School of Mathematical, Physical and Computational Sciences

On the treatment of correlated observation errors in data assimilation

by

Jemima M. Tabeart

Thesis submitted for the degree of Doctor of Philosophy

in Mathematics

School of Mathematical, Physical and Computational Sciences

August 2019

University of Reading

Declaration

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

Jemima Maple Tabeart

Publications

The work in chapters 5, 7 and 8 of this thesis are strongly based on the following publications:

> Tabeart J. M., Dance S. L., Haben S. A., Lawless A. S., Nichols N. K., Waller J. A. *The conditioning of least-squares problems in variational data assimilation.* Numerical Linear Algebra with Applications 2018;25:e2165. https://doi.org/10.1002/nla.2165
>
> Tabeart J. M., Dance S. L., Haben S. A., Lawless A. S., Nichols N. K., Waller J. A. *Improving the condition number of estimated covariance matrices.* In review Tellus A https://arxiv.org/abs/1810.10984
>
> Tabeart J. M., Dance S. L., Haben S. A., Lawless A. S., Migliorini, S. Nichols N. K., Smith, F., Waller J. A. *The impact of using reconditioned correlated observation error covariance matrices in the Met Office 1D-Var system.* In review Quarterly Journal Royal Meteorological Society http://arxiv.org/abs/1908.04071

All work undertaken in these publications was carried out by Jemima M. Tabeart with coauthors providing guidance and review.

Abstract

Data assimilation combines information from observations of a dynamical system with a previous forecast, with each term weighted by its respective uncertainty. An important recent area of research has been the introduction of correlated observation error covariance (OEC) matrices in numerical weather prediction systems. The benefits of correlated OEC matrices are multiple: they permit the use of high density observation networks, allow the capture of small scale processes and help make best use of available data. However, their use is often associated with convergence problems for iterative methods. In this thesis we study the theoretical impact of introducing correlated OEC matrices on the conditioning of variational data assimilation problems. We develop new bounds on the condition number of the Hessian for two data assimilation formulations and illustrate our findings with numerical examples in an idealised framework. The minimum eigenvalue of the OEC matrix is a key term for both problems, which motivates the use of reconditioning methods to reduce the condition number of correlation matrices. We develop theory for two reconditioning methods: ridge regression and the minimum eigenvalue method. We show for the first time that standard deviations are increased by both methods. Ridge regression reduces absolute correlations, whereas the minimum eigenvalue method makes smaller changes to correlations and variances. We then present the first in-depth study of the ridge regression method for an operational data assimilation system, using the Met Office 1D-Var system. Reconditioning improves convergence, but alters the quality control procedure, which is used to select appropriate observations for further assimilation. The results in this thesis provide guidance on how to include correlation information in general variational data assimilation problems while ensuring computational efficiency.

Acknowledgements

I would like to thank my team of supervisors: Prof. Sarah Dance, Dr Amos Lawless, Prof Nancy Nichols, Dr Joanne Waller and Dr David Simonin. I've had fun, learned lots and achieved more than I thought possible: none of which I could have done without your mentoring or support. The Satellite Applications and Ocean Data Assimilation teams were extremely enjoyable places to spend summers during my PhD. My collaborators at the Met Office, Stefano Migliorini and Fiona Hilton, have also been beyond helpful with mentorship, coding support, and tracking down obscure documentation.

My PhD was funded by the National Centre for Earth Observation through the EPSRC Centre for Doctoral Training in Mathematics of Planet Earth, without which I would not have been able to undertake this research. Many thanks go to staff within the CDT, particularly Anna Radomska, Sam Williams, Tristan Pryer and Jennifer Scott for administrative and organisational support. A huge thanks go to the other members of Cohort Bravo, with an extra special mention to Tasmin for continuous moral support. Maybe one day we will manage to live in the same city! Thank you also to members of the Department of Mathematics and Statistics and the Data Assimilation Research Centre, especially the administrators and other PhD students. Shout out to James, Maha, Hasen and Oliver for coffee breaks and trips to Sprinkles.

Outside academia, I am grateful to housemates old and new for providing relaxing and happy spaces to unwind: Darren, Kate and Ying Ying at 7URR and the Mimis: Hope, Katie, Rachel, Steph and MLadd both digitally and around the UK. You all inspire me and I couldn't have done it without you.

Finally, I want to thank my family, particularly my parents, siblings and growing menagerie of animals, who have never failed to provide limitless emotional and logistical support. Their constant encouragement has given me the freedom and self-belief to follow exciting opportunities all over the world.

Table of Contents

List of Tables

List of Figures

List of Symbols

$\mathbf{B}$	background error covariance matrix
$\mathbf{C}$	circulant error correlation matrix
$\mathbf{D}$	circulant error correlation matrix
$\mathbf{H}$	linearised observation operator
$\mathbf{I}_N$	$N \times N$ identity matrix
J	cost function
L	correlation lengthscale in the SOAR function
L_B	correlation lengthscale of circulant $\mathbf{B}$
L_R	correlation lengthscale of circulant $\mathbf{R}$
N	number of state variables
$\mathbf{R}$	observation error covariance matrix
$\mathbf{S}$	Hessian of cost function
$\mathbf{V}$	matrix of eigenvectors
h	observation operator
n	number of time steps
p	number of observations
$\mathbf{x}$	state vector
$\mathbf{x}_a$	optimal analysis state vector
$\mathbf{x}_b$	background state vector
$\mathbf{y}$	observation vector
$\mathbf{v}$	eigenvector of a matrix
$\mathbf{\Lambda}$	diagonal matrix of eigenvalues
Σ	observation error matrix of variances
$\kappa(\mathbf{M})$	condition number of $\mathbf{M}$
$\lambda_k(\mathbf{M})$	k^{th} eigenvalue of matrix $\mathbf{M}$ where $\lambda_1 \geq \cdots \geq \lambda_N$
σ_b	background error variance
σ_o	observation error variance

Abbreviations

OEC	Observation error covariance
NWP	Numerical weather prediction
1D-Var	One-dimensional variational data assimilation
3D-Var	Three-dimensional variational data assimilation
4D-Var	Four-dimensional variational data assimilation
ECMWF	European Centre for Medium-Range Weather Forecasts
NRL	U.S. Naval Research Laboratory
ME	Minimum eigenvalue method of reconditioning
RR	Ridge regression method of reconditioning
MVI	Minimum variance inflation
IASI	Infrared atmospheric sounding interferometer
CF	Cloud fraction
CTP	Cloud top pressure
IQR	Interquartile range
ST	Skin temperature

Chapter 1

Introduction

Weather forecasts are important for individuals, businesses, non-governmental organisations and governments [Kalnay, 2002, Bauer et al., 2015]. In order to initialise a forecast, information from recent observations is combined with a previous forecast, referred to as the background or prior, via a process known as data assimilation [Daley, 1991]. In order to find the most likely initial condition, or analysis, data assimilation methods weight the contribution of background and observation information by their respective uncertainties via error covariance matrices. Combining these two sources of information can be difficult, as observations may be at different locations, times, or of different quantities to meteorological forecast variables. Numerical weather prediction (NWP) is a high-dimensional and time-sensitive problem, meaning it is computationally and theoretically challenging [Bauer et al., 2015, Carrassi et al., 2018]. Therefore data assimilation methods need to be computationally efficient and able to cope with large volumes of data.

In this thesis we will study the impact of using correlated observation error covariance matrices in the variational data assimilation problem. In variational data assimilation, the analysis is found by minimising a nonlinear least squares objective function. In the linear setting the conjugate gradient method is often used to find the solution to the linear system associated with the variational problem. The unpreconditioned variational objective function consists of two terms which measure the misfit between the background and the observations, respectively. In the preconditioned formulation, a variable transform is used to decorrelate the background variables in the first term of the objective function.

The existence of correlated observation errors for satellite instruments in particular is well established [Stewart, 2010]. It has been shown that failing to account for correlated observation errors limits forecast skill [Rainwater et al., 2015] and that even accounting for approximate error statistics is better than not accounting for error correlations at all [Stewart et al., 2008b, Stewart, 2010, Stewart et al., 2013, 2014]. Another motivation for using correlated observation error covariance matrices is the increasing importance of high resolution forecasts e.g. for the prediction of hazardous weather such as flash flooding due to intense convective rainfall [Dance et al., 2019]. However, the presence of correlated observation errors prevented the operational use of highly spatially dense observations until recently for computational reasons [Simonin et al., 2019]. Observation error covariance matrices can be hard to estimate and can be expensive to implement [Stewart et al., 2008b]. In order to avoid using correlated observation error covariance matrices in the case that spatial error correlations are known to exist, observation information is often thinned [Simonin et al., 2019], limiting skill at high resolution. This results in large numbers of observations not being used: for example in the Metéo France convection-permitting forecast model, radar observations are thinned by a factor of 64 and infra-red satellite observations by a factor of 400 [Michel, 2018].

In recent years, there has been a growing interest in implementing correlated observation error covariance matrices in operational data assimilation routines at NWP centres (e.g. Weston [2011], Weston et al. [2014], Bormann et al. [2016], Campbell et al. [2017]). Many of these studies make use of the diagnostic of Desroziers et al. [2005] (henceforth referred to as DBCP) to estimate observation error covariance matrices. The DBCP approach makes use of samples, and can result in estimated correlation matrices that are rank deficient [Pourahmadi, 2013], or are not symmetric [Ménard, 2016]. This means that in order to use the correlated observation error information in practice, the matrices must be modified in some way. ‘Reconditioning’ methods which reduce the condition number of estimated covariance matrices are often used (e.g. Weston [2011], Weston et al. [2014], Bormann et al. [2015], Campbell et al. [2017]). In this thesis we study two commonly used methods of reconditioning theoretically for the first time: ridge regression and the minimum eigenvalue method. We compare both reconditioning methods with multiplicative variance inflation, which multiplies the estimated covariance matrix by a constant inflation factor. This method cannot change the condition number of the matrix, and is hence not a method of reconditioning, but is often used at NWP centres to mitigate for missing correlation information.

1.1 Research questions

In this thesis, we wish to understand how the introduction of correlated observation errors affects a general data assimilation problem in terms of convergence of the minimisation of the objective function. We will consider a number of research questions, which will allow us to focus on different aspects of the topic.

RQ 1: How does introducing correlated observation error affect the conditioning of the Hessian of the variational data assimilation problem?
Building on the work of Haben [2011], we will develop theoretical bounds on the condition number of the Hessian to understand the impact of the observation error covariance matrix. How are these bounds affected by changes to the observation error covariance matrix? How tight are these bounds for an idealised numerical framework? How well does the behaviour of the condition number of the Hessian represent convergence of the conjugate gradient method numerically?

RQ 2: What is the difference between the preconditioned and unpreconditioned data assimilation problem?
The control variable transform [Bannister, 2008], is used to precondition and decorrelate the background term of the variational data assimilation objective function. How does the importance of background and observation terms differ from the unpreconditioned case? Does the behaviour of the condition number of the Hessian represent convergence of the conjugate gradient method well for numerical experiments?

RQ 3: How do reconditioning methods alter covariance matrices?
Reconditioning methods have been used to mitigate problems with ill-conditioned estimated covariance matrices by increasing small eigenvalues of a sample covariance matrix. How do reconditioning methods alter correlations and standard deviations associated with the covariance matrix? How is the variational objective function changed by the use of reconditioning methods? How do two commonly-used reconditioning methods compare to multiplicative variance inflation?

RQ 4: What is the impact of using the ridge regression method of reconditioning on an operational data assimilation problem?
We present a case study using the operational Met Office 1D-Var system. How

do the qualitative theoretical conclusions from the linear case apply in a nonlinear, realistic setting? How are the quality control process and retrieved values affected by the introduction of correlated observation error and the use of reconditioning methods?

1.2 Outline

The thesis is structured as follows.

In Chapter 2 we introduce the variational data assimilation problem. We define three-dimensional variational data assimilation (3D-Var) and four-dimensional variational data assimilation (4D-Var), and describe the control variable transform (CVT) that is typically used at numerical weather prediction (NWP) centres. The Hessian of the objective function is defined for each formulation of the problem. We also discuss the importance of correctly specifying observation error statistics, and introduce the diagnostic of Desroziers et al. [2005] which is commonly-used to estimate correlated observation error covariances. In Chapter 3 we present theoretical results that will be used to develop bounds on the Hessian of the objective function. We introduce the concept of the condition number, and discuss results on the eigenvalues of matrix sums and products. We define the conjugate gradient method, and show that the condition number can be used to bound convergence of this algorithm. Finally we introduce specific matrix structures that will be used both theoretically and in numerical experiments in the rest of the thesis. In Chapter 4 we discuss previous work on the condition number of the Hessian as a proxy for convergence of the variational data assimilation problem. Of particular interest is the work of Haben et al. [2011a,b] and Haben [2011] who developed bounds on the condition number of the Hessian in terms of individual matrices in the case of uncorrelated observation errors. In this thesis we will extend these results to study the impact of using correlated observation error covariance matrices. We discuss numerical issues with the use of correlated observation errors at NWP centres, including convergence problems associated with ill-conditioned estimated covariance matrices. This motivates the study of reconditioning methods later in the thesis.

In Chapter 5 we address RQ 1. We develop new bounds on the condition number of the Hessian of the unpreconditioned 3D-Var objective function in terms of its constituent matrices, using similar techniques to those of Haben et al. [2011a,b] and Haben [2011]. We explicitly consider the case of correlated observation error covariance matrices for the first time numerically. These theoretical results are general and apply to any choice of covariance matrices with a linear observation operator. We study how these bounds differ from those of Haben [2011] with the introduction of correlated observation error covariance matrices. We test these bounds in a linear numerical framework, and compare how the bounds, the condition number and the convergence of the associated minimisation problem behave for different parameter choices. We demonstrate that the minimum eigenvalue of the observation error covariance matrix is an important term in both upper and lower bounds, and we show that relative importance of the background and observation error covariances is strongly dependent on the observation network. This chapter is based on the paper Tabeart et al. [2018].

In Chapter 6 we address RQ 2. We develop new bounds on the condition number of the Hessian of the preconditioned 4D-Var data assimilation problem. The minimum eigenvalue of the observation error covariance matrix appears in both upper and lower bounds. Numerical experiments reveal that reducing the condition number of either error covariance matrix does not guarantee a reduction in the condition number of the Hessian. The choice of observation operator is important in determining whether altering the background or observation error terms dominates the condition number of the Hessian. We also find that for an idealised spatial correlation problem the choice of background and observation parameters determines whether the condition number of the Hessian is a suitable proxy for convergence of a conjugate gradient problem.

In Chapter 7 we address RQ 3. We develop novel theory on reconditioning methods. Motivated by the importance of the minimum eigenvalue of the observation error covariance matrix we formalise two methods that are used in practice to increase small eigenvalues for a general correlation matrix: ridge regression and the minimum eigenvalue method. We compare theoretically how using both methods changes the standard deviations and correlations of a general covariance matrix, as well as the impact on the variational objective function. We find that ridge regression results in larger increases to standard deviations than the minimum eigenvalue method, and decreases the absolute value of all off-diagonal correlations. We contrast the two reconditioning methods with multiplicative variance inflation, which is frequently used at NWP centres to mitigate for missing correlation information. We implement both methods of reconditioning and multiplicative variance for two examples: a spatial covariance matrix, and an interchannel covariance matrix arising from a satellite based instrument. In the spatial setting, we find that the minimum eigenvalue method introduces spurious correlations at large distances. For the interchannel example the minimum eigenvalue method results in smaller changes to entries of the correlation matrix that the ridge regression method. An idealised data assimilation example reveals that both methods of reconditioning are able to change small scales of the analysis, whereas multiplicative variance inflation cannot reduce sample error on smaller scales. We provide guidance on which method of reconditioning is most suitable for different situations, and discuss aspects of the system that could be used to chose the reconditioning parameter. This chapter is based on the paper Tabeart et al. [2019a].

In Chapter 8 we address RQ 4. We study the impact of using the ridge regression method of reconditioning in a realistic operational system, using the Met Office one-dimensional variational data assimilation (1D-Var) system. This is the first time multiple levels of reconditioning have been compared systematically in an operational system. We investigate how the qualitative conclusions from Chapter 5 apply for a nonlinear problem. We consider how reconditioning affects the convergence of the minimisation routine, as well as changes to quality control, and the estimation of variables of meteorological interest. We find that increased use of reconditioning leads to improved convergence of the 1D-Var routine and reduces the differences between retrieved variables compared to the control. However, using correlated observation error covariance matrices increases the number of observations that pass the quality control step, emphasising that the quality control routine must be tuned when altering the data assimilation system. This chapter is based on a paper that is under review as Tabeart et al. [2019b].

Finally in Chapter 9 we present our main conclusions. We will summarise the key results that answer the research questions presented in Section 1.1. We consider the implications of our findings for the wider data assimilation community. The new bounds developed in this thesis provide users with guidance on how changes to their data assimilation system are likely to affect convergence. This will be useful when considering changes to the system, such as the introduction of a new observation type or improved estimate of the observation error covariance matrix. Combining this knowledge with the theoretical understanding of reconditioning methods presented in this thesis will ensure the use of an appropriate method for a given problem. Finally we suggest future work that could be undertaken to extend the conclusions of this thesis.

Chapter 2

Data assimilation

In this thesis we study how variational data assimilation problems are affected by the introduction of correlated observation error covariance matrices. In this chapter we derive the variational data assimilation problem, starting with Bayes' theorem in Section 2.1. We discuss the key variables of interest and introduce much of the notation that will be used for the remainder of this work. In particular, we introduce several formulations of the variational data assimilation problem including 4D-Var (Section 2.2), the incremental formulation (Section 2.2.1) and the control variable transform (Section 2.2.2). We discuss the sources and types of observation errors in Section 2.3. In Section 2.4 we introduce the diagnostic of Desroziers et al. [2005], which is used to estimate correlated observation error matrices at NWP centres.

2.1 Derivation of 3D-variational data assimilation methods from Bayes' Theorem

The aim of data assimilation is to find the most probable initial state of a dynamical system given a prior estimate, or background, and observations of the system. The background is given by a forecast from a previous time [Rawlins et al., 2007]. There are different types of data assimilation algorithms, which allow users to make best use of available resources. Alternative formulations often impose additional assumptions on the generic problem of interest. One common method of data assimilation in numerical weather prediction is variational data assimilation [Daley, 1991, Kalnay, 2002]. This makes use of a least squares objective function comprised of two terms: the background term and observation term, which measure the discrepancy from the background and observations respectively. Each term is weighted by its respective uncertainties. The objective function is then minimised to find the most probable initial state. This allows the analysis (most likely initial state) to pull close to the background in the absence of observations, or where observation values are uncertain, and to take advantage of 'good' observational information to improve on the prior.

We begin by presenting Bayes' Theorem. This result will then be used to derive the variational formulation of data assimilation.

Theorem 2.1.1 (Bayes' Theorem [Lewis et al., 2006])**.** *The probability density function of* $\mathbf{x}$ *given* $\mathbf{y}$

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}. \tag{2.1}$$

If $\mathbf{x}$ denotes the unknown state and $\mathbf{y}$ denotes observations of our system, then Bayes' theorem tells us that we can calculate the posterior density of the state conditioned on the observations, $P(\mathbf{x}|\mathbf{y})$, as the product of the likelihood of the observations conditioned on the model state, $P(\mathbf{y}|\mathbf{x})$, and the prior probability density function (pdf) of the model state, $P(\mathbf{x})$. This product is then normalised by the marginal density of the observations, $P(\mathbf{y})$. We note that, as this term is independent of the state $\mathbf{x}$, it is simply a normalising constant. As the aim of variational assimilation is to find the state that maximises $P(\mathbf{x}|\mathbf{y})$, called the maximum a posteriori (MAP) estimate, in this setting the normalising constant is typically neglected.

In the variational formulation with Gaussian errors, we can write each of the terms in Theorem 2.1.1 explicitly. The assumption of Gaussian errors is key for the variational data assimilation method, however it does not hold for all variables. In the case that model state variables and observations are normally distributed, and that our observation operator is linear, we have equivalence of the minimum variance estimate and maximum a posteriori (MAP) estimate as the mean and mode of a normal distribution are equal [Lewis et al., 2006].

Under these assumptions, the pdf of the prior is given by

$$P(\mathbf{x}) \propto \exp\{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\}, \tag{2.2}$$

[Lewis et al., 2006] where $\mathbf{x}_b \in \mathbb{R}^N$ is the background, or prior, state and $\mathbf{B} \in \mathbb{R}^{N\times N}$ is the background error covariance matrix. The conditional probability of the observations given the state can then be written as

$$P(\mathbf{y}|\mathbf{x}) \propto \exp\{-\frac{1}{2}(\mathbf{y}-h(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-h(\mathbf{x}))\}, \tag{2.3}$$

[Lewis et al., 2006] where $\mathbf{y} \in \mathbb{R}^p$ is the vector of observations, $\mathbf{R} \in \mathbb{R}^{p \times p}$ is the observation error covariance matrix, and $h : \mathbb{R}^N \to \mathbb{R}^p$ is the observation operator. The observation operator maps from state space into observation space in order to compare state variables with observations, and can be nonlinear (for example for satellite based instruments which measure top of the atmosphere radiances [Eyre, 1989]).

Under the additional assumption that observation errors are independent of background errors, and using the result of Theorem 2.1.1, we can combine (2.2) and (2.3) to obtain the posterior pdf of the state

$$P(\mathbf{x}|\mathbf{y}) \propto \exp\{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) - \frac{1}{2}(\mathbf{y}-h(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-h(\mathbf{x}))\}. \qquad (2.4)$$

As $P(\mathbf{y})$ is independent of the model state we can neglect it when computing the analysis, or most likely initial state, $\mathbf{x}_a$. The assumption that background and observation errors are independent does not always hold. For example, sometimes background fields are used for quality control purposes, such as cloud detection, to reject observations that are not permitted in the assimilation system. In this case, errors in the background and the observation will be correlated artificially by the quality control process [Bathmann, 2018]. However, in this thesis we will assume that observation and background errors are independent.

We wish to compute the analysis, that is the state $\mathbf{x}_a \in \mathbb{R}^n$ with MAP probability estimate given the observations and prior. We can reframe a maximisation problem as a minimisation problem by finding the state $\mathbf{x}$ that minimises the objective function

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}(\mathbf{y}-h(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-h(\mathbf{x})). \qquad (2.5)$$

The analysis, $\mathbf{x}_a \in \mathbb{R}^n$, minimises (2.5) and is the state that maximises (2.4); hence $\mathbf{x}_a$ is the most likely state given the observations and prior information. The function (2.5) is referred to as the cost function or objective function.

In order to study how the convergence of an iterative method used to solve the minimisation problem is affected by changes to the data assimilation system, we can consider the impact of those changes on the conditioning of the Hessian (matrix of second derivatives) of the objective function (2.5). The condition number will be defined formally in Section 3.1, but it can be used to study the sensitivity of the solution to small changes to background or observation data [Golub and Van Loan, 1996, Sec. 2.7]. The Hessian of the linearised objective function (2.5) is given by

$$\mathbf{S} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}, \tag{2.6}$$

where $\mathbf{H} \in \mathbb{R}^{p\times N}$ is the Jacobian of the observation operator h linearised about the current best estimate of the optimal solution of (2.5). In this thesis we will use the conditioning of (2.6) to study the sensitivity of the variational assimilation problem to changes to the observation error covariance matrix, $\mathbf{R}$.

2.2 Four-dimensional variational data assimilation

The variational objective function (2.5) has no time variable, and all observations are assumed to have been made at the same time. This formulation corresponds to 3D-Var [Lorenc et al., 2000]. The objective function is minimised over the relevant time window, with the assumption that all observations are made at the same time (typically halfway through the window). In reality, observations are made throughout the time window and allowing them to be fitted at the correct time is important to improve forecasts. For example, multiple observations could be made at the same location over one time window. The 4D-Var formulation permits the inclusion of all of these observations in the objective function and hence can account for dynamic evolution of the system over the time window. Many NWP centres now use 4D-Var, even for limited area models (e.g. Rawlins et al. [2007]). At some centres, 3D-Var with first guess at appropriate time (3D-Var FGAT), which we will also discuss briefly in this section, has replaced standard 3D-Var [Fisher and Andersson, 2001, Simonin et al., 2019].

2.2.1 Incremental 4D-Var

For a given time window $[t_0, t_n]$ let $\mathbf{x}_b$ be the background state and $\mathbf{B}_0 \in \mathbb{R}^{N\times N}$ be the background error covariance at the initial time t_0. Let $\mathbf{y}_i$ be observations taken at times $t_k, k = 0, 1, 2, \dots, n$, with corresponding observation error covariance matrix $\mathbf{R}_k \in \mathbb{R}^{p\times p}$, and let h_k be the, possibly nonlinear, observation operator that maps background state $\mathbf{x}_k$ at time t_k to the observation space at time t_k. Then, the full 4D-Var objective function for this system is given by

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T\mathbf{B}_0^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}\sum_{k=0}^{n}(\mathbf{y}_k - h_k[\mathbf{x}_k])^T\mathbf{R}_k^{-1}(\mathbf{y}_k - h_k[\mathbf{x}_k]) \tag{2.7}$$

subject to the nonlinear forecast model

$$\mathbf{x}_k = \mathcal{M}(t_{k-1}, t_k; \mathbf{x}_{k-1}). \tag{2.8}$$

The primal variational 4D-Var problem is typically solved in an incremental form. This general framework is used at both the Met Office and the European Centre for Medium-Range Weather Forecasts (ECMWF) [Rabier et al., 1998, Rawlins et al., 2007], where the nonlinear objective function (**??**) is solved via a series of inner and outer loops. This framework has been shown to be equivalent to a quasi-Newton method [Gratton et al., 2007, Lawless et al., 2005a]. A small number of outer loops solve the full nonlinear problem, and a larger number of inner loops solve the linearised problem. At the Met Office this inner loop is solved using the conjugate gradient method [Haben et al., 2011b]. The conjugate gradient method will be defined formally in Section 3.2.

The use of 4D-Var results in additional algorithmic complexities. The main computational difficulty is the need to run a dynamic model over the assimilation window in order to calculate the objective function. Tangent linear and adjoint models are formed to calculate the model trajectory and gradient of the objective function in the inner loop, and the full nonlinear model is used in the outer loop. Due to the high dimension and complexity of the state and observations, line-by-line adjoints are used at NWP centres [Errico and Raeder, 1999]. This means that any change to the model requires that the adjoint code be updated, making maintenance of adjoint codes costly.

Let $\mathbf{x}_i^b = \mathcal{M}(t_{i-1}, t_i; \mathbf{x}_{i-1}^b)$. Define $\delta\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_i^b$. We then consider the Taylor expansion of $\mathcal{M}(t_{i-1}, t_i; \mathbf{x}_{i-1})$ about $\mathbf{x}_i^b(t)$

$$\mathbf{x}_i^b + \delta\mathbf{x}_i = \mathcal{M}(t_{i-1}, t_i; \mathbf{x}_{i-1}^b) + \mathbf{M}_i\delta\mathbf{x}_{i-1} + \text{higher order terms} \tag{2.9}$$

$$\delta\mathbf{x}_i \approx \mathbf{M}_i\delta\mathbf{x}_{i-1} \tag{2.10}$$

where $\mathbf{M}_i \in \mathbb{R}^{N\times N}$ is the linearised model operator at time t_i, linearised about $\mathbf{x}_i^b$. Similarly, expanding $h_i[\mathbf{x}_i]$ about $\mathbf{x}_i^b$ we obtain

$$h_i[\mathbf{x}_i] \approx h_i[\mathbf{x}_i^b] + \mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}_i \tag{2.11}$$

where $\mathbf{H}_i \in \mathbb{R}^{N\times p_i}$ is the linearised observation operator at time t_i linearised around $\mathbf{x}_i^b$.

We then write the linearised objective function in terms of $\delta\mathbf{x}_0$

$$J(\mathbf{x}_0) = \frac{1}{2}\delta\mathbf{x}_0^T\mathbf{B}^{-1}\delta\mathbf{x}_0 + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{d}_i - \mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}_0)^T\mathbf{R}_i^{-1}(\mathbf{d}_i - \mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}_0) \tag{2.12}$$

where $\mathbf{d}_i = \mathbf{y}_i - h_i[\mathbf{x}_i^b]$ are the innovation vectors. These measure the misfit between the observations and the linearisation state, using the full nonlinear observation operator.

The 3D-Var first guess at appropriate time (3D-FGAT) algorithm can be obtained from (2.12) with $\mathbf{M}_k = \mathbf{I}$ in (2.9) [Fisher and Andersson, 2001, Lorenc and Rawlins, 2005]. 3D-FGAT propagates the background field forward to the time of the observations, but does not propagate the increment $\delta\mathbf{x}_k$, and is hence computationally cheaper than 4D-Var.

2.2.2 The control variable transform

One practical problem with implementing incremental 4D-Var is the cost of either explicitly forming the background error covariance $\mathbf{B}$, or evaluating matrix-vector products. This is partly due to the large size of this matrix; the number of state variables can be of the order of 10^9 [Carrassi et al., 2018]. This motivates the use of the control variable transform (CVT) to model the background error covariance matrix. The CVT uses the matrix square root of $\mathbf{B}$ to perform a variable transformation.

In order to simplyfy the notation in what follows, define the generalised observation operator as

$$\widehat{\mathbf{H}} = \left[\mathbf{H}_0^T, (\mathbf{H}_1\widehat{\mathbf{M}}_1)^T, \ldots, (\mathbf{H}_n\widehat{\mathbf{M}}_n)^T\right]^T, \tag{2.13}$$

where the linearised forward model from time t_0 to time t_i is given by

$$\widehat{\mathbf{M}}_i(\delta\mathbf{x}_i) = \mathbf{M}(t_i, t_0; \delta\mathbf{x}_{i-1}) = \mathbf{M}_i \ldots \mathbf{M}_1. \tag{2.14}$$

Similarly, we define

$$\widehat{\mathbf{d}}^T = \left[\mathbf{d}_o^T, \mathbf{d}_1^T, \ldots, \mathbf{d}_n^T\right] \tag{2.15}$$

is a vector made up of the innovation vectors.
Finally let $\widehat{\mathbf{R}} \in \mathbb{R}^{Q\times Q}$ denote the block diagonal matrix with the ith block consisting

of $\mathbf{R}_i$:

$$\delta\mathbf{x}_i = \mathbf{M}(t_{i-1}, t_i; \mathbf{x}_{i-1})\delta\mathbf{x}_{i-1} \equiv \mathbf{M}_i\delta\mathbf{x}_{i-1}. \tag{2.16}$$

The control variable transform (CVT) is then applied to the incremental form of the variational problem (6.6), via the change of variable $\delta\mathbf{z}_0 = \mathbf{B}^{-1/2}\delta\mathbf{x}_0$. The new variables $\delta\mathbf{z}_0$ are hence uncorrelated, with unit variances. This simplifies the background term in the updated objective function

$$J(\delta\mathbf{z}_0) = \frac{1}{2}\delta\mathbf{z}_0^T\delta\mathbf{z}_0 + \frac{1}{2}\left(\widehat{\mathbf{d}} - \widehat{\mathbf{H}}\mathbf{B}^{1/2}\delta\mathbf{z}_0\right)^T\widehat{\mathbf{R}}^{-1}\left(\widehat{\mathbf{d}} - \widehat{\mathbf{H}}\mathbf{B}^{1/2}\delta\mathbf{z}_0\right). \tag{2.17}$$

where $\mathbf{z}_0 = \mathbf{B}^{-1/2}\mathbf{x}_0$, $\mathbf{z}_b = \mathbf{B}^{-1/2}\mathbf{x}_b$.
It can be shown [Haben et al., 2011b] that use of the CVT is equivalent to pre- and post-multiplying the Hessian of the unpreconditioned data assimilation problem (6.10) by $\mathbf{B}^{1/2}$ (the uniquely defined, symmetric square root of $\mathbf{B}$). This yields a preconditioned Hessian for 4D-Var given by

$$\widehat{\mathbf{S}} = \mathbf{I}_N + \mathbf{B}^{1/2}\widehat{\mathbf{H}}^T\widehat{\mathbf{R}}^{-1}\widehat{\mathbf{H}}\mathbf{B}^{1/2}. \tag{2.18}$$

We note that with the additional assumption that $\mathbf{B}$ and $\mathbf{R}$ are strictly positive definite, then $\widehat{\mathbf{S}}$ is symmetric positive definite.

In this thesis, we will study the conditioning of the preconditioned incremental 4D-Var data assimilation problem (2.17) separately from the 3D-Var data assimilation problem (2.5). It can be shown [Haben et al., 2011b] that use of the CVT is equivalent to pre- and post-multiplying the Hessian of the unpreconditioned data assimilation problem by $\mathbf{B}^{1/2}$ (the uniquely defined, symmetric square root of $\mathbf{B}$). This yields a preconditioned Hessian for the 4D-Var problem (2.17) given by

$$\hat{\mathbf{S}} = \mathbf{I}_N + \mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2}, \tag{2.19}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix. We will consider the conditioning of (2.19) in Chapter 6.

2.3 Observation errors

In this thesis we study how the use of correlated observation errors will alter the variational data assimilation problems (2.5) and (**??**). In this section, we motivate our interest in correlated observation errors from a theoretical and practical perspective, and consider examples of instruments where errors between observations are likely to be related.

2.3.1 Definition of observation error

We begin by considering what is typically meant by ‘observation error’. The observation error covariance matrix, $\mathbf{R}$, accounts for uncertainty in the observations as well as uncertainty in the observation operator. In this way the error covariance matrix, $\mathbf{R}$, can be thought of as accounting for the statistics of all error associated with the observation term in (2.5) [Waller et al., 2014b]. For 4D-Var, model error also contributes to errors in the observation term of (2.7), but this term can be separated from observation errors (see Moodey et al. [2013], Howes et al. [2017]). We will not consider model error in this thesis.

A review of the different kinds of statistics of the errors that are included in the observation error covariance matrix is given by Janjić et al. [2018].We now describe some of the components that are included in the observation error term and whether they are likely to be correlated or uncorrelated.

- Error due to unresolved scales describes the difference between a perfect observation and a perfect observation of the scales resolved by the model. Depending on the synoptic situation, this type of error can be correlated between different observations [Janjić et al., 2018].

- Observation operator error arises due to the use of an approximate observation operator, rather than the true infinite-dimensional observation operator. In practical applications, observation operator error is increased as approximate observation operators are used to minimise computational cost (e.g. Waller et al. [2016c], Gauthier et al. [2018]). This can be correlated between different observations [Janjić et al., 2018].

- Preprocessing error may be correlated between different observations; one example is height assignment errors for atmospheric motion vectors [Bormann et al., 2003, Cordoba et al., 2017].

- Measurement error, also known as instrument noise. In the case of stochastic noise, measurement errors are uncorrelated between different observations. However, there are some instruments where this may not be the case, for example due to apodisation effects [Gambacorta and Barnet, 2013]. Typically instrument noise levels are well estimated by instrument manufacturers.

2.3.2 Treatment of observations and their uncertainties

In Chapter 8 we will present a case study using observations from a satellite-based instruments. Much previous research has considered the impact of correlated observation errors for hyperspectral instruments [Weston, 2011, Weston et al., 2014, Stewart et al., 2014, Campbell et al., 2017, Bormann et al., 2016]. Hyperspectral instruments are situated on satellites and measure top of the atmosphere radiances in the infrared, visible and microwave spectrum. Radiative transfer equations are used to compare measurements of brightness temperature with the meteorological variables of interest in the state vector. Although the radiative transfer equations can be solved very accurately using a line-by-line model, the need for timely forecasts mean that implementations such as RTTOV, which is used at the Met Office and ECMWF [Eyre, 1989, Matricardi et al., 2004], have to balance accuracy with speed in order to produce a fast solver.

For each instrument the relevant part of the electromagnetic spectrum is split into bands, with each band corresponding to a different ‘channel’. Instruments such as the Infrared Atmospheric Sounding Interferometer (IASI) can return measurements for 8461 channels across its spectral range [Collard, 2007]. As the channels are very narrow for hyperspectral instruments, it is can be difficult to ensure that channels are spectrally independent [Stewart et al., 2014]. In order to minimise the amount of duplicated information that is passed to the data assimilation system, NWP centres typically select a subset of around 300 channels [Stewart, 2010, Chalon et al., 2001] that provide independent information, and maximise the information content of the selected channels [Collard, 2007, Rabier et al., 2002, Fowler, 2017]. However, there is still overlap between weighting functions, which contain information about the sensitivity of a single channel to different pressure levels in the atmosphere [Stewart, 2010]. Overlapping weighting functions will lead to correlated observations, but not necessarily correlated observation errors. However, systematic errors, for example in the radiative transfer equation, may lead to correlated errors between different channels with overlapping weighting functions. Errors due to unresolved scales may also occur, as the instrument can measure on spatial scales which are too small to be represented well by the model. Therefore, for spectrally close channels, with similar weighting functions, the same feature might be misrepresented in a similar way, leading to correlated observation errors between channels [Stewart et al., 2014].

Although it was known that correlated observation errors existed, prior to the last decade, uncorrelated observation errors were assumed for all instruments. Partly this is due to the difficulty of estimating observation error statistics. Additionally, using non-diagonal correlation matrices increases the computational expense of inverting the observation error covariance matrix. For spatial correlations, thinning is one technique that can allow users to neglect correlated observation errors. For some instruments estimated correlation lengthscales are shorter than typical thinning distances [Bennitt et al., 2017], meaning that thinning can be a valid technique. However, for other instruments the correlation lengthscales have been found to be much longer than reasonable thinning distances [Waller et al., 2016a,c, Cordoba et al., 2017], meaning that correlations must be taken into account. The use of thinning results in a large number of observations being discarded: in the Metéo France convection-permitting limited area model radar observations are horizontally thinned by a factor of 64 and infrared satellite observations are horizontally thinned by a factor of 400 [Michel, 2018]. Thinning may also be necessary due to the large size of observation datasets, which can cause difficulties with storage and computational resource. However, alternative data compression methods, such as using a Fourier transform to retain only the largest modes of observation information, may allow a larger amount of information to be retained, whilst reducing the computational burden [Fowler, 2019].

The inclusion of correlated observation error information is crucial, particularly with the increasing desire for high resolution forecasts. In order to produce beneficial local forecasts, we need to exploit existing high density observation networks [Dance et al., 2019]. This means correlated observation errors must be taken into account in order to reduce observation thinning. Failing to account for correlated observation error information where it is present has been found to artificially cap forecast skill [Stewart et al., 2008b, 2013, Rainwater et al., 2015]. Work by Stewart [2010] and Healy and White [2005] found that including some correlation information is better than none, so that even approximate observation error statistics bring theoretical benefit to data assimilation systems.

The first study to estimate correlated observation error covariance matrices was carried out using the Met Office system for the IASI instrument [Stewart et al., 2008a]. Correlated observation error matrices were subsequently implemented operationally at the Met Office for IASI and the Atmospheric Infrared Sounder (AIRS) [Weston, 2011, Weston et al., 2014, Stewart et al., 2014]. Observation error covariance matrices were estimated separately for the 1D-Var and 4D-Var systems, and correlations were found to be much larger for the 4D-Var assimilation system. This is consistent with the fact that for 1D-Var error correlations arise from forward model and adjacent channel errors, whereas for 4D-Var errors of representativity are expected to dominate.

Observations from IASI are sensitive to different parts of the atmosphere and a variety of meteorological variables, including temperature and humidity. The largest correlations occur for channels that are sensitive to water vapour. The findings of other centres for IASI are very similar [Bormann et al., 2015, 2016, Campbell et al., 2017]. Large correlations for humidity sensitive channels have also been found for other infrared and microwave instruments (e.g. Waller et al. [2016a], Wang et al. [2018]).

Many of the early operational implementations of correlated observation error focused on interchannel errors. Parallelisation of code makes the use of spatial correlations potentially expensive: if two observations are correlated and assigned to different processors, then expensive communication will be required. However, placing all observations whose errors are correlated on one processor prevents full exploitation of a large number of processors. Michel [2018] proposed a Lanczos-based method which combines a reduced rank approximation to the observation error covariance matrix with regularisation to ensure matrix inversion could be performed by a sequence of linear operators rather than direct inversion. This method eliminates the need for global communication and could hence be parallelised. Alternatively, parameters for a specified correlation function or operator can be estimated in place of the full error covariance matrix. This is done by Guillet et al. [2019] using an implementation of a finite element method that is computationally efficient to invert. Simonin et al. [2019] proposed an alternative parallelisation scheme which groups observations with mutually correlated errors together for processing.

2.4 The diagnostic of Desroziers et al. [2005]

In the last decade, the use of correlated observation error covariance matrices in data assimilation has grown enormously. In Section 2.3, we discussed the benefits of including observation error correlations, such as the improved use of existing high-density observations, and the desire to move towards higher resolution forecasts. In this section we discuss one of the most popular methods that is used to estimate error correlations for NWP systems. We define the method and discuss some of its limitations.

The diagnostic of Desroziers et al. [2005] (henceforth referred to as DBCP) was originally designed to check whether the choice of background and observation error covariance matrices are consistent with the data assimilation system of interest. However, it was suggested that if users know that their error covariance matrices should be correlated, but are using diagonal matrices, the DBCP diagnostic could be used as a method to estimate the missing correlation information. We now provide a brief overview of the method.

Let $\mathbf{R}_t$, $\mathbf{B}_t$ be the 'true' error covariance matrices and $\mathbf{R}$, $\mathbf{B}$ be the assumed error statistics that are used in the variational data assimilation problem (2.5). We write the analysis, which minimises (2.5), as an update to the background, $\mathbf{x}_b$,

$$\mathbf{x}^a = \mathbf{x}_b + \mathbf{K}\mathbf{d}_b^o, \tag{2.20}$$

where

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \tag{2.21}$$

is the Kalman gain, and the innovation, or background residual, is given by

$$\mathbf{d}_b^o = \mathbf{y} - h(\mathbf{x}_b). \tag{2.22}$$

Let the analysis residual be given by

$$\mathbf{d}_a^o = \mathbf{y} - h(\mathbf{x}_a). \tag{2.23}$$

Desroziers et al. [2005] showed that

$$E[\mathbf{d}_a^o \mathbf{d}_b^{oT}] = \mathbf{R}(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{H}\mathbf{B}_t\mathbf{H}^T + \mathbf{R}_t) = \mathbf{R}^e, \tag{2.24}$$

i.e. that calculating the expectation of the analysis and background residuals provides an updated estimate for the observation error covariance matrix. In the case that the assumed error covariances are the exact error covariance matrices, i.e. $\mathbf{B} = \mathbf{B}_t$ and $\mathbf{R} = \mathbf{R}_t$, then (2.24) simplifies to

$$E[\mathbf{d}_a^o \mathbf{d}_b^{oT}] = \mathbf{R}_t. \tag{2.25}$$

In Desroziers et al. [2005], other update equations, similar to (2.24), are presented for the background error covariance matrix in observation space, $\mathbf{H}\mathbf{B}\mathbf{H}^T$, and the sum of the background and observation error covariance matrices in observation space. The method can either be performed as a single step, or the updated estimates can be used to run a new data assimilation procedure before repeating the algorithm above. This yields the iterative formulation.

Although practical implementations of the method fail to satisfy the assumptions of the original diagnostic, it is possible to obtain useful updated matrices $\mathbf{R}^e$ even if the input background and observation error covariance matrices are imperfect (e.g. Waller et al. [2016a,c]). Theoretical studies have considered how realistic implementations of the DBCP diagnostic are likely to perform in practice. In the case that both background and observation errors are homogeneous, Waller et al. [2016b] found that if the two error covariance matrices are structurally similar, the iterative method fails, but reasonable estimates may be obtained from a single iteration.

Bathmann [2018] found that convergence of the iterative version of the diagnostic depends on the input background error covariance matrix, with divergence occurring when $\mathbf{B}$ is overestimated. This is due to the fact that the updated observation error covariance matrix becomes rank deficient [Ménard, 2016]. Additionally convergence speed depends on the eigenvalues of both the background and observation error covariance matrix, as well as the distance between the input and true observation error covariance matrix [Ménard, 2016]. Although the iterative method may yield improved results compared to the single step iteration, it is extremely computationally expensive. Most NWP centres use the single step version of the diagnostic; a single iteration was performed in the U.S. Naval Research Laboratory (NRL) system and changes to the estimate of $\mathbf{R}$ were small [Campbell et al., 2017].

As well as providing improved estimates for observation error covariance matrices, the DBCP diagnostic can also be used to identify errors in the data assimilation and quality control routines. In Waller et al. [2016a] spatial variation in error statistics revealed that the quality control process was failing to remove mixed land-sea pixels. In Waller et al. [2016c], investigations into correlated observation error using the DBCP diagnostic led to the development of an improved observation operator.

For many instruments it can be difficult to validate the results of the DBCP diagnostic. However, expert knowledge of the instrument and observing system can provide insight into whether estimated OEC matrices are reasonable. For example standard deviation values can be compared against instrument noise information provided by the manufacturer, and in the case of spatially correlated observation errors correlation lengthscale can be compared against the previous best estimate. Consistency across different NWP centres using the same instruments but different assimilation systems can provide some confidence in the estimated correlation structure [Waller et al., 2019]. There has been recent work developing error inventories for specific instruments [Merchant et al., 2014] and metrologically derived uncertainties, so it may be possible to compare estimates from the DBCP diagnostic with alternative derivations of error in the future. In most previous work a single error covariance matrix is estimated and used globally. With moves towards all-sky assimilation of satellite observations [Geer, 2019] and increasing use of correlated observation errors, it is likely that situation dependent observation error matrices will be more common in the future.

2.5 Summary

In this chapter we defined the unpreconditoned and preconditioned variational data assimilation problems, and introduced the notation that will be used throughout this thesis. We discussed sources of observation error and why the inclusion of correlated observation error information is important for NWP centres. We presented the diagnostic of Desroziers et al. [2005], and discussed some of the challenges and benefits associated with its use in operational systems. In the next chapter we introduce the concept of conditioning, and some key results from numerical linear algebra.

Chapter 3

Results from numerical linear algebra

In this chapter we introduce key definitions and results from numerical linear algebra. In Section 3.1 we introduce the concept of matrix and vector norms and formally define what is meant by the condition number. We introduce the conjugate gradient method in Section 3.2 and discuss the relationship between convergence of this method and the condition number. In Section 3.3 we present results which allow us to write the eigenvalues of products and sums as products or sums of eigenvalues of individual matrices. In Section 3.4 we present a variety of matrix structures of interest to the data assimilation problem. These results will be used to develop theory on the conditioning of the variational data assimilation problem, and for numerical experiments in subsequent chapters. We begin by introducing a convention on the ordering of eigenvalues that will be used throughout this thesis.

Definition 3.0.1. *For* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *let the eigenvalues of* $\mathbf{A}$ *be given by*

$$\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq \cdots \geq \lambda_n(\mathbf{A}). \tag{3.1}$$

3.1 Vector and matrix norms and the condition number

In order to define the condition number, we must first introduce the concepts of vector and matrix norms, which we do in this section. We also present matrix structures that will be used in a particular characterisation of the condition number. We begin by defining the concept of a vector norm.

Definition 3.1.1 (Golub and Van Loan [1996])**.** *A vector norm on $\mathbb{R}^n$ is a function $f : \mathbb{R}^n \to \mathbb{R}$ that satisfies the following properties:*

$$f(\mathbf{x}) \geq 0, \quad \mathbf{x} \in \mathbb{R}^n \quad (f(\mathbf{x}) = 0 \iff \mathbf{x} = 0) \tag{3.2}$$

$$f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \tag{3.3}$$

$$f(\alpha\mathbf{x}) = |\alpha| f(\mathbf{x}), \quad \alpha \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n. \tag{3.4}$$

We now define a special and important class of vector norms: the $p-$norm.

Definition 3.1.2. *For $p \geq 1$ the $p-$norm on $\mathbb{R}^n$ is defined as*

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^2 + \cdots + |x_n|^p)^{1/p} \tag{3.5}$$

where x_i denotes the ith component of $\mathbf{x}$.

The most commonly used $p-$norms are given by $p = 1, 2, \infty$.

We now introduce the definitions of symmetric matrices and positive definite matrices.

Definition 3.1.3. *A square matrix, $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric if $\mathbf{A}^T = \mathbf{A}$, that is if $a_{ij} = a_{ji} \forall i, j = 1, 2, 3, \ldots, n$.*

Lemma 3.1.4. *The sum of two symmetric matrices of the same size is symmetric*

Proof. Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ be symmetric matrices. We consider their sum

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T = \mathbf{A} + \mathbf{B}. \tag{3.6}$$

Therefore the sum of two conformable symmetric matrices is symmetric. □

Definition 3.1.5. *A symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is called positive definite if for any non-zero vector $\mathbf{x} \in \mathbb{R}^n$*

$$\mathbf{x}^T \mathbf{A} \mathbf{x} > 0. \tag{3.7}$$

We denote the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ as $\|\mathbf{x}\|_{\mathbf{A}}^2$.

We now introduce characterisations of positive definite and positive semi-definite matrices in terms of their eigenvalues.

Theorem 3.1.6. *A symmetrix matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is positive definite if and only if all of its eigenvalues are strictly positive.*

Proof. See Gentle [2007, Sec 3.8.8] □

Definition 3.1.7. *A symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is called positive semidefinite if for any non-zero vector $\mathbf{x} \in \mathbb{R}^n$, the quadratic form is non-negative i.e.*

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0. \tag{3.8}$$

Theorem 3.1.8. *A symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is positive semidefinite if its eigenvalues are non-negative.*

Proof. See Gentle [2007, Sec 3.8.8] □

Symmetric positive definite (SPD) matrices have useful properties in terms of their eigendecomposition. We will show in Section 3.4 that correlation matrices are symmetric positive semi-definite (SPSD). In practice we often restrict ourselves to the case of strictly SPD matrices, so many of the properties presented in this section will apply directly. We now introduce some key properties of SPD matrices.

Lemma 3.1.9. *The sum of any two positive definite matrices of the same size is positive definite.*

Proof. Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ be positive definite matrices. We consider the quadratic form of the sum $\mathbf{A} + \mathbf{B}$:

$$\mathbf{x}^T (\mathbf{A} + \mathbf{B}) \mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{x}^T \mathbf{B} \mathbf{x} > 0. \tag{3.9}$$

For a non-zero choice of $\mathbf{x}$, both components of (3.9) are strictly positive as $\mathbf{A}, \mathbf{B}$ are positive definite matrices. Therefore the quadratic form of $\mathbf{A} + \mathbf{B}$ is strictly positive and the sum of two positive definite matrices is positive definite. □

Combining the results of Lemmas 3.1.4 and 3.1.9 we conclude that the sum of two SPD matrices is SPD. Similarly, we can prove that the sum of a positive definite matrix and a positive semi-definite matrix is positive definite.

Lemma 3.1.10. *The sum of a positive definite matrix and a positive semi-definite matrix is positive definite.*

Proof. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be positive definite and $\mathbf{B} \in \mathbb{R}^{n \times n}$ be positive semi-definite. Then, for any $\mathbf{x} \in \mathbb{R}^n$, consider $\mathbf{x}^T (\mathbf{A} + \mathbf{B}) \mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{x}^T \mathbf{B} \mathbf{x} > 0$. The first term is strictly greater than zero and the second term is greater than or equal to zero for any choice of $\mathbf{x}$ by definition. □

The results of Lemmas 3.1.4 and 3.1.10 guarantee that if the background error covariance matrix is strictly positive definite, the Hessians (2.6) and (2.19) are symmetric positive definite. We now briefly outline the proof: (2.6) and (2.19) are the sum of a SPD matrix and a SPSD matrix. Hence by the result of Lemma 3.1.10 the Hessians (2.6) and (2.19) are strictly positive definite. The result of Lemma 3.1.4 proves that both Hessians are symmetric, and hence (2.6) and (2.19) are SPD. Therefore the properties of SPD matrices presented in this section apply directly to the Hessian for both unpreconditioned and preconditioned formulations.

We now present some properties of general SPD matrices, beginning by characterising the eigendecomposition of a symmetric matrix.

Theorem 3.1.11. *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then we can decompose $\mathbf{A}$ as*

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T \tag{3.10}$$

where $\boldsymbol{\Lambda} \in \mathbb{R}^{n\times n}$ is a diagonal matrix of eigenvalues and $\mathbf{V} \in \mathbb{R}^{n\times n}$ is the corresponding orthogonal matrix of eigenvectors of $\mathbf{A}$, i.e. $\mathbf{V}^T\mathbf{V} = \mathbf{I}_n$.

Proof. See Gentle [2007, Sec 3.8.7] □

Using Theorem 3.1.11 we prove that the inverse of a SPD matrix is itself SPD.

Lemma 3.1.12. *The inverse of a symmetrix positive definite matrix is symmetric positive definite.*

Proof. Writing $\mathbf{A} \in \mathbb{R}^{n\times n}$ as in (3.10) we obtain $\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T$. We note that all the elements of $\boldsymbol{\Lambda}$ are positive by the result of Definition 3.1.5. We calculate the inverse of $\mathbf{A}$

$$\mathbf{A}^{-1} = (\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T)^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^T. \tag{3.11}$$

As the entries of $\boldsymbol{\Lambda}$ are all positive, so are the entries of $\boldsymbol{\Lambda}^{-1}$. Hence, all of the eigenvalues of $\mathbf{A}^{-1}$ are strictly positive and $\mathbf{A}$ is positive definite. We note that $\mathbf{A}^{-1}$ is symmetric by definition, and hence $\mathbf{A}^{-1}$ is symmetric positive definite. □

Lemma 3.1.13. *For $\mathbf{B} \in \mathbb{R}^{n\times n}$ positive definite and $\mathbf{A} \in \mathbb{R}^{n\times m}$ of rank m, the product $\mathbf{A}^T\mathbf{B}\mathbf{A}$ is positive definite.*

Proof. [Gentle, 2007, p89] □

We now define the concept of a matrix norm.

Definition 3.1.14. $\|\cdot\| : \mathbb{R}^{m\times n} \to \mathbb{R}$ *is a matrix norm for all* $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$*,* $\mathbf{C} \in \mathbb{R}^{n\times p}$*,* $c \in \mathbb{R}$ *if the following properties are satisfied*

$$\|\mathbf{A}\| \geq 0 \textit{ with equality if and only if } \mathbf{A} = 0, \tag{3.12}$$

$$\|c\mathbf{A}\| = |c|\|\mathbf{A}\|, \tag{3.13}$$

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|, \tag{3.14}$$

$$\|\mathbf{AC}\| \leq \|\mathbf{A}\|\|\mathbf{C}\|. \tag{3.15}$$

We now introduce some norms that will be used later in the thesis.

An important family of matrix norms are those that arise from vector norms. These are called induced, subordinate, or operator norms and are defined by the original vector $p-$norm that was introduced in Definition 3.1.2.

Definition 3.1.15. *For* $1 \leq p \leq \infty$ *the p-norm is defined as*

$$\|\mathbf{A}\|_p = \sup_{\mathbf{x}\neq 0} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p} \tag{3.16}$$

Commonly used induced norms, defined for $\mathbf{A} \in \mathbb{R}^{m\times n}$, are:

- The 2-norm: $\|\mathbf{A}\|_2 = \sigma_{max}(\mathbf{A})$, where $\sigma_{max}(\mathbf{A})$ represents the largest singular value of the matrix $\mathbf{A}$. In the case that $\mathbf{A}$ is SPD the 2-norm of $\mathbf{A}$ is given by its largest eigenvalue.
- The 1-norm: $\|\mathbf{A}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{i,j}|$ i.e. the maximum absolute column sum of $\mathbf{A}$.
- The ∞-norm: $\|\mathbf{A}\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{i,j}|$, i.e. the maximum absolute row sum of $\mathbf{A}$.

We note that for symmetric matrices the 1-norm and ∞-norm are equal.

Other matrix norms exist that do not arise from vector norms and are defined explicitly for matrices. Examples of these are the Frobenius norm and the Ky Fan norm, both of which will be used in this thesis.

Definition 3.1.16. *The Frobenius norm of a matrix* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *is given by*

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \tag{3.17}$$

Definition 3.1.17. *The Ky Fan p-k norm of* $\mathbf{A} \in \mathbb{C}^{m\times n}$ *is defined as:*

$$\|\mathbf{A}\|_{p,k} = \left(\sum_{i=1}^{k} \gamma_i(\mathbf{A})^p\right)^{1/p}, \tag{3.18}$$

where $\gamma_i(\mathbf{A})$ *denotes the i-th largest singular value of* $\mathbf{A}$*,* $p \geq 1$ *and* $k \in \{1, \ldots, \min\{m, n\}\}$*.*

Definition 3.1.18. *For a square matrix* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *we define the condition number of* $\mathbf{A}$ *in the* α*-norm to be*

$$\kappa_\alpha(\mathbf{A}) = \|\mathbf{A}\|_\alpha \|\mathbf{A}^{-1}\|_\alpha. \tag{3.19}$$

By convention we take $\kappa_\alpha(\mathbf{A}) = \infty$ *for a singular matrix* $\mathbf{A}$*.*

Corollary 3.1.19. *Any condition number is bounded below by one [Golub and Van Loan, 1996].*

We can interpret the condition number as a measure of how sensitive solutions of a linear equation $\mathbf{Ax} = \mathbf{b}$ are to perturbations in the data $\mathbf{b}$. A 'well-conditioned problem' will result in small perturbations to the solution with small changes to $\mathbf{b}$, whereas for an 'ill-conditioned problem', small perturbations to $\mathbf{b}$ can result in large changes to the solution. Whether a problem is well-conditioned or ill-conditioned is partly dependent on the application - for some problems a condition number of 100 will be acceptable, whereas for other problems this will be very large. The condition number can also provide an indication of how many digits of accuracy will be lost during computations [Gill et al., 1986, Cheney, 2005]. Similarly the condition number is a measure of the amplification of errors when inverting a matrix [Golub and Van Loan, 1996]. We will discuss further interpretations of the condition number in the next section when we introduce the conjugate gradient method.

For the remainder of this work we will focus on the condition number in the 2-norm given by

$$\kappa(\mathbf{A}) \equiv \kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2. \tag{3.20}$$

Theorem 3.1.20. *If* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *is a symmetric positive definite matrix with eigenvalues defined as in Definition 3.0.1 we can write the condition number in the* $2-$*norm as*

$$\kappa(\mathbf{A}) = \frac{\lambda_1(\mathbf{A})}{\lambda_n(\mathbf{A})}. \tag{3.21}$$

Proof. See [Golub and Van Loan, 1996, Sec. 2.7.2]. □

Many of the matrices that we consider in this thesis are SPD and hence the characterisation of the condition number given by Theorem 3.1.20 will be used throughout.

3.2 The conjugate gradient method

Suppose we want to solve the problem

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{3.22}$$

for some symmetric, positive definite matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, data $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^n$ unknown. The conjugate gradient method is a Krylov subspace method that can be used to iteratively solve the system of linear equations given by the problem (3.22). It makes use of gradient information in order to take optimal steps towards the minimum of a quadratic function. At each stage, the algorithm finds the direction that is orthogonal with respect to $\mathbf{A}$ to previous search directions. This means that convergence occurs in a maximum of n iterations in exact arithmetic, where n is the dimension of the problem [Gill et al., 1986]. We note however, that for computational implementations search directions may not be perfectly conjugate, and therefore more than n iterations may be required to reach a desired tolerance [Gill et al., 1986]. Convergence speed is affected by the eigenvalue structure of $\mathbf{A}$. For applications, typical values of n are large (e.g. 10^9 for NWP [Carrassi et al., 2018]), meaning that the permitted number of iterations must be much smaller than n for a tractable problem.

