Check for updates

Four-dimensional ensemble variational data assimilation and the unstable subspace

By MARC BOCQUET[1]* and ALBERTO CARRASSI[2], [1]*CEREA, joint laboratory École des Ponts ParisTech and EDF R&D, Université Paris-Est, Champs-sur-Marne, France;* [2]*Nansen Environmental and Remote Sensing Center, Bergen, Norway*

ABSTRACT

The performance of (ensemble) Kalman filters used for data assimilation in the geosciences critically depends on the dynamical properties of the evolution model. A key aspect is that the error covariance matrix is asymptotically supported by the unstable–neutral subspace only, i.e. it is spanned by the backward Lyapunov vectors with non-negative exponents. The analytic proof of such a property for the Kalman filter error covariance has been recently given, and in particular that of its confinement to the unstable–neutral subspace. In this paper, we first generalize those results to the case of the Kalman smoother in a linear, Gaussian and perfect model scenario. We also provide square-root formulae for the filter and smoother that make the connection with ensemble formulations of the Kalman filter and smoother, where the span of the error covariance is described in terms of the ensemble deviations from the mean. We then discuss how this neat picture is modified when the dynamics are nonlinear and chaotic, and for which analytic results are precluded or difficult to obtain. A numerical investigation is carried out to study the approximate confinement of the anomalies for both a deterministic ensemble Kalman filter (EnKF) and a four-dimensional ensemble variational method, the iterative ensemble Kalman smoother (IEnKS), in a perfect model scenario. The confinement is characterized using geometrical angles that determine the relative position of the anomalies with respect to the unstable–neutral subspace. The alignment of the anomalies and of the unstable–neutral subspace is more pronounced when observation precision or frequency, as well as the data assimilation window length for the IEnKS, are increased. These results also suggest that the IEnKS and the deterministic EnKF realize in practice (albeit implicitly) the paradigm behind the approach of Anna Trevisan and co-authors known as the assimilation in the unstable subspace.

Keywords: assimilation in the unstable subspace, ensemble Kalman filter, iterative ensemble Kalman smoother, four-dimensional ensemble variational method, Lyapunov vectors

1. Introduction

The unstable–neutral subspace of a dynamical system is defined as the vector space generated by the backward Lyapunov vectors with non-negative Lyapunov exponents (Legras and Vautard, 1996), i.e. the space of the small perturbations that are not growing exponentially under the action of the backward dynamics. Conversely, the stable subspace is defined here and differently from Legras and Vautard (1996) as the vector space spanned by the backward Lyapunov vectors with a negative Lyapunov exponent.

This splitting of the state space in terms of unstable and stable subspaces was found to play an important role in the understanding of data assimilation (DA) algorithms for nonlinear chaotic dynamics. The assimilation in the unstable subspace (AUS),

introduced by Anna Trevisan and collaborators (Trevisan and Uboldi, 2004; Carrassi et al., 2007, 2008a; Trevisan et al., 2010; Trevisan and Palatella, 2011; Palatella and Trevisan, 2015), has shed light on an efficient way to operate the assimilation of observations using the unstable subspace to describe the uncertainty in the state's estimate. AUS is based on two key properties of deterministic, typically dissipative, chaotic systems: (i) the perturbations tend to project on the unstable manifold of the dynamics, and (ii) the dimension of the unstable manifold is typically much smaller than the full state-space dimension. Applications to atmospheric and oceanic models (Uboldi and Trevisan, 2006; Carrassi et al., 2008b) showed that even for high-dimensional models, an efficient error control is achieved by monitoring only the unstable directions, making AUS a computationally efficient alternative to standard procedures in this

---

*Corresponding author. e-mail: marc.bocquet@enpc.fr

1

context. The AUS approach has recently motivated a research effort that has led to a proper mathematical formulation and assessment of its driving idea, i.e. the span of the estimation error covariance matrices asymptotically (in time) tends to the subspace spanned by the unstable and neutral backward Lyapunov vectors (Bocquet et al., 2017; Gurumoorthy et al., 2017).

The goal of this work is to study the extension of these convergence properties of the error uncertainty to the popular family of ensemble-based methods used in high-dimensional nonlinear geophysical applications. Our focus here is on both ensemble filter and smoother, with the latter gaining progressively more attention by virtue of their higher accuracy and suitability for reanalysis purposes. To that end, we will first look analytically at the Kalman filter (KF, Kalman, 1960) and smoother (KS, Anderson and Moore, 1979; Cohn et al., 1994), when the observation and evolution models are linear. We will study the general case when the initial error covariance matrix is of arbitrary rank. The KF will serve as a linear proxy for the ensemble Kalman filter (EnKF, Evensen, 2009) whereas the KS will do that for the iterative ensemble Kalman smoother (IEnKS, Bocquet and Sakov, 2014). The analytic results for the linear case will help to frame and interpret the nonlinear one, in which the convergence properties of the EnKF and IEnKS will be studied numerically.

In Section 2, we introduce an alternative derivation of some of the results of the degenerate (or rank-deficient) KF onto the unstable and neutral subspace, obtained in Bocquet et al. (2017) (later Boc17). The filter is presented in an ensemble square-root form to make the link with the ensemble square-root EnKF. In Section 3, we generalize the results obtained for the degenerate KF to the KS. Again, the smoother is presented in an ensemble square-root form to make the link with the ensemble square-root Kalman smoother (EnKS) or with the IEnKS. In Section 4, we recall the algorithm of the IEnKS, seen as an archetype of four-dimensional ensemble variational methods meant for filtering and smoothing with chaotic models. Elementary results about backward and covariant Lyapunov vectors are also recalled. They are put to use in our subsequent numerical investigation of the approximate convergence of the ensemble of the EnKF and IEnKS onto the unstable–neutral subspace using a low-order chaotic model. This convergence is characterized geometrically in regimes that deviate to different extents from the linear and Gaussian assumptions on which the analytic results are built. Conclusions are drawn in Section 5.

2. Linear case – convergence of the degenerate Kalman filter: analytic results

In this section, we recall and possibly re-derive some of the results obtained for the linear and perfect model case, by Gurumoorthy et al. (2017) and later generalized by Boc17, in particular to the degenerate KF, i.e. with an initial covariance matrix of arbitrary rank. In the present study, we will express them using the square-root formalism in order to make the

connection with the ensemble square-root KF typically used in high-dimensional systems (Tippett et al., 2003; Evensen, 2009).

The purpose is the estimation of the unknown state of a system based on partial and noisy observations. Throughout the entire text, the conventional notation $k = 0, 1, 2, \ldots$ is adopted to indicate that quantities are defined at times $t_0, t_1, t_2, \ldots$. The dynamical and observational models are both assumed to be linear, and expressible as

$$\mathbf{x}_k = \mathbf{M}_k \mathbf{x}_{k-1}, \tag{1}$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \tag{2}$$

with $\mathbf{x}_k \in \mathbb{R}^n$ and $\mathbf{y}_k \in \mathbb{R}^{d_k}$ being the system's state and observation vector, respectively, related by the linear observation operator $\mathbf{H}_k : \mathbb{R}^n \to \mathbb{R}^{d_k}$. The matrix $\mathbf{M}_{k:l}$ represents the linear action of the linear forward model from time t_l to time t_k, and is assumed to be non-singular throughout this paper. In particular $\mathbf{M}_{k:k} = \mathbf{I}_n$, where $\mathbf{I}_n$ is the n-by-n identity matrix. Moreover, $\mathbf{M}_k$ designates the 1-step action of the forward model from t_{k-1} to t_k: $\mathbf{M}_k \triangleq \mathbf{M}_{k:k-1}$ and, consequently $\mathbf{M}_{k:l} = \mathbf{M}_k \mathbf{M}_{k-1} \ldots \mathbf{M}_{l+1}$, with the symbol $\triangleq$ meaning that the expression is a definition. The Lyapunov spectrum of the dynamics defined by $\mathbf{M}_{k:0}$ is assumed to be non-degenerate, i.e. the Lyapunov exponents are all distinct. This assumption is made to avoid lengthy but not critical developments; in fact most of the results in this paper can be generalized to the degenerate spectrum case.

The observation noise vectors $\{\mathbf{v}_k\}_{k=0,1,2,\ldots}$ are assumed to form an unbiased Gaussian white sequence with statistics

$$\mathrm{E}[\mathbf{v}_k \mathbf{v}_l^{\mathrm{T}}] = \delta_{k,l} \mathbf{R}_k, \tag{3}$$

where $\mathbf{R}_k \in \mathbb{R}^{d_k \times d_k}$ is the observation error covariance matrix at time t_k. The error covariance matrix $\mathbf{R}_k$ can be assumed invertible without losing generality. The errors on the initial state, $\mathbf{x}_0$, are assumed Gaussian and unbiased as well, and they are independent from those of the observations.

2.1. Recurrence

The forecast error covariance matrix $\mathbf{P}_k$ of the KF satisfies the following dynamical Ricatti equation (Kalman, 1960)

$$\mathbf{P}_{k+1} = \mathbf{M}_{k+1} \left(\mathbf{I}_n + \mathbf{P}_k \mathbf{\Omega}_k \right)^{-1} \mathbf{P}_k \mathbf{M}_{k+1}^{\mathrm{T}}, \tag{4}$$

where

$$\mathbf{\Omega}_k \triangleq \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k \tag{5}$$

is the precision of the observations transferred to state space. The results of Gurumoorthy et al. (2017) and Boc17 have been

formulated in terms of $\mathbf{P}_k$. Instead, we seek here to derive them in terms of normalized anomalies, i.e. in terms of a matrix $\mathbf{X}_k$ in $\mathbb{R}^{n \times N}$ such that $\mathbf{P}_k = \mathbf{X}_k \mathbf{X}_k^T$. The term 'anomaly' is borrowed here from the jargon of ensemble DA where the columns of $\mathbf{X}_k$ contain the ensemble members deviation from the ensemble mean, normalized by $\sqrt{N-1}$. With this decomposition, Equation (4) can be reformulated as

$$
\begin{aligned}
\mathbf{X}_{k+1}\mathbf{X}_{k+1}^T &= \mathbf{M}_{k+1}\left(\mathbf{I}_n + \mathbf{X}_k\mathbf{X}_k^T\mathbf{\Omega}_k\right)^{-1}\mathbf{X}_k\mathbf{X}_k^T\mathbf{M}_{k+1}^T \\
&= \mathbf{M}_{k+1}\mathbf{X}_k\left(\mathbf{I}_N + \mathbf{X}_k^T\mathbf{\Omega}_k\mathbf{X}_k\right)^{-1}\mathbf{X}_k^T\mathbf{M}_{k+1}^T \quad (6)
\end{aligned}
$$

where we used the matrix shift lemma[1] to get the second line. From Equation (6), we have (Bishop et al., 2001)

$$
\mathbf{X}_{k+1} = \mathbf{M}_{k+1}\mathbf{X}_k\left(\mathbf{I}_N + \mathbf{X}_k^T\mathbf{\Omega}_k\mathbf{X}_k\right)^{-\frac{1}{2}}\mathbf{\Psi}_k, \quad (7)
$$

where $\mathbf{\Psi}_k$ is an arbitrary orthogonal matrix in $\mathbb{R}^{N \times N}$. We further impose that $\mathbf{\Psi}_k\mathbf{1} = \mathbf{1}$, where $\mathbf{1} = (1, \ldots, 1)^T$ is the column vector in $\mathbb{R}^n$ of entries 1, i.e. $\mathbf{1}$ is an eigenvector of $\mathbf{\Psi}_k$ of eigenvalue 1. Assuming $\mathbf{X}_0\mathbf{1} = \mathbf{0}$, i.e. the anomalies have zero mean, this additional condition enforces $\mathbf{X}_k\mathbf{1} = \mathbf{0}$ for $k \geq 1$ through the recursion Equation (7). In particular, the updated anomalies are centred on the analysis (Sakov and Oke, 2008), i.e. $\mathbf{X}_k^a\mathbf{1} = \mathbf{0}$, where

$$
\mathbf{X}_k^a = \mathbf{X}_k\left(\mathbf{I}_N + \mathbf{X}_k^T\mathbf{\Omega}_k\mathbf{X}_k\right)^{-\frac{1}{2}}\mathbf{\Psi}_k. \quad (8)
$$

2.2. Explicit covariance dependence on the initial condition and the observability matrix for the Kalman filter

One of the analytic results in Boc17 is the explicit formulation of $\mathbf{P}_k$ in terms of $\mathbf{P}_0$,

$$
\mathbf{P}_k = \mathbf{M}_{k:0}\mathbf{P}_0\left[\mathbf{I}_n + \mathbf{\Theta}_k\mathbf{P}_0\right]^{-1}\mathbf{M}_{k:0}^T \quad (9)
$$

where

$$
\mathbf{\Theta}_k \triangleq \sum_{l=0}^{k-1}\mathbf{M}_{l:0}^T\mathbf{\Omega}_l\mathbf{M}_{l:0}. \quad (10)
$$

Equation (9) is valid even if $\mathbf{P}_0$ is degenerate. The $\mathbf{\Theta}_k$ matrix is a measure of the observability of the state, but computed at the initial time t_0. Using again the matrix shift lemma, one can derive the square-root counterpart of Equation (9), which reads

$$
\mathbf{X}_k = \mathbf{M}_{k:0}\mathbf{X}_0\left[\mathbf{I}_N + \mathbf{X}_0^T\mathbf{\Theta}_k\mathbf{X}_0\right]^{-\frac{1}{2}}\mathbf{\Psi}_k \quad (11)
$$

where $\mathbf{\Psi}_k$ is an arbitrary orthogonal matrix which could differ from that used in Equation (7) and such that $\mathbf{\Psi}_k\mathbf{1} = \mathbf{1}$. Equation (11) is a right-transform formula from $\mathbf{X}_0$ to $\mathbf{X}_k$, since the transform matrix $\left[\mathbf{I}_N + \mathbf{X}_0^T\mathbf{\Theta}_k\mathbf{X}_0\right]^{-\frac{1}{2}}\mathbf{\Psi}_k$ acts to the right, i.e. in ensemble subspace, of the anomalies $\mathbf{X}_0$. Thus, the anomalies at t_k are the outcome of a long forecast from t_0 to t_k of the updated anomalies at t_0 that account for all observations. Alternatively, it is possible to use the matrix shift lemma to obtain the equivalent left-transform formula

$$
\mathbf{X}_k = \left[\mathbf{I}_n + \mathbf{M}_{k:0}\mathbf{P}_0\mathbf{M}_{k:0}^T\mathbf{\Gamma}_k\right]^{-\frac{1}{2}}\mathbf{M}_{k:0}\mathbf{X}_0\mathbf{\Psi}_k, \quad (12)
$$

where

$$
\mathbf{\Gamma}_k \triangleq \mathbf{M}_{k:0}^{-T}\mathbf{\Theta}_k\mathbf{M}_{k:0}^{-1} = \sum_{l=0}^{k-1}\mathbf{M}_{k:l}^{-T}\mathbf{\Omega}_l\mathbf{M}_{k:l}^{-1}, \quad (13)
$$

is known as the *information* matrix (Jazwinski, 1970)[2] and is a measure of the observability of the state at t_k. The transform matrix now acts in state-space rather than in ensemble subspace. Differently, although equivalently, the left-transform performs the analysis with $\mathbf{M}_{k:0}\mathbf{X}_0$ rather than $\mathbf{X}_0$, but does not make any forecast of the anomalies afterwards.

