A Unification of Ensemble Square Root Kalman Filters

Lars Nerger*, Tijana Janjić, Jens Schröter,

and Wolfgang Hiller

Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

* *Corresponding author address:* Lars Nerger, Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany.

E-mail: lars.nerger@awi.de

ABSTRACT

In recent years, several ensemble-based Kalman filter algorithms have been developed that have been classified as ensemble square-root Kalman filters. Parallel to this development, the SEIK (Singular "Evolutive" Interpolated Kalman) filter has been introduced and applied in several studies. Some publications note that the SEIK filter is an ensemble Kalman filter or even an ensemble square-root Kalman filter. This study examines the relation of the SEIK filter to ensemble square-root filters in detail. It shows that the SEIK filter is indeed an ensemble-square root Kalman filter. Furthermore, a variant of the SEIK filter, the Error Subspace Transform Kalman Filter (ESTKF), is presented that results in identical ensemble transformations to those of the Ensemble Transform Kalman Filter (ETKF) while having a slightly lower computational cost. Numerical experiments are conducted to compare the performance of three filters (SEIK, ETKF, and ESTKF) using deterministic and random ensemble transformations. The results show better performance for the ETKF and ESTKF methods over the SEIK filter as long as this filter is not applied with a symmetric square root. The findings unify the separate developments that have been performed for the SEIK filter and the other ensemble square-root Kalman filters.

1. Introduction

The original Ensemble Kalman Filter (EnKF, Evensen 1994) has been developed with the aim to enable the application of sequential data assimilation algorithms based on the Kalman filter with large-scale numerical models. Burgers et al. (1998) and Houtekamer and Mitchell (1998) clarified that the EnKF requires an ensemble of perturbed observations for statistical

consistency. The EnKF represents the state estimate by the mean of an ensemble of model state realizations while the ensemble covariance matrix represents the corresponding error covariance matrix. The prediction of the error covariance matrix is computed by propagating each model state of the ensemble with the full, usually nonlinear, numerical model.

Alternative filter algorithms have been developed that perform the analysis without perturbed observations. These filters use an explicit transformation of the state ensemble. Among these developments are the Ensemble Transform Kalman filter (ETKF, Bishop et al. 2001), the Ensemble Adjustment Kalman filter (EAKF, Anderson 2001) and the Ensemble Square-root Kalman filter with sequential processing of observations (EnSRF, Whitaker and Hamill 2002). These filters also have been reviewed by Tippett et al. (2003) in a uniform way as ensemble square-root Kalman filters. Another ensemble square-root Kalman filter has been derived by Evensen (2004).

The ensemble-based SEIK (Singular "Evolutive" Interpolated Kalman) filter has been introduced by Pham et al. (1998) a few years before the introduction of the ensemble square-root Kalman filters. The behavior of SEIK filter for nonlinear models was examined by Pham (2001). Comparison studies between the SEIK filter and the EnKF (Brusdal et al. 2003; Nerger et al. 2005a) argue that the SEIK filter can be more efficient than the EnKF because a smaller ensemble could be used to achieve comparable estimation errors. In addition, the computations used in the SEIK filter are much less costly than those of the EnKF (Nerger et al. 2007).

Overall, the developments in the SEIK filter and the ensemble square-root Kalman filters have been independent. In publications considering ensemble square-root filters, the SEIK filter is only occasionally mentioned. For example, Sakov and Oke (2008) note that the

SEIK and SEEK filters "essentially represent another flavor" of the ensemble square-root filter. Similarly, publications using the SEIK filter, describe it as an efficient alternative to the EnKF (e.g. Triantafyllou et al. 2003; Nerger et al. 2005a). Thus, while there are indications that the SEIK filter is an ensemble square-root filter, there is yet no clear classification of the SEIK filter or an identification of the square-root used in this algorithm.

The aim of this work is to examine the relation of the SEIK filter to the ensemble-square-root Kalman filters in detail. For this task, the ETKF and the SEIK filter will be reviewed in section 2. In section 3 it is shown that the SEIK filter is an ensemble square root filter and its relation to the ETKF is discussed. A variant of the SEIK filter that results in identical ensemble transformations to those of the ETKF, which we term the "Error Subspace Transform Kalman Filter" (ESTKF), is derived in section 4. The computational cost of the filters as well as a possible reduction of the cost of the ETKF are discussed in section 5. Numerical experiments are performed in section 6 to compare the filter behavior for different variants of the ensemble transformation matrix.

2. Filter Algorithms: ETKF and SEIK

In this section, the mathematical formulations of the ETKF and the SEIK filter are reviewed and the square-root in the ETKF is identified in analogy to Tippett et al. (2003). Only the global analysis formulation is considered. A localization (see Nerger et al. 2006; Hunt et al. 2007) can be formulated in an identical way for both filters.

The ETKF and the SEIK filter are ensemble-based Kalman filters. The state of a physical system, like the ocean or atmosphere, is estimated at time t_k by the state vector $\mathbf{x}_k$ of size

n and the corresponding error covariance matrix $\mathbf{P}_k$. An ensemble of m vectors $\mathbf{x}^{(\alpha)}$, $\alpha = 1, \ldots, m$, of model state realizations represents these quantities. The state estimate is given by the ensemble mean

$$\overline{\mathbf{x}}_k := \frac{1}{m} \sum_{i=1}^{m} \mathbf{x}_k^{(i)} \ . \tag{1}$$

With the ensemble matrix

$$\mathbf{X}_k := \left[\mathbf{x}_k^{(1)}, \ldots, \mathbf{x}_k^{(m)} \right], \tag{2}$$

$\mathbf{P}_k$ is given as the ensemble covariance matrix

$$\mathbf{P}_k := \frac{1}{m-1} \mathbf{X}'_k \left(\mathbf{X}'_k \right)^T \tag{3}$$

where $\mathbf{X}'_k := \mathbf{X}_k - \overline{\mathbf{X}_k}$ with $\overline{\mathbf{X}_k} = [\overline{\mathbf{x}_k}, \ldots, \overline{\mathbf{x}_k}]$ is the matrix of ensemble perturbations.

A forecast is computed by integrating the state ensemble using the numerical model until observations become available. The observations are available in form of the vector $\mathbf{y}_k^o$ of size p. The model state is related to the observations by $\mathbf{y}_k^o = \mathbf{H}_k(\mathbf{x}_k^f) + \epsilon_k$ where $\mathbf{H}$ is the observation operator, which is assumed to be linear. The vector of observation errors, ϵ_k, is assumed to be a white Gaussian distributed random process with covariance matrix $\mathbf{R}$.

