



# Application of a local attractor dimension to reduced space strongly coupled data assimilation for chaotic multiscale systems

Courtney Quinn, Terence J. O’Kane, and Vassili Kitsios

Decadal Climate Forecasting Project, CSIRO Oceans and Atmosphere, Hobart, TAS, Australia

**Correspondence:** Courtney Quinn (courtney.quinn@csiro.au)

Received: 1 October 2019 – Discussion started: 10 October 2019

Revised: 20 December 2019 – Accepted: 17 January 2020 – Published: 19 February 2020

**Abstract.** The basis and challenge of strongly coupled data assimilation (CDA) is the accurate representation of cross-domain covariances between various coupled subsystems with disparate spatio-temporal scales, where often one or more subsystems are unobserved. In this study, we explore strong CDA using ensemble Kalman filtering methods applied to a conceptual multiscale chaotic model consisting of three coupled Lorenz attractors. We introduce the use of the local attractor dimension (i.e. the Kaplan–Yorke dimension,  $\dim_{KY}$ ) to prescribe the rank of the background covariance matrix which we construct using a variable number of weighted covariant Lyapunov vectors (CLVs). Specifically, we consider the ability to track the nonlinear trajectory of each of the subsystems with different variants of sparse observations, relying only on the cross-domain covariance to determine an accurate analysis for tracking the trajectory of the unobserved subdomain. We find that spanning the global unstable and neutral subspaces is not sufficient at times where the nonlinear dynamics and intermittent linear error growth along a stable direction combine. At such times a subset of the local stable subspace is also needed to be represented in the ensemble. In this regard the local  $\dim_{KY}$  provides an accurate estimate of the required rank. Additionally, we show that spanning the full space does not improve performance significantly relative to spanning only the subspace determined by the local dimension. Where weak coupling between subsystems leads to covariance collapse in one or more of the unobserved subsystems, we apply a novel modified Kalman gain where the background covariances are scaled by their Frobenius norm. This modified gain increases the magnitude of the innovations and the effective dimension of the unobserved domains relative to the strength of the cou-

pling and timescale separation. We conclude with a discussion on the implications for higher-dimensional systems.

## 1 Introduction

As the world of climate modelling has moved towards coupled Earth system models which combine ocean, atmosphere, sea-ice, and biogeochemical effects, it is essential to understand how the respective domains of disparate spatio-temporal scales covary and influence each other. In the context of state estimation, strongly coupled data assimilation (CDA) in multiscale systems allows the quantification of cross-domain dynamics. Specifically, strong CDA uses the cross-covariances amongst all components to influence the analysis increment, meaning that unobserved subsystems are directly adjusted in the analysis step regardless of observation set and coupling strengths (Laloyaux et al., 2016; O’Kane et al., 2019). Strong CDA has shown the potential to outperform weakly coupled or uncoupled approaches in intermediate complexity atmosphere–ocean systems (Sluka et al., 2016; Penny et al., 2019); however, CDA has the additional requirements of increased ensemble sizes (Han et al., 2013) and well-observed state (prognostic) variables (Kang et al., 2011). Ensemble DA methods rely on properly sampling the variance of the model, implying that very large ensemble sizes are needed for high-dimensional systems. In practical implementation this is often not possible due to computational costs and limitations. It is therefore necessary to investigate and develop new methods of accurately representing error growth in multiscale systems.

One approach to reduce the requirement to adequately sample the background covariances is to perform CDA in the

reduced subspace of the unstable modes, known as assimilation in the unstable subspace (AUS)<sup>1</sup> (Trevisan and Uboldi, 2004; Carrassi et al., 2007; Trevisan et al., 2010; Trevisan and Palatella, 2011; Palatella et al., 2013; Bocquet and Carrassi, 2017). The concept behind AUS methods is that the analysis increment should lie in the unstable and neutral subspaces of the model, therefore retaining the spatial structure of dominant instabilities (Trevisan and Uboldi, 2004). Many implementations of AUS involve variational data assimilation methods, namely finding the model trajectory which best fits observations by solving an optimization cost function. On the other hand, the widely used ensemble Kalman filtering methods have been shown to best capture the unstable error growth in nonlinear systems (Evensen, 1997). For this reason, we focus on the ability to accurately represent the unstable and neutral subspace within the ensemble Kalman filtering framework.

The main motivation for this study comes from the conjecture that when applying ensemble Kalman filtering methods to high-dimensional nonlinear systems, the true time-dependent error covariance matrix collapses onto a subspace of the model domain which contains unstable, neutral, and sometimes weakly stable modes. While previous results prove the collapse of the error covariance matrix onto the unstable and neutral subspace for linear systems (Gurumoorthy et al., 2017; Bocquet et al., 2017), nonlinear systems have the additional complication that the unstable subspace is a function of the underlying trajectory and is not globally defined (Bocquet et al., 2017). As nonlinearity increases, short-term dynamics can cause some stable modes of the linearized system to experience significant growth. These additional modes are therefore important when considering local error growth in ensemble DA methods. While this has previously been addressed in the context of perfect nonlinear models (Ng et al., 2011), recent studies have shown that the additional transient unstable modes can also further amplify perturbations in the presence of model error (Grudzien et al., 2018a, b).

Such transient error growth has previously been explored in ocean–atmosphere models of varying complexity. One way of quantifying local error growth is through finite-time Lyapunov exponents (FTLEs), i.e. rates of expansion and contraction over finite intervals of time. Nese and Dutton (1993) utilized the largest (leading) FTLE to quantify predictability times along different parts of the attractor of a low-order ocean–atmosphere model. The statistical properties of FTLEs have been studied more recently in a range of atmosphere and ocean models with varying complexity, includ-

ing low-order models (Vannitsem, 2017) and intermediate-complexity models (Vannitsem, 2017; De Cruz et al., 2018; Kwasiok, 2019). While FTLEs provide local rates of error growth, one can also consider directions of local error growth. Early work in this area considered directed finite-time normal modes, which are generalized as the eigenvectors of tangent linear propagators over a given time interval, and studied their relation to blocking events in the atmosphere (Frederiksen, 1997, 2000; Wei and Frederiksen, 2005). More recent studies have focused on covariant Lyapunov vectors (CLVs) which give directions of asymptotic growth and decay in the tangent linear space. While these vectors have an average growth rate corresponding to the asymptotic Lyapunov exponents, their finite-time behaviour may differ. This finite-time behaviour of CLVs has been analysed across a range of quasi-geostrophic atmosphere (Schubert and Lucarini, 2015, 2016; Gritsun and Lucarini, 2017) and coupled atmosphere–ocean (Vannitsem and Lucarini, 2016) models.

In this study we utilize FTLEs and CLVs within the ensemble Kalman filtering framework applied to a low-dimensional chaotic model with spatio-temporal scale separation. The model was designed to represent the interaction between the ocean, extratropical atmosphere, and an ocean–atmosphere interface (referred to as the tropical atmosphere). The idea is that the ocean and extratropical atmosphere are only implicitly coupled through the interface, with the interface being strongly coupled to the ocean and weakly coupled to the extratropical atmosphere. We consider the performance of strong CDA on this paradigm model with different subsets of observations. We introduce the use of a varying number of CLVs to form the forecast error covariance matrix. The idea of AUS is incorporated through the use of the time-varying subspace defined by the local attractor dimension. The dimension is calculated through FTLEs and the error covariance matrix is constructed via a projection of the ensemble members onto a corresponding subset of the CLVs at each analysis step. We compare full-rank ensemble transform and square-root filters to “phase space” variants whose background covariances are defined in terms of either the finite-time or asymptotic attractor dimension. Another variant considered includes only the unstable, neutral, and weakest stable CLVs. We consider benchmark calculations compared to the recent work of Yoshida and Kalnay (2018) and then a comprehensive set of experiments where the various domains are partially or even completely unobserved.

The paper is organized as follows. Section 2 introduces the paradigm model and discusses the dynamical properties of a control simulation. Section 3 describes the method for calculating the CLVs and discusses the possibility of CLV alignment. Section 4 summarizes the Kalman filtering method and introduces the modification to the calculation of the error covariance matrix. The configurations and results of the DA experiments are presented in Sect. 5. The implications of the results of this study and future endeavours are discussed in Sect. 6.

<sup>1</sup>There is a computational cost accompanying the determination of the unstable and neutral subspaces which may or may not be less than the cost of adequately sampling the model covariances. We do not attempt to comment on numerical efficiency of AUS methods applied to high-dimensional systems. The exploration of cost-effective methods for determining the reduced subspace is left for future study.

## 2 Coupled Lorenz model for global circulation

### 2.1 Peña and Kalnay model

This section describes a low-dimensional chaotic model representing a coupled ocean and atmosphere. It is a system of three coupled Lorenz '63 models introduced by Peña and Kalnay (2004) to study unstable modes with a timescale separation. This model has previously been described in modified form by Norwood et al. (2013), who used it to investigate the instability properties of related dynamical vectors, i.e. bred vectors (BVs), singular vectors (SVs), and CLVs, and by Yoshida and Kalnay (2018) and O'Kane et al. (2019) in the context of strongly coupled data assimilation. The model is given as follows:

$$\dot{x}_e = \sigma(y_e - x_e) - c_e(Sx_t + k_1), \quad (1a)$$

$$\dot{y}_e = \rho x_e - y_e - x_e z_e + c_e(Sy_t + k_1), \quad (1b)$$

$$\dot{z}_e = x_e y_e - \beta z_e, \quad (1c)$$

$$\dot{x}_t = \sigma(y_t - x_t) - c(SX + k_2) - c_e(Sx_e + k_1), \quad (1d)$$

$$\dot{y}_t = \rho x_t - y_t - x_t z_t + c(SY + k_2) + c_e(Sy_e + k_1), \quad (1e)$$

$$\dot{z}_t = x_t y_t - \beta z_t + c_z Z, \quad (1f)$$

$$\dot{X} = \tau \sigma(Y - X) - c(x_t + k_2), \quad (1g)$$

$$\dot{Y} = \tau \rho X - \tau Y - \tau SXZ + c(y_t + k_2), \quad (1h)$$

$$\dot{Z} = \tau SXY - \tau \beta Z - c_z z_t. \quad (1i)$$

The model is proposed as representing the fast extratropical atmosphere (Eq. 1a–c), fast tropical atmosphere (Eq. 1d–f), and slow tropical ocean (Eq. 1g–i). The standard Lorenz parameter values of  $\sigma = 10$ ,  $\rho = 28$ , and  $\beta = 8/3$  are used. The coupling coefficients are  $c_e = 0.08$  and  $c = c_z = 1$ , representing a weak coupling of the extratropical to tropical atmosphere and a strong coupling of the tropical atmosphere and ocean. The scaling parameters are set as  $\tau = 0.1$  and  $S = 1$ , giving an explicit timescale separation (note that there is still a difference in the spatial scales through the dynamics). The parameters  $k_1 = 10$  and  $k_2 = -11$  are chosen as uncentering parameters.

More generally, these choices lead to a tropical subsystem that is dominated by changes in the ocean subsystem, but has an extratropical atmosphere, representative of weather noise, whose behaviour is similar to the original Lorenz model in that the system exhibits chaotic behaviour with two distinct regimes observed in the  $x_e$  and  $y_e$  coordinates of the extratropical atmosphere. The ocean exhibits significant deviations from its normal trajectory about every 2–8 years, where 3 years corresponds to approximately 10 time units. Figure 1 shows typical trajectories of the  $x_e$ ,  $x_t$ , and  $X$  components of the extratropical, tropical, and ocean subsystems respectively. Figure 2 shows the approximate phase space structure of each of the respective subsystem's attractors.

### 2.2 Lyapunov exponents and dimension

We are interested in analysing both the short-term and asymptotic dynamics of the system (1). We start by considering the asymptotic behaviour of trajectories, which can be understood through the Lyapunov exponents. Chaotic systems are characterized by one or more positive Lyapunov exponents (Benettin et al., 1976; Sano and Sawada, 1985), and the underlying attractor dimension can be related to the values of the Lyapunov exponents (Young, 1982; Eckmann and Ruelle, 1985).

We compute the Lyapunov exponents using a QR decomposition method (see e.g. Dieci et al., 1997). For the computation we run the model for 1000 time units to ensure convergence onto the attractor. We use a time step of 0.01 and compute the Lyapunov exponents from the last 500 time units, with an orthonormalization time step of 0.25 for the QR method. The convergence of the Lyapunov exponents is shown in Fig. 3. We observe that for these parameter values there are two unstable and two approximately neutral Lyapunov exponents. We list the values of all nine computed asymptotic exponents in Table 1.

Given the approximated asymptotic Lyapunov exponents in Table 1, we can compute the global attractor dimension, i.e. the number of active degrees of freedom. Note that there is a large spectral gap between the seventh and eighth Lyapunov exponents. This gives evidence that the effective dimension of the attractor will be less than eight. Throughout this study we will use the Kaplan–Yorke dimension (Kaplan and Yorke, 1979; Frederickson et al., 1983) to calculate the upper bound on the attractor dimension. It is defined as follows:

$$\dim_{\text{KY}} := j + \frac{\sum_{i=1}^j \lambda_i}{|\lambda_{j+1}|}, \quad (2)$$

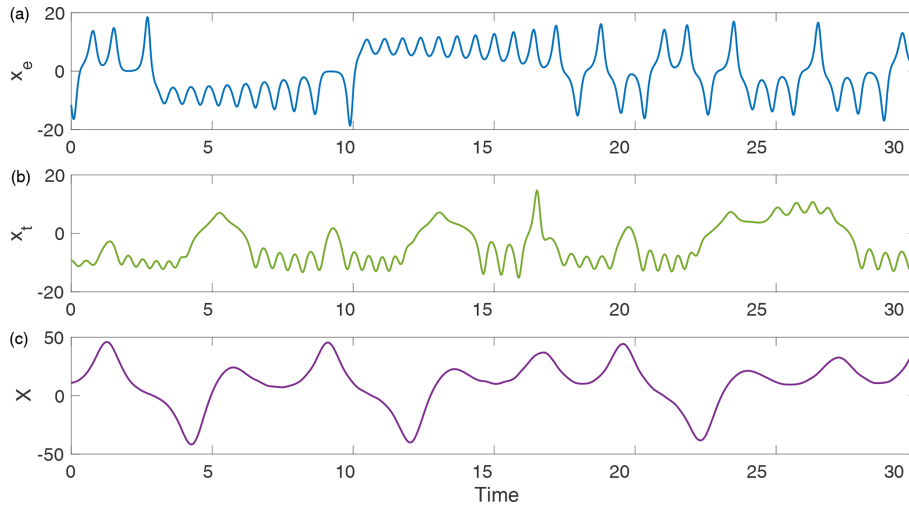
where  $j$  is the largest integer such that  $\sum_{i=1}^j \lambda_i \geq 0$  and

$\sum_{i=1}^{j+1} \lambda_i < 0$ . In addition we calculate the Kolmogorov–Sinai entropy  $\text{ent}_{\text{KS}}$  as a measure of the diversity of the trajectories generated by the dynamical system and determined through the Pesin formula:

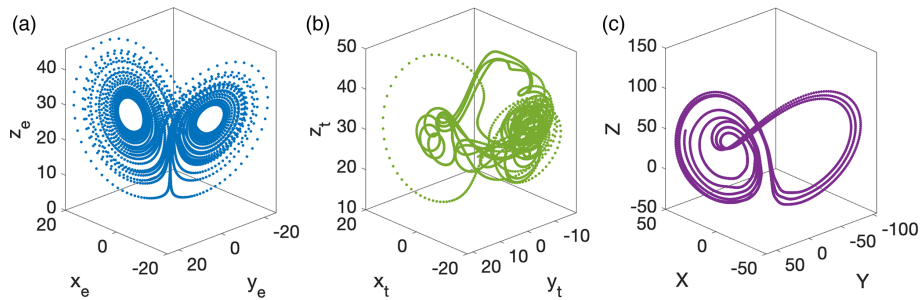
$$\text{ent}_{\text{KS}} = \sum_{\lambda > 0} \lambda_i, \quad (3)$$

which provides an upper bound to the entropy (Eckmann and Ruelle, 1985).

