Tellus **000**, 000–000 (0000) Printed June 14, 2010 (Tellus LaTeX style file v2.2)

Data assimilation with the Weighted Ensemble Kalman Filter

By Nicolas Papadakis[1], Etienne Mémin[2⋆], Anne Cuzol[3] and Nicolas Gengembre[2]
[1]*Barcelona Media/Av. Diagonal 177, 08017 Barcelona, Spain,*
[2]*INRIA, Campus de Beaulieu, 35042 Rennes cedex, France*
[3]*Université de Bretagne Sud, Campus de Tohannic BP 573, 56017 Vannes cedex, France*

(Manuscript received june 2009;)

ABSTRACT

In this paper, two data assimilation methods based on sequential Monte Carlo sampling are studied and compared: the ensemble Kalman filter and the particle filter. Each of these techniques has its own advantages and drawbacks. In this work we try to get the best of each method by combining them. The proposed algorithm, called the weighted ensemble Kalman filter, consists to rely on the Ensemble Kalman Filter updates of samples in order to define a proposal distribution for the particle filter that depends on the history of measurement. The corresponding particle filter reveals to be efficient with a small number of samples and does not rely anymore on the Gaussian approximations of the ensemble Kalman filter. The efficiency of the new algorithm is demonstrated both in terms of accuracy and computational load. This latter aspect is of the utmost importance in meteorology or in oceanography since in these domains, data assimilation processes involve a huge number of state variables driven by highly non-linear dynamical models. Numerical experiments have been performed on different dynamical scenarios. The performances of the proposed technique have been compared to the ensemble Kalman filter procedure, which has demonstrated to provide meaningful results in geophysical sciences.

1 Introduction

A major concern in earth sciences consists in predicting the future state of a set of variables characterizing atmospheric or oceanic flows. The accuracy of the forecast relies firstly on the quality of the physical model that describes the evolution of the state variables of the system but also on their initial state. The estimation of this initial state as a probabilistic inference problem, can be formulated as the estimation of an *a posteriori probability* distribution of the system state variables at a given time knowing an history of measurement data until that time. Formulated in this way, this estimation process referred as *analysis* in data assimilation community defines a stochastic filtering problem. In the following, the state variables at instant k will be represented by a vector $\mathbf{x}_k$ of dimension n. The sequence of measurements also called observations from time 1 to k will be denoted by a set of vectors of dimension m: $\mathbf{y}^o_{1:k} = \{\mathbf{y}^o_i,\, i \in \mathbb{N},\, 1 + i\Delta k \leqslant k\}$, where Δk is the latency between two successive measurements. Asynchrony between observation and analysis instants may arise in geophysical applications. They are due to observations availability or to specific applications requirements and have significant implications both on the method design and on the results (Fertig et al., 2007). Throughout this study, we will nevertheless consider that analysis is performed at measurement instants. The sought conditional distribution, called the filtering distribution, will be thus denoted $p(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k})$ and the corresponding conditional expectation will constitute the minimum variance estimate between the state variables trajectories at instant k and an history of observations until that time.

We will assume that the evolution of the state variables is described through a dynamical model of the form:

$$\mathbf{x}_t = M(\mathbf{x}_{t-\Delta t}) + \boldsymbol{\eta}_t, \tag{1}$$

where M is a deterministic function (which corresponds to the discrete scheme associated to a given numerical implementation of a physical conservation law with respect to a numerical integration time step Δt) and $\boldsymbol{\eta}_t$ is a white Gaussian noise with covariance $\mathbf{Q}_{\Delta t}$. The time step depends usually on the discrete scheme used for the state variables temporal derivative and is much smaller than the latency between two measurements – or two analysis instants – Δk.

The relation, at time k, between the state variables and the measurements will be assumed to be described by the following linear model:

$$\mathbf{y}^o_k = \mathbf{H}\mathbf{x}_k + \boldsymbol{\epsilon}_k, \tag{2}$$

where $\boldsymbol{\epsilon}_k$ is a white Gaussian noise with covariance $\mathbf{R}$.

When the dynamical model is linear, the corresponding filtering distribution is Gaussian and completely defined from its two first moments. The Kalman filter provides an

⋆ Corresponding author.
e-mail: etienne.memin@inria.fr

optimal analytic iterative formulation of their expression. When the dynamics is non-linear, extended formulations relying on linearization around the current estimates can be settled. This scheme thus has the inconvenience to rely on an approximate dynamics which appears to be conceptually problematic when dealing with geophysical models and forecasting applications. Furthermore, when applied to strongly non-linear dynamics the linearization generates instabilities that tend to make the filter diverge (Gauthier et al., 1993). Alternatives exist to deal with this problem. These schemes can be interpreted as Kalman filters or extended Kalman filters (although the strict equivalence is only valid for linear models) in which the update of the Kalman filter are simplified either considering constant error covariance and a functional minimization formulation (3Dvar or 4Dvar schemes (Courtier et al., 1998; Le Dimet and Talagrand, 1986)), or proceeding to windowing selection of observations (Optimal Interpolation (Lorenc, 1981)) or solving the assimilation problem in the observation space (3D/4D-PSASS (Bennet, 1992; Courtier, 1997)).

For highly non-linear dynamics, Evensen (Evensen, 1994) proposed a method based on a Monte Carlo sampling of the filtering law: the ensemble Kalman filter (EnKF)(see also (Anderson and Anderson, 1999; Bishop et al., 2001; Burgers et al., 1998; Evensen, 2003; 2006; Houtekamer and Mitchell, 1998; Whitaker and Hamill, 2002; Ott et al., 2004)). In this technique a set of hypothetic states (the ensemble members) are propagated through the dynamical model (prediction step) and corrected at a given time through a discrepancy between the forecast state and observations available at that time (correction step or analysis stage).

In Ensemble Kalman filtering the correction step is obtained through a Gaussian approximation of the predicted states distribution, so that classical formulas of Kalman filter may be used, and applied to each member of the set. These updating rules are obviously exact only when the filtering distribution remains Gaussian that is to say when the dynamics can be well approximated by a linear Gaussian model.

In parallel to these developments, Monte Carlo sequential non-linear filtering procedures, so called particle filters, have been proposed in signal processing for control and tracking problems (Arulampalam et al., 2002; Gordon et al., 1993; 2001). In this family of methods, the filtering law is sampled by hypothetic states (called particles) which are shifted in the prediction step and weighted in the corrective step. The particles plays formally the same role as the ensemble members used in the EnKF to sample the assumed Gaussian filtering distribution. In the prediction step, the particles are moved according to a sampling distribution (called the proposal distribution). The weights account for the deviation between the proposal distribution and the unknown true filtering distribution. These weights should not be confused with the coefficients used in some EnKF implementations where the analysis is performed in a reduced space and the state variable is expressed as a linear combination of the ensemble members (Hunt et al., 2007).

The central idea developed in this paper (that is suggested in (Bertino et al., 2003)) resides in the design of a proposal distribution defined from EnKF principles together with adapted updating rules for the particles weights. The goal sought here consists in defining a particle filter well suited to high dimensional assimilation problems with linear measurement model and that allows relaxing the Gaussian approximation assumption of the ensemble members in the ensemble Kalman filters. This idea is somehow reminiscent to the work done by (van der Merwe et al., 2001) to enhance the "unscented Kalman" filter through the particle filtering theory (see also (Julier and Uhlmann, 1997)).

After some necessary recalls about the two filtering methods, the EnKF and the particle filter, in section 2, our adaptation of the particle filter, called the weighted ensemble Kalman filter (WEnKF), is presented in section 3. In section 4, some experimental results on simple scalar scenarios (1-dimensional state space) are first presented. In section 5, special attention is paid to a high dimensional problem. The performances of EnKF and WEnKF are analyzed and compared on different scenarios.

2 Related data assimilation methods

In this section, the basic theory on ensemble Kalman filters and particle filters are recalled, and both methods are compared.

2.1 Kalman filter

In order to properly define the EnKF, we first remind the Kalman filter's formulas (Kalman, 1960; Kalman and Bucy, 1961) for a linear dynamics expressed by:

$$\mathbf{x}_k = \mathbf{M}\mathbf{x}_{k-1} + \boldsymbol{\eta}_k, \tag{3}$$

where $\mathbf{M}$ is a $n \times n$ matrix and $\boldsymbol{\eta}_k$ is a white Gaussian noise with covariance $\mathbf{Q}_k$. The linear measurement expressed by equation (2) is still assumed. The initialization of the filter is given by the distribution $p(\mathbf{x}_0)$, described by a Gaussian of mean $\mathbf{x}_0$ and covariance $\mathbf{P}_0$. The filter is updated at the successive instants k, through the two following steps:

- **Prediction:**

$$\mathbf{x}_k^f = \mathbf{M}\mathbf{x}_{k-1}^a, \tag{4}$$

$$\mathbf{P}_k^f = \mathbf{M}\mathbf{P}_{k-1}^a\mathbf{M}^T + \mathbf{Q}_k. \tag{5}$$

- **Correction:**

$$\mathbf{K}_k = \mathbf{P}_k^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_k^f\mathbf{H}^T + \mathbf{R})^{-1}, \tag{6}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{y}_k^o - \mathbf{H}\mathbf{x}_k^f), \tag{7}$$

$$\mathbf{P}_k^a = (\mathbb{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_k^f. \tag{8}$$

The result of the correction step, also called analysis, provides the vector $\mathbf{x}_k^a \triangleq \mathbb{E}(\mathbf{x}_k|\mathbf{y}_{1:k}^o)$ and the covariance $\mathbf{P}_k^a \triangleq \mathbb{E}((\mathbf{x}_k - \mathbf{x}_k^a)(\mathbf{x}_k - \mathbf{x}_k^a)^T|\mathbf{y}_{1:k}^o)$ characterizing the Gaussian filtering distribution at time k. These moments depend on the forecast mean $\mathbf{x}_k^a$ and covariance $\mathbf{P}_k^a$. The matrix $\mathbf{K}_k$, which takes place in the mean and covariance analysis, defines the so-called Kalman gain.

2.2 Ensemble Kalman filter

We now briefly present the ensemble extension of the Kalman filter for systems described by a non-linear dynamics (1) and a linear measurement model (2). The Ensemble Kalman filter aims at using the expressions provided by

the Kalman filter for the linear case, and requires a Gaussian approximation of the predictive law. The method, related to Monte Carlo approaches, relies on an ensemble of samples to describe the different probability distributions. The initial law is sampled by N members (or particles) $\mathbf{x}_0^{(i)}, i = 1, \ldots, N$. The forecast distribution and filtering distribution are respectively approximated through a prediction step and a correction step of the ensemble members.