The analysis equations of the ETKF and the SEIK filter are discussed separately below. As all operations are performed at the same time t_k, the time index k is omitted.

a. Analysis step of the ETKF

The ETKF has been introduced by Bishop et al. (2001). For the review of the analysis step of the ETKF, we follow Yang et al. (2009) and Hunt et al. (2007).

The computations performed in the ETKF are based on a square root of the state covari-

ance matrix given by the ensemble perturbations $\mathbf{X}'$. The analysis state covariance matrix $\mathbf{P}^a$ can be written as a transformation of the forecast ensemble perturbations as

$$\mathbf{P}^a = \mathbf{X}'^f \mathbf{A} (\mathbf{X}'^f)^T. \tag{4}$$

Here, $\mathbf{A}$ is an $m \times m$ matrix defined by

$$\mathbf{A}^{-1} := \gamma^{-1}(m-1)\mathbf{I} + (\mathbf{H}\mathbf{X}'^f)^T \mathbf{R}^{-1} \mathbf{H}\mathbf{X}'^f. \tag{5}$$

$\mathbf{A}$ is frequently denoted as 'transform matrix'. The factor γ is used to inflate the forecast covariance matrix to stabilize the filter performance.

The state estimate is updated according to

$$\overline{\mathbf{x}^a} = \overline{\mathbf{x}^f} + \mathbf{X}'^f \overline{\mathbf{w}}^{ETKF} \tag{6}$$

with the weight vector

$$\overline{\mathbf{w}}^{ETKF} := \mathbf{A} \left(\mathbf{H}\mathbf{X}'^f \right)^T \mathbf{R}^{-1} \left(\mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}^f} \right). \tag{7}$$

The square root of the forecast state covariance matrix is given by the perturbation matrix $\mathbf{X}'^f$ up to the scaling by $(m-1)^{-1}$. To obtain the square root of the analysis state covariance matrix, $\mathbf{X}'^f$ is transformed as

$$\mathbf{X}'^a = \mathbf{X}'^f \ \mathbf{W}^{ETKF}. \tag{8}$$

The weight matrix $\mathbf{W}^{ETKF}$ is computed from the square-root $\mathbf{C}$ with $\mathbf{C}\mathbf{C}^T = \mathbf{A}$ as

$$\mathbf{W}^{ETKF} := \sqrt{m-1}\mathbf{C}\mathbf{\Lambda}. \tag{9}$$

Here, $\mathbf{\Lambda}$ is an arbitrary orthogonal matrix of size $m \times m$ or the identity. To preserve the ensemble mean, the vector $(1, \ldots, 1)^T$ has to be an eigenvector of $\mathbf{\Lambda}$.

When the ETKF was introduced by Bishop et al. (2001), the form of the square-root $\mathbf{C}$ was not further specified. Studies about the properties of the ensemble transformation in different square-root filters (e.g., Wang et al. 2004; Sakov and Oke 2008) have shown that a symmetric matrix $\mathbf{C}$ ensures that the ensemble mean is preserved during the ensemble transformation. The use of the symmetric square root

$$\mathbf{C}_{sym} := \mathbf{U}\mathbf{S}^{-1/2}\mathbf{U}^T \tag{10}$$

has been proposed also for the localized version of the ETKF (LETKF, Hunt et al. 2007). Eq. (10) can be obtained from the singular value decomposition (SVD) $\mathbf{USV} = \mathbf{A}^{-1}$. The use of matrix $\mathbf{C}_{sym}$ from Eq. (10) provides a minimum transformation of the ensemble because the distance of the square-root from the identity matrix is minimized in the Frobenius norm (see Yang et al. 2009).

For efficiency, the analysis update of the state estimate (Eq. 6) and the ensemble transformation (Eq. 8) can be combined into a single transformation of $\mathbf{X}'^f$ as

$$\mathbf{X}^a = \overline{\mathbf{X}^f} + \mathbf{X}'^f \left(\overline{\mathbf{W}}^{ETKF} + \mathbf{W}^{ETKF} \right). \tag{11}$$

with $\overline{\mathbf{W}}^{ETKF} = \left[\overline{\mathbf{w}}^{ETKF}, \ldots, \overline{\mathbf{w}}^{ETKF} \right]$. This formulation leads directly to the analysis ensemble, without explicitly updating the state estimate by Eq. (6).

b. Analysis step of the SEIK filter

The SEIK filter has been introduced by Pham et al. (1998) and was described in more detail by Pham (2001). This review follows Nerger et al. (2006). The original separation of the analysis step into the state update ("analysis") and ensemble transformation ("re-

sampling") is followed here. The SEIK filter is then explicitly re-formulated as an ensemble square-root filter analogously to the ETKF in section 3. Quantities that are similar but not identical to those of the ETKF are marked using a tilde. It is assumed that the forecast ensemble is identical to that used in the ETKF.

Analysis: The computations of the analysis step update the state estimate and implicitly update the state covariance matrix from the forecast to the analysis matrix.

In the SEIK filter, the forecast covariance matrix $\mathbf{P}^f$ is treated in terms of the forecast state ensemble $\mathbf{X}^f$ by

$$\mathbf{P}^f = \mathbf{L}\mathbf{G}\mathbf{L}^T \tag{12}$$

with

$$\mathbf{L} := \mathbf{X}^f \, \tilde{\mathbf{T}}, \tag{13}$$

$$\mathbf{G} := (m-1)^{-1} \left(\tilde{\mathbf{T}}^T \tilde{\mathbf{T}} \right)^{-1}. \tag{14}$$

Here, $\tilde{\mathbf{T}}$ is an $m \times (m-1)$ matrix with full rank and zero column sums. Previous studies have always defined matrix $\tilde{\mathbf{T}}$ as

$$\tilde{\mathbf{T}} := \begin{pmatrix} \mathbf{I}_{(m-1)\times(m-1)} \\ \\ \mathbf{0}_{1\times(m-1)} \end{pmatrix} - \frac{1}{m} \left(\mathbf{1}_{m\times(m-1)} \right) \tag{15}$$

where $\mathbf{0}$ represents the matrix whose elements are equal to zero and $\mathbf{I}$ is the identity. The elements of the matrix $\mathbf{1}$ are equal to one. Matrix $\tilde{\mathbf{T}}$ implicitly subtracts the ensemble mean when the matrix $\mathbf{L}$ is computed. In addition, $\tilde{\mathbf{T}}$ removes the last column of $\mathbf{X}'^f$, thus $\mathbf{L}$ is an $n \times (m-1)$ matrix that holds the first $m-1$ ensemble perturbations.

