

On the Kalman Filter error covariance collapse into the unstable subspace

A. Trevisan¹ and L. Palatella²

¹Istituto di Scienze dell'Atmosfera e del Clima del CNR, via Gobetti 101, 40129 Bologna, Italy

²Istituto di Scienze dell'Atmosfera e del Clima del CNR, UdR di Lecce, via Lecce-Monteroni, km 1,200, 73100 Lecce, Italy

Received: 9 February 2011 – Revised: 10 March 2011 – Accepted: 21 March 2011 – Published: 28 March 2011

Abstract. When the Extended Kalman Filter is applied to a chaotic system, the rank of the error covariance matrices, after a sufficiently large number of iterations, reduces to $N^+ + N^0$ where N^+ and N^0 are the number of positive and null Lyapunov exponents. This is due to the collapse into the unstable and neutral tangent subspace of the solution of the full Extended Kalman Filter. Therefore the solution is the same as the solution obtained by confining the assimilation to the space spanned by the Lyapunov vectors with non-negative Lyapunov exponents. Theoretical arguments and numerical verification are provided to show that the asymptotic state and covariance estimates of the full EKF and of its reduced form, with assimilation in the unstable and neutral subspace (EKF-AUS) are the same. The consequences of these findings on applications of Kalman type Filters to chaotic models are discussed.

1 Introduction

This work is motivated by the problem of data assimilation in meteorology and oceanography. Atmospheric and oceanic observations are noisy and very scattered in space and time. Dynamical models describing the evolution of the state of the atmosphere and the ocean are chaotic and, moreover, the number of degrees of freedom is huge ($> 10^8$). In the Earth sciences, the classical problem of estimating the state from noisy and incomplete observations and the approximate knowledge of the equations governing the system's evolution is known as data assimilation.

Data assimilation methods can be classified in two categories: sequential and variational, the most notable in the two classes being Kalman Filters and four-dimensional

variational assimilation (4DVar), respectively (Ghil and Malanotte-Rizzoli (1991); Kalnay (2003) and references therein). 4DVar is an advanced technique that seeks the model trajectory that best fits the observations distributed within a given time interval with the dynamical constraint of the model equations (Talagrand and Courtier, 1987). The optimal control theory (Le Dimet and Talagrand, 1986) allows the minimization of the 4D-Var cost function to be made with respect to the state at the beginning of the interval. The solution of the minimization problem is obtained by forward integration of the model and backward integration of the adjoint of the tangent linear propagator that yields the expression of the cost function gradient.

The Kalman Filter was originally developed for linear systems. Based on information about the error covariance of the data and model forecast, the KF provides the optimal linear estimate of the state of the system. A straightforward way of extending the linear results to the nonlinear case is given by the Extended Kalman Filter (EKF) (Jazwinski, 1970; Miller et al., 1994): in the EKF the error covariance is propagated in time according to linearized model dynamics. In the forecast step the nonlinear model is integrated, starting from the previous analysis to obtain the background state, and the tangent linear equations propagate the analysis error covariance to obtain an estimate of the background error covariance. When observations become available they are assimilated in the analysis step that combines the background and the observations with the appropriate weights. The error covariance associated with the state estimate at the analysis step is also updated.

The problem of reducing the cost of the EKF that is prohibitive in many realistic circumstances has been addressed by a number of authors. In geophysical science literature reduced-rank approximations of the full EKF (Fukumori, 2002; Ghil and Malanotte-Rizzoli, 1991; Tippett et al., 2000; Todling and Cohn, 1994) confine the forecast error covariance matrix to a subspace of lower dimension by means of



Correspondence to: A. Trevisan
(a.trevisan@isac.cnr.it)

Singular Value Decomposition (SVD), Eigenvalue Decomposition (EVD) or projection on the leading Empirical Orthogonal Functions (Pham et al., 1998). A Monte Carlo approach, referred to as Ensemble Kalman Filter (EnKF) (Evensen, 1994), provides ensemble representations for the probability distribution of the state estimate. The EnKF has also proven effective in reducing the computational cost associated with the full EKF. In the EnKF, the ensemble-predicted error statistics are estimated from the ensemble perturbations. In the forecast step their evolution is computed as difference between nonlinear integrations of the model, a procedure similar to the breeding method (Toth and Kalnay, 1993, 1997). For sufficiently small analysis errors, this procedure would give the same result as using the tangent linear propagator.

The various flavors of the Ensemble Kalman Filter (EnKF) testify the importance of this problem for real world applications (see Blum et al., 2008; Kalnay, 2003 and references therein). These include several ad-hoc refinements as perturbing observations, covariance localization, additive or multiplicative inflation. In particular, covariance localization is beneficial to prevent filter divergence when the number of ensemble members is small (for a review see Evensen, 2003). In all these studies, how many members of the ensemble are needed is empirically estimated on a case to case basis.

Ott et al. (2004) estimate the local dimensionality (E-dimension) from the perturbations of the Local Ensemble Kalman Filter and the number of members of the ensemble necessary to represent the forecast error covariance in different geographical regions (tropics, extra-tropics, polar regions).

