Sampling Errors in Ensemble Kalman Filtering. Part I: Theory

WILLIAM SACHER AND PETER BARTELLO

McGill University, Montréal, Québec, Canada

(Manuscript received 27 July 2007, in final form 4 December 2007)

ABSTRACT

This paper discusses the quality of the analysis given by the ensemble Kalman filter in a perfect model context when ensemble sizes are limited. The overall goal is to improve the theoretical understanding of the problem of systematic errors in the analysis variance due to the limited size of the ensemble, as well as the potential of the so-called double-ensemble Kalman filter, covariance inflation, and randomly perturbed analysis techniques to produce a stable analysis—that is to say, one not subject to filter divergence. This is achieved by expressing the error of the ensemble mean and the analysis error covariance matrix in terms of the sampling noise in the background error covariance matrix (owing to the finite ensemble estimation) and by comparing these errors for all methods. Theoretical predictions are confirmed with a simple scalar test case. In light of the analytical results obtained, the expression of the optimal covariance inflation factor is proposed in terms of the limited ensemble size and the Kalman gain.

1. Introduction

This paper begins to address what is in our view the vital question facing those who employ the ensemble Kalman filter (EnKf; Evensen 1994): What are the systematic biases in the statistics on which it is based, resulting from a limited ensemble size? Ensemble forecasting is clearly an expensive proposition. The forecaster is faced with many choices to reduce the computational burden of the task. For example, is it better to have a small ensemble of high-resolution models (thereby potentially reducing model error at the risk of sampling error) or to have a large ensemble of lower-resolution models (thereby reducing sampling error at the risk of model error)? The answer depends on quite a number of things and in this study, the focus is much less ambitious. As a first step in the process, we attempt to quantify the sampling error resulting from the use of a relatively small (i.e., less than infinite) ensemble. Of course, we are not the first to be aware of these issues and our statistical error scaling arguments are employed to understand and improve previous attempts to deal with these biases, such as covariance inflation and the subdivision of the ensemble into smaller subensembles in order to remove "inbreeding" (see references below).

Blending of background and observation information is a necessary step in "data assimilation" (later referred to as DA), because neither model estimate nor observation is ever perfect. In this study, we consider a sequential (as opposed to variational) DA approach where the product of the analysis cycle is used as an initial state of a numerical prediction system. The basic tool in this approach is filtering; that is, all of the information is used up to and including the considered analysis time. We focus on one particular technique, the EnKf, and its numerical weather prediction (NWP) application. The EnKF has been the subject of studies in various meteorological and oceanographic contexts, many of them reviewed by Evensen (2007). However, to date, only the Canadian Meteorological Centre (CMC) has implemented it operationally (Houtekamer and Mitchell 2005). The Kalman filter (e.g., Gelb 1974), or its nonlinear extension the extended Kalman filter (EKF; e.g., Evensen 1992), are common examples of filtering methods. Nevertheless, it has been recognized that the EKF is impossible to implement in an operational NWP context, because of the computational burden (Evensen 1992; Keppenne 2000; Lawson and Hansen 2004). Moreover, unbounded error variance growth due to the linearization procedure has been observed (Evensen 1992). Alternatives to the EKF have been proposed, relying on the assumption that "the tra-

Corresponding author address: William Sacher, Dept. of Atmospheric and Oceanic Sciences, McGill University, 805 Sherbrooke St. W., Montreal, QC H3A 2K6, Canada.
E-mail: sacher@mail.mcgill.ca

DOI: 10.1175/2007MWR2323.1

jectory of the system state in a high-dimensional dynamical system typically lies on a smaller dimensional subspace of the entire phase space" (Farrell and Ioannou 2001) and, thus, can be called reduced-state filters (Lawson and Hansen 2004). One class of these employs Monte Carlo–based methods, where one generates an ensemble of N forecasts that samples the state space. They provide an estimation of the state's probability density function (pdf), whose quality is dependent on the size of the ensemble. The generated pdf can naturally be used to initiate an ensemble prediction system, where one performs "stochastic dynamic prediction" (Epstein 1969). This latter approach is possible with the atmosphere's deterministic laws because of the chaotic nature of the solutions. Monte Carlo filters can be divided into two subclasses: deterministic and stochastic, as stated by Tippett et al. (2003) or Lawson and Hansen (2004). In the stochastic filter (Houtekamer and Mitchell 1998; Burgers et al. 1998), observations are treated as random variables and thus perturbed to avoid an underestimation of the analysis variance (Burgers et al. 1998). This version has shown efficiency in various contexts (Houtekamer and Mitchell 1998; Hansen and Smith 2000; Reichle et al. 2002; Snyder and Zhang 2003). Deterministic formulations of Monte Carlo–based techniques, as reviewed by, for example, Tippett et al. (2003), do not require perturbed observations.

2. Sampling issues

Previous experiments dealing with realistic models have highlighted a few issues concerning EnKf's operational application: (a) The limited number of members [usually no more than $O(100)$] can cause spurious covariances between widely separated locations, (b) the unavoidable presence of model errors of different types (presence of unresolved scales, imperfectly known forcing, etc.) can also cause spurious covariances, and (c) the presence of observational error of representation and errors in the statistical description of observations. In the current study, we only address issue a above. However, issue a (as well as issue b) can potentially lead to "underestimation" of the analysis variance. In some cases, this is known to be fatal to the EnKf: the analysis loses track of the truth, whereas its spread remains "too small." As spread represents the uncertainty, the filter "believes" it performs better than it does in reality. This behavior is called filter divergence (e.g., Maybeck 1979, p. 338) and was observed by, for example, Houtekamer and Mitchell (1998) and Hamill et al. (2001) for the EnKf. We consider it important to clarify the words "too small" and "underestimation." For this purpose, we recall that "the goal of the assimilation is to produce a random sample of the conditional probability distribution that is consistent with the truth while minimizing the rms error of the ensemble mean from the truth" (Anderson and Anderson 1999). Hence, the aim of Monte Carlo filtering is to satisfy the following two criteria:

(i) to produce a sample with an ensemble mean, commonly used as the state's best estimate, having the smallest expected error, and
(ii) to produce a reliable ensemble of analysis, which is to say that the variations of the ensemble mean error, though not accessible, should be well represented by the ensemble of analyses given by the filter itself (in other words, the true state should be statistically indistinguishable from a randomly selected member of the ensemble).

In fact, criterion ii is a necessary condition for a perfect assimilation (Anderson and Anderson 1999). To be sure to satisfy it, a requirement is that the spread among the ensemble members remains representative of the difference between the ensemble mean and the true state, as the assimilation cycle proceeds (Anderson and Anderson 1999; Houtekamer and Mitchell 1998). We define the *optimal* analysis in the EnKf framework as the one obtained with an infinite number of members, and the *ideal* analysis as the one satisfying criteria i and ii when using a finite ensemble. Then, as only N members are used, and not an infinite number, the *ideal* value of the spread is not the one obtained with an infinite ensemble. It is rather represented by the mean squared error of the analysis mean calculated over the N ensemble members, which is itself different from the mean squared error (MSE) of an analysis using an ensemble of infinite size. Now, if the analysis spread calculated over the ensemble of size N under- or overestimates this *ideal* value, error statistics are not well represented (henceforth "misrepresented"), and criterion ii is not achieved. In fact, the analysis spread of a N-member ensemble systematically underestimates the MSE of the analysis mean calculated over these members.

To illustrate this, results of a twin-experiment simulation of one analysis cycle with the very simple case of a Gaussian process represented by a scalar are shown in Fig. 1. The curves show an average over 10^5 realizations. Forecasts and observations are taken from a normal distribution of zero mean and unit variance [$\mathcal{N}(0, 1)$] using a pseudorandom number generator, for various ensemble sizes. One can see that the EnKf analysis error variance converges toward the optimal value [in this case $\sigma_a^2 = (1 - k) = 0.5$] from below, that is, by negative error values. This result was also found by van Leeuwen (1999) in the multivariate case and by Whi-

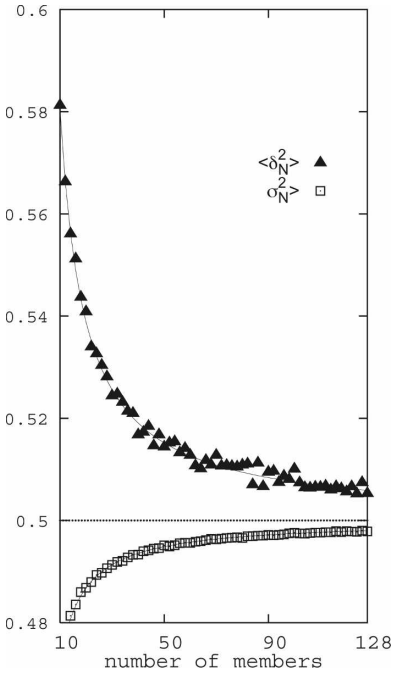

FIG. 1. Simulation of one analysis cycle with a Gaussian scalar, taken from $\mathcal{N}(0, 1)$. The observation ensemble is taken from $\mathcal{N}(0, 1)$ too; hence, the optimal gain $K = 0.5$, and the optimal analysis variance $\sigma_a^2 = 0.5$. Triangles show the MSE of the ensemble mean, and squares show the average analysis ensemble variance. Means are performed over 100 000 realizations. For each experimental curve, the theoretical curves $\langle \sigma_a^2(N) \rangle = \sigma_a^2[1 - 2k(1 - k)(N - 1)^{-1}]$ and $\langle \delta_a^2(N) \rangle = \sigma_a^2(1 + N^{-1})[1 + 2k(1 - k)(N - 1)^{-1}]$, the scalar versions of (13) and (15), are drawn in continuous lines.

taker and Hamill (2002) in a similar scalar case. On the other hand, the MSE of the ensemble mean converges toward the optimal value of the analysis error variance from above, with positive error values. This shows that, on average, the analysis generally does not satisfy criterion ii. Moreover, the analysis variance underestimates the *optimal* analysis variance. It subsequently underestimates the *ideal* value, which corresponds to the MSE of the ensemble mean.

Note that in experimental applications of the EnKf, underestimation of the analysis variance with respect to the optimal value is more often observed than overestimation, whereas it is known that an overestimate is "safer" than an underestimate (Maybeck 1979, p. 339; Daley 1991, section 4.9). Houtekamer and Mitchell (1998) claimed that besides sampling noise related to the limited size of the ensemble, "inbreeding" problems may come into play and also cause spread underestimation. This would refer to the fact that we produce the estimate of the analysis error variation with the same ensemble used to compute the analysis. This is another way to express that the gain and the analysis variance depend nonlinearly on the background error covariance matrix (Whitaker and Hamill 2002). Various techniques have been proposed to correct for inbreeding-sampling errors and stabilize the EnKf. They include covariance localization (Houtekamer and Mitchell 2001; Ott et al. 2004), a hybrid formulation of EnKf–three-dimensional variational data assimilation (3DVAR; Hamill and Snyder 2000), a double-ensemble Kalman filter (DEnKf) (Houtekamer and Mitchell 1998), and covariance inflation (Anderson and Anderson 1999; Hamill et al. 2001; Whitaker and Hamill 2002; Ott et al. 2004; Gillijns et al. 2005). Finally, Pham et al. (1998) used a "forgetting factor" to reduce the observation error covariance matrix **R**. We do not present "covariance localization" nor hybrid formulations here, though conclusions may be drawn from our analytic results. In this study, we want to improve the theoretical understanding of the problem of misrepresentation of the analysis variance of the EnKf, as well as investigate the potential of the DEnKf and covariance inflation techniques to produce stable analyses (i.e., not subject to filter divergence).

To quantitatively address the question, we need to compare representative measures of the spread and the error of the ensemble mean. The Gaussian hypothesis for the statistics leads one to naturally calculate the second-order moments of the pdf of the error of the ensemble mean and of the analysis error to verify whether the criteria are satisfied. We consider here an idealized context that intentionally does not include any localization procedure, though Houtekamer and Mitchell (2001) show that localization can be effective in forcing the filter to satisfy criterion ii. Moreover, signal contamination by errors present in an operational context (e.g., model or representation errors) is not considered either. Using the established results, we then propose a theoretically based covariance inflation technique. The technique presented, which modifies the analysis update, turns out to be easier to implement and less costly than the DEnKf, whose performance is also considered. Section 3 describes the formulation of the perturbed-observation EnKf used here. In section 4, we establish the analytical expressions for the average analysis error covariance matrix and the MSE of the ensemble mean for the EnKf. In section 5, similar analytic results are demonstrated for the DEnKf, and a new covariance inflation technique is described. For each analytical result, we give a simple application using the analysis of a Gaussian scalar process. This application, though very simplistic, gives a fairly good insight into the problem. Moreover, we will see in an upcoming second paper that our analytic results compare well to analyses made with a barotropic model.

