THE UNIVERSITY OF READING

DEPARTMENT OF MATHEMATICS AND STATISTICS

Conditioning and Preconditioning of the Minimisation Problem in Variational Data Assimilation

Stephen A. Haben

Thesis submitted for the degree of

Doctor of Philosophy

March 2011

Abstract

Many numerical weather prediction (NWP) centres around the world implement a variational data assimilation (Var) scheme to find the initial state of the atmosphere, called the analysis. The analysis is used as the initial conditions for a numerical forecast model. For an accurate weather forecast an accurate analysis is essential. Var is formulated as a optimization problem and is solved by a series of minimisations of linear least-square cost functions. The speed of convergence of these minimisations and the sensitivity of the analysis to perturbations are dependent on the condition number of the Hessian of the least-squares cost function. A small condition number of the Var Hessian is essential for an accurate forecast. Many NWP centres perform a control variable transform (CVT) in order to solve a preconditioned Var (PVar) scheme. In this thesis we consider the conditioning of Var and PVar in detail by deriving new theoretical bounds on the condition number of the Var and PVar Hessians. Using the bounds we show that the Var Hessian is ill-conditioned when the error covariance matrix of the prior estimate is ill-conditioned. We also show that preconditioning with the CVT produces a significant reduction in the condition number of Var. Additionally, we show using the theoretical bounds that the condition number of the PVar Hessian is reduced if we increase the spacing of observations, reduce the accuracy of the observations and reduce the number of observations. We demonstrate these results numerically for both a simple one-parameter periodic system and the Met Office PVar scheme. We also demonstrate that the CVT produces a significant increase in the convergence rate of the conjugate gradient method used to solve the Var scheme.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Stephen Haben

Acknowledgements

Firstly, I would like to thank my supervisors Prof. Nancy Nichols and Dr. Amos Lawless for all the guidance and support they have provided me over the course of this PhD. I am grateful for all their efforts and for helping me to constantly challenge and improve myself. I would also like to thank my Met Office supervisors Prof Mike Cullen, Dr Gordon Inverarity and Dr Tim Payne for all their patience and assistance. I would like to thank the Engineering and Physical Science Research Council (EPSRC) and the Met Office for the funding I have received. The funding was in the form of a Collaborative Award in Science and Engineering (CASE) studentship with sponsorship from the Met Office. Without this funding this research would not have been possible.

I would also like to thank my family for their constant encouragement and support. A special thank you goes to Catarina. Your positive attitude toward life and your inquisitive nature have been a constant inspiration to me over the past three and a half years.

I dedicate this thesis to my grandparents Robert and Kathleen.

Contents

9 Conclusions   **209**

List of Figures

Chapter 1

Introduction

Numerical weather prediction (NWP) centres produce forecasts of future weather states using a numerical model of the atmosphere to evolve an estimate of the initial state of the atmosphere forward in time. The accuracy of this estimate, called *the analysis*, is therefore a major factor in determining the accuracy of the resultant forecast. Variational data assimilation (Var) is one method popularly used in NWP centres for finding the analysis. In Var the analysis is the minimiser of a cost function. The cost function is essentially a weighted measure of the distance between the forecast states and the available observations within a fixed time window, weighted using the background (or forecast) and observation error covariance matrices. The resulting solution is the maximum likelihood best estimate of the state of the atmosphere under certain assumptions [40].

Var for NWP is a large and computationally expensive problem that involves minimising a highly non-linear cost function with respect to more than $10^7 - 10^8$ variables, given just 10^6 observations. To alleviate the computational expense, an incremental form of Var, first developed in [51], is usually implemented instead. In this version, a sequence of linearised cost functions, approximating the full non-linear problem, are solved. Each linearised cost

function is now linear least-squares and is minimised in an *inner-loop* using iterative gradient methods. The minimiser is then utilised in an *outer-loop* step to update the current best estimate of the analysis.

A measure of the sensitivity of the inner-loop solution to perturbations is given by the size of the condition number of the Hessian of the linearised cost function. If the system has a large condition number we call it *ill-conditioned*. A large condition number implies the solution is sensitive to small perturbations in the system. An ill-conditioned problem also means that convergence can be slow for the iterative solver used to find the solution. Slow convergence of the iterative gradient methods used to solve the inner-loop problem is assumed to be due to the ill-conditioning of the Hessian of the linearised cost function. It was suggested that an ill-conditioned background/forecast error covariance matrix is the source of the ill-conditioned Hessian [42].

In practice, operational NWP centres attempt to reduce the influence of the conditioning of the background covariance matrix by transforming to new variables with uncorrelated errors [10]. The aim is to reduce the condition number, or *precondition*, the Hessian. Comparisons have shown that the convergence rates can be significantly reduced by preconditioning [42], [18]. The factors that affect the condition number of the preconditioned system are only partially understood and have only been studied in very simplified circumstances [6], [63]. As computer processing power increases, the resolution of the numerical model increases, producing a considerable computational challenge for NWP centres both now and in the future. Understanding the conditioning of the unpreconditioned and preconditioned systems is increasingly important to measure the capabilities and limitations of the Var technique for NWP.

1.1 Aims of Thesis

The main aim of this thesis is to gain a greater understanding of the conditioning of the variational data assimilation problem. We aim to highlight the major factors which affect the conditioning of both the unpreconditioned and preconditioned Var systems. Specifically, we

- Show that common auto-correlation models, used to model the background error covariances matrix in Var, have condition numbers that are sensitive to correlation lengthscale. Hence for highly correlated background errors the background error covariance matrix will be ill-conditioned.

- Derive new theoretical bounds on the condition number of the Var Hessian. The bounds show that the conditioning of Var is dependent on the conditioning of the background covariance matrix. In particular we show that for highly correlated background errors the Hessian of the unpreconditioned system will be ill-conditioned.

- Derive new theoretical bounds on the condition number of the Hessian of the preconditioned Var system. We show in a simplified one-dimensional model that the system is less sensitive to the lengthscale and better conditioned than the Hessian of the unpreconditioned Var system.

- Show, using the bounds on the conditioning of the preconditioned system, that the conditioning is affected by the accuracy, number and positioning of the observations. In particular, we show that the condition number of the Hessian of the preconditioned Var system is reduced by thinning the observations, increasing the spacing between observations and making the observations less accurate

- Show using experiments on a one-parameter, periodic system that our preconditioned system offers a large increase in the convergence rate of

the conjugate gradient method compared to the unpreconditioned. In addition, increasing the spacing and making the observations less accurate also increases the convergence rate in the preconditioned system.

- Show that the theoretical results obtained for the preconditioned system also apply to the Met Office operational Var scheme. Namely we show that the condition number of the Met Office Var scheme is reduced by thinning the observations, increasing the spacing between observations and making the observations less accurate

1.2 Outline

The thesis structure is as follows.

In **Chapter 2** we introduce 3D and 4D variational data assimilation (denoted 3DVar and 4DVar respectively) and the incremental formulation and give an overview of the details of the methods as implemented in NWP centres. We also introduce an incremental Var formulation that includes the control variable transform (CVT). The CVT provides a method for modelling the background error covariance matrix and also simplifies the background term in the Var cost function. We show how the CVT is implemented in the Met Office Var operational scheme. This chapter provides the context and the motivation for the research in the subsequent chapters.

In **Chapter 3** we introduce the concept of condition number and the important role it plays in determining the accuracy of solutions to the Var scheme. We also show how the condition number can indicate the convergence rate of iterative gradient methods used to solve the Var minimisation. We briefly discuss the concept of preconditioning as a way of improving the conditioning of a system. Finally we provide some basic mathematical tools which are useful for establishing bounds on the conditioning of our theoretical problems

in later chapters.

In **Chapter 4** we relate the condition number to the data assimilation problem we introduced in Chapter 2. In particular we introduce the concept of conditioning with respect to Var and show how preconditioning can be applied to the Var problem. We also discuss previous literature which considered the conditioning and preconditioning of the Var problem and discuss the limitations in our current understanding. We then briefly describe how the CVT effectively preconditions the Var scheme in an operational setting.

In **Chapter 5** we discuss the background error covariance matrices and give a brief outline of their importance in variational data assimilation. We show that increasing the background error correlation lengthscale increases the condition number of some standard auto-covariance matrices. The conditioning of the background matrices is useful for understanding the results in the subsequent chapters.

In **Chapter 6** we consider the conditioning of the 3DVar problem in a theoretical setting. We derive new explicit bounds on the condition number of the Hessians of both the unpreconditioned and preconditioned 3DVar schemes. We show that the bounds on the condition number of the unpreconditioned system are proportional to the condition number of the background covariance matrix. Using results from Chapter 5 we demonstrate in experiments with a periodic, one-parameter 1D problem, that the unpreconditioned system is generally ill-conditioned. We also show how preconditioning significantly improves the condition number of the system and that the conditioning of the preconditioned system is affected by three main factors: the accuracy, thinning and spacing of the observations. Finally we demonstrate that preconditioning greatly improves the convergence rates of the conjugate gradient method applied to solve the Var problem.

In **Chapter 7** we extend the bounds of Chapter 6 to the 4DVar case by including the time parameter. We derive new bounds on the condition numbers of the new extended Hessians of the preconditioned and unpreconditioned systems and formulate hypotheses. We test these hypotheses using a simple advection model as our forecast equation. As in Chapter 6 we show that the condition number of the unpreconditioned Hessian is proportional to the condition number of the background error covariance matrix. Therefore when the background matrix is ill-conditioned the unpreconditioned system is ill-conditioned. We show that preconditioning with the CVT significantly reduces the conditioning of the Var problem. We also demonstrate that accuracy, thinning and spacing of the observations are again important for determining the conditioning of the preconditioned system. Finally we show that preconditioning improves the convergence rate of the conjugate gradient method when applied to our 1D system.

In **Chapter 8** we examine the conditioning of the Met Office Var scheme. We begin by giving an introduction to the scheme and investigating the conditioning of the current system design. We consider assimilating individual observation types and show that the conditioning of the Met Office Var scheme is dominated by the surface observations. Using pseudo and real observations we show that thinning and spacing the observations and reducing the accuracy of the observations reduces the condition number. Hence, we confirm that the results of Chapters 6 and 7 also hold in an operational system.

In **Chapter 9** we summarise the work, giving the main conclusions and the questions that have still to be answered. We finish by giving suggestions on the further work that could be done on this problem.

We now introduce variational data assimilation. This is the main context of the problem investigated in this thesis and motivates the research that follows.

Chapter 2

Variational Data Assimilation

In this chapter we introduce variational data assimilation (Var), specifically within the context of numerical weather prediction (NWP). The aim of this chapter is to provide motivation and context for the research presented in the subsequent chapters.

We begin by giving a brief overview of data assimilation and its application to NWP. We then focus on variational data assimilation which is the data assimilation method investigated in this thesis. We describe the practical implementation of Var within NWP by introducing the computationally easier to implement incremental Var method. This reduced version of Var was first described in [51] and marked the first realistic opportunity for using Var in an operational setting. We also describe some technical points of the minimisation used in incremental Var. Var requires the modelling of a large background error covariance matrix. In practice, construction of the background error covariance matrix can be simplified via a series of transformations which is implemented in the incremental Var cost function as a control variable transform (CVT). Finally we summarise this chapter.

2.1 Data Assimilation

Data assimilation is a method for combining past and present observations, an initial guess of the current state of the system and a suitable numerical model to find the 'best estimate' of the current state of the system, called the *analysis*. In NWP the resultant estimate is used as the initial condition of the numerical forecast model. The equations which model the atmospheric flow are highly chaotic, meaning small errors will rapidly grow and means an accurate analysis is vital for achieving an accurate forecast [32], [44]. Data assimilation is therefore an important component of NWP. It is also computationally very challenging. Typically, major NWP centres, including the European Centre for Medium Range Weather Forecasts (ECMWF) and the UK Met Office, are required to find an analysis consisting of millions to tens of millions of degrees of freedom. In contrast, observations are relatively sparse ($\approx 10^6$) and unevenly distributed. Fewer observations than unknowns mean the problem is underconstrained and therefore a prior guess, called the *background* state, must be included. In addition, both the observations and background states contain unknown errors that must be estimated and accurately accounted for in the assimilation scheme.

There are a variety of different data assimilation techniques that have been developed to tackle this state estimation problem including sequential methods, such as the Kalman filter and optimal interpolation, and variational methods [32]. Currently, in most major operational NWP centres including the Met Office [54] and the ECMWF [53], Var is the method of choice and for the remainder of this thesis is the focus of our research. We now present an introduction to the variational data assimilation method.

2.2 Variational Data Assimilation

In the 1980's NWP centres were moving away from optimal interpolation data assimilation methods toward a more variational approach [38], [16], [48], [32]. These approaches finally became what is now known as four-dimensional variational data assimilation (4DVar). The objective of 4DVar is to minimise the following weighted non-linear least squares cost function with respect to the initial state vector $\mathbf{x}_0$,

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{n}(\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i)^T \mathbf{R}_i^{-1}(\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i),$$

$$(2.1)$$

$$\equiv J_b + J_o, \tag{2.2}$$

subject to satisfying the nonlinear forecast model

$$\mathbf{x}_i = \mathcal{M}(t_{i-1}, t_i, \mathbf{x}_{i-1}). \tag{2.3}$$

The resultant minimiser of (2.1) is called the *analysis*, $\mathbf{x}_a$.

Within this notation $\mathbf{y}_i \in \mathbb{R}^{p_i}$ represent observations of the atmospheric variables at discrete time steps t_i for $i = 0, \ldots, n$ collected within a time window $[t_0, t_n]$, called the assimilation window. The vectors $\mathbf{x}_i \in \mathbb{R}^N$ are the model states defined at time t_i found by evolving forward the state vector from time t_{i-1} to t_i using the atmospheric forecast model $\mathcal{M}(t_i, t_{i-1}, \mathbf{x}_{i-1}) : \mathbb{R}^N \longrightarrow \mathbb{R}^N$. Since there are generally fewer observations than elements in the model state an *apriori* guess of the initial state called the *background*, $\mathbf{x}_b$, is required and is usually found using a previous short forecast or from climatology data. Since the observations and state vector are not necessarily of the same variable or at the same grid points, nonlinear observation operators, $\mathcal{H}_i : \mathbb{R}^N \longrightarrow \mathbb{R}^{p_i}$, are defined at each time step to map elements from state space to observation space. The matrices $\mathbf{B} \in \mathbb{R}^{N \times N}$ and $\mathbf{R}_i \in \mathbb{R}^{p_i \times p_i}$ are error covariance matrices and describe the variance and correlations of the background and
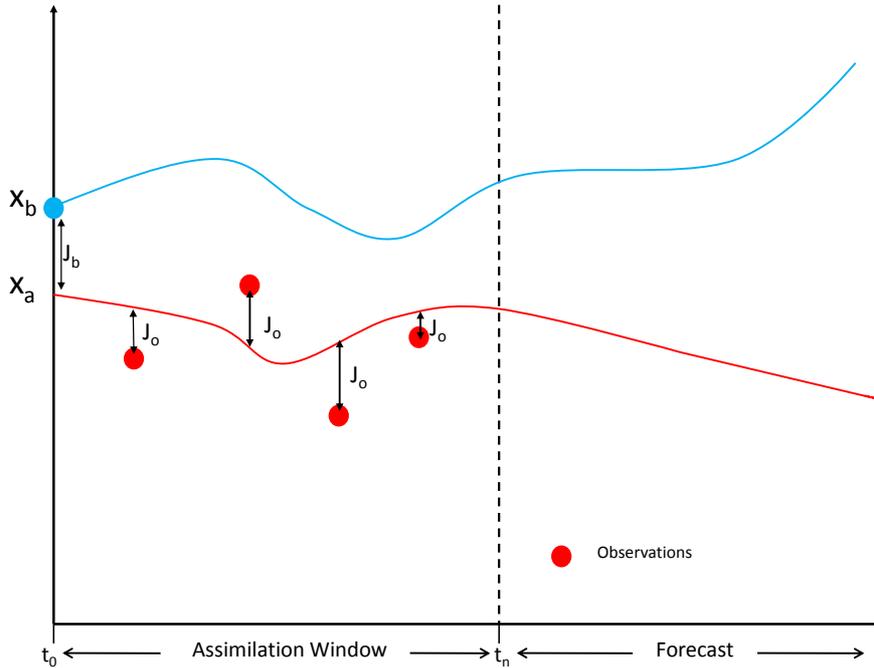
Figure 2.1: Schematic of 4 dimensional Variational data assimilation. Minimise the weighted distance between the analysis $\mathbf{x_a}$ and the background $\mathbf{x}_b$ (the J_b term) plus a weighted difference between the observations and the forecast of the analysis (the J_0 terms).

observation errors respectively.

Figure 2.1 gives a pictorial representation of 4DVar. The aim of Var is to combine the initial guess $\mathbf{x}_b$ with the observation data to give an improved initial guess, the analysis $\mathbf{x}_a$. The forecast from the analysis then fits the observation data more closely than the forecast from the background and is used to give future weather states beyond the assimilation window. Essentially 4DVar is a weighted non-linear least squares fit between observations (the J_o term) and a background state (the J_b term). If observations are only taken to be at one time point then the method is known as three-dimensional variational data assimilation (3DVar) and no forecast model is required in the assimilation. The cost function is solved iteratively using gradient optimization methods.

If we assume that the observation and model operators are linear then we write $\mathcal{H}_i = \mathbf{H}_i$ and $\mathcal{M}(t_{i-1}, t_i, \mathbf{x}_{i-1}) = \mathbf{M}_i \mathbf{x}_{i-1}$ where $\mathbf{H}_i$ and $\mathbf{M}_i$ are lin-

ear operators. With these linear assumptions the analysis has the following explicit form

$$\mathbf{x}_a = \mathbf{x}_b + (\mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}})^{-1} \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{d}} \qquad (2.4)$$

where

$$\hat{\mathbf{H}} = \left[\mathbf{H}_0^T, (\mathbf{H}_1 \mathbf{M}_1)^T, \ \ldots, \ (\mathbf{H}_n \mathbf{M}_n \ldots \mathbf{M}_1)^T \right]^T,$$

$$\hat{\mathbf{d}}^T = \left[\mathbf{d}_0^T, \ \mathbf{d}_1^T, \ \ldots, \ \mathbf{d}_n^T \right], \quad \text{with} \quad \mathbf{d}_i = \mathbf{y}_i - \mathbf{H}_i(\mathbf{x}_i),$$

and $\hat{\mathbf{R}}$ is the block diagonal matrix with blocks $\mathbf{R}_i$. Using (2.4) it can be shown that the errors in the analysis have the following covariance matrix [51]

$$\mathbf{A} = (\mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}})^{-1}. \qquad (2.5)$$

Due to the nonlinearity of the forecast and observation operators and the large number of variables in the solution vector the full variational data assimilation problem is computationally impractical to implement in a NWP system [46], [38]. We now describe a more computationally practical, incremental form of 4DVar.

2.3 Incremental 4DVar

The full 4DVar problem (2.1) cannot be solved efficiently enough to be used in an operational NWP system. It was not until the development of an incremental version of 4DVar, first proposed in [51], that the implementation of a variational data assimilation scheme in NWP became a realistic possibility. In incremental 4DVar a series of linearised quadratic cost functions are minimised subject to a linearised forecast model. The linearised cost function is solved iteratively at a lower resolution in the so-called *inner-loop* with the resultant minimiser then used to update the current, full resolution estimate of the analysis in a step known as the *outer-loop*. The reduced resolution in

the inner-loop produces a computationally more efficient algorithm. We now give an outline of the incremental 4DVar algorithm.

2.3.1 The 4DVar Algorithm

Let k represent the kth iteration of the outer loop.

1. If $k = 0$ then let $\mathbf{x}_0^{(k)} = \mathbf{x}_b$.

2. Using the non-linear forecast model (2.3) generate the $\mathbf{x}_i^{(k)}$ for each time step t_i.

3. Generate the *innovation* vectors at each time t_i,

$$\mathbf{d}_i^{(k)} = \mathbf{y}_i - \mathcal{H}_i(\mathbf{x}_i^{(k)}). \tag{2.6}$$

4. Define the increment, $\delta\mathbf{x}_0^{(k)} = \mathbf{x}_0^{(k+1)} - \mathbf{x}_0^{(k)}$.

5. Solve the inner-loop problem, i.e. minimise the following linearised cost function with respect to the increment $\delta\mathbf{x}_0^{(k)}$

$$\tilde{J}^{(k)}[\delta\mathbf{x}_0^{(k)}] = \frac{1}{2}[\delta\mathbf{x}_0^{(k)} - (\mathbf{x}_b - \mathbf{x}_0^{(k)})]^T\mathbf{B}^{-1}[\delta\mathbf{x}_0^{(k)} - (\mathbf{x}_b - \mathbf{x}_0^{(k)})]$$
$$+ \frac{1}{2}\sum_{i=0}^{n}(\mathbf{H}_i\delta\mathbf{x}_i^{(k)} - \mathbf{d}_i^{(k)})^T\mathbf{R}_i^{-1}(\mathbf{H}_i\delta\mathbf{x}_i^{(k)} - \mathbf{d}_i^{(k)},) \tag{2.7}$$

subject to the linearised model equations

$$\delta\mathbf{x}_i^{(k)} = \mathbf{M}(t_{i-1}, t_i)\,\delta\mathbf{x}_{i-1}^{(k)} \equiv \mathbf{M}_i\delta\mathbf{x}_{i-1}^{(k)}, \tag{2.8}$$

where $\mathbf{H}_i$ and $\mathbf{M}_i$ are linearisations of the observation and model forecast operators respectively, around $\mathbf{x}_i^{(k)}$.

6. Update the current estimate of the initial state

$$\mathbf{x}_0^{(k+1)} = \mathbf{x}_0^{(k)} + \delta\mathbf{x}_0^{(k)}.$$

7. Repeat steps 2 to 6 until desired convergence or maximum number of iterations has been reached.

Currently many operational NWP centres perform few if any outer loop updates and hence most of the computational expense is associated with the inner-loop minimisation in step 5 [54], [53]. Our research focuses on the inner-loop and therefore iteration superscripts are suppressed from here onwards. A convenient form of the linearised cost function which we make use of in subsequent chapters is

$$\tilde{J}[\delta\mathbf{x}_0] = \frac{1}{2}[\delta\mathbf{x}_0 - (\mathbf{x}_b - \mathbf{x}_0)]^T \mathbf{B}^{-1}[\delta\mathbf{x}_0 - (\mathbf{x}_b - \mathbf{x}_0)] + \frac{1}{2}(\hat{\mathbf{H}}\delta\mathbf{x}_0 - \hat{\mathbf{d}})^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{H}}\delta\mathbf{x}_0 - \hat{\mathbf{d}}), \tag{2.9}$$

where

$$\hat{\mathbf{H}} = \left[\mathbf{H}_0^T, (\mathbf{H}_1\hat{\mathbf{M}}_1)^T, \ \ldots, \ (\mathbf{H}_n\hat{\mathbf{M}}_n)^T\right]^T,$$

$$\hat{\mathbf{d}}^T = \left[\mathbf{d}_0^T, \ \mathbf{d}_1^T, \ \ldots, \ \mathbf{d}_n^T\right], \quad \text{with} \quad \mathbf{d}_i = \mathbf{y}_i - \mathcal{H}_i(\mathbf{x}_i),$$

and $\hat{\mathbf{R}}$ is the block diagonal matrix with blocks $\mathbf{R}_i$. Here we define

$$\hat{\mathbf{M}}_k = \mathbf{M}(t_k, t_0) = \mathbf{M}_k \ldots \mathbf{M}_1, \tag{2.10}$$

to be the linearised model equation mapping the state vector from time t_0 to t_k. In Chapter 6 we investigate the 3DVar system which can be written in the more simple form

$$\tilde{J}[\delta\mathbf{x}_0] = \frac{1}{2}[\delta\mathbf{x}_0 - (\mathbf{x}_b - \mathbf{x}_0)]^T \mathbf{B}^{-1}[\delta\mathbf{x}_0 - (\mathbf{x}_b - \mathbf{x}_0)] + \frac{1}{2}(\mathbf{H}_0\delta\mathbf{x}_0 - \mathbf{d}_0)^T \mathbf{R}_0^{-1}(\mathbf{H}_0\delta\mathbf{x}_0 - \mathbf{d}_0). \tag{2.11}$$

Of particular interest in this research is the Hessian, that is, the matrix of second derivatives, of (2.9) which has the following form

$$\tilde{J}'' = \mathbf{S} = \mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}}. \tag{2.12}$$

A comparison with (2.5) shows that the inverse of the Hessian is equal to the analysis error covariance matrix when the observation and model operators are exactly linear. More generally, the inverse of the Hessian is a close approximation to the analysis error covariance matrix when the observation and model operators are weakly non-linear [52].

2.3.2 The Inner-Loop Minimisation

In this section we discuss further details of the minimisation of the linearised cost function (2.7). Various methods can be used to minimise (2.7) including steepest descent, Newton's method, quasi-Newton and conjugate gradient methods [12]. Methods employed at the Met Office include the limited memory BFGS algorithm and the conjugate gradient method [5], [54]. In this thesis we use the conjugate gradient method due to its wide use in NWP and since it is the best compromise in terms of convergence rate and computer memory [12]. This scheme requires the calculation of the gradient of the cost function (2.7) given by

$$\nabla \tilde{J}[\delta \mathbf{x}_0] = \mathbf{B}^{-1}[\delta \mathbf{x}_0 - (\mathbf{x}_0^b - \mathbf{x}_0)] + \sum_{i=0}^{n} \mathbf{M}_1^T \ldots \mathbf{M}_i^T \mathbf{H}_i^T \mathbf{R}_i^{-1}(\mathbf{H}_i \mathbf{M}_i \ldots \mathbf{M}_1 \delta \mathbf{x}_0 - \mathbf{d}_i), \tag{2.13}$$

where $\mathbf{M}_i^T$ is the *adjoint* model of the linear operator $\mathbf{M}_i$ and propagates the input backwards in time from t_i to t_{i-1}. To calculate the gradient requires one forward run of the full non-linear model to calculate the innovations $\mathbf{d}_i$ $i = 0, \ldots, n$ and then a backward run of the adjoint model from time t_n to t_0 [46].

It should be noted that in practice the matrices $\mathbf{B}, \mathbf{H}_i, \mathbf{M}_i$ and the adjoints,

$\mathbf{H}_i^T, \mathbf{M}_i^T$, are too large to be stored and their action in matrix calculations, for instance, calculating the matrix-vector product $\mathbf{Bx}$, is not practical so other methods must be used. For instance a computationally cheaper alternative to calculating the adjoint via transposing the full matrix, $\mathbf{M}_i$, can be achieved directly from the forward model programming code statements [12]. Similarly, the storage and application of the matrix $\mathbf{B}$ in computations can be simplified by forming a control variable transform which we describe in the next section.

2.4   The Control Variable Transform

In this section we consider the control variable transform (CVT). In Var, since the background statistics are unknown the background covariance matrix $\mathbf{B}$ is also unknown. Additionally, due to the large dimension of the state vector the matrix $\mathbf{B}$ is too large (approximately $10^8 \times 10^8$) to implement or store explicitly [10]. To account for these difficulties many major NWP centres around the world implement a CVT in order to model the matrix $\mathbf{B}$ in practice [13], [43], [20], [11], [53], [54], [19] and [26].

In order to define the CVT we require the definition of the square root of a matrix.

Definition 2.4.1 *[39, Sec. 9.1] A matrix* $\mathbf{U}$ *is called the square-root of a symmetric matrix* $\mathbf{B} \in \mathbb{R}^{N \times N}$ *if*

$$\mathbf{B} = \mathbf{UU}^T, \tag{2.14}$$

We write $\mathbf{U} = \mathbf{B}^{1/2}$. *If* $\mathbf{U} = \mathbf{U}^T$ *then* $\mathbf{U}$ *is the unique symmetric square root of* $\mathbf{B}$.

The standard CVT uses a square root [39, Sec. 9.1], [10] of the background

covariance matrix to transform to new variables $\delta \mathbf{z}_0$

$$\delta \mathbf{x}_0 = \mathbf{B}^{1/2} \delta \mathbf{z}_0. \tag{2.15}$$

Substituting (2.15) into the linearised cost function (2.9) gives a new cost function

$$\hat{J}[\delta \mathbf{z}_0] = \frac{1}{2}[\delta \mathbf{z}_0 - (\mathbf{z}_0^b - \mathbf{z}_0)]^T [\delta \mathbf{z}_0 - (\mathbf{z}_0^b - \mathbf{z}_0)] + \frac{1}{2}(\hat{\mathbf{H}}\mathbf{B}^{1/2}\delta \mathbf{z}_0 - \hat{\mathbf{d}})^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{H}}\mathbf{B}^{1/2}\delta \mathbf{z}_0 - \hat{\mathbf{d}}), \tag{2.16}$$

where $\mathbf{z}_0 = \mathbf{B}^{-1/2}\mathbf{x}_0$ and $\mathbf{z}_0^b = \mathbf{B}^{-1/2}\mathbf{x}_0^b$. The application of the CVT results in a simplification of the background term. From the first term in (2.16) it is apparent that the errors of the new control variables are now uncorrelated with unit error variances, i.e. the error covariance matrix is now simply the identity matrix in terms of the new variables. Hence, the application of $\mathbf{B}^{-1/2}$ maps the original state vector to new variables with uncorrelated errors. We describe how the CVT can be constructed in practice in Chapter 8 where we consider the transform as it is implemented in the Met Office.

Most NWP centres include a form of the CVT based on the square root of the background error covariance matrix [13], [43], [20], [11], [53], [54], [19] and [26]. In this section we reformulate the incremental algorithm as introduced in Section 2.3 to incorporate the CVT.

Let k denote the kth outer-loop of the minimisation. The incremental 4DVar procedure that includes the CVT can be summarised as the following

1. Take a current guess of the atmosphere $\mathbf{x}_0^{(k)}$ with $\mathbf{x}_0^{(0)} = \mathbf{x}_0^b$.

2. Run the full non-linear model to calculate the evolved solution $\mathbf{x}_i^{(k)}$ at each time t_i and use these to calculate the innovations $\mathbf{d}_i^{(k)} = \mathbf{y}_i - \mathcal{H}_i[\mathbf{x}_i^{(k)}]$.

3. Define the increment $\delta \mathbf{x}_0^{(k)} = \mathbf{x}_0^b - \mathbf{x}_0^{(k)}$.

4. Apply the CVT to find

$$\mathbf{z}_0^b - \mathbf{z}_0^{(k)} = \mathbf{B}^{-1/2}(\mathbf{x}_0^b - \mathbf{x}_0^{(k)}) \qquad (2.17)$$

5. Find the inner loop solution $\delta\mathbf{z}_a^{(k)}$ by minimising (2.16).

6. Transform the inner loop solution $\delta\mathbf{z}_a^{(k)}$ back to the original variables using the CVT

$$\delta\mathbf{x}_a^{(k)} = \mathbf{B}^{1/2}\delta\mathbf{z}_a^{(k)}. \qquad (2.18)$$

7. Update the current estimate $\mathbf{x}_0^{(k+1)} = \mathbf{x}_0^{(k)} + \delta\mathbf{x}_0^{(k)}$.

8. Repeat the above steps until the desired stopping criteria has been obtained.

As in the original algorithm, the gradient must be calculated for each iteration of the minimisation. With the CVT the gradient now has the following slightly altered form

$$\nabla\hat{J}[\delta\mathbf{z}_0] = [\delta\mathbf{z}_0 - (\mathbf{z}_0^b - \mathbf{z}_0)] + \mathbf{B}^{T/2}\sum_{i=0}^{n}\hat{\mathbf{M}}_i^T\mathbf{H}_i^T\mathbf{R}_i^{-1}(\mathbf{H}_i\hat{\mathbf{M}}_i\mathbf{B}^{1/2}\delta\mathbf{z}_0 - \mathbf{d}_i). \quad (2.19)$$

where $\hat{\mathbf{M}}_i$ is given by (2.10). As before adjoint models $\mathbf{M}_i^T$ are required to calculate the gradient.

2.5 Summary

In this chapter we gave a brief introduction to data assimilation with the main focus on the variational method and its application to NWP. In practice, due to computational expense, an incremental form of Var is employed instead in which a series of reduced, linearised cost functions approximating the full problem are solved. We described another form of the incremental cost function which incorporates a control variable transform using a square

root of the background error covariance matrix. We showed that the CVT simplifies the background term in the incremental cost function and how the CVT is implemented in the Var minimisation scheme. We described some of the practical issues of the incremental implementation, in particular, the minimisation methods used and how the gradients are calculated for use in the algorithms. The main objective of this chapter has been to introduce the context of the thesis. The next chapter introduces the condition number which provides the necessary measure of the sensitivity of the analysis $\mathbf{x}_a$ to perturbations in the system.

Chapter 3

Condition Number

The condition number is a concept used in many different contexts to measure the sensitivity of a problem with respect to changes in certain quantities used in the problem [62], [23]. The condition number is a particularly useful quantity in numerical methods, where computational round off errors can accumulate and be magnified to produce inaccurate results [21, sec 4.8]. Problems with large condition numbers are called *ill-conditioned*. They can be highly sensitive to perturbations in the system and iterative methods used to solve them can be extremely slow to converge.

In this chapter we aim to introduce and highlight key properties of the conditioning associated to the problem of minimising quadratic functionals. Additionally, by the end of this chapter we aim to have the necessary tools for defining and estimating the condition number of the incremental Var cost function. We begin by stating some definitions and key properties of matrix and vector norms that are needed to define the condition number. Next, we define the condition number for our problem and show how this value indicates the sensitivity of the solution to small perturbations in the system. We then describe how ill-conditioning can be alleviated by *preconditioning* the problem. We introduce the conjugate gradient (CG) method and show how the condition number can impact the convergence rate of this iterative

scheme. Finally we introduce results for eigenvalues which are useful for estimating the conditioning of the Var minimisation in the later chapters.

3.1 Matrix and Vector Norms

Before we can define the condition number for a matrix we require some basic theory on matrix and vector norms. In this section we introduce normed vector spaces, in particular norms over the spaces of vectors $\mathbb{R}^N$ and matrices $\mathbb{R}^{N \times N}$ with real entries. In this thesis we are only concerned with square matrices. For further information of norms of vectors and matrices see [23, Chap 2].

Definition 3.1.1 *[70, Chap 2], [23, Chap 2] A normed vector space is a pair $(X, ||.||)$ consisting of a vector space, X, and a mapping, $||.|| : X \longrightarrow \mathbb{R}$, satisfying the following conditions,*

1. *$||\mathbf{x}|| \geq 0, \qquad \forall \mathbf{x} \in X$ and $||\mathbf{x}|| = 0 \Leftrightarrow \mathbf{x} = 0$.*

2. *$||\lambda \mathbf{x}|| = |\lambda| ||\mathbf{x}||$ for all $\lambda \in \mathbb{C}$ and $\mathbf{x} \in X$.*

3. *$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$ for all $\mathbf{x}, \mathbf{y} \in X$.*

In this section X is either the space of real-valued vectors or real-valued matrices, $\mathbb{R}^N$ and $\mathbb{R}^{N \times N}$ respectively. A commonly used class of vector norms are the *p-norms* on $\mathbb{R}^N$.

Definition 3.1.2 *The p-norm, $||.||_p$ (for $p \geq 1$), defined on $\mathbb{R}^N$, is a norm defined by*

$$||\mathbf{x}||_p = (|x_1|^p + |x_2|^p + \ldots + |x_N|^p)^{\frac{1}{p}}, \qquad (3.1)$$

where $\mathbf{x} = (x_1, \ldots, x_N)^T \in \mathbb{R}^N$, where T represents the transpose operator.

In this thesis we study the important class of positive-definite, symmetric matrices.

Definition 3.1.3 *Let* $\mathbf{S} \in \mathbb{R}^{N \times N}$ *with* $(i, j)^{th}$ *entry* $s_{i,j}$. *Then* $\mathbf{S}$ *is symmetric if* $s_{i,j} = s_{j,i}$ *for all* $i, j = 1, 2, \dots, N$.

Definition 3.1.4 *A matrix* $\mathbf{S} \in \mathbb{R}^{N \times N}$ *is positive-definite if* $\mathbf{x}^T \mathbf{S} \mathbf{x} > 0$ *for all nonzero vectors* $\mathbf{x} \in \mathbb{R}^N$.

We also use the eigenvalues and eigenvectors of a matrix in subsequent chapters.

Definition 3.1.5 *An eigenvalue of a matrix* $\mathbf{S} \in \mathbb{R}^{N \times N}$ *is a scalar* $\lambda \in \mathbb{R}$ *satisfying*

$$\mathbf{S}\mathbf{v} = \lambda\mathbf{v}, \tag{3.2}$$

for some non-zero vector $\mathbf{v} \in \mathbb{R}^N$ *which is the associated eigenvector.*

In the special case where $\mathbf{S}$ is symmetric and positive definite all eigenvalues are real and strictly positive [23, Sec. 8.1.1].

An important class of norms can be induced on $\mathbb{R}^N$ using symmetric positive-definite matrices.

Definition 3.1.6 *Let* $\mathbf{S} \in \mathbb{R}^{N \times N}$ *be a symmetric positive definite matrix then the S-norm,* $||.||_S$, *on* $\mathbb{R}^N$ *is defined by*

$$||\mathbf{x}||_S^2 = \mathbf{x}^T \mathbf{S} \mathbf{x} \tag{3.3}$$

for any $\mathbf{x} \in \mathbb{R}^N$.

We also require norms for matrices. Of particular interest in this thesis are the matrix norms induced from the vector p-norms (3.1).

Definition 3.1.7 *The matrix p-norm on $\mathbb{R}^{N \times N}$, $||.||_p$ ($p \geq 1$), is a norm defined by,*

$$||\mathbf{S}||_p = \sup_{\mathbf{x} \neq 0} \frac{||\mathbf{S}\mathbf{x}||_p}{||\mathbf{x}||_p}, \tag{3.4}$$

for $\mathbf{S} \in \mathbb{R}^{N \times N}$ and $\mathbf{x} \in \mathbb{R}^N$.

Explicit representations can be found for certain p. In this thesis we will concentrate on the matrix 2-norm applied to symmetric, square matrices. In this case, for a matrix $\mathbf{S} \in \mathbb{R}^{N \times N}$, we have [47, Appendix A]

$$||\mathbf{S}||_2 = |\lambda_{\max}(\mathbf{S})|, \tag{3.5}$$

where $\lambda_{\max}(\mathbf{S})$ is the eigenvalue of the matrix $\mathbf{S}$ with largest magnitude.

We will also make use of the ∞-norm defined by [47, Appendix A]

$$||\mathbf{S}||_\infty = \max_{1 \leq i \leq N} \sum_{j=1}^{N} |s_{i,j}|, \tag{3.6}$$

where $s_{i,j}$ is the $(i,j)^{th}$ element of the matrix $\mathbf{S} \in \mathbb{R}^{N \times N}$.

Having developed the necessary background we now use the definition of matrix norms to define the condition number.

3.2 Condition Number of the Hessian

In this thesis we are concerned with minimising the function (2.9) with respect to vector $\delta\mathbf{x}_0 \in \mathbb{R}^N$. This is clearly a quadratic function of the form,

$$\tilde{J}(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{S}\mathbf{w} - \mathbf{w}^T\mathbf{b} + c, \tag{3.7}$$

where

$$\mathbf{w} = \delta\mathbf{x}_0 \in \mathbb{R}^N, \tag{3.8}$$

$$\mathbf{S} = (\mathbf{B}^{-1} + \hat{\mathbf{H}}^T\hat{\mathbf{R}}\hat{\mathbf{H}}) \in \mathbb{R}^{N \times N}, \tag{3.9}$$

$$\mathbf{b} = \mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}_0) - \hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{d}} \in \mathbb{R}^N, \tag{3.10}$$

$$c = \frac{1}{2}((\mathbf{x}_b - \mathbf{x}_0)^T\mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}_0) + \hat{\mathbf{d}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{d}}) \in \mathbb{R}. \tag{3.11}$$

We note that the Hessian, $\mathbf{S}$, is a symmetric and positive-definite matrix and therefore there exists a unique minimum to (3.7).

Minimising (3.7) is therefore equivalent to finding the solution to the gradient equation $\nabla\tilde{J} = 0$, that is, solving the linear system

$$\mathbf{Sw} = \mathbf{b}. \tag{3.12}$$

When computing the solution to (3.7) or equivalently (3.12) we would like to know how sensitive the solution $\mathbf{w}$ is to small changes in the system components $\mathbf{S}$ and $\mathbf{b}$. The condition number is one way to measure the sensitivity of the solution and we consider this in the next section.

3.2.1   The Condition Number of a Linear System

Definition 3.2.1 *The condition number, $\kappa_p(\mathbf{S})$, of a matrix $\mathbf{S} \in \mathbb{R}^{N \times N}$ with respect to any p-norm $||.||$ is defined to be*

$$\kappa_p(\mathbf{S}) = ||\mathbf{S}||_p||\mathbf{S}^{-1}||_p. \tag{3.13}$$

Consider how changes in the linear system (3.12) produce changes in the solution $\mathbf{w}$. More precisely, suppose that small perturbations $\Delta\mathbf{S} \in \mathbb{R}^{N \times N}$ and $\Delta\mathbf{b} \in \mathbb{R}^N$ are applied to $\mathbf{S}$ and $\mathbf{b}$ respectively where $||\Delta\mathbf{S}||_p$ and $||\Delta\mathbf{b}||_p$

have order ϵ. The new linear system is now

$$(\mathbf{S} + \Delta\mathbf{S})\mathbf{w}_\epsilon = (\mathbf{b} + \Delta\mathbf{b}), \qquad (3.14)$$

with new solution $\mathbf{w}_\epsilon$.

The following bound on the relative change in the solution applies with respect to any matrix p-norm.

Theorem 3.2.2 *[23, sec 2.7]*

$$\frac{||\mathbf{w} - \mathbf{w}_0^\epsilon||_p}{||\mathbf{w}||_p} \leq \kappa_p(\mathbf{S}) \left(\frac{||\Delta\mathbf{S}||_p}{||\mathbf{S}||_p} + \frac{||\Delta\mathbf{b}||_p}{||\mathbf{b}||_p} \right) + O(\epsilon^2). \qquad (3.15)$$

Proof See [23, sec 2.7]

Theorem 3.2.2 shows that even if the relative perturbations, $\frac{||\Delta\mathbf{S}||_p}{||\mathbf{S}||_p}$, $\frac{||\Delta\mathbf{b}||_p}{||\mathbf{b}||_p}$ in $\mathbf{S}$ and $\mathbf{b}$ respectively, are small, the relative perturbation in the solution can be large if the condition number is large. The condition number therefore gives an indication of the sensitivity of the solution of (3.12) to small changes in the components $\mathbf{S}$, $\mathbf{b}$. If the condition number is large then we say that $\mathbf{S}$ is ill-conditioned which implies that the solution to (3.12) can be highly sensitive to small perturbations in the system.

The condition number is a norm dependent quantity. However, since matrix norms on $\mathbb{R}^{N \times N}$ are equivalent, if $\mathbf{S} \in \mathbb{R}^{N \times N}$ is ill-conditioned in one norm it is ill-conditioned in any norm [23, Sec 2.7.2]. In this thesis we will be restricting our attention to the 2-norm definition of the condition number. This is also known as the spectral condition number

Definition 3.2.3 *The spectral condition number of a matrix* $\mathbf{S} \in \mathbb{R}^{N \times N}$ *is defined to be*

$$\kappa_2(\mathbf{S}) = ||\mathbf{S}||_2 ||\mathbf{S}^{-1}||_2. \qquad (3.16)$$

For symmetric positive-definite matrices we have a useful representation of the spectral condition number.

Theorem 3.2.4 *If $\mathbf{S} \in \mathbb{R}^{N \times N}$ is a symmetric and positive-definite matrix then we can write*

$$\kappa_2(\mathbf{S}) = \frac{\lambda_{\max}(\mathbf{S})}{\lambda_{\min}(\mathbf{S})}, \tag{3.17}$$

where $\lambda_{\max}(\mathbf{S})$ and $\lambda_{\min}(\mathbf{S})$ represent the largest and smallest eigenvalue of $\mathbf{S}$.

Proof See [7, Appendix A.2] or [39, Appendix B.7]

From now on we assume that $\kappa(\mathbf{S})$ means the spectral condition number. The sensitivity of the solution to the linear system can be measured by the condition number of the matrix $\mathbf{S}$. Due to the equivalence of minimsing the quadratic function (3.7) and solving the linear system (3.12), the sensitivity of the solution that minimises (3.7) can therefore be measured by the condition number of the Hessian of the cost function.

When trying to solve a linear system such as (3.12) numerical round off errors can cause small perturbations in the components, $\mathbf{S}, \mathbf{b}$. As we have shown in this section, for an ill-conditioned problem small changes in the system can lead to a very different, and hence very inaccurate, solution. In the next section we consider one method for dealing with an ill-conditioned problem.

3.2.2 Preconditioning

In the numerical methods applied to solving systems such as (3.12), perfect arithmetic is not possible and rounding errors naturally arise into the computations. These errors can accumulate in the components and ultimately lead to an inaccurate solution. In the case of an ill-conditioned linear system

of the form (3.12) we showed in the previous section that even small changes in $\mathbf{S}$ and $\mathbf{b}$ can cause very different solutions. In practice this means that when applying numerical methods to an ill-conditioned problem the solution is unreliable. Additionally, as we shall see in the next section, ill-conditioning can also lead to slow convergence of the iterative methods used to solve the linear system (3.12).