We consider the application of the EKF to a chaotic system and we concentrate on issues related to the dimension of its unstable subspace. So et al. (1994) addressed a similar problem in control theory; in reconstructing the state vector of a chaotic system from the time series of an observed scalar, they showed that the control vector lies in the unstable space. The Assimilation in the Unstable Subspace (AUS), developed by Trevisan and co-authors, consists in confining the analysis update to the subspace spanned by the leading unstable directions (Trevisan and Uboldi, 2004). Applications to atmospheric and oceanic models (Uboldi and Trevisan, 2006; Carrassi et al., 2008b) showed that even dealing with high-dimensional systems, an efficient error control can be obtained by monitoring only a limited number of unstable directions. The forecast error in these directions was estimated with empirical techniques. More recently, Trevisan et al. (2010) formulated a reduced subspace 4-dimensional assimilation algorithm, 4DVar-AUS (Four-dimensional Variational Assimilation in the Unstable Subspace). The key result of this study is the existence of an optimal subspace dimension for the assimilation that is approximately equal to the unstable and neutral subspace dimension.

The methodology of the present paper goes back to the roots of the filtering theory, the EKF, to address questions regarding the number of degrees of freedom that describe the filter error evolution in a chaotic system of given unstable manifold dimension. In this work we first show how it is possible to define a mathematically rigorous algorithm with the following properties: the solution of the Kalman filter equations is reproduced when all degrees of freedom are considered; with assimilation increments limited to a subspace of the tangent space, a reduced order algorithm is obtained where the estimated errors are confined to the most unstable subspace of the system. We will discuss the equivalence of the full EKF with its reduced form (EKF-AUS) with assimilation in the unstable and neutral space, i.e. the span of the first $N^+ + N^0$ Lyapunov vectors, where N^+ and N^0 are the number of positive and null Lyapunov exponents. Then we compare numerical solutions obtained with the full EKF and the reduced algorithm EKF-AUS. Throughout the paper, the unified notation of (Ide et al., 1997) is used.

The results of this paper are based on the hypothesis that observations are accurate enough to ensure small analysis errors so that the errors evolution is correctly described by the tangent linear operator. In practice these are the same hypothesis of validity of the standard EKF algorithm. This hypothesis is crucial for all the considerations we will make regarding the Lyapunov exponents and vectors, since they are well-defined along the true trajectory while here we extend their properties to the pseudo-trajectory determined by the forecast-analysis cycle. This is reasonable if the pseudo-trajectory remains close enough to the true trajectory.

If the observations are so few and noisy that the analysis errors are large, the theoretical results of this paper do not apply directly. However, a straightforward generalization of the EKF-AUS algorithm to account for the nonlinear error behavior is possible, as discussed in the conclusions. Due to nonlinear error dynamics, the perturbations will not remain exactly in the unstable and neutral subspace and it may be necessary to use a larger subspace. Indeed the results obtained with the 4DVar-AUS algorithm (Trevisan et al., 2010) showed that, when observation errors are large enough that nonlinearity becomes important, the dimension of the subspace where errors live also increases; in such case a larger number of perturbations is needed.

2 The extended Kalman Filter and its reduction to the unstable space

2.1 The extended Kalman Filter

In the extended Kalman filter, the state evolves according to the full nonlinear equations and the tangent linear operator is used to predict the approximate error statistics. The estimate of the state, referred to as the analysis \mathbf{x}^a , is obtained by combining a forecast state \mathbf{x}^f with the possibly incomplete and

noisy p -dimensional observations $\mathbf{y}_k^0 = \mathcal{H}(\mathbf{x}_k) + \varepsilon_k^0$ given at discrete times $t_k > t_0$, $k \in \{1, 2, \dots\}$. The observation error ε_k^0 is assumed to be Gaussian with zero mean and known $p \times p$ covariance matrix \mathbf{R} . \mathcal{H} is the observation operator. The estimate update is given by the analysis equation:

$$\mathbf{x}_k^a = \mathbf{x}_k^f - \mathbf{K}_k \mathcal{H}(\mathbf{x}_k^f) + \mathbf{K}_k \mathbf{y}_k^0, \quad (1)$$

where \mathbf{x}_k^f is the *forecast* obtained by integrating the model equations from a previous analysis time:

$$\mathbf{x}_k^f = \mathcal{M}(\mathbf{x}_{k-1}^a), \quad (2)$$

\mathcal{M} being the nonlinear evolution operator. \mathbf{K}_k is the *gain matrix* at time t_k given by:

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}^T (\mathbf{H} \mathbf{P}_k^f \mathbf{H}^T + \mathbf{R})^{-1}, \quad (3)$$

where \mathbf{H} is the Jacobian of \mathcal{H} . The analysis error covariance update equation is given by:

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_k^f, \quad (4)$$

and \mathbf{P}_k^f , the forecast error covariance, is given by:

$$\mathbf{P}_k^f = \mathbf{M}_k \mathbf{P}_{k-1}^a \mathbf{M}_k^T, \quad (5)$$

where \mathbf{M} is the linearized evolution operator associated with \mathcal{M} . We have assumed that there is no model error.

2.2 The algorithm EKF-AUS

We introduce an algorithm that belongs to the family of square-root implementations of the EKF (Thornton and Bierman, 1976; Tippett et al., 2003). A reduced version is then obtained from the full rank algorithm. We perform the assimilation in a manifold of dimension m . When m is equal to the number n of degrees of freedom of the system, the algorithm solves the standard EKF equations.

We consider a chaotic system with a number N^+ of positive Lyapunov exponents and N^0 of null Lyapunov exponents. When $m = N^+ + N^0$ the reduced form, with Assimilation in the Unstable Subspace (EKF-AUS) is obtained.

At time $t = t_{k-1}$, let the $n \times m$ matrix \mathbf{X}^a be one of the square roots of \mathbf{P}^a , namely $\mathbf{P}^a = \mathbf{X}^a \mathbf{X}^{aT}$ and let the columns of $\mathbf{X}^a = [\delta \mathbf{x}_1^a, \delta \mathbf{x}_2^a, \dots, \delta \mathbf{x}_m^a]$ be orthogonal. At time $t = t_0$, the vectors $\delta \mathbf{x}_i^a$, $i = 1, 2, \dots, m$ are arbitrary independent initial perturbations. (Here and in the following we drop the time-step subscript from the equations since, unless otherwise stated, all terms refer to the same time step k).