The prediction step consists in propagating the ensemble of particles $\mathbf{x}_{k-1}^{a,(i)}$ through the non-linear dynamics (including the model noise simulation) in order to obtain the predicted particles, or forecast ensemble, denoted by:

$$\mathbf{x}_k^{f,(i)} = \sum_{t=k-1}^{k-\Delta t} \left(M(\mathbf{x}_t^{f,(i)}) + \boldsymbol{\eta}_{t+\Delta t}^{(i)} \right), \tag{9}$$

$$\mathbf{x}_{k-1}^{f,(i)} = \mathbf{x}_{k-1}^{a,(i)}. \tag{10}$$

This generally corresponds to the numerical integration of a non-linear stochastic partial differential evolution law. The second term of this sum corresponds to the simulation of a white Gaussian noise with covariance $\mathbf{Q}_{\Delta t}$ that is assumed to be known. Let us note that simpler dynamics defined only up to a random initial condition can be considered. In that case the knowledge of the model noise is not required. In this work we will only stick to the most general case and consider dynamics defined through a stochastic differential equation. Let us note however that the techniques proposed in this paper can be directly implemented for simplified dynamics defined up to a random initialization.

For systems such as (9), the repartition of the predicted particles $\{\mathbf{x}_k^{f,(i)}, i = 1, \ldots, N\}$ follows a Gaussian distribution only if two assumptions are satisfied:

- the analyzed set of particles $\{\mathbf{x}_{k-1}^{a,(i)}, i = 1, \ldots, N\}$ corresponds to a Gaussian repartition
- the dynamics operator M is linear.

Generally, we can consider that the less linear the dynamics, the less Gaussian the forecast particles distribution. Let us remark that departure from this Gaussian assumption depends obviously on the nonlinearity of the operator involved and on the noise covariance considered. It depends also on the duration of the temporal integration. For short time horizon, the Gaussian assumption may remain valid for some nonlinear operator which does not develop multimodalities on short time scale whereas it breaks very quickly for chaotic dynamics (such as the Lorentz models used as toy models to represent geophysical dynamics). The EnKF assumes this Gaussian approximation between assimilation times, in order to apply Kalman's formulas. The empirical mean of the forecast ensemble is firstly defined by:

$$\overline{\mathbf{x}}_k^f = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_k^{f,(i)}. \tag{11}$$

The empirical ensemble covariance matrix $\mathbf{P}_k^{fe}$ is then deduced from the following expression:

$$\mathbf{P}_k^{fe} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_k^{f,(i)} - \overline{\mathbf{x}}_k^f \right) \left(\mathbf{x}_k^{f,(i)} - \overline{\mathbf{x}}_k^f \right)^T. \tag{12}$$

This relation is the *non biased* estimator of the variance for data sets, taking into account that the mean is known up to its own variance.

Similarly to the Gaussian case presented in the previous section, an ensemble Kalman gain $\mathbf{K}_k^e$ can be computed from the ensemble variance. We have:

$$\mathbf{K}_k^e = \mathbf{P}_k^{fe} \mathbf{H}^T (\mathbf{H} \mathbf{P}_k^{fe} \mathbf{H}^T + \mathbf{R})^{-1}. \tag{13}$$

Let us now present the corrective step, in which the forecast ensemble members are moved toward the new observation. Two families of methods have been devised. The first type relies on an additional random noise variable that operates a perturbation on the observations whereas the second kind of techniques avoids the use of perturbed observations and introduces a deterministic corrective updating rule.

2.2.1 Techniques with perturbed observations In order to realize the correction step and to treat the observations as random measurements associated to each particle, (Evensen, 1994; Houtekamer and Mitchell, 1998; Burgers et al., 1998; Evensen, 2003) introduced a measurement process in which the observation is perturbed by N realizations denoted by $\{\boldsymbol{\epsilon}_k^{(i)}, i = 1, \ldots, N\}$ of the measurement noise probability density function. With such perturbed observations the empirical ensemble covariance keeps the same form as the one corresponding to the original Kalman updates (Burgers et al., 1998).

Each particle, $\mathbf{x}_k^{f,(i)}$, is then associated to the perturbed observation, $\mathbf{y}_k^o + \boldsymbol{\epsilon}_k^{(i)}$, and the corrective step of the Kalman filter is applied in order to obtain the analyzed ensemble $\{\mathbf{x}_k^{a,(i)}, i = 1, \ldots, N\}$. We have:

$$\mathbf{x}_k^{a,(i)} = \mathbf{x}_k^{f,(i)} + K_k^e (\mathbf{y}_k^o + \boldsymbol{\epsilon}_k^{(i)} - \mathbf{H}\mathbf{x}_k^{f,(i)}). \tag{14}$$

2.2.2 Methods without perturbed observations The other set of methods (Anderson, 2001; Bishop et al., 2001; Hunt et al., 2004; Ott et al., 2004; Tippett et al., 2003; Whitaker and Hamill, 2002), called square-root filters, confine the analysis to a subspace of vector perturbations that represents the analysis error covariance. These deterministic schemes avoid sampling issues (caused by a small number of members) by generating an ensemble with the desired sample mean and covariance. They implement an updating rule that reads:

$$\mathbf{x}_k^{a,(i)} = \sum_{j=1}^{N} (\mathbf{x}_k^{f,(j)} - \overline{\mathbf{x}}_k^f) W^{j,i} + \overline{\mathbf{x}}_k^f. \tag{15}$$

The matrix $\mathbf{W}$ is symetric and depends on ensemble covariance, the discrepancy between the forecast ensemble and the observation (also called innovation), and the observation error covariance.

2.2.3 Practical considerations for EnKF implementation In practice for both kinds of methods, the empirical ensemble covariance matrices, which dimensions are $n \times n$, are never computed nor stored. This point is essential for geophysical applications, as n corresponds to the dimension of the discrete spatial domain on which atmospheric states variables such as pressure, temperature or flow velocity have to be computed. Noticing that matrix vector product $\mathbf{P}_k^{fe}\mathbf{u}$ requires N scalar products (associated to a given

vector space definition), the products $\mathbf{P}_k^{f_e}\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}_k^{f_e}\mathbf{H}^T$ involved in the Kalman gain computation necessitate $N \times m$ scalar products, where m is the observation space dimension. The computation of the Kalman gain can be further reduced by assimilating the observations serially (Anderson, 2003; Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002) – with the assumption that they are uncorrelated. In this case the Kalman gain requires only scalar inverses. For non diagonal observation covariance, a SVD factorization enables to operate a change of coordinates allowing to come back to a serial assimilation (Anderson, 2003). Alternatively, several authors have proposed to define the analysis within reduced subspaces spanned by ensemble members (Bishop et al., 2001; Hunt et al., 2004; Ott et al., 2004).

Other important issues for practical implementations of EnKF concern covariance inflation and localization techniques. Covariance inflation (Anderson and Anderson, 1999; Anderson, 2007) consists to augment artificially the variance of each state vectors components. Inflation procedures allow to correct systematic underestimation of the ensemble error covariance matrix. Such a deficiency yields an over confidence on the model dynamics and hence leads to ignore in an extreme case the observations. Inflation can be applied either on the ensemble covariance or on the analysis covariance. Inflation is either multiplicative (a multiplicative factor is applied to the ensemble covariance (Anderson and Anderson, 1999; Anderson, 2007)) or additive (a small multiple of the identity is added to the ensemble covariance or to the analysis covariance (Ott et al., 2004)). Let us note that an appropriate tuning of the factor involved remains a difficult issue.

Finally, as noted in the introduction, in some applications the analysis time and the observation time may be different. This particularity leads however to significant changes in the method design, as the problem consists in a smoothing problem in which past and future measurements are used, whereas filtering refers only to the sequential assimilation of past and current observations. Schemes relying on ensemble Kalman smoother (Evensen and van Leeuwen, 2000), techniques introducing a linear combination of ensemble trajectories that best fits asynchronous observations (Hunt et al., 2004) or approaches combining variational assimilation and ensemble Kalman filtering may be used to cope with this problem (Fertig et al., 2007; Harlim and Hunt, 2007a;b; Zupanski, 2005).

2.3 Particle filter

In this subsection, we keep the same system composed of a linear measurement relation and a non-linear dynamics, both of them still including additive Gaussian noises. Even if particle filtering can be applied to much general cases, involving nonlinear measurement models, non additive and non Gaussian noises, in this study we stick to the kind of systems that are handled by ensemble Kalman filter techniques.

2.3.1 General problem To evaluate the general filtering distribution of the whole state variables trajectories $\mathbf{x}_{0:k}$ from the observations $\mathbf{y}^o_{1:k}$, a recursive expression is established from the Bayes' law:

$$\begin{aligned} p(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k}) &= \frac{p(\mathbf{y}^o_k|\mathbf{x}_k, \mathbf{y}^o_{1:k-1})p(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k-1})}{p(\mathbf{y}^o_k|\mathbf{y}^o_{1:k-1})} \\ &= \frac{p(\mathbf{y}^o_k|\mathbf{x}_k)p(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k-1})}{p(\mathbf{y}^o_k|\mathbf{y}^o_{1:k-1})}. \end{aligned} \tag{16}$$

From the Bayes' formula, the filtering law can be written as:

$$p(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k}) = p(\mathbf{x}_{0:k-1}|\mathbf{y}^o_{1:k-1})\frac{p(\mathbf{y}^o_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})}{p(\mathbf{y}^o_k|\mathbf{y}^o_{1:k-1})}. \tag{17}$$

We then get a recursive expression of the filtering distribution of the trajectories. For continuous dynamics, $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is obtained through Monte Carlo sampling of the dynamical evolution law between time $k-1$ and time k in the same way as for the ensemble Kalman filter. This distribution does not have to be explicitly known. It is important to note that the observations $\mathbf{y}^o_k$ are conditionally independent of the past trajectory $\mathbf{x}_{0:k-1}$ and the past observations $\mathbf{y}^o_{0:k-1}$ given the state variable $\mathbf{x}_k$:

$$p(\mathbf{y}^o_k|\mathbf{y}^o_{0:k-1}, \mathbf{x}_{0:k}) = p(\mathbf{y}^o_k|\mathbf{x}_k). \tag{18}$$

Particle filtering techniques propose to implement an approximation of the sought density $p(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k})$. This approximation consists of a finite weighted sum of N Diracs centered on hypothesized locations in the state space – called particles – of the initial system $\mathbf{x}_0$. At each particle $\mathbf{x}_k^{(i)}$ $(i = 1, \ldots, N)$ is assigned a weight $w_k^{(i)}$ describing its relevance. This approximation can be formulated with the following expression:

$$p(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k}) \approx \sum_{i=1}^{N} w_k^{(i)} \delta_{\mathbf{x}^{(i)}_{0:k}}(\mathbf{x}_{0:k}). \tag{19}$$

As the samples cannot be simulated directly from this distribution, particles are simulated from a proposal distribution $\pi(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k})$, (also called the importance distribution) that approximates the true filtering distribution. Each sample is then weighted with the ratio between the two distributions at the sample point. The importance weights $w_k^{(i)}$ account for the deviation with respect to the unknown true distribution.