We can consider the linearised variational objective function as a problem of the form (3.22), where the Hessian of the linearised objective function (2.6), (2.5) replaces $\mathbf{A}$ in (3.22). This formulation is derived explicitly in [Haben, 2011, Sec 3.2], and will be used for experiments studying convergence of a conjugate gradient method in Chapters 5, 6 and 7.

We now define the conjugate gradient method.

Definition 3.2.1 (Conjugate gradient method)**.** *[Trefethen and Bau, 1997] To apply the conjugate gradient method to the system* (3.22) *for a given tolerance, τ, we define the residual as $\mathbf{r}_k = \mathbf{A}\mathbf{x}_k - \mathbf{b}$, and the search direction as $\mathbf{p}_k$ for step k. For $k = 0$ let*

$\mathbf{x}_0 = 0$*,* $\mathbf{r}_0 = \mathbf{b}$ *and* $\mathbf{p}_0 = r_0$*. While* $\|\mathbf{r}_k\| > \tau$*:*

$$\begin{aligned}
\alpha_{k+1} &= \frac{\|\mathbf{r}_k\|_2^2}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k} \\
\mathbf{x}_{k+1} &= \mathbf{x}_k + \alpha_{k+1} \mathbf{p}_k \\
\mathbf{r}_{k+1} &= \mathbf{r}_k - \alpha_{k+1} \mathbf{A} \mathbf{p}_k \\
\beta_{k+1} &= \frac{\|\mathbf{r}_{k+1}\|_2^2}{\|\mathbf{r}_k\|_2^2} \\
p_{k+1} &= \beta_{k+1} \mathbf{p}_k + \mathbf{r}_{k+1} \\
k &= k+1.
\end{aligned} \tag{3.23}$$

The procedure given by Definition 3.2.1 enforces $\mathbf{A}-$conjugacy between search directions $\mathbf{p}_k$, which ensures rapid convergence of this method compared to other gradient descent methods. We can provide bounds on the convergence of the conjugate gradient method in terms of the condition number of $\mathbf{A}$.

Theorem 3.2.2. *Let* $\mathbf{A}$ *be a SPD matrix with condition number* κ*. Let* $\mathbf{e}_0 = \mathbf{A}\mathbf{x}_0 - \mathbf{b}$ *denote the initial error, and* $\mathbf{e}_k$ *denote the error at iteration* k *of the conjugate gradient method given by Definition 3.2.1. Then the* $\mathbf{A}$*-norms of the error satisfy*

$$\frac{\|\mathbf{e}_k\|_{\mathbf{A}}}{\|\mathbf{e}_0\|_{\mathbf{A}}} \leq 2 \Big/ \left[\left(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1} \right)^k + \left(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1} \right)^{-k} \right] \leq 2 \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \right)^k. \tag{3.24}$$

Proof. [Trefethen and Bau, 1997, Theorem 38.5] □

The bounds given by Theorem 3.2.2 are not tight. In particular, there are well known cases where convergence is much faster than the upper bound given by the condition number. These depend on other properties of eigenvalues that are not 'measured' by condition number e.g. repeated eigenvalues and eigenvalues that are closely grouped together.

We can obtain tighter bounds on the convergence of the conjugate gradient method. However, these typically require knowledge of the entire spectrum of $\mathbf{A}$ and are often harder to compute. An example of a sharp bound is given by the following theorem.

Theorem 3.2.3. *Let* P_k *be the set of polynomials of degree* k *with* $p(0) = 1$*. If the conjugate gradient algorithm given by Definition 3.2.1 with* $\mathbf{e}_0 = \mathbf{A}\mathbf{x}_0 - \mathbf{b}$ *has not converged before step* k*, then the problem*

$$\min_{p_k \in P_k} ||p_k(\mathbf{A})\mathbf{e}_0|| \tag{3.25}$$

has a unique solution, and the iterate x_k has error $\mathbf{e}_k = p_k(\mathbf{A})\mathbf{e}_0$ for this same polynomial p_k. Consequently we have

$$\frac{\|\mathbf{e}_k\|_{\mathbf{A}}}{\|\mathbf{e}_0\|_{\mathbf{A}}} = \inf_{p \in P_k} \frac{\|p(\mathbf{A})\mathbf{e}_0\|_{\mathbf{A}}}{\|\mathbf{e}_0\|_{\mathbf{A}}} \leq \inf_{p \in P_k} \max_{\lambda \in \Lambda(\mathbf{A})} |p(\lambda)|, \tag{3.26}$$

where $\Lambda(\mathbf{A})$ denotes the spectrum of $\mathbf{A}$.

Proof. Trefethen and Bau [1997, Theorem 38.3] □

From this result, we see that in the case of clustered eigenvalues, we obtain much faster convergence that might be predicted by the condition number alone given by the result of Theorem 3.2.2. The intuition behind this result is that for a matrix with eigenvalues that occur in r clusters we can construct a polynomial $\mathbf{P}_{r-1}$ such that $(1 + \lambda\mathbf{P}_{r-1}(\lambda))$ has zeroes inside each cluster. For repeated eigenvalues the polynomial will vanish, but for clustered eigenvalues the value of the polynomial will be small and hence the upper bound of (3.26) will be small for values of $k \geq r - 1$ [Axelsson, 1996].

Theorem 3.2.4. *If $\mathbf{A}$ has only n distinct eigenvalues, then the CG iteration converges in at most n steps in exact arithmetic.*

Proof. Gill et al. [1986, Theorem 38.4] □

This is a special case of Theorem 3.2.3.

3.3 Results on eigenvalues

In Chapters 5 and 6 we will develop bounds on the condition number of the Hessian of the variational cost function, separating the contribution of each of the constituent matrices. In order to simplify these bounds, we will exploit properties of the covariance matrices, $\mathbf{B}$ and $\mathbf{R}$. In Section 3.4 we will show that correlation matrices are symmetric positive semi-definite. For the theoretical and numerical results that follow we will restrict our attention to strictly positive definite covariance matrices. This allows us to use the formulation of the condition number given by Theorem 3.1.20, in terms of the eigenvalues of the matrix. It is therefore of interest to consider results which allow us to write the eigenvalues of products and sums of matrices in terms of the eigenvalues of the individual matrices.

We begin by presenting a result on the eigenvalues of the inverse of a matrix, $\mathbf{A}$. This result will be used to express the eigenvalues of inverse error covariance matrices in terms of the eigenvalues of the original error covariance matrices.

Theorem 3.3.1 (Eigenvalues of the matrix inverse)**.** *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *have eigenvalues* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$*. If* $\mathbf{A}$ *is non-singular then the eigenvalues of* $\mathbf{A}^{-1}$ *are given by* $1/\lambda_i$ *for* $i = 1, 2, 3, \ldots, n$*. In particular*

$$\lambda_n(\mathbf{A}^{-1}) = \frac{1}{\lambda_1(\mathbf{A})} \tag{3.27}$$

$$\lambda_1(\mathbf{A}^{-1}) = \frac{1}{\lambda_n(\mathbf{A})}. \tag{3.28}$$

Proof. [Bernstein, 2009, Fact 5.11.14] □

The next result shows that if two matrices are conformable, exchanging the order of multiplication preserves the values of the non-zero eigenvalues.

Theorem 3.3.2 (Eigenvalues of product)**.** *If* $\mathbf{B} \in \mathbb{R}^{m\times n}$ *and* $\mathbf{A} \in \mathbb{R}^{n\times m}$ *then* $\mathbf{AB}$ *and* $\mathbf{BA}$ *have the same non-zero eigenvalues.*

Proof. See [Harville, 1997, Theorem 21.10.1]. □

We now present bounds on the eigenvalues of the sum of two symmetric matrices in terms of the eigenvalues of the individual matrices. This result allows us to separate the contribution of the observation and background terms when bounding the condition numbers of (2.6) and (2.19).

Theorem 3.3.3. *Consider two symmetric matrices* $\mathbf{S}_1$*,* $\mathbf{S}_2 \in \mathbb{R}^{N\times N}$*. The* k^{th} *eigenvalue of the matrix sum* $\mathbf{S}_1 + \mathbf{S}_2$ *satisfies the following:*

$$\lambda_k(\mathbf{S}_1) + \lambda_N(\mathbf{S}_2) \leq \lambda_k(\mathbf{S}_1 + \mathbf{S}_2) \leq \lambda_k(\mathbf{S}_1) + \lambda_1(\mathbf{S}_2). \tag{3.29}$$

Proof. See [Wilkinson, 1965, Ch. 2 Thm 44]. □

We note that in the case that $\mathbf{S}_2$ is rank-deficient, as is the case for the second terms of (2.6) and (2.19), the lower bound of (3.29) will simplify to $\lambda_k(\mathbf{S}_1)$.

The following theorem bounds the eigenvalues of a matrix product above in terms of the eigenvalues of the constituent matrices. This result will be used to separate the contribution of the observation error covariance matrix from the observation operator in the second term of (2.6).

Theorem 3.3.4. *If* $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{N\times N}$ *are positive semi-definite Hermitian matrices, then*

$$\prod_{i=1}^{k} \lambda_i(\mathbf{FG}) \leq \prod_{i=1}^{k} \lambda_i(\mathbf{F})\lambda_i(\mathbf{G}), \quad k = 1, \ldots, N-1. \tag{3.30}$$

Proof. See [Marshall et al., 2011, Sec. 9 H.1.a.]. □

Two similar results allows us to bound the eigenvalues of a matrix product below in terms of products of the eigenvalues of the individual matrices.

Theorem 3.3.5. *If* $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{N\times N}$ *are positive semi-definite Hermitian and* $1 \leq i_1 < \cdots < i_k \leq N$*, then*

$$\prod_{t=1}^{k} \lambda_t(\mathbf{FG}) \geq \prod_{t=1}^{k} \lambda_{i_t}(\mathbf{F})\lambda_{N-i_t+1}(\mathbf{G}), \tag{3.31}$$

with equality for $k = N$*.*

Proof. See Wang and Zhang [1992, Theorem 2]. □

Theorem 3.3.6. *If* $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{N\times N}$ *are positive semi-definite Hermitian and* $1 \leq i_1 < \cdots < i_k \leq N$*, then*

$$\sum_{t=1}^{k} \lambda_{i_t}(\mathbf{FG}) \geq \sum_{t=1}^{k} \lambda_{i_t}(\mathbf{F})\lambda_{N-t+1}(\mathbf{G}). \tag{3.32}$$

Proof. See Wang and Zhang [1992, Theorem 4]. □

Finally we introduce the Rayleigh quotient, which can be used to estimate the eigenvalues of any symmetric matrix.

Definition 3.3.7. *For a symmetric matrix* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *the Rayleigh quotient is given by*

$$R_{\mathbf{S}}(\mathbf{x}) = \frac{\mathbf{x}^\dagger \mathbf{A}\mathbf{x}}{\mathbf{x}^\dagger \mathbf{x}}, \tag{3.33}$$

for $\mathbf{x} \in \mathbb{C}^n$*, where* $\mathbf{x}^\dagger$ *denotes the conjugate transpose of* $\mathbf{x}$*.*

An important property of the Rayleigh quotient (3.33) is the fact that it is bounded by the eigenvalues of $\mathbf{A}$.

Theorem 3.3.8. *Let* $\mathbf{A} \in \mathbb{R}^{N\times N}$ *be a symmetric matrix. For any value of* $\mathbf{x} \in \mathbb{C}^N$*, the Rayleigh quotient is bounded by*

$$\lambda_N(\mathbf{A}) \leq R_{\mathbf{S}}(\mathbf{x}) \leq \lambda_1(\mathbf{A}). \tag{3.34}$$

Proof. [Süli and Mayer, 2003, Sec 5.9] □

We can see that the lower and upper bounds are achieved when $\mathbf{x}$ is chosen in (3.33) such that it is the eigenvector corresponding to the smallest and largest eigenvalue of $\mathbf{A}$ respectively. This property will be used to develop tighter bounds for specific observation networks in Chapter 5.

3.4 Matrix structures

The theoretical results presented in later chapters will exploit the special properties of some matrix structures that are of particular interest for the variational data assimilation problem. We now introduce some of these structures and their properties. We begin by formally defining covariance matrices, which are the backbone of this thesis. For a given random vector the covariance matrix is defined probabilistically in Schott [2016, Sec 1.13]. The covariance matrix contains information about variances and correlations between different random variables.

Definition 3.4.1. *A covariance matrix, $\mathbf{R} \in \mathbb{R}^{p\times p}$, is a symmetric positive semi-definite matrix.*

Although a covariance matrix can be semi-definite, for practical applications we restrict our attention to strictly positive definite matrices. In the variational data assimilation problems (2.5) and (2.7), inverse error covariance matrices are used as weighting matrices. This means that the matrix must be strictly positive definite in order for its inverse to be well-defined.

We often wish to consider covariance information in terms of variances and correlations separately. These can be calculated from the original covariance matrix via the following formulae.

Definition 3.4.2. *Given a covariance matrix, the entries of the corresponding matrix of standard deviations are given by*

$$\Sigma(i,i) = \sqrt{(\mathbf{R}(i,i))}. \tag{3.35}$$

Variances are given by the square of standard deviations.

This definition implicitly requires that variances are non-negative.

Definition 3.4.3. *Given a covariance matrix, the entries of the corresponding correlation matrix are given by*

$$\mathbf{C}(i,j)\frac{\mathbf{R}(i,j)}{=}\sqrt{\mathbf{R}(i,i)}\sqrt{\mathbf{R}(j,j)}. \tag{3.36}$$

By this definition the diagonal entries of any correlation matrix must be units. This means that any symmetric, positive semi-definite matrix with units on the diagonal is a correlation matrix [Higham, 2002].

In the case that correlations are homogeneous and isotropic (i.e. only the distance between two points determines the correlation between them) circulant matrices arise naturally. These have a special structure where the matrix is fully determined by its first row. Each subsequent row is a cyclic permutation of the first row. Circulant matrices arise as correlation matrices for spatial statistics on an equally spaced periodic domain, and will be used in Chapter 5 to construct the covariance matrices for the numerical experiments.

Definition 3.4.4 (Davis [1979])**.** *A circulant matrix* $\mathbf{D} \in \mathbb{R}^{N\times N}$ *is a matrix of the form*

$$\mathbf{D} = \begin{pmatrix} d_0 & d_1 & d_2 & \cdots & d_{N-2} & d_{N-1} \\ d_{N-1} & d_0 & d_1 & \cdots & d_{N-3} & d_{N-2} \\ d_{N-2} & d_{N-1} & d_0 & \cdots & d_{N-4} & d_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ d_2 & d_3 & d_4 & \cdots & d_0 & d_1 \\ d_1 & d_2 & d_3 & \cdots & d_{N-1} & d_0 \end{pmatrix}.$$

Circulant matrices have various properties which are useful for applications and numerical experiments. Firstly, both eigenvalues and eigenvectors can be calculated via a discrete Fourier transform [Gray, 2006]. In practice, this means we can calculate the eigenvalues of $\mathbf{D}$ directly via the following formula using a fast Fourier transform.

Theorem 3.4.5. *The eigenvalues of a circulant matrix* $\mathbf{D} \in \mathbb{R}^{N\times N}$*, as given by Definition 3.4.4, are given by*

$$\gamma_m = \sum_{k=0}^{N-1} d_k \omega^{mk}, \tag{3.37}$$

with corresponding eigenvectors

$$\mathbf{v}_m = \frac{1}{\sqrt{N}}(1, \omega^m, \cdots, \omega^{m(N-1)}), \tag{3.38}$$

where $\omega = e^{-2\pi i/N}$ *is an* $N-$*th root of unity.*

Proof. See Gray [2006] for full derivation. □

To avoid confusion, for the remainder of this thesis the eigenvalues of a circulant matrix calculated using (3.37) will be denoted by γ_j rather than λ_j. This distinction is made as the ordering of eigenvalues given by (3.37) is given by wavenumber rather than size. We can see from (3.38) that the eigenvectors only depend on N, the dimension of the circulant matrix. Therefore any two circulant matrices of the same dimension will have the same set of eigenvectors.

Theorem 3.4.6. *The transpose, inverses, products and sums of circulant matrices are themselves circulant.*

Proof. Davis [1979, Thm 3.1.1, Thm 3.2.3, Thm 3.2.4] □

Circulant correlation matrices can be constructed using known correlation functions. One such function that will be used in this thesis is the second-order auto-regressive (SOAR) correlation function [Daley, 1991]. This is a homogeneous and isotropic function and naturally extends to a circulant form when we have equally spaced observations on a periodic domain, such as a latitude circle on the Earth. The long tails of this correlation function are suitable for estimating horizontal spatial correlations, and make SOAR a popular choice at NWP centres [Thiebaux, 1976, Stewart et al., 2013, Simonin et al., 2014, Waller et al., 2016c, Fowler et al., 2018, Tabeart et al., 2018].

We begin by defining the SOAR correlation function for two points on a real line separated by a distance, r.

Definition 3.4.7. *The second-order auto-regressive correlation function is given by*

$$\rho_S(r) = \left(1 + \frac{|r|}{L}\right) \exp\left(-\frac{|r|}{L}\right), \tag{3.39}$$

where $r \in \mathbb{R}$ is the distance between two points, and $L > 0$ is the correlation lengthscale.

This correlation function on the real line can then be transformed into an error correlation matrix on the circle. In particular we consider a 1D model with variables given by equally spaced gridpoints on the circle with radius $r = a$. In order to obtain a valid correlation model on the circle, we follow the procedure described in Haben [2011] and Waller et al. [2016b]. This substitutes a chordal distance for a 'great circle distance'; as discussed in Gaspari and Cohn [1999] and Jeong and Jun [2015]. This substitution is necessary to ensure the resulting covariance matrix is positive definite.

Definition 3.4.8. *The SOAR error correlation matrix on the finite domain is given by*

$$\mathbf{D}(i,j) = \left(1 + \frac{\left|2a\sin\left(\frac{\theta_{i,j}}{2}\right)\right|}{L}\right)\exp\left(\frac{-\left|2a\sin\left(\frac{\theta_{i,j}}{2}\right)\right|}{L}\right), \tag{3.40}$$

where $L > 0$ is the correlation lengthscale, $\theta_{i,j}$ denotes the angle between grid points i and j, and a is the radius of the domain. The chordal distance between adjacent grid points is given by

$$\Delta x = 2a\sin\left(\frac{\theta}{2}\right) = 2a\sin\left(\frac{\pi}{N}\right), \tag{3.41}$$

where N is the number of gridpoints and $\theta = \frac{2\pi}{N}$ is the angle between adjacent gridpoints.

The same procedure can be used to transform other correlation functions on the straight line into valid covariance models on circular domains.

3.5 Summary

In this chapter we introduced the concept of conditioning and the characterisation of the condition number in the case of symmetric positive definite matrices. We also presented a variety of results on the eigenvalues of products and sums of matrices. These results permit separation in terms of the eigenvalues of the constituent matrices, and will be used in the Chapters 5 and 6 to understand how each constituent matrix influences the conditioning of the Hessians (2.6) and (2.19). We introduced the conjugate gradient method, and showed how the condition number can be used to develop simple bounds on convergence. Finally we described key matrix structures that will be exploited in the numerical experiments that follow (see Chapters 5, 6 and 7). In the next chapter, we describe how the conditioning of the Hessian of the objective function can be used as a proxy to study convergence of the minimisation problem.

Chapter 4

The study of conditioning and introduction of correlated observation error

In this chapter we discuss how the conditioning of the Hessian of the data assimilation objective function can be studied as a proxy for convergence of its minimisation (Section 4.1). In Section 4.2 we present previous work on the conditioning of the Hessian in the case of uncorrelated observation error covariance matrices and describe how these results will be extended to consider the case of correlated observation error covariance matrices in subsequent chapters. In Chapter 2 we motivated the importance of including correlated observation error covariance matrices at NWP centres. In Section 4.3 we discuss computational issues related to convergence of the minimisation problem that have occurred when correlated observation errors have been introduced at operational NWP centres.

4.1 Using conditioning as a proxy for convergence

In operational systems, it is vital to ensure that convergence of the minimisation of the variational data assimilation problem is fast enough to ensure timely forecasts. At many meteorological centres, new forecasts are produced multiple times per day [Rawlins et al., 2007], with a very small proportion of the computing time being allocated to the data assimilation procedure. For example the first implementation of 4D-Var at the Met Office required 12 minutes to complete the global data assimilation routine [Rawlins et al., 2007], and a more recent implementation of 3D-Var with FGAT took around 4.5 minutes to complete for the UKV limited-area model [Simonin et al., 2019]. In applications, timeliness of forecasts is paramount, meaning that although improvements to data assimilation algorithms and the associated improvements to initial conditions are desirable, they are not prioritised over computational efficiency [Isaksen, 2012]. It is therefore of interest to study how changes to data assimilation algorithms are likely to alter convergence of the minimisation procedure, to ensure that proposed improvements will not result in slower convergence.

Studying convergence directly is difficult, because it is expensive to run a full data assimilation procedure multiple times. Additionally, complicated systems can make it difficult to isolate the impacts of changing a single component. In order to get a better understanding of the likely effects of broad changes to an algorithm, it therefore makes sense to consider simplified systems, and to develop theory that can be applied or extended to a specific system of interest. This motivates our use of the condition number of the Hessian of the variational data assimilation objective function as a proxy for convergence of the minimisation problem. In Chapter 3 we defined the condition number, and described how it can be used to bound the convergence of a conjugate gradient problem, as well as to understand how sensitive a system of interest is to perturbations of the initial condition. We also discussed the limitations of these bounds on convergence, particularly in the case of repeated or clustered eigenvalues. However, if the eigenvalues of the Hessian are available then the condition number is cheap to compute. Studying how the conditioning of the Hessian changes with alterations to the data assimilation system can provide insight into how convergence is likely to be affected.

4.2 Previous bounds on the condition number of the Hessian for variational data assimilation problems

Haben [2011], Haben et al. [2011a,b] studied the conditioning of the Hessian of the variational objective function as proxy for convergence of the minimisation problem. Haben [2011] developed bounds on the condition number of the Hessian of the 3D-Var and 4D-Var objective functions which separated the contribution of the background and observation error covariance matrices. Separate bounds were developed for the unpreconditioned and preconditioned problems, but similar techniques were used for both cases. General bounds which apply to any choice of background and observation error covariance matrix, and any linear observation operator were given, but not studied numerically. Tighter bounds were established and studied numerically for uncorrelated observation error covariances and direct observations.

We now present a bound on the condition number of the unpreconditioned 3D-Var Hessian (2.6) that was given in Haben et al. [2011a], Haben [2011]. This bound applies to the case of direct observations and uncorrelated observation errors, so will not apply to a general data assimilation problem. Further bounds on the condition number of the Hessian proven by Haben [2011] are presented in Chapters 5 and 6 for the unpreconditioned 3D-Var problem and the general preconditioned 3D-Var problem respectively.

Theorem 4.2.1. *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*

$$\left(\frac{1 + \frac{p}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_{\min}(\mathbf{C})}{1 + \frac{p}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_{\max}(\mathbf{C})}\right)\kappa(\mathbf{C}) \leq \kappa(\mathbf{S}) \leq \left(1 + \left(\frac{\sigma_b^2}{\sigma_o^2}\right)\lambda_{\min}(\mathbf{C})\right)\kappa(\mathbf{C}) \tag{4.1}$$

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

Proof. [Haben, 2011, Theorem 6.1.2] □

The behaviour of (4.1) was studied numerically in Haben [2011]. Experimental study of these bounds provided insight into how convergence is likely to change with alterations to the data assimilation system. In particular the impact of changing background error lengthscales, and the number and distribution of observations on (4.1) were considered. The key findings of Haben [2011] were that in the unpreconditioned data assimilation case:

- Increasing the lengthscale of the background error correlations increases the condition number of the Hessian.

Numerical experiments also revealed that in the preconditioned setting:

- Increasing the accuracy of observations increases the condition number of the Hessian.

- Increasing the number of observations increases the condition number of the Hessian.
- Increasing the density of observations increases the condition number of the Hessian.

Additionally, Haben [2011] studied the impact of changing the data assimilation system on convergence of the conjugate gradient method. Changes which increased the condition number of the Hessian also resulted in worse convergence of the conjugate gradient method. Therefore, in this specific setting studying the condition number of the Hessian is a good proxy for the qualitative changes to convergence of the associated minimisation with changes to the data assimilation system.

However, there are limitations to the results presented in Haben [2011], Haben et al. [2011a,b]. The bounds that were studied numerically exploited the particular structures of the specific experimental framework, meaning that they are no longer general. Some of these assumptions are not realistic for applications. In particular:

- Observation errors were assumed to be uncorrelated. We discussed why the use of correlated observation error covariance matrices is important for NWP applications in Section 2.3.
- The observation operator was restricted to direct observations only. Indirect observations occur frequently: observations of a meteorological variable of interest can be made at a different location to the state variables, and for satellite instruments there is a nonlinear relationship between measurements of brightness temperature and meteorological variables in state space.
- The background error covariance matrix was assumed to be circulant. This assumption holds for uniform spatial correlations on a uniform grid, which is often the case for horizontal background error covariances, but will not be true in the case of non uniform grids, such as for vertical correlations.

The numerical experiments presented in Chapters 5 and 6 will make use of circulant background error covariances matrices. Observation error covariance matrices will be assumed to be correlated, and our experiments will consider a range of observation operators which correspond to direct and indirect observations.

The numerical and theoretical investigation of Haben [2011] also focussed on spatially correlated background errors. While background errors are often spatially correlated in practice, observation error covariances can have spatial, temporal and interchannel correlations depending on the instrument. In this thesis we will consider both spatial and interchannel observation error covariances theoretically and numerically. We will develop explicit bounds for the case of correlated observation error covariance matrices, using similar techniques to those introduced in Haben [2011] to separate the contribution of observation and background error terms in (2.6) and (2.19) (see Chapters 5 and 6 respectively).

4.3 Convergence issues when using correlated OEC matrices at NWP centres

In Section 2.4 we discussed some of the known computational issues with estimates obtained using the DBCP diagnostic. Many of these issues are due to the fact that as a sampling method, it can fail to sample the full eigenspace of the true correlation matrix. Errors due to undersampling can manifest as very small eigenvalues, resulting in estimated covariance matrices that are numerically close to singular and hence very ill-conditioned [Higham et al., 2016]. Additionally, sampling methods recover matrices that do not satisfy the properties of covariance matrices, i.e. are not symmetric or positive definite [Ledoit and Wolf, 2004]. This has been show to occur for the DBCP diagnostic [Stewart et al., 2008a], and makes it difficult to use the iterative method in some situations [Ménard, 2016]. Thus, in order to use these sample-based estimates in the data assimilation system, they must be transformed by symmetrising and ensuring that all eigenvalues are strictly positive.

Typically users symmetrise the estimated observation error covariance matrix via $(\mathbf{R}+\mathbf{R}^T)/2$. To solve the problem of negative eigenvalues, Weston [2011] proposed making all small negative or zero eigenvalues small and positive. Other studies had similar problems, with Gauthier et al. [2018] finding negative variance values, and Bennitt et al. [2017] estimating correlation values larger than one. In the case of Gauthier et al. [2018] this motivated the use of the iterative diagnostic rather than performing a single step of the DBCP diagnostic. Bennitt et al. [2017] proposed comparing the results of the DBCP estimate with those from alternative diagnostic techniques, or using estimates to provide insight into error statistics rather than directly using the resulting covariance matrices in data assimilation systems.

Even matrices which are symmetric and positive definite, and hence are valid covariance matrices, can cause computational difficulties. Weston [2011] proposed two methods of 'reconditioning' to increase small eigenvalues, and mitigate the slow convergence associated with ill-conditioned sample covariance matrices. This reduces the condition number of an OEC matrix, and improves the convergence of a data assimilation procedure. Two potential methods of reconditioning were considered [Weston, 2011, Weston et al., 2014]. The first of these methods, which will be referred to as the ridge regression method in this thesis, applies additive inflation to the diagonal of the OEC matrix. The second method, which will be referred to as the minimum eigenvalue method, alters eigenvalues that are below a given threshold. Both of these methods will be studied theoretically for the first time in Chapter 7. Other methods of reconditioning that will not be considered in this work include thresholding [Bickel and Levina, 2008], localisation [Horn, 1991, Ménétrier et al., 2015, Smith et al., 2018] linear shrinkage [Ledoit and Wolf, 2004] and regularisation methods such as the Lasso penalty technique [Pourahmadi, 2013]. Many other centres have needed to adapt the results of the DBCP diagnostic prior to their use in operational data assimilation systems, (e.g. Bormann et al. [2015, 2016], Campbell et al. [2017]). This will be discussed further in Chapter 7.

To date, many of the studies of correlated OEC matrices in NWP systems have been empirical. This especially applies to the adjustment of estimated covariance matrices that is required when using the DBCP diagnostic [Weston, 2011, Weston et al., 2014, Bormann et al., 2015]. Often reconditioning methods have only been compared in terms of changes to convergence of the data assimilation algorithm, rather than the theoretical impact on the covariance matrices themselves. In Chapter 7 we will theoretically prove how standard deviations and correlations are changed by two commonly used methods of reconditioning, as well as comparing them against variance inflation, which is commonly used to account for missing correlation information.

4.4 Summary

In this chapter we described how the condition number of the Hessian can be used to study the convergence of a data assimilation problem. We discussed previous work where bounds on the condition number of the Hessian in terms of its constituent components were used to make qualitative conclusions about the effect of altering the data assimilation system. However many of these results considered the case of uncorrelated OEC matrices. In recent years the use of correlated observation errors at meteorological centres has expanded, with the popularisation of the DBCP diagnostic.

However, this diagnostic has theoretical and practical limitations. We wish to understand how the use of correlated OEC matrices alters the conclusions of Haben [2011], and whether theory can provide insights into methods that can be used to include correlated OEC matrices in data assimilation systems without negatively impacting convergence speed. In the next chapter we apply similar methods to those of Haben [2011] to develop new bounds on the condition number of the Hessian of the unpreconditioned variational data assimilation objective function in the case of correlated observation error covariance matrices.

Chapter 5

The conditioning of least squares problems in variational data assimilation

In this chapter we answer RQ 1 from Chapter 1 and consider how the introduction of correlated observation error affects the conditioning of the Hessian of the unpreconditioned variational data assimilation problem. We develop theoretical bounds on the condition number of the Hessian to understand the impact of changing the observation error covariance matrix. We wish to know

- How are these bounds affected by changes to the observation error covariance matrix? ?
- How tight are the new bounds for an idealised numerical framework?
- How well does the behaviour of the condition number of the Hessian represent convergence of the conjugate gradient method numerically?

The remainder of this chapter, excluding the chapter summary (Section 5.10) is strongly based on the paper: Tabeart J. M., Dance S. L., Haben S. A., Lawless A. S., Nichols N. K., Waller J. A. The conditioning of least-squares problems in variational data assimilation. Numerical Linear Algebra with Applications. 2018;25:e2165. https://doi.org/10.1002/nla.2165.

5.1 Abstract

In variational data assimilation a least squares objective function is minimised to obtain the most likely state of a dynamical system. This objective function combines

observation and prior (or background) data weighted by their respective error statistics. In numerical weather prediction (NWP), data assimilation is used to estimate the current atmospheric state, which then serves as an initial condition for a forecast. New developments in the treatment of observation uncertainties have recently been shown to cause convergence problems for this least squares minimization. This is important for operational NWP centres due to the time constraints of producing regular forecasts. The condition number of the Hessian of the objective function can be used as a proxy to investigate the speed of convergence of the least squares minimisation. In this chapter we develop novel theoretical bounds on the condition number of the Hessian. These new bounds depend on the minimum eigenvalue of the observation error covariance matrix, and the ratio of background error variance to observation error variance. Numerical tests in a linear setting show that the location of observation measurements has an important effect on the condition number of the Hessian. We identify that the conditioning of the problem is related to the complex interactions between observation error covariance and background error covariance matrices. Increased understanding of the role of each constituent matrix in the conditioning of the Hessian will prove useful for informing the choice of correlated observation error covariance matrix and observation location, particularly for practical applications.

5.2 Introduction

Data assimilation combines output from a numerical model of a dynamical system, the background or prior, with observations of the system to yield an accurate description of the current dynamical state (analysis). Contributions from observations and the background are weighted according to their relative uncertainty via error covariance matrices, meaning that assessing and quantifying observation error is crucial in order to obtain an accurate analysis sufficiently quickly [Buehner, 2010, Janjić et al., 2018]. One of the most well known applications of data assimilation is to numerical weather prediction (NWP), where observations of the atmosphere and ocean are combined with a prior model state of the atmosphere in order to produce the initial conditions for a weather forecast. Until recently, diagonal observation error covariance matrices have been used operationally at all major NWP centres [Weston, 2011], a choice that is only valid in the case that observation errors are uncorrelated. It has been shown that implementing diagonal error covariance matrices inappropriately, i.e. when error correlations are non-zero, may lead to suboptimal results [Rainwater et al., 2015, Stewart et al., 2008b, Stewart, 2010, Stewart et al., 2013, Waller et al., 2014a]. However, using diagnosed full observation error covariance matrices directly in the assimilation has been shown to cause problems with the speed of convergence of the assimilation scheme [Weston et al., 2014].

Variational assimilation, a popular data assimilation method [Haben et al., 2011b, Rawlins et al., 2007, Clayton et al., 2013], finds the analysis by minimising a nonlinear least squares objective function. This objective function, which is dependent on both observations and the background field, is minimised by an iterative method, such as the Gauss-Newton method [Lawless et al., 2005a, Gratton et al., 2007]. This consists of an outer loop that solves the full non-linear problem, and an inner loop that solves the linearised problem, often via a conjugate gradient method [Lawless et al., 2005b]. The conditioning of the Hessian matrix of the objective function provides a bound on the rate of convergence of the conjugate gradient minimization [Haben, 2011, Gill et al., 1986, Golub and Van Loan, 1996]. Hence it can be used as a rough estimate for the number of iterations needed to solve the inner loop problem. We note however, that this worst case bound on convergence can be improved on significantly in the case of clustered eigenvalues [Gill et al., 1986, Nocedal, 2006]. The magnitude of the condition number also provides an indication of the sensitivity of the system to perturbations in the data [Haben et al., 2011b]. Speed of convergence is critical in practice due to the need to provide timely forecasts. In this work we investigate how introducing correlated observation errors affects the condition number of the Hessian and examine the associated speed of convergence of a conjugate gradient method.

Correlated observation error statistics have been diagnosed for certain observation types e.g. Waller et al. [2016c], Bormann et al. [2016], Campbell et al. [2017], Waller et al. [2016a], Bormann et al. [2003, 2011], Stewart et al. [2014], Cordoba et al. [2017], although there are problems associated with their use. In particular, the methods used to diagnose observation error covariance matrices are imperfect, and the quality of these estimates is unclear. Due to unknown observation error statistics and in order to reduce the computational cost of operational assimilation, in practice the majority of observation errors are assumed uncorrelated. However, empirical evidence from simple model experiments indicate that even approximate correlation structures give significant benefit in terms of analysis accuracy [Stewart et al., 2013, Healy and White, 2005]. Similar conclusions can be drawn for practical implementations [Rainwater et al., 2015].

In Stewart [2010] and Stewart et al. [2014] it was shown that there were problems with the use of diagonal observation error covariance matrices in the variational data assimilation for certain instruments. Motivated by this work, in 2011 the UK Met Office first trialled the use of correlated observation errors in their operational system [Weston, 2011]. However, there were problems with the convergence of the minimisation algorithm which necessitated 'reconditioning' of observation error covariance matrices (by altering their eigenvalues), prior to their use in the system. In Weston [2011] and Weston et al. [2014] it was suggested that slow convergence was caused by the very small minimum eigenvalues of the diagnosed observation error covariance matrix. This work provides motivation to investigate further the role of the minimum eigenvalue of the observation error covariance matrix on the conditioning of the variational data assimilation problem; in turn, developing this crucial understanding will permit optimal use of correlated observation errors in data assimilation systems.

Even in the case of uncorrelated observation errors, the minimisation problem for any large system is very ill-conditioned. Preconditioning, where the original problem is transformed into an equivalent but less ill-conditioned problem, is used operationally to mitigate against slow convergence of the minimisation [Brown et al., 2016]. In data assimilation the most common method of preconditioning is the Control Variable Transform (CVT) [Haben, 2011, Bannister, 2008], where the preconditioner is based on the background error covariance matrix. The optimal choice of preconditioning depends on the formulation of the data assimilation problem [Dollar et al., 2010], and practical constraints may require the use of a less computationally intensive preconditioner [Pestana and Wathen, 2015]. In this work an unpreconditioned framework will be used, as it is unknown whether the introduction of correlated observation errors will alter the optimal choice of preconditioner. This framework also has practical relevance, as the UK Met Office uses an unpreconditioned 1D-Var routine, where each observation is assimilated individually, for quality control purposes. Hence, the bounds and conclusions presented here will apply directly to that case.

In this article we develop new theory for bounding the condition number of the Hessian of the least squares objective function. This theory applies to both uncorrelated and correlated choices of observation error. We investigate the impact of introducing these correlations via small-scale numerical tests which illustrate the influence of observation correlations associated with a physical lengthscale. We begin in Section 5.3 by defining notation common to data assimilation and the condition number. We explain why the conditioning of the system and the rate of convergence of the minimisation are linked and present results from linear algebra that will be used to construct the bounds discussed in Section 5.4. Three new sets of bounds will be introduced in Section 5.4; these will have a varying number of additional constraints on the constituent matrices. Bounds which separate the contribution of each of the constituent terms have been developed for both general matrices and matrices with additional assumptions on observation location and observation error correlations. In Section 5.5 we discuss our numerical framework for the experiments of Section 5.6. The results of these numerical tests support the theoretical conclusions presented in Section 5.4. In particular we see that the minimum eigenvalue of the observation error covariance matrix and the ratio of background variance to observation variance are important terms for controlling the conditioning of the variational problem for both the bounds in Section 5.4 and the numerical results from Section 5.6. We conclude in Section 5.7 that even in a simple linear setting, the choice of observation operator has a significant effect on the conditioning. The theoretical conclusions indicate how correlated error statistics in the observation and background can be expected to interact, as well as highlighting areas where reconditioning and similar techniques could be used to reduce the increased computational cost associated with using correlated observation errors operationally. Although the primary motivation for the investigation of the impact of correlated observation errors arises from their application in meteorology, the theory and conclusions presented here are very general and apply to any other application of variational data assimilation such as neuroscience [Nakamura and Potthast, 2015, Schiff, 2011] and ecology [Pinnington et al., 2016, 2017].

5.3 Variational assimilation and Condition number

5.3.1 Notation

In data assimilation, information from observations, $\mathbf{y} \in \mathbb{R}^p$, is combined with information from a background, or 'prior', field, $\mathbf{x}_b \in \mathbb{R}^N$. The analysis, $\mathbf{x}_a \in \mathbb{R}^N$, or posterior, is found by weighting each of the two components using their respective error statistics. It is assumed that observation errors and background errors are unbiased and mutually uncorrelated. The background and observation error covariance matrices are denoted by the symmetric positive semi-definite matrices $\mathbf{B} \in \mathbb{R}^{N\times N}$ and $\mathbf{R} \in \mathbb{R}^{p\times p}$ respectively (although in practice we assume $\mathbf{B}$ and $\mathbf{R}$ are positive definite matrices). Usually there are far fewer observations than state variables, i.e. $p \ll N$. Observation and background information may describe different variables or be situated at different locations in space. The observation operator $h : \mathbb{R}^N \to \mathbb{R}^p$, which may be nonlinear, is used to map from state space to observation space to allow comparison of observations with the background; in particular $\mathbf{y}$ will be compared to $h[\mathbf{x}]$.

For variational assimilation methods, the analysis is found by minimising an objective function. In this work we focus on 3D-Var, a particular variational assimilation method, which assimilates variables at a single fixed time in the assimilation window over the entire spatial domain [Apte et al., 2008]. In the case of 3D-Var the objective function is given by:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - h[\mathbf{x}])^T \mathbf{R}^{-1}(\mathbf{y} - h[\mathbf{x}]). \tag{5.1}$$

The state vector $\mathbf{x}_a$ that minimises this objective function is then used as the initial condition to produce a forecast. When h is linear this equation has an analytic solution [Haben, 2011, eq 2.4], but (5.1) is too expensive to be solved explicitly on an operational scale. In NWP, where observation operators can be nonlinear as well as high-dimensional, a gradient descent algorithm, such as the Gauss-Newton method, is used to solve a sequence of linearised problems, in order to converge iteratively to the solution, $\mathbf{x}_a$ [Haben et al., 2011b]. We note that $\mathbf{x}_a$ corresponds to the maximum a posteriori estimate under the assumption that all probability distributions are Gaussian [Cotter et al., 2012, Rodgers, 2000].

5.3.2 Condition Number

In practice, to solve the nonlinear problem, the Gauss Newton method is used to solve a sequence of linearised problems, often via a conjugate gradient method [Brown et al., 2016]. We will now consider the linearised problem, where the nonlinear problem given by (5.1) is linearised about $\mathbf{x}_a$, the optimal solution.

As the linearisation of (5.1) is a quadratic function [Apte et al., 2008], finding $\mathbf{x}_a$ is equivalent to solving a linear system of the form

$$\mathbf{S}\mathbf{w} = \mathbf{b}, \tag{5.2}$$

where $\mathbf{w} \in \mathbb{R}^N$ and $\mathbf{b} \in \mathbb{R}^N$ is given by (3.10) of [Haben, 2011, Sec 3.2]. (This formulation will be used in numerical experiments in Section 5.6). Here $\mathbf{S} \in \mathbb{R}^{N \times N}$ is the Hessian of the linearisation of the objective function (5.1) given by

$$\mathbf{S} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}, \tag{5.3}$$

where $\mathbf{H} \in \mathbb{R}^{p \times N}$ is the Jacobian of the observation operator h linearised about the optimal state. The Hessian can be used to study the sensitivity of the solution to small changes in observation or background data, by considering its condition number [Golub and Van Loan, 1996, Sec 2.7]. As $\mathbf{B}$ and $\mathbf{R}$ are symmetric positive definite, $\mathbf{S}$ is also symmetric positive definite and hence the L_2 condition number of $\mathbf{S}$ can be represented in terms of its eigenvalues.

5.3.3 Eigenvalue theory

For the remainder of the chapter the following ordering of eigenvalues of matrix $\mathbf{D}$ will be used: For a matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$, let

$$\lambda_{max}(\mathbf{D}) = \lambda_1(\mathbf{D}) \geq \lambda_2(\mathbf{D}) \geq \cdots \geq \lambda_N(\mathbf{D}) = \lambda_{min}(\mathbf{D}). \tag{5.4}$$

Theorem 5.3.1. *If $\mathbf{S} \in \mathbb{R}^{N \times N}$ is a symmetric and positive definite matrix then we can write the condition number in the L_2 norm as*

$$\kappa_2(\mathbf{S}) = \frac{\lambda_1(\mathbf{S})}{\lambda_N(\mathbf{S})}, \tag{5.5}$$

where $\lambda_1(\mathbf{S})$ and $\lambda_N(\mathbf{S})$ correspond to the largest and smallest eigenvalues of $\mathbf{S}$ respectively.

Proof. Golub and Van Loan [1996, Sec. 2.7.2] □

Henceforth $\kappa_2(\mathbf{S})$ will be referred to as the condition number of $\mathbf{S}$, and will be denoted $\kappa(\mathbf{S})$.

In order to determine bounds on the condition number of the Hessian we make use of the following result from linear algebra.

Theorem 5.3.2. *Consider two symmetric matrices $\mathbf{S}_1$, $\mathbf{S}_2 \in \mathbb{R}^{N \times N}$. The k^{th} eigenvalue of the matrix sum $\mathbf{S}_1 + \mathbf{S}_2$ satisfies the following:*

$$\lambda_k(\mathbf{S}_1) + \lambda_N(\mathbf{S}_2) \leq \lambda_k(\mathbf{S}_1 + \mathbf{S}_2) \leq \lambda_k(\mathbf{S}_1) + \lambda_1(\mathbf{S}_2). \tag{5.6}$$

Proof. See Wilkinson [1965, Ch. 2 Thm 44]. □

This result allows us to separate the contributions of $\mathbf{B}^{-1}$ and $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ when bounding the condition number of $\mathbf{S}$ given by (5.3) and is discussed in Section 5.4. A result bounding the eigenvalues of matrix products in terms of the eigenvalues of the constituent matrices is given by

Theorem 5.3.3. *If* $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{N\times N}$ *are positive semi-definite Hermitian matrices, then*

$$\prod_{i=1}^{k} \lambda_i(\mathbf{FG}) \leq \prod_{i=1}^{k} \lambda_i(\mathbf{F})\lambda_i(\mathbf{G}), \quad k = 1, \dots, N-1. \tag{5.7}$$

Proof. See Marshall et al. [2011, Sec. 9 H.1.a.]. □

Theorem 5.3.4. *If* $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{N\times N}$ *are positive semi-definite Hermitian and* $1 \leq i_1 < \cdots < i_k \leq N$*, then*

$$\prod_{t=1}^{k} \lambda_t(\mathbf{FG}) \geq \prod_{t=1}^{k} \lambda_{i_t}(\mathbf{F})\lambda_{N-i_t+1}(\mathbf{G}), \tag{5.8}$$

with equality for $k = N$*.*

Proof. See Wang and Zhang [1992]. □

5.4 Theoretical Results

We now present new bounds on the condition number of the Hessian given by (5.3). We begin in Section 5.4.1 by considering the general case: namely $\mathbf{B}$ and $\mathbf{R}$ are general covariance matrices, and $\mathbf{H}$ is any linear observation operator. In Section 5.4.2 we then introduce further assumptions that constrain $\mathbf{H}$ to only observe state variables. Finally in Section 5.4.3 we restrict the form of $\mathbf{B}$ and $\mathbf{R}$ to have a particular structure.

5.4.1 General bounds on the condition number

We begin by introducing bounds on the eigenvalues of $\mathbf{S}$ in terms of the eigenvalues of $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$,

Lemma 5.4.1. *For* $\mathbf{S} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ *where* $\mathbf{B} \in \mathbb{R}^{N\times N}$, $\mathbf{R} \in \mathbb{R}^{p\times p}$ *are symmetric positive definite covariance matrices, and* $\mathbf{H} \in \mathbb{R}^{p\times N}$ *with* $p < N$*, we can bound the eigenvalues of* $\mathbf{S}$ *below by*

$$\lambda_k(\mathbf{S}) \geq \max\{\lambda_k(\mathbf{B}^{-1}), \quad \lambda_N(\mathbf{B}^{-1}) + \lambda_k(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\}, \tag{5.9}$$

and above by

$$\lambda_k(\mathbf{S}) \leq \min\{\lambda_k(\mathbf{B}^{-1}) + \lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}), \quad \lambda_1(\mathbf{B}^{-1}) + \lambda_k(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\}, \tag{5.10}$$

where $\lambda_k(\mathbf{S})$ *is the* k*th eigenvalue of* $\mathbf{S}$*.*

Proof. The bounds follow immediately from the result of Theorem 5.3.2 by exchanging the order of addition. Note that $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ is not full rank, meaning that $\lambda_N(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) = 0$. □

As we wish to bound the condition number of $\mathbf{S}$, we are primarily interested in bounding $\lambda_1(\mathbf{S})$ and $\lambda_N(\mathbf{S})$. In this case, the bounds given by (5.9) and (5.10) then simplify to

$$\lambda_N(\mathbf{B}^{-1}) \leq \lambda_N(\mathbf{S}) \leq \min\left\{\lambda_N(\mathbf{B}^{-1}) + \lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}), \quad \lambda_1(\mathbf{B}^{-1})\right\}, \tag{5.11}$$

and

$$\max\left\{\lambda_1(\mathbf{B}^{-1}), \quad \lambda_N(\mathbf{B}^{-1}) + \lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\right\} \leq \lambda_1(\mathbf{S}) \leq \lambda_1(\mathbf{B}^{-1}) + \lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) \tag{5.12}$$

We note that this applies to any choice of correlation matrices $\mathbf{B}$ and $\mathbf{R}$ and for any linear choice of observation operator $\mathbf{H}$. This suggests that we expect the eigenvalues, and hence condition number, of $\mathbf{S}$ to vary based on the interactions between $\mathbf{B}$ and $\mathbf{R}$. We now introduce a new bound on the condition number of (5.3) for 3D-Var for the most general choice of $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$:

Theorem 5.4.2. *Let the background and observation error covariance matrices,* $\mathbf{B} \in \mathbb{R}^{N\times N}$ *and* $\mathbf{R} \in \mathbb{R}^{p\times p}$ *respectively, be symmetric positive definite covariance matrices, with* $p < N$*. 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 (given by* (5.3)*),*

$$\begin{aligned}\max\left\{\frac{1 + \lambda_1(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})}{\kappa(\mathbf{B})}, \quad \frac{\kappa(\mathbf{B})}{1 + \lambda_1(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})}\right\} &\leq \kappa(\mathbf{S}) \\ &\leq \Big(1 + \lambda_N(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\Big)\kappa(\mathbf{B}). \end{aligned}\tag{5.13}$$

(This is a slightly modified form of Haben [2011, (6.1.1)].)

Proof. To obtain an upper bound for the condition number of (5.3) we take the upper bound for $\lambda_1(\mathbf{S})$ in (5.12) and the lower bound (5.11) for $\lambda_N(\mathbf{S})$.

$$\kappa(\mathbf{S}) \leq \Big(1 + \lambda_N(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\Big)\kappa(\mathbf{B}), \tag{5.14}$$

using the fact that $(\lambda_1(\mathbf{B}^{-1}))^{-1} = \lambda_N(\mathbf{B})$. We can obtain a lower bound for the condition number similarly by taking the lower bound for $\lambda_1(\mathbf{S})$ in (5.12) and the upper bound for $\lambda_N(\mathbf{S})$ in (5.11). This gives two possible bounds for $\kappa(\mathbf{S})$ depending on which of the two terms is larger, giving

$$\kappa(\mathbf{S}) \geq \max\{\kappa(\mathbf{B})\Big(1 + \lambda_1(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\Big)^{-1}, (\kappa(\mathbf{B}))^{-1}\Big(1 + \lambda_1(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\Big)\} \tag{5.15}$$

using the fact that $(\lambda_1(\mathbf{B}))^{-1} = \lambda_N(\mathbf{B}^{-1})$. Combining these inequalities completes the proof. □

We note that the two terms in (5.15) are reciprocals. This means that the lower bound will always be greater than or equal to one. Any condition number is bounded below by one [Golub and Van Loan, 1996].

We now extend this result to write it in a form that explicitly separates the role of the observation error covariance matrices and the observation operator. This makes it easier to investigate how changes in $\mathbf{R}$, $\mathbf{B}$ and $\mathbf{H}$ affect the condition number of the Hessian.

Corollary 5.4.3. *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 (given by* (5.3)*)*

$$\max\left\{\frac{1 + \frac{\lambda_1(\mathbf{B})}{\lambda_p(\mathbf{R})}\lambda_p(\mathbf{H}\mathbf{H}^T)}{\kappa(\mathbf{B})}, \quad \frac{1 + \frac{\lambda_1(\mathbf{B})}{\lambda_1(\mathbf{R})}\lambda_1(\mathbf{H}\mathbf{H}^T)}{\kappa(\mathbf{B})}, \quad \frac{\kappa(\mathbf{B})}{1 + \frac{\lambda_1(\mathbf{B})}{\lambda_p(\mathbf{R})}\lambda_1(\mathbf{H}\mathbf{H}^T)}\right\} \leq \kappa(\mathbf{S})$$
$$\leq \left(1 + \frac{\lambda_N(\mathbf{B})}{\lambda_p(\mathbf{R})}\lambda_1(\mathbf{H}\mathbf{H}^T)\right)\kappa(\mathbf{B}). \tag{5.16}$$

Proof. Using Theorem 21.10.1 of Harville [1997], we see that $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ has precisely the same non-zero eigenvalues as $\mathbf{R}^{-1}\mathbf{H}\mathbf{H}^T$. Applying the same result, $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ also has the same non-zero eigenvalues as $\mathbf{H}\mathbf{H}^T\mathbf{R}^{-1}$. Therefore $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) = \lambda_1(\mathbf{R}^{-1}\mathbf{H}\mathbf{H}^T) = \lambda_1(\mathbf{H}\mathbf{H}^T\mathbf{R}^{-1})$. Applying Theorem 5.3.3 for $k = 1$

and $i_1 = 1$ yields the following bound:

$$\lambda_1(\mathbf{R}^{-1}\mathbf{H}\mathbf{H}^T) \leq \lambda_1(\mathbf{R}^{-1})\lambda_1(\mathbf{H}\mathbf{H}^T) = \frac{\lambda_1(\mathbf{H}\mathbf{H}^T)}{\lambda_N(\mathbf{R})}, \tag{5.17}$$

as $\lambda_1(\mathbf{R}^{-1}) = 1/\lambda_N(\mathbf{R})$. To bound $\lambda_1(\mathbf{R}^{-1}\mathbf{H}\mathbf{H}^T)$ below, we apply Theorem 5.3.4 for $k = 1$ and $i_1 = 1$ to obtain two lower bounds:

$$\lambda_1(\mathbf{R}^{-1}\mathbf{H}\mathbf{H}^T) \geq \max\{\lambda_1(\mathbf{R}^{-1})\lambda_p(\mathbf{H}\mathbf{H}^T), \lambda_p(\mathbf{R}^{-1})\lambda_1(\mathbf{H}\mathbf{H}^T)\} \tag{5.18}$$

$$\lambda_1(\mathbf{R}^{-1}\mathbf{H}\mathbf{H}^T) \geq \max\{\frac{\lambda_1(\mathbf{H}\mathbf{H}^T)}{\lambda_1(\mathbf{R})}, \frac{\lambda_p(\mathbf{H}\mathbf{H}^T)}{\lambda_p(\mathbf{R})}\} \tag{5.19}$$

Substituting (5.17) and (5.18) into the upper and lower bounds of Theorem 5.4.2 gives the desired result. □

We note that the upper bound in (5.16) increases as $\lambda_N(\mathbf{R})$ decreases. It is not immediately clear how the lower bound will change with $\mathbf{R}$. This will be discussed in Section 5.5.3, which provides a summary of how the bounds given by (5.16) vary with $\mathbf{R}$ and $\mathbf{B}$ for the numerical framework tested in Section 5.6.

5.4.2 Bounds on the condition number with additional restrictions on the choice of observation operator

We now develop a further bound which applies in the case that additional assumptions are made regarding the choice of observation operator. In particular we restrict the observation operator to direct observations of a single state variable. We note that if observations are restricted to direct observations of a single state variable then $\mathbf{H}^T\mathbf{H}$ is diagonal with $(\mathbf{H}^T\mathbf{H})_{i,i} = 1$ if variable i is observed and zero otherwise as shown by Haben et al. [2009]. Under this stricter assumption, we show that the value of $\lambda_1(\mathbf{H}\mathbf{H}^T)$ is the same irrespective of the choice of observations.

Lemma 5.4.4. *If $\mathbf{H}^T\mathbf{H} \in \mathbb{R}^{N\times N}$ is a diagonal matrix with $p < N$ units on the diagonal and the remaining elements zero, then $\mathbf{H}\mathbf{H}^T$ is the $p \times p$ identity matrix.*

Proof. As $\mathbf{H}^T\mathbf{H}$ is diagonal, we can calculate its eigenvalues directly; they are simply its diagonal elements. Hence $\mathbf{H}^T\mathbf{H}$ has p unit eigenvalues and $N - p$ zero eigenvalues. By Theorem 21.10.1 of Harville [1997], $\mathbf{H}\mathbf{H}^T$ has the same non-zero eigenvalues as $\mathbf{H}^T\mathbf{H}$, i.e. p units.

As $\mathbf{H}\mathbf{H}^T$ is symmetric, these eigenvalues correspond to p linearly independent eigenvectors. We now write $\mathbf{H}\mathbf{H}^T$ in terms of its eigendecomposition. Let

$\mathbf{\Lambda} = diag(\lambda_1, ..., \lambda_N) \in \mathbb{R}^{p \times p}$, be the matrix of eigenvalues of $\mathbf{HH}^T$, and $\mathbf{V} \in \mathbb{R}^{p \times p}$ be the corresponding matrix of eigenvectors of $\mathbf{HH}^T$. As the eigenvalues of $\mathbf{HH}^T$ are all units, $\mathbf{\Lambda} = \mathbf{I}_p$, the $p \times p$ identity. Then

$$\mathbf{HH}^T = \mathbf{V\Lambda V}^{-1} = \mathbf{VI}_p\mathbf{V}^{-1} = \mathbf{VV}^{-1} = \mathbf{I}_p. \tag{5.20}$$

Hence under the assumptions on $\mathbf{H}^T\mathbf{H}$, $\mathbf{HH}^T$ is the $p \times p$ identity matrix. □

Hence if observations are restricted to single state variables then $\mathbf{HH}^T = \mathbf{I}_p$. Eliminating the $\mathbf{H}$ and $\mathbf{H}^T$ terms from the bound given by Corollary 5.4.3 reduces the number of matrix multiplications required for evaluation. This result is now used to obtain a bound for the case where observation and background error covariances are correlated and observations are limited to model grid points. We additionally assume that observation variance, σ_o^2 and background variance, σ_b^2 are uniform variances, and hence the covariance matrices can be written as a scalar variance multiplied by a correlation matrix.

Corollary 5.4.5. *Let* $\mathbf{B} = \sigma_b^2\mathbf{C} \in \mathbb{R}^{N \times N}$ *and* $\mathbf{R} = \sigma_o^2\mathbf{D} \in \mathbb{R}^{p \times p}$ *where* $\mathbf{C}$ *and* $\mathbf{D}$ *are symmetric positive-definite correlation matrices, and* σ_b^2 *and* σ_o^2 *are positive scalars denoting the background and observation variances respectively. In addition let* $\mathbf{H}^T\mathbf{H}$ *be a diagonal matrix with* $p < N$ *units on the diagonal and the remaining elements zero. Then the following bound on the condition number of* $\mathbf{S}$ *(given by* (5.3)*) holds:*

$$\max\left\{\frac{1 + \frac{\sigma_b^2}{\sigma_o^2}\frac{\lambda_1(\mathbf{C})}{\lambda_N(\mathbf{D})}}{\kappa(\mathbf{C})}, \frac{\kappa(\mathbf{C})}{1 + \frac{\sigma_b^2}{\sigma_o^2}\frac{\lambda_1(\mathbf{C})}{\lambda_N(\mathbf{D})}}\right\} \leq \kappa(\mathbf{S}) \leq \left(1 + \frac{\sigma_b^2}{\sigma_o^2}\frac{\lambda_N(\mathbf{C})}{\lambda_N(\mathbf{D})}\right)\kappa(\mathbf{C}). \tag{5.21}$$

Proof. Using (5.16) with the definitions of $\mathbf{B}$ and $\mathbf{R}$ in the theorem statement along with the result of Lemma 5.4.4 yields the desired result immediately. □

The bounds given by (5.21) are equal to those given by (5.16) for the case of direct observations, so the comments concerning how the bounds change with $\mathbf{R}$ and $\mathbf{B}$ following Corollary 5.4.3 also apply here. In general, it is not possible to comment on how the lower bound given by (5.21) will behave with changing $\mathbf{B}$ and $\mathbf{R}$. In Section 5.6, we provide an overview for how the terms in (5.21) change for some specific choices of $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$.