In the standard EKF algorithm, the number of perturbations is equal to the total number of degrees of freedom of the system, $m = n$. In the reduced order algorithm referred to as EKF-AUS the number of perturbations is equal to the dimension of the unstable and neutral manifold: $m = N^+ + N^0$

2.2.1 The forecast step

In the forecast step, the tangent linear operator \mathbf{M} acts on the perturbations \mathbf{X}^a defined at (analysis) time $t = t_{k-1}$ (other terms in this and the following equations refer to time step $t = t_k$):

$$\mathbf{X}^f = \mathbf{M} \mathbf{X}^a \quad (6)$$

where $\mathbf{X}^f = [\delta \mathbf{x}_1^f, \delta \mathbf{x}_2^f, \dots, \delta \mathbf{x}_m^f]$. The $n \times n$ forecast error covariance matrix:

$$\mathbf{P}^f = \mathbf{X}^f \mathbf{X}^{fT} \quad (7)$$

can be cast in the form:

$$\mathbf{P}^f = \mathbf{E}^f \mathbf{\Gamma}^f \mathbf{E}^{fT} \quad (8)$$

where the m columns of \mathbf{E}^f are obtained by a Gram-Schmidt orthonormalization of the columns of \mathbf{X}^f . The $m \times m$ (in general non-diagonal) symmetric matrix $\mathbf{\Gamma}^f$ defined as:

$$\mathbf{\Gamma}^f = \mathbf{E}^{fT} \mathbf{X}^f \mathbf{X}^{fT} \mathbf{E}^f \quad (9)$$

represents the forecast error covariance matrix, confined to the subspace \mathcal{S}^m of the evolved perturbations. In the standard EKF algorithm ($m = n$), \mathbf{E}^f is $n \times n$ and its columns span the full space. In the reduced form algorithm, \mathbf{E}^f is $n \times m$ and its columns span an m -dimensional subspace \mathcal{S}^m of the entire phase space.

2.2.2 The analysis step

Using the definition of \mathbf{P}^f of Eq. (8) the Kalman gain expression becomes:

$$\mathbf{K} = \mathbf{E}^f \mathbf{\Gamma}^f (\mathbf{H} \mathbf{E}^f)^T \left[(\mathbf{H} \mathbf{E}^f) \mathbf{\Gamma}^f (\mathbf{H} \mathbf{E}^f)^T + \mathbf{R} \right]^{-1} \quad (10)$$

and the usual analysis error covariance update, Eq. (4) reads:

$$\begin{aligned} \mathbf{P}^a &= (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{E}^f \mathbf{\Gamma}^f \mathbf{E}^{fT} = \mathbf{E}^f \mathbf{\Gamma}^f \mathbf{E}^{fT} + \\ &\quad - \mathbf{E}^f \mathbf{\Gamma}^f (\mathbf{H} \mathbf{E}^f)^T \left[(\mathbf{H} \mathbf{E}^f) \mathbf{\Gamma}^f (\mathbf{H} \mathbf{E}^f)^T + \mathbf{R} \right]^{-1} \mathbf{H} \mathbf{E}^f \mathbf{\Gamma}^f \mathbf{E}^{fT} \\ &= \mathbf{E}^f \left\{ \mathbf{\Gamma}^f - \mathbf{\Gamma}^f \mathbf{E}^{fT} \mathbf{H}^T \left[(\mathbf{H} \mathbf{E}^f) \mathbf{\Gamma}^f (\mathbf{H} \mathbf{E}^f)^T + \mathbf{R} \right]^{-1} \right. \\ &\quad \left. \times \mathbf{H} \mathbf{E}^f \mathbf{\Gamma}^f \right\} \mathbf{E}^{fT} \equiv \mathbf{E}^f \mathbf{\Gamma}^{a'} \mathbf{E}^{fT} \end{aligned} \quad (11)$$

The analysis error covariance matrix \mathbf{P}^a can be written as:

$$\begin{aligned} \mathbf{P}^a &= \mathbf{E}^f \mathbf{\Gamma}^{a'} \mathbf{E}^{fT} = \mathbf{E}^f \mathbf{U} \mathbf{\Gamma}^a \mathbf{U}^T \mathbf{E}^{fT} \\ &\equiv \mathbf{E}^a \mathbf{\Gamma}^a \mathbf{E}^{aT}, \end{aligned} \quad (12)$$

where the columns of the $(m \times m)$ orthogonal invertible matrix \mathbf{U} are the eigenvectors of the symmetric matrix $\mathbf{\Gamma}^{a'} = \mathbf{U} \mathbf{\Gamma}^a \mathbf{U}^T$ (to numerically obtain the eigenvalues and the eigenvectors of $\mathbf{\Gamma}^{a'}$ we use the power iterations method (Golub et al., 1996)) and $\mathbf{\Gamma}^a = \text{diag}[\gamma_i^2]$ is diagonal. Therefore, the m columns of $\mathbf{E}^a = [\mathbf{e}_1^a, \mathbf{e}_2^a, \dots, \mathbf{e}_m^a]$ obtained by