The analysis update of the state estimate is given as a combination of the columns of the

matrix $\mathbf{L}$ by

$$\overline{\tilde{\mathbf{x}^a}} = \overline{\mathbf{x}^f} + \mathbf{L}\overline{\mathbf{w}}^{SEIK}. \tag{16}$$

Here, the vector $\overline{\mathbf{w}}^{SEIK}$ of size $m-1$ is given by

$$\overline{\mathbf{w}}^{SEIK} := \tilde{\mathbf{A}}\,(\mathbf{HL})^T\,\mathbf{R}^{-1}\left(\mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}^f}\right) \tag{17}$$

and the transform matrix $\tilde{\mathbf{A}}$ of size $(m-1)\times(m-1)$ is defined by

$$\tilde{\mathbf{A}}^{-1} := \tilde{\rho}\mathbf{G}^{-1} + (\mathbf{HL})^T\mathbf{R}^{-1}\mathbf{HL}. \tag{18}$$

In the SEIK filter, $\tilde{\rho}$ with $0 < \tilde{\rho} \leq 1$ is referred to as the "forgetting factor". It is the inverse of the inflation factor γ used in Eq. (5) of the ETKF. The analysis covariance matrix is given in factorized form by

$$\tilde{\mathbf{P}}^a = \mathbf{L}\tilde{\mathbf{A}}\mathbf{L}^T \tag{19}$$

but does not need to be explicitly computed.

For efficiency, the term $\mathbf{HL}$ is typically computed as $(\mathbf{HX}^f)\tilde{\mathbf{T}}$. Thus, $\tilde{\mathbf{T}}$ operates on the $p \times m$ matrix $\mathbf{HX}^f$, while $\mathbf{H}$ operates on each ensemble state.

Resampling: After the analysis step, the "resampling" of the ensemble is performed. Here, the forecast ensemble is transformed such that it represents $\tilde{\mathbf{x}}^a$ and $\tilde{\mathbf{P}}^a$. The transformation is performed according to

$$\tilde{\mathbf{X}}^a = \overline{\tilde{\mathbf{X}^a}} + \sqrt{m-1}\mathbf{L}\tilde{\mathbf{C}}\mathbf{\Omega}^T. \tag{20}$$

In previous studies, the SEIK filter was always described to use a Cholesky decomposition of the matrix $\tilde{\mathbf{A}}^{-1}$ to obtain $(\tilde{\mathbf{C}}^{-1})^T\tilde{\mathbf{C}}^{-1} = \tilde{\mathbf{A}}^{-1}$. However, other forms of the square-root, like the symmetric square root used in the ETKF, could be chosen. Section 6 will test the

influence of the chosen square root on the performance of the filter. The matrix $\mathbf{\Omega}$ is an $m \times (m-1)$ matrix whose columns are orthonormal and orthogonal to the vector $(1, \ldots, 1)^T$. Traditionally, $\mathbf{\Omega}$ is described to be a random matrix with these properties. However, using a deterministic $\mathbf{\Omega}$ is also valid. The procedure to generate a random $\mathbf{\Omega}$ (Pham 2001; Hoteit 2001) and a procedure for generating a deterministic variant are provided in the Appendix.

For efficiency, the matrix $\mathbf{L}$ can be replaced by $\mathbf{X}^f\tilde{\mathbf{T}}$ (Eq. 13). Then, the matrix $\tilde{\mathbf{T}}$ can be applied from the left to smaller matrices like the weight vector $\overline{\mathbf{w}}^{SEIK}$ or the matrix $\tilde{\mathbf{C}}$.

The original formulation of the SEIK filter used the normalization m^{-1} for the matrix $\mathbf{P}^f$ instead of using the sample covariance matrix that is normalized by $(m-1)^{-1}$. For consistency with other ensemble-based Kalman filters, Nerger and Gregg (2007) introduced the use of the sample covariance matrix in SEIK, which is also used here. In the SEIK filter, the ensemble is generated to be consistent with the normalization of $\mathbf{P}^f$. Hence, the normalization acts only as a scaling factor that influences the equations (3) and (20) as well as the definition of $\mathbf{G}$ in Eq. (14).

3. SEIK as an ensemble square-root filter

To identify the SEIK filter as an ensemble square-root filter, the analysis and resampling steps of SEIK are combined as a transformation of the square root of $\mathbf{P}^f$. Equation (20) can be written as

$$\tilde{\mathbf{X}}^a = \overline{\tilde{\mathbf{X}}^a} + \mathbf{L}\mathbf{W}^{SEIK} \tag{21}$$

with

$$\mathbf{W}^{SEIK} := \sqrt{m-1}\tilde{\mathbf{C}}\mathbf{\Omega}^T. \tag{22}$$

In addition, the state analysis update (16) can be combined with the ensemble transformation (21) to

$$\tilde{\mathbf{X}}^a = \overline{\mathbf{X}^f} + \mathbf{L}\left(\overline{\mathbf{W}}^{SEIK} + \mathbf{W}^{SEIK}\right), \tag{23}$$

with $\overline{\mathbf{W}}^{SEIK} = \left[\overline{\mathbf{w}}^{SEIK}, \ldots, \overline{\mathbf{w}}^{SEIK}\right]$.

Equation (23) performs a transformation of the matrix $\mathbf{L}$ analogous to the ensemble transformation of the ETKF (Eq. 11). Matrix $\mathbf{L}$ is the square root of the covariance matrix $\mathbf{P}^f$ used in the SEIK filter. With this, the SEIK filter is clearly an ensemble square-root filter.

It is particular for the SEIK filter that the matrix $\mathbf{L}$ has only $m-1$ columns, while other filters use a square-root with m columns. Using $m-1$ columns is possible because the rank of $\mathbf{P}^f$ is at most $m-1$. The SEIK filter utilizes this property by accounting for the fact that the sum of each row of the perturbation matrix $\mathbf{X}'^f$ is zero. Thus, while the columns of $\mathbf{X}'^f$ are linearly dependent, the columns of $\mathbf{L}$ are linearly independent if the rank of $\mathbf{P}^f$ is $m-1$. In this case, they build a basis of the error subspace estimated by the ensemble of model states (for a detailed discussion of the error subspace, see Nerger et al. (2005a)). In contrast, $\mathbf{X}'$ can be regarded as a transformation from its m-dimensional column space to the error subspace of dimension $m-1$ (see Hunt et al. 2007).