Equations (11) and (12) make explicit the analytic dependence of $\mathbf{X}_k$ at any time t_k on $\mathbf{X}_0$.

2.3. Variational correspondence

If $\mathbf{P}_k$ is full-rank, the analysis is also known to be the outcome of the minimization of the cost function

$$
\mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{y}_k - \mathbf{H}_k\mathbf{x}\|_{\mathbf{R}_k}^2 + \frac{1}{2}\|\mathbf{x} - \bar{\mathbf{x}}_k\|_{\mathbf{P}_k}^2 \quad (14)
$$

where $\|\mathbf{x}\|_{\mathbf{A}}^2 = \mathbf{x}^T\mathbf{A}^{-1}\mathbf{x}$ and $\bar{\mathbf{x}}_k$ is the forecast state vector at time t_k. The analysis state $\mathbf{x}_k^a$ is the argument of the minimum of $\mathcal{J}$ and the inverse of the analysis error covariance matrix $\mathbf{P}_k^a$ is the Hessian of $\mathcal{J}$. In the general case where $\mathbf{P}_k$ can be degenerate and is decomposed as $\mathbf{P}_k = \mathbf{X}_k\mathbf{X}_k^T$, the analysis is actually performed in the ensemble (affine) subspace

$$
\mathcal{V}_k = \left\{\bar{\mathbf{x}}_k + \mathbf{X}_k\mathbf{w}, \mathbf{w} \in \mathbb{R}^N\right\}. \quad (15)
$$

The analysis in $\mathcal{V}_k$ is then the outcome of the minimization of the reduced cost function (Hunt et al., 2007; Bocquet, 2011)

$$
\mathcal{L}(\mathbf{w}) = \frac{1}{2}\|\mathbf{y}_k - \mathbf{H}_k\left(\bar{\mathbf{x}}_k + \mathbf{X}_k\mathbf{w}\right)\|_{\mathbf{R}_k}^2 + \frac{1}{2}\|\mathbf{w}\|^2 \quad (16)
$$

where $\|\mathbf{w}\|^2 = \mathbf{w}^T \mathbf{w}$. The second term in the cost function represents the prior in the ensemble subspace $\mathcal{V}_k$. The analysis state is given by $\overline{\mathbf{x}}_k + \mathbf{X}_k \mathbf{w}_a$ where $\mathbf{w}_a = \operatorname{argmin} \mathcal{L}(\mathbf{w})$. The updated anomalies are given by

$$\mathbf{X}_k^a = \mathbf{X}_k \mathbf{T}^{-\frac{1}{2}} \mathbf{\Psi}_k, \qquad (17)$$

where $\mathbf{T} = \mathbf{I}_N + \mathbf{X}_k^T \mathbf{\Omega}_k \mathbf{X}_k \in \mathbb{R}^{N \times N}$ is the Hessian of $\mathcal{L}(\mathbf{w})$ and $\mathbf{\Psi}_k$ is an arbitrary orthogonal matrix in $\mathbb{R}^{N \times N}$ such that $\mathbf{\Psi}_k \mathbf{1} = \mathbf{1}$. This update formula coincides with Equation (8) and the inverse square root of the Hessian is pivotal in the computation of $\mathbf{P}_k^a$. This variational correspondence will also be exploited in Section 3 to easily get $\mathbf{P}_k^a$ for the KS.

2.4. Confinement of the Kalman filter to the unstable subspace

The main results of Boc17 are built upon Equation (9), which can be rewritten using the information matrix Equation (13), as

$$\mathbf{P}_k = \mathbf{M}_{k:0} \mathbf{P}_0 \mathbf{M}_{k:0}^T \left(\mathbf{I}_n + \mathbf{\Gamma}_k \mathbf{M}_{k:0} \mathbf{P}_0 \mathbf{M}_{k:0}^T \right)^{-1}. \qquad (18)$$

This shows that the dependence of $\mathbf{P}_k$ on $\mathbf{P}_0$ is also entirely determined through the rational dependence of $\mathbf{P}_k$ in the free – i.e. observationally-unconstrained – forecast of the initial covariances $\mathbf{M}_{k:0} \mathbf{P}_0 \mathbf{M}_{k:0}^T$ and in the information matrix $\mathbf{\Gamma}_k$.

Equation (18) was shown to imply (Boc17)

$$\mathbf{P}_k \leq \mathbf{M}_{k:0} \mathbf{P}_0 \mathbf{M}_{k:0}^T, \qquad (19)$$

where the order relationship is that of the cone of the positive semi-definite matrices (see for instance Appendix B of Boc17 for a definition of this manifold and useful properties). Moreover, Equation (18) also implies that $\operatorname{Im}(\mathbf{P}_k) = \mathbf{M}_{k:0} (\operatorname{Im}(\mathbf{P}_0))$ meaning that the column space of the error covariance matrix remains confined to the column space of the free forecast of the initial covariance matrix. Note that this is a patent property in the perturbations form given in Equation (11). Hence, because the dynamics are assumed non-degenerate, the rank of $\mathbf{P}_k$ will remain that of $\mathbf{P}_0$.

One of the effects of the dynamics is the convergence (or collapse) of $\mathbf{P}_k$ onto the unstable–neutral subspace $\mathcal{U}_k$ of dimension $n_0 \leq n$. In this context, it means that the projection of $\mathbf{P}_k$ onto the stable subspace $\mathcal{S}_k$ asymptotically vanishes, implying that $n - n_0$ of its eigenvalues vanish too, when k goes to infinity. If $\mathbf{P}_k$ is uniformly bounded, which can be guaranteed if the system is sufficiently observed, it was further shown that the stable backward Lyapunov vectors of the model are asymptotically in the null space of $\mathbf{P}_k$ (Boc17), which results in the asymptotic confinement of $\mathbf{P}_k$ to the unstable and neutral subspace.

Consequently, at machine precision, the rank of $\mathbf{P}_k$ follows the asymptotic inequality

$$\lim_{k \to \infty} \operatorname{rank}(\mathbf{P}_k) \leq \min(n_0, \operatorname{rank}(\mathbf{P}_0)). \qquad (20)$$

Equation (19) provides an upper bound for the eigenvalues of $\mathbf{P}_k$ along with an upper bound for the rate of convergence to zero of those associated to the stable modes. Denoting the ith singular value of $\mathbf{M}_{k:0}$ by $\exp(\lambda_i^k k)$, where λ_i^k converges to the Lyapunov exponent λ_i as k goes to infinity, and σ_i^k the ith eigenvalue of $\mathbf{P}_k$, Boc17 showed that there exists α_i such that for large enough k, $\sigma_i^k \leq \alpha_i \exp(2\lambda_i^k k)$, which proves that the eigenvalues of the stable modes converge to zero at least with an exponential rate. Hence, the convergence of $\mathbf{P}_k$ onto the unstable–neutral subspace is controlled by $\mathbf{M}_{k:0} \mathbf{P}_0 \mathbf{M}_{k:0}^T$, the free forecast of $\mathbf{P}_0$.

Finally, Boc17 found that the following three conditions are sufficient for the asymptotic convergence of $\mathbf{P}_k$ to a matrix sequence independent from the initial covariances $\mathbf{P}_0$, which shall hence be qualified as *universal*.

- **Condition 1** If $\mathbf{V}_{+,0}$ is a matrix whose columns are the forward Lyapunov vectors at t_0 of the dynamics $\mathbf{M}_{k:0}$ associated with the unstable and neutral directions, the condition reads $\operatorname{rank}\left(\mathbf{X}_0^T \mathbf{V}_{+,0}\right) = n_0$.
- **Condition 2** Let $\mathbf{U}_{+,k}$ be a matrix whose columns are the backward Lyapunov vectors at t_k of the dynamics $\mathbf{M}_{k:0}$ associated with the unstable and neutral directions. The model is sufficiently observed so that the unstable and neutral directions remain under control, that is

$$\mathbf{U}_{+,k}^T \mathbf{\Gamma}_k \mathbf{U}_{+,k} > \varepsilon \mathbf{I}_n \qquad (21)$$

where $\varepsilon > 0$ is a positive number.
- **Condition 3** In the presence of a neutral mode, additional assumptions are required because the asymptotic behaviour of neutral modes is poorly specified. They could grow or decay like unstable or stable modes albeit at a subexponential rate. A discussion on the issue and possible additional assumptions have been given in section 5.2 of Boc17. For instance, for an autonomous system (which guarantees the presence of at least one neutral mode), a possible condition is: for any neutral backward Lyapunov vector $\mathbf{u}_k$, we have

$$\liminf_{k \to \infty} \mathbf{u}_k^T \mathbf{\Gamma}_k \mathbf{u}_k = \infty. \qquad (22)$$

This means that the neutral modes are sufficiently observed. We refer to Boc17 for other assumptions on the neutral modes that would be valid alternatives to condition 3.
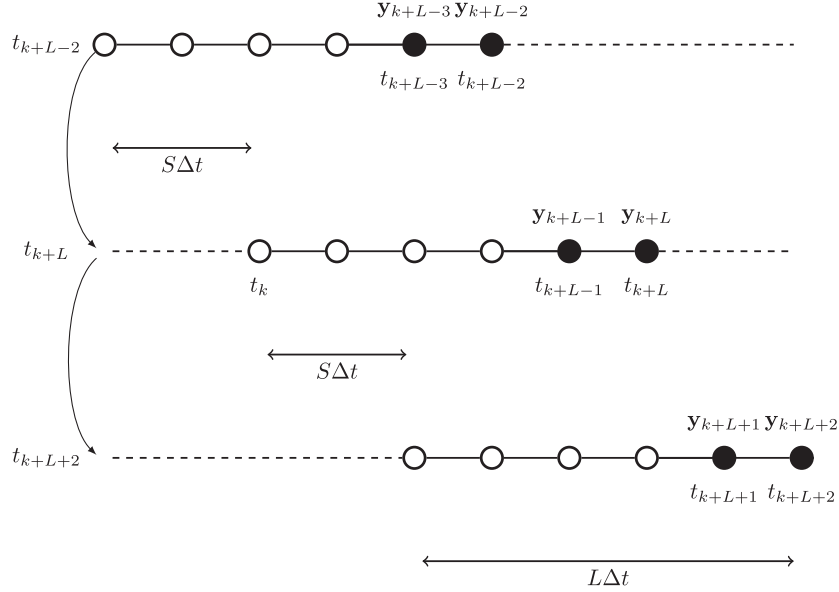
Fig. 1. Typical cycling of the smoothers considered in this study, in the specific case where $L = 5$ and $S = 2$, in units of $t_{l+1} - t_l$, the time interval between two updates. The method performs a smoothing update throughout the window but only assimilates the newest observations vectors (that have not been already assimilated) marked by black dots. Note that the time index of the dates and the observations are absolute for this schematic, not relative.

Other mild conditions usually met in practice, such as the uniform boundedness from above of the precision matrix $\mathbf{\Omega}_k$ defined in Equation (5), are also required to prevent pathological behaviours.