With the values in Table 1 we obtain a value of 5.9473 for the global attractor dimension of the nine-component system. As previously mentioned, asymptotically stable modes can experience transient periods of linear unstable growth. We therefore define local dimension through the computation of Eq. (2) using finite-time Lyapunov exponents (FTLEs).



**Figure 1.** Example trajectories of the coupled Lorenz model (Eq. 1) for  $x_e$  (a),  $x_t$  (b), and  $X$  (c).



**Figure 2.** Trajectories along attractors of the extratropical (a), tropical (b), and ocean (c) subsystems of Eq. (1).

The computation of FTLEs is similar to that of the asymptotic Lyapunov exponents, with the difference being a finite window of time over which the exponents are computed. The FTLEs then serve as a time-dependent measure of the local unstable, neutral, and stable growth rates of the evolving system (Abarbanel et al., 1991; Eckhardt and Yao, 1993; Yoden and Nomura, 1993). The temporal variability of FTLEs is highly dependent on the window size,  $\tau$ , over which they are computed. As  $\tau \rightarrow \infty$  the variability reduces and the FTLEs approach the asymptotic Lyapunov exponent values (Yoden and Nomura, 1993). We compute the FTLEs for a window size  $\tau = 4$  and the corresponding time-varying Kaplan–Yorke dimension which we will refer to as a “local” dimension. This will give a measure which reflects variations in the finite-time growth rates, with higher (lower) dimension signifying increased instability (stability) in the FTLEs. More specifically this is an upper bound on the true local dimension – the measure does not take into account the geometric degeneracy which can occur when many of the leading CLVs can become aligned. In practice this overestimation of dimension is actually beneficial within the DA framework (more discussion on this can be found in Sect. 4.2).

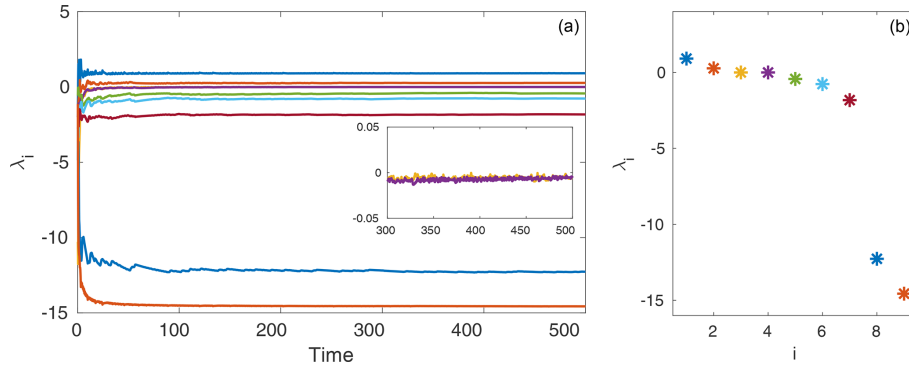
Figure 4 shows the local dimension plotted along the attractors of each of the subsystems. We see that the local dimension is lowest when the model is in the interior region of the ocean subsystem attractor. In contrast, the extratropical atmosphere subsystem attractor displays periods of low dimension largely uniformly confined to the centre of each lobe of the attractor. The tropical atmosphere also shows most of the measures of low dimension confined to the interior of the attractor, reflecting the strong ocean coupling.

### 3 Covariant Lyapunov vectors

The existence of Lyapunov vectors for a large class of dynamical systems was proven by Oseledets (1968). The multiplicative ergodic theorem states that there exists a relation between Lyapunov exponents,  $\lambda_i$ , and a (non-unique) set of vectors  $\phi$  such that

$$\lambda_i = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \log \|\mathcal{A}(x(t), \tau)\phi\| \quad \text{iff} \\ \phi \in \Phi_i(x(t)) \setminus \Phi_{i+1}(x(t)). \quad (4)$$

Here,  $\mathcal{A}(x(t), \tau)$  is the forward and backward mapping of solutions under the tangent dynamics of a given dynamical system.



**Figure 3.** (a) Convergence of Lyapunov exponents of Eq. (1) with an inset of zoom around zero to show the two neutral exponents. (b) Asymptotic values of Lyapunov exponents of Eq. (1).

**Table 1.** Asymptotic Lyapunov exponents of Eq. (1) computed over 500 time units using a QR decomposition method, time step of 0.01, and orthogonalization step of 0.25.

$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$	$\lambda_7$	$\lambda_8$	$\lambda_9$
0.9071	0.2670	-0.0056	-0.0060	-0.4326	-0.7706	-1.8263	-12.2691	-14.5640

cal system (also referred to as the cocycle). For system (1),  $\mathcal{A}(x(t), \tau) = e^{\tau Jf(x(t))}$ , where  $Jf$  denotes the Jacobian of  $f$ , the right-hand side of system (1). The subspaces ( $\Phi_i$ ) on which the growth rates ( $\lambda_i$ ) occur are covariant with the tangent dynamics and invariant under time reversal (Ginelli et al., 2007; Froyland et al., 2013). The covariant Lyapunov vectors (CLVs) are then defined as the set of directions at each point in phase space that satisfy Eq. (4) both backwards and forwards in time (Ginelli et al., 2007; Ng et al., 2011). In the last few decades there have been significant advances in the ability to numerically approximate such vectors for chaotic dynamical systems (Ginelli et al., 2007; Wolfe and Samelson, 2007; Froyland et al., 2013). In this work we will employ a numerical algorithm introduced by Froyland et al. (2013) (Algorithm 2.2 in the aforementioned study). We summarize this algorithm below.

It should be mentioned that the push forward step need not be equal to  $N$ ;  $M \neq N$  for  $\mathcal{A}(x^{i-N}, M)$ . However, for all calculations in this study we consider only the case  $M = N$ . The trajectory of the system is discretized with time step  $\Delta t$  such that  $x^i = x(t_i)$  and  $x^{i+m} = x(t_i + m \Delta t)$ .

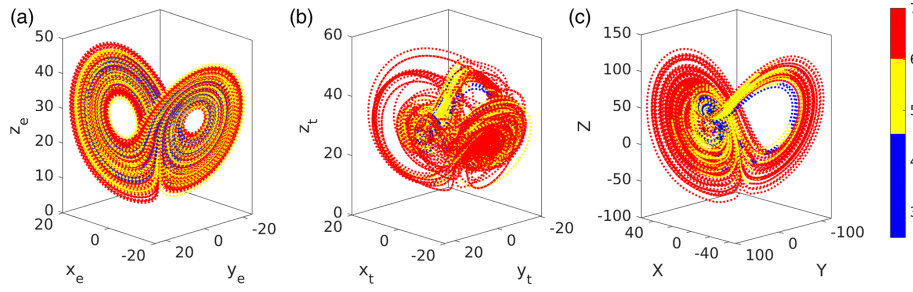
We examine the effectiveness of this algorithm on the Peña and Kalnay (2004) model introduced in Sect. 2. By definition, CLVs describe the directions in phase space corresponding to different growth rates. Two or more CLVs can align during a transition to a different regime in phase space. We calculate the alignment through

$$\theta_{i,j} = \frac{|\phi_i \cdot \phi_j|}{\|\phi_i\| \cdot \|\phi_j\|}. \quad (5)$$

Here,  $\theta_{i,j} = |\cos(\Theta_{i,j})|$ , where  $\Theta_{i,j}$  is the angle between the  $i$ th and  $j$ th CLVs. Larger values of  $\theta_{i,j}$  imply alignment of

the two CLVs. Figure 5 shows the alignment of the unstable CLVs ( $\theta_{1,2}$ ) and the neutral CLVs ( $\theta_{3,4}$ ) plotted against the  $X$  component of the system (1), along with the FTLEs and corresponding local time-varying Kaplan–Yorke dimension. Since the window to calculate the FTLEs was chosen based on the variability timescales of the ocean subsystem, the subsequent local dimension measure should reflect areas of increased or decreased instability along the ocean attractor (Fig. 4). We also expect alignment of the CLVs during a transition in this subsystem. The CLVs analysed in Fig. 5 are calculated from the first 35 time units of the previous model run with a time step  $\Delta t = 0.01$ . We start the calculation at  $t = 5$  to allow for initialization and a window size of  $\tau = 4$ . The parameters for the algorithm are set as  $N = 400$  and  $n = 25$ . It can be seen that the algorithm detects near alignment of either the unstable or neutral CLVs during the transitions in the ocean subsystem. The transitions here are between the inner part of the attractor with smaller oscillation amplitudes and the outer, large-amplitude excursions. In general, the alignment is most prominent in the outer, large-amplitude part of the attractor. This follows the changes in local dimension, shown in the lower panel of Fig. 5c. Higher local dimension tends to accompany alignment of the unstable and neutral CLVs. This relates to the inability of the Kaplan–Yorke dimension measure to account for finite-time geometric degeneracy.

At this point we will comment on the non-trivial relationship between the CLVs and FTLEs calculated here. As discussed in Vannitsem and Lucarini (2016), there are three different types of FTLEs one can compute: backward (FTBLEs), forward (FTFLEs), and covariant (FTCLEs). Each type of FTLE gives the local growth rate of the correspond-



**Figure 4.** Local Kaplan–Yorke dimension plotted along the trajectory in phase space for the extratropical atmosphere (a), tropical atmosphere (b), and ocean (c) subsystems.

---

**Algorithm 1** Summary of numerical algorithm for calculating CLVs introduced by Froyland et al. (2013)

---

1. Construct matrix cocycle  $\mathcal{A}(x^{i+m}, 0)$  for every  $m \in [-N, \dots, N]$ .
  2. Compute the eigenvectors  $e_j^{i-N}$  of  $\mathcal{A}(x^{i-N}, N)^* \mathcal{A}(x^{i-N}, N)$  [the right singular vectors of  $\mathcal{A}(x^{i-N}, N)$ ], where  $\mathcal{A}(x^{i-N}, N) = \mathcal{A}(x^i, 0) \cdot \dots \cdot \mathcal{A}(x^{i-N}, 0)$ . Note that  $\mathcal{A}(x^{i-N}, N)^*$  is the adjoint of  $\mathcal{A}(x^{i-N}, N)$ .
  3. Push forward by multiplication of matrix cocycle,  $\phi_j^i = \mathcal{A}(x^{i-N}, N) e_j^{i-N}$ .
  4. For each  $j$ , reorthogonalize  $\phi_j^i$  with subspace spanned by eigenvectors  $e_k^{i-n}$  for  $k = 1, \dots, j-1$  of  $\mathcal{A}(x^{i-n}, N)^* \mathcal{A}(x^{i-n}, N)$  every  $n$  time steps.
  5. The vector  $\phi_j^i$  is then an approximation of the  $j$ -th largest CLV at time  $t = t_i$ .
- 

ing Lyapunov vectors. Although all three converge to the asymptotic Lyapunov exponents as the computation window increases, the temporal variability for finite window size can be different depending on the model at hand. Vannitsem and Lucarini (2016) found that when calculating the growth rates of the CLVs, higher variability in the FTCLEs corresponding to neutral or near-zero modes occurred compared to the other two methods. This could have implications for the local Kaplan–Yorke dimension if one were to use the FTCLEs rather than the FTBLEs in the calculation. We remark here that the QR method produces backward Lyapunov vectors (BLVs) and their corresponding FTBLEs. We therefore compare the Kaplan–Yorke dimension as computed from FTCLEs and FTBLEs for the coupled Lorenz system (1) in Fig. 6. We see that for this model the dimension calculated from FTBLEs approximately bounds that calculated from FTCLEs. Since we will be using the local Kaplan–Yorke dimension as a lower bound within the framework of our experiments, the FTBLEs give a conservative estimate of dimension that is varying with our dynamics (see Sect. 4.2 for discussion of the implementation of dimension in our experiments).

In the following sections, we will utilize CLVs within the data assimilation framework of ensemble forecasting. The CLVs will be used to construct the forecast error covariance matrix, which informs the increment used on ensemble members to bring them closer to observations. Using CLVs in this context suggests a more accurate method of forming

the forecast error covariance matrix when the true covariance is undersampled due to an insufficient number of ensemble members (see e.g. Palatella and Trevisan, 2015, where the authors applied a similar approach using BLVs on the classical Lorenz, 1963, and Lorenz, 1996, systems).

