The Effect of Numerical Model Error on Data Assimilation ☆

S.E.Jenkins[a,b,1,*], C.J.Budd[b], M.A.Freitag[b], N.D.Smith[a]

[a]*Department of Electronic and Electrical Engineering, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom*
[b]*Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom*

Abstract

The effects of numerical model error on 4D-Variational data assimilation (4D-Var) are investigated. 4D-Var is a method used to create an initialisation for a numerical model, that best replicates subsequent observations of the system it models. The numerical error introduced by this model is considered in the form of numerical dissipation and dispersion. We find that a solely numerically dispersive model results in destructive interference and the loss of some wavenumber components in the initialisation.

We also determine upper bounds for the error in the initialisation due to numerical model error with and without observation errors. The bounds are found to depend on the regularity of the true initial condition. There is a critical number of discretisation points where both errors are minimised. Numerical results are presented to demonstrate the effectiveness of the upper bounds.

Even if observation errors are additive, uncorrelated white noise, the numerical scheme has the potential to introduce correlated noise structures into the initialisation and its subsequent forecast. However, this effect is reduced through the use of a numerically non-dissipative model.

Keywords: data assimilation, numerical model error, observation error

1. Introduction

This paper investigates the influence of numerical model error on the method of *Four-Dimensional Variational data assimilation* (4D-Var). This is a method widely used for comparing a forecast with data to initialise a system and is of particular interest due to its applications in numerical weather prediction (NWP).

The accuracy of the initialisation and its subsequent forecast is affected by many different sources of error [1]. Examples are observation errors due to instrument miscalibration and model error in the forward model. Model error can be viewed in two forms; inaccurate model equations and numerical model error. The former is introduced through a failure of the model equations to capture a property of the physical system, whilst the latter is due to errors introduced when numerically solving the model equations.

This paper aims to provide an insight into the affects of numerical model error on the accuracy of the initialisation. The impact of numerical model error is often ignored or treated as a random variable [2]. However the numerical model can result in significant and non-random errors.

Section 2 states the assumptions placed upon 4D-Var throughout this paper and introduces the linear advection equation as the forward problem. Three finite difference schemes are chosen as forward models; the Upwind, Preissman Box and Lax-Wendroff schemes. The effects of numerical dissipation and dispersion on the initialisation are then reviewed through these schemes. Section 3 develops a formulation for the initialisation which allows the effects of numerical dissipation and dispersion on the true initial condition to be analysed. Section 4 explores the error in the

☆ This document is a collaborative effort.

* Corresponding author.

Email addresses: S.E.Jenkins@bath.ac.uk (S.E.Jenkins), C.J.Budd@bath.ac.uk (C.J.Budd), M.A.Freitag@bath.ac.uk (M.A.Freitag), N.Smith@bath.ac.uk (N.D.Smith)

[1] Funded by a University of Bath research studentship.

initialisation measured in the l_2-norm, through the results of section 3. We find that in the absence of observation errors the rate of decay of the error in the initialisation, with respect to the number of spatial mesh points N, increases with the regularity of the true initial condition. In the presence of observation errors, the same result holds until a critical value of N is reached. At this point, the error begins to increase due to observation errors. Performing 4D-Var at this value of N minimises the effects of numerical model error and observation errors on the initialisation.

2. Problem formulation

4D-Var is a method used to solve an inverse problem; given a set of observations of a physical system taken over a period of time, a numerical model of the system and a priori information on the initial condition for the system, the method estimates an initial condition for the numerical model that best replicates the true state of the system. To find the best estimate for the initial condition $\mathbf{x}_0$, the following cost function is minimised with respect to $\mathbf{x}_0$:

$$J(\mathbf{x}_0) = (\mathbf{x}_0 - \mathbf{x}_b)^T B^{-1} (\mathbf{x}_0 - \mathbf{x}_b) + \sum_{l=0}^{L} \left[\mathbf{y}_l - \mathcal{H}_l(\mathbf{x}_l)\right]^T R_l^{-1} \left[\mathbf{y}_l - \mathcal{H}_l(\mathbf{x}_l)\right], \tag{1}$$

$$\mathbf{x}_{l+1} = \mathcal{M}_{l+1,l}(\mathbf{x}_l). \tag{2}$$

The solution will be termed the *analysis vector*, $\mathbf{x}_a \in \mathbb{R}^N$, following the convention of NWP literature, ie: $\nabla J(\mathbf{x}_a) = 0$. The period of time the observations are taken over is known as the *assimilation window*. Here the cost function uses $L + 1$ sets of observations and subscript l denotes the state of a variable at the time of the lth observation.

The variable $\mathbf{x}_b \in \mathbb{R}^N$ is the estimated initial condition termed the *background estimate*; $\mathcal{M}_{l+1,l} : \mathbb{R}^N \to \mathbb{R}^N$ is the forward model taking the lth state of the numerical model to its $(l + 1)$th state; $\mathbf{x}_l \in \mathbb{R}^N$ is the lth state of the numerical model; $\tilde{\mathbf{y}}_l \in \mathbb{R}^{m_l}$ is the lth perfect observation of the physical system; $\boldsymbol{\epsilon}_l \in \mathbb{R}^{m_l}$ is the observation error in the lth observation; $\mathbf{y}_l \in \mathbb{R}^{m_l}$ is the lth observation of the true physical system such that $\mathbf{y}_l = \tilde{\mathbf{y}}_l + \boldsymbol{\epsilon}_l$; $\mathcal{H}_l : \mathbb{R}^N \to \mathbb{R}^{m_l}$ is the lth observation operator; $B \in \mathbb{R}^{N \times N}$ is the background error covariance matrix; and $R_l \in \mathbb{R}^{m_l \times m_l}$ is the lth observation error covariance matrix. Here $N, m_l \in \mathbb{N} \ \ \forall l$. More details on these variables and 4D-Var can be found in [3, 4, 5, 6].

We choose the linear advection equation as the physical system, which can be solved using several numerical schemes. These introduce numerical model error through the approximation of derivatives [7]. In order to fully investigate the effects of numerical model error, all other errors present in the problem will be initially removed. Therefore, the background term of the cost function is neglected and observations will be taken at every spacial and temporal grid point of the numerical model. Hence $m_l = N$ and $\mathcal{H}_l = I_N \ \ \forall l$. Also, $\boldsymbol{\epsilon}_l$ is an iid Gaussian random variable, $\mathcal{N}(\mathbf{0}, \sigma_o^2 I_N), \ \ \sigma_o \in \mathbb{R}$, leading to $R_l = \sigma_o^2 I_N \ \ \forall l$. These assumptions result in the following cost function:

$$J(\mathbf{x}_0) = \frac{1}{\sigma_o^2} \sum_{l=0}^{L} \left[\mathbf{y}_l - \mathcal{M}_{l,0}(\mathbf{x}_0)\right]^T \left[\mathbf{y}_l - \mathcal{M}_{l,0}(\mathbf{x}_0)\right]. \tag{3}$$

Initially the problem will be investigated in the absence of observation errors. In this instance, σ_o^2 is set arbitrarily to one. Later observation errors will be reintroduced to the problem.

2.1. The physical system

Consider the linear advection equation, $u : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (d, t) \mapsto u(d, t)$, together with circulant boundary conditions and initial condition, $u_0 : \mathbb{R} \to \mathbb{R}$,

$$\begin{aligned} &u_t(d,t) + \eta u_d(d,t) = 0, && d \in [0,1), \ \ t > 0, \\ &u(d,t) = u(d+1,t), && d \in \mathbb{R}, \ \ t \geq 0, \\ &u(d,0) = u_0(d), && d \in [0,1). \end{aligned} \tag{4}$$

Here the *wave speed* $\eta \in \mathbb{R}$ remains constant. This problem is also considered in [8].

The solution to this problem, $u(d, t) = u_0(d - \eta t)$ [9], preserves the shape of the initial condition over time and propagates it through space with speed η. Problem (4) can be solved numerically using a finite difference scheme as the forward model. These find a numerical approximation to the analytic solution, introducing numerical model error, which will be considered in the form of *numerical dissipation* and *numerical dispersion* [10].

2.2. Numerical dissipation and dispersion

The initial condition $u_0(d)$, can be considered in the form of a Fourier series,

$$u_0(d) \approx \sum_{p=-\infty}^{\infty} c_p e^{2\pi i p d}, \quad \text{where } c_p = \int_0^1 u(d,0) e^{-2\pi i p d} \mathrm{d}d. \tag{5}$$

A finite difference scheme propagates these wavenumber components through time. Numerical dissipation occurs when the amplitude of at least one wavenumber component is not preserved over time [10]. This represents a loss of energy in the system. Numerical dispersion occurs when at least one wavenumber component is out of phase with the others [10]. This is due to the out of phase wavenumber component travelling with an incorrect wave speed.

This paper will consider three finite difference schemes which solve problem (4) numerically across $N+1$ mesh points in $[0,1]$, $N \in \mathbb{N}$, and are implemented by a matrix $M \in \mathbb{R}^{N\times N}$. Let $U_j^n \approx u(d_j, t^n)$, where $t^n = n\Delta t$, $n \in \mathbb{N}_0$ and $d_j = j\Delta d$, $j = 0, \ldots, N+1$, $\Delta d = 1/N$ and $\Delta t > 0$, the spatial and temporal step sizes respectively. Also let $h = \eta \Delta t/\Delta d \leq 1$, so that the finite difference scheme is numerically stable. Define $\mathbf{U}^n \in \mathbb{R}^N$ with jth entry U_{j-1}^n, $\quad j = 1, \ldots, N$. Applying the matrix M to $\mathbf{U}^n$, advances the numerical solution at $d_0, \ldots, d_{N-1}$, forward Δt in time, ie: $\mathbf{U}^{n+1} = M\mathbf{U}^n \quad \forall n$. The circulant boundary conditions mean that M is a circulant matrix [11] and $u(d_N, t^n) = u(d_0, t^n) \ \ \forall n$, hence $U_N^n = U_0^n \ \ n \in \mathbb{N}_0$.

As $\mathbf{U}^n$ is an N-dimensional vector, it can be constructed from the N vectors of the *Discrete Fourier Transform* (DFT) basis [12] $\{\mathbf{v}_p\}_{p=1}^N$, such that,

$$[\mathbf{v}_p]_q = \frac{1}{\sqrt{N}} e^{\frac{2\pi i (p-1)(q-1)}{N}} = \frac{1}{\sqrt{N}} e^{2\pi i (p-1) d_{q-1}}, \quad p, q = 1, \ldots, N, \tag{6}$$

is the qth element of the pth vector. This is the $(p-1)$th wavenumber component of the Fourier series, sampled at d_{q-1}, with amplitude $1/\sqrt{N}$. The numerical solution is constructed from N wavenumber components of the Fourier series, that are resolvable on the finite grid, represented by the basis vectors. The remaining wavenumber components of the Fourier series are indistinguishable on the discrete mesh, leading to *aliasing* [13]. As a result, the coefficients of the N resolvable wavenumber components are determined by the *Poisson summation*. There is no spectral leakage in this problem due to 1-periodicity, see [12]. This allows M to propagate all the wavenumber components of the Fourier series, by only directly acting on N of them.