We note that the ratio $\frac{\sigma_b^2}{\sigma_o^2}$ appears in both bounds, meaning that as the observations get more accurate, and the variance σ_o^2 decreases, we will see an increased upper bound. The effect of changing σ_o^2 on the lower bound depends on which term is the largest. For $\lambda_1(\mathbf{B}) < \lambda_N(\mathbf{B}) + \lambda_1(\mathbf{R})$ the first term in the lower bound of (5.21) is largest, meaning that decreasing σ_o^2 will increase the value of this term. For $\lambda_1(\mathbf{B}) > \lambda_N(\mathbf{B}) + \lambda_1(\mathbf{R})$, the second term is largest, leading to an increased lower bound for decreasing values of σ_o^2. This was also observed theoretically and numerically in Haben [2011] for the case that $\mathbf{R}$ is uncorrelated. Both of these results assume the same variance for all observations, which is not true in general. However, they indicate the general behaviour we would expect for an increase in accuracy across a wide range of observing systems.

5.4.3 Bounds on the condition number for circulant error covariance matrices

In this section we present a lower bound that is tighter than those of (5.16) for a given matrix framework. Improved bounds are obtained for this specific case by exploiting the eigenvalue and eigenvector properties of a particular matrix structure. It is feasible that for other matrix structures, similar properties could be used to compute tighter bounds for other classes of matrices. However, as the results from Section 5.4.1 are general and apply to any choice of covariance matrices, we do not consider other specialised bounds in this work.

It is often desirable for error correlations to be homogeneous and isotropic, meaning that the correlation between two points is determined solely by the distance between them [Haben et al., 2011a]. This makes circulant matrices a natural choice for correlation matrices on a one-dimensional periodic domain. For the numerical tests discussed in Section 5.6, both $\mathbf{B}$ and $\mathbf{R}$ will be chosen to be circulant matrices, although the bounds given by Theorem 5.4.2, Corollary 5.4.3 and Corollary 5.4.5 apply for any valid choice of correlation matrix.

Definition 5.4.6 (Davis [1979])**.** *A circulant matrix $\mathbf{D} \in \mathbb{R}^{N\times N}$ is a matrix of the form*

$$\mathbf{D} = \begin{pmatrix} d_0 & d_1 & d_2 & \cdots & d_{N-2} & d_{N-1} \\ d_{N-1} & d_0 & d_1 & \cdots & d_{N-3} & d_{N-2} \\ d_{N-2} & d_{N-1} & d_0 & \cdots & d_{N-4} & d_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ d_2 & d_3 s & d_4 & \cdots & d_0 & d_1 \\ d_1 & d_2 & d_3 & \cdots & d_{N-1} & d_0 \end{pmatrix}.$$

As described in Gray [2006], the structure of a circulant matrix of the form given by Definition 5.4.6 permits rapid calculation of eigenvalues and eigenvectors via a discrete Fourier transform. In practice, this means we can calculate the eigenvalues of $\mathbf{D}$ directly via the following formula.

Theorem 5.4.7. *The eigenvalues of a circulant matrix* $\mathbf{D}$*, as given by Definition 5.4.6, are given by*

$$\gamma_m = \sum_{k=0}^{N-1} d_k \omega^{mk}, \tag{5.22}$$

with corresponding eigenvectors

$$\mathbf{v}_m = \frac{1}{\sqrt{N}}(1, \omega^m, \cdots, \omega^{m(N-1)}), \tag{5.23}$$

where $\omega = e^{-2\pi i/N}$ *is an* $N-th$ *root of unity.*

Proof. See Gray [2006] for full derivation. □

To avoid confusion, the eigenvalues of a circulant matrix calculated using (5.22) will be denoted by γ_j rather than λ_j as they are ordered in terms of wavenumber rather than size. We can see from (5.23) that the eigenvectors only depend on N, the dimension of the circulant matrix. Therefore any $N \times N$ circulant matrix will have the same set of eigenvectors.

We now use this matrix structure to consider a further restriction to the case that observation error is assumed to be uncorrelated, and the background error covariance matrix is required to be circulant. In particular in the following theorem, $\mathbf{R}$ is taken to be a scalar multiple of the identity. We note that Theorem 5.4.8 was presented in Haben et al. [2011a] without proof.

Theorem 5.4.8. *Let* $\mathbf{B} = \sigma_b^2 \mathbf{C} \in \mathbb{R}^{N \times N}$ *where* $\mathbf{C}$ *is a symmetric positive-definite circulant matrix, and* $\mathbf{R} = \sigma_o^2 \mathbf{I}_p$ *where* $\mathbf{I}_p \in \mathbb{R}^{p \times p}$ *is the identity matrix. Both* 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. Then the following bounds on the condition number of* $\mathbf{S}$ *(given by* (5.3)*) hold*

$$\left(\frac{1 + \frac{p}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_N(\mathbf{C})}{1 + \frac{p}{N}\frac{\sigma_b^2}{\sigma_o^2}\lambda_1(\mathbf{C})} \right) \kappa(\mathbf{C}) \leq \kappa(\mathbf{S}) \leq \left(1 + \left(\frac{\sigma_b^2}{\sigma_o^2} \right) \lambda_N(\mathbf{C}) \right) \kappa(\mathbf{C}) \tag{5.24}$$

where $\lambda_1(\mathbf{C})$ *and* $\lambda_N(\mathbf{C})$ *are the largest and smallest eigenvalues of the matrix* $\mathbf{C}$ *respectively.*

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 $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) = \sigma_o^{-2}$. Additionally, we have

$\lambda_N(\mathbf{B}) = \sigma_b^2 \lambda_N(\mathbf{C})$. If we substitute these into the upper bound of (5.13) we obtain

$$\kappa(\mathbf{S}) \leq \left(1 + \frac{\sigma_b^2}{\sigma_o^2}\lambda_N(\mathbf{C})\right)\kappa(\mathbf{C}), \tag{5.25}$$

which establishes the upper bound. Rather then repeat this procedure with the lower bound we produce an improved estimate by applying the Rayleigh quotient, $R_{\mathbf{S}}(\mathbf{x}), \mathbf{x} \in \mathbb{C}^N$ (defined in [Süli and Mayer, 2003, Sec 5.9]). Let $\mathbf{v}_1 \in \mathbb{C}^N$ 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 (5.23)). In particular this implies that for an eigenvector, $\mathbf{v}_m$, of $\mathbf{C}^{-1}$

$$\mathbf{v}_m^\dagger \mathbf{H}^T\mathbf{H}\mathbf{v}_m = \frac{1}{N}\sum_{k\in K}\overline{e^{-2\pi ikm/N}}e^{-2\pi ikm/N} = \frac{1}{N}\sum_{k\in K}e^{2\pi ikm/N}e^{-2\pi ikm/N} = \frac{p}{N}, \tag{5.26}$$

where K are the positions of the non-zero diagonal elements of $\mathbf{H}^T\mathbf{H}$ and $\mathbf{v}^\dagger$ denotes the conjugate transpose of $\mathbf{v}$. The maximum value obtained by the Rayleigh quotient of $\mathbf{S}$ occurs at the eigenvector corresponding to the largest eigenvalue of $\mathbf{S}$ [Süli and Mayer, 2003, Sec 5.9]. Hence,

$$\lambda_1(\mathbf{S}) = \max_{\mathbf{v}\in\mathbb{C}^N}(R_{\mathbf{S}}(\mathbf{v})) \geq \mathbf{v}_1^\dagger(\mathbf{B}^{-1} + \sigma_o^{-2}\mathbf{H}^T\mathbf{H})\mathbf{v}_1 = \sigma_b^{-2}\lambda_1(\mathbf{C}^{-1}) + \sigma_o^{-2}\frac{p}{N}. \tag{5.27}$$

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

$$\lambda_N(\mathbf{S}) = \min_{\mathbf{v}\in\mathbb{C}^N}(R_{\mathbf{S}}(\mathbf{v})) \leq \mathbf{v}_N^\dagger(\mathbf{B}^{-1} + \sigma_o^{-2}\mathbf{H}^T\mathbf{H})\mathbf{v}_N = \sigma_b^{-2}\lambda_N(\mathbf{C}^{-1}) + \sigma_o^{-2}\frac{p}{N}. \tag{5.28}$$

Combining (5.27) and (5.28) we find

$$\kappa(\mathbf{S}) \geq \frac{\sigma_b^{-2}\lambda_1(\mathbf{C}^{-1}) + \sigma_o^{-2}\frac{p}{N}}{\sigma_b^{-2}\lambda_N(\mathbf{C}^{-1}) + \sigma_o^{-2}\frac{p}{N}} = \kappa(\mathbf{C})\left(\frac{1 + \frac{\sigma_b^2}{\sigma_o^2}\frac{p}{N}\lambda_N(\mathbf{C})}{1 + \frac{\sigma_b^2}{\sigma_o^2}\frac{p}{N}\lambda_1(\mathbf{C})}\right), \tag{5.29}$$

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

We note that the lower bound presented here is tighter than the others introduced in this section. This comes from the restriction on the form of $\mathbf{S}$ when additional assumptions are made on $\mathbf{R}$ and $\mathbf{H}$, and does not generalise to the other results presented in this work. We also observe that the lower bound (5.24) has an explicit dependence on the number of observations, p. As p increases, the lower bound of (5.24) decreases. Additionally, the ratio $\frac{\sigma_b^2}{\sigma_o^2}$ appears in both bounds, meaning that the discussion following the result of Corollary 5.4.5 also applies to the result of Theorem 5.4.8.

We now have bounds that separate the contributions of **B**, **R** and **H**. In the following section we will test these bounds numerically and discuss the impact of changing each of the constituent matrices in turn.

5.5 Numerical Framework

We now outline the experimental framework that will be used in Section 5.6 to numerically investigate the bounds presented in Section 5.4. In particular, in Section 5.5.1 we introduce specific matrix structures that will be used to generate covariance matrices. We note that these correlation structures illustrate the case where there is a physical lengthscale associated with our observation and background error correlations, as in the case of horizontal correlations. Different choices of observation operator will then be presented in Section 5.5.2. Finally, in Section 5.5.3 we define the experiments that will be studied in Section 5.6 and discuss the choice of parameters to be used in these tests in detail.

5.5.1 Correlation and SOAR Matrices

This work will make use of the second-order auto-regressive correlation (SOAR) function, which is used by the Met Office as a horizontal correlation function, as detailed in Simonin et al. [2014]. It is also commonly used to model background error correlations [Stewart et al., 2013] as its relatively long tails coincide well with estimates of correlation structure. Additionally these longer tails ensure that SOAR matrices are better conditioned for inversion than Gaussian matrices [Haben et al., 2011b, Haben, 2011].

The SOAR function, defined in Daley [1991], is homogeneous and isotropic and naturally extends to a circulant form when we have equally spaced observations on a periodic domain, such as a latitude circle on the Earth. We define the SOAR error correlation matrix for a 1D model with state variables (respectively observations) given by equally spaced gridpoints on a fixed domain on a unit circle (radius $a = 1$) following the procedure given in Haben [2011] and Waller et al. [2016b]. This makes

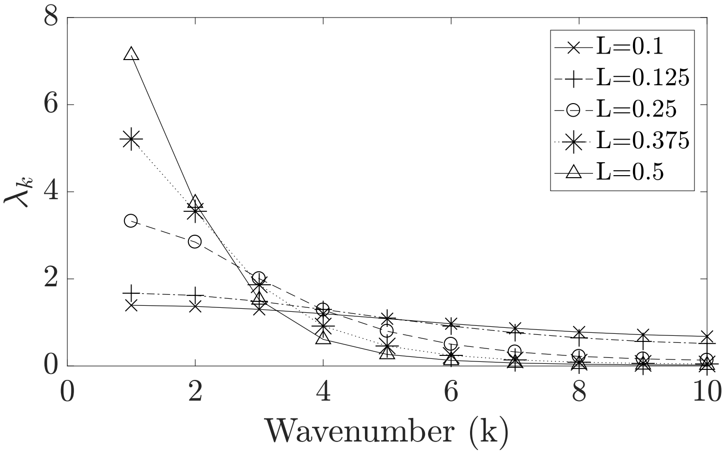


Figure 5.1: Eigenvalues of SOAR error correlation matrix given by (5.30) for $N = 20$ and $a = 1$.

use of a substitution of a chordal distance for a 'great circle distance' to ensure that we obtain a valid correlation model on the circle, as discussed in Gaspari and Cohn [1999] and Jeong and Jun [2015].

Definition 5.5.1. *The SOAR error correlation matrix on the finite domain is given by*

$$\mathbf{D}(i,j) = \left(1 + \frac{\left|2a\sin\left(\frac{\theta_{i,j}}{2}\right)\right|}{L}\right)\exp\left(\frac{-\left|2a\sin\left(\frac{\theta_{i,j}}{2}\right)\right|}{L}\right), \tag{5.30}$$

where $L > 0$ is the correlation lengthscale, $\theta_{i,j}$ denotes the angle between grid points i and j, and a is the radius of the domain. The chordal distance between adjacent grid points is given by

$$\Delta x = 2a\sin\left(\frac{\theta}{2}\right) = 2a\sin\left(\frac{\pi}{N}\right), \tag{5.31}$$

where N is the number of gridpoints and $\theta = \frac{\pi}{2N}$ is the angle between adjacent gridpoints.

As SOAR matrices are circulant by construction, we can calculate their eigenvalues directly using Equation (5.22). The distribution of eigenvalues is symmetric, and as shown in Figure 5.1, decreases monotonically towards the central value. This means that only two eigenvalues need to be calculated in order to obtain the maximum and minimum eigenvalues of any SOAR matrix; γ_1 and $\gamma_{N/2}$ (if N is even) or $\gamma_{(N+1)/2}$ (if N is odd) respectively. The circulant structure can hence be exploited to reduce the number of computations required for computing the bounds given by (5.16) and (5.21) for the condition number of the Hessian.

For the numerical experiments we alter the lengthscales of the SOAR matrices corresponding to background and observation error. Figures will be plotted in terms of the maximum eigenvalues of $\mathbf{B}^{-1}$ and $\mathbf{R}^{-1}$ (recalling that for any matrix $\mathbf{D} \in \mathbb{R}^{m\times m}$, $\lambda_1(\mathbf{D}^{-1}) = 1/\lambda_N(\mathbf{D})$). We note that this also means that $\lambda_1(\mathbf{D}^{-1}) = \gamma_{N/2}(\mathbf{D}^{-1})$ for N even (or $\lambda_1(\mathbf{D}^{-1}) = \gamma_{(N+1)/2}(\mathbf{D}^{-1})$ for N odd), using the notation established in Theorem 5.4.7. The relationship between increasing lengthscale and the spectrum of a SOAR matrix is shown in Figure 5.1 - namely that as the lengthscale, L, increases, the minimum eigenvalue of the SOAR matrix decreases and the maximum eigenvalue increases. This means that the maximum eigenvalue of the inverse of a SOAR matrix increases with lengthscale, and its minimum eigenvalue decreases.

Having described the choice of correlation matrices that will be used in the numerical tests in Section 5.6, in the next section we discuss the different choices of observation operator that will be tested in our experiments.

5.5.2 Choice of Observation Operator

Most previous research into the impact of correlated observation errors on the variational assimilation problem does not investigate the impact of using different observation operators systematically. Either the operational observation operator is used e.g. Weston et al. [2014], Bormann et al. [2016], or experiments are carried out in a simple linear case where $\mathbf{H}$ is taken to be a variant of the identity, as in Stewart et al. [2008b, 2013], Waller et al. [2014a], Ménard [2016]. In this chapter we compare how the condition number of the Hessian is affected by different choices of linear observation operator in order to gain some theoretical insight into the role played by this operator. We define three choices of observation operator that will be investigated in detail numerically. We are particularly interested in how important our choice of $\mathbf{H}$ is in determining both the true condition number of $\mathbf{S}$ and the value of the bounds given by (5.16). Firstly we note that all bounds presented in this work require the assumption that the observation operator, $\mathbf{H}$, is linear, and the bounds given by (5.21) and (5.24) have the restriction that observations are only of single state variables. All the choices of $\mathbf{H}$ that are tested in the numerical experiments presented in this work are linear, and two correspond to direct observations of single model variables.

Definition 5.5.2. *The observation operators* $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{H}_3 \in \mathbb{R}^{p\times N}$*, for* $N = 2p$*, are*

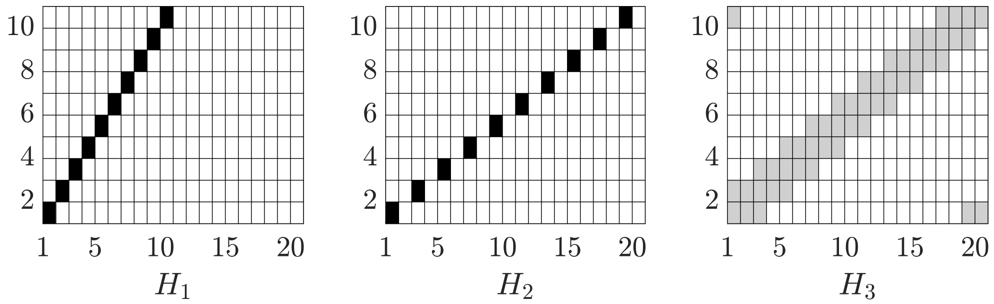


Figure 5.2: Visualisation of the observation operators described in Definition 5.5.2 for the case $p = 10$ and $N = 20$. Shading indicates the value of the entry in the matrix; in the case of $\mathbf{H}_1$ and $\mathbf{H}_2$ all non-zero entries are 1, and for $\mathbf{H}_3$ all non-zero entries are $\frac{1}{5}$.

defined as follows:

$$\mathbf{H}_1(i,j) = \begin{cases} 1, & j = i \textit{ for } i = 1,\dots,p \\ 0, & \textit{otherwise.} \end{cases} \tag{5.32}$$

$$\mathbf{H}_2(i,j) = \begin{cases} 1, & j = 2i \textit{ for } i = 1,\dots,p \\ 0, & \textit{otherwise.} \end{cases} \tag{5.33}$$

$$\mathbf{H}_3(i,j) = \begin{cases} \frac{1}{5}, & j \in \{2i-2, 2i-1, 2i, 2i+1, 2i+2 \pmod N\} \textit{ for } i = 1,\dots,p \\ 0, & \textit{otherwise.} \end{cases} \tag{5.34}$$

The choice of $\mathbf{H} = \mathbf{H}_1$ corresponds to observing the first p state variables, and making no observations in the second half of the state space. Choosing $\mathbf{H} = \mathbf{H}_2$ corresponds to making observations at alternate state variables over the entire model domain. The observation operator $\mathbf{H} = \mathbf{H}_3$ is a smoothed version of $\mathbf{H}_2$; state variables at alternate grid points are smoothed over 5 adjacent points in state space with equal weighting. This can be thought of as a simplified version of a satellite weighting function, [Stewart, 2010, Sec. 2.4.1] [Rodgers, 2000, Sec 2.1.3.], which measures average radiation over several model levels of the atmosphere. In Figure 5.2 these observation operators are depicted for a small scale example when $p = 10$ and $N = 20$.

The choice of $\mathbf{H}_1$ was made as a check to allow comparison with preliminary numerical tests with those from Chapter 6 of Haben [2011]. The bounds given by Corollary 5.4.5 in Section 5.4 require that $\mathbf{H}^T\mathbf{H}$ be a diagonal matrix with p units on the diagonal. The observation operator $\mathbf{H}_1$ satisfies this requirement, as does $\mathbf{H}_2$, meaning that we can apply the bounds of Corollary 5.4.5 for these two cases.

Additionally, by Lemma 5.4.4, $\mathbf{H}_1\mathbf{H}_1^T = \mathbf{H}_2\mathbf{H}_2^T$. This means that for fixed choices of $\mathbf{B}$ and $\mathbf{R}$, both $\mathbf{H} = \mathbf{H}_1$ and $\mathbf{H} = \mathbf{H}_2$ will yield the same upper and lower bounds. We wish to see whether there will be a significant difference in the true condition number of $\mathbf{S}$ for $\mathbf{H} = \mathbf{H}_1$ and $\mathbf{H} = \mathbf{H}_2$.

As $\mathbf{H} = \mathbf{H}_3$ does not satisfy the condition in the statement of Corollary 5.4.5, we must apply the more general bound given by (5.16) in Corollary 5.4.3. We would like to be able to use the same bounds to compare each of the three choices of observation operator. A short calculation reveals that we have equality of the bounds given by Corollaries 5.4.3 and 5.4.5 when observations are restricted to model grid points for the framework described here. Hence, for what follows we will be comparing the bounds given by (5.16) irrespective of the observation network chosen.

5.5.3 Experimental Design

We now discuss the experimental framework which will be used for the numerical tests presented in Section 5.6. In particular we motivate the range of parameters that will be investigated.

We fix the ratio between p, the number of observations, and N, the number of state variables, to be $N = 2p$ for all the experiments discussed below. The same ratio was used for numerical testing in Haben [2011] and is not representative of what is used in practice, where observations are much less dense. Unless stated otherwise, the values $N = 200$ and $p = 100$ were used for all the plots presented here. Other choices of p and N were studied in detail; as qualitative results were similar for all cases considered they will not be shown here. Both background error covariance matrix, $\mathbf{B} \in \mathbb{R}^{N\times N}$, and observation error covariance matrix, $\mathbf{R} \in \mathbb{R}^{p\times p}$, are chosen to be SOAR correlation matrices (see Section 5.5.1) with fixed variances $\sigma_b^2 = \sigma_o^2 = 1$.

The domain for the tests is the unit circle ($a = 1$). In the experiments that follow we will vary L_R, the correlation lengthscale of the SOAR matrix defining $\mathbf{R}$, and L_B, the correlation lengthscale of the the SOAR matrix defining $\mathbf{B}$, over a regular grid.In addition to studying the impact of changing the lengthscale of $\mathbf{B}$ and $\mathbf{R}$ for both sets of experiments, we also consider the effect of using the different choices of $\mathbf{H}$ presented in Section 5.5.2.

Table 5.1: Summary of how terms that appear in (5.16) change with the lengthscale L_R for $\mathbf{R} \in \mathbb{R}^{100\times 100}$.

	Lengthscale L_R				
	0.1	0.33	0.66	0.99	1
$\lambda_N(\mathbf{R})$	1.92×10^{-2}	5.74×10^{-4}	7.21×10^{-5}	2.14×10^{-5}	2.08×10^{-5}
$\lambda_1(\mathbf{R})$	6.40×10^{0}	2.26×10^{1}	4.67×10^{1}	6.36×10^{1}	6.40×10^{1}

Table 5.2: Summary of how terms that appear in (5.16) change with the lengthscale L_B for $\mathbf{B} \in \mathbb{R}^{200\times 200}$.

	Lengthscale L_B				
	0.1	0.33	0.66	0.99	1
$\lambda_N(\mathbf{B})$	2.54×10^{-3}	7.19×10^{-5}	8.99×10^{-6}	2.67×10^{-6}	2.59×10^{-6}
$\lambda_1(\mathbf{B})$	1.28×10^{1}	4.51×10^{1}	9.35×10^{1}	1.27×10^{2}	1.28×10^{2}
$\kappa(\mathbf{B})$	5.05×10^{3}	6.28×10^{5}	1.40×10^{7}	4.77×10^{7}	4.95×10^{7}

5.5.3.1 Condition number testing

In the numerical tests we consider how the condition number of $\mathbf{S}$ (calculated using the Matlab 2016b function *cond*) and the bounds given by (5.16) change as the minimum eigenvalues of both error covariance matrices change. Of particular interest is the interaction between changes to both $\mathbf{B}$ and $\mathbf{R}$. For the results presented in this chapter the lengthscales of both $\mathbf{B}$ and $\mathbf{R}$ were varied between 0.1 and 1. The equivalent eigenvalues of $\mathbf{R}$ and $\mathbf{B}$ for these parameters are given in Tables 5.1 and 5.2 respectively.

Tables 5.1 and 5.2 presents values of the terms that appear in (5.16) and depend on the background and observation error matrices for typical experimental values of L_B and L_R respectively. We observe that:

- As L_R increases $\lambda_N(\mathbf{R})$ decreases; hence the first term in the lower bound of (5.16) will increase with increasing L_R, and the third term in the lower bound of (5.16) will decrease with increasing L_R. It is therefore not possible in general to determine how the lower bound will change with increasing L_R.
- As L_R increases, $\lambda_1(\mathbf{R})$ increases, meaning that the second term in the lower bound of (5.16) will decrease with increasing L_R.
- As L_B increases, the difference between its minimum and maximum eigenvalues increases, meaning that the condition number of $\mathbf{B}$ increases with L_B.
- In this setting, the upper bound of (5.16) will increase as L_R or L_B increases, as

$\lambda_1(\mathbf{B})$ and $\kappa(\mathbf{B})$ increase with L_B and $\frac{1}{\lambda_N(\mathbf{R})}$ increases with L_R.

- As L_B increases, the ratio $\frac{\kappa(\mathbf{B})}{\lambda_1(\mathbf{B})} = \frac{1}{\lambda_N(\mathbf{B})}$ increases, meaning that for fixed L_R the first and second terms of the lower bound of (5.16) will decrease and the third term will increase.

Therefore increasing L_B for fixed L_R will cause both bounds to increase. It is not possible at this stage to say whether the upper and lower bound will move closer together or further apart as L_B increases. It is also not clear which term in the lower bound of (5.16) will be largest for a general choice of $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$. This means that we cannot say how the lower bound of (5.16) will change with L_R. We will investigate how the bounds change numerically with $\mathbf{B}$ and $\mathbf{R}$ in Section 5.6. Although we understand the effect of changing L_B and L_R on the bounds of the condition number, we now want to investigate their influence on the actual value of $\kappa(\mathbf{S})$.

5.5.3.2 Convergence of a conjugate gradient routine

In addition to studying how the condition number of the Hessian changes with $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$, it is of interest to determine the effect of these same changes on the rate of convergence of the minimisation of the objective function. In order to do this we consider the convergence rate of a conjugate gradient method applied to the linear system (5.2) associated with the 3D-Var cost function (5.1).

To do this, we follow the same method that is used in Chapter 6 of Haben [2011]; we construct a vector $\mathbf{w}$ that has small and large scale features, calculate $\mathbf{b} = \mathbf{S}\mathbf{w}$ and then recover $\mathbf{w}$ by applying a linear solver, in this case the conjugate gradient method, to $\mathbf{S}\mathbf{w} = \mathbf{b}$. In this case we used the Matlab conjugate gradient routine, *pcg.m* MATLAB [2016], to investigate the change in the number of iterations to convergence. In exact arithmetic the conjugate gradient method should converge to the true solution in exactly n iterations for an n-dimensional problem [Gill et al., 1986]. We note that in finite precision, convergence in n iterations may not occur as the search directions lose conjugacy due to round-off errors [Bardesley et al., 2013]. Operationally however, even n iterations is too many in order to obtain a solution in reasonable computational time. This problem is usually solved by preconditioning, but for this chapter we are interested in the unpreconditioned problem as discussed in Section 5.2. We use a tolerance of 1×10^{-6} on the relative residual for all results presented in the next section.

We expect that the impact of changing $\mathbf{B}$ and $\mathbf{R}$ on the condition number of the Hessian will be similar for both sets of experiments (condition number and conjugate gradient convergence) due to the theoretical link between the condition number and convergence of the conjugate gradient method [Nocedal, 2006, Golub and Van Loan, 1996]. As well as investigating the impact of changing lengthscale on the convergence of 3D-Var, we are interested in how the choice of observation operators introduced in Section 5.5.2 influences 3D-Var in terms of both the condition number and convergence of the conjugate gradient method.

5.6 Numerical Testing

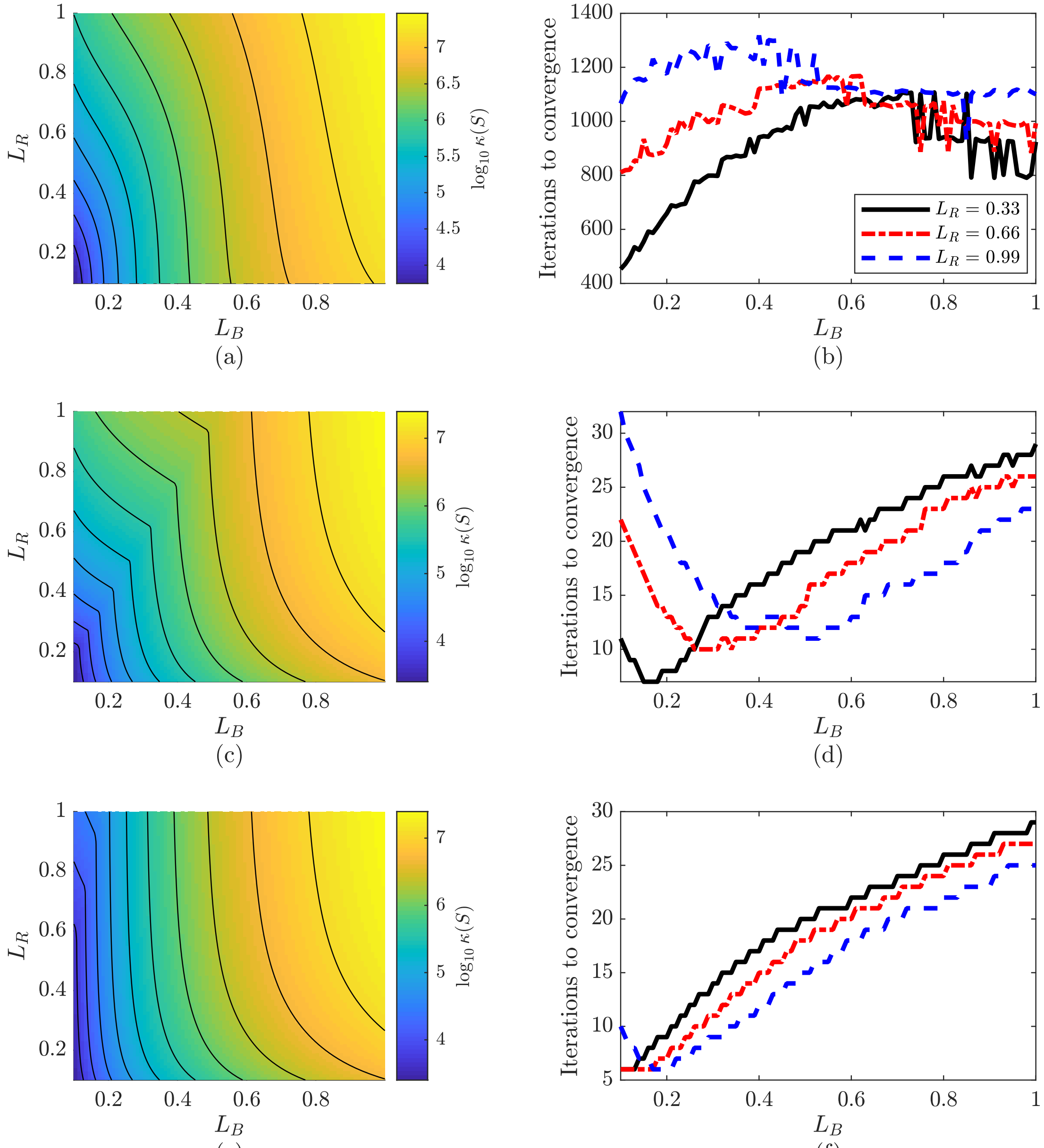
Our experiments focus on how $\kappa(\mathbf{S})$ changes with both L_R (for $\mathbf{R}$ correlated) and L_B for each of the choices of observation operator introduced in Section 5.5.2 (recalling that for any matrix $\mathbf{D} \in \mathbb{R}^{N\times N}$, $\lambda_1(\mathbf{D}^{-1}) = 1/\lambda_N(\mathbf{D})$). This extends the experiments of Haben [2011] where the effect of the lengthscale of $\mathbf{B}$ on the conditioning of the Hessian was considered for uncorrelated $\mathbf{R}$. We also investigate how correlations in $\mathbf{B}$ and in $\mathbf{R}$ interact in terms of both the bounds and the true conditioning of the Hessian. We then test our conclusions in terms of a minimisation problem, to assess the impact of changing correlation lengthscales on the number of iterations required for convergence of a conjugate gradient routine. We present and discuss the results for $\mathbf{H} = \mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$ separately before comparing the different cases.

5.6.1 Investigating changing lengthscales: observing the first p variables ($\mathbf{H} = \mathbf{H}_1$)

In Figure 5.3a we plot the condition number of $\mathbf{S}$ (colour) with L_B shown along the x-axis, and L_R shown on the y-axis, for the case $\mathbf{H} = \mathbf{H}_1$. Both axes and the colour values are shown with a logarithmic scale. We recall that as lengthscale increases, $\lambda_1(\mathbf{B}^{-1})$ and $\lambda_1(\mathbf{R}^{-1})$ both increase.

We observe that:

- For a fixed value of L_R, increasing L_B results in an increased value of $\kappa(\mathbf{S})$. This behaviour is also seen in Haben [2011] for an uncorrelated choice of $\mathbf{R}$. The effect of this increase depends on the size of L_R - larger values of L_R) lead to smaller gradients in the contours of $\kappa(\mathbf{S})$. The inclusion of correlated observation errors therefore results in a more complex dependence of $\kappa(\mathbf{S})$ on $\mathbf{B}$.



(a) (b)

(c) (d)

(e) (f)

Figure 5.3: Impact of different choices of observation operator $\mathbf{H}$ on $\kappa(\mathbf{S})$ (a, c, e) and convergence of the conjugate gradient algorithm (b, d, f) for: (a, b) $\mathbf{H} = \mathbf{H}_1$, (c, d) $\mathbf{H} = \mathbf{H}_2$ and (e, f) $\mathbf{H} = \mathbf{H}_3$. The matrices $\mathbf{B}$ and $\mathbf{R}$ are SOAR matrices (5.30) for $N = 200$ and $p = 100$. The x-axis denotes L_B. For (a, c, e) the y-axis shows L_R and the linestyle denotes $\log_{10}(\kappa(\mathbf{S}))$. Ten equally spaced contours (solid lines), and horizontal lines (corresponding to the lines plotted in Figures (b, d and f)) are also shown. The solid, dotted and dash-dotted lines represent $L_R = 0.33$, 0.66 and 0.99 respectively.

- For a small fixed value of L_B, increasing L_R results in an increased value of $\kappa(\mathbf{S})$, whereas for a large fixed value of L_B increasing L_R has minimal impact on the value of $\kappa(\mathbf{S})$.
- In general the impact of changing L_B on $\kappa(\mathbf{S})$ is larger than that of changing L_R.

We hence note that interactions of L_B and L_R have an important effect on the condition number of $\mathbf{S}$. This agrees with the results of Corollary 5.4.1 which showed that depending on the relationship between the largest eigenvalues of $\mathbf{B}^{-1}$ and $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$, there are two distinct bounds on the eigenvalues of $\mathbf{S}$, one in terms of $\lambda_1(\mathbf{B}^{-1})$ and one in terms of $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})$.

In Figure 5.3b we see the number of iterations required for the conjugate gradient method to solve the problem described in Section 5.5.3. The values of L_R plotted in Figure 5.3b are shown on Figure 5.3a as horizontal lines for 80 values of L_B. We note that the number of iterations required for convergence is extremely high, and in fact larger than the dimension of the problem. Although the conjugate gradient method converges in N iterations in exact arithmetic, iterates past iteration count N continue to approach the true solution for this problem. In particular, a large number of iterations of the conjugate gradient method applied to (5.2) are required in order to recover the large scale structure of $\mathbf{w}$. For applications, computational resource typically demands that much fewer than N iterations are used.

Firstly, for $L_B < 0.44$, increasing L_B for fixed L_R results in an increase in the number of iterations required for convergence. Additionally, for fixed L_B, increasing L_R results in a clear increase in the number of iterations. This behaviour agrees well with the qualitative conclusions from the condition number experiment in Figure 5.3b. For $L_B > 0.4$ we see a decrease in the number of iterations as L_B increases. In this range, the value of $\kappa(\mathbf{S})$ is similar across each of the horizontal lines shown in Figure 5.3a, so we could expect the number of iterations to convergence to be similar. Additionally, the Hessian is extremely ill-conditioned, which combined with a small tolerance in the conjugate gradient routine could explain the noisy values for large L_B.

5.6.2 Investigating changing lengthscales: observing p alternate state variables ($\mathbf{H} = \mathbf{H}_2$)

In Figure 5.3c we see how changing $\mathbf{B}$ and $\mathbf{R}$ affects the condition number of $\mathbf{S}$ for the case $\mathbf{H} = \mathbf{H}_2$. The changes in $\kappa(\mathbf{S})$ with L_R and L_B are qualitatively similar to the case $\mathbf{H}_1$ described in Section 5.6.1. Again we see interaction between $\mathbf{B}$ and $\mathbf{R}$ has an important effect on $\kappa(\mathbf{S})$, in agreement with the results of Corollary 5.4.1. However, for $\mathbf{H} = \mathbf{H}_2$ the change of behaviour of $\kappa(\mathbf{S})$ does not occur smoothly; we observe a discontinuity in the gradient of the contours. As L_R increases the value of L_B at which this 'kink' occurs also increases linearly. We will investigate this kink further in Section 5.6.6 and show that it is caused by a change in regime.

In Figure 5.3d we see the number of iterations required for the conjugate gradient method to converge for the case $\mathbf{H} = \mathbf{H}_2$.

- For fixed values of L_R we observe a change in behaviour as L_B increases; for smaller values of L_B we see a decrease in the number of iterations as L_B increases and for larger values of L_B the number of iterations increases with L_B. This does not agree with the results for the condition number of $\mathbf{S}$ in Figure 5.3c, where an increase in L_B causes an increase in $\kappa(\mathbf{S})$ for all values of L_R.
- For smaller values of L_B, increasing L_R leads to an increase in the number of iterations required for convergence. For larger values of L_B, that occur to the right of the kink, increasing L_R decreases the number of iterations. Again, this is unlike the results seen for the condition number, where increasing L_R leads to an increase in both the actual value and upper bound of $\kappa(\mathbf{S})$ for all values of L_B.

We note that the value of L_B where this change in behaviour occurs is the same as the value of $\lambda_1(\mathbf{B}^{-1})$ where the change in gradient of the contours occurs in Figure 5.3c, indicating that the kink is caused by an underlying change in regime. If we consider the eigenvalues of $\mathbf{S}$ (not shown here), clustering of eigenvalues increases as the kink is approached. The clustering of eigenvalues is important for convergence of a conjugate gradient method [Nocedal, 2006], and is not detected by the condition number. This explains the difference in behaviour between Figure 5.3c and 5.3 d with increasing L_B.

5.6.3 Investigating changing lengthscales: observing p alternate variables smoothed over 5 state variables ($\mathbf{H} = \mathbf{H}_3$)

In Figure 5.3e we see how changing $\mathbf{B}$ and $\mathbf{R}$ affects the condition number of $\mathbf{S}$ for the case $\mathbf{H} = \mathbf{H}_3$. The behaviour of $\kappa(\mathbf{S})$ with changing L_B and L_R is qualitatively similar to the case $\mathbf{H} = \mathbf{H}_2$. However, for $\mathbf{H} = \mathbf{H}_3$ and fixed L_B, only changes to very large values of L_R result in a significant change to $\kappa(\mathbf{S})$ and this is true for only the smallest values of L_B. Again, interaction between L_B and L_R has an important

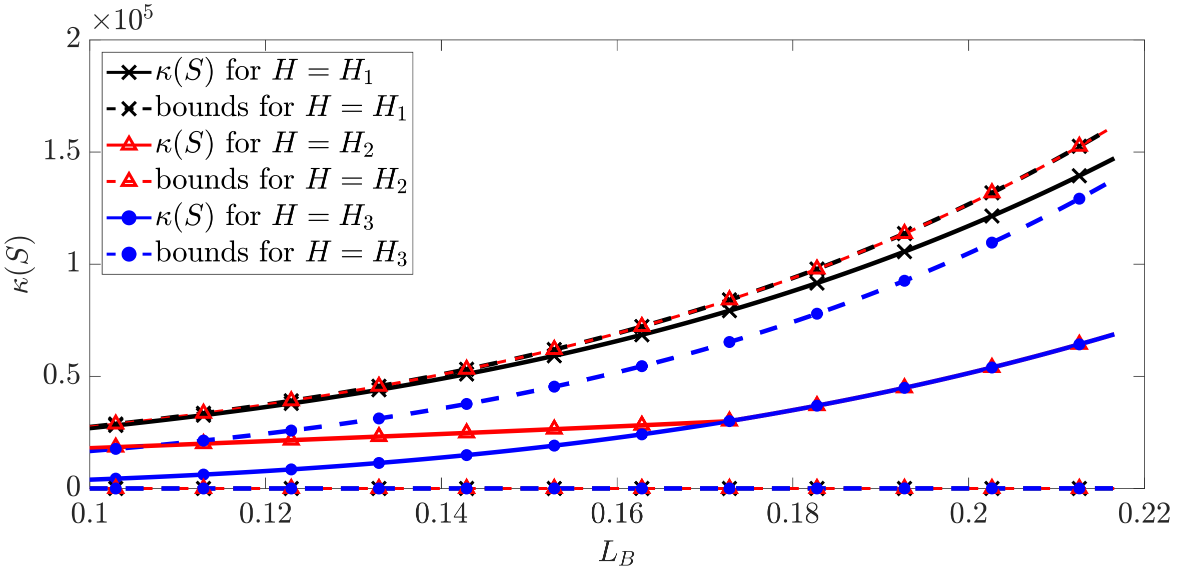


Figure 5.4: Bounds (dashed lines) and condition number (solid lines) of $\mathbf{S}$ for $\mathbf{H}_1$ (cross), $\mathbf{H}_2$ (triangle) and $\mathbf{H}_3$ (circle) for $L_R = 0.33$. The bounds are calculated using (5.16) for all choices of $\mathbf{H}$. We note that the bounds for the cases $\mathbf{H}_1$ and $\mathbf{H}_2$ are the same.

impact on $\kappa(\mathbf{S})$ but to much less of an extent than in the previous two cases. This agrees with the results of Corollary 5.4.1, as the value of $\lambda_1(\mathbf{H}_3^T\mathbf{R}^{-1}\mathbf{H}_3)$ is much smaller than $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})$ for $\mathbf{H} = \mathbf{H}_1$ or $\mathbf{H}_2$, and hence L_R will need to take a much larger value in order that $\lambda_1(\mathbf{H}_3^T\mathbf{R}^{-1}\mathbf{H}_3) + \lambda_N(\mathbf{B}^{-1}) > \lambda_1(\mathbf{B}^{-1})$. A discontinuity in gradient similar to the one observed for the case $\mathbf{H} = \mathbf{H}_2$ is seen here, but for much larger values of $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})$ than for Figure 5.3e.

In Figure 5.3f we see the number of iterations required for the conjugate gradient to converge for the problem described in Section 5.5.3 when $\mathbf{H} = \mathbf{H}_3$. Similarly to Figure 5.3b, we see an initial decrease in the number of iterations required for convergence, before a turning point where the number of iterations increases with L_B. This turning point occurs for the same values of L_B as the discontinuity in gradient that was seen in Figure 5.3e. As the value of L_B at which this kink occurs is much smaller than for the case $\mathbf{H} = \mathbf{H}_2$, for most values of L_B increasing L_R decreases the number of iterations. As in the case $\mathbf{H} = \mathbf{H}_2$, clustering of the eigenvalues of $\mathbf{S}$ increases as we approach the kink. The structure of the eigenvalues is more important in determining the convergence of a conjugate gradient method than the condition number in this case.

5.6.4 Investigating bounds and actual value of $\kappa(\mathbf{S})$ for different choices of observation operator

We now compare the effect of changing the observation operator on both the condition number of $\mathbf{S}$ and the bounds of $\mathbf{S}$ introduced in Section 5.4. Of particular interest is how tight the bounds are for different values of $\lambda_1(\mathbf{B}^{-1})$. For clarity, the Hessian for the cases $\mathbf{H} = \mathbf{H}_1$, $\mathbf{H} = \mathbf{H}_2$ and $\mathbf{H} = \mathbf{H}_3$ will be referred to as $\mathbf{S}_1$, $\mathbf{S}_2$ and $\mathbf{S}_3$ respectively. Figure 5.4 displays the actual value of the condition number and the bounds from (5.16) for a fixed choice of $\mathbf{R}$ with $L_R = 0.33$ for all three choices of $\mathbf{H}$. We recall (Section 5.5.2) that the bounds for the cases $\mathbf{H} = \mathbf{H}_1$ and $\mathbf{H} = \mathbf{H}_2$ are equal, with tighter bounds for the case $\mathbf{H} = \mathbf{H}_3$. This is because the maximum eigenvalue of $\mathbf{H}_3\mathbf{H}_3^T$, which appears in both upper and lower bounds, is 0.52 rather than 1.

- Figure 5.4 shows cases where both the upper and lower bound give by (5.16) are tight. The upper bound is close to the actual value of $\kappa(\mathbf{S})$ for $\mathbf{H}_1$, particularly when L_B is small. For small values of L_B the actual value of $\kappa(\mathbf{S})$ for $\mathbf{H}_3$ is much closer to the lower bound than the upper bound.

- The kink that was observed in Figure 5.3c for $\mathbf{H} = \mathbf{H}_2$ can also be seen in Figure 5.4. The kink occurs at the location where $\kappa(\mathbf{S}_2)$ coincides with $\kappa(\mathbf{S}_3)$. For values of L_B greater than the kink, $\kappa(\mathbf{S}_2)$ and $\kappa(\mathbf{S}_3)$ are very close to each other.

- For all choices of $\mathbf{H}$ shown in Figure 5.4, increasing L_B leads to the upper bound moving away from both the lower bound and the actual value of $\kappa(\mathbf{S})$.

We note that we have found different choices of $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$ where the actual values of $\mathbf{S}$ are close to both the upper and lower bounds given by (5.16). We now discuss the implications of changing $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$ in terms of the condition number of $\mathbf{S}$ and the number of iterations required for the conjugate gradient to converge.

5.6.5 Comparison of results

In this section we compare the results of the previous sections for different choices of observation operator $\mathbf{H}$, as well as different choices of $\mathbf{B}$ and $\mathbf{R}$. We recall that $\lambda_1(\mathbf{B}^{-1}) = 1/\lambda_N(\mathbf{B})$ and $\lambda_1(\mathbf{R}^{-1}) = 1/\lambda_N(\mathbf{R})$.

We begin by considering how the lower bounds given by Lemma 5.4.1 for $\lambda_1(\mathbf{S})$ change depending on whether $\lambda_1(\mathbf{B}^{-1})$ or $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) + \lambda_N(\mathbf{B}^{-1})$ is the larger term.

- For a fixed value of L_R and changing L_B: for small values of $\lambda_1(\mathbf{B}^{-1})$, the lower bound of $\lambda_1(\mathbf{S})$ from (5.9) is given by $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) + \lambda_N(\mathbf{B}^{-1})$, meaning that the maximum eigenvalue of $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ is most important for determining $\lambda_1(\mathbf{S})$.

- As L_B increases, at some point $\lambda_1(\mathbf{B}^{-1})$ will be larger than $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) + \lambda_N(\mathbf{B}^{-1})$, meaning that $\lambda_1(\mathbf{B}^{-1})$ will be the most important term for determining $\lambda_1(\mathbf{S})$.

- Alternatively, fixing L_B and changing L_R we observe similar behaviour: for smaller values of L_R, we see less impact on $\kappa(\mathbf{S})$ when changing L_R than for larger values of L_R, where a change in $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})$ has a significant effect on the value of $\kappa(\mathbf{S})$.

This behaviour is seen for all choices of $\mathbf{H}$ in Figure 5.3. This bound also provides justification for the variation with L_B and L_R in the gradient of the contours seen in Figures 5.3a, 5.3c and 5.3e.

We now consider the similarities between different choices of observation operator for the two experiments:

- For a fixed choice of $\mathbf{H}$ there are strong similarities between the effect of increasing $\lambda_1(\mathbf{B}^{-1})$ on the convergence of the conjugate gradient method and the effect on the condition number of the Hessian. In particular the kink in the condition number (Figures 5.3c and 5.3e) and the change in gradient for convergence (Figures 5.3d and 5.3f) occur at the same values of L_B and L_R for both $\mathbf{H} = \mathbf{H}_2$ and $\mathbf{H} = \mathbf{H}_3$. This indicates that the kink is due to a change in the underlying structure of $\mathbf{S}$.

- The effect of varying L_R and L_B for $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$ was broadly similar in terms of $\kappa(\mathbf{S})$, with the main difference being the discontinuity in the contours of $\kappa(\mathbf{S})$ seen for $\mathbf{H}_2$ and $\mathbf{H}_3$ but not for $\mathbf{H}_1$.

We also see some large differences between the two experiments. The main dissimilarity between the graphs for condition number (Figures 5.3a, 5.3c and 5.3e) and for convergence (Figures 5.3b, 5.3d and 5.3e) is that increasing L_B uniformly results in an increase in the condition number of $\mathbf{S}$, but is not always linked to an increase in the number of iterations required for convergence. This difference was explained in Sections 5.6.1-5.6.3 by the clustering of eigenvalues near the kink for $\mathbf{H}_2$ and $\mathbf{H}_3$.

For the conjugate gradient experiments, conclusions for the cases $\mathbf{H} = \mathbf{H}_2$ and $\mathbf{H} = \mathbf{H}_3$ were very different to the case $\mathbf{H} = \mathbf{H}_1$. Both $\mathbf{H} = \mathbf{H}_2$ and $\mathbf{H} = \mathbf{H}_3$ have block-circulant structures, meaning that in these cases $\mathbf{S}$ will have a block-circulant structure. We suggest that this is the reason for the difference in eigenvalue clustering behaviour compared to the case $\mathbf{H} = \mathbf{H}_1$. This was tested through the use of an additional non-circulant observation operator made by observing 100 random state variables. The behaviour in this case is very similar to that which was observed for $\mathbf{H} = \mathbf{H}_1$. The fact that qualitative behaviour for the case $\mathbf{H} = \mathbf{H}_1$ is the same as for the randomly selected observation operator supports the conjecture that the rapid convergence of the conjugate gradient seen for $\mathbf{H} = \mathbf{H}_2$ and $\mathbf{H} = \mathbf{H}_3$ is caused by the inherent block-circulant structure of $\mathbf{S}_2$ and $\mathbf{S}_3$.

5.6.6 Understanding the Discontinuity in the Gradient for $\mathbf{H} = \mathbf{H}_2$ and $\mathbf{H} = \mathbf{H}_3$

We now return to discuss the discontinuity in the gradient, or kink, that was observed for $\mathbf{H} = \mathbf{H}_2$ and $\mathbf{H} = \mathbf{H}_3$ for both the condition number of $\mathbf{S}$ (Figures 5.3c and 5.3e) and the convergence of the conjugate gradient method (Figures 5.3d and 5.3f). We now explain this theoretically, and discuss why the discontinuity in gradient is observed for $\mathbf{H}_2$ and $\mathbf{H}_3$ but not for $\mathbf{H}_1$. We begin by considering bounds for the eigenvalues of $\mathbf{S}$ in terms of the eigenvalues of $\mathbf{B}^{-1}$ and $\mathbf{R}^{-1}$, using the bounds given by Corollary 5.4.1 and the discussion that follows in Section 5.4.1.

Equations (5.9) - (5.12) explain the variation with $\lambda_1(\mathbf{B}^{-1})$ and $\lambda_1(\mathbf{R}^{-1})$ that was observed in Figure 5.3. However, as the bounds in (5.11) and (5.12) apply to all choices of $\mathbf{H}$, they do not explain the difference between the choices of $\mathbf{H}$ for which the kink is observed ($\mathbf{H}_2$ and $\mathbf{H}_3$) and the choices of $\mathbf{H}$ which have smoothly varying values of $\kappa(\mathbf{S})$, (namely $\mathbf{H}_1$).

In order to illustrate why the kink occurs for some choices of $\mathbf{H}$ but not for others we present a tighter upper bound for the specific framework used in the numerical experiments for two cases, beginning with $\mathbf{H}_1$. By expressing $\mathbf{S}$ in terms of the difference between a circulant matrix and a low-rank update, we use (5.22) to directly compute the eigenvalues of the circulant component via a direct Fourier decomposition. This allows us to show that the kink occurs when there is an significant change in the wavenumber corresponding to the largest eigenvalue of $\mathbf{S}$.

Lemma 5.6.1. *We define* $\mathbf{C}_1$ *as in Appendix 5.9. For* $\mathbf{H} = \mathbf{H}_1$ *we can bound the eigenvalues of* $\mathbf{S}$ *above by:*

$$\lambda_k(\mathbf{S}) \leq \lambda_k(\mathbf{C}_1) \tag{5.35}$$

where the eigenvalues of $\mathbf{C}_1$ *are given by:*

$$\gamma_m(\mathbf{C}_1) = \gamma_m(\mathbf{B}^{-1}) + \sum_{k=0}^{p-1} \omega^{mk} \mathbf{R}_{1,k}^{-1}, \quad m = 0, \ldots, N-1, \tag{5.36}$$

where $\omega = e^{-2\pi i/N}$*. Recall (using the notation introduced in Section 5.4.3) that the* $\gamma_j s$ *are ordered in terms of wavenumber rather than by decreasing eigenvalue.*

Proof. See Appendix 5.9. □

Lemma 5.6.1 yields an expression that is a sum of an eigenvalue of $\mathbf{B}^{-1}$, plus a term depending on the coefficients of $\mathbf{R}^{-1}$ and the structure of $\mathbf{H}_1$. The choice of $\mathbf{H} = \mathbf{H}_1$ is important in determining the wavenumber at which the maximum value of the second term of (5.36) is achieved. From Section 5.5.1 we recall that the largest eigenvalue of $\mathbf{B}^{-1}$ occurs for the pth wavenumber, $\gamma_{N/2}(\mathbf{B}^{-1})$, (for $N = 2p$) or $\gamma_{(N\pm1)/2}(\mathbf{B}^{-1})$ (for $N = 2p + 1$). The eigenvalues of $\mathbf{B}^{-1}$ ordered by wavenumber are shown by circles in Figure 5.5. The crosses in Figure 5.5 show the second term of (5.36) ordered by wavenumber. For $\mathbf{H}_1$, the largest value of the second term of (5.36) occurs for the same wavenumber as the largest eigenvalue of $\mathbf{B}^{-1}$. The maximum value of this term is equal to $\lambda_1(\mathbf{R}^{-1})$. This means that as $\lambda_1(\mathbf{S}_1)$ changes from being controlled by $\lambda_1(\mathbf{R}^{-1})$ to $\lambda_1(\mathbf{B}^{-1})$ the change appears smooth, as the wavenumber associated with the frequency of the largest eigenvalue remains constant. It is clear that increasing L_B will have a significant effect on the value of this bound, as changing L_B increases $\lambda_1(\mathbf{B}^{-1})$ significantly, and hence the upper bound given by (5.35). Therefore for both regimes, changing L_B has a large impact on both bounds for $\lambda_1(\mathbf{S})$.

We now present a similar bound for $\mathbf{H} = \mathbf{H}_2$.

Lemma 5.6.2. *For* $\mathbf{H} = \mathbf{H}_2$*, the eigenvalues of* $\mathbf{S}$ *are bounded above by:*

$$\lambda_k(\mathbf{S}) \leq \lambda_k(\mathbf{C}_2) \tag{5.37}$$

where the eigenvalues of $\mathbf{C}_2$ *are given by:*

$$\gamma_m(\mathbf{C}_2) = \gamma_m(\mathbf{B}^{-1}) + \sum_{k=0}^{p-1} \omega^{2mk} \mathbf{R}_{1,k}^{-1}, m = 1, 2, ..., p-1 \tag{5.38}$$

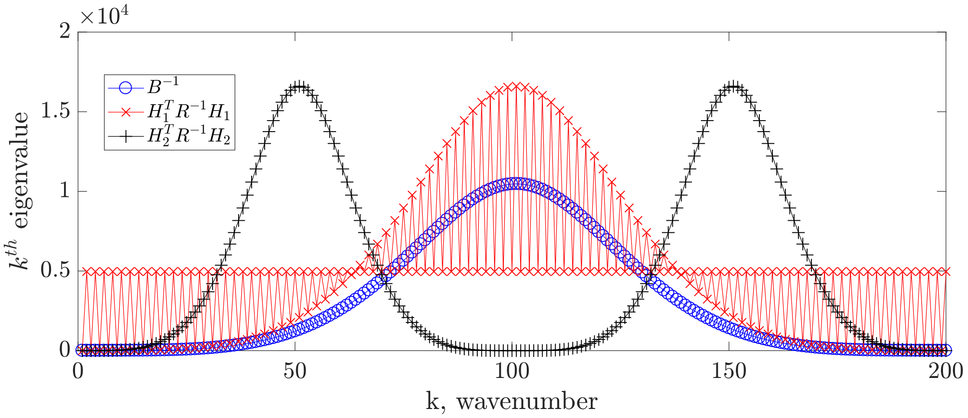


Figure 5.5: Plots of the contribution of the background and observation terms to the eigenvalues of the circulant matrix made up of the first row of $\mathbf{S}_1$ and $\mathbf{S}_2$ for $L_R = 0.7$ and $L_B = 0.3$. Circles denote the eigenvalues of $\mathbf{B}^{-1}$ (which is a term in both (5.36) and (5.38)), crosses denote the contribution of $\mathbf{R}^{-1}$ in the second term of (5.36) (i.e. for $\mathbf{H} = \mathbf{H}_1$) and pluses denotes the contribution of $\mathbf{R}^{-1}$ in the second term of (5.38) (i.e for $\mathbf{H} = \mathbf{H}_2$).

Recall (Section 5.4.3) that γ_js are ordered in terms of wavenumber rather than by maximum eigenvalue.

Lemma 5.6.2 also yields an upper bound that is the sum of an eigenvalue of $\mathbf{B}^{-1}$ and a term depending on $\mathbf{R}^{-1}$ and the choice of $\mathbf{H}_2$. We note that the values of the second term of (5.38) take the same values as the second term of (5.36) but in a different order. These are shown by the pluses in Figure 5.5, where we see that in order of wavenumber j, the second term of (5.38) yields the spectrum of $\mathbf{R}^{-1}$ twice. The second term of (5.38) is maximised for $j = p/2$ and $j = 3p/2$. These are different wavenumbers to the value of $j = p$ which maximises the first term.

Hence, we can bound $\lambda_1(\mathbf{S})$ above by $\lambda_1(\mathbf{R}^{-1}) + \lambda_{N/4}(\mathbf{B}^{-1})$ when $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) + \lambda_N(\mathbf{B}^{-1}) > \lambda_1(\mathbf{B}^{-1})$. In this case, increasing L_B has a very small effect on the upper bound for $\lambda_1(\mathbf{S})$, as $\lambda_{N/4}(\mathbf{B}^{-1})$ does not change significantly with L_B. However, when $\lambda_1(\mathbf{B}^{-1}) > \lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) + \lambda_N(\mathbf{B}^{-1})$, small changes to L_B will have a larger impact on $\lambda_1(\mathbf{B}^{-1})$ for all choices of $\mathbf{H}$. Similar behaviour is observed for fixed L_B and changing L_R. This change in the wavenumber of the largest eigenvalue explains why the kink occurs in the case of $\mathbf{H}_2$.