#### 4 Data assimilation with the Kalman filter

We now summarize the Kalman filter equations. For detailed derivations we refer the interested reader to the reviews by Evensen (2003), Houtekamer and Zhang (2016), or Carrassi et al. (2018). Here we follow the notation of Carrassi et al. (2018). Consider a deterministic or stochastic model defined by

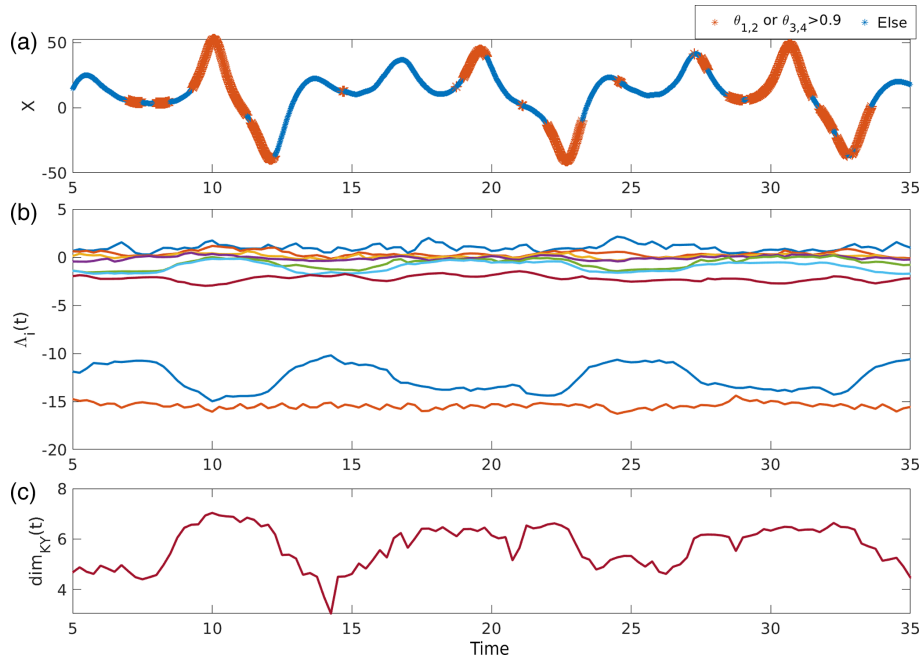
$$\mathbf{x}_k = \mathcal{M}_{k:k-1}(\mathbf{x}_{k-1}, \mathbf{p}) + \eta_k, \quad (6)$$

where  $\mathbf{x}_k \in \mathbb{R}^N$  is the model state at time  $t = t_k$ ,  $\mathbf{p} \in \mathbb{R}^p$  are the model parameters,  $\mathcal{M}_{k:k-1} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  is a function taking the model from time  $t_{k-1}$  to  $t_k$ , and  $\eta_k$  is the model error at time  $t_k$  (for deterministic systems, let  $\eta_k = 0$ ). Suppose that there exists a time-dependent set of observations  $\mathbf{y} \in \mathbb{R}^d$  which can be expressed as a function of the model state through

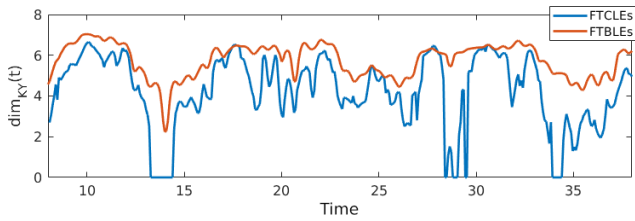
$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \epsilon_k. \quad (7)$$

The observation operator  $\mathcal{H}_k : \mathbb{R}^N \rightarrow \mathbb{R}^d$  can be linear or nonlinear, and  $\epsilon_k$  is the observational error.





**Figure 5.** Local dynamical properties of a segment of an example model run. (a) Alignment of CLVs associated with the unstable and neutral subspaces plotted along the  $x$  coordinate of the ocean subsystem. Large orange stars indicate high alignment of unstable or neutral CLVs ( $\theta_{1,2} > 0.9$  or  $\theta_{3,4} > 0.9$ ). (b) Time-varying FTLEs  $\Lambda_i(t)$  computed over window  $\tau = 4$ . (c) Local Kaplan–Yorke dimension calculated from FTLEs.



**Figure 6.** Comparison of the finite-time Kaplan–Yorke dimension calculated using the growth rates of CLVs (FTCLEs) and the QR method (FTBLEs).

In the Kalman filter method, Eqs. (6) and (7) are assumed to be linear, resulting in evolution and observation matrices  $\mathbf{M}_{k:k-1}$  and  $\mathbf{H}_k$  respectively. The model and observation errors,  $\eta_k$  and  $\epsilon_k$ , are taken to be uncorrelated in time (white noise) and from a Gaussian distribution with covariance matrices  $\mathbf{Q}_k \in \mathbb{R}^{N \times N}$  and  $\mathbf{R}_k \in \mathbb{R}^{d \times d}$  respectively. There are two basic steps to the Kalman filter method: forecast and analysis.

– Forecast equations

$$\mathbf{x}_k^f = \mathbf{M}_{k:k-1} \mathbf{x}_{k-1}^a, \quad (8a)$$

$$\mathbf{P}_k^f = \mathbf{M}_{k:k-1} \mathbf{P}_{k-1}^a \mathbf{M}_{k:k-1}^T + \mathbf{Q}_k. \quad (8b)$$

– Analysis equations

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k]^{-1}, \quad (9a)$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f), \quad (9b)$$

$$\mathbf{P}_k^a = (\mathbf{I}_k - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^f. \quad (9c)$$

There is a difficulty in finding accurate solutions to Eqs. (2)–(3) for realistic systems which have high dimensions and are nonlinear (as is the case in weather and climate forecasting). Within the Kalman filter class, various ensemble filter variants have been applied to tracking trajectories in nonlinear systems. The most popular are the deterministic filters (Tippett et al., 2003; Sakov and Oke, 2008; Sakov et al., 2012).

#### 4.1 Ensemble Kalman filtering

Ensemble Kalman filtering methods use Monte Carlo sampling to form the approximate error statistics of a model. An ensemble of model states  $\mathbf{x}^f \in \mathbb{R}^N$  with a finite number of ensemble members  $m$  produces an approximation to the true error covariance matrix as follows. The ensemble forecast anomaly matrix  $\mathbf{X}^f \in \mathbb{R}^{N \times m}$  is constructed with respect to the ensemble mean  $\bar{\mathbf{x}}^f \in \mathbb{R}^N$ :

$$\bar{\mathbf{x}}^f = \frac{1}{m} \sum_{n=1}^m \mathbf{x}_n^f, \quad (10a)$$

$$\mathbf{X}^f \equiv \frac{1}{\sqrt{m-1}} [\mathbf{x}_1^f - \bar{\mathbf{x}}^f, \dots, \mathbf{x}_m^f - \bar{\mathbf{x}}^f]. \quad (10b)$$

Note that we have dropped the time subscript  $k$ ; the subscripts used here refer to individual ensemble members. The

forecast error covariance matrices  $\mathbf{P}^f$  are then constructed through

$$\mathbf{P}^f = (\mathbf{X}^f)(\mathbf{X}^f)^T. \quad (11)$$

To preserve the variance of the ensemble through the analysis step, square-root (deterministic) schemes for ensemble Kalman filtering are often used. One such scheme is the ensemble transform Kalman filter (ETKF) developed by Bishop et al. (2001) and then further adapted for large, spatiotemporally chaotic systems by Hunt et al. (2007). In such schemes, the observations do not need to be perturbed to preserve the analysis covariance in Eq. (9c). The main idea is that a transform matrix  $\mathbf{T}$  can be used to adjust the ensemble analysis anomalies matrix,

$$\mathbf{X}^a \equiv \frac{1}{\sqrt{m-1}} \left[ \mathbf{x}_1^a - \bar{\mathbf{x}}^a, \dots, \mathbf{x}_m^a - \bar{\mathbf{x}}^a \right] = \mathbf{X}^f \mathbf{T}, \quad (12)$$

which ultimately forms the analysis error covariance matrix,

$$\mathbf{P}^a = (\mathbf{X}^a)(\mathbf{X}^a)^T. \quad (13)$$

This transform matrix  $\mathbf{T}$  is recovered by calculating the ensemble perturbations in normalized observation space:

$$\mathbf{S} = \mathbf{R}^{-1/2} \mathbf{H} \mathbf{X}^f. \quad (14)$$

The transform matrix  $\mathbf{T}$  is then defined as

$$\mathbf{T} = (\mathbf{I} + \mathbf{S}^T \mathbf{S})^{-1/2}, \quad (15)$$

where  $\mathbf{I}$  is the  $m \times m$  identity matrix. See Bishop et al. (2001) for the full derivation. This leads to the update of the ensemble mean and the individual ensemble members to their analysed state through the equations

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{K}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^f), \quad (16a)$$

$$\mathbf{x}_n^a = \bar{\mathbf{x}}^a + (\sqrt{m-1})[\mathbf{X}^f \mathbf{T}]_{*,n}. \quad (16b)$$

Following Bishop et al. (2001) we define the Kalman gain  $\mathbf{K}$  through Eq. (9a), and the subscript  $*, n$  denotes taking the  $n$ th column of the matrix.

Another deterministic scheme for ensemble Kalman filtering which uses a left-multiplied transform matrix was shown by Tippett et al. (2003) to be equivalent to ETKF:

$$\mathbf{X}^a = \mathbf{T} \mathbf{X}^f, \quad (17a)$$

$$\mathbf{T} = (\mathbf{I} - \mathbf{K} \mathbf{H})^{1/2}. \quad (17b)$$

We will refer to this left-multiplied transform filter as the ensemble square-root filter (ESRF). The ensemble mean is updated through Eq. (16a) and the individual ensemble members are then updated through

$$\mathbf{x}_n^a = \bar{\mathbf{x}}^a + (\sqrt{m-1})[\mathbf{T} \mathbf{X}^f]_{*,n}. \quad (18)$$

When using ensemble Kalman filtering methods like the ones introduced here, sampling errors can often occur. For

nonlinear models in particular, there is a systematic underestimation of analysis error covariances which eventually leads to filter divergence (Anderson and Anderson, 1999; Bocquet et al., 2015; Raanes et al., 2019). This is commonly avoided through the use of inflation. In other words, after each analysis step the ensemble anomalies are inflated through

$$\mathbf{x}_n^a = \bar{\mathbf{x}}^a + \lambda(\mathbf{x}_n^a - \bar{\mathbf{x}}^a), \quad \lambda > 1, \quad (19)$$

where  $(\lambda - 1)/100$  is the percentage inflation. Grudzien et al. (2018b) recently showed that the need for inflation tuning could potentially be compensated by including the asymptotic stable modes which produce transient instabilities. The following section introduces one way to account for such transient instabilities through a projection of the forecast error covariance matrix onto a subset of CLVs.

#### 4.2 Ensemble filtering in reduced subspace

Here we define the error covariance matrix  $\mathbf{P}^f$  based on the directions of growth and decay of errors associated with different timescales at the given analysis time. Specifically, we construct  $\mathbf{P}^f$  using the CLVs computed at each data assimilation time step where the number of CLVs is determined by the local attractor dimension  $\dim_{\text{KY}}$ . This allows for the inclusion of unstable, neutral and stable directions dependent on the local dynamics of the system. This differs from past approaches where the subspace was determined in terms of the asymptotic Lyapunov exponents, and therefore the rank of the error covariance matrix was kept fixed (Trevisan and Uboldi, 2004; Carrassi et al., 2008; Trevisan and Palatella, 2011; Palatella and Trevisan, 2015).

To determine the number of CLVs required to form the basis for  $\mathbf{P}^f$ , we use the time-dependent or local  $\dim_{\text{KY}}$  rounded up to an integer value. To determine how to weight the individual CLVs, we deconstruct the ensemble anomalies matrix defined in Eq. (10b),  $\mathbf{X}^f$ , into

$$\mathbf{X}^f = \Phi \mathbf{W}, \quad (20)$$

where  $\Phi$  is a matrix with columns equal to the CLVs ( $\phi_j$ ) and  $\mathbf{W}$  is a matrix of weights. The columns of  $\Phi$  are ordered according to the corresponding FTLEs in descending order (the first being the direction corresponding to the fastest unstable growth). In this formulation  $\Phi$  need not be square, i.e. the CLVs used do not need to span the entire space. We compute the CLVs at the assimilation step using Algorithm 1. Equation (20) can then be solved for  $\mathbf{W}$  in a least-squares sense through

$$\mathbf{W} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{X}^f. \quad (21)$$

The weights in  $\mathbf{W}$  combined with the directions in  $\Phi$  now define an object with dimension equal to the chosen number of CLVs whose covariance matrix is defined by

$$\mathbf{P}^f = \Phi \mathbf{W} \mathbf{W}^T \Phi^T. \quad (22)$$



We can then use the formulation of  $\mathbf{P}^f$  above in conjunction with the ensemble schemes of Sect. 4.1. We use the modified forecast covariance matrix (Eq. 22) in the calculation of the Kalman gain (Eq. 9a), which then also alters any subsequent calculations.

It is important that we span the local dimension of the attractor within the ensemble DA framework in order to avoid ensemble collapse. As mentioned in Sect. 3, the Kaplan–Yorke dimension computed from the FTLEs is an upper bound to the true local dimension. Figure 7 shows the alignment of the CLVs at two time instances of a model run where the leading FTLEs behave similarly and consequently the Kaplan–Yorke dimensions are approximately the same. In the first case ( $t = 306.24$ ), the leading five CLVs are strongly aligned as well as the two most stable (8 and 9). The dimension based on alignment would then intuitively be around four, and one would need to select the set of CLVs that are not aligned in order to avoid ensemble collapse. According to our method, since the local Kaplan–Yorke dimension is greater than seven, we would retain the leading eight CLVs, therefore retaining all necessary directions to maintain spread. The second case ( $t = 705.04$ ) shows very different alignment behaviour. Here the leading CLVs are not strongly aligned, but there is strong alignment of CLVs 4–6 and pairs 3, 7, 8, and 9. This would give an alignment-based dimension around five, but again we need to retain up to the eighth CLV to span the independent directions. While one could create a method based on alignment for selecting directions, we point out that the actual criterion for “strong alignment” is arbitrary, and one could risk excluding a significant direction. For this reason we argue that the local Kaplan–Yorke measure gives a conservative estimate for the number of CLVs to retain.

## 5 Results

We perform a collection of data assimilation experiments for the system (1) using a control run as observations (computed using a Runge–Kutta fourth-order scheme with  $\Delta t = 0.01$ ). Here we emphasize that we are interested in exploring the dynamical attributes of data assimilation across multiple timescales. In all the cases we are using standard strong CDA, meaning that the cross-covariances are used amongst all components regardless of the observation set (Laloyaux et al., 2016; O’Kane et al., 2019). This allows for the analysis increment of any unobserved subsystems to be influenced by the observations, even in cases of weak coupling between the subsystems. We experiment with different observation sets to explore the applicability and performance of the variable-rank strong CDA with incomplete or temporally correlated observations. Such observation sets are arguably more realistic representations of the types of observations used in climate applications.