One common solution to issues related to ill-conditioning is to change the system to one with a lower condition number and solve this equivalent problem. This is called *preconditioning*. In this thesis we symmetrically precondition the system (3.12) with a symmetric positive definite matrix $\mathbf{P}$ to give the following system [23, Sec 10.3],

$$\hat{\mathbf{S}}\mathbf{v} = \hat{\mathbf{b}}, \tag{3.18}$$

where $\hat{\mathbf{S}} = \mathbf{P}^{1/2}\mathbf{S}\mathbf{P}^{1/2}$ is the new preconditioned matrix, $\mathbf{v} = \mathbf{P}^{-1/2}\mathbf{w}$, $\hat{\mathbf{b}} = \mathbf{P}^{1/2}\mathbf{b}$ and $\mathbf{P}^{1/2}$ is the symmetric positive-definite square root of the matrix $\mathbf{P}$ [39, Sec. 9.1].

If the preconditioner is correctly chosen the new system (3.18) has a reduced condition number

$$\kappa(\hat{\mathbf{S}}) < \kappa(\mathbf{S}) \tag{3.19}$$

causing the solution to (3.18) to be less sensitive to perturbations. Once the solution $\mathbf{v}$ is found, the preconditioner can be applied again to find the solution to the original problem

$$\mathbf{w} = \mathbf{P}^{1/2}\mathbf{v}. \tag{3.20}$$

Choosing $\mathbf{P} = \mathbf{S}^{-1}$ gives the best reduction in the condition number since

$$\kappa(\hat{\mathbf{S}}) = \kappa(\mathbf{S}^{-1/2}\mathbf{S}\mathbf{S}^{-1/2}) = \kappa(\mathbf{I}) = 1. \tag{3.21}$$

However in this case constructing the preconditioner requires inverting $\mathbf{S}$,

which is equivalent to solving the original ill-conditioned problem (3.12). The goal in preconditioning is to construct a cost effective preconditioner which approximates the inverse of the original matrix $\mathbf{S}$. A compromise must be found between the cost of constructing the preconditioner and the reduction in cost of solving the preconditioned problem.

For the quadratic function (3.7), preconditioning is equivalent to performing a variable transform $\mathbf{v} = \mathbf{P}^{-1/2}\mathbf{w}$. The new cost function is now

$$\tilde{\tilde{J}}(\mathbf{v}) = \mathbf{v}^T\mathbf{P}^{1/2}\mathbf{S}\mathbf{P}^{1/2}\mathbf{v} - \mathbf{v}^T\mathbf{P}^{1/2}\mathbf{b} + c, \qquad (3.22)$$

and is minimised with respect to the new variable $\mathbf{v}$. It is easy to see by differentiation of the new cost function that minimising (3.22) is equivalent to solving (3.18).

In this section we have introduced the condition number for the minimisation of a quadratic cost function, or equivalently a linear matrix system, in the case of a symmetric, positive definite Hessian $\mathbf{S}$. We also showed how preconditioning can be used to alleviate the ill-conditioning of the system. We now introduce one of the minimisation algorithms that we use in this thesis for solving problems of the form (3.7) or (3.12).

3.3 Conjugate Gradient

3.3.1 The Conjugate Gradient Method

In this section we introduce the conjugate gradient (CG) method and show how the condition number relates to the convergence properties. The CG method is an iterative scheme which is widely used to solve linear problems such as (3.12), or equivalently (3.7), where $\mathbf{S}$ is a positive-definite symmetric matrix. We use the CG method to solve the inner-loop minimisation in

the incremental Var scheme (see Section 2.3.1), including for the Met Office operational experiments in Chapter 8.

The CG algorithm is theoretically a direct solver of (3.12). However due to round-off errors it is more commonly used as an iterative method. The following is a common form of the CG algorithm for minimising a function of the form (3.7), [23, Sec 10.2].

1. First initialize the system

$$\mathbf{w}^{(0)} = \mathbf{0}, \tag{3.23}$$

$$\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{S}\mathbf{w}^{(0)} \tag{3.24}$$

$$\mathbf{p}^{(0)} = \mathbf{r}^{(0)}, \tag{3.25}$$

2. With k representing the kth iteration, iterate over the following loop until the convergence criteria has been achieved

$$\alpha^{(k)} = \frac{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}{(\mathbf{p}^{(k)})^T \mathbf{S} \mathbf{p}^{(k)}}, \tag{3.26}$$

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}, \tag{3.27}$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha^{(k)} \mathbf{S} \mathbf{p}^{(k)}, \tag{3.28}$$

$$\beta^{(k)} = \frac{(\mathbf{r}^{(k+1)})^T \mathbf{r}^{(k+1)}}{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}, \tag{3.29}$$

$$\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta^{(k)} \mathbf{p}^{(k)}. \tag{3.30}$$

Here $\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{S}\mathbf{w}^{(k)}$ is called the residual at $\mathbf{w}^{(k)}$. At each iteration k, the conjugate gradient method finds the minimum of (3.7) over the subspace spanned by the search directions $\mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(k)} \in \mathbb{R}^N$ which are $\mathbf{S}$-conjugate i.e.

$$(\mathbf{p}^{(i)})^T \mathbf{S} \mathbf{p}^{(j)} = 0 \qquad i \neq j. \tag{3.31}$$

In particular this implies, for calculations using perfect arithmetic, that the CG method should converge in at most N iterations [21, Sec. 4.8].

Unfortunately rounding errors can lead to a loss of orthogonality, so the number of iterations may be larger in practice [62], [23, Sec. 10.2]. Additionally, in cases where N is very large, for instance in NWP applications, it is computationally impractical to run CG to the full N iterations. For these reasons suitable convergence criteria are required to terminate the CG iteration at the appropriate iterate or by stopping at a prescribed maximum number of iterations. The following theorem provides a useful indicator for the stopping criteria.

Theorem 3.3.1 *Suppose that* $\mathbf{w}^{(k)}$ *is the estimate to the true solution,* $\mathbf{w}$, *of (3.12) at the kth iteration of CG then the relative error satisfies,*

$$\frac{||\mathbf{w} - \mathbf{w}^{(k)}||_2}{||\mathbf{w}||_2} \leq \kappa(\mathbf{S}) \frac{||\mathbf{r}^{(k)}||_2}{||\mathbf{b}||_2}. \tag{3.32}$$

Proof See [33, Sec. 1.1]

Therefore a reduction in the relative residual $||\mathbf{r}_k||_2/||\mathbf{b}||_2$ is sufficient to guarantee a reduction in the relative error of the solution. The presence of the condition number in (3.32) implies that an ill-conditioned system may require much greater reduction in the relative residual compared to that of a well conditioned system in order to guarantee the same level of accuracy. In particular it means a more ill-conditioned system could have a less accurate solution if the iterative scheme is prematurely stopped.

The relevance of the condition number to the rate of convergence of CG can be further shown by the following error bound [23, Sec. 10.2], [47],

$$||\mathbf{w}^{(k)} - \mathbf{w}||_{\mathbf{S}} \leq 2||\mathbf{w}^{(0)} - \mathbf{w}||_{\mathbf{S}} \left(\frac{\sqrt{\kappa(\mathbf{S})} - 1}{\sqrt{\kappa(\mathbf{S})} + 1} \right)^k, \tag{3.33}$$

where $\mathbf{w}^{(0)}$ is the initial guess. This bound tends to be an overestimate but for small condition numbers ($\kappa(\mathbf{S}) \approx 1$) the right hand side of (3.33) will be approximately zero for very small k implying that the conjugate gradient

method will converge quickly to the truth. Other factors also affect the convergence rate of CG, in particular, the distribution of eigenvalues of $\mathbf{S}$, but this will not be investigated in this thesis [29].

One final note must be made about the conjugate gradient method as presented here with regard to the operational implementation in NWP. The number of elements in the Hessian of the incremental Var cost function (2.9) tends to be of the order of above $10^8 \times 10^8$. Hence, storing and applying the Hessian in CG is impossible. In practice explicit multiplication with the Hessian can be avoided when minimising a quadratic function (see [7, Sec 11.2.4, 11.2.5]). Define a trial point $\hat{\mathbf{w}}^{(k)} = \mathbf{w}^{(k)} + \tau^{(k)}\mathbf{p}^{(k)}$, where $\mathbf{w}^{(k)}$ is the current approximation at iteration k. Letting $\hat{\mathbf{g}}^{(k)} = \mathbf{S}\hat{\mathbf{w}}^{(k)} - \mathbf{b}$ and $\mathbf{g}^{(k)} = \mathbf{S}\mathbf{w}^{(k)} - \mathbf{b}$ represent the gradient of (3.7) at $\hat{\mathbf{w}}^{(k)}$ and $\mathbf{w}^{(k)}$ respectively then the multiplication by the Hessian in (3.26) and (3.28) can be calculated using

$$\frac{\hat{\mathbf{g}}^{(k)} - \mathbf{g}^{(k)}}{\tau^{(k)}} = \frac{\mathbf{S}(\mathbf{w}^{(k)} + \tau^{(k)}\mathbf{p}^{(k)}) - \mathbf{S}\mathbf{w}^{(k)}}{\tau^{(k)}} = \mathbf{S}\mathbf{p}^{(k)}. \qquad (3.34)$$

Thus the product $\mathbf{S}\mathbf{p}^{(k)}$ can be found from two runs of the adjoint model (2.13) to calculate the gradients $\hat{\mathbf{g}}^{(k)}$ and $\mathbf{g}^{(k)}$. This procedure is used within the Met Office inner-loop minimisation [5]. We now present a preconditioned version of CG.

3.3.2 Preconditioned CG

As seen from (3.33) the condition number can influence the convergence rate of the CG method. Therefore the CG algorithm should converge in fewer iterations when applied to (3.18) than to (3.12) if the preconditioning has reduced the condition number. An adjustment can be made to the CG algorithm so that the matrix $\hat{\mathbf{S}}$ does not need to be formed. Below is a modified conjugate gradient method which is equivalent to solving the preconditioned system (3.18) [23, Sec 10.3], [21]

1. First initialize the system

$$\mathbf{w}^{(0)} = \mathbf{0}, \tag{3.35}$$

$$\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{S}\mathbf{w}^{(0)}, \tag{3.36}$$

2. For the first iteration solve

$$\mathbf{P}^{-1}\mathbf{z}^{(0)} = \mathbf{r}^{(0)}, \tag{3.37}$$

and set $\mathbf{p}^{(0)} = \mathbf{z}^{(0)}$

3. Then iterate over the following loop until the convergence criteria has been achieved

$$\alpha^{(k)} = \frac{(\mathbf{r}^{(k)})^T \mathbf{z}^{(k)}}{(\mathbf{p}^{(k)})^T \mathbf{S}\mathbf{p}^{(k)}}, \tag{3.38}$$

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \alpha^{(k)}\mathbf{p}^{(k)}, \tag{3.39}$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha^{(k)}\mathbf{S}\mathbf{p}^{(k)}, \tag{3.40}$$

$$\text{Solve} \quad \mathbf{P}^{-1}\mathbf{z}^{(k+1)} = \mathbf{r}^{(k+1)} \tag{3.41}$$

$$\beta^{(k)} = \frac{(\mathbf{r}^{(k+1)})^T \mathbf{z}^{(k+1)}}{(\mathbf{r}^{(k)})^T \mathbf{z}^{(k)}}, \tag{3.42}$$

$$\mathbf{p}^{(k+1)} = \mathbf{z}^{(k+1)} + \beta^{(k)}\mathbf{p}^{(k)}. \tag{3.43}$$

The choice of $\mathbf{P}^{-1}$ is important in the effectiveness of the preconditioned conjugate gradient (PCG) method. In particular, each iteration of the PCG method requires solving $\mathbf{P}^{-1}\mathbf{z}^{(k+1)} = \mathbf{r}^{(k+1)}$. However, if the preconditioner is an ill-conditioned matrix, then the solution to this equation is potentially inaccurate.

In the next section we outline some results which we will use in calculating and estimating the eigenvalues in the later chapters.

3.4 Useful Results for Eigenvalues

In this thesis we aim to find estimates of the spectral condition number of the Hessian of the Var cost functions (2.9) and (2.16). This requires putting estimates on the eigenvalues of the Hessian. In this section we present useful results for estimating the eigenvalues of a matrix. We also introduce a useful group of matrices called circulant matrices which are used in this thesis and which have special eigenvalue properties.

3.4.1 Simple Bounds on Eigenvalues

In this section we give a couple of simple bounds that can be placed on the eigenvalues of a matrix. These are used to construct bounds on the condition number. Throughout this section we will assume that $\lambda_k(\mathbf{C})$ denotes the kth largest eigenvalue of the symmetric matrix $\mathbf{C}$.

Theorem 3.4.1 *[23, Sec. 8.1] Consider two symmetric matrices $\mathbf{S}_1, \mathbf{S}_2 \in \mathbb{R}^{N \times N}$. The kth largest eigenvalue of the matrix sum $\mathbf{S}_1 + \mathbf{S}_2$ satisfies the following*

$$\lambda_k(\mathbf{S}_1) + \lambda_n(\mathbf{S}_2) \leq \lambda_k(\mathbf{S}_1 + \mathbf{S}_2) \leq \lambda_k(\mathbf{S}_1) + \lambda_1(\mathbf{S}_2). \qquad (3.44)$$

Proof See [67].

We can show another simple result for the upper bound on the eigenvalues of a matrix $\mathbf{A}$

Theorem 3.4.2 *Consider a matrix $\mathbf{S} \in \mathbb{R}^{N \times N}$, then the following holds,*

$$|\lambda_k(\mathbf{S})| \leq ||\mathbf{S}||_p \qquad (3.45)$$

for any $p \geq 1$.

Proof See [7, Sec. A.1].

3.4.2 The Rayleigh Quotient

The Rayleigh quotient is an important function for estimating the eigenvalues of a matrix.

Definition 3.4.3 *[7, Sec. 4.4] The Rayleigh quotient for a symmetric matrix* $\mathbf{S} \in \mathbb{R}^{N \times N}$ *is defined to be*

$$R_{\mathbf{S}}(\mathbf{x}) = \frac{\mathbf{x}^H \mathbf{S} \mathbf{x}}{\mathbf{x}^H \mathbf{x}}, \tag{3.46}$$

for a vector $\mathbf{x} \in \mathbb{C}^N$.

Here, $\mathbf{x}^H$ means the complex-conjugate transpose of $\mathbf{x}$. Note that if $\mathbf{v}$ is an eigenvector of the matrix $\mathbf{S}$ then $R_{\mathbf{S}}(\mathbf{v})$ will be the corresponding eigenvalue. Let $\mathbf{S} \in \mathbb{R}^{N \times N}$ be a symmetric and positive definite matrix and $\mathbf{v}_{\max}$ and $\mathbf{v}_{\min}$ be the eigenvectors corresponding to the largest and smallest eigenvalues of $\mathbf{S}$ respectively. Then $R_{\mathbf{S}}(\mathbf{v}_{\max}) = \lambda_{\max}(\mathbf{S})$ and $R_{\mathbf{S}}(\mathbf{v}_{\min}) = \lambda_{\min}(\mathbf{S})$. We can also see that the Rayleigh quotient is bounded by these eigenvalues.

Theorem 3.4.4 *Let* $\mathbf{S} \in \mathbb{R}^{N \times N}$ *be a symmetric matrix. Then the following bounds hold on the Rayleigh quotient (3.46)*

$$\lambda_{\min}(\mathbf{S}) \leq R_{\mathbf{S}}(\mathbf{x}) \leq \lambda_{\max}(\mathbf{S}). \tag{3.47}$$

Proof See [57, Sec 5.9]

In the definition (3.46) we have allowed complex valued vectors, but the eigenvalues of a symmetric matrix are always real.

3.4.3 A Special System

In this section we introduce two matrix systems that have the same non-zero eigenvalues. In subsequent chapters this will allow us to write the eigenvalues of the Hessian of the Var cost function (2.16) in a simpler form.

Theorem 3.4.5 *Consider a matrix $\mathbf{E} \in \mathbb{R}^{N \times M}$ with $M < N$. Then the eigenvalues of $\mathbf{E}\mathbf{E}^T$ and $\mathbf{E}^T\mathbf{E}$ are equal. In addition $\mathbf{E}\mathbf{E}^T$ has $N - M$ extra eigenvalues equal to zero.*

Proof See [23, Sec 8.6]

3.4.4 Circulant Matrices

In this final section we introduce circulant matrices and some of their special properties. We find that a circulant matrix will be the natural structure for our one-dimensional covariance matrices when periodic boundary conditions are imposed. In this section one of the main advantages of circulant matrices is the simple form of the eigenvectors and eigenvalues.

Definition 3.4.6 *[24, Chap 3] A circulant matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$ is a matrix of the form*

$$\mathbf{C} = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 & \ldots & c_{N-2} & c_{N-1} \\ c_{N-1} & c_0 & c_1 & c_2 & \ldots & c_{N-3} & c_{N-2} \\ c_{N-2} & c_{N-1} & c_0 & & & & \vdots \\ & & & \ddots & \ddots & & \\ \vdots & & & & \ddots & & c_2 \\ c_2 & & & & \ddots & c_0 & c_1 \\ c_1 & c_2 & \ldots & & c_{N-2} & c_{N-1} & c_0 \end{pmatrix}. \tag{3.48}$$

Hence each row of a circulant matrix is a cyclic perturbation of the row above and knowledge of the entire matrix can be determined by specifying the top row alone. One special property of circulant matrices is that their eigenvalues can be written explicitly as a discrete Fourier transform of the top row.

Theorem 3.4.7 *[24] Let $\mathbf{C}$ be the matrix (3.48). Then the eigenvalues of $\mathbf{C}$ can be written*

$$\nu_m = \sum_{k=0}^{N-1} c_k e^{-2\pi i m k/N}, \tag{3.49}$$

with corresponding eigenvectors

$$\mathbf{v}_m = \frac{1}{\sqrt{N}}(1, e^{-2\pi i m/N}, \ldots, e^{-2\pi i m(N-1)/N})^T \tag{3.50}$$

for $m = 0, \ldots, N-1$.

Proof [24, Sec 3.1] Let $\mathbf{v} = (v_0, \ldots, v_{N-1})^T$ be an eigenvector of $\mathbf{C}$ with corresponding eigenvalue ν then we have

$$\mathbf{C}\mathbf{v} = \nu\mathbf{v}, \tag{3.51}$$

which can be written as the series of equations of the form

$$\sum_{k=0}^{m-1} c_{N-m+k} v_k + \sum_{k=m}^{N-1} c_{k-m} v_k = \nu v_m, \quad \text{with} \quad m = 0, \ldots N-1. \tag{3.52}$$

Changing the summation variable this is equivalent to

$$\sum_{k=N-m}^{N-1} c_k v_{k-(N-m)} + \sum_{k=0}^{N-1-m} c_k v_{k+m} = \nu v_m, \quad \text{with} \quad m = 0, \ldots N-1. \tag{3.53}$$

Now, let $v_k = \rho^k$, substitute into (3.53) and cancel the common factor of ρ^m to give

$$\rho^{-N} \sum_{k=N-m}^{N-1} c_k \rho^k + \sum_{k=0}^{N-1-m} c_k \rho^k = \nu, \quad \text{with} \quad m = 0, \ldots N-1. \tag{3.54}$$

36

Therefore if we choose ρ as one of the N distinct N^{th} roots of unity $(\rho^{-N} = 1)$ then we obtain the eigenvalue

$$\nu = \sum_{k=0}^{N-1} c_k \rho^k, \tag{3.55}$$

with corresponding eigenvector $\mathbf{v} = N^{-1/2}(1, \rho^1, \rho^2, \ldots, \rho^{N-1})^T$. Since ρ is of the form $e^{-2\pi i m k/N}$ this finalises the proof.

Circulant matrices have a convenient eigendecomposition. Let $\boldsymbol{\Lambda}$ be the diagonal matrix whose entries are the eigenvalues (3.49). Then a circulant matrix $\mathbf{C}$ can be written as

$$\mathbf{C} = \mathbf{F}\boldsymbol{\Lambda}\mathbf{F}^H, \tag{3.56}$$

where $\mathbf{F}^H$ is the complex conjugate transpose of $\mathbf{F}$ and $\mathbf{F}$ is the Fourier matrix whose columns are the eigenvectors (3.50),

$$\mathbf{F} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & 1 & \ldots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \ldots & \omega^{(N-1)} \\ 1 & \omega^2 & \omega^4 & \omega^6 & & \omega^{2(N-1)} \\ & & & & & \\ \vdots & & \vdots & \vdots & & \vdots \\ & & & & & \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & \ldots & \omega^{(N-1)^2} \end{pmatrix}, \tag{3.57}$$

where $\omega = \exp(-2\pi i/N)$. The Fourier matrix is unitary, that is it satisfies

$$\mathbf{F}^H\mathbf{F} = \mathbf{F}\mathbf{F}^H = \mathbf{I}_N. \tag{3.58}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix. Conversely, any matrix which has the eigendecomposition (3.56) is circulant [24]. The following theorem gives some other useful properties of a circulant matrix.

Theorem 3.4.8 *[24] Let* $\mathbf{C}$ *be a circulant matrix with eigendecomposition* $\mathbf{C} = \mathbf{F}\boldsymbol{\Lambda}\mathbf{F}^H$. *Then the inverse of the matrix* $\mathbf{C}$ *is also circulant with*

$$\mathbf{C}^{-1} = \mathbf{F}\boldsymbol{\Lambda}^{-1}\mathbf{F}^H. \tag{3.59}$$

Additionally, the product of any two circulant matrices $\mathbf{C}_1 = \mathbf{F}\boldsymbol{\Lambda}_1\mathbf{F}^H$ *and* $\mathbf{C}_2 = \mathbf{F}\boldsymbol{\Lambda}_2\mathbf{F}^H$ *is also circulant with*

$$\mathbf{C}_1\mathbf{C}_2 = \mathbf{F}\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2\mathbf{F}^H. \tag{3.60}$$

Proof The results follow directly from (3.58).

Using (3.56) we can find the symmetric square root of a circulant matrix $\mathbf{C} = \mathbf{F}\boldsymbol{\Lambda}\mathbf{F}^H$ with

$$\mathbf{C}^{1/2} = \mathbf{F}\boldsymbol{\Lambda}^{1/2}\mathbf{F}^H. \tag{3.61}$$

We now summarise this chapter.

3.5 Summary

In this chapter we have introduced the concept of condition number of a matrix and shown how it can be used to indicate the sensitivity of the solution to a linear system (3.12) to perturbations in $\mathbf{S}$ and $\mathbf{b}$. We showed the equivalence between solving the linear system (3.12) and minimising the quadratic cost function (3.7). We showed that an ill-conditioned problem can have a detrimental effect on both the accuracy of the solution of the linear system (3.12) (or equivalently the solution of (3.7)) and on the convergence rate of the CG method used to solve (3.12). We then introduced the concept of preconditioning which is one method of alleviating the difficulties associated with an ill-conditioned problem. Specifically, we showed that if an appropriate preconditioner is chosen, an equivalent system can be formed

which has a smaller condition number. The type of norm used to define the condition number was shown to be arbitary due to the equivalence of the real vector norms considered and so for the rest of the thesis we focus on the spectral condition number. To calculate the spectral condition number requires the eigenvalues of the Hessian and so in the final section of this chapter we considered important methods and results which enables us to analyse the eigenvalues and hence the conditioning of the Var Hessian in the later chapters. In the next section we show how the concept of conditioning applies to Var and how preconditioning is currently applied in operational centres.

Chapter 4

Conditioning and Preconditioning of Var

In Chapter 2 we showed that the Var method applied to NWP approximates the initial state of the atmosphere by minimising a nonlinear quadratic cost function. The initial state, called the analysis, is then used to predict the future weather states. Due to the inherently chaotic nature of the atmosphere the accuracy of the analysis is vital for creating an accurate forecast of the weather. In practice, most NWP centres solve a series of linearised cost functions approximating the full nonlinear cost function using an incremental form of Var. In Chapter 3 we showed that the condition number of the Hessian of a quadratic function (3.7) is an important quantity for estimating the sensitivity of the solution to perturbations in the system. We also showed that a large condition number can mean that the iterative methods used to minimise (3.7) can converge very slowly to the solution. In this chapter we present existing literature on the conditioning and preconditioning of variational data assimilation.

We begin by showing that early implementations of the full and incremental Var minimisation schemes as presented in Sections 2.2 and 2.3 were computationally expensive. This indicates that Var is an ill-conditioned problem.

We present existing literature which provide possible reasons for this ill-conditioning. We then consider early attempts at preconditioning the Var Hessian and show the benefits that this has on the convergence rate of the iterative methods used to solve Var. Finally, we consider the incremental version of Var which includes the control variable transform (CVT) as introduced in Section 2.4. We show how the CVT effectively preconditions Var and present research from the literature that shows the improvements in the convergence rate of the Var minimisation scheme.

The aim of this chapter is to highlight the importance of detailed knowledge of the conditioning of the Var problem. This chapter motivates our research in the subsequent chapters in which we derive theoretical bounds on the condition number of the Var Hessian.

4.1   Computational Cost of Var

As shown by equation (3.33) in Section 3.3 the condition number of the Hessian of the quadratic cost function (3.7) can be a useful indicator of the performance of the iterative methods used to solve the system. Conversely, the performance of the iterative method can be an indicator of the conditioning of the system. In this section we consider the computational cost and performance of the iterative solvers used in the 4DVar minimisation.

In most early implementations of 4DVar, the background term was excluded from the cost function (2.1) [50], [59]. Experiments were performed using a variety of models including the shallow water model, lower dimensional primitive equation models and vorticity equation models [50], [59], [51], [49]. The computational expense of these experiments tended to be large; for instance, using a shallow water model, with a state vector of size 2793, required about 10-15 iterations for convergence yet this was described as 'absolutely prohibitive in todays practice of weather prediction'. In this experiment,

described in [50], the system was overdetermined having a larger number of observations than state variables. Many other early 4DVar implementations tended to use around 30 iterations, for instance see [59] and references within [51]. In 1994 Courtier et al. reported that the cost of a 4DVar data assimilation with this number of iterations was computationally prohibitive. They also stated that to be able to implement 4DVar operationally required a 'significantly faster computer or a substantial algorithmic improvement or both' [51].

One step toward an algorithmic improvement was the introduction of the incremental version of Var that we described in Section 2.3. Incremental Var splits the 4DVar minimisation into a lower resolution inner loop and a full resolution outer-loop, reducing the overall cost of the algorithm [51]. Incremental Var also suffers from convergence issues as shown by Lorenc in [42]. In a single variable experiment on a 2D domain the Var algorithm failed to converge even after several thousand iterations.

We showed in Chapter 3 that one potential reason for slow convergence may be the ill-conditioning of the Hessian of the cost function (2.7). A variety of explanations have been given in the literature for the ill-conditioning of the minimisation problem in incremental 4DVar. Lorenc suggests that the background covariance matrix, $\mathbf{B}$, tends to be ill-conditioned since the background errors contain rough and smooth modes. These modes correspond to small and large eigenvalues of the background error covariance matrix and if the range of the eigenvalues is large enough $\mathbf{B}$ will have a large spectral condition number. If the background matrix dominates the conditioning of the Hessian, then this system will be ill-conditioned [42]. Another explanation is provided by Thacker in [58]. The inverse of the Hessian is the analysis error covariance matrix and hence explains the expected errors in the analysis. The eigenvalues of the analysis error covariance matrix describe the error variances in the directions of the corresponding eigenvectors. The reciprocals of these eigenvalues are the eigenvalues of the Hessian. Thus a poorly

determined state in the analysis corresponds to a small Hessian eigenvalue and a potentially ill-conditioned system.

As described in Chapter 3, one way of improving the convergence rates is by preconditioning the system. In the next section we briefly describe early attempts which were made at preconditioning Var.

4.2 Preconditioning Var

4.2.1 Early Preconditioning

As shown in the previous section, the cost of minimising the variational cost functions (2.1) and (2.7) are substantial and methods are required for reducing this expense. One such method is preconditioning. In this section we consider some of the early attempts to precondition the 4DVar minimisation scheme.

In Section 3.2.2 we showed that preconditioning can be considered as either a matrix multiplication or as a change of variables. Most early attempts considered using diagonal matrices as preconditioners, i.e. scaling the variables. Thepaut and Moll considered a 1DVar scheme to invert radiance data [60] with a state vector of size $N = 31$. In this paper, convergence results were given for different preconditioning matrices $\mathbf{P}$ equal either to the full Hessian, the diagonal of the Hessian, a diagonal matrix using climatological variances or the identity (i.e. no preconditioning). The methods which performed the best used the full or diagonal of the Hessian as preconditioner with less than 5 and 7 iterations required respectively. A disadvantage to these methods is the computational cost associated with forming the Hessian. The climatological preconditioner also performed reasonably well with less then 10 iterations for convergence. An advantage of this preconditioner is that it can be stored and computed separately from the minimisation. All

the preconditioners produced faster convergence than the system without preconditioning which took 15 iterations to converge.

Courtier et al. also used a diagonal preconditioner which was formed from the diagonal entries of an approximation to the Hessian [51]. Applying this preconditioning to 4DVar using the primitive model equations increased the convergence rate from 30 to 24 iterations compared to the unpreconditioned system. However, similar to the experiments in [60], the computational cost of approximating the Hessian was noted as a potential pitfall for the technique.

Zupanski attempted to improve the speed of convergence of the minimisation by rescaling the variables without computing the Hessian [71]. The rescaling was chosen so that the reduction in the cost function evaluated at consecutive iterations of the minimisation was as close as possible to the decrease found by preconditioning with the diagonal of the Hessian. In 4DVar experiments, without background terms, this technique produced quicker convergence of the minimisation compared to a control case without preconditioning. These improvements in convergence gave corresponding reductions in the condition number. Zupanski found similar improvements using this preconditioner in a full 4DVar cost function compared to simply preconditioning with the diagonal of the background matrix [72].

In this section we have seen that preconditioning techniques can have a positive impact on the convergence rates and condition number of the minimisation in 4DVar. In the next section we consider how the control variable transform alters the conditioning of the Var Hessian.

4.2.2 The CVT as a Preconditioner

In Section 3.2.2 we showed that a quadratic cost function could be preconditioned by transforming to new variables. In this section we present previous

literature which shows the effect of the CVT on the conditioning of the incremental Var problem (2.9).

The Var cost function with the CVT (2.16) has Hessian

$$\hat{\mathbf{S}} = \mathbf{I} + \mathbf{B}^{T/2}\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{B}^{1/2} = \mathbf{I} + \mathbf{B}^{T/2}\sum_{i=0}^{n}\mathbf{M}_1^T \ldots \mathbf{M}_i^T\mathbf{H}_i^T\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{M}_i\mathbf{M}_1\mathbf{B}^{1/2}.$$

$$(4.1)$$

Hence, the control variable transform is equivalent to symmetrically preconditioning the Hessian of the linear least squares cost function (2.9) with the square root of the background covariance matrix $\mathbf{B}$. One of the main advantages of the CVT is its potential effect on the conditioning of the Var scheme. Since there are generally fewer observations than elements in the state vector $\mathbf{x}_0$ the matrix $\mathbf{B}^{T/2}\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{B}^{1/2}$ is not full rank and thus has smallest eigenvalue equal to zero. Thus the smallest eigenvalue of (4.1) is equal to one. Since $\hat{\mathbf{S}}$ is symmetric and positive definite, by Theorem 3.2.4 the spectral condition number of the matrix (4.1) is

$$\kappa(\hat{\mathbf{S}}) = \frac{\lambda_{\max}(\hat{\mathbf{S}})}{\lambda_{\min}(\hat{\mathbf{S}})} = \lambda_{\max}(\hat{\mathbf{S}}) = 1 + \lambda_{\max}(\mathbf{B}^{T/2}\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{B}^{1/2}). \qquad (4.2)$$

The absence of small eigenvalues (< 1) implies that the new Hessian (4.1) has a potentially lower condition number than the original system (2.12). A reduction in the condition number of (4.1) compared to (2.12) is supported by experimental evidence in [42] and [18], which show substantial accelerations in the convergence rate for the preconditioned system compared to the unpreconditioned system.

How the CVT affects the condition number of the Var cost function has only been partially examined. In [6] the following expression was derived for the condition number of the Hessian of the preconditioned system for a simple 2 grid point domain with n observations at each grid point

$$\kappa(\hat{\mathbf{S}}) = 2n\frac{\sigma_b^2}{\sigma_o^2} + 1, \qquad (4.3)$$

where σ_b^2 and σ_o^2 are the background and observation error variances respectively. The expression (4.3) holds in the case of dense observations. Thus, for dense observations, an ill-conditioned system could be the result of accurate observations (small σ_o^2), an inaccurate background state (large σ_b^2) or excessive observations (large n). In the ECMWF Var system it was found that the surface pressure observations dominated the conditioning of the problem. In an experiment where the observation error variance on the pressure observations was doubled, the condition number of the Hessian decreased from 5474 to 1502 in rough agreement with the estimate (4.3) [63].

Further preconditioning strategies can be applied to the preconditioned Hessian (4.1). The most prominent is Hessian eigenvector preconditioning [64]. These methods are an extra level of preconditioning on top of the CVT. They use information from a previous outer-loop in order to eliminate larger eigenvalues from a later outer-loop. We do not consider this preconditioning in this thesis. We now finish by summarising the contents of this chapter.

4.3   Summary

We began this chapter by presenting research from existing literature that highlight the computational difficulties associated with the minimisation of both the standard (2.1) and incremental forms (2.7) of Var. Early Var schemes showed that an operational minimisation was impractical unless a sufficient reduction in the iteration count could be produced. The ill-conditioning of the Hessian of the Var cost function is one possible reason for the poor convergence. Earlier research suggested that an ill-conditioned Hessian could be the result of an ill-conditioned background error covariance matrix. One method for avoiding ill-conditioning is to find a way of preconditioning the problem.

We gave a brief review of early preconditioning techniques which focused

on rescaling the control variables. These examples showed the benefits of preconditioning by producing modest reductions in the computational cost. Finally, we considered the effect of the CVT on the conditioning of the system. We presented research from other literature that suggested that this preconditioned system was much better conditioned than the cost function without preconditioning. We also presented evidence from the literature, based on a 2 grid point system, that suggested that the conditioning of the preconditioned system with dense observations was largely affected by the number and accuracy of observations. The effect of the observation accuracy was confirmed in an operational experiment at the ECMWF.

Except for these simplified examples there is very little theory that explains what drives the conditioning of both the unpreconditioned and preconditioned Var schemes. The aim of this thesis is to provide solid theory describing the condition number of the Hessians of the incremental Var cost function, both with and without the CVT. In this thesis we confirm some of the experimental results on the conditioning of Var which was presented in this chapter. We confirm that the conditioning of the background covariance matrix does have a large affect on the conditioning of Var. We also show both theoretically and experimentally how the conditioning of Var is improved by including a CVT. Additionally, we identify what are the important factors which can affect the condition number of the preconditioned Var Hessian. In the next chapter we begin our exploration by considering the condition number of covariance matrices commonly used to model the auto-covariances of the background errors.

Chapter 5

Conditioning of Background Covariance Matrices

In major NWP centres it is common for the size of the state vector to be greater then the number of available observations. As computer processing power increases these centres have generally chosen to increase their model resolution. Hence, for the foreseeable future, despite the increases in observation data from satellites, Var will continue to solve an under-determined system. As mentioned in Chapter 2 this lack of information is accounted for in 4DVar by including an a priori guess of the initial state called the background. In Section 4.1 it was suggested that the background term may play an important role in the conditioning of the Var minimisation. The purpose of this chapter is to consider the condition number of some common background error covariance matrices. We use these results in later chapters to show how the background covariance matrix can influence the condition number of the Var Hessian.

We begin by introducing some basic properties of covariance functions and describe some common functions for modelling the background error auto-correlations on both the real line and the circle. The correlation functions are used to construct the background error correlation matrices. Finally we

analyse the condition number of the background error covariance matrices associated with the correlation functions introduced in this chapter before summarising the chapter.

5.1   Background Covariance Matrices

In this section we give a brief overview of background error covariance functions, some basic properties and how they are used to construct the background error covariance matrix, $\mathbf{B}$, used in Var. In Var, the first guess background state, $\mathbf{x}_b$, is usually found from a short forecast [32, Sec 5.1], [9]. The corresponding background error covariance matrix has an important influence on the final analysis [10]. For example, the background covariance matrix spreads information from the observations when forming the analysis $\mathbf{x}_a$ [10]. Therefore, to obtain an accurate forecast it is essential that the background errors are specified correctly. In this chapter we refer to *parameters* within the state vector, $\mathbf{x}_0$. In the context of meteorology and in this thesis a parameter means a physical quantity of the atmosphere such as pressure or wind velocity.

The covariance is a measure of how two variables change together and can be described by a covariance function. More formally, a function $f(x,y)$ is the covariance function of a random field X on $\mathbb{R}^n$ if

$$f(x,y) = <X(x)-<X(x)>, X(y)-<X(y)>> \qquad (5.1)$$

for $x,y \in \mathbb{R}^n$ and where $<>$ represents the expected value of a random field [17]. A direct consequence of the definition (5.1) is that the function is symmetric with $f(x,y) = f(y,x)$. Random variables that have a covariance equal to zero are said to be *uncorrelated*. The value of $f(x,x)$ is called the *variance* of the random variable at x and $(f(x,x))^{1/2}$ is called the *standard deviation*. By normalising the covariance function with the standard deviations

we obtain a *correlation function*

$$\rho(x, y) = \frac{f(x, y)}{(f(x, x)f(y, y))^{1/2}}. \tag{5.2}$$

Notice that a correlation function is a particular type of covariance function which has unit variances. In this chapter we assume that the variance of the variables are non-zero so that ρ is well defined. A correlation function is a dimensionless function whereas a covariance function has the units of $X(x)$ times $X(y)$. In this chapter and in Chapters 6 and 7 we primarily consider the errors in a single parameter (e.g. pressure) on a one-dimensional domain $(n = 1)$ so $X(x)$ represents the random field of errors at a position $x \in \mathbb{R}$. With no correlations between different parameters f is called an *auto-covariance* function and ρ a *auto-correlation* function. We also assume that the errors of our parameter are homogeneous. That is, the covariance, or correlation, function will only depend on the distance between errors and not on the position. In this case we may write the correlation function

$$\rho(x, y) = \hat{\rho}(|x - y|). \tag{5.3}$$

The following theorem is useful for showing whether a function is a correlation function.

Theorem 5.1.1 *[17] Let $\hat{\rho}$ be an even continuous function on $\mathbb{R}$ with $\hat{\rho}(0) = 1$ and*

$$\int_{\mathbb{R}} |\hat{\rho}(x)| dx < \infty, \tag{5.4}$$

then $\rho(x, y) = \hat{\rho}(|x - y|)$ is a homogeneous correlation function on $\mathbb{R}$ if and only if the Fourier transform of $\hat{\rho}$ is everywhere non-negative.

Proof See [17], [55, Sec. IX.2].

For a choice of points $s_1, s_2, \ldots, s_N \in \mathbb{R}$ an auto-correlation function $\rho(x, y)$ defines a positive-definite symmetric auto-correlation matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$ with

entries [17]

$$\mathbf{C}_{i,j} = \rho(s_i, s_j), \quad (5.5)$$

for $i, j = 1, \ldots, N$. We can use the error auto-correlation matrix (5.5) to define the background error auto-covariance matrix

$$\mathbf{B} = \boldsymbol{\Sigma}\mathbf{C}\boldsymbol{\Sigma}, \quad (5.6)$$

where $\mathbf{C}$ represents the background error correlation matrix and the diagonal matrix $\boldsymbol{\Sigma}$ contains the positive background error standard deviations [32, Sec 5.4]. From (5.6) the background covariance matrix is clearly symmetric and positive definite. In particular, since $\mathbf{B}$ is positive definite it automatically implies that all the eigenvalues of (5.6) are positive. We now describe some error auto-correlations functions which are commonly used for constructing the background auto-covariance matrices on the real line and the circle.

5.2 Auto-Correlation Functions and Matrices

In this thesis we consider a single, periodic, parameter on a one-dimensional domain and therefore require a error auto-covariance matrix to define the errors in the background. In this section we introduce some common auto-correlation functions and an auto-correlation matrix used to construct the background auto-covariance matrix. As shown in Section 5.1 the auto-correlation matrix can be combined with the error standard deviations to produce the background error covariance matrix (5.6).

We assume that errors are homogeneous, that is the correlations depend only on the distance between the errors and not on the position of the errors [10]. Since we use a periodic parameter we are concerned with modelling correlations on the circle. However, many well-defined correlation models

already exist on the real line and so we begin first by considering correlation functions on $\mathbb{R}$. We use these correlation functions to create valid correlation functions on the circle.

The *Gaussian* correlation function for points on the real line separated by a distance $|r|$ is defined by [14], [32],

$$\rho_G(r) = \exp\left(-\frac{|r|^2}{2L^2}\right), \tag{5.7}$$

where $L > 0$ is the correlation lengthscale and $r \in \mathbb{R}$. The Gaussian correlation function is commonly used in many applications including meteorology and remote sensing [35].

Another common model is the second-order auto-regressive correlation (SOAR) function [14], defined by

$$\rho_S(r) = \left(1 + \frac{|r|}{L}\right)\exp\left(-\frac{|r|}{L}\right). \tag{5.8}$$

The SOAR function is used in the Met Office system to model the horizontal error correlations and can be estimated using a spectral filter [41], [43], [27], [2].

Suppose we identify the values of our parameter at two points $-D$ and D. We have a periodic parameter on the real line and so the domain can be viewed as a circle. Although both the SOAR and Gaussian functions define correlation functions on the real line, they are not necessarily valid correlation functions on the circle since the Fourier transforms of (5.7) and (5.8) may no longer be positive and by Theorem 5.1.1 they are no longer correlation functions [68], [22], [66]. One alternative is to compactly support the correlation model (for instance see [13] and [2]); however, the approach we take is to transform to a valid correlation model on the circle by replacing the great circle distance r in (5.7) and (5.8) for the chordal distance

$$d = 2a\sin(\theta/2) \tag{5.9}$$

where θ is the angle between two points on the circle and a is the radius [66], [65]. This substitution guarantees that any valid correlation model on the real line is now a valid correlation model on the circle [69, Sec. 22.5]. In particular the Gaussian and SOAR functions (5.7) and (5.8) are now valid correlation functions on the circle [66], [17].

In this, and later, chapters we assume that the parameter is positioned at N equally spaced grid points on the circle $s_1, s_2, \ldots, s_N$. Hence, the angle $\Delta\theta$ between adjacent grid points on the circle is constant. We define a positive definite correlation matrix on the circle by sampling the correlation functions (5.7) and (5.8) at the points $s_1, s_2, \ldots, s_N$ [17]. The Gaussian correlation matrix C_G has components

$$(C_G)_{i,j} = \exp\left(-\frac{|2a\sin(\theta_{i,j}/2)|^2}{2L^2}\right), \tag{5.10}$$

where $i, j = 1, \ldots, N$ and $\theta_{i,j}$ is the angle between the points s_i and s_j on the circle. Similarly the SOAR error correlation matrix C_S has elements given by

$$(C_S)_{i,j} = \left(1 + \frac{|2a\sin(\theta_{i,j}/2)|}{L}\right)\exp\left(-\frac{|2a\sin(\theta_{i,j}/2)|}{L}\right) \tag{5.11}$$

where $i, j = 1, \ldots, N$ and $\theta_{i,j}$ is the angle between the points s_i and s_j on the circle.

