Implications of the Form of the Ensemble Transformation in the Ensemble Square Root Filters

PAVEL SAKOV AND PETER R. OKE

CSIRO Marine and Atmospheric Research and Wealth from Oceans Flagship Program, Hobart, Australia

(Manuscript received 7 September 2006, in final form 27 June 2007)

ABSTRACT

This paper considers implications of different forms of the ensemble transformation in the ensemble square root filters (ESRFs) for the performance of ESRF-based data assimilation systems. It highlights the importance of using mean-preserving solutions for the ensemble transform matrix (ETM). The paper shows that an arbitrary mean-preserving ETM can be represented as a product of the symmetric solution and an orthonormal mean-preserving matrix. The paper also introduces a new flavor of ESRF, referred to as ESRF with mean-preserving random rotations. To investigate the performance of different solutions for the ETM in ESRFs, experiments with two small models are conducted. In these experiments, the performances of two mean-preserving solutions, two non-mean-preserving solutions, and a traditional ensemble Kalman filter with perturbed observations are compared. The experiments show a significantly better performance of the mean-preserving solutions for the ETM in ESRFs compared to non-mean-preserving solutions. They also show that applying the mean-preserving random rotations prevents the buildup of ensemble outliers in ESRF-based data assimilation systems.

1. Introduction

There is an extensive literature on the ensemble square root filter (ESRF; see Tippett et al. 2003 and references therein) that can be viewed as a deterministic ensemble-based implementation of the Kalman filter. An ESRF updates an ensemble of forecasts in two steps. First, the ensemble mean is updated with the standard Kalman filter analysis equation. Second, the ensemble anomalies are transformed so that the ensemble-based analysis error covariance matches the theoretical solution given by Kalman filter theory.

It has recently been recognized (Wang et al. 2004) that some solutions described in the literature (e.g., Bishop et al. 2001; Evensen 2004) do not preserve the ensemble mean during this transformation. As a result, both the analysis and analysis error covariance effectively do not match the theoretical estimates as the filters intend. The significance of this has not received due prominence in the literature and has therefore been overlooked by some researchers in the data assimilation community (e.g., Evensen 2004; Leeuwenburgh 2005; Leeuwenburgh et al. 2005; Torres et al. 2006).

Note that preserving the mean is irrelevant for applications that do not require the calculation of analysis, such as adaptive sampling. This particularly applies to many studies using the ensemble transform Kalman filter (ETKF), starting with the original work by Bishop et al. (2001); however, the non-mean-preserving solutions used in these studies have often became associated with the ETKF (e.g., Tippett et al. 2003).

Despite the theoretical deficiency of using non-mean-preserving solutions in data assimilation, the importance of using a mean-preserving solution in practice remains unclear. Wang et al. (2004) investigated the performance of a mean-preserving and non-mean-preserving ETKF in the ensemble generation context. They compared forecast errors in terms of the total energy norm for ensembles generated with both solutions, and reported only a "small improvement" of the mean-preserving (spherical simplex) ETKF over the non-mean-preserving (one-sided) ETKF. The authors have not investigated implications of using a non-mean-preserving solution for data assimilation, in which the degradation of the filter performance may be stronger because of the repetitive subtraction of the ensemble mean.

Corresponding author address: Pavel Sakov, CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, TAS 7053, Australia.
E-mail: pavel.sakov@csiro.au

DOI: 10.1175/2007MWR2021.1

In this paper, we present an overview of the ESRF theory and show that any mean-preserving ESRF can be presented in a form of the symmetric solution for the ensemble transform matrix (ETM) multiplied by an arbitrary orthonormal mean-preserving matrix. Subsequently, we conduct numerical experiments with two small models that clearly demonstrate a major advantage of mean-preserving solutions over non-mean-preserving solutions in data assimilation. We then discuss different mean-preserving solutions and their implications for data assimilation with a nonlinear model.

This manuscript is organized as follows; relevant background information is presented in section 2; a description of the experiments presented is given in section 3 followed by the results in section 4. Section 5 contains a brief discussion of the significance of our results, and our conclusions are in section 6.

2. Background

Ensemble Kalman filter (EnKF) methods are based on the Kalman filter equations:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{d} - \mathbf{H}\mathbf{x}^f) \quad \text{and} \tag{1}$$

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f, \quad \text{where} \tag{2}$$

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}; \tag{3}$$

$\mathbf{x}^a$ is the analysis; $\mathbf{x}^f$ is the forecast; $\mathbf{d}$ is the vector of observations; $\mathbf{H}$ is the observation sensitivity matrix, $\mathbf{H} = \nabla_x \mathcal{H}(\mathbf{x}^f)$, where $\mathcal{H}(\mathbf{x})$ is the (nonlinear) operator mapping state vector space to observation space; $\mathbf{R}$ is the observation error covariance matrix; $\mathbf{P}^f$ is the forecast error covariance matrix; and $\mathbf{P}^a$ is the analysis error covariance matrix. Ensemble methods are typically used for applications with relatively high-dimensional discrete systems, when explicit storage and manipulation with the system error covariance $\mathbf{P}$ are impossible or not practical. In such cases, the model state and the system error covariance are stored and manipulated implicitly via an ensemble $\mathbf{X}$ of model states, $\mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_m]$, where m is the ensemble size. The covariance matrix $\mathbf{P}$ is typically assumed to be carried by the ensemble by means of the relation

$$\mathbf{P} = \frac{1}{m-1}\sum_{i=1}^{m}(\mathbf{X}_i - \mathbf{x})(\mathbf{X}_i - \mathbf{x})^{\mathrm{T}} = \frac{1}{m-1}\mathbf{A}\mathbf{A}^{\mathrm{T}}, \tag{4}$$

where $\mathbf{x}$ is the ensemble mean,

$$\mathbf{x} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{X}_i, \tag{5}$$

and $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_m]$ is the ensemble of anomalies, or perturbations, $\mathbf{A}_i = \mathbf{X}_i - \mathbf{x}$; and the model state is assumed to be represented by the ensemble mean.

The traditional EnKF (Evensen 1994; Houtekamer and Mitchell 1998) is a Monte Carlo method, in which the ensemble represents the system error covariance in a statistical way. During the analysis, each ensemble member is updated according to the analysis equation (1):

$$\mathbf{X}_i^a = \mathbf{X}_i^f + \mathbf{K}(\mathbf{d} - \mathbf{H}\mathbf{X}_i^f), \quad i = 1, \ldots, m. \tag{6}$$

Because the system state is represented by the ensemble mean, the average of this equation yields the analysis equation (1), which is therefore automatically satisfied; however, this is not the case for the covariance equation (2). A straightforward application of (6) is known to result in an ensemble collapse, when the ensemble spread reduces too rapidly (Burgers et al. 1998). The standard way to prevent the ensemble collapse in the EnKF is to update each ensemble member using independently perturbed observations (Burgers et al. 1998; Houtekamer and Mitchell 1998); the resulting algorithm has became known as the EnKF with perturbed observations and is currently commonly associated with the acronym EnKF. Because in the EnKF with perturbed observations the covariance equation (2) is satisfied in a statistical sense only, it results in a suboptimal filter behavior, which is particularly evident for small ensembles (Whitaker and Hamill 2002).

