Controlling Unpredictability with Observations in the Partially Observed Lorenz ’96 Model

K.J.H. Law¹, D. Sanz-Alonso², A. Shukla², A.M. Stuart²

¹ Division of Computer, Electrical and Mathematical Sciences and Engineering (CEMSE), 4700 King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia.
² Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

E-mail: abhi23iiser@gmail.com

Abstract.

The long-time behavior of filters for the partially observed Lorenz ’96 model is studied. It is proven that for both discrete-time and continuous-time observations the 3DVAR filter can recover the signal within a neighborhood defined by the size of the observational noise, as long as a sufficiently large proportion of the state vector is observed; an explicit form for a sufficient constant observation operator is given. Furthermore, non-constant adaptive observation operators, with data incorporated by use of both the 3DVAR and the extended Kalman filter, are studied numerically. It is shown that for carefully chosen adaptive observations, the proportion of state coordinates necessary to accurately track the signal is significantly smaller than the proportion proved to be sufficient for constant observation operator. Indeed it is shown that the necessary number of observations may even be significantly smaller than the total number of positive Lyapunov exponents of the underlying system.
1. Introduction

Data assimilation is concerned with the blending of data and dynamical mathematical models, often in an online fashion where it is known as filtering; motivation comes from applications in the geophysical sciences such as weather forecasting [7], oceanography [3] and oil reservoir simulation [12]. Over the last decade there has been a growing body of theoretical understanding which enables use of the theory of synchronization in dynamical systems to establish desirable properties of these filters. This idea is highlighted in the recent book [1] from a physics perspective and, on the rigorous mathematical side, has been developed from a pair of papers by Olson, Titi and co-workers [13, 6], in the context of the Navier-Stokes equation in which a finite number of Fourier modes are observed. This mathematical work of Olson and Titi concerns perfect (noise-free) observations, but the ideas have been extended to the incorporation of noisy data for the Navier-Stokes equation in the papers [4, 5]. Furthermore the techniques used are quite robust to different dissipative dynamical systems, and have been demonstrated to apply in the Lorenz ’63 model [6, 8], and also to point-wise in space and continuous time observations [2] by use of a control theory perspective similar to that which arises from the derivation of continuous time limits of discrete time filters [4]. A key question in the field is to determine relationships between the underlying dynamical system and the observation operator which are sufficient to ensure that the signal can be accurately recovered from a chaotic dynamical system, whose initialization is not known precisely, by the use of observed data. Our purpose is to investigate this question theoretically and computationally in the context of the Lorenz ’96 model, widely adopted as a useful test model in the atmospheric sciences data assimilation community [11]. Our theoretical results will demonstrate the robustness of the methodology proposed by Olson and Titi, by extending it to the Lorenz ’96 model, whilst our computational results will highlight the gap between the current theory and what can be achieved in practice. In particular we will conduct careful numerical experiments which demonstrate that generalizations of the idea of assimilation in the unstable subspace (AUS), proposed by Trevisan et al. [16], lead to desirable signal estimation properties with relatively sparse observation operators. Indeed this approach suggests highly efficient new algorithms if the observation operator is allowed to adapt to the current state of the dynamical system. The question of how to optimize the observation operator to maximize information was first addressed in the context of atmospheric science applications in [10].

The outline of this paper is as follows. In section 2 we set up the Lorenz ’96 model and discuss its properties relevant to this work. We also introduce a variety of observation models including a fixed observation operator, and an adaptive one which employs knowledge of the linearized dynamics over the assimilation window to align
the observation operator with the unstable directions of the dynamics. In sections 3 and 4 we study theoretical properties of various filters in continuous and discrete time, for a particular fixed observation operator which corresponds to observing two thirds of the signal. In both cases we start by studying the set-up of Olson and Titi, where the observations contain no noise. We then extend the results to deal with noisy observations by studying the 3DVAR filtering algorithm. In section 5 we numerically study the performance of a range of Kalman-based filtering schemes, with different choices of the observation operator. In subsection 5.1 we consider the 3DVAR filtering scheme, whilst subsection 5.2 focuses on the Extended Kalman Filter (ExKF). In subsection 5.2 we also compare the adaptive observation implementation of the ExKF with the AUS scheme [16], that motivates our adaptive observation operator. In section 6 we summarize the work and draw some brief conclusions.

In conclusion our work highlights the role of ideas from dynamical systems in the rigorous analysis of filtering schemes and, through computational studies, shows the gap between theory and practice, demonstrating the need for further theoretical developments. We emphasize that the adaptive observation operator methods may not be implementable in practice on the high dimensional systems arising in, for example, meteorological applications. However, they provide conceptual insights into the development of improved algorithms and it is hence important to understand their properties.

2. Set Up

We introduce a filtering framework for the Lorenz ’96 model. Subsection 2.1 deals with the forward model, and subsection 2.2 with the observation model. Finally, in subsection 2.3 we give a high-level explanation of the connection between synchronization and Kalman-based filters.

2.1. Forward Model: Lorenz ’96

The Lorenz ’96 model is a lattice-periodic system of coupled nonlinear ordinary differential equations whose solution \( u = (u^{(1)}, \ldots, u^{(J)})^T \in \mathbb{R}^J \) satisfies

\[
\frac{du^{(j)}}{dt} = u^{(j-1)}(u^{(j+1)} - u^{(j-2)}) - u^{(j)} + F \quad \text{for } j = 1, 2, \ldots, J,
\]

subject to the periodic boundary conditions

\[
u^{(0)} = u^{(J)}, \quad u^{(J+1)} = u^{(1)}, \quad u^{(-1)} = u^{(J-1)}. \]

Here \( F \) is a forcing parameter, constant in time. We will choose it, in the numerical results which follow, so that the dynamical system exhibits sensitive dependence on
initial conditions and positive Lyapunov exponents. For example, for \( F = 8 \) and \( J = 60 \) the system is chaotic. Our theoretical results apply to any choice of the parameter \( F \) and to arbitrarily large system dimension \( J \).

It is helpful to write the model in the following form, widely adopted in the analysis of geophysical models as dissipative dynamical systems [15]:

\[
d\mathbf{u}/dt + A\mathbf{u} + B(\mathbf{u}, \mathbf{u}) = \mathbf{f}, \quad \mathbf{u}(0) = \mathbf{u}_0
\]

(2.3)

where

\[
A = I_{J \times J}, \quad \mathbf{f} = \begin{pmatrix} F \\ \vdots \\ F \end{pmatrix}_{J \times 1}
\]

and for \( \mathbf{u}, \tilde{\mathbf{u}} \in \mathbb{R}^J \)

\[
B(\mathbf{u}, \tilde{\mathbf{u}}) = -\frac{1}{2} \begin{pmatrix} \tilde{u}^{(2)}(J) + u^{(2)}(J) - \tilde{u}^{(J)}(J-1) - u^{(J)}(J-1) \\ u^{(J-1)}(j+1) + \tilde{u}^{(j+1)}(J) - \tilde{u}^{(j-2)}(J-1) - u^{(j-2)}(J-1) \\ \vdots \\ \tilde{u}^{(J-1)}(1) + u^{(J-1)}(1) - \tilde{u}^{(J-2)}(J-1) - u^{(J-2)}(J-1) \end{pmatrix}_{J \times 1}
\]

We denote by \( \langle \cdot, \cdot \rangle \) and \( | \cdot | \) the standard Euclidean inner-product and norm. For positive-definite matrix \( C \) we define \( | \cdot |_C := |C^{-\frac{1}{2}} \cdot | \).

We will use the following properties of \( A \) and \( B \):

**Property 2.1.1.** For \( \mathbf{u}, \tilde{\mathbf{u}} \in \mathbb{R}^J \)

((1)) \( \langle A\mathbf{u}, \mathbf{u} \rangle = |\mathbf{u}|^2 \).

((2)) \( \langle B(\mathbf{u}, \mathbf{u}), \mathbf{u} \rangle = 0 \).

((3)) \( B(\mathbf{u}, \tilde{\mathbf{u}}) = B(\tilde{\mathbf{u}}, \mathbf{u}) \).

((4)) \( |B(\mathbf{u}, \tilde{\mathbf{u}})| \leq 2|\mathbf{u}||\tilde{\mathbf{u}}| \).

((5)) \( 2\langle B(\mathbf{u}, \tilde{\mathbf{u}}), \mathbf{u} \rangle = -\langle B(\mathbf{u}, \mathbf{u}), \tilde{\mathbf{u}} \rangle \).