The third periodic correlation matrix we consider we refer to as the *Laplacian* correlation matrix, $\mathbf{C}_{Lap}$. This is an explicit equation for the *inverse* of a correlation matrix of a periodic parameter [30, Chap. 2], [31], [37, Chap. 5]

$$\mathbf{C}_{Lap}^{-1} = \gamma^{-1}\left(\mathbf{I} + \frac{L^4}{2\Delta x^4}(\mathbf{L}_{xx})^2\right). \tag{5.12}$$

Here L is again the lengthscale, Δx is the great circle distance between grid points and $\gamma > 0$ is a constant that ensures that the maximum element of $\mathbf{C}$ is equal to one. The matrix $\mathbf{L}_{xx}$ is the second order derivative matrix given

by

$$\mathbf{L}_{xx} = \begin{pmatrix} -2 & 1 & 0 & 0 & \dots & 0 & 1 \\ 1 & -2 & 1 & 0 & \dots & 0 & 0 \\ & \ddots & \ddots & & & & \\ \vdots & & \ddots & \ddots & & & 0 \\ 0 & & & \ddots & \ddots & & 1 \\ 1 & 0 & \dots & & & 1 & -2 \end{pmatrix}. \tag{5.13}$$

Theorem 5.2.1 *For $N > 5$ the matrix (5.12) is positive definite.*

Proof To prove (5.12) is positive definite it is enough to show that all the eigenvalues are positive. Since L_{xx} is circulant then the matrix $\mathbf{C}_{Lap}^{-1}$ is also circulant. Assuming $N > 5$, the inverse of the Laplacian correlation model is therefore described by its top row

$$\hat{\mathbf{c}} = \gamma^{-1}[1 + 6q, -4q, q, 0, \dots, 0, q, -4q]. \tag{5.14}$$

where $q = \frac{L^4}{2\Delta x^4}$. As shown in Section 3.4.4 the eigenvalues of the circulant matrix, $\mathbf{C}_{Lap}^{-1}$, are simply the Fourier transform of $\hat{\mathbf{c}}$ (5.14). The eigenvalues of $\mathbf{C}_{Lap}^{-1}$ are

$$\nu_m = \gamma^{-1}(1 + 6q - 4q(e^{-2\pi im/N} + e^{-2\pi im(N-1)/N}) + q(e^{-4\pi im/N} + e^{-2\pi im(N-2)/N})), \tag{5.15}$$

where $m = 0, 1, \dots, N - 1$. Using the identity

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}, \tag{5.16}$$

and the fact that $e^{2\pi im} = 1$ for $m \in \mathbb{Z}$, equation (5.15) simplifies to

$$\nu_m = \gamma^{-1}\left(1 + 6q - 8q\cos\left(\frac{2\pi m}{N}\right) + 2q\cos\left(\frac{4\pi m}{N}\right)\right). \tag{5.17}$$

Treating equation (5.17) as a continuous function with respect to m we can

differentiate to give

$$\frac{d\nu_m}{dm} = \gamma^{-1} \left(\frac{16q\pi}{N}\right) \sin\left(\frac{2\pi m}{N}\right) \left(1 - \cos\left(\frac{2\pi m}{N}\right)\right). \tag{5.18}$$

For $m \in [0, N)$ the only extrema occur at $m = 0$ and $m = N/2$. Since

$$\nu_m \leq \gamma^{-1}(1 + 6q + 8q + 2q) = \gamma^{-1}(1 + 16q) = \nu_{N/2}, \tag{5.19}$$

$\nu_{N/2}$ is a global maximum of the function ν_m. For $m \in (0, N/2)$, $\sin(2\pi m/N) > 0$ and $-1 < \cos(x) < 1$ therefore by (5.18)

$$\frac{d\nu_m}{dm} > 0, \tag{5.20}$$

and ν_m is therefore strictly increasing for $m \in (0, N/2)$. Since ν_m is symmetric about the point $m = N/2$ then the minimum of ν_m for $m \in [0, N)$ must occur at $m = 0$,

$$\nu_0 = \gamma^{-1}(1 + 6q - 8q + 2q) = \gamma^{-1} > 0. \tag{5.21}$$

The eigenvalues occur at the discrete points $m = 0, 1, \ldots, N - 1$. Since the smallest eigenvalue ν_0 is positive all eigenvalues must be positive. Hence, we have shown $\mathbf{C}_{Lap}^{-1}$ is positive definite.

If we assume that the background error variances are equal to σ_b^2 at every grid point then by equation (5.6) the background covariance matrix is

$$\mathbf{B} = \sigma_b^2 \mathbf{C}, \tag{5.22}$$

where $\mathbf{C}$ is one of auto-covariance matrices (5.10) (5.11) or (5.12). The background error covariance matrix (5.22), $\mathbf{B}$, is circulant. In particular, it means we can explicitly find the eigenvalues, and therefore, the spectral condition number of the background error covariance matrix. When the system is not periodic the background covariance matrix will be Toeplitz (i.e. each diagonal of the matrix is constant). The circulant matrices can be used to approximate the Toeplitz matrices used to model the non-periodic

forms of the Gaussian and SOAR correlation matrices (see [25] for more details).
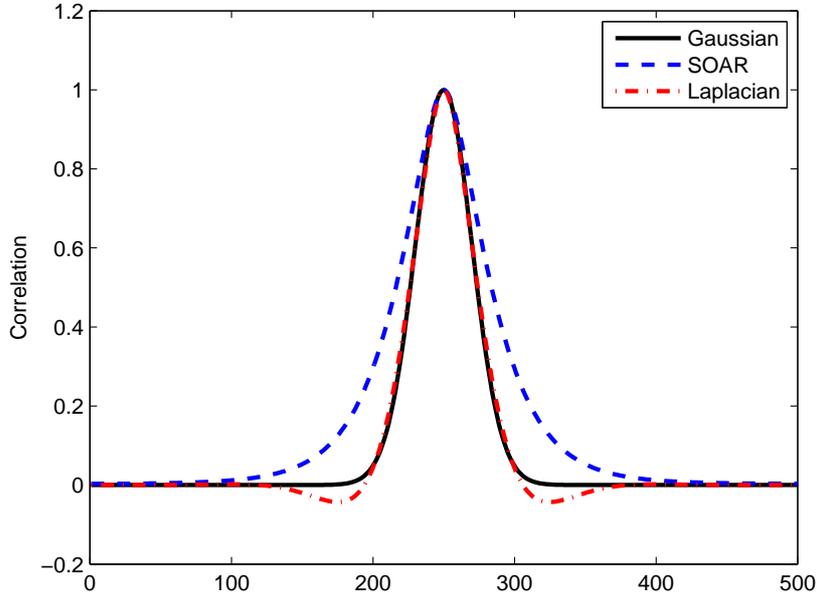
Figure 5.1: Plot showing the Gaussian correlation function (Solid line), the SOAR correlation function (dash line) and the Laplacian correlation function (dot-dash line) $L = 2$.

We now briefly compare the correlation matrices. Figure 5.1 shows the 250th row of the correlation matrices (5.10), (5.11) and (5.12) for $N = 500$ grid points on the circle with $\Delta x = 0.1$ (great circle distance between grid points) for lengthscale $L = 2$. The plot shows the close approximation of the positive components of the Laplacian correlation matrix and the Gaussian correlation matrix. The Laplacian has often been used instead of the Gaussian when the inverse is required since the inverse is explicitly known [30, Chap. 2], [37, Chap. 5]. Unlike the Gaussian and SOAR correlation matrices the Laplacian exhibits negative correlations. For the same lengthscale L, the SOAR function appears to have longer tails then the Gaussian function.

In the next section we consider the condition number of the **B** matrices of the form (5.22) using the Gaussian, SOAR and Laplacian correlation matrices defined in this section.

5.3 Conditioning of the Background Error Covariance Matrix

Since the background covariance matrix is considered to play a major role in the conditioning of the Hessian (2.12), we begin by examining the condition number of the $\mathbf{B} \in \mathbb{R}^{N \times N}$ matrix. Assuming that the background error states all have the same variance, $\sigma_b^2 > 0$, then we can write our covariance matrix in the simpler form

$$\mathbf{B} = \sigma_b^2 \mathbf{C}, \tag{5.23}$$

where $\mathbf{C}$ is a symmetric, positive definite, circulant matrix. Since $\mathbf{B}$ is symmetric and positive-definite, by Theorem 3.2.4 the condition number is

$$\kappa(\mathbf{B}) = \frac{\lambda_{\max}(\sigma_b^2 \mathbf{C})}{\lambda_{\min}(\sigma_b^2 \mathbf{C})} = \frac{\sigma_b^2 \lambda_{\max}(\mathbf{C})}{\sigma_b^2 \lambda_{\min}(\mathbf{C})} = \kappa(\mathbf{C}), \tag{5.24}$$

and hence is independent of the background error variance and is equal to the condition number of the correlation matrix $\mathbf{C}$. Since the correlation matrices we consider are circulant, the eigenvalues can be explicitly found via the discrete Fourier transform of the top row of $\mathbf{C}$. We denote this top row $[c_0, c_1, \ldots, c_{N-1}]$. The symmetry of $\mathbf{C}$ implies that $c_i = c_{N-i}$ for $i = 1, \ldots, N-1$, and hence combining with (3.49) we have

$$\nu_m = c_0 + 2 \sum_{k=0}^{N/2-1} c_k \cos\left(\frac{2\pi mk}{N}\right) + c_{N/2} \cos\left(m\pi\right), \tag{5.25}$$

for N even and

$$\nu_m = c_0 + 2 \sum_{k=0}^{(N-1)/2} c_k \cos\left(\frac{2\pi mk}{N}\right), \tag{5.26}$$

for N odd. An upper bound can be placed on the eigenvalues of $\mathbf{C}$ using Theorem 3.4.2 with $p = \infty$,

$$|\nu_m| \leq ||\mathbf{C}||_\infty = \sum_{k=0}^{N-1} |c_k|, \tag{5.27}$$

57

for $m = 0, 1, \ldots N - 1$. Since $c_k \in [-1, 1]$ then the eigenvalues are bounded above by N, the number of grid points (although this may be a large over-estimate of the largest eigenvalue). We now consider the conditioning of the three correlation models we introduced in the previous section, the Gaussian (5.10), the SOAR (5.11) and the Laplacian (5.12).

5.3.1 Gaussian Correlation Matrix
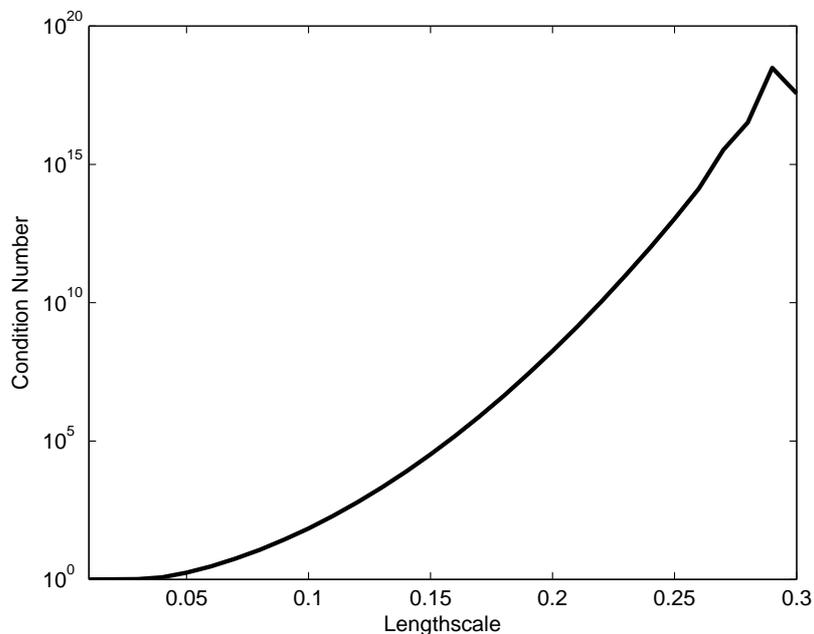
Figure 5.2: Condition number of the periodic Gaussian correlation matrix as a function of lengthscale.

In this section we consider the condition number of the Gaussian covariance matrix (5.10). Let $s_1, s_2, \ldots s_N$ be the positions on the circle with uniform spacing between grid points of great circle length Δx. Since all the coefficients of the Gaussian correlation matrix are positive, from (5.25) and (5.26) we can see that the largest eigenvalue of $\mathbf{C}$ is equal to

$$\nu_0 = \sum_{k=0}^{N-1} c_k = ||\mathbf{C}||_\infty, \qquad (5.28)$$

Figure 5.3: Largest (left) and smallest (right) eigenvalue of the periodic Gaussian correlation matrix as a function of lengthscale.

where $c_k = \exp\left(-\frac{|2a\sin(\theta_{0,k}/2)|^2}{2L^2}\right)$ and $\theta_{0,k}$ is the angle between points s_1 and s_{k+1}.

In order to examine the condition number we design a simple experiment. We define a grid of $N = 500$ grid points with $\Delta x = 0.1$. Figure 5.2 shows the condition number of the Gaussian covariance matrix as a function of the correlation lengthscale L on a logarithmic plot. The figure shows that the correlation matrix is extremely sensitive to lengthscale and when $L = 2\Delta x$ the condition number is of the order 10^7 and the matrix is very ill-conditioned. This confirms previous studies of the Gaussian correlation model in the non-periodic case [34], [35]. The ill-conditioning of the Gaussian auto-correlation model has been referred to as the *Gaussian anomaly* [34]. One possible explanation for this ill-conditioning is given by Kostinski in [35]. Consider a parametric family of functions of the form

$$\exp(-\nu^\alpha), \tag{5.29}$$

These are valid auto-correlation functions if $0 < \alpha \leq 2$ and give therefore non-negative Fourier coefficients. Since the Gaussian case $\alpha = 2$ lies on the boundary of the parameters which create a valid correlation model this has been suggested as the cause of the Gaussian anomaly [35]. By equation (5.28) the largest eigenvalue remains reasonably small as the lengthscale increases

thus the large increase in the condition number must be due to the rapid decrease in the smallest eigenvalue. This is confirmed by Figure 5.3. From the Figures 5.2 and 5.3 we also observe a sharp kink in the condition number and smallest eigenvalue plots when the lengthscale is approximately 0.28. This most likely occurs due to the computation of the smallest eigenvalue being around 10^{-17} which is beyond the machine precision of the Matlab program employed (Given as $2.2204e - 016$ for our version).

5.3.2 SOAR Correlation Matrix
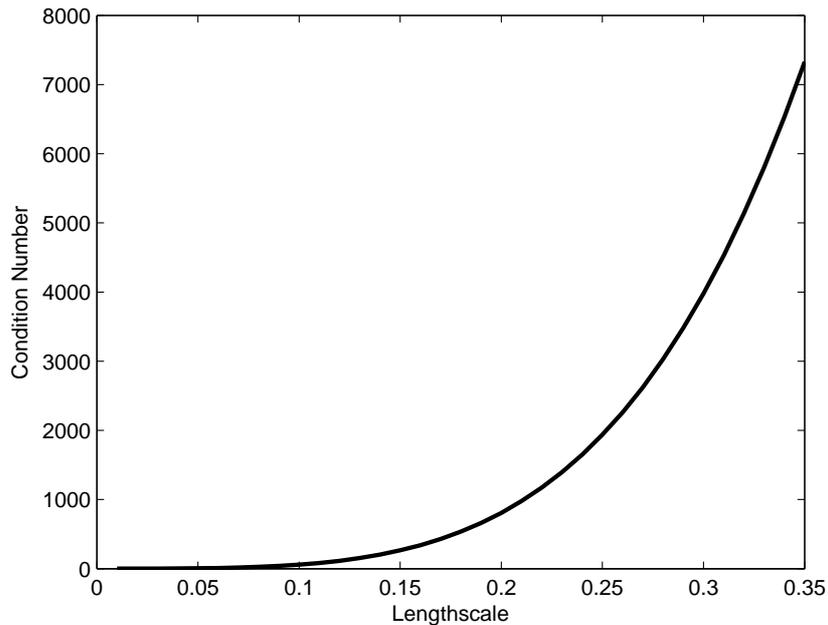
Figure 5.4: Condition number of the periodic SOAR correlation matrix as a function of lengthscale.

We now consider the SOAR correlation matrix (5.11). Like the Gaussian correlation matrix, each row of the SOAR correlation matrix is bell-shaped and contains only positive correlations as shown in Figure 5.1. Hence, assuming the grid is at equally spaced points $s_1, \ldots s_N$ on the circle, our maximum
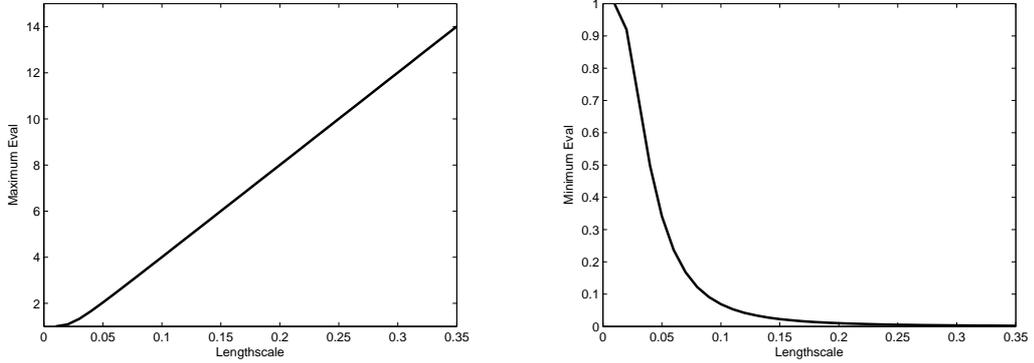
Figure 5.5: Largest (left) and smallest (right) eigenvalue of the periodic SOAR correlation matrix as a function of lengthscale.

eigenvalue is simply the sum of the top row as given by (5.28) but with

$$c_k = \left(1 + \frac{|2a\sin(\theta_{0,k}/2)|}{L}\right)\exp\left(-\frac{|2a\sin(\theta_{0,k}/2)|}{L}\right), \qquad (5.30)$$

where $\theta_{0,k}$ is the angle between point s_1 and s_{k+1}. We therefore expect the largest eigenvalue of (5.11) to increase slowly as the lengthscale increases.

The SOAR correlation function is commonly used to model the horizontal correlations in the Met Office operational system [42] [43]. Figure 5.4 shows a plot of the condition number of the SOAR background correlation matrix as a function of lengthscale with the same experimental design as in the Gaussian experiment shown in Figure 5.2 ($N = 500$, $\Delta x = 0.1$). As with the Gaussian correlation matrix the SOAR matrix is also sensitive to changes in the lengthscale. However, it is clearly not as sensitive and a comparison shows at lengthscale, $L = 0.2$, the condition number of the SOAR matrix is of order 10^3 compared to order 10^7 for the conditioning of the Gaussian matrix. Again the large reduction in the smallest eigenvalue as the lengthscale increases appears to be the main cause of the ill-conditioning, as shown in Figure 5.5. However in this case note that the smallest eigenvalue of the SOAR matrix has not reached machine precision here and at $L = 0.3$ is around 10^{-4}.

5.3.3 Laplacian Correlation Matrix
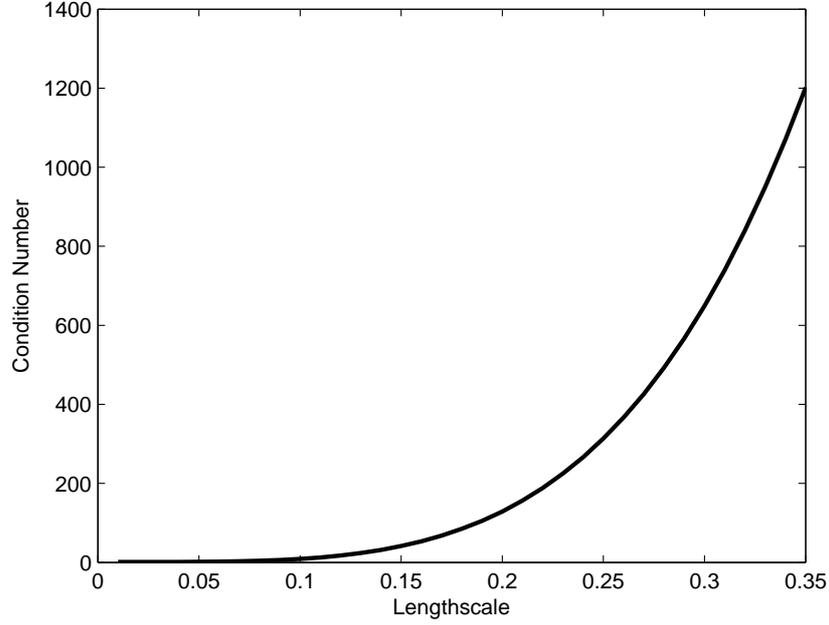
Figure 5.6: Condition number of the Laplacian correlation matrix as a function of lengthscale.

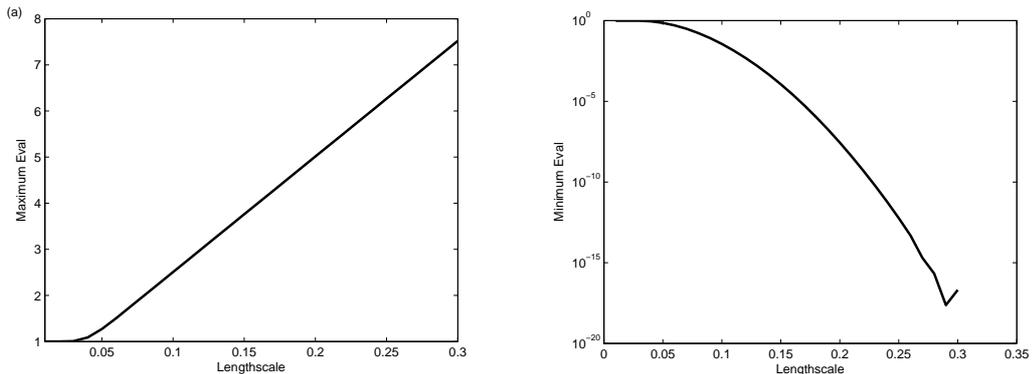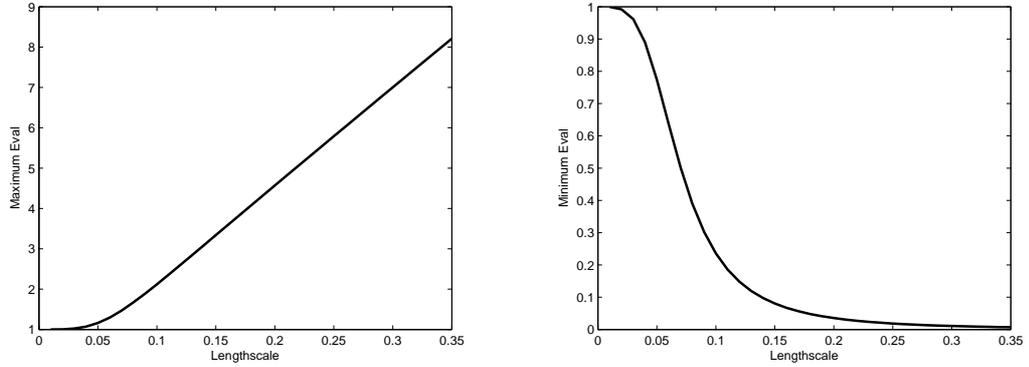Figure 5.7: Largest (left) and smallest (right) eigenvalue of the Laplacian correlation matrix as a function of lengthscale.

The final matrix we consider is the Laplacian correlation matrix $\mathbf{C}_{Lap}$ (5.12).

Theorem 5.3.1 *The condition number of* $\mathbf{C}_{Lap} \in \mathbb{R}^{N \times N}$ *is*

$$\kappa(\mathbf{C}_{Lap}) = \left(1 + 16\frac{L^4}{2\Delta x^4}\right), \tag{5.31}$$

if N is even or

$$\kappa(\mathbf{C}_{Lap}) = \left(1 + \frac{L^4}{2\Delta x^4}\left(6 - 8\cos\left(\frac{\pi(N-1)}{N}\right) + 2\cos\left(\frac{2\pi(N-1)}{N}\right)\right)\right),$$

$$(5.32)$$

if N is odd.

Proof In the proof to Theorem 5.2.1 we see that the eigenvalues of the Laplacian were given by

$$\nu_m = \gamma^{-1}\left(1 + 6q - 8q\cos\left(\frac{2\pi m}{N}\right) + 2q\cos\left(\frac{4\pi m}{N}\right)\right),$$

$$(5.33)$$

for $m = 0, 1, \ldots, N-1$, where $q = \frac{L^4}{2\Delta x^4}$. Considered as a continuous function with respect to $m \in [0, N)$ it was shown in the proof of Theorem 5.2.1 that the minimum value of (5.33) is at $m = 0$ giving $\nu_0 = \gamma^{-1}$ and the maximum is at $m = N/2$ giving

$$\nu_{N/2} = \gamma^{-1}\left(1 + 16\frac{L^4}{2\Delta x^4}\right).$$

$$(5.34)$$

Thus, if N is even then $\nu_{N/2}$ is an eigenvalue of (5.33) giving the maximum possible value of ν_m for $m \in [0, N)$. Therefore, since $\mathbf{C}_{Lap}$ is symmetric and positive definite, then by Theorem 3.2.4 we have

$$\kappa(\mathbf{C}_{Lap}) = \frac{\nu_{N/2}}{\nu_0} = \frac{\gamma^{-1}\left(1 + 16\frac{L^4}{2\Delta x^4}\right)}{\gamma^{-1}} = \left(1 + 16\frac{L^4}{2\Delta x^4}\right).$$

$$(5.35)$$

As shown in the proof of Theorem 5.2.1 the function (5.33) is monotonically increasing between $m = 0$ and $m = N/2$. Hence if N is odd the maximum eigenvalue must occur at $m = (N-1)/2$ giving

$$\nu_{N-1/2} = \gamma^{-1}\left(1 + \frac{L^4}{2\Delta x^4}\left(6 - 8\cos\left(\frac{\pi(N-1)}{N}\right) + 2\cos\left(\frac{2\pi(N-1)}{N}\right)\right)\right).$$

$$(5.36)$$

Therefore, the condition number of $\mathbf{C}_{Lap}$ is

$$\kappa(\mathbf{C}_{Lap}) = \left(1 + \frac{L^4}{2\Delta x^4}\left(6 - 8\cos\left(\frac{\pi(N-1)}{N}\right) + 2\cos\left(\frac{2\pi(N-1)}{N}\right)\right)\right).$$
(5.37)

This completes the proof.

If N is large then $(N-1)/N \approx 1$ and in the case for N odd

$$\cos\left(\frac{\pi(N-1)}{N}\right) \approx \cos(\pi) = -1,$$
(5.38)

and

$$\cos\left(\frac{2\pi(N-1)}{N}\right) \approx \cos(2\pi) = 1.$$
(5.39)

Hence, by equation (5.37) the condition number of $\mathbf{C}_{Lap}$ is

$$\kappa(\mathbf{C}_{Lap}) \approx \left(1 + \frac{L^4}{2\Delta x^4}(6 - 8(-1) + 2(1))\right) = 1 + 16\frac{L^4}{2\Delta x^4}.$$
(5.40)

Thus for large N the expression of the condition number of $\mathbf{C}_{Lap}$ is approximately the same for N odd or even.

Consider the same grid with $N = 500$ and $\Delta x = 0.1$ as in the Gaussian and SOAR examples in Sections 5.3.1 and 5.3.2 respectively. By Theorem 5.3.1 we would expect the condition number of the Laplacian to be sensitive to the lengthscale L, since the conditioning is approximately proportional to L^4. A plot of the condition number of $\mathbf{C}_{Lap}$ is shown in Figure 5.6. Indeed, the condition number becomes larger as the lengthscale increases as expected. The Laplacian appears to be the most well-conditioned of the three correlation matrices considered. For instance, the condition number of $\mathbf{C}_{Lap}$ is an order of magnitude 10 lower than the SOAR matrix at $L = 0.3$.

As in the previous examples Figure 5.7 shows that the large reduction in the smallest eigenvalue as the lengthscale increases causes the ill-conditioning of the correlation matrix as the lengthscale increases. Comparisons of Figures 5.3, 5.5 and 5.7 show that the increase in the largest eigenvalue is approx-

imately linear and relatively slow for all three correlation matrices, this is expected by considering the bound on the eigenvalues (5.27). As the length-scale increases the correlations slowly increase, as seen from the forms of the components of the correlation matrices (5.10) and (5.11). It is the rate of reduction in the smallest eigenvalue as the lengthscale increases that determines the conditioning of all three of the covariance matrices considered in this section. As the lengthscale increases the Gaussian correlation matrix (5.10), the SOAR correlation matrix (5.11) and the Laplacian correlation matrix (5.12) all become very ill conditioned with the condition number of the Gaussian increasing the most dramatically. Of the correlation matrices considered in this chapter the Gaussian is the most sensitive to the increase in the lengthscale and the Laplacian is the least sensitive. We now summarise the results of this chapter.

5.4 Summary

In this chapter we considered background covariance functions and matrices. In particular, we studied the condition number of some auto-correlation matrices which are commonly used to describe the horizontal error correlations in NWP. We began by introducing some basic definitions and results of covariance functions before introducing some commonly used correlation functions on the real line. Since in this thesis we are interested in correlation functions on the circle we showed how a change of variables could be applied to a correlation function on the real line to produce a valid correlation model on the circle. The correlation functions could then be used to produce co-variance matrices for periodic parameters on a finite domain on the real line. Finally, we considered the conditioning of three commonly used covariance matrices. It was found that all three matrices are sensitive to the correlation lengthscale and become more ill-conditioned as the lengthscales increases. In particular, it was found that the increase in the condition number of the matrices is caused by the rapid reduction in the size of the smallest eigenvalue

as the lengthscale increased. A comparison of the three covariance matrices showed a large variability in the conditioning of such matrices. Depending on the correlation matrix considered the condition number of the background covariance matrix is potentially several magnitudes larger or smaller. In particular, we generally found the Gaussian matrix to have a larger condition number than the Laplacian with the same lengthscale.

In the next chapter we consider the conditioning of the unpreconditioned and preconditioned 3DVar problems (2.11) and (2.16). In addition we show how the condition number of the background covariance matrix can influence the conditioning of the unpreconditioned Var problem (2.11).

Chapter 6

Conditioning of 3DVar

In this chapter we present new theoretical bounds which quantify the condition number of both the unpreconditioned and preconditioned Hessians of the 3DVar objective function (2.11). Using these bounds we show

- The conditioning of the Var Hessian is dependent on the conditioning of the background error covariance matrix. Hence, for an ill-conditioned background error covariance matrix the Var Hessian is also ill-conditioned.

- In a one parameter, periodic system, preconditioning with the control variable transform (CVT), as described in Section 2.4, reduces the condition number of Var.

- The condition number of the preconditioned Var system is reduced by three main factors: reducing the number of observations, using less accurate observations and increasing the spacing between observations.

- The convergence rate of the conjugate gradient (CG) method applied to Var is increased when we precondition with the CVT.

- The convergence rate of the CG method applied to the preconditioned

3DVar system increases when we increase the spacing between observations and decrease the observation accuracy.

We begin by deriving new bounds on the condition number of the 3DVar Hessian. These bounds show the dependency of the conditioning of 3DVar to the conditioning of the background error covariance matrix and we illustrate this with numerical experiments. Next we derive new theoretical bounds on the condition number of the Hessian of the 3DVar system which has been preconditioned via the CVT. We illustrate the benefits of preconditioning using numerical experiments and compare these results to the conditioning of the unpreconditioned system. We consider the conditioning of the preconditioned 3DVar scheme in more detail. In particular, we consider the effect of changing the observation accuracy and distribution on the condition number. Finally we emphasise the importance of the condition number in determining the convergence rate of gradient methods by performing numerical experiments using the conjugate gradient method.

We begin by investigating the conditioning of the unpreconditioned 3DVar system.

6.1   Conditioning of 3DVar

In this chapter we investigate the conditioning of 3DVar. Observations, $\mathbf{y}_0 \in \mathbb{R}^p$, are only taken at the beginning of the time window, t_0. Therefore, instead of minimising the full incremental 4DVar problem (2.9) we minimise the incremental 3DVar cost function

$$\tilde{J}[\delta\mathbf{x}_0] = \frac{1}{2}[\delta\mathbf{x}_0 - (\mathbf{x}_0^b - \mathbf{x}_0)]^T\mathbf{B}^{-1}[\delta\mathbf{x}_0 - (\mathbf{x}_0^b - \mathbf{x}_0)] + \frac{1}{2}(\mathbf{H}\delta\mathbf{x}_0 - \mathbf{d}_0)^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x}_0 - \mathbf{d}_0),$$

(6.1)

with respect to the initial increment $\delta\mathbf{x}_0 \in \mathbb{R}^N$. In this chapter, subsequent reference to 3DVar will mean the incremental form (6.1). In the next section

we derive theoretical bounds on the condition number of the Hessian of the 3DVar cost function.

6.1.1 Theory

In this section we consider the condition number of the Hessian of the 3DVar cost function (6.1)

$$\mathbf{S} = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}. \tag{6.2}$$

We derive theoretical bounds to analyse the condition number of (6.2). To begin we consider the general problem with minimal assumptions on the observations and the errors. The main result is contained in the following theorem.

Theorem 6.1.1 *Let $\mathbf{B} \in \mathbb{R}^{N \times N}$ and $\mathbf{R} \in \mathbb{R}^{p \times p}$, with $p < N$, be the background and observation error covariance matrices respectively. Additionally, let $\mathbf{H} \in \mathbb{R}^{p \times N}$ be the observation operator. Then the following bounds are satisfied by the condition number of the Hessian (6.2),*

$$\frac{\kappa(\mathbf{B})}{(1 + \lambda_{\max}(\mathbf{B})\lambda_{\max}(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}))} \leq \kappa(\mathbf{S}) \leq (1 + \lambda_{\min}(\mathbf{B})\lambda_{\max}(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}))\kappa(\mathbf{B}). \tag{6.3}$$

Proof By bounding the largest and smallest eigenvalues of $\mathbf{S}$ we can bound the condition number. By Theorem 3.4.1 we can put preliminary bounds on the eigenvalues

$$\lambda_k(\mathbf{B}^{-1}) + \lambda_{\min}(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \leq \lambda_k(\mathbf{S}) \leq \lambda_k(\mathbf{B}^{-1}) + \lambda_{\max}(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}). \tag{6.4}$$

However since $\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$ is not full rank its smallest eigenvalue is equal to zero. This means that

$$\lambda_{\max}(\mathbf{B}^{-1}) \leq \lambda_{\max}(\mathbf{S}) \leq \lambda_{\max}(\mathbf{B}^{-1}) + \lambda_{\max}(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}), \tag{6.5}$$

and

$$\lambda_{\min}(\mathbf{B}^{-1}) \leq \lambda_{\min}(\mathbf{S}) \leq \lambda_{\min}(\mathbf{B}^{-1}) + \lambda_{\max}(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}). \qquad (6.6)$$

Taking the upper bound on the largest eigenvalue in (6.5) and the lower bound on the smallest eigenvalue in (6.6) gives the following upper bound on the condition number of (6.2)

$$\kappa(\mathbf{S}) \leq \left(1 + (\lambda_{\max}(\mathbf{B}^{-1}))^{-1}\lambda_{\max}(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\right)\kappa(\mathbf{B}). \qquad (6.7)$$

Since $(\lambda_{\max}(\mathbf{B}^{-1}))^{-1} = \lambda_{\min}(\mathbf{B})$ we obtain the upper bound

$$\kappa(\mathbf{S}) \leq \left(1 + \lambda_{\min}(\mathbf{B})\lambda_{\max}(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\right)\kappa(\mathbf{B}). \qquad (6.8)$$

Similarly taking the lower bound on the largest eigenvalue in (6.5) and the upper bound on the smallest eigenvalue in (6.6) gives the following lower bound on the condition number

$$\kappa(\mathbf{S}) \geq \kappa(\mathbf{B})\left(1 + (\lambda_{\min}(\mathbf{B}^{-1}))^{-1}\lambda_{\max}(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\right)^{-1}. \qquad (6.9)$$

Using $\lambda_{\min}(\mathbf{B}^{-1}) = (\lambda_{\max}(\mathbf{B}))^{-1}$ we obtain the lower bound

$$\kappa(\mathbf{S}) \geq \kappa(\mathbf{B})\left(1 + \lambda_{\max}(\mathbf{B})\lambda_{\max}(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\right)^{-1}. \qquad (6.10)$$

This completes the proof.

The main assumption of Theorem 6.1.1 is that the 3DVar cost function is underdetermined $(p < N)$ with fewer observations then components in the increment $\delta\mathbf{x}_0$. Currently in NWP centres there are only 10^6 observations compared to the $10^7 - 10^8$ components in the state variable. Theorem 6.1.1 shows that the bounds on the condition number of (6.2) are related to the condition number of the background covariance matrix. A direct result of this theorem is that if the background covariance matrix is ill-conditioned then we can expect the Hessian (6.2) to be ill-conditioned.

We can produce tighter bounds on the condition number by restricting the forms of the components of the Hessian (6.2). We assume that our background covariance matrix is a circulant matrix of the form $\mathbf{B} = \sigma_b^2 \mathbf{C}$, where σ_b^2 is the background error variance. It is common practice in NWP centres to assume that the observation errors are uncorrelated [15]. Here we assume that the observation error covariance matrix is just a scalar multiple of the identitiy matrix $\mathbf{R} = \sigma_o^2 \mathbf{I}$, where σ_o^2 is the observation error variance. Additionally, we assume that the observations are made at grid points. In particular, this implies that $\mathbf{H}^T\mathbf{H}$ is diagonal with only ones on the diagonal if that position is observed and zeros otherwise. The bounds on this more specific system are derived in the following theorem.

Theorem 6.1.2 *Let* $\mathbf{B} = \sigma_b^2 \mathbf{C} \in \mathbb{R}^{N \times N}$ *and* $\mathbf{R} = \sigma_o^2 \mathbf{I}_p$ *where* $\mathbf{C}$ *is a symmetric positive-definite circulant matrix,* $\mathbf{I}_p \in \mathbb{R}^{p \times p}$ *is the identity matrix and* σ_b^2 *and* σ_o^2 *are positive scalars. In addition let* $\mathbf{H}^T\mathbf{H}$ *be a diagonal matrix with* $p < N$ *units on the diagonal and the remaining elements zero. Defining* $\mathbf{S} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$, *the following bounds on the condition number hold*

$$\left(\frac{1 + \frac{p}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_{\min}(\mathbf{C})}{1 + \frac{p}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_{\max}(\mathbf{C})} \right) \kappa(\mathbf{C}) \leq \kappa(\mathbf{S}) \leq \left(1 + \left(\frac{\sigma_b^2}{\sigma_o^2}\right)\lambda_{\min}(\mathbf{C}) \right)\kappa(\mathbf{C}), \quad (6.11)$$

where $\lambda_{\max}(\mathbf{C})$ *and* $\lambda_{\min}(\mathbf{C})$ *are the largest and smallest eigenvalues respectively of the matrix* $\mathbf{C}$.

Proof By the assumptions on the matrices in the theorem we can write $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} = \sigma_o^{-2}\mathbf{H}^T\mathbf{H}$ and therefore $\lambda_{\max}(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) = \sigma_o^{-2}$. Additionally, we have $\lambda_{\min}(\mathbf{B}) = \sigma_b^2\lambda_{\min}(\mathbf{C})$. If we substitute these into the upper bound of (6.3) we obtain

$$\kappa(\mathbf{S}) \leq \left(1 + \frac{\sigma_b^2}{\sigma_o^2}\lambda_{\min}(\mathbf{C}) \right)\kappa(\mathbf{C}), \quad (6.12)$$

which establishes the upper bound. Rather then repeat this procedure with the lower bound we produce an improved estimate by applying the Rayleigh

quotient (3.46). Let $\mathbf{v}_{\max} \in \mathbb{R}^N$ be the eigenvector corresponding to the largest eigenvalue of $\mathbf{C}^{-1}$. Since $\mathbf{C}^{-1}$ is circulant then all the components of the eigenvectors of $\mathbf{C}^{-1}$ lie on the unit circle in $\mathbb{C}$ (see (3.50)). In particular this implies that for any eigenvector, $\mathbf{v}_m$, of $\mathbf{C}^{-1}$

$$\mathbf{v}_m^H \mathbf{H}^T \mathbf{H} \mathbf{v}_m = \frac{1}{N} \sum_{k \in K} (e^{-2\pi i k m/N})^H e^{-2\pi i k m/N} = \frac{1}{N} \sum_{k \in K} e^{2\pi i k m/N} e^{-2\pi i k m/N} = \frac{p}{N},$$
(6.13)

where K are the positions of the non-zero diagonal elements of $\mathbf{H}^T \mathbf{H}$. By Theorem 3.4.4, the maximum value obtained by the Rayleigh quotient of $\mathbf{S}$ occurs at the eigenvector corresponding to the largest eigenvalue of $\mathbf{S}$. Hence,

$$\lambda_{\max}(\mathbf{S}) = \max_{\mathbf{v} \in \mathbb{R}^N}(R_{\mathbf{S}}(\mathbf{v})) \geq \mathbf{v}_{\max}^H(\mathbf{B}^{-1} + \sigma_o^{-2}\mathbf{H}^T\mathbf{H})\mathbf{v}_{\max} = \sigma_b^{-2}\lambda_{\max}(\mathbf{C}^{-1}) + \sigma_o^{-2}\frac{p}{N}.$$
(6.14)

Similarly the minimum value of the Rayleigh quotient occurs at the eigenvector corresponding to the smallest eigenvalue of $\mathbf{S}$. Let $\mathbf{v}_{\min}$ be the eigenvector corresponding to the smallest eigenvalue of $\mathbf{C}^{-1}$. Then again using the Rayleigh quotient we find

$$\lambda_{\min}(\mathbf{S}) = \min_{\mathbf{v} \in \mathbb{R}^N}(R_{\mathbf{S}}(\mathbf{v})) \leq \mathbf{v}_{\min}^H(\mathbf{B}^{-1} + \sigma_o^{-2}\mathbf{H}^T\mathbf{H})\mathbf{v}_{\min} = \sigma_b^{-2}\lambda_{\min}(\mathbf{C}^{-1}) + \sigma_o^{-2}\frac{p}{N}.$$
(6.15)

Combining (6.14) and (6.15) we find

$$\kappa(\mathbf{S}) \geq \frac{\sigma_b^{-2}\lambda_{\max}(\mathbf{C}^{-1}) + \sigma_o^{-2}\frac{p}{N}}{\sigma_b^{-2}\lambda_{\min}(\mathbf{C}^{-1}) + \sigma_o^{-2}\frac{p}{N}} = \kappa(\mathbf{C}) \left(\frac{1 + \frac{\sigma_b^2}{\sigma_o^2}\frac{p}{N}\lambda_{\min}(\mathbf{C})}{1 + \frac{\sigma_b^2}{\sigma_o^2}\frac{p}{N}\lambda_{\max}(\mathbf{C})} \right), \qquad (6.16)$$

giving the lower bound on the condition number. This completes the proof.

As with the more general theorem the bounds show that the condition number of the Hessian is related to the condition number of the background error covariance matrix. In Chapter 5 we introduced some common circulant background covariance matrices of the form $\mathbf{B} = \sigma_b^2 \mathbf{C}$ whose conditioning were sensitive to changes in the error correlation lengthscale. In particular, the correlation models considered were ill-conditioned. The theory presented

in this section implies that the 3DVar cost function (6.1) constructed using these background covariance matrices would also be sensitive to changes in the lengthscale. Hence, as the lengthscale increases we expect the condition number of 3DVar Hessian, $\mathbf{S}$, to be more ill-conditioned. In the next section we perform numerical experiments to illustrate the ill-conditioning of the 3DVar Hessian (6.2).

6.1.2   Numerical Experiments

In this section we confirm numerically the dominance of the conditioning of the background covariance matrix on the conditioning of the Hessian (6.2) as indicated by the theoretical bounds from Section 6.1.1. We consider a one-parameter periodic system defined on an equally spaced grid on the real line. For simplicity we assume that observations of the parameter are only made at grid points and have uncorrelated errors all with the same variance σ_o^2. Additionally, we assume that the background error covariance matrix is of the form $\sigma_b^2 \mathbf{C}$ with background error variance σ_b^2 and $\mathbf{C}$ is a circulant correlation matrix. With these assumptions the hypotheses of Theorem 6.1.2 are satisfied and the bounds (6.11) apply to the condition number of the Hessian (6.2).

For the remainder of this section we define a background covariance matrices on the circle using the Gaussian, SOAR and Laplacian correlation matrices defined in Section 5.2. We consider a domain consisting of $N = 500$ grid points with equal grid spacing $\Delta x = 0.1$ between adjacent nodes. We fix the observation and background error variances at $\sigma_o^2 = 1$ and $\sigma_b^2 = 1$ respectively. In this section we use two different configurations of observations. The first uses a set of 250 randomly distributed observations, we refer to this as observation configuration 1. The second, uses 250 observations positioned on the first 250 grid points of the domain. We call this observation configuration 2. We note that since both configurations contain the same number of

observations then the lower bound in (6.11) will be identical in both cases. In these experiments we consider the conditioning of the Hessian (6.2) as we vary the background error correlation lengthscale.

Figure 6.1: Condition number of the Hessian (solid line) and bounds (dashed line) as function of lengthscale with Gaussian background covariance matrix and observation configuration 1.

To begin with we consider the Gaussian background correlation matrix (5.10) introduced in Section 5.2 using observation configuration 1. Figure 6.1 is a log plot of the change in the conditioning of the Hessian as a function of background error correlation lengthscale (black solid line) together with the bounds derived in Theorem 6.1.2 (red dashed lines). Firstly, we note the close agreement between the condition number and the derived bounds. Although the difference between upper and lower bounds appears constant, since the y-axis is on a logarithmic scale the separation clearly increases with the lengthscale and the bounds are less tight. The important feature of this graph is the large increase in the condition number as a function of lengthscale. In particular, if we compare Figure 6.1 with Figure 5.2, showing the conditioning of the Gaussian covariance matrix, we see a close agreement in both the shape and magnitude of the curves. For instance when the

lengthscale is twice the grid length spacing $L = 0.2$ then the conditioning of both are of the order 10^7. This confirms Theorem 6.1.2 which predicts that the conditioning of the Hessian (6.2) is dependent on the conditioning of the background covariance matrix.
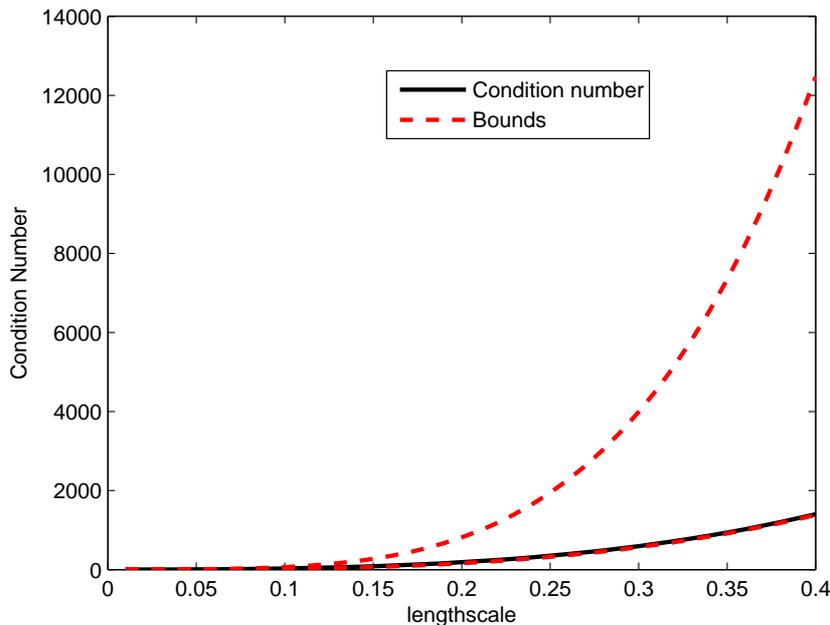
Figure 6.2: Condition number of the Hessian (solid line) and bounds (dashed line) as function of lengthscale with SOAR background covariance matrix and observation configuration 1.