Aliasing results in M applying the same phase shift to an unresolvable wavenumber component, as it applies to the resolvable wavenumber component it aliases to. This is not necessarily the correct phase shift for the considered unresolvable wavenumber component. This form of numerical dispersion is known as *sample error* [12]. M can also introduce numerical dissipation and dispersion into the resolvable wavenumber components of the solution.

2.3. Forward models for the linear advection equation

The Upwind, Preissman Box and Lax-Wendroff schemes are three finite difference schemes used to solve problem (4).

- Upwind scheme (explicit scheme) [9],

$$U_j^{n+1} = hU_{j-1}^n + (1-h)U_j^n. \tag{7}$$

- Preissman Box scheme (implicit scheme) [14],

$$(1-h)U_j^{n+1} + (1+h)U_{j+1}^{n+1} = (1+h)U_j^n + (1-h)U_{j+1}^n. \tag{8}$$

- Lax-Wendroff scheme (explicit scheme) [9],

$$U_j^{n+1} = \frac{h}{2}(h+1)U_{j-1}^n + (1-h^2)U_j^n + \frac{h}{2}(h-1)U_{j+1}^n. \tag{9}$$

Selecting $h = 0.5$ allows the effects of numerical dissipation and numerical dispersion on the resolvable wavenumber components to be investigated as individual processes and in combination. For simplicity, $\eta = 1$ is chosen in the following.

Let M denote the matrix used to implement either the Upwind, Preissman Box or Lax-Wendroff scheme as in section 2.2. As M is a circulant matrix, its eigenvalue decomposition can be constructed using the unitary DFT matrix, denoted $V \in \mathbb{C}^{N\times N}$ [11],

$$M = V\Lambda V^{-1} = V\Lambda V^{*}, \tag{10}$$

where $\cdot^{*}$, denotes Hermitian. The columns of V are the DFT basis vectors in (6), ie: $V_{p,q} = \left[\mathbf{v}_p\right]_q, \quad p, q = 1, \ldots, N$. The corresponding eigenvalues are found in $\Lambda = \text{diag}(\lambda_p) \in \mathbb{C}^{N\times N}$. The eigenvalues of M are scheme dependent, whereas the eigenvectors are scheme independent. In the case of a linear system, the eigenvalues affect the propagation of the initial state of the system through time [7] and introduce any numerical dissipation and dispersion into the state of the system (this is not physical dissipation or dispersion).

The eigenvalues $\lambda_p = |\lambda_p|e^{i\theta_p}, \quad \theta_p \in (-2\pi, 0], \quad p = 1, \ldots, N$, are such that the magnitude $|\lambda_p|$ and phase θ_p, affect the magnitude and phase of the corresponding wavenumber component of the DFT basis, respectively. The basis wavenumber components are complex; given an eigenpair of M, its complex conjugate is also an eigenpair of M [15]. Summing the conjugate pair of eigenvectors, scaled by their respective complex coefficients for the state of the system, results in a real wavenumber component for the linear system. We consider these wavenumber components in sections 3.1-3.4 of this paper.

2.4. Generating perfect observations

Consider the following finite difference scheme, the *Numerical Implementation of the Method of Characteristics* (NIMC):

$$U_j^{n+1} = hU_{j-1}^{n}.$$

This is an explicit finite difference scheme and can be implemented as described in section 2.2, via a circulant matrix, $M_{NIMC} \in \mathbb{R}^{N\times N}$. This scheme is always non-dispersive and numerically stable when $h \leq 1$, but is only non-dissipative when $h = 1$. The scheme can be used to generate perfect observations for the physical system in this instance. This leads to M_{NIMC} generating a perfect observation every $\Delta t = \Delta d/\eta$ in time. However, the imperfect schemes in this paper progress the forward model $\Delta t = h\Delta d/\eta$ in time with each application of M, where h is not necessarily equal to one. This means that given the same Δt, Δd and η, the NIMC cannot provide perfect observations at every time step of the imperfect forward model.

Hence another finite difference scheme is used in the form of the matrix $\tilde{M} = V\tilde{\Lambda}V^{*}$, where the eigenvalues in $\tilde{\Lambda}$ are given by $\tilde{\lambda}_p = e^{i\tilde{\theta}_p}$ such that, for only odd N,

$$\tilde{\theta}_p = \begin{cases} \frac{-2\pi(p-1)h}{N}, & \text{for } p \leq \frac{N+1}{2}, \\ 2\pi\left[(h-1) - \frac{(p-1)h}{N}\right], & \text{for } p > \frac{N+1}{2}. \end{cases} \tag{11}$$

This scheme is numerically non-dissipative and non-dispersive with respect to the resolvable wavenumber components for any $h \leq 1$, producing the state of the system every $\Delta t = h\Delta d/\eta$ in time. This will be known as the *Modified NIMC* (MNIMC). If the current time of the solution is divisible by $\Delta d/\eta$, then the solution of the system is exact. However, in between these times, the state of the system at each spatial mesh point is interpolated in time, ie: sample error. An additive correction term $\mathbf{r}_l \in \mathbb{R}^N$ can be used to correct the sample error in $\tilde{M}^l\mathbf{U}_0$,

$$\tilde{\mathbf{y}}_l = \tilde{M}^l\mathbf{U}_0 + \mathbf{r}_l, \tag{12}$$

where $\mathbf{y}_l = \tilde{\mathbf{y}}_l$ denotes a perfect observation. Choosing $h = 1$ results in $\tilde{M} = M_{NIMC}$ and consequently $\mathbf{r}_l = \mathbf{0} \ \ \forall l$.

3. The effect of numerical dissipation and dispersion on the analysis vector

Numerical dissipation and dispersion are introduced into the inverse problem through the forward model. This section explores how these effects impact the analysis vector. This is achieved by formulating the analysis vector in

terms of the true initial condition, allowing the direct impact of numerically dissipative and/or dispersive eigenvalues of the imperfect scheme, to be seen. Under the assumptions of section 2, $\mathcal{M}_{l+1,l} := M$, $\mathbf{x}_l \equiv \mathbf{U}^l \quad \forall l$, and the cost function becomes,

$$J(\mathbf{x}_0) = \sum_{l=0}^{L} [\mathbf{y}_l - M^l \mathbf{x}_0]^T [\mathbf{y}_l - M^l \mathbf{x}_0]. \tag{13}$$

Let $\mathcal{F} : \mathbb{R}^N \to \mathbb{C}^N$, $\mathbf{x} \to \mathcal{F}(\mathbf{x}) = V^* \mathbf{x}$, be the DFT and the analysis vector, $\mathbf{x}_a \in \mathbb{R}^N$, be the solution to the inverse problem, ie: $\nabla J(\mathbf{x}_a) = 0$. Then,

$$\mathbf{x}_a = \left[\sum_{k=0}^{L} (M^T M)^k\right]^{-1} \sum_{l=0}^{L} (M^T)^l \mathbf{y}_l = V \left[\sum_{k=0}^{L} (\Lambda^* \Lambda)^k\right]^{-1} \sum_{l=0}^{L} (\Lambda^*)^l V^* \mathbf{y}_l, \tag{14}$$

using (10). This can be re-written using the DFT,

$$\mathcal{F}(\mathbf{x}_a) = \left[\sum_{k=0}^{L} (\Lambda^* \Lambda)^k\right]^{-1} \left[\sum_{l=0}^{L} (\Lambda^*)^l \mathcal{F}(\mathbf{y}_l)\right] = \left[I_N + \sum_{k=1}^{L} (\Lambda^* \Lambda)^k\right]^{-1} \left[\sum_{l=0}^{L} (\Lambda^*)^l \mathcal{F}(\mathbf{y}_l)\right]. \tag{15}$$

Here the diagonal matrices Λ and Λ^* are known as the *forward* and *adjoint models* [4] respectively, in the Fourier basis. In the inverse problem, each set of observations is mapped back in time to $t = 0$, by the adjoint model, M^T. Once the observations have been mapped to the initial time, they are then summed. This process has the potential to create interference between the corresponding wavenumber components of each observation. The result is then normalised with respect to the eigenvalues of the scheme. The observation at $t = 0$ acts to regularise the solution of the inverse problem so that the matrix applying the normalisation is always invertible.

Expression (15) forms the coefficients of the complex wavenumber components in the construction of the analysis vector $\mathbf{x}_a$, ie: $\mathbf{x}_a = V\mathcal{F}(\mathbf{x}_a)$. Initially, consider perfect observations of the physical system ie: $\mathbf{y}_l = \tilde{\mathbf{y}}_l \quad \forall l$, given by (12). Let $h = q/a$, $q, a \in \mathbb{Z}$ such that $\gcd(q, a) = 1$. Then the analysis vector can be expressed as,

$$\mathbf{x}_a = A_L \tilde{\mathbf{x}}_0 + \boldsymbol{\rho}_L, \tag{16}$$

where the *model resolution matrix* $A_L \in \mathbb{R}^{N \times N}$ is such that,

$$A_L = V \left[\sum_{k=0}^{L} (\Lambda^* \Lambda)^k\right]^{-1} \left[\sum_{l=0}^{L} (\Lambda^* \tilde{\Lambda})^l\right] V^*, \tag{17}$$

and $\boldsymbol{\rho}_L \in \mathbb{R}^N$ is,

$$\boldsymbol{\rho}_L = V \left[\sum_{k=0}^{L} (\Lambda^* \Lambda)^k\right]^{-1} \left[\left\{\sum_{l=0}^{\frac{L-[L]_a}{a}-1} (\Lambda^* \tilde{\Lambda})^{la}\right\} \left\{\sum_{y=1}^{a-1} (\Lambda^*)^y V^* \mathbf{r}_y\right\} + \left(\Lambda^* \tilde{\Lambda}\right)^{L-[L]_a} \left\{\sum_{y=1}^{[L]_a} (\Lambda^*)^y V^* \mathbf{r}_y\right\}\right], \tag{18}$$

where $[\cdot]_a$ denotes modulo a. See [16] for details.

Expression (16) can be viewed as the sum of two analysis vectors created when solving the same problem but with two different sets of observations; $\mathbf{y}_l = \tilde{\mathbf{x}}_l$ and $\mathbf{y}_l = \mathbf{r}_l$. As a result the sample error in $\mathbf{y}_l = \tilde{\mathbf{y}}_l$ does not play a part in A_L, and is solely found in $\boldsymbol{\rho}_L$. Consequently, $\boldsymbol{\rho}_L$ acts as a correction term for sample errors introduced into the analysis vector by the MNIMC scheme.