As an alternative to the traditional EnKF, a number of methods have been proposed, identified as ESRFs, which allow a deterministic update of the ensemble anomalies, so that the analysis error covariance matches the theoretical value given by the Kalman filter (Tippett et al. 2003). In contrast to the EnKF, in an ESRF both the ensemble mean and the ensemble anomalies are updated explicitly. The ensemble mean is updated by using the analysis equation (1), while the ensemble anomalies are updated via an explicitly calculated transformation, represented by the ensemble transform matrix $\mathbf{T}$:

$$\mathbf{A}^a = \mathbf{A}^f\mathbf{T}, \tag{7}$$

such that the analyzed covariance matrix calculated from the analyzed anomalies $\mathbf{A}^a$ using (4) matches the theoretical value (2):

$$\mathbf{A}^f\mathbf{T}(\mathbf{A}^f\mathbf{T})^{\mathrm{T}} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{A}^f\mathbf{A}^{f\mathrm{T}}. \tag{8}$$

Because both the forecast and analyzed covariance belong to the same linear subspace spanned by the ensemble anomalies $\mathbf{A}^f$, all possible transformations of the ensemble can be represented in the form of (7).

The specific details of an ESRF depend on the flavor

and implementation; a brief description of four existing algorithms may be found in Tippett et al. (2003); see also Ott et al. (2004) and Evensen (2004). In this paper, we will use the formalism of the ETKF introduced by Bishop et al. (2001), for three reasons. First, the ETKF uses an ensemble transformation in the linear space of the ensemble (7), which seems quite natural, given that each analyzed anomaly is a linear combination of the forecast anomalies. Second, for many oceanographic and atmospheric applications, when the state vector dimension n and number of observations p are typically large compared with the ensemble size m, it represents the most numerically efficient flavor of the ESRF. When the inverse of the observation error covariance matrix $\mathbf{R}^{-1}$ is readily available, the computational cost of the ETKF is estimated as $O(m^2p + m^3 + m^2n)$ (Tippett et al. 2003), which, unlike some other solutions, is linear over n and p. Third, the ETKF may be shown to be formally equivalent to other flavors of ESRF in the sense that any two solutions for the ensemble transform matrix $\mathbf{T}$ that match (8) can be obtained one from another by a rotation in ensemble space.

A general form of an ETM that satisfies (8) is

$$\mathbf{T} = \mathbf{T}^s\mathbf{U}, \tag{9}$$

$$\mathbf{T}^s = \left[\mathbf{I} + \frac{1}{m-1}(\mathbf{H}\mathbf{A}^f)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{A}^f\right]^{-1/2} \tag{10}$$

(Bishop et al. 2001), and $[\cdot]^{-1/2}$ denotes the inverse of the unique positive definite square root of a positive definite matrix (Horn and Johnson 1985, p. 406); and $\mathbf{U}$ is an arbitrary orthonormal matrix, $\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I}$. Alternatively, a more complicated expression can be used:

$$\mathbf{T}^s = \left[\mathbf{I} - \frac{1}{m-1}(\mathbf{H}\mathbf{A}^f)^{\mathrm{T}}(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{H}\mathbf{A}^f\right]^{1/2} \tag{11}$$

(Evensen 2004), where $[\cdot]^{1/2}$ denotes the unique positive definite square root of a positive definite matrix, and $\mathbf{P}^f$ is used as an abbreviation for $\mathbf{A}^f\mathbf{A}^{f\mathrm{T}}/(m-1)$. The equivalence of (10) and (11) can be shown by using the Sherman–Morrison–Woodbury identity (e.g., Horn and Johnson 1985, p. 19). It is easy to see that the positive definite square root of a symmetric positive definite matrix is a symmetric matrix: such a matrix can be represented via its eigenvalue decomposition $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}}$, and $\mathbf{V}\mathbf{\Lambda}^{1/2}\mathbf{V}^{\mathrm{T}}$ is then its unique positive definite square root. Given the eigenvalue decomposition of the matrix in square brackets in (10),

$$\mathbf{I} + \frac{1}{m-1}(\mathbf{H}\mathbf{A}^f)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{A}^f = \mathbf{C}\mathbf{\Gamma}\mathbf{C}^{\mathrm{T}}, \tag{12}$$

the solution for $\mathbf{T}^s$ is

$$\mathbf{T}^s = \mathbf{C}\mathbf{\Gamma}^{-1/2}\mathbf{C}^{\mathrm{T}}. \tag{13}$$

Following Ott et al. (2004), we will refer to (13) as the "symmetric" solution [Wang et al. (2004) refer to it as the "spherical simplex" solution].

The symmetric solution (13) has been shown by Wang et al. (2004) to preserve the mean in the case when $\mathrm{rank}\{[(\mathbf{H}\mathbf{A}^f)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{A}^f]\} = m - 1$; this conclusion can be shown to be valid for the case of an arbitrary rank (see appendix A).

Despite the fact that any solution given by (9) results in the analyzed anomalies $\mathbf{A}^a$ such that the analyzed covariance $\mathbf{P}^a$ calculated by (4) matches the theoretical covariance (2), not all of these solutions are valid in the data assimilation context. Because the ensemble $\mathbf{X}$ carries *both* the state vector $\mathbf{x}$ (via ensemble mean) and the state error covariance $\mathbf{P}$ (via anomalies $\mathbf{A}$), it is necessary that the analyzed anomalies do not perturb the ensemble mean, or, in other words, remain zero centered:

$$\mathbf{A}^a\mathbf{1} = \mathbf{0}, \tag{14}$$

where $\mathbf{1}$ is a vector with each element equal to 1. Violation of the condition (14) is equivalent to an implicit alteration of the state vector, which, in turn, leads to an alteration of the error covariance matrix. More specifically, it can be shown that for a non-zero-centered ensemble the variance associated with each state vector element will, generally, be less than the variance calculated from the same ensemble without subtracting the ensemble mean. Suppose, for example, that the variance associated with some arbitrary vector $\mathbf{a}$ of length m is given by $\mathrm{var}(\mathbf{a}) = \mathbf{a}^{\mathrm{T}}\mathbf{a}/(m-1)$; $\mathbf{a}$ may be a row vector of the matrix $\mathbf{A}$ of ensemble anomalies, corresponding to some state vector element. If we then subtract the mean $\bar{\mathbf{a}}$ from each element of $\mathbf{a}$, than the actual variance represented by $\mathbf{a}$ becomes $\mathrm{var}(\mathbf{a} - \bar{\mathbf{a}}\mathbf{1}) = \mathrm{var}(\mathbf{a}) - m\bar{\mathbf{a}}^2/(m-1)$, which is, generally, smaller than the initially assumed value. The underestimation of the system error covariance is highly undesirable. Unlike the overestimation, it can lead to the divergence of the filter and therefore affects the robustness of the data assimilation system (Julier and Uhlmann 1997). Since applying an arbitrary rotation to the symmetric solution does not, generally, produce a zero-centered ensemble of analyzed anomalies, we will refer to (9) as non-mean-preserving general solution for the ETM.