We repeat the experiment, using the SOAR correlation matrix (5.11) introduced in Section 5.2 using observation configuration 1. As seen from the results in Figure 6.2 there is an increase in the condition number and bounds as the lengthscale increases. In this case the Hessian is ill-conditioned but much better conditioned than the Hessian using the Gaussian covariance matrix (we note that in this plot the x-axis has been extended due to the lower rate of growth of the condition number compared to the Gaussian.). Comparison with the conditioning of the background covariance matrix, shown in Figure 5.4, shows that in this case the Hessian has a lower condition number than the background matrix. If we look at lengthscale $L = 0.3$ we see that the condition number of the Hessian is around 500 compared to that of the background covariance matrix which is around 4000. We also note that
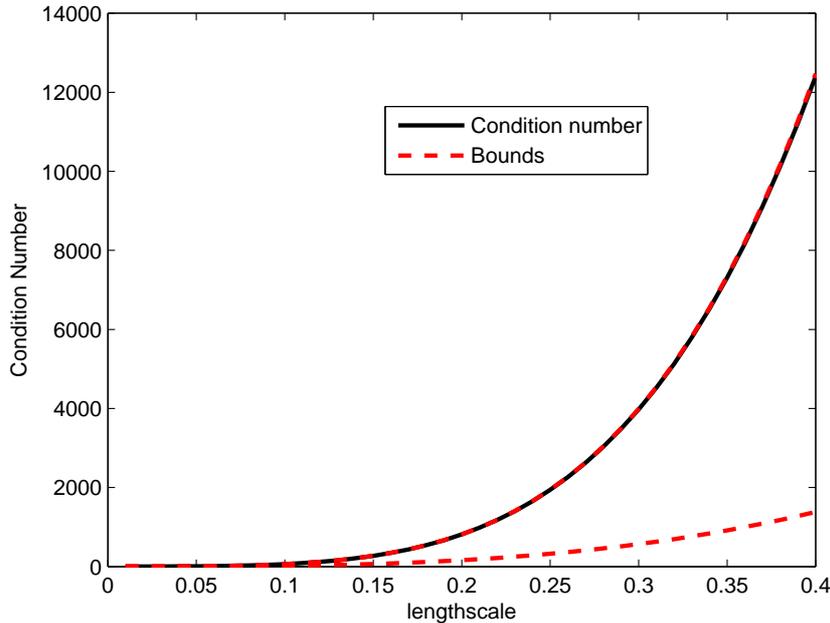
Figure 6.3: Condition number of the Hessian (solid line) and bounds (dashed line) as function of lengthscale with SOAR background covariance matrix and observation configuration 2.

the size of the upper bound in Figure 6.2 is similar to that of the condition number of background error covariance matrix. The condition number of $\mathbf{S}$ in this case has followed the lower bound but in other cases can follow the upper bound. We repeated the experiment using observation configuration 2. The condition number of the Hessian is of the same magnitude as the condition number of the SOAR background covariance matrix. In this case the condition number follows the upper bound as shown in Figure 6.3. The experiments show that the condition number of $\mathbf{S}$ can achieve both the lower and upper bounds presented in Theorem 6.1.2 depending on the choice of observations.

Finally we consider the conditioning of the Hessian constructed using the Laplacian background covariance matrix (5.12) introduced in Section 5.2. Figure 6.4 shows the condition number against lengthscale for observation configuration 1. Like the SOAR case the increase in the condition number with increase in correlation lengthscale is not as dramatic as for the Gaussian case. This is evident from the use of a linear y-axis used in Figure 6.4 com-
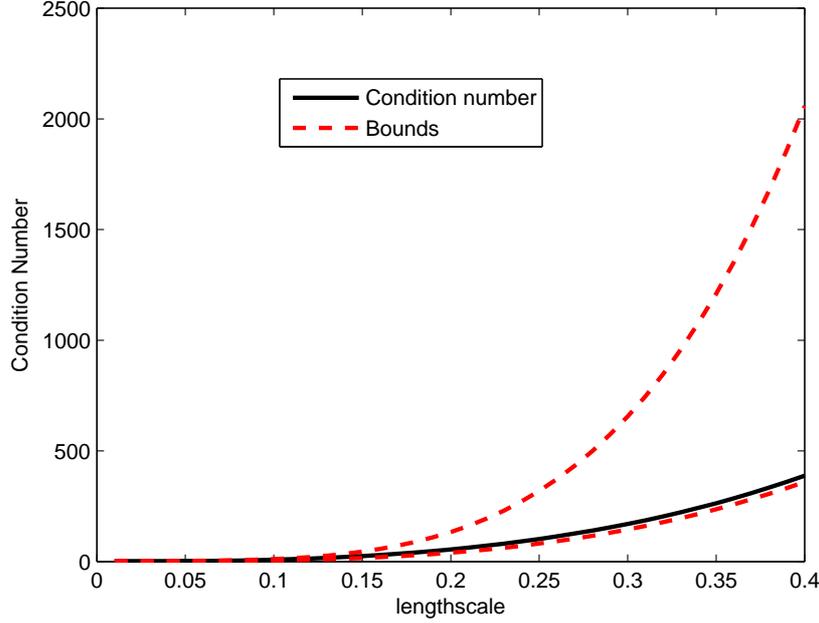
Figure 6.4: Condition number of the Hessian (solid line) and bounds (dashed line) as function of lengthscale with Laplacian background covariance matrix and observation configuration 1.

pared to the logarithmic y-axis used for Figure 6.1. As with the SOAR case, with this observation distribution, the condition number follows the lower bound. However, using observation configuration 2, the condition number of the Hessian using the Laplacian follows the upper bound as shown in Figure 6.5. As with the SOAR case, depending on the observations chosen, both the upper and lower bounds can be achieved by the condition number of the Hessian. Hence, the bounds (6.11) are a close approximation to the possible values of the condition number of the Hessian.

In these examples we have seen that for large lengthscales the Hessian (6.2) is ill-conditioned and this appears to be driven by the conditioning of the background error covariance matrices. In contrast, $\mathbf{C} \longrightarrow \mathbf{I}_N$ as $L \longrightarrow 0$ and therefore

$$(\mathbf{B}^{-1} + \sigma_o^{-2}\mathbf{H}^T\mathbf{H}) \longrightarrow (\sigma_b^{-2}\mathbf{I}_N + \sigma_o^{-2}\mathbf{H}^T\mathbf{H}), \qquad (6.17)$$

as $L \longrightarrow 0$ and the condition number of the Hessian becomes $1 + \frac{\sigma_b^2}{\sigma_o^2}$. This is confirmed by the lower and upper bounds (6.11) which are 1 and $1 + \frac{\sigma_b^2}{\sigma_o^2}$ for
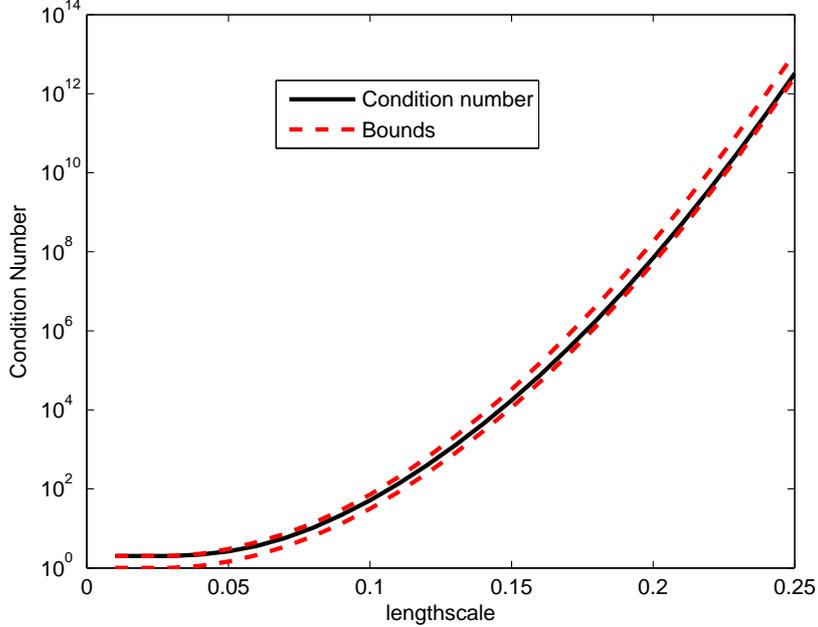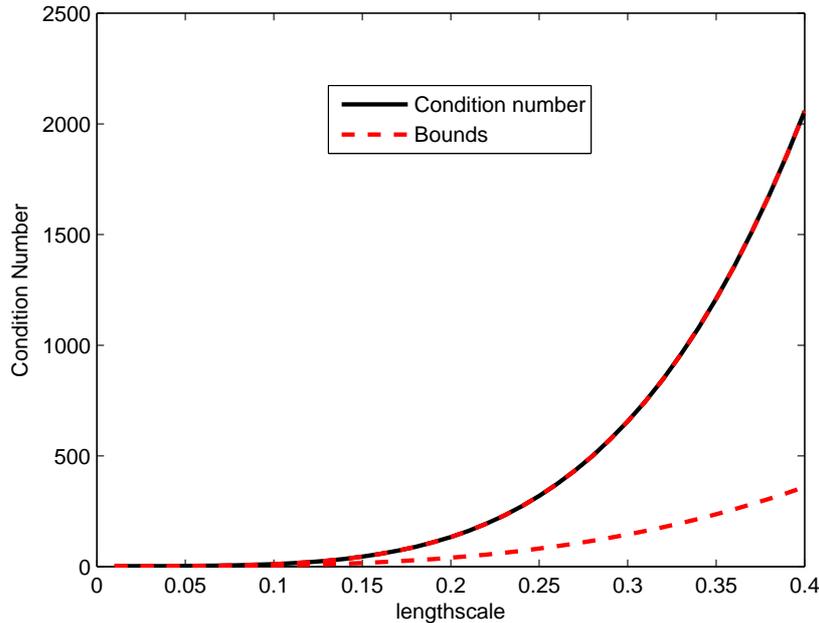
Figure 6.5: Condition number of the Hessian (solid line) and bounds (dashed line) as function of lengthscale with Laplacian background covariance matrix and observation configuration 2.

$\mathbf{B} = \sigma_b^2 \mathbf{I}_N$ respectively. As long as the ratio of the background and observation error variances is not excessively large we would expect the conditioning to be reasonably small when the lengthscale L is small. Hence from a conditioning perspective a small error correlation lengthscale is desirable. This is confirmed in the experiments presented here. In the next section we consider the effect of the control variable transform (CVT) on the conditioning of the problem.

6.2 Conditioning of Preconditioned 3DVar

In this section we consider the effect of the control variable transform (CVT), as described in Section 2.4, on the condition number of 3DVar cost function. As mentioned in Chapter 4 Lorenc showed in a simplified case that including the CVT in Var improved the convergence rate of the 3DVar minimisation scheme [42], indicating that the CVT improves the condition number of the

3DVar Hessian (6.2). In the next Section we consider the conditioning of the preconditioned 3DVar (which we refer to as P3DVar from here onwards) system in more detail by developing new theoretical bounds on the condition number of the P3DVar Hessian.

6.2.1 Theory

As shown in Chapter 4, including the CVT in Var is equivalent to symmetrically preconditioning the Hessian by the square root of the background covariance matrix $\mathbf{B}^{1/2}$. In this chapter we assume that we are using the uniquely defined, symmetric square root and hence $\mathbf{B}^{T/2} = \mathbf{B}^{1/2}$. The preconditioned 3DVar Hessian has the form

$$\hat{\mathbf{S}} = \mathbf{I}_N + \mathbf{B}^{T/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2} = \mathbf{I}_N + \mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2}. \qquad (6.18)$$

In this section we derive new theoretical bounds on the condition number of $\hat{\mathbf{S}}$. As in Section 6.1.1 we first provide bounds on a general Hessian (6.18) before developing more informative bounds on a more restrictive Hessian with tighter assumptions. We begin with the general problem.

Theorem 6.2.1 *Let $\mathbf{B} \in \mathbb{R}^{N \times N}$ be the background error covariance matrix and $\mathbf{R} = \in \mathbb{R}^{p \times p}$ be the observation error covariance matrix with $p < N$. Then the following bounds are satisfied by the condition number of the Hessian $\hat{\mathbf{S}} = \mathbf{I}_N + \mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2}$*

$$1 + \frac{1}{p}\sum_{i,j=1}^{p}(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2})_{i,j} \leq \kappa(\hat{\mathbf{S}}) \leq 1 + ||\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2}||_\infty,$$

$$(6.19)$$

where $\mathbf{A}_{i,j}$ represents the $(i,j)^{th}$ entry of matrix $\mathbf{A}$ and $\mathbf{A}^{1/2}$ is the symmetric square root of $\mathbf{A}$.

Proof By Theorem 3.4.5 we know that a matrix of the form $\mathbf{E}^T\mathbf{E}$ has the same non-zero eigenvalues as $\mathbf{E}\mathbf{E}^T$. If we let $\mathbf{E} = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}^{1/2}$ then by

Theorem 3.4.5 the Hessian (6.18) has the same eigenvalues (with the addition of $N - p$ unit eigenvalues) as the matrix,

$$\mathbf{I}_p + \mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}^{1/2}\mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1/2} = \mathbf{I}_p + \mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2}. \qquad (6.20)$$

Since $p < N$ then $\lambda_{\min}(\hat{\mathbf{S}}) = 1$ and using Theorem 3.4.2 with the ∞-norm we obtain the upper bound

$$\kappa(\hat{\mathbf{S}}) = \lambda_{\max}(\mathbf{I}_p + \mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2}) \leq 1 + ||\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2}||_\infty. \quad (6.21)$$

A lower bound is established by application of the Rayleigh quotient given in Definition 3.4.3. Let $\mathbf{v} = \frac{1}{\sqrt{p}}(1, 1, \ldots, 1)^T \in \mathbb{R}^p$ then we have

$$\kappa(\hat{\mathbf{S}}) = \lambda_{\max}(\mathbf{I}_p + \mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2}) \geq \mathbf{v}^T(\mathbf{I}_p + \mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2})\mathbf{v}$$

$$(6.22)$$

$$= 1 + \frac{1}{p}\sum_{i,j=1}^{p}(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2})_{i,j},$$

$$(6.23)$$

which completes the proof.

Theorem 6.2.1 shows that the conditioning of the Hessian (6.18) is now dependent on sums of the elements of the matrix $\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2}$. Since the matrices $\mathbf{H}, \mathbf{B}$ and $\mathbf{R}$ are general we cannot make specific comments on how they affect the conditioning of the P3DVar Hessian. However, the bounds can indicate factors which may cause ill-conditioning. If we assume that $\mathbf{R}^{-1/2} = \mathbf{D} = diag(d_1, d_2, \ldots, d_p)$, then the $(i, j)^{th}$ element of $\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2}$ is

$$(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2})_{i,j} = d_i(\mathbf{H}\mathbf{B}\mathbf{H}^T)_{i,j}d_j. \qquad (6.24)$$

The more accurate the observations, the larger the d_i are and by equation (6.24) the larger the components of $\mathbf{R}^{-1/2}\mathbf{HBH}^T\mathbf{R}^{-1/2}$ and hence the larger the bound in Theorem 6.2.1. Hence, with more accurate observations the Hessian $\hat{\mathbf{S}}$ is potentially more ill-conditioned.

We now make additional simplifying assumptions on the observations and errors to derive more explicit, descriptive bounds on the condition number of (6.25). We assume that the background and observation error covariance matrices can now be written $\mathbf{B} = \sigma_b^2 \mathbf{C}$ and $\mathbf{R} = \sigma_o^2 \mathbf{I}_p$ respectively. We also assume that observations are made only at grid points.

Theorem 6.2.2 *Let $\mathbf{B} = \sigma_b^2 \mathbf{C} \in \mathbb{R}^{N \times N}$ and $\mathbf{R} = \sigma_o^2 \mathbf{I}_p \in \mathbb{R}^{p \times p}$ be the error covariance matrices of the background and observations respectively with $p < N$ and scalars $\sigma_b^2, \sigma_o^2 > 0$. Also suppose that observations are made at grid points only. Then the following bounds on the condition number of $\hat{\mathbf{S}} = \mathbf{I}_N + \mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{HB}^{1/2}$ hold*

$$1 + \frac{1}{p}\frac{\sigma_b^2}{\sigma_o^2}\sum_{i,j \in K} c_{i,j} \leq \kappa(\hat{\mathbf{S}}) \leq 1 + \frac{\sigma_b^2}{\sigma_o^2}\|\mathbf{HCH}^T\|_\infty, \qquad (6.25)$$

where K are indices of the state variables which are observed and $\mathbf{B}^{1/2}$ is the symmetric square root of $\mathbf{B}$.

Proof From (6.21) in Theorem 6.2.1 we know that

$$\kappa(\hat{\mathbf{S}}) \leq 1 + \|\mathbf{R}^{-1/2}\mathbf{HBH}^T\mathbf{R}^{-1/2}\|_\infty = 1 + \frac{\sigma_b^2}{\sigma_o^2}\|\mathbf{HCH}^T\|_\infty, \qquad (6.26)$$

which proves the upper bound.

Similarly using (6.23) from Theorem 6.2.1 we know

$$\kappa(\hat{\mathbf{S}}) \geq 1 + \frac{1}{p}\sum_{i,j=1}^{p}(\mathbf{R}^{-1/2}\mathbf{HBH}^T\mathbf{R}^{-1/2})_{i,j} = 1 + \frac{1}{p}\frac{\sigma_b^2}{\sigma_o^2}\sum_{i,j=1}^{p}(\mathbf{HCH}^T)_{i,j}. \quad (6.27)$$

Since observations are only made at grid points, $\mathbf{HCH}^T$ is the matrix $\mathbf{C}$ with $N - p$ rows and columns removed at the positions which are not observed. This completes the proof.

We refer to $\mathbf{HCH}^T$ as the *reduced error correlation matrix* since it is the matrix $\mathbf{C}$ but containing only the row and columns that are observed. Since $\mathbf{C}$ is a correlation matrix, the entries have maximum value equal to unity and therefore from (6.25) we have an absolute upper bound on the condition number of the Hessian $\hat{\mathbf{S}}$ equal to

$$\kappa(\hat{\mathbf{S}}) \leq 1 + p\frac{\sigma_b^2}{\sigma_o^2}. \tag{6.28}$$

Now suppose that $\mathbf{C}$ is an auto-correlation matrix, for example, the SOAR correlation matrix (5.11). As we increase the correlation lengthscale the size of the entries of $\mathbf{C}$ increases and so do the terms of $\mathbf{HCH}^T$ and the bounds in (6.25). Therefore increasing the lengthscale increases the bounds (6.25) and potentially increases the condition number of the Hessian $\hat{\mathbf{S}}$.

From the bounds (6.25) other factors clearly effect the condition number of the Hessian of P3DVar. The condition number of $\hat{\mathbf{S}}$ is linearly related to the inverse of the observation variance. Hence, we expect more accurate observations to produce a more ill-conditioned system. In addition, the presence of the reduced error correlation matrix indicates that the number and distribution of the observations also plays an important role in determining the conditioning of P3DVar. We discuss the effect of the number of observations and the accuracy and positions of the observations in more detail in Section 6.3.

In the next section we investigate the conditioning of P3DVar using simple numerical experiments. We concentrate on the conditioning of $\hat{\mathbf{S}}$ as a function of lengthscale in order to illustrate the reduction achieved in the condition number of Var by preconditioning with the CVT.

6.2.2 Numerical Experiments
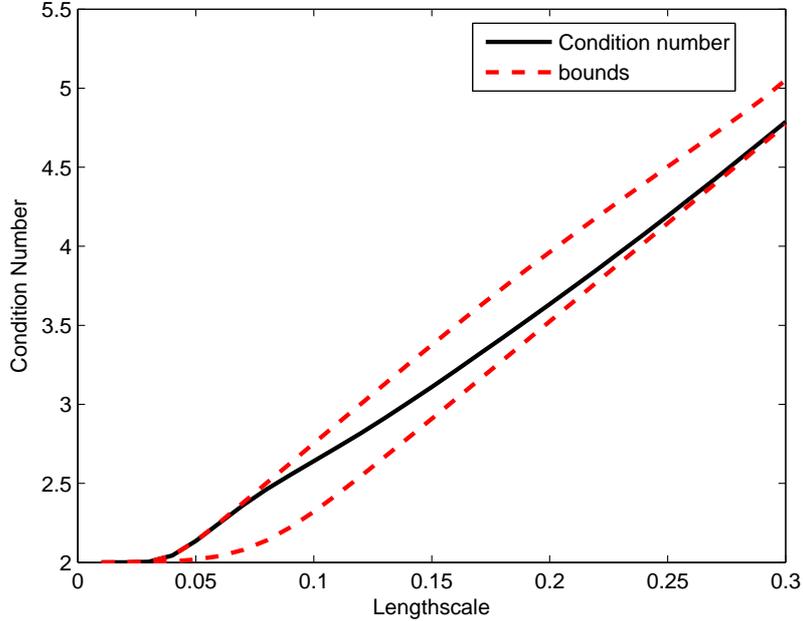
Figure 6.6: Condition number of the preconditioned Hessian (Solid line) and bounds (dashed line) as function of lengthscale with Gaussian background covariance matrix and observation configuration 1.

In this section we consider the condition of the P3DVar system numerically. As in Section 6.1.2 we consider a one-parameter, periodic system defined on an equally spaced grid. In this section we use a domain of $N = 500$ grid points with grid spacing $\Delta x = 0.1$. We observe 250 grid points and assume that the observations all have error variance $\sigma_o^2 = 1$. As in Section 6.1.2 the observations are distributed with observation configuration 1. We let $\mathbf{B} = \sigma_b^2 \mathbf{C}$, where $\sigma_b^2 = 1$ is the background error variance and $\mathbf{C}$ is either the Gaussian, SOAR or Laplacian correlation matrices defined on the circle as introduced in section 5.2. Hence, with this model design, the assumptions of Theorem 6.2.2 apply and the Hessian $\hat{\mathbf{S}}$ satisfies the bounds (6.25). By equation (6.28) the absolute maximum condition number of the Hessian is

$$1 + p\frac{\sigma_b^2}{\sigma_o^2} = 251. \tag{6.29}$$

A comparison with any of the numerical experiments in Section 6.1.2 shows

that P3DVar is likely to be better conditioned then the 3DVar system. For instance, consider the Hessian which uses the Laplacian correlation matrix (5.12) (This Hessian has the smallest condition number of the three considered: the Gaussian, SOAR and Laplacian). Comparison with Figure 6.4 shows that with lengthscale $L = 4\Delta x = 0.4$ the condition number of the Hessian, $\mathbf{S}$, of the unpreconditioned system has already exceeded the maximum possible condition number of the P3DVar system.

We begin our P3DVar experiments by considering the condition number of the Hessian $\hat{\mathbf{S}}$ with the Gaussian correlation matrix (5.10) as the correlation matrix $\mathbf{C}$. Figure 6.6 shows the condition number of P3DVar as a function of lengthscale. Firstly, as expected, the condition number of the Hessian (6.18) increases as we increase the error correlation lengthscale. This is clear from the bounds (6.25) since we are summing terms from the reduced error correlation matrix $\mathbf{HCH}^T$. As we increase the correlation lengthscale we increase the size of the elements of the reduced correlation matrix. Comparison of Figures 6.1 and 6.6 shows the clear benefits of preconditioning with the CVT. For instance, if we consider the condition number at lengthscale $L = 2\Delta x = 0.2$ then we can see a dramatic reduction in the size of the condition number of P3DVar compared to 3DVar from 10^7 to 3.7.

Figure 6.7 shows the condition number of the P3DVar system against lengthscale using the SOAR background error correlation matrix (5.11). As in the Gaussian case there is a slow linear increase in the condition number as we increase the correlation lengthscale as expected. Comparison with the unpreconditioned system in Figure 6.2 again shows a reduction in the condition number of the P3DVar Hessian compared to the 3DVar Hessian. For instance at lengthscale $L = 3\Delta x = 0.3$ we observe that the condition number has reduced from over 500 to approximately 7. Although less dramatic than the Gaussian case this amounts to a reduction of about 2 orders of magnitude.

Finally, we consider the P3DVar system using the Laplacian correlation matrix (5.12). Figure 6.8 shows the condition number of the P3DVar Hessian as
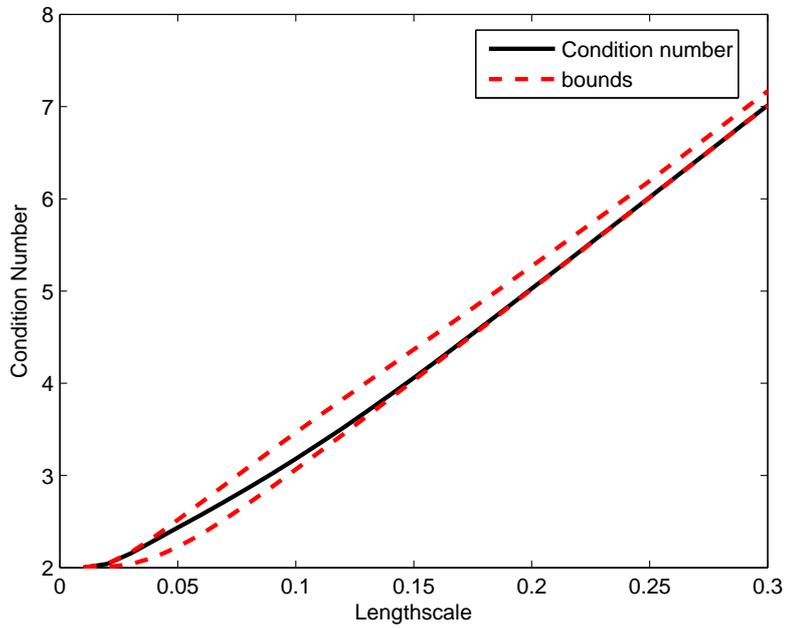
Figure 6.7: Condition number of the preconditioned Hessian (solid line) and bounds (dashed line) as function of lengthscale with SOAR background covariance matrix with observation configuration 1.
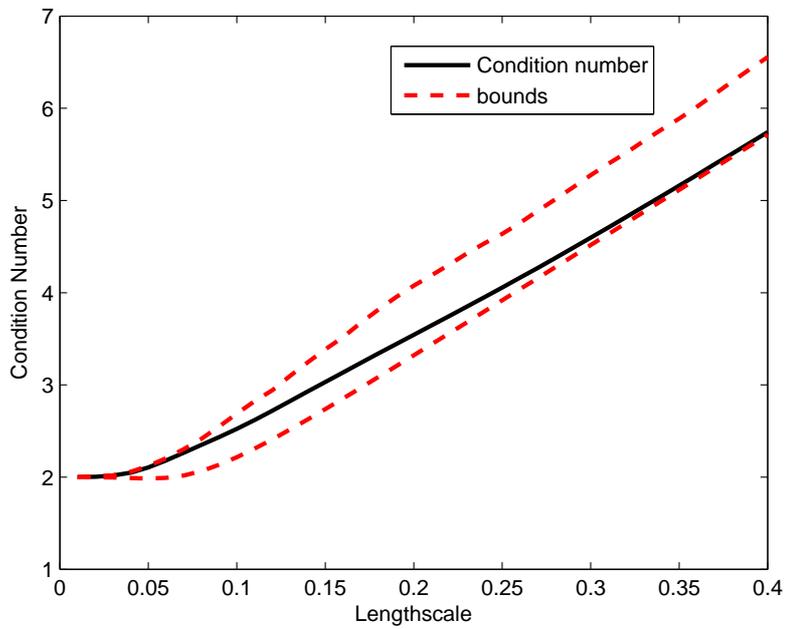
Figure 6.8: Condition number of the preconditioned Hessian (solid line) and bounds (dashed line) as function of lengthscale with Laplacian background covariance matrix with observation configuration 1.

a function of lengthscale. Again there is an increase in the condition number and bounds as the lengthscale increases.

In all cases the condition number increases in a linear fashion. Comparison of the unpreconditioned system in Figure 6.4 with the preconditioned in Figure 6.8 again shows an improvement in the condition number by applying the CVT. However, the Laplacian 3DVar Hessian has smaller reduction in the conditioning compared to the Gaussian and SOAR cases considered above. We note that in all cases the unpreconditioned system only matches the conditioning of the preconditioned system when the lengthscales are very small. By considering the upper bound on the condition number of $\hat{\mathbf{S}}$ in the experiments in this section we could predict that the conditioning of P3DVar was likely to be better than 3DVar. In fact, the upper bound (6.28) turned out to be a large overestimate of the conditioning of P3DVar in the numerical examples considered.

The bounds in Theorem 6.2.2 show that the position, number and accuracy of the observations are important factors in determining the condition number of the P3DVar Hessian. We investigate these factors in more detail in the next section.

6.3   Observations and Condition Number

In the previous section we found theoretical bounds on the condition number of the P3DVar Hessian (6.18) and showed that applying the CVT improves the conditioning of the 3DVar Hessian. In this section we consider some of the major factors that affect the conditioning of the P3DVar cost function as indicated by the theoretical bounds derived in Section 6.2.1. We begin by considering the effect of the observation accuracy on the condition number of (6.18).

6.3.1 Observation Accuracy

Obs Variance	Condition Number		
	Gaussian	SOAR	Laplacian
0.01	264.39	403.64	255.31
0.05	53.68	81.53	51.86
0.10	27.34	41.26	26.43
0.50	6.27	9.05	6.09
1.00	3.63	5.03	3.54
2.00	2.32	3.01	2.27
5.00	1.53	1.81	1.51
10.00	1.26	1.40	1.25

Table 6.1: Change in the condition number of the preconditioned Hessian with change in the observation error variance using observation configuration 1.

Here we focus on the effect of observation accuracy on the conditioning of the preconditioned 3DVar Hessian. Under the assumptions of Theorem 6.2.2 the bounds show that the conditioning of P3DVar is linearly related to the inverse of the observation error variance σ_o^2. In particular, assuming the Hessian satisfies the assumptions of Theorem 6.2.2, the condition number of the Hessian can be written explicitly as

$$\kappa(\hat{\mathbf{S}}) = \lambda_{\max}(\hat{\mathbf{S}}) = 1 + \frac{\sigma_b^2}{\sigma_o^2}\lambda_{\max}(\mathbf{HCH}^T). \tag{6.30}$$

If σ_b^2 is fixed then (6.30) implies that the preconditioned system will become more ill-conditioned as the accuracy of the observations increases (i.e. σ_o^2 decreases). We now perform a simple experiment to illustrate the relationship between the observation error variance and the condition number of $\hat{\mathbf{S}}$. A one parameter periodic system is defined on a domain of $N = 500$ grid points with $\Delta x = 0.1$ grid spacing. The correlation lengthscale is now fixed at $L = 2\Delta x = 0.2$ and the background variance is fixed at $\sigma_b^2 = 1$. The background covariance matrix is equal to $\mathbf{B} = \sigma_b^2\mathbf{C} = \mathbf{C}$, where $\mathbf{C}$ is one of the Gaussian, SOAR or Laplacian correlation matrices defined in Section 5.2. As in the previous section 250 observations are made at grid points with the same distribution as observation configuration 1 as introduced in Section 6.1.2. With these assumptions Theorem 6.2.2 applies and the Hessian $\hat{\mathbf{S}}$

satisfies the bounds (6.25).

The effect of the observation accuracy on the condition number can be seen in Table 6.1. The table shows the condition number is linearly related to the inverse of the observation error variance. For instance, the effect of increasing the observation variance from 0.01 to 0.1, using each of the different background covariance matrices, roughly decreases the condition number by a factor of 10. Next we look at the impact of the observation distribution.

6.3.2   Observation Spacing and Thinning

The bounds on the condition number of the P3DVar Hessian also depend on the sum of elements of the reduced correlation matrix $\mathbf{HCH}^T$. In this section we examine this matrix more closely. As we saw in Theorem 6.2.2 the lower bound on the condition depends on the average row sum $\frac{1}{p}\sum_{i,j \in K} c_{i,j}$ of the reduced correlation matrix and the upper bound depends on the maximum absolute row sum, $||\mathbf{HCH}^T||_\infty$, of the reduced correlation matrix. Thus the position of the observations, the size of the correlations between errors and the number of observations all play a part in the conditioning of P3DVar. We discuss each of these scenarios in turn.

Firstly, we consider the positioning of observations and correlation length-scale. In this section we assume that our Hessian $\mathbf{\hat{S}}$ satisfies the assumptions of Theorem 6.2.2. With many correlation matrices we would expect the error correlation $c_{i,j}$ between two points at position i and j to reduce the larger the distance $|i - j|$ between the errors. In contrast, the closer the points are the more correlated the errors are. In particular, we saw in chapter 5 three examples of matrices which have this property, namely the Gaussian, the SOAR and the Laplacian correlation matrices. To begin with we analyse the Gaussian and SOAR correlation matrices, (5.7) and (5.8) respectively, introduced in Section 5.2. Both these matrices only contain positive components. Since we are observing only at grid points, $\mathbf{HCH}^T$ is the original correla-

tion matrix except with rows and columns deleted at the positions which are unobserved. Since the correlation decreases the farther two points are positioned from each other then the components of $\mathbf{HCH}^T$ will be smaller the farther the observations are separated. Hence from the bounds (6.25) this implies that the sum of the elements in $\mathbf{HCH}^T$ and therefore, the condition number of the P3DVar Hessian, will be smaller as the distance between observations is increased.

We now consider the Laplacian correlation matrix (5.12). The correlation structure of this matrix is similar to the Gaussian except, unlike the Gaussian and SOAR correlation matrices, the Laplacian correlation matrix contains negative components as shown in Figure 5.1. Due to the negative correlations the magnitude of the components $c_{i,j}$ of the Laplacian correlation matrix are not monotonically decreasing with distance $|i - j|$ and therefore are not as simple to analyse as the SOAR and Gaussian correlation matrices. However if the observations are within a distance before $c_{i,j}$ become negative then the components of $\mathbf{HCH}^T$ will decrease monotonically. Hence, the farther the observations are the larger the reduction in the size of the bounds in (6.25). Hence an increase in the separation of the observations may result in a reduction in the condition number of $\hat{\mathbf{S}}$.

We also consider the effect of correlation lengthscale. As we increase the lengthscale the size of the elements of $\mathbf{C}$ increase. Hence with a fixed observation distribution increasing the lengthscale will increase the size of the elements in $\mathbf{HCH}^T$. We would then expect a larger condition number of $\hat{\mathbf{S}}$ the larger the correlation lengthscale.

We illustrate the effect of the spacing between observation and the error correlation of the background errors experimentally using a one-parameter one-dimensional, periodic system defined on an equally space grid of $N = 500$ grid points and a grid spacing of $\Delta x = 0.1$. We fix the observation and background error variances at $\sigma_o^1 = \sigma_b^2 = 1$. We choose a background error covariance matrix of the form $\mathbf{B} = \sigma_b^2 \mathbf{C}$ where $\mathbf{C}$ is the Gaussian, SOAR or
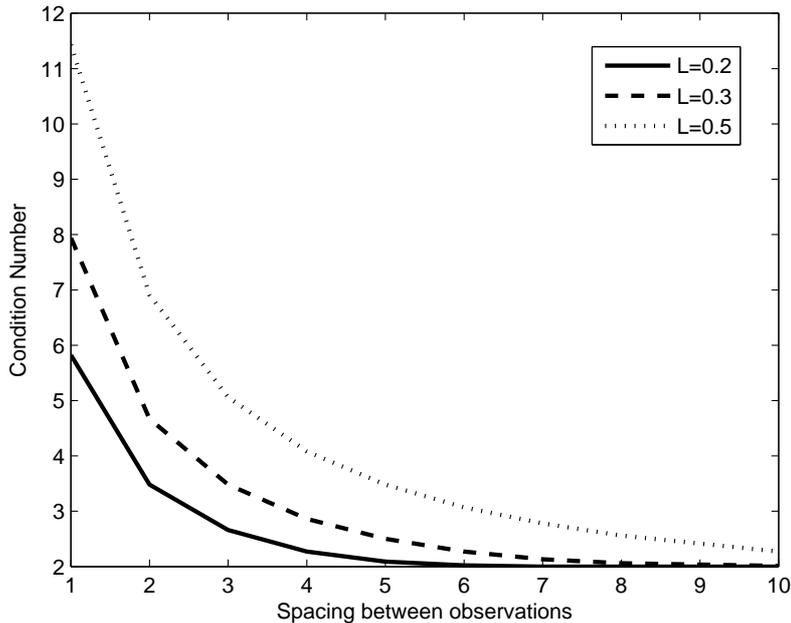
Figure 6.9: Condition Number of the preconditioned Hessian of single paramter system with observation spacing using the Gaussian background covariance matrix for different lengthscales. The solid line represents lengthscale $L = 0.2$, the dashed line represents a lengthscale $L = 0.3$ and a dotted line represents a lengthscale of $L = 0.5$.

Laplacian correlation matrices described in Section 5.2. We begin with $p = 20$ observations positioned on the first 20 grid points and calculate the condition number of the $\hat{\mathbf{S}}$ as the spacing between the observations is increased. The experiment was repeated for lengthscales $L = 0.2, 0.3$ and 0.5.

Figures 6.9, 6.10 and 6.11 show the condition number of $\hat{\mathbf{S}}$ constructed using the Gaussian, SOAR and Laplacian background covariance matrices respectively. In all three versions the results are as expected. The larger the spacing between observations the smaller the condition number. Additionally, the larger the error correlation lengthscale the larger the condition number for each fixed observation distribution. Both the Gaussian and Laplacian produce Hessians whose condition numbers have similar magnitude. This can be explained by the close approximation of the Laplacian correlation model with the Gaussian as seen in Figure 5.1. The comparison with the SOAR and Gaussian condition numbers in Figures 6.9 and 6.10 show that Hessian constructed using the SOAR has a larger condition number then the Gaus-
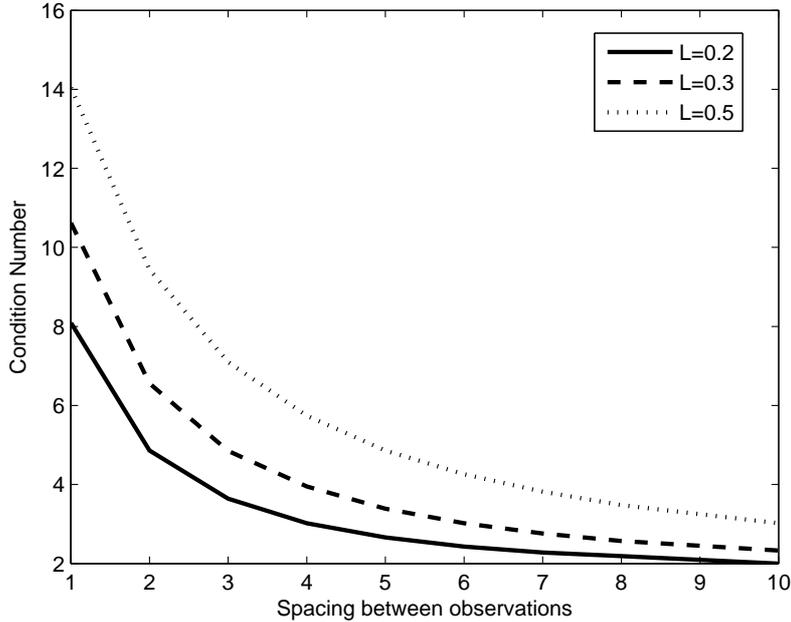
Figure 6.10: Condition Number of the preconditioned Hessian of single parameter system with observation spacing using the SOAR background covariance matrix for different lengthscales. The solid line represents lengthscale $L = 0.2$, the dashed line represents a lengthscale $L = 0.3$ and a dotted line represents a lengthscale of $L = 0.5$.

sian for the same lengthscale and observation spacing. However, both have a condition number of the same order of magnitude. This can be explained by Figure 5.1 which shows, for a fixed lengthscale, all components of the different $\mathbf{C}$ matrices are of similar size.

For the Laplacian case in Figure 6.11 there is a small increase in the condition number as the separation between the observations increases to 7 grid points. This could be explained by the bounds. Figure 5.1 shows that the Laplacian correlation function has negative correlations in the tails and therefore the magnitude of the coefficients $|c_{i,j}|$ of the Laplacian increase as the separation increases. This will result in an increase in the upper bound as negative coefficients are summed as the spacing increases. This could explain the small increase in the condition number of the Hessian shown in Figure 6.11.

Finally we consider the effect of removing observations on the condition number of $\hat{\mathbf{S}}$. As explained in Section 6.2.2 the reduced correlation matrix is the
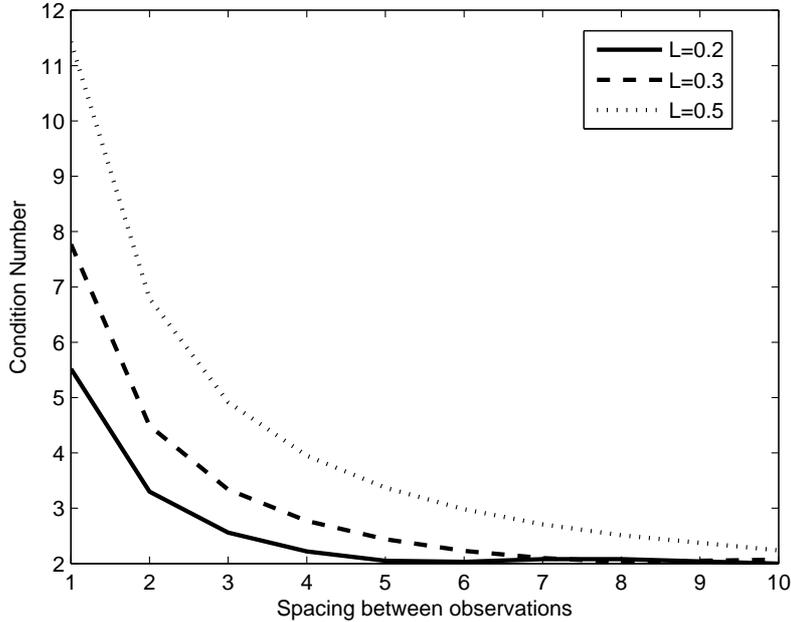
Figure 6.11: Condition Number of the preconditioned Hessian of single parameter system with observation spacing using the Laplacian background covariance matrix for different lengthscales. The solid line represents lengthscale $L = 0.2$, the dashed line represents a lengthscale $L = 0.3$ and a dotted line represents a lengthscale of $L = 0.5$.

original correlation matrix except with rows and columns deleted at the positions no observations are made. The bounds on the condition number of the P3DVar Hessian (6.25) will be smaller the fewer the observations since fewer correlation terms are summed. We can illustrate this using a one-parameter, one-dimensional periodic experiment. Still using the $N = 500$ grid point domain with $\Delta x = 0.1$ grid spacing and observation and background error variances fixed at $\sigma_o^2 = \sigma_b^2 = 1$ we remove the observations and analyse the effect on the condition number. We begin by observing every 2nd grid point and then every $10^{th}, 50^{th}$ and 250th hence we are observing $250, 50, 10$ and 2 observations in each case respectively. The condition number of the Hessians using all three choices of the background covariance matrix, Gaussian, SOAR and Laplacian, with thinning of the observations are shown in Table 6.2 for a fixed lengthscale of $L = 0.5$.

The table confirms our hypothesis. As the thinning is increased the condition number reduces. In the next section we consider the impact of the condition

	Gaussian	SOAR	Laplacian
No. Observations	Cond No.	Cond No.	Cond No.
250	7.27	11.02	6.91
50	2.72	3.04	2.24
10	2.00	2.00	2.00
2	2.00	2.00	2.00

Table 6.2: Effect of thinning the observation on the condition number of the preconditioned Hessian using a correlation lengthscale $L = 0.5$.

number on the convergence rate of an iterative solver.

6.4   Convergence results

In Chapter 3 we showed that the condition number is an important criterion for determining the convergence rate of some iterative solvers. The convergence rate of the iterative solvers applied to a linear system may be slower the more ill-conditioned the matrix is. For the majority of this chapter we have been analysing the condition number of the preconditioned and unpreconditioned Var schemes. In this section we now compare this to the convergence rate of the conjugate gradient (CG) method applied to solving the linear system (3.12) associated to the 3DVar cost function. We begin by considering the unpreconditioned 3DVar system.

