - y^a|$ for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta).

Figure 7.41: Nonlinear oscillator. Assimilated $m$ for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $m^t$ (black).
Chapter 7. Nonlinear experiments

Figure 7.42: Nonlinear oscillator. Assimilated $d$ for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $d^t$ (black).

Figure 7.43: Nonlinear oscillator. Assimilated $x$ for different observation variances and perfect observations, $\Delta t_{\text{obs}} = 10$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.05$ (red), $\sigma_o^2 = 0.1$ (cyan) and $\sigma_o^2 = 0.5$ (magenta). Observations for $\sigma_o^2 = 0.01$ (blue circles). $x^t$ (black).
Figure 7.44: Nonlinear oscillator. Assimilated $y$ for different observation variances and perfect observations, $\Delta t_{\text{obs}} = 10$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.05$ (red), $\sigma_o^2 = 0.1$ (cyan) and $\sigma_o^2 = 0.5$ (magenta). Observations for $\sigma_o^2 = 0.01$ (blue circles). $y^t$ (black).

Figure 7.45: Nonlinear oscillator. Assimilated $m$ for different observation variances and perfect observations, $\Delta t_{\text{obs}} = 10$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.05$ (red), $\sigma_o^2 = 0.1$ (cyan) and $\sigma_o^2 = 0.5$ (magenta). $m^t$ (black).
Chapter 7. *Nonlinear experiments*

139

![Figure 7.46: Nonlinear oscillator. Assimilated $d$ for different observation variances and perfect observations, $\Delta t_{\text{obs}} = 10$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.05$ (red), $\sigma_o^2 = 0.1$ (cyan) and $\sigma_o^2 = 0.5$ (magenta). $d^t$ (black).](image)

If $\sigma_o^2 = 0.5$, the scheme fails to estimate the state to any acceptable level of accuracy. In general the assimilated parameters converge faster to the true value as the variance value becomes smaller, as shown in Figures 7.45 and 7.46. Even for $\sigma_o^2 = 0.5$ the estimation of $m$ is very accurate. The assimilation fails because the scheme fails to recover the parameter $d$. Throughout all conducted experiments it was consistently observed that the relative error in the estimation of $m$ was smaller than the one in the estimation of $d$.

**Noisy observations**

Similar experiments were conducted, this time considering noisy observations.

In Figures 7.47-7.50 we show the results for different assimilation frequencies. Like for the perfect observation experiments, accurate estimates of $x$ and $y$ are obtained for observation frequencies up to $25\Delta t$, while for a $50\Delta t$ frequency the quality of the analysis is very poor. The analysis is as expected not so smooth as previously due to the effect of observations pulling the state out of the model trajectory. This is especially noticeable when observations are assimilated at every time step. The parameters do not converge to a specific value, but oscillate near the true solution. The relative error of the assimilated $m$ after $t = 80$ is at most 0.2, up to an observation frequency of $25\Delta t$.

A different situation is observed when the assimilated $d$ is analysed. For $\Delta t_{\text{obs}} = 1$ the estimation stays close to the true value until $t = 50$, after which it starts to augment.
Chapter 7. *Nonlinear experiments* 140

At the end of the time window the relative error is slightly above 1.5. A large relative error is also obtained for $\Delta t_{\text{obs}} = 25$.

We repeated the experiments averaging the parameters along a moving time window, as in the advection experiments. The results are shown in Figures 7.51 and 7.52. Averaging contributes to the improvement of the parameter estimation. The relative error of the estimation of $m$ is again smaller with respect to the one corresponding to the estimation of $d$. The best results are obtained for $\Delta t_{\text{obs}} = 10$. A reason could be that for this frequency of observations the observed values can reveal more about the dynamics of the system. This is closely related to the observability of the system in general and is an aspect to take care about in these experiments. For example, it can happen that the observation frequency is such that it is not possible to distinguish between two different sets of speed, position and parameters, one in which the mass oscillates 3 times between two consecutive observations and another where it oscillates only once (and therefore moves faster). This phenomenon is known as aliasing and refers to the existence of two signals (with different frequencies) which when sampled at a certain sample rate produce the same output and are therefore indistinguishable from one another. In this case it could not be identified for example, which is the maximum displacement and speed at different points and therefore the rate of decrease of the oscillation, which is closely related to the parameter $d$.

![Figure 7.47: Nonlinear oscillator. Assimilated $x$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). Observations for $\Delta t_{\text{obs}} = 50$ (blue circles). $x^2$ (black).](image-url)
Figure 7.48: Nonlinear oscillator. Assimilated $y$ for different observation frequencies and noisy observations, $\sigma_y^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). Observations for $\Delta t_{\text{obs}} = 50$ (blue circles). $y'$ (black).

Figure 7.49: Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma_y^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $m'$ (black).
Chapter 7. Nonlinear experiments

Figure 7.50: Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $d^t$ (black).

Figure 7.51: Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$, where the parameter estimates are averaged over a moving time window. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $m^t$ (black). The vertical dashed black line represents the point from which the parameter is averaged.
Figure 7.52: Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$, where the parameter estimates are averaged over a moving time window. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $d^t$ (black). The vertical dashed black line represents the point from which the parameter is averaged.

Figure 7.53: Nonlinear oscillator. Assimilated $x$ for different observation variances and noisy observations, $\Delta t_{\text{obs}} = 10$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.05$ (red), $\sigma_o^2 = 0.1$ (cyan) and $\sigma_o^2 = 0.5$ (magenta). Observations for $\sigma_o^2 = 0.5$ (blue circles). $x^t$ (black).
In Figures 7.53 and 7.54 the sequence of $x$ and $y$ estimations are shown for different observation variances and observations assimilated every $10\Delta t$. For very small variances, after a few assimilation steps, it cannot be distinguished the difference between assimilated and true state. They start being noticeable for $\sigma_o^2 = 0.01$ but stay under an acceptable level, except for $\sigma_o^2 = 0.5$. To give an idea of how noisy the corresponding observations are we have represented them in the aforementioned figures. It is expected that any assimilation scheme fails if the observations are so inaccurate. Looking at figures 7.55 and 7.56 it is clear that for $\sigma_o^2 = 0.5$ our method fails to properly estimate $d$, which causes the imprecise state estimation. The estimation of $m$ is nevertheless much better (the relative error is of 0.1 at the end of the assimilation window, which is much smaller than the relative error of 1.9 in the $d$ estimate). For the rest of the variances tested both parameters are recovered to a good level of accuracy, being more precise, as expected, for smaller observation variances.

![Figure 7.54: Nonlinear oscillator. Assimilated $y$ for different observation variances and noisy observations, $\Delta t_{obs} = 10$. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.05$ (red), $\sigma_o^2 = 0.1$ (cyan) and $\sigma_o^2 = 0.5$ (magenta). Observations for $\sigma_o^2 = 0.5$ (blue circles). $y^i$ (black).]
Figure 7.55: Nonlinear oscillator. Assimilated $m$ for different observation variances and noisy observations, $\Delta t_{\text{obs}} = 10$, where the parameter estimates are averaged over a moving time window. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.05$ (red), $\sigma_o^2 = 0.1$ (cyan) and $\sigma_o^2 = 0.5$ (magenta). $\bar{m}$ (black). The vertical dashed black line represents the point from which the parameter is averaged.

Figure 7.56: Nonlinear oscillator. Assimilated $d$ for different observation variances and noisy observations, $\Delta t_{\text{obs}} = 10$, where the parameter estimates are averaged over a moving time window. Assimilations for $\sigma_o^2 = 0.001$ (blue), $\sigma_o^2 = 0.01$ (green), $\sigma_o^2 = 0.05$ (red), $\sigma_o^2 = 0.1$ (cyan) and $\sigma_o^2 = 0.5$ (magenta). $\bar{d}$ (black). The vertical dashed black line represents the point from which the parameter is averaged.
Unobserved states

Experiments were conducted observing only one of the two state variables. In this case $\tilde{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$ or $\tilde{H} = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}$, depending on the variable observed.