Finally, we discuss why the kink occurs for different values of L_B and L_R for $\mathbf{H}_2$ and $\mathbf{H}_3$. We have shown that the kink occurs when $\lambda_1(\mathbf{B}^{-1})$ becomes larger than $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) + \lambda_N(\mathbf{B}^{-1})$. For all values of L_R, $\lambda_1(\mathbf{H}_2^T\mathbf{R}^{-1}\mathbf{H}_2) \gg \lambda_1(\mathbf{H}_3^T\mathbf{R}^{-1}\mathbf{H}_3)$. As the contribution of $\mathbf{B}^{-1}$ is not affected by the choice of observation operator, changing from $\mathbf{H}_2$ to $\mathbf{H}_3$ increases the value of L_R necessary for $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})$ to be greater than $\lambda_1(\mathbf{B}^{-1})$. Hence the kink is only visible (see Figure 5.3e) for $L_R \gg L_B$ for the choice $\mathbf{H} = \mathbf{H}_3$.

5.7 Conclusions

Data assimilation is an important technique for combining information from observations with model data for the purpose of state estimation. One application of this is in numerical weather prediction (NWP), where data assimilation is used to combine observations of the atmosphere with a numerical model, in order to obtain an accurate description of the current state of the atmosphere. In this case correct specification of the uncertainty of each term is needed to produce the best forecast. The introduction of correlated observation error terms at operational NWP centres motivates investigation into the influence of observation error covariance on the convergence of the data assimilation procedure. We emphasise that the results presented here are general, and are relevant for any application of variational data assimilation. Improved knowledge of the role of correlated observation error covariances will be of use in the context of engineering [Nakamura and Potthast, 2015], neuroscience [Nakamura and Potthast, 2015, Schiff, 2011] and ecology [Pinnington et al., 2016, 2017].

In this work we developed theoretical bounds on the condition number of the Hessian of the 3D-Var objective function, which can be studied as a bound on the speed of convergence of the minimisation. These bounds were then tested in a simple numerical framework. We found that

- The bounds separate the contributions of the (correlated) observation error, background error, and observation operator, allowing us to better understand the role played by each term. We note that Theorem 5.4.2 and Corollary 5.4.3 in particular are general bounds applying to any valid covariance matrices and any choice of observation network.
- Numerical experiments for simple linear choices of observation network revealed interaction between observation error and background error terms. This interaction was also demonstrated theoretically for any choice of observation network and error covariance matrices.

- The structure of the observation network was seen to be crucial for determining how the observation and background errors interact.
- Both bounds and experiments revealed that the minimum eigenvalue of the two error covariance matrices is important for determining the conditioning of the Hessian, as well as the number of iterations required for the convergence of a minimisation procedure. This agrees with the findings of Weston et al. [2014], where small minimum eigenvalues of the observation error covariance matrix caused convergence problems in an practical setting.
- The ratio of the variances was also shown to be influential, although this was not investigated in detail in this work. This was also seen in Haben [2011].

We emphasise that many of the theoretical results and conclusions presented in this work are general and apply to any valid choice of background and observation error covariance matrices, and any linear observation operator. In particular, although the theoretical results presented in this chapter focus on the 3D-Var problem, a natural extension to 4D-Var is obtained by replacing the observation operator $\mathbf{H}$ with the generalised 4D observation operator which incorporates dynamical model information [Haben, 2011]. It is therefore expected that the eigenvalues of this model will also be important for the conditioning of the Hessian in this framework.

The importance of the choice of observation operator was revealed by the numerical tests, both for the condition number of the Hessian, and in terms of interaction between observation and background error covariances. Even for two observation networks with identical theoretical bounds, very different behaviour was observed numerically. This was explained by the existence of underlying structures in the Hessian, induced by the structures of the constituent error covariance and observation operator matrices. Better understanding of these interactions will be important for predicting the response of operational systems to the introduction of correlated observation errors. This is particularly applicable in practical applications where diagnosed correlated observation error covariance matrices must be adapted prior to their use in order to ensure convergence of the minimisation of the objective function.

In the numerical experiments presented in this chapter in Section 5.6, the observation and background error covariance matrices were altered by changing the lengthscales of the underlying correlation functions. This approach is mainly applicable for spatial correlations, where correlation lengthscales have a physical interpretation. There is significant research investigating spatial correlations [Waller et al., 2016c,a, Cordoba et al., 2017, Waller et al., 2016b], but much current work concerns the practical implementation of interchannel correlations for satellite observations [Weston, 2011, Stewart, 2010, Weston et al., 2014, Bormann et al., 2016, Campbell et al., 2017, Stewart et al., 2014]. Although the theory presented in Section 5.4 applies directly to the case of interchannel correlations, it would be of interest to extend our numerical testing to the interchannel covariance case. In particular, practical experiments have revealed that the minimum eigenvalue of the observation error correlation matrix is important for the conditioning of the Hessian in the case of interchannel correlations [Weston, 2011, Weston et al., 2014], which coincides with the theoretical and experimental results presented in this work. This is of particular interest as the correlation structure used in Weston et al. [2014] is not circulant, and demonstrates that, even beyond the presented in this chapter, our qualitative conclusions provide useful insight.

An additional area of future interest is investigation into how the best choice of preconditioning changes with the introduction of correlated observation error. Bounds on conditioning for the preconditioned case could be found by extending the results presented here, using similar theoretical techniques to those used in this work. The numerical and theoretical results discussed in this chapter suggest that interactions between observation and background correlations are also likely in that framework. It is expected that understanding how the introduction of correlated observation error covariance affects the unpreconditioned 3DVar problem will provide insight for suitable preconditioning methods in the correlated setting. One question of particular interest is whether the use of the background error covariance term as a preconditioner, as is done for the Control Variable Transform (CVT) [Bannister, 2008], remains optimal. One example of an operational problem that is not preconditioned is the 1DVar used at the UK Met Office for quality control [Stewart et al., 2008b]; the conclusions from this chapter apply directly to that implementation. The application of these results to the UK Met Office system will be discussed in a future chapter.

5.8 Acknowledgements

This work is funded in part by the EPSRC Centre for Doctoral Training in Mathematics of Planet Earth, the NERC Flooding from Intense Rainfall programme (NE/K008900/1), the EPSRC DARE project (EP/P002331/1) and the NERC National Centre for Earth Observation.

5.9 Appendix: Proofs

In this section we present the proofs for Lemmas 5.6.1 and 5.6.2 (Section 5.6.6), in which we express $\mathbf{S}$ as the difference between a circulant matrix and a singular matrix in order to bound the eigenvalues of $\mathbf{S}$ above.

Proof of Lemma 5.6.1. We exploit the structure of $\mathbf{S}$ which arises from the choice of $\mathbf{H}$; entries from $\mathbf{R}^{-1}$ are only added to the top left $p \times p$ block of $\mathbf{B}^{-1}$. Let $\mathbf{C}_1$ be the circulant matrix generated by the first row of $\mathbf{S}_1$. Then for $i = 1, \dots, N$

$$\mathbf{C}_1(1,i) = \mathbf{B}^{-1}(1,i) + (\mathbf{H}_1^T\mathbf{R}^{-1}\mathbf{H}_1)(1,i) = \begin{cases} \mathbf{B}^{-1}(1,i) + \mathbf{R}^{-1}(1,i) \text{ for } i = 1, \dots, p \\ \mathbf{B}^{-1}(1,i) \text{ for } i = p+1, \dots, N. \end{cases} \tag{5.39}$$

Let $\widetilde{\mathbf{H}}_1$ be given by

$$\widetilde{\mathbf{H}}_1(i,j) = \begin{cases} 1 \text{ for } j = i, i = p+1, \dots, N \\ 0 \text{ otherwise.} \end{cases} \tag{5.40}$$

Then we can write $\mathbf{S}_1 = \mathbf{C}_1 - \widetilde{\mathbf{H}}_1^T\mathbf{R}^{-1}\widetilde{\mathbf{H}}_1$. Applying (5.6) we obtain

$$\lambda_k(\mathbf{S}) \le \lambda_k(\mathbf{C}_1) + \lambda_1(-\widetilde{\mathbf{H}}_1\mathbf{R}^{-1}\widetilde{\mathbf{H}}_1). \tag{5.41}$$

As $\widetilde{\mathbf{H}}_1^T\mathbf{R}^{-1}\widetilde{\mathbf{H}}_1$ is not full rank and is positive semidefinite, its smallest eigenvalue is 0. Hence $\lambda_1(-\widetilde{\mathbf{H}}_1^T\mathbf{R}^{-1}\widetilde{\mathbf{H}}_1) = -\lambda_N(\widetilde{\mathbf{H}}_1^T\mathbf{R}^{-1}\widetilde{\mathbf{H}}_1) = 0$ and we have that

$$\lambda_k(\mathbf{S}_1) \le \lambda_k(\mathbf{C}_1). \tag{5.42}$$

As $\mathbf{C}_1$ is circulant, we calculate its eigenvalues via a direct Fourier transform (5.22). In order of wavenumber, the eigenvalues of $\mathbf{C}_1$ are given by

$$\gamma_m(\mathbf{C}_1) = \sum_{k=0}^{p-1} \omega^{mk}(\mathbf{B}_{1,k}^{-1} + \mathbf{R}_{1,k}^{-1}) + \sum_{k=p}^{N-1} \omega^{km}\mathbf{B}_{1,k}^{-1} \quad m = 0, \dots, N \tag{5.43}$$

where $\omega = e^{2\pi i/N}$ is an N-th root of unity. Separating the contributions of $\mathbf{B}^{-1}$ and $\mathbf{R}^{-1}$ yields:

$$\gamma_m(\mathbf{C}_1) = \sum_{k=0}^{N-1} \omega^{mk}\mathbf{B}_{1,k}^{-1} + \sum_{k=0}^{p-1} \omega^{mk}\mathbf{R}_{1,k}^{-1}. \tag{5.44}$$

□

Proof of Lemma 5.6.2. Follows the same arguments as the proof for Lemma 5.6.1

above. □

5.10 Summary

In this chapter we developed general bounds on the condition number of the Hessian of the unpreconditioned variational data assimilation problem. The minimum eigenvalue of the observation error covariance matrix appeared in the denominator of both bounds, meaning that small eigenvalues are likely to yield ill-conditioned Hessians. Numerical experiments revealed cases where both the upper and lower bounds were tight. Both the background and observation error covariance matrix dominated the conditioning of the Hessian for different parameter choices. Notably, the choice of observation operator was important in determining whether the transition between these regimes was smooth or not. We found that the condition number of the Hessian represents convergence of the conjugate gradient method well for many examples. However, in some cases repeated eigenvalues led to faster convergence than could be expected by considering the condition number of the Hessian alone.

The importance of the minimum eigenvalue of the observation error covariance matrix motivates the study of reconditioning methods in Chapter 7. Other key terms in the bounds, such as the ratio of the background and observation variance, will be shown to be important for the conditioning and convergence of an unpreconditioned nonlinear data assimilation problem in Chapter 8. In the next chapter, we examine how the bounds and conclusions for the unpreconditioned problem alter when we consider the preconditioned data assimilation problem.

Chapter 6

Conditioning of the preconditioned variational data assimilation problem

In this chapter we address RQ 2 from Chapter 1 and study how the effects of introducing correlated observation error differ in the preconditioned variational assimilation problem compared to the unpreconditioned problem. We wish to know:

- How does the importance of the background and observation terms differ in the preconditioned case?
- Does the behaviour of the condition number of the Hessian represent convergence of the conjugate gradient method well for numerical experiments?

6.1 Abstract

Data assimilation combines prior and observation information, weighted by their respective uncertainties, to obtain the most likely initial state of a dynamical system. In the variational data assimilation framework, a very high dimensional nonlinear least squares problem is solved iteratively. Until recently, all numerical weather prediction centres used diagonal error covariance matrices for all observation types. The increasing use of full observation weighting matrices motivates theoretical study regarding how introducing correlated observation error covariance matrices affects convergence of the data assimilation routine. Previous work showed that the minimum eigenvalue of the observation error covariance matrix was important for the conditioning and convergence of the unpreconditioned data assimilation problem. In this chapter we study how the conditioning of the preconditioned data assimilation

problem is affected by the introduction of correlated observation error for the first time. The minimum eigenvalue of the observation error covariance matrix also appears in upper and lower bounds on the Hessian of the preconditioned objective function. Numerical experiments reveal that it is harder to separate the effects of changing each matrix than in the unpreconditioned cases. However, by considering bounds that exploit different matrix norms it is possible to obtain good estimates of how changes to parameters affect the conditioning of the Hessian. We find cases where the eigenstructures of the constituent matrices lead to much faster convergence of a conjugate gradient method than could be expected by just considering the conditioning of the Hessian. Practical implementations of data assimilation algorithms often require that estimated covariance matrices are modified prior to use to reduce sampling noise. The theory in this chapter can be used to select modifications that may lead to faster convergence.

6.2 Introduction

Data assimilation is the process by which observations of a dynamical system are combined with information from a model of the system to find the most likely state of the system at a given time. In the variational data assimilation framework, both the observation and prior, or background, terms are weighted by their respective uncertainties. The most mature application of data assimilation is to numerical weather prediction (NWP), where it is used to obtain the best estimate, or analysis, of the initial condition used to produce forecasts (see e.g. Carrassi et al. [2018], Daley [1991], Kalnay [2002]). However, data assimilation methods can be used for any dynamical system with observations, such as in ecology (e.g. Pinnington et al. [2016, 2017]), hydrology (e.g. Cooper et al. [2018]) and neuroscience (e.g. Nakamura and Potthast [2015], Schiff [2011]).

The variational data assimilation method finds the analysis, or most likely state of the dynamical system, by minimising a nonlinear least squares objective function using an iterative method. The commonly-used incremental formulation solves the data assimilation problem via a small number of nonlinear outer loops, and a larger number of inner loops which solve the linearised problem [Courtier et al., 1994]. This procedure has been shown to be equivalent to a Gauss-Newton method [Gratton et al., 2007, Lawless et al., 2005a,b]. As the inner loop is a linear least squares problem, the conjugate gradient method is often used for this minimisation [Fisher, 1998, Liu et al., 2018, Trémolet, 2007]. This method also has reasonable convergence and memory requirements for high dimensional problems [Chao and Chang, 1992].

One practical problem with implementing incremental 4D-Var is the cost of either explicitly forming the background error covariance, or evaluating matrix-vector products. This is partly due to the large size of this matrix; the number of state variables can be of the order of 10^9 [Carrassi et al., 2018]. This motivates the use of the control variable transform (CVT) to model the background error covariance matrix implicitly [Bannister, 2008]. The CVT uses the square root of the background error covariance matrix as a variable transform to obtain a modified objective function [Lewis et al., 2006, Sec 9.1] and can therefore be considered as a form of preconditioning. Statistically this equates to transforming the variables to states that are uncorrelated. This leads to the additional benefit that the Hessian of the objective function using the CVT is typically better conditioned than the Hessian corresponding to the unpreconditioned problem. We will refer to the incremental variational problem with the CVT as the preconditioned data assimilation problem for the remainder of this chapter.

The use of correlated observation error covariance matrices can bring benefit to applications [Stewart et al., 2013, Simonin et al., 2019], and indeed will be necessary to capture high resolution phenomena [Rainwater et al., 2015, Fowler et al., 2018]. However, the move from uncorrelated (diagonal) to correlated (full) covariance matrices, has been shown to cause problems with the convergence of the data assimilation procedure [Weston, 2011, Weston et al., 2014]. NWP is a high-dimensional application (typically 10^7 observations are assimilated every cycle [Carrassi et al., 2018]) and computational time is at a premium. In order to use correlated observation error covariance matrices in applications, methods must be developed to ensure computationally efficient implementation. Some recent examples of efficient methods include a diffusion-based covariance model where the inverse observation covariance operator is formulated directly [Guillet et al., 2019], and an alternative parallelisation scheme which groups together observations with mutually correlated errors for processing [Simonin et al., 2019].

It is well-known that convergence of a conjugate gradient method can be bounded by the condition number of the Hessian of the objective function [Gill et al., 1986, Golub and Van Loan, 1996, Haben, 2011]. This means that the condition number of the Hessian is often used as a proxy to study how changes to a data assimilation method are likely to affect convergence of the minimisation. The condition number of the

Hessian of an objective function is also of interest as it relates to the sensitivity of the problem to perturbations in the background or observations. We expect much faster convergence of the preconditioned problem than the unpreconditioned data assimilation formulation. However, it is important to note that there are circumstances, for example in the case of repeated eigenvalues, where the relationship between condition number and convergence of a conjugate gradient method is not so strong [Gill et al., 1986, Nocedal, 2006]. In this chapter we will therefore consider how the condition number of the Hessian relates to convergence of the conjugate gradient method in an idealised numerical framework.

Theoretical studies have considered how changes to the variational data assimilation problem will change its conditioning for the case of uncorrelated observation error covariance matrices and for the unpreconditioned problem. In the case of uncorrelated (diagonal) observation error covariance matrices, Haben et al. [2011b], Haben [2011] used the condition number of the Hessian as a proxy for convergence of the preconditioned variational data assimilation problem. The ratio between background and observation variances was found to play an important role in the conditioning of the Hessian, as well as the choice of observation network. For correlated (full) observation error covariance matrices, Tabeart et al. [2018] found cases where either observation or background error covariance matrices dominated conditioning of the unpreconditioned problem. It is expected that the role of each matrix will be more complicated for the preconditioned problem, where the background and observation error covariance matrices appear in the same product within the Hessian. Additionally, theoretical bounds and numerical results from Tabeart et al. [2018] showed that in the case of unpreconditioned variational assimilation, the minimum eigenvalue of the correlated observation error covariance matrix is important in determining the conditioning of the Hessian of the objective function. As estimated observation error covariance matrices often need to be modified prior to their use at NWP centres, e.g. by reconditioning [Weston, 2011, Tabeart et al., 2019a,b], understanding which properties of the observation error covariance matrix may lead to ill-conditioned data assimilation problems is of practical as well as theoretical interest.

In this chapter we consider the conditioning of the preconditioned variational data assimilation problem in the case of correlated observation error covariance matrices. We extend the analysis of Tabeart et al. [2018] to the preconditioned case. We begin in Section 6.3 by defining the problem and introducing the notation of data assimilation, alongside existing mathematical results relating to conditioning. In

Section 6.4 we present theoretical bounds on the condition number of the preconditioned Hessian in terms of its constituent matrices. We introduce the numerical framework that will be used for our experiments in Section 6.5, before presenting the results of these experiments and related discussion in Section 6.6. We find that increasing the correlation lengthscale of the background error covariance matrix, or decreasing the correlation lengthscale of the observation error covariance matrix reduces the condition number of the Hessian. We find cases where our bounds represent the qualitative behaviour of the conditioning well, as well as cases where other bounds from Haben [2011] are tighter. The effect of changing experimental parameters on the convergence of a conjugate gradient method is also studied. Finally, we present our conclusions in Section 6.7. The theoretical and numerical conclusions from this chapter will help users understand how changes to individual components of a data assimilation system are likely to affect its conditioning and convergence. One example of this is the use of reconditioning methods for estimated observation error covariance matrices [Weston, 2011, Weston et al., 2014, Campbell et al., 2017], where the results in this work may allow users to select a method or parameter value that will permit improved computational efficiency.

6.3 The preconditioned variational data assimilation problem

6.3.1 The control variable transform formulation of the data assimilation problem

We now define the preconditioned 4D-Var data assimilation problem and introduce the notation that will be used in this chapter. We remark that in the 4D-Var setting observations can be made at multiple times over a pre-defined observation window.

Let our time window begin at time t_0 and end at time t_n, with n being the total number of timesteps. Let our state have dimension N. We define the background, or prior, at time t_0 as $\mathbf{x}_b \in \mathbb{R}^N$, with corresponding background error covariance matrix $\mathbf{B} \in \mathbb{R}^{N\times N}$. Observations $y_i \in \mathbb{R}^{p_i}$ occur at time t_i, with corresponding observation error covariance matrix $\mathbf{R}_i \in \mathbb{R}^{p_i\times p_i}$. The total number of observations across the whole time window is given by $Q = \sum_{i=0}^{n} p_i$. In order to compare observations with state variables, we define an observation operator $h_i : \mathbb{R}^N \to \mathbb{R}^{p_i}$ which maps from state variable space to observation space at time t_i.

As $\mathbf{R}_i$ and $\mathbf{B}$ are covariance matrices, they are symmetric and positive semi-definite by definition. We additionally assume that all these matrices are strictly positive definite, and that observation and background errors are unbiased and mutually uncorrelated.

In order to compare an observation at time t_i with the state variable at time t_i, we propagate the state from the previous time using a nonlinear forecast model, $\mathcal{M}$ to obtain

$$\mathbf{x}_i = \mathcal{M}(t_{i-1}, t_i, \mathbf{x}_{i-1}). \tag{6.1}$$

This yields the full 4D-Var objective function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - h_i[\mathbf{x}_i])^T\mathbf{R}_i^{-1}(\mathbf{y}_i - h_i[\mathbf{x}_i]). \tag{6.2}$$

The state $\mathbf{x}_a$ which minimises (6.2) is the analysis. If $n = 0$ then (6.2) simplifies immediately to the 3D-Var objective function.

Let $\mathbf{x}_i^b = \mathcal{M}(t_{i-1}, t_i; \mathbf{x}_{i-1}^b)$. Define $\delta\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_i^b$. We then consider the Taylor expansion of $\mathcal{M}(t_{i-1}, t_i; \mathbf{x}_{i-1})$ about $\mathbf{x}_i^b(t)$

$$\mathbf{x}_i^b + \delta\mathbf{x}_i = \mathcal{M}(t_{i-1}, t_i; \mathbf{x}_{i-1}^b) + \mathbf{M}_i\delta\mathbf{x}_{i-1} + \text{higher order terms} \tag{6.3}$$

$$\delta\mathbf{x}_i \approx \mathbf{M}_i\delta\mathbf{x}_{i-1} \tag{6.4}$$

where $\mathbf{M}_i \in \mathbb{R}^{N\times N}$ is the linearised model operator at time t_i, linearised about $\mathbf{x}_i^b$. Similarly, expanding $h_i[\mathbf{x}_i]$ about $\mathbf{x}_i^b$ we obtain

$$h_i[\mathbf{x}_i] \approx h_i[\mathbf{x}_i^b] + \mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}_i \tag{6.5}$$

where $\mathbf{H}_i \in \mathbb{R}^{N\times p_i}$ is the linearised observation operator at time t_i linearised around $\mathbf{x}_i^b$.

We then write the linearised objective function in terms of $\delta\mathbf{x}_0$

$$J(\mathbf{x}_0) = \frac{1}{2}\delta\mathbf{x}_0^T\mathbf{B}^{-1}\delta\mathbf{x}_0 + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{d}_i - \mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}_0)^T\mathbf{R}_i^{-1}(\mathbf{d}_i - \mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}_0) \tag{6.6}$$

where $\mathbf{d}_i = \mathbf{y}_i - h_i[\mathbf{x}_i^b]$ are the innovation vectors. These measure the misfit between the observations and the linearisation state, using the full nonlinear observation operator.

We then define the generalised observation operator as

$$\widehat{\mathbf{H}} = \left[\mathbf{H}_0^T, (\mathbf{H}_1\widehat{\mathbf{M}}_1)^T, \ldots, (\mathbf{H}_n\widehat{\mathbf{M}}_n)^T\right]^T, \tag{6.7}$$

where the linearised forward model from time t_0 to time t_i is given by

$$\widehat{\mathbf{M}}_i(\delta\mathbf{x}_i) = \mathbf{M}(t_i, t_0; \delta\mathbf{x}_{i-1}) = \mathbf{M}_i \ldots \mathbf{M}_1. \tag{6.8}$$

Finally let $\widehat{\mathbf{R}} \in \mathbb{R}^{Q\times Q}$ denote the block diagonal matrix with the ith block consisting of $\mathbf{R}_i$:

$$\delta\mathbf{x}_i = \mathbf{M}(t_{i-1}, t_i; \mathbf{x}_{i-1})\delta\mathbf{x}_{i-1} \equiv \mathbf{M}_i\delta\mathbf{x}_{i-1}. \tag{6.9}$$

This allows us to write the Hessian of the linearised objective function, (6.6), in the simplified form

$$\mathbf{S} = \mathbf{B}^{-1} + \widehat{\mathbf{H}}^T\widehat{\mathbf{R}}^{-1}\widehat{\mathbf{H}}. \tag{6.10}$$

The formulation of the objective function given by (6.2) is too expensive to be used in practice both in terms of computation, but also storage. The number of state variables, N, is very large and typically $\mathbf{B}$ cannot be stored explicitly. One method used to combat this problem is preconditioning, where a mathematically equivalent but less expensive formulation is used. The most common preconditioner for the data assimilation problem in NWP applications makes use of the control variable transform (CVT) which is described in detail in Bannister [2008, 2017]. The CVT formulates the objective function in terms of alternative 'control variables', which means that the background matrix $\mathbf{B}$ does not need to be stored explicitly.

The control variable transform (CVT) is then applied to the incremental form of the variational problem (6.6), via the change of variable $\delta\mathbf{z}_0 = \mathbf{B}^{-1/2}\delta\mathbf{x}_0$. This yields the objective function

$$J(\delta\mathbf{z}_0) = \frac{1}{2}\delta\mathbf{z}_0^T\delta\mathbf{z}_0 + \frac{1}{2}\left(\widehat{\mathbf{d}} - \widehat{\mathbf{H}}\mathbf{B}^{1/2}\delta\mathbf{z}_0\right)^T\widehat{\mathbf{R}}^{-1}\left(\widehat{\mathbf{d}} - \widehat{\mathbf{H}}\mathbf{B}^{1/2}\delta\mathbf{z}_0\right). \tag{6.11}$$

where $\mathbf{z}_0 = \mathbf{B}^{-1/2}\mathbf{x}_0$, $\mathbf{z}_b = \mathbf{B}^{-1/2}\mathbf{x}_b$, and

$$\widehat{\mathbf{d}}^T = \left[\mathbf{d}_o^T, \mathbf{d}_1^T, \ldots, \mathbf{d}_n^T\right] \tag{6.12}$$

is a vector made up of the innovation vectors.

It can be shown [Haben et al., 2011b] that use of the CVT is equivalent to pre- and post-multiplying the Hessian of the unpreconditioned data assimilation problem (6.10) by $\mathbf{B}^{1/2}$ (the uniquely defined, symmetric square root of $\mathbf{B}$). This yields a preconditioned Hessian for 4D-Var given by

$$\widehat{\mathbf{S}} = \mathbf{I}_N + \mathbf{B}^{1/2}\widehat{\mathbf{H}}^T\widehat{\mathbf{R}}^{-1}\widehat{\mathbf{H}}\mathbf{B}^{1/2}. \tag{6.13}$$

We note that with the additional assumption that $\mathbf{B}$ and $\mathbf{R}$ are strictly positive definite, then $\widehat{\mathbf{S}}$ is symmetric positive definite.

The preconditioned Hessian (6.13) highlights the computational benefit of using the CVT. As there are fewer observations than state variables (typically difference of two orders of magnitude [Carrassi et al., 2018]), the second term in (6.13) is rank deficient. This means that the preconditioned Hessian is a low-rank update to the identity, and hence its minimum eigenvalue is 1. Therefore, the preconditioned Hessian will not suffer from small minimum eigenvalues that often result in ill-conditioning for the unpreconditioned problem. This improved conditioning is expected to lead to faster convergence of the associated data assimilation algorithm.

In this chapter we consider the conditioning of the Hessian of the CVT objective function (6.11) as a proxy for convergence of the preconditioned data assimilation problem. We will develop bounds on the condition number of (6.13) in terms of its constituent matrices and investigate the importance of correlated observation error covariance matrices for the preconditioned problem. For what follows we will write $\widehat{\mathbf{R}} \equiv \mathbf{R}$ and $\widehat{\mathbf{H}} \equiv \mathbf{H}$ in order to simplify notation.

6.3.2 Eigenvalue theory

For the remainder of this manuscript, we use the following order of eigenvalues: For a matrix $\mathbf{C} \in \mathbb{R}^{d\times d}$ let $\lambda_{max}(\mathbf{C}) = \lambda_1(\mathbf{C}) \geq \lambda_2(\mathbf{C}) \geq \cdots \geq \lambda_d(\mathbf{C}) = \lambda_{min}(\mathbf{C})$.

In this section we introduce theoretical results from linear algebra. These will be used in Section 6.4 to develop bounds on the condition number of the preconditioned Hessian in terms of its constituent matrices. We also present existing bounds on the Hessian of the preconditioned 4D-Var problem, which we will compare with the new bounds developed in Section 6.4 in the numerical experiments of Section 6.6.

We begin by formally defining the condition number.

Definition 6.3.1. *[Golub and Van Loan, 1996, Sec. 2.7.2] Let $\mathbf{C} \in \mathbb{R}^{d\times d}$ be a symmetric positive definite matrix. Then we characterise the condition number in the 2−norm of $\mathbf{C}$ as*

$$\kappa(\mathbf{C}) = \frac{\lambda_1(\mathbf{C})}{\lambda_d(\mathbf{C})} \tag{6.14}$$

This shall be referred to as the condition number for the remainder of this work.

Since $\widehat{\mathbf{S}}$ is symmetric positive definite we apply the characterisation of the condition number given by Definition 6.3.1 throughout this chapter.

The following result allows us to exchange the order of multiplication when computing the eigenvalues of a matrix product.

Theorem 6.3.2. *Let $\mathbf{F} \in \mathbb{R}^{m\times n}$ and $\mathbf{G} \in \mathbb{R}^{n\times m}$. Then, the nonzero distinct eigenvalues of $\mathbf{GF}$ are the same as those of $\mathbf{FG}$.*

Proof. Harville [1997, Theorem 21.10.1] □

We now present several results which bound the eigenvalues of matrix sums and products in terms of the eigenvalues of the individual matrices. These will be used in Section 6.4 to separate the contribution of the background and observation error covariance matrices to $\kappa(\widehat{\mathbf{S}})$. We begin by considering the eigenvalues of a matrix sum.

Theorem 6.3.3. *Consider two symmetric matrices $\mathbf{C}_1$, $\mathbf{C}_2 \in \mathbb{R}^{d\times d}$. The k^{th} eigenvalue of the matrix sum $\mathbf{C}_1 + \mathbf{C}_2$ satisfies*

$$\lambda_k(\mathbf{C}_1) + \lambda_d(\mathbf{C}_2) \leq \lambda_k(\mathbf{C}_1 + \mathbf{C}_2) \leq \lambda_k(\mathbf{C}_1) + \lambda_1(\mathbf{C}_2). \tag{6.15}$$

Proof. See [Wilkinson, 1965, Ch. 2 Theorem 44]. □

We now present three results which bound the eigenvalues of a matrix product in terms of the product of the eigenvalues of the individual matrices.

Theorem 6.3.4. *If $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{d\times d}$ are positive semi-definite Hermitian matrices, then*

$$\prod_{i=1}^{k} \lambda_i(\mathbf{FG}) \leq \prod_{i=1}^{k} \lambda_i(\mathbf{F})\lambda_i(\mathbf{G}), \quad k = 1, \ldots, d-1. \tag{6.16}$$

Proof. See [Marshall et al., 2011, Sec. 9 H.1.a.]. □

Theorem 6.3.5. *If* $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{d\times d}$ *are positive semi-definite Hermitian and* $1 \leq i_1 < \cdots < i_k \leq d$*, then*

$$\prod_{t=1}^{k} \lambda_t(\mathbf{FG}) \geq \prod_{t=1}^{k} \lambda_{i_t}(\mathbf{F})\lambda_{d-i_t+1}(\mathbf{G}), \tag{6.17}$$

with equality for $k = d$*.*

Proof. See Wang and Zhang [1992, Theorem 2]. □

Theorem 6.3.6. *If* $\mathbf{F}, \mathbf{G} \in \mathbb{C}^{d\times d}$ *are positive semidefinite Hermitian and* $1 \leq i_1 < \cdots < i_k \leq d$*, then*

$$\sum_{t=1}^{k} \lambda_{i_t}(\mathbf{FG}) \geq \sum_{t=1}^{k} \lambda_{i_t}(\mathbf{F})\lambda_{d-t+1}(\mathbf{G}). \tag{6.18}$$

Proof. See Wang and Zhang [1992, Theorem 4]. □

These results will be used to develop bounds on the condition number of the Hessian (6.13).

We first present an existing bound on $\kappa(\widehat{\mathbf{S}})$ from Haben [2011]. We note that this bound was developed for the 3D-Var case, although it extends naturally to the 4D-Var problem by replacing $\mathbf{R}$ with $\widehat{\mathbf{R}}$ and $\mathbf{H}$ with $\widehat{\mathbf{H}}$.

Theorem 6.3.7. *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* $\widehat{\mathbf{S}} = \mathbf{I}_N + \mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2}$

$$1 + \frac{1}{p}\sum_{i,j=1}^{p}\left(\mathbf{R}^{-1/2}\mathbf{HBH}^T\mathbf{R}^{-1/2}\right)_{i,j} \leq \kappa(\widehat{\mathbf{S}}) \leq 1 + \left\|\mathbf{R}^{-1/2}\mathbf{HBH}^T\mathbf{R}^{-1/2}\right\|_{\infty}. \tag{6.19}$$

Proof. See Haben [2011, Theorem 6.2.1] □

In this chapter, we want to develop bounds that separate the contribution of each constituent matrix. The bounds given by (6.19) do not separate out each term, meaning that they are likely to be tighter. In Section 6.6 we will numerically compare the bounds given by (6.19) with those given developed in Section 6.4.

6.4 Theoretical bounds on the Hessian of the preconditioned problem

In this section we develop new theoretical bounds on the condition number of the Hessian of the preconditioned variational data assimilation problem, following similar methods to the unpreconditioned case in Tabeart et al. [2018]. We begin by defining the problem of interest studied in this chapter.

Principal Theoretical Assumptions. *Let* $\mathbf{B} \in \mathbb{R}^{N\times N}$ *be the background error covariance matrix. For a given time window* $[t_0, t_n]$*, let observations* $\mathbf{y}_i \in \mathbf{R}^{p_i}$ *be made at times* t_i*, for* $0 \leq i \leq n$*. Let* $\mathbf{R}_i \in \mathbf{R}^{p_i\times p_i}$ *be the observation error covariance matrix corresponding to observations at time* t_i*. We define* $\mathbf{R} \in \mathbf{R}^{Q\times Q}$ *as the block diagonal matrix with* $\mathbf{R}_i$ *on the diagonal, where* $Q = \sum_{i=0}^{n} p_i < N$*. Let* $\mathbf{H} \in \mathbb{R}^{Q\times N}$ *be the generalised linearised observation operator given by* (6.7)*.*

The first result shows that the condition number can be calculated via the eigenvalues of the rank-p update.

Lemma 6.4.1. *Following the Principal Theoretical Assumptions we can express the condition number of* $\widehat{\mathbf{S}}$ *as*

$$\begin{aligned} \kappa(\widehat{\mathbf{S}}) &= 1 + \lambda_1(\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) && (6.20)\\ &= 1 + \lambda_1(\mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^T). && (6.21)\end{aligned}$$

Proof. We begin by showing that $\kappa(\widehat{\mathbf{S}}) = 1 + \lambda_1(\mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2})$, as was presented in [Haben, 2011, Equation (4.2)]. We define $\mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2} = \mathbf{C}$ and write $\widehat{\mathbf{S}} = \mathbf{I} + \mathbf{C}$. Let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N$ be the eigenvalues of $\mathbf{C}$, with corresponding eigenvectors v_i. As $Q < N$, $\mathbf{C}$ is rank-deficient and therefore $\sigma_N = 0$.
Then, we can calculate the eigenvalues of $\widehat{\mathbf{S}}$ via

$$\begin{aligned} \lambda_i(\widehat{\mathbf{S}})v_i &= (\mathbf{I} + \mathbf{C})v_i \\ &= v_i + \mathbf{C}v_i \qquad (6.22)\\ &= (1+\sigma_i)v_i. \end{aligned}$$

As $\sigma_N = 0$, we find that $\lambda_N(\widehat{\mathbf{S}}) = 1$. This means that $\kappa(\widehat{\mathbf{S}}) = \lambda_1(\widehat{\mathbf{S}}) = 1 + \sigma_1(\mathbf{C})$. We now use the result of Theorem 6.3.2 to exchange the order of multiplication. In order to obtain (6.20) let $\mathbf{G} = \mathbf{B}^{1/2}$ and $\mathbf{F} = \mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$. To obtain (6.21), we define $\mathbf{F} = \mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2}$ and $\mathbf{G} = \mathbf{B}^{1/2}\mathbf{H}^T$. □

The result of Lemma 6.4.1 shows that computing $\kappa(\widehat{\mathbf{S}})$ only requires the computation of the maximum eigenvalue of a single matrix product. We also note that the matrix products that appear in (6.20) and (6.21) are of different dimensions: $\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \in \mathbb{R}^{N\times N}$ and $\mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^T \in \mathbb{R}^{Q\times Q}$. Additionally, the first matrix product is rank deficient, whereas the second matrix product is full rank.

We now develop bounds on the condition number of $\widehat{\mathbf{S}}$ in terms of its constituent matrices.

6.4.1 General bounds on the condition number

Theorem 6.4.2. *Given the Principal Theoretical Assumptions we can bound* $\widehat{\mathbf{S}} = \boldsymbol{I}_N + \mathbf{B}^{1/2}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{1/2}$ *by*

$$\begin{aligned}&1+\max\left\{\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\lambda_N(\mathbf{B}),\quad \frac{\lambda_1(\mathbf{H}\mathbf{B}\mathbf{H}^T)}{\lambda_1(\mathbf{R})},\quad \frac{\lambda_Q(\mathbf{H}\mathbf{B}\mathbf{H}^T)}{\lambda_Q(\mathbf{R})}\right\}\\ &\le \kappa(\widehat{\mathbf{S}}) \le 1+\min\{\lambda_1(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}),\quad \frac{\lambda_1(\mathbf{H}\mathbf{B}\mathbf{H}^T)}{\lambda_Q(\mathbf{R})}\}.\end{aligned} \tag{6.23}$$

Proof. We write $\kappa(\widehat{\mathbf{S}})$ as in the statement of Lemma 6.4.1. To obtain the upper bound of (6.23), we use the result of Theorem 6.3.4 to separate the contribution of the background and observation term

$$\kappa(\widehat{\mathbf{S}}) = 1+\lambda_1(\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) \le 1+\lambda_1(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) \tag{6.24}$$

and similarly for the alternative formulation

$$\begin{aligned}\kappa(\widehat{\mathbf{S}}) &= 1+\lambda_1(\mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^T)\\ &\le 1+\lambda_1(\mathbf{R}^{-1})\lambda_1(\mathbf{H}\mathbf{B}\mathbf{H}^T) = 1+\frac{1}{\lambda_Q(\mathbf{R})}\lambda_1(\mathbf{H}\mathbf{B}\mathbf{H}^T).\end{aligned} \tag{6.25}$$

Combining these two expressions yields the upper bound in the theorem statement. To compute the lower bound of (6.23), we apply the result of Theorem 6.3.5 to (6.20). This yields

$$\begin{aligned}&\lambda_1(\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) \ge \max\left\{\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\lambda_N(\mathbf{B}),\quad \lambda_N(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\lambda_1(\mathbf{B})\right\}\\ &\lambda_1(\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) \ge \lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\lambda_N(\mathbf{B}).\end{aligned} \tag{6.26}$$

This last inequality is due to the fact that $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ is rank-deficient. We now apply the result of Theorem 6.3.5 to (6.21)

$$\begin{aligned}\lambda_1(\mathbf{R}^{-1}\mathbf{HBH}^T) &\geq \max\left\{\lambda_1(\mathbf{HBH}^T)\lambda_Q(\mathbf{R}^{-1}), \quad \lambda_Q(\mathbf{HBH}^T)\lambda_1(\mathbf{R}^{-1})\right\} \\ &\geq \max\left\{\frac{\lambda_1(\mathbf{HBH}^T)}{\lambda_1(\mathbf{R})}, \quad \frac{\lambda_Q(\mathbf{HBH}^T)}{\lambda_Q(\mathbf{R})}\right\}.\end{aligned} \tag{6.27}$$

Combining these bounds yields (6.23) as required.

□

We can separate the contribution of the observation error covariance matrix from the observation operator via the following result.

Corollary 6.4.3. *Under the same conditions as in Theorem 6.4.2, we can bound $\kappa(\widehat{\mathbf{S}})$ by*

$$1 + \max\left\{\frac{\lambda_Q(\mathbf{HH}^T)\lambda_N(\mathbf{B})}{\lambda_Q(\mathbf{R})}, \frac{\lambda_1(\mathbf{HH}^T)\lambda_N(\mathbf{B})}{\lambda_1(\mathbf{R})}\right\} \leq \kappa(\widehat{\mathbf{S}}) \leq 1 + \frac{\lambda_1(\mathbf{B})}{\lambda_Q(\mathbf{R})}\lambda_1(\mathbf{HH}^T) \tag{6.28}$$

Proof. We begin by considering the upper bound of (6.23). By Theorem 6.3.2 $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ has precisely the same nonzero eigenvalues as $\mathbf{R}^{-1}\mathbf{HH}^T$ and $\mathbf{HH}^T\mathbf{R}^{-1}$. Applying Theorem 6.3.4 for $k = 1$ to $\lambda_1(\mathbf{R}^{-1}\mathbf{HH}^T)$ yields:

$$\lambda_1(\mathbf{R}^{-1}\mathbf{HH}^T) \leq \lambda_1(\mathbf{R}^{-1})\lambda_1(\mathbf{HH}^T) = \frac{\lambda_1(\mathbf{HH}^T)}{\lambda_Q(\mathbf{R})}. \tag{6.29}$$

By Theorem 6.3.2 , $\mathbf{HBH}^T$ has precisely the same nonzero eigenvalues as $\mathbf{BH}^T\mathbf{H}$. Applying Theorem 6.3.4 for $k = 1$ yields:

$$\lambda_1(\mathbf{BH}^T\mathbf{H}) \leq \lambda_1(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{H}) = \lambda_1(\mathbf{B})\lambda_1(\mathbf{HH}^T). \tag{6.30}$$

The final equality arises as the nonzero eigenvalues of $\mathbf{HH}^T$ are equal to those of $\mathbf{H}^T\mathbf{H}$. Therefore the two cases from Theorem 6.4.2 yield the same 'factorised' upper bound, and gives the upper bound in (6.28).

We now consider the first term in the lower bound of (6.23) and bound $\lambda_1(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})$ below. We separate the contribution of $\mathbf{R}$ and $\mathbf{HH}^T$ using Theorem 6.3.5 for

$t = 1, i_1 = 1$. This yields

$$\lambda_1(\mathbf{R}^{-1}\mathbf{H}\mathbf{H}^T) \geq \max\left\{\lambda_Q(\mathbf{R}^{-1})\lambda_1(\mathbf{H}\mathbf{H}^T), \lambda_1(\mathbf{R}^{-1})\lambda_Q(\mathbf{H}\mathbf{H}^T)\right\} \tag{6.31}$$

$$\geq \max\left\{\frac{\lambda_1(\mathbf{H}\mathbf{H}^T)}{\lambda_1(\mathbf{R})}, \frac{\lambda_Q(\mathbf{H}\mathbf{H}^T)}{\lambda_Q(\mathbf{R})}\right\}. \tag{6.32}$$

Multiplying this by $\lambda_N(\mathbf{B})$ yields

$$\kappa(\widehat{\mathbf{S}}) \geq \max\left\{\frac{\lambda_1(\mathbf{H}\mathbf{H}^T)\lambda_N(\mathbf{B})}{\lambda_1(\mathbf{R})}, \frac{\lambda_Q(\mathbf{H}\mathbf{H}^T)\lambda_N(\mathbf{B})}{\lambda_Q(\mathbf{R})}\right\}. \tag{6.33}$$

This yields the two terms that appear in the lower bound of (6.28).

We now consider the second term of (6.23) and bound $\lambda_1(\mathbf{H}\mathbf{B}\mathbf{H}^T)$ below. We separate the contribution of $\mathbf{B}$ and $\mathbf{H}^T\mathbf{H}$ using Theorem 6.3.5 for $t = 1, k = 1, i_1 = 1, d = N$. This yields

$$\begin{aligned}\lambda_1(\mathbf{B}\mathbf{H}^T\mathbf{H}) &\geq \max\left\{\lambda_1(\mathbf{B})\lambda_N(\mathbf{H}^T\mathbf{H}), \lambda_N(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{H})\right\}\\ &\geq \lambda_N(\mathbf{B})\lambda_1(\mathbf{H}^T\mathbf{H}).\end{aligned} \tag{6.34}$$

The last inequality follows as $\mathbf{H}^T\mathbf{H}$ is not full rank and therefore $\lambda_N(\mathbf{H}^T\mathbf{H}) = 0$. Multiplying this result by $1/\lambda_1(\mathbf{R})$ gives the same value as the first term in (6.33).

Finally, we bound the third term of the lower bound in (6.23). As $\lambda_Q(\mathbf{H}\mathbf{B}\mathbf{H}^T) = \lambda_Q(\mathbf{B}\mathbf{H}^T\mathbf{H})$ by Theorem 6.3.2, we separate the contribution of $\mathbf{B}$ and $\mathbf{H}^T\mathbf{H}$ using Theorem 6.3.6 for $t = 1, k = 1, i_1 = Q, d = N$.

$$\begin{aligned}\lambda_Q(\mathbf{B}\mathbf{H}^T\mathbf{H}) &\geq \max\{\lambda_Q(\mathbf{B})\lambda_N(\mathbf{H}^T\mathbf{H}), \lambda_N(\mathbf{B})\lambda_Q(\mathbf{H}^T\mathbf{H})\}\\ &\geq \lambda_N(\mathbf{B})\lambda_Q(\mathbf{H}^T\mathbf{H}).\end{aligned} \tag{6.35}$$

Multiplying the second term of (6.35) by $1/\lambda_Q(\mathbf{R})$ gives the second term in (6.33), as $\lambda_Q(\mathbf{H}^T\mathbf{H}) = \lambda_Q(\mathbf{H}\mathbf{H}^T)$.

□

In general it is not possible to determine which term in the lower bound of (6.28) is larger, as this will depend on the choice of $\mathbf{B}, \mathbf{H}$ and $\mathbf{R}$. However, we are able to comment on how the bounds are likely to be altered by changes to individual matrices. As we increase $\lambda_Q(\mathbf{R})$ both the upper bound and first term in the lower bound decrease. Similarly, as $\lambda_1(\mathbf{H}\mathbf{H}^T)$ appears in the upper bound and second term of the lower bound, increases to this term will result in increases to both bounds. As smaller eigenvalues of $\mathbf{B}$ increase, the lower bound will increase but the upper bound will remain unchanged. Finally, as $\lambda_1(\mathbf{R})$ increases, the second term of the lower bound will decrease. In the experiments in Section 6.6 we will study how each of these terms changes with interacting parameters, and assess which lower bound is tighter for a variety of situations.

6.4.2 Bounds on the condition number in the case of circulant error covariance matrices

The theoretical bounds presented in Section 6.4.1 apply for any choice of observation and background error covariance matrices. However, for a given numerical framework, general bounds can typically be improved by exploiting specific structure of the matrices being used [Haben, 2011]. In this section we will show that under additional assumptions on the structure of the error covariance matrices and observation operator, the bounds given by (6.19) yield the exact value of $\kappa(\widehat{\mathbf{S}})$.

We begin by defining circulant matrices, which arise for spatial correlations for homogeneous, evenly distributed variables. We will make use of this structure in the numerical experiments in Section 6.6.

Definition 6.4.4 (Davis [1979])**.** *A circulant matrix $\mathbf{D} \in \mathbb{R}^{N\times N}$ is a matrix of the form*

$$\mathbf{D} = \begin{pmatrix} d_0 & d_1 & d_2 & \cdots & d_{N-2} & d_{N-1} \\ d_{N-1} & d_0 & d_1 & \cdots & d_{N-3} & d_{N-2} \\ d_{N-2} & d_{N-1} & d_0 & \cdots & d_{N-4} & d_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ d_2 & d_3 & d_4 & \cdots & d_0 & d_1 \\ d_1 & d_2 & d_3 & \cdots & d_{N-1} & d_0 \end{pmatrix}.$$

Circulant correlation matrices are computationally desirable to use as their eigenvalues can be calculated via a discrete Fourier transform.

Theorem 6.4.5. *The eigenvalues of a circulant matrix $\mathbf{D} \in \mathbb{R}^{N\times N}$, as given by Definition 5.4.6, are given by*

$$\gamma_m = \sum_{k=0}^{N-1} d_k \omega^{mk}, \tag{6.36}$$

with corresponding eigenvectors

$$\mathbf{v}_m = \frac{1}{\sqrt{N}}(1, \omega^m, \cdots, \omega^{m(N-1)}), \tag{6.37}$$

where $\omega = e^{-2\pi i/N}$ is an $N-$th root of unity.

Proof. See Gray [2006] for full derivation. □

We note that the result of Theorem 6.4.5 means that any circulant matrix of dimension N admits the same eigenvectors.

Our numerical experiments in Section 6.6 will use circulant background and observation error covariance matrices. In the case of circulant error covariance matrices, with some additional assumptions on the entries of matrix products, we can prove that the upper and lower bounds given by Theorem 6.3.7 are equal and yield the exact value of $\kappa(\widehat{\mathbf{S}})$.

Corollary 6.4.6. *If $\mathbf{HBH}^T$ and $\mathbf{R}$ are circulant matrices, and all of the entries of $\mathbf{R}^{-1/2}\mathbf{HBH}^T\mathbf{R}^{-1/2}$ are positive, then the upper and lower bounds in Theorem 6.3.7 are equal, and the bound is exact.*

Proof. The product of circulant matrices is a circulant matrix, the inverse of a circulant matrix is circulant [Gray, 2006], and the square root of a circulant matrix is also circulant [Mei, 2012]. Therefore if the product $\mathbf{HBH}^T$ is circulant, as $\mathbf{R}^{-1/2}$ is circulant by construction, then the product $\mathbf{R}^{-1/2}\mathbf{HBH}^T\mathbf{R}^{-1/2}$ is circulant.

The lower bound of (6.19) computes the average row sum of the product $\mathbf{R}^{-1/2}\mathbf{HBH}^T\mathbf{R}^{-1/2}$. As the product is circulant, each row has the same sum, given by $\sum_{k=1}^{p} d_k$, where d_i is the ith entry of the first row of the circulant matrix (as introduced in Definition 6.4.4).

The upper bound of (6.19) returns the maximum absolute row sum of the product. As the product is circulant with only positive entries, all absolute row sums are identically equal to $\sum_{k=1}^{p} |d_k| = \sum_{k=1}^{p} d_k$. Hence, we have equality of lower and upper bounds and hence the exact value for $\kappa(\widehat{\mathbf{S}})$. □

This results shows that, if the additional assumptions are satisfied, we can compute $\kappa(\widehat{\mathbf{S}})$ directly using the bounds (6.19). The bound given by (6.28) has the added benefit of separating the contributions of each matrix, allowing greater understanding of the influence of each term individually. Bounds which separate the contribution of $\mathbf{R}$ allow users to understand the implication of introducing new observation operators and new observation error covariance matrices on the conditioning of the preconditioned 4D-Var problem.

6.5 Numerical framework

In this section we describe the numerical framework that will be used to study how the bounds on the preconditioned Hessian (6.13) compare with the actual value of $\kappa(\widehat{\mathbf{S}})$. We use the framework that was introduced in Tabeart et al. [2018], which makes use of linear observation operators and circulant covariance matrices for the 3D-Var problem. We note that in the case of 3D-Var, variables with a hat in (6.11) and (6.13) simplify to the standard observation error covariance matrix, $\mathbf{R}$ and observation operator, $\mathbf{H}$.

We now define the different components of the numerical framework. Our domain is the unit circle, and we fix the ratio of state variables to observations as $N = 2p$, i.e. twice as many state variables as observations. Similarly to Tabeart et al. [2018] we define both the observation and background error covariance matrices to have a circulant structure with unit variances. This is a natural choice for correlations on a periodic domain with evenly distributed state variables. Circulant matrices also admit useful theoretical properties as was discussed in Section 6.4.2. The use of circulant error covariance matrices allow us to better understand the interaction between different terms in the Hessian, and to isolate the impact of parameter changes.

Specifically we will focus on circulant matrices arising from the SOAR [Daley, 1991] correlation functions [Johnson, 2003]. SOAR matrices are used in NWP applications as a horizontal correlation function [Simonin et al., 2014] and are fully defined by a correlation lengthscale for a given domain. We remark that we substitute the great circle distance in the SOAR correlation function with the chordal distance, as discussed in Gaspari and Cohn [1999] and Jeong and Jun [2015], to ensure that the properties of positive definiteness are satisfied and that we obtain a valid correlation matrix.

Definition 6.5.1. *The SOAR error correlation matrix on the unit circle is given by*

$$\mathbf{D}(i,j) = \left(1 + \frac{\left|2\sin\left(\frac{\theta_{i,j}}{2}\right)\right|}{L}\right)\exp\left(\frac{-\left|2\sin\left(\frac{\theta_{i,j}}{2}\right)\right|}{L}\right), \tag{6.38}$$

where $L > 0$ is the correlation lengthscale and $\theta_{i,j}$ denotes the angle between grid points i and j. The chordal distance between adjacent grid points is given by

$$\Delta x = 2\sin\left(\frac{\theta}{2}\right) = 2\sin\left(\frac{\pi}{N}\right), \tag{6.39}$$

where N is the number of gridpoints and $\theta = \frac{2\pi}{N}$ is the angle between adjacent gridpoints.

Both our background and observation error covariance matrices for the experiments presented in Section 6.6 will be SOAR with constant unit variance. We now introduce the observation operators that will be used for the 3D-Var experiments. Three of our observation operators are the same as those used in Tabeart et al. [2018] which we state again for clarity. Tabeart et al. [2018, Figure 2] shows a representation of a low dimensional version of the observation operator structure for $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$.

Definition 6.5.2. *The observation operators $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{H}_3 \in \mathbb{R}^{p\times N}$, for $N = 2p$, are defined as follows:*

$$\mathbf{H}_1(i,j) = \begin{cases} 1, & j = i \textit{ for } i = 1,\dots,p \\ 0, & \textit{otherwise.} \end{cases} \tag{6.40}$$

$$\mathbf{H}_2(i,j) = \begin{cases} 1, & j = 2i \textit{ for } i = 1,\dots,p \\ 0, & \textit{otherwise.} \end{cases} \tag{6.41}$$

$$\mathbf{H}_3(i,j) = \begin{cases} \frac{1}{5}, & j \in \{2i-2, 2i-1, 2i, 2i+1, 2i+2 \pmod N\} \textit{ for } i = 1,\dots,p \\ 0, & \textit{otherwise.} \end{cases} \tag{6.42}$$

The first choice of observation operator, $\mathbf{H}_1$ corresponds to direct observations of the first half of the domain. The second observation operator, $\mathbf{H}_2$, corresponds to direct observations of alternate state variables. The third observation operator, $\mathbf{H}_3$, is a smoothed version of $\mathbf{H}_2$. Observations of alternate state variables are smoothed equally over 5 adjacent state variables. The fourth choice of observation operator, $\mathbf{H}_4$ selects p random observations, which are then ordered. We considered a number of choices of random observation operator, and all choices yielded similar numerical results. In order to ensure a fair comparison, we fix the same choice of $\mathbf{H}_4$ for all of the results presented in Section 6.6. This choice of observation operator is shown in Figure 6.1. Observations are spread over the whole domain, but are clustered rather than evenly distributed. For the numerical experiments in Section 6.6, we will use $p = 100$ observations and $N = 200$ state variables. In the unpreconditioned case, structure in the observation operator, such as regularly spaced observations, was important for the tightness of bounds and convergence of a conjugate gradient method [Tabeart et al., 2018]. We therefore consider $\mathbf{H}_4$ as an operator without strict structure. This will allow us to see how structure (or the lack of it) affects the preconditioned problem.

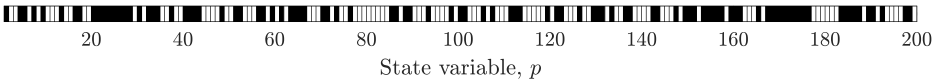


Figure 6.1: Representation of state variables that are observed for $\mathbf{H}_4$. Black denotes state variables that are observed directly and white denotes state variables that are not observed.

6.5.1 Changes to the condition number of the Hessian

Our first set of experiments consider how different combinations of parameters will alter the value of $\kappa(\widehat{\mathbf{S}})$ and the bounds given by (6.28). We compute the condition number of the Hessian (6.13) using the Matlab function *cond* [MATLAB, 2018b] and compare against the values given by our bounds. Tabeart et al. [2018, Table 1] shows how the maximum and minimum eigenvalues of SOAR matrices are affected by changes in lengthscale. The maximum and minimum eigenvalues of both error covariance matrices appear in (6.28). We can therefore predict how the bounds will change with varying parameter values.

- As L_R increases, $\lambda_Q(\mathbf{R})$ decreases. This means that both the upper bound and the first term in the lower bound of (6.28) will increase. However, $\lambda_1(\mathbf{R})$ increases with L_R meaning that the second term in the lower bound will decrease. It is therefore not possible to determine whether the lower bound will increase or decrease with increasing L_R in general.
- As L_B increases, $\lambda_1(\mathbf{B})$ increases. This means that the upper bound of (6.28) will increase with L_B.
- As L_B increases, $\lambda_N(\mathbf{B})$ decreases. This means that both terms in the lower bound of (6.28) will decrease with increasing L_B. Hence, the bounds (6.28) will diverge as L_B increases.

We wish to assess whether the qualitative behaviour of $\kappa(\widehat{\mathbf{S}})$ agrees with the qualitative behaviour of the bounds for our experimental framework. Additionally, we are interested in which of the lower bounds is better, and whether we can determine situations where one bound is tighter than the other. However, we note that the bounds (6.28) separate the contribution of different terms whereas these terms will interact in the value of the Hessian. This means the bounds (6.28) are likely to fail to account for interaction between $\mathbf{B}$ and $\mathbf{R}$.

6.5.2 Convergence of a conjugate gradient algorithm

Although conditioning of a problem is often used as a proxy to study convergence, there are well-known situations where the condition number provides a pessimistic indication of convergence speed. We therefore wish to assess how the convergence of a conjugate gradient method changes with the parameters of the data assimilation system, and whether situations where the bounds on the condition number are large suffer from poor convergence. Following a similar method to [Tabeart et al., 2018, Sec 5.3.2] we study how the speed of convergence of a conjugate gradient method applied to the linear system $\widehat{\mathbf{S}}\mathbf{x} = \mathbf{b}$ changes with the parameters of the system. We define $\mathbf{x}$ as a vector with features at a variety of scales, and then calculate $\mathbf{b} = \widehat{\mathbf{S}}\mathbf{x}$ before recovering $\mathbf{x}$. We use the MATLAB 2018b routine *pcg.m* to recover $\mathbf{x}$ using the conjugate gradient method. As we are studying a preconditioned system, convergence is fast. In order to make the differences between parameter choices more evident we use a strong tolerance of 1×10^{-10} on the relative residual.