Under these three conditions, we have

$$\lim_{k \to \infty} \left\{ \mathbf{P}_k - \mathbf{U}_{+,k} \left[\mathbf{U}_{+,k}^{\mathrm{T}} \mathbf{\Gamma}_k \mathbf{U}_{+,k} \right]^{-1} \mathbf{U}_{+,k}^{\mathrm{T}} \right\} = \mathbf{0}, \quad (23)$$

and the asymptotic universal sequence does not depend on $\mathbf{P}_0$. Furthermore, these conditions also imply that $\mathbf{P}_k$ and $\mathbf{U}_{+,k} \left[\mathbf{U}_{+,k}^{\mathrm{T}} \mathbf{\Gamma}_k \mathbf{U}_{+,k} \right]^{-1} \mathbf{U}_{+,k}^{\mathrm{T}}$ are bounded, and asymptotically share the same eigenvectors and eigenvalues. As a consequence $\mathbf{X}_k$ and $\mathbf{U}_{+,k} \left[\mathbf{U}_{+,k}^{\mathrm{T}} \mathbf{\Gamma}_k \mathbf{U}_{+,k} \right]^{-\frac{1}{2}}$ asymptotically share the same left singular vectors and the same singular values. Going further than Boc17, we can thus establish the existence of a universal sequence

$$\lim_{k \to \infty} \left\{ \mathbf{X}_k - \mathbf{U}_{+,k} \left[\mathbf{U}_{+,k}^{\mathrm{T}} \mathbf{\Gamma}_k \mathbf{U}_{+,k} \right]^{-\frac{1}{2}} \mathbf{\Psi}_k \right\} = \mathbf{0}, \quad (24)$$

which is unique up to the associated sequence of arbitrary orthogonal matrices $\mathbf{\Psi}_k$.

Our next step consists in exploring the extent to which these results generalize to a fixed-lag smoother.

3. Linear case – convergence of the degenerate Kalman smoother: analytic results

We consider the dynamical system Equations (1) and (2) with linear observation and evolution models and initial Gaussian statistics. In a linear, Gaussian framework, there are many acceptable and mathematically equivalent ways to define a fixed-lag smoother (Cosme et al., 2012). Here, we aim at defining a linear, Gaussian smoother that will serve as a proxy for the IEnKS with nonlinear dynamics, a correspondence that will be used in the numerical experiments in Section 4. We therefore follow its set-up, and in particular that of the so-called single DA IEnKS (SDA IEnKS, Bocquet and Sakov, 2014).

3.1. Fixed-lag approach

The smoother has a fixed lag of L time intervals $\Delta t = t_{l+1} - t_l$ between observations. The interval $[t_k, t_{k+L}]$ is the data assimilation window (DAW), composed of L such intervals between observations. The analysis takes place within this DAW where the state can be estimated a posteriori at any $t_l \in [t_k, t_{k+L}]$ within the DAW. However, the observations $\mathbf{y}_l$ within a given DAW are not necessarily all bound to be assimilated, since they could have already been done so in an earlier cycle. The outcome of the analysis is an updated set of perturbations at t_k, i.e. at the beginning of the DAW. In the forecast step, the ensemble of perturbations is propagated by S intervals, i.e. from t_k to t_{k+S}. The forecast is then followed by a new analysis in the

DAW $[t_{k+S}, t_{k+S+L}]$, and so on. The cycling of the analysis and forecast steps is illustrated in Fig. 1.

To avoid inducing correlations between forecast and observational errors, and hence statistical inconsistencies, it is natural to demand that the observations be assimilated once and for all in the DA run, as it is done in the SDA IEnKS. This constraint implies that only the latest S observation vectors are assimilated in the DAW $[t_k, t_{k+L}]$, i.e. $\mathbf{y}_{k+L-S+1}, \mathbf{y}_{k+L-S+2}, \ldots, \mathbf{y}_{k+L}$. We mention however that there are ways out of it (Bocquet and Sakov, 2014) that are only mathematically equivalent in the linear, Gaussian case, but they are not the focus of this study.

Because of the models' linearity, the update of the anomalies decouples from the update of the state vector, which allows to focus only on the anomaly update. We anticipate that this independence will no longer hold in the nonlinear model scenario considered in Section 4.

3.2. Variational correspondence

The updated anomalies can be derived using the variational correspondence introduced in Section 2. The reduced cost function associated to this smoother is

$$\mathcal{L}(\mathbf{w}) = \frac{1}{2} \sum_{l=k+L-S+1}^{k+L} \|\mathbf{y}_l - \mathbf{H}_l \mathbf{M}_{l:k} (\bar{\mathbf{x}}_k + \mathbf{X}_k \mathbf{w})\|_{\mathbf{R}_l}^2 + \frac{1}{2} \|\mathbf{w}\|^2, \tag{25}$$

whose Hessian is $\mathbf{I}_N + \mathbf{X}_k^{\mathrm{T}} \widehat{\mathbf{\Omega}}_k \mathbf{X}_k$ with

$$\widehat{\mathbf{\Omega}}_k \triangleq \sum_{l=k+L-S+1}^{k+L} \mathbf{M}_{l:k}^{\mathrm{T}} \mathbf{\Omega}_l \mathbf{M}_{l:k} \tag{26}$$

being the precision matrix in this fixed-lag smoother set-up. From the Hessian, following Section 2 and building on the variational correspondence, we infer the anomaly update:

$$\mathbf{X}_k^{\mathrm{a}} = \mathbf{X}_k \left[\mathbf{I}_N + \mathbf{X}_k^{\mathrm{T}} \widehat{\mathbf{\Omega}}_k \mathbf{X}_k \right]^{-\frac{1}{2}} \mathbf{\Psi}_k, \tag{27}$$

where $\mathbf{\Psi}_k$ is an arbitrary orthogonal matrix such that $\mathbf{\Psi}_k \mathbf{1} = \mathbf{1}$. This update is followed by the anomaly propagation

$$\mathbf{X}_{k+S} = \mathbf{M}_{k+S:k} \mathbf{X}_k^{\mathrm{a}}, \tag{28}$$

which yields the smoother anomaly matrix recurrence

$$\mathbf{X}_{k+S} = \mathbf{M}_{k+S:k} \mathbf{X}_k \left(\mathbf{I}_N + \mathbf{X}_k^{\mathrm{T}} \widehat{\mathbf{\Omega}}_k \mathbf{X}_k \right)^{-\frac{1}{2}} \mathbf{\Psi}_k. \tag{29}$$

3.3. Explicit covariance dependence on the initial condition and the observability matrix for the Kalman smoother

Equation (29) is similar to Equation (7), but with a forecast of S time units and $\mathbf{\Omega}_k$ replaced by $\widehat{\mathbf{\Omega}}_k$. Hence, it leads to an analogue of the right-transform Equation (11) for any k of the form $k = pS$, with $p = 0, 1, \ldots$

$$\mathbf{X}_{pS} = \mathbf{M}_{pS:0} \mathbf{X}_0 \left[\mathbf{I}_N + \mathbf{X}_0^{\mathrm{T}} \widehat{\mathbf{\Theta}}_{pS} \mathbf{X}_0 \right]^{-\frac{1}{2}} \mathbf{\Psi}_{pS}, \tag{30}$$

where

$$\widehat{\mathbf{\Theta}}_{pS} \triangleq \sum_{q=0}^{p-1} \mathbf{M}_{qS:0}^{\mathrm{T}} \widehat{\mathbf{\Omega}}_{qS} \mathbf{M}_{qS:0} \tag{31}$$

is an *observability* matrix for the fixed-lag smoother, analogous to that of the filter, Equation (10). In terms of the original precision matrices $\mathbf{\Omega}_k$, substituting Equation (26) into Equation (31), we obtain

$$\widehat{\mathbf{\Theta}}_{pS} = \sum_{l=1}^{pS} \mathbf{M}_{l+L-S:0}^{\mathrm{T}} \mathbf{\Omega}_{l+L-S} \mathbf{M}_{l+L-S:0}. \tag{32}$$

Similar to the filter case, there is a left-transform alternative for Equation (30), which reads

$$\mathbf{X}_{pS} = \left[\mathbf{I}_n + \mathbf{M}_{pS:0} \mathbf{P}_0 \mathbf{M}_{pS:0}^{\mathrm{T}} \widehat{\mathbf{\Gamma}}_{pS} \right]^{-\frac{1}{2}} \mathbf{M}_{pS:0} \mathbf{X}_0 \mathbf{\Psi}_{pS} \tag{33}$$

where

$$\widehat{\mathbf{\Gamma}}_{pS} = \sum_{q=0}^{p-1} \mathbf{M}_{pS:qS}^{-\mathrm{T}} \widehat{\mathbf{\Omega}}_{qS} \mathbf{M}_{pS:qS}^{-1} \tag{34}$$

is the information matrix, which can also be unfolded to yield

$$\widehat{\mathbf{\Gamma}}_{pS} = \sum_{l=1}^{pS} \mathbf{M}_{pS:l+L-S}^{-\mathrm{T}} \mathbf{\Omega}_{l+L-S} \mathbf{M}_{pS:l+L-S}^{-1}, \tag{35}$$

with the convention that $\mathbf{M}_{k:l} = \mathbf{M}_{l:k}^{-1}$ if $k < l$.

Equations (30) and (33) are the smoother-case counterparts of the filter's Equations (11) and (12), respectively. Analogously, they allow for the explicit analytic computation of the error covariance matrix at any time based on its initial condition and the information matrix.

3.4. Confinement of the Kalman smoother to the unstable subspace

Since Equation (18) for the KF – or equivalently its right- and left-transform square-root analogues Equations (11) and

(12) – implies Equation (19) (see Section 2.4), we analogously infer that Equation (33) for the KS implies

$$\mathbf{P}_{pS} \leq \mathbf{M}_{pS:0}\mathbf{P}_0\mathbf{M}_{pS:0}^{\mathrm{T}}. \tag{36}$$

As described in Section 2.4, Equation (19) yields an upper bound estimate of the rate of convergence onto the unstable subspace of the KF's covariance matrix $\mathbf{P}_k$. Similarly Equation (36) implies, along with the convergence onto the unstable subspace, the same upper bound rate for the convergence of the Kalman smoother covariance.

The asymptotic behaviours of the filter and smoother covariance $\mathbf{P}_k$ regarding the unstable subspace are quantitatively distinct. To see this, consider Equation (24) for the Kalman filter asymptotic covariance sequence, and apply it to the smoother's case; assuming $k = pS$, it leads to

$$\lim_{k \to \infty} \left\{ \mathbf{X}_k - \mathbf{U}_{+,k} \left[\mathbf{U}_{+,k}^{\mathrm{T}} \widehat{\mathbf{\Gamma}}_k \mathbf{U}_{+,k} \right]^{-\frac{1}{2}} \mathbf{\Psi}_k \right\} = \mathbf{0}. \tag{37}$$

The difference with the filter's case, Equation (24), lies in the information matrix $\widehat{\mathbf{\Gamma}}_k$.

We can use Equation (13) to compute the information matrix for the filter solution at the end of a L-long DAW

$$\mathbf{\Gamma}_k = \sum_{l=L-S+1}^{k-1} \mathbf{M}_{k:l}^{-\mathrm{T}} \mathbf{\Omega}_l \mathbf{M}_{k:l}^{-1}. \tag{38}$$

Note that the initial S batches of observations have been discarded to match those of the smoother. Practically, these observations are asymptotically irrelevant since their impact is exponentially dampened by dissipative dynamics.

Comparing the information matrices of the filter and smoother solutions, Equations (38) and (35), respectively, we have

$$\widehat{\mathbf{\Gamma}}_k = \mathbf{\Gamma}_k + \sum_{l=k}^{k+L-S} \mathbf{M}_{k:l}^{-\mathrm{T}} \mathbf{\Omega}_l \mathbf{M}_{k:l}^{-1}. \tag{39}$$

This shows that $\widehat{\mathbf{\Gamma}}_k \geq \mathbf{\Gamma}_k$: as expected the smoother's information matrix carries more information since it also accounts for future observations. Such an increased in observability generally implies better state's estimate, which in turn yields a smaller asymptotic sequence for the smoother covariance, $\widehat{\mathbf{P}}_k$, as compared to the filter case, $\mathbf{P}_k$. In fact

$$\mathbf{U}_{+,k} \left[\mathbf{U}_{+,k}^{\mathrm{T}} \widehat{\mathbf{\Gamma}}_k \mathbf{U}_{+,k} \right]^{-1} \mathbf{U}_{+,k}^{\mathrm{T}}$$
$$\leq \mathbf{U}_{+,k} \left[\mathbf{U}_{+,k}^{\mathrm{T}} \mathbf{\Gamma}_k \mathbf{U}_{+,k} \right]^{-1} \mathbf{U}_{+,k}^{\mathrm{T}}. \tag{40}$$

Equation (40) proves that the asymptotic precision of the smoother is greater than that of the filter, as expected. Finally, note that the gain in observability is nonetheless counteracted by the action of the unstable dynamics that sustain the error growth within the DAW, as can be seen in the right-hand side of Equation (39).

4. Nonlinear case – convergence of the ensemble-based Kalman filters and smoothers: numerical results

We now consider the nonlinear perfect dynamical system

$$\mathbf{x}_k = \mathcal{M}_{k:k-1}(\mathbf{x}_{k-1}), \tag{41}$$
$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \mathbf{v}_k, \tag{42}$$

which is the generalization of Equations (1) and (2), where the observational noise is kept additive, with the error statistics still specified as in Section 2. Both the model propagator, $\mathcal{M}_{k:k-1}$, and the observation operator, $\mathcal{H}_k$, can now be nonlinear.

Extending to this nonlinear scenario the analytic results derived in Sections 2 and 3 for the linear case is precluded or much more complicated: the unstable subspace is no longer globally defined and the observability conditions are difficult to infer. Our study will therefore rely on numerical experiments and will pertain to the convergence properties of the ensemble-based filters and smoothers typically used in nonlinear applications. We will describe the specific ensemble-based DA method used in this study in Section 4.1, the methods for estimating the unstable subspace in Section 4.2 and provide numerical results right afterward in Section 4.3.

The interpretation of the numerical results for the ensemble-based schemes will be guided by the following analogy. The degenerate KF and KS are proxies for the EnKF and EnKS, respectively, since they coincide when the model and observation operators are linear, the initial and observation error statistics are both Gaussian, and the ensemble anomalies describe the full range of uncertainty.