The eigenvalues of A_L in (17) determine the magnitude and phase change applied to each wavenumber component of $\tilde{\mathbf{x}}_0$, in the construction of $\mathbf{x}_a$. In this way, they can be described as *amplification factors* for the wavenumber components of $\tilde{\mathbf{x}}_0$. Let ν_p be an eigenvalue of A_L. Due to the diagonal structures of Λ and $\tilde{\Lambda}$, ν_p is constructed solely from λ_p and $\tilde{\lambda}_p$, the pth eigenvalues of M and $\tilde{M}$, respectively,

$$\nu_p = \frac{\sum_{l=0}^{L} \overline{\lambda}_p^l \tilde{\lambda}_p^l}{\sum_{k=0}^{L} |\lambda_p|^{2k}}. \tag{19}$$

The only possible source of numerical model error in ν_p, is λ_p. Let $\phi_p = \tilde{\theta}_p - \theta_p$, be the error in the phase shift applied by λ_p to the resolvable wavenumber components. Then,

$$\nu_p = \begin{cases} 1, & \text{for } |\lambda_p| = 1 \text{ and } \phi_p = 2\pi s, \quad s \in \mathbb{Z}, \\ \left(\frac{1+|\lambda_p|}{1+|\lambda_p|^{L+1}}\right), & \text{for } |\lambda_p| < 1 \text{ and } \phi_p = 2\pi s, \quad s \in \mathbb{Z}, \\ \frac{1}{L+1}\left|\frac{\sin\left[(L+1)\frac{\phi_p}{2}\right]}{\sin\left[\frac{\phi_p}{2}\right]}\right| e^{i\mu_p}, & \text{for } |\lambda_p| = 1 \text{ and } \phi_p \neq 2\pi s, \quad s \in \mathbb{Z}, \\ \left(\frac{[1-|\lambda_p|^{L+1}e^{i(L+1)\phi_p}][1-|\lambda_p|^2][1-|\lambda_p|e^{-i\phi_p}]}{[1-|\lambda_p|^{2(L+1)}][1+|\lambda_p|^2-2|\lambda_p|\cos(\phi_p)]}\right), & \text{for } |\lambda_p| < 1 \text{ and } \phi_p \neq 2\pi s, \quad s \in \mathbb{Z}, \end{cases} \tag{20}$$

by the sum of a geometric progression, where,

$$\tan(\mu_p) = \tan\left(\frac{L\phi_p}{2}\right), \quad \mu_p \in (-2\pi, 0].$$

When λ_p does not introduce numerical model error into the corresponding resolvable wavenumber component, $\nu_p = 1$, so the wavenumber component is preserved in $\mathbf{x}_a$. A numerically dissipative λ_p creates an amplification factor that only affects the amplitude of the corresponding resolvable wavenumber component.

In the case of a numerically dispersive eigenvalue of M, the amplification factor affects both the phase and amplitude of the corresponding resolvable wavenumber component. The affect on the magnitude is due to interference between the resolvable wavenumber components constructing $\mathbf{x}_a$, as discussed in section 3.

A numerically dissipative and dispersive eigenvalue of M, creates an amplification factor that is a combination of the numerically dissipative and dispersive amplification factors. However, it is not possible to isolate the dissipative and dispersive effects from one another. The magnitude and phase of the spectra of the model resolution matrix for each scheme are plotted in Figures 1, 3 and 5.

The contribution of $\boldsymbol{\rho}_L$ to the analysis vector is not as easy to analyse. The contribution of $\boldsymbol{\rho}_L$ to $\mathbf{x}_a$ can be reduced by choosing an $\tilde{\mathbf{x}}_0$ that is minimally constructed from unresolvable wavenumber components. As a result, a higher regularity initial condition will reduce $\boldsymbol{\rho}_L$. Here *regularity* is defined as the maximum number of times a function is continuously differentiable. Choosing $h = 1$ leads to $\boldsymbol{\rho}_L = \mathbf{0}$.

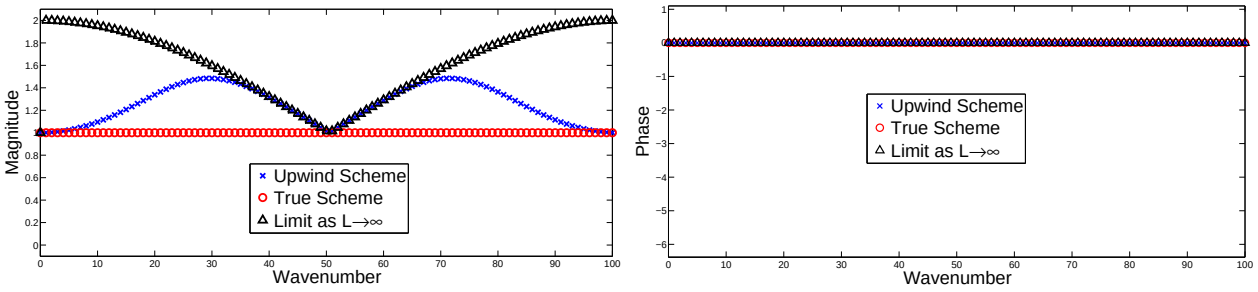
In order to understand the effects of the model resolution matrix on $\tilde{\mathbf{x}}_0$ and the contribution of $\boldsymbol{\rho}_L$ to the analysis vector, a low regularity initial condition in the form of a square wave is chosen. A square wave has regularity zero, requiring many high wavenumber components to resolve the edges of the wave. $\tilde{\mathbf{x}}_0$ is then a discrete sample of the square wave.

In sections 3.1-3.3, the magnitude and phase of the spectra of A_L are analysed for the three schemes, together with the result of applying A_L to $\tilde{\mathbf{x}}_0$, for the square wave when using $L = 4$. The corresponding $\boldsymbol{\rho}_L$ and $\mathbf{x}_a$ for the square wave are also shown for $L = 4$. The reader is reminded that A_L acts upon all wavenumber components of $\tilde{\mathbf{x}}_0$ through the effects of aliasing.

3.1. The Upwind scheme

The Upwind scheme is a numerically dissipative and non-dispersive scheme with respect to the resolvable wavenumber components, for $h = 0.5$. As the Upwind scheme is numerically non-dispersive with respect to the resolvable wavenumber components, the only dispersive effects applied to $\tilde{\mathbf{x}}_0$ by A_L are those due to sample error in the MNIMC scheme. The resolvable wavenumber components all remain in phase meaning there is no constructive or destructive interference effects. All but the lowest resolvable wavenumber component (ie: $p = 1$) are amplified by A_L. The greatest amplification effects are experienced by the medium resolvable wavenumber components, as can be seen in Figure 1(a). As L increases, the amplification of the lower resolvable wavenumber components increases. We remind the reader here the effects of A_L on the real wavenumber components of $\tilde{\mathbf{x}}_0$, can be seen in the first $(N + 1)/2$ (N odd) values of p in Figure 1, where increasing p represents increasing wavenumber.

The plots of $A_L\tilde{\mathbf{x}}_0$ and $\boldsymbol{\rho}_L$ in Figure (2) demonstrate oscillations at the locations of the corners of the square wave. The form of the oscillations is dictated by the numerically dissipative effects of the finite difference scheme. The oscillations in $A_L\tilde{\mathbf{x}}_0$ are created by the amplification matrix and contain the effects of sample error in the MNIMC



(a) Magnitude of the amplification factors.

(b) Phase of the amplification factors.

Figure 1: The magnitude and phase of the spectrum of the model resolution matrix, A_L for $L = 4$, together with their limit as $L \to \infty$, for the Upwind scheme when $N = 101$ and $h = 0.5$. The spectrum of the MNIMC scheme is included for comparison.

scheme. $\boldsymbol{\rho}_L$ possesses the greatest amplitude oscillations at the corners of the square wave, where the unresolvable wavenumber components have the greatest impact on the shape of the square wave. Adding $\boldsymbol{\rho}_L$ to $A_L\tilde{\mathbf{x}}_0$ results in an analysis vector which is visibly improved in comparison to $A_L\tilde{\mathbf{x}}_0$. This indicates how important accounting for the effects of sample error can be. The error in $\mathbf{x}_a$ is solely due to numerical model error in the forward model. Similar results follow for the remaining schemes in Figures (4) and (6).

3.2. The Preissman Box scheme

The Preissman Box scheme is a numerically non-dissipative and dispersive scheme with respect to the resolvable wavenumber components, for $h = 0.5$. As a result, A_L only introduces dispersive effects into $\tilde{\mathbf{x}}_0$, through the dispersion of the resolvable wavenumber components and the effects of sample error in the MNIMC scheme. The attenuation effects experienced by the medium to high resolvable wavenumber components of $\tilde{\mathbf{x}}_0$ in Figure 3(a) are solely due to destructive interference. The amplitude of the lowest resolvable wavenumber component is the only one not affected. In this instance, the low to medium resolvable wavenumber components experience a small attenuation effect, whilst the medium to high resolvable wavenumber components experience a much larger attenuation. The highest resolvable wavenumber components are almost attenuated to zero.

3.3. The Lax-Wendroff scheme

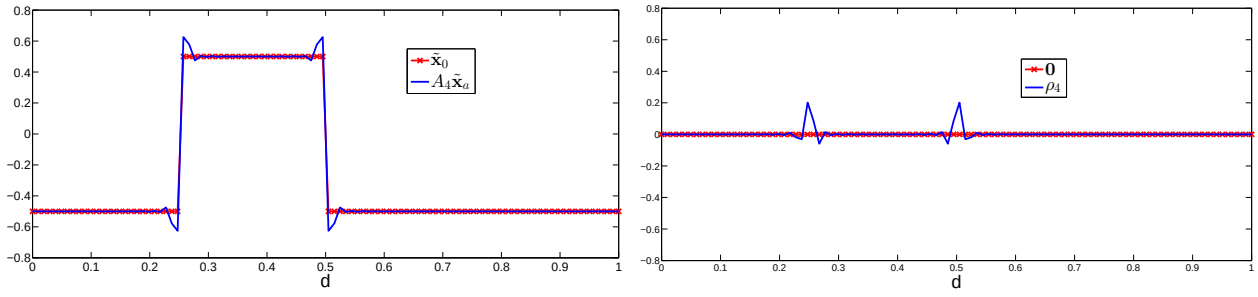
The Lax-Wendroff scheme is both numerically dissipative and dispersive with respect to the resolvable wavenumber components, for $h = 0.5$. As a result, $\tilde{\mathbf{x}}_0$ experiences a combination of the dissipative and dispersive effects seen in the application of A_L for the Upwind and Preissman Box schemes. This combination is seen in the amplification and attenuation effects on the resolvable wavenumber components, which appear to be a balance between the amplification effects seen for the Upwind scheme and the attenuation effects of the Preissman Box scheme.

3.4. The length of the assimilation window

Another factor that affects the behaviour of numerical model error is the length of the assimilation window. It is important to understand whether the extra time and processing power required to include more observations will yield an improvement in the solution.