![Figure 7.57: Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma^2 = 0.01$ and only $y$ is observed, where the parameter estimates are averaged over a moving time window. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $m^t$ (black). The vertical dashed black line represents the point from which the parameter is averaged.](image)

In Figures 7.57 and 7.58 the assimilated $m$ and $d$ are shown, when only $y$ is observed. The quality of the analysis is very similar to that when both state variables are observed. The parameters are recovered to an acceptable level of accuracy, except when observations are available only every $50\Delta t$ time steps. A very different situation occurs if only $x$ is observed. The estimation of $m$ is very similar for all observation frequencies, but the relative error never becomes smaller than 0.2 (see Figure 7.59). From Figure 7.60 it is clear that only for observations every $25\Delta t$ and $50\Delta t$ the estimation of $d$ is acceptable, though for those frequencies worse estimates of $m$ are obtained. From the experiments it can be concluded that the speed of the attached mass can reveal more about the system parameters than its position. This could be expected from analysing the observability of the state and parameters given the two different observational data sets.
Figure 7.58: Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$ and only $y$ is observed, where the parameter estimates are averaged over a moving time window. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $d^t$ (black). The vertical dashed black line represents the point from which the parameter is averaged.

Figure 7.59: Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$ and only $x$ is observed, where the parameter estimates are averaged over a moving time window. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $m^t$ (black). The vertical dashed black line represents the point from which the parameter is averaged.
Chapter 7. Nonlinear experiments

Figure 7.60: Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma_\theta^2 = 0.01$ and only $x$ is observed, where the parameter estimates are averaged over a moving time window. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $d^t$ (black). The vertical dashed black line represents the point from which the parameter is averaged.

Figure 7.61: Nonlinear oscillator. Assimilated $m$ for different observation frequencies and noisy observations, $\sigma_\theta^2 = 0.01$ and only $x$ is observed, where the parameter estimates are averaged over a moving time window. The state variables are assumed to be correlated with correlation $\sigma_{xy} = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $m^t$ (black). The vertical dashed black line represents the point from which the parameter is averaged.
Chapter 7. Nonlinear experiments

Figure 7.62: Nonlinear oscillator. Assimilated $d$ for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$ and only $x$ is observed, where the parameter estimates are averaged over a moving time window. The state variables are assumed to be correlated with correlation $\sigma_{xy} = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $25\Delta t$ (cyan) and $50\Delta t$ (magenta). $d^t$ (black). The vertical dashed black line represents the point from which the parameter is averaged.

We repeated the experiment, this time considering a positive cross-covariance between the two state variables in $B_{xx}^b$ ($\sigma_{xy} = B_{xx1,2}^b = B_{xx2,1}^b = 0.01$). The results are shown in Figures 7.61 and 7.62. The estimation of $m$ gets closer to the true value, but the real improvement is the estimation of $d$. Only for $\Delta t_{\text{obs}} = 25$ and $\Delta t_{\text{obs}} = 25$ the estimation is not precise. As before, the estimation of $d$ when observations are available at every time step is worse than for observations every $5\Delta t$ and $10\Delta t$. By adding the correlation between state variables the unobserved variable "knows" in which direction and proportion its estimate should move given the update of the other variable according to the observation available. Considering the right structure of the state-state error covariance the assimilation results are improved. Nevertheless, it is not clear why if the same $B_{xx}^b$ matrix is used when both state variables are observed there is no improvements in the quality of the results.

Selection of $\alpha$

We investigated how the selection of $\alpha$ influences the quality of the state assimilation. For this purpose 20 independent experiments were conducted, combining different $\alpha$’s and observation frequencies (a similar analysis considering different observation variances can be found in Appendix C.1). For each combination the mean and the variance of the RMSE for the state estimates were computed (Figure 7.63). We also computed, for each
Chapter 7. Nonlinear experiments

If $\Delta t_{\text{obs}} \geq 35$, the state RMSE is larger than 0.3, regardless of the size of $\alpha$. This corresponds to a variance 9 times larger than the observation variance. In general, the longer the time between assimilations, the larger $\alpha$ must be set to obtain accurate results. The large RMSE is caused by a very imprecise estimation of the parameter $d$ (the relative error is larger than 1) for large assimilation frequencies. Notice the similarities between the white regions for the mean values in figures 7.63 and 7.65. Furthermore, the variance of the relative error of the assimilated $d$ is considerably large, which means that the results are in general not consistent.

For assimilation frequencies between 1 and 10, smaller state RMSE and parameter relative errors were obtained for values of $\alpha$ between 0.1 and 0.5 (see the predominant blue
Figure 7.65: Nonlinear oscillator. Relative error for the assimilated parameter $\alpha$ for different $\alpha$’s and observation variances. Left: $|\frac{d-a}{d} - a|$ mean. Right: $|\frac{d-a}{d} - a|$ variance.

colour near the lower left corner of the mean plots). On the other hand, for $\Delta t_{\text{obs}} = 20$ or $\Delta t_{\text{obs}} = 30$ the best results are obtained setting $\alpha = 1$.

In general, the parameter $m$ is easier to recover than the parameter $d$ no matter which $\alpha$ is used. Difficulties in estimating the parameter $d$ also arose when the full EKF update was considered (see Appendix C.2). In this case the quality of the assimilation was much worse than the one obtained with the LRSSC. An explanation can be probably found analysing the observability conditions for $d$. In order to have useful information to estimate the rate of damping of the system it is necessary to observe the extreme positions (or near) of the mass along the time window. For example, if the mass is observed only when the displacement is 0 or close to 0, then little can be said about the change in the amplitude of the oscillations, which is equivalent to how the system is damped over time. Observing these extreme positions depends on where the observations are taken, therefore the frequency of observations influences the assimilation. On the other hand, if observations are very noisy then the location of the mass is vaguely known and the information coming from the observational data is not useful to estimate the damping rate, as shown for $\sigma_d^2 = 0.5$.

7.2.5 Summary

In this section we presented the experimental results for the LRSSC applied to the nonlinear oscillator problem. Like for the advection model, the method performed very well in most of the situations tested for perfect and noisy observations. The assimilation results were consistent with the error statistics, obtaining state estimates whose errors were smaller than the standard deviation of observations.
In general there was a marked difference between the accuracy in the estimation of both parameters. The difficulties in recovering the parameter \(d\) are associated with more restrictive observability conditions for this parameter. The same difficulties were observed when the EKF was applied.

If only the displacement of the mass is observed it is not possible to recover both parameters unless extra correlation between the state variables is included. In order to recover the parameter information about the speed of the mass is needed (related to the observability conditions for the parameters), which is not available in case the speed is not directly observed or its value cannot be inferred from observing the displacement by using the speed-displacement correlations.

A substantial difference with respect to what was observed in the advection experiments is that the optimality of the scheme parameter \(\alpha\) depends on the frequency of the observations. In general, for frequent observations smaller \(\alpha\)’s were preferable, while for observations very sparse in time better results were obtained for larger \(\alpha\)’s. It is not unreasonable to think that if observations are frequent, then short time-scale errors play an important role in the correction of the state error statistics, while if the observations are sparse in time then depending on the model long time-scale errors can be more relevant than time-specific errors.

\section{Lorenz 63}

The Lorenz 63 model was first introduced by Lorenz in 1963 \cite{lorenz1963}. It arises from a simplified system of differential equations, to describe a two-dimensional flow of fluid of uniform depth, and is defined by the 3 coupled nonlinear differential equations

\begin{align}
\frac{dx}{dt} &= \sigma(y - x), \quad (7.30) \\
\frac{dy}{dt} &= x(\rho - z) - y, \quad (7.31) \\
\frac{dz}{dt} &= xy - \beta z, \quad (7.32)
\end{align}

where \(x = x(t)\) represents the convective intensity, \(y = y(t)\) the difference in temperatures between the ascending and descending currents and \(z = z(t)\) the deviation of the temperature profile from linearity. The model parameters are: the Prandtl number \(\sigma\), the normalized Rayleigh number \(\rho\) and a geometrical parameter \(\beta\). For more insight about the physical meaning and characteristics of the model the interested reader is remitted to \cite{lorenz1982}, \cite{guckenheimer1982} and \cite{hering2005}. Because of its strong non-linearity, chaotic behaviour
and its instability properties, this model has been used as test problem to analyse the performance of DA methods ([81], [99], [42]).