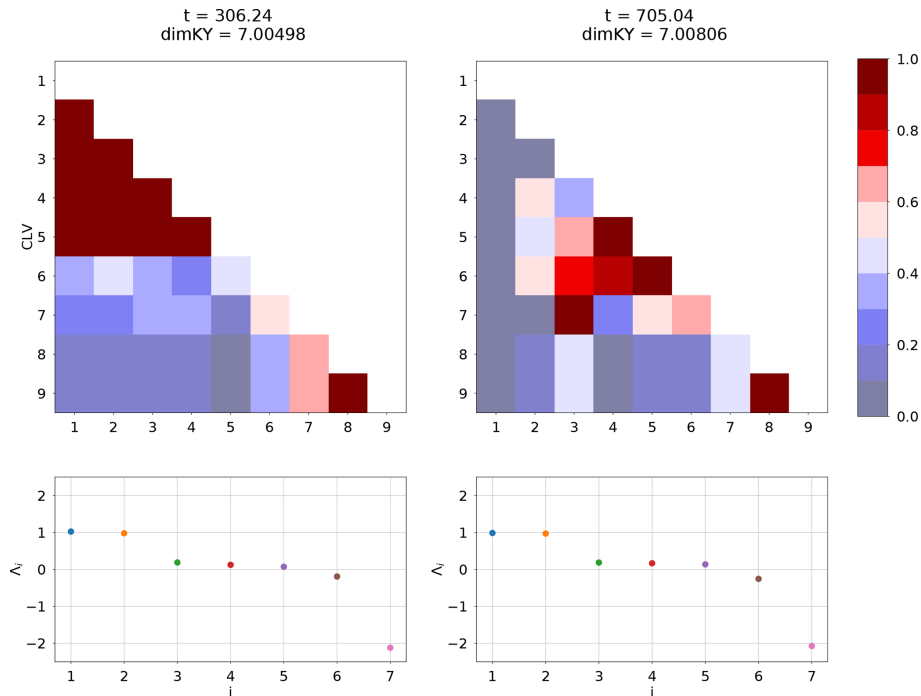
The initialization settings for the DA experiments are as follows, unless otherwise stated. We use the settings from

Yoshida and Kalnay (2018): assimilation window 0.08, inflation factor 1 % (optimal for all experiments unless noted otherwise), and 10 ensemble members. We run the model for 75 000 time steps (9375 assimilation windows) and use 50 000 time steps (6250 assimilation windows) for calculating analysis error statistics. The model is allowed to spin up for 400 time steps before starting the assimilation cycles as we are using a window of  $\tau = 4$  for the calculation of the FTLEs and CLVs. We note that for the CLV method, the system must be sufficiently tracking the control to accurately calculate the initial CLVs. For this reason we start the assimilation before there is significant ensemble divergence. The ensemble members are initialized as perturbations from the control initial condition, taken from a uniform distribution defined on  $[-0.025, 0.025]$ .

The dynamical properties of the experiments are calculated with respect to the ensemble mean trajectory. The FTLEs are computed using the QR decomposition over the previous 400 time steps leading up to the assimilation time step. The local Kaplan–Yorke dimension is then calculated from the FTLEs. The CLVs are then calculated using a slight modification to Algorithm 1 – due to the absence of an accurate future trajectory of the ensemble mean, we do not perform the reorthogonalization to the eigenvectors of  $\mathcal{A}(x^{i-n}, N)^* \mathcal{A}(x^{i-n}, N)$  (step 4). This is equivalent to using Algorithm 2.1 of Froyland et al. (2013). We use the previous 400 time steps of the ensemble mean trajectory leading up to the assimilation time step to compute the matrix cocycle and subsequently the CLVs.

### 5.1 Constructing the observations

In the Kalman filtering method introduced in Sect. 4, there is an underlying assumption that the observations have some error with variance  $\mathbf{R}$ . This error variance is typically unknown and chosen a priori. If we consider the observations in a statistical sense, we can deconstruct them at each assimilation step into a mean field and perturbation value:  $\mathbf{y} = \bar{\mathbf{y}} + \hat{\mathbf{y}}$ . In such a formulation,  $\bar{\mathbf{y}}$  would be the truth at a given point in time and the observation error variance would be the average of the variance of the perturbations,  $\mathbf{R} = \hat{\mathbf{y}} \hat{\mathbf{y}}^T$ . To emulate this in deterministic models where the truth  $\bar{\mathbf{y}}$  is known (i.e. from a control run), it is common practice to construct the observations by adding to the truth a random value  $\hat{\mathbf{y}}$  taken from a normal distribution with variance given by the diagonals of  $\mathbf{R}$ . However, this produces uncorrelated observation errors which have the same variance at any given point in phase space. In many applications the true variance of the observation error can be spatially dependent and errors are often correlated in time (O’Kane and Frederiksen, 2008). We therefore also consider the case where there is error in the observations, but it is consistent with the underlying nonlinear dynamics by constructing a trajectory that “shadows” the truth. Both types of observational errors, with the addi-



**Figure 7.** We compare the alignment of the CLVs at two time steps of a model run. The two time steps have a similar Kaplan–Yorke measure with the distribution of the leading FTLEs, shown in the bottom panels. We observe that the behaviour of the alignment can be vastly different for similar FTLE behaviour; however, the method of retaining CLVs based on the Kaplan–Yorke measure gives a reliable way of reflecting the true local dimension regardless of alignment.

tional case of perfectly observing a subset of variables, are explored.

The main differences in our subsequent experiments are in the subset of observations used and their corresponding observational errors. We aim to assess the performance of the reduced-rank strong CDA within the different configurations. We first present a benchmark test on the CLV method which is identical to an experiment presented in Yoshida and Kalnay (2018), where the  $y$  component of each subsystem is observed. The perturbations to the control run are from a normal distribution with error variance  $\mathbf{R} = \text{diag}([1, 1, 25])$ . This case is referred to as *benchmark observations*. We then consider two experiments where the observations are less sparse within the subsystems; however, one subsystem is completely unobserved: *atmosphere observations* ( $y_e, z_e, y_t, z_t$ ) and *ENSO (El Niño–Southern Oscillation) observations* ( $y_t, z_t, Y, Z$ ). For the atmosphere observations case we sample the perturbations from error variance  $\mathbf{R} = \mathbf{I}_4$ , while for the ENSO observations we experiment with different error variances. We then consider correlated observation errors through *shadowed observations*. For these experiments the observation error variances  $\mathbf{R}$  are set to the standard values from Yoshida and Kalnay (2018), but the actual perturbations are constructed through a model trajectory which is initialized close to the control run and forced by a relaxation term back to the control. The observation errors

then also reflect the local nonlinear growth. We repeat the benchmark, atmosphere, and ENSO observation cases with this type of observation error. Finally, we reduce the observation space to only the extratropical subsystem ( $x_e, y_e, z_e$ ). This extends upon the work in O’Kane et al. (2019), where the authors considered only ocean observations ( $X, Y, Z$ ). Due to the difficulty of constraining a system through only fast, weakly coupled dynamics, we decrease the assimilation window and assume perfect observations.

## 5.2 Benchmark observations

The first DA experiment we consider is a benchmark case with observations ( $y_e, y_t, Y$ ). We reproduce the results in Yoshida and Kalnay (2018) using the CLV method introduced in Sect. 4.2. We first perform a DA experiment using a full-rank covariance matrix (equivalent to the ETKF method introduced in Sect. 4.1) and then compare it to using reduced subspace methods. The first reduced subspace method uses a fixed number of CLVs defined by spanning the asymptotic unstable and neutral subspace plus the first stable mode as in Ng et al. (2011). The second reduced subspace method also uses a fixed number of CLVs, except the number is defined by the asymptotic Kaplan–Yorke dimension as suggested in Carrassi et al. (2008) (note in that study that the authors discuss the number of ensemble members which is equivalent to the rank of the covariance matrix). These fixed numbers

are five and six CLVs respectively. Finally we analyse our novel reduced subspace method which uses a variable number of CLVs based on the local Kaplan–Yorke dimension. We note that all experiments perform similarly when using the BLVs instead of the CLVs; however, BLVs do not provide the same local phase-space information. While we focus on the performance of the CLV method in this work, we include a comparison using BLVs for the experiment utilizing local dimension.

The error statistics of all the experiments are listed in Table 2. The analysis RMSE is calculated for each subsystem individually at every assimilation step and then averaged over the steps, in line with the error statistics produced in Yoshida and Kalnay (2018). We also calculate the average RMSE of the full system. The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{N}(\bar{\mathbf{x}}^a - \mathbf{x})^T(\bar{\mathbf{x}}^a - \mathbf{x})}, \quad (23)$$

where  $N$  is the number of states in the analysed system (either 3 or 9) and  $\mathbf{x}$  is the truth (control run in our case). We also calculate the spread and average increment for each state variable through

$$\begin{aligned} \text{spread} &= \sqrt{\text{diag}(\mathbf{P}^f)}, \\ \text{average increment} &= \overline{(\mathbf{x}^a - \mathbf{x}^f)}. \end{aligned} \quad (24)$$

Finally, we calculate the bias with respect to our observations  $\mathbf{y}$ :

$$\text{bias} = \overline{(\mathbf{y} - \mathbf{H}\mathbf{x}^f)}. \quad (25)$$

We observe that all experiments generally succeed at constraining the full system. The trajectories, spread, increments, and error metrics of the variable CLV experiments are shown in Fig. 8; all other experiments behave qualitatively similarly. When comparing the reduced subspace experiments to the full-rank case in Table 2, we find that using five CLVs (unstable, neutral, and one stable) performs worse than the other CLV experiments. While using the asymptotic Kaplan–Yorke dimension (six CLVs) shows improvement over five CLVs, the most improvement is in the variable CLV (and BLV) case. Although all the methods produce a comparable average dimension (last column of Table 2), our experiments show that taking into account local variations in dimension is most effective.

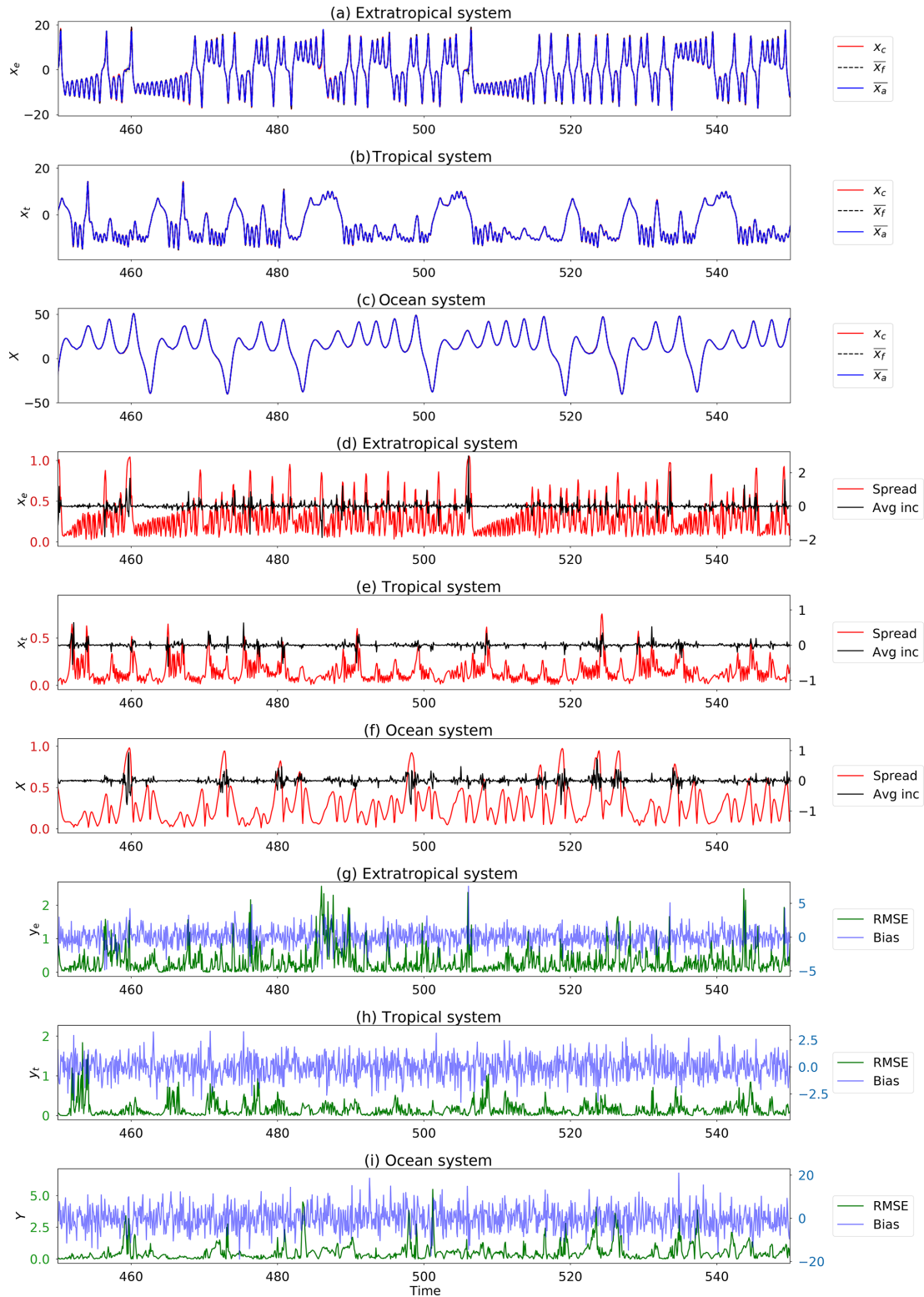
To take a closer inspection of the dynamics during the assimilation, Fig. 9 shows the corresponding dynamical properties of the variable CLV experiment in time. Panel (a) shows the FTLEs computed from the ensemble mean. The first thing to note is the correlation of the temporal FTLE variations. The most unstable and most stable modes have the same frequency of variability and remain highly correlated throughout the whole experiment. The lower-frequency variations

seen in the second most stable mode are often correlated with some of the weakly stable modes. These seem to have the biggest impact on the local dimension (shown in panel b). We can compare the changes in FTLEs and local dimension to the rank of the covariance matrix (panel c). A decrease in local dimension typically occurs when one of the weakly stable modes becomes more stable, and therefore fewer CLVs are needed to span the local unstable and neutral subspace. In contrast, when the dimension increases, the weakly stable modes become more unstable, at times even becoming positive. This implies that more CLVs are needed to span the local unstable and neutral subspace, and therefore the rank of the covariance matrix increases. The local Kolmogorov–Sinai entropy is also shown with the local dimension. We see that directly before a full-dimension collapse ( $\text{dim}_{\text{KY}} = 0$ ), there is a spike in local entropy. A collapse in dimension here occurs when all the FTLEs become negative. This does not impact the effectiveness of the method, however, as it just enforces the covariance matrix and Kalman gain to be zero for that analysis step. The ensemble members are not adjusted and are therefore left to evolve accordingly. Figure 9d shows the statistics of the FTLEs. The mean values, shown as a full circle, are used to calculate the average dimension in Table 2. These can be compared to the asymptotic values computed in Sect. 2.2 which are shown as open circles to the right of the FTLE statistics. We also show the standard deviation for each of the FTLEs. The largest standard deviations are found in the fifth and sixth FTLEs, which correspond to the first two stable modes of the system. This supports the hypothesis that the weakly stable modes are most influential in the variation of local dimension. We also see that the maximum values are greater than (or equal to) zero. This implies that at a given time the fifth and sixth modes have moved into the unstable/neutral subspace and additional modes are therefore needed to account for nonlinear error growth.