Due to the importance density sampling, the target distribution will be thus fairly sampled by the particles $\mathbf{x}_{0:k}^{(i)}$ weighted by weights $w_k^{(i)}$, defined as:

$$w_k^{(i)} = \frac{p(\mathbf{x}^{(i)}_{0:k}|\mathbf{y}^o_{1:k})}{\pi(\mathbf{x}^{(i)}_{0:k}|\mathbf{y}^o_{1:k})}. \tag{20}$$

The closer the approximation to the true distribution, the more efficient the filter. However, any importance function can be chosen, with the only restriction that its support contains the one of the target density. Furthermore, so as to obtain a sequential formulation of the samples, the importance law is chosen with the following recursive form:

$$\pi(\mathbf{x}_{0:k}|\mathbf{y}^o_{1:k}) = \pi(\mathbf{x}_{0:k-1}|\mathbf{y}^o_{1:k-1})\pi(\mathbf{x}_k|\mathbf{y}^o_{1:k}, \mathbf{x}_{0:k-1}). \tag{21}$$

By combining this relation with the recursive expression of the objective law given by relation (17), we get a general recursive evaluation for the weights as the new measurement $\mathbf{y}^o_k$ becomes available:

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{y}^o_k|\mathbf{x}^{(i)}_k)p(\mathbf{x}^{(i)}_k|\mathbf{x}^{(i)}_{k-1})}{\pi(\mathbf{x}^{(i)}_k|\mathbf{x}^{(i)}_{0:k-1}, \mathbf{y}^o_{1:k})}, \tag{22}$$

with weights normalized to 1, and thus removing the factor $p(\mathbf{y}_k^o|\mathbf{y}_{1:k-1}^o)$ of relation (17), which is the same for all the particles. Marginals of the complete filtering density $p(\mathbf{x}_k|\mathbf{y}_{1:k}^o)$ can be then easily obtained from the normalized new weights and (19):

$$p(\mathbf{x}_k|y_{1:k}^o) \approx \sum_{i=1}^{N} w_k^{(i)} \delta_{\mathbf{x}_k^{(i)}}(\mathbf{x}_k). \tag{23}$$

Thus, by propagating the particles from time $k-1$ through the proposal density $\pi(\mathbf{x}_k^{(i)}|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k}^o)$, and by weighting the sampled states with $w_k^{(i)}$, we obtain a sampling of the filtering law. Different choices are possible for this proposal density (Doucet et al., 2000), provided the weights in the correction step are computed accordingly and that its support includes the support of the filtering distribution. The most common ones are discussed in the following. The adaptation of the particle filter we propose in this paper, and that is described in details in the next section, relies on a particular formulation of this proposal density.

Asymptotically, for a number of particles tending to infinity, convergence toward the Bayesian filtering distribution of various classes of particle filters have been demonstrated (Crisan and Doucet, 2002; Del Moral, 2004) with a rate of $1/\sqrt{N}$. The same convergence result applies for the ensemble Kalman filter. However for non-linear systems, EnKF does not converge toward the optimal limiting filtering distribution (Le Gland et al., 2010). In practical implementations the number of particles is difficult to fix. The number of required particles to ensure the filter convergence depends on the state space dimension but also on the ability we have to draw samples in meaningful areas of the state space.

Several versions of particle filters have been devised in the statistical community or in the geophysical community. In a recent review, Van Leeuwen provides a complete description of the different variants proposed so far and gives some new research directions that would be interesting to follow in order to improve the global performances of particle filters (in term of accuracy or in term of filter degeneracy) (van Leeuwen, 2009). Among all the routes of possible improvements, in this study we focus on the definition of a sound proposal distribution and on an adaptation of the resampling scheme. In the following sections we describe briefly both procedures.

2.3.2 Resampling Limiting ourselves to the prediction and correction steps for updating the particles induces an increase over time of the weight variance (Kong et al., 1994). In practice, this degeneracy problem makes the number of significant particles decreases dramatically over time, implying an impoverishment of the estimate. To avoid degeneracy of the particles, the filtering is supplemented with a resampling step of the particles with respect to their distribution. This procedure aims at discarding particles with weak normalized weights and at duplicating particles associated to strong weights. Consequently, resampled particles (with equal weights $1/N$) tend to be concentrated in the significant areas of the state space. Various resampling strategies have been described in the literature and will not be developed in this paper. More details can be found in (Doucet et al., 2000). The accuracy of the sampling can be estimated by a quantity called the effective sample size (Doucet et al., 2000), or by an entropy term as in the information theory (Pham, 2001) (see also (Kong et al., 1994)). These criteria give an indication when the resampling should be performed. Examples and comparisons of several resampling schemes used in geophysical data assimilation problems are thoroughly described in (van Leeuwen, 2009).

2.3.3 Bootstrap filter If the proposal distribution is set to the dynamics:

$$\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k}^o) = p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}), \tag{24}$$

the weights updating rule (22) thus simplifies to the data likelihood $p(\mathbf{y}_k^o|\mathbf{x}_k^{(i)})$. This particle filter version is called *Bootstrap* filter (Gordon et al., 1993), and is the most frequently used due to the simplicity of its implementation. This is the kind of filter that has been considered in several studies in order to compare the respective performances of ensemble Kalman filters and particle filters in geophysical sciences (Kivman, 2003; van Leeuwen, 2003; Pham, 2001).

In this approach the particles are sampled on the sole basis of the dynamics and ignore completely the future measurement. When the dynamics is highly non-linear a great number of particles is required in order to cover a sufficiently large domain of the state space. In case of bad prediction all the particles become insignificant and the filter will diverge if future random explorations of the state space do not recover any significant configurations at a given time. For huge state spaces driven by fluid flow conservation laws, this ability to recover meaningful areas of the state space is very unlikely. In addition, for geophysical flows analysis problems, the number of particles is severely constrained by the forecast computational cost which further penalizes this scheme.

2.3.4 Optimal importance filter In order to limit the degeneracy of the particles, Doucet *et al.* have introduced the notion of *optimal importance function* which minimizes the variance of the weights conditioned upon $\mathbf{x}_{0:k-1}$ and $\mathbf{y}_{1:k}^o$. This optimal choice corresponds to the proposal density (Doucet et al., 2000):

$$\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k}^o) = p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k^o). \tag{25}$$

From relation (18), observing that

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k^o) = \frac{p(\mathbf{y}_k^o|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)})}{p(\mathbf{y}_k^o|\mathbf{x}_{k-1}^{(i)})}, \tag{26}$$

the weight recursive formulation (22) becomes in that case:

$$w_k^{(i)} \propto w_{k-1}^{(i)} p(\mathbf{y}_k^o|\mathbf{x}_{k-1}^{(i)}). \tag{27}$$

The corresponding filter thus requires to evaluate $p(\mathbf{y}_k^o|\mathbf{x}_{k-1})$, called the predicted likelihood, up to a proportionality constant. Analytic expression of this function can be written for models composed of a discrete non-linear dynamics and a likelihood defined as a Gaussian mixture (Arnaud and Mémin, 2007). For continuous dynamical models, this density can only be approximated at the expense of additional integrations of the dynamical models. To see that, we consider an augmented state vector gathering the states between instants $k-1+\Delta t$ and k:

$\mathbf{x}'_k = (\mathbf{x}_{(k-1)+\Delta t}, \mathbf{x}_{(k-1)+2\Delta t}, \ldots, \mathbf{x}_k)^T$. For this augmented state vector the optimal importance density reads:

$$p(\mathbf{x}'_k|\mathbf{x}'_{k-1}, \mathbf{y}^o_k) \propto \prod_{t=k-1+\Delta t}^{k} p(\mathbf{x}_t|\mathbf{x}_{t-\Delta t}) p(\mathbf{y}^o_k|\mathbf{x}_t).$$

The predictive likelihood $p(\mathbf{y}^o_k|\mathbf{x}_t)$ is not explicitly known for the different intermediate states. This probability can be evaluated for a given state through a backward Kolmogorov equation (the adjoint of a Fokker planck equation) associated to the infinitesimal generator of the dynamical model (written as a stochastic differential equation). This integration whose price is of the same order as the original dynamical model must be realized for each intermediate state $\mathbf{x}_t^{(i)}$ from a pseudo-measurement at time t and the likelihood $p(\mathbf{y}^o_t|\mathbf{x}_t^{(i)})$. This procedure is in practice totally unfeasible for geophysical data assimilation process.

2.4 Comparison of EnKF and particle filter

The two methods, the particle filter and the EnKF, share some similarities: stochastic sampling and processing organized in two steps (the predictive and the corrective steps). Technically, the main difference lays on how the corrective step is done. The computation of a weight for each particle is necessary in the particle filtering procedure, whereas the particles are moved a second time in the ensemble Kalman filter. From a conceptual point of view, particle filtering aims at approximating through a combination of Dirac masses centered on the particles the complete filtering distribution. On the other hand, like the original Kalman filter, the ensemble Kalman filter intends only to estimate the two first moments of the distribution. From a theoretical point of view, for a number of particles tending to infinity, the limiting distribution of the particle filter converges to the optimal filtering distribution (Crisan and Doucet, 2002; Del Moral, 2004). This is not true anymore for the ensemble Kalman filter with non-linear dynamics, which converges toward a different limiting distribution that has still to be characterized (Le Gland et al., 2010; 2009).

From this discussion, we see that both methods have clearly their own advantages and drawbacks:

- The EnKF places the particles in significant areas of the state space by shifting them according to the dynamics and its uncertainty but also taking into account the new observation and the past measurements. This process and its *in fine* limited goal (the estimation of the a posteriori two first moments) allow, as demonstrated in practice in numerous geophysical data assimilation problems, to reduce efficiently the number of necessary particles. The accuracy of the filter and eventually that it does not diverge remains circumvented by the validity of the Gaussian approximation of the filtering distribution (Harlim and Majda, 2009). Despite of this limitation, EnKF has been successfully applied to a wide range of geophysical systems in operational settings (Houtekamer et al., 2005; Houtekamer and Mitchell, 2006; Szunyogh et al., 2008; Miyoshi and Yamane, 2007; Liu et al., 2008; Whitaker et al., 2008).

Let us note that recent ensemble Kalman filters (Fertig et al., 2007; Harlim and Hunt, 2007a;b; Hunt et al., 2004; Zupanski, 2005) are relaxing the Gaussian assumption through a combination of variational data assimilation and ensemble Kalman filters. These methods formulate the problem as a deterministic optimization where the solution is expressed as a linear combination of the ensemble members. The optimization is conducted with respect to the coefficients of these linear combinations. Those coefficients are weighting the contribution of the different members and are in nature completely different from the weights involved in particle filtering. Furthermore, these techniques express the data assimilation problem as a smoothing deterministic optimization problem that may converge through a local minimum. Eventual multimodality of the pdf may pose problem, as only a single (eventually local) mode will be retained. In theory, particle filters allow to tackle this situation as they may keep a trace along time of the most pregnant modes of a multimodal distribution.