Because by definition the forecast anomalies are zero centered, $\mathbf{A}^f\mathbf{1} = \mathbf{0}$, to preserve the ensemble mean it is sufficient for the ETM to retain (up to an arbitrary

multiple) the sum of elements of each row vector of the forecast ensemble $\mathbf{A}^f$, which may be written as

$$\mathbf{T1} = \alpha\mathbf{1}, \tag{15}$$

where α is a scalar constant. If the ensemble is of full rank, so that $\mathrm{rank}(\mathbf{A}) = m - 1$, then up to an arbitrary multiple $\mathbf{1}$ can be the only vector orthogonal to each row vector of $\mathbf{A}$; and the condition that the analyzed ensemble is zero centered, $\mathbf{A}^f\mathbf{T1} = \mathbf{0}$, can only be satisfied if $\mathbf{T1} = \alpha\mathbf{1}$. Consequently, for a full-rank ensemble condition (15) is not only a sufficient but also a necessary condition for an ETM $\mathbf{T}$ to retain zero centering of the ensemble anomalies; and to preserve the ensemble mean only those solutions need to be left out of the general non-mean-preserving solution (9) that satisfy this condition.

Because the symmetric solution $\mathbf{T}^s$ is a mean-preserving solution, it satisfies condition (15). Consequently, for a solution $\mathbf{T}$ in representation (9) to be mean preserving, it is sufficient for the orthonormal matrix $\mathbf{U}$ in (9) to satisfy (15); and for a full-rank ensemble (15) becomes a necessary and sufficient condition. Hence, for a full-rank ensemble the general mean-preserving solution for the ETM may be written as

$$\mathbf{T} = \mathbf{C}\boldsymbol{\Gamma}^{-1/2}\mathbf{C}^{\mathrm{T}}\mathbf{U}^p, \tag{16}$$

where $\mathbf{U}^p$ is an arbitrary orthonormal mean-preserving matrix,

$$\mathbf{U}^p\mathbf{U}^{p\mathrm{T}} = \mathbf{I}, \quad \mathbf{U}^p\mathbf{1} = \mathbf{1}. \tag{17}$$

Note that for preserving the mean it would be sufficient to require $\mathbf{U}^p\mathbf{1} = \mathrm{const} \times \mathbf{1}$; however, the constant here can only be equal to 1, because if $\mathbf{Ua} = \mathbf{b}$, where $\mathbf{U}$ is an orthonormal matrix and $\mathbf{a}$ and $\mathbf{b}$ are vectors, then $|\mathbf{a}| = |\mathbf{b}|$.

To obtain a general representation for an arbitrary mean-preserving orthonormal matrix, we note that the condition $\mathbf{U}^p\mathbf{1} = \mathbf{1}$ means that the linear transformation $\mathbf{U}^p$ preserves the unit vector $\mathbf{1}/\sqrt{m}$. Consequently, in an orthonormal basis $\mathbf{B} = [\mathbf{B}_1, \ldots, \mathbf{B}_m]$ such that the first basis vector coincides with this vector, $\mathbf{B}_1 = \mathbf{1}/\sqrt{m}$, this transformation will have the following form:

$$[\mathbf{U}^p]_{\mathbf{B}} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{U} \end{bmatrix}, \tag{18}$$

where $\mathbf{U}$ is an arbitrary orthonormal matrix of size $(m - 1) \times (m - 1)$. Given the matrix $[\mathbf{U}^p]_{\mathbf{B}}$ of the linear transformation $\mathbf{U}^p$ in the basis $\mathbf{B}$, it is possible to find its matrix in the original basis:

$$\mathbf{U}^p = \mathbf{B}[\mathbf{U}^p]_{\mathbf{B}}\mathbf{B}^{\mathrm{T}} \tag{19}$$

(e.g., Horn and Johnson 1985, p. 32). This representation can be used for generating a random mean-preserving matrix $\mathbf{U}^p$ of size $m \times m$ from a random orthonormal matrix $\mathbf{U}$ of size $(m - 1) \times (m - 1)$. The exact composition of the orthonormal matrix $\mathbf{B}$ is not important, as long as $\mathbf{B}_1 = \mathbf{1}/\sqrt{m}$.

Some authors (e.g., Anderson 2001; Whitaker and Hamill 2002) seek a solution for the ensemble transformation in the linear space of the system state rather than in the linear space of the ensemble by left multiplying the ensemble anomalies by the ETM:

$$\mathbf{A}^a = \tilde{\mathbf{T}}\mathbf{A}^f. \tag{20}$$

We use the tilde over $\mathbf{T}$ here to underline the difference between the ETMs in the left-multiplied and right-multiplied forms. Written in this form, the ensemble transformation automatically preserves the ensemble mean, which may be shown by right multiplying (20) by $\mathbf{1}$; but apart from the trivial case of assimilating a single observation, this representation is, generally, more computationally expensive because it involves calculation of a matrix of size $n \times n$, rather than $m \times m$. It may be shown (see appendix B) that the symmetric solution can be rearranged into the left-multiplied form (20), with the ETM given by

$$\tilde{\mathbf{T}}^s = (\mathbf{I} - \mathbf{KH})^{1/2}, \tag{21}$$

where $\mathbf{K}$ is the Kalman gain (3). The superscript s in $\tilde{\mathbf{T}}^s$ here underlines the equivalence of this solution with the symmetric solution in the right-multiplied form (7), given by either (10) or (11), although $\tilde{\mathbf{T}}^s$ itself is, generally, nonsymmetric (but positive definite).

In the singular evolutive extended Kalman filter (SEEK) and singular evolutive interpolated Kalman filter (SEIK; Pham et al. 1998; Pham 2001), the analyzed error covariance is calculated via updating of the eigenvalues of the covariance matrix with subsequent redrawing of the ensemble anomalies. Because the analyzed ensemble matches the theoretical expressions for the mean and covariance given by the Kalman filter, these filters essentially represent another flavor of the ESRF, so that the equivalent ensemble transformation can be written in form of (9). The particular form of the ensemble transformation in these filters depends on the implementation of the redrawing procedure, which must ensure that the redrawn ensemble remains zero centered (Hoteit et al. 2002; Nerger et al. 2005).

The original formulation of the ETKF (Bishop et al. 2001) used the simplest (one sided) solution:

$$\mathbf{T} = \mathbf{C}\boldsymbol{\Gamma}^{-1/2}. \tag{22}$$

Leeuwenburgh et al. (2005) showed that, for a scalar model, assimilation of a single observation using the one-sided solution results in all ensemble anomalies but

one equal to zero. This result can be easily generalized for the arbitrary dimension of the state vector: if a small number of observations are assimilated using the one-sided solution, the variance of the ensemble at each observation location will be provided by a constant set of p ensemble members, with the anomalies of all other $m - p$ members being equal to zero at these locations. This deficiency of the one-sided solution has been recognized by Evensen (2004) and Leeuwenburgh et al. (2005). They raised concern about the tendency of the one-sided solution to produce outliers that contain most of the variance in the ensemble noticed by Lawson and Hansen (2004). To counter this deficiency of the one-sided solution, Evensen (2004) and Leeuwenburgh et al. (2005) proposed an ensemble transformation with a random rotation in ensemble space. Such a rotation is carried out by right multiplying the one-sided ETM (22) by a random orthonormal matrix and has a form

$$\mathbf{T} = \mathbf{C}\boldsymbol{\Gamma}^{-1/2}\mathbf{U}, \tag{23}$$

where $\mathbf{U}$ is a random orthonormal matrix. This solution does not, generally, preserve the ensemble mean but does tend to improve the performance compared to the one-sided solution by spreading the variance between the ensemble members.

Below we present experiments with two small models in which we compare the performance of a number of solutions for the ETM in the ETKF as well as the performance of the traditional EnKF with perturbed observations. We demonstrate a significantly better performance of the mean-preserving solutions for the ETM and conclude that only such solutions should be used in practical data assimilation.