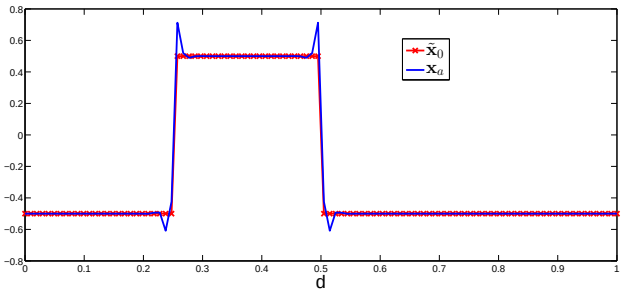
Figures 1(a) and 5(a) show that the amplification factors corresponding to the smallest magnitude eigenvalues of the respective schemes, approach their limit for $L \to \infty$, when $L = 4$. In the cases of the Upwind and Lax-Wendroff schemes, these eigenvalues correspond to the highest wavenumber components. Hence, increasing the length of the assimilation window will not affect the contribution of the high wavenumber components to the analysis vector and its forecast. The amplification factor of the remaining wavenumber components can be altered by increasing the length of the assimilation window.

In the case of a numerically non-dissipative and dispersive eigenvalue, λ_p, such as those found in the Preissman Box scheme, $\nu_p \to 0$ as $L \to \infty$. The result can be seen in Figure 3(a) where increasing the length of the assimilation window, by adding extra observations, decreases the contribution of the corresponding wavenumber components from $\tilde{\mathbf{x}}_0$ to $\mathbf{x}_a$. As more observations are included, destructive interference increases, leading to a loss of information in $\mathbf{x}_a$



(a) $A_L\tilde{\mathbf{x}}_0$

(b) $\boldsymbol{\rho}_L$



(c) $\mathbf{x}_a = A_L\tilde{\mathbf{x}}_0 + \boldsymbol{\rho}_L$

Figure 2: The analysis vector, $\mathbf{x}_a = A_L\tilde{\mathbf{x}}_0 + \boldsymbol{\rho}_L$, for a square wave initial condition, when using the Upwind scheme and perfect observations, $\mathbf{y}_l = \tilde{\mathbf{y}}_l = \tilde{\mathbf{x}}_l + \mathbf{r}_l$, for $N = 101$ and $L = 4$.

and its subsequent forecast. Hence for a solely numerically dispersive scheme, increasing the number of observations does not necessarily improve the accuracy of the analysis vector and its forecast.

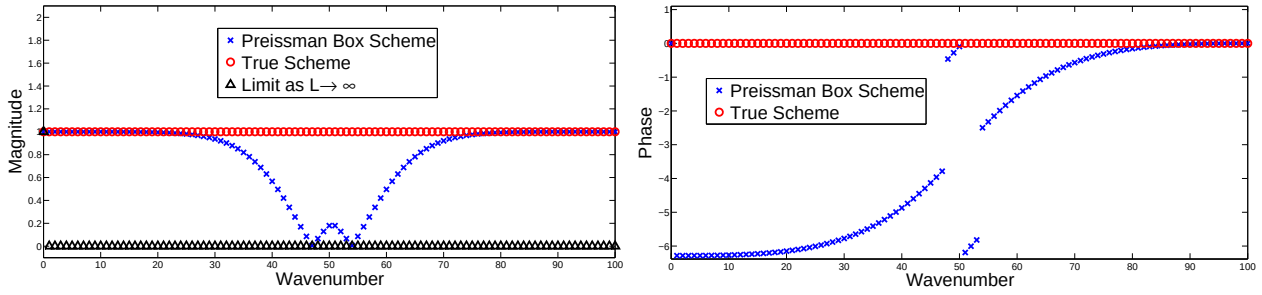
4. Error analysis

Numerical model error can be measured through the direct error on the analysis vector or by its effect on the subsequent forecast. Both quantities are important in different applications of the inverse problem. These errors can be shown to converge to zero for sufficiently smooth initial conditions, when measured in the l_2-norm, by considering the global and truncation errors associated with the forward model of each scheme [16]. To investigate these errors for any regularity initial condition, a spectral approach is taken, using the formulation for the analysis vector found in section 3.

4.1. Spectral approach in the absence of observation errors

A spectral approach can be used to provide an upper bound for the l_2-norm of the error in the analysis vector for any regularity initial conditions. Here we remind the reader that regularity is defined as the maximum number of times a function is continuously differentiable.

Lemma 1. *Let the assumptions detailed so far in this paper allowing* $\mathbf{x}_a$ *to be stated as in (16), hold true. Additionally, let* $u_0(d)$ *have regularity* $r \in \mathbb{N}_0$ *over* $(0, 1)$ *such that* $u_0^{(s)}(d)$ *is bounded and piecewise monotone over* $(0, 1)$ *and*



(a) Magnitude of the amplification factors

(b) Phase of the amplification factors

Figure 3: The magnitude and phase of the spectrum of the model resolution matrix, A_L for $L = 4$, together with their limit as $L \to \infty$, for the Preissman Box scheme when $N = 101$ and $h = 0.5$. The spectrum of the MNIMC scheme is included for comparison.

$\lim_{d\to 0^+} u_0^{(s)}(d) = \lim_{d\to 1^-} u_0^{(s)}(d)$, *for all* $s = 0, \dots r$. *Then,*

$$\begin{aligned} \|\tilde{\mathbf{x}}_0 - \mathbf{x}_a\|_2^2 \quad \leq \quad & N\left\{|1-\nu_1|D_1 + (|1-\nu_1| - 2\xi_1)\frac{D_3}{N^{r+1}}\right\}^2 \\ & + N\sum_{p=2}^{\frac{N+1}{2}}\left\{|1-\nu_p|\frac{D_2}{(p-1)^{r+1}} + (|1-\nu_p| - 2\xi_p)\frac{D_3}{N^{r+1}}\right\}^2 \\ & + N\sum_{p=\frac{N+3}{2}}^{N}\left\{|1-\nu_p| - 2\xi_p\right\}^2\left\{\frac{D_2}{(p-1)^{r+1}} + \frac{D_3}{N^{r+1}}\right\}^2, \end{aligned} \tag{21}$$

where D_1, D_2 *and* D_3 *are constants independent of* p *and* N *and*

$$\xi_p = \frac{\left|\sum_{l=0}^{\frac{L-[L]_a}{a}-1}\left[|\lambda_p|^a e^{ia\phi_p}\right]^l\right|\left\{\sum_{y=1}^{a-1}|\lambda_p|^y\right\} + |\lambda_p|^{L-[L]_a}\sum_{y=1}^{[L]_a}|\lambda_p|^y}{\sum_{k=0}^{L}|\lambda_p|^{2k}}. \tag{22}$$

Proof. We refer to [16] for the proof where use has been made of [17] and [12]. □

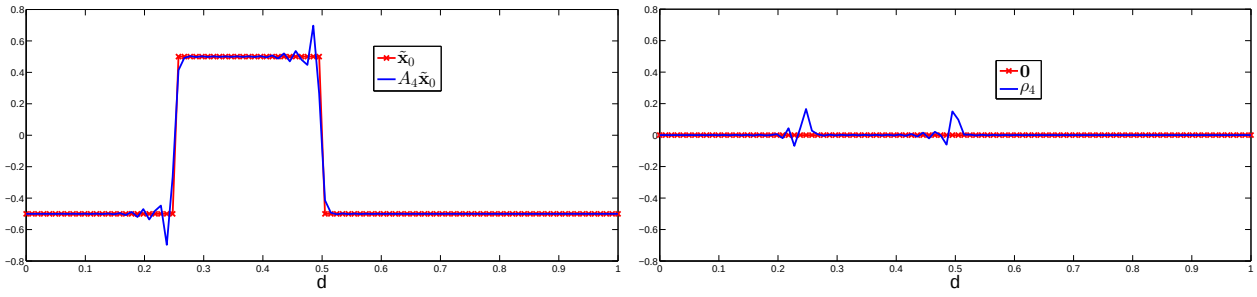
The upper bound can be used to analyse the order of convergence of the error to zero, with respect to either N, the number of discretisation points or L, the number of observations. Examining (21) indicates that the order of convergence of the upper bound will be explicitly dependent on the regularity of the initial condition, given by r. It should be noted here that the order of convergence with respect to either N or L is found for constant h, resulting in the length of the assimilation window varying in time, $T = L\Delta t = Lh/(\eta N)$, as N and L are varied.

In the case of the Upwind, Preissman Box and Lax-Wendroff schemes, $\nu_1 = 1$. Hence the terms relating to ν_1 in the upper bound of (21), are zero. In the case of the MNIMC scheme, $A_L = I$, so the only contribution to the error is from sample error. The upper bound is consistent with this as only the upper bound on the sample error remains. Choosing $h = 1$ results in both the error and its upper bound becoming zero.

The order of convergence of the upper bound with respect to either N or L, is in part determined by the order of convergence of the $|1-\nu_p|$ and ξ_p terms. $|1-\nu_p|$ has a direct impact on the analysis vector, whilst ξ_p is a consequence of applying an upper bound to the error in the analysis vector. $|1-\nu_p|$ and ξ_p are both dependent on N and L; N determines the number of points, whilst L determines the shape, of the plots in Figure 7.

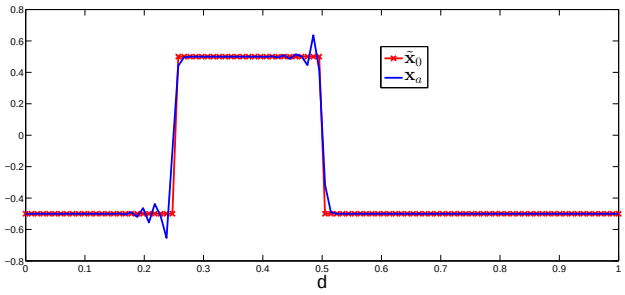
The number of mesh points and observations used in NWP are typically $O(10^7)$ [18] and $O(10^5 - 10^6)$ [6] respectively. As a result, it is realistic to consider the order of convergence of the upper bound in (21) when L is small in comparison to N (ie: a small assimilation window).

The order of convergence of $|1-\nu_p|$ to zero with respect to N, for fixed p, was found numerically using fixed $L = 4$. This was found to be less than or equal to zero for all p and remained constant for small p ie: $(p-1)/N \ll 1$, where $|1-\nu_p| = O(N^{-2})$ for the Upwind scheme and $|1-\nu_p| = O(N^{-3})$ for the Preissman Box and Lax-Wendroff schemes, for such p, $p \neq 1$. As a result, $|1-\nu_p|$ is either decaying to zero or remaining constant as N increases.