- Particle filtering relies on an exact sequential probabilistic formulation of the filtering law. It thus gives an exact result up to sampling errors for a large number of particles. This technique is as a consequence quite sensitive to the curse of dimensionality and works well in practice only for state spaces of reduced dimension. When the importance sampling is based solely on the dynamics (*i.e.* the bootstrap filter), the state space exploration ignores completely the new measurement. Such a sampling, "blind" to the different available observations, requires therefore a great number of particles to be efficient. This constitutes a serious limitation for their use in geophysical applications. From a computational point of view, for an individual element, particle filters have lower computational cost compared to EnKF, as no matrix inversions are required. The efficiency of the simple particle filter scheme is however completely ruined by the great number of particles needed to ensure the convergence of the filter.

The EnKF has been mainly used for huge dimensional state spaces associated to forecast of high computational cost (as in meteorology or oceanography). In such contexts, this process has shown to provide a correct approximation of the two first moments of the filtering distribution addressed, even for a small set of particles (if the Gaussian assumption appears to be valid). When the state space dimension is much larger than the ensemble size, EnKF may also suffer from the curse of dimensionality and exhibit degeneracy. A common practice used to tackle this problem in EnKF consists to perform a local analysis in considering for a given grid points only observations that lie into a given neighborhood. This process called spatial "localization" may be implemented in different ways (Houtekamer and Mitchell, 1998; Whitaker et al., 2008; Anderson, 2003; Ott et al., 2004; Hunt et al., 2007) and is justified by the suppression of spurious long-range correlations that may be introduced through a small ensemble size (Houtekamer and Mitchell, 1998). For a large ensemble, if achievable, localization would not be needed. On the other hand, localization also improves the computational efficiency of EnKF. In addition, it provides a broader state space exploration, as the analysis consists in that case to choose locally linear combinations of members. Arguments on intrinsic low dimensional manifold representation in local regions of atmospheric dynamics is also advanced in order to justify this localization procedure (Patil et al., 2001; Kuhl et al., 2007).

In comparison, particle filtering has demonstrated to be very efficient in a complete non-linear setting for problems with a state space of reduced dimension. However, the curse of dimensionality is more pregnant for the *Bootstrap* filters as the new measurements are not taken into account in the proposal distribution. In case of large innovation (*i.e.* if the predicted state is far from the observations), the filter requires a great number of particles to explore a sufficiently large region of the state space. This drawback makes it inadequate for data assimilation processes in geophysical sciences, which, for computational reasons, are restricted to a very small number of particles for the exploration of state spaces of huge dimensions. Formal arguments are also given in (Snyder et al., 2008) for an intrinsic deficiency of the bootstrap particle filters in the case of peaked likelihood function (such as those that result from a large number of independent observations). Even in the simple case of the bootstrap particle filter there exist for the moment no theoretical bounds allowing indicating, with respect to the state dimension, the number of particle that should be drawn.

Experimental comparisons between ensemble Kalman filtering and particle filtering have only been made on the basis of the *Bootstrap* particle filters (Kivman, 2003; Pham, 2001; van Leeuwen, 2003; Zhou et al., 2006; Rémy et al., 2010). In all these studies it is reported that for highly non-linear systems this simplest particle filter provided superior results compared to EnKF but required a much higher number of particles than the number of members needed for the EnKF. It has been noted that as the number of observations was increased the need of even a greater number of particles was required (Zhou et al., 2006), confirming hence the formal statement developped in (Snyder et al., 2008) for bootstrap filter. Divergence was observed when to few particles were used. In some experiments convergence of the EnKF to wrong parameter values was also observed (Harlim and Majda, 2009; Kivman, 2003; van Leeuwen, 2003). This confirms the asymptotic analysis done in (Le Gland et al., 2010) which states that EnKF for a non-linear state dynamics with linear observations model converges toward a limiting distribution that differs from the sought filtering distribution. To avoid particle filter divergence, (van Leeuwen, 2003) considered likelihood with broader tail than a Gaussian distribution. In (van Leeuwen, 2002), it has been also proposed to draw the samples of the state vectors using future observations in order to guide the particles toward the new observations. With the same purpose, we propose in the following section a scheme that enables through an ensemble Kalman filtering procedure to propose samples that integrate an history of past and current observations.

3 The weighted ensemble Kalman filter

Starting from the descriptions of the previous section, a hybrid filtering procedure that takes advantage of both the particle filter and the EnKF can be devised.

The importance sampling principle indicates that a wide range of proposal distributions can be considered. We claim and we will experimentally demonstrate that a proposal distribution defined by the two successive particle shifts in the EnKF procedure constitutes an efficient proposal mechanism. Moreover, the insertion of the EnKF shift mechanisms into a particle filter scheme enables a better approximation of the filtering law and thus increases the filter accuracy. Two cases corresponding to the two EnKF variants identified in section 2.2 can be distinguished.

3.1 Proposal distribution from EnKF with perturbed observations

The expression (14) defines the displacement of the ensemble members; it can be rewritten as:

$$\mathbf{x}_k^{f,(i)} = \sum_{t=k-1}^{k-\Delta t} \left(M(\mathbf{x}_t^{f,(i)}) + \boldsymbol{\eta}_{t+\Delta t}^{(i)} \right) \tag{28}$$

$$\mathbf{x}_{k-1}^{f,(i)} = \mathbf{x}_{k-1}^{(i)} \tag{29}$$

$$\mathbf{x}_k^{(i)} = (\mathbb{I} - \mathbf{K}_k^e \mathbf{H}) \mathbf{x}_k^{f,(i)} + \mathbf{K}_k^e \mathbf{y}_k^o + \mathbf{K}_k^e \boldsymbol{\epsilon}_k^{(i)} \tag{30}$$

$$= \boldsymbol{\mu}_k^{(i)} + \boldsymbol{\gamma}_k^{(i)} \tag{31}$$

where

$$\boldsymbol{\mu}_k^{(i)} = (\mathbb{I} - \mathbf{K}_k^e \mathbf{H}) \mathbf{x}_k^{f,(i)} + \mathbf{K}_k^e \mathbf{y}_k^o, \tag{32}$$

$$\boldsymbol{\gamma}_k^{(i)} = \mathbf{K}_k^e \boldsymbol{\epsilon}_k^{(i)}. \tag{33}$$

The first term constitutes a drift term that depends respectively on the integration of the dynamics for a given particle between $k-1$ and k, the new measurement and the ensemble error covariance (through the Kalman gain computed from the forecast ensemble). Relying on the usual assumption of the EnKF (*i.e.* considering the dynamics as a discrete Gaussian system), its conditional distribution given the data is described as the following Gaussian distribution:

$$\boldsymbol{\mu}_k^{(i)} \sim \mathcal{N}\left(\overline{\boldsymbol{\mu}}_k^{(i)}, \mathbf{V} \right), \tag{34}$$

with

$$\overline{\boldsymbol{\mu}}_k^{(i)} = (\mathbb{I} - \mathbf{K}_k^e \mathbf{H}) \sum_{t=k-1}^{k-\Delta t} M(\mathbf{x}_t^{f,(i)}) + \mathbf{K}_k^e \mathbf{y}_k^o \tag{35}$$

$$\mathbf{V} = (\mathbb{I} - \mathbf{K}_k^e \mathbf{H}) \mathbf{Q} (\mathbb{I} - \mathbf{K}_k^e \mathbf{H})^T. \tag{36}$$

Here $\mathcal{N}(\boldsymbol{y}, \boldsymbol{\Upsilon})$ denotes a Gaussian distribution with mean $\boldsymbol{y}$ and variance $\boldsymbol{\Upsilon}$. In what follows, a notation $\mathcal{N}(\mathbf{x}; \boldsymbol{y}, \boldsymbol{\Upsilon})$ will denote the value of this Gaussian at $\mathbf{x}$. The covariance $\mathbf{Q}$ depends on the model noise and on the propagation of the incertitude by the dynamics. Assuming it is independent of a given particle (which is true only for linear dynamics or discrete Gaussian systems), its precise characterization will not be needed as shown below.

The second term is a noise term that depends on the measurement noise but also on the accumulated forecast ensemble error through the Kalman gain. It is a zero mean Gaussian variable with covariance $\mathbf{K}_k^e \mathbf{R} \mathbf{K}_k^{e\,T}$. Both terms are uncorrelated and as a consequence, the conditional distribution of the state $\mathbf{x}_k$ given the new measurement and the particle at time $k-1$ is approached by:

$$\mathbf{p}(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k^o) = \mathcal{N}\left(\overline{\boldsymbol{\mu}}_k^{(i)}, \boldsymbol{\Sigma}_k^e \right) \tag{37}$$

with

$$\boldsymbol{\Sigma}_k^e = (\mathbb{I} - \mathbf{K}_k^e \mathbf{H}) \mathbf{Q} (\mathbb{I} - \mathbf{K}_k^e \mathbf{H})^T + \mathbf{K}_k^e \mathbf{R} \mathbf{K}_k^{e\,T},$$

$$= (\mathbb{I} - \mathbf{K}_k^e \mathbf{H}) \mathbf{P}_k^{f_e} \tag{38}$$

In the establishment of this distribution it can be noticed that we have neglected the dependence on the other particles (that appears through the forecast ensemble covariance) in considering the forecast ensemble covariance as a mean field variable, intrinsic to the system and that does not depend on system realizations –, which is only true asymptotically for a number of particles tending to infinity.

3.2 Proposal distribution from EnKF without perturbed observations

As remarked in (van Leeuwen, 2009) a similar construction can be applied to deterministic EnKF procedures that do not introduce random perturbations of the observations. In that case we have from (15) and the dynamical model (9):

$$\mathbf{x}_k^{(i)} = \sum_{j=1}^{N} (\sum_{t=k-1}^{k-\Delta t} \left(M(\mathbf{x}_t^{f,(j)}) + \boldsymbol{\eta}_{t+\Delta t}^{(j)} \right) - \overline{\mathbf{x}}_k^f) W^{j,i} + \overline{\mathbf{x}}_k^f. \tag{39}$$

Relying on the same Gaussian approximation as previously for the forecast distribution we have now:

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k^o) = \mathcal{N}\left(\overline{\boldsymbol{\mu}}_k^{(i)}, \boldsymbol{\Sigma}_k^e \right), \tag{40}$$

with

$$\overline{\boldsymbol{\mu}}_k^{(i)} = \sum_{j=1}^{N} (\sum_{t=k-1}^{k-\Delta t} \left(M(\mathbf{x}_t^{f,(j)}) \right) - \overline{\mathbf{x}}_k^f) W^{j,i} + \overline{\mathbf{x}}_k^f \tag{41}$$

$$\boldsymbol{\Sigma}_k^e = \mathbf{W}\mathbf{Q}\mathbf{W}^T \tag{42}$$

Square-root filters are usually built in order that the mean and covariance of the ensemble respect the analysis mean and covariance given by the Kalman update equations. The Kalman analysis mean and covariance of the ensemble are directly verified in EnKF with perturbed observations (37-38). Up to sampling errors these two procedures sample ensembles distributions with identical two first moments.