6.4.1   Convergence Rate of 3DVar

In this section we consider the convergence rate of the CG method, introduced in Section 3.3.1, to find the solution, $\mathbf{w}$, of a linear system

$$\mathbf{Sw} = \mathbf{b}, \tag{6.31}$$

where $\mathbf{S}$ is the Hessian (6.2) of the 3DVar cost function. The linear system is solved using the Matlab CG method pcg.m and is equivalent to the CG
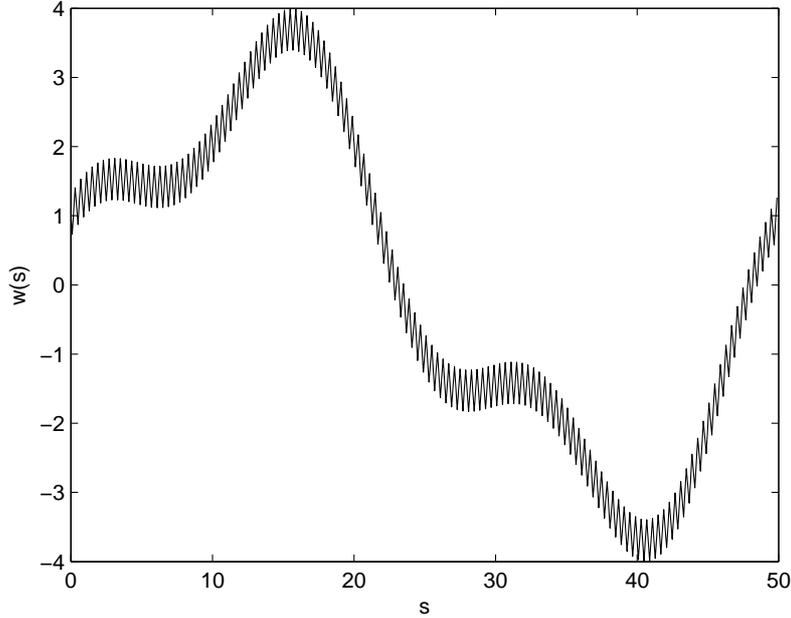
Figure 6.12: Function used to define the true solution for the CG experiments.

algorithm given in Section 3.3.1. In the experiments in this section the CG algorithm stops when either a pre-defined number of iterations have been computed (in our case 2000) or by a sufficient reduction in the relative residual

$$\frac{||\mathbf{b} - \mathbf{S}\mathbf{w}_k||}{||\mathbf{b}||} = \frac{||\mathbf{r}_k||}{||\mathbf{b}||}, \tag{6.32}$$

where $\mathbf{w}_k$ is the estimate to the true solution to (6.31) found from the kth iteration of CG. In this section we require the relative residual to reduce below 10^{-6} for the CG method to stop.

To perform the CG experiments we consider a single parameter, periodic system on the real line. We choose $N = 500$, equally spaced grid points with spacing $\Delta x = 0.1$. The background error covariance matrix is of the form $\mathbf{B} = \sigma_b^2 \mathbf{C}$ where we choose the background error variance $\sigma_b^2 = 1$ and $\mathbf{C}$ is either the Gaussian, SOAR or Laplacian correlation matrices on the circle introduced in Section 5.2. We choose observations with error variances $\sigma_o^2 = 1$. In this section we use observation configuration 1, used in Section 6.1.2, which consists of 250 randomly distributed observations made at grid

points. With these assumptions the Hessian, $\mathbf{S}$, satisfies the bounds given in Theorem 6.1.2.

For the true solution we choose

$$\mathbf{w}^*(s) = 2\sin(qs) + \cos(3qs) - 0.3\sin(125qs), \qquad (6.33)$$

defined at the grid points $s = k\Delta x$ for $k = 0, 1, \ldots, (N-1)$ where $q = \frac{2\pi}{N\Delta x}$. The true solution is displayed in Figure 6.12 and clearly shows (6.33) is a mixture of large and small scales.

We will also compute the relative error in the solution to test that our system has converged. This is simply the ratio of the magnitudes of the error and the truth

$$\frac{||\mathbf{w}^* - \mathbf{w}^a||}{||\mathbf{w}^*||}, \qquad (6.34)$$

where $\mathbf{w}^a$ represents the approximate solution found by the CG solver. We will refer to this relative error as the *accuracy* throughout the remainder of this chapter.

In Section 6.1.2 we considered the condition number of the unpreconditioned 3DVar system. In the numerical experiments in Section 6.1.2 we constructed three different Hessians using different background covariance matrices and showed that in all cases the condition number of 3DVar was sensitive to the lengthscales. The larger the correlation lengthscale on the background errors the larger the condition number of the 3DVar Hessian. According to the discussion in Section 3.3.1 we would expect the CG method to converge more slowly for an ill-conditioned system than one with a smaller condition number. Therefore we would expect the rate of convergence of the CG method to be slower the larger the correlation lengthscale of the background errors.

To begin with we consider the Hessian constructed using the Gaussian correlation matrix (5.10). Table 6.3 shows the number of iterations required for the CG method to converge for several different lengthscales. As expected

Lengthscale	Cond No.	Iterations	Accuracy
0.02	2.00	2	6.4041e-007
0.04	2.18	6	5.9725e-007
0.06	3.60	11	3.3604e-007
0.08	1.05e+001	20	3.2377e-007
0.10	5.15e+001	45	2.9223e-007
0.12	3.92e+002	118	2.1064e-007
0.14	4.44e+003	340	1.7969e-007
0.16	7.48e+004	1123	1.5188e-007

Table 6.3: Convergence rates of CG against lengthscale using the Unpreconditioned Hessian with the Gaussian correlation matrix

with smaller lengthscales the number of iterations is relatively small and the relative accuracy of the final solution is good being an order of 10^{-7} this can be explained by the small condition number when the lengthscales is small. However, as the lengthscales increase the number of iterations required to solve the system increases rapidly due to the rapid increase in condition number (See Figure 6.1). For lengthscales larger than $L = 0.16$ the CG method reached the maximum number of 2000 iterations without converging. This can be explained by the ill-conditioning of the Hessian. Since $\mathbf{S}$ is very ill-conditioned for lengthscales above $L = 0.16$ (greater then 10^5) the system is very sensitive to round-off errors and results in the CG method not converging.

We repeated the same experiments using the Hessians constructed using the SOAR (5.11) and Laplacian (5.12) background error covariance matrix. In these cases and for all lengthscales considered the CG method fully converged to the solution giving an accuracy of order 10^{-7}. The convergence rates for both Hessians are shown in Figure 6.13. The figure shows that the convergence becomes slower as the lengthscale increases. This is as expected since the condition numbers of the Hessian increases as the lengthscale increases as shown in Figures 6.2 and 6.4. Additionally, the conditioning of the Hessians was, in general, better when the SOAR and the Laplacian were used in comparison to the Gaussian and this is shown by the faster convergence rates of the solver for the SOAR and Laplacian cases. For instance at lengthscale
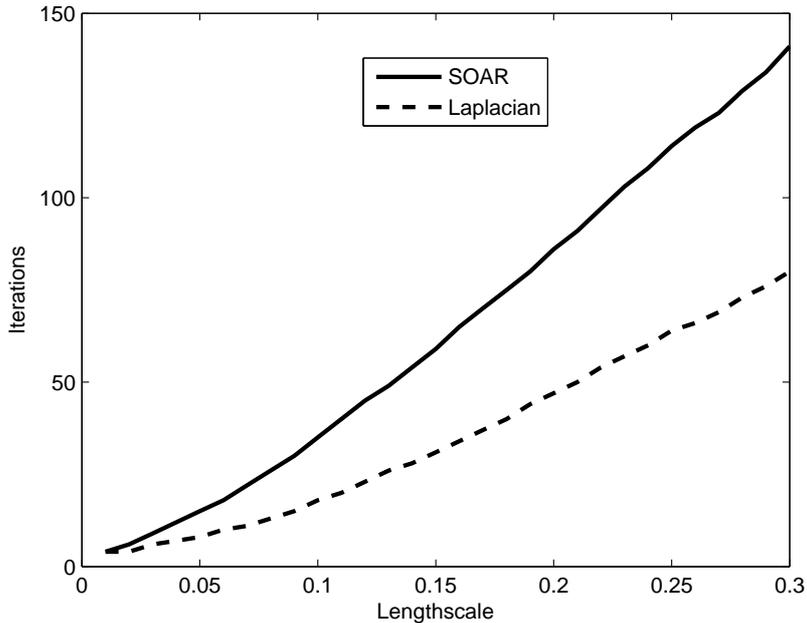
Figure 6.13: The effects of different lengthscales on the convergence of CG for unpreconditioned 3DVar for the SOAR correlation matrix (solid line) and the Laplacian correlation matrix (dashed line).

$L = 0.12$ the CG method takes 118 iterations to converge to the solution in the Gaussian case compared to the SOAR and Laplacian cases which take under 50 iterations as shown by comparing Table 6.3 with Figure 6.13. We now turn our attention to the convergence rate of the CG method applied to the preconditioned system.

6.4.2 Convergence Rate of P3DVar

In this section we consider the convergence rate of the CG method, applied to the preconditioned Var system. Minimising the preconditioned 3DVar scheme is equivalent to solving the linear system

$$\hat{\mathbf{S}}\mathbf{v} = \hat{\mathbf{b}}, \tag{6.35}$$

for $\mathbf{v}$ where $\hat{\mathbf{S}} = \mathbf{I} + \mathbf{B}^{1/2}\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{B}^{1/2}$ is the Hessian of the preconditioned 3DVar cost function (2.16) introduced in Section 2.4, $\mathbf{v} = \mathbf{B}^{-1/2}\mathbf{w}$ and $\hat{\mathbf{b}} =$

$\mathbf{B}^{1/2}\mathbf{b}$. Once the estimate to the solution, $\mathbf{v}_a$, of (6.35) is found the estimate of the original linear system (6.31) can be found by transforming back to the original variables, $\mathbf{w}_a = \mathbf{B}^{1/2}\mathbf{v}_a$.

We solve (6.35) using the Matlab CG routine pcg.m which uses a preconditioned CG method equivalent to the algorithm given in Section 3.3.2. The CG algorithm stops when either the maximum number of iterations of the method have been completed (2000 in this Section) or when the relative residual (6.32) has been reduced below 10^{-6}. We note that the residual of the unpreconditioned system is used as the stopping criteria instead of the residuals of the preconditioned system (6.35). This is because of the form of the preconditioned CG algorithm. Instead of solving (6.35) with the standard CG method as given in Section 3.3.1 an equivalent form is used (given in Section 3.3.2) which is in terms of the solution to the unpreconditioned system (6.31).

To test the performance of CG on the preconditioned Var scheme we consider a single parameter, periodic system on the real line. We choose $N = 500$, equally spaced grid points with spacing $\Delta x = 0.1$. The background error covariance matrix $\mathbf{B} = \sigma_b^2 \mathbf{C}$ where $\sigma_b^2 = 1$ and $\mathbf{C}$ is one of either the Gaussian, SOAR or Laplacian correlation matrices on the circle introduced in Section 5.2. We choose observations with error variances σ_o^2. As in the previous Section, we use observation configuration 1, which consists of 250 randomly distributed observations made at grid points. With these criteria the Hessian $\hat{\mathbf{S}}$, satisfies the bounds given in Theorem 6.2.2.

The true solution for the preconditioned system is $\mathbf{v}^* = \mathbf{B}^{-1/2}\mathbf{w}^*$ where $\mathbf{w}^*$ is (6.33) defined at the grid points. We measure the accuracy of the final estimate, $\mathbf{w}^a = \mathbf{B}^{1/2}\mathbf{v}_a$, using the relative error (6.34).

In Section 6.2 we showed preconditioning 3DVar via the CVT produces a large reduction in the condition number of $\hat{\mathbf{S}}$ compared to the condition number of $\mathbf{S}$. We therefore expect the convergence rate of the CG method

applied to the preconditioned system (6.35) to be much quicker than solving the original problem (6.31).

Lengthscale	Cond No.	Iterations	Accuracy
0.02	2.000	2	4.5063e-007
0.04	2.044	4	2.1813e-010
0.06	2.249	5	1.4359e-008
0.08	2.462	6	2.4823e-007
0.10	2.641	8	1.1140e-007
0.12	2.819	9	1.9990e-007
0.14	3.010	10	1.5867e-007
0.16	3.213	11	1.3690e-007
0.18	3.421	11	3.0299e-007
0.20	3.634	11	5.5338e-007
0.22	3.852	11	7.0100e-007

Table 6.4: Convergence rate of CG against lengthscale to solve the preconditioned Hessian using the Gaussian correlation matrix.

We begin by considering the preconditioned Gaussian Hessian. The results for selected lengthscales are shown in Table 6.4. It is apparent from the Table 6.4 that the convergence rate has been substantially reduced by using the preconditioning. For small lengthscales (≤ 0.1) both the unpreconditioned and preconditioned systems have small condition numbers and therefore a small iteration count. As the lengthscale is increased the number of iterations required for convergence of P3DVar only gradually increases compared to that of 3DVar. The accuracy of the solution for the P3DVar system remains good with an order of 10^{-7} for larger lengthscales. Hence, a reduction in the condition number via the CVT has resulted in a faster convergence of the CG algorithm. However, for lengthscales above $L = 0.25$ the CG method began to stagnate (two consecutive iterates were the same) without achieving the desired reduction in the relative residual. Additionally, for lengthscales above $L = 0.28$ it was found that the CG method failed to converge within 2000 iterations. Even though the condition number of $\hat{\mathbf{S}}$ is small, at approximately 4, the CG method is failing to converge. This poor performance of the algorithm can be explained by the fact that in each iteration of the preconditioned CG method a vector $\mathbf{z}$ must be found using the preconditioner as seen in equation (3.41). In our case the preconditioner is $\mathbf{P} = \mathbf{B}$ so we

must solve the equation

$$\mathbf{B}^{-1}\mathbf{z}_{k+1} = \mathbf{r}_{k+1} \qquad (6.36)$$

in terms of $\mathbf{z}_{k+1}$. In Chapter 5 we have seen that for large lengthscales the Gaussian background covariance matrix is extremely ill-conditioned. For instance at $L = 0.25$ the condition number of the Gaussian correlation matrix is an order of magnitude of 10^{12} as seen in Figure 5.2. This implies the solution to (6.36) is likely to be solved inaccurately for large lengthscales and therefore the CG method will not converge.
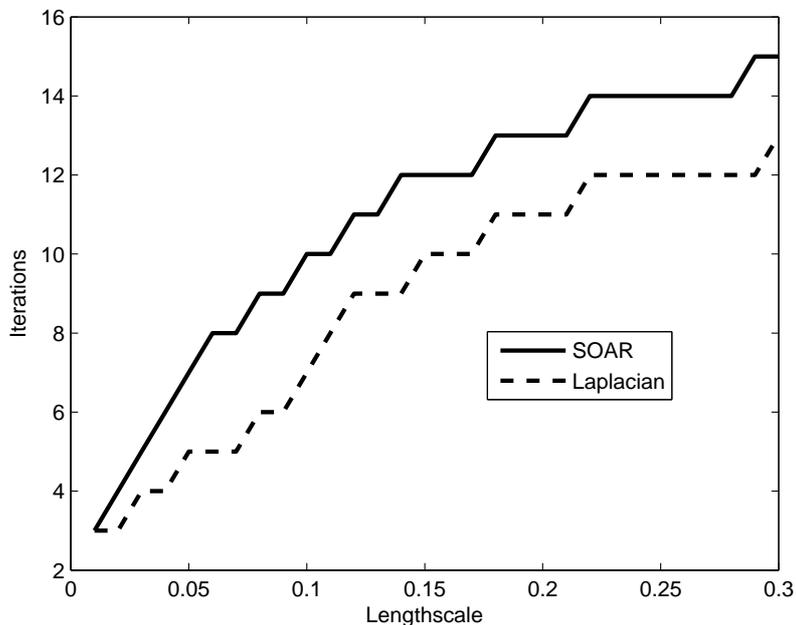
Figure 6.14: The convergence of CG for preconditioned 3DVar as a function of lengthscale for the SOAR correlation matrix (Solid line) and the Laplacian correlation matrix (dashed line).

We repeat the CG experiments on the P3DVar Hessian for the SOAR and Laplacian case. As in the unpreconditioned case the accuracy of the solutions all are of the order 10^{-7}. The convergence results are given in Figure 6.14 as a function of lengthscale. Since the condition number in both cases is proportional to the lengthscale we expect the convergence rate to decrease as the lengthscale increases. The figure shows an increase in the number of iterations required for the CG method to converge as the lengthscale increases. Preconditioning with the CVT reduces the condition number of the Hessian

in both the SOAR and Laplacian case and this explains the increase in the convergence rate compared to the unpreconditioned case (Figure 6.13). In fact there is a reduction in the number of iterations by a factor of 10.

The last two section have shown the benefit of a small condition number in order to quickly solve a linear system using the CG method. We have also seen the importance of the condition number of the preconditioner itself. Using the Gaussian background error covariance we showed that even though the Hessian, $\hat{\mathbf{S}}$, has a small condition number the CG method fails to converge. We now consider the effect of observation position and accuracy on the convergence rate.

6.4.3 Convergence Rate and Observations
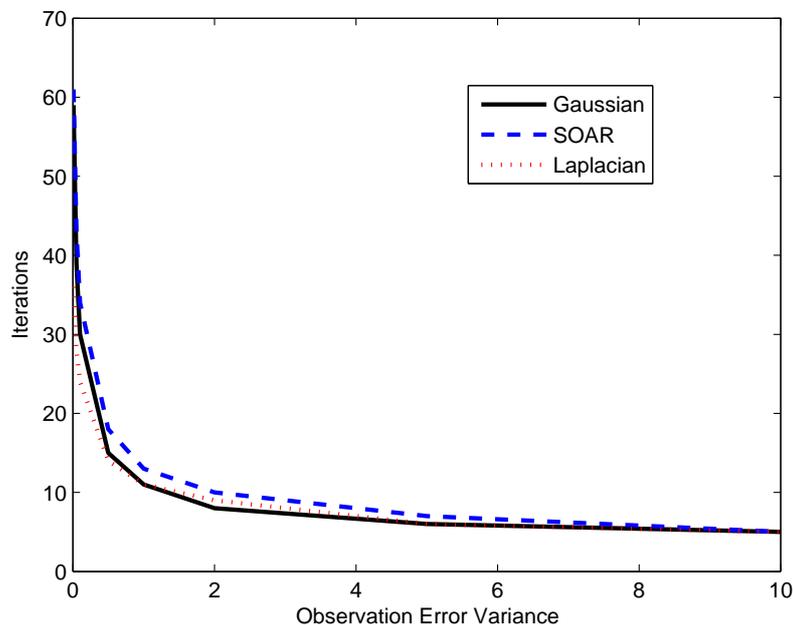
Figure 6.15: The effects of Variance on the convergence of CG for preconditioned 3DVar for the Gaussian correlation matrix (solid line), the SOAR correlation matrix (dashed line) and the Laplacian correlation matrix (dotted).

In Section 6.3 we saw that two of the main factors influencing the conditioning of the P3DVar system were the observation position and the observation

accuracy. In this section we briefly consider the impact of these elements on the convergence rate of the CG method used to solve (6.35). We test the Matlab preconditioned conjugate gradient routine, pcg.m, on a periodic one-parameter system on a one dimensional domain consisting of $N = 500$ grid points with $\Delta x = 0.1$. The observations are made at 250 grid points distributed according to observation configuration 1 used in the previous section. The background covariance matrix is of the form $\mathbf{B} = \sigma_b^2 \mathbf{C}$ with variances on the background errors are $\sigma_b^2 = 1$ and we use the Gaussian, SOAR and Laplacian correlation matrices, as defined in Section 5.2, to define $\mathbf{C}$. We consider the convergence performance of the CG method to solve the preconditioned system (6.35) as we change the observation error variance. As in the previous section the CG method stops after 2000 iterations or when the relative residual is reduced to below 10^{-6}.

Fixing the background error correlation at $L = 0.2$ we show the convergence rates as a function of σ_o^2 in Figure 6.15. All solutions had an accuracy of 10^{-7}. The solid black line shows the Hessian for the Gaussian background matrix, the blue dashed line is the SOAR and the red dotted line is the Laplacian case. As the graph shows, in all cases, the less accurate the observations the faster the convergence. This is unsurprising as this corresponds to the fact that the condition number is lower in all cases as σ_o^2 increases.

Next we consider the effect of observation spacing on the convergence rate of the CG method to solve (6.35). We now fix the observation error variance at $\sigma_o^2 = 1$ and the background error variance at σ_b^2. We take 20 observations on the first 20 grid points. We change the spacing between observations keeping the same separation between adjacent observations. The results are shown in Figures 6.16, 6.17 and 6.18 for the P3DVar Hessians using the Gaussian, SOAR and Laplacian background covariance matrices respectively. The experiments where repeated with the lengthscales $L = 0.2, 0.3$ and 0.5, except in the Gaussian case since the CG method failed to converge for $L = 0.3, 0.5$. As explained in the previous section this may be a result of the
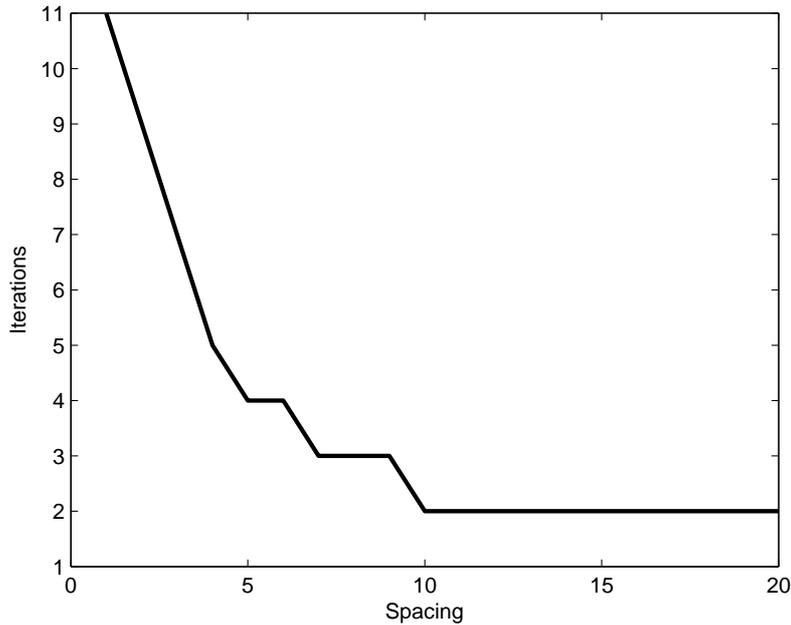
Figure 6.16: The effects of spacing on the convergence rate of CG for the preconditioned 3DVar using the Gaussian background covariance matrix for different lengthscales. The solid line represents lengthscale $L = 0.2$.

ill-conditioning of the preconditioner $\mathbf{B}$. In all cases where the CG method converged the figures show that as the spacing between the observations increases the convergence rate increases. This is as expected since as the spacing increases the condition number of $\hat{\mathbf{S}}$ decreases. Additionally, since for a larger lengthscale the condition number of all the Hessians increases then, there is a general increase in the number of iterations required for the CG solver to converge as the lengthscale increases.

6.5 Summary

In this chapter we have presented a first analysis of the conditioning of 3DVar. We have derived new theoretical bounds on the condition number of both the unpreconditioned and preconditioned Hessians of the 3DVar cost function. Using these bounds we have been able to identify key features that determine the conditioning of the system and have tested these factors in an experimen-
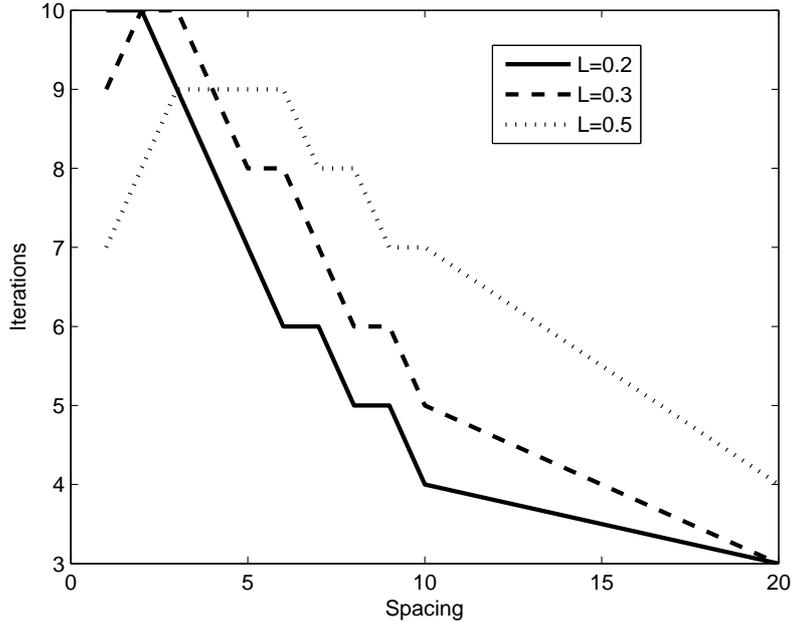
Figure 6.17: The effects of spacing on the convergence rate of CG for the preconditioned 3DVar using the SOAR background covariance matrix for different lengthscales. The solid line represents lengthscale $L = 0.2$, the dashed line represents a lengthscale $L = 0.3$ and a dotted line represents a lengthscale of $L = 0.5$.

tal setting. In the unpreconditioned system we found that the background covariance matrix can make a significant contribution to the condition number. This hypothesis was tested experimentally and it was confirmed that if the background matrix is ill-conditioned then so is the 3DVar Hessian. These results confirm previous experimental results, summarised in Section 4.1, which suggested that 3DVar is ill-conditioned due to the presence of an ill-conditioned background error covariance matrix [42]. Additionally, we found that the Hessian using the Gaussian background matrix was in general significantly more ill-conditioned then the Hessians using the Laplacian or SOAR background matrices.

In the preconditioned Var system we found that the condition number was dependent on the observation error variances and on the sum of the coefficients of a reduced background error correlation matrix. We showed experimentally that the preconditioning can reduce the condition number by several orders of magnitude. In addition, we identified that the closer the observations were
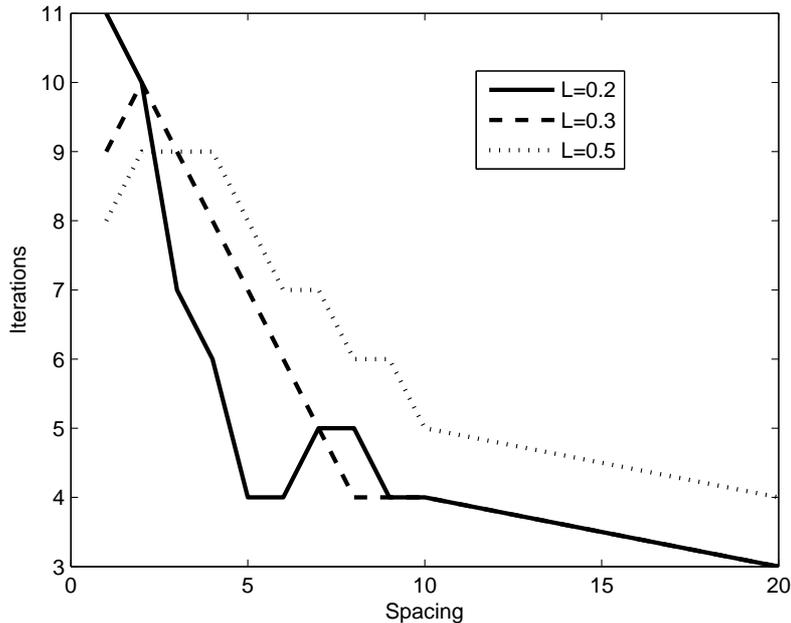
Figure 6.18: The effects of spacing on the convergence of CG for preconditioned 3DVar with the Laplacian background covariance matrix for different lengthscales. The solid line represents lengthscale $L = 0.2$, the dashed line represents a lengthscale $L = 0.3$ and a dotted line represents a lengthscale of $L = 0.5$.

positioned, the more observations we made and the more accurate the observations were, the larger the condition number of the P3DVar Hessian. These results seem counter-intuitive. According to these results, by having more accurate observations and more observations, we solve the P3DVar problem less accurately (as indicated by the larger condition number). This can be explained by the fact that a larger number of observations and more accurate observations put tighter restrictions on the optimization problem and it becomes more difficult to solve the problem.

Finally, we examined the convergence rate of the conjugate gradient method applied to our preconditioned and unpreconditioned Var systems. We found that the results were as expected with a larger condition number leading to a more slowly converging CG method. In particular, we found that the preconditioned system outperformed the unpreconditioned system in terms of number of iterations. In addition, we found that by increasing the observation error variance and increasing the spacing between the observations that a

faster convergence of the CG method was obtained when applied to P3DVar. We also found that preconditioning should be applied with some caution. In CG experiments for solving the P3DVar system we found the algorithm failed to converge for large lengthscales. The preconditioned version of the CG method requires solving an equation of the form $\mathbf{P}^{-1}\mathbf{z} = \mathbf{r}$ at each iteration where $\mathbf{P}$ is the preconditioner. In our case the preconditioner is the very ill-conditioned Gaussian background error covariance matrix and hence we expect $\mathbf{z}$ to found inaccuarately. This could account for the lack of convergence in our experiments. In the next chapter we consider the conditioning of a theoretical 4DVar model.

Chapter 7

Conditioning of 4DVar

In Chapter 6 we considered the conditioning of 3DVar in a theoretical setting. The theory highlighted key factors which affect the condition number of both the unpreconditioned (6.2) and preconditioned (6.18) 3DVar Hessians. We illustrated the theory experimentally using a one-parameter 1D periodic system. In this chapter we extend these results to the 4DVar Hessian (2.12) by incorporating a forecast model into our system. We derive new theoretical bounds on the condition number of both the unpreconditioned and preconditioned Hessians of the 4DVar cost function (2.9). Using the bounds we show

- The conditioning of the 4DVar Hessian is dependent on the conditioning of the background error covariance matrix. Hence, for an ill-conditioned background error covariance matrix the Var Hessian is also ill-conditioned.

- In a one parameter, periodic system that preconditioning with the control variable transform (CVT), as described in Section 2.4, reduces the condition number of Var.

- The condition number of the preconditioned Var system is reduced by three main factors: reducing the number of observations, using less

accurate observations and increasing the spacing between observations.

- The convergence rate of the conjugate gradient (CG) method applied to Var is increased when we precondition with the CVT.

- The convergence rate of the CG method applied to the preconditioned 4DVar system increases when we increase the spacing between observations and decrease the observation accuracy.

We begin by deriving new theoretical bounds on the condition number of the 4DVar Hessian. In 4DVar we require a forecast model. In this chapter we use the 1D linear-advection model. We derive theoretical bounds on the condition number in the specific case when the forecast model is the linear-advection model. We show that the bounds on the condition number of the 4DVar Hessian are dependent on the condition number of the background error covariance matrix and we illustrate this using numerical experiments on a periodic, single parameter on a one dimensional domain.

We derive new theoretical bounds on the condition number of the preconditioned system and examine the condition number with numerical experiments on our 1D periodic system. Our new bounds on the preconditioned system show that the condition number of the 4DVar Hessian is dependent on the *4D background error covariance matrix*. We briefly introduce this matrix and consider its structure in the case of the linear advection equation. Next we consider the effect of the observation error variances and distribution on the condition number of the preconditioned Hessian. As with the preconditioned 3DVar scheme we show, using numerical experiments on a one-parameter periodic system, that reducing the accuracy of the observations and thinning and spacing the observations of the parameter can reduce the condition number of the preconditioned 4DVar Hessian. Finally we show the importance of the condition number in terms of the convergence rate of gradient methods used to solve the 4DVar systems with numerical experiments using the conjugate gradient method. We begin by considering the conditioning of

4DVar without any preconditioning.

7.1 Conditioning of 4DVar

7.1.1 Theory

In this section we investigate the conditioning of 4DVar theoretically. Recall from Section 2.3.1 that we can write the incremental 4DVar Hessian in the compact form

$$\mathbf{S} = \tilde{J}'' = \mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}}, \tag{7.1}$$

where

$$\hat{\mathbf{H}} = \left[\mathbf{H}_0^T, (\mathbf{H}_1 \mathbf{M}_1)^T, \dots, (\mathbf{H}_n \mathbf{M}_n \dots \mathbf{M}_1)^T \right]^T, \tag{7.2}$$

and $\hat{\mathbf{R}}$ is the block diagonal matrix with the error covariance matrices of the observations at time t_k, $\mathbf{R}_k \in \mathbb{R}^{q_k \times q_k}$, down the diagonal for $k = 0, \dots n$.

In the following theorem we extend the theoretical bounds on the condition number of the 3DVar Hessian derived in Theorem 6.1.1 to the 4DVar Hessian (7.1).

Theorem 7.1.1 *Let $\mathbf{B} \in \mathbb{R}^{N \times N}$ be the background error covariance matrix and $\mathbf{R}_j \in \mathbb{R}^{q_j \times q_j}$ be the observation error covariance matrix at time t_j for $j = 0, 1 \dots n$. Let $\hat{\mathbf{R}} \in \mathbb{R}^{Q \times Q}$ be the block diagonal matrix consisting of the blocks $\mathbf{R}_j$ down the diagonal where $Q = \sum_{j=0}^{n} q_j$ and $Q < N$. In addition, define $\hat{\mathbf{H}} \in \mathbb{R}^{Q \times N}$ to be the generalised observation operator (7.2). Then the following bounds hold on the condition number of $\mathbf{S} = \mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}}$*

$$\frac{\kappa(\mathbf{B})}{(1 + \lambda_{\max}(\mathbf{B})\lambda_{\max}(\hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}}))} \leq \kappa(\mathbf{S}) \leq \kappa(\mathbf{B})(1 + \lambda_{\min}(\mathbf{B})\lambda_{\max}(\hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}})), \tag{7.3}$$

where, $\lambda_{\max}(\mathbf{A})$ and $\lambda_{\min}(\mathbf{A})$ represent the largest and smallest eigenvalue of the matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ respectively.

Proof We put basic bounds on the spectral condition number of (7.1) by putting bounds on the maximum and minimum eigenvalues of $\mathbf{S}$. Since $Q < N$, $\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}} \in \mathbb{R}^{N \times N}$ has a minimum eigenvalue equal to zero. Thus using Theorem 3.4.1 we have the following bounds on the maximum eigenvalue

$$\lambda_{\max}(\mathbf{B}^{-1}) \leq \lambda_{\max}(\mathbf{S}) \leq \lambda_{\max}(\mathbf{B}^{-1}) + \lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}), \qquad (7.4)$$

and the minimum eigenvalue

$$\lambda_{\min}(\mathbf{B}^{-1}) \leq \lambda_{\min}(\mathbf{S}) \leq \lambda_{\min}(\mathbf{B}^{-1}) + \lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}). \qquad (7.5)$$

Combining (7.4) and (7.5) we obtain the following bounds on the condition number of $\mathbf{S}$

$$\frac{\lambda_{\max}(\mathbf{B}^{-1})}{\lambda_{\min}(\mathbf{B}^{-1}) + \lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}})} \leq \kappa(\mathbf{S}) \leq \frac{\lambda_{\max}(\mathbf{B}^{-1}) + \lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}})}{\lambda_{\min}(\mathbf{B}^{-1})}.$$
$$(7.6)$$

First, consider the lower bound

$$\kappa(\mathbf{S}) \geq \frac{\lambda_{\max}(\mathbf{B}^{-1})}{\lambda_{\min}(\mathbf{B}^{-1})}(1 + (\lambda_{\min}(\mathbf{B}^{-1}))^{-1}\lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}))^{-1}. \qquad (7.7)$$

Since $\kappa(\mathbf{B}) = \kappa(\mathbf{B}^{-1}) = \frac{\lambda_{\max}(\mathbf{B}^{-1})}{\lambda_{\min}(\mathbf{B}^{-1})}$ and $(\lambda_{\min}(\mathbf{B}^{-1}))^{-1} = \lambda_{\max}(\mathbf{B})$ we obtain

$$\kappa(\mathbf{S}) \geq \kappa(\mathbf{B})(1 + \lambda_{\max}(\mathbf{B})\lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}))^{-1}. \qquad (7.8)$$

Next consider the upper bound in (7.6)

$$\kappa(\mathbf{S}) \leq \frac{\lambda_{\max}(\mathbf{B}^{-1})}{\lambda_{\min}(\mathbf{B}^{-1})}(1 + (\lambda_{\max}(\mathbf{B}^{-1}))^{-1}\lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}})). \qquad (7.9)$$

Since $(\lambda_{\max}(\mathbf{B}^{-1}))^{-1} = \lambda_{\min}(\mathbf{B})$ we obtain

$$\kappa(\mathbf{S}) \leq \kappa(\mathbf{B})(1 + \lambda_{\min}(\mathbf{B})\lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}})), \qquad (7.10)$$

as required, which completes the proof.

The bounds presented in Theorem 7.1.1 are very general. However, in common with most NWP centres, we assume that we have fewer observations than variables in the state space $(Q < N)$. Although the bounds are very general we can infer from the presence of the $\kappa(\mathbf{B})$ term in both the upper and lower bounds that the conditioning of the background covariance matrix plays an important role in the conditioning of the Hessian. Theorem 7.1.1 shows, like Theorem 6.1.1 found for the 3DVar case, that the bounds on the condition number of the 4DVar Hessian are related to the background error covariance matrix $\mathbf{B}$. Therefore if the background error covariance matrix is ill-conditioned then we expect the 4DVar Hessian to be ill-conditioned. The effect of $\kappa(\mathbf{B})$ on $\kappa(\mathbf{S})$ was demonstrated in Section 6.1.2 in the 3DVar case with numerical experiments on a one parameter periodic system. It was confirmed that the ill-conditioned background error covariance matrix determined an ill-conditioned Hessian.

In the next theorem we extend Theorem 6.1.2 to the 4DVar case by putting tighter restrictions on the components of the Hessian to produce more precise bounds on the condition number of the 4DVar Hessian.

Theorem 7.1.2 *Let* $\mathbf{B} = \sigma_b^2 \mathbf{C} \in \mathbb{R}^{N \times N}$ *be the background error covariance matrix where* $\mathbf{C}$ *a symmetric, positive definite, circulant, correlation matrix and* $\sigma_b^2 > 0$ *is the background error variance. We assume that* q *observations are taken with the same error variance at each time step* t_j *and therefore* $\mathbf{R}_j = \sigma_o^2 \mathbf{I}_q \in \mathbb{R}^{q \times q}$ *for* $j = 0, 1, \ldots n$ *where* $\mathbf{I}_q$ *is the q-by-q identity matrix,* $\sigma_o^2 > 0$ *is the observation error variance and* $q(n+1) < N$. *Let* $\hat{\mathbf{R}} \in \mathbb{R}^{q(n+1) \times q(n+1)}$ *be the block diagonal matrix consisting of the matrices* $\mathbf{R}_j$ *down the diagonal. Assume that observations of the parameter are made only at grid points at the same positions at each time step* t_j; *then* $\mathbf{H}_j^T \mathbf{H}_j = \mathbf{H}^T \mathbf{H} \in \mathbb{R}^{N \times N}$ *where* $\mathbf{H}^T \mathbf{H}$ *is a diagonal matrix with ones on the diagonal if the point is observed and zeros otherwise. Finally, assume* $\mathbf{M}_j = \mathbf{M} \in \mathbb{R}^{N \times N}$ *for* $j = 1, \ldots, n$ *and* $\mathbf{M}_0 = \mathbf{I}_N$ *where* $\mathbf{M}$ *is a circulant matrix. Then the following bounds hold on*

the condition number of the Hessian $\mathbf{S} = \mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}}$

$$\left(\frac{1 + \frac{q}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_{\min}(\mathbf{C})\gamma_{\min}}{1 + \frac{q}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_{\max}(\mathbf{C})\gamma_{\max}} \right) \kappa(\mathbf{C}) \leq \kappa(\mathbf{S}) \leq \kappa(\mathbf{C})\left(1 + \frac{\sigma_b^2}{\sigma_o^2}\lambda_{\min}(\mathbf{C})\lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{H}}) \right),$$

$$(7.11)$$

where, given that $\lambda_{max}(\mathbf{M})$ and $\lambda_{\min}(\mathbf{M})$ are the eigenvalues of $\mathbf{M}$ with largest and smallest magnitude respectively, we have $\gamma_{\max} = \sum_{j=0}^{n} |\lambda_{\max}(\mathbf{M})|^{2j}$ and $\gamma_{\min} = \sum_{j=0}^{n} |\lambda_{\min}(\mathbf{M})|^{2j}$.

Proof Since $\mathbf{R}_j = \sigma_o^2 \mathbf{I}_q$ for $j = 0, \ldots n$ then in particular $\hat{\mathbf{R}} = \sigma_o^2 \mathbf{I}_{q(n+1)}$ and $\lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}) = \sigma_o^{-2}\lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{H}})$. If we substitute this into the upper bound found in Theorem 7.1.1 in equation (7.3) then using $\lambda_{\min}(\mathbf{B}) = \lambda_{\min}(\sigma_b^2\mathbf{C}) = \sigma_b^2\lambda_{\min}(\mathbf{C})$ we obtain the upper bound

$$\kappa(\mathbf{S}) \leq \kappa(\mathbf{C})\left(1 + \frac{\sigma_b^2}{\sigma_o^2}\lambda_{\min}(\mathbf{C})\lambda_{\max}(\hat{\mathbf{H}}^T\hat{\mathbf{H}}) \right). \tag{7.12}$$

To find the lower bound requires application of the Rayleigh quotient. Let $\mathbf{v}_k$ be an eigenvector of $\mathbf{B}$ (hence it is of the form (3.50)) and consider the Rayleigh quotient of $\hat{\mathbf{H}}^T\hat{\mathbf{H}}$ at $\mathbf{v}_k$. Since $\mathbf{M}_1 \ldots \mathbf{M}_j = \mathbf{M}^j$ and $\mathbf{H}_j^T\mathbf{H}_j = \mathbf{H}^T\mathbf{H}$ for $j = 0, 1, \ldots n$ we have

$$\mathbf{v}_k^H(\hat{\mathbf{H}}^T\hat{\mathbf{H}})\mathbf{v}_k = \mathbf{v}_k^H\left(\sum_{j=0}^{n} (\mathbf{M}^j)^T\mathbf{H}^T\mathbf{H}\mathbf{M}^j \right)\mathbf{v}_k, \tag{7.13}$$

where the superscript H denotes the complex conjugate transpose. Since we have assumed that $\mathbf{M}$ is circulant then $\mathbf{M}^j$ is also circulant by Theorem 3.4.8. All circulant matrices share the same eigenvectors and since $\mathbf{C}$ and $\mathbf{B}$ are also circulant it follows that

$$\mathbf{M}^j\mathbf{v}_k = (\lambda_k(\mathbf{M}))^j\mathbf{v}_k, \tag{7.14}$$

and

$$\mathbf{v}_k^H(\mathbf{M}^j) = \mathbf{v}_k^H(\bar{\lambda}_k(\mathbf{M}))^j, \tag{7.15}$$

112

where $\bar{\lambda}_k(\mathbf{M})$ represents the complex conjugate of $\lambda_k(\mathbf{M})$. We note that $\mathbf{H}^T\mathbf{H} \in \mathbb{R}^{N \times N}$ is a diagonal matrix with ones on the diagonal if the point is observed or a zero otherwise and so by (6.13)

$$\mathbf{v}_k^H\mathbf{H}^T\mathbf{H}\mathbf{v}_k = \frac{q}{N}, \qquad (7.16)$$

since we make observations of the parameter at q grid points at each timestep. Substituting (7.14) and (7.15) into (7.13) and using (7.16) gives

$$\mathbf{v}_k^H \sum_{j=0}^{n} (\mathbf{M}^j)^T\mathbf{H}^T\mathbf{H}\mathbf{M}^j\mathbf{v}_k = \sum_{j=0}^{n}(\bar{\lambda}_k)^j(\lambda_k)^j\mathbf{v}_k^H\mathbf{H}^T\mathbf{H}\mathbf{v}_k \qquad (7.17)$$

$$= \frac{q}{N}\sum_{j=0}^{n}|\lambda_k|^{2j}. \qquad (7.18)$$

Let $\mathbf{v}_{\max}$ denote the eigenvector associated with $\lambda_{\max}(\mathbf{B}^{-1})$ and $\lambda_\alpha(\mathbf{M})$. Applying the Rayleigh quotient to $\mathbf{S}$ and using Theorem 3.4.4 we obtain

$$\lambda_{\max}(\mathbf{S}) \geq \mathbf{v}_{\max}^H\mathbf{S}\mathbf{v}_{\max} = \mathbf{v}_{\max}^H\mathbf{B}^{-1}\mathbf{v}_{\max} + \sigma_o^{-2}\mathbf{v}_{\max}^H(\hat{\mathbf{H}}^T\hat{\mathbf{H}})\mathbf{v}_{\max} \qquad (7.19)$$

$$= \sigma_b^{-2}\lambda_{\max}(\mathbf{C}^{-1}) + \frac{q}{N}\sigma_o^{-2}\sum_{j=0}^{n}|\lambda_\alpha(\mathbf{M})|^{2j} \qquad (7.20)$$

$$\geq \sigma_b^{-2}\lambda_{\max}(\mathbf{C}^{-1}) + \frac{q}{N}\sigma_o^{-2}\sum_{j=0}^{n}|\lambda_{\min}(\mathbf{M})|^{2j}, \quad (7.21)$$

where $\lambda_{\min}(\mathbf{M})$ is the eigenvalue of $\mathbf{M}$ with the smallest magnitude. Similarly, let $\mathbf{v}_{\min}$ be the eigenvector corresponding to $\lambda_{\min}(\mathbf{B}^{-1})$ and $\lambda_\beta(\mathbf{M})$. Then applying the Rayleigh quotient and using Theorem 3.4.4 gives

$$\lambda_{\min}(\mathbf{S}) \leq \sigma_b^{-2}\lambda_{\min}(\mathbf{C}^{-1}) + \frac{q}{N}\sigma_o^{-2}\sum_{j=0}^{n}|\lambda_{\max}(\mathbf{M})|^{2j}, \qquad (7.22)$$

where $\lambda_{\max}(\mathbf{M})$ is the eigenvalue of $\mathbf{M}$ with the largest magnitude. Now

define

$$\gamma_{\max} = \sum_{j=0}^{n} |\lambda_{\max}(\mathbf{M})|^{2j} \tag{7.23}$$

$$\gamma_{\min} = \sum_{j=0}^{n} |\lambda_{\min}(\mathbf{M})|^{2j} \tag{7.24}$$

We can combine the bounds (7.21) and (7.22) to give

$$\kappa(\mathbf{S}) \geq \left(\frac{1 + \frac{q}{N} \frac{\sigma_b^2}{\sigma_o^2} \lambda_{\min}(\mathbf{C}) \gamma_{\min}}{1 + \frac{q}{N} \frac{\sigma_b^2}{\sigma_o^2} \lambda_{\max}(\mathbf{C}) \gamma_{\max}} \right) \kappa(\mathbf{C}), \tag{7.25}$$

as required which completes the proof.