While the equations of the SEIK filter are very similar to those of the ETKF this does not automatically imply that their state and error estimates are identical, in particular because the analyses use matrices of different size. However, if the same forecast ensembles are used

in the ETKF and the SEIK filter, the analysis state $\mathbf{x}^a$ and the analysis state covariance matrix $\mathbf{P}^a$ will be identical. This identity is due to the fact that the analysis formulations of both methods refer to the same error subspace to compute the optimal combination of ensemble perturbations. A basis of this space is given by $\mathbf{L}$. It is used directly by the SEIK filter. In contrast, the ETKF utilizes the ensemble representation of the error subspace given by $\mathbf{X}'^f$. Nonetheless, the matrices $\mathbf{A}$ (Eq. 5) and $\tilde{\mathbf{A}}$ (Eq. 18) both describe the same quantity - an error covariance matrix - in the same space represented by either $\mathbf{X}'^f$ or $\mathbf{L}$. Therefore, the optimization computed in the analysis steps results in the same state and error estimates.

While the identity of $\mathbf{x}^a$ and $\mathbf{P}^a$ for both filters can be established by the argumentation above, the ensembles that represent these quantities are only unique up to a unitary matrix $\mathbf{B}$, i.e. $\mathbf{X}'^a = \tilde{\mathbf{X}}'^a\mathbf{B}$ (see, e.g. Livings et al. 2008). For example, this is the case when random rotations are used to generate $\mathbf{\Omega}$ or $\mathbf{\Lambda}$. However, for deterministic transformations and in the use of the symmetric square root of $\tilde{\mathbf{A}}$, the experiments discussed in section 6 indicate that the differences between the transformation matrices of the SEIK filter and the ETKF are very small with differences in the matrix entries below 2%.

4. Identical transformations in SEIK and ETKF

The ensemble transformation in the square-root formulation of SEIK, which was discussed in section 3, generally exhibits very small deviations from the transformation performed by the ETKF. As the transformation in the ETKF has been described to be the minimum transformation, it should be desirable to obtain the same transformation with the SEIK filter. This goal is achieved by a modification of the SEIK filter that is described in this

section.

The modification of the SEIK filter is motivated by the properties of the matrix $\mathbf{\Omega}$. In general, $\mathbf{\Omega}$ is an $m \times (m-1)$ matrix that re-generates m ensemble perturbations in combination with an ensemble transformation matrix of size $(m-1) \times (m-1)$. For a deterministic ensemble transformation, a deterministic form $\hat{\mathbf{\Omega}}$ can be used whose elements are defined by:

$$\hat{\mathbf{\Omega}}_{i,j} = \begin{cases} 1 - \frac{1}{m}\frac{1}{\frac{1}{\sqrt{m}}+1} & \text{for } i = j, i < m \\ -\frac{1}{m}\frac{1}{\frac{1}{\sqrt{m}}+1} & \text{for } i \neq j, i < m \\ -\frac{1}{\sqrt{m}} & \text{for } i = m \end{cases} \tag{24}$$

Geometrically, $\hat{\mathbf{\Omega}}$ is the Householder matrix associated with the vector $m^{-1/2}(1,\ldots,1)^T$ (see Appendix). Thus, $\hat{\mathbf{\Omega}}$ projects vectors in the ensemble space spanned by $\mathbf{X}^f$ onto the error subspace spanned by $\mathbf{L}$. Like $\tilde{\mathbf{T}}$, $\hat{\mathbf{\Omega}}$ has a full rank and zero column sums. In addition, the columns of $\hat{\mathbf{\Omega}}$ are orthonormal, which is not the case for $\tilde{\mathbf{T}}$. Using $\hat{\mathbf{\Omega}}$, one can replace Eqns. (12) – (14) by

$$\mathbf{P}^f = \mathbf{L}_\Omega \mathbf{G}_\Omega \mathbf{L}_\Omega^T \tag{25}$$

and

$$\mathbf{L}_\Omega := \mathbf{X}^f \hat{\mathbf{\Omega}}, \tag{26}$$

$$\mathbf{G}_\Omega := (m-1)^{-1}\left(\hat{\mathbf{\Omega}}^T\hat{\mathbf{\Omega}}\right)^{-1} = (m-1)^{-1}\mathbf{I}_{(m-1)\times(m-1)}\ . \tag{27}$$

Now, matrix $\tilde{\mathbf{A}}^{-1}$ from Eq. (18) is computed as:

$$\tilde{\mathbf{A}}_\Omega^{-1} := \tilde{\rho}(m-1)\mathbf{I} + (\mathbf{H}\mathbf{L}_\Omega)^T\mathbf{R}^{-1}\mathbf{H}\mathbf{L}_\Omega. \tag{28}$$

Finally, the ensemble transformation (Eq. 20) becomes

$$\tilde{\mathbf{X}}^a = \overline{\tilde{\mathbf{X}}^a} + \sqrt{m-1}\mathbf{X}^f\hat{\mathbf{\Omega}}\tilde{\mathbf{C}}_\Omega\hat{\mathbf{\Omega}}^T \tag{29}$$

where $\tilde{\mathbf{C}}_\Omega$ is the square-root of $\tilde{\mathbf{A}}_\Omega$. $\hat{\mathbf{\Omega}}\tilde{\mathbf{C}}_\Omega\hat{\mathbf{\Omega}}^T$ is the projection of $\tilde{\mathbf{C}}_\Omega$ from the error space onto the ensemble space. If the symmetric square root is used to compute $\tilde{\mathbf{C}}_\Omega$, the projected transformation matrix is identical to the matrix $\mathbf{C}$ used in the ETKF. In case of random ensemble transformations, only the rightmost $\hat{\mathbf{\Omega}}$ in Eq. (29) is replaced by the random matrix $\mathbf{\Omega}$, while $\hat{\mathbf{\Omega}}$ is used at all other places.

This reformulation of the SEIK filter is consistent with its original motivation to compute the ensemble transformation matrix in the error space and to project the required matrices onto this space and finally back onto the ensemble space. The choice of $\tilde{\mathbf{T}}$ is arbitrary as long as its column sums are zero and the matrix is of full rank. However, only the application of $\hat{\mathbf{\Omega}}$ results in consistent projections, because it is symmetrically applied in the computation of $\mathbf{A}$ as well as in the ensemble transformation (Eq. 29). Because the ensemble transformation is performed in the error subspace, the new filter variant is referred to as Error Subspace Transform Kalman filter (ESTKF). The main difference between the SEIK filter and the ESTKF is that the application of $\tilde{\mathbf{T}}$ in Eq. (13) subtracts the ensemble mean and drops the last ensemble member. The resulting matrix $\mathbf{L}$ actually depends on the order of the ensemble members in the ensemble matrix $\mathbf{X}^f$, which is arbitrary. In contrast, matrix $\mathbf{L}_\Omega$ defined by Eq. (26) will be independent of the order of the ensemble members. This is evident from the action of $\hat{\mathbf{\Omega}}$ when computing $\mathbf{L}_\Omega$: $\hat{\mathbf{\Omega}}$ not only subtracts the ensemble mean, but also subtracts the value of the last column of $\mathbf{X}^f$ divided by $\sqrt{m}$ from each column. The columns of $\hat{\mathbf{\Omega}}$ are then normalized by an additional division by $\sqrt{m}^{-1}+1$. These operations ensure that the value of the last column of $\mathbf{X}^f$ is implicitly contained in matrix $\mathbf{L}_\Omega$.