$$\mathbf{E}^a = \mathbf{E}^f \mathbf{U} \quad (13)$$

span the same subspace \mathcal{S}^m as the columns of \mathbf{E}^f . Consequently, the analysis step preserves the subspace \mathcal{S}^m , an important point in the subsequent discussion (see Sec. 2.3) on the comparison of our scheme with the Benettin et al. (1980) algorithm. The square root of \mathbf{P}^a , written as $\mathbf{P}^a = \mathbf{E}^a \mathbf{\Gamma}^a \mathbf{E}^{aT} \equiv \mathbf{X}^a \mathbf{X}^{aT}$, provides a set of orthogonal vectors:

$$\mathbf{X}^a = \mathbf{E}^a (\mathbf{\Gamma}^a)^{1/2}. \quad (14)$$

The columns of

$$\mathbf{X}^a = [\gamma_1 \mathbf{e}_1^a, \gamma_2 \mathbf{e}_2^a, \dots, \gamma_m \mathbf{e}_m^a] = [\delta \mathbf{x}_1^a, \delta \mathbf{x}_2^a, \dots, \delta \mathbf{x}_m^a] \quad (15)$$

are the new set of perturbation vectors defined after the analysis step at time t_k that enter the forecast step (6) at the next time step t_{k+1} . Their amplitude is consistent with the analysis error covariance in the subspace \mathcal{S}^m

Notice that, as in other Kalman square root filters, with the introduction of \mathbf{E}^a and \mathbf{E}^f , forming the full forecast and analysis error covariance matrices can be avoided. The analysis equation is the usual Eq. (1) with \mathbf{K} given by (10). When $m = n$, \mathbf{K} is the usual Kalman gain. In EKF-AUS ($m = N^+ + N^0$) the analysis increment is confined to the subspace \mathcal{S}^m spanned by the m columns of \mathbf{E}^f in view of the form of \mathbf{K} .

2.2.3 Numerical Implementation

In the numerical implementation of the algorithm one can start with a number of perturbations m that is at least as large as $N^+ + N^0$. If an independent estimate of $N^+ + N^0$ is not available one can start with an initial guess for m that is an overestimate of $N^+ + N^0$. We now summarize the different steps of the EKF-AUS algorithm as they are performed in the numerical implementation:

1. we evolve the state vector \mathbf{x}_{k-1}^a with the full nonlinear model equations and the $n \times m$ matrix \mathbf{X}_{k-1}^a using the tangent linear operator \mathbf{M}_k , as in Eq. (6), obtaining \mathbf{X}_k^f at time step k
2. we orthonormalize the m columns of \mathbf{X}_k^f obtaining the columns of \mathbf{E}^f
3. we calculate the $m \times m$ covariance matrix $\mathbf{\Gamma}^f$ through Eq. (9)
4. we perform the analysis on the state vector as in Eq. (1) with \mathbf{K} given by Eq. (10)
5. we calculate $m \times m$ covariance matrix $\mathbf{\Gamma}^{a'}$ that after Eq. (11) can be written as

$$\mathbf{\Gamma}^{a'} = \mathbf{\Gamma}^f - \mathbf{\Gamma}^f \mathbf{E}^{fT} \mathbf{H}^T \times [(\mathbf{H} \mathbf{E}^f) \mathbf{\Gamma}^f (\mathbf{H} \mathbf{E}^f)^T + \mathbf{R}]^{-1} \mathbf{H} \mathbf{E}^f \mathbf{\Gamma}^f. \quad (16)$$

(notice that the matrix to be inverted has the dimension p of the number of observations)

6. we diagonalize the $m \times m$ matrix $\mathbf{\Gamma}^{a'}$. We put the normalized eigenvectors of $\mathbf{\Gamma}^{a'}$ in the orthogonal $m \times m$ matrix \mathbf{U} and then we obtain $\mathbf{E}^a = \mathbf{E}^f \mathbf{U}$. Both \mathbf{E}^f and \mathbf{E}^a are $n \times m$.
7. we multiply the i -th column of \mathbf{E}^a by the square root γ_i of i -th eigenvalue of $\mathbf{\Gamma}^a$ obtaining the new set of perturbations \mathbf{X}_k^a .

2.3 Further discussion on the EKF-AUS algorithm

In summary, the EKF-AUS algorithm is obtained by reducing the dimension of the subspace, \mathcal{S}^m , where the analysis update, $\mathbf{x}^a - \mathbf{x}^f$ and the estimated analysis and forecast error covariance \mathbf{P}^a and \mathbf{P}^f are confined. It will soon become clear why m is set equal to the dimension of the unstable and neutral subspace.

During the forecast step the error evolution is governed by Eq. (6). Suppose that at a certain time, t_k the first $N^+ + N^0$ of the m perturbations are confined to the unstable and neutral subspace, that is the columns of \mathbf{E}^a span the same subspace as the leading $N^+ + N^0$ Lyapunov vectors. After the forecast step, this subspace is mapped into the unstable subspace referred to the state \mathbf{x}^f at time t_{k+1} . The new perturbations \mathbf{X}^f are amplified but they are still confined in the subspace of the leading $N^+ + N^0$ Lyapunov vectors. These perturbations will survive throughout the long time assimilation process. The remaining $m - N^+ - N^0$ perturbations, being recurrently orthogonalized, are deprived of the component along the unstable manifold like in the Benettin et al. (1980) algorithm, while the component along the stable manifold is damped during the forecast (the validity of this reasoning does not require the orthogonality of the stable and unstable manifolds).

The main difference between our algorithm and that of Benettin et al. (1980) is introduced in the analysis step. In our algorithm, use is made of Eq. (11) to redefine the perturbations. These are confined by construction to the subspace spanned by the columns \mathbf{E}^f and after the analysis step the span of the columns of \mathbf{E}^a and of the columns of \mathbf{X}^a will be the same space spanned by the columns of \mathbf{E}^f as shown in Sect. 2.2.2. For a detailed proof of this fact we refer to Appendix A.