In Theorem 7.1.2 we assume that the same observations positions are used at each time step. In NWP centres this can be interpreted as fixing the observation stations. For satellite data however this may be a less realistic assumption. In Theorem 7.1.2 we have observation errors uncorrelated with each other. This is a common practice in current NWP centres [15]. Another important comment is that the model is assumed to be circulant in order to simplify the bounds (7.11). For periodic models discretised on a uniform grid, examples can be formulated which obey this assumption, for instance the advection model discretised using the upwind scheme which we introduce in the next section and its extension the diffusion advection equation using the central difference method. However we note that this will not be true when the grid spacing is not uniform or when the system is not periodic.

As with the more general bounds (7.3), the tighter bounds (7.11) in Theorem 7.1.2 are dependent on the condition number of the background error covariance matrix. As shown in Chapter 6 this resulted in the ill-conditioning of the 3DVar scheme since our background error covariance matrices (see Chapter 5) was ill-conditioned. In order to test the influence of the condition number of the background error covariance matrix on the condition number of the Hessian $\mathbf{S}$ we require a forecast model in order to construct $\mathbf{M}$. In the next

section we introduce the 1D linear advection equation which we use as our forecast model in this chapter.

7.1.2 Advection Equation

In this section we construct a simplified model $\mathbf{M}$ which is used to construct the 4DVar Hessian (7.1) used in the numerical experiments later in this chapter. We consider a single periodic parameter on a one-dimensional domain. To construct the Hessian (7.1) we require a forecast model which evolves our state vector from t_{i-1} to t_i. The model we use is the one-dimensional linear advection equation. This differential equation describes the motion of a scalar quantity $u(x,t)$ advected in a velocity field, where x and t represent the spatial and temporal domains respectively. Formally, the advection equation is a partial differential equation of the form

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0, \tag{7.26}$$

where the scalar quantity is transported within a vector field traveling with velocity $a(x,t)$. In our case we restrict ourselves to the case where we have a constant speed $a > 0$. The solutions of (7.26) are of the form $u(x,t) = u(x-at)$. We must specify boundary conditions in order to have a *well-posed* (or unique) solution that depends continuously on the boundary data [32, Chap 3.1]. Since our model is periodic we use periodic boundary conditions, hence

$$u(x) = u(x + x_0), \tag{7.27}$$

for some $x_0 \in \mathbb{R}$. Within the context of the atmosphere the advection equation can be viewed as a simple model describing the transportation of a passive tracer in the atmosphere.

In general, dynamical systems that describe the atmosphere are highly nonlinear and cannot be solved directly. The model variables are approximated

on a discrete domain and evolved forward from one time step to the next using a discretised model. We define a uniform grid on the real line consisting of N grid points with an equal spacing of Δx between adjacent points. In addition, we define $n+1$ time steps $t_0, t_1, \dots t_n$ with step length Δt, between adjacent time-steps. Let $U_k^j = U(k\Delta x, j\Delta t)$ be the numerical approximation to $u(x,t)$ at the point $(k\Delta x, j\Delta t)$ for $k = 0, \dots N-1$, $j = 0, \dots n$. Since the system is periodic we have $U_N^j = U_0^j$ for $j = 0, \dots n$. The derivatives $\frac{\partial u}{\partial t}$ and $\frac{\partial u}{\partial x}$ can be approximated using finite difference methods. We use the *upwind scheme* to approximate (7.26) [45, Chap. 4] giving

$$U_k^{j+1} = U_k^j - a\frac{\Delta t}{\Delta x}(U_k^j - U_{k-1}^j) \tag{7.28}$$

for $k = 0, \dots N-1$ and $j = 0, \dots n-1$.

We can write the finite difference system (7.28) as the matrix-vector system

$$\mathbf{U}^j = \mathbf{M}_j \mathbf{U}^{j-1} \tag{7.29}$$

where $\mathbf{U}^j = (U_0^j, U_1^j, \dots, U_{N-1}^j)^T$ and we have the matrix

$$\mathbf{M}_j = \begin{pmatrix} 1-\nu & 0 & 0 & 0 & \dots & 0 & \nu \\ \nu & 1-\nu & 0 & 0 & \dots & 0 & 0 \\ 0 & \nu & 1-\nu & 0 & 0 & & \vdots \\ & & \ddots & \ddots & & & \\ \vdots & & & \ddots & & 0 \\ 0 & & & \ddots & 1-\nu & 0 \\ 0 & 0 & \dots & & 0 & \nu & 1-\nu \end{pmatrix}, \tag{7.30}$$

where $\nu = a\frac{\Delta t}{\Delta x}$. Since the evolution between any time steps is the same we can write $\mathbf{M}_j = \mathbf{M}$ for $j = 1, \dots n$. Due to the periodic boundary conditions $\mathbf{M}$ is circulant. With $\nu \in (0,1)$ the finite difference system (7.28) is a one-step finite difference scheme which is consistent, stable and convergent [45, Sec. 5.4]. Since the matrix $\mathbf{M}$ is circulant it satisfies the assumptions of

Theorem 7.1.2. Before we can calculate the bounds (7.11) in terms of the advection model (7.28) we require a result on the maximum and minimum magnitudes of the eigenvalues of the linear model $\mathbf{M}$.

Theorem 7.1.3 *Let $\mathbf{M}$ be the discretised model of the advection equation discretised using the upwind scheme defined by (7.30). Then for $\nu \in (0, 1)$*

$$|\lambda_{\max}(\mathbf{M})|^2 = 1, \tag{7.31}$$

and

$$|\lambda_{\min}(\mathbf{M})|^2 \geq (1 - 2\nu)^2, \tag{7.32}$$

where $\lambda_{\max}(\mathbf{M})$ and $\lambda_{\min}(\mathbf{M})$ represent the eigenvalues of $\mathbf{M}$ with the largest and smallest magnitude respectively. We have equality in (7.32) if N is even.

Proof Since $\mathbf{M}$ is a circulant matrix then the eigenvalues are the discrete Fourier transform of the top row of $\mathbf{M}$ (see Section 3.4.4 for more details). Hence, the eigenvalues are of the form

$$\lambda_m = (1 - \nu) + \nu e^{-2\pi i m(N-1)/N} = 1 - \nu + \nu e^{2\pi i m/N}, \tag{7.33}$$

for $m = 0, 1, \ldots, N - 1$. Thus

$$|\lambda_m|^2 = (\bar{\lambda}_m)(\lambda_m) = (1 - \nu + \nu e^{(-2\pi i m/N)})(1 - \nu + \nu e^{(2\pi i m/N)}) \tag{7.34}$$

$$= (1 - \nu)^2 + 2\nu(1 - \nu)\cos(2\pi m/N) + \nu^2. \tag{7.35}$$

If we treat $|\lambda_m|^2$ as a continuous function in $m \in [0, N)$ we can differentiate (7.35) to give

$$\frac{d|\lambda_m|^2}{dm} = -2\nu(1 - \nu)(2\pi/N)\sin(2\pi m/N), \tag{7.36}$$

which implies the extrema of $|\lambda_m|^2$ occur at $m = 0$ and $m = N/2$. A second

differentiation gives

$$\frac{d^2|\lambda_m|^2}{dm^2} = -2\nu(1-\nu)(2\pi/N)^2\cos(2\pi m/N). \tag{7.37}$$

Evaluating (7.37) at $m = 0$ and $m = N/2$ give a negative and a positive value respectively. This implies that $|\lambda_0(\mathbf{M})|^2$ is the largest value, and therefore eigenvalue, of $|\lambda_m|^2$ giving

$$|\lambda_0(\mathbf{M})|^2 = (1-\nu)^2 + 2\nu(1-\nu) + \nu^2 = 1. \tag{7.38}$$

The minimum value of $|\lambda_m|^2$ occurs at $m = N/2$ giving

$$|\lambda_{N/2}(\mathbf{M})|^2 = (1-\nu)^2 - 2\nu(1-\nu) + \nu^2 = 1 - 4\nu + 4\nu^2 = (1-2\nu)^2. \tag{7.39}$$

Therefore we have $|\lambda_m|^2 \geq (1-2\nu)^2$ for any eigenvalue λ_m of $\mathbf{M}$. If N is even then $\lambda_{N/2}(\mathbf{M})$ is an eigenvalue of $\mathbf{M}$. This completes the proof.

We notice from (7.35), that $|\lambda_m|^2$ is symmetric around $m = N/2$ and by (7.36) $\frac{d|\lambda_m|^2}{dm} < 0$ for $m \in (0, \frac{N}{2}), \nu \in (0, 1)$ and so the function is monotonically decreasing between $m = 0$ and $m = N/2$. This means if N is odd then the minimum eigenvalue will be at $m = (N-1)/2$ giving

$$|\lambda_{\min}(\mathbf{M})|^2 = (1-\nu)^2 + 2\nu(1-\nu)\cos(\pi(N-1)/N) + \nu^2. \tag{7.40}$$

We can now calculate the bounds given in Theorem 7.1.2 in the case were $\mathbf{M}$ is the matrix (7.30).

Theorem 7.1.4 *Suppose we have the same assumptions as given in Theorem 7.1.2 but now $\mathbf{M}$ is the matrix (7.30). Then we have the following lower bound on the condition number of (7.1),*

$$\kappa(\mathbf{S}) \geq \left(\frac{1 + \frac{q}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_{\min}(\mathbf{C})\gamma_{adj}}{1 + \frac{q(n+1)}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_{\max}(\mathbf{C})} \right) \kappa(\mathbf{C}) \tag{7.41}$$

where

$$\gamma_{adj} = \frac{(1 - |1 - 2\nu|^{2(n+1)})}{1 - |1 - 2\nu|^2}. \tag{7.42}$$

Proof From Theorem 7.1.2 it remains to find $\gamma_{\max}$ and $\gamma_{\min}$. In Theorem 7.1.3 we showed $|\lambda_{\max}| = 1$ hence

$$\gamma_{\max} = \sum_{j=0}^{n} |\lambda_{\max}(\mathbf{M})|^{2j} = (n+1), \tag{7.43}$$

Similarly we showed that $|\lambda_{\min}(\mathbf{M})|^2 \geq (1 - 2\nu)^2$ and therefore

$$\gamma_{\min} = \sum_{j=0}^{n} |\lambda_{\min}(\mathbf{M})|^{2j} \geq \sum_{j=0}^{n} |1 - 2\nu|^{2j}, \tag{7.44}$$

which is a geometric sum giving

$$\gamma_{\min} \geq \sum_{j=0}^{n} |1 - 2\nu|^{2j} = \frac{(1 - |1 - 2\nu|^{2(n+1)})}{1 - |1 - 2\nu|^2}, \tag{7.45}$$

as required.

Notice that if N is even then we have equality in (7.45). In the next section we investigate the condition number of the Hessian using the bounds derived in this section together with numerical experiments on a periodic system.

7.1.3   Numerical Experiments

In Section 7.1.2 we derived bounds on the condition number of a 4DVar Hessian using a simple 1D discretised advection model (7.28) as the forecast model $\mathbf{M}$. We found that the lower and upper bounds are dependent on the condition number of the background error covariance matrices. As we saw in Chapter 6 the sensitivity of the condition number of the background error covariance matrices to the correlation lengthscale of the background errors gave rise to a corresponding sensitivity in the condition number of

the Hessian. From the bounds derived in Sections 7.1.1 and 7.1.2 we expect the same behaviour in the 4DVar system. In this section we present a small experiment to test this hypothesis.

We consider a single periodic parameter defined on a one-dimensional domain with $N = 500$ grid points and $\Delta x = 0.1$ grid spacing. Let $\mathbf{M}$ be the discrete linear advection forecast model defined by (7.30). We use a time-step $\Delta t = 0.1$ and $a = 0.3$, giving $\nu = a\frac{\Delta t}{\Delta x} = a \in (0,1)$ and therefore the numerical model (7.28) is stable. Observations are made at three time steps at intervals of length $3\Delta t$ so we observe at $t_0 = 0, t_1 = 3\Delta t$ and $t_2 = 6\Delta t$. A random distribution of $q = 20$ observations of the parameter are made at grid points at each time step giving a total of $q(n+1) = 60$ observations. Our observation operators are given by $\mathbf{H}_j = \mathbf{H}$ for $j = 0, 1, 2$. Observation errors are assumed to be spatially and temporally uncorrelated with observation variance $\sigma_o^2 = 1$ and therefore $\mathbf{R}_j = \sigma_o^2\mathbf{I}_q$ for $j = 0, 1, 2$. In this section we consider the background error covariance matrix $\mathbf{B} = \sigma_b^2\mathbf{C}$ where $\mathbf{C}$ is the SOAR correlation matrix (5.11) introduced in Section 5.2 and the background error variance $\sigma_b^2 = 1$. We concentrate on the SOAR case in this chapter; similar results apply to the Gaussian and Laplacian forms of the Hessian. The SOAR correlation matrix is used to model the horizontal background error correlations of the Met Office operational system which we study in the next chapter.

With the above assumptions, the hypothesis of Theorems 7.1.2 and 7.1.4 are satisfied and the bounds (7.11) and (7.41) hold on the Hessian $\mathbf{S}$. Figure 7.1 shows the condition number of the 4DVar Hessian together with the bounds as a function of background error correlation lengthscale. As predicted by the bounds derived in the previous section the conditioning of the background error covariance matrix appears to drive the conditioning of the Hessian matrix (7.1). This is apparent from the similarity in the magnitudes of the condition numbers of the Hessian and the SOAR background error covariance matrix
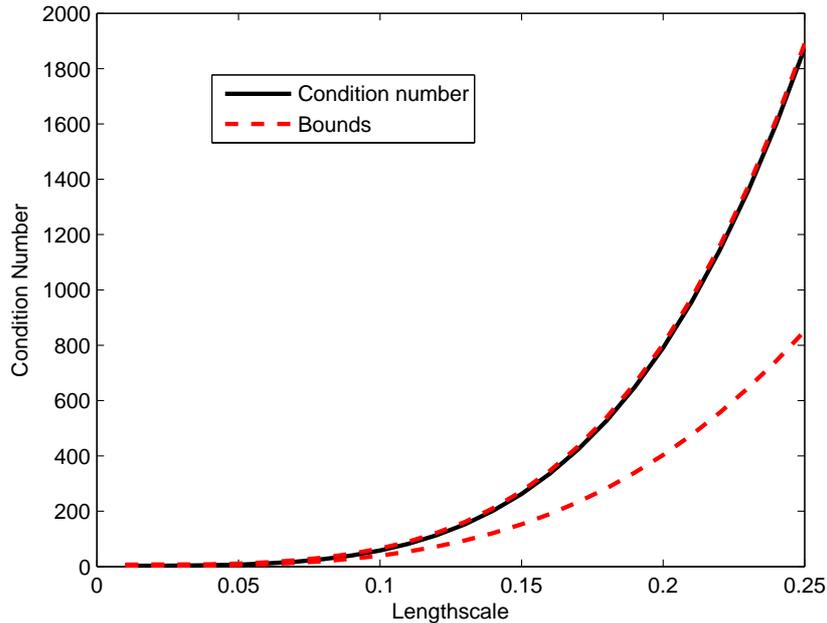
Figure 7.1: The condition number of the unpreconditioned 4DVar Hessian (solid line) and the bounds (dashed line) as a function of lengthscale using the SOAR background correlation matrix.

at fixed lengthscales. For instance, at lengthscale $L = 0.2$ the condition number of the Hessian (Figure 7.1) and the condition number of the SOAR background matrix (Figure 5.4) are both around 800. Comparison between the 3DVar Hessian in Figures 6.2 and 6.3 and the 4DVar Hessian shows that both systems have very similar condition numbers despite the differences in the 4DVar and 3DVar systems. At lengthscale $L = 0.25$ the 3DVar Hessian in Figure 6.3 has the same condition number as the 4DVar Hessian in Figure 7.1 of about 2000. This is further evidence that the background covariance matrix is driving the condition number. Similarly it is possible to show that the condition number of the 4DVar Hessians using the Gaussian and the Laplacian background error covariance matrices are determined by the condition number of the respective background error covariance matrices and have similar plots with respect to lengthscale as their 3DVar counterparts shown in Figures 6.1 and 6.5 respectively. We now consider the conditioning of the preconditioned 4DVar Hessian.

7.2 Preconditioned 4DVar

Having considered the preconditioned 3DVar Hessian in Chapter 6 we now consider the condition number of the 4DVar Hessian, preconditioned using the control variable transform (CVT) described in Section 2.4. From here onwards we refer to this preconditioned 4DVar system as P4DVar. In Chapter 6 we showed that preconditioning with the CVT significantly reduced the condition number of the 3DVar Hessian. In this section we aim to show that the same advantages can be achieved in the condition number of the 4DVar Hessian by including the control variable transform.

7.2.1 Theory

In this subsection we derive new theoretical bounds on the conditioning of P4DVar. As shown in Chapter 4 the Hessian of P4DVar can be written

$$\hat{\mathbf{S}} = \mathbf{I} + \mathbf{B}^{T/2}\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{B}^{1/2}, \tag{7.46}$$

where $\hat{\mathbf{H}}$ is defined by (7.2) and $\hat{\mathbf{R}}$ is the block matrix with the error covariances $\mathbf{R}_i$ down the diagonal. We use the symmetric square root of the background covariance matrix, hence throughout the remainder of this chapter we assume $\mathbf{B}^{T/2} = \mathbf{B}^{1/2}$. In the following theorem we extend the theoretical bounds on the condition number of the P3DVar Hessian (6.18) found in Section 6.2.1 to the P4DVar Hessian (7.46).

Theorem 7.2.1 *Let $\mathbf{B} = \sigma_b^2\mathbf{C} \in \mathbb{R}^{N \times N}$ be the background error covariance matrix where $\sigma_b^2 > 0$ is the background error variance and $\mathbf{C}$ is the error correlation matrix. We assume that q observations are taken with the same error variance at each time step t_j and therefore $\mathbf{R}_j = \sigma_o^2\mathbf{I}_q \in \mathbb{R}^{q \times q}$ for $j = 0, \ldots n$ where $\mathbf{I}_q$ is the q-by-q identity matrix, $\sigma_o^2 > 0$ is the observation error variance and $q(n+1) < N$. Let $\hat{\mathbf{R}} \in \mathbb{R}^{q(n+1) \times q(n+1)}$ be the block diagonal*

matrix with the matrices $\mathbf{R}_j$ down the diagonal. Let $\mathbf{H}_j \in \mathbb{R}^{q \times N}$ and $\mathbf{M}_j \in \mathbb{R}^{N \times N}$ be the observation and the linear model operators respectively at time t_j for $j = 0, 1 \ldots n$ with $\mathbf{M}_0 := \mathbf{I}_N$. Additionally, let $\hat{\mathbf{H}} \in \mathbb{R}^{(n+1)q \times N}$ be the generalised observation operator defined by (7.2). Then the following bounds hold on the condition number of the Hessian $\hat{\mathbf{S}} = \mathbf{I} + \mathbf{B}^{1/2} \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}} \mathbf{B}^{1/2}$

$$1 + \frac{1}{q(n+1)} \frac{\sigma_b^2}{\sigma_o^2} \sum_{k,l=1}^{q(n+1)} (\hat{\mathbf{H}} \mathbf{C} \hat{\mathbf{H}}^T)_{k,l} \leq \kappa(\hat{\mathbf{S}}) \leq 1 + \frac{\sigma_b^2}{\sigma_o^2} ||\hat{\mathbf{H}} \mathbf{C} \hat{\mathbf{H}}^T||_\infty, \qquad (7.47)$$

where $(\mathbf{A})_{i,j}$ represents the $(i,j)^{th}$ entry of the matrix $\mathbf{A}$.

Proof Since there are fewer observations than variables in the state space ($q(n+1) < N$) then the smallest eigenvalue of (7.46) is simply unity and the condition number of the Hessian is simply the largest eigenvalue of $\hat{\mathbf{S}}$. By Theorem 3.4.5 with ($\mathbf{E} = \hat{\mathbf{R}}^{-1/2} \hat{\mathbf{H}} \mathbf{B}^{1/2}$) we know that $\mathbf{E}^T \mathbf{E} = \mathbf{B}^{1/2} \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}} \mathbf{B}^{1/2}$ and $\mathbf{E} \mathbf{E}^T = \hat{\mathbf{R}}^{-1/2} \hat{\mathbf{H}} \mathbf{B} \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1/2}$ have the same non-zero eigenvalues. Hence the Hessian (7.46) has the same non-unit eigenvalues as

$$\tilde{\mathbf{S}} = \mathbf{I}_{q(n+1)} + \hat{\mathbf{R}}^{-1/2} \hat{\mathbf{H}} \mathbf{B} \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1/2} = \mathbf{I}_{q(n+1)} + \frac{\sigma_b^2}{\sigma_o^2} \hat{\mathbf{H}} \mathbf{C} \hat{\mathbf{H}}^T. \qquad (7.48)$$

By Theorem 3.4.2 applied to (7.48) with the ∞-norm (3.6) we find

$$\kappa(\hat{\mathbf{S}}) = \lambda_{\max}(\tilde{\mathbf{S}}) \leq 1 + \frac{\sigma_b^2}{\sigma_o^2} ||\hat{\mathbf{H}} \mathbf{C} \hat{\mathbf{H}}^T||_\infty, \qquad (7.49)$$

which establishes the upper bound.

The lower bound can be established by the calculating the Rayleigh quotient of $\tilde{\mathbf{S}}$ with the unit vector $\mathbf{y} = \frac{1}{\sqrt{q(n+1)}} (1, 1, \ldots, 1)^T \in \mathbb{R}^{q(n+1)}$,

$$R_{\tilde{\mathbf{S}}}(\mathbf{y}) = \mathbf{y}^T \tilde{\mathbf{S}} \mathbf{y} = 1 + \frac{1}{q(n+1)} \frac{\sigma_b^2}{\sigma_o^2} \sum_{k,l=1}^{q(n+1)} (\hat{\mathbf{H}} \mathbf{C} \hat{\mathbf{H}}^T)_{k,l}. \qquad (7.50)$$

By Theorem 3.4.4 we have $\lambda_{\max}(\tilde{\mathbf{S}}) \geq R_{\tilde{\mathbf{S}}}(\mathbf{y})$ for any $\mathbf{y} \in \mathbb{R}^{q(n+1) \times q(n+1)}$, which completes the proof.

The bounds (7.47) are dependent on the sum of what we call the *reduced 4D background error covariance matrix* $\hat{\mathbf{H}}\mathbf{C}\hat{\mathbf{H}}^T$, this can be compared to the reduced background error correlation matrix $\mathbf{H}\mathbf{C}\mathbf{H}^T$ introduced in Section 6.2.1. We can write the reduced 4D background error covariance matrix in the form

$$\hat{\mathbf{H}}\mathbf{C}\hat{\mathbf{H}}^T = \tilde{\mathbf{H}}\tilde{\mathbf{C}}\tilde{\mathbf{H}}^T, \tag{7.51}$$

where $\tilde{\mathbf{H}}$ is the block diagonal matrix

$$\tilde{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 & \mathbf{0} & \\ & & \ddots & \\ & & \mathbf{0} & \mathbf{H}_n \end{pmatrix}, \tag{7.52}$$

and $\tilde{\mathbf{C}}$ is the 4D-background covariance matrix

$$\tilde{\mathbf{C}} = \begin{pmatrix} \mathbf{C} & \mathbf{C}\hat{\mathbf{M}}_1^T & & \dots & \mathbf{C}\hat{\mathbf{M}}_n^T \\ \hat{\mathbf{M}}_1\mathbf{C} & \hat{\mathbf{M}}_1\mathbf{C}\hat{\mathbf{M}}_1^T & & \dots & \hat{\mathbf{M}}_1\mathbf{C}\hat{\mathbf{M}}_n^T \\ \hat{\mathbf{M}}_2\mathbf{C} & \hat{\mathbf{M}}_2\mathbf{C}\hat{\mathbf{M}}_1^T & & & \hat{\mathbf{M}}_2\mathbf{C}\hat{\mathbf{M}}_n^T \\ & & \ddots & \ddots & \\ \vdots & & & \ddots & \\ & & & & \ddots \\ \hat{\mathbf{M}}_n\mathbf{C} & \hat{\mathbf{M}}_n\mathbf{C}\hat{\mathbf{M}}_1^T & & & \hat{\mathbf{M}}_n\mathbf{C}\hat{\mathbf{M}}_n^T \end{pmatrix}, \tag{7.53}$$

where $\hat{\mathbf{M}}_j = \mathbf{M}_j\mathbf{M}_{j-1}\dots\mathbf{M}_1$. Suppose that ϵ_0^b describes the background errors at time t_0, with covariance matrix $\mathbf{C}$, and ϵ_k^b describes the background errors evolved to time t_k, i.e. $\epsilon_k^b = \mathbf{M}_k\mathbf{M}_{k-1}\dots\mathbf{M}_1\epsilon_0^b$. We can write the error covariance matrix describing the covariance between the errors evolved to

times t_k and t_l as

$$< \epsilon_k^b(\epsilon_l^b)^T >= \mathbf{M}_k\mathbf{M}_{k-1}\ldots\mathbf{M}_1 < \epsilon_0^b(\epsilon_0^b)^T > \mathbf{M}_1^T\ldots\mathbf{M}_{l-1}^T\mathbf{M}_l^T \qquad (7.54)$$

$$= \mathbf{M}_k\mathbf{M}_{k-1}\ldots\mathbf{M}_1\mathbf{C}\mathbf{M}_1^T\ldots\mathbf{M}_{l-1}^T\mathbf{M}_l^T, \qquad (7.55)$$

$$= \hat{\mathbf{M}}_k\mathbf{C}\hat{\mathbf{M}}_l, \qquad (7.56)$$

where $<>$ represents expected value over the random variables. Hence each block of the 4DVar covariance matrix $\tilde{\mathbf{C}}$ can be viewed as the covariance of the 4D background errors vector $(\epsilon_0^{bT}, \epsilon_1^{bT}, \ldots, \epsilon_n^{bT})^T$. If observations of the parameter are only made at grid points then the $(k,l)^{th}$ block of $\tilde{\mathbf{H}}\tilde{\mathbf{C}}\tilde{\mathbf{H}}^T$ is the matrix $\hat{\mathbf{M}}_k\mathbf{C}\hat{\mathbf{M}}_l$ but minus those rows and columns at the positions which are not observed. From the form of the 4D background error covariance matrix (7.53) the model $\mathbf{M}_k$ clearly has an influence on the conditioning of the P4DVar system. Due to the wide variety of possible models this is difficult to analyse in a general way and so in this chapter we concentrate on a simple linear advection equation for the model $\mathbf{M}_k$ for $k = 1, \ldots n$, introduced in the Section 7.1.2.

Theorem 7.2.1 shows that the bounds (7.47) are linearly related to the inverse of the observation error variance σ_o^2. We therefore expect the condition number of the Hessian $\hat{\mathbf{S}}$ to increase as the observations become more accurate (σ_o^2 decrease). In Chapter 6 we showed that changing the number and spacing of the observations changes the condition number of the preconditioned 3DVar Hessian (6.18). We found that an increase in the spacing of the observations and the reduction in the number of observations, produced a reduction in the condition number of the Hessian $\hat{\mathbf{S}}$. From the bounds (7.47) given in Theorem 7.2.1 clearly the observation operators $\mathbf{H}_k$ $(k = 0, 1, \ldots, n)$ is an important influence on the condition number of the Hessian.

If we assume that observations of the parameter are only made at grid points, $\tilde{\mathbf{H}}\tilde{\mathbf{C}}\tilde{\mathbf{H}}^T$ is the matrix formed by deleting rows and columns of the 4D covariance matrix corresponding to the positions at which the parameters are not

observed. Let us assume that the background error correlation matrix $\mathbf{C}$ has positive coefficients $c_{i,j}$ which monotonically decrease with increased separation $|i - j|$. If the action of the model on the background error covariance matrix preserves the positivity of the coefficients and ensures that the coefficients remain a monotonically decreasing function of separation, then we expect that the condition number of the Hessian to decreases as the spacing between observations increases and as observations are removed. We analyse the effect of the observation accuracy and distribution in more detail in a later section. In the next section we perform some numerical experiments to analyse the conditioning of P4DVar in a simple one-parameter, periodic system.

7.2.2   Numerical Experiments

In this section we examine the P4DVar system for single parameter, one-dimensional, periodic system. We define $N = 500$ grid points on the real line with uniform spacing $\Delta x = 0.1$. In our experiments we use the linear advection model discretised using the upwind scheme (7.28), introduced in Section 7.1.2, as our model $\mathbf{M}$ and use the SOAR correlation matrix $\mathbf{C}$ given by equation (5.11) to construct our background error covariance matrix $\mathbf{B} = \sigma_b^2 \mathbf{C}$. Here we fix the background error variance as $\sigma_b^2 = 1$ and so $\mathbf{B} = \mathbf{C}$.

To describe the coefficients of the 4D-background error covariance matrix (7.53) we require knowledge of the coefficients of the matrices $\hat{\mathbf{M}}_k \mathbf{C} \hat{\mathbf{M}}_k^T$ and $\hat{\mathbf{M}}_k \mathbf{C}$. In the advection equation case this is simplified since $\hat{\mathbf{M}}_k = \mathbf{M}^k$ where $\mathbf{M}$ is the matrix (7.30). We consider a time step of $\Delta t = 0.1$ and $a = 0.3$; hence $\nu = a \frac{\Delta t}{\Delta x} = 0.3$ and the finite difference system (7.28) is stable. We consider observations at 3 time steps $t_0 = 0, t_1 = 3\Delta t$ and $t_2 = 6\Delta t$. We can

write the 4D background error covariance matrix for our system as

$$\tilde{\mathbf{C}} = \begin{pmatrix} \mathbf{C} & \mathbf{C}(\mathbf{M}^T)^3 & \mathbf{C}(\mathbf{M}^T)^6 \\ \mathbf{M}^3\mathbf{C} & \mathbf{M}^3\mathbf{C}(\mathbf{M}^T)^3 & \mathbf{M}^3\mathbf{C}(\mathbf{M}^T)^6 \\ \mathbf{M}^6\mathbf{C} & \mathbf{M}^6\mathbf{C}(\mathbf{M}^T)^3 & \mathbf{M}^6\mathbf{C}(\mathbf{M}^T)^6 \end{pmatrix}. \qquad (7.57)$$

Since $\mathbf{M}$ and $\mathbf{C}$ are circulant, each block of the matrix (7.57) is circulant by Theorem 3.4.8. Since the blocks are circulant to describe the coefficents of each matrix $\mathbf{M}^k\mathbf{C}(\mathbf{M}^T)^l$ $k, l = 0, 3, 6$ we can consider a single row. We

Figure 7.2: The 250th row of the $\mathbf{B}$ matrix (solid line), $\mathbf{M}^3\mathbf{B}(\mathbf{M}^T)^3$ (dashed line) and $\mathbf{M}^6\mathbf{B}(\mathbf{M}^T)^6$ (dotted line) where $\mathbf{M}$ is the discretised advection model.

first consider the diagonal blocks of (7.57). These are matrices of the form $\mathbf{M}^k\mathbf{C}(\mathbf{M}^T)^k$ for $k = 0, 3, 6$. Figure 7.2 shows the 250th rows of these matrices. The plot shows that the coefficients $(\mathbf{M}^k\mathbf{C}(\mathbf{M}^T)^k)_{i,j}$ for $i, j = 1, \ldots N$ are all positive and are monotonically decreasing as a function of distance $|i-j|$. The matrix $\mathbf{M}$ applied to $\mathbf{C}$ results in a reduction of the maximum components $(\mathbf{M}^k\mathbf{C}(\mathbf{M}^T)^k)_{i,i}$ compared to the components of the matrix $\mathbf{C}$. This indicates that there is a reduction in the variance of the errors $\epsilon_k^b = \mathbf{M}^k\epsilon_0^b$ by evolving the background error ϵ_0^b at time $t = 0$ with the model $\mathbf{M}$. The reduction in

the error variance is accompanied by an increase in the size of the coefficients $(\mathbf{M}^k \mathbf{B} (\mathbf{M}^T)^k)_{i,j}$ at larger distances $|i-j|$. For instance, Figure 7.2 shows that the matrix $\mathbf{M}^6 \mathbf{B} (\mathbf{M}^T)^6$ has slightly larger coefficients than the matrix $\mathbf{C}$ at column number 245. This reduction in the variance can be explained by the damping effect of the upwind scheme [45, Sec. 4.2].

We now consider the off diagonal blocks of (7.57). Figure 7.3 shows the components of the columns for the 250th row of the $\mathbf{M}^k \mathbf{C}$ matrices for $k = 0, 3, 6$. Evolution of the background error covariance matrix with $\mathbf{M}$ shows that the coefficients of the matrices $(\mathbf{M}^k \mathbf{C})_{i,j}$ remain positive and the 'bell-shape' of the SOAR correlation matrix is preserved but with the components with the largest magnitude below the diagonal of the matrix. This can be explained by the dynamics of the advection model. With a constant positive velocity a, the finite difference scheme (7.28) approximates the advection model (7.26) and therefore the matrix $\mathbf{M}^k$ transports a scalar quantity by a distance of approximately $ak\Delta t$ or by $ak\Delta t/\Delta x$ grid points. For the matrix $(\mathbf{M}^k \mathbf{C})$, $\mathbf{M}$ transports the coefficients of each column of $\mathbf{C}$ by a distance $ak\Delta t$ and this corresponds to a shift of $ak\Delta t$ of the components of the rows as seen in Figure 7.3. The figure shows that the largest component of the matrix $(\mathbf{M}^k \mathbf{C})$ have been advected by approximately $ak\Delta t/\Delta x \approx 0.9, 1.8$ grid points distance for $k = 3, 6$ respectively, from the position of the largest component of the matrix $\mathbf{C}$ at column number 250. We now consider the conditioning of the P4DVar Hessian using the 4D background error covariance matrix (7.57).

We consider a periodic parameter on a one-dimensional domain. We use the model $\mathbf{M}$ and the background error covariance matrix $\mathbf{B}$ defined in this section. We fix the observation error variance at $\sigma_o^2 = 1$ and assume that observations are only made at grid points. We use the same observation configuration of 20 random observations made at the same points of each time step $0, 3\Delta t, 6\Delta t$ as used in the experiments for the 4DVar system in Section 7.1.3. Hence, in total we use 60 observations. With this system design we satisfy the assumptions of Theorem 7.2.1 and therefore the bounds
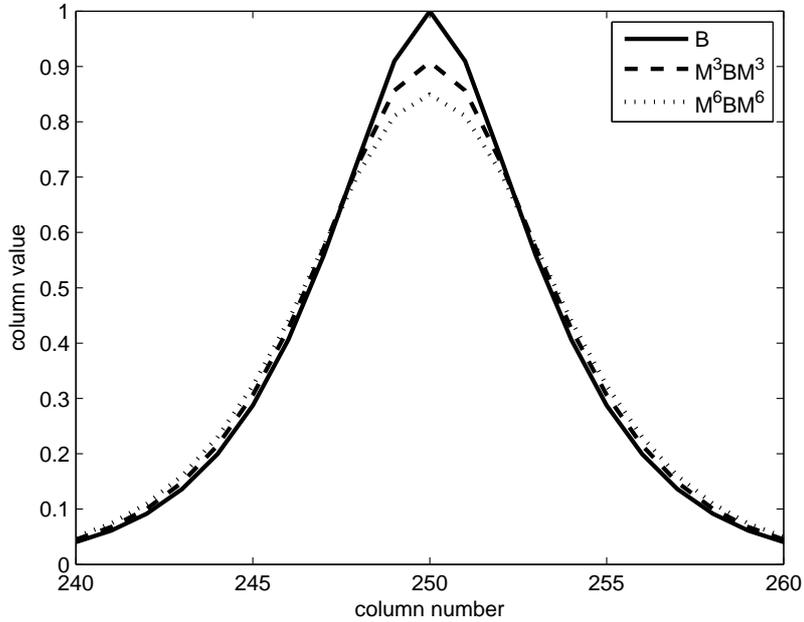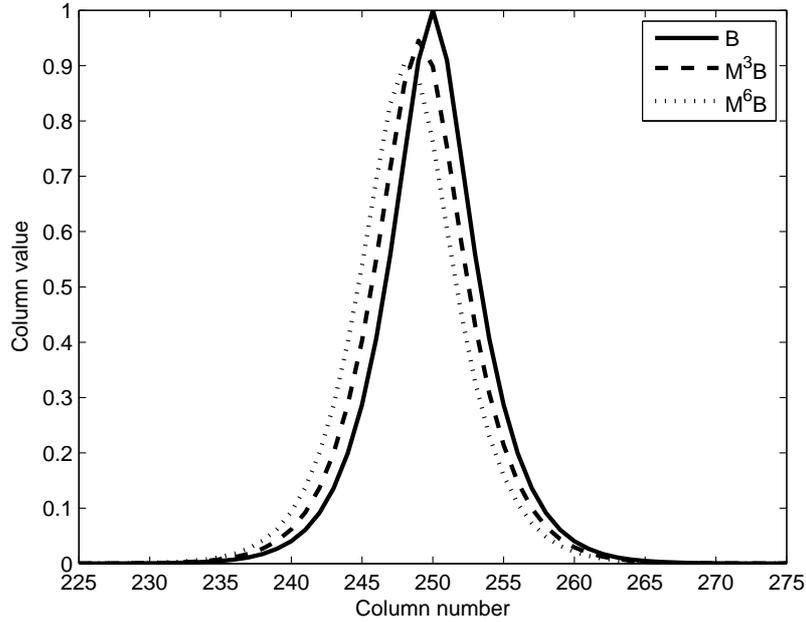
Figure 7.3: The 250th row of the $\mathbf{B}$ matrix (solid line), $\mathbf{M}^3\mathbf{B}$ (dashed line) and $\mathbf{M}^6\mathbf{B}$ (dotted line) where $\mathbf{M}$ is the discretised advection model.

(7.47) hold for our P4DVar Hessian $\hat{\mathbf{S}}$.

Figure 7.4 shows the conditioning of the preconditioned 4DVar Hessian as a function of lengthscale. The plot also shows the theoretical bounds derived in Theorem 7.2.1 with the change in background error correlation lengthscale. As the lengthscale increases so does the condition number of $\hat{\mathbf{S}}$. This can be explained by the theoretical bounds (7.47). The bounds on the condition number of $\hat{\mathbf{S}}$ are dependent on the sums of the reduced 4D background error covariance matrix. As the lengthscale increases the size of the components of the matrix (7.57) also increase because the components of the $\mathbf{C}$ increase with lengthscale. Therefore the condition number also increases.

As seen by Figures 7.3 and 7.2, the size of the components of the 4D background error covariance matrix are no bigger then the maximum component of the matrix $\mathbf{C}$. In this case the largest component of $\mathbf{C}$ is one. Therefore, since we use 60 observations in total, the maximum row sum of the reduced
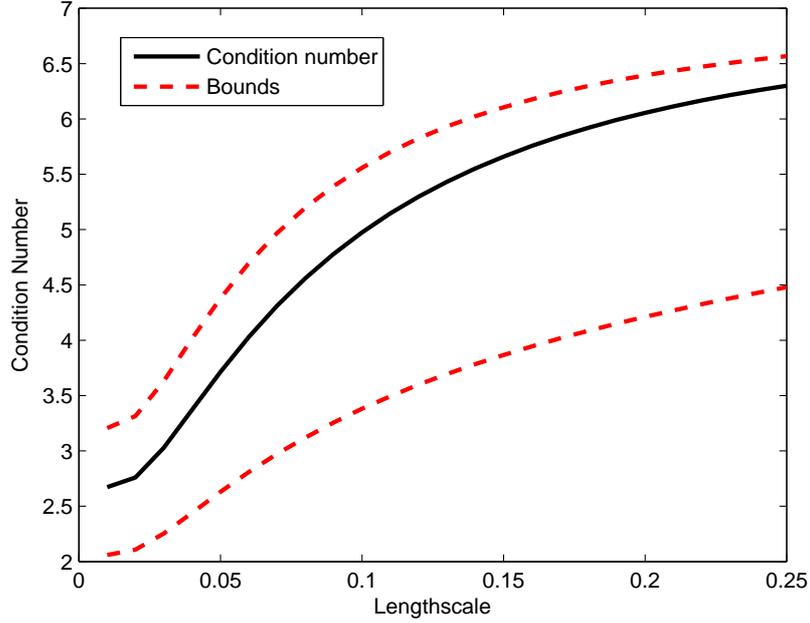
Figure 7.4: The condition number of the preconditioned Hessian (solid line) and the bounds (dotted) as a function of lengthscale using the SOAR background correlation matrix

4D background error covariance matrix satisfies

$$||\hat{\mathbf{H}}\mathbf{C}\hat{\mathbf{H}}^T||_\infty = ||\tilde{\mathbf{H}}\tilde{\mathbf{C}}\tilde{\mathbf{H}}^T||_\infty \leq 60, \tag{7.58}$$

and the by equation (7.47) the condition number of the Hessian satisfies

$$\kappa(\hat{\mathbf{S}}) \leq 1 + \sigma_o^{-2}||\hat{\mathbf{H}}\mathbf{C}\hat{\mathbf{H}}^T||_\infty \leq 1 + 60 = 61. \tag{7.59}$$

For the unpreconditioned system given in Section 7.1.3, the condition number of the 4DVar Hessian exceeded 61 after $L = 0.15 = 1.5\Delta x$ as shown by Figure 7.1. Therefore we expect the P4DVar system to be much better conditioned then the unpreconditioned system. Comparing Figures 7.1 and 7.4 shows that preconditioning produces a significant reduction in the condition number of 4DVar Hessian. For example, at lengthscale $L = 0.25$ the condition number has been reduced from approximately 1800 for the 4DVar Hessian to about 6 for the P4DVar Hessian. Repeating the numerical experiments using the P4DVar Hessians derived using the Laplacian and Gaussian matrices give

similar results with a significant reduction in the condition number when 4DVar is preconditioned with the background covariance matrix, we do not present these results here.

The bounds indicate that the condition number of the P4DVar Hessian (7.46) depends on other factors in the system: the observation accuracy, the observation distribution and the number of observations. We consider these factors in more detail in the next section.

7.3 Observation Accuracy

In this section we consider the role of observation accuracy on the conditioning of P4DVar. The bounds (7.47) in Theorem 7.2.1 show that the conditioning of P4DVar is linearly related to the inverse of the observation error variance σ_o^2. Assuming that the Hessian satisfies the assumptions of Theorem 7.2.1 then the condition number of the P4DVar Hessian can be written explicitly as

$$\kappa(\hat{\mathbf{S}}) = \lambda_{\max}(\tilde{\mathbf{S}}) = 1 + \frac{\sigma_b^2}{\sigma_o^2}\lambda_{\max}(\hat{\mathbf{H}}\mathbf{B}\hat{\mathbf{H}}^T), \qquad (7.60)$$

where $\tilde{\mathbf{S}}$ is the matrix defined by equation (7.48). The equation (7.60) indicates that increasing the accuracy of the observations (reducing σ_o^2) will cause the condition number of P4DVar to increase. We illustrate this with a simple numerical experiment. We consider a single periodic parameter defined on the real line. We choose $N = 500$ uniformly spaced grid points with spacing of length $\Delta x = 0.1$. Observations of the parameters are taken at three time points $t_0 = 0, t_1 = 3\Delta t$ and $t_2 = 6\Delta t$ where $\Delta t = 0.1$ and is the time step for the discretised advection model defined by (7.28) with $a = 0.3$. Observations are taken only at grid points with 20 observations at each time step using the same observation configuration as used in Section 7.2.2. We assume that the observation errors are uncorrelated with error variance σ_o^2 and therefore the condition number of the P4DVar Hessian is of the form given in (7.60).