The use of $\hat{\mathbf{\Omega}}$ instead of $\tilde{\mathbf{T}}$ does not change the computational cost of the SEIK filter. The matrix $\hat{\mathbf{\Omega}}$ needs also to be initialized in the previous formulation of the SEIK filter. In

addition, the multiplication of a matrix by $\hat{\mathbf{\Omega}}$ has the same cost as the multiplication by $\tilde{\mathbf{T}}$.

5. Comparison of the computational costs and algorithmic enhancement of the ETKF

The computational cost of the SEIK filter is very similar to that of the ETKF. The leading costs of both filters are summarized in Table 1. The leading computational cost of both filter algorithms scales in the same way. However, the cost of the SEIK filter is slightly lower because of the use of matrix $\mathbf{L}$ with $m-1$ columns instead of $\mathbf{X}'^f$ with m columns.

One second-order term that does not appear explicitly in Table 1 is the computation of $\mathbf{X}'^f$ in the ETKF with a cost of $O(nm)$. The SEIK filter applies the matrix $\tilde{\mathbf{T}}$ to $\mathbf{HX}^f$ and to $\mathbf{W}^{SEIK}$ (Eq. 21). In the ESTKF, the matrix $\hat{\mathbf{\Omega}}$ is applied analogously. These operations have a cost of $O(p(m-1)+m(m-1)^2)$. In the typical situation, where the state dimension n is much larger than the observation dimension p and the ensemble size m is smaller than p, this alternative will be computationally less costly.

The ETKF can be modified to use an analog to matrix $\tilde{\mathbf{T}}$. The computation of the perturbation matrix can be formulated as

$$\mathbf{X}' = \mathbf{X}\overline{\mathbf{T}} \tag{30}$$

where the $m \times m$ matrix $\overline{\mathbf{T}}$ is defined by

$$\overline{\mathbf{T}} = \mathbf{I}_{m\times m} - \frac{1}{m}\mathbf{1}_{m\times m}. \tag{31}$$

Now, the equations of the ETKF that involve $\mathbf{X}^{'}$ can be reformulated. Eq. (5) becomes

$$\mathbf{A}^{-1} = \gamma^{-1}(m-1)\mathbf{I} + \left((\mathbf{H}\mathbf{X}^f)\overline{\mathbf{T}}\right)^T \mathbf{R}^{-1} \left((\mathbf{H}\mathbf{X}^f)\overline{\mathbf{T}}\right)^T \tag{32}$$

and Eq. (7) is written as

$$\overline{\mathbf{w}}^{ETKF} = \mathbf{A}\left((\mathbf{H}\mathbf{X}^f)\overline{\mathbf{T}}\right)^T \mathbf{R}^{-1}\left(\mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}^f}\right). \tag{33}$$

Further, the transformation equation (11) becomes

$$\mathbf{X}^a = \overline{\mathbf{X}^f} + \mathbf{X}^f\overline{\mathbf{T}}\left(\overline{\mathbf{W}}^{ETKF} + \mathbf{W}^{ETKF}\right). \tag{34}$$

As in the SEIK filter, this formulation avoids the explicit computation and storage of the ensemble perturbation matrix $\mathbf{X}'^f$. Instead, the matrix $\overline{\mathbf{T}}$ is applied to $\mathbf{H}\mathbf{X}^f$ of size $p \times m$ and to the sum of the weight matrices in Eq. (34) of size $m \times m$. This changes the computational cost to $O(pm + m^3)$ instead of $O(nm)$ for the direct computation of $\mathbf{X}'^f$. This formulation can also be applied with domain localization, but here $(\mathbf{H}\mathbf{X}^f)\overline{\mathbf{T}}$ should be computed globally, before performing the local analyses.

6. Numerical experiments

a. Experimental setup

In this section, the behavior of the ETKF will be compared with the explicit square-root formulation of the SEIK filter using the symmetric square-root introduced in section 3 (referred to as SEIK-sqrt) and with the ESTKF. In addition, the original SEIK filter with a square-root based on Cholesky decomposition from section 2b is applied (referred to as SEIK-orig). To compare the filters in the standard configuration of the ETKF, experiments with

deterministic ensemble transformations are conducted. Experiments including a random rotation are then performed to compare the filters in the standard configuration of the SEIK filter.

The algorithms are applied in identical twin experiments using the model by Lorenz (1996), denoted below as L96, that has been further discussed by Lorenz and Emanuel (1998). The L96 model is a simple nonlinear model that has been used in several studies to examine the behavior of different ensemble-based Kalman filters (e.g. Anderson 2001; Whitaker and Hamill 2002; Ott et al. 2004; Sakov and Oke 2008). Here, the same configuration as used by Janjić et al. (2011) is applied. The model state dimension is set to 40. It is small enough to allow for the successful application of the filters without localization for reasonably small ensemble sizes (see e.g. Sakov and Oke 2008). In our experiments, the localization mainly allowed for the use of smaller ensemble sizes compared to the global analysis, while the relative behavior of the filters was the same as without localization. Thus, for simplicity, only results for global filters are discussed below. The model as well as the filter algorithms are part of the release of the Parallel Data Assimilation Framework (PDAF, Nerger et al. 2005b, available online at http://pdaf.awi.de).

For the twin experiments, a trajectory over 60000 time steps is computed from the initial state of constant value of 8.0, but with $x_{20} = 8.008$ (see Lorenz and Emanuel 1998). This trajectory represents the "truth" for the data assimilation experiments. Observations of the full state are assimilated, which are generated by adding uncorrelated random normal noise of unit variance to the true trajectory. The observations are assimilated at each time step with an offset of 1000 time steps to omit the spin-up period of the model.

The initial ensemble for all experiments is generated by second-order exact sampling from

the variability of the true trajectory (see Pham 2001). Identical initial ensembles are used for all filter variants.

All experiments are performed over 50000 time steps. The ensemble size as well as the forgetting factor are varied in the experiments. For the ETKF, the covariance inflation is also expressed in terms of the forgetting factor (i.e. $\gamma = \rho^{-1}$ is used in Eq. (5)). Following the motivation of the SEIK filter as a low-rank filter, the ensembles used here are of a size that is at most equal to the state dimension.