We consider how changes to lengthscale and observation operator alter the convergence of the conjugate gradient method. In the case that convergence behaves differently to conditioning, we study the eigenstructure of $\widehat{\mathbf{S}}$ to understand why these differences occur.

6.6 3D-Var experiments

In this section we present the results of our numerical experiments. Figures will be plotted as a function of changes to correlation lengthscales for both $\mathbf{B}$ and $\mathbf{R}$. We recall that increasing the lengthscale of a SOAR correlation matrix will reduce its smallest eigenvalue and increase its largest eigenvalue, [Waller et al., 2016b, Tabeart et al., 2018].

Figure 6.2 shows how the condition number of the preconditioned Hessian (6.13) changes with the lengthscales of $\mathbf{B}$ and $\mathbf{R}$ for different choices of $\mathbf{H}$. For $\mathbf{H}_1$, increasing L_R increases the value of $\kappa(\widehat{\mathbf{S}})$, whereas changes with L_B are much smaller. For $\mathbf{H}_2$, large values of $\kappa(\widehat{\mathbf{S}})$ occur for very large values of L_R and small values of L_B. For a fixed value of L_R, increasing L_B results in a rapid decrease in the value of $\kappa(\widehat{\mathbf{S}})$. For small fixed values of L_R ($L_R < 0.1$), this decrease is followed by a slow increase to $\kappa(\widehat{\mathbf{S}})$ with increasing L_B for $L_B > 0.2$. The minimum value of $\kappa(\widehat{\mathbf{S}})$ occurs when $L_R = L_B$; in this case $\mathbf{H}\mathbf{B}\mathbf{H}^T = \mathbf{R}$ to machine precision for both $\mathbf{H}_2$ and $\mathbf{H}_3$. The

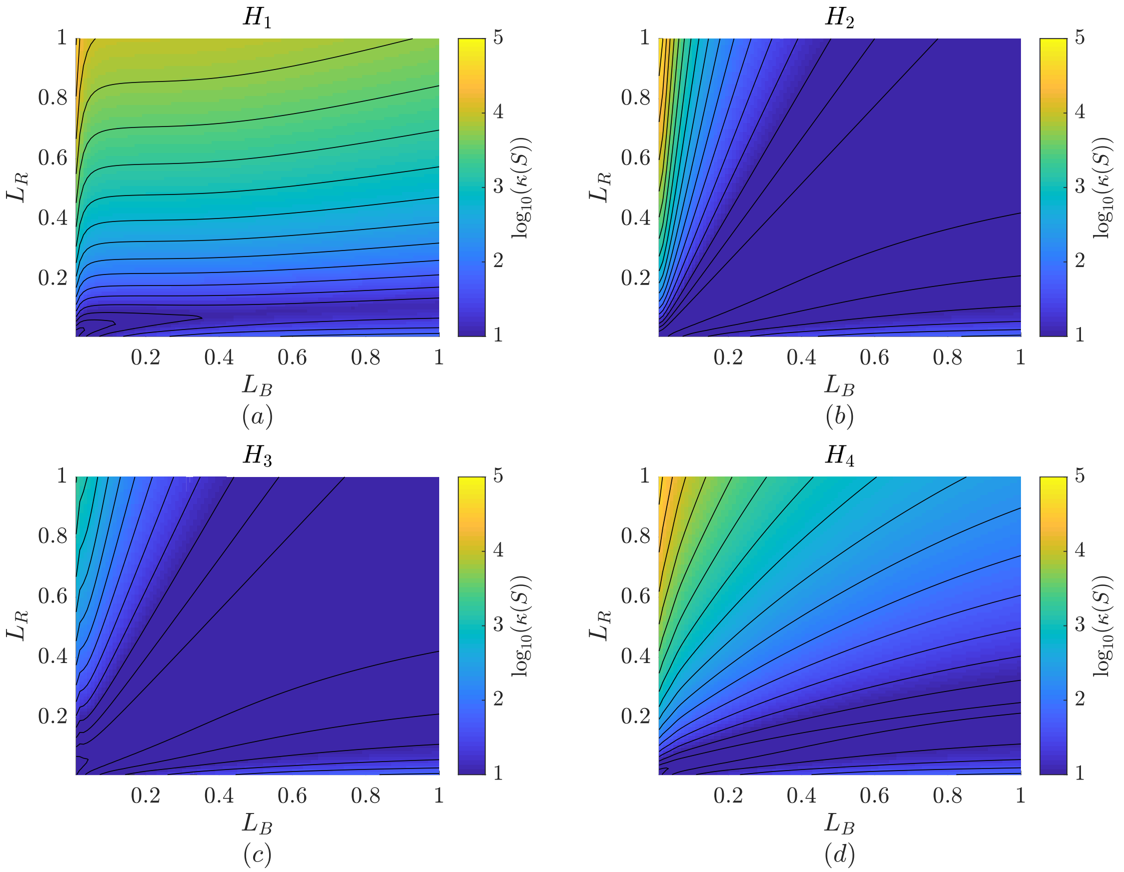


Figure 6.2: Change to $\kappa(\widehat{\mathbf{S}})$ with changes in L_R, L_B for (a) $\mathbf{H}_1$, (b) $\mathbf{H}_2$, (c) $\mathbf{H}_3$ and (d) $\mathbf{H}_4$. The colour map is shown on a logarithmic scale which is standardised for all figures. Contours range from $\kappa(\widehat{\mathbf{S}}) = 0.25$ to $\kappa(\widehat{\mathbf{S}}) = 5$ with a contour interval of 0.25.

qualitative behaviour for $\mathbf{H}_2$ and $\mathbf{H}_3$ is very similar, with smaller values of $\kappa(\widehat{\mathbf{S}})$ for $\mathbf{H}_3$ than $\mathbf{H}_2$. This is also the case in the unpreconditioned setting [Tabeart et al., 2018], and occurs as $\mathbf{H}_3$ can be considered as a smoothed version of $\mathbf{H}_2$. Qualitatively the behaviour for $\mathbf{H}_4$ is a compromise between $\mathbf{H}_1$ and $\mathbf{H}_2$; we can reduce $\kappa(\widehat{\mathbf{S}})$ by increasing L_B or decreasing L_R. In the unpreconditioned case it is always beneficial (in terms of reducing $\kappa(\widehat{\mathbf{S}})$) to decrease either L_R or L_B. However, in the preconditioned setting, for $\mathbf{H}_2$ and $\mathbf{H}_3$ there are cases where $\kappa(\widehat{\mathbf{S}})$ could be reduced by increasing L_B or L_R.

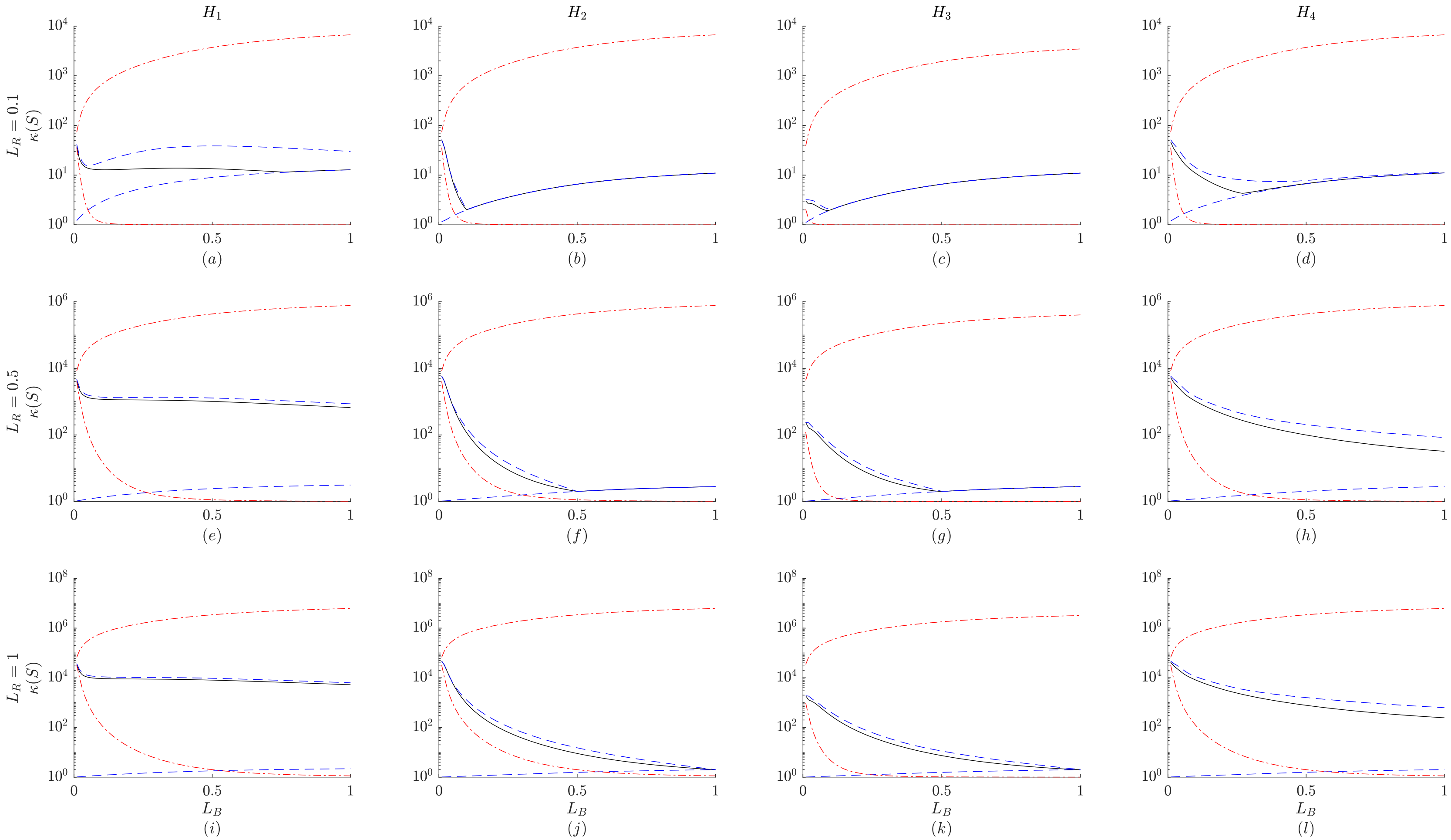


Figure 6.3: Bounds and value of $\kappa(\widehat{\mathbf{S}})$ for (a,e,i) $\mathbf{H}_1$, (b,f,j) $\mathbf{H}_2$, (c,g,k) $\mathbf{H}_3$ and (d,h,l) $\mathbf{H}_4$ as a function of L_B. Blue dashed lines denote the bounds given by (6.19), red dot-dashed lines denote the upper bound and first term in the lower bound of (6.28). The solid black line denotes the value of $\kappa(\widehat{\mathbf{S}})$ calculated using the *cond* command in MATLAB [2018b]. The different rows correspond to different values of L_R.

Figure 6.3 shows the value of $\kappa(\widehat{\mathbf{S}})$, terms in the bounds (6.28) and the bounds (6.19) for various combinations of $\mathbf{H}$, $\mathbf{R}$ and $\mathbf{B}$. The second term in the lower bound (6.28), given by $1 + \lambda_1(\mathbf{H}\mathbf{H}^T)\lambda_N(\mathbf{B})(\lambda_1(\mathbf{R}))^{-1}$, is not shown, as it performs worse than the first term of (6.28), given by $1 + \lambda_Q(\mathbf{H}\mathbf{H}^T)\lambda_N(\mathbf{B})(\lambda_Q(\mathbf{R}))^{-1}$, for all parameter combinations studied. For all choices of $\mathbf{H}$, L_R and L_B, the upper bound of (6.28) is much larger than the actual value of $\kappa(\widehat{\mathbf{S}})$ and does not represent the qualitative behaviour well: the bound increases with L_R which is not the case for $\kappa(\widehat{\mathbf{S}})$ for any of the experiments. This shows that considering the effect of changing each term individually may be no longer be appropriate in the preconditioned case; it is more complicated to separate the effects of changing parameters within a product than within a sum (as in the unpreconditioned setting). The first term in the lower bound of (6.28) represents the qualitative behaviour of $\kappa(\widehat{\mathbf{S}})$ well, capturing the decrease of $\kappa(\widehat{\mathbf{S}})$ with an increase in L_B. However, the value given by the bound is much smaller than the value of $\kappa(\widehat{\mathbf{S}})$, particularly for $L_B < L_R$.

The upper bound of (6.19) represents the qualitative and quantitative behaviour well for all parameter choices. For smaller values of L_R the lower bound of (6.19) also performs well. However, this bound increases monotonically and fails to capture the decrease in $\kappa(\widehat{\mathbf{S}})$ with increasing L_B for values of $L_B < L_R$. For $\mathbf{H}_2$ and $\mathbf{H}_3$ the upper and lower bounds of (6.19) are equal for $L_B > L_R$. This results from Corollary 6.4.6 as $\mathbf{H}\mathbf{B}\mathbf{H}^T$ is circulant when $\mathbf{H} = \mathbf{H}_2$ or $\mathbf{H} = \mathbf{H}_3$ and all entries in the product $\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-1/2}$ are positive for $L_B \geq L_R$.
Comparing the bounds given by (6.28) and (6.19), we find that the upper bound of (6.19) performs better for all parameters studied. The best lower bound depends on the choice of L_B and L_R: for lower values of L_B and larger values of L_R the first term of (6.28) is the tightest. Otherwise the bound given by (6.19) yields the tightest bound in this setting. Although the bounds given by (6.19) represent the behaviour of $\kappa(\widehat{\mathbf{S}})$ well, we note that the numerical framework considered here has a very specific structure that is unlikely to occur in practice. Observation operators are likely to be much less smooth and have less regular structure: e.g. observations may not occur at the location of state variables, observation and state variables may not be evenly spaced, data may be missing, leading to different observation networks at different times or time windows. This may make a difference to the performance of both sets of bounds.

We now consider how altering the data assimilation system affects the convergence of a conjugate gradient method for the problem introduced in Section 6.5.2. Figure 6.4

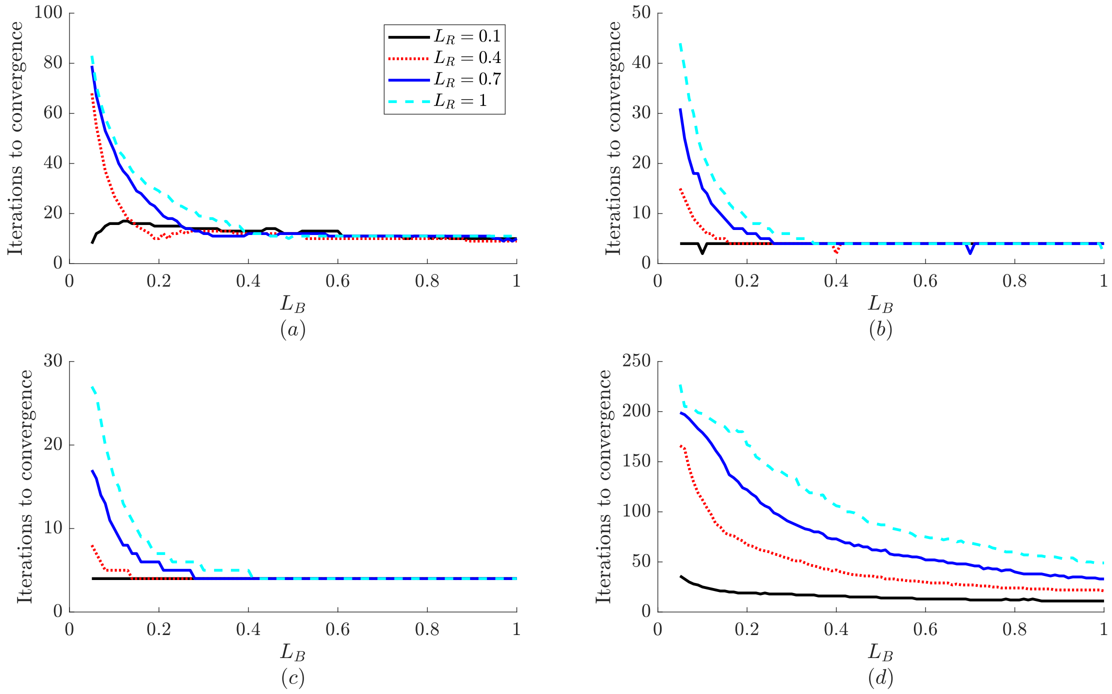


Figure 6.4: Number of iterations required for a conjugate gradient method to converge for changing values of L_R and L_B for (a) $\mathbf{H}_1$, (b) $\mathbf{H}_2$, (c) $\mathbf{H}_3$ and (d) $\mathbf{H}_4$.

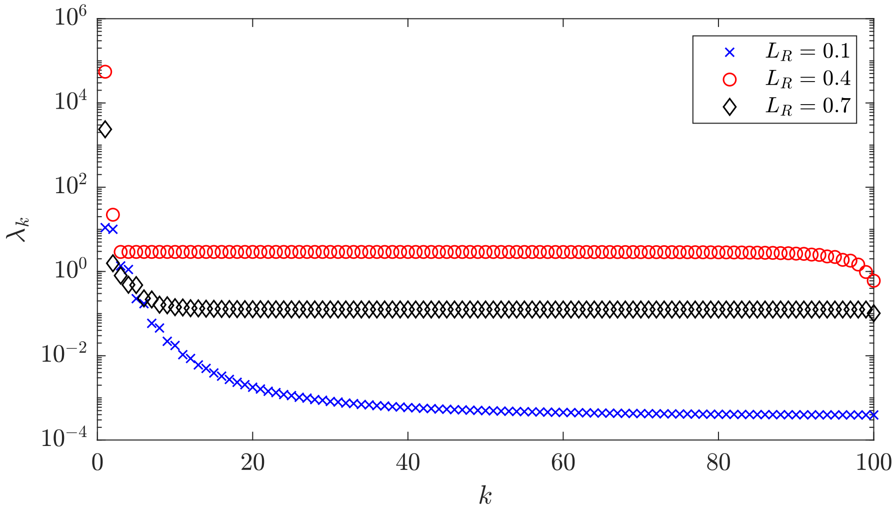


Figure 6.5: Eigenvalues of $\mathbf{R}^{-1}\mathbf{H}_1\mathbf{B}\mathbf{H}_1^T$ for $L_B = 0.8$ and $L_R = 0.1$, $L_R = 0.4$, $L_R = 0.7$. Note the y-axis is plotted with a logarithmic scale

shows how convergence of the conjugate gradient problem changes with L_B, L_R and $\mathbf{H}$. We see that for many cases $\kappa(\widehat{\mathbf{S}})$ is a good proxy for convergence: for $\mathbf{H}_2$, $\mathbf{H}_3$ and $\mathbf{H}_4$ reductions in $\kappa(\widehat{\mathbf{S}})$ and the number of iterations required for convergence occur for the same changes to L_R and L_B. The main difference in behaviour is seen for $\mathbf{H}_1$, where increasing L_R increases $\kappa(\widehat{\mathbf{S}})$ for all choices of L_B, but makes no difference to the number of iterations required for convergence for $L_B > 0.4$.

Figure 6.5 shows the full spectrum of eigenvalues of $\mathbf{R}^{-1}\mathbf{H}_1\mathbf{B}\mathbf{H}_1^T$ for $L_B = 0.8$ and $L_R = 0.1, 0.4, 0.7$. We recall that convergence of the conjugate gradient method depends on the distribution of the entire spectrum, whereas the condition number is sensitive only to the two extreme eigenvalues. In particular, we expect faster convergence to occur where eigenvalues are repeated or clustered [Trefethen and Bau [1997, Theorems 38.3, 38.5], Gill et al. [1986, Theorem 38.4]]. Figure 6.5 shows that increasing L_R leads to an increase in clustering of the eigenvalues of $\mathbf{R}^{-1}\mathbf{H}_1\mathbf{B}\mathbf{H}_1^T$ as well as an increase in $\kappa(\widehat{\mathbf{S}})$. It can be shown numerically that for $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$ and a fixed value of L_R, the number of distinct clusters decreases with increasing L_B until a limiting value is reached. We note that for $\mathbf{H}_4$ the number of distinct clusters is larger than for other choices of observation operator and does not reach a limiting value with L_B for values of L_B studied in our experiments. This explains why convergence of the conjugate gradient method was slowest for this choice of $\mathbf{H}$, and why increasing L_B or decreasing L_R leads to faster convergence for this choice of $\mathbf{H}$.

We conclude that the condition number is a good proxy for convergence in this framework. However, due to the specific structure of the observation network, which leads to repeated and clustered eigenvalues, we obtain faster convergence than can be predicted by $\kappa(\widehat{\mathbf{S}})$ for some parameter choices.

We note that the experiments presented in this section have used constant variances for both background and observation error covariance matrices. Previous work reveals that in the unpreconditioned case the ratio of background and observation variances is important for the conditioning of the unpreconditioned assimilation problem [Haben, 2011]. Further work which studies the effect of changing observation variance on conditioning and convergence of the preconditioned data assimilation problem would therefore be of interest. We recall that all of the experiments presented in this section used constant unit variances for both the background and observation error covariance matrices.

6.7 Conclusions

The inclusion of correlated observation errors in data assimilation is important for high resolution forecasts [Fowler et al., 2018, Rainwater et al., 2015], and to ensure we make the best use of existing data [Michel, 2018, Stewart et al., 2013, Simonin et al., 2019]. However, multiple studies have found issues with convergence of data assimilation routines when introducing correlated observation error covariance (OEC) matrices [Weston, 2011, Campbell et al., 2017, Bormann et al., 2015]. In this chapter, we study the effect of introducing correlated OEC matrices on the convergence of the preconditioned variational data assimilation problem. This extends the theoretical and numerical results of a previous study [Tabeart et al., 2018] that considered the unpreconditioned formulation.

In this chapter, we developed bounds on the condition number of the Hessian of the preconditioned variational data assimilation problem. We then studied these bounds numerically in an idealised framework. We found that:

- The minimum eigenvalue of the OEC matrix appears in both the upper and lower bounds. This was also true for the unpreconditioned case.
- Decreasing the lengthscale of the observation error covariance matrix or increasing the lengthscale of the background error covariance matrix reduced the condition number of the Hessian. Our new lower bound represented the qualitative behaviour better than an existing bound for many cases.
- For most cases the conditioning of the Hessian performed well as a proxy for the convergence of a conjugate gradient method. However in other cases, repeated eigenvalues (induced by the specific structure of the numerical framework) meant that convergence was much faster than predicted by the conditioning. The ratio between background and observation lengthscales was a determining factor for this.

We remark that our findings about repeated eigenvalues occur as our numerical framework has very specific structure. In particular, the eigenvectors of the background and observation error covariance matrices are strongly related. Other experiments not discussed in this chapter considered the use of the Laplacian correlation function [Haben, 2011]. Qualitative conclusions were very similar to those presented in Section 6.6. Although the additional assumptions of Corollary 6.4.6 are not satisfied for the Laplacian correlation function, due to negative entries in

Laplacian correlation function, the bounds presented in Haben [2011] were still tight. In applications, we are likely to have more complex observation operators, and the background and observation error covariance matrices are less likely to have complementary structures. One example is for NWP and the use of observations from satellite based instruments. These have interchannel correlation structures that are different from the typical spatial correlations of background error covariance matrices. We also note that our state variables were evenly distributed and homogeneous, which will not be the case for non-uniform meshes.

In the unpreconditioned case using a similar numerical framework Tabeart et al. [2018] found that improving the conditioning of the background or observation error covariance matrix separately would always decrease $\kappa(\widehat{\mathbf{S}})$. The preconditioned system is more complicationed, with some cases where making the conditioning of the background or observation error covariance matrix worse resulting in smaller values of $\kappa(\widehat{\mathbf{S}})$. We expect the relationship between each of the constituent matrices to be complicated for more general problems. This is relevant for practical applications, as estimated observation error covariance matrices typically need to be treated via reconditioning methods before they can be used [Weston, 2011, Bormann et al., 2015]. Currently the use of reconditioning methods is heuristic [Tabeart et al., 2019a] meaning that there may be flexibility to select a treated matrix that will result in faster convergence in some cases. Theoretical knowledge about the contribution of observation error covariance matrices to the conditioning of the Hessian will allow users to choose between reconditioning methods or select parameters in a more informed manner.

Acknowledgements

This work is funded in part by the EPSRC Centre for Doctoral Training in Mathematics of Planet Earth, the NERC Flooding from Intense Rainfall programme (NE/K008900/1), the EPSRC DARE project (EP/P002331/1) and the NERC National Centre for Earth Observation.

6.8 Summary

In this chapter we developed bounds on the Hessian of the preconditioned data assimilation problem. We found that the minimum eigenvalue of the observation error covariance matrix appears in both upper and lower bounds, and that small eigenvalues of the observation error covariance matrix are likely to lead to larger bounds on the condition number. This agrees with the qualitative conclusions for the unpreconditioned problem that was considered in Chapter 5. Numerical experiments revealed that reducing the condition number of either the background or observation error covariance matrix does not guarantee a reduction in the condition number of the Hessian. This behaviour was not well represented by our bounds, which separate the contribution of each term. We also studied the convergence of a conjugate gradient problem, and found cases where the eigenstructure of the preconditioned Hessian led to much faster convergence than would be expected by simply considering its conditioning. This is a similar finding to the unpreconditioned problem in Chapter 5. However, clustering of eigenvalues occurred for all choices of observation operator in the unpreconditioned case, due to the complementary spatial structures of the background and observation error covariance matrices.

The conclusions from this chapter will not apply directly to the case study in Chapter 8, as the Met Office 1D-Var system is unpreconditioned. However, the results from this chapter will apply to the Met Office implementation of 4D-Var, which is preconditioned. In the next chapter we consider the use of reconditioning methods, which alter the eigenvalues of a covariance matrix in order to reduce its condition number. Our findings that the minimum eigenvalue of the observation error covariance appears in both upper and lower bounds developed in this chapter indicates that reconditioning methods are likely to also be beneficial to the preconditioned data assimilation problem.

Chapter 7

Improving the condition number of estimated covariance matrices

This chapter will address RQ 3 from Chapter 1 and consider how covariance matrices are altered by the application of reconditioning methods. Reconditioning methods have been used to mitigate problems with ill-conditioned estimated covariance matrices by modifying small eigenvalues of a sample covariance matrix. We wish to know:

- How do reconditioning methods alter correlations and standard deviations associated with the covariance matrix?
- How is the variational objective function altered by the use of reconditioning methods?
- How do two commonly-used reconditioning methods compare to multiplicative variance inflation?

The remainder of this chapter, excluding the chapter summary (Section 7.9), is strongly based on the paper: Tabeart J. M., Dance S. L., Haben S. A., Lawless A. S., Nichols N. K., Waller J. A. Improving the condition number of estimated covariance matrices. Tellus A (in press). We also include the supplementary material to the paper in Section 7.8. The submitted paper can be found at https://arxiv.org/abs/1810.10984.

7.1 Abstract

High dimensional error covariance matrices and their inverses are used to weight the contribution of observation and background information in data assimilation procedures. As observation error covariance matrices are often obtained by sampling methods, estimates are often degenerate or ill-conditioned, making it impossible to invert an observation error covariance matrix without the use of techniques to reduce its condition number. In this chapter we present new theory for two existing methods that can be used to 'recondition' any covariance matrix: ridge regression, and the minimum eigenvalue method. We compare these methods with multiplicative variance inflation, which cannot alter the condition number of a matrix, but is often used to account for neglected correlation information. We investigate the impact of reconditioning on variances and correlations of a general covariance matrix in both a theoretical and practical setting. Improved theoretical understanding provides guidance to users regarding method selection, and choice of target condition number. The new theory shows that, for the same target condition number, both methods increase variances compared to the original matrix, with larger increases for ridge regression than the minimum eigenvalue method. We prove that the ridge regression method strictly decreases the absolute value of off-diagonal correlations. Theoretical comparison of the impact of reconditioning and multiplicative variance inflation on the data assimilation objective function shows that variance inflation alters information across all scales uniformly, whereas reconditioning has a larger effect on scales corresponding to smaller eigenvalues. We then consider two examples: a general correlation function, and an observation error covariance matrix arising from interchannel correlations. The minimum eigenvalue method results in smaller overall changes to the correlation matrix than ridge regression, but can increase off-diagonal correlations. Data assimilation experiments reveal that reconditioning corrects spurious noise in the analysis but underestimates the true signal compared to multiplicative variance inflation.

7.2 Introduction

The estimation of covariance matrices for large dimensional problems is of growing interest [Pourahmadi, 2013], particularly for the field of numerical weather prediction (NWP) [Bormann et al., 2016, Weston et al., 2014] where error covariance estimates are used as weighting matrices in data assimilation problems, (e.g. Daley [1991], Ghil [1989], Ghil and Malanotte-Rizzoli [1991]). At operational NWP centres there are typically $\mathcal{O}(10^7)$ measurements every 6 hours [Bannister, 2017], meaning that observation error covariance matrices are extremely high-dimensional. In nonlinear least squares problems arising in variational data assimilation, the inverse of correlation matrices are used, meaning that well-conditioned matrices are vital for practical applications [Bannister, 2017]. This is true in both the unpreconditioned and preconditioned variational data assimilation problem using the control variable transform, as the inverse of the observation error covariance matrix appears in both formulations. The convergence of the data assimilation problem can be poor if either the background or observation variance is small; however, the condition number and eigenvalues of background and observation error covariance matrices have also been shown to be important for convergence in both the unpreconditioned and preconditioned case in Haben et al. [2011b,a], Haben [2011], Tabeart et al. [2018]. Furthermore, the conditioning and solution of the data assimilation system can be affected by complex interactions between the background and observation error covariance matrices and the observation operator [Tabeart et al., 2018, Johnson et al., 2005]. The condition number of a matrix, $\mathbf{A}$, provides a measure of the sensitivity of the solution $\mathbf{x}$ of the system $\mathbf{Ax} = \mathbf{b}$ to perturbations in $\mathbf{b}$. The need for well-conditioned background and observation error covariance matrices motivates the use of 'reconditioning' methods, which are used to reduce the condition number of a given matrix.

In NWP applications, observation error covariance matrices are often constructed from a limited number of samples Cordoba et al. [2017], Waller et al. [2016a,c]. This can cause problems with sampling error, leading to sample covariance matrices, or other covariance matrix estimates, that are very ill-conditioned or can fail to satisfy required properties of covariance matrices (such as symmetry and positive semi-definiteness) [Higham et al., 2016, Ledoit and Wolf, 2004]. In some situations it may be possible to determine which properties of the covariance matrix are well estimated. One such instance is presented in Skøien and Blöschl [2006], which considers how well we can expect the mean, variance and correlation lengthscale of a sample correlation to represent the true correlation matrix depending on different properties of the measured domain (e.g. sample spacing, area measured by each observation). However, this applies only to direct estimation of correlations and will not apply to diagnostic methods, e.g. Desroziers et al. [2005], where transformed samples are used and covariance estimates may be poor. We note that in this chapter, we assume that the estimated covariance matrices used in our experiments represent the desired correlation information matrix well and that differences are due to noise rather than neglected sources of uncertainty. This may not be the case for practical situations, where reconditioning may need to be performed in conjunction with other techniques to compensate for the underestimation of some sources of error.

Depending on the application, a variety of methods have been used to combat the problem of rank deficiency of sample covariance matrices. In the case of spatially correlated errors it may be possible to fit a smooth correlation function or operator to the sample covariance matrix as was done in Simonin et al. [2019] and Guillet et al. [2019] respectively. Another approach is to retain only the first k leading eigenvectors of the estimated correlation matrix and to add a diagonal matrix to ensure the resulting covariance matrix has full rank [Michel, 2018, Stewart et al., 2013]. However, this has been shown to introduce noise at large scales for spatial correlations and may be expensive in terms of memory and computational efficiency [Michel, 2018]. Although localisation can be used to remove spurious correlations, and can also be used to increase the rank of a degenerate correlation matrix [Hamill et al., 2001], it struggles to reduce the condition number of a matrix without destroying off-diagonal correlation information [Smith et al., 2018]. A further way to increase the rank of a matrix is by considering a subset of columns of the original matrix that are linearly independent. This corresponds to using a subset of observations, which is contrary to a key motivation for using correlated observation error statistics: the ability to include a larger number of observations in the assimilation system [Janjić et al., 2018]. Finally, the use of transformed observations imay result in independent observation errors [Migliorini, 2012, Prates et al., 2016]; however, problems with conditioning will manifest in other components of the data assimilation algorithm, typically the observation operator. Therefore, although other techniques to tackle the problem of ill-conditioning exist, they each have limitations. This suggests that for many applications the use of reconditioning methods, which we will show are inexpensive to implement and are not limited to spatial correlations, may be beneficial.

We note that small eigenvalues of the observation error covariance matrix are not the only reason for slow convergence: if observation standard deviations are small, the observation error covariance matrix may be well-conditioned, but convergence of the minimisation problem is likely to be poor [Haben, 2011, Tabeart et al., 2018]. In this case reconditioning may not improve convergence and performance of the data assimilation routine.

Two methods in particular, referred to in this work as the minimum eigenvalue method and ridge regression, are commonly used at NWP centres. Both methods are used by Weston [2011], where they are tested numerically. Additionally in Campbell et al. [2017] a comparison between these methods is made experimentally and it is shown that reconditioning improves convergence of a dual four-dimensional variational assimilation system. However, up to now there has been minimal theoretical investigation into the effects of these methods on the covariance matrices. In this chapter we develop theory that shows how variances and correlations are altered by the application of reconditioning methods to a covariance matrix.

Typically reconditioning is applied to improve convergence of a data assimilation system by reducing the condition number of a matrix. However, the convergence of a data assimilation system can also be improved using multiplicative variance inflation, a commonly used method at NWP centres such as the European Centre for Medium-Range Weather Forecasts (ECMWF) [Liu and Rabier, 2003, McNally et al., 2006, Bormann et al., 2015, 2016] to account for neglected error correlations or to address deficiencies in the estimated error statistics by increasing the uncertainty in observations. It is not a method of reconditioning when a constant inflation factor is used, as it cannot change the condition number of a covariance matrix. In practice multiplicative variance inflation is often combined with other techniques, such as neglecting off-diagonal error correlations, which do alter the conditioning of the observation error covariance matrix.

Although it is not a reconditioning technique, in Bormann et al. [2015] multiplicative variance inflation was found to yield faster convergence of a data assimilation procedure than either the ridge regression or minimum eigenvalue methods of reconditioning. This finding is likely to be system-dependent; the original diagnosed error covariance matrix in the ECMWF system has a smaller condition number than the corresponding matrix for the same instrument in the Met Office system [Weston et al., 2014]. Additionally, in the ECMWF system the use of reconditioning methods only results in small improvements to convergence, and there is little difference in convergence speed for the two methods. This contrasts with the findings of Weston [2011], Weston et al. [2014], Campbell et al. [2017] where differences in convergence speed when using each method of reconditioning were found to be large. Therefore, it is likely that the importance of reducing the condition number of the observation error covariance matrix compared to inflating variances will be sensitive to the data assimilation system of interest. Aspects of the data assimilation system that may be important in determining the level of this sensitivity include: the choice of preconditioning and minimisation scheme [Bormann et al., 2015], quality of the covariance estimate, interaction between background and estimated observation error covariance matrices within the data assimilation system [Fowler et al., 2018, Tabeart et al., 2018], the use of thinning and different observation networks. We also note that

Stewart et al. [2008b], Stewart [2010], Stewart et al. [2013] consider changes to the information content and analysis accuracy corresponding to different approximations to a correlated observation error covariance matrix (including an inflated diagonal matrix). Stewart et al. [2013], Healy and White [2005] also provide evidence in idealized cases to show that inclusion of even approximate correlation structure gives significant benefit over diagonal matrix approximations, including when variance inflation is used.

In this work we investigate the minimum eigenvalue and ridge regression methods of reconditioning as well as multiplicative variance inflation, and analyse their impact on the covariance matrix. We compare both methods theoretically for the first time, by considering the impact of reconditioning on the correlations and variances of the covariance matrix. We also study how each method alters the objective function when applied to the observation error covariance matrix. Other methods of reconditioning, including thresholding [Bickel and Levina, 2008] and localisation [Horn, 1991, Ménétrier et al., 2015, Smith et al., 2018] have been discussed from a theoretical perspective in the literature but will not be included in this work. In Section 7.3 we describe the methods more formally than in previous literature before developing new related theory in detail in Section 7.4. We show that the ridge regression method increases the variances and decreases the correlations for a general covariance matrix and the minimum eigenvalue method increases variances. We prove that the increases to the variance are bigger for the ridge regression method than the minimum eigenvalue method for any covariance matrix. We show that both methods of reconditioning reduce the weight on observation information in the objective function in a scale dependent way, with the largest reductions in weight corresponding to the smallest eigenvalues of the original observation error covariance matrix. In contrast, multiplicative variance inflation using a constant inflation factor reduces the weight on observation information by a constant amount for all scales. In Section 7.5 the methods are illustrated via numerical experiments for two types of covariance structures. One of these is a simple general correlation function, and one is an interchannel covariance arising from a satellite based instrument with observations used in NWP. We provide physical interpretation of how each method alters the covariance matrix, and use this to provide guidance on which method of reconditioning is most appropriate for a given application. We present an illustration of how all three methods alter the analysis of a data assimilation problem, and relate this to the theoretical conclusions concerning the objective function. We finally present our conclusions in Section 7.6. The methods are very general and, although their initial application was to observation error covariances arising from numerical weather prediction, the results presented here apply to any sampled covariance matrix, such as those arising in finance [Higham, 2002, Qi and Sun, 2010] and neuroscience [Nakamura and Potthast, 2015, Schiff, 2011].

7.3 Covariance matrix modification methods

We begin by defining the condition number, noting that all covariance matrices are positive semi-definite by definition. The condition number provides a measure of how sensitive solutions of a linear equation $\mathbf{Ax} = \mathbf{b}$ are to perturbations in the data $\mathbf{b}$. A 'well-conditioned problem' will result in small perturbations to the solution with small changes to $\mathbf{b}$, whereas for an 'ill-conditioned problem', small perturbations to $\mathbf{b}$ can result in large changes to the solution. We distinguish between the two cases of strictly positive definite covariance matrices, and covariance matrices with zero minimum eigenvalue. Symmetric positive definite matrices admit a definition for the condition number in terms of their maximum and minimum eigenvalues. For the remainder of the work, we define the eigenvalues of a symmetric positive semi-definite matrix $\mathbf{S} \in \mathbb{R}^{d\times d}$ via:

$$\lambda_1(\mathbf{S}) \geq \dots\ \geq \lambda_d(\mathbf{S}) \geq 0. \tag{7.1}$$

Theorem 7.3.1. *If* $\mathbf{S} \in \mathbb{R}^{d\times d}$ *is a symmetric positive definite matrix with eigenvalues defined as in 7.1 we can write the condition number in the* L_2 *norm as* $\kappa(\mathbf{S}) = \frac{\lambda_1(\mathbf{S})}{\lambda_d(\mathbf{S})}$.

Proof. See [Golub and Van Loan, 1996, Sec. 2.7.2]. □

For a singular covariance matrix, $\mathbf{S}$, the convention is to take $\kappa(\mathbf{S}) = \infty$ [Trefethen and Bau, 1997, Sec. 12]. We also note that real symmetric matrices admit orthogonal eigenvectors which can be normalised to produce a set of orthonormal eigenvectors.

Let $\mathbf{R} \in \mathbb{R}^{d\times d}$ be a positive semi-definite covariance matrix with condition number $\kappa(\mathbf{R}) = \kappa$. We wish to recondition $\mathbf{R}$ to obtain a covariance matrix with condition number κ_{max}

$$1 \leq \kappa_{max} < \kappa, \tag{7.2}$$

where the value of κ_{max} is chosen by the user. We denote the eigendecomposition of $\mathbf{R}$ by

$$\mathbf{R} = \mathbf{V_R}\mathbf{\Lambda}\mathbf{V_R}^T \tag{7.3}$$

where $\mathbf{\Lambda} \in \mathbb{R}^{d\times d}$ is the diagonal matrix of eigenvalues of $\mathbf{R}$ and $\mathbf{V_R} \in \mathbb{R}^{d\times d}$ is a corresponding matrix of orthonormal eigenvectors.

In addition to considering how the covariance matrix itself changes with reconditioning, it is also of interest to consider how the related correlations and standard deviations are altered. We decompose $\mathbf{R}$ as $\mathbf{R} = \mathbf{\Sigma C \Sigma}$, where $\mathbf{C}$ is a correlation matrix, and $\mathbf{\Sigma}$ is a non-singular diagonal matrix of standard deviations. We calculate $\mathbf{C}$ and $\mathbf{\Sigma}$ via:

$$\mathbf{\Sigma}(i,i) = \sqrt{\mathbf{R}(i,i)}, \quad \mathbf{C}(i,j) = \frac{\mathbf{R}(i,j)}{\sqrt{\mathbf{R}(i,i)}\sqrt{\mathbf{R}(j,j)}}. \tag{7.4}$$

We now introduce the ridge regression method and the minimum eigenvalue method; the two methods of reconditioning that will be discussed in this work. We then define multiplicative variance inflation. This last method is not a method of reconditioning, but will be used for comparison purposes with the ridge regression and minimum eigenvalues methods.

7.3.1 Ridge regression method

The ridge regression method (RR) adds a scalar multiple of the identity to $\mathbf{R}$ to obtain the reconditioned matrix $\mathbf{R}_{RR}$. The scalar δ is set using the following method.

- Define
$$\delta = \frac{\lambda_1(\mathbf{R}) - \lambda_d(\mathbf{R})\kappa_{max}}{\kappa_{max} - 1}. \tag{7.5}$$
- Set $\mathbf{R}_{RR} = \mathbf{R} + \delta \mathbf{I}$

We note that this choice of δ yields $\kappa(\mathbf{R}_{RR}) = \kappa_{max}$.

In the literature [Hoerl and Kennard, 1970, Ledoit and Wolf, 2004], ‘ridge regression’ is a method used to regularise least squares problems. In this context, ridge regression can be shown to be equivalent to Tikhonov regularisation [Hansen, 1998]. However, in this chapter we apply ridge regression as a reconditioning method directly to a covariance matrix. For observation error covariance matrices, the reconditioned matrix is then inverted prior to its use as a weighting matrix in the data assimilation objective function. As we are only applying the reconditioning to a single component matrix in the variational formulation, the implementation of the ridge regression method used in this chapter is not equivalent to Tikhonov regularisation applied to the variational data assimilation problem [Budd et al., 2011, Moodey et al., 2013].

This is shown in Section 7.4.5 where we consider how applying ridge regression to the observation error covariance matrix affects the variational data assimilation objective function. The ridge regression method is used at the Met Office [Weston et al., 2014].

7.3.2 Minimum eigenvalue method

The minimum eigenvalue method (ME) fixes a threshold, T, below which all eigenvalues of the reconditioned matrix, $\mathbf{R}_{ME}$, are set equal to the threshold value. The value of the threshold is set using the following method.

- Set $\lambda_1(\mathbf{R}_{ME}) = \lambda_1(\mathbf{R})$
- Define $T = \lambda_1(\mathbf{R})/\kappa_{max} > \lambda_d(\mathbf{R})$, where κ_{max} is defined in (7.2).
- Set the remaining eigenvalues of $\mathbf{R}_{ME}$ via

$$\lambda_k(\mathbf{R}_{ME}) = \begin{cases} \lambda_k(\mathbf{R}) & \text{if } \lambda_k(\mathbf{R}) > T \\ T & \text{if } \lambda_k(\mathbf{R}) \leq T \end{cases}. \tag{7.6}$$

- Construct the reconditioned matrix via $\mathbf{R}_{ME} = \mathbf{V_R}\mathbf{\Lambda}_{ME}\mathbf{V}_{\mathbf{R}}^T$, where $\mathbf{\Lambda}_{ME}(i,i) = \lambda_i(\mathbf{R}_{ME})$.

This yields $\kappa(\mathbf{R}_{ME}) = \kappa_{max}$. The updated matrix of eigenvalues can be written as $\mathbf{\Lambda}_{ME} = \mathbf{\Lambda} + \mathbf{\Gamma}$, the sum of the original matrix of eigenvalues and $\mathbf{\Gamma}$, a low-rank diagonal matrix update with entries $\mathbf{\Gamma}(k,k) = \max\{T - \lambda_k, 0\}$. Using (7.3) the reconditioned $\mathbf{R}_{ME}$ can then be written as:

$$\mathbf{R}_{ME} = \mathbf{V_R}(\mathbf{\Lambda} + \mathbf{\Gamma})\mathbf{V}_{\mathbf{R}}^T = \mathbf{R} + \mathbf{V_R}\mathbf{\Gamma}\mathbf{V}_{\mathbf{R}}^T. \tag{7.7}$$

Under the condition that $\kappa_{max} > d - l + 1$, where l is the index such that $\lambda_l \leq T < \lambda_{l-1}$, the minimum eigenvalue method is equivalent to minimising the difference $\mathbf{R} - \mathbf{R}_{ME} \in \mathbb{R}^{d \times d}$ with respect to the Ky Fan 1-d norm (The proof is provided in Appendix 7.7). The Ky Fan $p-k$ norm (also referred to as the trace norm) is defined in Fan [1959], Horn [1991], and is used in Takana and Nakata [2014] to find the closest positive definite matrix with condition number smaller than a given constant. A variant of the minimum eigenvalue method is applied to observation error covariance matrices at ECMWF [Bormann et al., 2016].

7.3.3 Multiplicative variance inflation

Multiplicative variance inflation (MVI) is a method that increases the variances corresponding to a covariance matrix. Its primary use is to account for neglected error correlation information, particularly in the case where diagonal covariance matrices are being used even though non-zero correlations exist in practice. However, this method can also be applied to non-diagonal covariance matrices.

Definition 7.3.2. *Let $\alpha > 0$ be a given variance inflation factor and*

$$\mathbf{R} = \mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma}$$

be the estimated covariance matrix. Then multiplicative variance inflation is defined by

$$\mathbf{\Sigma}_{MVI} = \alpha\mathbf{\Sigma}. \tag{7.8}$$

This is equivalent to multiplying the estimated covariance matrix by a constant. The updated covariance matrix is given by

$$\mathbf{R}_{MVI} = (\alpha\mathbf{\Sigma})\,\mathbf{C}\,(\alpha\mathbf{\Sigma}) = \alpha^2\mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma} = \alpha^2\mathbf{R}. \tag{7.9}$$

The estimated covariance matrix is therefore multiplied by the square of the inflation constant. We note that the correlation matrix, $\mathbf{C}$, is unchanged by application of multiplicative variance inflation.

Multiplicative variance inflation is used at NWP centres including ECMWF [Bormann et al., 2016] to counteract deficiencies in estimated error statistics, such as underestimated or neglected sources of error. Inflation factors are tuned to achieve improved analysis or forecast performance, and are hence strongly dependent on the specific data assimilation system of interest. Aspects of the system that might influence the choice of inflation factor include observation type, known limitations of the covariance estimate, and observation sampling or thinning.

Although variance inflation with a fixed inflation factor is not a method of reconditioning, as it is not able to alter the condition number of a covariance matrix, we include it in this chapter for comparison purposes. This means that variance inflation can only be used in the case that the estimated matrix can be inverted directly, i.e. is full rank. Multiplicative variance inflation could also refer to the case where the constant inflation factor is replaced with a diagonal matrix of inflation factors. In this case the condition number of the altered covariance matrix would change. An example of a study where multiple inflation factors are used is given by Heilliette and Garand [2015], where the meteorological variable to which an observation is sensitive determines the choice of inflation factor. However, this is not commonly used in practice, and will not be considered in this chapter.

7.4 Theoretical considerations

In this section we develop new theory for each method. We are particularly interested in the changes made to $\mathbf{C}$ and $\mathbf{\Sigma}$ for each case. Increased understanding of the effect of each method may allow users to adapt or extend these methods, or determine which is the better choice for practical applications.

We now introduce an assumption that will be used in the theory that follows.

Main Assumption: Let $\mathbf{R} \in \mathbb{R}^{d\times d}$ be a symmetric positive semi-definite matrix with $\lambda_1(\mathbf{R}) > \lambda_d(\mathbf{R})$.

We remark that any symmetric, positive semi-definite matrix with $\lambda_1 = \lambda_d$ is a scalar multiple of the identity, and cannot be reconditioned since it is already at its minimum possible value of unity. Hence in what follows, we will consider only matrices $\mathbf{R}$ that satisfy the Main Assumption.

7.4.1 Ridge Regression Method

We begin by discussing the theory of RR. In particular we prove that applying this method for any positive scalar, δ, results in a decreased condition number for any choice of $\mathbf{R}$.

Theorem 7.4.1. *Under the conditions of the Main Assumption, adding any positive increment to the diagonal elements of* $\mathbf{R}$ *decreases its condition number.*

Proof. We recall that $\mathbf{R}_{RR} = \mathbf{R} + \delta\mathbf{I}$. The condition number of $\mathbf{R}_{RR}$ is given by

$$\kappa(\mathbf{R}_{RR}) = \frac{\lambda_1(\mathbf{R}_{RR})}{\lambda_d(\mathbf{R}_{RR})} = \frac{\lambda_1(\mathbf{R}) + \delta}{\lambda_d(\mathbf{R}) + \delta}. \tag{7.10}$$

It is straightforward to show that for any $\delta > 0$, $\kappa(\mathbf{R}_{RR}) < \kappa(\mathbf{R})$, completing the proof.

□

We now consider how application of RR affects the correlation matrix $\mathbf{C}$ and the diagonal matrix of standard deviations $\mathbf{\Sigma}$.

Theorem 7.4.2. *Under the conditions of the Main Assumption, the ridge regression method updates the standard deviation matrix Σ_{RR}, and correlation matrix $\mathbf{C}_{RR}$ of $\mathbf{R}_{ME}$ via*

$$\Sigma_{RR} = (\mathbf{\Sigma}^2 + \delta \mathbf{I}_d)^{1/2}, \quad \mathbf{C}_{RR} = \Sigma_{RR}^{-1}\mathbf{R}\Sigma_{RR}^{-1} + \delta\Sigma_{RR}^{-2}. \tag{7.11}$$

Proof. Using (7.4), $\mathbf{\Sigma}(i,i) = (\mathbf{R}(i,i))^{1/2}$. Substituting this into the expression for $\mathbf{R}_{RR}$ yields:

$$\Sigma_{RR}(i,i) = (\mathbf{R}_{RR}(i,i))^{1/2} = (\mathbf{R}(i,i) + \delta)^{1/2} = (\mathbf{\Sigma}(i,i)^2 + \delta)^{1/2}. \tag{7.12}$$

Considering the components of $\mathbf{C}_{RR}$ and the decomposition of Σ_{RR} given by (7.4):

$$\Sigma_{RR}\mathbf{C}_{RR}\Sigma_{RR} = \mathbf{R} + \delta \mathbf{I}_d, \quad \mathbf{C}_{RR} = \Sigma_{RR}^{-1}\mathbf{R}\Sigma_{RR}^{-1} + \delta\Sigma_{RR}^{-2} \tag{7.13}$$

as required. □

Theorem 7.4.2 shows how we can apply RR to our system by updating $\mathbf{C}$ and $\mathbf{\Sigma}$ rather than $\mathbf{R}$. We observe, from (7.11), that applying RR leads to a constant increase to variances for all variables. However, the inflation to standard deviations is additive, rather than the multiplicative inflation that occurs for multiplicative variance inflation. We now show that RR also reduces all non-diagonal entries of the correlation matrix.

Corollary 7.4.3. *Under the conditions of the Main Assumption, for $i \neq j$, $|\mathbf{C}_{RR}(i,j)| < |\mathbf{C}(i,j)|$.*

Proof. Writing the update equation for $\mathbf{C}$, given by (7.11), in terms of the variance and correlations of $\mathbf{R}$ yields:

$$\mathbf{C}_{RR} = \Sigma_{RR}^{-1}\mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma}\Sigma_{RR}^{-1} + \delta\Sigma_{RR}^{-2}. \tag{7.14}$$

We consider $\mathbf{C}_{RR}(i,j)$ for $i \neq j$. As Σ_{RR} and $\mathbf{\Sigma}$ are diagonal matrices, we obtain

$$\mathbf{C}_{RR}(i,j) = \Sigma_{RR}^{-1}(i,i)\mathbf{\Sigma}(i,i)\mathbf{C}(i,j)\mathbf{\Sigma}(j,j)\Sigma_{RR}^{-1}(j,j). \tag{7.15}$$

From the update equation (7.11), $\Sigma_{RR}(i,i) > \mathbf{\Sigma}(i,i)$ for any choice of i. This means that $\Sigma_{RR}^{-1}(i,i)\mathbf{\Sigma}(i,i) < 1$ for any choice of i. Using this in (7.15) yields that for all values of i, j with $i \neq j$, $|\mathbf{C}_{RR}(i,j)| < |\mathbf{C}(i,j)|$ as required.

For $i = j$, it follows from (7.14) that $\mathbf{C}_{RR}(i,i) = 1$ for all values of i. □

7.4.2 Minimum Eigenvalue Method

We now discuss the theory of ME as introduced in Section 7.3.2. Using the alternative decomposition of $\mathbf{R}_{ME}$ given by (7.7) enables us to update directly the standard deviations for this method.

Theorem 7.4.4. *Under the conditions of the Main Assumption, the minimum eigenvalue method updates the standard deviations, $\mathbf{\Sigma}_{ME}$, of $\mathbf{R}$ via*

$$\mathbf{\Sigma}_{ME}(i,i) = \left(\mathbf{R}(i,i) + \sum_{k=1}^{d} \mathbf{V}_R(i,k)^2 \mathbf{\Gamma}(k,k)\right)^{1/2}. \tag{7.16}$$

This can be bounded by

$$\mathbf{\Sigma}(i,i) \le \mathbf{\Sigma}_{ME}(i,i) \le \left(\mathbf{\Sigma}(i,i)^2 + T - \lambda_d(\mathbf{R})\right)^{1/2}. \tag{7.17}$$

Proof.

$$\mathbf{\Sigma}_{ME}(i,i) = \left(\mathbf{R}(i,i) + \left(\mathbf{V}_R \mathbf{\Gamma} \mathbf{V}_R^T\right)(i,i)\right)^{1/2} \tag{7.18}$$

$$= \left(\mathbf{R}(i,i) + \sum_{k=1}^{d} \mathbf{V}_R(i,k)^2 \mathbf{\Gamma}(k,k)\right)^{1/2}. \tag{7.19}$$

Noting that $\mathbf{\Gamma}(k,k) \ge 0$ for all values of k, we bound the second term in this expression by

$$0 \le \sum_{k=1}^{d} \mathbf{V}_R(i,k)^2 \mathbf{\Gamma}(k,k) \le \max_k\{\mathbf{\Gamma}(k,k)\} \sum_{k=1}^{d} \mathbf{V}_R(i,k)^2 \tag{7.20}$$

$$\le (T - \lambda_d(\mathbf{R})) \sum_{k=1}^{d} \mathbf{V}_R(i,k)^2. \le T - \lambda_d(\mathbf{R}) \tag{7.21}$$

This inequality follows from the orthonormality of $\mathbf{V_R}$, and by the fact that $T > \lambda_d(\mathbf{R})$ by definition.

□

Due to the way the spectrum of $\mathbf{R}$ is altered by ME it is not evident how correlation entries are altered in general for this method of reconditioning.

7.4.3 Multiplicative variance inflation

We now discuss theory of MVI that was introduced in Section 7.3.3. We prove that MVI is not a method of reconditioning, as it does not change the condition number of a covariance matrix.

Theorem 7.4.5. *Multiplicative variance inflation with a constant inflation parameter cannot change the condition number or rank of a matrix.*

Proof. Let $\alpha^2 > 0$ be our multiplicative inflation constant such that $\mathbf{R}_{MVI} = \alpha^2\mathbf{R}$. The eigenvalues of $\mathbf{R}_{MVI}$ are given by $\alpha^2\lambda_1, \alpha^2\lambda_2, \dots, \alpha^2\lambda_d$.

If $\mathbf{R}$ is rank-deficient, then $\lambda_{min}(\mathbf{R}_{MVI}) = \alpha^2\lambda_d = 0$ and hence $\mathbf{R}_{MVI}$ is also rank deficient. If $\mathbf{R}$ is full rank then we can compute the condition number of $\mathbf{R}_{MI}$ as the ratio of its eigenvalues, which yields

$$\kappa(\mathbf{R}_{MVI}) = \frac{\alpha^2\lambda_1}{\alpha^2\lambda_d} = \kappa(\mathbf{R}). \tag{7.22}$$

Hence the condition number and rank of $\mathbf{R}$ are unchanged by multiplicative inflation. □

7.4.4 Comparing ridge regression and minimum eigenvalue methods

Both RR and ME change $\mathbf{R}$ by altering its eigenvalues. In order to compare the two methods, we can consider their effect on the standard deviations. We recall from Sections 7.4.1 and 7.4.2 that RR increases standard deviations by a constant and the changes to standard deviations by ME can be bounded above and below by a constant.

Corollary 7.4.6. *Under the conditions of the Main Assumption, for a fixed value of $\kappa_{max} < \kappa$, $\mathbf{\Sigma}_{ME}(i,i) < \Sigma_{RR}(i,i)$ for all values of i.*

Proof. From Theorems 7.4.2 and 7.4.4 the updated standard deviation values are given by

$$\Sigma_{RR} = \left(\mathbf{\Sigma}^2 + \delta\mathbf{I}_d\right)^{1/2} \quad \text{and} \quad \mathbf{\Sigma}_{ME}(i,i) \leq \left(\mathbf{\Sigma}(i,i)^2 + T - \lambda_d(\mathbf{R})\right)^{1/2}. \tag{7.23}$$

From the definitions of δ and T we obtain that

$$\delta = \frac{\lambda_1(\mathbf{R}) - \lambda_d(\mathbf{R})\kappa_{max}}{\kappa_{max} - 1} > \frac{\lambda_1(\mathbf{R}) - \lambda_d(\mathbf{R})\kappa_{max}}{\kappa_{max}} = T - \lambda_d(\mathbf{R}). \tag{7.24}$$

We conclude that the increment to the standard deviations for RR is always larger than the increment for ME. □

7.4.5 Comparison of methods of reconditioning and multiplicative variance inflation on the variational data assimilation objective function

We demonstrate how RR, ME and MVI alter the objective function of the variational data assimilation problem when applied to the observation error covariance matrix. We consider the 3D-Var objective function here for simplicity of notation, although the analysis extends naturally to the 4D-Var case. We begin by defining the 3D-Var objective function of the variational data assimilation problem.