The described dynamical system has 1 stationary point corresponding to no convection at \((0, 0, 0)\). If \(\rho > 1\), another two stationary points arise: \((\sqrt{\beta(\rho - 1)}, \sqrt{\beta(\rho - 1)}, \rho - 1)\) and \((-\sqrt{\beta(\rho - 1)}, -\sqrt{\beta(\rho - 1)}, \rho - 1)\), corresponding to steady convection. A classic parameter set-up is the one proposed by Lorenz: \(\sigma = 10, \rho = 28\) and \(\beta = \frac{8}{3}\). For these set of parameters none of the three stationary points is stable and the behaviour of the system is chaotic. In our experiments we assume that these are the true parameters.

### 7.3.1 Discretization

For discretizing the model we use a simple Runge Kutta of second order scheme. The resulting discrete model is

\[
\begin{align*}
x_{k+1} &= x_k + \frac{\Delta t}{2} \left(h_k^{1,x} + h_k^{2,x}\right), \\
y_{k+1} &= y_k + \frac{\Delta t}{2} \left(h_k^{1,y} + h_k^{2,y}\right), \\
z_{k+1} &= z_k + \frac{\Delta t}{2} \left(h_k^{1,z} + h_k^{2,z}\right),
\end{align*}
\]

with

\[
\begin{align*}
h_k^{1,x} &= \sigma (y_k - x_k), \\
h_k^{1,y} &= \rho x_k - y_k - x_k z_k, \\
h_k^{1,z} &= x_k y_k - \beta z_k, \\
h_k^{2,x} &= \sigma \left[\left(y_k + \Delta th_k^{1,y}\right) - \left(x_k + \Delta th_k^{1,x}\right)\right], \\
h_k^{2,y} &= \rho \left(x_k + \Delta th_k^{1,x}\right) - \left(y_k + \Delta th_k^{1,y}\right) h_k^{2,y} - \left(x_k + \Delta th_k^{1,x}\right) \left(z_k + \Delta th_k^{1,z}\right), \\
h_k^{2,z} &= \left(x_k + \Delta th_k^{1,x}\right) \left(y_k + \Delta th_k^{1,y}\right) - \beta \left(z_k + \Delta th_k^{1,z}\right),
\end{align*}
\]

and the step-size is denoted by \(\Delta t\). The resulting augmented state vector and forecast operator are:

\[
w_k = \begin{pmatrix} x_k \\ y_k \\ z_k \\ \sigma_k \\ \rho_k \\ \beta_k \end{pmatrix}, \quad w_{k+1} = \tilde{f}(w_k),
\]

\[
de fining the system using the following equations: \(\frac{\Delta t}{2} \left(h_k^{1,x} + h_k^{2,x}\right), \frac{\Delta t}{2} \left(h_k^{1,y} + h_k^{2,y}\right), \frac{\Delta t}{2} \left(h_k^{1,z} + h_k^{2,z}\right), \sigma (y_k - x_k), \rho x_k - y_k - x_k z_k, x_k y_k - \beta z_k, \sigma \left[\left(y_k + \Delta th_k^{1,y}\right) - \left(x_k + \Delta th_k^{1,x}\right)\right], \rho \left(x_k + \Delta th_k^{1,x}\right) - \left(y_k + \Delta th_k^{1,y}\right) h_k^{2,y} - \left(x_k + \Delta th_k^{1,x}\right) \left(z_k + \Delta th_k^{1,z}\right), \left(x_k + \Delta th_k^{1,x}\right) \left(y_k + \Delta th_k^{1,y}\right) - \beta \left(z_k + \Delta th_k^{1,z}\right),
\end{align*}
\]
with \( \tilde{f} : \mathbb{R}^6 \rightarrow \mathbb{R}^6 \) resulting from coupling the parameter evolution equations

\[
\begin{align*}
\sigma_{k+1} &= \sigma_k \tag{7.43} \\
\rho_{k+1} &= \rho_k \tag{7.44} \\
\beta_{k+1} &= \beta_k \tag{7.45}
\end{align*}
\]

to the state forecast \( f : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) represented by the system (7.33)-(7.35).

### 7.3.2 State-parameter and state-state error covariances.

To obtain the state-parameter error covariance \( \mathbf{B}_{xp} \), it is necessary to compute the derivatives of \( f \) with respect to the parameters:

\[
\mathbf{N}_k = \left( \begin{array}{ccc}
\frac{\delta f_1}{\delta \sigma} & \frac{\delta f_1}{\delta \rho} & \frac{\delta f_1}{\delta \beta} \\
\frac{\delta f_2}{\delta \sigma} & \frac{\delta f_2}{\delta \rho} & \frac{\delta f_2}{\delta \beta} \\
\frac{\delta f_3}{\delta \sigma} & \frac{\delta f_3}{\delta \rho} & \frac{\delta f_3}{\delta \beta}
\end{array} \right)
\left( \begin{array}{c}
x^a_k \\
y^a_k \\
z^a_k
\end{array} \right)
\]

(7.46)

We assume that the parameters are uncorrelated and therefore

\[
\mathbf{B}_{pp} = \begin{pmatrix}
\sigma^2_{\sigma} & 0 & 0 \\
0 & \sigma^2_{\rho} & 0 \\
0 & 0 & \sigma^2_{\beta}
\end{pmatrix},
\]

(7.47)

where \( \sigma^2_{\sigma} \), \( \sigma^2_{\rho} \) and \( \sigma^2_{\beta} \) are the variances of the parameters \( \sigma \), \( \rho \) and \( \beta \) respectively. The state-parameter and state-state error covariance are updated according to equations (5.1) and (5.20), respectively.

### 7.3.3 Experimental set-up

Twin experiments were conducted considering that the discretized augmented-model \( \tilde{f} \) is perfect. As step-size we used \( \Delta t = 0.01 \). Like in [101] and [15], the initial true state variables are \( x^t_0 = -5.4458 \), \( y^t_0 = -5.4841 \) and \( z^t_0 = -22.5606 \). The initial background state and parameter were generated by perturbing the true state and parameters, respectively, using Gaussian noise generated with the statistical distribution specified by the respective error covariances. In the experiments presented in this section \( x^b_0 = -5.08 \), \( y^b_0 = -6.11 \), \( z^b_0 = 23.07 \), \( \sigma^b = 10.97 \), \( \rho^b = 30.15 \) and \( \beta^b = 1.72 \). The parameter variances are set to \( \sigma^2_{\sigma} = 2 \), \( \sigma^2_{\rho} = \sigma^2_{\beta} = 1 \). We generate pseudo-observations for the three state
variables adding random noise with variance \( \sigma_o^2 = 0.01 \) to the truth’s trajectory, so

\[
R = \sigma_o^2 I_{3 \times 3}
\]

(7.48)

and

\[
\tilde{H} = \begin{pmatrix} I_{3 \times 3} & 0_{3 \times 3} \end{pmatrix}
\]

(7.49)

The initial state errors are assumed to be uncorrelated with variance \( \sigma_x^2 \):

\[
B_{xx}^b = \sigma_x^2 I_{3 \times 3}.
\]

(7.50)

In our experiments \( \sigma_x^2 = 1 \).

### 7.3.4 Experimental results

#### Perfect observations

Experiments were conducted for different observation frequencies and variances, considering perfect observations. Though various values of \( \alpha \) were tested, we show the results obtained for \( \alpha = 0.3 \) to illustrate the performance of the method. Later in this section we will analyse the results for other values of \( \alpha \).

![Figure 7.66: Lorenz 63. State assimilation for perfect observations, \( \Delta t = 20 \) and \( \sigma_o^2 = 0.01 \). Assimilated value (magenta) and true value of the state (black).](image)

We conducted assimilation experiments for observation frequencies between 1 and 20. In Figure 7.66 the assimilation of the state variables is shown just for \( \Delta t_{\text{obs}} = 20 \), as
Figure 7.67: Lorenz 63. Parameter assimilation for different observation frequencies and perfect observations, with $\sigma_0^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $15\Delta t$ (cyan) and $20\Delta t$ (magenta). True parameters (black).