Ten sets of experiments with different random numbers for the initial ensemble generation are performed for each combination of ensemble size and forgetting factor to assess the dependence of the results on the initial ensemble. The performance of the filters is assessed using the root mean square (RMS) error averaged over the 50000 time steps of each experiment. The RMS errors are then averaged over each set of ten experiments with different random numbers for the ensemble generation. We refer to this mean error as MRMSE. Note that the full length of the true trajectory is only used to generate the initial ensemble. For the computation of the RMS errors, only the time steps 1000 to 51000 of the true trajectory are used.

b. Results with deterministic ensemble transformations

First, the performance of the filters is studied when deterministic ensemble transformations are used. This is the common configuration for the ETKF. In this case, the rotation matrix $\mathbf{\Lambda}$ in Eq. (9) of the ETKF is the identity. In the SEIK-orig, SEIK-sqrt and ESTKF formulations, the deterministic matrix $\hat{\mathbf{\Omega}}$ defined by Eq. (24) is used. For the SEIK-orig

filter, a Cholesky decomposition of $\tilde{\mathbf{A}}^{-1}$ in Eq. (20) is applied, while the other filters use the symmetric square root.

The left column of figure 1 shows the MRMSE for the four filter variants as a function of the forgetting factor and the ensemble size. Filter divergence is defined for an MRMSE larger than one. A white field indicates a parameter set for which the filter diverges in at least one of the ten experiments.

The ETKF and SEIK-sqrt methods provide almost identical results, with some differences mostly close to the edge to filter divergence. The differences between the results from the ETKF and the ESTKF are even smaller. While mathematically, both variants are identical, the numerical results differ slightly close to the edge to filter divergence. Here, the results of each set of 10 experiments with different random numbers show a larger variability. Thus, the behavior of the filters is less stable in this region and small differences can lead to significant differences. For example in the case with $m = 40$ and a forgetting factor of 0.99, the ESTKF still converges while the ETKF diverges. However, the divergence occurs only in three of the ten experiments, which is counted as divergence in the computation of the mean MRMSE. The differences in the MRMSE for the ETKF and ESTKF result from the distinct analysis formulations of both filters. These become visible with the finite numerical precision of the computations over the long assimilation experiments of 50000 analysis steps. When one considers only the first analysis step, the difference between the transformation matrices is of order $\mathcal{O}(10^{-15})$. The differences in the ensemble transformation matrices of ETKF and SEIK-sqrt are of order $\mathcal{O}(10^{-3})$. While these differences are small with a difference up to 2% of the actual values of the transformation matrix, they can lead to a slightly larger deviation of the MRMSE for the SEIK-sqrt from the MRMSE of ESTKF and ETKF.

The behavior of the SEIK-orig is distinct from the other filters. The filter diverges in most cases with a forgetting factor of 0.97 and above. In contrast, the other filters diverge only for a forgetting factor of at least 0.99. In addition, the minimum MRMSE obtained with SEIK-orig using the deterministic $\hat{\mathbf{\Omega}}$ is 0.192 in contrast to the MRMSE of about 0.180 obtained with ETKF and SEIK-sqrt. This difference is statistically significant.

c. Results with random ensemble transformations

The original SEIK filter was always described using a random transformation matrix $\mathbf{\Omega}$ that preserves the ensemble mean and covariance matrix. Here, the performance of the four filter methods is examined using random rotations. Thus, $\mathbf{\Lambda}$ in Eq. (9) is now used as a mean-preserving random matrix. In SEIK-orig and SEIK-sqrt, a random matrix $\mathbf{\Omega}$ is used (see Appendix for its construction). In the ESTKF a random matrix $\mathbf{\Omega}$ is only used for the computation of the weight matrix $\mathbf{W}^{SEIK}$ in Eq. (22). Because $\mathbf{\Lambda}$ and $\mathbf{\Omega}$ have distinct sizes and are generated by different schemes, the random rotations applied in the ETKF will be distinct from those used in the SEIK filters and the ESTKF.

The MRMSE for the four filter variants with random transformations is shown in the right column of figure 1. The randomization results in almost identical MRMSE for all four methods. This indicates that the ensembles of the four methods are statistically of equal quality. Significant differences between the four filters only occur close to the edge to filter divergence, where the filters' behavior is less stable. The fact that the results of SEIK-orig are comparable to those of the other filters shows that the traditional use of the Cholesky decompostion of $\tilde{\mathbf{A}}^{-1}$ in Eq. (20) in SEIK-orig does not deteriorate the state estimate.

The smallest obtained MRMSE is 0.1754. Thus, the MRMSE is slightly smaller with random than with deterministic transformations. This behavior is consistent with the findings by Sakov and Oke (2008). The difference to the MRMSE obtained with deterministic transformations is statistically significant.

d. Ensemble quality

The inferior behavior of SEIK-orig in case of deterministic ensemble transformations can be related to a suboptimal representation of the ensemble. The analysis equations of the filter algorithms based on the Kalman filter assume that the errors are Gaussian distributed. Lawson and Hansen (2004) discussed the effects of nonlinearity on the example of the classical EnKF with perturbed observations and the deterministic ensemble square-root filter (Whitaker and Hamill 2002). They found that the ensemble distributions remain closer to Gaussian in the case of the stochastic EnKF.

The ensemble quality can be assessed on the basis of the skewness and kurtosis of the ensembles. These statistical moments will be non-zero if the ensembles are non-Gaussian. Table 2 shows the median and the semi-interquartile range (SIQR) of the skewness and kurtosis for experiments with $m = 40$ and a forgetting factor of $\rho = 0.97$. The median of the skewness is about equal for all four filters. However, the SIQR is larger for SEIK-orig than for the other filters. Thus, it is more likely that the ensemble is skewed when applying SEIK-orig. Further, the median and SIQR of the kurtosis are much larger for SEIK-orig than for the filters using the symmetric square root. Thus, the ensemble distributions of SEIK-sqrt, ESTKF, and ETKF are closer to Gaussian distributions than the distribution

of SEIK-orig. The stronger deviation from Gaussianity of the ensemble for SEIK-orig is frequently caused by outliers.

When random ensemble rotations are applied, the statistics of skewness and kurtosis are almost identical for all four methods. The median of the skewness is about zero with an SIQR of 0.24. The kurtosis has a median of -0.26 with an SIQR of 0.37. Thus, the values of SIQR and median are closer to zero than in the case of deterministic transformations. This behavior can be attributed to the removal of ensemble outliers by the random rotation (see Sakov and Oke 2008; Anderson 2010).

7. Conclusion

This study examined the Singular "Evolutive" Interpolated Kalman (SEIK) filter. It was shown that the SEIK filter belongs to the class of ensemble square-root Kalman filters. In addition, a variant of the SEIK filter was developed that results in ensemble transformations that are identical to those of the ETKF, but has at a slightly lower computational cost. The variant is referred to as Error Subspace Transform Kalman Filter (ESTKF) because it explicitly projects the ensemble onto the error subspace and computes the ensemble transformation in this space.