(a) $A_L\tilde{\mathbf{x}}_0$

(b) $\boldsymbol{\rho}_L$



(c) $\mathbf{x}_a = A_L\tilde{\mathbf{x}}_0 + \boldsymbol{\rho}_L$

Figure 4: The analysis vector, $\mathbf{x}_a = A_L\tilde{\mathbf{x}}_0 + \boldsymbol{\rho}_L$, for a square wave initial condition, when using the Preissman Box scheme and perfect observations, $\mathbf{y}_l = \tilde{\mathbf{y}}_l = \tilde{\mathbf{x}}_l + \mathbf{r}_l$, for $N = 101$ and $L = 4$.

Corollary 2. *It is important to note that in the case of $r = \infty$, it is appropriate to consider the limit of the upper bound in (21), which retains two terms (assuming $N > 1$),*

$$\|\tilde{\mathbf{x}}_0 - \mathbf{x}_a\|_2^2 \leq N|1-\nu_1|^2 D_1^2 + N\,|1-\nu_2|^2\, D_2^2. \tag{23}$$

These components originate from the upper bound on $A_L\tilde{\mathbf{x}}_0$, as the sample error is negligible for an initial condition where $r = \infty$. As a result, the limit of the upper bound is zero when $r = \infty$ and $M = \tilde{M}$. The numerical results for the three considered schemes indicate that for an initial condition such that $r = \infty$,

$$\|\tilde{\mathbf{x}}_0 - \mathbf{x}_a\|_2^2 = O(N^{1+2\gamma}), \tag{24}$$

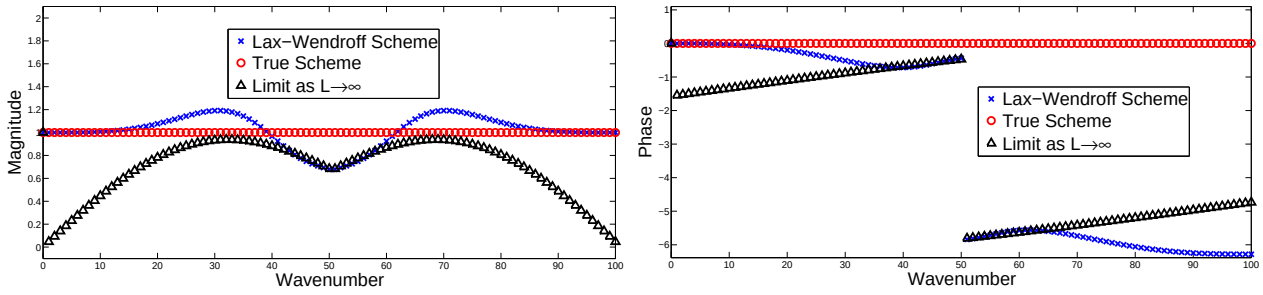
where γ is the numerical order of convergence for $|1-\nu_p|$ to zero with respect to N, for small p, $p \neq 1$ ie: for the Upwind $\gamma = -2$.

Similarly, the order of convergence of $|1-\nu_p|$ to zero with respect to L, for fixed p, was found numerically using fixed $N = 3^7$. This was found to be positive for all p and at most $|1-\nu_p| = O(L)$ for all three schemes, for each p, $p \neq 1$. This results show that increasing the value of L causes the value of $|1-\nu_p|$ to diverge from zero for $p \neq 1$. The identical order of convergence with respect to L for each scheme ($p \neq 1$) is not surprising, as the dependence of $|1-\nu_p|$ on L is similar for each scheme, unlike the dependence on N.

For the Upwind scheme, the numerical order of convergence for $|1-\nu_p|$ to zero, with respect to both N and L, can be explained through its asymptotic expansion as $N \to \infty$, for fixed p,

$$|1-\nu_p| \sim \frac{\pi^2 L}{4}\left(\frac{p-1}{N}\right)^2, \tag{25}$$

and is valid for $1 \leq p \leq N$. This expansion indicates that $|1-\nu_p|$ has orders of convergence $O(N^{-2})$ and $O(L)$ for the Upwind scheme. These match the numerical orders of convergence found for $|1-\nu_p|$ to zero with respect to N when p is small ($p \neq 1$) and the maximum order of convergence with respect to L, for the Upwind scheme.



(a) Magnitude of the amplification factors

(b) Phase of the amplification factors

Figure 5: The magnitude and phase of the spectrum of the model resolution matrix, A_L for $L = 4$, together with their limit as $L \to \infty$, for the Lax Wendroff scheme when $N = 101$ and $h = 0.5$. The spectrum of the MNIMC scheme is included for comparison.

r	α			β		
	Upwind	Preissman Box	Lax-Wendroff	Upwind	Preissman Box	Lax-Wendroff
0	-6.7708×10^{-15}	-2.6329×10^{-10}	-3.7260×10^{-9}	5.7945×10^{-1}	3.6188×10^{-1}	4.0297×10^{-1}
1	−2.0000	−2.0000	−2.0000	1.5053	9.394×10^{-1}	1.0230
2	−3.0000	−4.0000	−4.0232	1.9957	1.6588	1.6731
3	−3.0000	−4.9565	−5.0600	2.0000	1.9990	1.9955
4	−3.0000	−4.9377	−5.0000	2.0000	2.0000	1.9999
5	−3.0000	−4.9345	−5.0000	2.0000	2.0000	2.0000
6	−3.0000	−4.9338	−5.0000	2.0000	2.0000	2.0000
7	−3.0000	−4.9336	−5.0000	2.0000	2.0000	2.0000
∞	−3.0000	−5.0000	−5.0000	2.0000	2.0000	2.0000

Table 1: Numerical orders of convergence to zero, with respect to N and L, for (26), $O(N^\alpha L^\beta)$, using the Upwind, Preissman Box and Lax-Wendroff schemes, given to 5sf. The results for N were identified using fixed $L = 4$ and larger values of N than those used to produce the results for α in Table 2. As a result, the values of α displayed here have a greater accuracy than those displayed in Table 2. The results for L were identified using fixed $N = 3^7$. The results for $r = \infty$ were identified using (23).

As $|1 - \nu_1| = 0$ for the Upwind, Preissman Box and Lax-Wendroff schemes, the upper bound in (21) can be considered as the sum of ten distinct summations. Each summation has an order of convergence to zero with respect to N and L, which influences the overall order of convergence for the upper bound. The order of convergence of each summation was identified numerically, in order to identify the dominant summation in the upper bound.

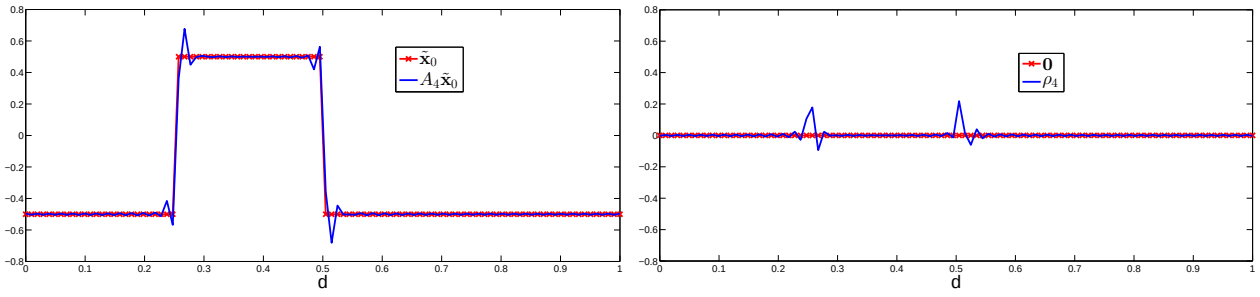
The summation with the dominant order of convergence for each considered scheme and regularity initial condition, with respect to both N and L, was found to be,

$$N \sum_{p=2}^{N} \frac{|1 - \nu_p|^2}{(p-1)^{2(r+1)}}. \tag{26}$$

This summation is composed from the amplification factors and the upper bound on the continuous Fourier coefficients of $u_0(d)$ [16, 17, 12], giving it an explicit dependence on the regularity of $u_0(d)$. The orders of convergence to zero for this summation, with respect to both N and L, are given in Table 1.

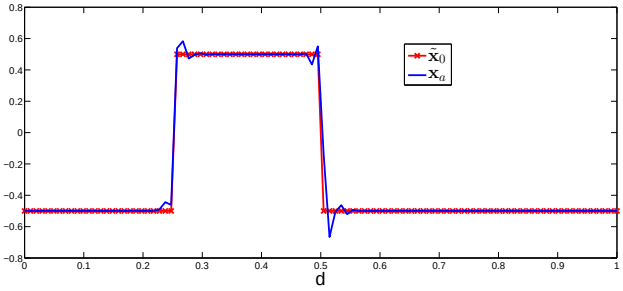
Table 1 shows that the order of convergence of (26) to zero with respect to N, for an initial condition such that $r = 0$, is $O(N^0)$. This indicates that the error in the analysis vector does not decay as N is increased. This is due to the error that always exists when a Fourier series is used to approximate a discontinuous function.

Expression (26) decays to zero for initial conditions with larger values for r, as N increases. As the regularity is increased, the order of convergence to zero with respect to N is initially $O(N^{-2r})$. However, once a critical regularity is achieved the order of convergence saturates; $O(N^{-3})$ for the Upwind scheme when $r \geq 2$ and $O(N^{-5})$ for the Preissman Box and Lax-Wendroff schemes when $r \geq 3$. The orders of convergence at saturation point match the orders of convergence for the upper bound, when considering its limit for an initial condition with $r = \infty$, found in (24).



(a) $A_L\tilde{\mathbf{x}}_0$

(b) $\boldsymbol{\rho}_L$



(c) $\mathbf{x}_a = A_L\tilde{\mathbf{x}}_0 + \boldsymbol{\rho}_L$

Figure 6: The analysis vector, $\mathbf{x}_a = A_L\tilde{\mathbf{x}}_0 + \boldsymbol{\rho}_L$, for a square wave initial condition, when using the Lax-Wendroff scheme and perfect observations, $\mathbf{y}_l = \tilde{\mathbf{y}}_l = \tilde{\mathbf{x}}_l + \mathbf{r}_l$, for $N = 101$ and $L = 4$.