Figure 7.68: Lorenz 63. Norm of the assimilated state error vector for perfect observations and $\sigma_0^2 = 0.01$. Assimilations for $\Delta t_{\text{obs}} = 1$ (blue), $\Delta t_{\text{obs}} = 5$ (green), $\Delta t_{\text{obs}} = 10$ (red), $\Delta t_{\text{obs}} = 15$ (cyan) and $\Delta t_{\text{obs}} = 20$ (magenta).

For more frequent observations the assimilated state cannot be distinguished from the true state by looking at the plot. Only at the beginning of the assimilation window differences between the true and assimilated states are perceptible. In Figure 7.67 we show the evolution of the parameter estimation along the time window. The more often
assimilations are observed the faster the convergence of the parameter estimates, but in general the parameter estimates at the end of the assimilation window are very accurate for all the observation frequencies, being the largest relative error of the order of $10^{-2}$ for the parameter $\sigma$ when observing every 20 time steps.

A good approximation of the state depends on a precise estimation of the parameter. This is why the norm of the state estimation error vector decreases faster for shorter assimilation periods as shown in Figure 7.68. Nevertheless, after $t = 12$ both state and parameter estimation are very precise for all the observation frequencies shown. We consistently observed that it takes longer for the $\sigma$ estimates to converge to the true solution. This can be explained by looking at the sensitivity of the model with respect to the parameter $\sigma$. In [79] Marzban shows that the state variable $z$ is practically insensitive to $\sigma$. We cannot expect to recover a parameter accurately if changes in its value produce no change in the output of the system, or a change which is smaller than the observation noise. In our experiments setting a larger variance for this parameter in comparison to the variance of the other two, contributes to a faster convergence of the scheme. When longer periods without observations were considered, the method failed to recover parameters and states with an acceptable precision. The results could be improved when larger values of $\alpha$ were considered, as we will discuss at the end of this section.

![Figure 7.69: Lorenz 63. Parameter assimilation for different observation variances and perfect observations assimilated every $10 \Delta t$. Assimilations for $\sigma_0^2 = 0.001$ (blue), $\sigma_0^2 = 0.01$ (green), $\sigma_0^2 = 0.05$ (red), $\sigma_0^2 = 0.1$ (cyan) and $\sigma_0^2 = 0.5$ (magenta). True parameters (black).](image)
Other experiments were conducted for different observation variances in the range 0.001 to 0.5 (though using perfect observations) and observations available every 10 time steps. From Figures 7.69 and 7.70 can be concluded that the less we trust in the observations the slower the convergence. Again, after $t = 12$ the precision of parameter and state estimation is very accurate.

Noisy observations

Similar experiments were conducted, this time for noisy observations. First we considered observation frequencies $\Delta t_{\text{obs}} = 1, 5, 10, 15, 20$. Like in previous experiments, in the presence of noisy observations the parameter estimates oscillate close to the corresponding true values (see Figure 7.71). Better and more stable estimates are obtained for $\rho$, while it takes longer to obtain a good approximation for $\sigma$. Figure 7.72 reflects the big impact of the observation frequency on the assimilation quality. There is a huge difference between the norm of the error for $\Delta t_{\text{obs}} = 1, 5, 10$ and that for $\Delta t_{\text{obs}} = 15$ or $\Delta t_{\text{obs}} = 20$.

Again, much better results are obtained when the parameters are averaged after a tune-up phase, as shown in Figure 7.73. Not only the parameter estimations are more stable and precise, but the estimation error is considerably decreased (see Figure 7.74). For larger observation frequencies the assimilation fails to improve the initial state and
parameter estimation. For such cases better results are obtained if $\alpha$ is augmented, as discussed later in this section.

**Figure 7.71:** Lorenz 63. Parameter assimilation for different observation frequencies and noisy observations, with $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $15\Delta t$ (cyan) and $20\Delta t$ (magenta). True parameter (black).

**Figure 7.72:** Lorenz 63. Norm of the assimilated state error vector for noisy observations and $\sigma_o^2 = 0.01$. Assimilations for $\Delta t_{\text{obs}} = 1$ (blue), $\Delta t_{\text{obs}} = 5$ (green), $\Delta t_{\text{obs}} = 10$ (red), $\Delta t_{\text{obs}} = 15$ (cyan) and $\Delta t_{\text{obs}} = 20$ (magenta).
Figure 7.73: Lorenz 63. Parameter assimilation for different observation frequencies and noisy observations, with $\sigma_0^2 = 0.01$, where the parameter estimates are averaged over a moving time window. Assimilations every $\Delta t$ (blue), $5\Delta t$ (green), $10\Delta t$ (red), $15\Delta t$ (cyan) and $20\Delta t$ (magenta). The true value of the parameters is given by the black line. The dashed black line indicates the point from which the parameter is averaged.

Figure 7.74: Lorenz 63. Norm of the assimilated state error vector for noisy observations and $\sigma_0^2 = 0.01$, where the parameter estimates are averaged over a moving time window. Assimilations for $\Delta t_{\text{obs}} = 1$ (blue), $\Delta t_{\text{obs}} = 5$ (green), $\Delta t_{\text{obs}} = 10$ (red), $\Delta t_{\text{obs}} = 15$ (cyan) and $\Delta t_{\text{obs}} = 20$ (magenta). The dashed black line indicates the point from which the parameter is averaged.
Figure 7.75: Lorenz 63. State assimilation for $\sigma_o^2 = 0.5$ and noisy observations assimilated every $10\Delta t$, where the parameter estimates are averaged over a moving time window. Assimilated state (magenta), true value of the states (black) and observations (blue circles). The dashed black line indicates the point from which the parameter is averaged.

Other experiments were conducted considering different observation variances and observation frequency $\Delta t_{\text{obs}} = 10$. We only show the results obtained when the parameter estimations are averaged after $t = 10$, as the quality of the assimilation is considerably better than in the case without averaging. Only for very noisy observations ($\sigma_o^2 = 0.5$) the difference between the assimilated and true states is noticeable (Figure 7.75). The best parameter estimate is obtained for $\beta$, regardless of how large the observation variance is (see Figure 7.76). The estimates for $\sigma$ and $\rho$ are extremely precise for variances up to 0.05 and very good for larger variances. The largest relative error for the parameter estimations at the end of the time windows is smaller than 0.017, which shows the effectiveness of the scheme. In Figure 7.77 it can be appreciated how noisier observations affect the quality of the state assimilation. The error norm increases considerably when the observation variance is increased, though it is kept under acceptable levels after the tune-up phase.
Figure 7.76: Lorenz 63. Parameter assimilation for noisy observations assimilated every $10\Delta t$, where the parameter estimates are averaged over a moving time window. Assimilations for $\sigma_0^2 = 0.001$ (blue), $\sigma_0^2 = 0.01$ (green), $\sigma_0^2 = 0.05$ (red), $\sigma_0^2 = 0.1$ (cyan) and $\sigma_0^2 = 0.5$ (magenta). The true value of the parameters is given by the black line. The dashed black line indicates the point from which the parameter is averaged.

Figure 7.77: Lorenz 63. Norm of the assimilated state error vector for noisy observations assimilated every $10\Delta t$, where the parameter estimates are averaged over a moving time window. Assimilations for $\sigma_0^2 = 0.001$ (blue), $\sigma_0^2 = 0.01$ (green), $\sigma_0^2 = 0.05$ (red), $\sigma_0^2 = 0.1$ (cyan) and $\sigma_0^2 = 0.5$ (magenta). The true value of the parameters is given by the black line. The dashed black line indicates the point from which the parameter is averaged.
Selection of $\alpha$

A large number of experiments were conducted aiming to determine an optimal value for the scheme parameter $\alpha$, depending on the observation variance and frequency. In Figure 7.78 we show the RMSE for different combinations of $\alpha$-$\Delta t_{\text{obs}}$. This RMSE is measured including state and parameter variables, from $t = 10$ to $t = 15$. Though the data shown corresponds to the results for single runs of the assimilation scheme, it is representative to the behaviour observed during the extensive experimenting phase.