3. Formulation of the EnKf

The Kalman filter is a "recursive state estimation technique" (Lawson and Hansen 2004). It consists of two steps: (i) a propagation step using the system dynamics to evolve the first two moments of the state's pdf (characterizing completely a Gaussian distribution) and (ii) an analysis step using a solution derived from Bayes's rule, giving "either a maximum likelihood estimate (e.g., Lorenc 1986) or a minimum error variance estimate (e.g., within estimation theory; Cohn 1997)." It can also be identified "as a recursive least squares problem properly weighted by the inverses of the relevant error covariance matrices (e.g., Wunsch 1996)," as reviewed by Lawson and Hansen (2004).

Conceptually, the EnKf differs from other Kalman filters only by the way it treats the propagation step. It uses a Monte Carlo method to generate an ensemble of N forecast members ψ_j, $j \in [\![1, N]\!]$, that are independent realizations of the process, to sample the state space. Their evolutions between times t_{k-1} and t_k are calculated by using the fully nonlinear model:

$$\psi_j^f(t_k) = \mathcal{M}[\psi_j^a(t_{k-1})], \tag{1}$$

where ψ_j^f (respectively, ψ_j^a) is the jth member of the forecast (respectively, analysis) ensemble of state vectors. Each member is a state vector of length n, the dimension of the state space. In the general case, a model error term should be present on the right-hand side of (1). In this study, we consider the model as perfect.

This ensemble can subsequently be used to estimate the evolution of any moment of the pdf with respect to time. Particularly,

$$\mathbf{P}_N^f = \frac{1}{N-1}\sum_{j=1}^{N}(\psi_j^f - \overline{\psi^f})(\psi_j^f - \overline{\psi^f})^{\mathrm{T}} = \frac{1}{N-1}\sum_{j=1}^{N}\boldsymbol{\xi}_{j^f}^f\boldsymbol{\xi}_{j^f}^{f\mathrm{T}}, \tag{2}$$

where the overbar stands for the ensemble-based estimate of the mean and $\boldsymbol{\xi}_{j^f}^f = \psi_j^f - \overline{\psi^f}$. In addition, $\mathbf{P}_N^f$ is an estimate (using N ensemble members) of the background error covariance matrix $\mathbf{P}^f = \mathbf{P}_{N=\infty}^f$, defined as

$$\mathbf{P}^f = E\{[\psi^f - E(\psi^f)][\psi^f - E(\psi^f)]^{\mathrm{T}}\}, \tag{3}$$

where $E\{\}$ stands for an expectation value. This expression is itself an estimate of the true background error covariance matrix $\mathbf{P} = E\{(\psi^f - \psi^t)(\psi^f - \psi^t)^{\mathrm{T}}\}$, where ψ^t is the true state. In the EnKf, one sees that the background error is approximated by the departure from the ensemble mean. To simplify the analytic expressions, the notation $\boldsymbol{\xi}_a^b = \psi^a - \psi^b$ is extensively used henceforth. Moreover, any matrix refers to $\mathbf{X}_N$, the approximation to $\mathbf{X} = \mathbf{X}_{N=\infty}$ using N ensemble members. Finally, the true mean value of the process is noted $\psi^\mu = E\{\psi^f\}$. In the stochastic version of the ensemble Kalman filter used in this paper, the analysis step at time t_k consists of the following equations:

$$\psi_j^a = (\mathbf{I} - \mathbf{K}_N\mathbf{H})\psi_j^f + \mathbf{K}_N\mathbf{d}_j \quad \text{and} \tag{4}$$

$$\mathbf{K}_N = \mathbf{P}_N^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_N^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \tag{5}$$

where $\mathbf{R}$, an $m \times m$ matrix, is the observation error covariance matrix, m being the number of observations, and $\mathbf{I}$ is the $n \times n$ identity matrix. Also, $\mathbf{H}$, an $m \times n$ matrix, is an interpolation operator, which projects from model space onto observation space, and $\mathbf{K}_N$, an $n \times m$ matrix, is the so-called Kalman gain matrix. Here, $\forall_j \in [\![1, N]\!]$, and $\mathbf{d}_j = \mathbf{d} + \boldsymbol{\delta}_j$, where $\boldsymbol{\delta}_j$ is a random vector taken from $\mathcal{N}(0, \mathbf{R})$, and $\mathbf{d}$ is the actual measurement vector. Then, as in (2), we have for the analysis error covariance matrix:

$$\mathbf{P}_N^a = \frac{1}{N-1}\sum_{j=1}^{N}(\psi_j^a - \overline{\psi^a})(\psi_j^a - \overline{\psi^a})^{\mathrm{T}} = \frac{1}{N-1}\sum_{j=1}^{N}\boldsymbol{\xi}_{j^a}^a\boldsymbol{\xi}_{j^a}^{a\mathrm{T}}. \tag{6}$$

The ensemble of maximum likelihood (or minimum error variance) estimates is thus a linear combination of the two available fields, namely background and observation fields. Note that in this formulation, one keeps a second-order moment closure. This is not mandatory with Monte Carlo filters, because theoretically one has access to higher-order-moment estimates. Moreover, a linearized version $\mathbf{H}$ of the generally nonlinear interpolation operator $\mathcal{H}$ has been used here.

This definition leads to an interpretation of the EnKf as a Monte Carlo method where the ensemble of model states evolves in state space, with the mean as the best estimate and the error variance as an estimate of the spread. Observations are similarly represented by another ensemble, generated by the addition of random noise sampled from the observational error distribution (Burgers et al. 1998). Hence, at measurement times, each ensemble member is updated using a different observation taken from this ensemble of perturbations. Here, $\mathbf{K}_N$ is the Kalman gain evaluated over the ensemble. Note the nonlinear dependence in (5) of $\mathbf{K}_N$ on $\mathbf{P}_N^f$.

4. On sampling errors in the EnKf

In this section we want to establish a general expression for the average analysis error covariance matrix

and for the MSE of the ensemble mean in terms of the optimal values $\mathbf{P}^a$ and $\mathbf{K}$ (obtained for an ensemble of infinite size). The MSE is the second-order moment of the pdf of the difference between the ensemble mean and the truth and, therefore, is a measure of the analysis accuracy [as used in Houtekamer and Mitchell (1998) and Charron et al. (2006)]. We introduce here the new matrix $\mathbf{\Delta}^a_N$:

$$\mathbf{\Delta}^a_N = (\psi^t - \overline{\psi^a})(\psi^t - \overline{\psi^a})^{\mathrm{T}} = \boldsymbol{\xi}^a_t \boldsymbol{\xi}^{a\mathrm{T}}_t, \tag{7}$$

where ψ^t is the unknown true state, assumed to be a freely evolving solution of the system. We will also use the equivalent definition for the forecast, using superscript f instead of a. As suggested by Houtekamer and Mitchell (1998), the EnKf satisfies criterion ii (i.e., produces a representative analysis ensemble of the actual error variation) if $\mathrm{tr}(\mathbf{\Delta}^a_N) \approx \mathrm{tr}(\mathbf{P}^a_N)$. We proceed to this comparison in order to discuss the performance of the EnKf.

a. Averaged analysis error covariance matrix

Following van Leeuwen (1999), let us now assume that the ensemble estimates are not too far off; hence, we express the statistical error with respect to the unknown true background covariance matrix $\mathbf{P}^f$ as $\boldsymbol{\epsilon}$: $\mathbf{P}^f_N = \mathbf{P}^f + \boldsymbol{\epsilon}$, with $\|\boldsymbol{\epsilon}\| \ll \|\mathbf{P}^f\|$. We want to express the analysis error covariance matrix $\mathbf{P}^a_N$ as a series expansion in $\boldsymbol{\epsilon}$. Similarly, we suppose $\overline{\boldsymbol{\delta}_j \boldsymbol{\delta}_j^{\mathrm{T}}} = \mathbf{R} + \boldsymbol{\rho}$ with $\|\boldsymbol{\rho}\| \ll \|\mathbf{R}\|$. We shall assume $E\{\boldsymbol{\epsilon}\} = \mathbf{0}_n$ and $E\{\boldsymbol{\rho}\} = \mathbf{0}_m$, where $\mathbf{0}_n$ and $\mathbf{0}_m$ are, respectively, the null matrix for the state and observation spaces.

Henceforth, in order to simplify the algebra, we define $\mathbf{L}_N = \mathbf{I} - \mathbf{K}_N\mathbf{H}$, as well as $\mathbf{\Pi}^a_N = \mathbf{L}_N\mathbf{P}^f_N$. Also, $\mathbf{\Theta} = (\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})$ and $\mathbf{\Phi}$ is the inverse of $\mathbf{\Theta}$. Then, the optimal Kalman gain, given by (5) when taking an infinite ensemble, can be written $\mathbf{K} = \mathbf{P}^f\mathbf{H}^{\mathrm{T}}\mathbf{\Phi}$, and its transpose $\mathbf{K}^{\mathrm{T}} = (\mathbf{P}^f\mathbf{H}^{\mathrm{T}}\mathbf{\Phi})^{\mathrm{T}} = \mathbf{\Phi}\mathbf{H}\mathbf{P}^f$. We finally define the gainlike matrix $\boldsymbol{\kappa} = \boldsymbol{\epsilon}\mathbf{H}^{\mathrm{T}}\mathbf{\Phi}$.

Substituting (4) and the mean of (4) into (6), one obtains (Burgers et al. 1998; Whitaker and Hamill 2002; Furrer and Bengtsson 2007)

$$\mathbf{P}^a_N = \mathbf{\Pi}^a_N + \mathbf{K}_N\boldsymbol{\rho}\mathbf{K}^{\mathrm{T}}_N + \mathbf{L}_N\overline{\boldsymbol{\xi}^f_{j^f}\boldsymbol{\delta}_j^{\mathrm{T}}}\mathbf{K}^{\mathrm{T}}_N + (\mathbf{L}_N\overline{\boldsymbol{\xi}^f_{j^f}\boldsymbol{\delta}_j^{\mathrm{T}}}\mathbf{K}^{\mathrm{T}}_N)^{\mathrm{T}}. \tag{8}$$

Hence, while the perturbed observations avoid systematic error covariance underestimation with respect to the optimal analysis, as shown by Burgers et al. (1998), they introduce sampling errors represented by the last three terms in (8). Then, since $\|\boldsymbol{\epsilon}\| \ll \|\mathbf{P}^f\|$, we have $\|\mathbf{H}\boldsymbol{\kappa}\| \ll \|\mathbf{I}_m\|$ and orders bigger than 2 in $\mathbf{H}\boldsymbol{\kappa}$ and/or $\boldsymbol{\rho}$ will be systematically discarded below. Using this approximation in the Taylor expansions (A1) and (A2), given in appendix A, leads to

$$\mathbf{K}_N = \mathbf{K} + \mathbf{L}(\boldsymbol{\kappa} - \boldsymbol{\kappa}\mathbf{H}\boldsymbol{\kappa}) + O(\|\boldsymbol{\kappa}\mathbf{H}\boldsymbol{\kappa}\|) \quad \text{and} \tag{9}$$

$$\mathbf{\Pi}^a_N = \mathbf{P}^a + \mathbf{L}(\boldsymbol{\epsilon} - \boldsymbol{\epsilon}\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}} - \boldsymbol{\kappa}\mathbf{\Theta}\boldsymbol{\kappa}^{\mathrm{T}} + \boldsymbol{\kappa}\mathbf{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}) + O(\|\boldsymbol{\kappa}\mathbf{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}\|), \tag{10}$$

where we have considered $\mathbf{P}^a = \mathbf{L}\mathbf{P}^f$ as the optimal analysis error covariance matrix and $\|\,\|$ is the matrix norm of interest.