**Proof.** Properties (1), (2) and (3) are straightforward and we omit the proofs. We start showing (4). For any \( \mathbf{u} \in \mathbb{R}^J \) set

\[
|\mathbf{u}|_\infty = \max_{1 \leq j \leq J} |u^{(j)}|
\]

and recall that \( |\mathbf{u}|^2 \geq |\mathbf{u}|_\infty^2 \). Then, for \( \mathbf{u}, \tilde{\mathbf{u}} \in \mathbb{R}^J \), and for \( 1 \leq j \leq J \), we have that

\[
2|B(\mathbf{u}, \tilde{\mathbf{u}})^{(j)}| \leq |\mathbf{u}|_\infty (|\tilde{u}^{(j+1)}| + |\tilde{u}^{(j-2)}|) + ||\tilde{\mathbf{u}}||_\infty (|u^{(j+1)}| + |u^{(j-2)}|),
\]
and so

$$4|B(u, \tilde{u})|^2 \leq 2\|u\|^2 \sum_{j=1}^{J} (|\tilde{u}^{(j+1)}| + |\tilde{u}^{(j-2)}|)^2 + 2\|\tilde{u}\|^2 \sum_{j=1}^{J} (|u^{(j+1)}| + |u^{(j-2)}|)^2$$

$$\leq 8\|u\|_{\infty}^2 |\tilde{u}|^2 + 8\|\tilde{u}\|_{\infty}^2 |u|^2$$

$$\leq 16|u|^2 |\tilde{u}|^2.$$  

Hence

$$|B(u, \tilde{u})| \leq 2|u||\tilde{u}|.$$  

For (5) we use rearrangement and periodicity of indices under summation as follows:

$$2\langle B(u, \tilde{u}), u \rangle = \sum_{j=1}^{J} \left( u^{(j)}(u^{(j-1)}\tilde{u}^{(j+1)} + \tilde{u}^{(j-1)}u^{(j+1)} - \tilde{u}^{(j-1)}u^{(j-2)} - u^{(j-1)}\tilde{u}^{(j-2)}) \right)$$

$$= \sum_{j=1}^{J} \left( u^{(j)}(u^{(j-1)}\tilde{u}^{(j+1)} - \tilde{u}^{(j)}u^{(j-1)}u^{(j-2)} - \tilde{u}^{(j-1)}u^{(j-2)}) \right)$$

$$= \sum_{j=1}^{J} \left( u^{(j-1)}(u^{(j-2)}\tilde{u}^{(j)} - \tilde{u}^{(j+1)}u^{(j-1)}u^{(j-1)}) \right)$$

$$= \sum_{j=1}^{J} \left( \tilde{u}^{(j)}(u^{(j-1)}u^{(j-2)} - u^{(j+1)}u^{(j-1)}) \right)$$

$$= -\langle B(u, u), \tilde{u} \rangle.$$  

The existence of an absorbing, forward-invariant ball for equation (2.3) follows in a straightforward fashion from the preceding identities, as shown in the following proposition.

**Proposition 2.1.2.** Let $K = 2JF^2$ and define $B := \{ u \in \mathbb{R}^J : |u|^2 \leq K \}$. Then $B$ is an absorbing, forward-invariant ball for equation (2.3).

**Proof.** Taking the Euclidean inner product of $u(t)$ with equation (2.3) and using properties (1) and (2) we get

$$\frac{1}{2} \frac{d|u|^2}{dt} = -|u|^2 + \langle f, u \rangle.$$  

Using Young’s inequality for the last term gives

$$\frac{d|u|^2}{dt} + |u|^2 \leq JF^2.$$
Therefore, using Gronwall’s lemma,

\[ |u(t)|^2 \leq |u_0|^2 e^{-t} + JF^2(1 - e^{-t}), \]

and the result follows. \[\square\]

We close this subsection by introducing some notation. First, we rewrite (2.3) as

\[ \frac{du}{dt} = \mathcal{F}(u), \quad u(0) = u_0, \]

where \( \mathcal{F}(u) := f - Au - B(u,u) \).

The solution to (2.4) is referred to as the signal. We denote by \( \Psi : \mathbb{R}^J \times \mathbb{R}^+ \rightarrow \mathbb{R}^J \) the solution operator for the equation (2.4), so that \( u(t) = \Psi(u_0; t) \). In our discrete time filtering developments we assume that, for some fixed \( h > 0 \), the signal is subject to observations at times \( t_k := kh, \ k \geq 1 \). We then write \( \Psi(\cdot) := \Psi(\cdot; h) \) and \( u_k := u(kh) \), with slight abuse of notation to simplify the presentation.

2.2. Observation Operator and Filtering

2.2.1. General Setting

Our main interest is in using partial observations of the discrete time dynamical system

\[ u_{k+1} = \Psi(u_k), \quad k \geq 0, \]  

(2.5)

to make estimates of the state of the system. To this end we introduce the family of linear observation operators \( \{H_k\}_{k \geq 1} \), where \( H_k : \mathbb{R}^J \rightarrow \mathbb{R}^J \) is assumed to have rank (which may change with \( k \)) less than or equal to \( M \leq J \), and consider data \( \{y_k\}_{k \geq 1} \) given by

\[ y_k = H_k u_k + \nu_k, \quad k \geq 1, \]

(2.6)

where we assume that the random and/or systematic error \( \nu_k \) (and hence also \( y_k \)) is contained in \( H_k \mathbb{R}^J \). If \( Y_k = \{y_\ell\}_{\ell=1}^k \) then the objective of filtering is to estimate \( u_k \) from \( Y_k \) given incomplete knowledge of \( u_0 \). We are most interested in the case where \( M < J \), so that the observations are partial, and \( H_k \mathbb{R}^J \) is a strict subset of \( \mathbb{R}^J \); in particular we address the question of how small \( M \) can be chosen whilst still allowing accurate recovery of the signal over long time-intervals. Let \( m_k \) denote some estimate of \( u_k \) given \( Y_k \). All the discrete time algorithms we consider proceed inductively in the sense that the estimate \( m_{k+1} \) is determined by the previous one, \( m_k \), and the observed data \( y_{k+1} \).

It is convenient to see this update \( m_k \mapsto m_{k+1} \) as a two-step process. In the first one, known as forecast step, the estimate \( m_k \) is evolved with the dynamics of the underlying model yielding a prediction for the current state of the system. In the second step,
known as *analysis step*, the forecast is used in conjunction with the observed data \( y_{k+1} \) to produce the estimate \( m_{k+1} \) of the true state of the underlying system \( u_{k+1} \).

In section 3 we study the continuous time filtering problem, where the goal is to estimate the value of a continuous time signal

\[
 u(t) = \Psi(u_0, t), \quad t \geq 0,
\]

at time \( T > 0 \). As in the discrete case, it is assumed that only incomplete knowledge of \( u_0 \) is available. In order to estimate the signal \( u(T) \) we have access, at each time \( 0 < t \leq T \), to a (perhaps noisily perturbed) projection of the signal given by a time-independent observation matrix \( H \).

### 2.2.2. Fixed Observation Operator

In our discrete and continuous time theoretical developments we will consider a specific choice of fixed observation matrix \( P \) (so that \( H_k = H = P \)) that we now introduce. First, we let \( \{e_j\}_{j=1}^J \) be the standard basis for the Euclidean space \( \mathbb{R}^J \) and assume that \( J = 3J' \) for some \( J' \geq 1 \). Then the projection matrix \( P \) is defined by replacing every third column of the identity matrix \( I_{J \times J} \) by the zero vector:

\[
P = \left( e_1, e_2, 0, e_4, e_5, 0, \ldots \right)_{J \times J}. \tag{2.7}
\]

Thus \( P \) has rank \( M = 2J' \). We also define its complement \( Q \) as

\[
Q = I_{J \times J} - P.
\]

**Remark 2.2.1.** Note that in the definition of the projection matrix \( P \) we could have chosen either the first or the second column to be set to zero periodically, instead of choosing every third column this way; the theoretical results in the next two sections would be unaltered by doing this.