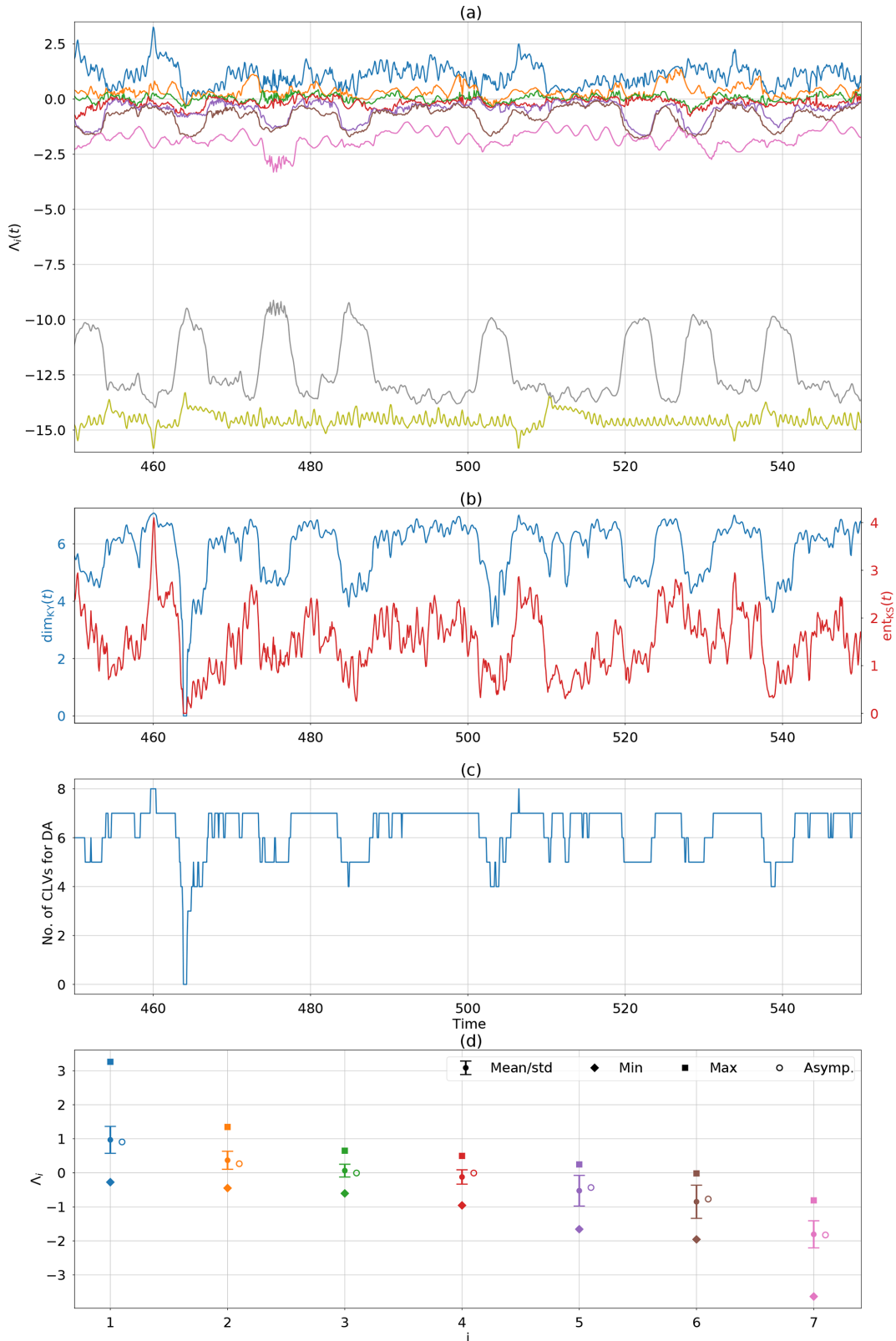
### 5.3 Atmosphere observations

We now consider the case where only the two atmosphere subsystems are observed; however, the observations are less sparse within each subsystem in that we take both the  $y$  and  $z$  components. Yoshida and Kalnay (2018) considered a similar case where only  $y$  component observations were used with the cross-covariances in the atmosphere, but the ocean was also observed and assimilated separately. Here we rely only on the cross-covariances to recover the ocean subsystem, as the ocean is strongly coupled to the tropical atmosphere. For this case we use the following settings: assimilation window 0.08, observation error covariance matrix  $\mathbf{R} = \mathbf{I}_4$ , inflation factor 1 %, and 10 ensemble members. The model is run for the same amount of time as in the benchmark observation experiments.

Table 3 shows the summary statistics of performing full-rank DA (nine CLVs), rank equal to the number of CLVs corresponding to asymptotic dimension, and a variable rank



**Figure 8.** Segment of DA using the variable CLV method and benchmark observations ( $y_e$ ,  $y_t$ ,  $Y$ ). **(a–c)** Trajectories shown are control run (red), ensemble mean (blue), and individual ensemble members (faint). **(d–f)** Metrics shown are the size of the ensemble spread (red) and ensemble mean increment (black). **(g–i)** Metrics shown are root mean square error (RMSE, green) and bias (blue). For conciseness we only show the results for one coordinate of each subsystem. The other two coordinates behave similarly. Parameters: assimilation window 0.08, inflation factor 1 %, 10 ensemble members, and observation error covariance matrix  $\mathbf{R} = \text{diag}([1, 1, 25])$ .



**Figure 9.** Local attractor properties of DA using the variable CLV method and benchmark observations ( $y_e, y_t, Y$ ). **(a)** FTLEs calculated from the ensemble mean trajectory over a window of  $\tau = 4$ . **(b)** Local Kaplan–Yorke dimension and Kolmogorov–Sinai entropy computed through Eqs. (2) and (3) using the FTLEs at the given time. **(c)** Rank of the covariance matrix. **(d)** Statistics of the first seven FTLEs compared to asymptotic values.

**Table 2.** Summary metrics of DA experiments using a right-transform matrix (Eq. 15) and benchmark observations ( $y_e, y_t, Y$ ). The angle brackets  $\langle \cdot \rangle$  denote the average over assimilation steps. Compare to results in Yoshida and Kalnay (2018). Parameters: assimilation window 0.08, inflation factor 1 %, 10 ensemble members.

Method	Observations [error variance]	$\langle \text{RMSE} \rangle$ extratropical	$\langle \text{RMSE} \rangle$ tropical	$\langle \text{RMSE} \rangle$ ocean	$\langle \text{RMSE} \rangle$ full	$\langle \text{dim}_{KY} \rangle$
CLVs – nine (full rank)	$y_e, y_t, Y$ [1, 1, 25]	0.3142	0.1598	0.4948	0.4027	5.8928
CLVs – five (unstable/neutral subspace + 1)	$y_e, y_t, Y$ [1, 1, 25]	0.3123	0.1843	0.5920	0.4550	5.8870
CLVs – six (global dimension)	$y_e, y_t, Y$ [1, 1, 25]	0.3156	0.1674	0.5513	0.4310	5.8892
CLVs – variable (local dimension)	$y_e, y_t, Y$ [1, 1, 25]	0.3215	0.1688	0.5346	0.4272	5.8863
BLVs – variable (local dimension)	$y_e, y_t, Y$ [1, 1, 25]	0.3149	0.1658	0.5122	0.4141	5.8895

**Table 3.** Summary metrics of DA experiments using a right-transform matrix (Eq. 15) and atmosphere observations ( $y_e, z_e, y_t, z_t$ ). Parameters: assimilation window 0.08, inflation factor 1 %, 10 ensemble members.

Method	Observations [error variance]	$\langle \text{RMSE} \rangle$ extratropical	$\langle \text{RMSE} \rangle$ tropical	$\langle \text{RMSE} \rangle$ ocean	$\langle \text{RMSE} \rangle$ full	$\langle \text{dim}_{KY} \rangle$
CLVs – nine	$y_e, z_e, y_t, z_t$ [1, 1, 1, 1]	0.1734	0.1332	0.5782	0.4515	5.8672
CLVs – six	$y_e, z_e, y_t, z_t$ [1, 1, 1, 1]	0.1715	0.1386	0.5807	0.4524	5.8718
CLVs – variable	$y_e, z_e, y_t, z_t$ [1, 1, 1, 1]	0.1697	0.1383	0.5659	0.4433	5.8681

equal to the number of CLVs corresponding to local dimension. All the experiments have comparable summary statistics; however, the variable CLV experiment slightly outperforms both fixed-rank methods (Table 3). We also analysed the dynamical properties of the variable CLV experiment and did not find any significant differences to those of the benchmark experiment (Fig. 9).

#### 5.4 ENSO observations

The strong coupling and low-frequency variation in the ocean and tropical atmosphere subsystems represent an ENSO-like variability. We therefore refer to the case of observing the  $y$  and  $z$  components from the tropical and ocean subsystems as ENSO observations. Again, this is similar to a case studied in Yoshida and Kalnay (2018), except that we do not assimilate the extratropical atmosphere at all and attempt to recover the variability solely through the cross-covariances. We do not expect to track the control trajectory of the extratropical system, but we are interested to see whether we can avoid collapse (i.e. the loss of variability) in the ensemble mean of the extratropical attractor.

Since the extratropical subsystem is likely to be unconstrained with these observations, the ensemble mean will not be accurately estimated. In such a case, the variable CLV method fails due to the fact that the first CLV (which corresponds to the directions of fastest error growth) is inaccurately calculated. Therefore the true directions of unstable growth are inaccurately sampled in the reduced space experiments. This is amplified by the fact that we are using uncorrelated observation errors; if the observation errors have a temporal correlation, the dominant direction of nonlinear unstable growth can be ascertained even without tracking the extratropical subsystem (see the following section on shadowed observations). The inaccurate dimension reduction leads to exponential growth in the system and numerical instability. For this reason we turn our focus only on the full-rank (nine CLVs) method and the accuracy of the observations.

We use the following settings for all the DA experiments with ENSO observations: assimilation window 0.08, inflation factor 1 %, and 10 ensemble members. The model is run for the same amount of time as in the benchmark observation experiments. We study the effect of reducing the observation error variances in  $\mathbf{R}$ : standard observation er-



ror ( $\mathbf{R} = \text{diag}([1, 1, 25, 25])$ ), reduced tropical atmosphere error ( $\mathbf{R} = \text{diag}([0.01, 0.01, 25, 25])$ ), reduced ocean error ( $\mathbf{R} = \text{diag}([1, 1, 0.25, 0.25])$ ), and reduced overall error ( $\mathbf{R} = \text{diag}([0.01, 0.01, 0.25, 0.25])$ ). The summary statistics are listed in Table 4 for all the experiments.

We find that for the standard observation errors, there is a collapse in the variance of the ensemble mean for the extratropical subsystem (Fig. 10). However, when the observation error for the tropical subsystem is reduced, that variance is significantly increased. This can be seen through the increase in the average local Kaplan–Yorke dimension (Table 4). There is also a slight increase in the ability to track the extratropical subsystem of the control run. We note that decreasing the ocean observation error alone does not provide any improvements over the total error statistics or dimension, actually making them slightly worse. When the observation error of both subsystems is reduced, there is only a small improvement to the overall error statistics and dimension in comparison to the reduced tropical error case. The improvement is most notable for the cases with reduced tropical observation errors due to the tropical system’s weak direct coupling to the extratropical system.

### 5.5 Shadowed observations

In this section we explore a different type of observation error. Rather than randomly perturbing the control run to form our observation points, we use a trajectory that shadows the control run which produces correlated observational errors. In other words, we construct a new trajectory with a relaxation to the control run. This is implemented in the model as follows:

$$\dot{\tilde{x}}_e = \sigma(\tilde{y}_e - \tilde{x}_e) - c_e(S\tilde{x}_t + k_1), \quad (26a)$$

$$\dot{\tilde{y}}_e = \rho\tilde{x}_e - \tilde{y}_e - \tilde{x}_e\tilde{z}_e + c_e(S\tilde{y}_t + k_1) + \alpha_1(y_e - \tilde{y}_e), \quad (26b)$$

$$\dot{\tilde{z}}_e = \tilde{x}_e\tilde{y}_e - \beta\tilde{z}_e, \quad (26c)$$

$$\dot{\tilde{x}}_t = \sigma(\tilde{y}_t - \tilde{x}_t) - c(S\tilde{X} + k_2) - c_e(S\tilde{x}_e + k_1), \quad (26d)$$

$$\dot{\tilde{y}}_t = \rho\tilde{x}_t - \tilde{y}_t - \tilde{x}_t\tilde{z}_t + c(S\tilde{Y} + k_2) + c_e(S\tilde{y}_e + k_1) + \alpha_2(y_t - \tilde{y}_t), \quad (26e)$$

$$\dot{\tilde{z}}_t = \tilde{x}_t\tilde{y}_t - \beta\tilde{z}_t + c_z\tilde{Z}, \quad (26f)$$

$$\dot{\tilde{X}} = \tau\sigma(\tilde{Y} - \tilde{X}) - c(\tilde{x}_t + k_2), \quad (26g)$$

$$\dot{\tilde{Y}} = \tau\rho\tilde{X} - \tau\tilde{Y} - \tau S\tilde{X}\tilde{Z} + c(\tilde{y}_t + k_2) + \alpha_3(Y - \tilde{Y}), \quad (26h)$$

$$\dot{\tilde{Z}} = \tau S\tilde{X}\tilde{Y} - \tau\beta\tilde{Z} - c_z\tilde{z}_t. \quad (26i)$$

It is sufficient to constrain the trajectory using the relaxation term only in the  $y$  coordinates. The parameters  $\alpha_1 = 2.75$ ,  $\alpha_2 = 0.8$ , and  $\alpha_3 = 0.8$  are the relaxation strengths and  $y_e(t)$ ,  $y_t(t)$ , and  $Y(t)$  are taken from the control trajectory at the given time.