3. Experimental setup

We will compare the performance of a number of different solutions for the ETM in the ETKF using two small models: the linear advection model (LA model) of Evensen (2004) and the strongly nonlinear model of Lorenz and Emanuel (1998; L40 model). These two models expose different aspects of the performance of the filter. Specifically, the LA model is a stationary, linear model. It allows investigation of the scenario where the analysis gradually approaches the known stationary solution. In contrast, the L40 model is a strongly nonlinear model characterized by instability in regard to small perturbations. We investigate the performance of the ETKF with the symmetric solution for the ETM (10); the ETKF with the one-sided solution (22); the ETKF with the solution with random rotations (9); the ETKF with the solution with mean-preserving random rotations (16); and the traditional EnKF with perturbed observations.

a. LA model

The LA model, described below, is based on that of Evensen (2004). The dimension of the state vector $\mathbf{x}$ is 1000; the signal propagates (advects) in the positive direction by one element at each time step without changing its shape; and the model domain is periodic:

$$\mathbf{x}(t) = [\mathbf{x}_1(t), \ldots, \mathbf{x}_{1000}(t)], \quad t = 1, 2, \ldots;$$

$$\mathbf{x}_i(t+1) = \begin{cases} \mathbf{x}_{i-1}(t), & i = 2, \ldots, 1000; \\ \mathbf{x}_{1000}(t), & i = 1 \end{cases}$$

where $\mathbf{x}_i(t)$ is the ith component of the state vector at the tth time step.

To generate a state vector sample, a sum of a 25 sine curves with random amplitude and phase and a random offset is calculated:

$$\mathbf{s}_i = \sum_{k=0}^{25} a_k \sin\left(\frac{2\pi k}{1000} i + \varphi_k\right), \quad i = 1, \ldots, 1000,$$

where a_k and φ_k are random numbers uniformly distributed in the intervals (0, 1) and (0, 2π), correspondingly. This sample is then normalized to have a variance of 1:

$$\mathbf{x}(1) = \frac{\mathbf{s}}{[(\mathbf{s} - \bar{\mathbf{s}})^T(\mathbf{s} - \bar{\mathbf{s}})]^{1/2}},$$

where $\bar{\mathbf{s}}$ denotes the state average. To generate the initial ensemble, a specified number of samples are generated using this procedure. The ensemble mean field is subsequently subtracted from each member, and another random sample (climatology) is added to each member. The true field is defined as a sum of yet another random sample and the climatology.

Therefore, by construction all possible model state vectors belong to a subspace with dimension of 51 (referred to hereafter as the model subspace dimension) of the full model state vector space with dimension of 1000. Also by construction, the initial root-mean-square error (RMSE) of the ensemble mean is equal to 1. Note that because the ensemble mean has been subtracted from each member *after* the normalization, each ensemble anomaly has an initial variance that is slightly different from 1.

Four observations of the true field are conducted and assimilated into the model at every fifth time step, t = 1, 6, 11, . . . , at equidistant locations i = {125, 375, 625, 875}. Each observation is contaminated with random normally distributed uncorrelated noise with variance of 0.01.

b. L40 model

The L40 model (Lorenz and Emanuel 1998) is a strongly nonlinear model with a state vector dimension of 40. Lorenz and Emanuel (1998) argue that it roughly imitates the evolution of an unspecified scalar meteorological quantity (such as temperature or vorticity) along the latitude circle. This model has been used for testing ensemble-based assimilation methods in a number of earlier studies (Anderson 2001; Whitaker and Hamill 2002; Ott et al. 2004; Lawson and Hansen 2004). It contains 40 coupled ordinary differential equations in a domain with cyclic boundary condition:

$$\dot{\mathbf{y}}_i = (\mathbf{y}_{i+1} - \mathbf{y}_{i-2})\mathbf{y}_{i-1} - \mathbf{y}_i + 8, \quad i = 1, \ldots, 40;$$

$$\mathbf{y}_0 = \mathbf{y}_{40}, \quad \mathbf{y}_{-1} = \mathbf{y}_{39}, \quad \mathbf{y}_{41} = \mathbf{y}_1.$$

The consecutive model states are obtained by integrating these equations forward by intervals of $\Delta t = 0.05$, so that the model states $\mathbf{x}(t)$ relate to the solution of the above system as

$$\mathbf{x}(t) = \mathbf{y}(0.05t), \quad t = 1, 2, \ldots.$$

The model has an estimated fractal dimension of 27.1, the doubling time of the leading Lyapunov exponent of 0.42, the mean of 2.34, and standard deviation of 3.66 (Lorenz and Emanuel 1998).

Following Lorenz and Emanuel (1998), each model time step in our tests is conducted by a single step of the standard fourth-order Runge–Kutta integrator. The ensemble members are initialized by random sampling from a set of 10 000 model states obtained during one continuous integration at $t = 1000, 1001, \ldots, 11\,000$. The true field is initialized by one more randomly chosen state from this set. Following Whitaker and Hamill (2002) and Ott et al. (2004), at every time step we conduct 40 observations of the true field at each grid point, $i = 1, \ldots, 40$; each observation is contaminated with random normally distributed uncorrelated noise with a variance of 1.

Common for assimilation with nonlinear models, an ensemble inflation is applied to the ensemble at the end of each assimilation step by multiplying the ensemble anomalies by the inflation factor δ with typical values between 1.00 and 1.10: $\mathbf{A}^a \leftarrow \delta \cdot \mathbf{A}^a$.

4. Results

a. LA model

Figure 1 shows the RMSE versus the ensemble size for four different filters applied to the LA model. The filters involved include the symmetric ETKF, the one-sided ETKF, the ETKF with random rotations, and the EnKF. Because of the linearity of the model, for a given initial ensemble and forward observations all mean-preserving solutions yield the same covariance matrix (within the round-up error). Consequently, all mean-preserving solutions produce the same results and are represented by the symmetric ETKF. The RMSE values represent an average over 50 independent model runs; the RMSE of a particular run is calculated as a mean RMSE of analyses at time steps $t = 901, 906, \ldots, 1001$. The same set of 50 initial ensembles has been used for each filter presented in Fig. 1. The figure also contains the average over all 50 realization RMSE $\sigma_{\min}$ of the best possible fit in the least squares sense for a given ensemble, referred to as "best fit," and calculated as follows:

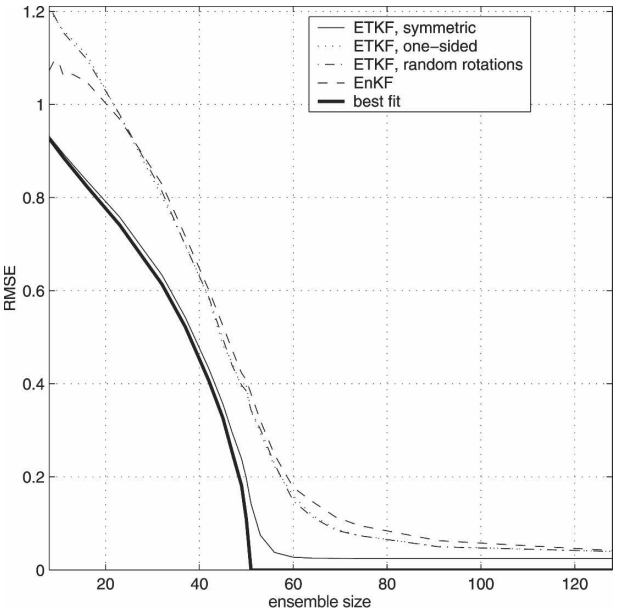

FIG. 1. RMSE of different flavors of ETKF and EnKF for the LA model, averaged for the time interval $t = [900, 1000]$, and over 50 realizations.

$$\sigma_{\min} = \|\mathbf{X}\mathbf{s} - \mathbf{x}^t\|, \quad \mathbf{s} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{x}^t,$$

where $\mathbf{X}$ is the ensemble matrix, and $\mathbf{x}^t$ is the true solution. No inflation has been applied in these experiments with the linear model.