One may calculate the expectation of (8), or equivalently a mean over many independent realizations of an analysis cycle using N ensemble members, denoted as $\langle . \rangle$. If we suppose $\langle \overline{\boldsymbol{\xi}^f_k\boldsymbol{\delta}_k^{\mathrm{T}}} \rangle = \mathbf{0}_{n,m}$, that is, that the observation and the forecast error are uncorrelated (a common assumption in DA), we are left with

$$\langle \mathbf{K}_N \rangle = \mathbf{K} - \mathbf{L}\langle \boldsymbol{\kappa}\mathbf{H}\boldsymbol{\kappa} \rangle + O(\|\boldsymbol{\kappa}\mathbf{H}\boldsymbol{\kappa}\|) \quad \text{and} \tag{11}$$

$$\langle \mathbf{P}^a_N \rangle = \mathbf{P}^a - \mathbf{L}\langle \boldsymbol{\kappa}\mathbf{\Theta}\boldsymbol{\kappa}^{\mathrm{T}} \rangle\mathbf{L}^{\mathrm{T}} + O(\|\boldsymbol{\kappa}\mathbf{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}\|), \tag{12}$$

that is to say, the sum of the optimal value and a negative term arising from the limited-size ensemble, implying that the analysis variance and the gain experience a systematic negative bias, as noted by van Leeuwen (1999) and Furrer and Bengtsson (2007). This result is consistent with the general observation that ensemble spread is too low and that the filter can ultimately diverge by putting too much weight on the background. Here, $\langle \boldsymbol{\kappa}\mathbf{\Theta}\boldsymbol{\kappa}^{\mathrm{T}} \rangle = \langle \boldsymbol{\epsilon}\mathbf{H}^{\mathrm{T}}\mathbf{\Phi}\mathbf{H}\boldsymbol{\epsilon} \rangle$ represents the variance of the error variance, weighted by $\mathbf{\Phi} = \mathbf{\Theta}^{-1}$. For a Gaussian process, it can be shown (see appendix B) that

$$\mathrm{tr}(\mathbf{L}\langle \boldsymbol{\kappa}\mathbf{\Theta}\boldsymbol{\kappa}^{\mathrm{T}} \rangle\mathbf{L}^{\mathrm{T}}) = \frac{1}{N-1}[\mathrm{tr}(\mathbf{P}^a\mathbf{L}\mathbf{K}\mathbf{H}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{K}\mathbf{H})];$$

then,

$$\mathrm{tr}(\langle \mathbf{P}^a_N \rangle) \approx \mathrm{tr}(\mathbf{P}^a) - \frac{1}{N-1}[\mathrm{tr}(\mathbf{P}^a\mathbf{L}\mathbf{K}\mathbf{H}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{K}\mathbf{H})]. \tag{13}$$

Note that for the straightforward case of a scalar, we have $\mathrm{tr}(\mathbf{P}^a\mathbf{L}\mathbf{K}\mathbf{H}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{K}\mathbf{H}) = 2\sigma_a^2(1 - k)k$, where σ_a and k are the optimal analysis variance and Kalman gain, respectively. Then, $\langle \sigma_a^2(N) \rangle \approx \sigma_a^2[1 - 2(1 - k)k(N - 1)^{-1}]$, which is plotted in Fig. 1 and matches with the simulation quite closely.

b. Mean squared error of the ensemble mean

We consider here a twin experiment in which one of the realizations, ψ^t, is arbitrarily chosen to be the truth, and an ensemble set of N realizations, ψ_j, $j \in [\![1, N]\!]$, is used to calculate the statistical properties of the process. Taking the expectation of $\mathbf{\Delta}^a_N = \boldsymbol{\xi}^a_t\boldsymbol{\xi}^{a\mathrm{T}}_t$, assuming

a Gaussian process, and using (9), and (C1) from appendix C, we get (details can be found in appendix D)

$$\langle \boldsymbol{\Delta}_N^a \rangle = \mathbf{P}^a + \frac{\mathbf{P}^a}{N} + \left(1 + \frac{1}{N}\right)\mathbf{L}\langle \boldsymbol{\kappa}\boldsymbol{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}\rangle\mathbf{L}^{\mathrm{T}} + O(\|\boldsymbol{\kappa}\boldsymbol{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}\|), \tag{14}$$

which is the sum of three terms: (i) the optimal analysis variance $\mathbf{P}^a$, (ii) the sampling variance $\mathbf{P}^a/N$, and (iii) a systematic bias similar but opposite to the one in (12). The first two terms give the regular variance of the error variance term [as seen in (C1) for the background analysis variance], and the third one comes from the nonlinear dependence of $\mathbf{K}_N$ on $\mathbf{P}_N^f$. Unlike for the analysis covariance in (12), we note that the MSE is converging toward the optimal value $\mathbf{P}^a$ with a *positive* error. In absolute value, this error is on average larger than the error made on the analysis covariance by the additional factor $\mathbf{P}^a/N$. Taking the trace of (14) leads to

$$\mathrm{tr}(\langle \boldsymbol{\Delta}_N^a \rangle) \approx \left(1 + \frac{1}{N}\right)\left\{\mathrm{tr}(\mathbf{P}^a) + \frac{1}{N-1}[\mathrm{tr}(\mathbf{P}^a\mathbf{L}\mathbf{K}\mathbf{H}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{K}\mathbf{H})]\right\}. \tag{15}$$

When compared with (13), we see that criterion ii is generally not satisfied and that the average analysis variance underestimates the MSE of the ensemble mean. For a scalar process, we have $\langle \delta_a^2(N) \rangle \approx (1 + N^{-1})\sigma_a^2[1 + 2(1 - k)k(N - 1)^{-1}]$, which again matches with the result of the simulation presented in Fig. 1.

Order 2 truncation is a very good approximation for $\langle \mathbf{P}_N^a \rangle$ even for small values of N. In this case, higher-order terms almost cancel each other out, though they are not negligible individually. Although it is almost not perceptible in the scalar case, this is not exactly true in the expression of $\langle \boldsymbol{\Delta}_N^a \rangle$, for which higher-order terms in ϵ may have to be added.

5. Proposed solutions: DEnKf and covariance inflation

The results of the previous section show the impossibility for the EnKf to satisfy criterion ii, and therefore that it is naturally subject to divergence for small ensembles. The double-ensemble Kalman filter (DEnKf; Houtekamer and Mitchell 1998), where the N-member ensemble is split into two $N/2$-member subsets is an attempt to fix this problem. The covariance information from one of the subsets is used in the data assimilation of the other subset. By using one subensemble to calculate the background error covariance needed to update the forecast of the other subensemble, the DEnKf is expected to remove the negative bias in the analysis error variance described in the last section, because the dominant term in (10), $-\mathbf{L}\boldsymbol{\kappa}\boldsymbol{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}$, responsible for the systematic negative bias in the analysis error variance described in the last section, will average to zero.

The authors further suggested that the ensemble of size N can be divided in l subensembles of sizes N/l, updating each N/l-member subset by using the other one [of size $(l - 1)N/l$] to calculate $\mathbf{K}_N$. As it still implies two ensembles, we will later refer to this as lDEnKf ($l \in [\![2, N]\!]$), a filter using l independent subensembles. To attain criterion ii perfectly, Houtekamer and Mitchell (1998) also suggested the use of a NDEnKf, therefore updating each member with a gain matrix estimated from all other members of the ensemble (though they did not test this formulation). Note that a lDEnKf theoretically requires the calculation of l Kalman gain matrices and may therefore be more expensive than the single-ensemble version.

Alternatively, some authors correct directly for the spread misrepresentation by using the so-called covariance inflation technique, where the ensemble-based covariances are multiplied by a tunable factor $r \geq 1$ (Anderson and Anderson 1999; Whitaker and Hamill 2002; Ott et al. 2004; Anderson 2001). As seen previously, in order to satisfy criterion ii, these methods should seek an enhanced analysis variance that is as close as possible to the *ideal* value (and not as close as possible to the *optimal* value), although it is necessarily suboptimal due to the use of an ensemble of limited size. We stress here that this inflation is a priori dependent on N as well as on $\mathbf{K}$, as verified below.

In section 5a, we examine the ability of the DEnKf to produce an *ideal* analysis variance in a perfect model context by extending our theoretical analysis. In section 5b, this is employed to formulate covariance inflation and randomly perturbed analysis methods giving, on average, an analysis that satisfies criteria i and ii.

a. Sampling errors in the DEnKf

Here, we examine separately the quality of the partitioned and the merged analysis. This distinction only stands for the way the analysis is used once produced. The former considers the whole ensemble given by merging all of the subensembles, whereas the latter considers each subensemble separately. In a partitioned analysis, each subensemble is defined at the first analysis cycle and evolves independently. In a merged analysis, the l subensembles of size N/l are unified in a single one of size N at the end of each analysis, and then randomly redivided. We chose to do so because authors have used both types of presentations in previous stud-

ies. We treat the 2DEnKf fully, and infer the result for the general case of an *l*DEnKf.

1) PARTITIONED ANALYSIS

The equivalent of (8) for the analysis error covariance matrix of the first ensemble in the 2DEnKf is

$$\mathbf{P}_1^a = \mathbf{\Pi}_1^a + (\mathbf{K}_2 - \mathbf{K}_1)(\mathbf{I}_m + \mathbf{H}\boldsymbol{\kappa}_1)\boldsymbol{\Theta}\mathbf{K}_2^{\mathrm{T}} + \mathbf{K}_2\boldsymbol{\rho}_1\mathbf{K}_2^{\mathrm{T}} + \mathbf{L}_2\overline{\boldsymbol{\xi}_{j^f}^f\boldsymbol{\delta}_j^{\mathrm{T}}}\mathbf{K}_2^{\mathrm{T}} + (\mathbf{L}_2\overline{\boldsymbol{\xi}_{j^f}^f\boldsymbol{\delta}_j^{\mathrm{T}}}\mathbf{K}_2^{\mathrm{T}})^{\mathrm{T}}, \quad (16)$$

where $\mathbf{P}_1^a$, $\mathbf{P}_1^f$, $\mathbf{P}_2^f$, $\mathbf{K}_1$, $\mathbf{K}_2$ are, respectively, the analysis, and the forecast error covariance matrices and Kalman gains calculated over the indicated subensemble, and $\mathbf{I}$ is the $m \times m$ identity matrix. The j in $\boldsymbol{\xi}_{j^f}^f$ and $\boldsymbol{\delta}_j$, as well as the overbar refer to the first ensemble. Furthermore, $\mathbf{L}_2 = \mathbf{I} - \mathbf{K}_2\mathbf{H}$, $\mathbf{\Pi}_1^a = \mathbf{L}_2\mathbf{P}_1^f$, and $\overline{\boldsymbol{\delta}_j\boldsymbol{\delta}_j^{\mathrm{T}}} = \mathbf{R} + \boldsymbol{\rho}_1$. An equivalent expression can be written for $\mathbf{P}_2^a$. We follow here van Leeuwen (1999) in his attempt to write an equivalent of (12) for this case, as well as Houtekamer and Mitchell (1999), considering $\mathbf{P}_1^f = \mathbf{P}^f + \boldsymbol{\epsilon}_1$ and $\mathbf{P}_2^f = \mathbf{P}^f + \boldsymbol{\epsilon}_2$. Neglecting terms of orders higher than two in (A3) and (A1), we have the following expressions:

$$\mathbf{K}_2 = \mathbf{K} + \mathbf{L}(\boldsymbol{\kappa}_2 - \boldsymbol{\kappa}_2\mathbf{H}\boldsymbol{\kappa}_2) + O(\|\boldsymbol{\kappa}_2\mathbf{H}\boldsymbol{\kappa}_2\|) \quad \text{and} \quad (17)$$

$$\mathbf{\Pi}_1^a = \mathbf{P}^a + \mathbf{L}(\boldsymbol{\epsilon}_1 - \boldsymbol{\epsilon}_2\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}} - \boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_1^{\mathrm{T}} + \boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}\mathbf{H}^T\mathbf{K}^T) + O(\|\boldsymbol{\kappa}_{1,2}\boldsymbol{\Theta}\boldsymbol{\kappa}_{1,2}^{\mathrm{T}}\|), \quad (18)$$