We initialize the shadowed trajectory with a small perturbation to the control trajectory initial condition. We then propagate the shadowed trajectory along with the control tra-

jectory, taking the observations from the shadowed trajectory at each assimilation cycle. Figure 11 shows the difference between the observations constructed using random perturbations and those from the shadowed trajectory. One benefit of the shadowed trajectory which is clearly visible in these figures is that it much more closely maintains the structure of the attractor. This is not the case when using randomized perturbations, where the structure is not as discernible.

We repeat the observation experiments of the previous three sections: benchmark observations, atmospheric observations, and ENSO observations. We only focus on the full-rank and variable CLV methods. When using correlated observational errors in any ensemble Kalman filtering method, a larger inflation value and ensemble size are needed to avoid ensemble collapse. We find that for the standard ETKF method increasing the ensemble size to 11 members is sufficient. To facilitate a direct comparison, we therefore also use this ensemble size for the CLV experiments. The inflation value varies slightly with the different observation cases. The setup and results of the experiments are shown in Table 5.

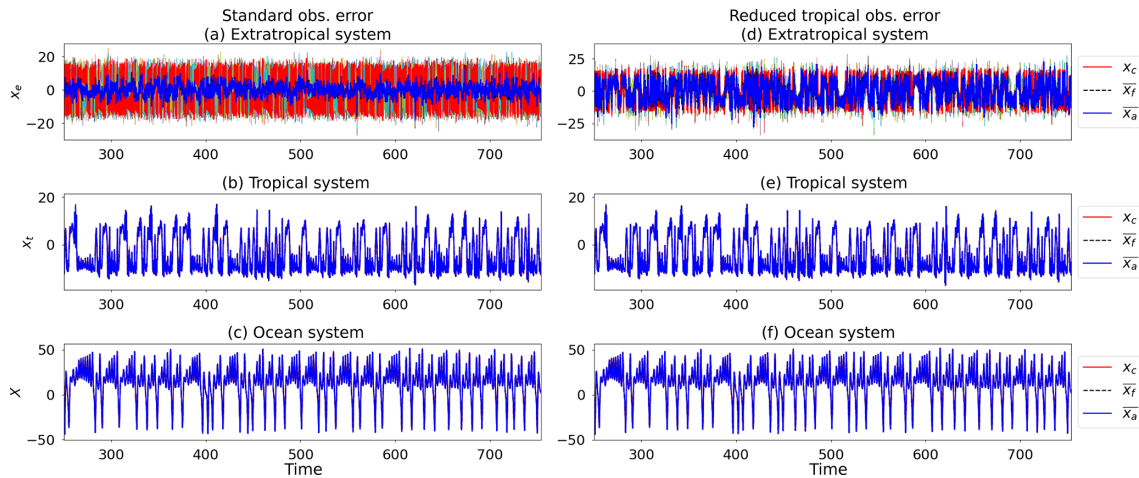
When using the benchmark observation set from the shadowed trajectory, the variable CLV method outperforms the full-rank method. This shows improvement over the case with random observation errors discussed in Sect. 5.2, where the variable CLV method performed slightly worse than the full rank. For the atmosphere observation case, we again observe similar performance when using the variable CLV method and the full rank, with the largest reduction of average RMSE in the ocean subsystem. The major change in the ENSO observation case from the random observational error experiments in Sect. 5.4 is that when using the shadowed trajectory observations the variable CLV method no longer fails. There is still difficulty in accurately calculating the first CLV from the extratropical subsystem being unconstrained; however, the correlation in the errors of the tropical and ocean subsystems provides additional information about the underlying nonlinear error growth. We observe that the variable CLV method performs comparably to the full rank, with a slightly larger average RMSE in the ocean subsystem (and subsequently the full). The inability to track the extratropical subsystem can once again be seen through the decrease in average local dimension.

### 5.6 Extratropical observations

We finally consider the case where one subsystem is fully observed and the others are completely unobserved. We choose to observe the extratropical subsystem, as it is the most extreme case with weakly coupled fast dynamics. Due to the difficulty of constraining the full system with such minimal observations, the assimilation window is reduced to 0.02 and we use *perfect observations*. In other words, we do not add any perturbations to the control run when making the observation. Having more accurate (mean) observations was

**Table 4.** Summary metrics of DA experiments using right-transform matrix (15) and ENSO observations ( $y_t, z_t, Y, Z$ ). Parameters: assimilation window 0.08, inflation factor 1 %, 10 ensemble members.

Method	Observations [error variance]	(RMSE) extratropical	(RMSE) tropical	(RMSE) ocean	(RMSE) full	(dim $\kappa_Y$ )
CLVs – nine	$y_t, z_t, Y, Z$ [1, 1, 25, 25]	7.0467	0.1571	0.3675	4.7182	4.5449
CLVs – nine	$y_t, z_t, Y, Z$ [0.01, 0.01, 25, 25]	5.5203	0.0480	0.1764	3.6860	5.4052
CLVs – nine	$y_t, z_t, Y, Z$ [1, 1, 0.25, 0.25]	7.1308	0.1386	0.0873	4.7566	4.4959
CLVs – nine	$y_t, z_t, Y, Z$ [0.01, 0.01, 0.25, 0.25]	5.1697	0.0431	0.0822	3.42478	5.4899



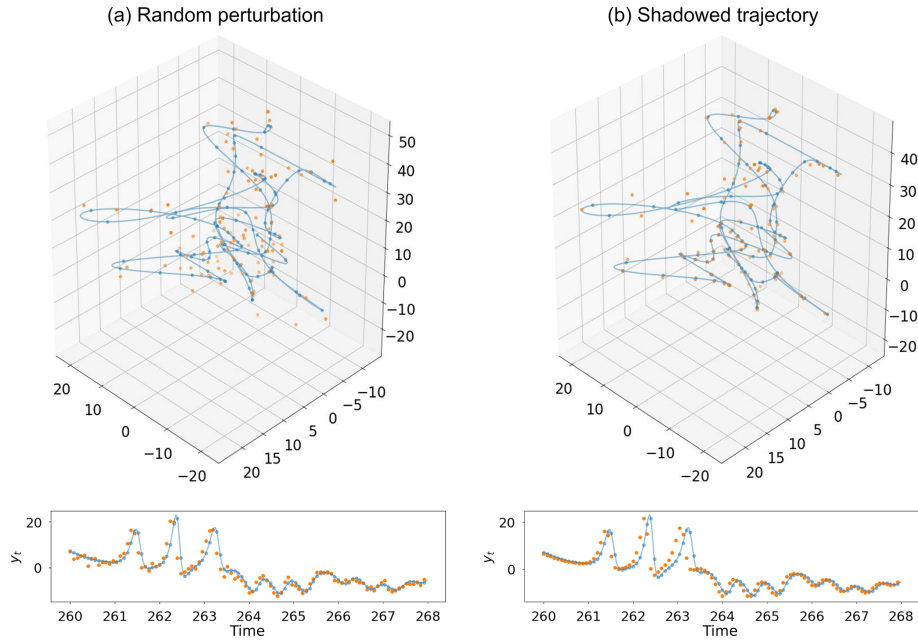
**Figure 10.** Trajectories of DA using nine CLVs,  $\mathbf{R} = \text{diag}([1, 1, 25, 25])$  (a–c), and  $\mathbf{R} = \text{diag}([0.01, 0.01, 25, 25])$  (d–f) with ENSO observations ( $y_t, z_t, Y, Z$ ). Trajectories shown are control run (red), ensemble mean (blue), and individual ensemble members (multicoloured). For conciseness we only show the results for the  $x$  coordinate of each subsystem. The other two coordinates behave similarly. Parameters: assimilation window 0.08, inflation factor 1 %, 10 ensemble members.

shown in Sect. 5.4 to improve the performance of the unobserved subsystems in the assimilation.

It is also clear from the previous experiments that the inability to constrain unobserved subsystems leads to a collapse in dimension and correspondingly a collapse in the covariances. A collapse in covariance is commonly avoided through the use of inflation (Anderson and Anderson, 1999; Hamill et al., 2001; Carrassi et al., 2008; Raanes et al., 2019). While a static background inflation avoids full covariance collapse, we are interested here in the covariance collapse related only to specific subsystems which are not being constrained. For such a case, we argue that the forecast error covariance matrix should be scaled by a factor relating to the ensemble performance at each analysis step. In other words, when an individual subsystem is not being constrained, the covariances should be increased in the calculation of the Kalman gain, analogous to the approach outlined by Miller et al. (1994) for the strongly nonlinear Lorenz '63 system.

In their study the authors find that the forecast error covariance is often underestimated in such highly nonlinear systems, particularly when the model is in a region of phase space subject to transitions. Subsequently, the Kalman gain is underestimated. The authors account for this by including the third and fourth moments of anomalies in the Kalman gain calculation.

Here we introduce a new method for adaptive scaling of the Kalman gain. Rather than explicitly calculating higher moments of the anomalies, we account for the underestimation through a spread-dependent factor which balances our forecast error covariance  $\mathbf{P}^f$  and observation error variance  $\mathbf{R}$  accordingly, the idea being that a larger spread implies we have underestimated the covariances, and vice versa. We note that this adaptive scaling is different to traditional inflation as it does not directly adjust the underlying ensemble spread. We use this in conjunction with the static background inflation of 1 % to avoid total ensemble collapse.



**Figure 11.** Comparison of the two different types of observations used in the data assimilation experiments. We show the observation space for the benchmark observations case ( $y_e$ ,  $y_t$ ,  $Y$ ). A selection of 800 time steps (100 observations) of the control model (blue line) and values at the observation times (blue dots) are shown in both plots, along with the actual observations used (orange dots). The top figures show behaviour along the attractor and the bottom panels show  $y_t$  in time. **(a)** Observations formed by random perturbations to control run points. **(b)** Observations taken from a trajectory which shadows the control run.

**Table 5.** Summary metrics of DA experiments using a right-transform matrix (Eq. 15) and shadowed trajectory as observations. We set the observation error covariances to the standard values as in Yoshida and Kalnay (2018). Parameters: assimilation window 0.08, 11 ensemble members, inflation as noted in the table.

Method	Observations [error variance]	$\langle \text{RMSE} \rangle$ extratropical	$\langle \text{RMSE} \rangle$ tropical	$\langle \text{RMSE} \rangle$ ocean	$\langle \text{RMSE} \rangle$ full	$\langle \text{dim}_{\text{KY}} \rangle$
CLVs – nine 3% inflation	$\tilde{y}_e, \tilde{y}_t, \tilde{Y}$ [1, 1, 25]	0.5200	0.1635	0.4255	0.4552	5.9576
CLVs – variable 3% inflation	$\tilde{y}_e, \tilde{y}_t, \tilde{Y}$ [1, 1, 25]	0.4848	0.1568	0.4095	0.4263	5.9371
CLVs – nine 3% inflation	$\tilde{y}_e, \tilde{z}_e, \tilde{y}_t, \tilde{z}_t$ [1, 1, 1, 1]	0.4258	0.1342	0.4300	0.4687	5.9237
CLVs – variable 3% inflation	$\tilde{y}_e, \tilde{z}_e, \tilde{y}_t, \tilde{z}_t$ [1, 1, 1, 1]	0.4337	0.1373	0.4027	0.4649	5.9191
CLVs – nine 4% inflation	$\tilde{y}_t, \tilde{z}_t, \tilde{Y}, \tilde{Z}$ [1, 1, 25, 25]	6.8613	0.1431	0.3025	4.5961	4.5309
CLVs – variable 4% inflation	$\tilde{y}_t, \tilde{z}_t, \tilde{Y}, \tilde{Z}$ [1, 1, 25, 25]	6.8722	0.1463	0.3370	4.6054	4.5541

We scale the Kalman gain in the following way:

$$\hat{\mathbf{K}} = \frac{\|\mathbf{P}^f\| \mathbf{P}^f \mathbf{H}^T}{\|\mathbf{P}^f\| \mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R}}, \quad (27)$$

or equivalently

$$\hat{\mathbf{K}} = \mathbf{P}^f \mathbf{H}^T \left[ \mathbf{H} \mathbf{P}^f \mathbf{H}^T + \frac{\mathbf{R}}{\|\mathbf{P}^f\|} \right]^{-1}. \quad (28)$$

Here the scaling factor is the Frobenius norm of  $\mathbf{P}^f = (\mathbf{X}^f)(\mathbf{X}^f)^T$ , where  $\mathbf{X}^f$  is the ensemble spread matrix defined

by Eq. (10b). This rescaling factor is mathematically similar to the  $\mathbf{K}$ -factor adaptive quality control procedure introduced by Sakov and Sandery (2017) and the  $\beta$ -factor rescaling of the background error covariances introduced by Bowler et al. (2013). The  $\mathbf{K}$ -factor method was used to account for inconsistencies in observations and therefore uses an adaptive observation error covariance  $\mathbf{R}$  that takes into account innovation size at each analysis step, while the  $\beta$ -factor is a deflation to the forecast error covariance matrix to avoid the underestimation of the ensemble spread ( $0 < \beta < 1$ ). Both the  $\mathbf{K}$ -factor procedure and the  $\beta$ -factor multiplier can be shown to have the same scaling effect on the Kalman gain  $\mathbf{K}$  defined by Eq. (9a) as the adaptive scaling presented here, with the difference in that the modified  $\hat{\mathbf{K}}$  in Eq. (28) takes into account both effects: small  $\|(\mathbf{X}^f)(\mathbf{X}^f)^T\|$  behaves like the  $\beta$ -factor and large  $\|(\mathbf{X}^f)(\mathbf{X}^f)^T\|$  behaves like the  $\mathbf{K}$ -factor. We discuss the limiting behaviour of this adaptive scaling method in terms of the increment size and analysis error covariance in Appendix A.

Due to the fact that only the Kalman gain is being adjusted, for ease of implementation we use the ESRF method introduced in Sect. 4.1. This allows for the left-transform matrix to be calculated with the modified Kalman gain:

$$\mathbf{T} = (\mathbf{I} - \hat{\mathbf{K}}\mathbf{H})^{1/2}. \quad (29)$$

We apply the adaptive gain to both the full-rank (nine CLVs) and variable-rank formulations of the covariance matrix. The results of the variable-rank experiments with and without adaptive gain are shown in Fig. 12 and the error statistics of all four experiments are listed in Table 6.