The numerical results for N, for the Upwind scheme in Table 1, are seen when (25) is substituted into (26),

$$N\sum_{p=2}^{N}\frac{|1-\nu_p|^2}{(p-1)^{2(r+1)}} = \begin{cases} O(N^{-2r}), & \text{for } r = 0, 1, \\ O(N^{-3}), & \text{for } r \geq 2. \end{cases} \tag{27}$$

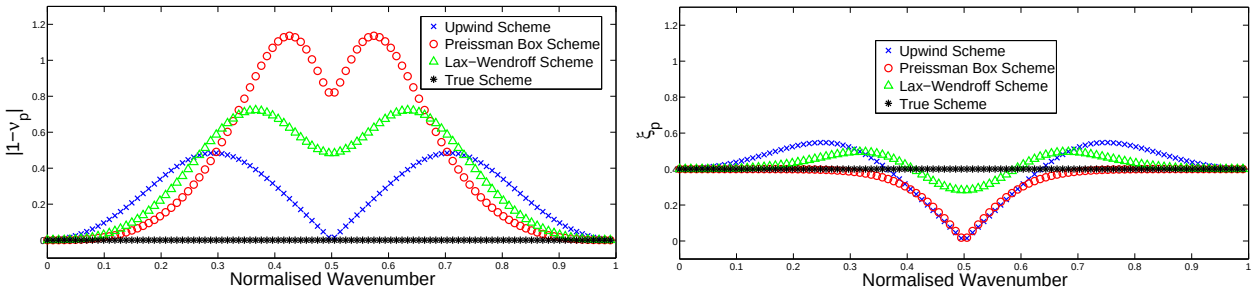
Table 1 shows that the order of convergence of (26) to zero, with respect to L, is positive for all values of r. This indicates that the error in the analysis vector increases as the length of the assimilation window is increased. The order of convergence also increases with the value of r associated with the initial condition, until a critical value is reached, where the order of convergence saturates at $O(L^2)$.

Numerical experiments were performed to investigate whether (21) was a good indicator for the growth and decay of the error in the analysis vector. 4D-Var was performed using the same conditions as in the above analysis, for the same finite difference schemes. $\|\tilde{\mathbf{x}}_0 - \mathbf{x}_a\|_2^2$ was then determined numerically and its order of convergence to zero with respect to N and L, was found. The functions and their value of r chosen for $\tilde{\mathbf{x}}_0$ were the square wave ($r = 0$), the triangular wave ($r = 1$) and the Gaussian wave, $\mathcal{N}(0.5, 0.01)$ ($r = \infty$).

The numerical experiments were executed using the Newton method, a zero first guess and a tolerance of 10^{-10}, to minimise the cost function. The Newton method for the linear advection equation converges in 1 step. For comparison with the results in Table 1, the same fixed values of N and L were chosen.

Initially consider the results for the order of convergence with respect to N in Table 2. The results are close to those in Table 1 for initial conditions with the same value of r. Figure 8(a) plots the numerical order of convergence with respect to N, as N is increased in powers of three. It shows that the order of convergence fluctuates about the order of convergence shown in Table 1, for each value of r. This explains why the results in Tables 1 and 2 do not match exactly for N. Table 2 also shows that the order of convergence with respect to L is a good match to those found in Table 1. These results indicate that (21) is an appropriate upper bound for the considered error in the analysis vector, with respect to N and L.

The next step is to re-introduce observation errors and understand how numerical model error and observation errors interact.



(a) $|1 - \nu_p|$ (b) ξ_p

Figure 7: The values of $|1 - \nu_p|$ and ξ_p plotted against the corresponding normalised wavenumber ie: $(p-1)/N$, of the corresponding eigenvalue, for the Upwind, Preissman Box, Lax-Wendroff and MNIMC schemes, for $N = 101$ and $L = 4$.

Variable	Upwind		
	$r = 0$	$r = 1$	$r = \infty$
α	1.4148×10^{-15}	-2.2612	-3.0000
β	5.6939×10^{-1}	1.5096	2.0000
Variable	**Preissman Box**		
	$r = 0$	$r = 1$	$r = \infty$
α	-6.5427×10^{-1}	-1.2809	-4.9178
β	3.7952×10^{-1}	9.8836×10^{-1}	2.0662
Variable	**Lax-Wendroff**		
	$r = 0$	$r = 1$	$r = \infty$
α	5.5724×10^{-1}	-2.0836	-4.9946
β	3.1248×10^{-1}	1.0187	2.0194

Table 2: Numerical orders of convergence to zero, with respect to N and L, for the error in the analysis vector from 4D-Var experiments, given to 5sf, $\|\tilde{\mathbf{x}}_0 - \mathbf{x}_a\|_2^2 = O(N^\alpha L^\beta)$. The results for N and L were identified using fixed $L = 4$ and fixed $N = 3^7$, respectively.

4.2. Spectral approach with observation errors

Section 4.1 provides an upper bound for the numerical model error in the analysis vector, for differing regularities of initial condition, $u(d, 0)$. It is possible to apply a similar bound to the error in the analysis vector when observation errors are also included.

Consider the case where each observation contains observation errors, $\mathbf{y}_l = \tilde{\mathbf{y}}_l + \boldsymbol{\epsilon}_l$, as described in section 2. Specifically, let us consider the random error known as *white noise* [15]. The cost function becomes,

$$J(\mathbf{x}_0) = \frac{1}{\sigma_o^2} \sum_{l=0}^{L} \left[\mathbf{y}_l - M^l \mathbf{x}_0\right]^T \left[\mathbf{y}_l - M^l \mathbf{x}_0\right]. \tag{28}$$

Minimising (28) with respect to $\mathbf{x}_0$,

$$\mathbf{x}_a = \left[\sum_{k=0}^{L} (M^T M)^k\right]^{-1} \sum_{l=0}^{L} (M^T)^l \left[\tilde{\mathbf{y}}_l + \boldsymbol{\epsilon}_l\right]. \tag{29}$$

Using the eigenvalue decomposition of M and $\tilde{M}$ as well as (16),

$$\mathbf{x}_a = A_L \tilde{\mathbf{x}}_0 + \boldsymbol{\rho}_L + V \left[\sum_{k=0}^{L} (\Lambda^* \Lambda)^k\right]^{-1} \left[\sum_{l=0}^{L} (\Lambda^*)^l V^* \boldsymbol{\epsilon}_l\right]. \tag{30}$$

The analysis vector in the presence of observation errors, is expressed in part by $\mathbf{x}_a$ without observation errors, as in (16). The observation errors form a separate term. If the errors did not possess the same variance, (30) would not have this property.

The term containing the observation errors in (30) would be the analysis vector when considering observations of the form $\mathbf{y}_l = \boldsymbol{\epsilon}_l$. The effect of numerical model error on the white noise may lead to correlations within the observation noise component of (30). If correlations have been introduced, then this will create artifacts in the analysis vector which will be propagated into its forecast. The autocorrelation function is used to determine if the observation noise contribution to $\mathbf{x}_a$ is still white noise.

The autocorrelation function is defined as in Mitra (2011) . The autocorrelation of an N-periodic sample $\mathbf{x} \in \mathbb{R}^N$, at lag $j = 0, \ldots, N-1$, is defined as $y_j : \mathbb{R}^N \to \mathbb{R}$, such that $\mathbf{x} \mapsto y_j(\mathbf{x})$ where,

$$y_j(\mathbf{x}) = \frac{1}{N} \sum_{p=1}^{N} [\mathbf{x}]_p [\mathbf{x}]_{(p-j)(\mathrm{mod}N)}, \tag{31}$$

where $[\mathbf{x}]_p$ denotes the pth element of $\mathbf{x}$. Also, define $\mathbf{y} \in \mathbb{R}^N$ such that the jth element of $\mathbf{y}(\cdot)$ is $y_{j-1}(\cdot)$ and let $\mathcal{F}_p(\cdot)$ denote the pth element of $\mathcal{F}(\cdot)$, $1 \leq p \leq N$. Then by the Wiener-Khintchine Theorem [15], the DFT of the autocorrelation of $\mathbf{x}$ is defined as,

$$\mathcal{F}[\mathbf{y}(\mathbf{x})] = \frac{1}{\sqrt{N}} \left[|\mathcal{F}_1(\mathbf{x})|^2, |\mathcal{F}_2(\mathbf{x})|^2, \ldots, |\mathcal{F}_N(\mathbf{x})|^2 \right]^T.$$

Using (30), the autocorrelation of the noise component of the analysis vector is given by,

$$y_{j-1}\left(\left[\sum_{k=0}^{L} (\Lambda^* \Lambda)^k \right]^{-1} \sum_{l=0}^{L} (\Lambda^*)^l V^* \boldsymbol{\epsilon}_l \right) = \frac{1}{N} \sum_{p=1}^{N} \left| \frac{\sum_{l=0}^{L} \overline{\lambda}_p^l \mathcal{F}_p(\boldsymbol{\epsilon}_l)}{\sum_{k=0}^{L} |\lambda_p|^{2k}} \right|^2 e^{\frac{2\pi i (j-1)(p-1)}{N}}, \tag{32}$$

for $j = 1, \ldots, N$. Hence,

$$\mathbb{E}\left[y_{j-1}\left(\left[\sum_{k=0}^{L} (\Lambda^* \Lambda)^k \right]^{-1} \sum_{l=0}^{L} (\Lambda^*)^l V^* \boldsymbol{\epsilon}_l \right) \right] = \frac{\sigma_o^2}{N} \sum_{p=1}^{N} \frac{e^{\frac{2\pi i (j-1)(p-1)}{N}}}{\sum_{k=0}^{L} |\lambda_p|^{2k}}, \quad \forall j = 1, \ldots, N, \tag{33}$$

which relies upon the values of j, N, L and σ_o^2, together with the dissipative properties of the considered finite difference scheme. It does not utilise the dispersive properties of the scheme. Expression (33) is potentially non-zero for all j, for a numerically dissipative finite difference scheme, indicating that the noise component of the analysis vector may no longer be random white noise. However, in the case of a non-dissipative scheme ie: $|\lambda_p| = 1 \quad \forall p$, only $j = 1$ is non-zero. Using a non-dissipative scheme such as the Preissman Box scheme, means that the noise component of the analysis vector will retain the white noise structure implicit in the observations.

A spectral approach as in section 4.1 is now used to provide an upper bound for the l_2-norm of the error in the analysis vector, for any regularity initial condition, in the presence of observation errors.