The matrix \( P \) provides sufficiently rich observations to allow the accurate recovery of the signal in the long-time asymptotic regime. This will be shown in a continuous time setting in section 3, and in a discrete time setting in section 4. The following property of \( P \) will be used:

**Property 2.2.2.** The bilinear form satisfies \( B(Qu, Qu) = 0 \) and, furthermore, there is a constant \( c > 0 \) such that

\[
|\langle B(u, u), \tilde{u} \rangle| \leq c|u||\tilde{u}||Pu|.
\]
Proof. The first part is automatic since, if \( q := Qu \), then for all \( j \) either \( q^{(j-1)} = 0 \) or \( q^{(j-2)} = q^{(j+1)} = 0 \). Since \( B(Qu,Qu) = 0 \) and \( B(\cdot,\cdot) \) is a bilinear operator we can write
\[
B(u,u) = B(Pu + Qu, Pu + Qu) \\
= B(Pu, Pu) + 2B(Pu, Qu).
\]
Now using property (4) and the fact that there is \( c > 0 \) such that \( |Pu| + 2|Qu| \leq \frac{c}{2} |u| \),
\[
|\langle B(u,u), \tilde{u} \rangle| \leq |B(u,u)||\tilde{u}| \\
\leq |B(Pu, Pu) + 2B(Pu, Qu)||\tilde{u}| \\
\leq 2|Pu||\tilde{u}|(|Pu| + 2|Qu|) \\
\leq c|Pu||\tilde{u}||u|.
\]

2.2.3. Adaptive Observation Operator  Successful filtering of chaotic models is driven by observing enough of the unstable dynamics to control the exponential separation of trajectories in the dynamics. If the observation operator is fixed at \( P \) then, as we will see, this is sufficient to effect this control. However, by adapting the observations to the dynamics, we will be able to obtain the same quality of reconstruction with far fewer observations, that is with a smaller value of \( M \).

The variational equation for the dynamical system (2.4) is given by
\[
\frac{d}{dt}D\Psi(u,t) = DF(\Psi(u,t)) \cdot D\Psi(u,t), \quad D\Psi(u,0) = I_{J \times J},
\]
using the chain rule. The solution of the variational equation gives the derivative matrix of the solution operator \( \Psi \), which in turn yields how \( \Psi \) acts under small variations in the initial value \( u \).

Let \( L_{k+1} := L(t_{k+1}) \) be the solution of the variational equation (2.8) over the assimilation window \((t_k, t_{k+1})\), initialized at \( I_{J \times J} \), given as
\[
\frac{dL}{dt} = DF(\Psi(m_{k-1}, t - t_{k-1}))L, \quad L(t_k) = I_{J \times J}.
\]

Let \( \{\lambda^j_k, \psi^j_k\}_{j=1}^J \) denote eigenvalue/eigenvector pairs of the matrix \( L_{k+1}^T L_{k+1} \), where the eigenvalues (which are, of course, real) are ordered to be non-decreasing, and the eigenvectors are orthonormalized with respect to the Euclidean inner-product \( \langle \cdot, \cdot \rangle \). We define the adaptive observation operator \( H_k \) to be
\[
H_k := H_0(\psi^1_k, \cdots, \psi^J_k)^T
\]
where
\[
H_0 = \begin{pmatrix} 0 & 0 \\ 0 & I_{M \times M} \end{pmatrix}.
\]
Thus $H_0$ and $H_k$ both have rank $M$. Defined in this way we see that $H_k v$ is given by the vector

$$(0, \ldots, 0, \langle \psi^j_{k+1-M+1}, v \rangle, \ldots, \langle \psi^j_k, v \rangle)$$

that is the projection of $v$ onto the $M$ eigenvectors of $L_{k+1}^T L_{k+1}$ with largest modulus.

**Remark 2.2.3.** In the following work we consider the leading eigenvalues and corresponding eigenvectors of the matrix $L_k^T L_k$ to track the unstable (positive Lyapunov growth) directions. To leading order in $h$ it is equivalent to consider the matrix $L_k L_k^T$ in the case of frequent observations (small $h$) as can be seen by the following expressions

$$L_k^T L_k = (I + h D F_k)^T (I + h D F_k) + O(h^2)$$

and

$$L_k L_k^T = (I + h D F_k)(I + h D F_k)^T + O(h^2)$$

where $D F_k = D F(m_k)$.

Of course for large intervals $h$, the above does not hold, and the difference between $L_k^T L_k$ and $L_k L_k^T$ may be substantial. It is however clear that these operators have the same eigenvalues, with the eigenvectors of $L_k^T L_k$ corresponding to $\lambda^j_k$ given by $L_k \psi^j_k$ for the corresponding eigenvector $\psi^j_k$ of $L_k^T L_k$. That is to say, for the linearized deformation map $L_k$, the direction $\psi^j_k$ is the pre-deformation principle direction corresponding to the principle strain $\lambda^j_k$ induced by the deformation. The direction $L_k \psi^j_k$ is the post-deformation principle direction corresponding to the principle strain $\lambda^j_k$. The dominant directions chosen in Eq. (2.10) are those directions corresponding to the greatest growth over the interval $(t_k, t_{k+1})$ of infinitesimal perturbations to the predicting trajectory, $\Psi(m_{k-1}, h)$ at time $t_k$. This is only one sensible option. One could alternatively consider the directions corresponding to the greatest growth over the interval $(t_{k-1}, t_k)$, or over the whole interval $(t_{k-1}, t_{k+1})$. Investigation of these alternatives is beyond the scope of this work and is therefore deferred to later investigation.

2.3. Synchronization and Data Assimilation

A fundamental idea underlying successful filtering is synchronization, as we now explain. Slightly abusing notation, we write $\Psi(v) = \Psi(Pv, Qv)$. Consider a true signal governed by (2.5) and written as $u_k = (p^\dagger_k, q^\dagger_k)$ and a filter $m_k = (p_k, q_k)$ defined as follows:

$$p_{k+1}^\dagger = P \Psi(p_k^\dagger, q_k^\dagger), \quad p_{k+1} = p_{k+1}^\dagger,$$

$$q_{k+1}^\dagger = Q \Psi(p_k^\dagger, q_k^\dagger), \quad q_{k+1} = Q(p_{k+1}^\dagger).$$
On the left the true signal is simply evolving according to (2.5) whilst on the right the synchronization filter is constructed by feeding the observed $p$-component of the signal into the dynamical model projected by $Q$. The paper [6] shows that, for the Navier-Stokes equation with sufficiently rich observations, and for the Lorenz ’63 model observed in only the first component,

$$|q_k - q^\dagger_k| \to 0 \text{ as } k \to \infty$$

so that the filter asymptotically recovers the true signal. This is a synchronization property.

The discrete time assimilation filters we study in this paper build on this property as follows. For any positive definite operator $A$ define $| \cdot |_A = |A^{-\frac{1}{2}} \cdot |$. The filters have the form

$$m_{k+1} = \arg\min_m \left\{ \frac{1}{2}|m - \Psi(m_k)|_{\tilde{C}_{k+1}}^2 + \frac{1}{2}|y_{k+1} - H_{k+1}m|_\Gamma^2 \right\}. \quad (2.12)$$

The norm in the second term is only applied within the $M$-dimensional image space of $H_{k+1}$, where $y_{k+1}$ lies; then $\Gamma$ is realized as a positive-definite $M \times M$ matrix in this image space, and $\tilde{C}_{k+1}$ is a positive-definite $J \times J$ matrix. The minimization represents a compromise between respecting the model and respecting the data, with the covariance weights $\tilde{C}_{k+1}$ and $\Gamma$ determining the relative size of the two contributions; see [9] for more details.

The filter where the model covariance $\tilde{C}_{k+1}$ is constant, $\tilde{C}_{k+1} = C_0$ is known as the 3DVAR filter. It can be seen as a perturbation (which allows for noisy data) of the noise free synchronization filter introduced in [6] and described above. For example, if $H_k = P$ and $C_0$ is suitably chosen in terms of $\Gamma$ (see subsection 4.2 for details) the solution to the variational problem (2.12) can be written, for some $\eta \geq 0$, as

$$m_{k+1} = S\Psi(m_k) + (I - S)y_{k+1}, \quad (2.13a)$$
$$S = \frac{\eta}{1 + \eta} P + Q. \quad (2.13b)$$

The limit $\eta \ll 1$ is termed variance inflation in the applied literature since it may be derived by increasing the covariance weight $C_0$. In this regime the weight on the observations is comparatively large and the 3DVAR is close to the synchronization filter. Indeed for the extreme value $\eta = 0$, the 3DVAR filter is given by

$$m_{k+1} = Q\Psi(m_k) + Py_{k+1}.$$ 

This may correspond to infinite $C_0$, or noiseless observations. In the latter case, $y_{k+1} = Pu_{k+1}$, and if the above is initialized with the same $m_0 = (p_0, q_0)$ then it is
Partially Observed Lorenz '96

exactly the synchronization filter: the unobserved part of the signal is estimated using
the model, and the observed part using only the observations.