We see from Fig. 12 that there is a remarkable improvement when using the adaptive gain method. Not only is the ensemble spread reduced in the unobserved subsystems, but the ensemble mean is also able to track the control run. This improvement in tracking the control run is exemplified in Table 6 with a significant reduction in the average RMSE of individual subsystems. As expected, the average dimension is also significantly increased.

## 6 Concluding remarks

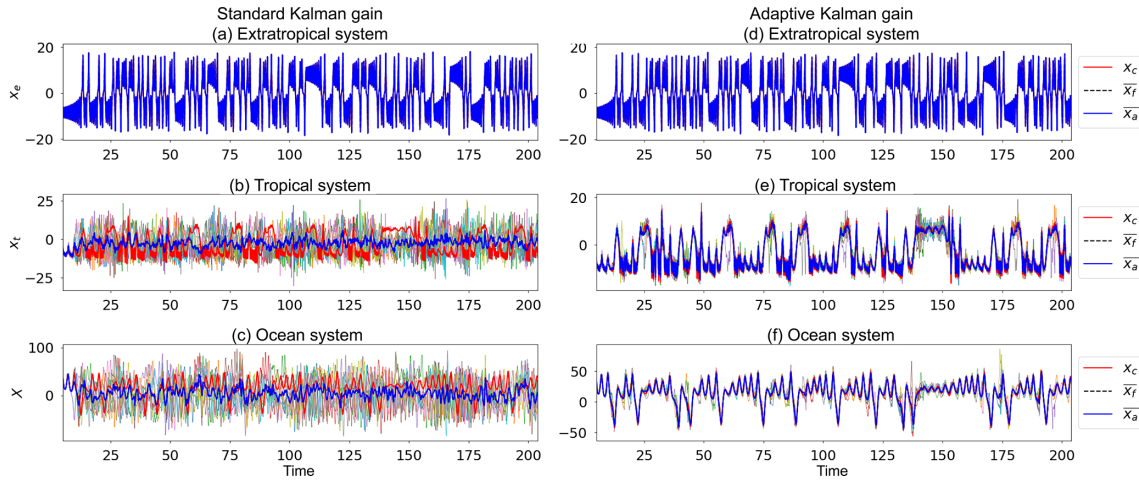
This study presents an initial understanding of the transient dynamics associated with the Kalman filter forecast error covariance matrix for nonlinear multiscale coupled systems. We have explored the varying rank of the error covariance matrix related to the transient growth in the stable modes of the system, and in particular the applicability of this varying rank to different configurations of strong CDA. Additionally, we have shown the large impact of using isolated observations and cross-domain covariances in such a coupled system. The cross-covariances are significantly underestimated when the observed subsystems are weakly coupled to the unobserved ones; however, this can be compensated through

either reduced observational error or the use of an adaptive scaling of the Kalman gain.

The dynamical properties of strongly coupled DA in a multiscale system were investigated through a low-dimensional nonlinear chaotic model to represent the interactions between the extratropical atmosphere, ocean, and tropical atmosphere–ocean interface. The model contains significant spatio-temporal scale separations between the subsystems, as well as varying coupling strengths. We introduced a local dimension measure, namely the Kaplan–Yorke dimension calculated using FTLEs, to specify the appropriate rank of the forecast error covariance matrix at each analysis step. We have shown that through using time-varying CLVs to form a reduced-rank forecast error covariance matrix, comparable results to the full-rank ETKF and ESRF schemes are achievable.

We considered a benchmark experiment previously explored in Yoshida and Kalnay (2018) to examine the most effective number of CLVs needed to form the forecast error covariance matrix. We found that when using less than full rank, the variable amount based on local dimension performed the best. We also found there was not significant improvement when increasing to full rank. In particular, we found that spanning the space comprised of the asymptotic unstable, neutral, and first weakly stable mode (five CLVs in this case) performed worse than using either dimension measure (asymptotic and local). This suggests that significant growth occurring in more than one weakly stable mode is important when capturing short-term dynamics of highly nonlinear systems. We therefore see improvement when implementing a rank based on local dimension over asymptotic dimension; however, all methods produce successful results in this case, where all subsystems are sampled in the observations.

We then tested the effectiveness of the reduced-rank forecast error covariance matrix in strong CDA when a subsystem is completely unobserved, i.e. using only cross-covariances to determine the increment of the unobserved system. The first set of these experiments used observations from the two atmosphere subsystems, extratropical and tropical, while the ocean was left completely unobserved. In this case we found that the DA succeeded in constraining the system to the observations when using the full rank, asymptotic dimension, and local dimension to determine the rank of the covariance matrix; however, the variable-rank method performed better than the fixed-rank methods. The second set of experiments consisted of ENSO observations or observations from the strongly coupled tropical and ocean subsystems only. In this case, the observational errors and weak coupling to the extratropical subsystem caused the reduced-rank experiment to fail. The full-rank experiment succeeded in tracking the tropical and ocean subsystems but left the extratropical subsystem unconstrained. This resulted in a collapse of the variance in the ensemble mean and a subsequent reduction in the average dimension. However, we found that



**Figure 12.** Trajectories of DA experiments using variable CLVs, left-transform matrix (Eq. 17b), and perfect observations from the extratropical subsystem of a control run ( $x_e, y_e, z_e$ ), with (a–c) the standard Kalman gain (Eq. 9a) and (d–f) the adaptive Kalman gain (Eq. 28). Trajectories shown are control run (red), ensemble mean (blue), and individual ensemble members (multicoloured). For conciseness we only show the results for the  $x$  coordinate of each subsystem. The other two coordinates behave similarly. Parameters: assimilation window 0.02, inflation factor 1 %, 10 ensemble members, observation error covariance  $\mathbf{R} = \mathbf{I}_4$ .

**Table 6.** Summary metrics of DA experiments using a left-transform matrix (Eq. 17b) and the full extratropical subsystem as observations ( $x_e, y_e, z_e$ ). We use perfect observations (no random error added to the control run) with the observation error covariances set to the standard values as in Yoshida and Kalnay (2018). Parameters: assimilation window 0.02, inflation factor 1 %, 10 ensemble members.

Method	Observations [error variance]	$\langle \text{RMSE} \rangle$ extratropical	$\langle \text{RMSE} \rangle$ tropical	$\langle \text{RMSE} \rangle$ ocean	$\langle \text{RMSE} \rangle$ full	$\langle \text{dim}_{\text{KY}} \rangle$
CLVs – nine	$x_e, y_e, z_e$ [1, 1, 1]	0.0640	8.4752	36.3662	21.7108	4.1332
CLVs – var	$x_e, y_e, z_e$ [1, 1, 1]	0.0670	9.4872	40.4894	24.1528	4.1152
CLVs – nine adaptive gain	$x_e, y_e, z_e$ [1, 1, 1]	0.0032	0.7241	3.5757	2.1504	5.9991
CLVs – var adaptive gain	$x_e, y_e, z_e$ [1, 1, 1]	0.0034	0.8350	4.0632	2.4501	5.8888

reducing the observational error variance of the tropical subsystem provided an increase in ensemble mean variance of the extratropical subsystem and therefore an increase in the average dimension. Reducing the observational error variance of the ocean did not provide a significant improvement since it is only indirectly coupled to the extratropical subsystem.

The effect of correlated observational errors was also explored. We constructed a trajectory which shadowed the control run and used this as our observational set, repeating all the previous experiments with different observation subsets. Since the correlated errors preserve the underlying dynamical structure of the system, we found that the reduced-rank method based on local dimension was the most successful in all the experiments when compared to those using random observational error. This included the ENSO observa-

tions case, where the extratropical subsystem remained unconstrained.

Finally, we showed that when only observing the extratropical subsystem, the unobserved subsystems could not be constrained due to their weak or indirect coupling to the observations. This manifested as an overall reduction in dimension as well as a collapse in the cross-domain covariances. In order to counter the covariance and dimension collapse, we introduced a novel scheme for adaptive Kalman gain scaling. This adaptive scaling is based on a measure of the overall spread of the system, therefore accounting for unobserved subsystems that have become unconstrained. Through use of the adaptive scaling the weakly coupled unobserved subsystems were able to be relatively constrained, and moreover the ensemble mean of the unobserved subsystems was able to track the control run. The adaptive scaling introduced here

should be tested on additional systems with weak coupling in order to assess its general applicability, although care may need to be taken in the choice of the norm.

We now turn to the implications for more realistic high-dimensional systems. It has been shown that when using a finer model resolution (increasing dimension) there is an increase in near-zero asymptotic Lyapunov exponents (De Cruz et al., 2018). We observed through the examination of the dynamical properties of the coupled Lorenz system that the stable yet near-zero exponents have the largest temporal variability, which affects the local dimension. As the number of near-zero exponents increases, we may expect that the temporal variability in dimension will increase further. This would have strong implications for the necessary rank of the forecast error covariance and the subsequent number of ensemble members. It is not implausible that the number of ensemble members could vary significantly in time. In such a case where the model's degree of freedom is much larger than its effective dimension, the projection onto CLVs becomes even more effective. This would ensure the ensemble perturbations lie in subspaces associated with error growth at the given time and that the directions of error growth are sufficiently sampled. Such improvements in modelling error growth of high-dimensional atmospheric systems have already been seen through the use of finite-time normal modes (Wei and Frederiksen, 2005). There is still more work to be done on how CLVs relate to meteorological and climatic events in such models, similar to the blocking studies of Schubert and Lucarini (2016). Future work should also consider the numerical cost of CLV calculation and methods to increase efficiency for high-dimensional systems.

The adaptive gain result presented here highlights the utility of ensemble filtering methods. While the ensemble mean of the subsystems manages to track the control run, the individual members are not so constrained. The variability in the spread of the ensemble members can provide a measure for uncertainty of the corresponding subsystem at a given region in phase space. Additionally, the ability to constrain the ensemble mean of the system from only observations of the weakly coupled fast subsystem is a new result for strong CDA. While dynamically this is intuitive (accurate knowledge of the fast dynamics of a system is sufficient to reconstruct the full attractor), this has not previously been shown to be achievable in DA experiments. If such a scheme could be shown to scale to high-dimensional climate models, then accurate and frequent atmospheric observations could potentially be sufficient to constrain the full system. For this reason it is important that the scheme be analysed for general applicability and tested on a variety of coupled dynamical systems.



### Appendix A: Limiting behaviour of adaptive Kalman gain scaling

We address the implication of the adaptive Kalman gain scaling for the two extreme cases: ensemble collapse ( $\|(\mathbf{X}^f)(\mathbf{X}^f)^T\| \rightarrow 0$ ) and ensemble divergence ( $\|(\mathbf{X}^f)(\mathbf{X}^f)^T\| \rightarrow \infty$ ).

1.  $\|(\mathbf{X}^f)(\mathbf{X}^f)^T\| \rightarrow 0$

We consider the Kalman gain in the form

$$\mathbf{K} = \delta \mathbf{P}^f \mathbf{H}^T [\delta \mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R}]^{-1}, \quad (\text{A1})$$

where  $\delta = \|(\mathbf{X}^f)(\mathbf{X}^f)^T\|$  and the subscript from Eq. (9a) have been dropped for brevity. For  $\delta \ll 1$ , Eq. (A1) can be expanded to

$$\begin{aligned} \mathbf{K} &= \delta \mathbf{P}^f \mathbf{H}^T [\mathbf{R}^{-1} - \delta \mathbf{R}^{-1} \mathbf{H} \mathbf{P}^f \mathbf{H}^T \mathbf{R}^{-1} + \mathcal{O}(\delta^2)] \\ &= \delta \mathbf{P}^f \mathbf{H}^T \mathbf{R}^{-1} - \delta^2 \mathbf{P}^f \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{P}^f \mathbf{H}^T \mathbf{R}^{-1} \\ &\quad + \mathcal{O}(\delta^3). \end{aligned} \quad (\text{A2})$$

Letting  $\delta \rightarrow 0$ , Eq. (A2) simplifies to  $\mathbf{K} = \mathbf{0}$ , the zero matrix. In this case the mean analysis increment and error covariance Eqs. (9b–9c) simplify to

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f, \quad (\text{A3})$$

$$\mathbf{P}^a = \mathbf{P}^f. \quad (\text{A4})$$

In other words, in the case of collapsed spread, the analysis is equal to the forecast.

2.  $\|(\mathbf{X}^f)(\mathbf{X}^f)^T\| \rightarrow \infty$

We consider the Kalman gain in the form

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T [\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \delta \mathbf{R}]^{-1}, \quad (\text{A5})$$

where  $\delta = \frac{1}{\|(\mathbf{X}^f)(\mathbf{X}^f)^T\|}$ . For  $\delta \ll 1$ , Eq. (A5) can be expanded to

$$\begin{aligned} \mathbf{K} &= \mathbf{P}^f \mathbf{H}^T [(\mathbf{H} \mathbf{P}^f \mathbf{H}^T)^{-1} - \delta (\mathbf{H} \mathbf{P}^f \mathbf{H}^T)^{-1} \\ &\quad \mathbf{R} (\mathbf{H} \mathbf{P}^f \mathbf{H}^T)^{-1} + \mathcal{O}(\delta^2)]. \end{aligned} \quad (\text{A6})$$

Letting  $\delta \rightarrow 0$ , Eq. (A6) simplifies to

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T)^{-1}. \quad (\text{A7})$$

In this case the mean analysis increment and error covariance Eqs. (9b–9c) become

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T)^{-1} (\mathbf{y} - \mathbf{H} \bar{\mathbf{x}}^f), \quad (\text{A8})$$

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T)^{-1} \mathbf{H}) \mathbf{P}^f. \quad (\text{A9})$$

Recall that  $\mathbf{y}$  are the observations defined by  $\mathbf{y} = \mathbf{H}\mathbf{x}$ , where  $\mathbf{x}$  is the truth (the observation error has implicitly been set to zero in the case of large spread). For a function  $\mathcal{F}$  with a Taylor series expansion operating on

two arbitrary matrices  $\mathbf{A}$  and  $\mathbf{B}$ , we have the following identity:

$$\mathcal{F}(\mathbf{A}\mathbf{B})\mathbf{A} = \mathbf{A}\mathcal{F}(\mathbf{B}\mathbf{A}). \quad (\text{A10})$$

Taking  $\mathbf{A} = \mathbf{H}$ ,  $\mathbf{B} = \mathbf{P}^f \mathbf{H}^T$ , and  $\mathcal{F}$  the inverse function, we can simplify (A8–A9) to

$$\bar{\mathbf{x}}^a = \mathbf{x}, \quad (\text{A11})$$

$$\mathbf{P}^a = \mathbf{0}. \quad (\text{A12})$$

In such a case, all ensemble members are adjusted to the same value as inferred from the observations.