The “best” $\alpha$ varies according to the observation frequency. When observations are available at every time step, there is almost no difference when using different $\alpha$ values. Then, up to $\Delta t_{\text{obs}}=15$ almost any selection of $\alpha$ improve the results with respect to $\alpha = 1$. For example, notice that if $\alpha = 0.6$ is used instead of $\alpha = 1$, the reduction of the estimation error for observations assimilated every $\Delta t$, $5\Delta t$, $10\Delta t$, $15\Delta t$, $20\Delta t$ is of $4\%$, $14\%$, $28\%$, $44\%$ and $27\%$, respectively. Nevertheless, if observations are assimilated every 25 time steps, then the best results are obtained for $\alpha = 1$ (the corresponding error is almost half of the one for the second best choice: $\alpha = 0.8$). For values of $\alpha$ smaller than 0.8 the results are completely wrong. If observations are assimilated every 30 time steps or less frequently the assimilation process fails regardless of the value of $\alpha$.

We also analysed different combinations of $(\alpha, \sigma_o^2)$ for $\Delta t_{\text{obs}} = 10$. Comparing Figure 7.79 with the previous one, it can be concluded that it is more critical for the assimilation scheme the number and distribution of observations than their accuracy. Only for $\alpha =
0.2 and very noisy observations ($\sigma^2_o = 0.5$) the RMSE is above 0.85. This speaks about the accuracy of the method when observations are available often enough. In general, smaller $\alpha$’s are more effective when the observations are very accurate, while for larger variances better results are obtained for larger values of $\alpha$. The differences on the RSME are not caused by a large mismatch of a single assimilated value but by a consistent improvement of the parameter and state assimilation along the time window, as shown in Figures 7.80 and 7.81.

![Graph](image)

**Figure 7.79:** Lorenz 63. RMSE for different $\alpha$’s and observation variances.

### 7.3.5 Summary

The LRSSC performed very well when applied to the Lorenz 63 model. For perfect observations the convergence speed of the estimated parameters and state variables to their true values was strongly related to the size of the variance and frequency of observations. When noisy observations were considered the quality of the observational data influenced as expected the quality of the assimilation, though in all cases the initial parameter estimates were considerably improved.

In general, the assimilation scheme had more difficulties recovering the parameter $\sigma$. We suppose that this is due to the fact that the output of the system is less sensitive to this parameter, therefore, the inverse problem of estimating the parameter given the observational data is much more difficult to solve. Notice that because observations are not perfect, in order to recover a parameter we need that its update causes a change in the output much larger than the standard deviation of the errors in the observations.
Experiments were conducted for different scheme parameter $\alpha$’s and observation frequencies and variances. Here similar results as for the nonlinear oscillator experiments were observed. For frequent observations, or observations with a low level of noise smaller $\alpha$’s produced better assimilation results, while for infrequent or very noisy observations larger $\alpha$’s were more effective.
Chapter 7. \textit{Nonlinear experiments} \hfill 166

7.4 Summary and discussion

In this chapter we have tested our proposed LRSSC applied to several simple dynamical systems with different number of parameters. Our goal was to assess the performance of the method for different values of $\alpha$ and to investigate whether results can be improved, when compared to the results obtained for $\alpha = 1$ (in this case the LRSSC is identical to the FDSPC method).

Experiments were conducted for different observation frequencies and variances, as well as various spatial distributions of observations (only for the advection model), considering perfect and noisy observations. We showed the results for an arbitrary choice of $\alpha$, but the same relations between the quality of the assimilation and the different assimilation parameters summarized below were found when other values of $\alpha$ were considered.

The hybrid method performed extremely well in most of the observation configurations tested. The quality of the assimilation, in two of the three models tested, was even better than when applying the computationally more expensive EKF (see Section 4.3 for the advection experiments and Appendix C.2 for the nonlinear oscillator experiments). Only for very noisy observations or when very few observations were available, the LRSSC failed to recover state and parameters to an acceptable level of accuracy. We found that the more precise, frequent and dense the observations, the faster the parameters converge to the true value, even for perfect observations.

Especially when observations are noisy, the parameter estimation is improved if it is averaged after a tune-up phase along a moving time window. Nevertheless, the point from which the averaging starts, and the size of the averaging time window must be carefully chosen. In general, without the averaging strategy, the parameters approach monotonically to the true value and then start oscillating in a certain neighbourhood of the truth. The averaging results in smaller parameter updates, smoothing the trajectory of the parameter estimates. If at the time point where the averaging starts, the parameter’s oscillation phase has not started yet, then the first average value will be still far from the truth. This can lead to the parameter estimates converging to a wrong value. On the other hand, let us assume that the size of the averaging time window is $S$. After the tune-up phase, each update of the parameter can be expressed as

\begin{equation}
\hat{p}_{k+1}^{\text{ave}} = \frac{S - 1}{S} \hat{p}_{k+1}^{\text{ave}} + \frac{p_k^2}{S},
\end{equation}

where $\hat{p}_{k}^{\text{ave}}$ represents the final update of the parameter and $p_k^2$ the parameter update before averaging. If $S$ is very large, the contribution of the new assimilated information
to the parameter final value is very small and the new updated value of the parameter
does not differ much from the previous. This can lead to an imprecise estimation of the
parameter at the end of the time window, unless the starting average value is already
close enough to the truth. On the contrary, if $S$ is very small the parameter increments
are relatively large and the oscillation effect prevails over the damping effect, leading to
a slower convergence.

Improvements were made by choosing $\alpha$’s different from $1$. The optimal choice of $\alpha$
depends on the model, as well as on the observation frequency and accuracy. While for
the advection model better results are obtained the smaller the $\alpha$ value is chosen, for
the other two models this is only true for frequent and relatively precise observations.
These results held for several combinations of parameter-parameter and state-state error
covariances tested. For the nonlinear oscillator and the Lorenz 63 model, if observations
are very sparse in time or very noisy, then there is not a clear pattern for the choice
of $\alpha$ and it depends on the exact observation frequency and variance. The experiments
show that by considering a time-dependent state-state error covariance, the speed of the
parameter convergence can be increased, resulting in a faster decrease of the assimilated
state error. This allows us to obtain a sufficiently accurate analysis in shorter assimila-
tion time windows. In some cases, improvements in the state assimilation were observed
when considering $\alpha < 1$, even if the trajectories of the assimilated parameter for a static
state-state error covariance and the time-dependent one were similar. This is due to
the fact that the structure of the time-dependent covariance reflects more accurately
the statistics of the state error. Accurate estimates were obtained as well, when our
approach was tested using other simple models, like the Two Box Model ([64], [76]).
Chapter 8

Conclusions and Further Work

Data assimilation (DA) is a well known technique that combines model dynamics, state observations and the error statistics to obtain good estimations of the system state. It is common that models depend on unknown parameters, which are crucial for a precise forecast of the state, for simulation and control. The manual tuning of these parameters is a time-consuming and prone-to-error process. An alternative is to estimate state and parameters at the same time following the augmented state approach. In this thesis we investigated the role played by the state-parameter covariance matrix $B_{xp}$ in the state-parameter joint estimation. Moreover, we showed that the derivatives of the model with respect to the parameters can be used for a low-rank update of the state-state covariance matrix $B_{xx}$, contributing to improve the assimilation results at a relatively low computational cost. In the following sections we present a summary of the main results of this work and discuss possible further lines of investigation.

8.1 Conclusions

In Chapter 2 we presented basic results concerning the general DA problem that were used in the framework of this investigation. Special attention was given to the KF and EKF formulas, which describe the propagation of the error statistics along the assimilation process, and to the 3D-Var, known for its robustness, simplicity and computational efficiency. Our ultimate goal was to combine ideas from these two approaches in the context of parameter estimation, to obtain good estimates of state and parameters at low computational cost.

The augmented state approach, introduced in Chapter 3, is a strategy which obtains at the same time estimates of the state and parameter variables, by solving a new DA
problem. States and parameters are combined into a new augmented state vector, and the forecast and observation operators are also modified, reflecting the time-invariant property of the parameters and the fact that they are not observed, respectively.

Nevertheless, in Chapter 4 we showed that following the augmented state approach, even for a simple example, a classical method like the EKF can fail to recover state and parameter to an acceptable level of accuracy.