Motivated by this perturbation result for \( \eta \ll 1 \) we will, in what follows, study the
noise-free case first, and then perturb this setting to deal with the more complicated
case of noisy data and the 3DVAR filter.

Our discussion in this subsection has been based on a formulation of the filtering
problem in discrete time. In the case of noise-free observations, the continuous time
analogue was, in fact, analyzed before the discrete case, in the paper [13]. Again
we slightly abuse notation by writing \( \mathcal{F}(v) = \mathcal{F}(Pv,Qv) \) and consider a true signal
governed by (2.4) written as \( u = (p^\dagger,q^\dagger) \) and a filter \( m = (p,q) \) defined as follows:

\[
\frac{dp^\dagger}{dt} = P\mathcal{F}(p^\dagger,q^\dagger), \quad p = p^\dagger,
\]

\[
\frac{dq^\dagger}{dt} = Q\mathcal{F}(p^\dagger,q^\dagger). \quad dq = Q\mathcal{F}(p^\dagger,q).
\]

On the left the true signal is simply evolving according to (2.4) whilst on the right the
filter is constructed by feeding the observed \( p \) component into the dynamical system
projected by \( Q \). The continuous time limit of 3DVAR with constant observation
operator \( H \), is obtained by setting \( \Gamma = h^{-1}\Gamma_0 \) and \( C_{k+1} = \tilde{C}_{k+1} = C \) and letting
\( h \to 0 \). The resulting filter, derived in [4], is given by

\[
\frac{dm}{dt} = \mathcal{F}(m) + CH^*\Gamma_0^{-1}\left(\frac{dz}{dt} - Hm\right), \tag{2.14}
\]

where the observed data is now \( z \) —formally the time-integral of \( y \)— and satisfies the
SDE

\[
\frac{dz}{dt} = Hv^\dagger + HT_0^\dagger \frac{dw}{dt}, \tag{2.15}
\]

for \( w \) a unit Wiener process. This filter has the effect of nudging the solution towards
the observed data in the \( H \)-projected direction. A similar idea is used in [2] to assimilate
pointwise observations of the Navier-Stokes equation.

3. Continuous Data Assimilation

In this section we consider the observation operator to be fixed at \( P \), and assume that
the data arrives continuously in time. Subsection 3.1 deals with noiseless data, and the
more realistic noisy scenario is studied in subsection 3.2. We aim to show that, in the
large time asymptotic, the filter is close to the truth. In the absence of noise our results
are analogous to those for the partially observed Lorenz '63 and Navier-Stokes models
in [13]; in the presence of noise the results are similar to those proved in [4] for the
Navier-Stokes equation and in [8] for the Lorenz '63 model, and generalize the work in
[14] to non-globally Lipschitz vector fields.
3.1. Noiseless Observations

The true solution $v$ satisfies the following equation

$$\frac{dv}{dt} + v + B(v, v) = f, \quad v(0) = v_0. \tag{3.1}$$

Suppose that the projection of the true solution, $Pv$, is perfectly observed and continuously assimilated into the approximate solution $m$. The approximate solution $m$ has the following form

$$m = Pv + q, \tag{3.2}$$

where $v$ is the true solution given by (3.1) and $q$ satisfies the equation (2.3) projected by $Q$ to obtain

$$\frac{dq}{dt} + q + QB(Pv + q, Pv + q) = Qf \tag{3.3}$$

with the initial condition $q(0) = q_0$. Equations (3.2) and (3.3) form the continuous time synchronization filter. The following theorem shows that the approximate solution converges to the true solution asymptotically as $t \to \infty$.

**Theorem 3.1.1.** Let $m$ be given by the equations (3.2), (3.3) and let $v$ be the true underlying solution of the equation (3.1) with initial data $v_0 \in B$, the absorbing ball in Proposition 2.1.2, so that $\sup_{t \geq 0} |v(t)|^2 \leq K$. Then

$$\lim_{t \to \infty} |m(t) - v(t)|^2 = 0.$$

**Proof.** Define the error in the approximate solution as $\delta = m - v = q - Qv$. Note that $Q\delta = \delta$. The error satisfies the following equation

$$Q\frac{d\delta}{dt} + Q\delta + Q(B(Pv + q, Pv + q) - B(v, v)) = 0.$$

Splitting $v = Pv + Qv$ and noting, from Properties 2.2.2 that $B(Qv, Qv) = 0$ and $B(q, q) = 0$, yields

$$\frac{dQ\delta}{dt} + Q\delta + 2QB(Pv, Q\delta) = 0.$$

Taking the inner product with $Q\delta$ gives

$$\frac{1}{2} \frac{d|Q\delta|^2}{dt} + |Q\delta|^2 + 2\langle B(Pv, Q\delta), Q\delta \rangle = 0.$$

Note that from the Properties 2.1.1, (3) and (5), and Property 2.2.2, we have

$$2\langle B(u, Q\delta), Q\delta \rangle = -\langle B(Q\delta, Q\delta), u \rangle = 0.$$
Thus since $Q\delta = \delta$ we have
\[
\frac{d|\delta|^2}{dt} + 2|\delta|^2 = 0,
\]
and so
\[
|\delta(t)|^2 = |\delta(0)|^2 e^{-2t}.
\]
As $t \to \infty$ the error $\delta(t) \to 0$.

The result establishes that in the case of high frequency in time observations the approximate solution converges to the true solution even though the signal is observed partially at frequency $2/3$ in space. We now extend this result by allowing for noisy observations.

### 3.2. Noisy Observations: Continuous 3DVAR

Recall that the continuous time limit of 3DVAR is given by (2.14) where the observed data $z$, the integral of $y$, satisfies the SDE (2.15). We study this filter in the case where $H = P$ and under small observation noise $\Gamma_0 = \epsilon^2 I$. The 3DVAR model covariance is then taken to be of the size of the observation noise. We choose $C = \sigma^2 I$, where $\sigma^2 = \sigma^2(\epsilon) = \eta^{-1}\epsilon^2$, for some $\eta > 0$. Then equations (2.14) and (2.15) can be rewritten as
\[
\frac{dm}{dt} = F(m) + \frac{1}{\eta} \left( \frac{dz}{dt} - Pm \right)
\] (3.4)
where
\[
\frac{dz}{dt} = Pv + \epsilon P \frac{dw}{dt},
\] (3.5)
and $w$ is a unit Wiener process. Note that the parameter $\epsilon$ represents both the size of the 3DVAR observation covariance and the size of the noise in the observations. The next result shows that the approximate solution $m$ converges to a neighbourhood of the true solution $v$ where the size of the neighbourhood depends upon $\epsilon$. Similarly as in [8] and [4] it is required that $\eta$, the ratio between the size of observation and model covariances, is sufficiently small.

**Theorem 3.2.1.** Let $m$ solve the equation (3.4) and let $v$ solve the equation (3.1) with the initial data $v(0) \in B$, the absorbing ball of Proposition 2.1.2, so that $\sup_{t \geq 0} |v(t)|^2 \leq K$. Then for the constant $c$ as given in the Property 2.2.2, given $\eta < \frac{4}{c^2 K}$ we obtain
\[
\mathbb{E}|m(t) - v(t)|^2 \leq e^{-\lambda t}|m(0) - v(0)|^2 + \frac{2Jc^2}{3\eta^2 \lambda}(1 - e^{-\lambda t}),
\] (3.6)
where $\lambda$ is defined by
\[
\lambda = 2 \left( 1 - \frac{c^2 \eta K}{4} \right).
\] (3.7)
Thus
\[
\limsup_{t \to \infty} \mathbb{E}|m(t) - v(t)|^2 \leq a \epsilon^2,
\]
where \(a = \frac{2L^2}{3 \lambda} \) does not depend on the strength of the observation noise, \(\epsilon\).