From the KF and EnKF literature, we can anticipate a few consequences of the nonlinearity that alter the AUS picture. Firstly, the moments of the error statistics are no longer independent as in Equations (4) and (7) for the linear case, and the error covariance (second-order moment) now depends on the error mean (first-order moment). Indeed, as opposed to the linear model case, the tangent linear evolution depends on the trajectory around which the model is linearized. Secondly, despite the true error statistics being generally non-Gaussian, the EnKF and EnKS still truncate their description to second-order only. This generates an error, usually of small magnitude (but dependent on the degree of deviation from the linear framework) which can accumulate over the cycles. It may lead to the filter divergence if not properly
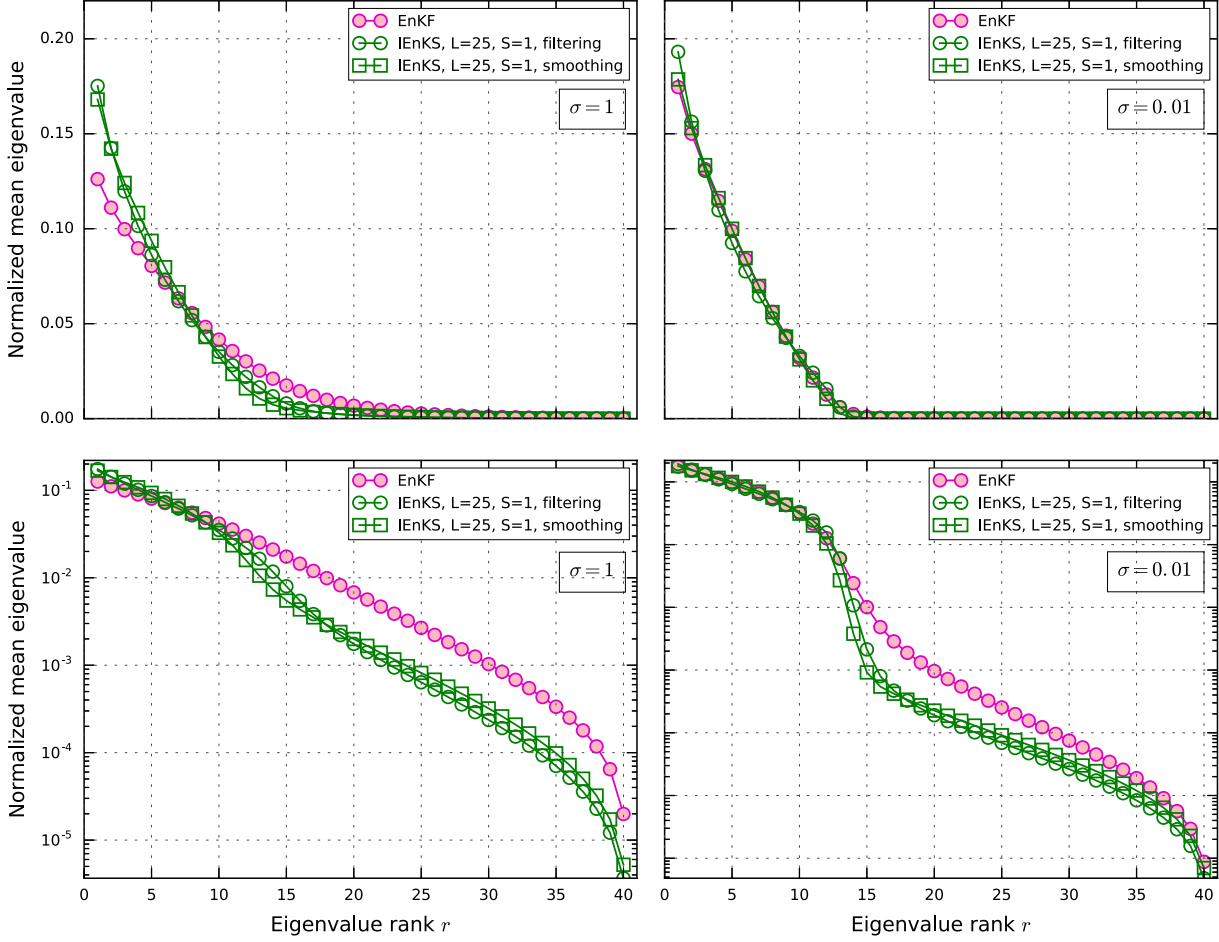
Fig. 2.   Mean eigenvalues of the analysis error covariance matrix normalized by the total variance, in linear (top panels) and in logarithmic scale (bottom panels) for the EnKF and the IEnKS with $L = 25$ and $S = 1$, and for an observation error standard deviation $\sigma = 1$ (left panels) and $\sigma = 0.01$ (right panels). For the IEnKS, the average normalized spectra of both $\mathbf{P}_k^a$ (smoothing covariances) and $\mathbf{P}_{k+L}^a$ (filtering covariances) are shown. The set-up is $\mathbf{H} = \mathbf{I}_d$, $\mathbf{R} = \sigma^2 \mathbf{I}_d$, $N = 41$ and $\Delta t = 0.05$.

accounted for in the DA set-up. These two facts certainly spoil the picture of a perfect confinement to the unstable subspace.

It has been recently argued that the modes of the dynamics that are mostly affected by the nonlinearity are those close to neutrality, in the region of the Lyapunov spectrum where the error is transferred from the unstable to the stable subspace (Ng et al., 2011; Bocquet et al., 2015; Palatella and Trevisan, 2015). Note that even in the linear framework, the convergence of the covariances in the neutral directions is expected to be much slower (see (Boc17) for an extended discussion) than in the other directions.

This implies that the filter covariance confinement for some of the stable modes close to neutral may fail. In the context of the extended KF and for a nonlinear chaotic model, Palatella and Trevisan (2015) have looked at how the AUS picture is affected by nonlinearity by considering the interactions among the backward Lyapunov vectors, and have exposed how

nonlinearity transfers errors in between modes. This picture was shown to support the widespread use of multiplicative inflation which acts on the filter in this region of the Lyapunov spectrum (Bocquet et al., 2015).

4.1. The iterative ensemble Kalman smoother

When dealing with nonlinear dynamics and/or nonlinear observation operator, the KF and KS are usually replaced by their ensemble-based counterparts, the EnKF and EnKS, respectively. By virtue of its improved skill over the EnKF and EnKS, we hereafter adopt the iterative ensemble Kalman smoother (IEnKS, Bocquet and Sakov, 2014).

The IEnKS is a nonlinear four-dimensional ensemble variational method (Sakov et al., 2012; Bocquet and Sakov, 2013, 2014; Bocquet, 2016). It provides both the filter solution (at the end of the DAW) and the smoother one (at any time within the
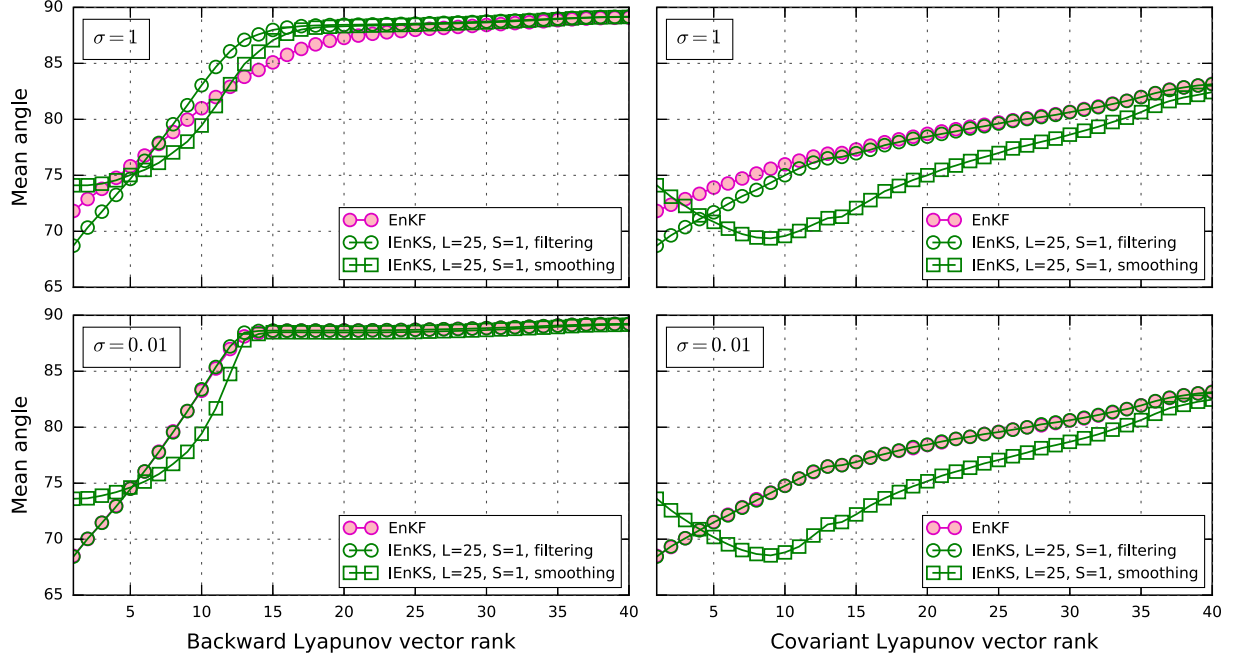
Fig. 3. Time- and ensemble-averaged angle θ_p (in degree) defined by Equation (54), of the anomalies of the ensemble onto each of the backward (left panels) and covariant (right panels) Lyapunov vectors for the EnKF and the IEnKS with $L = 25$ and $S = 1$, and for an observation error standard deviation $\sigma = 1$ (upper panels) and $\sigma = 0.01$ (lower panels). For the IEnKS, the average normalized spectra of both $\mathbf{P}_k^{\mathrm{a}}$ (smoothing covariances) and $\mathbf{P}_{k+L}^{\mathrm{a}}$ (filtering covariances) are shown. The se-tup is $\mathbf{H} = \mathbf{I}_d$, $\mathbf{R} = \sigma^2 \mathbf{I}_d$, $N = 20$ and $\Delta t = 0.05$.

DAW). The EnKF, the maximum likelihood ensemble filter – or MLEF – (Zupanski, 2005), the 4D-ETKF (Hunt et al., 2004), and 4D-En-Var (Liu et al., 2008) can all be seen as particular cases of the IEnKS. In the following, we will use the single DA variant of the IEnKS, in which the observations are assimilated once.

We follow the same set-up as in Section 3. At t_k (i.e. $L\Delta t$ in the past), the prior distribution is estimated from an ensemble of N state vectors of $\mathbb{R}^n$: $\mathbf{x}_{k,[1]}, \ldots, \mathbf{x}_{k,[N]}$. Index k refers to time while $[n]$ refers to the ensemble member index. They can be gathered into the ensemble matrix $\mathbf{E}_k = [\mathbf{x}_{k,[1]}, \ldots, \mathbf{x}_{k,[N]}] \in \mathbb{R}^{n \times N}$. The ensemble can equivalently be given in terms of its mean $\bar{\mathbf{x}}_k = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_{k,[i]}$ and its (normalized) anomaly matrix $\mathbf{X}_k = \frac{1}{\sqrt{N-1}}[\mathbf{x}_{k,[1]} - \bar{\mathbf{x}}_k, \ldots, \mathbf{x}_{k,[N]} - \bar{\mathbf{x}}_k]$.

As with the EnKF, this prior is modelled, within the ensemble subspace, as a Gaussian distribution of mean $\bar{\mathbf{x}}_k$, and covariance matrix $\mathbf{X}_k\mathbf{X}_k^{\mathrm{T}}$, the first- and second-order sampled moments of the ensemble. The background error covariance matrix is rarely full-rank since the anomalies of the ensemble span a vector space of dimension smaller or equal to $N - 1$ and in a realistic context $N \ll n$. One seeks an analysis state vector $\mathbf{x}_k$ in $\mathcal{V}_k = \left\{\bar{\mathbf{x}}_k + \mathbf{X}_k\mathbf{w}, \mathbf{w} \in \mathbb{R}^N\right\}$, identical to Equation (15).

The variational analysis of the IEnKS over $[t_k, t_{k+L}]$ stems from a cost function. The restriction of this cost function to the ensemble subspace reads

$$\mathcal{L}(\mathbf{w}) = \frac{1}{2}\sum_{l=k+L-S+1}^{k+L}\|\mathbf{y}_l - \mathcal{F}_{l:k}(\bar{\mathbf{x}}_k + \mathbf{X}_k\mathbf{w})\|_{\mathbf{R}_l}^2 \quad (43)$$

$$+ \frac{1}{2}\|\mathbf{w}\|^2, \quad (44)$$

where $\mathcal{F}_{l:k} = \mathcal{H}_l \circ \mathcal{M}_{l:k}$ with the $\circ$ symbol representing the composition of operators. S batches of observations are assimilated in this analysis. For instance, the case $L = 0$, $S = 1$, where $\mathcal{M}_{k:k}$ is defined as the identity, corresponds to an *ensemble transform* variant of the MLEF, while a 4D variational analysis is performed when $L \geq 1$ and $1 \leq S \leq L + 1$.