*Data availability.* All simulation data are available from the corresponding author by request.

*Author contributions.* All the authors designed the study. The schemes for the calculation of Lyapunov exponents and CLVs were adapted and implemented by CQ in both Python and Matlab. The Python codes for the ensemble Kalman filtering methods were produced by VK, with modifications by CQ. All the figures were produced by CQ. All the authors contributed to the direction of the study, discussion of results, and the writing and approval of the manuscript.

*Competing interests.* The authors declare that they have no conflicts of interest.

*Acknowledgements.* The authors would like to thank Dylan Harries, Pavel Sakov, and Paul Sandery for their valuable input throughout the preparation of this paper, as well as the two anonymous reviewers for their thorough and insightful comments.

*Financial support.* The authors were supported by the Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO) Decadal Climate Forecasting Project (<https://research.csiro.au/dfp>, last access: 12 February 2020).

*Review statement.* This paper was edited by Alberto Carrassi and reviewed by two anonymous referees.

## References

- Abarbanel, H. D., Brown, R., and Kennel, M. B.: Variation of Lyapunov exponents on a strange attractor, *J. Nonlinear Sci.*, 1, 175–199, 1991.
- Anderson, J. L. and Anderson, S. L.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Mon. Weather Rev.*, 127, 2741–2758, 1999.
- Benettin, G., Galgani, L., and Strelcyn, J.-M.: Kolmogorov entropy and numerical experiments, *Phys. Rev. A*, 14, 2338, <https://doi.org/10.1103/PhysRevA.14.2338>, 1976.
- Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, *Mon. Weather Rev.*, 129, 420–436, 2001.
- Bocquet, M. and Carrassi, A.: Four-dimensional ensemble variational data assimilation and the unstable subspace, *Tellus A*, 69, 1304504, <https://doi.org/10.1080/16000870.2017.1304504>, 2017.
- Bocquet, M., Raanes, P. N., and Hannart, A.: Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation, *Nonlin. Processes Geophys.*, 22, 645–662, <https://doi.org/10.5194/npg-22-645-2015>, 2015.
- Bocquet, M., Gurumoorthy, K. S., Apte, A., Carrassi, A., Grudzien, C., and Jones, C. K.: Degenerate Kalman filter error covariances and their convergence onto the unstable subspace, *SIAM/ASA Journal on Uncertainty Quantification*, 5, 304–333, 2017.
- Bowler, N. E., Flowerdew, J., and Pring, S. R.: Tests of different flavours of EnKF on a simple model, *Q. J. Roy. Meteor. Soc.*, 139, 1505–1519, 2013.
- Carrassi, A., Trevisan, A., and Uboldi, F.: Adaptive observations and assimilation in the unstable subspace by breeding on the data-assimilation system, *Tellus A*, 59, 101–113, 2007.
- Carrassi, A., Vannitsem, S., Zupanski, D., and Zupanski, M.: The maximum likelihood ensemble filter performances in chaotic systems, *Tellus A*, 61, 587–600, 2008.
- Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, *Wiley Interdisciplinary Reviews: Climate Change*, 9, e535, 2018.
- De Cruz, L., Schubert, S., Demaeyer, J., Lucarini, V., and Vannitsem, S.: Exploring the Lyapunov instability properties of high-dimensional atmospheric and climate models, *Nonlin. Processes Geophys.*, 25, 387–412, <https://doi.org/10.5194/npg-25-387-2018>, 2018.
- Dieci, L., Russell, R. D., and Van Vleck, E. S.: On the computation of Lyapunov exponents for continuous dynamical systems, *SIAM J. Numer. Anal.*, 34, 402–423, 1997.
- Eckhardt, B. and Yao, D.: Local Lyapunov exponents in chaotic systems, *Physica D*, 65, 100–108, 1993.
- Eckmann, J.-P. and Ruelle, D.: Ergodic theory of chaos and strange attractors, in: *The theory of chaotic attractors*, 273–312, Springer, 1985.
- Evensen, G.: Advanced data assimilation for strongly nonlinear dynamics, *Mon. Weather Rev.*, 125, 1342–1354, 1997.
- Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dynam.*, 53, 343–367, 2003.
- Frederiksen, J. S.: Adjoint sensitivity and finite-time normal mode disturbances during blocking, *J. Atmos. Sci.*, 54, 1144–1165, 1997.
- Frederiksen, J. S.: Singular vectors, finite-time normal modes, and error growth during blocking, *J. Atmos. Sci.*, 57, 312–333, 2000.
- Frederickson, P., Kaplan, J. L., Yorke, E. D., and Yorke, J. A.: The Liapunov dimension of strange attractors, *J. Differ. Equations*, 49, 185–207, 1983.
- Froyland, G., Hüls, T., Morriss, G. P., and Watson, T. M.: Computing covariant Lyapunov vectors, Oseledets vectors, and dichotomy projectors: A comparative numerical study, *Physica D*, 247, 18–39, 2013.
- Ginelli, F., Poggi, P., Turchi, A., Chaté, H., Livi, R., and Politi, A.: Characterizing dynamics with covariant Lyapunov vectors, *Phys. Rev. Lett.*, 99, 130601, <https://doi.org/10.1103/PhysRevLett.99.130601>, 2007.
- Gritsun, A. and Lucarini, V.: Fluctuations, response, and resonances in a simple atmospheric model, *Physica D*, 349, 62–76, 2017.
- Grudzien, C., Carrassi, A., and Bocquet, M.: Asymptotic forecast uncertainty and the unstable subspace in the presence of additive model error, *SIAM/ASA Journal on Uncertainty Quantification*, 6, 1335–1363, 2018a.
- Grudzien, C., Carrassi, A., and Bocquet, M.: Chaotic dynamics and the role of covariance inflation for reduced rank Kalman fil-

- ters with model error, *Nonlin. Processes Geophys.*, 25, 633–648, <https://doi.org/10.5194/npg-25-633-2018>, 2018b.
- Gurumoorthy, K. S., Grudzien, C., Apte, A., Carrassi, A., and Jones, C. K.: Rank deficiency of Kalman error covariance matrices in linear time-varying system with deterministic evolution, *SIAM J. Control Optim.*, 55, 741–759, 2017.
- Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Mon. Weather Rev.*, 129, 2776–2790, 2001.
- Han, G., Wu, X., Zhang, S., Liu, Z., and Li, W.: Error covariance estimation for coupled data assimilation using a Lorenz atmosphere and a simple pycnocline ocean model, *J. Climate*, 26, 10218–10231, 2013.
- Houtekamer, P. L. and Zhang, F.: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation, *Mon. Weather Rev.*, 144, 4489–4532, 2016.
- Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D*, 230, 112–126, 2007.
- Kang, J.-S., Kalnay, E., Liu, J., Fung, I., Miyoshi, T., and Ide, K.: “Variable localization” in an ensemble Kalman filter: Application to the carbon cycle data assimilation, *J. Geophys. Res.-Atmos.*, 116, <https://doi.org/10.1029/2010JD014673>, 2011.
- Kaplan, J. L. and Yorke, J. A.: Chaotic behavior of multidimensional difference equations, in: *Functional Differential equations and approximation of fixed points*, 204–227, Springer, 1979.
- Kwasniok, F.: Fluctuations of finite-time Lyapunov exponents in an intermediate-complexity atmospheric model: a multivariate and large-deviation perspective, *Nonlin. Processes Geophys.*, 26, 195–209, <https://doi.org/10.5194/npg-26-195-2019>, 2019.
- Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K., and Janssen, P.: A coupled data assimilation system for climate reanalysis, *Q. J. Roy. Meteor. Soc.*, 142, 65–78, 2016.
- Lorenz, E. N.: Deterministic nonperiodic flow, *J. Atmos. Sci.*, 20, 130–141, 1963.
- Lorenz, E. N.: Predictability: A problem partly solved, in: *Proc. Seminar on predictability*, ECMWF, Reading, UK, vol. 1, 1996.
- Miller, R. N., Ghil, M., and Gauthiez, F.: Advanced data assimilation in strongly nonlinear dynamical systems, *J. Atmos. Sci.*, 51, 1037–1056, 1994.
- Nese, J. M. and Dutton, J. A.: Quantifying predictability variations in a low-order ocean-atmosphere model: a dynamical systems approach, *J. Climate*, 6, 185–204, 1993.
- Ng, G.-H. C., McLaughlin, D., Entekhabi, D., and Ahanin, A.: The role of model dynamics in ensemble Kalman filter performance for chaotic systems, *Tellus A*, 63, 958–977, 2011.
- Norwood, A., Kalnay, E., Ide, K., Yang, S.-C., and Wolfe, C.: Lyapunov, singular and bred vectors in a multi-scale system: an empirical exploration of vectors related to instabilities, *J. Phys. A*, 46, 254021, <https://doi.org/10.1088/1751-8113/46/25/254021>, 2013.
- Oseledets, V. I.: A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems, *Trudy Moskovskogo Matematicheskogo Obschestva*, 19, 179–210, 1968.
- O’Kane, T. and Frederiksen, J.: Comparison of statistical dynamical, square root and ensemble Kalman filters, *Entropy*, 10, 684–721, 2008.
- O’Kane, T. J., Sandery, P. A., Monselesan, D. P., Sakov, P., Chamberlain, M. A., Matear, R. J., Collier, M. A., Squire, D. T., and Stevens, L.: Coupled Data Assimilation and Ensemble Initialization with Application to Multiyear ENSO Prediction, *J. Climate*, 32, 997–1024, 2019.
- Palatella, L. and Trevisan, A.: Interaction of Lyapunov vectors in the formulation of the nonlinear extension of the Kalman filter, *Phys. Rev. E*, 91, 042905, <https://doi.org/10.1103/PhysRevE.91.042905>, 2015.
- Palatella, L., Carrassi, A., and Trevisan, A.: Lyapunov vectors and assimilation in the unstable subspace: theory and applications, *J. Phys. A*, 46, 254020, <https://doi.org/10.1088/1751-8113/46/25/254020>, 2013.
- Peña, M. and Kalnay, E.: Separating fast and slow modes in coupled chaotic systems, *Nonlin. Processes Geophys.*, 11, 319–327, <https://doi.org/10.5194/npg-11-319-2004>, 2004.
- Penny, S., Bach, E., Bhargava, K., Chang, C.-C., Da, C., Sun, L., and Yoshida, T.: Strongly coupled data assimilation in multiscale media: experiments using a quasi-geostrophic coupled model, *J. Adv. Model. Earth Sy.*, 11, 1803–1829, 2019.
- Raanes, P. N., Bocquet, M., and Carrassi, A.: Adaptive covariance inflation in the ensemble Kalman filter by Gaussian scale mixtures, *Q. J. Roy. Meteor. Soc.*, 145, 53–75, 2019.
- Sakov, P. and Oke, P. R.: Implications of the form of the ensemble transformation in the ensemble square root filters, *Mon. Weather Rev.*, 136, 1042–1053, 2008.
- Sakov, P. and Sandery, P.: An adaptive quality control procedure for data assimilation, *Tellus A*, 69, 1318031, <https://doi.org/10.1080/16000870.2017.1318031>, 2017.
- Sakov, P., Oliver, D. S., and Bertino, L.: An iterative EnKF for strongly nonlinear systems, *Mon. Weather Rev.*, 140, 1988–2004, 2012.
- Sano, M. and Sawada, Y.: Measurement of the Lyapunov spectrum from a chaotic time series, *Phys. Rev. Lett.*, 55, 1082, <https://doi.org/10.1103/PhysRevLett.55.1082>, 1985.
- Schubert, S. and Lucarini, V.: Covariant Lyapunov vectors of a quasi-geostrophic baroclinic model: analysis of instabilities and feedbacks, *Q. J. Roy. Meteor. Soc.*, 141, 3040–3055, 2015.
- Schubert, S. and Lucarini, V.: Dynamical analysis of blocking events: spatial and temporal fluctuations of covariant Lyapunov vectors, *Q. J. Roy. Meteor. Soc.*, 142, 2143–2158, 2016.
- Sluka, T. C., Penny, S. G., Kalnay, E., and Miyoshi, T.: Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation, *Geophys. Res. Lett.*, 43, 752–759, 2016.
- Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S.: Ensemble square root filters, *Mon. Weather Rev.*, 131, 1485–1490, 2003.
- Trevisan, A. and Palatella, L.: On the Kalman Filter error covariance collapse into the unstable subspace, *Nonlin. Processes Geophys.*, 18, 243–250, <https://doi.org/10.5194/npg-18-243-2011>, 2011.
- Trevisan, A. and Uboldi, F.: Assimilation of standard and targeted observations within the unstable subspace of the observation–forecast cycle system, *J. Atmos. Sci.*, 61, 103–113, 2004.
- Trevisan, A., D’Isidoro, M., and Talagrand, O.: Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension, *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences*, *Appl. Meteorol. Phys. Oceanogr.*, 136, 487–496, 2010.

- Vannitsem, S.: Predictability of large-scale atmospheric motions: Lyapunov exponents and error dynamics, *Chaos: An Interdisciplinary J. Nonlinear Sci.*, 27, 032101, <https://doi.org/10.1063/1.4979042>, 2017.
- Vannitsem, S. and Lucarini, V.: Statistical and dynamical properties of covariant Lyapunov vectors in a coupled atmosphere-ocean model—multiscale effects, geometric degeneracy, and error dynamics, *J. Phys. A*, 49, 224001, <https://doi.org/10.1088/1751-8113/49/22/224001>, 2016.
- Wei, M. and Frederiksen, J. S.: Finite-time normal mode disturbances and error growth during Southern Hemisphere blocking, *Adv. Atmos. Sci.*, 22, 69–89, 2005.
- Wolfe, C. L. and Samelson, R. M.: An efficient method for recovering Lyapunov vectors from singular vectors, *Tellus A*, 59, 355–366, 2007.
- Yoden, S. and Nomura, M.: Finite-time Lyapunov stability analysis and its application to atmospheric predictability, *J. Atmos. Sci.*, 50, 1531–1543, 1993.
- Yoshida, T. and Kalnay, E.: Correlation-Cutoff Method for Covariance Localization in Strongly Coupled Data Assimilation, *Mon. Weather Rev.*, 146, 2881–2889, 2018.
- Young, L.-S.: Dimension, entropy and Lyapunov exponents, *Ergod. Theor. Dyn. Syst.*, 2, 109–124, 1982.