Lemma 3. *Let the assumptions detailed so far in this paper allowing $\mathbf{x}_a$ to be stated as in (30), hold true. Additionally, let $u_0(d)$ have regularity $r \in \mathbb{N}_0$ over $(0,1)$ such that $u_0^{(s)}(d)$ is bounded and piecewise monotone over $(0,1)$ and $\lim_{d\to 0^+} u_0^{(s)}(d) = \lim_{d\to 1^-} u_0^{(s)}(d)$, for all $s = 0, \ldots r$. Then,*

$$\|\tilde{\mathbf{x}}_0 - \mathbf{x}_a\|_2^2 \leq E_M + E_O + E_C, \tag{34}$$

where

$$\begin{aligned} E_M \quad = \quad & N \left\{ |1-\nu_1| D_1 + (|1-\nu_1| - 2\xi_1) \frac{D_3}{N^{r+1}} \right\}^2 \\ & + N \sum_{p=2}^{\frac{N+1}{2}} \left\{ |1-\nu_p| \frac{D_2}{(p-1)^{r+1}} + (|1-\nu_p| - 2\xi_p) \frac{D_3}{N^{r+1}} \right\}^2 \\ & + N \sum_{p=\frac{N+3}{2}}^{N} \left\{ |1-\nu_p| - 2\xi_p \right\}^2 \left\{ \frac{D_2}{(p-1)^{r+1}} + \frac{D_3}{N^{r+1}} \right\}^2, \end{aligned} \tag{35}$$

Variable	Upwind	Preissman Box	Lax-Wendroff
α	1.0000	1.0000	1.0000
β	-3.3207×10^{-4}	-9.9719×10^{-1}	-2.0866×10^{-3}

Table 3: Numerical orders of convergence to zero, with respect to N and L, for $\mathbb{E}[E_O]$ in (38), given to 5sf, $\mathbb{E}[E_0] = O(N^\alpha L^\beta)$. The results for N and L were identified using fixed $L = 4$ and fixed $N = 3^7$, respectively.

$$E_O = Ny_0 \left(\left[\sum_{k=0}^{L} (\Lambda^* \Lambda)^k \right]^{-1} \sum_{l=0}^{L} (\Lambda^*)^l V^* \boldsymbol{\epsilon}_l \right), \tag{36}$$

$$\begin{aligned} E_C \quad = \quad & -2\sqrt{N} \left\{ |1 - \nu_1| D_1 + (|1 - \nu_1| - 2\xi_1) \frac{D_3}{N^{r+1}} \right\} \left| \frac{\sum_{l=0}^{L} \overline{\lambda}_1^l \mathcal{F}(\boldsymbol{\epsilon}_l)}{\sum_{k=0}^{L} |\lambda_1|^{2k}} \right| \\ & -2\sqrt{N} \sum_{p=2}^{\frac{N+1}{2}} \left\{ |1 - \nu_p| \frac{D_2}{(p-1)^{r+1}} + (|1 - \nu_p| - 2\xi_p) \frac{D_3}{N^{r+1}} \right\} \left| \frac{\sum_{l=0}^{L} \overline{\lambda}_p^l \mathcal{F}(\boldsymbol{\epsilon}_l)}{\sum_{k=0}^{L} |\lambda_p|^{2k}} \right| \\ & -2\sqrt{N} \sum_{p=\frac{N+3}{2}}^{N} \left\{ |1 - \nu_p| - 2\xi_p \right\}^2 \left\{ \frac{D_2}{(p-1)^{r+1}} + \frac{D_3}{N^{r+1}} \right\} \left| \frac{\sum_{l=0}^{L} \overline{\lambda}_p^l \mathcal{F}(\boldsymbol{\epsilon}_l)}{\sum_{k=0}^{L} |\lambda_p|^{2k}} \right|, \end{aligned} \tag{37}$$

and D_1, D_2 and D_3 are constants independent of p and N and ξ_p is defined as in (22).

Proof. We refer to [16] for the proof where use has been made of the result from Lemma 1. □

The upper bound is formed from the equivalent upper bound in the absence of observation errors (E_M), together with the autocorrelation at lag 0 of the noise component in the analysis vector (E_O) and cross terms (E_C). It can be used to analyse the order of convergence of the error to zero, with respect to either N or L.

E_O and E_C are dependent on the random variables $\{\boldsymbol{\epsilon}_l\}_{l=0}^{L}$. However by the strong law of large numbers [19], if the experiments could be repeated, then as the number of experiments increased, the sample means of E_O and E_C would tend toward their expected values. As a consequence, we consider the expected values of E_O and E_C:

$$\mathbb{E}[E_O] = \sigma_o^2 \sum_{p=1}^{N} \frac{1}{\sum_{k=0}^{L} |\lambda_p|^{2k}}, \quad \text{and} \quad \mathbb{E}[E_C] = 0, \tag{38}$$

in order to identify the orders of convergence of the upper bound with respect to both N and L.

The expected value of E_C is zero whilst the expected value of E_O is dependent upon N, L, σ_o^2 and the dissipative properties of the scheme. Hence the expected value is independent of both the regularity of the initial condition and the dispersive properties of the finite difference scheme. A non-dissipative scheme leads to $\mathbb{E}[E_O] = \sigma_o^2 N/(L+1)$, so that the order of convergence for $\mathbb{E}[E_O]$ to zero with respect to N and L, is $O(N)$ and $O(L^{-1})$ respectively.

The order of convergence of the upper bound in (34) to zero, with respect to N or L, is determined by the dominant order of convergence possessed by either E_M or $\mathbb{E}[E_0]$. The orders of convergence to zero for E_M were analysed in section 4.1. Table 3 displays the numerical orders of convergence to zero, with respect to N and L, for $\mathbb{E}[E_O]$.

Initially consider the order of convergence of the upper bound in (34) with respect to N. The results of section 4.1 show that E_M remains constant or decays to zero, whilst Table 3 shows that $\mathbb{E}[E_O]$ increases, as N is increased. A similar property is seen for the order of convergence of each variable to zero, with respect to L. E_M increases, and $\mathbb{E}[E_O]$ decreases, as L is increased. Subsequently, the dominant order of convergence of (34) to zero, for both N and L, will be determined by the order of magnitude of the coefficients of each term.

4D-Var experiments similar to those in section 4.1 were run to investigate the appropriateness of (34) as an upper bound for the error in the analysis vector in the presence of observation errors. The results can be seen in Figures 9(a) and 9(b), for N and L respectively. Numerical results for N were generated using a tolerance of 10^{-8} for the convergence of the Newton method. Otherwise the experiments were performed using the same conditions as for the 4D-Var experiments in section 4.1.

Figure 9(a) shows that initially the error in the analysis vector decreases as N increases, according to the order of convergence seen for E_M. Once a critical value of N has been reached, the order of convergence then begins to increase with the order of convergence demonstrated by $\mathbb{E}[E_O]$. This provides a critical value for N at which the effect of both numerical model error and observation errors on the accuracy of the analysis vector is minimised. L and σ_o^2 form part of the coefficient for N in $\mathbb{E}[E_O]$. Increasing L or decreasing σ_o^2 will result in the critical value of N increasing, whilst decreasing L or increasing σ_o^2 will result in the critical value of N decreasing.

Figure 9(b) shows a similar picture to that seen in Figure 9(a). However in this instance, the initial decrease in the error in the analysis vector corresponds to the order of convergence of $\mathbb{E}[E_O]$. As L is increased further, a critical value is reached where E_M becomes dominant over $\mathbb{E}[E_O]$, and the error begins to increase with L. As with N, this critical value of L is determined by the coefficients of $\mathbb{E}[E_O]$. Decreasing either N or σ_o^2 will result in the critical value of L decreasing, whilst increasing either N or σ_o^2 will result in the critical value of L increasing.

When considering the orders of convergence with respect to either N or L, reducing σ_o^2 corresponds to reducing the error in the observations. As a result, it is not surprising that reducing σ_o^2 results in E_M becoming the dominant order of convergence, for a given value of N or L.

This analysis suggests that (34) is an appropriate bound to demonstrate the order of convergence for the error in the analysis vector. As a result, given a fixed value for σ_o^2 and either N or L, it is possible to choose a value for L and N respectively, that minimises the error in the analysis vector due to numerical model error and observation errors.

5. Conclusion

In 4D-Var, using a purely dissipative finite difference scheme, wavenumber components of the analysis vector can be completely lost due to destructive interference. As a consequence these components will not be present in the forecast.

The rate at which the error in the analysis vector decays to zero, as the number of discretisation points is increased, was found to depend on the regularity of the initial condition. The error was constant for a zero regularity initial condition, whilst higher regularity initial conditions allowed the error to decay. An upper bound for the error was determined and numerical experiments support the outcome.

Including observation errors resulted in a similar upper bound for the error in the analysis vector. As the number of discretisation points was increased, the error in the analysis vector decayed due to the decay of the numerical model error. However, past a critical number of grid points, the error begins to increase due to observation errors. This suggests there is a critical value of N at which the effects of numerical model error and observation errors are both minimised. The same trend is seen as the length of the assimilation window is increased, however the initial decay is due to observation errors, whilst the subsequent increase is due to numerical model error.

Observation errors in the form of white noise were found to have the potential to introduce correlated noise structures into the analysis vector, possibly leading to artifacts in the analysis vector and its subsequent forecast. Using a non-dissipative scheme would reduce the effects of these artifacts.

References

[1] M. Freitag, R. Potthast, Synergy of inverse problems and data assimilation techniques, in: M. Cullen, M. Freitag, S. Kindermann, R. Scheichl (Eds.), Large Scale Inverse Problems. Computational Methods and Applications in the Earth Sciences, volume 13 of *Radon Ser. Comput. Appl. Math.*, Walter de Gruyter, Berlin, 2013, pp. 1–40.

[2] N. Nichols, Treating model error in 3-D and 4-D data assimilation, in: R. Swinbank, V. Shutyaev, W. Lahoz (Eds.), Data Assimilation for the Earth System, volume 26 of *NATO science series*, Springer Netherlands, 2003, pp. 127–135.

[3] A. Lorenc, Analysis methods for numerical weather prediction, Quarterly Journal of the Royal Meteorological Society 112 (1986) 1177–1194.

[4] F. Bouttier, P. Courtier, Data assimilation concepts and methods, Meteorological Training Course Lecture Series (1999).

[5] R. Daley, Atmospheric Data Analysis, Cambridge Atmospheric and Space Science Series, Cambridge University Press, Cambridge, UK, 1st edition, 1999.

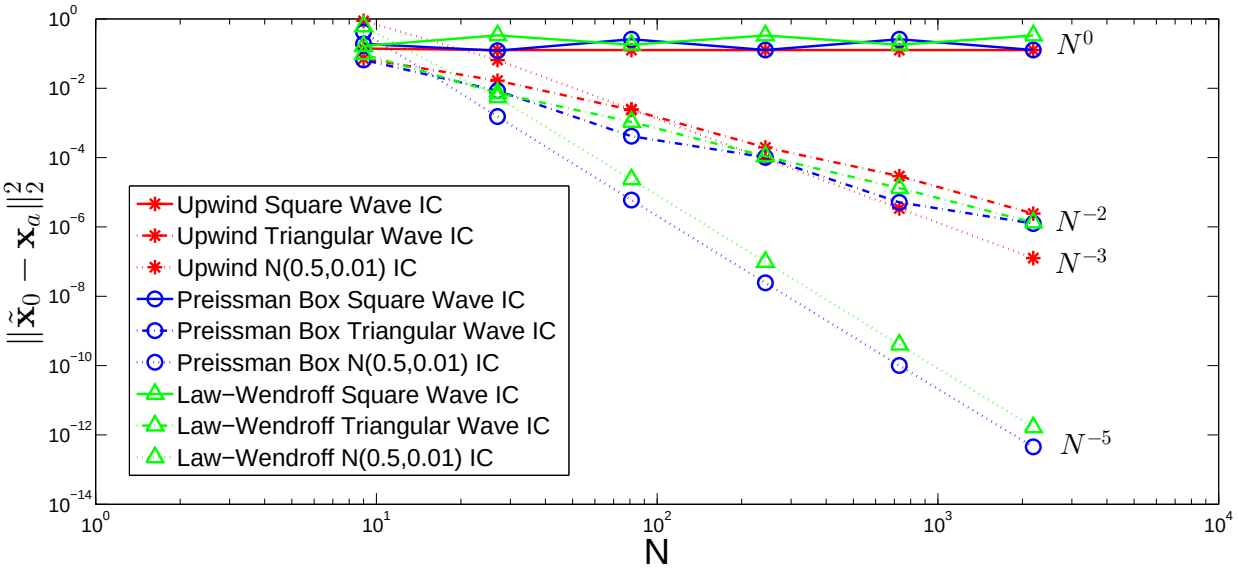
[6] N. Nichols, Mathematical concepts of data assimilation, in: W. Lahoz, B. Khattatov, B. Menard (Eds.), Data Assimilation, Making Sense of Observations, Springer Berlin Heidelberg, 2010, pp. 13–39.