Figure 1 shows a remarkable performance of the symmetric ETKF, with RMSEs that are very close to the best possible values. The other three filters produce very similar results, with both the one-sided ETKF and the ETKF with random rotations being very close to each other and only slightly better than the EnKF.

One interesting feature of the performance of the symmetric ETKF in Fig. 1 is the early "saturation" of

the RMSE versus the ensemble size, with almost no improvement after the ensemble size exceeds the dimension of the model subspace (equal to 51). This may be related to the orthogonalization of the ensemble anomalies over time that is achieved by the symmetric ETKF.

To demonstrate this phenomenon, we show the normalized singular value decomposition (SVD) spectra of the ensemble anomalies $\mathbf{A}^f$ after 1000 time steps, averaged over 50 runs, in Fig. 2. The first example (Fig. 2a) shows the normalized SVD spectra (i.e., singular values divided by the largest singular value) for the case when the ensemble size is smaller than the model subspace dimension ($m = 30 < 51$), while the second example (Fig. 2b) shows the SVD spectra for the case when the ensemble size is larger than the model subspace dimension ($m = 60 > 51$). In both cases the SVD spectrum of the ensemble anomalies for the symmetric ETKF is nearly flat. This means that all modes of the model are equally represented in the ensemble. In the case when the ensemble size is smaller than the model subspace dimension ($m = 30 < 51$), the flatness of the SVD spectrum also means that the ensemble anomalies are orthogonal to each other and represent the eigenvectors of the covariance matrix (which are defined in this case up to an arbitrary rotation in the system's subspace). In contrast, the normalized SVD spectra of the non-mean-preserving one-sided ETKF and the ETKF with random rotations are nearly identical and both show a tendency toward the loss of rank. This may be a result of the repetitive subtraction of the ensemble mean that has been noted by Wang et al. (2004). The spectra of the EnKF ensemble anomalies show neither a tendency toward loss of rank or toward orthogonalization.

b. L40 model

Figure 3 shows the average RMSE over a long run of the L40 model versus ensemble size and inflation factor for five different filters: the symmetric ETKF, the ETKF with mean-preserving random rotations, the one-sided ETKF, the ETKF with random rotations, and the EnKF. Each run uses the same true field, and for a given ensemble size each filter starts from the same ensemble. The RMSE values are averaged over time steps $t = 1000, \ldots, 3 \times 10^5$. The white cells correspond to experiments in which the filter did not converge, which are defined as the runs with RMSE greater than 1.

Figure 3 clearly shows a superior performance of the two mean-preserving filters: the symmetric ETKF and the ETKF with mean-preserving random rotations; however, in contrast with the case of the LA model, the specific characteristics of their performance are different. The ETKF with mean-preserving random rotations is able to achieve a marginally smaller mean RMSE than the symmetric ETKF; but at the same time it requires more inflation. This difference in performance characteristics of the two mean-preserving filters occurs because of the nonlinearity of the L40 model. Unlike the results from the LA model, the one-sided ETKF performs substantially worse than the ETKF with random rotations, which in turn performs close to the EnKF.

Note that for all filters involved, other than the symmetric ETKF, the best performance is often achieved with the minimal or close to minimal value of the inflation factor necessary for the convergence of the filter. This means that achieving the best performance with these filters in practice may require a significant tuning of the inflation factor to avoid divergence.

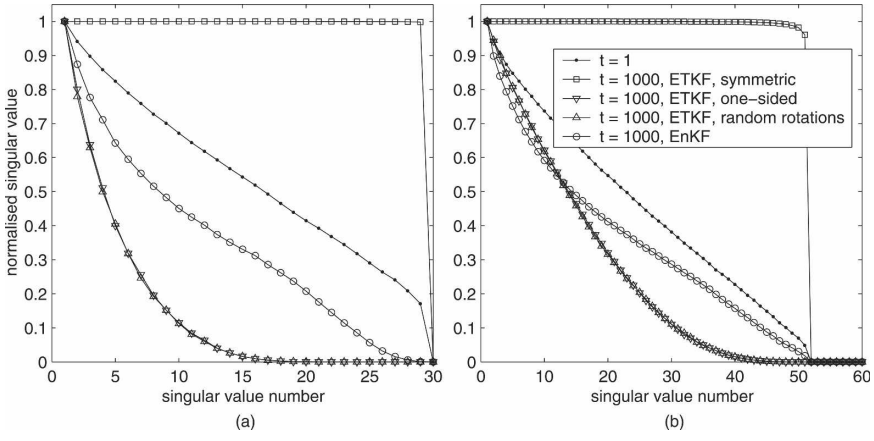

FIG. 2. Normalized singular-value spectra of the ensemble anomalies for the LA model, averaged over 50 realizations: (a) $m = 30$; (b) $m = 60$.

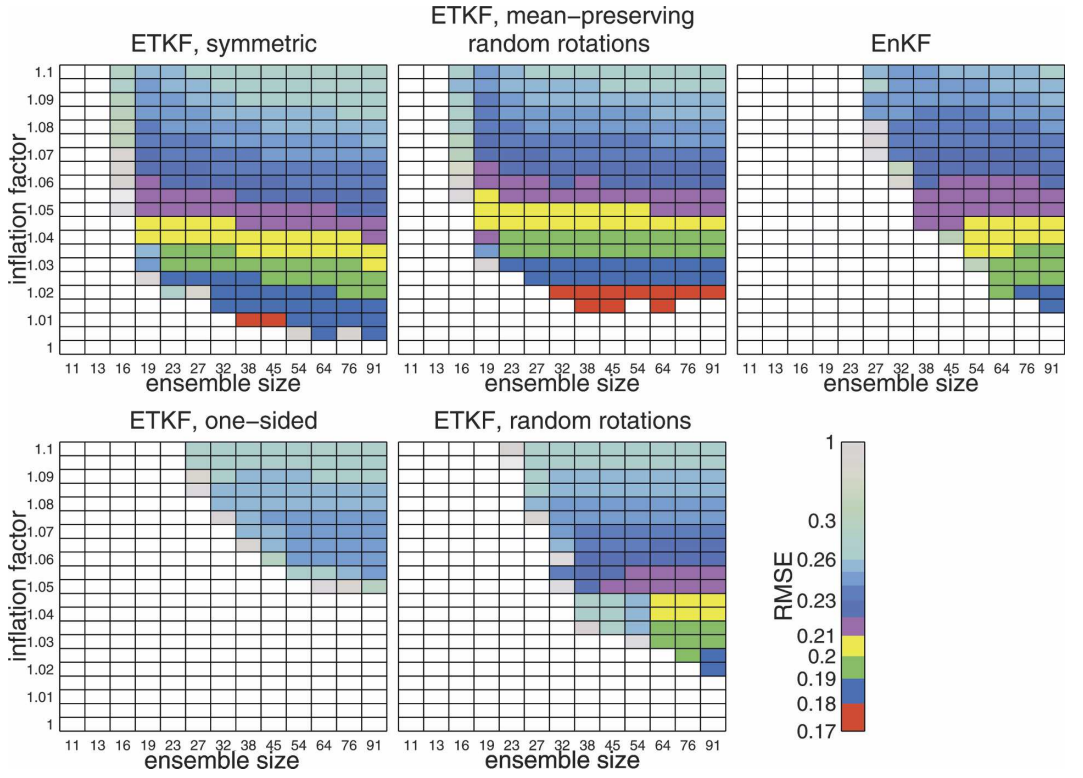

FIG. 3. RMSE of different flavors of ETKF and EnKF for the L40 model averaged over a long model run.