(see proof in appendix A). When comparing (18) with (10), we can see that the term $\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_1^{\mathrm{T}}$ may explain the inbreeding effect suggested by Houtekamer and Mitchell (1998). When assuming the two sets of members to be independent, $\langle -\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_1^{\mathrm{T}}\rangle = \mathbf{0}_m$ and the average effect disappears, as was already predicted by van Leeuwen (1999). Van Leeuwen (1999) showed that the error term in (18) is partly due to the errors in the ensemble covariance in combination with those in the gain, and partly due to the nonlinearity of the gain itself. The first type of errors, represented by the term $\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_1^{\mathrm{T}}$, leads to inbreeding, while the latter represented by the term $\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}$ are of different sign and partly compensate the inbreeding effect. In the DEnKF the inbreeding part disappears, but the second effect is still present. Then, taking into account the other terms in (16), and averaging leads to (see appendix E)

$$\langle\mathbf{K}_2\rangle = \mathbf{K} - \mathbf{L}\langle\boldsymbol{\kappa}_2\mathbf{H}\boldsymbol{\kappa}_2\rangle + O(\|\boldsymbol{\kappa}_2\mathbf{H}\boldsymbol{\kappa}_2\|^2) \quad \text{and} \quad (19)$$

$$\langle\mathbf{P}_1^a\rangle = \mathbf{P}^a + \mathbf{L}\langle\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}\rangle\mathbf{L}^T + O(\|\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}\|). \quad (20)$$

Equivalent and symmetric expressions can be written for $\langle\mathbf{K}_1\rangle$ and $\langle\mathbf{P}_2^a\rangle$. By comparing this expression with (12), one can see that the sign of the error with respect to the optimal analysis covariance matrix $\mathbf{P}^a$ is reversed. Hence, by cutting the ensemble in two, one reverses the systematic bias (with respect to optimality) of the analysis error covariance matrix from negative to positive. On the other hand, the gain still has a systematic negative bias. For the 2DEnKf, each subensemble has $N/2$ members. Hence, using (B1), it follows that

$$\mathrm{tr}(\langle\mathbf{P}_1^a\rangle) \approx \mathrm{tr}(\mathbf{P}^a) + \frac{2}{N-2}[\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})]. \quad (21)$$

In the more general case of an *l*DEnKf, each subensemble is of size N/l. Equation (20) remains valid with $\mathrm{tr}(\mathbf{L}\langle\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}\rangle\mathbf{L}^{\mathrm{T}}) = [l^{-1}(l-1)N - 1]^{-1}[\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})]$, leading to

$$\mathrm{tr}(\langle\mathbf{P}_p^a\rangle) \approx \mathrm{tr}(\mathbf{P}^a) + \frac{l}{(l-1)N-l}[\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})], \quad (22)$$

with $l \in [\![2, N]\!]$ and $p \in [\![1, l]\!]$. Then, the bigger l is, the closer to the optimal analysis variance $\mathbf{P}^a$ is.

Now, to evaluate the MSE $\langle\mathbf{\Delta}_N^a\rangle$ for the first subensemble, we note that the third term of (14) comes from the gain used to compute the analysis. With the 2DEnKf, it follows that

$$\langle\mathbf{\Delta}_1^a\rangle = \mathbf{P}^a + \frac{2}{N}\mathbf{P}^a + \left(1 + \frac{2}{N}\right)\mathbf{L}\langle\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}\rangle\mathbf{L}^{\mathrm{T}} + O(\|\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}\|) \quad \text{and} \quad (23)$$

$$\mathrm{tr}(\langle\mathbf{\Delta}_1^a\rangle) \approx \left(1 + \frac{2}{N}\right)\left\{\mathrm{tr}(\mathbf{P}^a) + \frac{2}{N-2}[\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})]\right\}. \quad (24)$$

Comparing (15) and (24), we see that the analysis is less accurate than with the EnKf. This means that criterion i is not satisfied. Then, comparing (21) and (24), we see that criterion ii is generally not satisfied either, because the average analysis variance underestimates the MSE of the ensemble mean. These two results are consistent with the simulation results of Houtekamer and Mitchell (1998).

When generalized to the *l*DEnKf, $l, p \in [\![2, N]\!] \times [\![1, l]\!]$, (24) becomes

$$\mathrm{tr}(\langle\mathbf{\Delta}_p^a\rangle) \approx \left(1 + \frac{l}{N}\right)\left\{\mathrm{tr}(\mathbf{P}^a) + \frac{l}{(l-1)N-l}[\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})]\right\}. \quad (25)$$

Because of the term $[l\ \mathrm{tr}(\mathbf{P}^a)]/N$, the accuracy can be substantially degraded. Moreover, the discrepancy between (22) and (25) increases with l. Consequently, a partitioned analysis is not likely to take full advantage of a multiple division of the ensemble. The analysis error variance remains smaller than the mean squared error of the ensemble mean.

We stress here that for small values of N, the second-order truncation is still a good one for $\mathrm{tr}\langle\mathbf{P}_1^a\rangle$, but not as precise as for the EnKf. Indeed, higher-order terms will not compensate because those containing products of $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ average to zero, whereas the others do not.

2) MERGED ANALYSIS

We define $\mathbf{P}_N^{a(1,2)}$ as the analysis error covariance matrix using the whole ensemble given by the 2DEnKf analysis (i.e., merging the two subensembles to get only one). It can be proven (see appendix F) that

$$\langle\mathbf{P}_N^{a(1,2)}\rangle = \mathbf{P}^a + \frac{3}{4}\mathbf{L}(\langle\boldsymbol{\kappa}_1\boldsymbol{\Theta}\boldsymbol{\kappa}_1^{\mathrm{T}}\rangle + \langle\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}\rangle)\mathbf{L}^{\mathrm{T}} + O(\|\boldsymbol{\kappa}_{1,2}\boldsymbol{\Theta}\boldsymbol{\kappa}_{1,2}^{\mathrm{T}}\|), \tag{26}$$

where the second and third terms look like the second term in (20), but weighted differently. Taking the trace leads to

$$\mathrm{tr}[\langle\mathbf{P}_N^{a(1,2)}\rangle] \approx \mathrm{tr}(\mathbf{P}^a) + \frac{3}{N-2}[\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})]. \tag{27}$$

The generalized version for an lDEnKf has the following form:

$$\mathrm{tr}[\langle\mathbf{P}_N^{a(1,l)}\rangle] \approx \mathrm{tr}(\mathbf{P}^a) + \frac{l+1}{(l-1)N-l}[\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})]. \tag{28}$$

We now define $\boldsymbol{\Delta}_N^{a(1,2)} = \boldsymbol{\xi}_t^a\boldsymbol{\xi}_t^{a\mathrm{T}}$ as the error of the mean calculated over the whole ensemble for the 2DEnKf. Considering $\overline{\psi^a} = 0.5(\overline{\psi_1^a} + \overline{\psi_2^a})$, and taking the average, it follows after some calculation (details can be found in appendix G) that

$$\langle\boldsymbol{\Delta}_N^{a(1,2)}\rangle = \left(1+\frac{1}{N}\right)\left[\mathbf{P}^a + \frac{1}{4}\mathbf{L}(\langle\boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^{\mathrm{T}}\rangle + \langle\boldsymbol{\kappa}_1\boldsymbol{\Theta}\boldsymbol{\kappa}_1^{\mathrm{T}}\rangle)\mathbf{L}^{\mathrm{T}}\right] + O(\|\boldsymbol{\kappa}_{1,2}\boldsymbol{\Theta}\boldsymbol{\kappa}_{1,2}^{\mathrm{T}}\|), \tag{29}$$

$$\mathrm{tr}[\langle\boldsymbol{\Delta}_N^{a(1,2)}\rangle] \approx \left(1+\frac{1}{N}\right)\left\{\mathrm{tr}(\mathbf{P}^a) + \frac{1}{N-2}[\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})]\right\}. \tag{30}$$

Recalling (15), one notices that for N is not too small $\mathrm{tr}[\langle\boldsymbol{\Delta}_N^{a(1,2)}\rangle] \approx \mathrm{tr}(\langle\boldsymbol{\Delta}_N^a\rangle)$. This means that cutting the ensemble in two has almost no effect on the precision of the filter when using the mean over the whole ensemble. This result can be generalized to the lDEnKf:

$$\mathrm{tr}[\langle\boldsymbol{\Delta}_N^{a(1,l)}\rangle] \approx \left(1+\frac{1}{N}\right)\left\{\mathrm{tr}(\mathbf{P}^a) + \frac{l-1}{(l-1)N-l} \times [\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})]\right\}. \tag{31}$$

By comparing with (28), we see that criterion ii can theoretically be verified if there exists an $l \in [\![2, N]\!]$ such that

$$l \approx 1 + 2\frac{\mathrm{tr}(\mathbf{P}^a\mathbf{LKH}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{KH})}{\mathrm{tr}(\mathbf{P}^a)}. \tag{32}$$

Nevertheless, the solution is problem dependent, and it is not clear how to find a general solution for the optimal value of l.

Figure 2 shows the results obtained with a scalar process and a 2DEnKf. One can see that in this case the analysis variance given by the merged analysis matches very well the error of the ensemble mean on average. This means that in this case ($k = 0.5$), 2 is the optimal value for l, a value effectively given by solving the scalar version of (32). On the other hand, the partitioned analysis (the first ensemble in this case) gives an analysis with a degraded accuracy, whose variance is still underestimating the MSE of the ensemble mean. However, as predicted above, it overestimates the optimal value. Furthermore, the experimental results match the theoretical curves quite closely, as in the case of the EnKf. This is a good indication of the accuracy of the approximations made. Higher values of l have also been tested, namely $l = 4$ and $l = 8$. The results shown in Figs. 3 and 4 are also quite consistent with the theory. Figure 5 shows the simulation for the extreme case of an NDEnKf. As predicted, the analysis variance underestimates the MSE of the ensemble mean. The error with respect to the optimal value is almost exactly equal and opposite to the one given by the EnKf.

We see that the use of a DEnKf is a very good alternative to the EnKf. It can give an analysis that satisfies criteria i and ii, which is impossible with the EnKf. Nevertheless, it appears that the improvement is not as great when one divides the ensemble in more than two in this case.

b. Optimal covariance inflation

To correct for systematic sampling errors in $\mathbf{P}_N^a$ and to satisfy criterion ii, some authors multiply the forecast

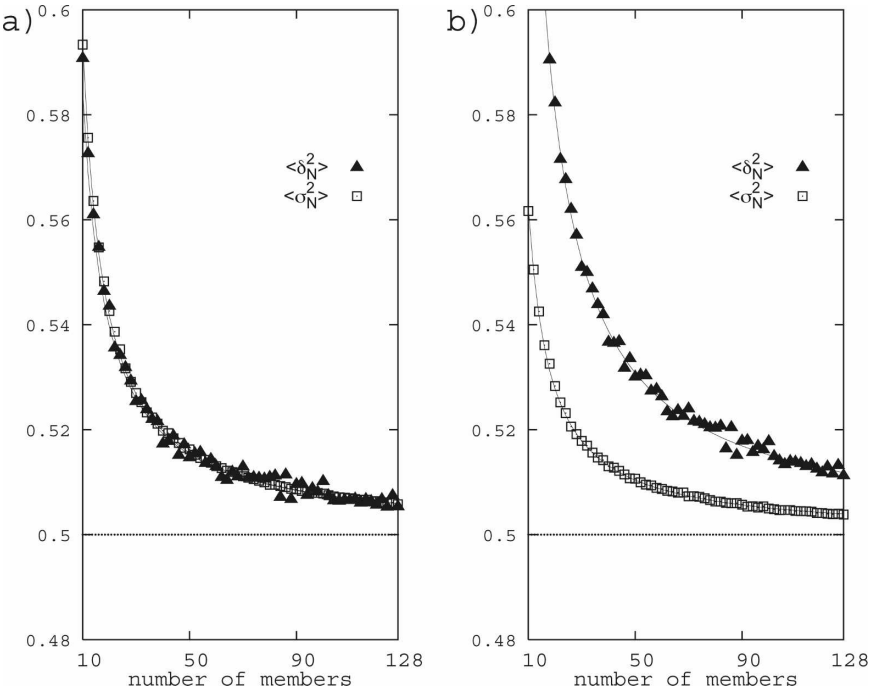

FIG. 2. Same simulation as in Fig. 1 but for a 2DEnKf. (a) Statistics for the merged analysis. The scalar versions of (27) and (30) are drawn in continuous lines. (b) Statistics for the analysis of the first ensemble. The scalar versions of (21) and (24) are drawn in continuous lines.