In the EKF, during the forecast step and under perfect model assumptions, the covariances between parameters are unchanged. On the other hand, during the assimilation step the parameter variance is always reduced. This means that an eventual excessive reduction of the parameter variance during an assimilation step (an overconfidence of the filter in the parameter estimate) can cause the divergence of the scheme, as shown for the linear advection model. We proposed simple modifications of the EKF formulas to overcome this difficulty. The first two are based on controlling the size of the norm of state-parameter covariance $B_{xp}$ and the parameter-parameter covariance $B_{pp}$. This can be done by either fixing $B_{pp}$ or resizing the norm of $B_{xp}$, such that it is kept constant. Very good estimates were obtained when these simple modifications were combined with an averaging of the parameter along the assimilation window. The third strategy discussed considers the wrong parameter value as a source of model error and incorporates the resulting model error covariance matrix in the update of the augmented state covariance. The ideas of fixing the parameter-parameter covariance and using the parameter sensitivities in the update of the state-state covariance were incorporated in the novel method proposed in this thesis.

In the context of DA some schemes combine ideas of two or more methods to improve either the accuracy of the state estimation or the computational efficiency of the assimilation scheme. In Chapter 5 we focused on the method proposed by Smith [101], to which we refer as the FDSPC, and that was designed especially for the state and parameter estimation problem using the augmented state approach. The author proposes a 3D-Var-like method, in which the $B_{xp}$ is updated using the derivative of the model with respect to the parameters and the parameter-parameter covariance $B_{pp}$. Here we showed that this update is related to the computation of an empirical covariance matrix using state samples generated with different parameter values. The main strength of the method is that a flow-dependent $B_{xp}$ is essential to obtain good estimates of the state and parameters. At the same time, the proposed update can be computed very efficiently, supposing that the number of parameters is small. Based on this approach we proposed a modification, which we named the LRSSC, consisting in incorporating a flow-dependent $B_{xx}$, while keeping the same update for $B_{xp}$. For the sake of clarity we repeat the update formula presented in Section 5.2 here:
\[ B_{xx} = \alpha B_{xx}^b + (1 - \alpha)NB_{pp}N^\top. \] (8.1)

The new update uses a convex combination of a time-invariant background covariance, which accounts for the long-time-scale state errors and contributes to preserving the positive-definite property of the covariance, and a flow-dependent matrix, typically low-ranked (if the number of parameters is small), which also depends on the sensitivities of the model with respect to the parameter and a static \( B_{pp} \). This flow-dependent term accounts for the short-time-scale errors and is related to a model biased caused by a wrong estimation of the parameters. In Section 5.2 we showed that for the linear advection model the uncertainties on the parameter do have a huge impact on the propagation of the state error statistics, so it was expected that the new approach could contribute to improve the results obtained by Smith. In fact, the FDSPC can be considered as a particular case of our more general method, setting \( \alpha = 1 \). Furthermore, the order of the computational cost of the LRSSC is equivalent to that of the FDSPC, which means that it is possible to get a flow-dependent \( B_{xx} \) without significant extra computational effort.

Convergence conditions for the sequential 3D-Var concerning the covariance \( B_{xp} \) were obtained in Chapter 6, for the linear, time-invariant state-parameter dynamical system. Despite the fact that in practical applications the forecast or the observation operator are in general nonlinear, most of the basic theory is based on the assumption that both operators are linear, so the study of this case is of crucial importance. We showed that in this case the LRSSC, and therefore the FDSPC, are equivalent to a sequential 3D-Var with a special \( B_{xx} \) and \( B_{xp} \) structure. This is because the derivatives of the model with respect to the parameters are time-invariant. It implies that convergence conditions imposed to \( B_{xp} \) can be translated into conditions imposed to \( B_{pp} \). We characterized the 3D-Var iteration process as a new dynamical system, for which the expected value of the state and parameter estimation error tends to zero if the eigenvalues of the corresponding iteration matrix are inside the unit circle.

Special scenarios were analysed and we were able to find conditions on \( B_{xp} \) to ensure that the eigenvalues of the aforementioned iteration matrix are inside the unit circle when the 3D-Var is applied. Furthermore, equivalent conditions on \( B_{pp} \) were found for the FDSPC and the LRSSC methods. One of these special scenarios is when the state is completely observed and the observations are very precise (\( R \approx 0 \)). We showed that for the single parameter case, the convergence condition is more restrictive the larger \( \alpha \) is set. Necessary and sufficient conditions were deduced for the case, where only one state variable is observed, as well as for the case where the state estimation is very precise (\( B_{xx} \approx 0 \)).
An important theoretical result was the proof that for the linear and time-invariant problem convergence is not attained if the number of parameters is larger than the dimension of the observation vector. This is a critical result concerning the application of the method to estimate model bias. In such cases each parameter represents the bias of a single state variable forecast. This means that unless the state is completely observed, which is never the case in large dimensional systems, the sequential 3D-Var method diverges. We also showed a relation between convergence speed and the condition number of the covariance $B_{xx}$ for the both FDSPC and LRSSC methods, depending on the structure of $B_{pp}$. The parameter statistics are typically only vaguely known, so the theoretical results obtained in this chapter can be used to decide on the structure of $B_{pp}$, or even on an inflation factor of $B_{pp}$ and $B_{xx}$.

In Chapter 7 we tested our approach using different simple nonlinear models. In general we obtained very good results, approximating parameter and state vectors to a high level of accuracy. Only when observations were too sparse, noisy or infrequent, the scheme failed to recover the parameter and therefore the state, with a reasonable precision. In general, better results were obtained when the parameters were averaged along a moving time window after a tune-up phase during which the system stabilizes. This was also observed by Smith in [101] when applying the FDSPC method. For $\alpha < 1$, in many cases a faster convergence of the parameter was achieved, compared to assimilations for $\alpha = 1$. This contributed to reduce the duration of the tune-up phase. Furthermore, even in situations where the quality of the parameter estimation and the convergence speed of the method were similar for $\alpha = 1$ and for smaller value of $\alpha$, better estimations of the state, reflected in a smaller RMSE, were obtained for the flow-dependent structure of the state-state covariance. We observed that the optimal choice of $\alpha$ depends on the specific model, as well as on the precision and frequency of observations. In general, better results were obtained for small values of $\alpha$ when observations were relatively accurate or often assimilated.

Ideally the error statistics should be transported forward by the model. Nevertheless, the computational cost of such an approach is prohibitive for large scale problems. The LRSSC method can be seen as a compromise between a precise representation of the error statistics and a low computational cost. The results suggest that updating $B_{xx}$ by combining a static covariance with a low-rank, flow-dependent term related to the model derivatives with respect to the parameters, can contribute to improve the estimations of parameters and state. Based on the theoretical and practical results of this investigation we expect that the LRSSC can be applied to larger and more complex models.
8.2 Future work

In Chapter 6 we obtained theoretical results for the convergence of the sequential 3D-Var when applied to models which are linear with respect to both, state and parameter. It would be interesting to find convergence results for more complex models. The next level of complexity would be to consider models which are linear with respect to the state, if the parameter is fixed, but where the forecast matrix depends on the parameter, such that the augmented state problem is no longer linear. This is the case, for example, of the linear advection model. Such dynamical systems can be described by an equation of the form $x_{k+1} = A(p)x_k$. In this case neither a standard sequential 3D-Var, nor any of the studied approaches can be reduced to a time-invariant dynamical system as in the pure linear case. This implies extra difficulties to find convergence conditions and requires therefore a completely different analysis.

In our experiments we only considered small dimensional models to test our strategy. Though the results were promising, a next step should be to investigate, how the method performs when applied to more complex, high dimensional problems. Real applications differ from our test models not only in the dimension of the problem. Many other aspects add extra complexity to the assimilation process, like for example, non-homogeneously distributed observations, correlations between the observed variables or parameters, amongst others. It would be then necessary to design and implement a more robust computational software, based on our strategy, that can solve these more realistic problems.