Further study of Fig. 3 shows that comparing the performance of different filters in absolute terms is not straightforward because different values of the inflation factor are required to achieve the best performance for a particular filter. For example, while the symmetric ETKF clearly outperforms the one-sided ETKF in our tests with the L40 model, testing them with a common inflation factor of 1.09 (suitable for the one-sided ETKF) would lead us to the conclusion that these filters perform nearly equally. Such a comparison is further complicated by the stochastic nature of the L40 model, which only makes it possible to characterize the filter performance in a statistical sense. In particular, we found it necessary to use longer runs (3×10^5 steps) compared to Whitaker and Hamill (2002) (5×10^4 steps) or Ott et al. (2004) (4×10^4 steps) to account for rare but long periods of divergence that happen for smaller inflation factors. This is particularly essential for assessing the difference in the performance between the symmetric ETKF and the ETKF with mean-preserving random rotations for small inflation factors; however, even runs of 3×10^5 steps do not exclude the possibility of filter divergence in longer runs. For example, the run of the ETKF with random mean-preserving rotations and inflation factor of 1.015 has diverged for ensemble size $m = 54$ but produced three best values of RMSE for $m = 38$, 45, and 64.

To compare the performance of different filters in absolute terms, we plot the best RMSE achieved for a given filter and ensemble size over all inflation factors involved (Fig. 4). The plot shows substantially better performance achieved by the symmetric ETKF and the ETKF with mean-preserving random rotations compared to other filters; this is particularly evident for small and intermediate ensemble sizes. Similar to the behavior of the symmetric ETKF in the case of the LA model, both these mean preserving solutions show an early "saturation" of the performance versus the ensemble size. Interestingly, the symmetric ETKF shows a slight degradation in the performance for ensemble sizes greater than 40.

To investigate a possible reason for this degradation in performance of the symmetric ETKF for larger ensembles, we investigate the statistics of the ensemble spread by using rank histograms. Figure 5 contains three ensemble rank histograms that summarize our findings. Each of these histograms shows statistics collected over 2×10^4 steps after a spinup of 1000 steps. At each assimilation, the number of ensemble members with the value of the first element smaller than that of

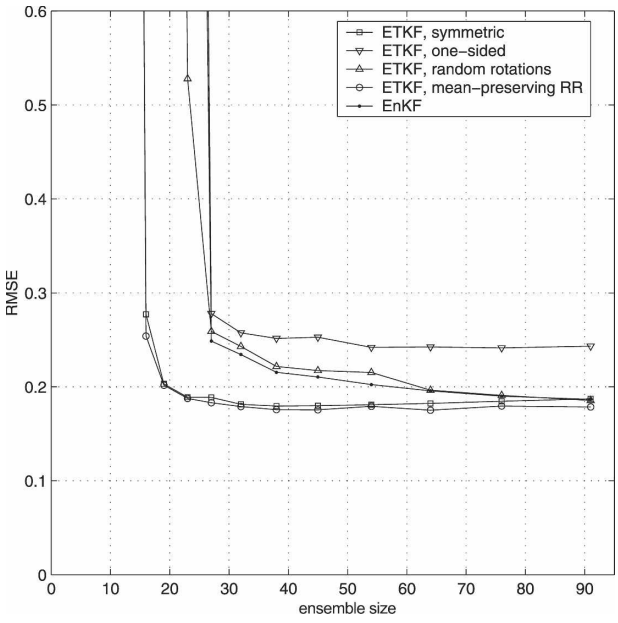

FIG. 4. The best RMSE from Fig. 3 for a given ensemble size over all inflation factors.

the true field is calculated, after which the counter for the corresponding bin is increased by one. Figure 5 presents histograms for the symmetric ETKF with the ensemble size of $m = 38$ and the ensemble inflation of $\delta = 1.01$, for the symmetric ETKF with $m = 91$ and $\delta = 1.01$, and for the ETKF with mean-preserving random rotations with $m = 91$ and $\delta = 1.02$. For each configuration, we used the value of the inflation that yields the best RMSE in the long model runs presented in Fig. 3. We see that the symmetric ETKF with the ensemble size of $m = 91$ has a distinct cup shaped ensemble rank histogram, similar to observed in Lawson and Hansen (2004); however, its histogram for a smaller ensemble of size $m = 38$ is much flatter. In contrast, the ensemble rank histogram for the ETKF with mean-preserving random rotations is cap shaped, which we attribute to the ensemble inflation: the cap-shaped histograms become more pronounced for bigger ensemble inflations and flatter for smaller ensemble inflations. This happens because the ensemble inflation makes the true field more likely to be positioned in the middle of the ensemble sorted by the observed element. The histogram for the ETKF with random rotations becomes flat when $\delta = 0$, but in that case the system becomes prone to divergence.

If follows from Fig. 5 that the degradation in the performance of the symmetric ETKF for larger ensembles is accompanied by the deterioration in the ensemble statistics. A cup-shaped rank histogram can be caused by a nonnormal distribution of the ensemble or by its bias (Hamill 2001). Because the symmetric ETKF and the ETKF with mean-preserving random rotations have very close RMSE, the observed difference in the ensemble rank histograms for these filters is most likely to be caused by different distributions of the ensemble members. Figure 6 shows the statistics of the distribution of the ensemble relative to the ensemble mean for these filters. The presented histograms were calculated in the same way as the ensemble rank histograms in Fig. 5, but the ensemble mean has been used instead of the true field for obtaining the rank value at each step.

The much broader distribution in Fig. 6 for the symmetric ETKF indicates that it is more likely to have asymmetric distribution of the ensemble members relative to the ensemble mean than the ETKF with mean-preserving random rotations. For example, for the symmetric ETKF the number of ensemble members ranked below the ensemble mean is smaller than 41 or greater than 50 in 61.9% of all cases, while for the ETKF with mean-preserving random rotations this fraction is only 10.0%.