3.3 WEnKF algorithm

The distribution $p(\mathbf{x}_k | \mathbf{x}_k^{(i)}, \mathbf{y}_k^o)$ (defined from the two previous cases (37) or (40)) provides us a natural expression for the proposal distribution:

$$\pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k}^o) = p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k^o) = \mathcal{N}(\overline{\boldsymbol{\mu}}_k^{(i)}, \boldsymbol{\Sigma}_k^e).$$

In order to make the estimation of the filtering distribution exact (up to the sampling), the formulas of the particle filter are applied. As a consequence, each member of the ensemble must be weighted at each instant, k, with a weight, $w_k^{(i)}$, defined iteratively by:

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{y}_k^o | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{\mathcal{N}\left(\mathbf{x}_k^{(i)} - \overline{\boldsymbol{\mu}}_k^{(i)}; 0, \boldsymbol{\Sigma}_k^e \right)}, \text{ and } \sum_{i=1}^{N} w_k^{(i)} = 1. \tag{43}$$

The proposal distribution depends on the ensemble Kalman gain estimated from the mean and the covariance of the predicted particles $\mathbf{x}_k^{f,(i)}$ where the mean and covariance of the forecast particles in relations (11) and (12) must be here computed according to the particles weights at instant $k-1$. When a resampling scheme is applied systematically at each measurement instant every particle weights are reinitialized to $1/N$, the previous formula simplifies as:

$$w_k^{(i)} \propto \frac{p(\mathbf{y}_k^o | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{\mathcal{N}\left(\mathbf{x}_k^{(i)} - \overline{\boldsymbol{\mu}}_k^{(i)}; 0, \boldsymbol{\Sigma}_k^e \right)}, \text{ and } \sum_{i=1}^{N} w_k^{(i)} = 1, \tag{44}$$

and the classical empirical formulas can be used to compute the mean and covariance of the ensemble.

The probability distribution $p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})$ is known in a discrete setting (or when the dynamics discrete scheme time step corresponds to the measurement lapse time). It is usually analytically and numerically unavailable in a continuous setting:

$$\begin{aligned} p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}) &= \int_{t=k-1+\Delta t}^{k} \prod p(\mathbf{x}_t^{(i)} | \mathbf{x}_{t-\Delta t}^{(i)}) d\mathbf{x}_{k-\Delta t} \cdots d\mathbf{x}_{k-1+\Delta t}, \\ p(\mathbf{x}_t^{(i)} | \mathbf{x}_{t-\Delta t}^{(i)}) &= \mathcal{N}\left(M(\mathbf{x}_{t-\Delta t}^{(i)}), \mathbf{Q}_{\Delta t} \right). \end{aligned} \tag{45}$$

We will see however that the calculation of this expression will not be required when the Gaussian transition densities (45) are represented on the basis of a very small number of particles compared to the state space dimension.

The Weighted ensemble Kalman filter (WEnKF) procedure can be simply summarized by the algorithm 1. Let us note that in this synoptic description, we described the simplest resampling scheme, which consists to draw the samples directly from the discrete weight distribution. This resampling consists to draw with replacement N particles according to their normalized weights. Other efficient schemes reducing the sampling noise can be used instead (see (van Leeuwen, 2009) and references therein for the description of different resampling schemes and some comparison elements on geophysical applications).

Algorithm 1 The WEnKF algorithm, one iteration.

Require: Ensemble at instant $k-1$: $\{\mathbf{x}_{k-1}^{(i)}, i = 1, \ldots, N\}$ observations $\mathbf{y}_k^o$

Ensure: Ensemble at time k: $\{\mathbf{x}_k^{(i)}, i = 1, \ldots, N\}$

EnKF step: Get $\mathbf{x}_k^{(i)}$ from the assimilation of $\mathbf{y}_k^o$ with an EnKF procedure (28-33) or (28,39);

Compute the weights $w_k^{(i)}$ according to (44);

Resample: For $j = 1 \ldots N$, sample with replacement index $a(j)$ from discrete probability $\{w_k^{(i)}, i = 1, \ldots, N\}$ over $\{1, \ldots, N\}$ and set $\mathbf{x}_k^{(j)} = \mathbf{x}_k^{a(j)}$;

3.4 Practical considerations for implementation in high dimension

For high dimensional problems, we first have to check if the WEnKF does not imply some additional and prohibitive computational costs compared to traditional ensemble filters. The classical EnKF has been very carefully designed in order to allow its application in geophysical assimilation issues. This has been done by taking care of problems that arise when facing high dimensional covariance matrices and noise simulation. Based on similar recipes (see (Bishop et al., 2001; Evensen, 2003) for a detailed description), we explain in what follows how the weights can be computed on similar computational basis.

3.4.1 Weights computation The three probabilities involved in equation (44) have to be evaluated in order to compute the weights of the ensemble members after the prediction step. These conditional probability distributions, namely, the *likelihood* $p(\mathbf{y}_k^o|\mathbf{x}_k^{(i)})$, the transition distribution $p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-\Delta t}^{(i)})$ (45) and the *proposal* distributions $\mathcal{N}\left(\mathbf{x}_k^{(i)} - \overline{\boldsymbol{\mu}}_k^{(i)}; 0, \boldsymbol{\Sigma}_k^e\right)$, are related to centered Gaussian distributions:

$$p(\mathbf{a}_t^{(i)}|\mathcal{F}_t) \propto \exp(-\frac{1}{2}\mathbf{a}_t^{(i)T}\Upsilon_t^{-1}\mathbf{a}_t^{(i)}), \tag{46}$$

with covariance matrices of size $n \times n$ and where $\mathbf{a}_t^{(i)}$ stand for samples given respectively by the innovation ($\mathbf{d}_{t=k}^{(i)} = \mathbf{x}_{t=k}^{(i)} - \mathbf{y}_{t=k}^o$) for the likelihood, the model noise variables $\boldsymbol{\eta}_t^{(i)}$ for the transition distribution, and the analyzed particles resulting from the ensemble Kalman process for the proposal distribution. These three conditional distributions are formally defined according to respective σ-algebras $\mathcal{F}_t$. In the following we drop the time index for sake of notations simplification.

The inversion of $n \times n$ full rank covariance matrices cannot be reasonably performed for high dimensional problems. To deal with this recurrent problem, covariance matrices are represented on the basis of a reduced set of pseudo-random fields under-sampling the high dimensional state space (Bishop et al., 2001; Evensen, 2003). This consists in representing the matrix $\boldsymbol{\Upsilon}$ by the set of associated samples $\mathbf{a}^{(i)}, i = 1, \ldots, N$ previously introduced.

Given $\mathbf{A}$ the $n \times N$ matrix whose columns are the $\mathbf{a}^{(i)}$, the matrix Υ is approximated by its empirical estimation:

$$\boldsymbol{\Upsilon} \sim \boldsymbol{\Upsilon}^e = \frac{\mathbf{A}\mathbf{A}^T}{N-1}. \tag{47}$$

The rank of the matrix $\boldsymbol{\Upsilon}^e$ is lower or equal to N. Hence, $\boldsymbol{\Upsilon}^e$ is not invertible. A pseudo inverse can nevertheless be defined from a reduced Singular Value Decomposition of matrix $\mathbf{A}$ called the Thin SVD. The thin SVD decomposes the $n \times N$ matrix $\mathbf{A}$ as the product $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{W}^T$ where only the N first column vector of the $n \times n$ orthogonal matrix $\mathbf{U}$ are computed. Thus $\mathbf{U}$ is $n \times N$, $\mathbf{W}$ is an orthogonal matrix of size $N \times N$, and $\mathbf{S}$ is a $N \times N$ diagonal matrix gathering the N first eigenvalues of $\mathbf{A}$. Defining the pseudo inverse $\mathbf{S}^+$ of the matrix $\mathbf{S}$ such as:

$$\begin{aligned} S_{jj}^+ &= 1/S_{jj} \quad \text{if } S_{jj} \neq 0, \\ S_{jj}^+ &= 0 \quad \text{otherwise,} \end{aligned}$$

an approximation of the Gaussian distribution can be defined from the pseudo inverse of matrix Υ^{e+}:

$$\begin{aligned} p(\mathbf{a}^{(i)}|\mathcal{F}) &\propto \exp(-\frac{1}{2}\mathbf{a}^{(i)T}\boldsymbol{\Upsilon}^{e+}\mathbf{a}^{(i)}) \\ &\propto \exp(-\frac{N-1}{2}\mathbf{a}^{(i)T}\mathbf{U}\mathbf{S}^+\mathbf{S}^+\mathbf{U}^T\mathbf{a}^{(i)}). \end{aligned} \tag{48}$$

The evaluation of the probability above is thus performed through the following steps:

(i) Singular Value Decomposition of A.
(ii) Computation of $\mathbf{S}^+$.
(iii) Computation of $\mathbf{a}_1^{(i)} = \mathbf{U}^T\mathbf{a}^{(i)}$ for all i ($N \times 1$ vectors).
(iv) Computation of $\mathbf{a}_2^{(i)} = \mathbf{S}^+\mathbf{a}_1^{(i)}$ ($N \times 1$ vectors).
(v) Computation of the scalar $a_3^{(i)} = \mathbf{a}_2^{(i)T}\mathbf{a}_2^{(i)}$
(vi) The required quantity is $\exp(-\frac{N-1}{2}a_3^{(i)})$.

In this process, none step has a complexity greater than $n \times N$, and none matrix greater than $n \times N$ needs to be stored in memory. The likelihood $p(\mathbf{y}_k^o|\mathbf{x}_k^{(i)})$ can then be evaluated through this procedure. The covariance is approximated with samples of the observation noise and the likelihood is then evaluated at point $\mathbf{a}_k^{(i)} \triangleq \mathbf{x}_k^{(i)} - \mathbf{y}_k^o$.

In high dimensional problems, it is important to note that usually the sample dimension is much lower than the state space dimension. In this case, if the probability is evaluated from the same samples as those used to approximate the covariance matrix (and drawn from the target Gaussian distribution), the obtained distribution is uniform for the different particles. The demonstration of this statement is given in the appendix. As opposed to the likelihood, we fall in this case when evaluating the *a priori* and the proposal distribution. As a matter of fact, these two distributions are evaluated and approximated from Gaussian i.i.d. samples. Indeed $\boldsymbol{\eta}_t^{(i)}$ and $\mathbf{x}_k^{(i)}$ are used for the *a priori* and the proposal distributions respectively. Their evaluations are not necessary as their values have no influence on the weights after their normalization.