In the framework of this investigation we only considered fixed convex combinations of the static and flow-dependent matrices involved in the update of $B_{xx}$. We showed that for the models tested, the best choice of the scheme parameter $\alpha$ depends on the observation variance or frequency, amongst other factors. However, the relation between these factors and $\alpha$ should be more carefully investigated. Besides, it would be important to have a deeper insight of the role of $\alpha$ in the formulas for the covariance inverse update. Further research should also consider time-dependent linear combinations. We believe that the sensitivity of the model with respect to the parameters could be used as a measure of importance to decide on the coefficient of the flow-dependent term of the covariance update. Another possibility would be to investigate the suitability of $B_{xx}$ updates of the form

$$
(B_{xx})_{k+1} = \alpha_1 B_{xx}^b + \sum_{i=1}^T \alpha_{i+1}N_{k+1-i}B_{pp}N_{k+1-i}^T, \quad (8.2)
$$
i.e., not only considering short-time-scale errors 1 time step prior the assimilation step, but $T$ time steps. Of course, this generalization of the LRSSC approach would result computationally more expensive the more terms in the sum are considered.
Appendix A

Steady State of the Linear Advection Discrete Model

The discrete linear advection model presented in Section 4.3 is described by

\[ x_{k+1} = Fx_k \]  \hspace{1cm} (A.1)

where

\[
F = \begin{pmatrix}
  c\mu & 0 & \ldots & 0 & 1 - c\mu \\
  1 - c\mu & c\mu & & & 0 \\
  0 & \ddots & \ddots & \ddots & \ddots \\
  \vdots & 1 - c\mu & c & & 0 \\
  0 & \ldots & 0 & 1 - c\mu & c\mu
\end{pmatrix}
\]  \hspace{1cm} (A.2)

with \( \mu = \frac{\Delta t}{\Delta z} \) and \( 0 < c < \frac{\Delta z}{\Delta t} \). We will show that the only steady state of (A.1) is \( x^* = \bar{x} \left(1, 1, \ldots, 1\right) \), with \( \bar{x} = \sum_{i=1}^{n} x_0 \), for any initial state \( x_0 \). Matrix \( F \) is a circulant matrix, which means that each row vector is rotated an element to the right in relation to the preceding row. For a deeper understanding and application cases see [96], [97], [46] and [55]. This property means that all the matrix coefficients can be inferred from the first row. Therefore, a common way of representing such matrices is by the notation \( F = \text{circ}[c_0 = f_{11}, c_1 = f_{12}, \ldots, c_{n-1} = f_{1n}] \) (in this case \( F = \text{circ}[c\mu, 0, 0, \ldots, 0, 1 - c\mu] \)).

On the other hand

\[ 0 < c < \frac{\Delta z}{\Delta t} \iff 0 < c \frac{\Delta t}{\Delta z} < 1 \iff 0 < c\mu < 1. \]  \hspace{1cm} (A.3)

From (A.3) it is clear that the entries of the matrix \( F \) are all positive and smaller than 1. Moreover, the sum of the elements of each row (each column) is 1. Then \( F \) is not only circulant but stochastic. From [112] we know that if \( A \) is a circulant stochastic matrix.
the eigenvalues are given by evaluating the polynomial \( p_A(x) = \sum_{i=0}^{n-1} c_i x^i \) at the \( n \)-th roots of unity \( w^j, j = 0, 1, \ldots, n-1 \). In our case \( p_F(x) = c\mu + (1 - c\mu)x^{n-1} \). The \( n \) eigenvalues of \( F \) are

\[
p_F(w^j) = c\mu + (1 - c\mu)w^j(n-1), \quad j = 0, 1, \ldots, n-1 \quad (A.4)
\]

\[
=w^\frac{1}{j=0} \begin{cases} c\mu + (1 - c\mu)w^{(j-1)n+j} & j = 1, \ldots, n-1 \\ 1 & j = 0 \end{cases} \quad (A.5)
\]

\[
(w^n)^{j-1} = \begin{cases} c\mu + (1 - c\mu)w^{n-j} & j = 1, \ldots, n-1 \\ 1 & j = 0 \end{cases} \quad (A.6)
\]

All the \( n \) eigenvalues are different. Notice that they are the result of dividing the circumference of radius \( 1 - c\mu \) and center in \((c\mu, 0)\) into \( n \) equally spaced points. The eigenvalue with the largest absolute value is 1 and it is the only eigenvalue with that absolute value. The eigenspace associated to it has therefore dimension 1. It can be checked that if \( v = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix} \), then \( Fv = v \), so \( v \) is the eigenvector associated to 1 and

\[
\lim_{k \to \infty} F^k = vv^\top = \frac{1}{n} \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 1 & 1 \\ \vdots & \ddots \\ 1 & 1 \end{pmatrix}, \quad (A.7)
\]

from where

\[
\lim_{k \to \infty} x^k = \lim_{k \to \infty} F^k x_0 = x^* . \quad (A.8)
\]
Appendix B

Computational Cost of the Hybrid Approach Covariance Update

Here we analyse the cost of computing our hybrid approach update presented in Section 5.2. As before, \( n \) and \( \ell \) refer to the dimensions of the state and parameter vectors, respectively. We assume that a computational code \( b(x) \) for computing the matrix-vector multiplication \( B_{xx}^{-1}x \) is available, and that its computational cost is \( O(r) \). If \( B_{xx}^{-1} \) is a dense matrix with no special structure then \( r \approx n^2 \), where \( n \) is the dimension of the state vector \( x \). Nevertheless, for many high dimensional problems \( r \ll n^2 \). We also assume that \( B_{pp}^{-1} \) is given. Here we show that evaluating \( B^{-1}z \) according to the update of the LRSSC presented in section 5.2 has a computational cost of \( O(rl) \), which is equivalent to the computational complexity of the FDSPC update.

The update of the inverse of the state-state covariance matrix times a matrix \( A \in \mathbb{R}^{n \times q} \) according to 5.26 can be written as:

\[
B_{xx}^{-1}A = \left( \frac{B_{xx}^{-1}}{a_1} - \frac{B_{xx}^{-1}N}{a_1} \left( \frac{B_{pp}^{-1}}{a_2} + \frac{N^\top B_{xx}^{-1}N}{a_1} \right)^{-1} \frac{N^\top B_{xx}^{-1}}{a_1} \right) A, \tag{B.1}
\]
Appendix B. Computational cost of the hybrid approach covariance update

and is equivalent to the three-steps computation:

\[
X = \frac{B_{xx}^{-1} N}{\alpha_1}, \quad (B.2)
\]

\[
S = \left( \frac{B_{pp}^{-1}}{\alpha_2} - N^T X \right)^{-1}, \quad (B.3)
\]

\[
B_{xx}^{-1} A = \frac{B_{xx}^{-1}}{\alpha_1} A - X S X^T A. \quad (B.4)
\]

Below we analyse the cost of this update, taking into account that the cost of the matrix multiplication \(M_1 M_2\), with \(M_1 \in \mathbb{R}^{m \times q}\) and \(M_2 \in \mathbb{R}^{q \times n}\) is \(O(mnq)\):

1. Compute \(X\): the cost is \(O(\ell)\), as requires to run \(b(N_i)\) for \(i = 1, 2, \ldots \ell\). The resulting matrix has dimension \(n \times \ell\).

2. Compute \(S\): we compute the product \(V_1 = N^T X \ (O(n\ell^2), V_1 \in \mathbb{R}^{\ell \times \ell})\), sum the two matrices \((O(\ell^2))\) and invert the resulting matrix \(S = \left( \frac{B_{pp}^{-1}}{\alpha_1} - V_1 \right)^{-1} \ (O(\ell^3)) \).

3. Evaluating \((B.4)\): \(V_2 = X^T A \ (O(nq\ell), V_2 \in \mathbb{R}^{\ell \times q})\). The remaining matrix multiplications \(V_3 = SV_2\) and \(V_4 = XV_3\) have cost \((O(\ell^2q))\) and \((O(nq\ell))\) respectively.

4. \(\frac{B_{xx}^{-1}}{\alpha_1} A\) has cost \(O(q\ell)\), as requires to run \(b(A_i)\) for \(i = 1, 2, \ldots q\).

5. The matrix subtraction has order \(O(nq)\).

Overall the computational cost is then \(O(\max\{\ell, q\ell, \ell^3, nq\ell, q\ell^2\})\), which is equivalent to \(O(\max\{\ell, nq\ell, q\ell^2\})\) for \(\ell \ll n\).