**Proof.** From (3.4) and (3.5) we have
\[
\frac{dm}{dt} = \mathcal{F}(m) + \frac{1}{\eta} \left( P v + \epsilon P \frac{dw}{dt} - P m \right).
\]
Thus
\[
\frac{dm}{dt} = -m - B(m, m) + f + \frac{1}{\eta} P \left( v - m \right) + \frac{\epsilon}{\eta} P \frac{dw}{dt}.
\]
We also have
\[
\frac{dv}{dt} = -v - B(v, v) + f.
\]
Defining \(\delta = m - v\) gives
\[
\frac{d\delta}{dt} = -\delta - 2B(v, \delta) - B(\delta, \delta) - \frac{1}{\eta} P \delta + \frac{\epsilon}{\eta} P \frac{dw}{dt}.
\]
Now taking inner product with \(\delta\), using Lemma 3.2.2, Properties 2.1.1 and applying Itô’s formula
\[
\frac{1}{2} d|\delta|^2 + \left(1 - \frac{\epsilon^2 K}{4}\right)|\delta|^2 dt \leq \frac{\epsilon}{\eta} \langle P dw, \delta \rangle + \frac{J}{3} \epsilon^2 dt.
\]
Integrating the inequality and taking expectations shows that \(\mathbb{E}|\delta|^2\) satisfies the integrated form of the differential inequality
\[
\frac{d\mathbb{E}|\delta|^2}{dt} \leq -\lambda \mathbb{E}|\delta|^2 + \frac{2J \epsilon^2}{3 \eta^2}.
\]
Use of the Gronwall inequality gives the desired result. \(\square\)

**Lemma 3.2.2.** Let \(v \in \mathcal{B}\). Then
\[
\langle \delta + 2B(v, \delta) + B(\delta, \delta) + \frac{1}{\eta} P \delta, \delta \rangle \geq \left(1 - \frac{\epsilon^2 K}{4}\right)|\delta|^2.
\]

**Proof.** Use of Property 2.1.1 items (3) and (5), together with Property 2.2.2 shows that
\[
\langle \delta + 2B(v, \delta) + B(\delta, \delta) + \frac{1}{\eta} P \delta, \delta \rangle = |\delta|^2 + 2\langle B(v, \delta), \delta \rangle + \langle B(\delta, \delta), \delta \rangle + \langle \frac{1}{\eta} P \delta, \delta \rangle
\]
\[
= |\delta|^2 - \langle B(\delta, \delta), v \rangle + \langle \frac{1}{\eta} P \delta, \delta \rangle
\]
\[
\geq |\delta|^2 - cK \delta \langle P \delta \rangle + \frac{1}{\eta} |P \delta|^2
\]
\[
\geq |\delta|^2 - \frac{\theta |\delta|^2}{2} - \frac{c^2 K |P \delta|^2}{2\theta} + \frac{1}{\eta} |P \delta|^2.
\]
4. Discrete Data Assimilation

We now turn to discrete data assimilation. Recall that filters in discrete time can be split into two steps: forecast and analysis. In this section we establish conditions under which the corrections made at the analysis steps overcome the divergence inherent due to nonlinear instabilities of the model. As in the previous section we study first the case of noiseless data, generalizing the work of [9] from the Navier-Stokes and Lorenz ’63 models to include the Lorenz ’96 model, and then study the case of 3DVAR, generalizing the work in [5, 8], which concerns the Navier-Stokes and Lorenz ’63 models respectively, to the Lorenz ’96 model.

4.1. Noiseless Observations

Let $h > 0$, and set $t_k := kh, k \geq 0$. For any function $g : \mathbb{R}^+ \to \mathbb{R}^J$, continuous in $[t_k, t_{k+1})$, we denote $g(t_k^-) := \lim_{t \to t_k^-} g(t)$. Let $v$ be a solution lying on the absorbing forward-invariant ball $B$ and satisfying the equation (3.1). For $t \in (t_k, t_{k+1})$ the discrete time synchronization filter $m$ of [6] may be expressed as follows:

$$\frac{dm}{dt} + m + B(m, m) = f, \quad t \in (t_k, t_{k+1}),$$

$$m(t_k) = Pv(t_k) + Qm(t_k^-).$$

Define the error $\delta = m - v$. Subtracting the equation (2.1) from the equation (4.1) gives

$$\frac{d\delta}{dt} + \delta + 2B(v, \delta) + B(\delta, \delta) = 0, \quad t \in (t_k, t_{k+1}),$$

$$\delta(t_k) = Q\delta(t_k^-).$$

We now show that the filter $m$ converges to the true underlying solution. Before establishing the main result we state and prove the following lemma.

**Lemma 4.1.1.** For any $F > 0$ and $J \geq 3$, there exists $\beta \in \mathbb{R}$ such that the following result holds

$$|\delta(t)|^2 \leq |\delta(t_k)|^2 e^{\beta(t-t_k)} \quad t \in [t_k, t_{k+1}).$$

**Proof.** Taking the inner product with $\delta$ in the equation (4.2) we obtain, for $t \in (t_k, t_{k+1})$,

$$\frac{1}{2} \frac{d|\delta|^2}{dt} + |\delta|^2 + 2\langle B(v, \delta), \delta \rangle + \langle B(\delta, \delta), \delta \rangle = 0$$

(4.3)
so that, by Property 2.1.1 item (2),
\[ \frac{1}{2} \frac{d|\delta|^2}{dt} + |\delta|^2 - 2|\langle B(v, \delta), \delta \rangle| \leq 0. \]

Using Properties 2.1.1(4) and (5) gives
\[ |\langle B(v, \delta), \delta \rangle| \leq K \frac{1}{2} |\delta|^2, \]
where \( K \) is defined in Proposition 2.1.2, so that
\[ \frac{1}{2} \frac{d|\delta|^2}{dt} \leq (2K^{\frac{1}{2}} - 1)|\delta|^2. \]
Integrating the differential inequality gives
\[ |\delta(t)|^2 \leq |\delta(t_k)|^2 e^{\beta(t-t_k)}. \]  \( (4.4) \)

Note if \( F < \frac{1}{2\sqrt{2}J} \) then \( \beta = 2(2K^{\frac{1}{2}} - 1) < 0 \) and the subsequent analysis may be significantly simplified. Thus we assume in what follows that \( F \geq \frac{1}{2\sqrt{2}J} \) so that \( \beta \geq 0 \).

Thus Lemma 4.1.1 gives an estimate on the growth of the error in the forecast step. Our aim now is to show that this growth can be controlled by observing \( Pv \), provided \( h \) is sufficiently small. To ease the notation of the proofs we introduce three functions that will be used in this section and in the following one. Namely we define, for \( t > 0 \),
\[ A_1(t) := \frac{16K}{\beta} (e^{\beta t} - 1) + \frac{4R_0^2}{2\beta} (e^{2\beta t} - 1), \]  \( (4.5) \)
\[ B_1(t) := \frac{16c^2K^2}{\beta} \left[ \frac{e^{\beta t} - e^{-t}}{\beta + 1} - (1 - e^{-t}) \right] + e^{-t} + \frac{4c^2KR_0^2}{2\beta} \left[ \frac{e^{2\beta t} - e^{-t}}{2\beta + 1} - (1 - e^{-t}) \right], \]  \( (4.6) \)
and
\[ B_2(t) := c^2K \{ 1 - e^{-t} \}. \]  \( (4.7) \)

Here and in what follows \( c, \beta \) and \( K \) are as in Property 2.2.2, Lemma 4.1.1 and Proposition 2.1.2. In order to carry out our theoretical developments two different norms in \( \mathbb{R}^J \) will be used. In each case the constant \( R_0 > 0 \) quantifies the size of the initial error \( \delta(0) \), measured in the relevant norm for the result at hand. Finally we denote \( \delta_k := \delta(t_k) \).

**Theorem 4.1.2.** Let \( v \) be a solution of the equation (3.1) with \( v(0) \in B \). Then there exists \( h^* > 0 \) such that for any \( h \in (0, h^*) \) the approximating solution \( m \) given by (4.1) converges to \( v \) as \( t \to \infty \).

**Proof.** Notice that \( B_1(0) = 1 \) and \( B_1'(0) = -1 \), so that there is \( h^* > 0 \) with the property that \( B_1(h) \in (0, 1) \) for all \( h \in (0, h^*) \). Fix any such assimilation time \( h \) and
Partially Observed Lorenz ’96

\[ \gamma = B_1(h) \in (0,1). \] Let \( R_0 := |\delta_0| \). We show by induction that, for every \( k \), \( |\delta_k|^2 \leq \gamma^k R_0^2 \). We suppose that it is true for \( k \) and we prove it for \( k + 1 \).