[7] K. Morton, D. Mayers, Numerical solution of partial differential equations, Cambridge University Press, Cambridge, UK, 2nd edition, 2005.

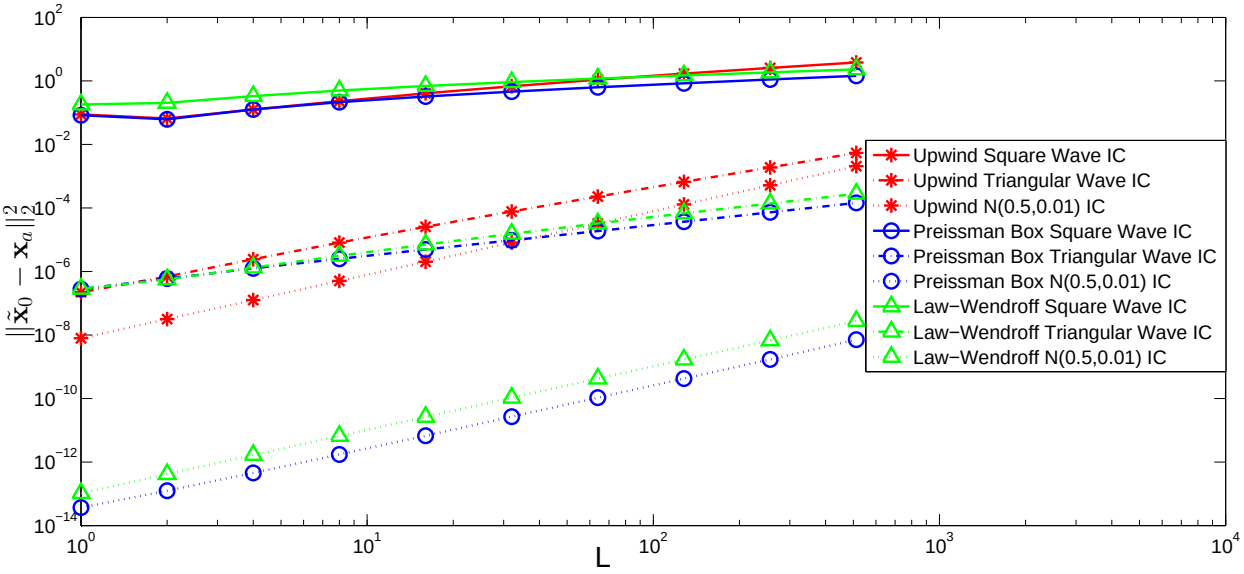
[8] M. Freitag, N. Nichols, C. Budd, Resolution of sharp fronts in the presence of model error in variational data assimilation, Quarterly Journal of the Royal Meteorological Society (2012).

[9] R. LeVeque, Numerical Methods for Conservation Laws, Lectures in Mathematics, Birkhäuser Verlag, Basel, Switzerland, 1st edition, 1999.

[10] D. Durran, Numerical Methods for Wave Equations in Geophysical Fluid Dynamics, volume 32 of *Texts in Applied Mathematics Series*, Springer New York, New York, 1st edition, 1999.
[11] P. Hansen, N. J.G., D. O'Leary, Deblurring Images: Matrices, Spectra, and Filtering, Fundamentals of Algorithms, SIAM, Philadelphia, 1st edition, 2006.
[12] W. Briggs, V. Henson, The DFT : an owner's manual for the discrete Fourier transform, SIAM, Philadelphia, 1st edition, 1995.
[13] M. Boas, Mathematical Methods in the Physical Sciences, John Wiley & Sons Inc., USA, 3rd edition, 2006.
[14] M. Freitag, Transcritical Flow Modelling with the Box Scheme, Dissertation, MSc in modern applications of mathematics, University of Bath, UK, 2003.
[15] S. Mitra, Digital Signal Processing: A Computer Based Approach, McGraw Hill Higher Education, New York, 3rd edition, 2006.
[16] S. Jenkins, Numerical Model Error in Data Assimilation, Thesis, PhD in engineering and mathematics, University of Bath, UK, 2013.
[17] H. Carslaw, Introduction to the Theory of Fourier's Series and Integrals, Dover Publications, Inc., New York, 3rd edition, 1950.
[18] E. Hólm, Lecture notes on assimilation algorithms, Meteorological Training Course Lecture Series (2003).
[19] G. Grimmett, D. Stirzaker, Probability and random processes, Clarendon Press, Oxford, 1st edition, 1982.

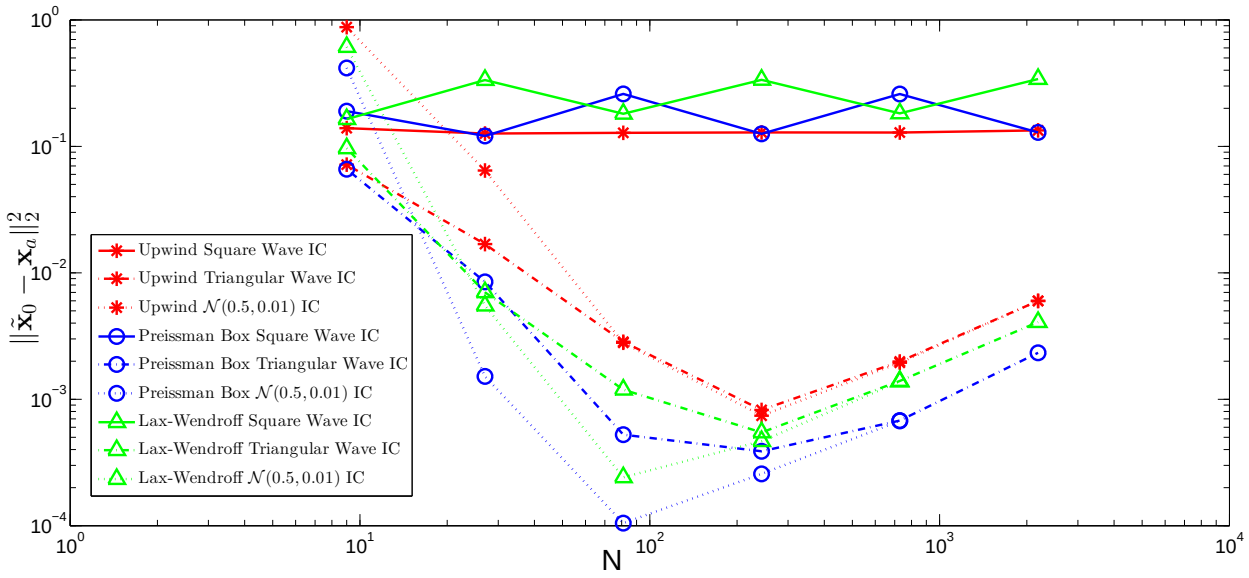


(a) The order of convergence to zero, with respect to N, using fixed $L = 4$.

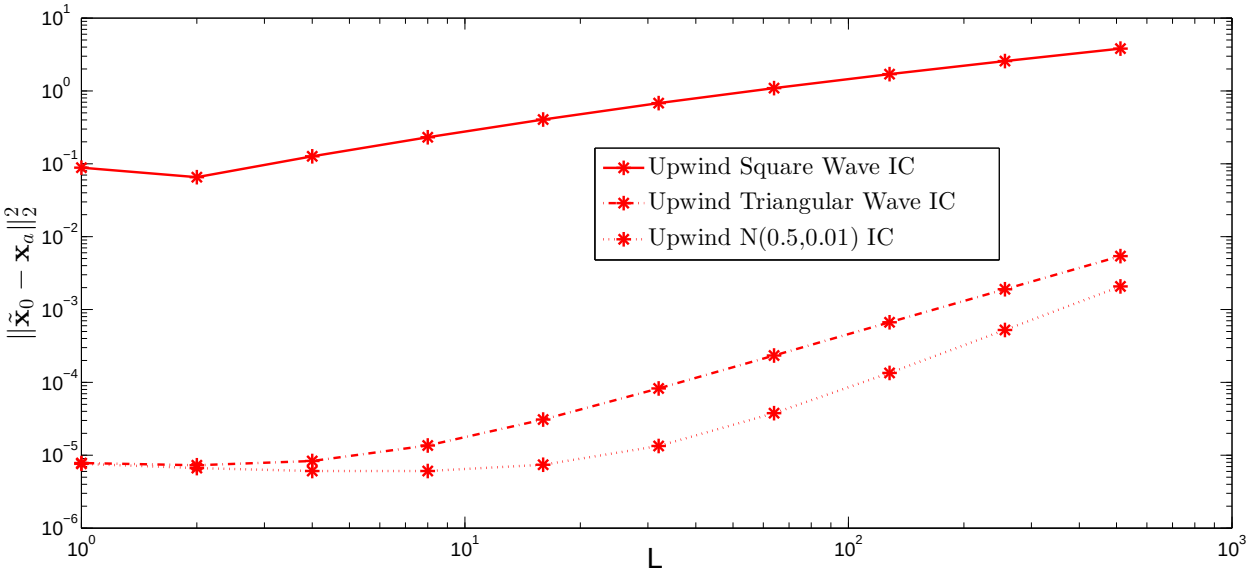


(b) The order of convergence to zero, with respect to L using fixed $N = 3^7$.

Figure 8: Numerical orders of convergence to zero, with respect to N and L, for $||\tilde{\mathbf{x}}_0 - \mathbf{x}_a||_2^2$. Numerical model error is the only error affecting the accuracy of $\mathbf{x}_a$. The results were obtained using 4D-Var data assimilation experiments.



(a) The order of convergence to zero, with respect to N, using fixed $L = 4$ and $\sigma_o^2 = 5 \times 10^{-6}$.



(b) The Order of convergence to zero, with respect to L, using fixed $N = 3^7$ and $\sigma_o^2 = 5 \times 10^{-9}$.

Figure 9: Numerical orders of convergence to zero, with respect to N and L, for $\|\tilde{\mathbf{x}}_0 - \mathbf{x}_a\|_2^2$. Numerical model error and observation error are the only errors affecting the accuracy of $\mathbf{x}_a$. The results were obtained using 4D-Var data assimilation experiments.