3.4.2 Particles resampling and smoothing The resampling stage performed by the WEnKF method leads to discard particles with low weights and to duplicate particles associated to high weight values. As a consequence several groups of identical particles may be generated. These groups constitute sets of identical initial conditions for the next forecast. As the aim of particles ensemble is to explore configurations of a gigantic state space in comparison to the ensemble cardinality, such a crude discretization of the initial condition yields an immediate loss of efficiency. To attenuate this potential problem, we smooth the ensemble members distribution by adding a zero mean Gaussian field perturbation whose covariance depends on the mean discrepancy between the estimate and the current measurement given the whole history of measurements. The smoothed initial conditions $\tilde{\mathbf{x}}_k^{(i)}$ for the forecast are defined as:

$$\tilde{\mathbf{x}}_k^{(i)} = \mathbf{x}_k^{(i)} + \boldsymbol{\eta}^{(i)}, \tag{49}$$

$$\boldsymbol{\eta}^{(i)} \sim \mathcal{N}(0, \Lambda), \tag{50}$$

$$\Lambda = (E(\|\mathbf{x}_k - \mathbf{y}_k^o\|^2|\mathbf{y}_{1:k}^o) + \alpha)^{1/2}\overline{\mathbf{Q}}, \tag{51}$$

$$\simeq (\sum_i w_k^{(i)}\|\mathbf{x}_k^{(i)} - \mathbf{y}_k^o\|^2 + \alpha)^{\frac{1}{2}}\overline{\mathbf{Q}}. \tag{52}$$

In this expression $\overline{\mathbf{Q}}$ is a covariance matrix with variance 1 and that has the same structrure as the model noise covariance matrix $\mathbf{Q}_{\Delta t}$. The parameter α, set in practice to a small value avoids degeneracy problems caused by a null discrepancy of a single particle associated to unity weight. The proposed adaptation of the resampling comes to introduce an uncertainty on the forecast's initial states, which depends on the filtering distribution but also on the sample cloud's spread around the current measurements. Here the additional Gaussian variable constitutes a "mean-field" variable as the covariance matrix is defined from the filtering distribution. Each draw of the Gaussian variable is nearly independent from the individual particles and the law of the

perturbed variable is such that:

$$\tilde{\mathbf{x}}_k^{(i)} \sim p(\mathbf{x}_k^{(i)}|\mathbf{y}_{1:k}^o) \star \mathcal{N}(0,\Lambda), \tag{53}$$

where $\star$ denotes the convolution product. The considered perturbation corresponds thus to a smoothing of the filtering distribution by a Gaussian kernel. A similar strategy has been proposed and analyzed in (Musso et al., 2001). It is also reminiscent to the adaptation proposed in (Xiong et al., 2006) which consists in applying a Gaussian resampling to the particle nest. We will show in the following section that the proposed smoothing allows stabilizing efficiently the filter in a high dimensional experiment. This smoothing may be also related to the technique proposed in (Hamill and Snyder, 2000), which consider a Gaussian perturbation to the ensemble members (and therefore to smooth the filtering distribution) through the definition of a background covariance as a weighted mean of the ensemble covariance and a typical background covariance used in 3DVAR assimilation of Quasi-Geostrophic model. The perturbation considered is however constant in time and requires the tuning of a weighting coefficient. Let us remark that for dynamics defined only up to a random initial condition and in which the model noise covariance does not have to be specified, covariance $\overline{\mathbf{Q}}$ may be settled to the empirical covariance of the particles. In that case this comes thus to implement an inflation procedure similar to those carried out in EnKF implementations with a factor defined from the innovation.

In the following sections, the proposed adaptation of the particle filter is assessed on different examples. We first study its performance on linear or non-linear scalar problems for which measurements are available at every discrete instant. The technique will be afterward assessed on a higher dimensional state space for the assimilation of a 2D turbulent flow. For all these experiments we choose to rely on an implementation of EnKF that incoporates perturbed observations (Evensen, 1994; Houtekamer and Mitchell, 1998; Burgers et al., 1998; Evensen, 2003).

4 Experiments on scalar problems

In this section, the EnKF and WEnKF are compared on 1D examples with linear or non-linear dynamics. Comparisons are done from the following metrics:

- For the linear case, both methods are compared to the Kalman filter which gives the optimal solution. Denoting x_k^a the Kalman mean estimate, the error is computed between the mean of the distribution given by the Kalman filter and the one obtained by both ensemble techniques (EnKF or WEnKF), and reads $\mathrm{e}(k) = (x_k^a - \sum_{i=1}^{N} w_k^{(i)} x_k^{(i)})^2$ at a given time k, where $w_k^{(i)} = \frac{1}{N}$ for the EnKF. A similar error metric is used to compare the variances estimated by both methods to the reference Kalman variance P_k^a.
- For the non-linear case, since the Kalman optimal filtering distribution is unknown, estimated means of both methods are simply compared to the reference state x_k^{true} (ground truth): $\mathrm{e}(k) = (x_k^{true} - \sum_{i=1}^{N} w_k^{(i)} x_k^{(i)})^2$. Estimated variances are also computed in order to compare the resulting dispersion of both methods, but can not be compared to a ground truth.

4.1 Linear dynamics

As a first example, we study scalar scenarios with a linear dynamics, for which we know the optimal filtering distribution, as it is provided by the classical Kalman filter. The state x_k is evolving along the dynamical model:

$$x_k = x_{k-1} + \eta_k,$$

where η_k is a white Gaussian noise with variance σ_Q. The measurement model is given by:

$$z_k = x_k + \epsilon_k, \tag{54}$$

where ϵ_k is a white Gaussian noise with variance σ_R. We also assume a Gaussian distribution for the initial state (at time $k = 0$), centered at point 0 with a variance σ_B.

Different random realizations of the dynamics and the measures have been obtained from different values of σ_Q, σ_R and σ_B, in order to get different scenarios for 30 successive time steps. The assimilation of these scenarios have been then carried out using the EnKF and the WEnKF. For the WEnKF a systematic resampling has been considered. The weights updating rule is thus defined through equation (44), with Σ_k^e given by (38).

As the simulation of the random variables may play an important role (for each particle, there are two random draws for the dynamics noise and the measurement noise), the two filtering techniques are performed in parallel with the same random realizations. In order to have mean comparisons between both techniques a great number of trials have been realized (the experiments have been run 5000 times). The errors have then been averaged over the 30 time steps and all trials. Moreover, we indicate the number of occurences the WEnKF gives better results than the EnKF (closer to the analytical result given by Kalman) for the mean and variance estimations. The results obtained for 10 particles (or members) on different scenarios are summarized in Table 1, where the first table represents the results related to the estimation of the mean, and the second table compares the results for the variance estimation.

In these first experiments, as we deal with linear Gaussian systems, the filtering distribution is also Gaussian. The assumptions of the Ensemble Kalman filter are thus fully satisfied. This case should be thus favorable to the Ensemble Kalman filter. We observe, however, that the introduction of particles weighting enables to enhance the results of the mean estimation in most experiments. In order to improve the readability of Table 1, let us describe more precisely the results associated to a specific scenario. For instance, in the third line corresponding to the scenario $\sigma_Q = 1$, $\sigma_R = 1$ and $\sigma_B = 1$, the WEnKF gives better results (*i.e.* closer to the optimal result given by the Kalman filter) in 56.1% of the runs with a mean squared error of 0.106. Weighting the particles enables to take into account the fact that the estimated ensemble Kalman gain is only an approximation of the true gain. It also allows to take into account the probability of forecast realizations. On the second table of Table 1, we observe however that the error associated to the variance estimation tends to be higher for the WEnKF method. It can also be noticed that the empirical variance of the WEnKF is smaller than the EnKF one for all scenarios.

Note that a significant enhancement of the results of the WEnKF can be obtained by an empirical evaluation of the

Gaussian proposal distribution associated to the variables $\delta\mathbf{x}_k^{(i)} = \mathbf{x}_k^{(i)} - \boldsymbol{\mu}^{(i)}$ in (43) instead of relying on its analytical expression (37). The empirical mean $\overline{\delta\mathbf{x}}_k$ and the empirical variance $\hat{\Sigma}_k^e$ are computed through relations:

$$\overline{\delta\mathbf{x}}_k = \frac{1}{N}\sum_{i=1}^{N} \delta\mathbf{x}_k^{(i)},$$

and

$$\hat{\Sigma}_k^e = \frac{1}{N-1}\sum_{i=1}^{N}\left(\delta\mathbf{x}_k^{(i)} - \overline{\delta\mathbf{x}}_k\right)\left(\delta\mathbf{x}_k^{(i)} - \overline{\delta\mathbf{x}}_k\right)^T,$$

and the weights are then updated according to:

$$w_k^{(i)} \propto \frac{p(\mathbf{y}_k^o|\mathbf{x}_k^{(i)})p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{\mathcal{N}\left(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k^{(i)}; \overline{\delta\mathbf{x}}_k, \hat{\Sigma}_k^e\right)}. \tag{55}$$

The actual proposal distribution used is then more correctly described with respect to the particles and allows a compensation of the simulation errors of the random realizations[1].

The Table 2 gives the results obtained for the same scenarios as Table 1 but using this empirical proposal distribution. The results obtained by the WEnKF for the mean estimation are significantly improved for all the scenarios (the ratio of favorable tests increased and reach a ratio greater than 90% in most scenarios). The variance estimation is also improved, leading to better results for the WEnKF method in most cases. We observe also that the empirical variance of the WEnKF is smaller for all configurations.

Let us remark that in both techniques the same inflation procedure as the one defined in section 3.4.2 has been implemented. For the linear dynamics, its influence revealed however to be minor. The beneficial of this smoothing is much more noticeable in the non-linear case (see table 3 and 4).

By increasing the number of particles, the mean squared error is decreasing for both methods, and both converge to the analytic result of the Kalman filter. Figure 1 illustrates this decrease for four different scenarios of Table 2. We observe that the error for the WEnKF always converges faster toward zero, except for the case in which the model noise σ_Q is very small with respect to the measurement noise σ_R, where both performances are close. This case is the most defavorable configuration for the WEnKF, already observed in Tables 1 and 2 for $N = 10$ particles).

4.2 Non-linear dynamics

We now perform an evaluation of the WEnKF for a non-linear dynamics associated to a linear Gaussian measurement. The following scalar dynamics has been chosen:

$$x_k = \sin(3x_{k-1}) + \eta_k,$$

where η_k is a white Gaussian noise with variance σ_Q. This dynamics is associated to the same measurement model (54) as before. The initial state is still defined as a zero mean Gaussian distribution with variance σ_B.

Five scenarios with different choices of parameters have been created. As previously, a great number of independent trials (5000) have been realized and their results have been averaged.

Two different tables of results are presented (Table 3 and 4), that sum up the results obtained for the five non-linear scenarios studied with $N = 10$ particles. Both have been obtained computing the weights of the WEnKF particles according to expression (55), where the empirical form of the proposal distribution is used in order to compensate the approximation error and improve results. The difference in the two sets of results consists in the use of the inflation procedure (smoothing of the particles distribution, described in section 3.4.2) for Table 4, but not for Table 3. We can note that the introduction of the smoothing procedure tends to reduce the mean squared error for both methods. This highlights the fact that an improvement of the EnKF procedure may lead to an improvement of the results of the WEnKF, through a better description of the proposal distribution. We also observe that the empirical variances are slightly increased by the smoothing procedure for both methods, but the empirical variance computed from the WEnKF distribution remains smaller, showing that the WEnKF tends to concentrate more the particles than the EnKF.