Definition 7.4.7. *The objective function of the variational data assimilation problem is given by*

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}(\mathbf{y}-h[\mathbf{x}])^T\mathbf{R}^{-1}(\mathbf{y}-h[\mathbf{x}]) := J_b + J_o \tag{7.25}$$

where $\mathbf{x}_b \in \mathbb{R}^n$ is the background or prior, $\mathbf{y} \in \mathbb{R}^d$ is the vector of observations, $h : \mathbf{R}^n \to \mathbf{R}^d$ is the observation operator mapping from control vector space to observation space, $\mathbf{B} \in \mathbb{R}^{n\times n}$ is the background error covariance matrix, and $\mathbf{R} \in \mathbb{R}^{d\times d}$ is the observation error covariance matrix. Let J_o denote the observation term in the objective function and J_b denote the background term in the objective function.

In order to compare the effect of using each method, they are applied to the observation error covariance matrix in the variational objective function (7.25). We note that analogous results hold if all methods are applied to the background error covariance matrix in the objective function.

We begin by presenting the three updated objective functions, and then discuss the similarities and differences for each method together at the end of Section 3.5. We first consider how applying RR to the observation error covariance matrix alters the variational objective function (7.25).

Theorem 7.4.8. *By applying RR to the observation error covariance matrix we alter the objective function* (7.25) *as follows:*

$$J_{RR}(\mathbf{x}) = J(\mathbf{x}) - (\mathbf{y}-h[\mathbf{x}])^T V_{\mathbf{R}}\mathbf{\Lambda}_\delta V_{\mathbf{R}}^T(\mathbf{y}-h[\mathbf{x}]), \tag{7.26}$$

where $\mathbf{\Lambda}_\delta$ is a diagonal matrix with entries given by $(\mathbf{\Lambda}_\delta)_{ii} = \frac{\delta}{\lambda_i(\lambda_i+\delta)}$.

Proof. We denote the eigendecomposition of $\mathbf{R}$ as in (7.3). Applying RR to the observation error covariance matrix, $\mathbf{R}$, we obtain

$$\mathbf{R}_{RR} = \mathbf{V}_{\mathbf{R}}(\mathbf{\Lambda} + \delta \mathbf{I}_p)\mathbf{V}_{\mathbf{R}}^T. \tag{7.27}$$

We then calculate the inverse of $\mathbf{R}_{RR}$ and express this in terms of $\mathbf{R}^{-1}$ and an update term:

$$\begin{aligned}
\mathbf{R}_{RR}^{-1} &= \mathbf{V}_{\mathbf{R}}(\mathbf{\Lambda} + \delta \mathbf{I}_p)^{-1}\mathbf{V}_{\mathbf{R}}^T && (7.28)\\
&= \mathbf{V}_{\mathbf{R}}\text{Diag}\Big(\frac{1}{\lambda_i} - \frac{\delta}{\lambda_i(\lambda_i + \delta)}\Big)\mathbf{V}_{\mathbf{R}}^T && (7.29)\\
&= \mathbf{R}^{-1} - \mathbf{V}_{\mathbf{R}}\text{Diag}\Big(\frac{\delta}{\lambda_i(\lambda_i + \delta)}\Big)\mathbf{V}_{\mathbf{R}}^T. && (7.30)
\end{aligned}$$

Substituting (7.30) into (7.25), and defining $\mathbf{\Lambda}_\delta$ as in the theorem statement we can write the objective function using the reconditioned observation error covariance matrix as (7.26). □

The effect of RR on the objective function differs from the typical application of Tikhonov regularisation to the variational objective function [Budd et al., 2011, Moodey et al., 2013]. In particular, we subtract a term from the original objective function rather than adding one, and the term depends on the eigenvectors of $\mathbf{R}$ as well as the innovations (differences between observations and the background field in observation space). Writing the updated objective function as in (7.26) shows that the size of the original objective function (7.25) is decreased when RR is used. Specifically, as we discuss later, the contribution of small-scale information to the observation term, J_o, is reduced by the application of RR.

We now consider how applying ME to the observation error covariance matrix alters the objective function (7.25).

Theorem 7.4.9. *By applying ME to the observation error covariance matrix we alter the objective function* (7.25) *as follows:*

$$J_{ME}(\mathbf{x}) = J(\mathbf{x}) - (\mathbf{y} - h[\mathbf{x}])^T V_{\mathbf{R}}\widetilde{\mathbf{\Gamma}} V_{\mathbf{R}}^T(\mathbf{y} - h[\mathbf{x}]), \tag{7.31}$$

where

$$\widetilde{\mathbf{\Gamma}}(i,i) = \begin{cases} 0 & if \lambda_i \geq T \\ \frac{T-\lambda_i}{T\lambda_i} & if \lambda_i < T. \end{cases} \tag{7.32}$$

Proof. We begin by applying ME and decomposing $\mathbf{R}_{ME}$ as in (7.7):

$$\mathbf{R}_{ME} = \mathbf{V}_{\mathbf{R}}(\boldsymbol{\Lambda} + \boldsymbol{\Gamma})\mathbf{V}_{\mathbf{R}}^T. \tag{7.33}$$

Therefore calculating the inverse of the reconditioned matrix yields

$$\mathbf{R}_{ME} = \mathbf{V}_{\mathbf{R}}(\boldsymbol{\Lambda} + \boldsymbol{\Gamma})^{-1}\mathbf{V}_{\mathbf{R}}^T. \tag{7.34}$$

As this is full rank we can calculate the inverse of the diagonal matrix $\boldsymbol{\Lambda} + \boldsymbol{\Gamma}$

$$(\boldsymbol{\Gamma} + \boldsymbol{\Lambda})^{-1}(i,i) = \begin{cases} \frac{1}{\lambda_i} & if \lambda_i \geq T \\ \frac{1}{\lambda_i + (T - \lambda_i)} & if \lambda_i < T \end{cases} \tag{7.35}$$

$$= \boldsymbol{\Lambda}^{-1} - \begin{cases} 0 & if \lambda_i \geq T \\ \frac{T - \lambda_i}{T \lambda_i} & if \lambda_i < T. \end{cases} \tag{7.36}$$

Defining $\tilde{\boldsymbol{\Gamma}}$ as in the theorem statement, and we can write $\mathbf{R}_{ME}^{-1}$ as

$$\mathbf{R}_{ME}^{-1} = \mathbf{R}^{-1} - \mathbf{V}_{\mathbf{R}}\tilde{\boldsymbol{\Gamma}}\mathbf{V}_{\mathbf{R}}^T. \tag{7.37}$$

Substituting this into the definition of the objective function (7.25) we obtain the result given in the theorem statement. □

As $\tilde{\boldsymbol{\Gamma}}$ is non-zero only for eigenvalues smaller than the threshold, T, the final term of the updated objective function (7.31) reduces the weight on eigenvectors corresponding to those small eigenvalues. As all the entries of $\tilde{\boldsymbol{\Gamma}}$ are non-negative, the size of the observation term in the original objective function (7.25) is decreased when ME is used.

Finally we consider the impact on the objective function of using MVI. We note that this can only be applied in the case that the estimated error covariance matrix is invertible as, by the result of Theorem 7.4.5, variance inflation cannot change the rank of a matrix.

Theorem 7.4.10. *In the case that* $\mathbf{R}$ *is invertible, the application of MVI to the observation error covariance matrix alters the objective function* (7.25) *as follows*

$$J_{MVI}(\mathbf{x}) = J_b + \frac{1}{\alpha^2}J_o. \tag{7.38}$$

Proof. By Definition 7.3.2, $\mathbf{R}_{MVI} = \alpha^2\mathbf{R}$ for inflation parameter α. The inverse of

$\mathbf{R}_{MVI}$ is given by

$$\mathbf{R}_{MVI}^{-1} = \frac{1}{\alpha^2}\mathbf{R}^{-1}. \tag{7.39}$$

Substituting this into (7.25) yields the updated objective function given by (7.38). □

For both reconditioning methods, the largest relative changes to the spectrum of $\mathbf{R}$ occur for its smallest eigenvalues. In the case of positive spatial correlations, small eigenvalues are typically sensitive to smaller scales. For spatial correlations, weights on scales of the observations associated with smaller eigenvalues are reduced in the variational objective function, increasing the relative sensitivity of analysis to information content from the observations at large scales.

We also see that for RR and ME smaller choices of κ_{max} yield larger reductions to the weight applied to small scale observation information. For RR, a smaller target condition number results in a larger value of δ, and hence larger diagonal entries of $\mathbf{\Lambda}_\delta$. For ME, a smaller target condition number yields a larger threshold, T, and hence larger diagonal entries of $\tilde{\mathbf{\Gamma}}$. This means that the more reconditioning that is applied, the less weight the observations will have in the analysis. This reduction in observation weighting is different for the two methods; RR reduces the weight on all observations, although the relative effect is larger for scales corresponding to the smallest eigenvalues, whereas ME only reduces weight for scales corresponding to eigenvalues smaller than the threshold T. In ME, the weights on scales for eigenvalues larger than T are unchanged.

Applying MVI with a constant inflation factor also reduces the contribution of observation information to the analysis. In contrast to both methods of reconditioning, the reduction in weight is constant for all scales and does not depend on the eigenvectors of $\mathbf{R}$. This means that there is no change to the sensitivity to different scales using this method. The analysis will simply pull closer to the background data with the same relative weighting between different observations as occurred for analyses using the original estimated observation error covariance matrix.

We have considered the impact of RR, ME and MVI on the unpreconditioned 3D-Var objective function. For the preconditioned case, Johnson et al. [2005] showed how, when changing the relative weights of the background and observation terms by inflating the ratio of observation and background variances, it is the complex interactions between the error covariance matrices and the observation operator that affects which scales are present in the analysis. This suggests that in the preconditioned setting MVI will also alter the sensitivity of the analysis to different scales.

7.5 Numerical experiments

In this section we consider how reconditioning via RR and ME and application of MVI affects covariance matrices arising from two different choices of estimated covariance matrices. Both types of covariance matrix are motivated by numerical weather prediction, although similar structures occur for other applications.

7.5.1 Numerical framework

The first covariance matrix is constructed using a second-order auto-regressive (SOAR) correlation function [Yaglom, 1986] with lengthscale 0.2 on a unit circle. This correlation function is used in NWP systems [Fowler et al., 2018, Stewart et al., 2013, Tabeart et al., 2018, Waller et al., 2016b, Thiebaux, 1976] where its long tails approximate the estimated horizontal spatial correlation structure well. In order to construct a SOAR error correlation matrix, S, on the finite domain, we follow the method described in Haben [2011], Tabeart et al. [2018]. We consider a one-parameter periodic system on the real line, defined on an equally spaced grid with $N = 200$ grid points. We restrict observations to be made only at regularly spaced grid points. This yields a circulant matrix where the matrix is fully defined by its first row. To ensure the corresponding covariance matrix is also circulant, we fix the standard deviation value for all variables to be $\sigma = \sqrt{5}$.

One benefit of using this numerical framework is that it allows us to calculate a simple expression for the update to the standard deviations for ME. We recall that RR updates the variances by a constant, δ. We now show that in the case where $\mathbf{R}$ is circulant, ME also updates the variances of $\mathbf{R}$ by a constant.

Circulant matrices admit eigenvectors which can be computed directly via a discrete Fourier transform [Gray, 2006] (via $\mathbf{R} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\dagger$, where † denotes conjugate transpose). This allows the explicit calculation of the ME standard deviation update given by (7.16) as

$$\boldsymbol{\Sigma}_{ME}(i,i) = \left(\boldsymbol{\Sigma}(i,i) + \frac{1}{d}\sum_{k=1}^{d}\boldsymbol{\Gamma}(k,k)\right)^{1/2}. \tag{7.40}$$

This follows from (7.16) because the circulant structure of the SOAR matrix yields

$\sum_{k=1}^{d} \mathbf{V}(i,k)^2 = 1/d$.

We therefore expect both reconditioning methods to increase the SOAR standard deviations by a constant amount. As the original standard deviations were constant, this means that reconditioning will result in constant standard deviations for all variables. These shall be denoted σ_{RR} for RR and σ_{ME} for ME. Constant changes to standard deviations also means that an equivalent MVI factor that corresponds to the change can be calculated. This will be denoted by α.

Our second covariance matrix comprises interchannel error correlations for a satellite-based instrument. For this we make use of the Infrared Atmospheric Sounding Interferometer (IASI) which is used at many NWP centres within data assimilation systems. A covariance matrix for IASI was diagnosed in 2011 at the Met Office, following the procedure described in Weston [2011], Weston et al. [2014] (shown in Online Resource 1). The diagnosed matrix was extremely ill-conditioned and required the application of the ridge regression method in order that the correlated covariance matrix could be used in the operational system. We note that we follow the reconditioning procedure of Weston et al. [2014], where the reconditioning method is only applied to the subset of 137 channels that that are used in the Met Office 4D-Var system. These channels are listed in Stewart et al. [2008a, Appendix A]. As the original standard deviation values are not constant across different channels, reconditioning will not change them by a constant amount, as is the case for Experiment 1. We note that the 137 × 137 matrix considered in this chapter corresponds to the covariance matrix for one 'observation' at a single time and spatial location. The observation error covariance matrix for all observations from this instrument within a single assimilation cycle is a block diagonal matrix, with one block for every observation, each consisting of a submatrix of the 137 × 137 matrix .

In the experiments presented in Section 7.5.2 we apply the minimum eigenvalue and the ridge regression methods to both the SOAR and IASI covariance matrices. The condition number before reconditioning of the SOAR correlation matrix is 81121.71 and for the IASI matrix we obtain a condition number of 2005.98. We consider values of κ_{max} in the range $100 - 1000$ for both tests. We note that the equivalence of the minimum eigenvalue method with the minimiser of the Ky Fan $1 - d$ norm is satisfied for the SOAR experiment for $\kappa_{max} \geq 168$ and the IASI experiment for $\kappa_{max} \geq 98$.

Table 7.1: Change in standard deviation for the SOAR covariance matrix for both methods of reconditioning. Columns 4 and 6 show α, the multiplicative inflation factor corresponding to the values for σ_{RR} and σ_{ME} respectively.

κ_{max}	σ	σ_{RR}	α corr. RR	σ_{ME}	α corr. ME
1000	2.23606	2.26471	1.013	2.25439	1.008
500	2.23606	2.29340	1.026	2.27599	1.018
100	2.23606	2.51306	1.124	2.45737	1.099

7.5.2 Results

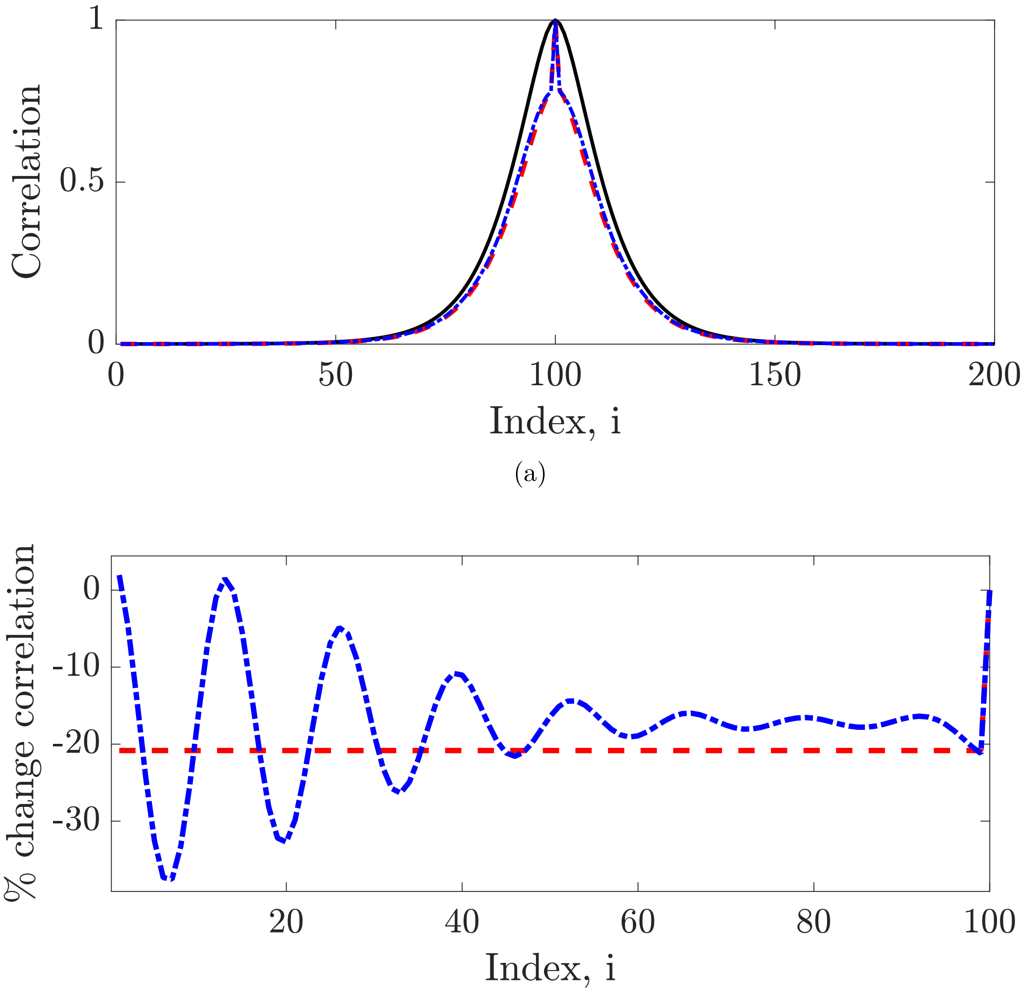
7.5.2.1 Changes to the covariance matrix

Example 1: Horizontal correlations using a SOAR correlation matrix
Due to the specific circulant structure of the SOAR matrix and constant value of standard deviations for all variables, (7.11) and (7.40) indicate that we expect increases to standard deviations for both methods of reconditioning to be constant. This was found to be the case numerically. In Table 7.1 the computed change in standard deviation for different values of κ_{max} is given as an absolute value and as α, the multiplicative inflation constant that yields the same change to the standard deviation as each reconditioning method. We note that in agreement with the result of Corollary 7.4.6 the variance increase is larger for the RR than the ME for all choices of κ_{max}. Reducing the value of κ_{max} increases the change to standard deviations for both methods of reconditioning. The increase to standard deviations will result in the observations being down-weighted in the analysis. As this occurs uniformly across all variables for both methods, we expect the analysis to pull closer to the background. Nevertheless, we expect this to be a rather small effect. For this example, even for a small choice of κ_{max} the values of the equivalent multiplicative inflation constant, α, is small, with the largest value of $\alpha = 1.124$ occurring for RR for $\kappa_{max} = 100$.

As the SOAR matrix is circulant, we can consider the impact of reconditioning on its correlations by focusing on one matrix row. In Figure 7.1 the correlations and percentage change for the 100th row of the SOAR matrix are shown for both methods for $\kappa_{max} = 100$. These values are calculated directly from the reconditioned matrix. We note that by definition of a correlation matrix, $\mathbf{C}(i, i) = 1 \; \forall \; i$ for all choices of reconditioning. This is the reason for the spike in correlation visible in the centre of Figure 7.1a and on the right of Figure 7.1b. As multiplicative variance inflation does not change the correlation matrix, the black line corresponding to the correlations of the original SOAR matrix also represents the correlations in the case of multiplicative inflation. We also remark that although ME is not equivalent to the minimiser of the Ky Fan $1 - d$ norm for $\kappa_{max} = 100$, the qualitative behaviour in terms of correlations and standard deviations is the same for all values in the range $100 - 1000$. It is important to note that ME is still a well-defined method of reconditioning even if it is not equivalent to the minimiser of the Ky Fan $1 - d$ norm.

Figure 7.1a shows that for both methods, application of reconditioning reduces the value of off-diagonal correlations for all variables, with the largest absolute reduction occurring for variables closest to the observed variable. Although there is a large change to the off-diagonal correlations, we notice that the correlation lengthscale, which determines the rate of decay of the correlation function, is only reduced by a small amount. This shows that both methods of reconditioning dampen correlation values but do not significantly alter the overall pattern of correlation information. Figure 7.1b shows the percentage change to the original correlation values after reconditioning is applied. For RR, although the difference between the original correlation value and the reconditioned correlation depends on the index i, the relative change is constant across all off-diagonal correlations. As MVI does not alter the correlation matrix, it would correspond to a horizontal line through 0 for Figure 7.1b.

When we directly plot the correlation values for the original and reconditioned matrices in Figure 7.1a, the change to correlations for ME appears very similar to changes for RR. However, when we consider the percentage change to correlation in Figure 7.1b we see oscillation in the percentage differences of the ME correlations, showing that the relative effect on some spatially distant variables can be larger than for some spatially close variables. The spatial impact on individual variables differs significantly for this method. We also note that ME increases some correlation values. These are not visible in Figure 7.1, due to entries in the original correlation matrix that are close to zero. Although the differences between $\mathbf{C}$ and $\mathbf{C}_{ME}$ far from the diagonal are small, small correlation values in the tails of the original SOAR matrix mean that when considering the percentage difference we obtain large values, as seen in Figure 7.1b. This suggests that RR is a more appropriate method to use in this context, as the reconditioned matrix represents the initial correlation function better than ME, where spurious oscillations are introduced. These oscillations occur as ME changes the weighting of eigenvectors of the covariance matrix. As the eigenvectors of circulant matrices can be expressed in terms of discrete Fourier modes, ME has the effect of amplifying the eigenvalues corresponding to the highest frequency eigenvectors. This results in the introduction of spurious oscillations in correlation



(a)

(b)

Figure 7.1: Changes to correlations between the original SOAR matrix and the reconditioned matrices for $\kappa_{max} = 100$. (a) shows $\mathbf{C}(100,:) = \mathbf{C}_{MVI}(100,:)$ (black solid), $\mathbf{C}_{RR}(100,:)$ (red dashed), $\mathbf{C}_{ME}(100,:)$ (blue dot-dashed) (b) shows $100 \times \frac{\mathbf{C}(100,:)-\mathbf{C}_{RR}(100,:)}{\mathbf{C}(100,:)}$ (red dashed) and $100 \times \frac{\mathbf{C}(100,:)-\mathbf{C}_{ME}(100,:)}{\mathbf{C}(100,:)}$ (blue dot-dashed). As the SOAR matrix is symmetric, we only plot the first 100 entries for (b).

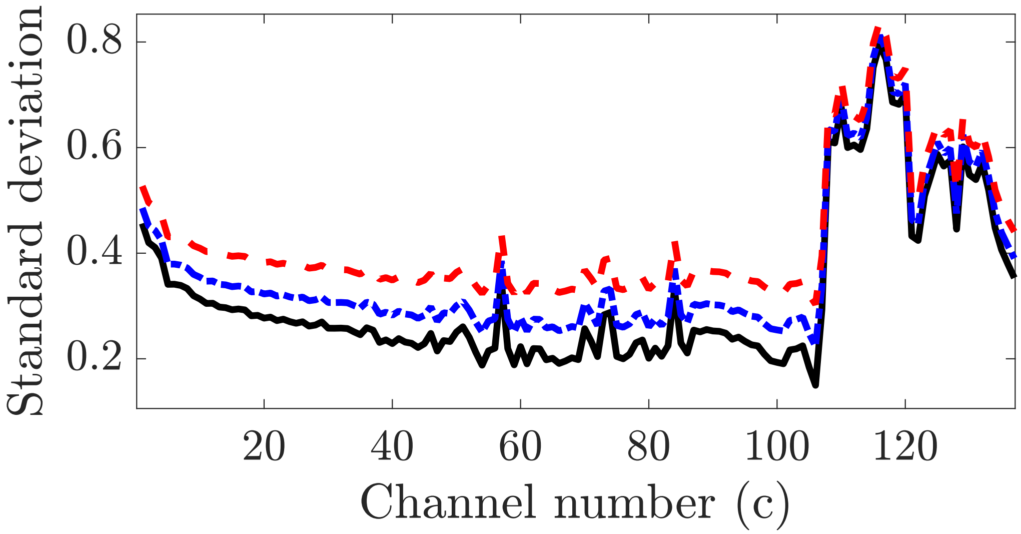


Figure 7.2: Standard deviations for the IASI covariance matrix $\mathbf{\Sigma}$ (black solid), $\mathbf{\Sigma}_{RR}$ (red dashed), $\mathbf{\Sigma}_{ME}$ (blue dot-dashed) for $\kappa_{max} = 100$.

space.

Both methods reduce the correlation lengthscale of the error covariance matrix. In Tabeart et al. [2018], it was shown that reducing the lengthscale of the observation error covariance matrix decreases the condition number of the Hessian of the 3D-Var objective function and results in improved convergence of the minimisation problem. Hence the application of reconditioning methods to the observation error covariance matrix is likely to improve convergence of the overall data assimilation problem. Fowler et al. [2018] studied the effect on the analysis of complex interactions between the background error correlation lengthscale, the observation error correlation lengthscale and the observation operator in idealised cases. Their findings for a fixed background error covariance, and direct observations, indicate that the effect of reducing the observation error correlation lengthscale (as in the reconditioned cases) is to increase the analysis sensitivity to the observations at larger scales. In other words, more weight is placed on the large-scale observation information content and less weight on the small scale observation information content. This corresponds with the findings of Section 7.4.5, where we proved that both methods of reconditioning reduce the weight on small scale observation information in the variational objective function. However, the lengthscale imposed by a more complex observation operator could modify these findings.

Example 2: Interchannel correlations using an IASI covariance matrix

We now consider the impact of reconditioning on the IASI covariance matrix. We note that there is significant structure in the diagnosed correlations (see Stewart et al. [2014, Fig. 8] and Section 7.8), with blocks of highly correlated channels in the lower right hand region of the matrix. We now consider how RR, ME and MVI change the variances and correlations of the IASI matrix.

Figure 7.2 shows the standard deviations $\mathbf{\Sigma}$, $\mathbf{\Sigma}_{RR}$ and $\mathbf{\Sigma}_{ME}$. These are calculated from the reconditioned matrices, but the values coincide with the theoretical results of Theorems 7.4.2 and 7.4.4. Standard deviation values for the original diagnosed case have been shown to be close to estimated noise characteristics of the instrument for each of the different channels [Stewart et al., 2014]. We note that the largest increase to standard deviations occurs for channel 106 only and corresponds to a multiplicative inflation factor for this channel of 2.02 for RR and 1.81 for ME. Channel 106 is sensitive to water vapour and is the channel in the original diagnosed covariance matrix with the smallest standard deviation. The choice of $\kappa_{max} = 100$ is of a similar size to the value of the parameters used at NWP centres [Weston, 2011, Weston et al., 2014, Bormann et al., 2016]. This means that in practice, the contribution of observation information from channels where instrument noise is low is being substantially reduced.

Channels are ordered by increasing wavenumber, and are grouped by type. We expect different wavenumbers to have different physical properties, and therefore different covariance structures. In particular larger standard deviations are expected for higher wavenumbers due to additional sources of error [Weston et al., 2014], which is observed on the right hand side of Figure 7.2. For RR, larger increases to standard deviations are seen for channels with smaller standard deviations for the original diagnosed matrix than those with large standard deviations. This also occurs to some extent for ME, although we observe that the update term in (7.16) is not constant in this case. This means that the reduction in weight in the analysis will not be uniform across different channels for ME. The result of Corollary 7.4.6 is satisfied; the increase to the variances is larger for RR than ME. This is particularly evident for channels where the variance from the original diagnosed covariance matrix is small. As MVI increases standard deviations by a constant factor, the largest changes for this method would occur for channels with large standard deviations in the original diagnosed matrix. This is in contrast to RR, where the largest changes occur for the channels in the original diagnosed matrix with the smallest standard deviation.

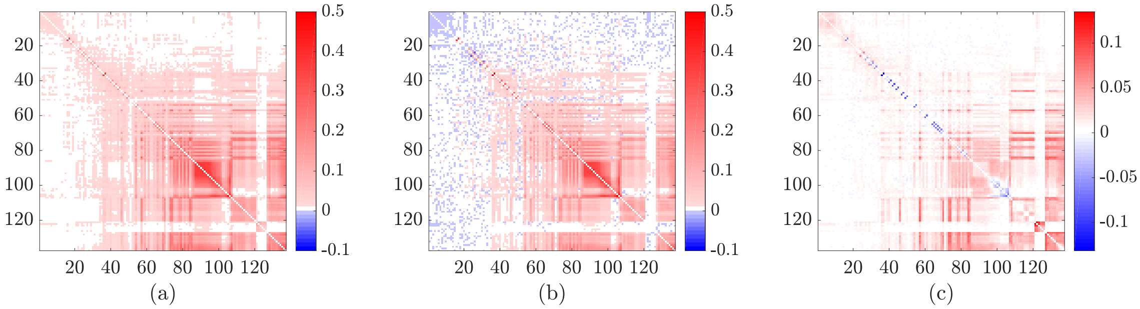


Figure 7.3: Difference in correlations for IASI (a) $(\mathbf{C}-\mathbf{C}_{RR})\circ sign(\mathbf{C})$, (b) $(\mathbf{C}-\mathbf{C}_{ME})\circ sign(\mathbf{C})$, and (c) $(\mathbf{C}_{ME}-\mathbf{C}_{RR})\circ sign(\mathbf{C})$, where $\circ$ denotes the Hadamard product. Red indicates that the absolute correlation is decreased by reconditioning and blue indicates the absolute correlation is increased. The colourscale is the same for (a) and (b) but different for (c). Condition numbers of the corresponding covariance matrices are given by $\kappa(\mathbf{R}) = 2005.98$, $\kappa(\mathbf{R}_{RR}) = 100$ and $\kappa(\mathbf{R}_{ME}) = 100$.

Figure 7.3 shows the difference between the diagnosed correlation matrix, $\mathbf{C}$, and the reconditioned correlation matrices $\mathbf{C}_{RR}$ and $\mathbf{C}_{ME}$. As some correlations in the original IASI matrix are negative, we plot the entries of $(\mathbf{C}-\mathbf{C}_{RR})\circ sign(\mathbf{C})$ and $(\mathbf{C}-\mathbf{C}_{ME})\circ sign(\mathbf{C})$ in Figures 7.3a and b respectively. Here $\circ$ denotes the Hadamard product, which multiplies matrices of the same dimension elementwise. This allows us to determine whether the magnitude of the correlation value is reduced by the reconditioning method; a positive value indicates that the reconditioning method reduces the magnitude of the correlation, whereas a negative value indicates an increase in the correlation magnitude. For RR, all differences are positive, which agrees with the result of Theorem 7.4.2. As MVI does not change the correlation matrix, an equivalent figure for this method is not given. We also note that there is a recognisable pattern in Figure 7.3a, with the largest reductions occurring for the channels in the original diagnosed correlation matrix which were highly correlated. This indicates that this method of reconditioning does not affect all channels equally.

For ME, we notice that there are a number of entries where the absolute correlations are increased after reconditioning. There appears to be some pattern to these entries, with a large number occurring in the upper left hand block of the matrix for channels with the smallest wavenumber [Weston et al., 2014]. However, away from the diagonal for channels 0-40, where changes by RR are very small, the many entries where absolute correlations are increased by ME are much more scattered. This more noisy change to the correlations could be due to the fact that 96 eigenvalues are set to be equal to a threshold value by the minimum eigenvalue method in order to attain $\kappa_{max} = 100$. One method to reduce noise was suggested in Smith et al. [2018], which showed that applying localization methods (typically used to reduce spurious long-distance correlations that arise when using ensemble covariance matrices via the Schur product) after the reconditioning step can act to remove noise while retaining covariance structure.

For positive entries, the structure of $\mathbf{C}_{ME}$ appears similar to that of $\mathbf{C}_{RR}$. There are some exceptions however, such as the block of channels 121-126 where changes in correlation due to ME are small, but correlations are changed by quite a large amount for RR. The largest elementwise difference between RR and the original diagnosed correlation matrix is 0.138, whereas the largest elementwise difference between ME and the original diagnosed correlation matrix is 0.0036. The differences between correlations for ME and RR are shown in Figure 7.3c.

For both methods, although the absolute value of all correlations is reduced, correlations for channels 1-70 are eliminated. This has the effect of emphasising the correlations for channels that are sensitive to water vapour. Weston et al. [2014], Bormann et al. [2016] argue that much of the benefit of introducing correlated observation error for this instrument can be related to the inclusion of correlated error information for water vapour sensitive channels. Therefore, although the changes to the original diagnosed correlation matrix are large it is likely that a lot of the benefit of using correlated observation error matrices is retained.

We also note that it is more difficult to choose the best reconditioning method in this setting, due to the complex structure of the original diagnosed correlation matrix. In particular, improved understanding of how each method alters correlations and standard deviations is not enough to determine which method will perform best in an assimilation system. One motivation of reconditioning is to improve convergence of variational data assimilation algorithms. Therefore, one aspect of the system that can be used to select the most appropriate method of reconditioning is the speed of convergence. As ME results in repeated eigenvalues, we would expect faster convergence of conjugate gradient methods applied to the problem $\mathbf{R}\mathbf{x} = \mathbf{b}$ for $\mathbf{x}, \mathbf{b} \in \mathbb{R}^d$ for ME than RR. However, Campbell et al. [2017], Weston [2011], Weston et al. [2014], Bormann et al. [2015] find that RR results in faster convergence than ME for operational variational implementations. This is likely due to interaction between the reconditioned observation error covariance matrix and the observation operator, as the eigenvalues of $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ are shown to be important for the conditioning of the variational data assimilation problem in Tabeart et al. [2018].

Another aspect of interest is the influence of reconditioning on the analysis and forecast performance. We note that this is likely to be highly system and metric dependent. For example, Campbell et al. [2017] studies the impact of reconditioning on predictions of meteorological variables (temperature, geopotential height, precipitable water) over lead times from 0 to 5 days. In the U.S. Naval Research Laboratory system, ME performed slightly better at short lead times, whereas RR had improvements at longer lead times [Campbell et al., 2017]. Differences in forecast performance were mixed, whereas convergence was much faster for RR. This meant that the preferred choice was RR. However, in the ECMWF system, Bormann et al. [2015] studied the standard deviation of first-guess departures against independent observations. Using this metric of analysis skill, ME was found to out-perform RR. The effect of RR on the analysis of the Met Office 1D-Var system is studied in Tabeart et al. [2019b], where changes to retrieved variables sensitive to water vapour (humidity, variables sensitive to cloud) are found to be larger than for other meteorological variables such as temperature.

7.5.2.2 Changes to the analysis of a data assimilation problem

In Section 7.4.5 we considered how the variational objective function is altered by RR, ME and MVI. We found that the two methods of reconditioning reduced the weight on scales corresponding to small eigenvalues by a larger amount than MVI, which changes the weight on all scales uniformly. In this section we consider how the analysis of an idealised data assimilation problem is altered by each of the three methods. We also consider how changing κ_{max} alters the analysis of the problem.

In order to compare the three methods, we study how the solution of a conjugate gradient method applied to the linear system $\mathbf{S}\mathbf{x} = \mathbf{b}$ changes for RR, ME and MVI, where $\mathbf{S} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ is the linearised Hessian associated with the 3D-Var objective function (7.25). Haben [2011] showed that this is equivalent to solving the 3D-Var objective function in the case of a linear observation operator. We define a background error covariance matrix, $\mathbf{B} \in \mathbb{R}^{200\times 200}$, which is a SOAR correlation matrix on the unit circle with correlation lengthscale 0.2 and a constant variance of 1. Our observation operator is given by the identity, meaning that every state variable is observed directly.

We construct a ‘true’ observation error covariance $\mathbf{R}_{true}$, given by a 200 dimensional SOAR matrix on the unit circle with standard deviation 1 and lengthscale 0.7. We

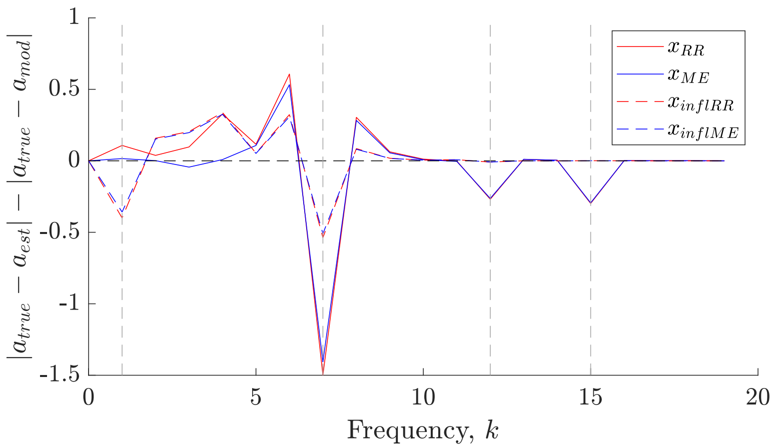


Figure 7.4: Change in pointwise difference of Discrete Fourier Transform (DFT) from $\mathbf{x}_{est}$ to $\mathbf{x}_{mod}$ where a_{est} denotes the vector of coefficients of the imaginary part of $DFT(\mathbf{x}_{est})$. A positive (negative) value indicates that $\mathbf{x}_{mod}$ is closer to (further from) $\mathbf{x}_{true}$ than $\mathbf{x}_{est}$ and the amplitude shows how large this change is. Vertical dashed lines show the locations of non-zero values for the true signal.

then take 250 random samples of $\mathbf{R}_{true}$ to construct an estimated sample covariance matrix $\mathbf{R}_{est}$ with $\kappa(\mathbf{R}_{est}) = 3.95 \times 10^8$. The largest estimated standard deviation is 1.07 and the smallest is 0.90, compared to the true constant standard deviation of 1. RR, ME and MVI are then applied to $\mathbf{R}_{est}$ with $\kappa_{max} = 100$. When applying MVI, we use two choices of α which correspond to changes to the standard deviations $(\mathbf{R}_{RR}(1,1))^{1/2}$, $\alpha_{RR} = 1.41$, and $(\mathbf{R}_{ME}(1,1))^{1/2}$, $\alpha_{ME} = 1.39$. The modified error covariance matrices will be denoted $\mathbf{R}_{inflRR} = \alpha_{RR}^2\mathbf{R}_{est}$ and $\mathbf{R}_{inflME} = \alpha_{ME}^2\mathbf{R}_{est}$.

We define a true state vector,

$$\mathbf{x}(k) = 4\sin(k\pi/100) - 5.1\sin(7k\pi/100) + 1.5\sin(12k\pi/100) - 3\sin(15k\pi/100) + 0.75\sin(45k\pi/100), \tag{7.41}$$

which has five scales. We then construct $\mathbf{b} \equiv \mathbf{S}\mathbf{x}$ using $\mathbf{R}_{true}$, and apply the Matlab 2018b *pcg.m* routine to the problem $(\mathbf{B}^{-1} + \mathbf{R}^{-1})\mathbf{x} = \mathbf{b}$ for each choice of $\mathbf{R}$. We recall that $\mathbf{S} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} = \mathbf{B}^{-1} + \mathbf{R}^{-1}$ as $\mathbf{H} = \mathbf{I}$. Let $\mathbf{x}_{est}$ denote the solution that is found using $\mathbf{R}_{est}$ and $\mathbf{x}_{mod}$ refer to a solution found using a modified version of $\mathbf{R}_{est}$, namely $\mathbf{R}_{RR}$, $\mathbf{R}_{ME}$, $\mathbf{R}_{inflRR}$ or $\mathbf{R}_{inflME}$. The maximum number of iterations allowed for the conjugate gradient routine is 200, and convergence is reached when the relative residual is less than 1×10^{-6}.

From Section 7.4.5 we expect RR, ME and MVI to behave differently at small and large scales. We therefore analyse how using each method alters the solution $\mathbf{x}$ at different scales using the discrete Fourier transform (DFT). This allows us to assess how well each scale of $\mathbf{x}_{true}$ is recovered for each choice of $\mathbf{R}$. As $\mathbf{x}_{true}$ is the sum of sine functions, only the imaginary part of the DFT will be non-zero. We therefore define $a_{true} = \text{imag}(DFT(x_{true}))$; similarly $a_{est} = \text{imag}(DFT(x_{est}))$ and $a_{mod} = \text{imag}(DFT(x_{mod}))$.

By construction, as $\mathbf{x}$, given by (7.41), is the sum of sine functions of period $2\pi n/200$ for $n = 1, 7, 12, 15, 45$, a_{true} returns a signal with 5 peaks, one for each value of n at frequency $k = n$. The amplitude for all other values of k is zero. For frequencies larger than 20, all choices of estimated and modified $\mathbf{R}$ recover a_{true} well. Figure 7.4 shows the correction that is applied by the modified choices of $\mathbf{R}$ compared to $\mathbf{R}_{est}$ for the first 20 frequencies. A positive (negative) value shows that a_{mod} moves closer to (further from) a_{true} than a_{est}. The distance from 0 shows the size of this change. For the first true peak ($k = 1$) RR is able to move closer to a_{true} than a_{est} for the first true peak. However, both reconditioning methods move further from the truth at the location of true signals $k = 7, 12, 15$. For frequencies where a_{est} has a spurious non-zero signal RR and ME are able to move closer to a_{true} than a_{est}. At the location of true signals $k = 7, 12, 15$, MVI makes smaller changes compared to a_{est} than either method of reconditioning. As all modifications to $\mathbf{R}_{est}$ move a_{mod} further from a_{true} than a_{est} for $k = 7, 12, 15$, MVI is therefore better able to recover the value of a_{true} than RR or ME at these true peaks. However, MVI introduces a larger error for the first peak at $k = 1$ than RR or ME, and changes for frequencies $k > 5$ are much smaller than for reconditioning. This agrees with the findings of Section 7.4.5, that the weight on all scales is changed equally by MVI, whereas both methods of reconditioning result in larger changes to smaller scales and are hence able to make larger changes to amplitudes for higher frequencies. We recall from Section 7.4.5 that ME changes only the smaller scales, whereas RR also makes small changes to the larger scales. This behaviour is seen in Figure 7.4: for frequencies $k = 0$ to 5 ME results in very small changes, with much larger changes for frequencies $5 \leq k \leq 15$. RR makes larger changes for larger values of k, but also moves closer to a_{true} for $1 \leq k \leq 3$.

We now consider how changing κ_{max} alters the quality of $\mathbf{x}_{RR}$. As the behaviour for $\kappa_{max} = 100$ shown in Figure 7.4 was similar for both RR and ME, we only consider changes to RR. Figure 7.5 shows the difference between a_{true} and a_{RR} for different choices of κ_{max}. Firstly we consider the true signal that occurs at frequencies $k = 1, 7, 12, 15$. For $k = 1$ the smallest error occurs for $\kappa_{max} = 50$ and the largest error

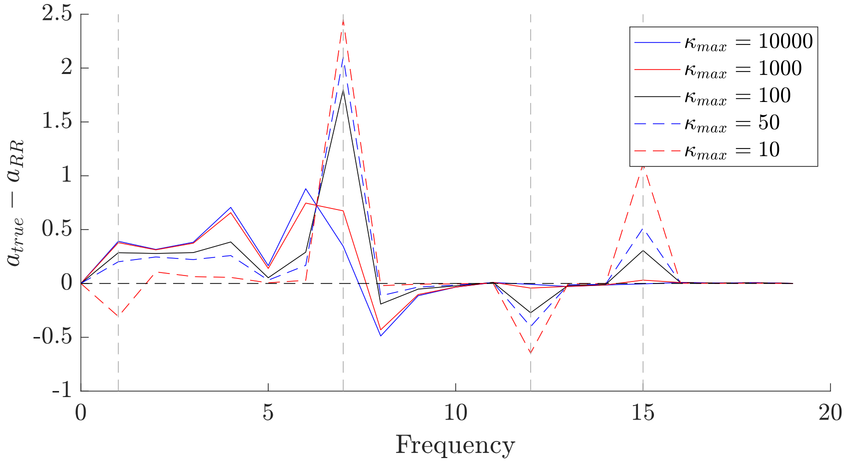


Figure 7.5: Difference between a_{true} and a_{RR} for different choices of κ_{max}. Vertical dashed lines show the locations of non-zero values for the true signal.

Table 7.2: Changes to convergence of RR, MI and MVI for different values of κ_{max}. For all choices of κ_{max}, convergence for $\mathbf{R}_{true}$ occurs in 17 iterations and $\mathbf{R}_{est}$ occurs in 244 iterations.

κ_{max}	$10,000$	$1,000$	100	50	10
RR	245	244	170	141	73
ME	240	239	193	145	76
Infl RR	244	244	238	233	199

occurs for $\kappa_{max} = 10000$. For $k = 7, 12, 15$ the error increases as κ_{max} decreases. For all other frequencies, reducing κ_{max} reduces the error in the spurious non-zero amplitudes. For very large values of κ_{max} we obtain small errors for the true signal, but larger spurious errors elsewhere. Very small values of κ_{max} can control these spurious errors, but fail to recover the correct amplitude for the true signal. Therefore a larger reconditioning constant will result in larger changes to the analysis. This means that there is a balance to be made in ensuring the true signal is captured, but spurious signal is depressed. For this framework a choice of $\kappa_{max} = 100$ provides a good compromise between recovering the true peaks well and suppressing spurious correlations.

Finally, Table 7.2 shows how convergence of the conjugate gradient method is altered by the use of reconditioning and MVI. Using a larger inflation constant does lead to slightly faster convergence compared to $\mathbf{R}_{est}$. However, reducing κ_{max} leads to a much larger reduction in the number of iterations required for convergence for both RR and ME. This agrees with results in operational data assimilation systems, where the choice of κ_{max} and reconditioning method makes a difference to convergence Weston [2011], Tabeart et al. [2019b].

7.6 Conclusions

Applications of covariance matrices often arise in high dimensional problems [Pourahmadi, 2013], such as numerical weather prediction (NWP) [Bormann et al., 2016, Weston et al., 2014]. In this chapter we have examined two methods that are currently used at NWP centres to recondition covariance matrices by altering the spectrum of the original covariance matrix: the ridge regression method, where all eigenvalues are increased by a fixed value, and the minimum eigenvalue method, where eigenvalues smaller than a threshold are increased to equal the threshold value. We have also considered multiplicative variance inflation, which does not change the condition number or rank of a covariance matrix, but is used at NWP centres [Bormann et al., 2016].

For both reconditioning methods we developed new theory describing how variances are altered. In particular, we showed that both methods will increase variances, and that this increase is larger for the ridge regression method. We also showed that applying the ridge regression method reduces all correlations between different variables. Comparing the impact of reconditioning methods and multiplicative variance inflation on the variational data assimilation objective function we find that all methods reduce the weight on observation information in the analysis. However, reconditioning methods have a larger effect on smaller eigenvalues, whereas multiplicative variance inflation does not change the sensitivity of the analysis to different scales. We then tested both methods of reconditioning and multiplicative variance inflation numerically on two examples: Example 1, a spatial covariance matrix, and Example 2, a covariance matrix arising from numerical weather prediction. In Section 7.5.2 we illustrated the theory developed earlier in the work, and also demonstrated that for two contrasting numerical frameworks, the change to the correlations and variances is significantly smaller for the majority of entries for the minimum eigenvalue method.

Both reconditioning methods depend on the choice of κ_{max}, an optimal choice of which will depend on the specific problem in terms of computational resource and required precision. The smaller the choice of κ_{max}, the more variances and correlations are altered, so it is desirable to select the largest condition number that the system of interest can deal with. Some aspects of a system that could provide insight into reasonable choices of κ_{max} are:

- For conjugate gradient methods, the condition number provides an upper bound on the rate of convergence for the problem $A\mathbf{x} = \mathbf{b}$ [Golub and Van Loan, 1996], and can provide an indication of the number of iterations required to reach a particular precision [Axelsson, 1996]. Hence κ_{max} could be chosen such that a required level of precision is guaranteed for a given number of iterations.

- For more general methods, the condition number can provide an indication of the number of digits of accuracy that are lost during computations [Gill et al., 1986, Cheney, 2005]. Knowledge of the error introduced by other system components, such as approximations in linearised observation operators and linearised models, relative resolution of the observation network and state variables, precision and calibration of observing instruments, may give insight into a value of κ_{max} that will maintain the level of precision of the overall problem.

- The condition number measures how errors in the data are amplified when inverting the matrix of interest [Golub and Van Loan, 1996]. Again, the magnitude of errors resulting from other aspects of the system may give an indication of a value of κ_{max} that will not dominate the overall precision.

For our experiments we considered choices of κ_{max} in the range $100 - 1000$. For Experiment 2 these values are similar to those considered for the same instrument at different NWP centres e.g. 25, 100, 1000 [Weston, 2011], 67 [Weston et al., 2014], 54 and 493 [Bormann et al., 2015], 169 [Campbell et al., 2017]. We note that the dimension of this interchannel error covariance matrix in operational practice is small and only forms a small block of the full observation error covariance matrix. Additionally, the matrix considered in this chapter corresponds to one observation type; there are many other observation types with different error characteristics.

In this work we have assumed that our estimated covariance matrices represent the desired correlation matrix well, in which case the above conditions on κ_{max} can be used. This is not true in general, and it may be that methods such as inflation and localisation are also required in order to constrain the sources of uncertainty that are underestimated or mis-specified. In this case, the guidance we have presented in this chapter concerning how to select the most appropriate choice of reconditioning method and target condition number will need to be adapted. Additionally, localisation alters the condition number of a covariance matrix as a side effect; the user does not have the ability to choose the target condition number κ_{max} or control changes to the distribution of eigenvalues [Smith et al., 2018]. This indicates that reconditioning may still be needed in order to retain valuable correlation information whilst ensuring that the computation of the inverse covariance matrix is feasible.

The choice of which method is most appropriate for a given situation depends on the system being used, and knowledge of its 'true' error statistics. The ridge regression method preserves eigenstructure by increasing the weight of all eigenvalues by the same amount, compared to the minimum eigenvalue method which only increases the weight of small eigenvalues and introduces a large number of repeated eigenvalues. We have found that ridge regression results in constant changes to variances and strict decreases to absolute correlation values, whereas the minimum eigenvalue method makes smaller, non-monotonic changes to correlations and non-constant changes to variances. In the spatial setting, the minimum eigenvalue method introduced spurious correlations, whereas ridge regression resulted in a constant percentage reduction for all variables. In the inter-channel case, changes to standard deviations and most correlations were smaller for the minimum eigenvalue method than for ridge regression.

Another important property for reconditioning methods is the speed of convergence of minimisation of variational data assimilation problems. It is well-known that other aspects of matrix structure, such as repeated or clustered eigenvalues, are important for the speed of convergence of conjugate gradient minimisation problems. As the condition number is only sensitive to the extreme eigenvalues, conditioning alone cannot fully characterise the expected convergence behaviour. In the data assimilation setting, complex interactions occur between the constituent matrices [Tabeart et al., 2018], which can make it hard to determine the best reconditioning method a priori. One example of this is seen for operational implementations in Campbell et al. [2017], Weston [2011] where the ridge regression method results in fewer iterations for a minimisation procedure than the minimum eigenvalue method, even though the minimum eigenvalue method yields observation error covariance matrices with a large number of repeated eigenvalues. Furthermore, Tabeart et al. [2018] found cases in an idealised numerical framework where increasing the condition number of the Hessian of the data assimilation problem was linked to faster convergence of the minimisation procedure. Again, this was due to interacting eigenstructures between observation and background terms, which could not be measured by the condition number alone.

Additionally, Haben [2011], Tabeart et al. [2018] find that the ratio of background to observation error variance is important for the convergence of a conjugate gradient problem. In the case where observation errors are small, poor performance of conjugate gradient methods is therefore likely. This shows that changes to the analysis of data assimilation problems due to the application of reconditioning methods are likely to be highly system dependent, for example due to: quality of estimated covariance matrices, interaction between background and observation error covariance matrices, specific implementations of the assimilation algorithm, and choice of preconditioner and minimisation routine. However, the improved understanding of alterations to correlations and standard deviations for each method of reconditioning provided here may allow users to anticipate changes to the analysis for a particular system of interest using the results from previous idealised and operational studies (e.g. Tabeart et al. [2018], Fowler et al. [2018], Simonin et al. [2019], Weston et al. [2014], Bormann et al. [2016]).

7.7 Appendix: Equivalence of the minimum eigenvalue method with the Ky Fan 1-d norm method

We introduce the Ky-Fan $p-k$ norm. We show that the solution to a nearest matrix problem in the Ky-Fan $1-d$ norm, where $\mathbf{X} \in \mathbb{R}^{d\times d}$, is equivalent to the minimum eigenvalue method of reconditioning introduced in Section 7.3.2 with an additional assumption.

Definition 7.7.1. *The Ky Fan p-k norm of* $\mathbf{X} \in \mathbb{C}^{m\times n}$ *is defined as:*

$$||\mathbf{X}||_{p,k} = \left(\sum_{i=1}^{k} \gamma_i(\mathbf{X})^p\right)^{1/p}, \tag{7.42}$$

where $\gamma_i(\mathbf{X})$ *denotes the i-th largest singular value of* $\mathbf{X}$*,* $p \geq 1$ *and* $k \in \{1, \dots, \min\{m, n\}\}$*.*

As covariance matrices are positive semi-definite by definition, the singular values of a covariance matrix $\mathbf{X} \in \mathbb{R}^{d\times d}$ are equal to its eigenvalues.

Theorem 7.7.2. *Let* $\mathbf{X} \in \mathbb{R}^{d\times d}$ *be a symmetric positive semi-definite matrix, with eigenvalues* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq 0$ *and corresponding matrix of eigenvectors given by*

$\mathbf{V_R}$. *The choice of* $\hat{\mathbf{X}}$ *that minimises*

$$||\mathbf{X} - \hat{\mathbf{X}}||_{1,p}, \tag{7.43}$$

subject to the condition $\kappa(\hat{\mathbf{X}}) = \hat{\kappa}$, *for* $\hat{\kappa} \geq d - l + 1$, *is given by* $\hat{\mathbf{X}} = \mathbf{V_R} Diag(\lambda^*) \mathbf{V_R^T}$, *where* λ^* *is defined by*

$$\lambda_k^* = \begin{cases} \mu^* := \frac{\lambda_1}{\hat{\kappa}} if\ \lambda_k < \mu^* \\ \lambda_k^* = \lambda_k \ otherwise. \end{cases} \tag{7.44}$$

and where l *is the index such that* $\lambda_l \leq \mu^* < \lambda_{l-1}$.

Proof. We apply the result given in Theorem 4 of Takana and Nakata [2014] for the trace norm (defined as $p = 1$ and $k = d$) to find the optimal value of μ^*. Theorem 2 of the same work yields the minimising solution $\hat{\mathbf{X}}$ for the value of μ^*.

We remark that the statement of Theorem 4 of Takana and Nakata [2014] uses the stronger assumption that $\hat{\kappa} \geq d$. However, a careful reading of the proof of this theorem indicates that a weaker assumption is sufficient: we assume that $\hat{\kappa} > d - l + 1$ where l is the index such that $\lambda_l \leq \mu^* < \lambda_{l-1}$. □

We note that this optimal value of μ^* is the same as the threshold $T = \frac{\lambda_1}{\hat{\kappa}}$ defined for the minimum eigenvalue method in (7.6) and hence the minimum eigenvalue method is equivalent to the Ky Fan 1-d minimizer of (7.43)in the case that $\kappa \geq d - l + 1$.

The minimum eigenvalue method is still a valid method of reconditioning when the additional assumption on the eigenvalues of $\mathbf{X}$ is not satisfied. In particular, in the experiments considered in Section 7.5 we see qualitatively similar behaviour for the choices of T that satisfy the assumption, and those that do not. It is possible that the lower bound on the condition number imposed by the additional constraint on κ_{max} could provide guidance on the selection of the target condition number.

Acknowledgements

Thanks to Stefano Migliorini and Fiona Smith at the UK Met Office for the provision of the IASI matrix. This work is funded in part by the EPSRC Centre for Doctoral Training in Mathematics of Planet Earth, the NERC Flooding from Intense Rainfall programme (NE/K008900/1), the EPSRC DARE project (EP/P002331/1) and the NERC National Centre for Earth Observation.

7.8 Supplementary material

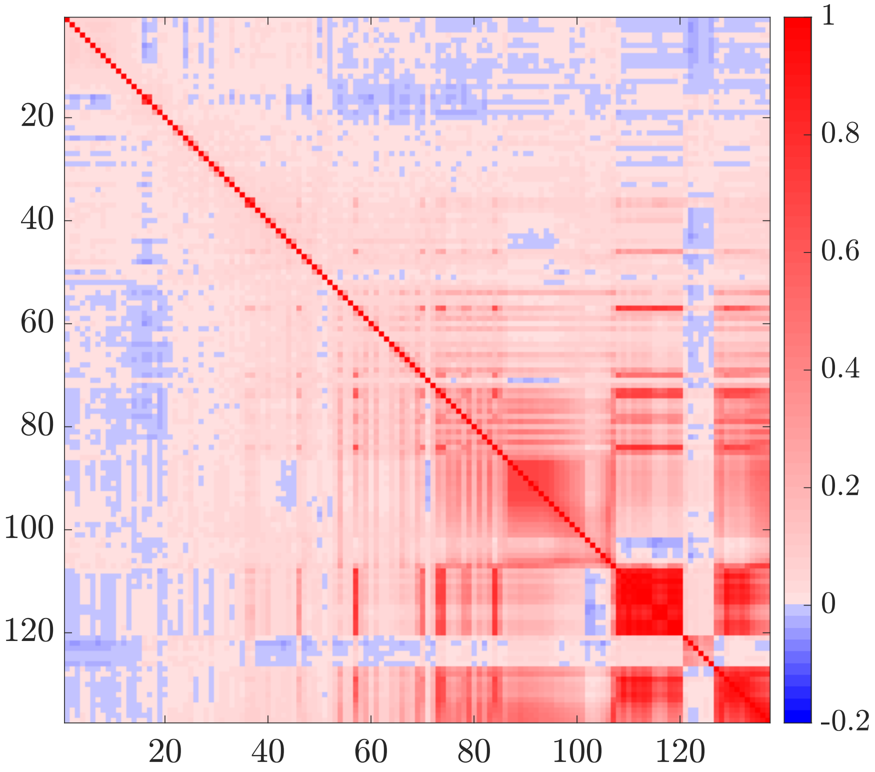


Figure 7.6: Diagnosed correlation matrix for IASI for the subset of 137 channels with non-zero off-diagonal entries.

Figure 7.6 shows the estimated and symmetrised observation error covariance matrix for IASI, obtained using the diagnostic of Desroziers et al. [2005] using the Met Office 4D-Var routine. The effects of reconditioning on this matrix were studied in Section 7.5.

7.9 Summary

In this chapter we developed theory on two methods of reconditioning, ridge regression and the minimum eigenvalue method, and compared them against multiplicative variance inflation. We found that both methods of reconditioning increase variances, with ridge regression resulting in larger increases to variances that the minimum eigenvalue method for any choice of covariance matrix. We proved that the ridge regression method strictly decreases the absolute value of off-diagonal correlations, whereas the minimum eigenvalue method can increase the absolute value of correlations. Both methods of reconditioning reduce the weight on scales associated with small eigenvalues of the observation error covariance matrix in the variational objective function, whereas multiplicative variance inflation reduces the weight on all scales equally. Numerical experiments revealed that for spatial correlations the minimum eigenvalue method can introduce spurious oscillations, but for an inter-channel example the minimum eigenvalue method made smaller changes to correlations and variances. An illustrative example showed that both methods of reconditioning are able to make changes to the analysis of a data assimilation problem on smaller scales, whereas multiplicative variance inflation cannot reduce spurious sample error on smaller scales.