Another difference from (Benettin et al., 1980) is that in our case the vectors at the end of the analysis step are orthogonal but not orthonormal, so that perturbations that are damped during the forecast are not artificially “kept alive” (as it happens with the orthonormalization step of Benettin et al., 1980). The $(\mathbf{I} - \mathbf{K}\mathbf{H})$ term has a stabilizing effect that tends to reduce the amplitude of perturbations in view of the existence of observations (Carrassi et al., 2008). As a consequence, when the amplitude of perturbations is redefined after the analysis step the decaying modes are further damped while the unstable modes are kept from diverging.

In conclusion, arbitrary initial perturbation vectors, the columns of \mathbf{X}^a at $t = t_0$, subject to successive forecast and

analysis steps will, in the long run, be confined to the unstable and neutral subspace. In view of this reasoning, based on the linearization of the error dynamics, it is argued that the EKF and EKF-AUS algorithms will asymptotically produce the same estimate of error covariances. The numerical results confirm the validity of these statements.

From previous considerations, if we choose an initial number of perturbations $m > N^+ + N^0$, $m - N^+ - N^0$ perturbations will be damped by the assimilation process. On the contrary, if we choose $m < N^+ + N^0$, some unstable directions will be ignored and, consequently, the filter will eventually diverge.

3 Numerical results

We now compare the numerical results obtained with the reduced EKF-AUS algorithm with those obtained with the full EKF, when the EKF algorithm itself is stable and no filter divergence is observed. Experiments are based on Lorenz96 model (Lorenz, 1996), that has been widely used for testing data assimilation algorithms. The governing equations are:

$$\frac{d}{dt}x_j = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F \quad (17)$$

with $j = 1, \dots, n$. The variables x_j represent the values of a scalar meteorological quantity at n equally spaced geographic sites on a periodic longitudinal domain. The model has chaotic behavior for the value of the forcing, $F = 8$, used in most studies. The number of variables n of the model is varied to obtain different systems with a different number of degrees of freedom and, consequently, a different number of positive Lyapunov exponents. With $n = 40, 60, 80$ the systems have 13, 19, 25 positive Lyapunov exponents, respectively.

All simulations are performed in a perfect model scenario, that is a trajectory on the attractor of the system is assumed to represent the *true* state evolution. Observations are created by adding uncorrelated random noise with Gaussian distribution (zero mean, variance σ_0^2) to the *true* state. The assimilation experiments with EKF and EKF-AUS use the same reference trajectory and the same observations. The performance of the assimilation is measured by the analysis error, the root mean square (*rms*) of the difference between the *true* and the estimated state vectors. Observations are taken at discrete times corresponding to the assimilation times t_k . Results shown refer to the following parameters but different choices of observational configurations gave qualitatively similar results. The time interval between observations is 0.05 (= 4 time integration steps); every other grid point is observed and the observation points are shifted by one at each observation time.

Figure 1 shows the *rms* estimation error for the three systems ($n = 40, 60, 80$) as a function of the observation error standard deviation, σ_0 . Values are in the range $0.002 \leq \sigma_0 \leq$

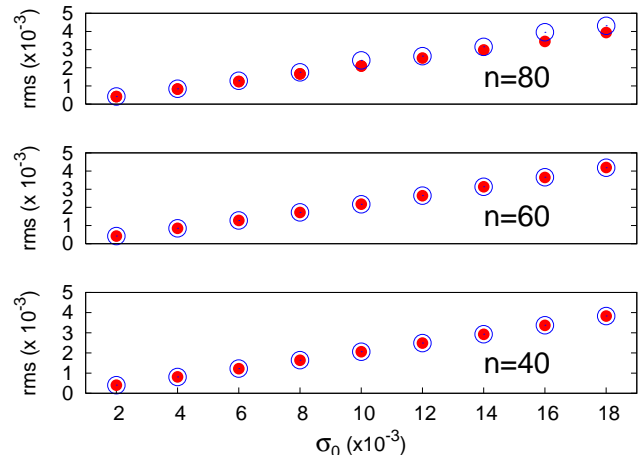


Fig. 1. Average root mean square error of the algorithms: EKF (full circles), EKF-AUS (empty circles) for systems with $n = 40, 60, 80$ degrees of freedom, as a function of σ_0 , the observation error standard deviation. The average was performed over 1000 assimilation steps (50 time units) after a stabilization period of the same length.

0.018 where filter divergence is not observed; for values of $\sigma_0 > 0.02$, divergence possibly occurs in both EKF and EKF-AUS algorithms. The *rms* error was averaged over 1000 successive assimilation times or 50 time units, after waiting a transient time of 50 time units for the filter error to stabilize. The sets of points displayed in the figure refer to EKF and EKF-AUS, ($m = 14, 20, 26$) and show that the average error is statistically the same. The results show that, in agreement with the theory, for sufficiently small values of σ_0 , error dynamics is linear and the estimation error grows linearly with the observation error. Notice also that the asymptotic average error is about the same in the three model configurations with $n = 40, 60, 80$. This is because the number of unstable directions is proportional to the extension of the spatial domain and the number of observations that can detect them increases by the same proportion.