The cost function, Equation (43), is iteratively minimized in the reduced space of the ensemble using a Gauss–Newton algorithm. The jth iteration of the minimization reads

$$\mathbf{w}^{(j+1)} = \mathbf{w}^{(j)} - \mathcal{G}_{(j)}^{-1}\nabla\mathcal{L}_{(j)}(\mathbf{w}^{(j)}), \quad (45)$$

using the gradient $\nabla\mathcal{L}_{(j)}$ and an approximation $\mathcal{G}_{(j)}$ of the Hessian in ensemble subspace

$$\mathbf{x}_k^{(j)} = \bar{\mathbf{x}}_k + \mathbf{X}_k\mathbf{w}^{(j)}, \quad (46)$$

$$\nabla\mathcal{L}_{(j)} = \mathbf{w}^{(j)} - \sum_{l=k+L-S+1}^{k+L}\mathbf{Y}_{l,(j)}^{\mathrm{T}}\mathbf{R}_l^{-1}\left[\mathbf{y}_l - \mathcal{F}_{l:k}(\mathbf{x}_k^{(j)})\right],$$

$$(47)$$

$$\mathcal{G}_{(j)} = \mathbf{I}_N + \sum_{l=k+L-S+1}^{k+L} \mathbf{Y}_{l,(j)}^{\mathrm{T}} \mathbf{R}_l^{-1} \mathbf{Y}_{l,(j)}. \qquad (48)$$

At the first iteration, one sets $\mathbf{w}^{(0)} = \mathbf{0}$ so that $\mathbf{x}_k^{(0)} = \bar{\mathbf{x}}_k$. The observation anomaly matrix, $\mathbf{Y}_{l,(j)} = [\mathcal{F}_{l:k}]'_{|\mathbf{x}_k^{(j)}} \mathbf{X}_k$, is the tangent linear of the operator from ensemble subspace to the observation space. The estimation of these sensitivities based on the ensemble avoids the use of the adjoint model. Note that other iterative minimization schemes could be used in place of Gauss–Newton such as quasi-Newton methods or Levenberg–Marquardt (Bocquet and Sakov, 2012) with distinct convergence properties.

These sensitivities can be computed using finite differences (the so called *bundle* IEnKS variant). The ensemble is rescaled closer to the mean trajectory by a factor $\varepsilon > 0$, typically $10^{-6} < \varepsilon < 10^{-1}$. It is then subject to the combined action of the model propagator and of the observation operator, using the composed operator $\mathcal{F}_{l:k}$, after which it is rescaled back by the inverse factor ε^{-1}. The operation reads

$$\mathbf{Y}_{l,(j)} \approx \frac{1}{\varepsilon} \mathcal{F}_{l:k} \left(\mathbf{x}_k^{(j)} \mathbf{1}^{\mathrm{T}} + \varepsilon \mathbf{X}_k \right) \left(\mathbf{I}_N - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N} \right), \qquad (49)$$

where $\mathbf{1} = (1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^N$. The last factor is meant to substract the mean of the ensemble observation forecast. Note that the $\sqrt{N-1}$ factor needed to normalize the anomalies has been incorporated in ε.

The iterative minimization is stopped when $\|\mathbf{w}^{(j)} - \mathbf{w}^{(j-1)}\|$ reaches a predetermined numerical threshold. Let us denote $\mathbf{w}_k^{\star}$ the solution of the cost function minimization, with the $\star$ symbol identifying any quantity obtained at the minimum. Subsequently, the posterior ensemble can be generated at t_k:

$$\mathbf{E}_k^{\star} = \mathbf{x}_k^{\star} \mathbf{1}^{\mathrm{T}} + \sqrt{N-1} \mathbf{X}_k \mathbf{W}_k^{\star} \mathbf{\Psi}_k, \qquad (50)$$

where

$$\mathbf{W}_k^{\star} = \left[\mathbf{I}_N + \sum_{l=k+L-S+1}^{k+L} \left(\mathbf{Y}_l^{\star} \right)^{\mathrm{T}} \mathbf{R}_l^{-1} \mathbf{Y}_l^{\star} \right]^{-\frac{1}{2}}, \qquad (51)$$

and $\mathbf{\Psi}_k$ is an arbitrary orthogonal matrix satisfying $\mathbf{\Psi}_k \mathbf{1} = \mathbf{1}$ so that the posterior ensemble is centred on the fixed-lag smoother analysis, $\mathbf{x}_k^{\star} = \bar{\mathbf{x}}_k + \mathbf{X}_k \mathbf{w}^{\star}$. If filtering is considered, it additionally requires the free model forecast, initialized on $\mathbf{x}_k^{\star}$, throughout the DAW or beyond present time for forecasting purposes.

During the forecast step of the IEnKS, the ensemble is propagated for $S\Delta t$ time units

$$\mathbf{E}_{k+S} = \mathcal{M}_{k+S:k}(\mathbf{E}_k^{\star}), \qquad (52)$$

and will stand as the prior for the next analysis at time t_{k+S}.

Techniques such as 4D-Var and the IEnKS smooth the estimation within a DAW by fitting a trajectory of the model to a series of observations. Because the model is used within the DAW, its dynamics is expected to force the covariance matrix onto the unstable subspace. This is implicitly achieved with the 4D-Var and explicitly so, through the ensemble, with the IEnKS. This convergence onto the unstable subspace, in terms of both its temporal rate and the amplitude of residual error on the stable modes, is expected to depend on the length of the DAW and on the DA system's deviations from linearity. We can furthermore foresee that the confinement to the unstable subspace is to be more pronounced in the IEnKS than in the EnKF, given that the former can use a much longer DAW than the forecast step in the EnKF, thus facilitating the alignment of the ensemble anomalies along the unstable directions. This is one of the properties studied in the following numerical experiments.

4.2. Numerical representation of the unstable subspace: backward and covariant Lyapunov vectors

The unstable subspace is numerically estimated and characterized using either the backward or the covariant Lyapunov vectors, hereafter denoted BLV and CLV, respectively. The BLVs are an orthogonal set of singular vectors related to the tangent linear operator on infinite time, from the far past to present time t_0 (Legras and Vautard, 1996). These vectors can be computed concurrently with the Lyapunov exponents. We use here a traditional QR decomposition to periodically re-orthonormalize a set of P perturbations meant to compute the first P BLVs and associated Lyapunov exponents. After a long enough spinup, the orthonormalized perturbations coincide with the BLVs. Unlike the exponents, the BLVs depend on time, i.e. they are different for different points on a trajectory solution of the model dynamics, and are norm-dependent, i.e. they depend on the chosen inner products in the tangent space. The BLVs enable to identify the unstable–neutral subspace. However, because they are repeatedly orthonormalized, each individual direction, with the exception of the first one, which is used to start the orthogonalization procedure, does not represent a natural stable/unstable mode of the dynamics. Moreover, BLVs are not covariant with the tangent linear dynamics, nor invariant under time reversal. The former property means that BLVs at a given point are not mapped by tangent propagators to the BLVs at the mapped point. Such a map only holds for the subspaces they span.

While covariant, norm-independent, Lyapunov vectors (CLV) have been known for quite some time (Eckmann and Ruelle

1985; Legras and Vautard 1996; Trevisan and Pancotti 1998), the recent development of efficient numerical techniques for their computation (Wolfe and Samelson 2007; Ginelli et al. 2013), have awakened much interest in studies with complex dynamics, including in the geosciences (see e.g. Vannitsem and Lucarini, 2016, and references therein]. CLVs are not orthogonal and belong to the intersection of the Oseledec subspaces (Legras and Vautard, 1996). They are invariant under time reversal and covariant with the tangent linear dynamics in the sense that they may, in principle, be computed once and then determined for all times using the tangent propagator (Kuptsov and Parlitz, 2012). The pth CLV at time t_l, $\mathbf{u}_l^p$, can be characterized by

$$||\mathbf{M}_{k:l}\mathbf{u}_l^p|| \sim \exp(\lambda^p(k-l)) \quad 1 \leq p \leq P \leq n \quad (53)$$

for any l and k such that $|k - l| \to \infty$. CLVs are thus the natural directions of growth/decay on the tangent space at rate given by the corresponding Lyapunov exponent. In the present study, CLVs are computed using the method developed by Ginelli et al. (2013).

For the sake of a proper interpretation of the convergence numerical results, it is worth highlighting some key differences between the CLVs (and BLVs) used to define the unstable subspace, and the anomalies $\mathbf{X}_k$ describing the ensemble subspace. While the CLV (and the BLV) subspaces belong to the tangent space and are propagated by the corresponding tangent linear dynamics, the anomalies do not live in the tangent spaces and are propagated using the full nonlinear model. Note as well that the Lyapunov vectors are ordered by decreasing Lyapunov exponent, but there is no such ordering for the anomalies in $\mathbf{X}_k$, and they are in principle indistinguishable from each other.

In the following numerical experiments, we use $P = n$, so that the full spectrum of the Lyapunov exponents and vectors is computed.

4.3. Numerical experiments

Most of the analytic results have been formulated for the forecast error covariance matrix $\mathbf{P}_k$. Yet, many of the numerical results that follow will be obtained for the analysis error covariance matrix $\mathbf{P}_k^a$. This is not an issue since the qualitative convergence behaviour of $\mathbf{P}_k$ is equivalent to that of $\mathbf{P}_k^a$ (in fact $\mathbf{P}_k^a \leq \mathbf{P}_k$) and the quantitative results can easily be adapted, due to $\mathbf{P}_{k+1} = \mathbf{M}_{k+1}\mathbf{P}_k^a\mathbf{M}_{k+1}^T$.

We consider the one-dimensional Lorenz model (Lorenz and Emanuel, 1998) with $n = 40$ space variables, denoted L95, well documented and studied in the fields of numerical weather prediction and DA. Its equations can be found in Appendix 1. Assuming the standard forcing value $F = 8$, it has 13 positive and one neutral Lyapunov exponents. Hence, the dimension of the unstable–neutral subspace $\mathcal{U}_k$ is $n_0 = 14$.

In the following, we shall apply the EnKF and IEnKS DA methods to this model. We will assume that the state vector is fully observed with observation operators $\mathbf{H}_k \triangleq \mathbf{I}_d$ where $d = 40$. Synthetic observations are generated from the model and $\mathbf{H}_k$ at each t_k, every $\Delta t = 0.05$ time step, unless otherwise stated. These observations are independently (in time and space) perturbed by a random Gaussian noise of mean 0 and variance σ^2, so that the time-independent observation error covariance matrices are $\mathbf{R}_k \triangleq \sigma^2\mathbf{I}_d$. Localization is not used in the experiments given that the ensemble size is set to $N \geq 15$, which renders localization unnecessary in this model and observational configuration.

The need for inflation is accounted for using the finite-size filters and smoothers as described in Bocquet (2011), Bocquet and Sakov (2012), and Bocquet et al. (2015). In the absence of model error as assumed here, this technique prevents the need to tune multiplicative inflation and avoids its very high numerical cost. The additional computational cost of using the finite-size filters and smoothers is marginal.

4.3.1. Eigenvalues of the analysis error covariance matrix.
The linear theory predicts that the eigenvalues of $\mathbf{P}_k^a$ related to the stable modes should asymptotically vanish. Given that $\mathcal{U}_k$ has dimension 14, a transition in the eigenspectrum of $\mathbf{P}_k^a$ is expected to be seen in the fifteenth eigenvalue, reflecting a total of 14 independent anomalies (recall that one degree of freedom is consumed in subtracting the ensemble mean). We look at the eigenspectrum of $\mathbf{P}_k^a$ to identify any signature of $\mathcal{U}_k$ on the performance of the DA and possible deviations from the linear theory prognosis.

An ensemble of size $N = 41$ is chosen so that $\mathbf{P}_k^a$ has full rank and its spectrum has generically 40 non-zero eigenvalues. The mean eigenvalues, averaged over time, for the EnKF and for the IEnKS with $L = 25$ and $S = 1$, are computed. The relative contributions of each eigenmode to the total variance is obtained by normalizing each eigenvalue by the total variance (sum of the eigenvalues). In the IEnKS case, both the spectra of the error covariance matrix at the beginning (smoothing) and at the end of the DAW (filtering) are computed.

The contributions to the total variance are plotted in Fig. 2 in linear (upper panels) and logarithmic scale (lower panels). The sensitivity to the degree of nonlinearity in the DA system is explored using two different observational errors, $\sigma = 1$ (left panels) and $\sigma = 0.01$ (right panels). The former case corresponds to an error of about 5% of the L95 variables' range, and it is consistent with the typical error level in synoptic-scale meteorology (see e.g. Kalnay, 2002), while the latter is closer to a linear/Gaussian regime. Note, however, that the regime where σ vanishes is not necessarily linear if Δt is fixed. Linearity can only be ensured by reducing Δt, as explained in Appendix 1.

For the larger observational error, $\sigma = 1$ (left panels), the signature of the $\mathcal{U}_k$ dimension is only moderately visible in the EnKF case. The transition around eigenvalue rank $r = 15$ gets
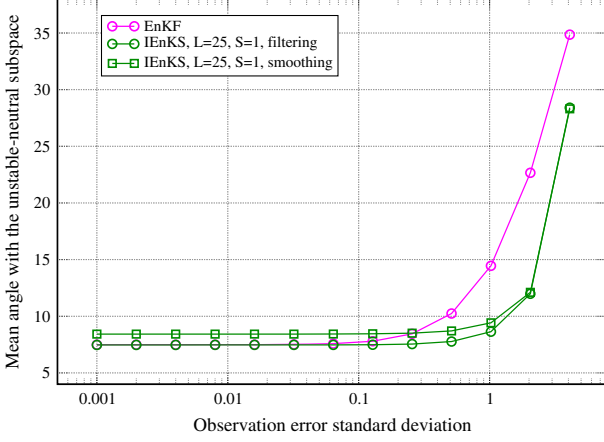
Fig. 4. Time- and ensemble-averaged angle θ_k^i (in degree) from Equation (55), between an anomaly from the EnKF ensemble and from the IEnKS ensemble ($L = 25$, $S = 1$) and the unstable–neutral subspace, when the observation error standard deviation σ is varied from 10^{-3} to 4. The set-up is $\mathbf{H} = \mathbf{I}_d$, $\mathbf{R} = \sigma^2 \mathbf{I}_d$, $N = 20$ and $\Delta t = 0.05$.