To compute \(B^{-1}z\) we can write the augmented state \(z\) as \(z = \begin{pmatrix} x \\ p \end{pmatrix}\), where \(x \in \mathbb{R}^n\) and \(p \in \mathbb{R}^\ell\). Using 5.29:

\[
\begin{pmatrix}
B_{xx} & B_{xp} \\
B_{xp}^T & B_{pp}
\end{pmatrix}
\begin{pmatrix}
x \\
p
\end{pmatrix}
= 
\begin{pmatrix}
B_{xx}^{-1} + B_{xx}^{-1} B_{xp} S B_{xp}^T B_{xx}^{-1} & -B_{xx}^{-1} B_{xp} S \\
-S B_{xp}^T B_{xx}^{-1} & S
\end{pmatrix}
\begin{pmatrix}
x \\
p
\end{pmatrix}, \quad (B.5)
\]
with $S := (B_{pp} - B_{xp}^\top B_{xx}^{-1} B_{xp})^{-1}$. This is equivalent to the computation of the following steps:

$$
Y = B_{xx}^{-1} B_{xp}, 
$$

(B.6)

$$
S = \left( B_{pp} - B_{xp}^\top Y \right)^{-1},
$$

(B.7)

$$
B_{-1} z = \begin{pmatrix}
B_{xx}^{-1} + Y S Y^\top - Y S

\newline
-S Y^\top & S
\end{pmatrix}
\begin{pmatrix}
x
\newline
p
\end{pmatrix},
$$

(B.8)

$$
= \begin{pmatrix}
B_{xx}^{-1} x + Y S Y^\top x - Y S p

\newline
-S Y^\top x + S p
\end{pmatrix}.
$$

(B.9)

Such update is the same for the FDSPC and the LRSSC. The only difference is the state-state covariance used. Attending to the dimensions of the matrices involved, we calculate the total computational cost:

1. Computing $Y$: $O(\text{max}(r \ell, n \ell^2))$ for the LRSSC and $O(r \ell)$ for the FDSPC.

2. Computing $S$: Obtaining $B_{xp}^\top Y$ is $O(n \ell^2)$ and the cost of the matrix inversion is $O(\ell^3)$.

3. Computing the upper block of vector $B_{-1} z$: $V_5 = Y^\top x$ ($O(n \ell)$, $V_5 \in \mathbb{R}^\ell$), $V_6 = S V_5$ ($O(\ell^2)$, $V_6 \in \mathbb{R}^\ell$), $V_7 = Y V_6$ ($O(n \ell)$, $V_7 \in \mathbb{R}^n$), $V_8 = B_{xx}^{-1} x$ ($O(r)$ for the FDSPC and $O(\text{max}(r \ell, n \ell))$ for the LRSSC, $V_8 \in \mathbb{R}^n$), $V_9 = S p$ ($O(\ell^2)$, $V_9 \in \mathbb{R}^\ell$), $V_{10} = Y V_9$ ($O(n \ell)$, $V_{10} \in \mathbb{R}^n$). The sum of the three resulting vectors is $O(n)$. This makes an overall cost of $O(\text{min}(n \ell, r \ell))$.

4. Computing the lower block of vector $B_{-1} z$: This is equivalent to $V_9 - V_6$, which has cost $O(\ell)$.

The most expensive steps of the sequential process and therefore of the overall computation are calculating the matrices $Y$ and $S$. For the FDSPC, as well as for the LRSSC, the computational complexity of completing these two steps is the same: $O(\text{max}(n \ell^2, r \ell))$. It is expected that in general $r > n \ell$, unless the inverse of the background state covariance matrix $B_{xx}^{-1}$ has a very special and simple structure, for example diagonal or tri-diagonal. Even in those cases the cost $n \ell^2$ is affordable if the number of parameters is not too large. So we can conclude that the computational complexity of the update for both approaches is $O(r \ell)$ in most of our problems of interest.
Appendix C

Nonlinear Oscillator Experiments

Here we present experimental results obtained for the nonlinear oscillator introduced in Section 7.2. First we show the results related to finding an optimal value for the scheme parameter $\alpha$ of the LRSSC, according to the observation variance $\sigma_o^2$. Later we show the results when the standard EKF was applied.

C.1 Choosing $\alpha$ according to $\sigma_o^2$.

We present the results of experimenting with different observation variances and $\alpha$'s. For each combination 20 independent experiments were conducted, computing the state RMSE and the relative error of the parameters at the end of the time window. The average and variance of these quantities are shown in Figures C.1, C.2 and C.3. The quality of the state assimilation strongly depends, as expected, on the size of the observation variance. Small values lead to accurate estimations of the state and the parameters for $\alpha \geq 0.2$ and $\sigma_o^2 \leq 0.01$. In this case, there is no $\alpha$ for which the accuracy of the parameter and state recovery is clearly better than for the rest.

C.2 EKF experiments.

In the following we show the results of EKF experiments for different observation frequencies ($\Delta t_{obs} = 1, 5, 10$) and $\sigma_o^2 = 0.01$. Even for observations assimilated at every time step the scheme struggles to accurately estimate $x$ and $y$. For $\Delta t_{obs} = 10$ the state is damped too fast due to an overestimation of both parameters (see figures C.4-C.7).

If the parameter-parameter error covariance $B_{pp}$ is fixed and the parameters are averaged after $t = 25$ then the results are drastically improved (see figures C.8-C.11).
Appendix C. Nonlinear oscillator experiments.

Figure C.1: Nonlinear oscillator. RMSE for different $\alpha$’s and observation variances. Left: RMSE mean. Right: RMSE variance.

Figure C.2: Nonlinear oscillator. Relative error for the assimilated parameter $m$ for different $\alpha$’s and observation variances. Left: $|m^a - m^t|/m^t$ mean. Right: $|m^a - m^t|/m^t$ variance.

Figure C.3: Nonlinear oscillator. Relative error for the assimilated parameter $d$ for different $\alpha$’s and observation variances. Left: $|d^a - d^t|/d^t$ mean. Right: $|d^a - d^t|/d^t$ variance.
Appendix C. Nonlinear oscillator experiments.

Figure C.4: Nonlinear oscillator. EKF assimilation of $x$ for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue line), $5\Delta t$ (green line) and $10\Delta t$ (red line). The true value of $x$ is given by the black line.

Figure C.5: Nonlinear oscillator. EKF assimilation of $y$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue line), $5\Delta t$ (green line) and $10\Delta t$ (red line). The true value of $y$ is given by the black line.
Appendix C. Nonlinear oscillator experiments.

Figure C.6: Nonlinear oscillator. EKF assimilation of $m$ for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue line), $5\Delta t$ (green line) and $10\Delta t$ (red line). The true value of $m$ is given by the black line.

Figure C.7: Nonlinear oscillator. Assimilation of $d$ for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue line), $5\Delta t$ (green line) and $10\Delta t$ (red line). The true value of $d$ is given by the black line.
Figure C.8: Nonlinear oscillator. EKF assimilation of $x$ fixing $B_{pp}$, for different observation frequencies and perfect observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue line), $5\Delta t$ (green line) and $10\Delta t$ (red line). The true value of $x$ is given by the black line. The black dotted line represents the point from which the parameter is averaged.

Figure C.9: Nonlinear oscillator. EKF assimilation of $y$ fixing $B_{pp}$, for different observation frequencies and noisy observations, $\sigma_o^2 = 0.01$. Assimilations every $\Delta t$ (blue line), $5\Delta t$ (green line) and $10\Delta t$ (red line). The true value of $y$ is given by the black line. The black dotted line represents the point from which the parameter is averaged.
Figure C.10: Nonlinear oscillator. EKF assimilation of $m$ fixing $B_{pp}$, for different observation frequencies and perfect observations, $\sigma^2_o = 0.01$. Assimilations every $\Delta t$ (blue line), $5\Delta t$ (green line) and $10\Delta t$ (red line). The true value of $m$ is given by the black line. The black dotted line represents the point from which the parameter is averaged.

Figure C.11: Nonlinear oscillator. Assimilation of $d$ fixing $B_{pp}$, for different observation frequencies and noisy observations, $\sigma^2_o = 0.01$. Assimilations every $\Delta t$ (blue line), $5\Delta t$ (green line) and $10\Delta t$ (red line). The true value of $d$ is given by the black line. The black dotted line represents the point from which the parameter is averaged.
Bibliography