To help interpret the finding that the two algorithms give statistically equivalent results, we study the behavior of the error covariance in the EKF. At time $t = 0$, the EKF is initialized with an arbitrary estimate of \mathbf{P}^a whose rank is the same as the total number of degrees of freedom. The rank of \mathbf{P}^a is shown in Fig. 2 for the case $\sigma_0 = 0.01$, as a function of the number of iterations and for three system configurations, ($n = 40, 60$ and 80). By rank of \mathbf{P}^a we mean in practice the number of eigenvalues that exceed a given threshold value. In all cases we observe that, changing this value in the range 10^{-8} – 10^{-11} , the asymptotic rank estimate varies by one unit. To obtain the eigenvalues of \mathbf{P}^a we use the power iterations method (Golub et al., 1996). As shown in Fig. 2, the rank value, initially equal to the total number of degrees of freedom, decreases until it gradually approaches a value very near, in each system, to the dimension of its unstable and neutral subspace.

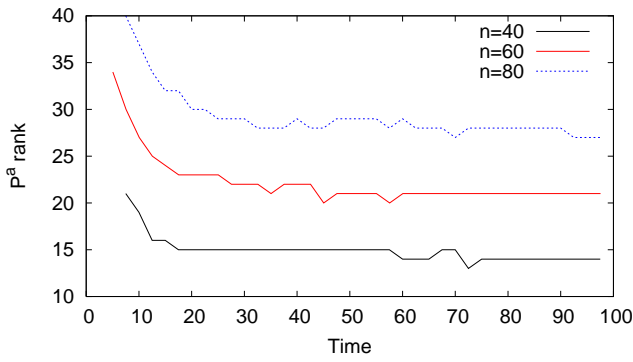


Fig. 2. Rank of the analysis error covariance matrix \mathbf{P}^a in the EKF algorithm. For $n = 40, 60, 80$, the dimension of the unstable and neutral space is 14, 20, 26, respectively. $\sigma_0 = 0.01$.

The analysis error covariance matrix was explicitly written in Eq. (12) in terms of the orthonormal basis \mathbf{E}^a of the analysis error, the eigenvalues of $\mathbf{\Gamma}^a$ being the error variances in this subspace; this formulation was useful to interpret the behavior of the error covariance in the stable and unstable directions. The estimated errors in the stable directions are damped during the forecast step and are reduced by the effect of the $(\mathbf{I} - \mathbf{KH})$ term at analysis step. We conclude that only errors in the unstable directions, that amplify during the forecast step, can survive along the sequence of successive forecast and analysis iterations.

The numerical results confirm this argument: the rank of the error covariance matrix of the EKF decreases until it reaches a value consistent with the expectation that, in the long run, only the unstable or weekly stable error components remain active. At the same time, the reduction of the rank of \mathbf{P}^a in the EKF explains the long-term equivalence in the performance of EKF and EKF-AUS: the subspace dimension of the covariance matrices in the two algorithms become asymptotically the same. Further numerical evidence of validity of the theory is provided by the spectrum of eigenvalues of the error covariance matrices. Fig. 3, obtained in the same numerical settings of Figs. 1 and 2, shows the eigenvalues of \mathbf{P}^a for EKF and of $\mathbf{\Gamma}^a$ for EKF-AUS computed at the final time of Fig. 2. The eigenvalues obtained with the two algorithms, in logarithmic scale in the figure, are the same within numerical accuracy.

The above results explain why the two algorithms, EKF and EKF-AUS, after a transient stage start to behave in a similar fashion and asymptotically give the same numerical analysis solution.

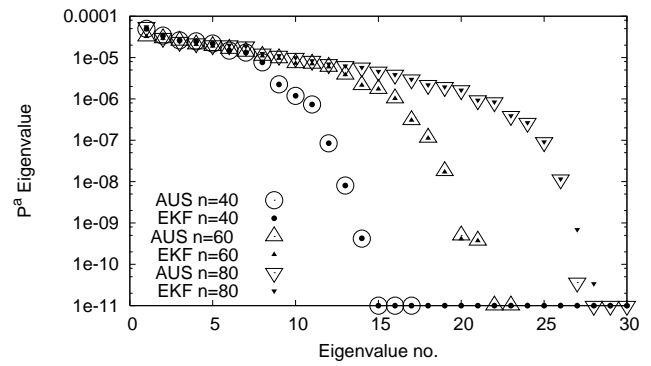


Fig. 3. Eigenvalues of the analysis error covariance matrix at final time of assimilation with: EKF (full) and EKF-AUS (empty) for systems with $n = 40, 60, 80$. $\sigma_0 = 0.01$.

4 Conclusions

We considered a system with a chaotic attractor and studied the consequences of the existence of an unstable manifold on the evolution of filter solution error. The algorithm we presented reproduces exactly the EKF equations when all degrees of freedom are considered. The idea of confining the assimilation in the invariant unstable and neutral subspace is not new and was exploited in a series of papers, including the formulation of the 4DVar-AUS algorithm (Trevisan et al., 2010); it was the original purpose of the present work to apply the same idea to the Kalman Filter. The most important and new result of the paper, not foreseen by the authors themselves, is that the extended Kalman Filter solution collapses into this invariant subspace so that its solution is not different from the solution of the reduced form of the algorithm (EKF-AUS). More specifically, the EKF algorithm and EKF-AUS algorithm with a number of degrees of freedom equal to the number of positive and neutral exponents produce the same asymptotic state and error covariance estimates. In a sense the EKF solution converges to the EKF-AUS solution, not viceversa. This happens because the rank of the full EKF error covariance matrices asymptotically becomes as small as the dimension of the unstable and neutral manifold of the original system equations. Theoretical arguments providing a rationale for this behavior were corroborated by numerical results.