Reconditioning methods were also found to result in faster convergence than multiplicative variance inflation. This agrees with the results of Chapter 5 that increasing the minimum eigenvalue of the observation error covariance matrix is likely to improve convergence of a conjugate gradient method. In this chapter we implemented reconditioning methods for an idealised data assimilation problem. In the next chapter, we will consider how the ridge regression method of reconditioning performs for an operational nonlinear data assimilation problem.

Chapter 8

The impact of using reconditioned correlated observation error covariance matrices in the Met Office 1D-Var system

In this chapter we answer RQ 4 from Chapter 1, and study how the use of the ridge regression method of reconditioning affects an operational data assimilation problem. We present a case study using the operational Met Office 1D-Var retrieval system. We wish to know

- How do the qualitative theoretical conclusions from the linear case apply in a non-linear, realistic setting?
- How are the quality control process and retrieved values affected by the introduction of correlated observation error and the use of reconditioning methods?

The work in this chapter, excluding the chapter summary (Section 8.8), has been strongly based on a paper submitted to the Quarterly Journal of the Royal Meteorological Society as: Tabeart J. M., Dance S. L., Lawless A. S., Migliorini, S. Nichols N. K., Smith, F., Waller J. A. The impact of using reconditioned correlated observation error covariance matrices in the Met Office 1D-Var system. Quarterly Journal of the Royal Meteorological Society. The submitted paper can be found at http://arxiv.org/abs/1908.04071.

8.1 Abstract

Recent developments in numerical weather prediction have led to the use of correlated observation error covariance (OEC) information in data assimilation and forecasting systems. However, diagnosed OEC matrices are ill-conditioned and may cause convergence problems for variational data assimilation procedures. Reconditioning methods are used to improve the conditioning of covariance matrices while retaining correlation information. In this chapter we study the impact of using the 'ridge regression' method of reconditioning to assimilate Infrared Atmospheric Sounding Interferometer (IASI) observations in the Met Office 1D-Var system. This is the first systematic investigation of how changing target condition numbers affects convergence of a 1D-Var routine. This procedure is used for quality control, and to estimate key variables (skin temperature, cloud top pressure, cloud fraction) that are not analysed by the main 4D-Var data assimilation system. Our new results show that the current (uncorrelated) OEC matrix requires more iterations to reach convergence than any choice of correlated OEC matrix studied. This suggests that using a correlated OEC matrix in the 1D-Var routine would have computational benefits for IASI observations. Using reconditioned correlated OEC matrices also increases the number of observations that pass quality control. However, the impact on skin temperature, cloud fraction and cloud top pressure is less clear. As the reconditioning parameter is increased, differences between retrieved variables for correlated OEC matrices and the operational diagonal OEC matrix reduce. These retrieval differences are smaller than retrieved standard deviation values for over 75% of IASI observations. Up to 5% of retrievals have large differences for alternative choices of the OEC matrix. As correlated choices of OEC matrix yield faster convergence, using stricter convergence criteria along with these matrices may further increase efficiency and improve quality control.

8.2 Introduction

In numerical weather prediction (NWP) a data assimilation procedure is used to combine observations of the atmosphere with a model description of the system in order to obtain initial conditions for forecasts. The contribution of each component is weighted by its respective error statistics. In recent years, interest in the understanding and use of correlated observation error statistics has grown (e.g. Janjić et al. [2018]). This increased interest has been motivated by results showing that neglecting correlated observation errors hinders forecasts [Rainwater et al., 2015, Stewart et al., 2008b], and that even including poorly approximated correlation structures is better than using uncorrelated error statistics in the presence of correlated errors [Stewart et al., 2013, Healy and White, 2005].

Previously, uncorrelated observation error statistics were used for all observations, even when it was known that non-zero error correlations were present. Determining error statistics is a non-trivial problem, as they cannot be observed directly and must be estimated in a statistical sense. It was also thought that it would not be possible to use correlated observation error covariance (OEC) matrices operationally due to the increased computational cost associated with inverting a dense matrix rather than a diagonal matrix [Stewart et al., 2013]. The development of a new method to check error consistency by Desroziers et al. [2005] was first applied to explicitly diagnose error correlations using the Met Office system [Stewart et al., 2008a]. Since then, the diagnostic introduced in Desroziers et al. [2005] (henceforth referred to as DBCP) has been used widely at operational centres [Weston, 2011, Weston et al., 2014, Stewart et al., 2014, Bennitt et al., 2017, Bormann et al., 2011, 2016, Campbell et al., 2017, Gauthier et al., 2018, Wang et al., 2018], although uncorrelated OEC matrices are still used operationally for most instruments. Although much of the initial use of the diagnostic to estimate observation errors focussed on interchannel correlations, this has been extended to spatial correlations [Waller et al., 2014b, 2016a,c, Cordoba et al., 2017, Michel, 2018]. Theoretical work has also demonstrated how well the diagnostic is expected to perform depending on either the accuracy of the initial choice of background and OEC matrices for the single step [Waller et al., 2016b] and the iterative form of the diagnostic [Ménard, 2016, Bathmann, 2018]. The use of the diagnostic in data assimilation schemes using localization has also been considered [Waller et al., 2017].

The output of the diagnostic cannot be used directly in the assimilation procedure. Diagnosed matrices are asymmetric, and some are not positive definite [Stewart et al., 2014, Weston et al., 2014] and are therefore not valid covariance matrices. Typically, the matrices are symmetrised, and negative and zero eigenvalues are set to be small and positive [Weston, 2011]. Additionally, diagnosed OEC matrices are often ill-conditioned. This means that small perturbations to the observations will result in large changes to the analysis, and that iterative methods are likely to converge slowly. Indeed, the direct use of diagnosed matrices has led to problems with non-convergence of the minimisation of the data assimilation procedure [Weston, 2011, Weston et al., 2014]. Weston [2011] suggested that part of these problems were due to small minimum eigenvalues of the diagnosed OEC matrix, $\mathbf{R}$.

One way to study the effect of changes to the assimilation system on the convergence of the objective function minimisation is by using the condition number of the Hessian of the variational objective function as a proxy for convergence. This was done in Haben [2011] for the case of a linear observation operator. In Tabeart et al. [2018] the minimum eigenvalue of the OEC matrix, $\mathbf{R}$, appears in bounds on the condition number of the Hessian of the variational assimilation problem, indicating that this term will also be important for convergence of the objective function minimisation.

Increased understanding of how the eigenvalues of $\mathbf{R}$ affect the convergence of the data assimilation problem motivated investigation into ‘reconditioning’ methods [Weston, 2011, Weston et al., 2014, Campbell et al., 2017, Tabeart et al., 2018]. These methods increase eigenvalues of the matrix $\mathbf{R}$ to improve the conditioning of the OEC matrix, while maintaining much of the existing correlation structure of the diagnosed matrix. Two methods are commonly used by NWP centres: ‘ridge regression’ which increases all eigenvalues of $\mathbf{R}$ by the same amount, and the ‘minimum eigenvalue’ method which changes only the smallest eigenvalues. These methods were investigated theoretically in Tabeart et al. [2019a] where it was found that both methods increase standard deviations, and that the ridge regression method strictly reduces all off-diagonal correlations. Both methods were compared in an operational system in Campbell et al. [2017], where the sensitivity of forecasts to the choice of method was found to be small, but the ridge regression outperformed the minimum eigenvalue method in terms of convergence. A method similar to the minimum eigenvalue method is used at the European Centre for Medium Range Weather Forecasts (ECMWF) [Bormann et al., 2016], but will not be discussed further in this chapter.

The aim of this chapter is to investigate the use of the ridge regression method within the Met Office system. At the Met Office, in addition to the 4D-Variational data assimilation routine (4D-Var) that is used to produce the initial conditions for weather forecasts, a 1D-Variational data assimilation routine (1D-Var) is used for quality control and pre-processing purposes [Eyre, 1989]. The 1D-Var routine assimilates observations individually, and is used to remove observations that are likely to cause problems with convergence in the 4D-Var routine, as well as to estimate model variables that are not included in the 4D-Var state vector [Pavelin and Candy, 2014, Pavelin et al., 2008]. After the work of Weston [2011], Weston et al. [2014], correlated OEC matrices were introduced in the 4D-Var routine for IASI (Infrared Atmospheric

Sounding Interferometer) and other hyperspectral IR sounders. However, this was not the case for the 1D-Var routine, where a diagonal OEC matrix continues to be used. Previous work found that diagnosed observation error correlations were small for most channels for the 1D-Var routine [Weston, 2011, Stewart et al., 2014] and the proportional increase in computational cost was estimated to be large compared with using correlated OEC matrices in 4D-Var [Weston et al., 2014].

In this chapter we study how the use of reconditioning methods affects the 1D-Var routine when applied to interchannel OEC matrices for the Infrared Atmospheric Sounding Interferometer (IASI). We examine whether the ridge regression method of reconditioning allows us to include correlated observation error information more efficiently than the diagnosed OEC matrix. This method of reconditioning is used at the Met Office to recondition OEC matrices that are used in the 4D-Var routine. We compare a selection of reconditioned OEC matrices with the current diagonal operational error covariance matrix, and an inflated diagonal OEC matrix. This is the first time that multiple levels of reconditioning have been compared systematically in an operational system. We study the impact of reconditioning in terms of the computational efficiency as well as the effect on important meteorological variables.

In Section 8.3 the data assimilation problem is defined and the ridge regression method of reconditioning is introduced. In Section 8.4 we provide an overview of the experimental design. In Sections 8.5 and 8.6 we discuss the impact of changing the OEC matrix on the 1D-Var procedure, and alterations to the quality control and pre-processing for the 4D-Var routine respectively. We find that convergence is improved for any of the choices of reconditioning compared to the current operational choice of OEC matrix. Additionally, increasing the amount of reconditioning results in faster convergence - which corresponds to theoretical results for the linear variational data assimilation problem in Tabeart et al. [2018]. However, the quality control procedure is altered by changing the OEC matrix, with a larger number of observations being accepted for reconditioned correlated OEC matrices compared to the current diagonal choice of OEC matrix. We also find that for most variables, the difference between retrieved values for different choices of OEC matrix are small compared to retrieved standard deviations. However, there are a significant minority of observations for which differences are very large. Finally, in Section 8.7 we summarise our results and conclusions.

8.3 Variational data assimilation and reconditioning

8.3.1 Data assimilation

In data assimilation, a weighted combination of observations, $\mathbf{y} \in \mathbb{R}^p$, with a background, or 'prior', field, $\mathbf{x}_b \in \mathbb{R}^n$, is used to obtain the analysis, or posterior, $\mathbf{x}_a \in \mathbb{R}^n$. The weights are the respective error statistics of the two components. The matrix $\mathbf{R} \in \mathbb{R}^{p\times p}$ is the observation error covariance (OEC) matrix and $\mathbf{B} \in \mathbb{R}^{n\times n}$ is the background error covariance matrix. In order to compare observations with the background field, the, possibly non-linear, observation operator $H : \mathbb{R}^n \to \mathbb{R}^p$ is used to map from state space to observation space. The weighted combination is written in the form of an objective function in terms of $\mathbf{x} \in \mathbb{R}^n$, the model state vector. In the case of 3D-Var the objective function is given by:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - H[\mathbf{x}])^T\mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}]). \tag{8.1}$$

The value of $\mathbf{x}$ that minimises (8.1) is given by $\mathbf{x}_a$.

The first order Hessian, or matrix of second derivatives, of the objective function (8.1) is given by

$$\nabla^2 J \equiv \mathbf{S} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}, \tag{8.2}$$

where $\mathbf{H} \in \mathbb{R}^{p\times N}$ is the Jacobian of the observation operator, $H[\mathbf{x}]$, linearised about the current best estimate of the optimal solution of (8.1).

We now define the condition number of a matrix. Let $\lambda_{max}(\mathbf{S}) = \lambda_1(\mathbf{S}) \geq \cdots \geq \lambda_N(\mathbf{S}) = \lambda_{min}(\mathbf{S})$ be the eigenvalues of $\mathbf{S}$. We note that this ordering convention will be used for the remainder of the chapter. Although covariance matrices are symmetric positive semi-definite by definition, in practice $\mathbf{B}$ and $\mathbf{R}$ are required to be strictly positive definite in order that they can be inverted in (8.1). This means that $\mathbf{S}$ is symmetric positive definite, and its condition number is given by

$$\kappa(\mathbf{S}) = \frac{\lambda_1(\mathbf{S})}{\lambda_N(\mathbf{S})}. \tag{8.3}$$

We note that the minimum possible value of the condition number of any matrix is one. The condition number of the Hessian is of interest because it can be used to study the sensitivity of the solution to small changes in the background or observation data [Golub and Van Loan, 1996, Sec 2.7]. As (8.1) is non-linear, it is solved using a sequence of Gauss-Newton iterations with an inner linearised problem solved using the conjugate gradient method [Haben et al., 2011b]. The rate of convergence of the minimisation of the linearised problem by a conjugate gradient function can also be bounded by $\kappa(\mathbf{S})$ [Golub and Van Loan, 1996], although this bound is quite pessimistic. In particular, clustering of eigenvalues can result in much faster convergence than is predicted by $\kappa(\mathbf{S})$ [Nocedal, 2006].

8.3.2 Reconditioning: motivation and definition

In Weston [2011], observations from IASI were used at the Met Office for an initial study investigating the feasibility of using correlated observation error matrices in their 4D-Var system. A first guess of the OEC matrix was obtained using the DBCP diagnostic.

One problem that was encountered in Weston [2011] and Weston et al. [2014] was the ill-conditioning of the matrix resulting from the DBCP diagnostic. The use of an ill-conditioned OEC matrix can result in slower convergence of a variational scheme [Weston et al., 2014, Tabeart et al., 2018]. Similar problems were encountered at ECMWF where a degradation in the forecast was seen when the raw output of the DBCP diagnostic was tested [Lupu et al., 2015]. Weston [2011] suggested that the convergence problems were caused by very small minimum eigenvalues of the diagnosed observation error covariance matrix.

Tabeart et al. [2018] developed bounds for the condition number of the Hessian in terms of its constituent matrices in the case of a linear observation operator. This provides an indication of the role of each matrix in the conditioning of $\mathbf{S}$, and therefore the convergence of the associated minimisation problem. The bound which separates the role of each matrix is given by

$$\begin{aligned}\max\Bigg\{&\frac{1+\frac{\lambda_{max}(\mathbf{B})}{\lambda_{min}(\mathbf{R})}\lambda_{max}(\mathbf{H}\mathbf{H}^T)}{\kappa(\mathbf{B})},\frac{1+\frac{\lambda_{max}(\mathbf{B})}{\lambda_{max}(\mathbf{R})}\lambda_{max}(\mathbf{H}\mathbf{H}^T)}{\kappa(\mathbf{B})},\\ &\frac{\kappa(\mathbf{B})}{1+\frac{\lambda_{max}(\mathbf{B})}{\lambda_{min}(\mathbf{R})}\lambda_{max}(\mathbf{H}\mathbf{H}^T)}\Bigg\} \qquad (8.4)\\ &\leq \kappa(\mathbf{S}) \leq \left(1+\frac{\lambda_{min}(\mathbf{B})}{\lambda_{min}(\mathbf{R})}\lambda_{max}(\mathbf{H}\mathbf{H}^T)\right)\kappa(\mathbf{B}).\end{aligned}$$

These bounds show that the minimum eigenvalue, $\lambda_{min}(\mathbf{R})$, of the OEC matrix is a key term in the upper bound for $\mathbf{S}$, meaning that increasing the minimum eigenvalue of $\mathbf{R}$ is a reasonable heuristic for reducing the condition number of $\mathbf{S}$ and improving the conditioning of the problem (8.1).

In the case that the error covariance matrices can be written as the product of a scalar variance with a correlation matrix, e.g. $\mathbf{R} = \sigma_o^2\mathbf{D}$ and $\mathbf{B} = \sigma_b^2\mathbf{C}$, and observations are restricted to model variables, we can simplify the bound (8.4) to

$$\begin{aligned} \max\left\{ \frac{1+\frac{\sigma_b^2}{\sigma_o^2}\frac{\lambda_{max}(\mathbf{C})}{\lambda_{min}(\mathbf{D})}}{\kappa(\mathbf{C})}, \frac{\kappa(\mathbf{C})}{1+\frac{\sigma_b^2}{\sigma_o^2}\frac{\lambda_{max}(\mathbf{C})}{\lambda_{min}(\mathbf{D})}} \right\} \\ \leq \kappa(\mathbf{S}) \leq \left(1+\frac{\sigma_b^2}{\sigma_o^2}\frac{\lambda_{min}(\mathbf{C})}{\lambda_{min}(\mathbf{D})}\right)\kappa(\mathbf{C}). \end{aligned} \tag{8.5}$$

The qualitative conclusions of Tabeart et al. [2018] can be summarised as follows.

- The minimum eigenvalue of $\mathbf{R}$ was shown to be important for determining both the conditioning of the Hessian, and the speed of convergence of a minimisation procedure. This can be seen in (8.4) and (8.5).

- The ratio of the background and observation variances was also shown to be important for conditioning of the Hessian. This can be seen in (8.5) explicitly for the case of direct observations where variances are homogeneous for both background and small scale matrices. However, we expect the conclusion to hold more broadly, for example in the case where all standard deviation values corresponding to an OEC matrix were larger than those corresponding to another OEC matrix, then the bounds would be smaller for the first choice of OEC matrix.

- Although (8.4) and (8.5) separate the contribution of each term, numerical experiments revealed that the level of interaction between observation error and background error statistics depends on the choice of observation network. Examples of observation operators which yield identical bounds for (8.4) but different dependence of $\kappa(\mathbf{S})$ on $\mathbf{B}$ and $\mathbf{R}$ were found experimentally in Tabeart et al. [2018].

These conclusions motivate the use of reconditioning methods. In order to make operational implementation of correlated observation error matrices feasible, it is necessary to reduce the impact of the very small eigenvalues of the matrix $\mathbf{R}$ by increasing its condition number. To achieve this, different methods of inflation, or reconditioning are used to improve conditioning of correlation matrices for a variety of applications. The ridge regression method is used to recondition OEC matrices at the Met Office [Weston, 2011, Weston et al., 2014], and hence will be the reconditioning method that is considered in the remainder of this chapter. The ridge regression method adds a scalar multiple of the identity to $\mathbf{R}$ to obtain the reconditioned matrix $\mathbf{R}_{RR}$. This scalar, δ, is chosen such that $\kappa(\mathbf{R}_{RR}) = \kappa_{max}$, a user-specified condition number. The method for calculating δ for a given choice of κ_{max} was formally defined in Tabeart et al. [2019a] as follows:

Definition 8.3.1. ***Ridge regression reconditioning constant,*** δ ***[Tabeart et al., 2019a]***
Define $\delta = (\lambda_{max}(\mathbf{R}) - \lambda_{min}(\mathbf{R})\kappa_{max})/(\kappa_{max} - 1)$.
Set $\mathbf{R}_{RR} = \mathbf{R} + \delta \boldsymbol{I}$.

We note that this choice of δ yields $\kappa(\mathbf{R}_{RR}) = \kappa_{max}$. Mathematical theory describing the effect of this reconditioning method on the correlations and variances of any covariance matrix was developed in Tabeart et al. [2019a], which showed that the ridge regression method increases error variances for all observations, and decreases all off-diagonal correlations. In this chapter we investigate whether the qualitative conclusions from Tabeart et al. [2018] hold in the case of a non-linear observation operator, and we study the impact of reconditioning methods in an operational system.

8.4 Experimental Overview

8.4.1 Met Office System

The experiments carried out in this manuscript will use observations from the IASI instrument on the EUMETSAT MetOp constellation. IASI is an infrared Fourier transform spectrometer, and measures infrared radiation emissions from the atmosphere and surface of the earth [Chalon et al., 2001]. We note that the observation operator for this instrument, a radiative transfer model, is highly non-linear so the conclusions from Tabeart et al. [2018] will not necessarily apply to this problem. The infrared spectrum is split into channels corresponding to different wavelengths; this means that an observation at a single location will provide information for up to 8641 channels. An early use of the DBCP diagnostic focused on observations from IASI implemented in the Met Office system [Stewart et al., 2008a]. Much of the subsequent research on correlated observation error uses IASI observations [Weston, 2011, Weston et al., 2014, Stewart et al., 2014, Bormann et al., 2016]. In particular IASI has channels that are sensitive to water vapour which have been found to have errors with large correlations [Stewart et al., 2014, Weston et al., 2014, Bormann et al., 2016].

One attraction of IASI, and other hyperspectral instruments, is the large number of available channels, which provides high vertical resolution. However, using all of these channels is not feasible in current operational NWP systems for reasons including computational expense, and not requiring too many observations of a similar type. Additionally, when IASI was first used, there was a reluctance to include correlated observation errors so an effort was made to choose channels that are spectrally different and hence less likely to have correlated errors [Stewart et al., 2014]. This means that of the 8461 available channels, only a few hundred are used at most NWP centres [Stewart et al., 2014]. At the time of the experiments, the Met Office stored a subset of 314 channels with a maximum of 137 being used in the 4D-Var system. A list of these channels is given by Stewart [2010, Appendix A]. As there is a large degree of redundancy between channels [Collard et al., 2010], directly assimilating a larger number of channels is likely to make the conditioning of the OEC matrix worse. This has motivated alternative approaches such as principal component compression [Collard et al., 2010] and the use of transformed retrievals [Prates et al., 2016], which will not be considered in this work.

A larger number of channels is used for the 1D-Var assimilation than for the Met Office 4D-Var assimilation; standard deviation values for these channels are filled in from the current operational (diagonal) OEC matrix. We chose to focus on the channels used in the 4D-Var system in order to be consistent between both assimilation systems. We also note that not all channels are used for each assimilation; for example, some channels are not used in the presence of cloud. In this case, rows and columns corresponding to channels that are affected by cloud are deleted from the OEC matrix. As the submatrix chosen from the full OEC matrix used could change at each observation time, there may be a difference in the condition number of the OEC matrix used in practice and the OEC matrices presented in this work. However, the Cauchy interlacing theorem [Bernstein, 2009, Lemma 8.4.4] states that the condition number will not be increased by deleting rows and columns of a symmetric positive definite matrix. This means that the values given here are upper bounds for $\kappa(\mathbf{S})$ even if the quality control procedure excludes some channels.

We test the impact of using correlated OEC matrices in the Met Office 1D-Var system and consider the effect of using the ridge regression method of reconditioning with different choices of target condition number. At the Met Office, 1D-Var is run prior to every 4D-Var assimilation procedure, meaning that retaining current computational efficiency and speed of convergence is desirable. We note that a single IASI observation consists of brightness temperature values for each of the channels that are used in the assimilation. A 1D-Var procedure takes observations separately at each location to retrieve variables such as temperature and humidity over a 1D column of the atmosphere. This procedure is much cheaper and more parallelisable than a 4D-Var algorithm.

The 1D-Var routine performs two main functions:

1. Quality control (QC): Observations that require more than 10 iterations for the 1D-Var minimisation to reach convergence are not passed to the 4D-Var routine. This is because it is assumed that observations for which the retrieval procedure takes too long to converge for the 1D-Var minimisation will also result in slow convergence for a 4D-Var minimisation. A Marquardt-Levenberg minimisation algorithm is used, with convergence criteria is based on the value of the cost function and normalised gradient [Pavelin et al., 2008]. Changing the OEC matrix will alter the speed of convergence of 1D-Var, and hence affect which observations are accepted.

2. Estimation of values for certain variables that are not included in the 4D-Var state vector: values for skin temperature, cloud fraction, cloud top pressure and emissivity over land are fixed by the 1D-Var procedure. Altering the OEC matrix will change retrieved values for these variables.

Changing the OEC matrix is therefore likely to have two main effects on results of the 1D-Var procedure: changing the observations that are accepted by the quality control, and changing the values of those variables not included in the 4D-Var state vector. Skin temperature (ST), cloud fraction (CF) and cloud top pressure (CTP) are retrieved as scalar values at each observation location. In contrast, surface emissivity is retrieved as a spectrum, which is represented as a set of leading principal components [Pavelin and Candy, 2014]. As we expect the interactions between the choice of $\mathbf{R}$ and the retrieved values to be complex, in this work we only consider the effect of changing $\mathbf{R}$ on the three scalar variables: skin temperature, cloud top pressure and cloud fraction.

We finish by commenting briefly on the implementation of Marquardt-Levenberg that is used at the Met Office for 1D-Var. As the nonlinear Jacobian is recomputed at every iteration, 1D-Var can be thought of as a fully nonlinear model with only outer loops and no inner loops (see Rodgers [1998] for further discussion). Therefore the experimental results presented in the following section will allow us to test how well previous linear theory from Tabeart et al. [2018] applies to a nonlinear data assimilation algorithm.

8.4.2 Experimental Design

We now describe the experimental framework and key areas of interest that will be investigated in Sections 8.5 and 8.6. We use the operational Met Office 1D-Var framework at the time of the experiments (July 2016), and consider how the results change for different choices of OEC matrix. Background profiles are obtained from the Unified Model (UM) background files for the corresponding configuration. A number of different times and dates for the six months between December 2015 and June 2016 were considered, but as results were similar across all trials we only present results from experiments for 16th June 2016 0000Z.

The correlated choices of $\mathbf{R}$ are calculated using the method introduced in Weston et al. [2014]; applying the ridge regression method of reconditioning to the diagnosed matrix for a variety of choices of κ_{max}. The matrices estimated by the DBCP diagnostic depend considerably on the choice of background and observation error matrices. For all OEC matrices produced, the same 4 days of IASI and background NWP data (03/12/15-06/12/15), were used as input data. We note that the estimated OEC matrix was obtained using background and OEC matrices from the 4D-Var assimilation routine rather than the 1D-Var routine. Although this is not theoretically consistent with the smaller error correlations that have been estimated for the 1D-Var problem in previous studies [Stewart et al., 2014, Weston, 2011], the use of 4D-Var error statistics allows us to better understand the impact that our changes are likely to have on 4D-Var. We are using 1D-Var as a pre-processing step for 4D-Var to remove observations that are likely to cause convergence issues in the main assimilation algorithm.

We use the operational background error covariance matrix, $\mathbf{B}$, at the time of the

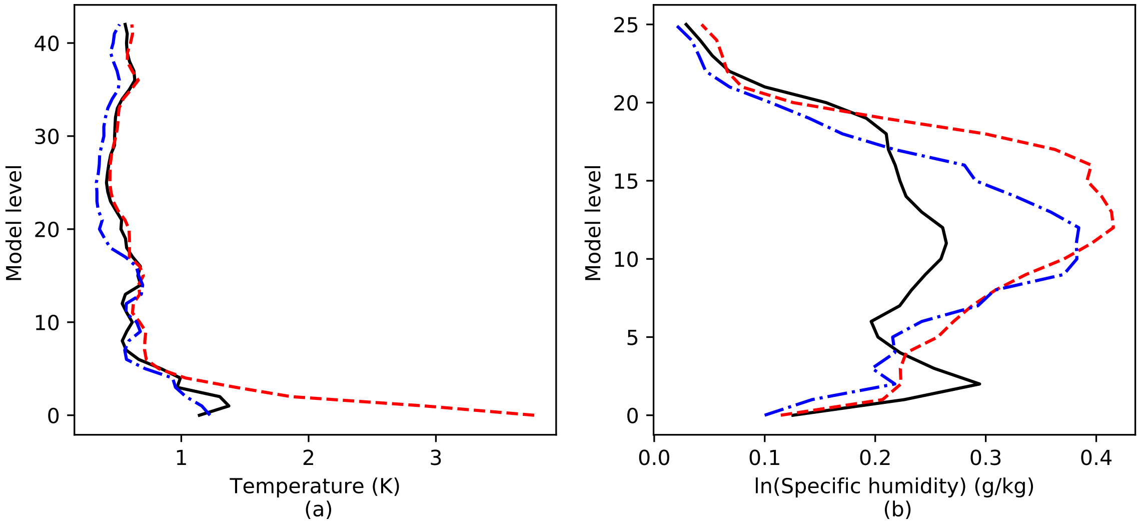


Figure 8.1: Standard deviation values for the operational background error covariance matrices, **B**, for the northern hemisphere (solid line), tropics (dot-dashed line) and southern hemisphere (dashed line) for temperature (a) and ln(specific humidity) (b). 43 model levels are determined by the 43 evenly distributed pressure levels in the radiative transfer retrieval algorithm, where model level 0 corresponds to the surface and model level 42 the top of the atmosphere. ln(specific humidity) is only computed for the lowest 26 model levels.

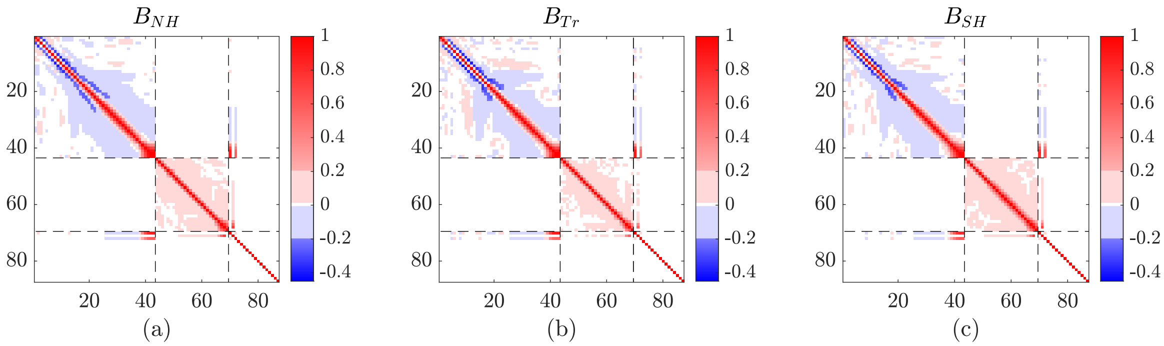


Figure 8.2: Correlation matrices for the operational background error covariance matrices, **B**, for the northern hemisphere (a), tropics (b) and southern hemisphere (c). Dashed vertical and horizontal lines separate inter and cross correlations between temperature, ln(specific humidity) and other variables (from left to right). ST is variable 72, CTP is 74 and CF is 75.

Variable	CF	CTP (hPa)	ST (NH) (K)	ST (Tr) (K)	ST (SH) (K)
Standard deviation	1	1000	2.24	1.92	2.02

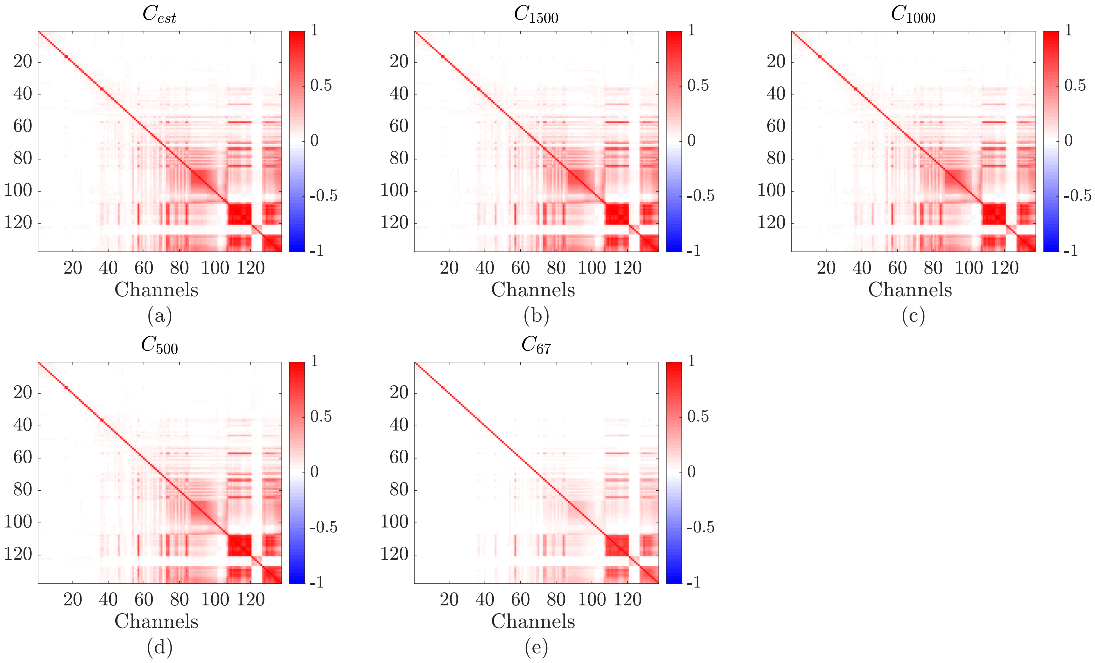
Table 8.1: Background standard deviation values for variables not included in the 4D-Var state vector.

experiments. This consists of three different choices of **B** for the northern hemisphere (30N:90N), the tropics (30S:30N) and the southern hemisphere (90S:30S). Figure 8.1 shows background standard deviation (BSD) values for temperature and humidity variables, and Table 8.1 gives BSD values for CF, CTP and ST for each of the choices of **B**. The 43 model levels are determined by the 43 evenly distributed pressure levels in the radiative transfer retrieval algorithm. Figures in this paper are plotted with model level 0 corresponding to the surface, and model level 42 the top of the atmosphere. This is the opposite ordering to that which is used by RTTOV (Radiative Transfer for TOVS (Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder)). We note that standard deviations for cloud variables are assumed to be very large so that the background is ignored for these variables [Pavelin et al., 2008]. In Sections 8.5 and 8.6 we will compare the standard deviations from the background error covariance matrix against retrieved standard deviations for the observations as well as differences between observations for different choices of **R**. Figure 8.2 shows that correlations corresponding to the three choices of **B** are qualitatively very similar. Cross-correlations between variables are quite weak, with no correlations between temperature and specific humidity. Most correlations larger than 0.2 occur for adjacent model levels for temperature and specific humidity. Correlations greater than 0.2 also occur between surface temperature and ST and temperature for larger model level numbers, and surface specific humidity and specific humidity at larger model level numbers. CTP and CF are uncorrelated with all other variables.

We apply the DBCP diagnostic to the subset of 137 channels that are assimilated in the 4D-Var routine. The 1D-Var routine uses additional channels [Hilton et al., 2009], with a total of 183 channels being assimilated. Observation errors for these additional channels are assumed to be uncorrelated, and filled in with values from the diagonal error covariance matrix $\mathbf{R}_{diag}$. If additional channels are included in future versions of the operational system, it would be advisable to recompute the DBCP diagnostic applied to all channels.

The seven different choices of the matrix **R** that were tested are now listed:

- $\mathbf{R}_{infl}$ which is an inflated diagonal matrix. This matrix was used prior to the introduction of correlated observation error in the 4D-Var assimilation scheme Weston et al. [2014]. In particular variances are inflated to account for the fact that the assumption of uncorrelated errors is incorrect. The standard deviations corresponding to $\mathbf{R}_{infl}$ are shown by the black solid line in Figure 8.4. The largest value entry of $\mathbf{R}_{infl}$ is 16, and the smallest entry is 0.25. The



(a) (b) (c) (d) (e)

Figure 8.3: The correlation matrices corresponding to each correlated choice of $\mathbf{R}$. Recall that the subscript defines the choice of κ_{max} used in the ridge regression method of reconditioning, which was applied to the covariance matrix $\mathbf{R}_{est}$.

construction of $\mathbf{R}_{infl}$ is described in Hilton et al. [2009].

- $\mathbf{R}_{diag}$, the current operational matrix for 1D-Var retrievals, which is diagonal. The standard deviations are calculated as instrument noise plus $0.2K$ forward-model noise [Collard, 2007]. The variances of $\mathbf{R}_{diag}$ are shown by the solid red line in Figure 8.4. The variances are much smaller than for $\mathbf{R}_{infl}$; for the first 120 channels, the diagonal elements of $\mathbf{R}_{diag}$ are all less than 0.27 and the largest value of $\mathbf{R}_{diag}$ is given by 0.49.

- $\mathbf{R}_{est}$, the symmetrised raw output of the code that produces the DBCP diagnostic. This is computed by $\mathbf{R}_{est} = \frac{1}{2}(\mathbf{R}_{DBCP} + \mathbf{R}_{DBCP}^T)$, where $\mathbf{R}_{DBCP} \in \mathbb{R}^{137\times137}$ is the output of the DBCP diagnostic.

- Reconditioned versions of $\mathbf{R}_{est}$ so that the correlated submatrix has a condition number of 1500, 1000, 500 and 67, referred to respectively as $\mathbf{R}_{1500}, \mathbf{R}_{1000}, \mathbf{R}_{500}$ and $\mathbf{R}_{67}$.

The correlations and standard deviations corresponding to $\mathbf{R}_{est}, \mathbf{R}_{1500}, \mathbf{R}_{1000}, \mathbf{R}_{500}$ and $\mathbf{R}_{67}$ are shown in Figures 8.3 and 8.4 respectively. We refer to the experiments using each choice of OEC matrix as E with subscript corresponding to that of the OEC

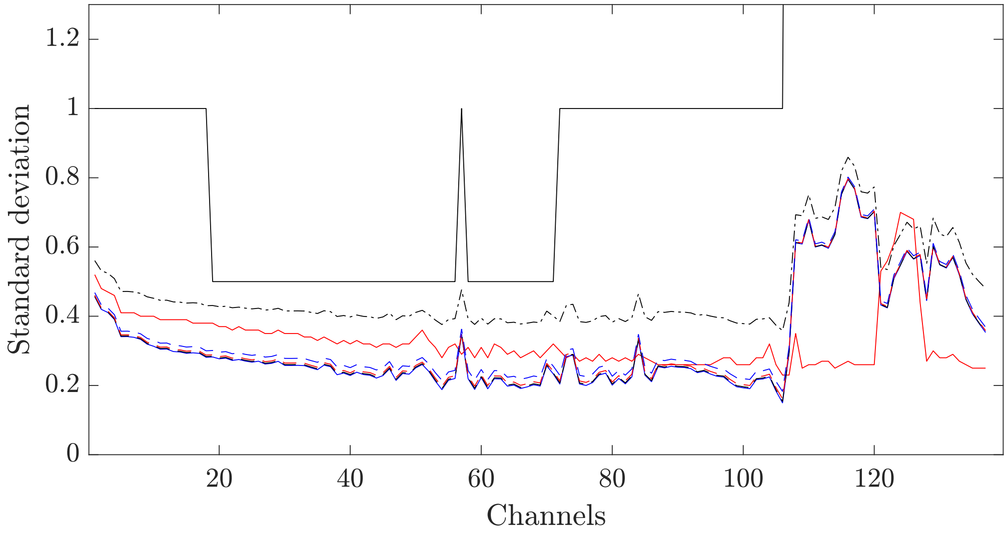


Figure 8.4: Standard deviations values corresponding to each choice of $\mathbf{R}$. Black solid line denotes $\mathbf{R}_{infl}$, red solid line denotes $\mathbf{R}_{diag}$, blue solid line denotes $\mathbf{R}_{est}$, black dashed line denotes $\mathbf{R}_{1500}$, red dashed line denotes $\mathbf{R}_{1000}$, blue dashed line denotes $\mathbf{R}_{500}$ and black dot-dashed line denotes $\mathbf{R}_{67}$. We note that the standard deviation value for $\mathbf{R}_{infl}$ for channels 106-137 is 4K.

Experiment name	E_{diag}	E_{est}	E_{1500}	E_{1000}	E_{500}	E_{67}	E_{infl}
Choice of $\mathbf{R}$	$\mathbf{R}_{diag}$	$\mathbf{R}_{est}$	$\mathbf{R}_{1500}$	$\mathbf{R}_{1000}$	$\mathbf{R}_{500}$	$\mathbf{R}_{67}$	$\mathbf{R}_{infl}$
$\lambda_{min}(\mathbf{R})$	0.025	0.00362	0.00482	0.007244	0.0145	0.1010	0.0625
$\kappa(\mathbf{R})$	9.263	2730	1500	1000	500	67	64

Table 8.2: Minimum eigenvalues and condition number of $\mathbf{R}$ for each experiment.

matrix (i.e. $E_{diag}, E_{est}, E_{1500}E_{1000}, E_{500}, E_{67}$ and E_{infl}).

Details of the conditioning, and minimum eigenvalues of each of the choices of $\mathbf{R}$ can be found in Table 8.2. We see that for the non-diagonal matrices, as we decrease the target condition number, we increase the minimum eigenvalue of $\mathbf{R}$. This agrees with the theoretical results of Tabeart et al. [2019a]. We also see that of the two diagonal choices of $\mathbf{R}$, $\mathbf{R}_{infl}$ has the larger value of $\lambda_{min}(\mathbf{R})$, suggesting that we might expect better convergence compared to $\mathbf{R}_{diag}$. We also notice that the largest value of $\lambda_{min}(\mathbf{R})$ occurs for $\mathbf{R}_{67}$. It will be of interest to consider whether the introduction of correlations has more effect on convergence and conditioning than the value of $\lambda_{min}(\mathbf{R})$. We note that the inclusion of 46 extra channels in the 1D-Var algorithm, in addition to the 137 channels used in the 4D-Var algorithm, could change the condition numbers presented in Table 8.2 by the introduction of very small or very large eigenvalues.

Our numerical experiments will be broadly split into two groups. Firstly we will consider the effect of changing the OEC matrix, **R**, on the 1D-Var procedure itself in Section 8.5. This includes the impact on retrieved values and the convergence of the 1D-Var assimilation. Secondly, in Section 8.6, we will consider the impact of these changes on the 4D-Var procedure, by looking at how the number of accepted observations varies, and how the retrieved values of skin temperature, cloud top pressure, and cloud fraction retrievals are altered.

8.5 Impact on Met Office 1D-Var routine

In this section we consider the impact of changing the OEC matrix used in the Met Office 1D-Var system on the conditioning of the Hessian and on individual retrievals of temperature and humidity. In particular, the conditioning of the Hessian is important in terms of speed of convergence of the minimisation procedure. We recall (Section 8.4.1) that in 1D-Var information for each observation location is assimilated separately. Here a single observation corresponds to information from a column of IASI channels valid at one location. This corresponds to 97330 observations over the 4 days of data discussed in Section 8.4.2 with objective functions that converge in 10 or fewer iterations for all choices of E_{exp}. For much of the discussion that follows we will consider statistics of this set of 97330 observations to understand how changing the OEC matrix affects 1D-Var for IASI observations.

8.5.1 Influence of observation error covariance matrix on convergence and conditioning of the 1D-Var routine

We begin by investigating explicitly the effect of changing the OEC matrix, **R**, on the 1D-Var routine. We consider two variables: the number of iterations required for convergence for the minimisation routine and the condition number of the Hessian of the 1D-Var cost function.

Firstly we consider the number of iterations required for the minimisation of the 1D-Var cost function to reach convergence for each assimilated observation. For NWP centres, this is a variable of significant interest, as the extra expense of introducing correlated error predominantly comes from the increase in the number of iterations needed before convergence in the case of interchannel errors [Weston, 2011]. We note that this may not be the case for other types of error correlation such as spatial and temporal correlations (where the computation of matrix-vector products may require additional communication between processors [Simonin et al., 2019]). The minimisation is deemed to have converged when the absolute value of the difference between each component of two successive estimates of the state vector is smaller than $0.4\sigma_{\mathbf{B}}$, where $\sigma_{\mathbf{B}}$ is the vector whose components are the background error variances for each retrieved variable. Values deemed to be unphysical, such as temperature components falling out of the range $70K - 340K$, are discarded.

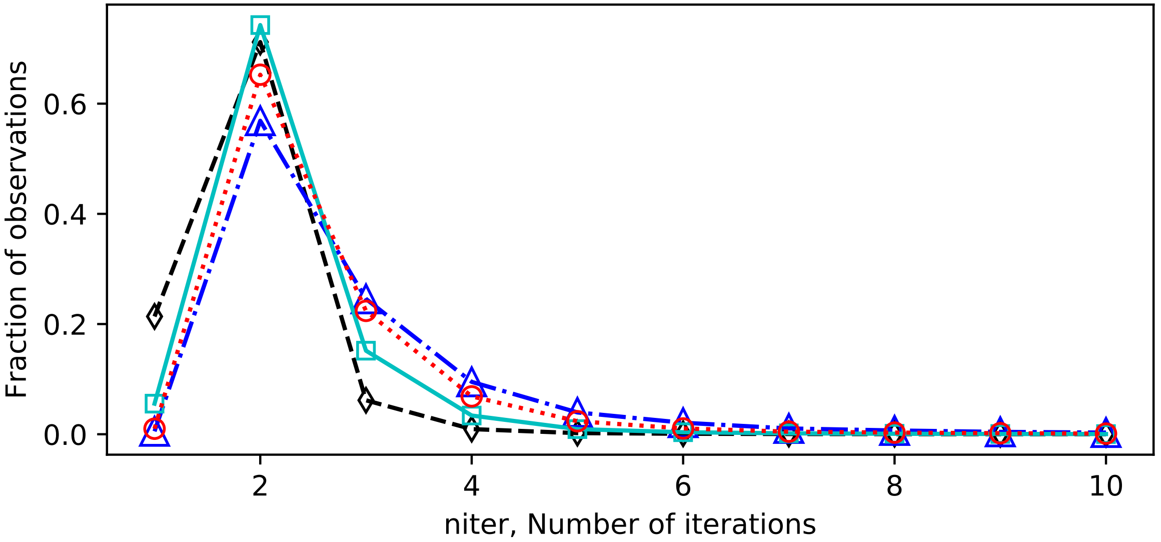


Figure 8.5: Number of iterations required for convergence of the minimization of the 1D-Var cost function as a fraction of the total number of observations common to all choices of $\mathbf{R}$. Symbols correspond to: $\mathbf{R}_{diag}$ ($\triangle$), $\mathbf{R}_{est}$ ($\circ$), $\mathbf{R}_{67}$ ($\square$) and $\mathbf{R}_{infl}$ ($\diamond$).

	E_{diag}	E_{est}	E_{1500}	E_{1000}	E_{500}	E_{67}	E_{infl}
$\max \kappa(\mathbf{S})$	3.01×10^{12}	7.546×10^{11}	7.469×10^{11}	7.30×10^{11}	7.02×10^{11}	3.71×10^{11}	1.74×10^{11}
mean $\kappa(\mathbf{S})$	2.78×10^{10}	6.71×10^{9}	6.62×10^{9}	6.43×10^{9}	6.00×10^{9}	4.01×10^{9}	2.83×10^{9}
median $\kappa(\mathbf{S})$	2.09×10^{8}	1.31×10^{8}	1.32×10^{8}	1.33×10^{8}	1.37×10^{8}	1.78×10^{8}	2.89×10^{8}

Table 8.3: Maximum, mean and median values of $\kappa(\mathbf{S})$ for E_{diag} and experiments.

For each observation, we store the number of iterations required for the corresponding 1D-Var objective function to converge, *niter*. Figure 8.5 shows the fraction of observations that have objective functions that converge in *niter* iterations for four choices of $\mathbf{R}$. We note that the behaviour for the other correlated experiments is similar to the behaviour for E_{est} and hence only the distributions for E_{est} and E_{67} are shown. We see that for all experiments $niter = 2$ is the modal class and contains over 50% of the observations. We begin by considering experiments corresponding to correlated choices of the matrix $\mathbf{R}$. Our results show that as the minimum eigenvalue of the matrix $\mathbf{R}$ increases, there is a decrease in the required number of iterations. This agrees with the theoretical conclusions of Tabeart et al. [2018]. However, the overall effect of reconditioning on convergence speed is less for 1D-Var than was observed in the case of 3D-Var or 4D-Var as described in Weston [2011]. It is likely that this is because the average number of iterations is greater in 3D and 4D-Var, and the maximum permitted number of iterations is much larger than the 10 allowed for the 1D-Var minimisation.

We now consider the two diagonal choices of OEC matrix, $\mathbf{R}_{infl}$ and $\mathbf{R}_{diag}$. The distribution corresponding to E_{diag} is more heavily weighted towards a higher number of iterations than any of the correlated cases. This is not what we might expect from an uncorrelated choice of OEC matrix, particularly as it is well-conditioned compared to most other choices of OEC matrix. In particular, $\lambda_{min}(\mathbf{R}_{diag})$ is greater than the minimum eigenvalue for all choices of correlated OEC matrix apart from $\mathbf{R}_{67}$ (see Table 8.2). In contrast, for the experiment E_{infl} convergence is faster than for any of the other experiments.

As we noted in Section 8.3.2, the minimum eigenvalue of the matrix $\mathbf{R}$ is not the only important property for determining the speed of convergence. The distribution of standard deviations for E_{diag} and E_{infl} is shown in Figure 1 of Weston et al. [2014]. As the standard deviations for $\mathbf{R}_{infl}$ are much larger than the standard deviations for any other choice of $\mathbf{R}$, the ratio of background variance to observation variance will be smaller for E_{infl} than other experiments, resulting in smaller condition numbers of the Hessian and hence faster convergence of the 1D-Var minimisation. We recall from Section 8.3.2 that the ratio of background to observation error variances appears in the bounds on the condition number of the Hessian given by (8.5) in Tabeart et al. [2018] and similar bounds in Haben [2011]. It is clear from these bounds that decreasing the observation error variance will increase the value of the bounds. We can therefore explain the worse convergence seen for E_{diag} by considering channels

107-121 and 128-137, where variances for $\mathbf{R}_{diag}$ are smaller than the variances for correlated choices of $\mathbf{R}$. These channels are sensitive to water vapour, and also correspond to the strongest positive correlations in $\mathbf{R}_{est}$. Typically, inflation is used when correlated errors are not accounted for; here we have the opposite effect with smaller variances for uncorrelated $\mathbf{R}_{diag}$. In terms of the minimisation of the 1D-Var objective function, this means that E_{diag} is pulling much closer to observations for those channels than any of the correlated experiments. This makes it harder to find a solution, resulting in slower convergence.

We now consider how the condition number of the Hessian of the 1D-Var cost function, $\kappa(\mathbf{S})$, changes with the experiment E. From theoretical results developed in Tabeart et al. [2018], in particular the result of Corollary 1, we expect $\kappa(\mathbf{S})$ to decrease as $\lambda_{min}(\mathbf{R})$ increases. The minimum eigenvalues for each choice of OEC matrix, $\mathbf{R}$, discussed here can be seen in Table 8.2. The condition number of $\mathbf{S}$ is computed separately for each objective function. We can therefore consider the maximum, mean and median value of $\kappa(\mathbf{S})$ over the 97330 observations for each experiment. This information is shown in Table 8.3. As discussed in Section 8.3.1 the condition number of any matrix is bounded below by one. We therefore do not include the minimum values of $\kappa(\mathbf{S})$ in the table. We firstly note that the maximum values of $\kappa(\mathbf{S})$ are extremely large, with the largest value occurring for the matrix $\mathbf{R}_{diag}$. For experiments with correlated OEC matrices, increasing $\lambda_{min}(\mathbf{R})$ results in a decrease in the maximum value of $\kappa(\mathbf{R})$. We note that the changes to the condition number for $\mathbf{R}_{67}$ compared to $\mathbf{R}_{500}$ are much larger than the difference in conditioning between other experiments. The maximum value of $\kappa(\mathbf{R})$ for the OEC matrix $\mathbf{R}_{infl}$ is the smallest of all choices of OEC matrix. A decrease in the maximum value of $\kappa(\mathbf{S})$ corresponds to a distribution that has increased weight at the lower end of the spectrum for the iteration count distribution shown in Figure 8.5.

We now consider the mean and the median of $\kappa(\mathbf{S})$. Firstly we note that the values of the mean and median differ by at least one order of magnitude. The distribution of $\kappa(\mathbf{S})$ is not symmetric: it is bounded below by 1, with very large maximum values. The mean is skewed by such outliers and we note that for a boxplot of this data (not shown) the mean does not lie within the interquartile range (IQR) of the data for all experiments other than E_{67} and E_{infl}. Both the maximum and mean of $\kappa(\mathbf{S})$ decrease with increasing $\lambda_{min}(\mathbf{R})$, for correlated OEC matrices. The largest values occur for the experiment E_{diag}, and the smallest for the experiment E_{infl}. In contrast, the median is largest for the experiment E_{infl}, and decreasing $\lambda_{min}(\mathbf{R})$ increases the median value of $\kappa(\mathbf{S})$ for experiments with correlated choices of OEC matrix. Considering the deciles indicates that the spread of $\kappa(\mathbf{S})$ across all observations reduces as more reconditioning is applied.

We have seen that introducing correlated OEC matrices improves convergence and reduces $\kappa(\mathbf{S})$ compared to the current operational choice. Additionally, reducing the target condition number results in further improvements. This behaviour agrees with the theoretical conclusions of Tabeart et al. [2018] that were summarised in Section 8.3.2. For a linear observation operator we expect the upper bound on the condition number of the Hessian to decrease as the minimum eigenvalue of the OEC matrix, $\mathbf{R}$, increases. This is shown in (8.4). This inequality also shows that the ratio between background and observation variance is important for the conditioning of $\mathbf{S}$. The final column of Table 8.3 shows that range of $\kappa(\mathbf{S})$ for the experiment E_{infl} is less than the range for any experiment with a correlated choice of OEC matrix. The variances for $\mathbf{R}_{infl}$ are much larger than the variances for any other OEC matrix considered in this work. We therefore conclude that the qualitative conclusions of Tabeart et al. [2018], as presented in Section 8.3.2, hold in this framework, even in the case of a non-linear observation operator.

8.5.2 Effect of changing the observation error covariance matrix on 1D-Var Retrievals

In this section we consider how changing the OEC matrix impacts the retrieved values of physical variables. In particular we focus on temperature and specific humidity, as we obtain profiles that occur across multiple model levels rather than individual values. We note that the retrieved temperature and humidity values are not passed to the 4D-Var assimilation procedure. However, studying how these variables change for different choices of OEC matrix helps us understand the impact of changing the OEC matrix, $\mathbf{R}$, on the 1D-Var assimilation. Additionally, as part of the 1D-Var assimilation procedure, retrieved standard deviation (RSD) values for each of the retrieval values are derived. The RSD values are calculated as the square root of the diagonal entries of the inverse of the Hessian given by (8.2), i.e. the retrieved analysis error covariance in state variable space. For each 1D-Var assimilation we obtain a different value for RSD for each retrieved variable. We therefore consider the average RSD value for a given experiment and retrieved variable. For temperature and specific humidity this means that we obtain different RSD values for each model level. Comparing the range of differences between retrievals to the RSD values will allow us

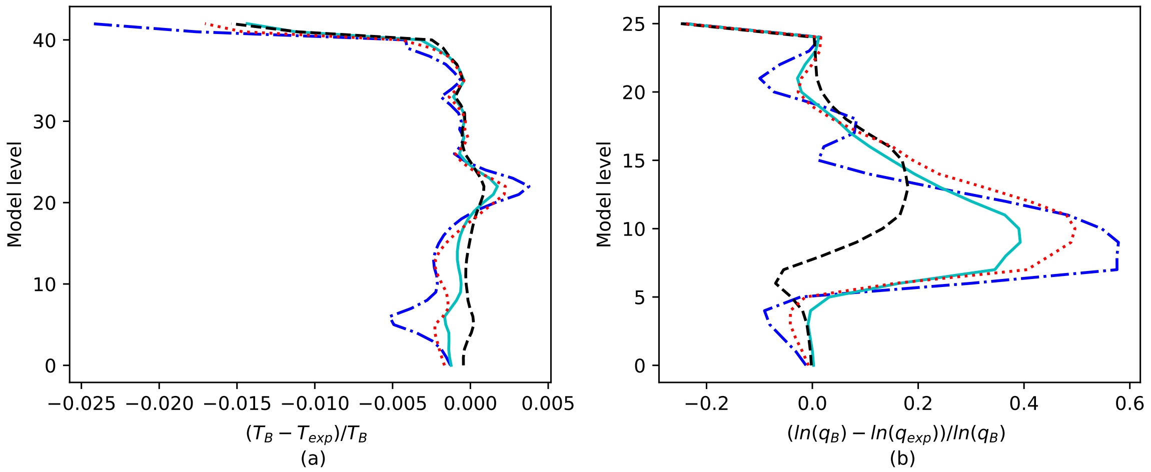


Figure 8.6: Relative difference between background and retrieved profiles from observation at (-33.16N,-32.70E) for 16th June 2016 0000Z for (a) temperature (b) ln(specific humidity). Differences are shown for E_{diag} (blue dot-dashed line), E_{est} (red dotted line), E_{67} (cyan solid line) and E_{infl} (black dashed line).

to determine whether the difference made when changing the OEC matrix, $\mathbf{R}$, is of a similar order to expected variation, or much larger (and hence results in significant differences). We will also compare RSD and differences against BSD values as shown in Figure 8.1.

Figure 8.6 shows the relative difference between background profiles and retrieved profiles for temperature and humidity for observations at the location (-33.16N,-32.70E). Retrievals are shown at pressure levels in the atmosphere. These model levels are determined by the 43 evenly distributed pressure levels in the radiative transfer retrieval algorithm. Specific humidity is only calculated for the lowest 26 model levels. We note that this is the configuration that was used at the time of the experiments (July 2016). Relative differences from the background are much larger for specific humidity profiles than for temperature. However qualitative behaviour is similar for both variables. In both cases E_{diag} is the most different from the background, implying that the use of correlated observation errors increases the weighted importance of the background. Increasing the amount of reconditioning used decreases the norm of the difference between the retrieved profile and the background for all correlated OEC matrices. Hence, applying a larger amount of reconditioning results in a retrieved profile that is closer to the background. Finally, the retrieval corresponding to E_{infl} is closest to the background for both variables. For this case, standard deviations have been inflated, meaning that we expect the retrieved profile

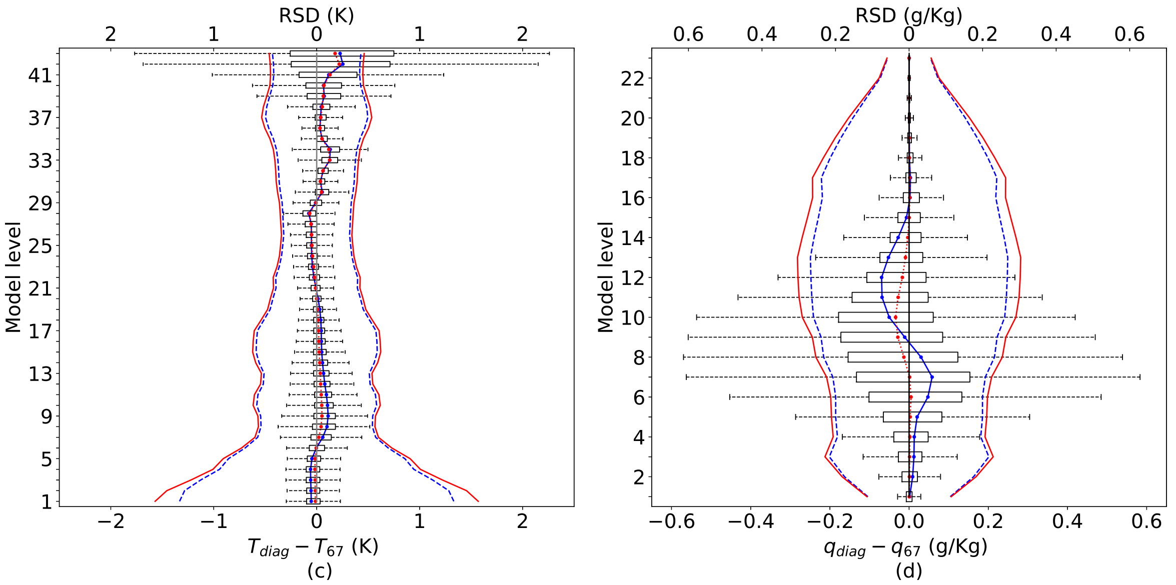


Figure 8.7: Differences in retrievals between E_{diag} and E_{67} for trial on 16th June 0000Z for (a) temperature and (b) ln(specific humidity) for 97330 observations. Dashed lines and solid lines give the mean RSD values for E_{diag} and E_{67} respectively. Dashed lines with dots denote the median and solid lines with dots denote the mean for each pressure level. The solid box contains the middle 50% of the data, and the whiskers (dashed horizontal lines) extend to the quartiles plus/minus 1.5 times the interquartile range (IQR) - the difference between the third and first quartiles. Outliers, which lie outside the range of the whiskers, are not shown.

to fit closer to the background. This is particularly evident for specific humidity where there is a large relative difference between background and retrieved values for model level 7 for E_{diag}, E_{est} and E_{67}. This occurs due to large differences between background and retrieved brightness temperature for channels 128-137, which have water vapour mixing ratio Jacobians that peak at pressure level 7 [Stewart, 2010]. We recall that these channels are sensitive to water vapour, and have the strongest positive correlations in $\mathbf{R}_{est}$. This explains why specific humidity is particularly affected by changes to the OEC matrix for this model level, although we note that noticeable changes also occur for temperature for this model level.