We choose the background error covariance matrix to be $\mathbf{B} = \sigma_b^2 \mathbf{C}$ where $\mathbf{C}$ is the SOAR correlation matrix (5.11) and $\sigma_b^2 = 1$ is the background error variance. We fix the correlation lengthscale as $L = 0.2$ and consider the condition number of $\hat{\mathbf{S}}$ as we vary the observation error variance. The change in condition number as the observation accuracy changes is shown in Table 7.1.

Obs Variance	Condition Number
0.01	506.53
0.05	102.11
0.10	51.55
0.50	11.11
1.00	6.06
2.00	3.53
5.00	2.01
10.00	1.51

Table 7.1: Change in the condition of the preconditioned Hessian with change observation error variance using the SOAR correlation matrix.

The table confirms that the condition number is linearly related to the inverse of the observation error variance. For instance, halving the variance from 0.1 to 0.05 roughly doubles the condition number from 51.55 to 102.11. As in the P3DVar case considered in Section 6.3.1, the more accurate the observations the more ill-conditioned the Hessian is. The same results apply for the preconditioned Hessian using the Gaussian and Laplacian correlation matrices but we do not present these results here. We now turn our attention to the observation distribution and its influence on the conditioning.

7.4 Observation Distribution

The other main factor affecting the condition number of P4DVar are sums of the elements of the reduced 4D background error covariance matrix $\hat{\mathbf{H}}\mathbf{B}\hat{\mathbf{H}}^T$ derived in the bounds (7.47) in Theorem 7.2.1. As shown in Section 7.2.2,

each block of the reduced covariance matrix is a matrix of the form $\mathbf{H}_i \hat{\mathbf{M}}_i \mathbf{B} \hat{\mathbf{M}}_j^T \mathbf{H}_j^T$. We consider the change in the condition number of the preconditioned Hessian (7.46) as a function of the change in observation distribution in two ways

- Keeping the number of observations constant but changing the spacing of the observations.

- Changing the number of observations.

In this section we consider the advection equation $\mathbf{M}$ discretised using the upwind scheme (7.28) as our forecast model. We consider a periodic parameter on the real line and fix $N = 500$ grid points with uniform spacing $\Delta x = 0.1$. We let the background $\mathbf{B} = \sigma_b^2 \mathbf{C}$ where $\sigma_b^2 = 1$ and $\mathbf{C}$ is the SOAR correlation matrix (5.11). We use a time step of $\Delta t = 0.1$ for the advection model and let $a = 0.3$. Therefore, $\nu = a \frac{\Delta t}{\Delta x} = 0.3$ and the upwind scheme is numerically stable. Additionally, we choose observations only at grid points and choose the same spatial positions for the observations for each time step, therefore $\mathbf{H}_i = \mathbf{H}$ for $i = 0, \dots n$.

7.4.1 Observation Spacing

We begin by considering the effect of observation spacing on the condition number of the P4DVar Hessian. Since observations of the parameter are only made at grid points then the reduced 4D background error covariance matrix $\hat{\mathbf{H}} \mathbf{C} \hat{\mathbf{H}}^T$ is simply the 4D background error covariance matrix (7.53) but with rows and columns removed at the positions where no observations are made. We assume that observations are made at time steps $t_0 = 0, t_1 = 3\Delta t, t_2 = 6\Delta t$ and we initially take 20 observations of the parameter on the first 20 grid points on the domain. We consider the effect of increasing the spacing of the observations on the condition number of $\hat{\mathbf{S}}$. Using this experimental design

we satisfy the criteria of Theorem 7.2.1 and therefore the bounds (7.47) hold on the Hessian $\hat{\mathbf{S}}$.

In order to understand the effect of the observation spacing on the condition number of $\hat{\mathbf{S}}$ we need to understand the shape of the coefficients of the 4D background error covariance matrix, which for our experimental design is given by the matrix

$$
\tilde{\mathbf{C}} = \begin{pmatrix}
\mathbf{C} & \mathbf{C}(\mathbf{M}^T)^3 & \mathbf{C}(\mathbf{M}^T)^6 \\
\mathbf{M}^3\mathbf{C} & \mathbf{M}^3\mathbf{C}(\mathbf{M}^T)^3 & \mathbf{M}^3\mathbf{C}(\mathbf{M}^T)^6 \\
\mathbf{M}^6\mathbf{C} & \mathbf{M}^6\mathbf{C}(\mathbf{M}^T)^3 & \mathbf{M}^6\mathbf{C}(\mathbf{M}^T)^6
\end{pmatrix}.
\tag{7.61}
$$

In Section 7.2.2 we showed that since the matrices $\mathbf{M}$ and $\mathbf{C}$ are circulant then the coefficients of the matrices in each block of (7.61) can be described by considering just a single row. The 250th row of the diagonal matrices and the off diagonal matrices of (7.61) are shown in Figures 7.3 and 7.2. Figure 7.2 shows that all the coefficients $g_{i,j}$ of the matrices of the form $\mathbf{G} = \mathbf{M}^k\mathbf{C}(\mathbf{M}^T)^k$ for $k = 0, 3, 6$ are positive and monotonically decrease as the separation $|i-j|$ increases. Hence, as the observations of the parameter become more spaced out, we would expect the coefficients of $\mathbf{H}\mathbf{M}^k\mathbf{C}(\mathbf{M}^T)^k\mathbf{H}^T$ to become smaller.

For the off-diagonal matrices of (7.61), $\mathbf{F} = \mathbf{M}^k\mathbf{C}$ $k = 3, 6$, the coefficients $f_{i,j}$ are monotonically decreasing as the separation, $|i - j|$, increases with $|i - j| > d$ where

$$
d = \frac{ak\Delta t}{\Delta x}.
\tag{7.62}
$$

In our experiment $d = ak\Delta t/\Delta x = 0.3 \times 6 = 1.8$ grid points. Hence, as long as the separation between observations is larger than 2 grid points, increasing the spacing between observations decreases the size of the coefficients of $\mathbf{H}\mathbf{M}^k\mathbf{C}(\mathbf{M}^T)^l\mathbf{H}^T$ $k, l = 0, 3, 6$. Since the coefficients of the reduced 4D background error covariance matrix $\hat{\mathbf{H}}\mathbf{C}\hat{\mathbf{H}}^T$ decrease with increasing separation of the observations, by the bounds (7.47) in Theorem 7.2.1, we expect the condition number of $\hat{\mathbf{S}}$ to decrease as the observation spacing increases. We illustrate this numerically.
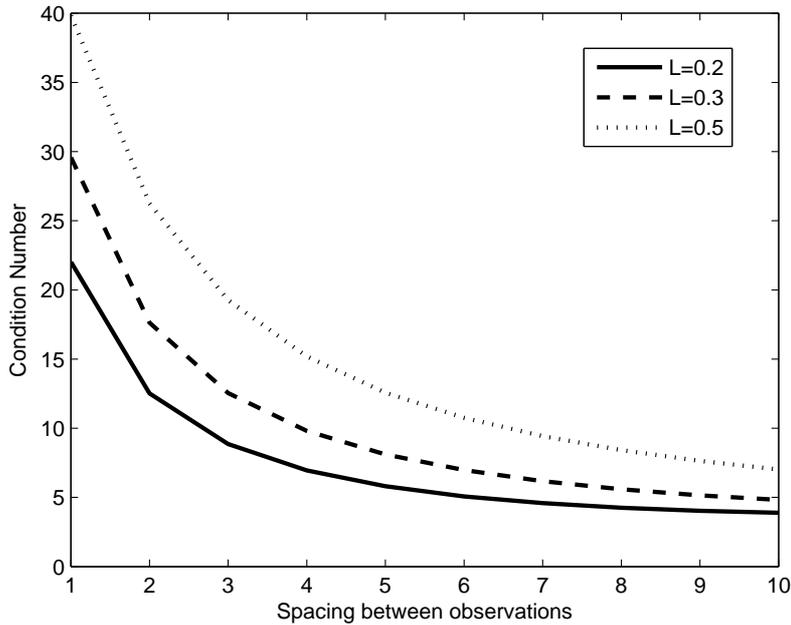
Figure 7.5: Condition number of preconditioned Hessian with changing observation spacing using the SOAR background correlation matrix with different spacing. The solid line represents lengthscale $L = 0.2$, the dashed line represents lengthscale $L = 0.3$ and the dotted line represents the lengthscale $L = 0.5$.

The effect of changing the spacing between these observations is shown in Figure 7.5 for three different lengthscales $L = 0.2, 0.3, 0.5$. As expected the condition number of the P4DVar Hessian decreases as the spacing between the 20 observation increases. The same was also found for the P4DVar Hessian using the Gaussian (5.10) and Laplacian (5.12) correlation matrices but we do not show these results here. We now consider the effect of changing the number of observations on the conditioning of P4DVar.

7.4.2 Number of Observations

There are two main ways of changing the number of observations. We can change the number of grid points we observe or we change the number of time steps we observe. First consider changing the number of grid points that we observe at. Let us assume that we take p observations of the parameter at grid points only and that we observe the parameter at the same grid points

for each time step $t_0, t_1, \ldots, t_n$. Therefore $\mathbf{H}_j = \mathbf{H}$ for $j = 0, 1, \ldots n$. Now if we remove r observations at each time step (at the same positions at each time step) then each $p \times p$ matrix $\mathbf{HM}^k\mathbf{C}(\mathbf{M}^T)^l\mathbf{H}^T$ will be reduced to a $(p - r) \times (p - r)$ matrix and the reduced background error covariance matrix $\hat{\mathbf{H}}\mathbf{C}\hat{\mathbf{H}}^T$ will be reduced from a $p(n + 1) \times p(n + 1)$ matrix to a $((p - r)(n + 1)) \times ((p - r)(n + 1))$. If the entries of the 4D background covariance matrix are positive this will result in a reduction in the sums of the components of the reduced background covariance matrix and therefore a reduction in the bounds (7.47) on the condition number of the P4DVar Hessian $\hat{\mathbf{S}}$.

Now consider the effect of changing the number of time steps that we observe. If we stop taking observations at a certain time t_m then any matrix block of the form $\mathbf{HM}^k\mathbf{C}(\mathbf{M}^T)^l\mathbf{H}^T$, where either k or l is equal to m, is removed from the reduced 4D background error covariance matrix. If we assume that the entries of the 4D background covariance matrix are positive then reducing the number of times that we observe results in a reduction in the sums of the components of the reduced background covariance matrix and therefore a reduction in the bounds (7.47) on the condition number of the P4DVar Hessian $\hat{\mathbf{S}}$. Conversely by observing at more time steps will produce a increase in the condition number of the bounds on $\hat{\mathbf{S}}$. We therefore expect that the condition number of $\hat{\mathbf{S}}$ to decrease when fewer observations, are made, either temporally or spatially. We now test these hypothesis experimentally.

We consider a single periodic parameter on the real line. We use the same experimental design as in the previous section except we fix the lengthscale as $L = 0.2$ and we change the observation configuration. For our discrete model, $\mathbf{M}$, we see that the 4D background error covariance matrix only has positive components as shown by Figures 7.3 and 7.2 in Section 7.2.2. We begin by testing the effect of removing observations at grid points. We observe the same observation distribution at each time step $t_0 = 0, t_1 = 3\Delta t, 6\Delta t$. We begin by observing 160 (giving a total of 480 observations) of the 500 grid points with 2 grid points between the observations. We then observe every

other of these grid points (Giving a total of $320 = 80 \times 3$ observations), then every other of the remainder ($120 = 40 \times 3$ observations) and so on. Table 7.2 shows the results. As we expected, the fewer grid points we observed the smaller the condition number. The same result was found when the Gaussian (5.10) and Laplacian (5.12) correlation matrices were used to construct the matrix $\mathbf{B}$ but we do not present the results here.

Number of Observations	Condition number
480	13.00
320	7.01
120	4.26
60	3.61
30	3.59

Table 7.2: Change in the condition number of the preconditioned Hessian using the SOAR correlation matrix as a function of the number of observations.

We now consider the effect of changing the number of time steps that we observe. We observe at 20 random grid points using the same observation configuration as the numerical experiments performed in Section 7.2.2. We either observe at three time steps $0, 3\Delta t, 6\Delta t$ giving a total of 60 observations or at seven time steps, $0, \Delta t, 2\Delta t, 3\Delta t, 4\Delta t, 5\Delta t, 6\Delta t$ giving a total of 140 observations. With a lengthscale of $L = 0.2$ the condition number increased from 6.06 for observations at three time steps to 12.97 with observations at seven time steps. Similarly, for lengthscale $L = 0.5$ the condition number increased from 6.87 to 14.71. This confirms that thinning the observations in both time and space reduces the condition number of the P4DVar Hessian (7.46) as expected. In the next section we briefly consider the implications of the condition number on the convergence rate of the conjugate gradient method applied to the unpreconditioned and preconditioned Var schemes.

7.5 Convergence Rates

In Chapter 6 we showed that the condition number was a reasonable indicator of the convergence rate of the conjugate gradient (CG) method applied to solve the unpreconditioned and preconditioned 3DVar systems. In this section consider the convergence rate of CG applied to the 4DVar and P4DVar systems.

As shown in Section 3.2, minimising a quadratic cost function is equivalent to solving a linear system. Minimising the 4DVar cost function (2.11) is equivalent to finding the solution, $\mathbf{w}$, of the linear system

$$\mathbf{S}\mathbf{w} = \mathbf{b}, \tag{7.63}$$

where $\mathbf{S} = \mathbf{B}^{-1} + \hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}$ is the Hessian of the 4DVar cost function. Similarly minimising the preconditioned 4DVar scheme is equivalent to solving the linear system

$$\hat{\mathbf{S}}\mathbf{v} = \hat{\mathbf{b}}, \tag{7.64}$$

for $\mathbf{v} = \mathbf{B}^{-1/2}\mathbf{w}$ where $\hat{\mathbf{S}} = \mathbf{I} + \mathbf{B}^{1/2}\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{B}^{1/2}$ is the Hessian of the preconditioned 4DVar cost function (2.16) introduced in Section 2.4, $\mathbf{v} = \mathbf{B}^{-1/2}\mathbf{w}$ and $\hat{\mathbf{b}} = \mathbf{B}^{1/2}\mathbf{b}$. We solve (7.63) using the Matlab CG method pcg.m which is equivalent to the CG algorithm given in Section 3.3.1. We also solve (7.64) using the Matlab CG method pcg.m but using a preconditioned version of the CG method for (7.63) similar to that given in Section 3.3.2. This is equivalent to solving (7.64) directly using the normal CG algorithm given in Section 3.3.1.

The stopping criteria for both algorithms is when either the maximum number of iterations of the CG method have been used (in our case 2000) or when the relative residual

$$\frac{||\mathbf{b} - \mathbf{S}\mathbf{w}_k||}{||\mathbf{b}||} = \frac{||\mathbf{r}_k||}{||\mathbf{b}||}, \tag{7.65}$$

has reduced to less then 10^{-6}. Here $\mathbf{w}_k$ is the estimate of the true solution to (7.63) at the kth iteration of the CG method. We will also compute the relative error in the solution $\mathbf{w}^*$ to test that our system has converged. This is simply the ratio of the magnitudes of the error and the truth

$$\frac{||\mathbf{w}^* - \mathbf{w}^a||}{||\mathbf{w}^*||}, \tag{7.66}$$

where $\mathbf{w}^a$ is the final solution found by the CG method.We refer to the relative error as the accuracy throughout this section.

We consider a single periodic parameter defined on a 1D domain for the remainder of this section. Consider $N = 500$ uniformly spaced grid points on the real line with spacing $\Delta x = 0.1$. For the true solution we choose

$$\mathbf{w}^*(s) = 2\sin(qs) + \cos(3qs) - 0.3\sin(125qs), \tag{7.67}$$

defined at the grid points $s = k\Delta x$ for $k = 0, 1, \ldots, (N-1)$ where $q = \frac{2\pi}{N\Delta x}$. The true solution is the same as used in the 3DVar experiments in Section 6.4.1 and is displayed in Figure 6.12. We begin by comparing the convergence rates of the CG to solve 4DVar with and without the control variable transform (CVT).

7.5.1 Convergence 4DVar vs P4DVar

In this section we compare the convergence rates of the CG methods applied to solve the unpreconditioned and preconditioned systems (7.63) and (7.64) respectively. We choose the background error covariance $\mathbf{B} = \sigma_b^2\mathbf{C}$ where $\mathbf{C}$ is the SOAR correlation matrix (5.11) and $\sigma_b^2 = 1$ is the background error variance. For our numerical forecast model we use the advection model $\mathbf{M}$ defined by (7.30) using a time step $\Delta t = 0.1$ and $a = 0.3$, therefore $\nu = a\frac{\Delta t}{\Delta x} \in (0, 1)$ and as mentioned in Section 7.1.2, this means the numerical scheme is stable. We take observations at 20 grid points and at the same

positions at the time steps $0, 3\Delta t, 6\Delta t$. We use the same configuration of observations as in Sections 7.1.3 and 7.2.2 and choose an observation error variance of $\sigma_o^2 = 1$. We showed in Sections 7.1.3 and 7.2.2 that with this configuration the preconditioned system was much better conditioned then the unpreconditioned system. We also found that the condition number of both of the 4DVar and P4DVar Hessians, $\mathbf{S}$ and $\hat{\mathbf{S}}$ respectively, increased as the lengthscale increased. We therefore expect the CG method to converge much faster for the preconditioned system than for the unpreconditioned. We also expect that the CG method converges slower as the lengthscale increases.
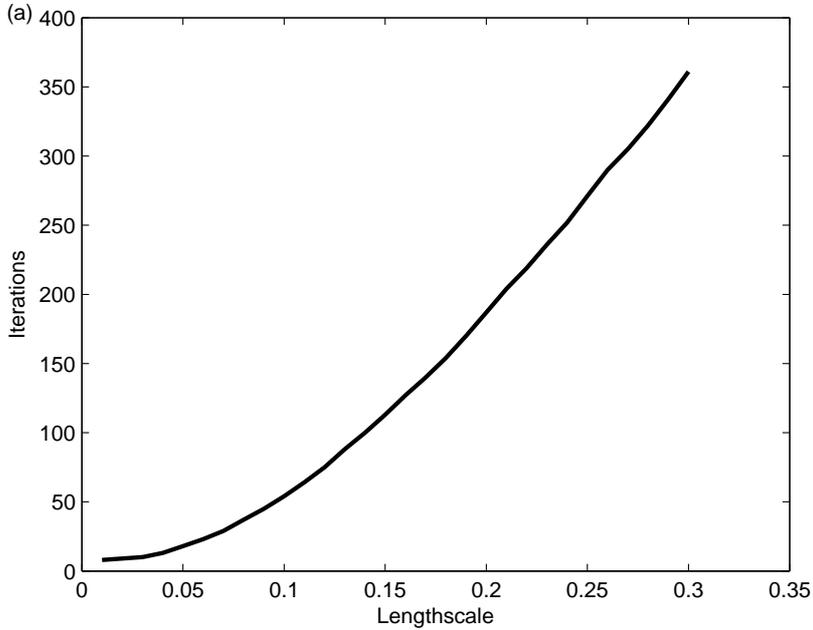
Figure 7.6: The effects of different lengthscales on the convergence of CG for unpreconditioned 4DVar for the SOAR correlation matrix.

We tested this hypothesis on our simple periodic system. In the experiments the final solutions had relative errors of $10^{-6} - 10^{-7}$ and have therefore sufficiently converged. Comparing, Figures 7.6 and 7.7 confirm that the unpreconditioned system converges slower then the unpreconditioned system. For instance, at lengthscale $L = 0.2$, the CG method takes around 150 iterations when applied to converge for the unpreconditioned system but 12 iterations for the preconditioned system. Additionally, in both systems, increasing the lengthscale increases the condition number and corresponds with the gen-
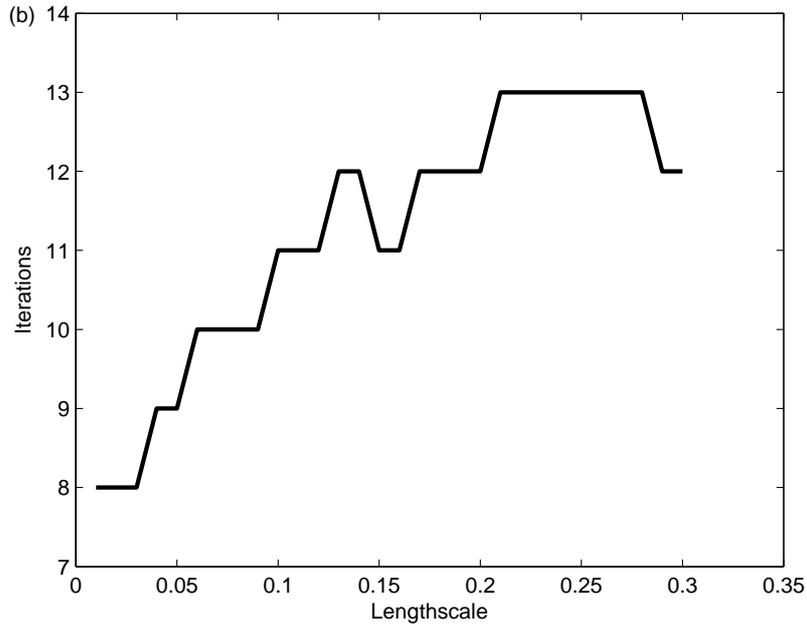
Figure 7.7: The effects of different lengthscales on the convergence of CG for preconditioned 4DVar for the SOAR correlation matrix.

eral increasing in the number of iterations required for CG to converge. We now consider how the convergence rate of the CG method varies for solving the preconditioned system when we change the accuracy and spacing of observations.

7.5.2 Observation Variance

In Section 7.3 we showed that the condition number of the Hessian $\hat{\mathbf{S}}$ decreased as the error variance increased. In this section we briefly consider the effect of the error variance on the convergence rate of the CG method applied to solve (7.64). We use the same design as the previous section except now we fix the lengthscale at $L = 0.2$ and allow the observation error variance to vary. As before the relative error in the final solution was of the magnitude of $10^{-6} - 10^{-7}$. Figure 7.8 shows that the convergence rate increases as we reduce the accuracy of the observations. This is as expected since a decrease in accuracy resulted in a increase in the condition number.
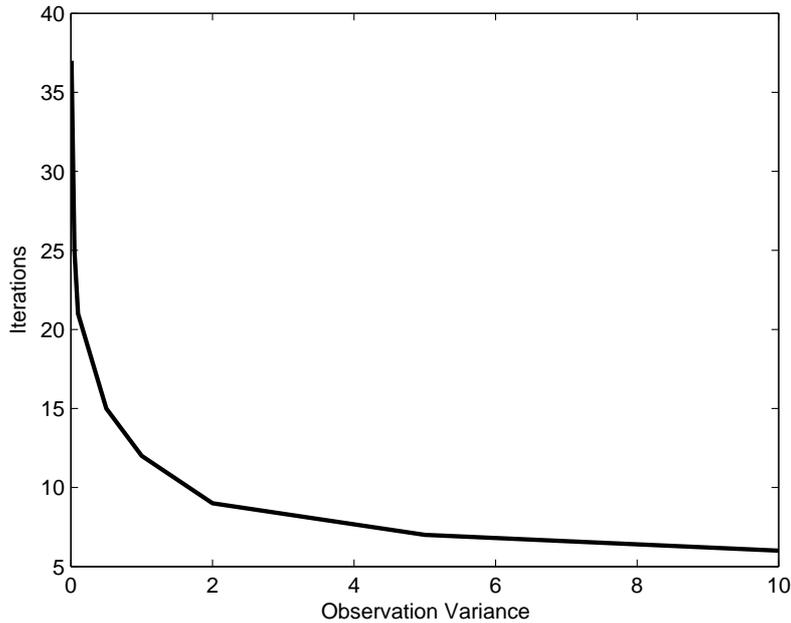
Figure 7.8: The effects of the observation error variance on the convergence rate of the CG method for preconditioned 4DVar using the SOAR correlation matrix.

We now consider the effect of the observation spacing on the convergence rate of the CG method applied to solve the P4DVar system.

7.5.3 Observation Distribution

In Section 7.4 we showed that the observation distribution, both spatially and temporally, is a major factor in determining the condition number of the P4DVar Hessian $\hat{\mathbf{S}}$. In particular, we showed as we increased the spacing of the observations the condition number of $\hat{\mathbf{S}}$ decreased. Additionally we found that by reducing the number of time steps that we observe we also reduce the condition number of $\hat{\mathbf{S}}$. We therefore expect that the convergence rate of the CG method used to solve the linear system (7.64) will increase as we increase the spacing of the observations and observe at fewer time steps. To test this hypothesis we perform a simple experiment using a single periodic parameter on a 1D domain. We fix $N = 500$ grid points with equal spacing $\Delta x = 0.1$. We use the advection model discretised using the upwind scheme

142

as the forecast model $\mathbf{M}$ with a time step of $\Delta t = 0.1$ and $a = 0.3$. We choose the background error covariance $\mathbf{B} = \sigma_b^2 \mathbf{C}$, where $\mathbf{C}$ is the SOAR correlation matrix (5.11) and $\sigma_b^2 = 1$ is the background error variance. We assume that our observation errors are uncorrelated with observation error variance $\sigma_o^2 = 1$. First we make observations of the parameter at three times steps $0, 3\Delta t, 6\Delta t$ and take 20 observations on the initial 20 grid points of the domain. We consider the effect on the convergence rate of the CG method as we increase the spacing between the observations. Figure 7.9 shows the
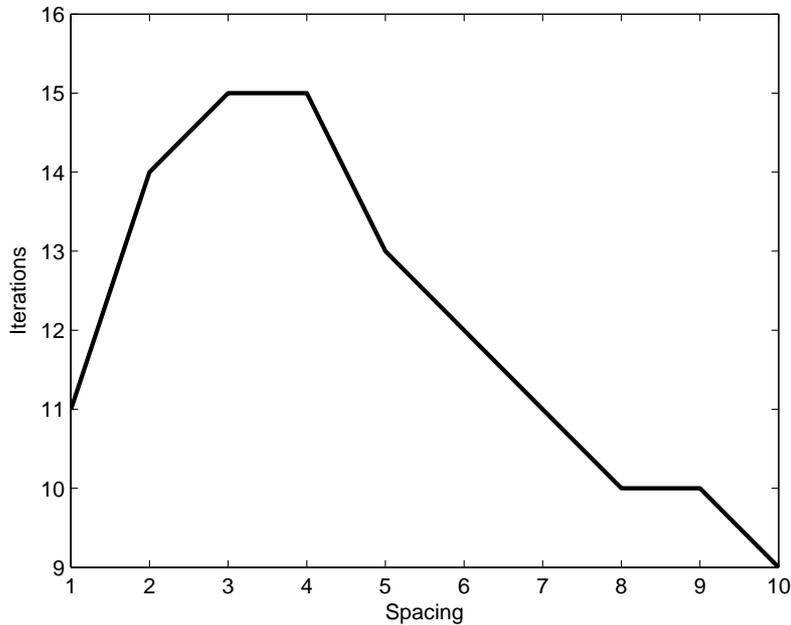
Figure 7.9: The effects of the observation spacing on the convergence rate of the CG method for preconditioned 4DVar using the SOAR correlation matrix.

effect on the convergence rate of the CG method applied to solving the linear system (7.64) as we change the spacing between these observations using a lengthscale $L = 0.2$. For other lengthscales the pattern is similar. As the spacing increases the convergence rate initially decreases but then begins to increase as the spacing becomes larger then $3\Delta x$. The initial decrease in the convergence rates of CG as the spacing increases from Δx to $3\Delta x$ is contrary to what we would expect since from Δx to $3\Delta x$ the condition number decreases as seen from Figure 7.5. We have not been able to explain this behaviour.

We also consider the effect of changing the number of time steps that we observe. We use the same experimental design as for the spacing experiment used in this section but now use a random distribution of 20 observations as used in Section 7.5.1. We consider either observing at three time steps $0, 3\Delta x, 6\Delta x$ giving a total of 60 observation or at 7 time steps $0, \Delta x, 2\Delta x, 3\Delta x, 4\Delta x, 6\Delta x, 6\Delta x$, giving a total of 140 observations. For a lengthscale $L = 0.2$ the condition number increased from 6.06 to 12.97 when we observed more frequently. We found the number of iterations of the CG method required for convergence increased from 12 to 15 when we increased the number of observations. Similarly, for $L = 0.5$ the number of iterations required for the CG method to converge increased from 12 to 16 when we observed more frequently with the corresponding change in the condition number of 6.87 to 14.71. Hence, in general increasing the spacing of the observations and reducing the number of observations increases the convergence rate of the CG method. We now summarise this chapter.

7.6 Summary

In this chapter we investigated the condition number of the 4DVar system. We have sought to analyse the effect of the forecast model on the conditioning. In particular, we derived new theoretical bounds on preconditioned and unpreconditioned 4DVar. In the unpreconditioned systems we confirmed that, just like the 3DVar system, the conditioning is related to the conditioning of the background covariance matrix. We confirmed this experimentally using a linear advection equation as our forecast model. We showed that the as the lengthscale increased, the conditioning of the background error covariance matrix dominated the conditioning of the 4DVar Hessian. Therefore for large lengthscales the 4DVar Hessian is ill-conditioned.

We derived theoretical bounds on the preconditioned 4DVar system, preconditioned using the control variable transform. We showed that the condition

number of the P4DVar Hessian was dependent on the sum of components of a reduced 4D background error covariance matrix. This matrix is very dependent on the forecast model employed and therefore is difficult to analyse in general. We considered the matrix using a discretised advection model as our forecast equation. In order to understand the conditioning of the P4DVar Hessian we experimented using a single parameter periodic system. In our numerical experiments we found that preconditioning with the background error covariance matrix produced a large reduction in the condition number of the 4DVar Hessian.

By analysing the action of the advection model on the background error covariance matrix we could also predict the behaviour of the conditioning of P4DVar as we changed the observation errors and positions. Using numerical experiments on a periodic one-parameter 1D domain we found that the condition number of the P4DVar Hessian was reduced by reducing the number of observations, increasing the observation error variance and increasing the spacing between observations. These results appear to be counter-intuitive since we have shown that by increasing the accuracy of the observations, and having more observations, the solution to the 4DVar problem is less accurate (indicated by the increase in the condition number). This may be explained by the fact that highly accurate observations and more observations put tighter restrictions on the optimisation problem and so the problem becomes more difficult to solve and accurately satisfy the constraints. It is difficult to predict the effect of the observation accuracy and distribution on the condition number of the Hessian for a general forecast model, this will depend on the form of the coeffiecients of the 4D background error covariance matrix.

Finally we analysed the convergence rate of the CG method as applied to linear problems associated with solving the 4DVar and P4DVar minimisations. In general, the convergence rate results coincided with the condition number results. That is, a reduction in the condition number caused a increase in

the convergence rate of the CG method. A large improvement was found in the convergence rate of the CG method applied to the 4DVar when the system was preconditioned with the CVT. In the preconditioned system the most significant changes in the convergence rates are a result of changing the accuracy of the observations.

Now we have analysed the system theoretically and in simple numerical experiments for both 3DVar and 4DVar we now consider the conditioning of the Met Office operational Var scheme.

Chapter 8

Operational results

Many operational numerical weather prediction (NWP) centres utilise a variational data assimilation scheme in order to find the analysis. The Var scheme is usually implemented with a control variable transform (CVT) in order to model the background error covariance matrix. As shown in Chapters 6 and 7 the CVT has the added beneficial effect of improving the conditioning of the Var problem. The CVT has been shown to produce a significant reduction in the computational cost of the minimisation [42]. In Chapters 6 and 7 we derived new algebraic bounds on the condition number of the preconditioned Var Hessian. These bounds, together with simple numerical experiments, showed that the condition number of the preconditioned Hessian is reduced by

1. Using less accurate observations.

2. Increasing the spacing between observations.

3. Using less observations.

In this chapter we show that these results also apply in an operational setting in the Met Office global Var scheme. We show this with experiments using both real and pseudo observations.

In the ECMWF operational system it was found that doubling the error in the surface observations halved the condition number of the Var Hessian [63]. This is in rough agreement with the approximation to the condition number given in [6] for a simple system and with our theoretical results in Chapters 6 and 7 which showed that the conditioning of preconditioned Var problem is linearly related to the accuracy of the observations. Tremolet showed that the surface observations dominated the conditioning of the ECMWF Var scheme. It was suggested that this was caused by the large number of dense accurate surface observations around Europe [63]. This hypothesis was supported by a plot of the components of the leading eigenvector which showed that the components with largest magnitude were centred over Europe. This conclusion is also supported by the theory in Chapters 6 and 7 were we showed that reducing the spacing between observations produced a larger condition number of the preconditioned Var Hessian. In this chapter we also consider the condition of the Var Hessian when different observations types are assimilated in the Met Office scheme and show that the surface observations dominate the conditioning of the scheme. We begin this chapter with an overview of the Met Office (MO) system including the minimisation schemes, the observations and the covariance matrices. We then consider the condition number of the system using pseudo observations. Next we look at real observations. We begin by considering how different observations types affect the condition number and then we consider how thinning observations can improve the Var conditioning.

8.1 The Met Office Variational Data Assimilation Scheme

In this section we give an overview of the Met Office variational data assimilation scheme.

8.1.1 Constructing the Cost Function

The UK Met Office currently implements an incremental form of the variational data assimilation scheme using a control variable transform (CVT) (see Section 2.4) [54], [43]. The background state, $\mathbf{x}_b$ is used as an initial guess and is usually found from an earlier 6 hour forecast [54]. The CVT transforms the increment $\delta\mathbf{x}_0$, with background error covariance matrix $\mathbf{B}$, to new variables, $\delta\mathbf{z}_0$, which have uncorrelated errors. This provides a method for modelling the background error covariance matrix without forming it explicitly. The Met Office CVT, denoted $\mathbf{U}$, can be described by a series of transforms [10], [43]

$$\mathbf{U} = \mathbf{U}_p\mathbf{U}_v\mathbf{U}_h, \tag{8.1}$$

where the subscript p represents a parameter transform, and v and h represent vertical and horizontal spatial transformations respectively. In order to understand the CVT it is useful to consider the inverse of the CVT (8.1), which describes the transformation from the increment $\delta\mathbf{x}_0$ to the new control variables $\delta\mathbf{z}_0$

$$\delta\mathbf{z}_0 = \mathbf{U}^{-1}\delta\mathbf{x}_0 = \mathbf{U}_h^{-1}\mathbf{U}_v^{-1}\mathbf{U}_p^{-1}\delta\mathbf{x}_0. \tag{8.2}$$

We briefly describe each of the transforms used to construct the CVT. In the Met Office CVT, the parameter transform $\mathbf{U}_p^{-1}$ transforms from parameters of vertical and horizontal wind velocity, potential temperature, moisture density, pressure and specific humidity to new control variables of stream function, velocity potential, unbalanced pressure and relative humidity which are considered to have uncorrelated errors between parameters (see [10] and [3]). With respect to these new parameters the background error covariance matrix can be written in block diagonal form since there exist no inter-parameter correlations between the errors. For each of the new parameters a vertical transform, $\mathbf{U}_v^{-1}$, is applied and removes spatial correlations in the vertical for each fixed horizontal (latitude-longitude) position on the globe. This is achieved by projecting onto the orthogonal, uncorrelated modes of the vertical error covariance matrix of each parameter. Similarly, the horizontal

transform $\mathbf{U}_h^{-1}$ projects the errors in the horizontal directions at each constant vertical mode onto the orthogonal, uncorrelated modes of the horizontal error covariance matrix. These spatial transforms are essentially spectral transforms onto the eigenvectors of the error covariance matrix and must be calibrated so as to fit with the assumed errors within the background state (see [10] for further details). In the Met Office system a 2D version of the SOAR background correlation model described in Chapter 5 is used to model the horizontal correlations [1], [4] [28], [8]. Since the transforms are approximations to the actual error correlations then they provide only an estimate of the true background error covariance matrix

$$\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{T/2} \approx \mathbf{U}\mathbf{U}^T = \mathbf{U}_p\mathbf{U}_v\mathbf{U}_h\mathbf{U}_h^T\mathbf{U}_v^T\mathbf{U}_p^T. \qquad (8.3)$$

This estimate $\mathbf{U}\mathbf{U}^T$ is called the *implied background error covariance matrix.*

Observations are collected from a variety of sources including surface observations, aircraft, sondes and satellites with various instruments (E.g. ATOVS, SSMI, IASI etc.) and are assimilated over a 6 hour data window [54]. The errors are almost all considered to be uncorrelated and therefore the observation error covariance matrices $\mathbf{R}_i$ are diagonal with observation error variances as the diagonal elements.

8.1.2   Minimising the Cost Function

In this section we outline the minimisation scheme used within the Met Office. Currently the Met Office performs only one outer loop update and therefore the operational cost of Var is associated with the inner-loop minimisation (see Section 2.4). There are two iterative methods available for the Var minimisation: the limited Memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) or the conjugate gradient (CG) method [5]. In this chapter all minimisation experiments will involve the CG method equivalent to the algorithm described in Section 3.3. An advantage of this method is that we

can obtain estimates of the leading eigenvector and eigenvalue of the Hessian
for little extra cost via the Lanczos method [23, Sec. 10.2]. Since there are
fewer observations than the number of elements in the state vector the small-
est eigenvalue of the Hessian is unity and therefore the condition number is
simply the largest eigenvalue of the Hessian. Thus the Lanczos method can
be used to calculate the condition number of the Hessian.

The stopping criteria for the minimisation is determined either by achieving
a sufficient reduction in the cost function gradient

$$\frac{|\nabla^{(k)} J|}{|\nabla^{(0)} J|} < \epsilon, \tag{8.4}$$

for some tolerance $\epsilon > 0$ at the kth iteration as suggested in [36] or after a
prescribed maximum number of iterations of the gradient solver have been
applied. The experiments are performed on the NEC SX6 supercomputer and
some on the more recent IBM POWER6. We now consider the conditioning
of the operational Var scheme in the case of assimilating pseudo observations.

8.2 Pseudo Observation Results

In Chapters 6 and 7 we showed that the condition number of the Hessian of
P3DVar and P4DVar is linearly related to the accuracy of the observations
and the spacing of the observations. In this section we show that the results of
the theory of Sections 6.3, 7.3 and 7.4 also apply in the Met Office operational
system in the case of pseudo observations. The Met Office variational schemes
allow the inclusion of a maximum of 20 pseudo observations whose position,
values and errors can be defined by the user. Hence unlike real observations
we have the ability to adjust key properties of the observations assimilated.
Pseudo observations can be positioned at any point of the global grid and
at any 40 minute interval throughout a 6 hour time window together with
their observation error variances σ_o^2. In our experiments we use the Met

	Condition Number	
Error Variance	3DVar	4DVar
0.01	15242150	18078035
0.1	1524483	1808122
1.0	152422	180781
10	15243	18078
25	6098	7232
50	3050	3618
75	2033	2412
100	1525	1809

Table 8.1: Change in condition number of the preconditioned Var Hessian with change in the error variance of pseudo observations in 3DVar and 4DVar, see text for observation distributions.

Office $N108L38$ grid which means we consider $L = 38$ vertical levels, $2N$ latitudinal points and $\frac{3}{2}N - 1$ longitudinal points where $N = 108$. This equates to about 5/3 degrees between grid points in the East-West direction and 10/9 degrees in the North-South direction.

The value of the innovation vectors $\mathbf{d}_i = \mathbf{y}_i - \mathcal{H}_i(\mathcal{M}(t_0, t_i, \mathbf{x}_b))$ at each time observed can also be specified, where $\mathcal{M}(t_0, t_i, \mathbf{x}_b)$ is the full non-linear model which evolves the background state from time t_0 to t_i as introduced in Section 2.2. By specifying the innovations explicitly the background state $\mathbf{x}_b$ is not necessary. We begin our investigation of the operational system by considering how the observation errors can affect the condition number of the preconditioned 3D and 4DVar Hessians.

8.2.1   Observation Error Variance

In Sections 6.3.1 and 7.3 we showed that the Hessians of the preconditioned 3DVar and 4DVar systems are linearly related to the inverse of the observation error variances. In this section we show that these results also apply to the Met Office Var Hessian. In the 3DVar case we define a set of 16 surface pressure observations arranged in a 4-by-4 grid with one grid spacing between each row and column. The observations are positioned to be approximately

over the U.K.

In the 4DVar case, 6 surface pressure observations are observed at the start, middle and end of a 160 minute time window with the observations in a 2-by-3 grid with one grid spacing between each row and column. These are similarly placed close to the U.K. In both 3D and 4DVar the innovation vectors are fixed at 100 Pa but are not part of the Hessian and thus have no affect on the condition number. In our experiments the observation positions are kept fixed and the observation error variance is varied in the range 1 Pa^2 to 100 Pa^2. Table 8.1 shows the change in the condition number of the preconditioned Hessian of 3D and 4DVar as the observation error varies. The table confirms that the condition number is approximately inversely proportional to the observation accuracy. For example, increasing the observation variance by a factor of 10 from 1 Pa^2 to 10 Pa^2 reduces the 3DVar condition number by a factor of 10 from 152422 to 15243 and in 4DVar reduces the condition number from 180781 to 18078.

Since the inverse of the Hessian is the analysis covariance matrix, the reciprocal of the eigenvalues of the Hessian are eigenvalues of the covariance matrix. The eigenvalues of the covariance matrix correspond to the error variances in the direction of the corresponding eigenvector. Hence the most accurate data (smallest error variance) is in the direction of the eigenvector corresponding to the largest eigenvalue of the Hessian. This is illustrated by a plot of the surface pressure components of the leading eigenvector of the Hessian for the above 3DVar observations distribution and observation error variance $\sigma_o^2 = 1$ Pa^2. Figure 8.1 shows that the components around the observations have the largest magnitude and correspond to the most accurate (and in this case only) observations. We now consider the effect of observation spacing on the condition number.
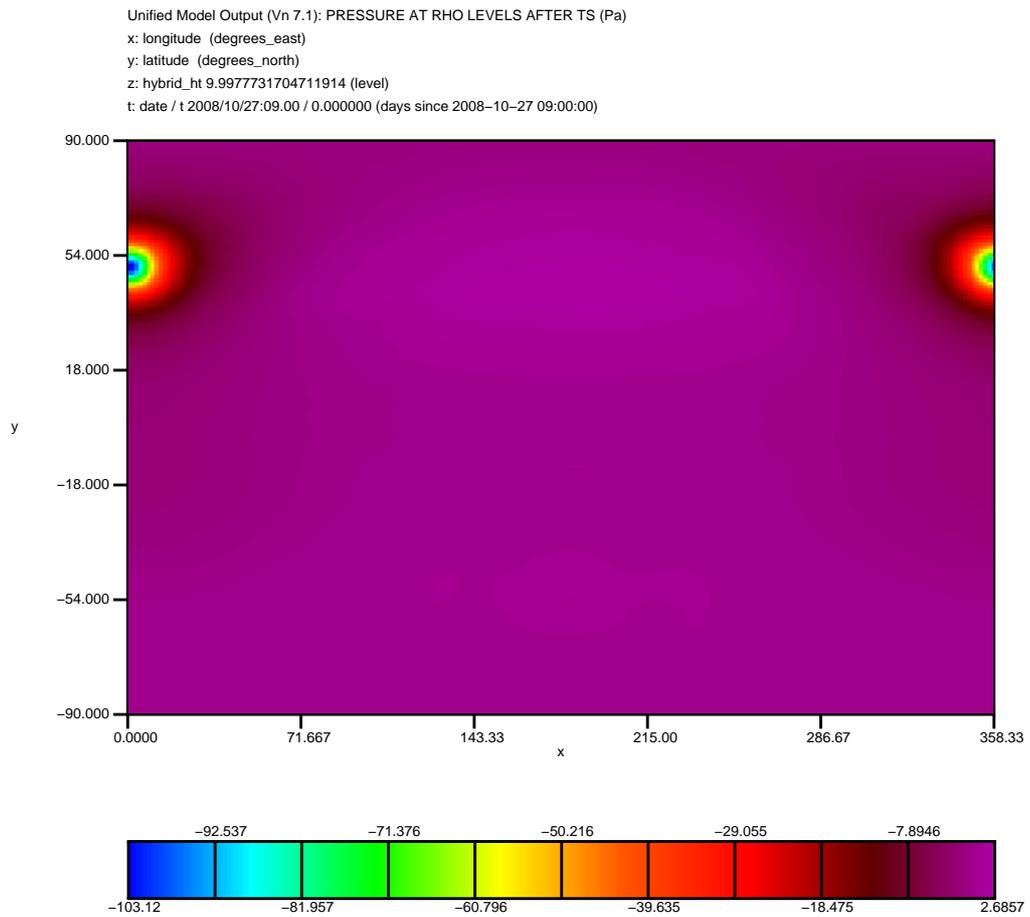
Figure 8.1: Surface pressure components of the leading eigenvector of the 3DVar Hessian using pseudo observations in a 16-by-16 grid over Europe with error variance 1 Pa^2.

	Condition Number	
Obs Config.	$\sigma_o^2 = 25$	$\sigma_o^2 = 100$
1	6098	1525
2	4615	1155
3	2756	690
4	1285	322
5	825	207
6	652	163

Table 8.2: Condition number of the P3DVar Hessian as function of the change in spacing of pseudo observations. The details of the observation configuration are given in the main text.

	Condition Number	
Obs Config.	$\sigma_o^2 = 25$	$\sigma_o^2 = 100$
1	7232	1809
2	6218	1555
3	4240	1061
4	2260	566
5	1607	402
6	1376	344

Table 8.3: Condition number of P4DVar Hessian as a function of the change in spacing of pseudo observations. Details of the observation configuration are given in the main text.