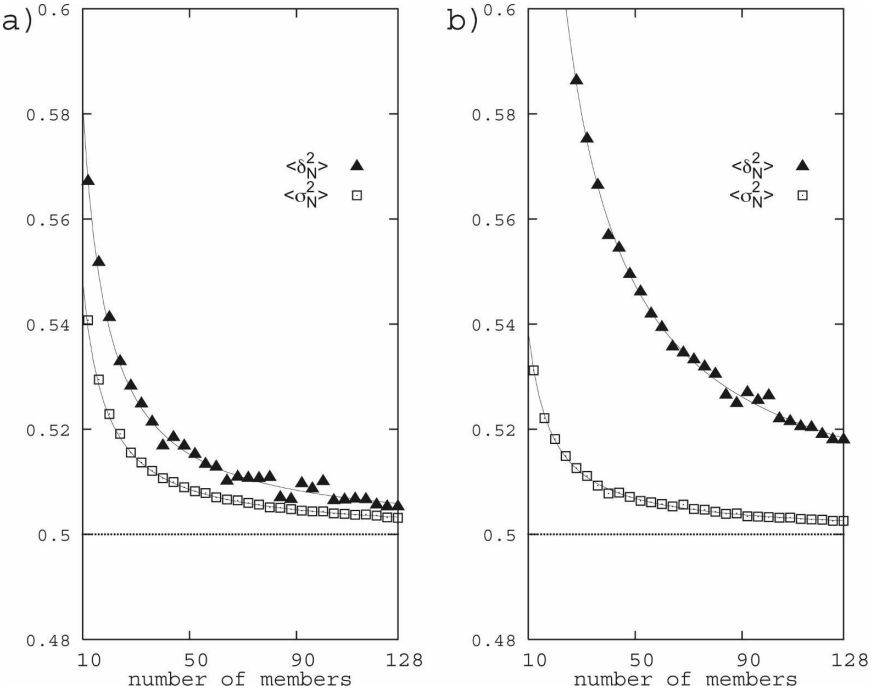

FIG. 3. Same simulation as in Fig. 1 but for a 4DEnKf. (a) Statistics for the merged analysis. The scalar versions of (28) and (31) for $l = 4$ are drawn in continuous lines. (b) Statistics for the analysis of the first ensemble. The scalar versions of (22) and (25) for $l = 4$ are drawn in continuous lines.

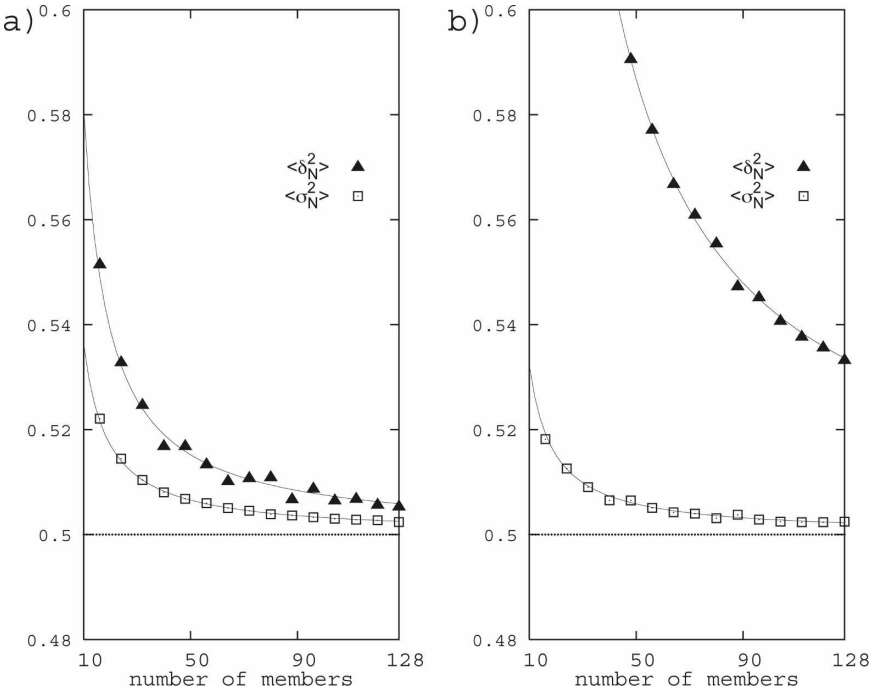

FIG. 4. Same simulation as in Fig. 1 but for an 8DEnKf. (a) Statistics for the merged analysis. The scalar versions of (28) and (31) for $l = 8$ are drawn in continuous lines. (b) Statistics for the analysis of the first ensemble. The scalar versions of (22) and (25) for $l = 8$ are drawn in continuous lines.

or analysis covariances by an inflation factor r. To properly satisfy criterion ii, we have seen that this factor is necessarily dependent on the ensemble size. Instead of using a blindly chosen factor, we propose here two inflation techniques, namely inflation of the deviations from the background before analysis, and addition of random errors after analysis.

The analyses given by the two methods on average satisfy criteria i and ii. They do not lead to underestimated error covariances. We therefore expect these analyses to be less subject to the problem of filter divergence.

We follow here Hamill et al. (2001), inflating the deviations from the ensemble-mean first guess by a constant factor r for each member of the ensemble at the end of the propagation step, that is, $\widetilde{\psi_j^f} = \psi_j^f + (r - 1)(\psi_j^f - \overline{\psi^f})$, implying $\boldsymbol{\xi}_{j^a}^{\widetilde{a}} = \boldsymbol{\xi}_{j^a}^{a} + (r - 1)\mathbf{L}_N \boldsymbol{\xi}_{j^f}^{f}$, which is equivalent to a multiplication of the background error covariance matrix $\mathbf{P}_N^f$ by r^2. This inflation technique does not affect the forecast and analysis ensemble means; therefore, $\langle \widetilde{\boldsymbol{\Delta}_N^a} \rangle = \langle \boldsymbol{\Delta}_N^a \rangle$. The average of the new mean analysis covariance matrix is

$$\langle \widetilde{\mathbf{P}_N^a} \rangle = \langle \mathbf{P}_N^a \rangle + (r^2 - 1)\langle \mathbf{P}_N^a \mathbf{L}_N^{\mathrm{T}} \rangle$$

Satisfying criterion ii requires $\mathrm{tr}(\langle \widetilde{\mathbf{P}_N^a} \rangle) \approx \mathrm{tr}(\langle \widetilde{\boldsymbol{\Delta}_N^a} \rangle)$; that is,

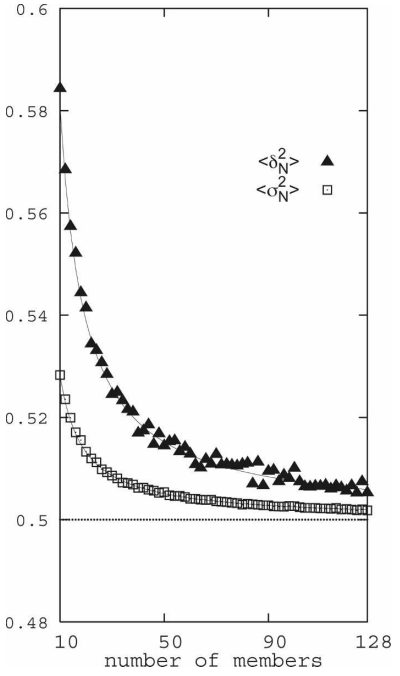

FIG. 5. Same simulation as in Fig. 1 but for an NDEnKf, N being the number of ensemble members. The scalar versions of (28) and (31) for $l = N$ are drawn in continuous lines.

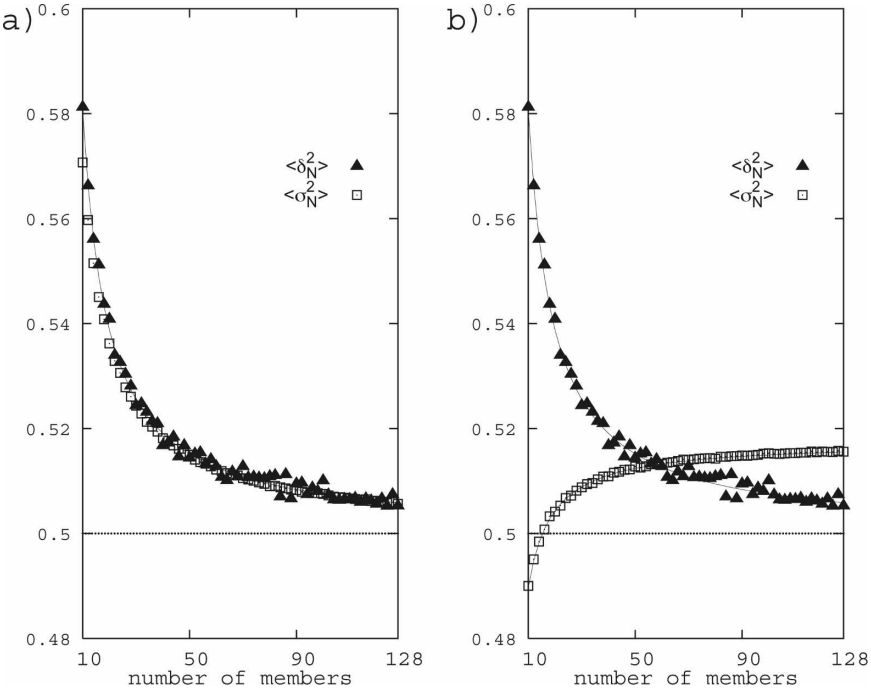

FIG. 6. Same simulation as in Fig. 1 but for a sampling-error-corrected EnKf, using an inflation factor. (a) The inflation coefficient r_N, which depends on the number of ensemble members, has been used. (b) The constant value $r = 1.035$ has been used. The curves match closely the corresponding analytical curve $\langle \sigma_a^2(N) \rangle = \sigma_a^2[1.035 - 2k(1-k)(N-1)^{-1}]$, represented by the continuous line.

$$r \approx \left\{ 1 + \frac{1}{N} \frac{\mathrm{tr}(\mathbf{P}^a)}{\mathrm{tr}(\langle \mathbf{L}_N \mathbf{P}^a_N \rangle)} + \frac{\left(2 + \frac{1}{N}\right)[\mathrm{tr}(\mathbf{P}^a \mathbf{L}\mathbf{K}\mathbf{H}) + \mathrm{tr}(\mathbf{P}^a \mathbf{L})\mathrm{tr}(\mathbf{K}\mathbf{H})]}{(N-1)\mathrm{tr}(\langle \mathbf{L}_N \mathbf{P}^a_N \rangle)} \right\}^{1/2}, \tag{33}$$

which can be approximated by judiciously using the value $\mathbf{P}^a_N$, $\mathbf{L}_N$, $\mathbf{K}_N$ as well as an expression for $\langle \mathbf{L}_N \mathbf{P}^a_N \rangle$ similar to (12). In the scalar case, r_N takes the simple form

$$r_N \approx \left[1 + \frac{1}{N(1-k)} + \frac{2k\left(2 + \frac{1}{N}\right)}{(N-1)} \right]^{1/2}. \tag{34}$$

This expression gives a "flow dependent" covariance inflation, which is a priori different for each analysis cycle. In Fig. 6a, we see that the application of (34) at $k = \frac{1}{2}$ gives an analysis that satisfies criterion ii on average. Figure 6b shows that a constant value for r is optimal for one particular value of N. Figure 7 displays isolines of r given by (34). Typical inflation values of a few percent are observed for $k \leq 0.8$. This is consistent with the values used by, for example, Anderson and Anderson (1999) and Whitaker and Hamill (2002).