Fig. 5. Time- and ensemble-averaged angle θ_k^i (in degree) from Equation (55), between an anomaly from the EnKF and IEnKS ($L = 5$, $S = 1$) ensembles and the unstable–neutral subspace, when the time interval between updates is varied from $\Delta t = 0.01$ to $\Delta t = 0.50$. The se-tup is $\mathbf{H} = \mathbf{I}_d$, $\mathbf{R} = \sigma^2 \mathbf{I}_d$ with $\sigma = 10^{-2}$ and $N = 20$.

steeper in the IEnKS, particularly in the filtering case. This latter behaviour is due to the fact that a free forecast is performed within the DAW which submits the anomalies to the dynamics over a possibly long time interval. Diminishing the observation error to $\sigma = 0.01$ yields a steeper transition of the spectrum around $r = 15$ in the linear scale and a significant inflection in the logarithmic scale, for both the EnKF and IEnKS.

Thus, the picture of the confinement proven analytically in the linear/Gaussian case seems to remain approximately valid in this nonlinear scenario. Most importantly, the deviations from the linear predictions grow along with the degree of nonlinearity in the DA system. The latter is controlled by acting on Δt, or on a combination of σ and F as discussed in Appendix 1. Nevertheless, the spectrum of $\mathbf{P}_k^a$ alone does not suffice to characterize the alignment of the anomalies with $\mathcal{U}_k$. To ascertain this picture, we look at the geometry of the ensemble of anomalies in the following sections. This calls for a geometrical characterization of the relative position of the anomalies with respect to the Lyapunov vectors.

4.3.2. Alignment of the anomalies with each Lyapunov vector. We begin by studying how the ensemble anomalies are oriented with respect to each of the unstable–neutral mode independently. The projection of the anomaly $\mathbf{v}_k^i = [\mathbf{X}_k]_i$ with $1 \leq i \leq N$, onto any of the Lyapunov vectors $\left\{\mathbf{u}_k^p\right\}_{1 \leq p \leq P}$ at t_k, is given by $\left\|\mathbf{v}_k^i\right\| \cos(\theta_k^{i,p})$ where $\theta_k^{i,p} \in [0, \pi]$ is the angle between both vectors (recall that the Lyapunov vectors are normalized to 1). All the anomalies are projected onto each Lyapunov vector yielding an angle θ_k^p ($1 \leq p \leq P$) defined by
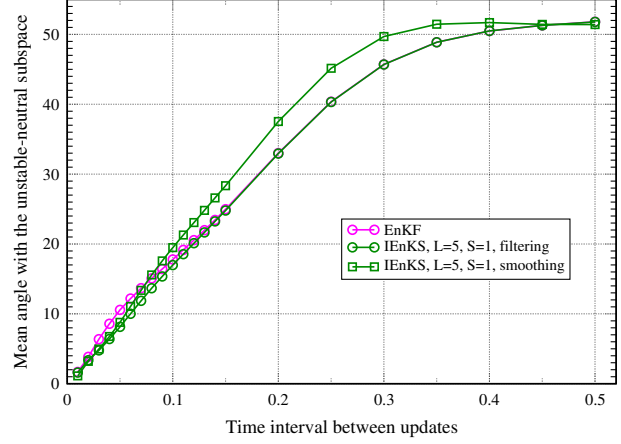
$$\theta_k^p = \frac{1}{N} \sum_{i=1}^{N} \theta_k^{i,p} = \frac{1}{N} \sum_{i=1}^{N} \arccos \left\{ \frac{(\mathbf{u}_k^p)^{\mathrm{T}} \mathbf{v}_k^i}{\left\|\mathbf{v}_k^i\right\|} \right\}. \quad (54)$$

These angles are further averaged over time and the mean values are reported in Fig. 3, using either BLVs (left panels) or CLVs (right panels) to estimate $\mathbf{u}_k^p$, and two observation error standard deviations, $\sigma = 1$ (upper panels) and $\sigma = 0.01$ (lower panels). Results are for the EnKF and for the IEnKS with $L = 25$ and $S = 1$.

If the span of the error covariance matrix, $\mathbf{P}_k^a$ (or $\mathbf{P}_{k+L}^a$ for the filter solution of the IEnKS), is the unstable–neutral subspace, then the mean angle between the anomalies and the stable Lyapunov vectors will consistently tend to 90° with increasing r.

According to Fig. 3, such a prediction is overall confirmed by the EnKF (magenta circles) and by the filter solution of the IEnKS (green circles). The transition around rank $r = 15$ is visible.

The effect of the observation accuracy is more apparent in the EnKF with the BLVs used to estimate $\mathcal{U}_k$, for which the transition from the unstable–neutral to the stable portion of the Lyapunov spectrum gets much sharper when $\sigma = 0.01$. Such an effect is practically undetected in the IEnKS filtering solution, suggesting that the IEnKS is already able to keep the error sufficiently small, and thus of linear-like type, even when $\sigma = 1$.

Another key aspect from Fig. 3 is the distinct behaviour when BLVs or CLVs are chosen for the Lyapunov vectors. In both cases, the angles tend continuously to 90 when the rank is increased, but while the trend is smooth for the CLVs, the BLVs clearly display an abrupt transition around a rank $r = n_0 + 1 = 15$. This difference is due to the fact that in contrast to the BLVs, CLVs are not orthogonal. Indeed each CLV will generally
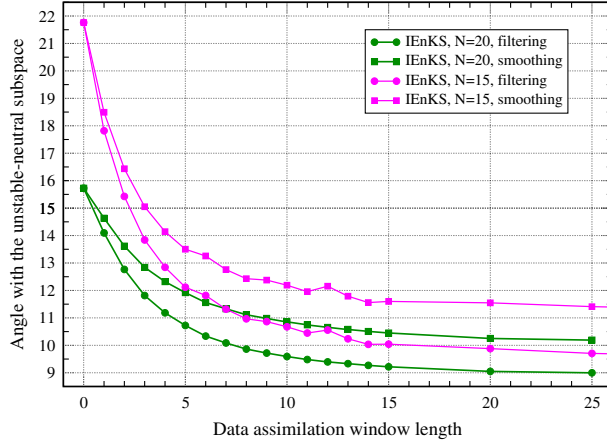
Fig. 6. Time- and ensemble-averaged angle θ_k^i (in degree) from Equation (55), between an anomaly from the IEnKS ensemble and the unstable–neutral subspace, when the DAW length is varied from $L = 0$ (corresponding to the EnKF) to $L = 25$, and $S = 1$. The set-up is $\mathbf{H} = \mathbf{I}_d$, $\mathbf{R} = \mathbf{I}_d$ and $\Delta t = 0.05$.

have a non-zero projection on other CLVs, usually those in its spectrum proximity. Figure 3 reveals that, when $r \geq 2$, the projections on the CLVs are systematically larger than those on the corresponding BLVs, indicating that the angle between adjacent CLVs is smaller than 90° (possibly much smaller given the smooth decrease in projection along $\mathbf{u}_k^p$ with increasing rank). Exploring whether, or to which extent, this CLVs feature is generic for chaotic dissipative systems, or identifying the underlying dynamical mechanisms that cause it, are both relevant research directions but are not the central focus in this study (see Vannitsem and Lucarini, 2016, for a study on the CLVs along these lines].

The discussion in relation to Fig. 3 has so far pertained to the EnKF and to the filter solution of the IEnKS alone; the IEnKS smoothing solution (green squares) deserves special attention. Interestingly, the projections are smaller (larger) along Lyapunov vectors with $r \leq 5$ ($r > 5$) than their EnKF and IEnKS filter counterparts. This behaviour is so pronounced in the CLVs case that the peak of projection is unexpectedly towards the end of the unstable spectrum ($8 \leq r \leq 11$). The smoothing procedure appears thus to have reduced the uncertainty along the leading unstable modes, and to have also operated a rotation of the subspace on which this uncertainty is confined (the span of $\mathbf{P}_k^a$) resulting in an increased alignment along other (than the leading ones) Lyapunov modes. However, we will see from the study of the global alignment of the unstable–neutral and anomaly subspaces described in Section 4.3.3 (as opposed to the individual direction approach of Fig. 3), that this projection onto the unstable and neutral subspace is always slightly smaller for the smoothing solution, $\mathbf{P}_k^a$, compared to the filter one, $\mathbf{P}_{k+L}^a$.

4.3.3. Global alignment of the anomalies and unstable–neutral subspaces. In this section, we propose a first approach to globally characterize the alignment between the two subspaces. We consider the angle between any anomaly of the ensemble and the unstable–neutral subspace $\mathcal{U}_k$. Because $\mathcal{U}_k$ is spanned by the orthonormal basis composed of the n_0 BLVs with non-negative Lyapunov exponents, the square cosine of $\theta_k^i \in [0, \pi]$, the angle between the anomaly $\mathbf{v}_k^i$ and $\mathcal{U}_k$ is obtained as ($1 \leq i \leq N$)

$$\cos^2(\theta_k^i) = \sum_{p=1}^{n_0} \cos^2(\theta_k^{i,p}) = \sum_{p=1}^{n_0} \frac{\left\{ (\mathbf{u}_k^p)^{\mathrm{T}} \mathbf{v}_k^i \right\}^2}{\left\| \mathbf{v}_k^i \right\|^2}, \qquad (55)$$

where $\mathbf{u}_k^p$ is the pth BLV.

The degree of alignment, measured using θ_k^i from Equation (55), is studied as a function of several parameters of the DA system controlling the degree of nonlinearity: observational error and frequency, as well as the length of the DAW.

Impact of observation error

In Fig. 4, the time- and ensemble-averaged angle θ_k^i from Equation (55) for the EnKF and IEnKS (smoothing and filtering solution) with $L = 25$ and $S = 1$, is plotted as a function of the observation error in the range $\sigma \in [10^{-3}, 4]$.

Figure 4 evidences how the covariance subspaces projection on $\mathcal{U}_k$ increases along with the observations accuracy. The more accurate the observations the smaller the estimation error, which gets closer to that of the linear approximation under which the perfect convergence onto the unstable–neutral subspace holds. Remarkably, the increased accuracy of the IEnKS over the EnKF is reflected into a slower reduction of the projection of its error covariance matrix on $\mathcal{U}_k$ for increasing σ. The deviation from the pure linear framework is manifested by the small, but non-zero, limiting angle value when $\sigma \to 0$ and when Δt is fixed. This is heuristically discussed in Appendix 1 using a simple dimensional analysis. Finally note that, as anticipated in Section 4.3.2, the projection of the smoothing solution of the IEnKS is slightly smaller than that of the filtering one.

Impact of observation frequency

We now address the impact of nonlinearity by adjusting the time interval Δt between observations. In Fig. 5, the time- and ensemble-averaged angle θ_k^i from Equation (55), (in degree) for the EnKF and the IEnKS (smoothing and filtering solution) with $L = 5$ and $S = 1$, is displayed as a function of Δt in the range $[0.01, 0.50]$.

The angles clearly converge to zero as Δt goes to zero too. This is consistent with the heuristic dimensional analysis discussed in Appendix 1, which shows that the DA system truly becomes linear in this limit in contrast to what occurs by acting only on the observation accuracy. The rate of this convergence is the fastest, and of the same amplitude, for the EnKF and IEnKS (filter solution), for $\Delta t \gtrsim 0.15$, with the IEnKS being faster
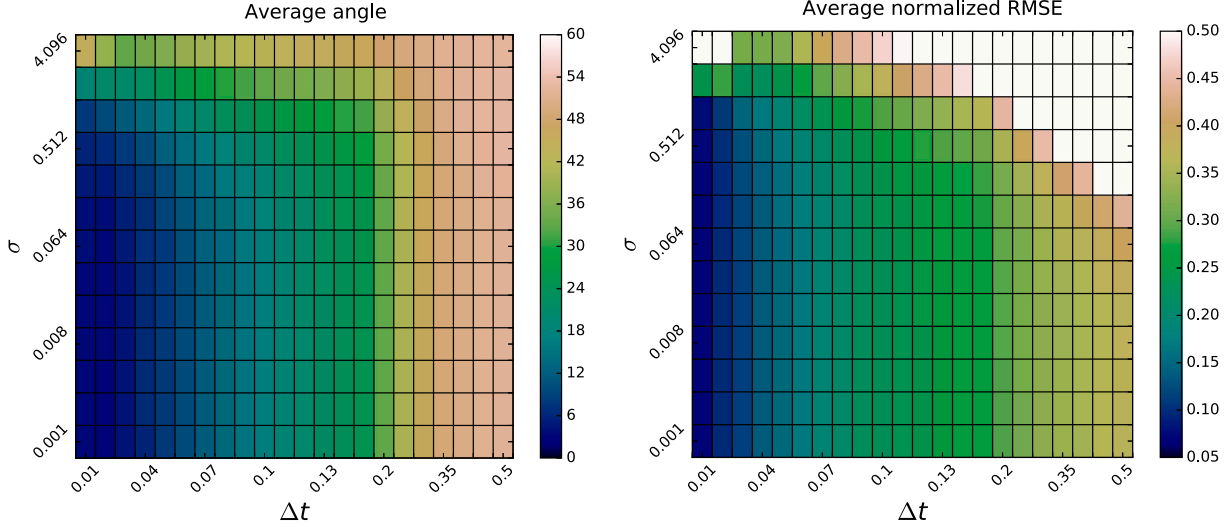
Fig. 7.   Time- and ensemble-averaged angle θ_k^i from Equation (55) (left panel, shadow colours in degree) between an anomaly of the EnKF and the unstable–neutral subspace, and the EnKF RMSE normalized by σ (right panel, shadow colours), on the plane $(x, y) = (\Delta t, \sigma)$. The set-up is $\mathbf{H} = \mathbf{I}_d$, $\mathbf{R} = \sigma^2 \mathbf{I}_d$ and $N = 20$.

in the remaining part of the explored range of Δt. The exact behaviour of the angle curve when Δt goes to 0 depends on the chosen inflation adaptive scheme, which, for the three curves, is taken from Bocquet et al. (2015). As expected, the angle for smoothing is greater than for filtering.