8.2.2 Observation Spacing

In Chapters 6 and 7 we showed that increasing the spacing of observations reduced the condition number of the P3DVar and P4DVar Hessians. We now show that the same results apply to the operational system in the case of assimilating pseudo observations.

Consider the same 4-by-4 3DVar and 2-by-3 4DVar pseudo surface pressure observation distribution as described in Section 8.2.1 and consider experiments using two fixed observation error variances $\sigma_o^2 = 25$ Pa^2 or $\sigma_o^2 = 100$ Pa^2 but varying the observation spacing. Keeping one observation fixed roughly above the U.K. we then increase spacing between observations in adjacent rows and columns. We refer to the original position of the observations as observation configuration 1. We then consider spacings of 2, 4, 8, 12 and 16 grid lengths which we label Configuration 2,3,4,5 and 6 respectively.
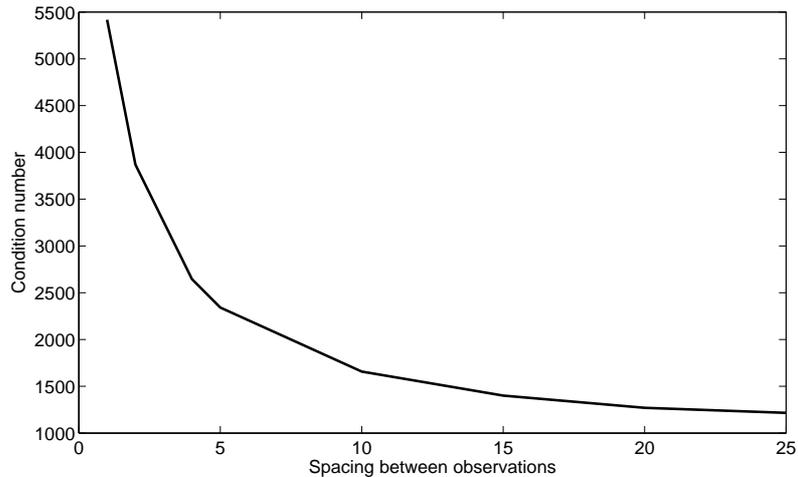
Figure 8.2: Condition number of the Hessian of the Met Office 3DVar scheme as a function of spacing of 8 equatorial pseudo observations.

Tables 8.2 and 8.3 shows the impact of the changing the observation spacing for 3D and 4DVar respectively for two different observation error variances. In agreement with the results in sections 6.3.2 and 7.4 an increase in the spacing gives a reduction in the condition number. For instance, with $\sigma_o^2 = 25$ Pa^2, increasing the spacing from 4 to 12 gridlengths (configuration 3 to 5) results in the condition number reducing from 2756 to 825 in the 3DVar case and from 4240 to 1607 in the 4DVar case.

A further 3DVar experiment was performed but this time using eight equally spaced pseudo surface pressure observations positioned along the equator. The observation error variance was fixed at 1 Pa^2 and the spacing was increased. Figure 8.2 shows the change in the condition number of the Hessian as a function of lengthscale. Comparison with Figure 5.1 shows the similarity of the reduction in condition number with spacing and the reduction in the error correlations of the SOAR correlation matrix (5.11) with distance between errors. The reduction in the condition number is rapid but tails off as the spacing increases further. Once again the increase in the spacing of the observations coincides with the decrease in the condition number of the Hessian. We now consider the conditioning of the Met Office operational system using real observations.

156

8.3   Real Observation Results

In this section we consider the condition number of the Hessian of the Met Office variational assimilation scheme using real observations. We begin by considering the conditioning of the Met Office scheme with respect to individual observation types and then we investigate the effect of observation thinning.

8.3.1   Conditioning and Observation Types

In the Met Office 4DVar minimisation scheme observational data is assimilated during a six hour time window from a variety of different sources including aircraft, satellite, sondes and surface data [54]. In [63] it was suggested that the conditioning of the ECMWF Var scheme was dominated by the accurate surface observations which are densely distributed around Europe. In this section we show that the conditioning of the Met Office Var scheme is also dominated by surface observations. We also link this result to the theory in Sections 6.3 and 7.4.

In the experiments in this section the model time steps are at 40min intervals in a 6 hour time window. The minimisation is performed on the Met Office N108L38 global grid. We began by considering observational data from two different dates, 12Z on 27th October 2008 and 12Z on 14th July 2009. Different assimilations were performed in 3DVar and 4DVar using all the observations and then single observation types only.

Figures 8.3 and 8.4 show, for the 3DVar and 4DVar respectively, the condition numbers of the Hessian when all the observations or when only single observation types are assimilated. Some of the data types, such as GPSRO and SSMI are not shown. When assimilated on their own these observation types gave a Hessian with a condition number of less then 50 and are thus not
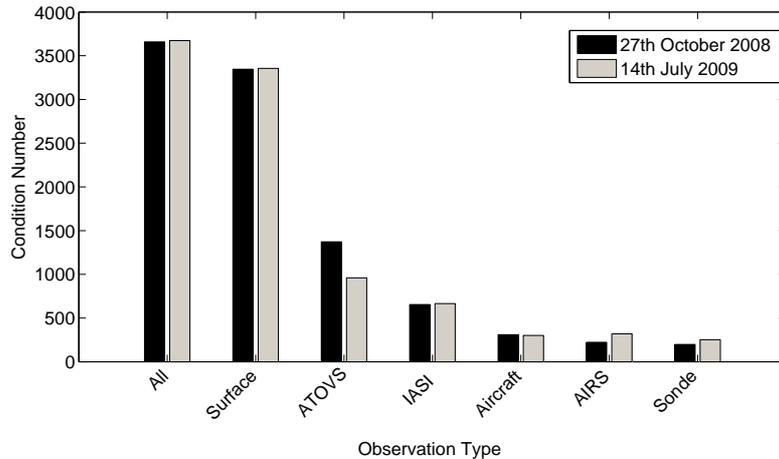
Figure 8.3: Condition number of the Met Office 3DVar scheme when assimilating different observation types.

considered to be major contributors to the conditioning. For the 4DVar case observations are assimilated around the six hour time window centred around 12Z and therefore the analysis is given at the start of the time window at 9Z. We first notice that in both the 3DVar and 4DVar cases the magnitudes of the condition numbers are similar for both days for each observations type. The similarity for the different days indicates that we can expect consistent results regardless of the days we consider.

From both Figures 8.3 and 8.4 it is evident that when assimilating single observation types it is the surface observations that give a Hessian with the largest condition number and that the magnitude of this condition number is similar to the condition number of the Hessian when all the observational data is assimilated. This suggests that, like the ECMWF system, it is the surface observations which dominate the condition number of the Hessian of the Met Office operational Var scheme. To further test this hypothesis further experiments were performed with observation types removed. Tables 8.4 and 8.5 show the conditioning of 3DVar and 4DVar respectively and the effect of removing the surface observations. We also performed experiments with all the observation data assimilated except the ATOVS data, which when assimilated alone, produces a Hessian which is the second most ill-
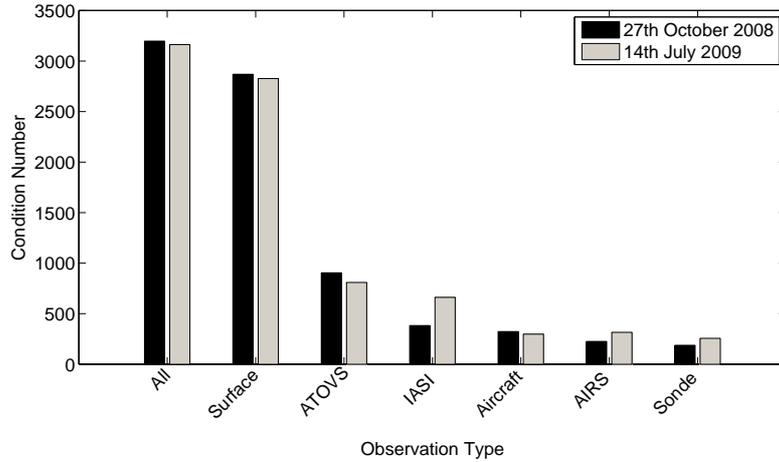
Figure 8.4: Condition number of the Met Office 4DVar scheme when assimilating different observation types.

	Condition Number	
Observation	27/10/2008	14/07/2009
All	3658	3673
Only Surface	3345	3355
Only ATOVS	1372	960
No Surface	1431	1215
No ATOVS	3667	3624

Table 8.4: Conditioning of the Met Office 3DVar Scheme when assimilating different observation types.

conditioned when only single observation types are used. As seen in the tables the removal of the surface observations from the assimilation results in a reduction of around 70% and 60 % in the condition number for the 3D and 4DVar cases respectively compared to when all the observations are assimilated. The condition number of the Hessian is now similar in magnitude to the condition number of the Hessian when only ATOVS data is assimilated. However, in both the 3DVar and 4DVar cases the change caused by removing the ATOVS data is minimal producing no significant effect on the condition number.

According to the theory presented in Chapters 6 and 7, poor conditioning in the preconditioned system can be caused by the inclusion of accurate observations and dense observations (small spacing between the positions of

**Data Coverage: Surface
Pressure, Level 1 (14/7/2009, 12 UTC)
Total number of observations assimilated: 26380**

MISSING (4986) NOT FLAGGED (20911) FLAGGED (483)
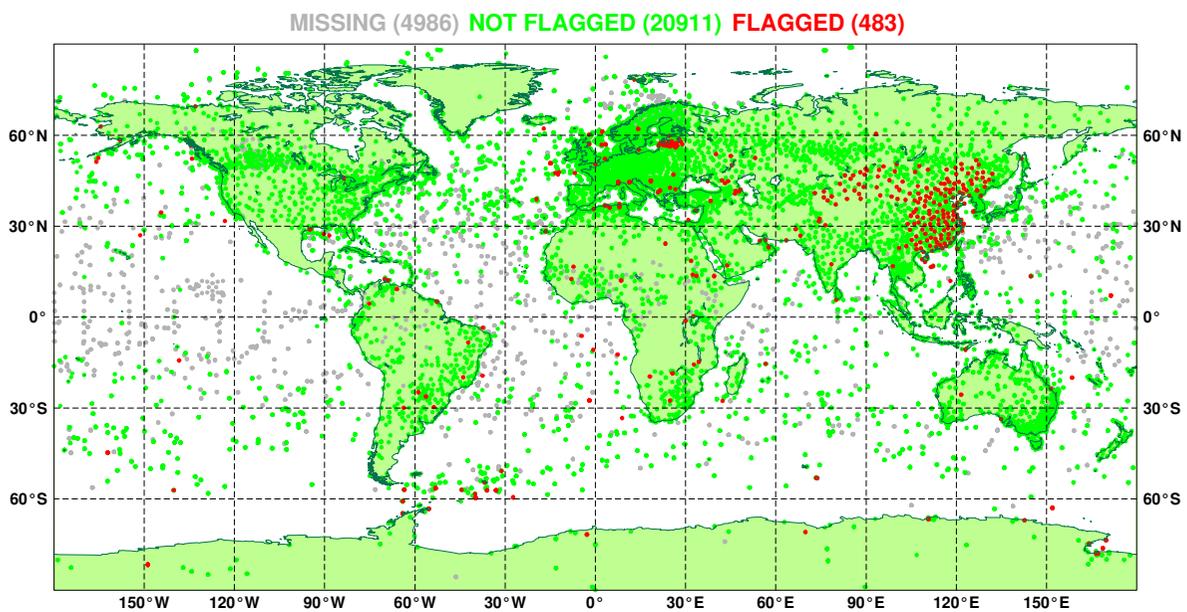
Figure 8.5: Surface Pressure Observations for July 14 2009. Flagged data is not assimilated into the Met Office Var scheme.

	Condition Number	
Observation	27/10/2008	14/07/2009
All	3197	3163
Only Surface	2869	2827
Only ATOVS	905	809
No Surface	972	976
No ATOVS	3146	-

Table 8.5: Conditioning of the Met Office 4DVar Scheme when assimilating different observation types.
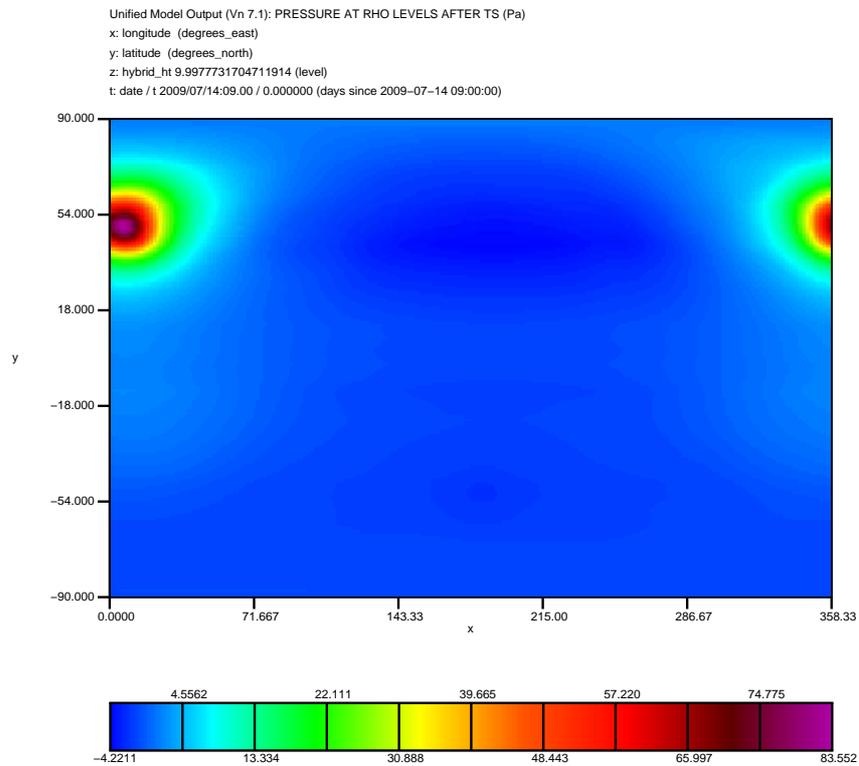
Figure 8.6: Surface pressure components of the leading eigenvector of the 4DVar Hessian produced from assimilating observation data from 14th July 2009.

Time	Condition Number
00Z	2662
06Z	2737
12Z	2770

Table 8.6: Conditioning of the Met Office 4DVar Scheme when assimilating surface observations at different times of 16/07/2009

the observations) in the assimilation. Figure 8.5 shows the distribution of surface pressure observations for the date 14th July 2009. We notice that a high concentration of observations are centred over Europe. Plotting the components of the leading eigenvector of the Hessian (corresponding to the direction with the smallest analysis error variance) shows that the most accurately analysed data coincides with the dense surface data over Europe. For instance, Figure 8.6 shows the distribution of the pressure components of the leading eigenvector of the 4DVar Hessian at the surface of the Earth. The plot clearly shows the components with largest magnitude are centred at approximately 0 degree Latitude and 50 degree Longitude and are concentrated around Europe. A comparison with the leading eigenvector of the Hessian where only the surface observations are assimilated shows the same pattern as Figure 8.6 giving further evidence that the surface observations dominate the condition number of the Hessian.

Finally it is worth mentioning that the magnitude of the condition numbers are consistent when assimilating data from different times of the same day. Table 8.6 shows the condition number of the 4DVar Hessian produced from assimilating surface pressure observational data from the 16th of July 2009. As can be seen from the table the condition number varies very little between the three assimilation times used.

8.3.2 Thinning Observations

In the previous section we concluded that the surface observations dominate the conditioning of the Met Office operational Var scheme. In this section

162

**Data Coverage: Surface
Pressure, Level 1 (11/3/2009, 12 UTC)
Total number of observations assimilated: 25645**

MISSING (4484)　NOT FLAGGED (20627)　FLAGGED (534)

Figure 8.7: The Unthinned surface pressure data from 11 March 2009. Flagged data is not assimilated into the Met Office Var scheme.
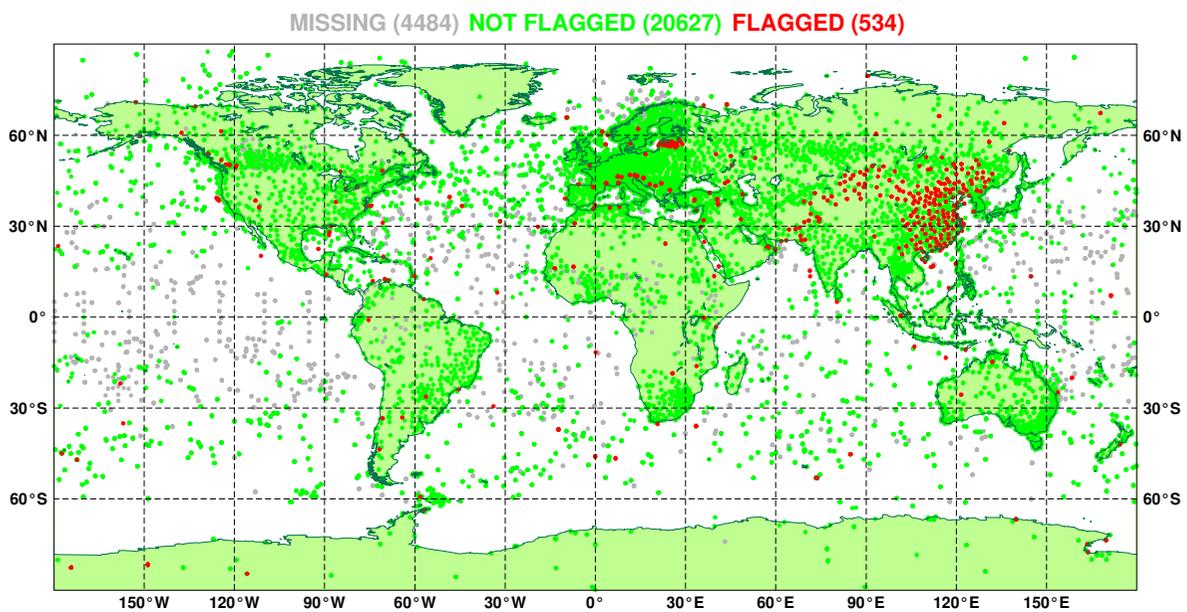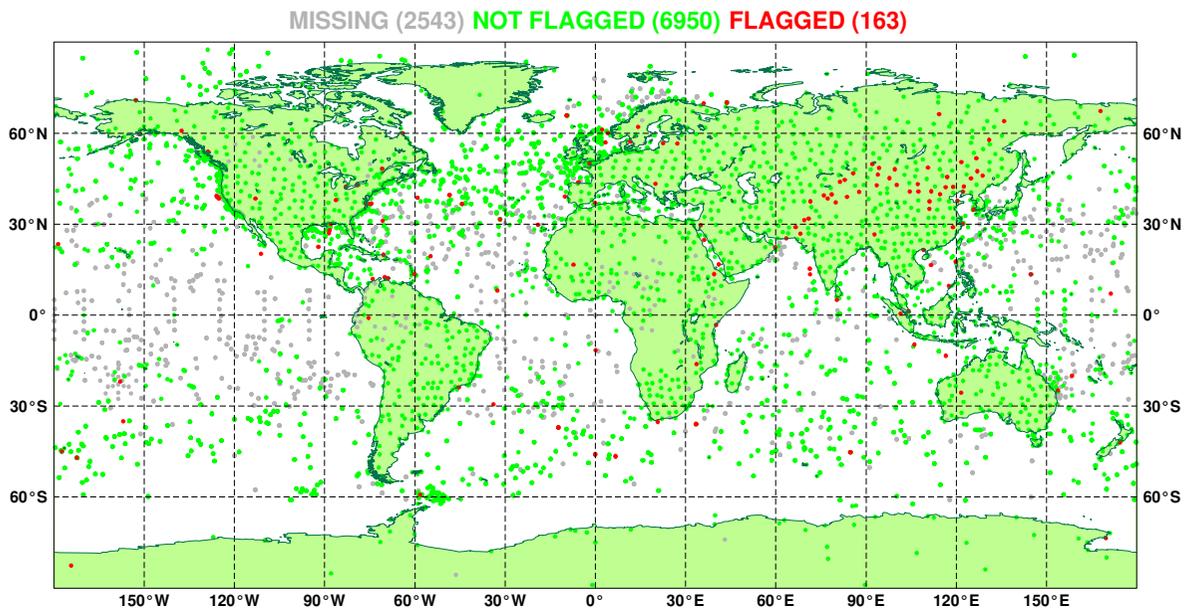
Figure 8.8: The Thinned surface pressure data from 11 March 2009. Flagged data is not assimilated into the Met Office Var scheme.

Experiment		Condition Number	
Scheme	Observations	No Thinning	Thinning
3DVar	Surface	3395	501
3DVar	All	3590	1406
3DVar	No Surface	1406	-
4DVar	Surface	2778	456
4DVar	All	2975	928

Table 8.7: Condition number of Hessian of the Met Office Var scheme with and without thinning for data from 11/03/2009.

we consider the effect of thinning the observations on the condition number of the Hessian. In this section *thinning* will refer to the process of removing observations so that there is increased spacing around the observation. According to the theory and simple experiments of Sections 6.3.2 and 7.4 the conditioning of the system should improve if observations are thinned. Thinning observations is common practice in NWP centres to ensure that no observation error correlations remain which are not accounted for in the assimilation scheme [15]. In this section we show that thinning the surface observations produces a reduction in the condition number of the Met Office Var scheme.

We assimilate observational data from 12Z on 11th March 2009. The surface pressure data is shown in Figure 8.7 on the Met Office N108L38 global grid. We reduce the density of the surface observations by applying a thinning of 300km. Hence, observations are removed so that there is a spacing of $300km$ between the remaining observations. The thinned data in Figure 8.8 shows the large reduction in land observations. The 3DVar and 4DVar Met Office minimisation schemes are applied using all the observations with both the original and then the thinned surface data. We also calculated the condition number for Var experiments where we only assimilated the unthinned and thinned surface data and an additional experiment in 3DVar where all but the surface data was assimilated. The results are shown in Table 8.7.

The table shows the condition number of the Hessians for both the 3DVar

and 4DVar assimilations. When only the surface observations are assimilated the thinning of the data shows a dramatic reduction in the condition number of the Hessian. The condition number of the thinned Hessian is reduced by over 80% of the condition number of the unthinned Hessian, from 3395 to 501 in the 3DVar case and 2778 to 456 in the 4DVar case.

Now consider the case where all the observations are assimilated. In the 3DVar case, the condition number of the Hessian which assimilated the thinned surface observations is 60% of the size of the condition number of the Hessian which assimilated the unthinned surface observations. Similarly in 4DVar, the condition number of the Hessian which assimilated the thinned data is 70% the size of the condition number of the Hessian for the unthinned. In the 3DVar case, the Table 8.7 shows that the condition number of the Hessian that assimilated all the observations including the thinned surface observations is exactly the same magnitude as the condition number of the Hessian that assimilated all the observations minus the surface observations. Hence we conclude that the thinned surface observations now have no affect on the condition number of the system. As predicted by the theory, in Sections 6.3.2 and 7.4 we have shown that by thinning the observations we improve the conditioning of the Met Office Var scheme.

We remark that although thinning the data gives an improved condition number it also implies we must remove information from the assimilation. This may result in a less accurate analysis. In reality, in order to obtain an accurate analysis, a balance is necessary between the numerical accuracy of the problem (provided by the measure of the condition number) and the information content. We now summarise this chapter.

8.4 Summary

The Met Office implement a preconditioned Var scheme, preconditioned via the control variable transform. A well-conditioned system is necessary for an accurate forecast. In Sections 6.3 and 7.4 we showed that as the observations become less accurate, the spacing between observations increases, and the number of observations are reduced the condition number of the preconditioned Hessian decreases. In this chapter we showed that these results also apply in the Met Office operational Var scheme.

We began by giving an overview of the Met Office operational system including outlines of how the Var cost function is constructed and minimised. Next we considered the conditioning of the Met Office system when only assimilating pseudo observations. This gave us the ability to define the errors and positions of observations which is not possible with real observations. We showed in both 3DVar and 4DVar experiments that the condition number of the Met Office Var scheme was improved by assimilating more inaccurate pseudo observations with greater spacing between them. We then looked at the conditioning of the Met Office scheme using real observations. First we looked at individual observation types. We found that surface observations dominated the conditioning of the Met Office Var minimisation. We then considered the effect of thinning the surface observations. We found that by sufficiently thinning the observations we reduced the overall condition number of the system and removed the dominance of the surface observations on the conditioning of the Met Office Var scheme. Hence we showed that the results predicted by the theory and experiments in Chapters 6 and 7 also apply in an operational system. We now finish by summarising this thesis and considering possible future work.

Chapter 9

Conclusions

In numerical weather prediction (NWP), data assimilation techniques are applied which combine observation data, a prior estimate and a numerical forecast model in order to generate a best guess of the initial state of the atmosphere, called the analysis. The accuracy of this analysis is vital in order to ensure an accurate future weather forecast. Variational data assimilation (Var) techniques are currently employed in many major NWP centres around the world. In practice, Var is implemented in an incremental form (see Section 2.3) and commonly in conjunction with a control variable transform (CVT) in order to model the background error covariance matrix (see Section 2.4). The incremental formulation requires minimising a linear least-squares cost function linearised around the current best estimate in a so-called inner-loop and then an update of the current best estimate in an outer-loop step (see Section 2.3). In many NWP centres, few if any, outer-loop updates are performed and so the main computational cost is associated with the inner-loop minimisation.

The condition number of the Hessian of the incremental cost function gives a measure of the sensitivity of the analysis to perturbations in the system. Additionally the condition number can give an indication of the rate of convergence of the iterative schemes used to solve the inner-loop step. Thus, in

an NWP context, a small condition number is desirable in order to obtain an accurate analysis in real-time. In this thesis we have considered the conditioning of the linearised cost functions associated to the incremental version of Var without the CVT (2.9) denoted Var and with the CVT (2.16) denoted PVar.

There exists very little theoretical research in the literature explaining the conditioning of Var or PVar. Most results come from either simple models or experimental evidence [6], [18], [42], [63]. Without proper information on the conditioning of the Var and PVar problems it can be difficult to,

- Identify sources of ill-conditioning in Var or PVar.

- Explain why PVar seems to have better conditioning than Var.

- Predict what effect future modifications may have on the condition number of Var and PVar.

In this thesis we expanded on the current existing body of research in the conditioning of variational data assimilation by deriving theoretical bounds on the condition number of the Var and PVar Hessians. In Chapter 2 we introduced data assimilation with the main focus on variational data assimilation. We also described an incremental version of Var which included the CVT based on a square root of the background error covariance matrix. We also gave an overview of how the CVT is implemented within the Met Office operational Var scheme.

In Chapter 3 we introduced the concept of condition number of a matrix. We showed how the sensitivity of the analysis to perturbations is related to the condition number of the Hessian of the Var/PVar cost functions. A larger condition number implies a sensitive solution and we call the system ill-conditioned. We also introduced the concept of preconditioning which changes the condition number of a matrix. We showed that the CVT effectively acts as a preconditioner. We then showed how an ill-conditioned

system can imply a slow convergence rate of the gradient methods used to minimise a linear least-squares cost function. Finally, we presented theoretical results for bounds on the eigenvalues of square matrices which we used to put theoretical bounds on the condition number of the Var/PVar Hessians in Chapters 6 and 7.

In Chapter 4 we presented previous results which indicated that the convergence of the Var minimisation was slow. This suggested that Var may be an ill-conditioned system. In [42], it was suggested that if the background error covariance matrix was ill-conditioned then this may cause ill-conditioning if the Var scheme. Comparison with previous experimental work indicated that the preconditioned system performed much better than the Var scheme without the CVT [42]. We also presented prior research from the literature on the conditioning of the PVar scheme in a simple 2-grid system that showed that the conditioning of the preconditioned scheme deteriorated in the case of dense observations when the accuracy and number of observations was increased [6].

In Chapter 5 we considered the conditioning of three common auto-covariance matrices defined for a periodic system on a 1D domain. In each of the cases we showed that there is a large increase in the condition number of the background error covariance matrix as the error correlation lengthscale was increased.

In Chapters 6 and 7 we derived new theoretical bounds on the condition number of the 3DVar and 4DVar Hessians respectively which showed the conditioning of Var was related to the conditioning of the background error covariance matrix. Using the results from Chapter 5 we illustrated this relationship using numerical experiments for a one-parameter, periodic system on a 1D domain. We showed that if the background covariance matrix was ill-conditioned then so was the Var scheme.

In Chapters 6 and 7 we also derived new theoretical bounds on the condi-

tion number of the 3DVar and 4DVar Hessians preconditioned via the control variable transform. We showed that the CVT significantly reduced the condition number of Var scheme and we illustrated it using numerical experiments on a periodic system define on a 1D domain using the background covariance matrices defined in Chapter 5 to construct the Hessians. The bounds also indicated that the observation distribution and accuracy played an important role in the condition number of the Hessian. We showed that the condition number was reduced by thinning the observations, reducing the spacing between observations and by increasing the observation error variance. In Chapter 8 we showed that the same results also applied to the Met Office operational Var scheme with experiments using both real and pseudo observations.

Results for the convergence rate of the conjugate gradient method were presented in Chapters 6 and 7. We found that the convergence rate was signifcantly increased by preconditioning using the CVT. Additionally in the preconditioned system an increase in the convergence rate of the conjugate gradient method was found by reducing the number of observations, increasing the spacing between the observations or by reducing the accuracy of the observations. We now summarise our main conclusions.

9.1 Conclusions

In this thesis we have developed new algebraic bounds on the condition number on the Var and PVar Hessians. The aim is to provide a more theoretical understanding of the conditioning of the Var system. More specifically, in this thesis

1. We showed that common auto-correlation models, used to model the background error covariance matrices in Var, have condition numbers

which are sensitive to correlation lengthscale and hence, for highly correlated background errors, will be ill-conditioned (see Chapter 5).

2. We derived new theoretical bounds on the condition number of the Var Hessian and showed that the conditioning of Var is dependent on the conditioning of the background error covariance matrix. Using results from Chapter 5, we showed that for highly correlated background errors the Hessian of the unpreconditioned system is generally ill-conditioned (see Sections 6.1.1 and 7.1.2). We illustrated this with experiments using a periodic system on a one-dimensional domain.

3. We derived new theoretical bounds on the condition number of the Hessian of the preconditioned Var system. We showed using a periodic system on a one-dimensional domain that the system is less sensitive to the lengthscale and better conditioned than the Hessian of the unpreconditioned Var system (see Sections 6.2 and 7.2.1).

4. Using the bounds on the conditioning of the preconditioned system we showed in experiments on a periodic system on a one-dimensional domain that the conditioning is affected by the accuracy, number and positioning of the observations (see Sections 6.3, 7.3 and 7.4).

5. We showed using experiments on a one-parameter, periodic system that our preconditioned system produces a large increase in the convergence rate of the conjugate gradient method compared to the unpreconditioned system. In addition, increasing the spacing and making the observations less accurate also increases the convergence rate in the preconditioned system (see Sections 6.4 and 7.5).

6. We showed that the effect of the observation accuracy and distribution found in Sections 6.3 7.3 and 7.4 on the PVar system also apply to the Met Office operational Var scheme. We showed that the condition number of the Met Office Var scheme could be reduced by thinning the observations, increasing the spacing between observations and making the observations less accurate (see Chapter 8).

9.2 Further Work

In the derivation of the theoretical bounds on both the preconditioned and unpreconditioned Hessian clearly many assumptions were made. The bounds on the unpreconditioned case for 3DVar given in (6.3) and for the 4DVar case in (7.3) are very general with minimimal assumptions made on the observations and errors compared to the more specific bounds given in (6.11) for 3DVar and (7.11) for 4DVar. The results shown in the Chapters 6 and 7 indicate that it is the conditioning of the background error covariance matrix that determines the conditioning of the unpreconditioned system. However, the bounds show that the condition number is also dependent on a number of other factors which we have not fully explored in this thesis. For instance, it was shown in numerical experiments on a 1D periodic system that for the SOAR version of the 3DVar Hessian the conditioning of Var could either follow the upper bound (Figure 6.3) or the lower bound (Figure 6.2) depending on the distribution of the observations. Hence a difference of an order of magnitude of 10 can result in the condition number by choosing different observation distributions. A similar result is found for the 3DVar Hessian constructed using the Laplacian background error covariance matrix (Figures 6.5 and 6.4). The reason for this change has not been explained in this thesis. Similarly, the bounds on the unpreconditioned Hessians also depend on the background and observation error variances but their affect on the condition number has not been explored in this research.

For both the unpreconditioned and the preconditioned systems the theory of Chapters 6 and 7 used observations at grid points only in order to simplify the theory. In practice this is an unrealisitic assumption and the observation operator may be a combination of interpolations and transformations to different variables. In order to understand the full operational system, more realistic obervations operators must be investigated where the observations are not assumed to be at grid points. A first step towards this would be to consider incorporating a simple linear interpolation from grid points to

observation locations into the observation operator.

In Chapter 7 the condition number of the P4DVar Hessian was shown to be dependent on the 4D background error covariance matrix introduced in Section 7.2.2. We described this matrix in the case of a simple discretised advection model which provided a convenient form for analysis. In general, this matrix will be much more complicated but further analysis has been beyond the scope of this thesis. Investigating more realistic models of the atmosphere and their action on the background error covariance matrix will be particularly useful for grasping a more complete understanding of the conditioning of operational Var schemes that use a CVT.

In this thesis we saw that the conditioning of PVar was in general, significantly better than that of Var in the experiments considered. However, for small correlation lengthscales we can see, for example in, Figures 6.4 and 6.8, that both systems have Hessians with similar magnitude of condition number. Thus it is feasible that in some circumstances PVar may be more ill-conditioned than Var. The question of when this may occur has not been fully investigated.

Throughout this thesis correlation in the observation errors has been ignored. However, currently research has begun to investigate the effect of incorporating observations error correlations [56], [25]. The impact of correlations on the conditioning of the Var problem, and hence the accuracy of analysis, is unknown. The background error covariance matrix includes correlations and has been shown to be very ill-conditioned hence incorporating correlations in the observation covariance matrix may imply that the Hessian becomes more ill-conditioned. Understanding the effect of observation error correlations may be important for understanding the accuracy of the analysis in the future.

In this thesis we have concentrated on the CVT as a preconditioner. A useful and natural extension is to consider additional preconditioners. A common

extra level of preconditioning that can be applied on top of preconditioning the Hessian with the CVT is eigenvector preconditioning. If the Hessian does not vary much from one assimilation cycle to the next then we can generate eigenvectors of the Hessian from a previous cycle to precondition the Hessian of a later assimilation cycle [64]. In Chapter 8 we showed that the surface observations consistently dominated the conditioning of the Hessian of the preconditioned system. Furthermore the leading eigenvector appears to be have a consistent structure regardless of the different assimilation cycles. If this is the case then the leading eigenvectors may be calculated before the minimisation and used to precondition the Hessian. Some preliminary tests using eigenvector preconditioning has already been attempted using the Met Office operational system resulting in a reduction in the conditioning of the system [61].

In Chapter 8 we found that thinning the surface observations produced a better condition system. However this came at the price of removing information from the assimilation. An important question remains in what is the optimal compromise between having a more ill-conditioned system, and hence a more inaccurate numerical solution, and removing observations and having an unrealistic solution? If this optimal thinning can be found then the most accurate analysis for predicting future weather states can be realised.

Bibliography

[1] Control variable transforms, var scientific documentation paper 13. Technical report, UK Met. Office, 1999.

[2] Global horizontal spectra for 3d-var. Technical report, UK Met. Office, 2002.

[3] Control variable transform: Parameter transforms, var scientific documentation paper 11. Technical report, UK Met. Office, 2006.

[4] Control variable transform: Horizontal transformations, var scientific documentation paper 14. Technical report, UK Met. Office, 2007.

[5] Minimisation and convergence, var scientific documentation paper 5. Technical report, UK Met. Office, 2007.

[6] Erik Anderson, Mike Fisher, Rosemary Munro, and Anthony McNally. Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Quarterly Journal of the Royal Meteorological Society*, 126:1455–1472, 2000.

[7] Owe Axelsson. *Iterative Solution Methods*. Cambridge University Press, 1996.

[8] Ross Bannister. Darc internal report no. 5: On control variable transforms in the met office 3d and 4d var., and a description of the proposed waveband summation transformation. Technical report, DARC, 2003.

[9] Ross N. Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. i: Characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, 134:1951–1970, 2008.

[10] Ross N. Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. ii: Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, 134:1971–1996, 2008.

[11] D. M. Barker, W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q. N. Xiao. A three-dimensional variational data assimilation system for mm5: Implementation and initial results. *Monthly Weather Review*, 132:897–914, 2004.

[12] Winston C. Chao and Lang-Ping Chang. Development of a four-dimensional variational analysis system using the adjoint method at gla. part i: Dynamics. *Monthly Weather Review*, 120:1661–1673, 1992.

[13] P. Courtier, E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, and M. Fisher. The ecmwf implementation of three-dimensional variational assimilation (3d-var). i: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124:1783–1808, 1998.

[14] Roger Daley. *Atmospheric Data Analysis*. Cambridge University Press, first edition edition, 1993.

[15] M. L. Dando, A. J. Thorpe, and J. R. Eyre. The optimal density of atmospheric sounder observations in the met office nwp system. *Quarterly Journal of the Royal Meteorological Society*, 133:1933–1943, 2007.

[16] Francois Xavier Le Dimet and Olivier Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, 38A:97–110, 1986.

[17] G. Gaspari and S. E. Cohn. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125:723–758, 1999.

[18] P. Gauthier, C. Charette, L. Fillion, P. Koclas, and S. Larocher. Implementation of a 3d variational data assimilation system at the canadian meteorological centre. part i: the global analysis. *Atmsophere-Ocean*, 37:103–156, 1999.

[19] Pierre Gauthier, Monique Tanguay, Stephane Laroche, Simon Pellerin, and Josee Morneau. Extension of 3dvar to 4dvar: Implementation of 4dvar at the meteorological service of canada. *Monthly Weather Review*, 135:2339–2354, 2007.

[20] Pierre Gauthier and Jean-Noel Thepaut. Impact of the digital filter as a weak constraint in the preoperational 4dvar assimilation system of meteo-france. *Monthly Weather Review*, 129:2089–2102, 2001.

[21] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical Optimization*. Academic Press, 1982.

[22] Tilmann Gneiting. Simple tests for the validity of correlation function models on the circle. *Statistics and Probability Letters*, 39:119–122, 1998.

[23] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The John Hopkins University Press, third edition edition, 1997.

[24] Robert M. Gray. *Toeplitz and circulant matrices: A Review (Foundations and Trends(R) in Communications and Information Theory)*. now publishers Inc., 2006.

[25] S. B. Healy and A. A. White. Use of discrete fourier transforms in the 1d-var retrival problem. *Quarterly Journal of the Royal Meteorological Society*, 131:63–72, 2005.

[26] Xiang-Yu Huang, Qingnong Xiao, Dale M. Barker, Xin Zhang, John Michalakes, Wei Huang, Tom Henderson, John Bray, Yongsheng Chen,

Zaiahong Ma, Jimy Dudhia, Yongrun Guo, Xiaoyan Zhang, Duk-Jin Won, Hui-Chuan Lin, and Ying-Hwa Kuo. Four-dimensional variational data assimilation for wrf: Formulation and preliminary results. *Monthly Weather Review*, 137:299–314, 2009.

[27] Bruce Ingleby. The statistical structure of forecast errors and its representation in the met. office global 3-d variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 127:209–231, 2001.

[28] N. B. Ingleby. Global horizontal spectra for 3d-var, available via var. scientific documentation paper 20. Technical report, UK Met. Office, 2002.

[29] Jennings. Influence of the eigenvalue spectrum on the convergence rate of the conjugate gradient method. *IMA Journal of Applied Mathematics*, 20:61–72, 1977.

[30] Christine Johnson. *Information Content of Observations in Variational Data Assimilation*. PhD thesis, University of Reading, UK, 2003.

[31] Christine Johnson, Brian J. Hoskins, and Nancy K. Nichols. A singular vector perspective of 4d-var. filtering and interpolation. *Quarterly Journal of the Royal Meteorological Society*, 131:1–19, 2005.

[32] Eugenia Kalnay. *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, first edition edition, 2003.

[33] C. T. Kelly. *Iterative Methods for Linear and Nonlinear Equations (Frontiers in Applied Mathematics)*. SIAM, 1995.

[34] A. C. Koivunen and A. B. Kostinski. The feasibilty of data whitening to improve performance of weather radar. *Journal of Applied Meterology*, 38:741–749, 1999.

[35] A. B. Kostinski and A. C. Koivunen. On the condition number of gaussian sample covariance matrices. *IEEE Transactions on Geoscience and Remote Sensing*, 38:329–332, 2000.

[36] Amos S. Lawless and Nancy K. Nichols. Inner-loop stopping creiteria for incremental four-dimensional variational data assimilation. *Monthly Weather Review*, 134:3425–3435, 2006.

[37] Daniel Lea. *Joint assimilation of sea surface temperature and sea surface height.* PhD thesis, University of Oxford, UK, 2001.

[38] John M. Lewis and John C. Derber. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, 37A:309–322, 1985.

[39] John M. Lewis, S. Lakshmivarahan, and S. K. Dhall. *Dynamic Data Assimilation: A Least Squares Approach.* Cambridge University Press, 2006.

[40] Andrew Lorenc. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 112:1177–1194, 1986.

[41] Andrew Lorenc. Iterative analysis using covariance functions and filters. *Quarterly Journal of the Royal Meteorological Society*, 118:569–591, 1992.

[42] Andrew Lorenc. Development of an operational variational assimilation scheme. *Journal of the Meteorological Society of Japan*, 75:339–346, 1997.

[43] Andrew Lorenc. The met. office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126:2991–3012, 2000.

[44] E. N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20:130–141, 1963.

[45] K. W. Morton and D. F. Mayers. *Numerical Solution of Partial Differential Equations.* Cambridge University Press, 1994.

[46] N. K. Nichols. *Data Assimilation for the Earth System.* 2003.

[47] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization.* 2000.

[48] David F. Parrish and John C. Derber. The national meteorological center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120:1747–1763, 1992.

[49] P.Courtier and O. Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equations. ii: Numerical results. *Tellus*, 42A:531–549, 1990.

[50] P.Courtier and O. Talagrand. Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, 42A:531–549, 1990.

[51] P.Courtier, J. N. Thepaut, and A. Hollingsworth. A strategy for operational implementation of 4d-var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120:1367–1388, 1994.

[52] F. Rabier and P. Courtier. Four-dimensional assimilation in the presence of baroclinic instability. *Quarterly Journal of the Royal Meteorological Society*, 118:649–672, 1992.

[53] F. Rabier, H. Jarvinen, E. Klinker, J. F. Mahfouf, and A. Simmons. The ecmwf operational implementation of four-dimensional variational assimilation. i: Experimental results with simplified physics. *Quarterly Journal of the Royal Meteorological Society*, 126:1143–1170, 2000.

[54] F. Rawlins, S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne. The met office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 133:347–362, 2007.

[55] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics II. Fourier Analysis, Self Adjointnes.* Academic Press, 1975.

[56] Laura Stewart. *Correlated observation errors in data assimilation.* PhD thesis, University of Reading, UK, 2010.

[57] Endre Suli and David Mayers. *An Introduction to Numerical Analysis.* Cambridge University Press, 2003.

[58] William C. Thacker. The role of the hessian matrix in fitting models to measurements. *Journal of Geophysical Research*, 94:6177–6196, 1989.

[59] Jean Noel Thepaut and Philippe Courtier. Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model. *Quarterly Journal of the Royal Meteorological Society*, 117:1225–1254, 1991.

[60] Jean Noel Thepaut and Patrick Moll. Variational inversion of simulated tovs readiances using the adjoint technique. *Quarterly Journal of the Royal Meteorological Society*, 116:1425–1448, 1990.

[61] U.K. Met Office Tim Payne. personal communication.

[62] Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra.* SIAM, 1997.

[63] Y. Tremolet. Incremental 4d-var convergence study. *Tellus A*, 59:706–718, 2007.

[64] J. Tshimanga, S. Gratton, A. T. Weaver, and A. Sartenaer. Limited-memory preconditioners, with application to incremental four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 134:751–769, 2008.

[65] Anthony Weaver and Philippe Courtier. Correlation modelling on the sphere using a generalized equation. *Quarterly Journal of the Royal Meteorological Society*, 127:1815–1846, 2001.

[66] Rudolf O. Weber and Peter Talkner. Some remarks on spatial correlation function models. *Monthly Weather Review*, 121:2611–2617, 1993.

[67] J. H. Wilkinson. *The Algebraic Eigenvalue Problem,*. Clarendon Press, Oxford, 1965.

[68] Andrew T. A. Wood. When is a truncated covariance function on the line a covariance function on the circle? *Statistics and Probability Letters*, 24:157–164, 1995.

[69] A. M. Yaglom. *Correlation Theory of Stationary and Related Random Functions I. Basic Results.* Springer-Verlag, 1986.

[70] Nicholas Young. *An introduction to Hilbert space.* Cambridge University Press, 2001.

[71] Milija Zupanski. A preconditioning algorithm for large-scale minimization problems. *Tellus*, 45A:478–492, 1993.

[72] Milija Zupanski. A preconditioning algorithm for four-dimensional variational data assimilation. *Monthly Weather Review*, 124:2562–2573, 1996.