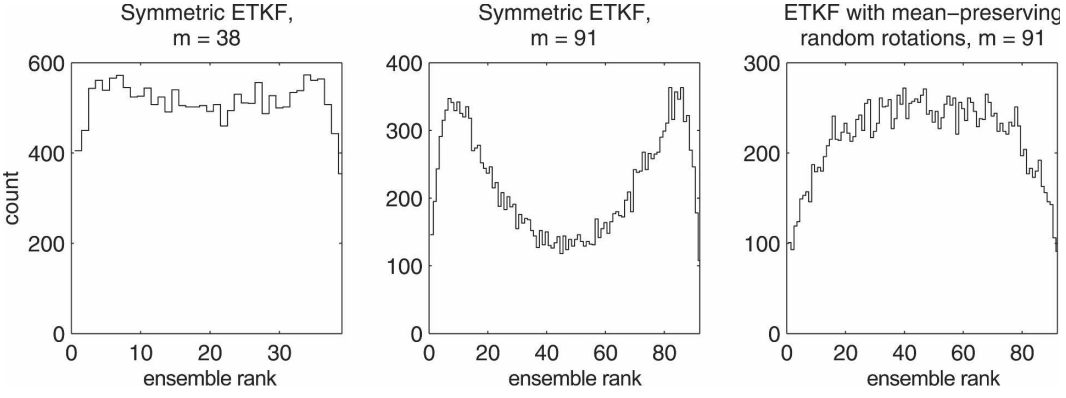

FIG. 5. Ensemble rank histograms for three different configurations.

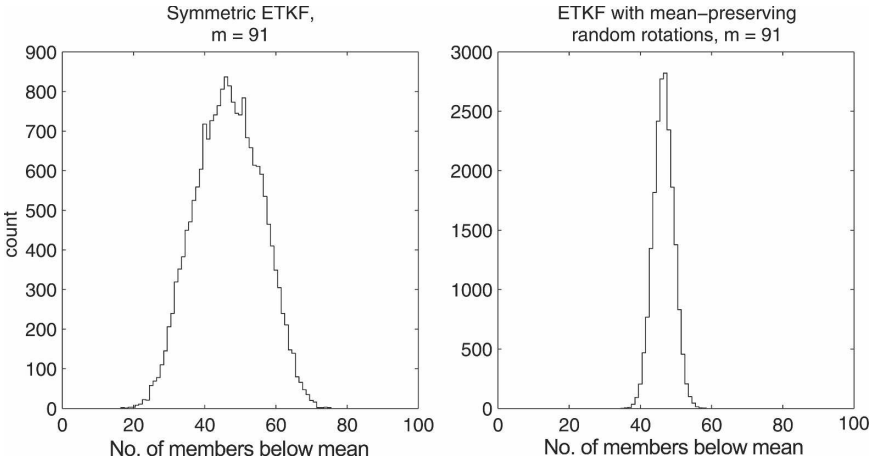

FIG. 6. Distribution of the ensemble relative to the ensemble mean for the symmetric ETKF and the ETKF with mean-preserving random rotations.

These statistics show that the symmetric ETKF is prone to the existence of the ensemble outliers, or ensemble members with particularly big deviations from the mean, and that this effect increases with the increase of the size of the ensemble. Figure 4 shows that this asymmetry of the ensemble results only in a very marginal deterioration in the performance of the symmetric ETKF with the increase of the ensemble size. It becomes more pronounced in systems with bigger nonlinearity or larger ensembles, such as used by Lawson and Hansen (2004).

5. Discussion

The results presented in the previous section demonstrate that the mean-preserving solutions for the ETM perform substantially better than either the one-sided solution or the solution with random rotations. We also note that the ESRFs with the mean-preserving solutions for the ETM achieved the best performance with the ensembles of size equal to, or slightly larger than, the model subspace dimension, unlike the traditional EnKF with perturbed observations, which always performed better with larger ensembles. This is evident from the saturation of the RMSE for the mean-preserving ETMs in Figs. 1 and 4, which occurs when the ensemble size exceeds the dimension of the model subspace (equal to 51 for the LA model and approximately 27 for the L40 model). These results indicate to us that the ensemble anomalies in ESRFs may be interpreted as a factorization of the system's error covariance rather than its Monte Carlo approximation.

The observed difference in the performance characteristics of the symmetric ETKF and the ETKF with mean-preserving random rotations for the L40 model (Figs. 3 and 4) can only occur in a nonlinear system and is therefore a consequence of the nonlinearity of the L40 model. We note that, in the case of the symmetric ETKF operating with the L40 model, an increase of the ensemble size over the model dimension produces a counterintuitive result of slightly degrading the filter performance (Fig. 4). We attribute this to the buildup of the ensemble outliers observed in Lawson and Hansen (2004). The analysis of the rank histograms presented in section 4 indicates that the effect of outliers only occurs in systems with large ensembles; therefore, it is not clear to us whether this effect ever manifests itself in large-scale applications like NWP. Importantly, using ETMs with mean-preserving random rotations prevents the buildup of outliers when such a possibility exists.

6. Summary and conclusions

Based on a series of data assimilation experiments with two small models, we demonstrated that mean-preserving solutions for the ETM in ESRFs yield a significantly better performance than that of non-mean-preserving solutions. For the LA model, the mean-preserving ESRFs produce the RMSEs that are close to the best possible RMSE for a given initial ensemble, while non-mean-preserving solutions and the EnKF with perturbed observations result in a much bigger errors. For the nonlinear L40 model the two mean-preserving solutions used in the experiments achieve convergence with a smaller ensemble, require less ensemble inflation, and result in smaller RMSEs, in particular for ensembles of modest size, compared to non-mean-preserving solutions. Concerning the non-mean-preserving solutions, we believe that the performance

of the one-sided solution and the solution with random rotations in our tests justify a conclusion that these solutions should not be used in practice for data assimilation, although they can still yield a similar or marginally better performance than the traditional EnKF for some systems. The traditional EnKF is found to require substantially larger ensembles than mean-preserving ESRFs to achieve convergence or to achieve a similar value of the RMSE. Given a sufficiently large ensemble, its performance becomes as good as that of any other scheme we tried in our experiments.

Considering the performance of particular mean-preserving solutions for the ETM, we note that, although these solutions produce different ensembles, they are otherwise equivalent for a linear system. For a nonlinear system, the symmetric solution preserves the identity of the ensemble members during the analysis in the best possible way, but it may also lead to the existence of ensemble outliers that cause deterioration in the system performance.

Acknowledgments. This research is funded by Australia's CSIRO through appropriation funding and by the U.S. Office of Naval Research Ocean Modeling Program through Grant N000140410345. We thank Stephen E. Cohn, Olwijn Leeuwenburgh, and the anonymous reviewer, as well as the anonymous reviewers of the manuscript MWR 1808, for useful comments and suggestions.

APPENDIX A

The Symmetric Solution Always Preserves the Ensemble Mean

We will show that, for any zero-centered ensemble $\mathbf{A}$, $\mathbf{A1} = 0$, the symmetric solution (10) for the ETM satisfies

$$\mathbf{T}^s\mathbf{1} = \mathbf{1}, \tag{A1}$$

which is a sufficient condition for preserving the ensemble mean. Using (13), this condition can be expanded as

$$\mathbf{C}\boldsymbol{\Gamma}^{-1/2}\mathbf{C}^{\mathrm{T}}\mathbf{1} = \mathbf{1},$$

where $\mathbf{C}$ and Γ are defined by the eigenvalue decomposition (12). Let us assume that $\mathbf{C}$ and Γ are organized so that the first r vectors in $\mathbf{C}$ span the range of $\mathbf{B}$, where

$$\mathbf{B} = \frac{1}{m-1}(\mathbf{HA})^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{HA}$$

is the matrix added to the identity matrix in (10):

$$\mathrm{range}(\mathbf{B}) = \langle \mathbf{C}_1, \ldots, \mathbf{C}_r \rangle, \quad \mathrm{null}(\mathbf{B}) = \langle \mathbf{C}_{r+1}, \ldots, \mathbf{C}_m \rangle.$$