Taking the inner product of \( P\delta \) with the equation (4.2) gives

\[
\frac{1}{2} \frac{d|P\delta|^2}{dt} + |P\delta|^2 + 2\langle B(v, \delta), P\delta \rangle + \langle B(\delta, \delta), P\delta \rangle = 0
\]

so that, by Property 2.1.1, item (4),

\[
\frac{1}{2} \frac{d|P\delta|^2}{dt} + |P\delta|^2 \leq 4|v||\delta||P\delta| + 2|\delta|^2|P\delta|.
\]

By the inductive hypothesis we have \(|\delta_k|^2 \leq R_0^2\) since \( \gamma \in (0,1) \). Shifting the time origin by setting \( \tau := t - t_k \) and using Lemma 4.1.1 gives

\[
\frac{1}{2} \frac{d|P\delta|^2}{d\tau} + |P\delta|^2 \leq 4K^\frac{1}{2}|\delta||P\delta| + 2|\delta_k|e^{\beta \tau}|\delta||P\delta|
\]

\[
\leq 4K^\frac{1}{2}|\delta||P\delta| + 2R_0e^{\frac{\beta}{2}}|\delta||P\delta|.
\]

(4.8)

Applying Young’s inequality to each term on the right-hand side we obtain

\[
\frac{d|P\delta|^2}{d\tau} \leq 16K|\delta|^2 + 4R_0^2e^{\beta \tau}|\delta|^2.
\]

(4.9)

Integrating from 0 to \( s \), where \( s \in (0, h) \), gives

\[
|P\delta(s)|^2 \leq A_1(s)|\delta_k|^2.
\]

(4.10)

Now again consider the equation (4.3) using Property 2.1.1(5) to obtain

\[
\frac{1}{2} \frac{d|\delta|^2}{d\tau} + |\delta|^2 - |\langle B(\delta, \delta), v \rangle| \leq 0.
\]

Using Property 2.2.2 and Young’s inequality yields

\[
\frac{1}{2} \frac{d|\delta|^2}{d\tau} + |\delta|^2 \leq c|\delta||P\delta|
\]

\[
\leq cK^\frac{1}{2}|\delta||P\delta|
\]

\[
\leq \frac{|\delta|^2}{2} + \frac{c^2K}{2} |P\delta|^2.
\]

(4.11)

Employing the bound (4.10) then gives

\[
\frac{d|\delta|^2}{d\tau} + |\delta|^2 \leq \left( \frac{16c^2K^2}{\beta} (e^{\beta \tau} - 1) + \frac{4c^2KR_0^2}{2\beta} (e^{2\beta \tau} - 1) \right) |\delta_k|^2.
\]

Therefore, upon using Gronwall’s lemma,

\[
|\delta(s)|^2 \leq B_1(s)|\delta_k|^2.
\]

It follows that

\[
|\delta_{k+1}|^2 \leq \gamma|\delta_k|^2 \leq \gamma^{k+1} R_0^2,
\]

and the induction (and hence the proof) is complete.
4.2. Noisy Observations: Discrete 3DVAR

Now we consider the situation where the data is noisy and $H_k = P$. We employ the 3DVAR filter which results from the minimization principle (2.12) in the case where $\hat{C}_{k+1} = \sigma^2 I$ and $\Gamma = \epsilon^2 I$. Recall the true signal is determined by the equation (2.5) and the observed data by the equation (2.6), now written in terms of the true signal $v_k = v(t_k)$ solving the equation (2.3) with $v_0 \in B$. Thus

$$v_{k+1} = \Psi(v_k), \quad v_0 \in B$$
$$y_{k+1} = Pv_{k+1} + \nu_{k+1}.$$

If we define $\eta := \frac{\epsilon^2}{\sigma^2}$ then the filter, given in (2.13), can be written as

$$m_{k+1} = \left( \frac{\eta}{1 + \eta} P + Q \right) \Psi(m_k) + \frac{1}{1 + \eta} y_{k+1},$$

after noting that $Py_{k+1} = y_{k+1}$ because $P$ is a projection and $\nu_{k+1}$ is assumed to lie in the image of $P$. In fact the data has the following form:

$$y_{k+1} = Pv_{k+1} + P\nu_{k+1}$$
$$= P\Psi(v_k) + \nu_{k+1}.$$

Combining the two equations gives

$$m_{k+1} = \left( \frac{\eta}{1 + \eta} P + Q \right) \Psi(m_k) + \frac{1}{1 + \eta} (P\Psi(v_k) + \nu_{k+1}). \tag{4.12}$$

We can also write the equation for the true solution $v_k$, given by (2.5), in the following form:

$$v_{k+1} = \left( \frac{\eta}{1 + \eta} P + Q \right) \Psi(v_k) + \frac{1}{1 + \eta} P\Psi(v_k). \tag{4.13}$$

Note that $v_k = v(t_k)$ where $v(\cdot)$ solves (3.1).

We are interested in comparing $m_k$, the output of the filter, with $v_k$ the true signal which underlies the data. We define the error process $\delta(t)$ as follows:

$$\delta(t) = \begin{cases} 
\delta_k := m_k - v(t) & \text{if } t = t_k \\
\Psi(m_k, t - t_k) - v(t) & \text{if } t \in (t_k, t_{k+1}).
\end{cases}$$

Observe that $\delta$ is discontinuous at times $t_k$ which are multiples of $h$, since $m_{k+1} \neq \Psi(m_k; h)$. Subtracting (4.13) from (4.12) we obtain

$$\delta_{k+1} = \delta(t_{k+1}) = \left( \frac{\eta}{1 + \eta} P + Q \right) \delta(t_{k+1}) + \frac{1}{1 + \eta} \nu_k. \tag{4.14}$$

In what follows we will assume that the $\nu_k$ are independent random variables that satisfy the bound $|\nu_k| \leq \epsilon$, thereby linking the scale of the covariance $\Gamma$ employed in 3DVAR to the size of the noise. Before stating the next theorem we introduce a norm $\| \cdot \|$ defined by $\|z\| := |z| + |Pz|$, $z \in \mathbb{R}^J$. 

Theorem 4.2.1. Let \( v \) be the solution of the equation (3.1) with \( v(0) \in B \). Assume that \( \{\nu_k\}_{k \geq 1} \) is a sequence of independent bounded random variables such that, for every \( k \), \( |\nu_k| \leq \epsilon \). Then there are choices (detailed in the proof) of assimilation step \( h > 0 \) and parameter \( \eta > 0 \) such that, for some \( \alpha \in (0, 1) \) and provided that the noise \( \epsilon > 0 \) is small enough, the error satisfies

\[
\| \delta_{k+1} \| \leq \alpha \| \delta_k \| + 2\epsilon. \tag{4.15}
\]

Thus, there is \( a > 0 \) such that

\[
\limsup_{k \to \infty} \| \delta_k \| \leq a \epsilon.
\]

Proof. Let \( A_1(\cdot), B_1(\cdot) \) and \( B_2(\cdot) \) be as in (4.5, 4.6, 4.7), and set

\[
M_1(t) := \frac{2\eta}{1 + \eta} \sqrt{A_1(t) + B_1(t)},
\]

\[
M_2(t) := \frac{2\eta}{1 + \eta} \sqrt{B_2(t)}.
\]

Since \( A_1(0) = 0, B_1(0) = 1, B_2(0) = 0 \) and

\[
\frac{d}{dt} \sqrt{B_1(t)} \bigg|_{t=0} = -1/2 < 0
\]

it is possible to find \( h, \eta > 0 \) small such that

\[
M_2(h) < M_1(h) =: \alpha < 1.
\]

Let \( R_0 = \| \delta_0 \| \). We show by induction that for such \( h \) and \( \eta \), and provided that \( \epsilon \) is small enough so that

\[
\alpha R_0 + 2\epsilon < R_0,
\]

we have that \( \| \delta_k \| \leq R_0 \) for all \( k \). Suppose for induction that it is true for \( k \). Then \( |\delta_k| \leq \| \delta_k \| \leq R_0 \) and we can apply (after shifting time as before) Lemma [4.2.2] below to obtain that

\[
|P\delta(t)| \leq \sqrt{A_1(t)}|\delta_k|^2 + |P\delta_k|^2 \leq \sqrt{A_1(t)}|\delta_k| + |P\delta_k|
\]

and

\[
|\delta(t)| \leq \sqrt{B_1(t)}|\delta_k|^2 + B_2(t)|P\delta_k|^2 \leq \sqrt{B_1(t)}|\delta_k| + \sqrt{B_2(t)}|P\delta_k|.
\]

Therefore,

\[
|P\delta_{k+1}| + |\delta_{k+1}| \leq \left( \frac{2\eta}{1 + \eta} \sqrt{A_1(h) + B_1(h)} \right) |\delta_k| + \left( \frac{2\eta}{1 + \eta} \sqrt{B_2(h)} \right) |P\delta_k| + 2\epsilon = M_1(h)|\delta_k| + M_2(h)|P\delta_k| + 2\epsilon.
\]
Since $M_2(h) < M_1(h) = \alpha$ we deduce that
\[
\|\delta_{k+1}\| \leq \alpha \|\delta_k\| + 2\epsilon,
\]
which proves (4.15). Furthermore, the induction is complete, since
\[
\|\delta_{k+1}\| \leq \alpha \|\delta_k\| + 2\epsilon \leq \alpha R_0 + 2\epsilon \leq R_0.
\]

\[\square\]

**Lemma 4.2.2.** In the setting of Theorem 4.2.1, for $t \in [0, h)$ and $R_0 := \|\delta_0\|$ we have
\[
|P\delta(t)|^2 \leq A_1(t)|\delta_0|^2 + |P\delta_0|^2
\]
and
\[
|\delta(t)|^2 \leq B_1(t)|\delta_0|^2 + B_2(t)|P\delta_0|^2,
\]
where $A_1, A_2$ and $B_1$ are given by (4.5, 4.6, 4.7).