Numerical twin experiments with the Lorenz-96 model and deterministic ensemble transformations showed very similar results for the SEIK filter and the ETKF. The differences in the results of the ESTKF and the ETKF are significantly smaller except in the parameter region where both filters exhibit unstable behavior. The variations in the results are related to the ensemble transformations performed in the filters. The differences in the ensemble

transformations of SEIK and ETKF are very small. The transformations of the ESTKF and ETKF are analytically identical and at the initial time of the experiments also identical up to numerical precision. However, in the full twin experiments the tiny differences grow due to the finite precision of the computations in combination with the nonlinearity of the model.

Using a Cholesky decomposition in the original SEIK filter with deterministic ensemble transformation resulted in higher errors than the application of the symmetric square root. This effect was caused by an inferior ensemble quality. Accordingly, the experiments indicate that for deterministic ensemble transformations, the symmetric square-root should be used in the SEIK filter.

The assimilations with random ensemble transformations provided results that were superior to those using deterministic transformations. This effect was caused by the fact that with randomization the ensemble statistics were closer to Gaussian distributions, which are assumed in the analysis step of the Kalman filter. In the case of random transformations, the original SEIK filter with Cholesky decomposition provided state estimates of the same quality as the other filter methods. The numerical results are particular for the specific implementation of the filter algorithms as well as the Lorenz-96 model. However, following the analytical considerations, other implementations of the SEIK filter, the ESTKF, and the ETKF should provide similar results.

The findings of this study unify the developments of the SEIK filter with the class of ensemble square-root Kalman filters. Further, the newly introduced ESTKF variant of the SEIK filter provides consistent projections between the ensemble space and the error subspace. Together with the ETKF, the ESTKF has the advantage to provide minimum trans-

formations of the ensemble members. If the minimum transformation is not required, the original SEIK filter is also well suited for practical data assimilation applications.

Acknowledgments.

We are grateful to the editor, Dr. Herschel Mitchell, as well as three anonymous reviewers whose comments helped to improve the text. Also, we like to thank Dr. Marc Taylor for carefully proofreading the manuscript.

APPENDIX

Generation of matrix $\mathbf{\Omega}$

The generation of the matrix $\mathbf{\Omega}$ based on random numbers has been discussed by Hoteit (2001) and Pham (2001) as "second order exact sampling". With respect to generating a particular deterministic form $\hat{\mathbf{\Omega}}$ of $\mathbf{\Omega}$, we review its proposed generation. Note that the algorithm to generate $\mathbf{\Omega}$ results in spherical sigma points discussed by Wang et al. (2004).

Matrix $\mathbf{\Omega}$ is required to have orthonormal columns. In addition, the columns need to be orthogonal to the vector whose elements are all one. A Householder matrix associated with the vector $\mathbf{a}_i = (a_{i,1}, \ldots, a_{i,i})^T$ of size i can be used to generate $\mathbf{\Omega}$. It is given by

$$h(\mathbf{a}_i) = \mathbf{I}_{i\times i} - \frac{1}{|a_{i,i}| + 1}\mathbf{a}_i^{sign}\left(\mathbf{a}_i^{sign}\right)^T. \tag{A1}$$

Here, $\mathbf{a}_i^{sign}$ is identical to $\mathbf{a}_i$ except for the last element, which is $a_{i,i}^{sign} = a_{i,i} + sign(a_{i,i-1})$.

Using $h(\mathbf{a}_i)$, the following recursion (see Hoteit 2001) generates a random matrix $\mathbf{\Omega}$:

1. Set $\mathbf{\Omega}_1 = \mathbf{a}_1$ where $\mathbf{a}_1$ is 1 or -1 with equal probability.

2. Recursion: For $i = 2, \ldots, m-1$ initialize a random vector $\mathbf{a}_i$ of unit norm. Then use the first $i-1$ columns of the Householder matrix $h(\mathbf{a}_i)$ in Eq. (A1), denoted by h^-, to compute the $i \times i$ matrix

$$\mathbf{\Omega}_i = \begin{pmatrix} h^-(\mathbf{a}_i)\mathbf{\Omega}_{i-1} & \mathbf{a}_i \end{pmatrix} \tag{A2}$$

3. For $\mathbf{a}_m = m^{-1/2}(1, \ldots, 1)^T$ compute the final $m \times (m-1)$ matrix $\mathbf{\Omega}$ as

$$\mathbf{\Omega} = h^-(\mathbf{a}_m)\mathbf{\Omega}_{m-1}. \tag{A3}$$

A simple deterministic variant of $\mathbf{\Omega}$ can be obtained by taking

$$\hat{\mathbf{\Omega}} = h^-(\mathbf{a}_m) \tag{A4}$$

with $\mathbf{a}_m = m^{-1/2}(1, \ldots, 1)^T$. This is equivalent to choosing $\mathbf{\Omega}_{m-1} = \mathbf{I}_{(m-1)\times(m-1)}$ in (A3).

REFERENCES

Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.

Anderson, J. L., 2010: A non-gaussian ensemble filter update for data assimilation. *Mon. Wea. Rev.*, **138**, 4186–4198.

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.

Brusdal, K., J. M. Brankart, G. Halberstadt, G. Evensen, P. Brasseur, P. J. van Leeuwen, E. Dombrowsky, and J. Verron, 2003: A demonstration of ensemble based assimilation methods with a layered OGCM from the perspective of operational ocean forecasting systems. *J. Mar. Syst.*, **40-41**, 253–289.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: On the 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 (C5)**, 10 143–10 162.

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

Hoteit, I., 2001: Filtres de kalman réduits et efficaces pour l'assimilation de données en océanographie. Ph.D. thesis, l'Université de Joseph Fourier, Grenoble, France.

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

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

Janjić, T., L. Nerger, A. Albertella, J. Schröter, and S. Skachko, 2011: On domain localization in ensemble based Kalman filter algorithms. *Mon. Wea. Rev.*, **139**, 2046–2060.

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.

Livings, D. M., S. L. Dance, and N. K. Nichols, 2008: Unbiased ensemble square root filters. *Physica D*, **237**, 1021–1028.

Lorenz, E. N., 1996: Predictability - a problem partly solved. *Proceedings Seminar on Predictability*, ECMWF, Reading, UK, 1–18.

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

Nerger, L., S. Danilov, W. Hiller, and J. Schröter, 2006: Using sea level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter. *Ocean Dynamics*, **56**, 634–649.