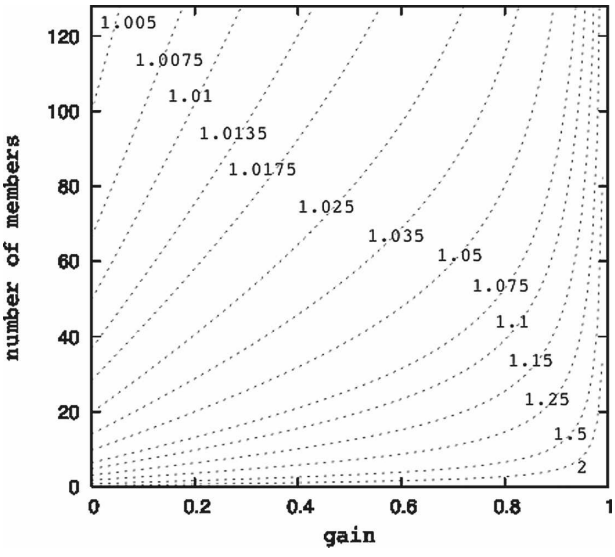

FIG. 7. Isolines of the scalar version of the inflation coefficient r_N, as a function of the number of ensemble members and the gain k.

Instead of using the inflation factor r, some authors (Gillijns et al. 2005; Corazza et al. 2002) suggest perturbing the analysis by adding a random vector, in the way that perturbed observations are added to increase the analysis variance in Burgers et al. (1998). Added vectors can be chosen such that criterion ii is satisfied. Then, we suggest

$$\widetilde{\psi}_j^a = \psi_j^a + \frac{\mu_j}{N^{1/2}}\boldsymbol{\xi}_{j^a}^a + \nu_j\left(\frac{2+\frac{1}{N}}{N-1}\right)^{1/2}\mathbf{L}_{\mathcal{N}}\{\mathbf{K}_{\mathcal{N}}(\boldsymbol{\delta}_j - \mathbf{H}\boldsymbol{\xi}_{j^f}^f) + [\mathrm{tr}(\mathbf{K}_N\mathbf{H})]^{1/2}\boldsymbol{\xi}_{j^f}^f\}, \quad (35)$$

where μ, υ are two independent random scalars of zero mean and unit variance. Then, we have $\overline{\widetilde{\psi}^a} = \overline{\psi^a}$; that is to say $\langle\widetilde{\boldsymbol{\Delta}}_N^a\rangle = \langle\boldsymbol{\Delta}_N^a\rangle$. With this formulation, we have $\langle\widetilde{\boldsymbol{\Delta}}_N^a\rangle \approx \langle\widetilde{\mathbf{P}}_N^a\rangle$ and therefore criterion ii is satisfied. Simulations applying (35) to the scalar problem (not shown) yield a similar quality as in Fig. 6. This formulation may be easier to implement than the covariance inflation factor because of the absence of additive matrix calculations.

6. Conclusions

Theoretical expressions have been established for the analysis error covariance matrices and the mean squared error of the ensemble mean given by various versions of a perturbed-observation EnKf. The performance of each version is examined by verifying whether it produces a sample with an ensemble mean having the smallest error (a requirement termed criterion i) and if the MSE of the ensemble mean is well represented by the spread drawn from the ensemble of analyses given by the filter (termed criterion ii). For the EnKf used with a finite ensemble, our results show that it never satisfies criterion ii fully. That is, the analysis spread among the ensemble members always underestimates the mean squared difference between the ensemble mean and the true state and the filter is therefore naturally subject to divergence. The possibility for the double-ensemble Kalman filter technique to properly correct this problem was studied. We showed that an *l*DEnKf (i.e., cutting the ensemble in l subensembles of size N/l) can potentially give adequate analyses when using the ensemble resulting from the merging of all the subensembles. Nevertheless, the best value of l is problem dependent, typically depending on the observation density and frequency. In addition, using an *l*DEnKf requires the computation of l Kalman gains, whose cost may be prohibitive. For the case of a partitioned *l*DEnKf, that is using each subensemble independently, we showed that the analysis never satisfies criterion ii, though performing relatively better than the EnKf for small values of l, but at the expense of a degradation of the accuracy of the ensemble mean (criterion i is not satisfied) whose level increases with l. To correct for the misrepresentation problems and satisfy criteria i and ii on average, we presented flow-dependent techniques such as an optimal covariance inflation, and a randomly perturbed analysis. They are motivated by the present theoretical analysis, which takes into account the dependence of criterion ii on the number of members and the Kalman gain. We expect these methods to be less subject to the problem of filter divergence. It should be stressed here that all of the calculations have been done in an idealized context, ignoring model or representativity errors. This can virtually invalidate some of our results and would affect our theoretical value for the inflation factor r. It is also important to keep in mind that the analytic results concern the means over a large number of realizations of analyses. For one given analysis, they might not be so relevant—especially if the variance of the squared error of the ensemble mean is relatively high. In NWP, as one works with state spaces of high dimensions, we are confident that the averages will have some significance. We should also stress that the whole theory developed here follows the traditional assumption of a Gaussian process. As a result, we may expect correction to r in a strongly nonlinear context. Nevertheless, we expect our results to better guide the user in the determination of the r factor in an operational context, previously chosen by trial and error (Anderson and Anderson 1999). We have performed simulations using a very nonlinear two-dimensional turbulence model that have led to very encouraging results regarding the capacity of these techniques to avoid filter divergence. They will be presented in the forthcoming Part II of this paper.

Acknowledgments. This work would not have been possible without the contributions of Peter Houtekamer and Herschel Mitchell, with whom we had hours of stimulating and fruitful discussions. We thank two anonymous reviewers for their very helpful and productive comments. We acknowledge financial support from the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) in the form of Grant GR-500B.

APPENDIX A

Taylor Expansion of $\mathbf{K}_N$ and $\boldsymbol{\Pi}_N^a$ in the EnKf

Substituting $\mathbf{P}_N^f = \mathbf{P}^f + \boldsymbol{\epsilon}$ in (5) leads to $(\mathbf{I}_m + \mathbf{H}\boldsymbol{\kappa})_N = (\mathbf{P}^f + \boldsymbol{\epsilon})\mathbf{H}^{\mathrm{T}}\boldsymbol{\Phi}\,(\mathbf{I}_m + \mathbf{H}\boldsymbol{\kappa})^{-1}$. We suppose $\|\mathbf{H}\boldsymbol{\kappa}\| \ll$

$\|\mathbf{I}_m\|$, then $\rho(\mathbf{H}\boldsymbol{\kappa}) < 1$ (where ρ is the spectral radius), and the Taylor expansion for $(\mathbf{I}_m + \mathbf{H}\boldsymbol{\kappa})^{-1}$ is

$$(\mathbf{I} + \mathbf{H}\boldsymbol{\kappa})^{-1} = \mathbf{I} - \mathbf{H}\boldsymbol{\kappa} + (\mathbf{H}\boldsymbol{\kappa})^2 + \cdots + (-1)^M(\mathbf{H}\boldsymbol{\kappa})^M + O(\|\mathbf{H}\boldsymbol{\kappa}\|^M),$$

where $\|\ \|$ is the matrix norm of interest. It leads to the following expressions for $\mathbf{K}_N$ and $\boldsymbol{\Pi}_N^a$:

$$\mathbf{K}_N = \mathbf{K} + \mathbf{L}\sum_{k=0}^{M}(-1)^k\boldsymbol{\kappa}(\mathbf{H}\boldsymbol{\kappa})^k + O[\|(\boldsymbol{\kappa}\mathbf{H})^M\boldsymbol{\kappa}\|] \quad \text{(A1)}$$

and

$$\boldsymbol{\Pi}_N^a = \mathbf{L}\mathbf{P}^f + \mathbf{L}\left[\boldsymbol{\epsilon} - \sum_{k=0}^{M-1}(-1)^k(\boldsymbol{\kappa}\mathbf{H})^k\boldsymbol{\kappa}\boldsymbol{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}\right]\mathbf{L}^{\mathrm{T}} + O[\|(\boldsymbol{\kappa}\mathbf{H})^M\boldsymbol{\epsilon}\|]. \quad \text{(A2)}$$

For the DEnKf, we have $\mathbf{K}_2 = \mathbf{P}_2^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_2^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$ and $\boldsymbol{\Pi}_1^a = \mathbf{L}_2\mathbf{P}_1^f$, and then

$$\boldsymbol{\Pi}_1^a = \mathbf{L}\mathbf{P}^f + \mathbf{L}\left[\boldsymbol{\epsilon}_1 - \boldsymbol{\epsilon}_2\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}} - \sum_{k=0}^{M-1}(-1)^k(\boldsymbol{\kappa}_2\mathbf{H})^k\boldsymbol{\kappa}_2\boldsymbol{\Theta}(\boldsymbol{\kappa}_1^{\mathrm{T}} - \boldsymbol{\kappa}_2^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}})\right] + O[\|(\boldsymbol{\kappa}_2\mathbf{H})^M\boldsymbol{\epsilon}_{1,2}\|]. \quad \text{(A3)}$$

Here, $\mathbf{K}_2$ (respectively, $\mathbf{K}_1$) has the same expression as $\mathbf{K}_N$, replacing $\boldsymbol{\kappa}$ by $\boldsymbol{\kappa}_2$ (respectively, $\boldsymbol{\kappa}_1$). The expression for $\boldsymbol{\Pi}_2^a = \mathbf{L}_1\mathbf{P}_2^f$ is obtained by symmetry.

APPENDIX B

Calculation of $\mathrm{tr}(\langle\mathbf{L}\boldsymbol{\kappa}\boldsymbol{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\rangle)$ for a Gaussian Multivariate Process

We know (e.g., Maybeck 1979, p. 130) that $\mathrm{var}(\sigma^2 - s^2) = 2\sigma^4(N-1)^{-1}$ for a Gaussian scalar process, if σ^2 and s^2 are the true and the unbiased estimated variances. To generalize this result to the multivariate case, as was done in Furrer and Bengtsson (2007), we use the spectral decomposition of $\mathbf{P}^f = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^{\mathrm{T}}$, where $\boldsymbol{\Lambda}$ is an $n \times n$ diagonal matrix with terms $\lambda_i i \in [\![1, n]\!]$, and define $\boldsymbol{\Lambda}_N = \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}_N^f\boldsymbol{\Gamma}$ (in general not diagonal) as an unbiased estimator of $\boldsymbol{\Lambda}$. For any matrix $\mathbf{A}$ and $\mathbf{B}$, we have

$$\langle\boldsymbol{\epsilon}\mathbf{A}\mathbf{B}\boldsymbol{\epsilon}\rangle = \langle\boldsymbol{\Gamma}(\boldsymbol{\Lambda}_N - \boldsymbol{\Lambda})\mathbf{A}\mathbf{B}(\boldsymbol{\Lambda}_N - \boldsymbol{\Lambda})^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\rangle = \boldsymbol{\Gamma}\langle\boldsymbol{\Lambda}_N\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{A}\mathbf{B}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_N\rangle\boldsymbol{\Gamma}^{\mathrm{T}} - \mathbf{P}^f\mathbf{A}\mathbf{B}\mathbf{P}^f.$$

We consider now $(\mathbf{X})_{i,j}$ as the general term of $\mathbf{X}$ and define $(\boldsymbol{\Gamma})_{i,j} = \gamma_{i,j}$, $(\boldsymbol{\Gamma}^{\mathrm{T}})_{i,j} = \gamma'_{i,j}$, $(\mathbf{A})_{i,j} = a_{i,j}$, $(\mathbf{B})_{i,j} = b_{i,j}$, $(\mathbf{P}_N^f)_{i,j} = \widehat{p_{i,j}}$, $(\mathbf{P}^f)_{i,j} = p_{i,j}$, $(\boldsymbol{\Lambda})_{i,j} = \lambda_{i,j}$, $(\boldsymbol{\Lambda}_N)_{i,j} = \widehat{\lambda_{i,j}}$, and $(\boldsymbol{\Lambda}_N\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{A}\boldsymbol{\Gamma})_{ij}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{B}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_N)_{kl} = f_{i,j,k,l}$. This leads to

$$\langle f_{i,j,k,l}\rangle = \sum_{t,p,q,a,b,u,c,d,r,s}\gamma'_{i,q}\gamma_{p,t}\gamma'_{t,a}a_{a,b}\gamma_{b,j}\gamma'_{k,c}b_{c,d}\gamma_{d,u}\gamma'_{u,r}\gamma_{s,l} \times\langle\widehat{p_{q,p}}\widehat{p_{r,s}}\rangle \quad \text{and}$$

$$\langle f_{i,j,k,l}\rangle = (\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{P}^f\mathbf{A}\mathbf{B}\mathbf{P}^f\boldsymbol{\Gamma})_{i,l} + \frac{1}{N-1}[(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}^f\mathbf{B}^{\mathrm{T}}\boldsymbol{\Gamma})_{j,k}\lambda_{i,l} + (\boldsymbol{\Lambda}\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\boldsymbol{\Gamma})_{i,k}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Lambda})_{j,k}],$$

where we have used $\langle\widehat{p_{q,p}}\widehat{p_{r,s}}\rangle = p_{q,p}p_{r,s} + (N-1)^{-1}(p_{q,s}p_{p,r} + p_{q,r}p_{p,s})$. This leads to

$$\langle\boldsymbol{\epsilon}\mathbf{A}\mathbf{B}\boldsymbol{\epsilon}\rangle = \frac{1}{N-1}[\mathbf{P}^f\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}^f + \mathrm{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{P}^f\mathbf{B}^{\mathrm{T}})\mathbf{P}^f]$$

and

$$\mathrm{tr}(\langle\mathbf{L}\boldsymbol{\kappa}\boldsymbol{\Theta}\boldsymbol{\kappa}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\rangle) = \frac{1}{N-1}[\mathrm{tr}(\mathbf{P}^a\mathbf{L}\mathbf{K}\mathbf{H}) + \mathrm{tr}(\mathbf{P}^a\mathbf{L})\mathrm{tr}(\mathbf{K}\mathbf{H})], \quad \text{(B1)}$$

where we have replaced $\mathbf{A}$ by $\mathbf{I}$ and $\mathbf{B}$ by $\mathbf{H}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{H}$, and have used the trace properties of a product of matrices.