Proof. As in equation (4.9) we have
\[
\frac{d|P\delta|^2}{dt} \leq 16K|\delta|^2 + 4R_0^2e^{\beta t}|\delta|^2.
\]
On integrating from 0 to $t$ as before, and noting that now $P\delta_0 \neq 0$ in general, we obtain
\[
|P\delta(t)|^2 \leq \left(\frac{16K}{\beta} \{e^{\beta t} - 1\} + \frac{4R_0^2}{2\beta} \{e^{2\beta t} - 1\}\right)|\delta_0|^2 + |P\delta_0|^2,
\]
which proves (4.16).

For the second inequality recall the bound (4.11)
\[
\frac{1}{2} \frac{d|\delta|^2}{dt} + |\delta|^2 \leq \frac{|\delta|^2}{2} + \frac{c^2K}{2}|P\delta|^2,
\]
and combine it with (4.16) to get
\[
\frac{d|\delta|^2}{dt} + |\delta|^2 \leq \left(\frac{16c^2K^2}{\beta} \{e^{\beta t} - 1\} + \frac{4c^2KR_0^2}{2\beta} \{e^{2\beta t} - 1\}\right)|\delta_0|^2 + c^2K|P\delta_0|^2.
\]
Applying Gronwall’s inequality yields (4.17).

\[\square\]
5. Numerical Results

The previous two sections demonstrate that, for fixed observation operator $P$, the noiseless data filters of Olson and Titi, and the 3DVAR filter from atmospheric sciences applications, can, asymptotically in time, overcome an incorrect initialization of the (chaotic) Lorenz '96 model and accurately recover the signal. This theory requires observation of $2/3$ of the signal vector at each assimilation time. In this section we explore how carefully chosen observation operators can significantly decrease the fraction of points which need to be observed. We make a small shift of notation and now consider the observation operator $H_k$ as a linear mapping from $\mathbb{R}^J$ into $\mathbb{R}^M$, rather than as a linear operator from $\mathbb{R}^J$ into itself, with rank $M$; the latter perspective was advantageous for the presentation of the analysis, but differs from the former which is sometimes computationally advantageous and more widely used for the description of algorithms. Recall the minimization principle (2.12), noting that now the first norm is in $\mathbb{R}^J$ and the second in $\mathbb{R}^M$.

5.1. 3DVAR

Here we consider the minimization principle (2.12) with the choice $\hat{C}_{k+1} = C_0 \in \mathbb{R}^{J \times J}$, a strictly positive-definite matrix, for all $k$. Assuming that $\Gamma \in \mathbb{R}^{M \times M}$ is also strictly positive-definite, the filter may be written as

$$m_{k+1} = \Psi(m_k) + C_0 H^T_{k+1} (H_{k+1} C_0 H^T_{k+1} + \Gamma)^{-1} (y_{k+1} - H_{k+1} \Psi(m_k))$$

(5.1)

As well as using the choice of $H_k$ defined in (2.10), we also employ the fixed observation operator where $H_k = H$, including the choice $H = P$ analyzed in the previous section. In the last case $J = 3J'$, $M = 2J'$ and $P$ is realized as a $2J' \times 3J'$ matrix.

In section 5 we make the choices $C_0 = \sigma^2 I_{J \times J}$, $\Gamma = \epsilon^2 I_{M \times M}$ and define $\eta = \epsilon^2/\sigma^2$. Throughout our experiments we take $h = 0.1$, $\epsilon^2 = 0.01$ and fix the parameter $\eta = 0.01$ (i.e. $\sigma = 1$). We use the Lorenz '96 model (2.1) to define $\Psi$, with the parameter choices $F = 8$ and $J = 60$. The system then has 19 positive Lyapunov exponents. The observational noise is i.i.d Gaussian with respect to time index $k$, with distribution $\nu_1 \sim N(0, \epsilon^2)$. Figures 5.1 and 5.2 exhibit, for fixed observation 3DVAR and adaptive observation 3DVAR, the root mean square error (RMSE), over spatial degrees of freedom, between the truth $v_k$ and the estimate $m_k$, of a single trajectory for a single observation process realization. This is then averaged over time, ignoring initial transients, and again averaged for $10^4$ realizations of the observational noise (with the identical truth realization).

The Figure 5.1 shows the RMSE for fixed observation operator where the observed space is of dimension 60, 40 and 24 respectively; the observation operator with $M = 40$ is defined as in the equation (2.7), as analyzed in section 4. For both $M = 60$ and
$M = 40$ the error decreases rapidly and the approximate solution converges to a
neighbourhood of the true solution where the size of the neighbourhood depends upon
the variance of the observational noise. The root mean square error, averaged over
the trajectory, after ignoring the initial transients, is $1.27 \times 10^{-2}$ when $M = 60$ and
$1.11 \times 10^{-2}$ when $M = 40$; note that this is on the scale of the observational noise.
The rate of convergence of the approximate solution to the true solution in the case
of partial observations is lower than the rate of convergence when full observations
are used. We also consider the case when $24 = 40\%$ of the modes are observed;
this percentage is motivated by the work reported in [1] where it was demonstrated
that observing $40\%$ of the modes, with the observation directions chosen carefully
and with observations sufficiently frequent in time, is sufficient for the approximate
solution to converge to the true underlying solution. The Figure 5.1c shows that, in
our observational set-up, observing 24 of the modes only allows marginal reconstruction
of the signal, asymptotically in time; the root mean square error makes regular large
excursions to $O(1)$ and the time-averaged root mean square error over the trajectory is
$(6.69 \times 10^{-2})$, which is larger than for 40 or 60 observations.

The Figure 5.1 shows the RMSE for adaptive observation 3DVAR. In this case we
notice that the error is consistently small, uniformly in time, with just 9 or more modes
observed. When $M = 9$ (15\% observed modes) the root mean square error averaged
over the trajectory is $1.3794 \times 10^{-2}$ which again is of the order of the observational
noise variance. For $M \geq 9$ the error is similar – see Figure 5.2b. On the other hand,
for smaller values of $M$ the error is not controlled as shown in Figure 5.2a, where the
root mean square error for $M = 7$ is compared with that for $M = 9$. It is noteworthy
that the number of observations sufficient for accurate reconstruction is approximately
half the number of positive Lyapunov exponents.

5.2. Extended Kalman Filter

In the Extended Kalman Filter (ExKF) the approximate solution evolves according to
the minimization principle (2.12) with $C_k$ chosen as a covariance matrix evolving in
the forecast step according to the linearized dynamics, and in the assimilation stage
updated according to Bayes’ rule based on a Gaussian observational error covariance.
This gives the method

\[ m_{k+1} = \Psi(m_k) + \hat{C}_{k+1} H^T_{k+1} (H_{k+1} \hat{C}_{k+1} H^T_{k+1} + \Gamma)^{-1} (y_{k+1} - H_{k+1} \Psi(m_k)) \]
\[ \hat{C}_{k+1} = D \Psi(m_k) C_k D \Psi(m_k)^T, \]
\[ G_{k+1} = \hat{C}_{k+1} H^T_{k+1} (H_{k+1} \hat{C}_{k+1} H^T_{k+1} + \Gamma)^{-1}, \]
\[ C_{k+1} = (I_{J \times J} - G_{k+1} H_{k+1}) \hat{C}_{k+1}. \]
(a) Percentage of components observed = 100%. Average error over the trajectory = \(1.27 \times 10^{-2}\).

(b) Percentage of components observed = 66.67%. Average error over the trajectory = \(1.11 \times 10^{-2}\).

(c) Percentage of components observed = 40%. Average error over the trajectory = \(6.69 \times 10^{-2}\).