We observe in these two tables that the results obtained by the WEnKF are better or comparable to the results of the EnKF. The performances of the WEnKF are better in average when the model noise σ_Q is not too small with respect to the initial noise σ_B or the measurement noise σ_R.

Figure 2 clearly shows that the two methods do not always converge toward the same filtering distribution when the number of particles increases. These experimental results tend to confirm the asymptotical results described in (Le Gland et al., 2010) which state that the EnKF has a limiting distribution that differs from the optimal Bayesian filtering distribution.

5 Experiments on a high dimensional non-linear problem

The high-dimensional experiments are performed on an example with non-linear dynamics. As for the non-linear scalar experiment, the optimal filtering distribution is unknown. The performances of EnKF and WEnKF methods are then compared by means of the error between the mean of the estimated filtering distributions and the ground truth at each time step: $\mathrm{e}(k) = (\mathbf{x}_k^{true} - \overline{\mathbf{x}_k})^T(\mathbf{x}_k^{true} - \overline{\mathbf{x}_k})$, where $\overline{\mathbf{x}_k} = \sum_{i=1}^{N} w_k^{(i)}\mathbf{x}_k^{(i)}$, and $w_k^{(i)} = \frac{1}{N}$ for the EnKF method. Moreover, a dispersion criteria is computed for both methods as $d(k) = \sum_{i=1}^{N} w_k^{(i)}(\mathbf{x}_k^{(i)} - \overline{\mathbf{x}_k})^T(\mathbf{x}_k^{(i)} - \overline{\mathbf{x}_k})$.

The high dimensional experiments are realized on scenarios representing the evolution of a 2D turbulent flow. The flow evolution is described through the 2D incompressible vorticity-velocity formulation of the Navier-Stokes equations

[1] (Pham, 2001) proposes to control the random realizations so as to ensure they have the required mean and variance, through a "second order accuracy" sampling method. Weighting and using the empirical propagation in the WEnKF enables to compensate the simulation errors in the same way. We will see in the next paragraph that contrary to the second order accuracy sampling, the WEnKF also compensates the approximation error in the non-linear case.

with a stochastic forcing:

$$d\xi = -\nabla\xi \cdot \mathbf{w}dt + \frac{1}{Re}\Delta\xi dt + \sigma_Q dW, \tag{56}$$

where $\xi = v_x - u_y$ denotes the vorticity of the velocity field $\mathbf{w} = [u, v]^T$, and Re is the flow Reynolds number. The operators $\nabla\xi$ and $\Delta\xi$ denote the 2D gradient and Laplacian operators applied on the vorticity ξ. The random forcing term dW is an isotropic Gaussian fields correlated in space and uncorrelated in time. Its covariance operator is defined as:

$$Q(\mathbf{r}, \tau) = \mathbb{E}(dW(\mathbf{x}, t)dW(\mathbf{x}+\mathbf{r}, t+\tau)) = g_\lambda(\mathbf{r})dt\delta(\tau), \tag{57}$$

where $g_\lambda(\mathbf{r}) = \exp(-\|\mathbf{r}\|^2/\lambda^2)$ describes the spatial correlation structure and the delta Dirac function stands for the temporal correlation structure. The Gaussian function plays here the role of a cut-off function. Different numerical methods have been proposed to generate such random fields with additional power law constraints on the power spectrum (Elliot et al., 1997). In this work we used a simplified procedure similar to the one proposed in (Evensen, 1994).

Furthermore, let us recall that introducing the Biot-Savart integral, the velocity field can be recovered from its vorticity:

$$\mathbf{w} = \nabla G * \xi, \tag{58}$$

where $G(\mathbf{x}) = \frac{\ln(|\mathbf{x}|)}{2\pi}$ is the Green kernel associated to the Laplacian operator.

To achieve an accurate and stable discretization of the advective term $\nabla\xi \cdot \mathbf{w}$, we have used a conservative numerical scheme (Shu, 1998). Such schemes respect exactly the conservation law within the discretization cell by integrating the flux value at cell boundaries. Total Variation Diminishing (TVD) schemes (which are monotonicity preserving flux) prevent from an increase of oscillations over time and enable to transport shocks. All these methods are well detailed in (Shu, 1998). The time integration is realized with a third-order Runge Kutta scheme, which also respects the TVD property (Shu, 1998). The motion field is updated at each time step in the Fourier domain with equation (58). With this whole non-oscillatory scheme, the vorticity-velocity equations can be integrated on a domain of interest. In this work we considered a squared 2D spatial domain of 64×64 and a temporal range $k \in [0, 100]$. This model has been simulated starting from a fixed initial condition built from the simulation of a Gaussian random field of covariance $\sigma_B^2 B$, with $B(\mathbf{x}, \mathbf{y}) = g_\lambda(\mathbf{x} - \mathbf{y})$. The samples have been drawn with the same procedure as for the model noise. Examples of some vorticity map realizations obtained for a given "eddy diffusivity" value, σ_Q, and different initial noise amplitudes, σ_B, are presented in figure 3.

The linear measurement model (54) is still used. However we considered that these observations where available only at discrete measurement instants $k \in \mathbb{N}, k + \Delta k \leqslant 100$ with the measurements latency $\Delta k = 1, 2, 3, 4, 5$. The length of the time integration interval Δt is given by the Runge-Kutta process (a typical value is $\Delta t = 0.1 << \Delta k$) and is fixed as the inverse of the velocity field infinite norm in order to garantee a stable scheme (Shu, 1998). It is thus not directly defined as a fixed portion of the measurement interval Δk. The observations, with spatial resolution 64×64, have been generated applying different levels of Gaussian noise of covariance $R(\mathbf{x}, \mathbf{y}) = \sigma_R g_\lambda(\mathbf{x} - \mathbf{y})$ to the simulated velocity fields. As we used three different values for $\sigma_B = \sigma_R$ and σ_Q, this provides us a benchmark of 9 scenarios to assess our filtering technique. A value of $\lambda = 3$ has been used for all these scenarios. Table 5 sums up the different parameters used to build this benchmark.

The two filters have been run and compared for the very same number of particles (or ensemble members). Reduced sets of particles have been used in the experiments. We have fixed the number N of particles according to the measurement latency as follows: $N = 20 \times \Delta k$, since both methods need more particles to avoid divergence when the measurement latency gets larger. For both methods, 100 independent trials have been run. The smoothing procedure of the particles distribution (presented in section 3.4.2 and which can be related to covariance inflations approaches usually applied in EnKF methods) has been applied to both methods, aiming at avoiding the divergence of the filters when the number of particles is small and the measurements latency gets larger.

Before examining in detail the results, let us first point out that with the smoothing procedure introduced both techniques converged in 100% of the trials. Without this smoothing, for a small number of particles and large measurements latency the WEnKF was unstable and diverged for a great proportion of the trials. As an example, for $N = 60$ particles and $\Delta k = 5$ we observed filter divergence in 87% of the trials, whereas the filter converges in 100% of the trials in the same configuration and the smoothed resampling scheme. This smoothing appears thus to be essential for the WEnkF.

The quantitative results of experimentations are presented in Table 5. The error values given for each method represent the errors $e(k)$ averaged over all time steps and trials, and similarly for the dispersion criteria $d(k)$. The behaviors of the filtering techniques, for different measurement latencies $\Delta k \in \{1, 2, 3, 4, 5\}$, are presented in figures 4 to 8. For each measurement latency Δk, the associated nine plots illustrate along time the mean squared error levels reached by the EnKF and WEnKF for the 9 studied scenarios. These results have been averaged over 100 independent trials.

As it can be observed from Table 5 and the different plots, the WEnKF provided better results than the EnKF filter in most cases. Let us describe the conclusions that can be made from these experiments:

- When the latency Δk between observations is small both filters give globally comparable results. This may be observed in Table 5 where the errors and the dispersion are quasi identical. For small measurement time intervals the implicit linear assumption of the EnKF seems to work well for the dynamics carried out here. For small integration steps (in between two observations) the system does not suffer from strong pertubations and behave more like a deterministic dynamics with an uncertainty on the initial condition.
- When Δk increases, both methods need some time steps to stabilize and start to converge, due to the fact that particles are not guided nor corrected toward the observations on greater time intervals.
- However, when the measurement latency Δk gets larger, the WEnKF method converges more quickly than

the EnKF, and always gives a smaller error, for all parameters configurations tested. Note that the greater Δk, the faster the WEnKF converges in comparison with the EnKF. For large Δk an ineffective Gaussian assumption of the ensemble members distribution probably penalizes the EnKF and slows down its convergence rate.

- Finally, we can observe from Table 5 that the dispersion criteria is smaller for the WEnKF method for almost all scenarios. This has already been observed in the scalar case when comparing estimated variances, and confirms that the WEnKF method tends to concentrate more the particles.

All these results allow us to conclude that the WEnKF strategy leads globally to a gain of efficiency compared to the traditional EnKF implementation. This gain is provided with a comparable computational cost. Let us remark that convergences over different limiting distribution have not been observed in this last experiment opposite the non linear scalar case. This may be an indication that both filters still converge toward an inexact filtering distribution. Several directions of improvements could probably be investigated in the future. First of all, the approximation of the covariance matrix through random draws shows some limitations as the filters are in a way blind to the quality of the forecasts and proposals with respect to the previous state and the new measurements. Other approximations will have on the other hand to tackle the difficult issue of a relevant and efficient approximations/evaluations of the weights computation. Otherwise, as an augmentation of the number of particles cannot be a solution for geophysical applications, efficient dimension reduction approaches allowing a clever repartition of the particles cloud have to be investigated. Strategies coupling variational assimilation and ensemble based technique may provide some improvements for small measurements latencies or asynchronous observations and would be worth investigating.

6 Conclusions

In this paper two sequential Monte Carlo filters have been studied: the ensemble Kalman filter and a particular version of a particle filter adapted to high dimensional problems. The advantages and drawbacks of both methods have been described and analyzed.

The particle filter proposed in this paper borrows to the ensemble Kalman filter its ensemble members sampling strategy. This modification consisted in defining the particle filter proposal density from this ensemble members sampling scheme. The weights corresponding to the ratio between the Bayesian filtering and the proposal distribution are then updated according to this "ensemble Kalman" proposal density. The sequential Bayesian filter resulting from this strategy constitutes *in fine* a simple extension of Ensemble Kalman filters.

It has been shown experimentally in this paper that such an extension allows improving the results obtained by a traditional implementation of ensemble Kalman filtering. Also as a by product of the resampling used in the particle filter, the extended ensemble filter proposed, provides in general a faster error decrease along time.

The weighted ensemble Kalman filter has been tested on linear and non linear scalar scenarios. We also applied the method for the assimilation of 2D flow scenarios. For this last example, we showed that with a smoothed resampling scheme and an adapted implementation strategy the proposed modification of the particle filter was well suited to a high dimensional problem. In several cases, this filter has shown to outperform the traditional ensemble Kalman filtering with a comparable computational load. The good behavior of the WEnKF technique for large measurements latency could probably be beneficial for geophysical applications for which the time interval between two observations is long.