Let us represent $\mathbf{C}$ as a sum of two components, corresponding to the range and null space of $\mathbf{B}$:

$$\mathbf{C} = \tilde{\mathbf{C}} + \tilde{\tilde{\mathbf{C}}}, \quad \tilde{\mathbf{C}} = \mathbf{C}\tilde{\mathbf{I}}, \quad \tilde{\tilde{\mathbf{C}}} = \mathbf{C}\tilde{\tilde{\mathbf{I}}}$$

where $\tilde{\mathbf{I}}$ is a diagonal matrix containing ones for the first r diagonal entries and zeros for the rest, and $\tilde{\tilde{\mathbf{I}}} = \mathbf{I} - \tilde{\mathbf{I}}$. We now notice that, because $\boldsymbol{\Gamma}^{\alpha}\tilde{\tilde{\mathbf{I}}} = \tilde{\tilde{\mathbf{I}}}$ for an arbitrary power α,

$$\mathbf{C}\boldsymbol{\Gamma}^{\alpha}\mathbf{C}^{\mathrm{T}} = \tilde{\mathbf{C}}\boldsymbol{\Gamma}^{\alpha}\tilde{\mathbf{C}}^{\mathrm{T}} + \tilde{\tilde{\mathbf{C}}}\tilde{\tilde{\mathbf{C}}}^{\mathrm{T}},$$

so that the condition (A1) becomes

$$\tilde{\mathbf{C}}\boldsymbol{\Gamma}^{-1/2}\tilde{\mathbf{C}}^{\mathrm{T}}\mathbf{1} + \tilde{\tilde{\mathbf{C}}}\tilde{\tilde{\mathbf{C}}}^{\mathrm{T}}\mathbf{1} = \mathbf{1}.$$

For a zero-centered ensemble we have $\mathbf{A1} = \mathbf{0}$, and therefore, from the definition of $\mathbf{B}$, $\mathbf{B1} = \mathbf{0}$. This means that

$$\mathbf{1} \subseteq \mathrm{null}(\mathbf{B}^{\mathrm{T}}) = \mathrm{null}(\mathbf{B}).$$

Consequently, from the definition of $\tilde{\mathbf{C}}$ and $\tilde{\tilde{\mathbf{C}}}$, first, $\tilde{\mathbf{C}}^{\mathrm{T}}\mathbf{1} = \mathbf{0}$, and second, because $\tilde{\tilde{\mathbf{C}}}^{\mathrm{T}}\mathbf{1}$ represents the vector of coefficients of expansion of $\mathbf{1}$ in the basis of columns of $\tilde{\tilde{\mathbf{C}}}$, $\tilde{\tilde{\mathbf{C}}}\tilde{\tilde{\mathbf{C}}}^{\mathrm{T}}\mathbf{1} = \mathbf{1}$.

Note that the above arguments can be applied to any solution of the form $\mathbf{T} = (\mathbf{I} + \mathbf{A}^{\mathrm{T}}\mathbf{DA})^{\alpha}$, where $\mathbf{D}$ is an arbitrary symmetric matrix, and α is an arbitrary power.

APPENDIX B

The Left-Multiplied Form of the Symmetric Solution

We will show that the ensemble transformation represented by the symmetric solution (11) in the right-multiplied form (7) can be rearranged to the left-multiplied form

$$\mathbf{A}^a = \tilde{\mathbf{T}}^s\mathbf{A}^f,$$

with the ETM given by

$$\tilde{\mathbf{T}}^s = (\mathbf{I} - \mathbf{KH})^{1/2};$$

in other words, we will show that

$$\mathbf{A}^f\left(\mathbf{I} - \frac{\mathbf{A}^{f\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{HA}^f}{m-1}\right)^{1/2} = (\mathbf{I} - \mathbf{KH})^{1/2}\mathbf{A}^f,$$

where $\mathbf{M} \equiv \mathbf{HP}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R}$, and $\mathbf{K}$ is the Kalman gain (3).

Because the left-hand side of this equation can be rearranged to the equivalent form (10), in which the expression in square brackets is positive definite, and because the inverse of a positive definite matrix is a positive definite matrix, the expression in brackets in the left-hand side is positive definite. Consequently, it

can be expanded into Taylor series about $\mathbf{I}$ (hereafter we omit the superscript for $\mathbf{A}^f$):

$$\mathbf{A}^a = \mathbf{A}\left(\mathbf{I} - \frac{\mathbf{A}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{H}\mathbf{A}}{m-1}\right)^{1/2} = \mathbf{A}\sum_{k=0}^{\infty} a_k\left(\frac{\mathbf{A}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{H}\mathbf{A}}{m-1}\right)^k,$$

where a_k are the coefficients of the expansion. Because

$$\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{N}\mathbf{A})^k = (\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{N})^k\mathbf{A},$$

where $\mathbf{N}$ is an arbitrary matrix, this expression can be further rearranged as follows:

$$\mathbf{A}^a = \left[\sum_{k=0}^{\infty} a_k\left(\frac{\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{H}}{m-1}\right)^k\right]\mathbf{A} = \left[\sum_{k=0}^{\infty} a_k(\mathbf{K}\mathbf{H})^k\right]\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})^{1/2}\mathbf{A}.$$

REFERENCES

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.,* **129,** 2884–2903.

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

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.,* **126,** 1719–1724.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99,** 10 143–10 162.

——, 2004: Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dyn.,* **54,** 539–560.

Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.,* **129,** 550–560.

Horn, R. A., and C. R. Johnson, 1985: *Matrix Analysis.* Cambridge University Press, 561 pp.

Hoteit, I., D.-T. Pham, and J. Blum, 2002: A simplified reduced order Kalman filtering and application to altimetric data assimilation in Tropical Pacific. *J. Mar. Syst.,* **36,** 101–127.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126,** 796–811.

Julier, S. J., and J. K. Uhlmann, 1997: New extension of the Kalman filter to nonlinear systems. *Signal Processing, Sensor Fusion, and Target Recognition VI,* I. Kadar, Ed., International Society for Optical Engineering (SPIE Proceedings, Vol. 3068), 182–193.

Lawson, W. G., and J. A. Hansen, 2004: Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Wea. Rev.,* **132,** 1966–1981.

Leeuwenburgh, O., 2005: Assimilation of along-track altimeter data in the tropical Pacific region of a global OGCM ensemble. *Quart. J. Roy. Meteor. Soc.,* **131,** 2455–2472.

——, G. Evensen, and L. Bertino, 2005: The impact of ensemble filter definition on the assimilation of temperature profiles in the tropical Pacific. *Quart. J. Roy. Meteor. Soc.,* **131,** 3291–3300.

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

Nerger, L., W. Hiller, and J. Schröter, 2005: A comparison of error subspace Kalman filters. *Tellus,* **57A,** 715–735.

Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus,* **56A,** 415–428.

Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.,* **129,** 1194–1207.

——, J. Verron, and M. C. Roubaud, 1998: A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.,* **16,** 323–340.

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

Torres, R., J. I. Allen, and F. G. Figueiras, 2006: Sequential data assimilation in an upwelling influenced estuary. *J. Mar. Syst.,* **60,** 317–329.

Wang, X., C. H. Bishop, and S. J. Julier, 2004: Which is better, an ensemble of positive–negative pairs or a centered spherical simplex ensemble? *Mon. Wea. Rev.,* **132,** 1590–1605.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.,* **130,** 1913–1924.