APPENDIX C

Calculation of $\boldsymbol{\Delta}_N^f$ for a Gaussian Process

We consider $\boldsymbol{\Delta}_N^f = \boldsymbol{\xi}_t^f\boldsymbol{\xi}_t^{f^{\mathrm{T}}}$, where $\boldsymbol{\xi}_t^f = \psi^t - \overline{\psi^f}$, and introduce $\psi^\mu = E\{\psi\} \approx \langle\psi\rangle$, the expectation value of the process (or the value obtained when the ensemble is of infinite size). Then, if $\boldsymbol{\xi}_t^\mu = \psi^t - \psi^\mu$, and $\boldsymbol{\xi}_\mu^f = \psi^\mu - \overline{\psi^f}$, it follows that $\boldsymbol{\Delta}_N^f = (\boldsymbol{\xi}_t^\mu + \boldsymbol{\xi}_\mu^f)(\boldsymbol{\xi}_t^\mu + \boldsymbol{\xi}_\mu^f)^{\mathrm{T}}$. Taking the average yields $\langle\boldsymbol{\Delta}_N^f\rangle = \langle\boldsymbol{\xi}_t^\mu\boldsymbol{\xi}_t^{\mu^{\mathrm{T}}}\rangle + \langle\boldsymbol{\xi}_t^\mu\boldsymbol{\xi}_\mu^{f^{\mathrm{T}}}\rangle + \langle\boldsymbol{\xi}_\mu^f\boldsymbol{\xi}_t^{\mu^{\mathrm{T}}}\rangle + \langle\boldsymbol{\xi}_\mu^f\boldsymbol{\xi}_\mu^{f^{\mathrm{T}}}\rangle$. Considering that the truth ψ^t is independent of each realization, and that its deviation with respect to the expectation value is independent of the deviation of each realization with respect to the expectation value, we have $\langle\boldsymbol{\xi}_t^\mu\boldsymbol{\xi}_\mu^{f^{\mathrm{T}}}\rangle = \langle\boldsymbol{\xi}_\mu^f\boldsymbol{\xi}_t^{\mu^{\mathrm{T}}}\rangle = 0_n$, and, finally, for a Gaussian process,

$$\langle\boldsymbol{\Delta}_N^f\rangle = \mathbf{P}^f + \frac{\mathbf{P}^f}{N}. \quad \text{(C1)}$$

APPENDIX D

Taylor Expansion for $\boldsymbol{\Delta}_N^a$

Taking the mean of (4), and providing $\mathbf{d} - \mathbf{H}\overline{\boldsymbol{\psi}^f} = \mathbf{H}\boldsymbol{\xi}_t^f - \boldsymbol{\delta}_t$, with $\boldsymbol{\delta}_t = \mathbf{H}\boldsymbol{\psi}^t - \mathbf{d}$, it follows that $\boldsymbol{\Delta}_N^a = [\boldsymbol{\xi}_t^f - \mathbf{K}_N(\mathbf{H}\boldsymbol{\xi}_t^f - \boldsymbol{\delta}_t)][\boldsymbol{\xi}_t^f - \mathbf{K}_N(\mathbf{H}\boldsymbol{\xi}_t^f - \boldsymbol{\delta}_t)]^{\mathrm{T}}$. Considering $\langle\boldsymbol{\xi}_t^f\boldsymbol{\delta}_t^{\mathrm{T}}\rangle = \mathbf{0}_{nm}$ leads to

$$\langle \boldsymbol{\Delta}_N^a \rangle = \langle \boldsymbol{\xi}_t^f \boldsymbol{\xi}_t^{f\mathrm{T}} - \boldsymbol{\xi}_t^f \boldsymbol{\xi}_t^{f\mathrm{T}} \mathbf{H}^\mathrm{T} \mathbf{K}_N^\mathrm{T} - \mathbf{K}_N \mathbf{H} \boldsymbol{\xi}_t^f \boldsymbol{\xi}_t^{f\mathrm{T}} + \mathbf{K}_N (\boldsymbol{\delta}_t \boldsymbol{\delta}_t^\mathrm{T} + \mathbf{H} \boldsymbol{\xi}_t^f \boldsymbol{\xi}_t^{f\mathrm{T}} \mathbf{H}^\mathrm{T}) \mathbf{K}_N^\mathrm{T} \rangle.$$

We assume that the background and observation errors are not correlated with the sampling error $\boldsymbol{\epsilon}$ (e.g., $\langle \boldsymbol{\kappa} \boldsymbol{\delta}_t \boldsymbol{\delta}_t^\mathrm{T} \rangle = \langle \boldsymbol{\kappa} \rangle \langle \boldsymbol{\delta}_t \boldsymbol{\delta}_t^\mathrm{T} \rangle$, $\langle \boldsymbol{\kappa} \boldsymbol{\xi}_t^f \boldsymbol{\xi}_t^{f\mathrm{T}} \rangle = \langle \boldsymbol{\kappa} \rangle \langle \boldsymbol{\Delta}_N^f \rangle$). We recall that $\langle \boldsymbol{\kappa} \rangle = \mathbf{0}_m$. We use (C1) and $\langle \boldsymbol{\delta}_t \boldsymbol{\delta}_t^\mathrm{T} \rangle = R(1 + N^{-1})$. Then, replacing $\mathbf{K}_N$ by the expression found in (A1), and noticing that sums cancel out each other because of the identity $\mathbf{KR} = \mathbf{LP}^f \mathbf{H}^\mathrm{T}$, we finally have

$$\langle \boldsymbol{\Delta}_N^a \rangle = \left(1 + \frac{1}{N}\right) [\mathbf{P}^a + \mathbf{L} \sum_{k=0}^{M-1} \sum_{p=0}^{M-1} (-1)^{k+p} \boldsymbol{\kappa} (\mathbf{H}\boldsymbol{\kappa})^k \times \boldsymbol{\Theta} (\boldsymbol{\kappa}^\mathrm{T} \mathbf{H}^\mathrm{T})^p \boldsymbol{\kappa}^\mathrm{T} \mathbf{L}^\mathrm{T}] + O[\|(\boldsymbol{\kappa}\mathbf{H})^M \boldsymbol{\epsilon}\|]. \quad \text{(D1)}$$

APPENDIX E

Taylor Expansion for $\mathbf{P}_N^a$ in the Partitioned DEnKf Method

In (16), the last three terms average to zero. Then, using expressions for $\mathbf{K}_1$ and $\mathbf{K}_2$ given by (17) and using $(\mathbf{H}\boldsymbol{\kappa}\boldsymbol{\Theta})^\mathrm{T} = \boldsymbol{\Theta}\boldsymbol{\kappa}^\mathrm{T}\mathbf{H}^\mathrm{T}$ in the second term leads to

$$(\mathbf{K}_2 - \mathbf{K}_1)(\mathbf{I}_m + \mathbf{H}\kappa_1)\boldsymbol{\Theta}\mathbf{K}_2^\mathrm{T} = \mathbf{L}(\boldsymbol{\kappa}_2 - \boldsymbol{\kappa}_1)\boldsymbol{\Theta}(\mathbf{K}^\mathrm{T} + \boldsymbol{\kappa}_1^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{K}^\mathrm{T} + \boldsymbol{\kappa}_2^\mathrm{T}\mathbf{L}^\mathrm{T}) + \mathbf{L}(\boldsymbol{\kappa}_1\boldsymbol{\Theta}\boldsymbol{\kappa}_1^\mathrm{T} - \boldsymbol{\kappa}_2\boldsymbol{\Theta}\boldsymbol{\kappa}_2^\mathrm{T})\mathbf{H}^\mathrm{T}\mathbf{K}^\mathrm{T} + O(\|\kappa_2\boldsymbol{\Theta}\boldsymbol{\kappa}_{1,2}^\mathrm{T}\|).$$

Using this last expression, then (16) and (18), taking the mean over a large number of realizations, assuming that $\langle \overline{\boldsymbol{\xi}_j^f} \overline{\boldsymbol{\delta}_j^\mathrm{T}} \rangle = \mathbf{0}_{nm}$, $\langle \boldsymbol{\kappa}_1 \boldsymbol{\Theta} \boldsymbol{\kappa}_2^\mathrm{T} \rangle = \langle \boldsymbol{\rho}_1 \rangle = \mathbf{0}_m$, and finally that $\langle \boldsymbol{\kappa}\boldsymbol{\rho} \rangle = \mathbf{0}_m$, one obtains the following for $\mathbf{P}_1^a$:

$$\langle \mathbf{P}_1^a \rangle = \mathbf{P}^a + \mathbf{L}\langle \boldsymbol{\kappa}_2 \boldsymbol{\Theta} \boldsymbol{\kappa}_2^\mathrm{T} \rangle \mathbf{L}^\mathrm{T} + O(\|\boldsymbol{\kappa}_2 \boldsymbol{\Theta} \boldsymbol{\kappa}_2^\mathrm{T}\|).$$

APPENDIX F

Taylor Expansion for $\mathbf{P}_N^a$ in the Merged DEnKf Method

Isolating members coming from each ensemble and taking $\boldsymbol{\xi}_t^{a,p}$ as in appendix G yields $\mathbf{P}_N^{a(1,2)} = (N - 1)^{-1}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)$, with

$$\boldsymbol{\Sigma}_q = \sum_{j=(1-q)N/2+1}^{qN/2} \frac{1}{4} (\boldsymbol{\xi}_{j^a}^{a,1} \boldsymbol{\xi}_{j^a}^{a,1\mathrm{T}} + \boldsymbol{\xi}_{j^a}^{a,2} \boldsymbol{\xi}_{j^a}^{a,1\mathrm{T}} + \boldsymbol{\xi}_{j^a}^{a,1} \boldsymbol{\xi}_{j^a}^{a,2\mathrm{T}} + \boldsymbol{\xi}_{j^a}^{a,2} \boldsymbol{\xi}_{j^a}^{a,2\mathrm{T}}).$$

Expression of, for example, $\boldsymbol{\Sigma}_1$ can be obtained by considering that $\boldsymbol{\xi}_{j^a}^{a,p} \boldsymbol{\xi}_{j^a}^{a,l\mathrm{T}} = \boldsymbol{\xi}_{j^a}^{a,m} \boldsymbol{\xi}_{j^a}^{a,m\mathrm{T}} + \boldsymbol{\xi}_{a,p}^{a,m} \boldsymbol{\xi}_{a,l}^{a,m\mathrm{T}} - \boldsymbol{\xi}_{j^a}^{a,m} \boldsymbol{\xi}_{a,l}^{a,m\mathrm{T}} - \boldsymbol{\xi}_{a,p}^{a,m} \boldsymbol{\xi}_{j^a}^{a,m\mathrm{T}}$, and $\sum_{j=1}^{N/2} \boldsymbol{\xi}_{j^a}^{a,1} = 0$. This gives

$$\boldsymbol{\Sigma}_1 = \sum_{j=1}^{N/2} \boldsymbol{\xi}_{j^a}^{a,1} \boldsymbol{\xi}_{j^a}^{a,1\mathrm{T}} + \frac{N}{8} (\boldsymbol{\xi}_t^{a,2} \boldsymbol{\xi}_t^{a,2\mathrm{T}} + \boldsymbol{\xi}_t^{a,1} \boldsymbol{\xi}_t^{a,1\mathrm{T}} - \boldsymbol{\xi}_t^{a,1} \boldsymbol{\xi}_t^{a,2\mathrm{T}} - \boldsymbol{\xi}_t^{a,2} \boldsymbol{\xi}_t^{a,1\mathrm{T}}).$$