It is worth discussing at this point why the 4DVar-AUS algorithm turns out to be superior to standard 4DVar, while EKF-AUS gives the same results as the standard EKF. The interpretation is simple. The full EKF performs exactly as EKF-AUS because the estimated asymptotic error covariances are the same. Standard 4DVar instead looks for the minimizing solution in the full space; 4DVar-AUS is superior to 4DVar because it looks for the solution in the unstable manifold without introducing observation errors in the stable one.

The present results were obtained in the framework of the EKF, but the important message is that we expect to find the same behavior, namely the collapse of the covariance matrices, in any Kalman type Filter whenever the estimation error is sufficiently small.

In an ensemble approach, when observations are sufficiently dense and accurate that error dynamics is approximately linear, we expect the necessary and sufficient number of ensemble members to be equal to the number of positive and null Lyapunov exponents, $N^+ + N^0$. If the number of ensemble members is too small or the ensemble members do not adequately span the subspace of the true error covariance, in many applications it has been found necessary to apply ad hoc fixes, such as additive or multiplicative covariance inflation or localization. However, these fixes, by arbitrarily modifying the perturbations subspace and disrupting their structure can have detrimental effects on the forecast (Hamill and Whitaker, 2011). Our algorithm ensures the independence of the ensemble perturbations and spanning of the unstable and neutral subspace.

In the Lorenz (1996) model, a good performance of the filter was obtained without the need of numerical fixes. In this, as well as in more complex meteorological models, the stability properties have small phase space variability and the local exponents associated to a globally unstable (stable) direction are generally positive (negative). Roughly speaking, the number of (locally) unstable directions varies little in phase space. Consequently, the eigenvalues of the error covariance matrix associated with the unstable directions remain numerically bounded from zero most of the time. If, on the contrary, a globally unstable direction becomes locally very stable for a sufficiently long period of time, the corresponding eigenvalue will become zero to numerical precision. When the local exponent becomes again positive, the eigenvalue will remain zero and the filter will diverge. This can explain why additive covariance inflation, that prevents rank reduction was found to be beneficial for the performance of ensemble Kalman Filters.

The present results suggest that the development of data assimilation schemes should exploit the chaotic properties of the forecast model. Regarding weather forecasting applications, we report the results of (Carrassi et al., 2007) where the authors find that a quasi-geostrophic model with 7 levels and 14784 ($= 64 \times 33 \times 7$) degrees of freedom has only 24 positive Lyapunov exponents. This result shows how a model with such large number of degrees of freedom, making the direct application of EKF unfeasible, becomes treatable with EKF-AUS with a number of perturbations of less than 1/600 of the number of the original degrees of freedom. The present arguments are corroborated by the successful application to operational forecasting of Ensemble Kalman Filters with a number of ensemble members that is orders of magnitude smaller than the number of degrees of freedom of the model.

It is worth noting that the EKF-AUS algorithm does not require an a-priori knowledge of the spectrum of the Lyapunov

exponents but only a reasonable upper limit for $N^+ + N^0$. Since in the long run only the unstable directions survive, the algorithm can be useful for another application: it can be exploited as an alternative, numerically efficient, methodology to obtain an approximate estimate of the unstable manifold dimension.

Appendix A

Proof of the equivalence of the span of the columns of \mathbf{E}^f , \mathbf{E}^a and \mathbf{X}^a

After the forecast step and orthonormalization we obtain the matrix \mathbf{E}^f . As said in the text, the estimate of error covariance after the analysis is obtained calculating the non-diagonal matrix $\mathbf{\Gamma}^{a'}$ that represents the error covariance matrix in the subspace spanned by the columns of \mathbf{E}^f . After that the matrix $\mathbf{\Gamma}^{a'}$ is diagonalized with a change of basis given by the orthogonal matrix \mathbf{U} . The matrix \mathbf{U} is then applied to \mathbf{E}^f in Eq. (12) to obtain \mathbf{E}^a . The i -th column of the matrix \mathbf{E}^a are then rescaled by the square root γ_i of the i -th eigenvalue of $\mathbf{\Gamma}^a$. To prove that the span \mathcal{S}^m of the columns of \mathbf{E}^a is the same as the span of the columns of \mathbf{E}^f we first observe that the subspace \mathcal{S}^m is m -dimensional and thus it can be defined by $n - m$ independent linear equations. An orthogonal basis \mathbf{v}_j , $j \in [1, n - m]$ of the orthogonal complement of \mathcal{S}^m can be used to identify the subspace \mathcal{S}^m by means of

$$\mathbf{v}_j^T \mathbf{e} = 0, \quad \forall j \in [1, n - m], \quad (\text{A1})$$

where $\mathbf{e} \in \mathcal{S}^m$ is the generic vector of \mathcal{S}^m . After the analysis the error matrix has the form $\mathbf{P}^a = \mathbf{E}^f \mathbf{U} \mathbf{\Gamma}^a (\mathbf{U} \mathbf{E}^f)^T \equiv \mathbf{E}^a \mathbf{\Gamma}^a \mathbf{E}^{aT}$ with $\mathbf{\Gamma}^a$ diagonal. The eigenvectors of this matrix are the columns of $\mathbf{E}^a = \mathbf{E}^f \mathbf{U}$ that are linear combinations of the columns of \mathbf{E}^f . These vectors still fulfill the condition (A1) and, consequently, still span \mathcal{S}^m . The next algorithm step of Eq. (14) only re-scales the columns of \mathbf{E}^a obtaining \mathbf{X}^a . Of course, after the rescaling, the columns of \mathbf{X}^a still span the same subspace. We thus prove that the span of the columns of \mathbf{E}^f is the same as that of the columns of \mathbf{E}^a and \mathbf{X}^a .