Impact of the DAW length

We now address the impact of the length, L, of the DAW in the IEnKS. In Fig. 6, the time- and ensemble-averaged angle θ_k^i from Equation (55), is plotted as a function of L in the range $[0, 25]$. Recall that the case $L = 0$ corresponds to the EnKF. Two experiments are carried out, with ensemble size $N = 15$ and 20, respectively. The former case is taken to represent a critical set-up for the IEnKS, in term of the sampling error, so as to enhance the impact of the DAW length on the DA performance.

As expected, the alignment of the anomaly from the IEnKS and the unstable–neutral subspace increases continuously with the DAW, and the rate of this process is faster for the filtering solution of the IEnKS than for the smoothing one. In agreement with the previous results, Fig. 6 is also consistent with the linear prediction in the sense that the angle between the two subspaces appears inversely proportional to the accuracy of the DA method: the EnKF, as well as all the experiments with the smallest ensemble size, $N = 15$, systematically display smaller projections.

4.3.4. Alignment and accuracy of the filter. The convergence of the span of the ensemble-based covariance onto the unstable–neutral subspace is explored further in Figs. 7 and 8. We show here only the results for the EnKF, but we checked that the same qualitative conclusions hold for the IEnKS too.

Figure 7 shows the time- and ensemble-averaged angle θ_k^i from Equation (55) (left panel) and the root mean square error

normalized by σ (normalized RMSE, right panel), in shadow colours on the plane $(\Delta t, \sigma)$.

The patterns of the two panels are very similar: in general areas of small angles (strong alignment) corresponds to areas of lower normalized RMSEs (high accuracy), and conversely. Note however that, while the filter maintains a satisfactory accuracy even when Δt is large, as long as the observation error is small, the same behaviour is not found in the degree of alignment. In this case, in fact the time between two updates, Δt, seems to play a stronger role in pushing the system out of the linear/Gaussian framework for which the full alignment holds. This is supported by the dimensional analysis of Appendix 1. Similarly, the weak dependence of the normalized RMSEs in σ for fixed Δt, and the limiting behaviour of these quantities as a function of Δt when σ vanishes is also in agreement with the analysis of Appendix 1.

We now compute the time- and ensemble-averaged angle between an anomaly and $\mathcal{U}_k$ as a function of the ensemble size in the range $N \in [5, 40]$ for the EnKF. The RMSE is plotted as well for a qualitative comparison. The curves are displayed in Fig. 8, that illustrates intelligibly the direct relation between the geometrical properties of the ensemble covariance and the filter performance. This result also suggests a guideline to properly size the ensemble: a number of members at least as big as the dimension of the unstable–neutral subspace are necessary to attain high filter accuracy, and avoid the imperative need for localization (Bocquet, 2016).

4.4. Anomaly simplex and the unstable–neutral subspace

A complete characterization of the relative orientation between the subspace spanned by the anomalies and the unstable–neutral
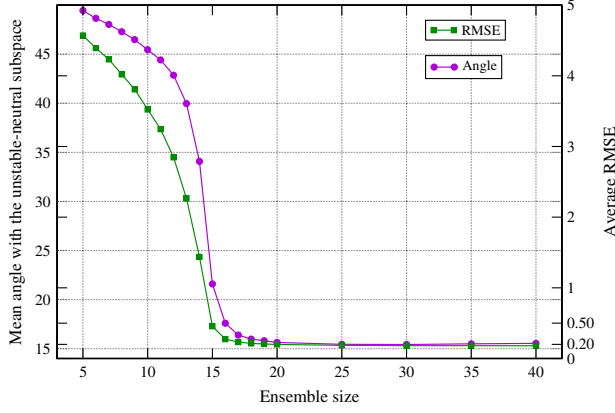
Fig. 8.   Time- and ensemble-averaged angle (in degree) between an anomaly from the EnKF ensemble and the unstable–neutral subspace as a function of the ensemble size N (left Oy axis) and corresponding time-averaged RMSE of the EnKF (right Oy axis). The set-up is $\mathbf{H} = \mathbf{I}_d$, $\mathbf{R} = \mathbf{I}_d$, and $\Delta t = 0.05$.

Fig. 9.   Time-averaged principal angle (in degree) between the anomaly simplex (see text) and the unstable–neutral subspace, as a function of its index for several EnKF/IEnKS configurations (see legend).

subspace cannot be achieved by the angle we have previously defined alone. It requires the use of a set of angles, called *principal*, whose number is given by the minimum of the dimensions of both subspaces. Considering ensemble sizes greater than the dimension of $\mathcal{U}_k$, there are 14 principal angles for the L95 model. These angles are intrinsic and do not depend on the parameterization of both subspaces.

They can be computed through the singular values of the product of two orthonormal matrices formed by the orthonormal basis of both subspaces (Björck and Golub, 1973). An orthonormal basis of $\mathcal{U}_k$ is given by the set of unstable and neutral BLVs, and assembled column-wise into the matrix $\mathbf{Z}_k = [\mathbf{u}_k^1, \ldots, \mathbf{u}_k^{n_0}]$. An orthonormal basis of the subspace spanned by the error covariance matrix can be obtained from a QR decomposition of the anomaly matrix $\mathbf{X}_k = \mathbf{Q}_k \mathbf{T}_k$ where $\mathbf{Q}_k \in \mathbb{R}^{M \times n_0}$ is an orthonormal matrix and $\mathbf{T}_k \in \mathbb{R}^{n_0 \times N}$ is upper triangular. The cosines of the principal angles are then given by the singular values of $\mathbf{Q}_k^{\mathrm{T}} \mathbf{Z}_k$ whose number is n_0 if $N \geq n_0 + 1$. Note that the dimension of the intersection of the two subspaces is given by the number of zero angles which suggests that these angles are a good measure of the degree of alignment of the subspaces.

The time-averaged 14 principal angles computed for several configurations of the EnKF and IEnKS are shown in Fig. 9. As expected, the more accurate the DA method, the flatter and lower the profile of the principal angles is, which reveals a stronger alignment of the two subspaces. Remarkably, the smallest angles are obtained with the EnKF using the most accurate observations ($\sigma = 0.01$). Nevertheless, the IEnKS attains a similarly good alignment using much less accurate observations ($\sigma = 1$).

A complementary view is given in Fig. 10 which shows the principal angles of the smoothing analysis from an IEnKS run with $N = 15$ and $S = 1$ as a function of the DAW length L.
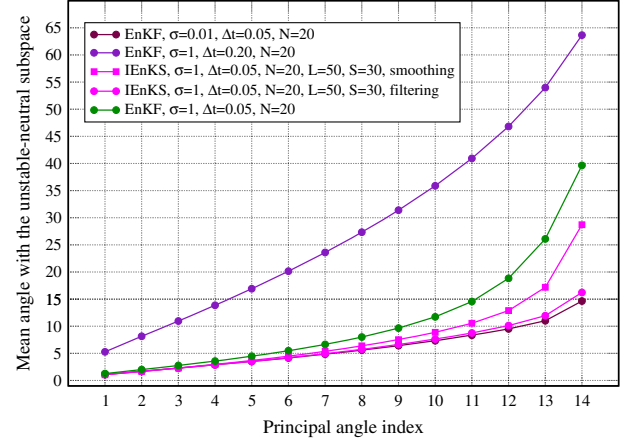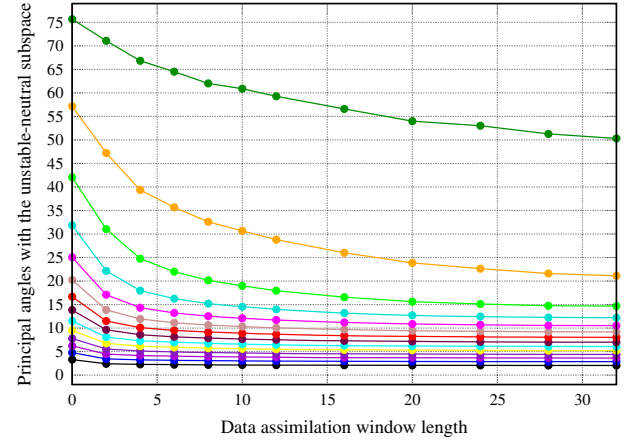

Fig. 10.   The 14 time-averaged principal angles (in degree) between the IEnKS anomaly subspace of the smoothing analysis and the unstable–neutral subspace, as a function of the DAW length. The set-up is $\mathbf{H} = \mathbf{I}_d$, $\mathbf{R} = \mathbf{I}_d$, $N = 15$, $\Delta t = 0.05$ and $S = 1$.

The 14 principal angles all converge towards smaller values as the DAW is increased, as expected. Nevertheless, several of them seem to converge to substantially high values, between 20 and 50°, pointing to an incomplete alignment of the two subspaces. This is due to the relative low accuracy of the IEnKS which uses here only $N = 15$ members. In analogy with the experiments depicted in Fig. 6, this choice has been done intentionally to enhance the visibility of the impact of the DAW.

5. Conclusions

Gurumoorthy et al. (2017) and Boc17, have recently provided analytic proofs for the convergence of the solution of the Kalman

filter covariance equation onto the unstable–neutral subspace of the dynamics, when the model and observation operators are both linear, the model is perfect and the initial and observation error statistics are Gaussian.

The present study is rooted in the same stream of research and has been prompted by both methodological and applied goals. In the first part, on the methodological side, we have extended the convergence results of Gurumoorthy et al. (2017) and Boc17, to degenerate (i.e. for initial error covariance of any rank) linear Gaussian smoothers and have proved analytically that their error covariance projection onto the stable subspace goes asymptotically to zero. We have also provided an upper bound for the rate of such convergence, derived the explicit covariance dependence on the initial condition and on the information matrix, and proved the existence of a universal sequence of covariance matrices to which the solutions converge if the system is sufficiently observed and the initial error covariances have non-zero projections onto each of the unstable–neutral forward Lyapunov vectors. All these results have been obtained using the square-root formulation of the filter and smoother, which complements Gurumoorthy et al. (2017) and Boc17 in which the proofs for the filter case were obtained using the full error covariance form.

The choice to work with the square-root formulation was not driven by the sole mathematical appeal, but by the interest in the nonlinear dynamics scenario, addressed in the second part of this work. DA in this case is usually carried out using EnKFs or smoothers (Evensen, 2009), and the reduced-rank covariance description is implemented using an ensemble of model realizations. In this context, the ensemble anomaly matrix is the analogue of the square root of the error covariance matrix in the linear/Gaussian framework. Obtaining similar rigorous convergence results in the nonlinear case is extremely difficult. We have thus turned to numerical experimentation, and have used the theoretical findings of the linear case as guidance. The IEnKS (which reduces to the EnKF when the assimilation window is set to zero) behaves similarly, in this nonlinear scenario, to the rank-deficient Kalman linear filter and smoother. The IEnKS is an exemplar of a four-dimensional ensemble variational DA method and it belongs to the general class of deterministic, square-root, ensemble-based algorithms. Because, like 4D-Var, it maintains dynamical consistency throughout the DA window, the IEnKS could be prone to an even stronger confinement of its ensemble anomalies to the unstable and neutral subspace than the EnKF.

We have studied the alignment of the IEnKS anomalies along the unstable–neutral subspace in a number of data assimilation set-ups, by acting on different experimental parameters, including the frequency and accuracy of the observations or the length of the DA window. The unstable–neutral subspace is computed using either the backward or the covariant Lyapunov vectors. The alignment with the anomalies is measured using three different approaches aiming at characterizing it along individual directions or globally as the relative position of two subspaces.

The anomalies clearly have a dominant projection onto the unstable–neutral subspace, with only a marginal portion of the estimated uncertainty on the stable one. Results reveal that the degree of this projection scales with the level of nonlinearity in the DA system, suggesting that linear-like convergence is achievable if DA is able to keep the estimation error small enough so that it behaves quasi-linearly. The converse is also supported by the numerical results: the more the error covariance spans the full range of unstable–neutral modes, the better is the skill of the DA (in terms of root-mean-square-error). This implies, and it is numerically confirmed, that the dimension of the unstable–neutral subspace is a lower bound for the sizes of the ensemble for which the filter/smoother performs satisfactorily. Smaller ensemble sizes would necessarily require localization. While such a relation between minimum ensemble size and filter performance was already argued in previous studies (Carrassi et al., 2009; Ng et al., 2011; Bocquet and Sakov, 2014), this is the first time they are corroborated based on a thorough geometrical characterization.

These results and their implications for filters/smoothers design are only ensured for deterministic, square-root, formulations and for which no rotation or other manipulation of the ensemble anomalies is performed (see e.g. Tippett et al., 2003). This is indeed the case for the IEnKS for which we can further conclude that it *implicitly* tracks the unstable–neutral subspace and uses it to confine the analysis correction. It thus realizes in practice the paradigm behind the approach known as the AUS (Palatella et al., 2013).