Nerger, L., S. Danilov, G. Kivman, W. Hiller, and J. Schröter, 2007: Data assimilation with the ensemble Kalman filter and the SEIK filter applied to a finite element model of the North Atlantic. *J. Mar. Syst.*, **65**, 288–298.

Nerger, L. and W. W. Gregg, 2007: Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. *J. Mar. Syst.*, **68**, 237–254.

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

Nerger, L., W. Hiller, and J. Schröter, 2005b: PDAF - the Parallel Data Assimilation Framework: Experiences with Kalman filtering. *Use of High Performance Computing in Meteorology - Proceedings of the 11. ECMWF Workshop*, W. Zwieflhofer and G. Mozdzynski, Eds., World Scientific, 63–83.

Ott, E., et al., 2004: A local ensemble Kalman filter for atmospheric data asimilation. *Tellus*, **56A**, 415–428.

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

Pham, D. T., J. Verron, and L. Gourdeau, 1998: Singular evolutive Kalman filters for data assimilation in oceanography. *C. R. Acad. Sci., Ser. II*, **326 (4)**, 255–260.

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

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.

Triantafyllou, G., I. Hoteit, and G. Petihakis, 2003: A singular evolutive interpolated Kalman filter for efficient data assimilation in a 3-D complex physical-biogeochemical model of the Cretan sea. *J. Mar. Syst.*, **40-41**, 213–231.

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**, 1913–1927.

Yang, S.-C., E. Kalnay, B. R. Hunt, and N. E. Bowler, 2009: Weight interpolation for efficient data assimilation with the local ensemble transform kalman filter. *Q. J. Roy. Meteor. Soc.*, **135**, 251–262.

List of Tables

TABLE 1. Summary of the leading computational cost of the ensemble transformations as a function of ensemble size m, number of observations p, and state dimension n.

Filter	Cost
ETKF	$O(pm^2 + m^3 + nm^2)$
SEIK	$O\left(p(m-1)^2 + m(m-1)^2 + nm(m-1)\right)$

TABLE 2. Skewness and kurtosis for the case of deterministic ensemble transformations. Shown are the median and the semi-interquartile range (SIQR) for an experiment with 5000 analysis steps for $m = 40$ and a forgetting factor of 0.97.

	skewness		kurtosis	
Filter	median	SIQR	median	SIQR
ETKF	0.025	0.456	0.2	0.79
SEIK-orig	0.025	0.630	2.1	2.46
SEIK-sqrt	0.024	0.441	0.2	0.69
ESTKF	0.023	0.445	0.2	0.76

List of Figures

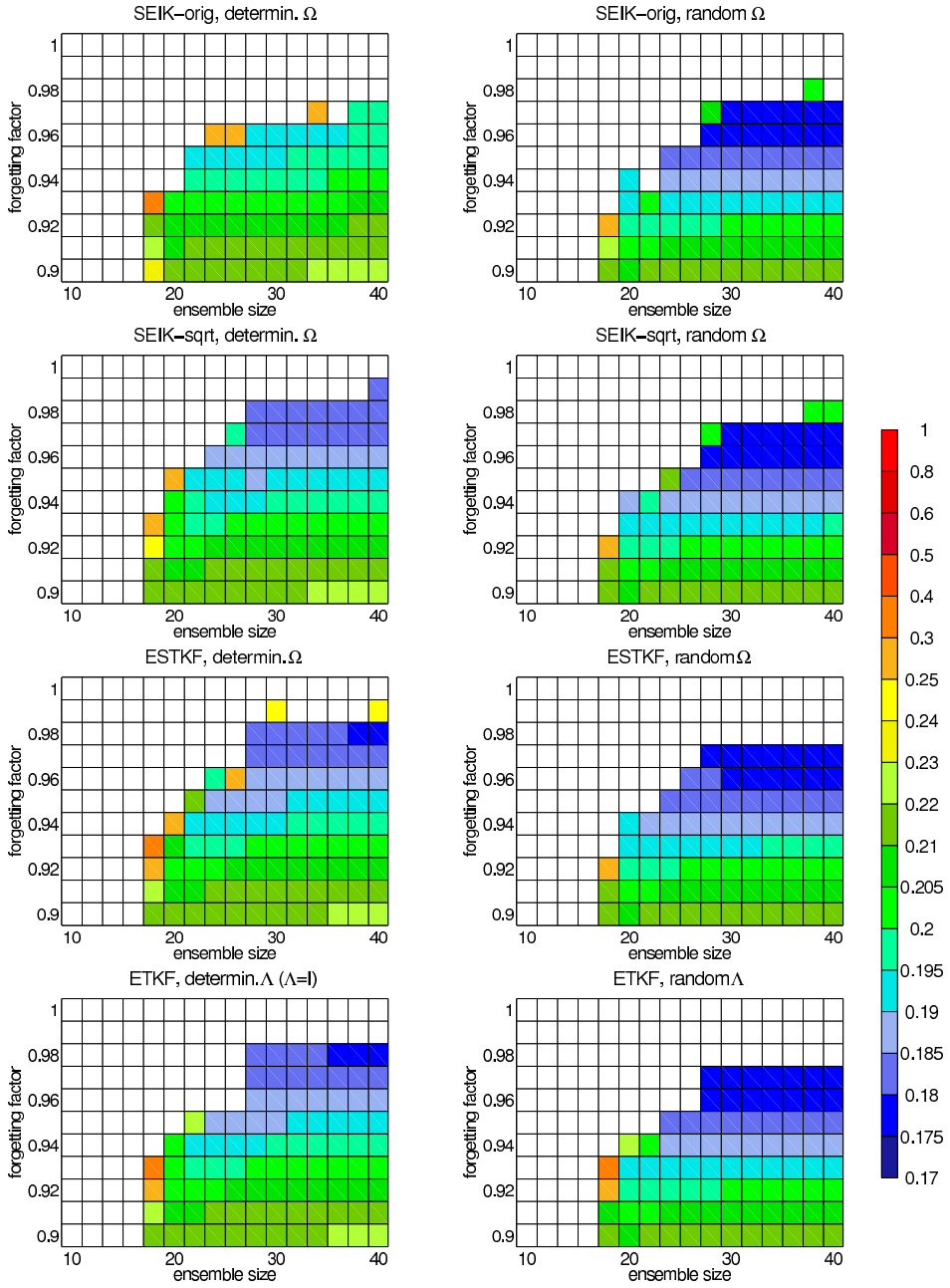

FIG. 1. RMS mean errors for the SEIK-orig (top), SEIK-sqrt (second row), ESTKF (third row), and ETKF (last row). Left: errors obtained using deterministic ensemble transformation matrices; Right: error obtained using random transformation matrices.