Then, taking the average, using (20), (23), and $\langle \boldsymbol{\xi}_t^{a,1} \boldsymbol{\xi}_t^{a,2\mathrm{T}} \rangle = \mathbf{P}^a + O(\|\boldsymbol{\kappa}_{1,2}\boldsymbol{\Theta}\boldsymbol{\kappa}_{1,2}\|)$ (see appendix G) gives

$$\langle \boldsymbol{\Sigma}_1 \rangle = \frac{1}{2}(N-1)\mathbf{P}^a \left(\frac{5N}{8} - 1\right) \mathbf{L}\langle \boldsymbol{\kappa}_2 \boldsymbol{\Theta} \boldsymbol{\kappa}_2 \rangle \mathbf{L}^\mathrm{T} + \frac{N}{8} \mathbf{L}\langle \boldsymbol{\kappa}_1 \boldsymbol{\Theta} \boldsymbol{\kappa}_1 \rangle \mathbf{L}^\mathrm{T} + O(\|\boldsymbol{\kappa}_{1,2} \boldsymbol{\Theta} \boldsymbol{\kappa}_{1,2}^\mathrm{T}\|),$$

and a symmetric expression for $\langle \boldsymbol{\Sigma}_2 \rangle$. Further, if we suppose $(N-1)^{-1}\langle \boldsymbol{\kappa}_i \boldsymbol{\Theta} \boldsymbol{\kappa}_i^\mathrm{T} \rangle = O(\|\boldsymbol{\kappa}_i \boldsymbol{\Theta} \boldsymbol{\kappa}_i^\mathrm{T}\|)$, then

$$\langle \mathbf{P}_N^{a(1,2)} \rangle = \mathbf{P}^a + \frac{3}{4} \mathbf{L}(\langle \boldsymbol{\kappa}_1 \boldsymbol{\Theta} \boldsymbol{\kappa}_1^\mathrm{T} \rangle + \langle \boldsymbol{\kappa}_2 \boldsymbol{\Theta} \boldsymbol{\kappa}_2^\mathrm{T} \rangle) \mathbf{L}^\mathrm{T} + O(\|\boldsymbol{\kappa}_{1,2} \boldsymbol{\Theta} \boldsymbol{\kappa}_{1,2}^\mathrm{T}\|). \quad \text{(F1)}$$

The generalized version for an *l*DEnKF has the following form:

$$\langle \mathbf{P}_N^{a(1,l)} \rangle = \mathbf{P}^a + \mathbf{L} \frac{l+1}{l^2(l-1)} \sum_{p=1}^{l} \sum_{q=1}^{l} \langle \boldsymbol{\kappa}_p' \boldsymbol{\Theta} \boldsymbol{\kappa}_q'^\mathrm{T} \rangle \mathbf{L}^\mathrm{T} + O(\|\boldsymbol{\kappa}_{[1,l]}' \boldsymbol{\Theta} \boldsymbol{\kappa}_{[1,l]}'^\mathrm{T}\|), \quad \text{(F2)}$$

where $\boldsymbol{\kappa}_p' = \boldsymbol{\epsilon}_p' \mathbf{H}^\mathrm{T} \boldsymbol{\Phi}$, $\boldsymbol{\epsilon}_p'$ is the error of a background error covariance matrix calculated over all the subensembles but the pth.

APPENDIX G

Taylor Expansion for $\boldsymbol{\Delta}_N^a$ in the Merged DEnKf Method

Considering $\overline{\psi^a} = 0.5(\overline{\psi_1^a} + \overline{\psi_2^a})$ and $\boldsymbol{\xi}_t^{a,p} = \psi^t - \overline{\psi_p^a}$, we have

$$\langle \boldsymbol{\Delta}_N^{a(1,2)} \rangle = \frac{1}{4} (\langle \boldsymbol{\xi}_t^{a,1} \boldsymbol{\xi}_t^{a,1\mathrm{T}} \rangle + \langle \boldsymbol{\xi}_t^{a,2} \boldsymbol{\xi}_t^{a,2\mathrm{T}} \rangle + \langle \boldsymbol{\xi}_t^{a,1} \boldsymbol{\xi}_t^{a,2\mathrm{T}} \rangle + \langle \boldsymbol{\xi}_t^{a,2} \boldsymbol{\xi}_t^{a,1\mathrm{T}} \rangle).$$

The first two terms of this sum are given by (23). By using the expression of $\overline{\psi_1^a}$ and $\overline{\psi_2^a}$, assuming observation and forecast error are uncorrelated, and $\langle \boldsymbol{\xi}_\mu^{f,1} \boldsymbol{\xi}_\mu^{f,2} \rangle = 0_n$, $\langle \boldsymbol{\xi}_\mu^{j^f} \boldsymbol{\xi}_\mu^{k^f} \rangle = 0_n$, $\forall (j,k) \in [\![1,N]\!]^2$ and $j \neq k$ ($\boldsymbol{\psi}^\mu$ being the expectation value of the process), it follows that

$$\langle \boldsymbol{\xi}_t^{a,1} \boldsymbol{\xi}_t^{a,2\mathrm{T}} \rangle = \langle \boldsymbol{\xi}_t^\mu \boldsymbol{\xi}_t^{\mu\mathrm{T}} - \boldsymbol{\xi}_t^\mu \boldsymbol{\xi}_t^{\mu\mathrm{T}} \mathbf{H}^\mathrm{T} \mathbf{K}_1^\mathrm{T} - \mathbf{K}_2 \mathbf{H} \boldsymbol{\xi}_t^\mu \boldsymbol{\xi}_t^{\mu\mathrm{T}} + \mathbf{K}_2 \boldsymbol{\delta}_t \boldsymbol{\delta}_t^\mathrm{T} \mathbf{K}_1^\mathrm{T} + \mathbf{K}_2 \mathbf{H} \boldsymbol{\xi}_t^\mu \boldsymbol{\xi}_t^{\mu\mathrm{T}} \mathbf{H}^\mathrm{T} \mathbf{K}_1^\mathrm{T} \rangle.$$

Then, we use (19) and the symmetric expression for $\mathbf{K}_1$, as well as $\langle \boldsymbol{\xi}_t^\mu \boldsymbol{\xi}_t^{\mu\mathrm{T}} \rangle = \mathbf{P}^f$ and $\langle \boldsymbol{\delta}_t \boldsymbol{\delta}_t^\mathrm{T} \rangle = \mathbf{R}$. We also assume

that $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ are uncorrelated, and as is done in appendix D, we assume that the background and observation errors are not correlated with $\boldsymbol{\epsilon}$. It follows that $\langle \boldsymbol{\xi}_t^{a,1} \boldsymbol{\xi}_t^{a,2^{\mathrm{T}}} \rangle = \langle \boldsymbol{\xi}_t^{a,2} \boldsymbol{\xi}_t^{a,1^{\mathrm{T}}} \rangle = \mathbf{P}^a + O(\|\boldsymbol{\kappa}_{1,2} \boldsymbol{\Theta} \boldsymbol{\kappa}_{1,2}^{\mathrm{T}}\|)$; therefore,

$$\langle \boldsymbol{\Delta}_N^{a(1,2)} \rangle = \left(1 + \frac{1}{N}\right) \left[\mathbf{P}^a + \frac{1}{4} \mathbf{L} (\langle \boldsymbol{\kappa}_2 \boldsymbol{\Theta} \boldsymbol{\kappa}_2^{\mathrm{T}} \rangle + \langle \boldsymbol{\kappa}_1 \boldsymbol{\Theta} \boldsymbol{\kappa}_1^{\mathrm{T}} \rangle) \mathbf{L}^{\mathrm{T}} \right] + O(\|\boldsymbol{\kappa}_{1,2} \boldsymbol{\Theta} \boldsymbol{\kappa}_{1,2}^{\mathrm{T}}\|). \tag{G1}$$

Generalization to the *l*DEnKf leads to

$$\langle \boldsymbol{\Delta}_N^{a(1,l)} \rangle = \left(1 + \frac{1}{N}\right) \left(\mathbf{P}^a + \mathbf{L} \frac{1}{l^2} \sum_{p=1}^{l} \sum_{q=1}^{l} \langle \boldsymbol{\kappa}_p'^{\mathrm{T}} \boldsymbol{\Theta} \boldsymbol{\kappa}_q'^{\mathrm{T}} \rangle \mathbf{L}^{\mathrm{T}} \right) + O(\|\boldsymbol{\kappa}'_{[1,l]} \boldsymbol{\Theta} \boldsymbol{\kappa}_{[1,l]}'^{\mathrm{T}}\|), \tag{G2}$$

where $\boldsymbol{\kappa}_p'$ is as in appendix F.

REFERENCES

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

——, and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.,* **127,** 2741–2758.

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

Charron, M., P. L. Houtekamer, and P. Bartello, 2006: Assimilation with an ensemble Kalman filter of synthetic radial wind data in anisotropic turbulence: Perfect model experiments. *Mon. Wea. Rev.,* **134,** 618–637.

Cohn, S., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan,* **75,** 257–288.

Corazza, M., E. Kalnay, D. J. Patil, E. Ott, J. Yorke, B. R. Hunt, I. Szunyogh, and M. Cai, 2002: Use of the breeding technique in the estimation of the background error covariance matrix for a quasi-geostrophic model. Preprints, *Symp. on Observations, Data Assimilation and Probabilistic Prediction,* Orlando, FL, Amer. Meteor. Soc., 6.4. [Available online at http://ams.confex.com/ams/pdfpapers/28755.pdf.]

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

Epstein, E., 1969: Stochastic dynamic prediction. *Tellus,* **21,** 739–759.

Evensen, G., 1992: Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *J. Geophys. Res.,* **97** (C11), 17 905–17 924.

——, 1994: Inverse methods and data assimilation in nonlinear ocean models. *Physica D,* **77,** 108–129.

——, 2007: *Data Assimilation: The Ensemble Kalman Filter.* Springer, 280 pp.

Farrell, B., and P. Ioannou, 2001: State estimation using a reduced-order kalman filter. *J. Atmos. Sci.,* **58,** 3666–3680.

Furrer, R., and T. Bengtsson, 2007: Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *J. Multivariate Anal.,* **98,** 227–255.

Gelb, A., 1974: *Applied Optimal Estimation.* The MIT Press, 382 pp.

Gillijns, S., D. S. Bernstein, and B. D. Moor, 2005: The reduced rank transform square root filter for data assimilation. *Proc. 14th IFAC Symp. on System Identification (SYSID-2006),* Newcastle, Australia, Int. Federation on Automatic Control, FrB2.5.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.,* **128,** 2905–2919.

——, J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.,* **129,** 2776–2790.

Hansen, J. A., and L. A. Smith, 2000: The role of operational constraints in selecting supplementary observations. *J. Atmos. Sci.,* **57,** 2859–2871.

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

——, and ——, 1999: Reply. *Mon. Wea. Rev.,* **127,** 1378–1379.

——, and ——, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129,** 123–137.

——, and ——, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.,* **131,** 3269–3289.

Keppenne, C., 2000: Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. *Mon. Wea. Rev.,* **128,** 1971–1981.

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.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177–1194.

Maybeck, P. S., 1979: *Stochastic Models, Estimation, and Control.* Vol. 1. Mathematics in Science and Engineering, Vol. 141-1, Academic Press, 442 pp.

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

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

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002: Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Wea. Rev.,* **130,** 103–114.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.,* **131,** 1663–1677.

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

van Leeuwen, P. J., 1999: Comment on "Data assimilation using an ensemble Kalman filter technique." *Mon. Wea. Rev.,* **127,** 1374–1377.

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

Wunsch, W., 1996: *The Ocean Circulation Inverse Problem.* Cambridge University Press, 458 pp.