Two natural extensions of the present work the authors are currently pursuing are the inclusion of model error and the applicability of the error convergence onto the unstable–neutral subspace in the design of efficient Bayesian data assimilation methods. A third extension concerns localization. Since localization affects the dynamical balance of the ensemble update, it is unclear how the AUS picture is modified. The numerical tools developed in this study could help clarify this question.

Acknowledgements

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. Assuming $x \mapsto f(x)$ can be written as a formal power series, i.e. $f(x) = \sum_{i=0}^{\infty} a_i x^i$, one has $\mathbf{A} f(\mathbf{BA}) = \sum_{i=0}^{\infty} a_i \mathbf{A}(\mathbf{BA})^i = \sum_{i=0}^{\infty} a_i (\mathbf{AB})^i \mathbf{A} = f(\mathbf{AB})\mathbf{A}$. This proves the matrix shift lemma, i.e. $\mathbf{A} f(\mathbf{BA}) = f(\mathbf{AB})\mathbf{A}$. The application of this formal identity to matrices $\mathbf{A} \in \mathbb{R}^{n \times m}$ and $\mathbf{B} \in \mathbb{R}^{m \times n}$ further requires that $f(\mathbf{AB})$ and $f(\mathbf{BA})$ exist.

2. Our definition of $\mathbf{\Gamma}_k$ slightly differs – notably in the index range – from that in Jazwinski (1970) which is based on the analysis error matrix $\mathbf{P}_k^{\mathrm{a}}$ while we focus on the forecast error covariance matrix $\mathbf{P}_k$.

References

Anderson, B. D. O. and Moore, J. B. 1979. *Optimal Filtering*. Prentice-Hall Inc, Englewood Cliffs, NJ.

Bishop, C. H., Etherton, B. J. and Majumdar, S. J. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.* **129**, 420–436.

Björck, Å. and Golub, G. H. 1973. Numerical methods for computing angles between linear subspaces. *Math. Comput.* **27**, 579–594.

Bocquet, M. 2011. Ensemble Kalman filtering without the intrinsic need for inflation. *Nonlinear Processes Geophys.* **18**, 735–750.

Bocquet, M. 2016. Localization and the iterative ensemble Kalman smoother. *Q. J. R. Meteorol. Soc.* **142**, 1075–1089.

Bocquet, M., Gurumoorthy, K. S., Apte, A., Carrassi, A., Grudzien, C. and co-authors. 2017. Degenerate Kalman filter error covariances and their convergence onto the unstable subspace. *SIAM/ASA J. Uncertainty Quantification*. Accepted for publication; arXiv preprint arXiv:1604.02578.

Bocquet, M., Raanes, P. N. and Hannart, A. 2015. Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation. *Nonlinear Processes Geophys.* **22**, 645–662.

Bocquet, M. and Sakov, P. 2012. Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems. *Nonlinear Processes Geophys.* **19**, 383–399.

Bocquet, M. and Sakov, P. 2013. Joint state and parameter estimation with an iterative ensemble Kalman smoother. *Nonlinear Processes Geophys.* **20**, 803–818.

Bocquet, M. and Sakov, P. 2014. An iterative ensemble Kalman smoother. *Q. J. R. Meteorol. Soc.* **140**, 1521–1535.

Carrassi, A., Ghil, M., Trevisan, A. and Uboldi, F. 2008a. Data assimilation as a nonlinear dynamical systems problem: Stability and convergence of the prediction-assimilation system. *Chaos* **18**, 023112.

Carrassi, A., Trevisan, A., Descamps, L., Talagrand, O. and Uboldi, F. 2008b. Controlling instabilities along a 3DVar analysis cycle by assimilating in the unstable subspace: A comparison with the EnKF. *Nonlinear Processes Geophys.* **15**, 503–521.

Carrassi, A., Trevisan, A. and Uboldi, F. 2007. Adaptive observations and assimilation in the unstable subspace by breeding on the data-assimilation system. *Tellus A* **59**, 101–113.

Carrassi, A., Vannitsem, S., Zupanski, D. and Zupanski, M. 2009. The maximum likelihood ensemble filter performances in chaotic systems. *Tellus A* **61**, 587–600.

Cohn, S. E., Sivakumaran, N. S. and Todling, R. 1994. A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.* **122**, 2838–2867.

Cosme, E., Verron, J., Brasseur, P., Blum, J. and Auroux, D. 2012. Smoothing problems in a Bayesian framework and their linear Gaussian solutions. *Mon. Wea. Rev.* **140**, 683–695.

Eckmann, J.-P. and Ruelle, D. 1985. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57**, 617–656.

Evensen, G. 2009. *Data Assimilation: The Ensemble Kalman Filter*, 2nd ed. Springer-Verlag, Berlin.

Ginelli, F., Chaté, H., Livi, R. and Politi, A. 2013. Covariant Lyapunov vectors. *J. Phys. A: Math. Theor.* **46**, 254005.

Gurumoorthy, K. S., Grudzien, C., Apte, A., Carrassi, A. and Jones, C. K. R. T. 2017. Rank deficiency of Kalman error covariance matrices in linear time-varying system with deterministic evolution. *SIAM J. Control Optim*. Accepted for publication, arXiv preprint arXiv:1503.05029.

Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J. and co-authors. 2004. *Four-dimensional ensemble Kalman filtering. Tellus A* **56**, 273–277.

Hunt, B. R., Kostelich, E. J. and Szunyogh, I. 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D* **230**, 112–126.

Jazwinski, A. H. 1970. *Stochastic Processes and Filtering Theory* Academic Press, New York.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *J. Fluids Eng.* **82**, 35–45.

Kalnay, E. 2002. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, New York.

Kuptsov, P. V. and Parlitz, U. 2012. Theory and computation of covariant Lyapunov vectors. *J. Nonlinear Sci.* **22**, 727–762.

Legras, B. and Vautard, R. 1996. A guide to lyapunov vectors. In: *ECMWF Workshop on Predictability*, ECMWF, Reading, UK, 135–146.

Liu, C., Xiao, Q. and Wang, B. 2008. An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. *Mon. Wea. Rev.* **136**, 3363–3373.

Lorenz, E. N. and Emanuel, K. A. 1998. Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.* **55**, 399–414.

Ng, G.-H. C., McLaughlin, D., Entekhabi, D. and Ahanin, A. 2011. The role of model dynamics in ensemble Kalman filter performance for chaotic systems. *Tellus A* **63**, 958–977.

Palatella, L., Carrassi, A. and Trevisan, A. 2013. Lyapunov vectors and assimilation in the unstable subspace: Theory and applications. *J. Phys. A: Math. Theor.* **46**, 254020.

Palatella, L. and Trevisan, A. 2015. Interaction of Lyapunov vectors in the formulation of the nonlinear extension of the Kalman filter. *Phys. Rev. E* **91**, 042905.

Sakov, P. and Oke, P. R. 2008. Implications of the form of the ensemble transformation in the ensemble square root filters. *Mon. Wea. Rev.* **136**, 1042–1053.

Sakov, P., Oliver, D. S. and Bertino, L. 2012. An iterative EnKF for strongly nonlinear systems. *Mon. Wea. Rev.* **140**, 1988–2004.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2003. Ensemble square root filters. *Mon. Wea. Rev.* **131**, 1485–1490.

Trevisan, A., D'Isidoro, M. and Talagrand, O. 2010. Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension. *Q. J. R. Meteorol. Soc.* **136**, 487–496.

Trevisan, A. and Palatella, L. 2011. On the Kalman filter error covariance collapse into the unstable subspace. *Nonlinear Processes Geophys.* **18**, 243–250.

Trevisan, A. and Pancotti, F. 1998. Periodic orbits, Lyapunov vectors, and singular vectors in the Lorenz system. *J. Atmos. Sci.* **55**, 390–398.

Trevisan, A. and Uboldi, F. 2004. Assimilation of standard and targeted observations within the unstable subspace of the observation-analysis-forecast cycle. *J. Atmos. Sci.* **61**, 103–113.

Uboldi, F. and Trevisan, A. 2006. Detecting unstable structures and controlling error growth by assimilation of standard and adaptive observations in a primitive equation ocean model. *Nonlinear Processes Geophys.* **16**, 67–81.

Vannitsem, S. and Lucarini, V. 2016. Statistical and dynamical properties of covariant Lyapunov vectors in a coupled atmosphere-ocean model-multiscale effects, geometric degeneracy, and error dynamics. *J. Phys. A: Math. Theor.* **49**, 224001.

Wolfe, C. L. and Samelson, R. M. 2007. An efficient method for recovering Lyapunov vectors from singular vectors. *Tellus A* **59**, 355–366.

Zupanski, M. 2005. Maximum likelihood ensemble filter: Theoretical aspects. *Mon. Wea. Rev.* **133**, 1710–1726.

Appendix 1.

In this appendix, an elementary dimensional analysis is performed on an ensemble-based DA system applied to the L95 model. The model is governed by the following set of ordinary differential equations. For $i = 1, \ldots, n$:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = b(x_{i+1} - x_{i-2})x_{i-1} - ax_i + F \qquad (A1)$$

where F is the forcing parameter. The dimensional parameters b and a are introduced here to represent the impacts of nonlinear convection and linear dampening, respectively; $F = 8$, $b = a = 1$ in the reference configuration of Lorenz and Emanuel (1998) as well as in the experiments described in this study. Periodic boundary conditions are assumed: $x_{n+1} = x_1, x_0 = x_n$ and $x_{-1} = x_{n-1}$. The DA system is also determined by an observation error scale σ (typically its standard deviation), and a time interval between updates Δt during which the ensemble is forecast.

Assuming a dimension for all these parameters, three sufficient adimensional factors stand out:

$$\alpha = a\Delta t, \qquad \beta = \sigma b \Delta t, \qquad \phi = Fb(\Delta t)^2. \qquad (A2)$$

Hence, two form factors, f and g, can be associated to the adimensional normalized RMSE of the state estimate and to a typical (adimensional) angle that characterizes the relative position of the perturbations of the DA system and the unstable–neutral subspace, respectively, such that

$$\frac{\text{RMSE}}{\sigma} = f\left(\alpha, \beta, \phi\right), \qquad (A3)$$

$$\theta = g\left(\alpha, \beta, \phi\right). \qquad (A4)$$

In a regime not too far from Gaussianity, the form factor f can be approximated as follows. The linearization of the model Equation (A1) around some mean value $\overline{x}$ yields the following evolution equation for the perturbations:

$$\frac{\mathrm{d}\delta x_i}{\mathrm{d}t} = \overline{x}b\left(\delta x_{i+1} - \delta x_{i-2}\right) - a\delta x_i, \qquad (A5)$$

which can be written in Fourier space ($1 \leq k \leq n$):

$$\frac{\mathrm{d}\delta \widehat{x}_k}{\mathrm{d}t} = (c_k \overline{x}b - a)\, \delta \widehat{x}_k, \qquad (A6)$$

where c_k is wavenumber dependent. The mean state $\overline{x}$ is a pure dynamical quantity that depends on a, b and F only. From dimensional analysis, it can be represented as

$$\overline{x} = \frac{F}{a}h\left(\frac{Fb}{a^2}\right), \qquad (A7)$$

where h is a one-variable form factor. If $\overline{x}b/a > 9/8$, which can be numerically seen to happen when $F > 3.10$ and $a = b = 1$, there is a non-zero bandwidth in Fourier space for which $c_k \overline{x}b - a > 0$ (Lorenz and Emanuel, 1998), which generates an asymptotic non-zero analysis RMSE. If the DA system is close to Gaussian, a one-variable asymptotic analysis shows (e.g. Bocquet et al., 2015) that if c is the maximum of c_k, the normalized average RMSE of the analysis scales as:

$$\frac{\text{RMSE}}{\sigma} \simeq \sqrt{1 - e^{-2(c\overline{x}b-a)\Delta t}}. \qquad (A8)$$

This is quite consistent with the normalized RMSEs shown in the left panel of Fig. 7. In particular, in the regime where the perturbation around $\overline{x}$ is valid, the normalized RMSE is independent from σ, i.e. f in Equation (A3) does not depend on β.

Because b, which scales like the magnitude of nonlinearity, always comes with Δt in β as well as in ϕ (see Equations (A2) and (A3)), the limit $\Delta t \to 0$ ensures that the system becomes linear and that

$$\frac{\text{RMSE}}{\sigma} \propto \sqrt{(c\overline{x}b - a)\,\Delta t}. \qquad (A9)$$

This binding of b to Δt also applies to Equation (A4) implying that when $\Delta t \to 0$, the angle θ should also vanish since it does for linear models (Sections 2 and 3). This matches the trend observed in Fig. 5. We also numerically observed that, in this limit, all principal angles slowly vanish (not shown).

Differently from Δt, σ multiplies b only in β as seen from Equation (A2). Hence, there is no guarantee that a linear regime is reached when $\sigma \rightarrow 0$ if $F > 0$ and $\Delta t > 0$ are fixed, in spite of the fact that the dispersion of the ensemble and the (unnormalized) RMSE vanish linearly with σ. This supports the convergence of the angle θ to a fixed non-zero value as seen in Fig. 4. We also numerically observed that, in this limit, all principal angles converge to a finite value (not shown).

In the limit $F \rightarrow 0$, or at least when F decreases below the regime where the unstable bandwidth disappear, and if $\Delta t > 0$ and $b > 0$ remain fixed, there is no guarantee that θ vanishes. Indeed, F multiplies b in ϕ but not in β. The model could become non-chaotic, and even stable, but would remain non-linear. Numerically, we observed that, in this limit, only a few principal angles vanish while the others remain at finite values (not shown).