Acknowledgements. We thank E. Kalnay and O. Talagrand for their insightful comments and suggestions. This work has been funded by the Strategic Project: *Nowcasting con l'uso di tecnologie GRID e GIS*, PS080.

Edited by: O. Talagrand

References

- Benettin, G., Galgani L., Giorgilli, A., and Strelcyn, J. M.: Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing them, *Meccanica*, 15, 9–21, 1980.
- Blum, J., Le Dimet, F. X., and Navon, I.: Data Assimilation for Geophysical Fluids, *Handbook of Numerical Analysis, Special Volume on Computational Methods for the Ocean and the Atmosphere*, edited by: Temam, R. and Tribbia, J., Elsevier, New York, 1–63, 2008.
- Carrassi, A., Trevisan, A., and Uboldi, F.: Adaptive observations and assimilation in the unstable subspace by breeding on the data-assimilation system, *Tellus*, 59, 101–113, 2007.
- Carrassi, A., Ghil, M., Trevisan, A., and Uboldi, F.: Data assimilation as a nonlinear dynamical systems problem: Stability and convergence of the prediction-assimilation system, *Chaos*, 18, 023112, 2008.
- Carrassi, A., Trevisan, A., Descamps, L., Talagrand, O., and Uboldi, F.: Controlling instabilities along a 3DVar analysis cycle by assimilating in the unstable subspace: a comparison with the EnKF, *Nonlin. Proc. Geophys.*, 15, 503–521, 2008b.
- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99, 10143–10162, 1994.
- Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, *Ocean Dynamics*, 53, 343–367, 2003.
- Fukumori, I.: A Partitioned Kalman Filter and Smoother, *Mon. Wea. Rev.*, 130, 1370–1383, 2002.
- Ghil, M. and Malanotte Rizzoli, P.: Data assimilation in meteorology and oceanography, *Adv. Geophys.*, 33, 141–266, 1991.
- Golub, G. H. and von Loan, C. F.: *Matrix Computation*, John Hopkins University Press, Baltimore and London, 3rd Edition, ISBN 0-8018-5413-X, 1996.
- Hamill, T. M. and Whitaker, J. S.: What constrains Spread Growth in forecasts initialized from ensemble Kalman Filters?, *Mon. Weather Rev.*, 139, 117–131, 2011.
- Ide, K., Courtier, P., Ghil, M., and Lorenc, A.: Unified notation for data assimilation: Operational, sequential and variational, *J. Meteorol. Soc. Jpn.*, 75, 181–189, 1997.
- Jazwinski, A.H.: *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.
- Kalnay, E.: *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge, UK, 2003.
- Le Dimet, F. X. and Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations – theoretical aspects, *Tellus*, 38/2, 97–110, 1986.
- Lorenz, E. N.: Predictability: A problem partly solved. *Proc. Seminar on Predictability. European Center for Medium-Range Weather Forecasting: Shinfield Park, Reading, UK*, 1–18, 1996.
- Miller, R. N., Ghil, M., and Gauthier, F.: Advanced data assimilation in strongly nonlinear dynamical systems, *J. Atmos. Sci.*, 51, 1037–1056, 1994.
- Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A.: A local ensemble Kalman filter for atmospheric data assimilation, *Tellus A*, 56/5, 415–428, 2004.
- Pham, D. T., Verron, J., and Roubaud, M. C.: A singular evolutive extended Kalman filter for data assimilation in oceanography, *J. Marine Syst.*, 16, 323–340, 1998.
- So, P., Ott, E., and Dayawansa, W. P.: Observing chaos: Deducing and tracking the state of a chaotic system from limited observation, *Phys. Rev. E*, 49, 2650–2660, 1994.
- Talagrand, O., and Courtier, P.: Variational assimilation of observations with the adjoint vorticity equations, *Q. J. Roy. Meteorol. Soc.*, 113, 1311–1328, 1987.
- Thornton, C. L. and Bierman, G. J.: A numerical Comparison of Discrete Kalman Filtering Algorithms: An Orbit Determination Case Study, *JPL Technical Memorandum 33-771*, Pasadena, 1976.
- Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, C. H., and Whitaker, J. S.: Ensemble square root filters, *Mon. Weather Rev.*, 131, 1485–1490, 2003.
- Tippett, M. K., Cohn, S. E., Todling, R., and Marchesin, D.: Conditioning of the Stable, Discrete-Time Lyapunov Operator, *SIAM J. Matrix Anal. Appl.*, 22, 56–65, 2000.
- Todling, R. and Cohn, S. E.: Suboptimal Schemes for Atmospheric Data Assimilation Based on the Kalman Filter, *Mon. Wea. Rev.* 122, 2530–2557, 1994.
- Toth, Z. and Kalnay, E.: Ensemble Forecasting at NMC: The Generation of Perturbations, *Bull. Am. Met. Soc.* 74/12, 2317–2330, 1993.
- Toth, Z. and Kalnay, E.: Ensemble forecasting at NCEP and the breeding method, *Mon. Wea. Rev.*, 125/12, 3297–3319, 1997.
- Trevisan, A., D’Isidoro, M., and Talagrand, O.: Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension, *Q. J. Roy. Meteorol. Soc.*, 136, 487–496, 2010.
- Trevisan, A. and Uboldi, F.: Assimilation of Standard and Targeted Observations within the Unstable Subspace of the Observation-Analysis-Forecast Cycle System, *J. Atmos. Sci.*, 61, 103–113, 2004.
- Uboldi, F., and Trevisan, A.: Detecting unstable structures and controlling error growth by assimilation of standard and adaptive observations in a primitive equation ocean model. *Nonlin. Proc. Geophys.*, 16, 67–81, 2006.