Figure 5.1: Fixed Observation Operator 3DVAR
(a) Comparison of RMSE between $M = 7$ and $M = 9$. Average error over the trajectory $1.9547 \times 10^{-1}$, $1.3794 \times 10^{-2}$ respectively.

(b) Averaged RMSE for different choices of $M$

Figure 5.2: Adaptive Observation 3DVAR
Note that
\[ m_{k+1} = \Psi(m_k) + G_{k+1}(y_{k+1} - H_{k+1}\Psi(m_k)). \] (5.6)

We first consider the ExKF scheme with a fixed observation operator \( H_k = H \). We make two choices for \( H \): the full rank identity operator and a partial observation operator with 40\% of the modes observed. For the first case the filtering scheme is the standard ExKF with all the modes being observed. The approximate solution converges to the true solution and the error decreases rapidly as can be seen in the Figure 5.3a. The root mean square error is \( 9.1495 \times 10^{-4} \), which is an order of magnitude smaller than the analogous error for the 3DVAR algorithm when fully observed which is, recall, \( 1.27 \times 10^{-2} \). For the partial observations case with \( M = 24 \) we see that again the approximate solution converges to the true underlying solution as shown in the Figure 5.3b. Furthermore the solution given by the ExKF with \( M = 24 \) is far more robust than for 3DVAR with this number of observations. The root mean square error is also lower for ExKF (2.178 \times 10^{-3}) when compared with the 3DVAR scheme (6.69 \times 10^{-2}).

We now turn to adaptive observation within the context of the ExKF. The Figure 5.4 shows that it is possible to obtain an RMS error which is of the order of the observational error, and is robust over long time intervals, using only a 7 dimensional observation space, improving marginally on the 3DVAR situation where 9 dimensions were required to attain a similar level of accuracy.

The AUS scheme, as proposed by Trevisan and co-workers [16], is an ExKF method which operates by confining the analysis update to the subspace spanned by a finite number of directions, ideally designed to capture the instabilities in the dynamics. This is typically achieved by choosing to work in the subspace of the linear dynamics spanned by \( M \) largest growth directions where \( M \) is fixed as the number of non-negative Lyapunov exponents. Asymptotically this method with \( H_k = I_{J \times J} \) behaves similarly to the adaptive ExKF with the same \( M \). To understand the intuition behind the AUS method we plot in Figure 5.5a the rank of the covariance matrix \( C_k \) from standard ExKF based on observing 60 and 24 modes. Notice that in both cases the rank approaches a value of 19 or 20 and that 19 is the number of positive Lyapunov exponents. This means that the covariance is effectively zero in 40 of the observed dimensions and that, as a consequence of the minimization principle (2.12), data will be ignored in the 40 dimensions where the covariance is negligible. It is hence natural to simply confine the update step to the subspace of dimension 19 given by the number of positive Lyapunov exponents, right from the outset. This is what AUS does by reducing the rank of the error covariance matrix \( C_k \). Numerical results are given in Figure 5.5b which shows the root mean square error over the trajectory for the ExKF-AUS assimilation scheme with time. After initial transience the error is mostly of the numerical order of the observational noise. Occasional jumps outside this error bound are observed but the
(a) Percentage of components observed = 100%. Average error over the trajectory = 9.1495 × 10^{-4}.

(b) Percentage of components observed = 40%. The average error over the trajectory = 2.178 × 10^{-3}.

Figure 5.3: Fixed Observation ExKF. The zoomed in figures shows the variability in RMSE between time $t = 20$ and $t = 70$. 
(a) Comparison of RMSE between $M = 5$ and $M = 7$. Average error over the trajectory $5.2973 \times 10^{-1}$, $2.7514 \times 10^{-2}$ respectively.

(b) Averaged RMSE for different choices of $M$.

Figure 5.4: Adaptive Observation ExKF
approximate solution converges to the true solution each time. The root mean square error for ExKF-AUS is $1.075 \times 10^{-2}$. However, if the rank of the error covariance matrix $C_0$ in AUS is chosen to be less than the number of unstable modes for the underlying system, then the approximate solution does not converge to the true solution.

6. Conclusions

In this paper we have studied the long-time behaviour of filters for partially observed dissipative dynamical systems, using the Lorenz ’96 model as a canonical example. We have highlighted the connection to synchronization in dynamical systems, and shown that this synchronization theory, which applies to noise-free data, is robust to the addition of noise, in both the continuous and discrete time settings. In so doing we are studying the 3DVAR algorithm. In the context of the Lorenz ’96 model we have identified a fixed observation operator, based on observing $2/3$ of the components of the signal’s vector, which is sufficient to ensure desirable long-time properties of the filter. However it is important to realize that it is to be expected that, within the context of fixed observation operators, considerably fewer observations may be needed to ensure such desirable properties. Ideas from nonlinear control theory will be relevant in addressing this issue. We also studied adaptive observation operators, targeted to observe the directions of maximal growth within the local linearized dynamics. We demonstrated that with these adaptive observers, considerably fewer observations are required. We also made a connection between these adaptive observation operators, and the AUS methodology which is also based on the local linearized dynamics, but works by projecting within the model covariance operators of ExKF, whilst the observation operators themselves are fixed; thus the model covariances are adapted. Our method and the AUS exhibit very similar behaviour in terms of the dimension of the observation operator/model covariance operator required to get desirable long-time properties of the filter.

Acknowledgments

AbS and DSA are supported by the EPSRC-MASDOC graduate training scheme. AMS is supported by EPSRC, ERC and ONR. KJHL is supported by King Abdullah University of Science and Technology, and is a member of the KAUST SRI-UQ Center.
(a) Standard ExKF with 60 and 24 observed modes. The rank of the error covariance matrix $C_k$ decays to (approximately) the number of unstable Lyapunov modes in the underlying system, namely 19.

(b) Average error over the trajectory = $1.075 \times 10^{-2}$. The zoomed in figures shows the variability in RMSE between time $t = 20$ and $t = 70$.

Figure 5.5: Rank of error covariance and ExKF-Assimilation in Unstable Space
[1] H.D.I. Abarbanel. *Predicting the Future: Completing Models of Observed Complex Systems*. Springer. Series: Understanding Complex Systems, 2013.

[2] A. Azouani, E. Olson, and E.S. Titi. Continuous data assimilation using general interpolant observables. *Journal of Nonlinear Science*, 24:277–304, 2014.

[3] A. Bennett. *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, 2003.

[4] D. Bloemker, K.J.H. Law, A.M. Stuart, and K.C. Zygalakis. Accuracy and stability of the continuous-time 3DVAR filter for the navier-stokes equation. *Nonlinearity*, 2014.

[5] C.E.A. Brett, K.F. Lam, K.J.H. Law, D.S. McCormick, M.R. Scott, and A.M. Stuart. Accuracy and stability of filters for dissipative pdes. *PhysicaD: Nonlinear Phenomena*, 2013.

[6] K. Hayden, E. Olson, and E.S. Titi. Discrete data assimilation in the Lorenz and 2d Navier-Stokes equations. *Physica D: Nonlinear Phenomena*, pages 1416–1425, 2011.

[7] E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 2003.

[8] K.J.H. Law, A. Shukla, and A.M. Stuart. Analysis of the 3dvar filter for the partially observed lorenz ’63 model. *Discrete and Continuous Dynamical Systems A*, 34:1061–1078, 2014.

[9] K.J.H. Law, A.M. Stuart, and K.C. Zygalakis. *Data Assimilation: A Mathematical Introduction*. Lecture Notes, 2014.

[10] E.N. Lorenz and K.A. Emanuel. Optimal sites for supplementary weather observations: Simulation with a small model. *Journal of the Atmospheric Sciences*, 55:399–414, 1998.

[11] A. Majda and J. Harlim. *Filtering Complex Turbulent Systems*. Cambridge University Press, 2012.

[12] D. Oliver, A. Reynolds, and N. Liu. *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press, 2008.

[13] E. Olson and E. Titi. Determining modes for continuous data assimilation in 2d turbulence. *Journal of Statistical Physics*, 113:799–840, 2003.

[14] T. Tarn and Y. Rasis. Observers for nonlinear stochastic systems. *Automatic Control, IEEE Transactions*, 21(4):441–488, 1976.

[15] R. Temam. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, volume 68 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1997.

[16] A. Trevisan and F. Uboldi. Assimilation of standard and targeted observations within the unstable subspace of the observation analysis forecast cycle system. *Journal of the Atmospheric Sciences*, 61(1):103–113, 2004.