We now consider the differences between retrieved values for E_{diag} and E_{67} for all 97330 observations that were accepted by the 1D-Var routine for all choices of OEC matrix. Figures 8.7a and b are box plots showing the distribution of these differences across each model level for temperature and ln(specific humidity profiles) respectively. The qualitative behaviour for other experiments was very similar and is not shown here. Figures 8.7a shows that for most model levels the whiskers are contained within the average RSD values, and for model levels $1-41$ the central 50% of differences lie within the averaged RSD. This indicates that changing from an uncorrelated to correlated choice of OEC matrix has a generally small impact on temperatures for the majority of model levels compared to RSD.

Mean RSD values for E_{67} and E_{diag} are also very similar, with larger mean RSD values for E_{67} than E_{diag} for all model levels. This is observed for all correlated choices of OEC matrix; the mean RSD is increased for all model levels compared to E_{diag}. This suggests that using a correlated choice of OEC matrix increases the mean RSD for temperature i.e. by introducing correlations we have less confidence in the retrieved values, or 1D-Var analysis. This increase to standard deviations is expected from theoretical and idealised studies [Stewart et al., 2008b, Rainwater et al., 2015, Fowler et al., 2018]. We also note that by including correlations we put less weight on the individual channels but allow more freedom to fit multivariate information arising from the combination of channels. Comparing the RSD values to the BSD values given by Figure 8.1 we find that across all three choices of $\mathbf{B}$, the standard deviation values are similar to RSD values for most model levels. For model levels where the BSDs are smaller than both E_{diag} and experimental RSD, differences are small in comparison to all standard deviation values.

Figure 8.7b shows the differences between retrieved values of specific humidity for E_{diag} and E_{67} for 23 model levels. As was the case for temperature, the mean RSD values for all other choices of experiment are larger than those for E_{diag}. We note that differences for model levels 1 and $18-23$ are very small compared to RSD. However, for model levels $5-12$, the whiskers lie outside the values for mean RSD. This means there is a large proportion of model levels where changing the OEC matrix has a larger impact on retrieved specific humidity values than we would expect due to instrument noise and other quantified types of uncertainty. We also note that for these model levels we have non-zero and non-equal means and medians. This suggests that the distribution of differences is not symmetric. Again, BSD values are larger than RSD values for the majority of model levels for specific humidity. However, whiskers still extend past the BSD values for levels 5 - 10 for all choices of $\mathbf{B}$.

The effect of changing the OEC matrix, $\mathbf{R}$, seems to affect a larger proportion of the retrieved specific humidity values than temperature values. This coincides with the findings of Bormann et al. [2016], Weston et al. [2014]. They found large changes to humidity fields with the introduction of correlated OEC matrices in 4D-Var assimilation procedures, which resulted in improved NWP skill scores.

8.6 Impact on variables that influence 4D-Var routine

In Section 8.5 we showed that the choice of OEC matrix, $\mathbf{R}$, does make a difference to the 1D-Var routine in terms of convergence, and the individual retrieval values. We now consider variables that directly impact the main 4D-Var procedure that is used to initialise forecasts. Changes to the OEC matrix in the 1D-Var routine affect 4D-Var in two main ways: firstly by altering the observations that are accepted by the quality control procedure, and secondly via retrieved values of variables that are not analysed in the 4D-Var state vector. We will consider these two aspects in turn.

8.6.1 Changes to the quality control procedure

In Section 8.5.1 we showed that increasing $\lambda_{min}(\mathbf{R})$ increases the speed of convergence of the 1D-Var routine. We now investigate whether changing the OEC matrix, $\mathbf{R}$, alters the number of observations that pass the quality control step that was described in Section 8.4.1. We also consider how the number of observations accepted by experiment (respectively E_{diag}) and rejected by E_{diag} (respectively experiment) changes for different choices of OEC matrix. This information is presented in Table 8.4.

We begin by considering in more detail why changing the OEC matrix would result in changes to the number of observations that pass quality control. Observations are rejected if the minimisation of the 1D-Var procedure requires more than 10 iterations to converge. In Section 8.5.1 we found that introducing correlated observation error reduces the number of iterations required for convergence, and that decreasing the target condition number increases convergence speed further. This suggests that introducing correlated OEC matrices and using reconditioning will result in a larger number of observations that converge fast enough to pass this aspect of quality control. We therefore expect the use of reconditioning methods to result in a larger number of accepted observations.

The first row of Table 8.4 shows that the number of accepted observations increases as $\lambda_{min}(\mathbf{R})$ (see Table 8.2) increases and the largest number of accepted observations occurs for experiment E_{infl}. This coincides with what we would expect due to alterations in the quality control procedure. However, we note that the number of accepted observations is slightly larger for E_{diag} than E_{est} even though convergence for

	E_{exp}					
Experiment	E_{est}	E_{1500}	E_{1000}	E_{500}	E_{67}	E_{infl}
No. of accepted obs (T)	100655	100795	101002	101341	102333	102859
Accepted by E_{diag} and E_{exp}	99039	99175	99352	99656	100382	100679
Accepted by E_{exp}, rej. by E_{diag}	1616	1620	1650	1685	1951	2180
Accepted by E_{diag}, rej. by E_{exp}	1647	1511	1334	1030	304	7

Table 8.4: Number of observations accepted by the 1D-Var quality control for each experiment (E_{exp}) compared to E_{diag}. For E_{diag} the total number of accepted observations is 100686. Here T refers to the total number of distinct observations (defined in Section 8.4.1) accepted by E_{exp} for each experiment. The number of observations accepted by all experiments is 97330.

E_{est} was faster than for E_{diag} across the set of common observations. The second row of Table 8.4 shows that most observations are accepted by both E_{exp} and E_{diag}. We see that the number of accepted observations increases with $\lambda_{min}(\mathbf{R})$ for correlated choices of $\mathbf{R}$. The largest number of observations is accepted by E_{infl}. The third and fourth rows of Table 8.4 shows the number of observations that are accepted by E_{exp} (respectively E_{diag}) and rejected by E_{diag} (respectively E_{exp}). However, this number is smaller than 2.2% of the total number of observations all choices of E_{exp}. For what follows we shall consider the large majority of observations that are accepted by both E_{diag} and E_{exp}. Although observations that are accepted by only one of E_{diag} and E_{exp} are of interest, the fact that there are very few observations in either of these sets makes it hard to study their properties statistically.

8.6.2 Changes to retrieved values for variables that are not included in the 4D-Var control vector

In this Section we consider how altering the OEC matrix used in the 1D-Var routine alters the retrieved values of variables that are not included in the 4D-Var control vector. For all three variables, Figure 8.8 shows that the majority of retrievals are changed by a small amount for each choice of experiment. The largest differences occur between E_{diag} and E_{exp} for ST, CF and CTP, where the IQR and whiskers are much larger than for any correlated choice of OEC matrix. For correlated OEC matrices, we see a reduction in IQR and whisker length as $\lambda_{min}(\mathbf{R})$ increases. This indicates that as we increase the amount of reconditioning that is applied, the differences between E_{diag} and E_{exp} reduce. However, there are some differences between the variables.

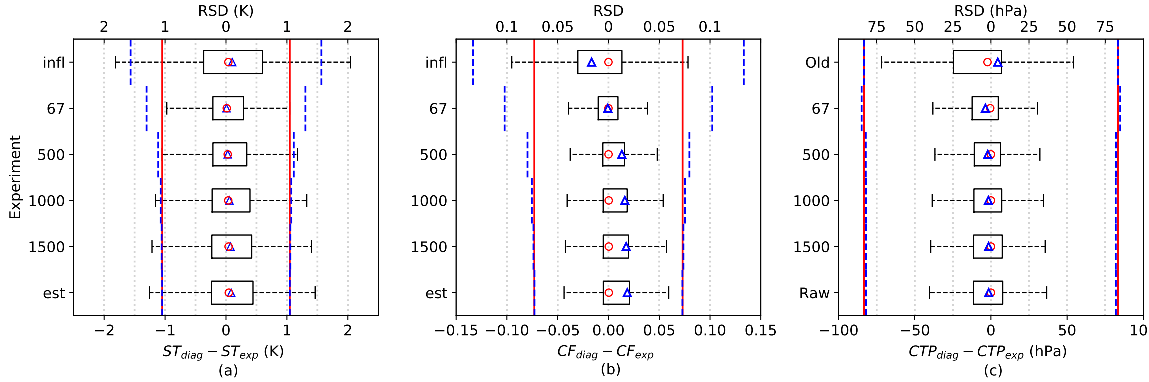


Figure 8.8: Box plot showing differences between retrieved variables for $E_{diag} - E_{exp}$ for (a) ST (skin temperature) (b) CF (cloud fraction) and (c) CTP (cloud top pressure). The circle shows the median, the triangle depicts the mean, the solid box contains the central 50% of data (the interquartile range), and the dashed horizontal lines show the whiskers which extend to the quartile $\pm 1.5 \times IQR$. Vertical dashed lines show the mean retrieved standard deviation (RSD) values for the experiment, and the solid vertical lines shows the mean RSD values for E_{diag}. Outliers (not shown) lie in the range (a) $\pm 33.52K$, (b) ± 1 and (c) $\pm 913.25hPa$. The number of outliers and extreme outliers for these experiments is presented in Tables 8.5 and 8.6.

Firstly, for ST all choices of OEC matrix yield whiskers that are equal to or exceed the RSD values corresponding to E_{diag} (solid line), and all except E_{67} exceed the RSD values for the corresponding experiment (dashed line). In contrast, the whiskers for correlated choices of OEC matrix are well within both RSD values for CF and CTP, as well as the BSD values given in Table 8.1. This shows that compared to expected observation variability, differences between CF and CTP retrievals are small for correlated choices of OEC matrix. However, we recall that BSD values for cloud variables were artificially inflated [Pavelin et al., 2008].

For ST and CTP the values of the mean and median are close for all correlated choices of E_{exp}, and the box and whiskers are fairly symmetric about 0. In contrast, for CF, differences between the mean and median occur, and the box extends further into the positive axis. Cloud errors are expected to vary greatly with the cloud state, meaning that it is difficult to interpret gross statistics [Eyre, 1989]. We include them here for completeness.

For all variables, the majority of retrievals change by a small amount, relative to RSD, when comparing the experiment to E_{diag}. However, Table 8.5 shows that over 15% of observations are classed as outliers for all three variables. These outliers are defined as

	E_{est}	E_{1500}	E_{1000}	E_{500}	E_{67}	E_{infl}
% outliers (ST)	15.1	15.3	15.6	16.3	17.6	15.9
% of outliers (CF)	23.9	24.02	24.2	24.6	25.3	21.4
% outliers (CTP)	22.8	22.8	23.0	22.9	21.4	18.8
Maximum difference (ST (K))	21.67	21.12	21.14	22.38	21.03	26.83
Minimum difference (ST (K))	-33.52	-33.01	-32.14	-29.76	-23.82	-20.88

Table 8.5: Percentage of outliers for cloud fraction, cloud top pressure and skin temperature. Outliers are differences which fall outside the whiskers shown in Figure 8.8. Maximum and minimum differences are shown for skin temperature only; maximum differences for cloud fraction and cloud top pressure are ±1 and $\pm 913.25hPa$ respectively, for all choices of $\mathbf{R}$.

	E_{est}	E_{1500}	E_{1000}	E_{500}	E_{67}	E_{infl}
% large outliers ($\|ST\| > 5K$)	1.6	1.5	1.5	1.4	1.4	3.6
% large outliers ($\|CF\| > 0.25$)	4.9	4.7	4.4	3.9	3.2	7.5
% large outliers ($\|CTP\| > 225hPa$)	3.3	3.3	3.3	3.3	2.7	4.4

Table 8.6: Number of large outliers for cloud fraction, cloud top pressure and skin temperature for each experiment. Large outliers are defined as observations with absolute differences greater than 0.25 for CF, $225hPa$ for CTP and $5K$ for ST. This corresponds to absolute differences greater than approximately 25% of the maximum differences presented in Table 8.5.

observations with retrieval differences that are not between $Q_1 - 1.5IQR$ and $Q3 + 1.5IQR$, where Q_1 and Q_3 denote the first and third quartiles of the data respectively, and are not shown in Figure 8.8. Not all of these outliers represent large differences between retrieved values. Instead we consider 'large' outliers, which we define in this setting as differences larger than 25% of the maximum differences for each variable. For cloud variables the maximum difference is defined by the possible range of values: ±1 and $\pm 913.25hPa$ for CF and CTP respectively. For ST we use the maximum difference between retrievals from the data set. These values are given in Table 8.5.

Table 8.6 shows the percentage of large outliers for each variable, which is much smaller than the total number of outliers for all variables and experiments. For all variables, the number of large outliers decreases with $\lambda_{min}(\mathbf{R})$ for correlated experiments. The experiment E_{infl} has a much greater number of large outliers than any experiment with a correlated choice of OEC matrix, agreeing with earlier findings that the qualitative and quantitative differences between E_{infl} and E_{diag} are much larger than for any other experiment.

As background information has almost no weight for cloud variables, due to inflated BSD values, changing the OEC matrix could result in much larger differences between retrieved values for CF and CTP than for other variables. However, this is not the case for ST, where the maximum differences given in Table 8.5 are extremely large compared to RSD and BSD values. The number of observations with extremely large retrievals is small: for correlated experiments fewer than 10 observations yield absolute differences larger than 20K. These observations can be considered as failures of the 1D-Var algorithm and should be removed by the quality control procedure. This emphasises that when altering the OEC matrix, the quality control procedure needs to be altered as well.

Previous studies by Stewart et al. [2014], Weston et al. [2014], Bormann et al. [2016], Campbell et al. [2017] have shown that the largest impacts of applying the DBCP diagnostic to IASI occur for humidity sounding channels, which will affect clouds and retrieved values associated with clouds. Skin temperature is also sensitive to cloud; although in partly overcast conditions it is possible to retrieve estimates of skin temperature, errors in the modelling of cloud effects are likely to dominate the surface signal [Stewart et al., 2014, Pavelin and Candy, 2014]. In terms of impact, under cloudy conditions the 4D-Var assimilation procedure is less sensitive to skin temperature [Pavelin and Candy, 2014], so it is possible that these large changes to retrievals will not result in large impacts when passed to 4D-Var. However, further work is needed to understand the origin and consequences of these extreme differences fully.

8.7 Conclusions

It is widely known that many observing systems in numerical weather prediction (NWP) have errors that are correlated [Janjić et al., 2018] for reasons including scale mismatch between observation and model resolution, approximations in the observation operator or correlations introduced by preprocessing. However, diagnosed error covariance matrices have been found to be extremely ill-conditioned, and cause convergence problems when used in existing NWP computer systems [Campbell et al., 2017, Weston et al., 2014]. Tabeart et al. [2018] established that increasing the minimum eigenvalue of the OEC matrix improves bounds on the conditioning of the associated linear variational data assimilation problem. This provided insights into possible reconditioning methods which could permit the inclusion of correlation information while ensuring computational efficiency [Tabeart et al., 2019a].

In this chapter we have investigated the impact of changing the OEC matrix for the IASI instrument in the Met Office 1D-Var system, an operational non-linear assimilation system. In particular we have considered how reconditioning methods could permit the implementation of correlated observation error matrices. The 1D-Var system is used for quality control purposes and to retrieve values of variables that are not included in the 4D-Var state vector. As each observation is assimilated individually, it is more straightforward to understand and isolate the effects of using different choices of OEC matrix on retrieved variables and convergence compared to the more complicated 4D-Var procedure.

We found that:

- The current operational choice of observation error covariance (OEC) matrix for IASI results in the slowest convergence of the 1D-Var routine of all OEC matrices considered. Increasing the amount of reconditioning applied to correlated OEC matrices improves convergence of the 1D-Var routine, in accordance with the qualitative theoretical conclusions of Tabeart et al. [2018, 2019a].
- Most experimental choices of correlated OEC matrix resulted in a larger number of IASI observations that were accepted by the 1D-Var routine than the current diagonal operational choice. Increasing the amount of reconditioning applied to correlated OEC matrices increases the number of IASI observations that converge in fewer than 10 iterations, and hence pass the quality control component of 1D-Var.
- Retrieval differences for skin temperature, cloud fraction and cloud top pressure are smaller than retrieved standard deviation values for over 75% of IASI observations for all choices of correlated OEC matrix. Up to 5% of retrievals have large differences relative to the retrieved standard deviation.
- As the minimum eigenvalue of the OEC matrix is increased, the difference between E_{diag} (using the current operational diagonal OEC matrix) and experimental retrieved values reduces.

We also find that for most variables studied RSD values are of a similar size to BSD values. We note that the BSD values for cloud variables are artificially inflated, and are hence an order of magnitude larger than the corresponding RSD values. This indicates that observation information has as large or a larger weight in the 1D-Var objective function than background profiles.

The qualitative conclusions from this work agree with the theoretical results of Tabeart et al. [2018], which prove that for a linear observation operator, increasing the minimum eigenvalue of the OEC matrix is important in terms of convergence of a variational data assimilation routine.

We emphasise that the specific choice of correlated observation error covariance matrices studied in this work are not necessarily more optimal than the current choice of uncorrelated observation error covariance. Although we have studied the effect of changing the OEC matrices within the 1D-Var routine, we have not assessed whether these changes lead to improvement or degradation of either the 1D-Var assimilation system, or the 4D-Var assimilation system and subsequent forecasts. However, our results clearly show that the analysis and speed of convergence of the 1D-Var assimilation problem are sensitive to the choice of observation error covariance matrix, and the use of reconditioning methods.

In particular these convergence results contradict the common assumption that the use of correlated OEC matrices in a variational data assimilation scheme will cause convergence problems. In fact, one key benefit of using correlated OEC matrices in a 1D-Var framework is the increase in convergence speed, particularly when combined with reconditioning methods. At the Met Office the 1D-Var routine is run every 6 hours for the global model so reducing the cost of the routine would save significant computational effort. Additionally, the faster convergence that is achieved by correlated choices of OEC could permit stricter convergence criteria, e.g. reducing the maximum number of iterations from 10 to 8, which would also result in computational savings. However, care needs to be taken to consider how this will interact with other aspects of the quality control procedure and ensure that ‘good’ observations are not rejected.

Changes to OEC matrices also alter the quality control aspect of the 1D-Var procedure, so care needs to be taken to ensure that these changes to the system are well understood. In particular, reducing the number of iterations required for convergence of the 1D-Var routine means that a larger number of observations were accepted by our tests and passed to the 4D-Var routine. For observations that were accepted by all experiments, we considered changes to retrieved estimates for skin temperature, cloud top pressure and cloud fraction. Although changes to the retrieved values with different OEC matrices were small for the majority of observations, for a small percentage of observations, the differences between retrieved values were very large. As ST, CTP and CF are not estimated as part of the 4D-Var procedure, such large changes may have significant effects on the analysis for 4D-Var. The most extreme of these differences (particularly for ST) are unrealistic and can be viewed as 1D-Var failures. This highlights that changes to the 1D-Var system, such as with the introduction of correlated OEC matrices, must be made in conjunction with tuning of the quality control procedures. In general, improvements to convergence need to be be balanced with impacts on other aspects of the assimilation system, such as changes to quality control, analysis fit and forecast skill.

Acknowledgements

This work is funded in part by the EPSRC Centre for Doctoral Training in Mathematics of Planet Earth, the NERC Flooding from Intense Rainfall programme (NE/K008900/1), the EPSRC DARE project (EP/P002331/1) and the NERC National Centre for Earth Observation.

8.8 Summary

In this chapter we presented the first detailed case study of the effects of using correlated observation error and the ridge regression method of reconditioning in a 1D-Var data assimilation framework. This method of reconditioning was studied theoretically in Chapter 7. We find that using correlated observation error covariance matrices results in faster convergence than the operational diagonal covariance matrix. This agrees with the qualitative conclusions of Chapter 5, that increasing small eigenvalues of the observation error covariance matrix will improve the conditioning of the unpreconditioned data assimilation problem. Changes to retrieved variables are mostly small, although a very small number of observations resulted in extremely large changes to retrieved variables. Changes to the quality control procedure need to be taken into account when introducing a new observation error covariance matrix. Using more reconditioning via a smaller target condition number results in further improvements to convergence, and smaller differences for retrieved variables. In the next chapter we summarise the conclusions of this thesis, and suggest ideas for future work.

Chapter 9

Conclusions

The use of correlated observation error covariance (OEC) matrices in data assimilation algorithms for numerical weather prediction (NWP) is of increasing importance [Stewart et al., 2008b, Stewart, 2010] with a growing quantity of satellite data, and the move towards higher resolution forecasts [Rainwater et al., 2015]. However, case studies have found that using correlated OEC matrices degrades the convergence of operational NWP data assimilation schemes (e.g. Weston [2011], Weston et al. [2014], Campbell et al. [2017], Bormann et al. [2015]). Much of the prior work on this topic has been empirical, but small eigenvalues of estimated OEC matrices were thought to be a primary cause of these convergence issues [Weston, 2011, Weston et al., 2014]. 'Reconditioning' techniques have been popular ad hoc methods to mitigate the problems caused by ill-conditioning of estimated OEC matrices [Bormann et al., 2015, Campbell et al., 2017, Weston et al., 2014]. However, it was not well understood how the use of these methods altered the properties of the covariance matrices themselves, or what impact the use of reconditioning techniques had on the overall data assimilation problem. Previous work by Haben et al. [2011a,b] and Haben [2011] used the conditioning of the variational data assimilation problem as a proxy for convergence of conjugate gradient methods and studied the case of uncorrelated observation errors in detail for the unpreconditioned and preconditioned 3D-Var and 4D-Var problems. In this thesis we have developed a theory of conditioning of the Hessian of both the unpreconditioned and preconditioned data assimilation problems for the case of correlated OEC matrices. We have designed a numerical framework that allows us to study interactions between terms of the Hessian. We developed theory for two methods of reconditioning that are used at NWP centres. A case study implementing one of these methods of reconditioning in the Met Office 1D-Var system shows how qualitative conclusions from a more restrictive theory can be applied in practice to realistic applications. Improved understanding of the contribution of correlated OEC matrices to the convergence of data assimilation problems permits the design of computationally efficient methods which allow the inclusion of correlated observation error information. We now provide an overview of the key results of this thesis, and answer the research questions that were proposed in the first chapter. We then discuss ideas for further work.

9.1 Key questions and conclusions

In Chapter 1 we presented the key research questions that we would study in the thesis.

RQ 1: How does introducing correlated observation error affect the conditioning of the Hessian of the variational data assimilation problem?

How are these bounds affected by changes to the observation error covariance matrix? How tight are these bounds for an idealised numerical framework? How well does the behaviour of the condition number of the Hessian represent convergence of the conjugate gradient method?

RQ 2: What is the difference between the preconditioned and unpreconditioned case?

How does the importance of background and observation terms differ from the unpreconditioned case? Does the behaviour of the condition number of the Hessian represent convergence of the conjugate gradient method well for numerical experiments?

RQ 3: How do reconditioning methods modify covariance matrices?

How do reconditioning methods modify correlations and standard deviations associated with the covariance matrix? How is the variational objective function modified by the use of reconditioning methods? How do two commonly-used reconditioning methods compare to multiplicative variance inflation?

RQ 4: What is the impact of using the ridge regression method of reconditioning on an operational data assimilation problem?

How do the qualitative theoretical conclusions from the linear case apply in a nonlinear, realistic setting using the Met Office 1D-Var system? How are the quality control process and retrieved values affected by the introduction of correlated observation error and the use of reconditioning methods?

We now discuss the main conclusions of this thesis, and explicitly consider how the key research questions have been addressed.

In the initial chapters we provided background material for this thesis. In Chapter 2 we defined the variational data assimilation problem. We discussed different sources of observation error, and introduced the diagnostic of Desroziers et al. [2005]. Numerical linear algebra results were presented in Chapter 3. In particular, the condition number was formally defined, and its relationship to the convergence of the conjugate gradient method was established. In Chapter 4 we motivated the use of conditioning as a proxy for convergence of a conjugate gradient method, and highlighted some of the numerical issues associated with the introduction of correlated OEC matrices at NWP centres.

RQ 1: In Chapter 5 we developed general bounds on the condition number of the Hessian of the unpreconditioned variational data assimilation problem. We found that:

- The minimum eigenvalue of the OEC matrix appeared in the denominator of both bounds, meaning that small eigenvalues are likely to yield ill-conditioned Hessians.
- We found experimental cases where both upper and lower bounds are tight.
- Numerical experiments revealed that for different experimental cases, conditioning of the Hessian is dominated by either the background or observation error covariance matrix. The choice of observation network determined the smoothness of the transition between these two regimes.
- In our numerical framework, the conditioning of the Hessian represented the convergence of a conjugate gradient method well in many examples. For instances where the behaviour was different, repeated eigenvalues of the Hessian led to rapid convergence of the conjugate gradient method. This is a well-known case where the condition number provides a very pessimistic upper bound on convergence.

RQ 2: In Chapter 6 we developed bounds on the Hessian of the preconditioned data assimilation problem. We found that:

- The minimum eigenvalue of the OEC matrix appears in both upper and lower bounds, meaning that small eigenvalues of the OEC matrix are likely to lead to ill-conditioned Hessians.

- Numerical experiments revealed that, unlike in the unpreconditioned case, reducing the condition number of the background or observation error covariance matrix did not always decrease the condition number of the Hessian. This behaviour was not well-represented by our bounds, which separate the contribution of each term.
- For many cases, the condition number gives a good indication of how changes to the data assimilation problem are likely to affect convergence of a conjugate gradient method. However, there were also cases where clustered eigenvalues of the preconditioned Hessian led to much faster convergence that would be expected by simply considering its conditioning.

RQ 3: In Chapter 7 we developed theory of two methods of reconditioning, ridge regression and the minimum eigenvalue method, and compared them against multiplicative variance inflation. We found that:

- Both methods of reconditioning increase variances, with ridge regression resulting in larger increases to variances for any choice of covariance matrix. We proved that the ridge regression method strictly decreases the absolute value of off-diagonal correlations. Numerical experiments revealed that the minimum eigenvalue method can increase the absolute value of correlations, but results in smaller absolute changes to correlation entries. However, it can introduce spurious correlations at large distances.
- Both methods of reconditioning reduce the weight on small eigenvalues of the OEC matrix, whereas multiplicative variance inflation reduces the weight on all scales equally.
- An illustrative example showed that both methods of reconditioning are able to make changes to the analysis of a data assimilation problem on smaller scales, whereas multiplicative variance inflation cannot reduce spurious sample error on smaller scales. Reconditioning methods were also found to result in faster convergence of a conjugate gradient method than multiplicative variance inflation.

RQ 4: In Chapter 8 we presented the first detailed case study of the effects of using correlated observation error and the ridge regression method of reconditioning in an operational 1D-Var data assimilation framework. We found that:

- Using correlated OEC matrices resulted in faster convergence than the operational diagonal covariance matrix. Applying ridge regression, which increases the smallest eigenvalue of the OEC matrix, improves convergence of the data assimilation algorithm further.

- Changes to retrieved variables were mostly small, although for a very small number of observations changes to the OEC matrix resulted in extremely large changes to retrieved variables. Using more reconditioning via a smaller target condition number results in further improvements to convergence, and smaller differences from the control for retrieved variables.

We now consider the practical implications for the findings of these results. Firstly, the theoretical importance of the minimum eigenvalue of the OEC matrix coincides with empirical results from operational experiments [Weston, 2011]. As reconditioning techniques were designed to mitigate for small eigenvalues, users can have more confidence that increasing small eigenvalues of the OEC matrix will improve convergence for their system. Additionally, theoretical study of the impact of reconditioning methods on correlations and standard deviations will help users pick most appropriate method for their application, and have a better understanding of how these techniques are likely to change their analysis. Comparison of the two methods of reconditioning with multiplicative variance inflation emphasises the difference between these two methods, in particular regarding their effect on the solution of the variational objective function.

The bounds for both the unpreconditioned and preconditioned case also provide further insight into the interactions between terms. For example, in the unpreconditioned setting, examples were found where the observation operator determined whether the background or observation error covariance matrix dominated the conditioning of the Hessian. In the preconditioned problem it was harder to separate the effect of changing one error covariance matrix on the conditioning of the Hessian. For both formulations the choice of observation network, particularly whether observations were regularly distributed, had a large effect on conditioning and convergence of a conjugate gradient method. In Chapter 8 the ratio between background and observation error variance that occurs in the bounds on the condition number in Chapter 5 was used to understand differences between experiments in an operational system. This shows how the bounds in this thesis can be informative from a qualitative perspective as well as quantitatively for idealised experiments.

9.2 Future work

We begin by considering some major questions of interest to the data assimilation community.

- Is the Control Variable Transform the optimal choice of preconditioning in the presence of correlated OEC matrices? An improved preconditioner may depend on the structure of both the observation and background error covariance matrices; in the numerical experiments presented in this thesis, preconditioning with the background error covariance matrix acted as a good preconditioner due to the circulant structures of both error covariance matrices. However, in the case when the background and observation error covariance matrices have very different correlation structures, an additional or improved preconditioner based on the OEC matrix could be beneficial.

- The theoretical and numerical examples in this thesis have considered linear observation operators. However, for many observation types, such as satellite observations, the observation operator is highly nonlinear [Eyre, 1989]. In Chapter 8 we showed that the qualitative conclusions from the linear setting hold for experiments using a nonlinear observation operator. Many theoretical studies use idealised observation operators (e.g. all state variables observed [Fowler et al., 2018], regularly-spaced observations [Haben et al., 2011a, Waller et al., 2016b]). A more thorough examination of how more realistic choices of observation operator modify the properties of the data assimilation problem could assess the extent to which the conclusions from the linear setting hold for nonlinear choices of observation operator.

We now consider some more specific research questions that arise directly from the work presented in this thesis.

- The conditioning analysis in Chapters 5 and 6 could be extended in a number of ways:
  - For specific problems of interest, exploiting the relevant structure of the data assimilation framework is likely to lead to improved, tighter bounds. This could include: making use of different norms, and taking advantage of circulant error covariance matrices and direct observations to simplify the bounds in this thesis.
  - Further experiments using non-homogeneous or non-spatial correlations would be particularly useful for the preconditioned setting, to assess

whether the conclusions of Chapter 6 hold for the case that the eigenvectors of the background and observation error covariances are unrelated.

- In Chapter 7 we compared two methods of reconditioning. Future lines of research could include
  - Designing new methods of reconditioning, by making use of different norms, or combining aspects of both existing methods. Alternative metrics include finding the nearest correlation matrix using a modified Cholesky factorisation [Higham and Strabić, 2016] or using the entropy loss function [Lin et al., 2014].
  - Developing more rigorous and specific methods to select a value of κ_{max}.
- The case study using the Met Office 1D-Var system in Chapter 8 could be extended to assess the impact of changing the OEC matrix used in the 1D-Var routine on forecast performance. Preliminary tests (not presented in this thesis) showed that passing the results of the 1D-Var assimilation to the operational 4D-Var routine degrades forecast performance, even though most differences between the control and experimental 1D-Var outputs were small. However, changes to the quality control procedure in the 1D-Var routine meant that a small number of observations with gross errors were passed to the 4D-Var routine. Changes to the quality control procedure would be necessary in order to perform a true comparison.

Finally, the theory of conditioning developed in this thesis could be extended to consider a broader class of problems. One such problem of interest is comparing different optimisation techniques for data assimilation problems. The inner loop of the variational data assimilation problem is often solved on a lower-dimension subspace in order to reduce the cost of each iteration. A natural approach is to solve the inner loop on a lower-resolution grid. However, it has been shown that more accurate low order approximations to the linearised objective function can be obtained using model reduction methods [Lawless et al., 2008]. A theoretical comparison of model reduction methods (for example the use of ensembles, or projections of the existing problem into a lower dimensional space) on conditioning and convergence could provide insight into their relative computational efficiency and accuracy.

Bibliography

A. Apte, C. K. R. T. Jones, A. M. Stuart, and J. Voss. Data assimilation: Mathematical and statistical perspectives. *International Journal for Numerical Methods in Fluids*, 56(8):1033–1046, 2008.

O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, 1996.

R. N. Bannister. Review: 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.

R. N. Bannister. A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143(703):607–633, 2017.

J. M. Bardesley, A. Parker, A. Solonen, and M. Howard. Krylov space approximate Kalman filtering. *Numerical Linear Algebra with Applications*, 20:171–184, 2013.

K. Bathmann. Justification for estimating observation-error covariances with the desroziers diagnostic. *Quarterly Journal of the Royal Meteorological Society*, 144 (715):1965–1974, 2018.

P. Bauer, A. Thorpe, and G. Brunet. The quiet revolution of numerical weather prediction. *Nature*, 525:47–55, 2015.

G. V. Bennitt, H. R. Johnson, P. P. Weston, J. Jones, and E. Pottiaux. An assessment of ground-based GNSS Zenith Total Delay observation errors and their correlations using the Met Office UKV model. *Quarterly Journal of the Royal Meteorological Society*, 143(707):2436–2447, 2017.

D. S. Bernstein. *Matrix mathematics : theory, facts, and formulas*. Princeton University Press, Princeton, N.J. ; Oxford, 2nd ed. edition, 2009.

P. J. Bickel and E. Levina. Covariance regularization by thresholding. *The Annals of Statistics*, 36(6):2577–2604, December 2008.

N. Bormann, S. Saarinen, G. Kelly, and J.-N. Thépaut. The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Monthly Weather Review*, 131:706–718, 2003.

N. Bormann, A. J. Geer, and P. Bauer. Estimates of observation-error characteristics in clear and cloudy regions for microwave imager radiances from numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 137:2014–2023, 2011.

N. Bormann, M. Bonavita, R. Dragani, R. Eresmaa, M. Matricardi, and T. McNally. Enhancing the impact of IASI observations through an updated observation error covariance matrix. Technical Memoranda 756, ECMWF, Reading, UK, 2015.

N. Bormann, M. Bonavita, R. Dragani, R. Eresmaa, M. Matricardi, and A. McNally. Enhancing the impact of IASI observations through an updated observation error covariance matrix. *Quarterly Journal of the Royal Meteorological Society*, 142(697): 1767–1780, 2016.

K. L. Brown, I. Gejadze, and A. Ramage. A multilevel approach for computing the limited-memory hessian and its inverse in variational data assimilation. *SIAM Journal on Scientific Computing*, 38(5):2934–2963, 2016.

C. J. Budd, M. A. Freitag, and N. K. Nichols. Regularization techniques for ill-posed inverse problems in data assimilation. *Computers & Fluids*, 46(1):168–173, 2011.

M. Buehner. Error Statistics in Data Assimilation: Estimation and Modelling. In W. Lahoz, B. Khattotov, and R. Menard, editors, *Data Assimilation - Making Sense of Observations*, pages 93–112. Springer-Verlag-Heidelberg, 2010.

W. F. Campbell, E. A. Satterfield, B. Ruston, and N. L. Baker. Accounting for correlated observation error in a dual-formulation 4D variational data assimilation system. *Monthly Weather Review*, 145(3):1019–1032, 2017.

A. Carrassi, M. Bocquet, L. Bertino, and G. Evensen. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):e535, 2018.

G. Chalon, F. Cayla, and D. Diebel. IASI: An advanced sounder for operational meteorology. In *Proceedings of IAF, Toulouse, France*, 2001.

W. C. Chao and L.-P. Chang. Development of a Four-Dimensional Variational Analysis System Using the Adjoint Method at GLA. Part 1: Dynamics. *Monthly Weather Review*, 120(8):1661–1674, 1992.

W. Cheney. *Numerical mathematics and computing*. Brooks/Cole Cengage Learning, Pacific Grove, Calif., 5th ed., international ed. edition, 2005.

A. M. Clayton, A. C. Lorenc, and D. M. Barker. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, 139:1445–1461, 2013.

A. D. Collard. Selection of IASI channels for use in numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 133(629):1977–1991, 2007.

A. D. Collard, A. P. McNally, F. I. Hilton, S. B. Healy, and N. C. Atkinson. The use of principal component analysis for the assimilation of high-resolution infrared sounder observations for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 136(653):2038–2050, 2010.

E. S. Cooper, S.L. Dance, J. Garcia-Pintado, N.K. Nichols, and P.J. Smith. Observation impact, domain length and parameter estimation in data assimilation for flood forecasting. *Environmental Modelling & Software*, 104:199 – 214, 2018.

M. Cordoba, S. L. Dance, G. A. Kelly, N. K. Nichols, and J. A. Waller. Diagnosing Atmospheric Motion Vector observation errors for an operational high resolution data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 143(702):333–341, 2017.

S. L. Cotter, M. Dashti, and A. M. Stuart. Variational data assimilation using targetted random walks. *International Journal for Numerical Methods in Fluids*, 68 (4):403–421, 2012.

P. Courtier, J.-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4d-var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519):1367–1387, 1994.

R. Daley. *Atmospheric Data Analysis*. Cambridge University Press, New York, 1991.

S. L. Dance, S. P. Ballard, R. N. Bannister, P. Clark, H. L. Cloke, T. Darlington, D. L. A. Flack, S. L. Gray, L. Hawkness-Smith, N. Husnoo, A. J. Illingworth, G. A. Kelly, H. W. Lean, D. Li, N. K. Nichols, J. C. Nicol, A. Oxley, R. S. Plant, N. M. Roberts, I. Roulstone, D. Simonin, R. J. Thompson, and J. A. Waller.

Improvements in forecasting intense rainfall: Results from the franc (forecasting rainfall exploiting new data assimilation techniques and novel observations of convection) project. *Atmosphere*, 10(3):125, 2019.

P. J. Davis. *Circulant Matrices.* New York: Wiley, 1979.

G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131:3385–3396, 2005.

H. S. Dollar, N. I. M. Gould, M. Stoll, and A. J. Wathen. Preconditioning saddle-point systems with applications in optimization. *SIAM Journal on Scientific Computing*, 32(1):249–270, 2010.

Ronald M. Errico and Kevin D. Raeder. An examination of the accuracy of the linearization of a mesoscale model with moist physics. *Quarterly Journal of the Royal Meteorological Society*, 125(553):169–195, 1999.

J. R. Eyre. Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory and simulation for TOVS. *Quarterly Journal of the Royal Meteorological Society*, 115(489):1001–1026, 1989.

K. Fan. Convex sets and their applications. Lecture notes, Applied Mathematics Division, Argonne National Laboratory, 1959.

M. Fisher. Minimization algorithms for variational data assimilation. In *Developments in Numerical Methods for Atmospheric Modelling*, pages 364–385. ECMWF, Reading,UK, 7-11 September 1998, 1998.

M. Fisher and E. Andersson. Developments in 4D-Var and Kalman Filtering. ECMWF Technical Memorandum 347, 2001.

A. M. Fowler. A sampling method for quantifying the information content of IASI channels. *Monthly Weather Review*, 145(2):709–725, 2017.

A. M. Fowler, S. L. Dance, and J. A. Waller. On the interaction of observation and prior error correlations in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 144(710):48–62, 2018.

A.M. Fowler. Data compression in the presence of observational error correlations. *Tellus A: Dynamic Meteorology and Oceanography*, 71(1):1–16, 2019.

A. Gambacorta and C. D. Barnet. Methodology and information content of the noaa nesdis operational channel selection for the cross-track infrared sounder (cris). *IEEE Transactions on Geoscience and Remote Sensing*, 51(6):3207–3216, June 2013.

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

P. Gauthier, P. Du, S. Heilliette, and L. Garand. Convergence issues in the estimation of interchannel correlated observation errors in infrared radiance data. *Monthly Weather Review*, 146(10):3227–3239, 2018.

A. J. Geer. Correlated observation error models for assimilating all-sky infrared radiances. *Atmospheric Measurement Techniques*, 12(7):3629–3657, 2019.

J. E. Gentle. *Matrix Algebra: Theory, Computations, and Applications in Statistics.* Springer, 2007.

M. Ghil. Meteorological data assimilation for oceanographers. part i: Description and theoretical framework. *Dynamics of Atmospheres and Oceans*, 13(3):171 – 218, 1989.

M. Ghil and P. Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. volume 33 of *Advances in Geophysics*, pages 141 – 266. Elsevier, 1991.

P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization.* Academic Press, Amsterdam, London, 1986.

G. H. Golub and C. F. Van Loan. *Matrix Computations.* The John Hopkins University Press, third edition, 1996.

S. Gratton, A. S. Lawless, and N. K. Nichols. Approximate Gauss-Newton methods for nonlinear least squares problems. *SIAM Journal on Optimization*, 18(1): 106–132, 2007.

R. M. Gray. Toeplitz and circulant matrices: A review. *Foundations and Trends in Communications and Information Theory*, 2(3):155–239, 2006.

O. Guillet, A. T. Weaver, X. Vasseur, Y. Michel, S. Gratton, and S. Gürol. Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh. *Quarterly Journal of the Royal Meteorological Society*, 145(722):1947–1967, 2019.

S. A. Haben. *Conditioning and preconditioning of the minimisation problem in variational data assimilation*. PhD thesis, University of Reading, 2011.

S. A. Haben, A. S. Lawless, and N. K. Nichols. Conditioning of the 3DVAR Data Assimilation Problem. Mathematics Report Series 03/2009, University of Reading, 2009.

S. A. Haben, A. S. Lawless, and N. K. Nichols. Conditioning and preconditioning of the variational data assimilation problem. *Computers & Fluids*, 46(1):252–256, 2011a.

S. A. Haben, A. S. Lawless, and N. K. Nichols. Conditioning of incremental variational data assimilation, with application to the Met Office system. *Tellus A: Dynamic Meteorology and Oceanography*, 64(4):782–792, 2011b.

T. M. Hamill, J. S. Whitaker, and C. Snyder. Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Monthly Weather Review*, 129(11):2776–2790, 2001.

P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Society for Industrial and Applied Mathematics, Philadelphia, 1998.

D. A. Harville. *Matrix Algebra from a Statistician's Point of View*. Springer-Verlag, New York., 1997.

S. B. Healy and A. A. White. Use of discrete Fourier transforms in the 1D-Var retrieval problem. *Quarterly Journal of the Royal Meteorological Society*, 131(605): 63–72, 2005.

S. Heilliette and L. Garand. Impact of accounting for inter-channel error covariances at the canadian meteorological centre. In *Proc. 2015 EUMETSAT Meteorological Satellite Conf.*, page 8. Toulouse, France, EUMETSAT, 2015. URL `https://www.eumetsat.int/website/home/News/ConferencesandEvents/PreviousEvents/DAT_2305526.html`.

N. J. Higham. Computing the nearest correlation matrix - a problem from finance. *IMA Journal of Numerical Analysis*, 22(3):329–343, 2002.

N. J. Higham and N. Strabić. Bounds for the distance to the nearest correlation matrix. *SIAM Journal on Matrix Analysis and Applications*, 37(3):1088–1102, 2016.

N. J. Higham, N. Strabić, and V. Šego. Restoring definiteness via shrinking, with an application to correlation matrices with a fixed block. *SIAM Review*, 58(2):245–263, 2016.

F. Hilton, N. C. Atkinson, S. J. English, and J. R. Eyre. Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments. *Quarterly Journal of the Royal Meteorological Society*, 135:495–505, 2009.

A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

C. R. Horn, R. A. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, 1991.

K. E. Howes, A. M. Fowler, and A. S. Lawless. Accounting for model error in strong-constraint 4d-var data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143(704):1227–1240, 2017.

L. Isaksen. Data assimilation on future computer architectures. In *Seminar on Data assimilation for atmosphere and ocean, 6-9 September 2011*, pages 301–322, Shinfield Park, Reading, 2012. ECMWF. URL `https://www.ecmwf.int/node/10118`.

T. Janjić, N. Bormann, M. Bocquet, J. A. Carton, S. E. Cohn, S. L. Dance, S. N. Losa, N. K. Nichols, R. Potthast, J. A. Waller, and P. Weston. On the representation error in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1257–1278, 2018.

J. Jeong and M. Jun. Covariance models on the surface of a sphere: when does it matter. *Stat*, 4:167–182, 2015.

C. Johnson. *Information Content of Observations in variational data assimilation.* PhD thesis, University of Reading, 2003.

C. Johnson, B. J. Hoskins, and N. K. Nichols. A singular vector perspective of 4d-var: Filtering and interpolation. *Quarterly Journal of the Royal Meteorological Society*, 131 (605):1–19, 2005.

E. Kalnay. *Atmospheric Modeling, Data Assimilation, and Predictability.* Cambridge University Press, 2002.

A. S. Lawless, S. Gratton, and N. K. Nichols. An investigation of incremental 4D-Var using non-tangent linear models. *Quarterly Journal of the Royal Meteorological Society*, 131:459–476, 2005b.

A. S. Lawless, N. K. Nichols, C. Boess, and A. Bunse-Gerstner. Approximate Gauss-Newton methods for optimal state estimation using reduced-order models. *International Journal for Numerical Methods in Fluids*, 56(8):1367–1373, 2008.

A. S. Lawless, S. Gratton, and N. K. Nichols. Approximate iterative methods for variational data assimilation. *International Journal for Numerical Methods in Fluids*, 47(10-11):1129–1135, 2005a.

O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88:365–411, 2004.

J. M. Lewis, S. Lakshmivarahan, and S. K. Dhall. *Dynamic Data Assimilation: A Least Squares Approach*. Cambridge University Press, 2006.

L. Lin, N. J. Higham, and J. Pan. Covariance structure regularization via entropy loss function. *Computational Statistics & Data Analysis*, 72:315–327, 2014.

Y. Liu, L. Zhang, and Z. Lian. Conjugate gradient algorithm in the four-dimensional variational data assimilation system in grapes. *Journal of Meteorological Research*, 32(6):974–984, 2018.

Z.-Q. Liu and F. Rabier. The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Quarterly Journal of the Royal Meteorological Society*, 129(594):3013–3035, 2003.

A. C. Lorenc, S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, and F. W. Saunders. The Met. Office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126:2991–3012, 2000.

Andrew C. Lorenc and F. Rawlins. Why does 4d-var beat 3d-var? *Quarterly Journal of the Royal Meteorological Society*, 131(613):3247–3257, 2005.

C. Lupu, C. Cardinali, and A. P. McNally. Adjoint-based forecast sensitivity applied to observation-error variance tuning. *Quarterly Journal of the Royal Meteorological Society*, 141:3157–3165, 2015.

A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer, second edition, 2011.

MATLAB. *(R2016b)*. The MathWorks Inc., Natick, Massachusetts, 2016.

MATLAB. *(R2018b)*. The MathWorks Inc., Natick, Massachusetts, 2018b.

M. Matricardi, F. Chevallier, G. Kelly, and J.-N. Thépaut. An improved general fast radiative transfer model for the assimilation of radiance observations. *Quarterly Journal of the Royal Meteorological Society*, 130:153–173, 2004.

A. P. McNally, P. D. Watts, J. A. Smith, R. Engelen, G. A. Kelly, J. N. Thépaut, and M. Matricardi. The assimilation of airs radiance data at ecmwf. *Quarterly Journal of the Royal Meteorological Society*, 132(616):935–957, 2006.

Y. Mei. Computing the square roots of a class of circulant matrices. *Journal of Applied Mathematics*, 2012. URL `https://doi.org/10.1155/2012/647623`.

R. Ménard. Error covariance estimation methods based on analysis residuals: theoretical foundation and convergence properties derived from simplified observation networks. *Quarterly Journal of the Royal Meteorological Society*, 142: 257–273, 2016.

B. Ménétrier, T. Montmerle, Y. Michel, and L. Berre. Linear filtering of sample covariances for ensemble-based data assimilation. Part I: Optimality criteria and application to variance filtering and covariance localization. *Monthly Weather Review*, 143(5):1622–1643, 2015.

C. J. Merchant, O. Embury, J. Robert-Jones, E. Fiedler, C. E. Bulgin, G. K. Corlett, S. Good, A. McLaren, N. Rayner, S. Morak-Bozzo, and C. Donlon. Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). *Geoscience Data Journal*, 1(2): 179–191, 2014.

Y. Michel. Revisiting Fisher's approach to the handling of horizontal spatial correlations of the observation errors in a variational framework. *Quarterly Journal of the Royal Meteorological Society*, 144(716):2011–2025, 2018.

S. Migliorini. On the equivalence between radiance and retrieval assimilation. *Monthly Weather Review*, 140(1):258–265, 2012.

A. J. F. Moodey, A. S. Lawless, R. W. E. Potthast, and P. J. van Leeuwen. Nonlinear error dynamics for cycled data assimilation methods. *Inverse Problems*, 29(2): 025002, 2013.

G. Nakamura and R. Potthast. *Inverse Modeling: an introduction to the theory and methods of inverse problems and data assimilation.* 2053-2563. IOP Publishing, 2015.

J. Nocedal. *Numerical optimization.* Springer series in operations research and financial engineering. Springer, New York ; London, 2nd edition, 2006.

E. G. Pavelin and B. Candy. Assimilation of surface-sensitive infrared radiances over land: Estimation of land surface temperature and emissivity. *Quarterly Journal of the Royal Meteorological Society*, 140(681):1198–1208, 2014.

E. G. Pavelin, S. J. English, and J. R. Eyre. The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 134(632):737–749, 2008.

J. Pestana and A. J. Wathen. Natural preconditioning and iterative methods for saddle point systems. *SIAM Review*, 57(1):71–91, 2015.

E. M. Pinnington, E. Casella, S. L. Dance, A. S. Lawless, J. I. L. Morison, N. K. Nichols, M. Wilkinson, and T. L. Quaife. Investigating the role of prior and observation error correlations in improving a model forecast of forest carbon balance using four-dimensional variational data assimilation. *Agricultural and Forest Meteorology*, 228-229:299–314, 2016.

E. M. Pinnington, E. Casella, S. L. Dance, A. S. Lawless, J. I. L. Morison, N. K. Nichols, M. Wilkinson, and T. L. Quaife. Understanding the effect of disturbance from selective felling on the carbon dynamics of a managed woodland by combining observations with model predictions. *Journal of Geophysical Research: Biogeosciences*, 122(4):886–902, 2017.

M. Pourahmadi. *High-dimensional covariance estimation.* Wiley series in probability and statistics. Wiley, Hoboken, NJ, 2013.

C. Prates, S. Migliorini, L. Stewart, and J. Eyre. Assimilation of transformed retrievals obtained from clear-sky IASI measurements. *Quarterly Journal of the Royal Meteorological Society*, 142:1697–1712, 2016.

H. Qi and D. Sun. Correlation stress testing for value-at-risk: an unconstrained convex optimization approach. *Computational Optimization and Applications*, 45 (2):427–462, 2010.

F. Rabier, J.-N. Thépaut, and P. Courtier. Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1861–1887, 1998.

F. Rabier, N. Fourrié, D. Chafäi, and P. Prunet. Channel selection methods for infrared atmospheric sounding interferometer radiances. *Quarterly Journal of the Royal Meteorological Society*, 128(581):1011–1027, 2002.

S. Rainwater, C. H. Bishop, and W. F. Campbell. The benefits of correlated observation errors for small scales. *Quarterly Journal of the Royal Meteorological Society*, 141:3439–3445, 2015.

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.

C. D. Rodgers. Information content and optimisation of high spectral resolution remote measurements. *Advances in Space Research*, 21(3):361 – 367, 1998.

C. D. Rodgers. *Inverse methods for atmospheric sounding : theory and practice.* Series on atmospheric, oceanic and planetary physics ; v.2. World Scientific, Singapore ; [River Edge, N.J.], 2000.

S. J. Schiff. *Neural Control Engineering The Emerging Intersection Between Control Theory and Neuroscience.* Computational Neuroscience. MIT Press, Cambridge, 2011.

J. R. Schott. *Matrix Analysis for Statistics.* John Wiley & Sons, Incorporated, New York, 2016.

D. Simonin, S. P. Ballard, and Z. Li. Doppler radar radial wind assimilation using an hourly cycling 3d-var with a 1.5 km resolution version of the Met Office Unified Model for nowcasting. *Quarterly Journal of the Royal Meteorological Society*, 140: 2298–2314, 2014.

D. Simonin, J. A. Waller, S. P. Ballard, S. L. Dance, and N. K. Nichols. A pragmatic strategy for implementing spatially correlated observation errors in an operational system: an application to Doppler radar winds. *Quarterly Journal of the Royal Meteorological Society*, 2019. doi: https://doi.org/10.1002/qj.3592.

J. O. Skøien and G. Blöschl. Sampling scale effects in random fields and implications for environmental monitoring. *Environmental Monitoring and Assessment*, 114(1): 521–552, 2006.

P. J. Smith, A. S. Lawless, and N. K. Nichols. Treating sample covariances for use in strongly coupled atmosphere-ocean data assimilation. *Geophysical Research Letters*, 45(1):445–454, 2018.

L. M. Stewart. *Correlated observation errors in data assimilation.* PhD thesis, University of Reading, 2010.

L. M. Stewart, J. Cameron, S. L. Dance, S. English, J. Eyre, and N. K. Nichols. Observation error correlations in IASI radiance data. Mathematics report series, University of Reading, Reading, UK, 2008a.

L. M. Stewart, S. L. Dance, and N. K. Nichols. Correlated observation errors in data assimilation. *International Journal for Numerical Methods in Fluids*, 56:1521–1527, 2008b.

L. M. Stewart, S. L. Dance, and N. K. Nichols. Data assimilation with correlated observation errors: experiments with a 1-D shallow water model. *Tellus A: Dynamic Meteorology and Oceanography*, 65:19546 (14pp), 2013.

L. M. Stewart, S. L. Dance, N. K. Nichols, J. R. Eyre, and J. Cameron. Estimating interchannel observation-error correlations of IASI radiance data in the Met Office system. *Quarterly Journal of the Royal Meteorological Society*, 140:1236–1244, 2014.

E. Süli and D. F. Mayer. *An Introduction to Numerical Analysis.* Cambridge University Press, 2003.

J. M. Tabeart, S. L. Dance, S. A. Haben, A. S. Lawless, N. K. Nichols, and J. A. Waller. The conditioning of least squares problems in variational data assimilation. *Numerical Linear Algebra with Applications*, 25(5):e2165, 2018.

J. M. Tabeart, S. L. Dance, A. S. Lawless, N. K. Nichols, and J. A. Waller. Improving the conditioning of estimated covariance matrices. 2019a. In press, Tellus A: Dynamic Meteorology and Oceanography., https://arxiv.org/abs/1810.10984.

J. M. Tabeart, S. L. Dance, A. S. Lawless, N. K. Nichols, J. A. Waller, S. Migliorini, and F. Hilton. The impact of reconditioning of the correlated observation error covariance matrix on the Met Office system. 2019b. Submitted, Quarterly Journal Royal Meteorological Society, http://arxiv.org/abs/1908.04071.

M. Takana and K. Nakata. Positive definite matrix approximation with condition number constraint. *Optimisation Letters*, 8:939–947, 2014.

H. J. Thiebaux. Anisotropic correlation functions for objective analysis. *Monthly Weather Review*, 104(8):994–1002, 1976.

L. N. Trefethen and D. Bau. *Numerical linear algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 1997.

Y. Trémolet. Incremental 4d-var convergence study. *Tellus A: Dynamic Meteorology and Oceanography*, 59(5):706–718, 2007.

J. A. Waller, S. L. Dance, A. S. Lawless, and N. K. Nichols. Estimating correlated observation error statistics using an ensemble transform Kalman filter. *Tellus A: Dynamic Meteorology and Oceanography*, 66(1):23294, 2014a.

J. A. Waller, S. L. Dance, A. S. Lawless, N. K. Nichols, and J. R. Eyre. Representativity error for temperature and humidity using the Met Office high-resolution model. *Quarterly Journal of the Royal Meteorological Society*, 140: 1189–1197, 2014b.

J. A. Waller, S. P. Ballard, S. L. Dance, G. Kelly, N. K. Nichols, and D. Simonin. Diagnosing Horizontal and Inter-Channel Observation Error Correlations for SEVIRI Observations Using Observation-Minus-Background and Observation-Minus-Analysis Statistics. *Remote Sensing*, 8 (7):851, 2016a.

J. A. Waller, S. L. Dance, and N. K. Nichols. Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. *Quarterly Journal of the Royal Meteorological Society*, 142:418–431, 2016b.

J. A. Waller, D. Simonin, S. L. Dance, N. K. Nichols, and S. P. Ballard. Diagnosing Observation Error Correlations for Doppler Radar Radial Winds in the Met Office UKV Model Using Observation-Minus-Background and Observation-Minus-Analysis Statistics. *Monthly Weather Review*, 144(10):3533–3551, 2016c.

J. A. Waller, S. L. Dance, and N. K. Nichols. On diagnosing observation-error statistics with local ensemble data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143(708):2677–2686, 2017.

J. A. Waller, E. Bauernschubert, S. L. Dance, N. K. Nichols, R. Potthast, and D. Simonin. Observation error statistics for Doppler Radar radial wind superobservations assimilated into the DWD COSMO-KENDA system. *Monthly Weather Review*, 2019. URL `https://doi.org/10.1175/MWR-D-19-0104.1`.

B. Wang and F. Zhang. Some inequalities for the eigenvalues of the product of positive semidefinite Hermitian matrices. *Linear Algebra and Its Applications*, 160: 113–118, 1992.

T. Wang, J. Fei, X. Cheng, X. Huang, and J. Zhong. Estimating the correlated observation-error characteristics of the chinese fengyun microwave temperature sounder and microwave humidity sounder. *Advances in Atmospheric Sciences*, 35 (11):1428–1441, 2018.

P. Weston. Progress towards the implementation of correlated observation errors in 4D-Var. Forecasting research technical report 560, Met Office, Exeter, UK, 2011.

P. P. Weston, W. Bell, and J. R. Eyre. Accounting for correlated error in the assimilation of high-resolution sounder data. *Quarterly Journal of the Royal Meteorological Society*, 140:2420–2429, 2014.

J. H. Wilkinson. *The Algebraic Eigenvalue Problem.* Clarendon Press, 1965.

A. M. Yaglom. *Correlation Theory of Stationary and Related Random Functions, Volume I: Basic Results.* Springer-Verlag, 1986.