In future works, we aim at directly considering image observations, as it has been done in (Corpetti et al., 2009) in the context of a variational assimilation strategy or introducing image based non-linear measurements and sampling relying on the dynamics stable and unstable directions (Cuzol and Memin, 2009). We also plan to test this methodology on realistic meteorological or oceanic forecast models.

Acknowledgements

The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant PREVASSEMBLE (ANR- 08-COSI-012) "Méthodes d'Ensemble pour l'Assimilation des Observations et la Prévision en Météorologie et Océanographie".

APPENDIX A: A simplification of the particle weights computation

The weights computation (44) requires to evaluate the proposal density distribution $\mathcal{N}\left(\mathbf{x}_k^{(i)} - \overline{\boldsymbol{\mu}}_k^{(i)}; 0, \boldsymbol{\Sigma}_k^e\right)$ and the *a priori* distribution $p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})$ (45). These distributions require to evaluate Gaussian distributions through a small number of samples. We are going to prove that for ensembles of dimension much smaller than the size of the state space, all the samples are equiprobable with respect to these distributions.

Both distributions require to evaluate Gaussian distributions $\mathcal{N}(\boldsymbol{\gamma}_k^{(i)}; 0, \Sigma_k^e)$ centered on $\overline{\boldsymbol{\mu}}_k^{(i)}$ (32) and $\mathbf{x}_{t-\Delta t}^{(i)}$ with covariance matrices $\mathbf{K}_k^e \mathbf{R} \mathbf{K}_k^{eT}$ and $\mathbf{Q}_{\Delta t}$ respectively for the proposal and the transition distributions involved in the expression of the *a priori* distribution (45).

Of course, in the latter case, successive products at every time steps and huge integrals have to be performed in addition (45). Nevertheless, let us focus first on the inner Gaussian expressions involved in the evaluation of this distribution.

As the process will be the same for both distributions, we here rely on a general formulation. For samples of Gaussian distribution $\mathbf{y}^{(i)}$, $i = 1 \cdots N$, of mean 0 and variance Σ, we have:

$$\mathcal{N}(\mathbf{y}^{(i)}; 0, \boldsymbol{\Sigma}) \propto \exp\left(-\frac{1}{2}\mathbf{y}^{(i)T} \left(\boldsymbol{\Sigma}\right)^{-1} \mathbf{y}^{(i)}\right). \qquad \text{(A1)}$$

Let $\boldsymbol{\Gamma}$ be a $n \times N$ matrix whose columns are samples of the target density $\mathbf{y}^{(i)}$, the covariance matrix is approximated

by its empirical extension, through the method previously described in paragraph 3.4.1:

$$\mathbf{\Sigma}^e \simeq \frac{\mathbf{\Gamma}\mathbf{\Gamma}^T}{N-1}. \tag{A2}$$

In order to keep simple notations, the thin Singular Value Decomposition of $\mathbf{\Gamma}$ is written with the same notations as those used previously (§3.4.1). We then have $\mathbf{\Gamma} = \mathbf{U}\mathbf{S}\mathbf{W}^T$ and the previous expression reads:

$$\mathbf{\Sigma}^e \simeq \frac{\mathbf{U}\mathbf{S}\mathbf{S}^T\mathbf{U}^T}{N-1}. \tag{A3}$$

We also denote by $\mathbf{\Omega}$ the matrix whose components Ω_{ij} are defined as:

$$\mathbf{\Omega} = \frac{1}{2}\mathbf{\Gamma}^T \left(\mathbf{\Sigma}^e\right)^{-1} \mathbf{\Gamma}, \tag{A4}$$

and such that the weight contribution for each sample i, defined by the equation (A1), is proportional to $\exp(-\Omega_{ii})$. By replacing the matrix by its approximation (A3), we do not need to realize an inversion but a pseudo inversion:

$$\begin{aligned}
\mathbf{\Omega} &= \frac{1}{2}\mathbf{\Gamma}^T \left(\frac{\mathbf{U}\mathbf{S}\mathbf{S}^T\mathbf{U}^T}{N-1}\right)^+ \mathbf{\Gamma}, &\text{(A5)}\\
&= \frac{N-1}{2}\mathbf{\Gamma}^T\mathbf{U}\left(\mathbf{S}\mathbf{S}^T\right)^+\mathbf{U}^T\mathbf{\Gamma}, &\text{(A6)}\\
&= \frac{N-1}{2}\mathbf{W}\mathbf{S}\mathbf{S}^+\mathbf{S}^+\mathbf{S}\mathbf{W}^T. &\text{(A7)}
\end{aligned}$$

The matrix $\mathbf{\Gamma}$ being of size $n \times N$, with $N << n$, its rank (and its number of non null singular values) is lower than or equal to N. As the columns of $\mathbf{\Gamma}$ are provided by the simulation of a centered Gaussian distribution, the sum of the columns will be close to zero and the rank lower than or equal to $N-1$. We assume that the rank is equal to $N-1$ since the set of samples is assumed to represent a space of dimension much greater than N and particles alignment would constitute a loss of efficiency.

We then have $\mathbf{S}\mathbf{S}^+ = \mathbf{1}_N^{N-1}$, where $\mathbf{1}_n^m$ defines a matrix of size $n \times n$ with the m first diagonal terms equal to 1 and the others are 0. One has:

$$\mathbf{\Omega} = \frac{N-1}{2}\mathbf{W}\mathbf{1}_N^{N-1}\mathbf{W}^T, \tag{A8}$$

and we deduce that

$$\Omega_{ii} = \frac{N-1}{2}\sum_{k=1}^{N-1} W_{ik}^2. \tag{A9}$$

As the columns of $\mathbf{W}$ are normalized vectors, this expression reads:

$$\Omega_{ii} = \frac{N-1}{2}\left[1 - W_{iN}^2\right]. \tag{A10}$$

This quantity depends on the components of the last vector of the orthogonal matrix $\mathbf{W}$, *i.e.* the one corresponding to the N-th singular value, which is equal to zero. This vector is thus in the kernel of the matrix $\mathbf{\Gamma}$. The kernel is spanned by vector $(1, 1, \ldots, 1)^T$, noting the set of samples has a null mean ($(1, 1, \ldots, 1)\mathbf{\Gamma} = 0$) and according to the kernel definition.

Furthermore, as the columns of W are normalized, $W_{iN}^2 = 1/N$. We then conclude that $\Omega_{ii} = \Omega_{jj}, \forall (i,j) \in [1, N]^2$.

In relation (44), the proposal density distribution $\mathcal{N}\left(\mathbf{x}_k^{(i)} - \overline{\boldsymbol{\mu}}_k^{(i)}; 0, \mathbf{\Sigma}_k^e\right)$ and the inner Gaussian transition distributions $p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-\Delta t}^{(i)})$ involved in the *a priori* distribution (45) are thus independent of the particle index (when the number of particles is inferior to the size of the state space). For the different ensemble members the proposal distribution and the *a priori* distribution have equiprobable values, their computations in the weight updating rule can be thus ignored.

REFERENCES

Anderson, J., 2001: An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129(12), 2884–2903.

Anderson, J., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, 59A, 210–224.

Anderson, J. and S. Anderson, 1999: A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127(12), 2741–2758.

Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Monthly Weather Review*, 131(4), 634–642.

Arnaud, E. and E. Mémin, 2007: Partial linear gaussian model for tracking in image sequences using sequential monte carlo methods. *International Journal of Computer Vision*, 74(1), 75–102.

Arulampalam, M., S. Maskell, N. Gordon, and T. Clapp, 2002: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Processing*, 50(2).

Bennet, A., 1992: *Inverse Methods in Physical Oceanography*. Cambridge University Press.

Bertino, L., G. Evensen, and H. Wackernagel, 2003: Sequential data assimilation techniques in oceanography. *Int. Statist. Rev.*, 71(2), 223–241.

Bishop, C., B. Etherton, and S. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. part i: Theoretical aspects. *Monthly weather review*, 129(3), 420–436.

Burgers, G., P. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126, 1719–1724.

Corpetti, T., P. Heas, E. Memin, and N. Papadakis, 2009: Pressure image asimilation for atmospheric motion estimation. *Tellus*, 61A, 160–178.

Courtier, P., 1997: Dual formulation of four-dimensional variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, 123, 2249–2461.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingswort, F. Rabier, and M. Fisher, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-VAR). part 1: formulation. *Quarterly Journal of the Royal Meteorological Society*, 124, 1783–1807.

Crisan, D. and A. Doucet, 2002: Survey of convergence results on particle filering methods for practitioners. *IEEE Trans. Signal Processing*, 50(3), 736–746.

Cuzol, A. and E. Memin, 2009: A stochastic filter technique for fluid flows velocity fields tracking. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 31(7), 1278–1293.

Del Moral, P., 2004: *Feynman-Kac Formulae Genealogical and Interacting Particle Systems with Applications*. Springer, New York; Series: Probability and Applications.

Doucet, A., S. Godsill, and C. Andrieu, 2000: On sequential monte carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3), 197–208.

Elliot, F., D. Horntrop, and A. Majda, 1997: A Fourier-wavelet

Monte Carlo method for fractal random fields. *Journal of Computational Physics*, 132(2), 384–408.

Evensen, G., 1994: Sequential data assimilation with a non linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99 (C5)(10), 143–162.

Evensen, G., 2003: The ensemble Kalman filter, theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4), 343–367.

Evensen, G., 2006: *Data assimilation: The ensemble Kalman filter*. Springer-Verlag, New-york.

Evensen, G. and P. van Leeuwen, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Monthly Weather Review*, 128(6), 1852–1867.

Fertig, E., J. Harlim, and B. Hunt, 2007: A comparative study of 4D-VAR and a 4D ensemble Kalman filter : perfect model simulations with lorenz-96. *Tellus*, 59A, 96–100.

Gauthier, P., P. Courtier, and P. Moll, 1993: Assimilation of simulated wind lidar data with a Kalman filter. *Monthly Weather Review*, 121(6), 1803–1820.

Gordon, N., A. Doucet, and J. D. Freitas, 2001: *Sequential Monte Carlo methods in practice*. Springer-Verlag.

Gordon, N., D. Salmond, and A. Smith, 1993: Novel approach to non-linear/non-gaussian bayesian state estimation. *IEEE Processing-F*, 140(2).

Hamill, T. and C. Snyder, 2000: A hybrid ensemble Kalman filter-3d variational analysis scheme. *Monthly Weather Review*, 128(8), 2905–2919.

Harlim, J. and B. Hunt, 2007a: Four-dimensional local ensemble transform Kalman filter : numerical experiments with a global circulation model. *Tellus*, 59A(5), 731–748.

Harlim, J. and B. Hunt, 2007b: A non-gaussian ensemble filter for assimilating infrequent noisy observations. *Tellus*, 59A, 225–237.

Harlim, J. and A. Majda, 2009: Catastrophic filter divergence in filtering nonlinear dissipative systems. *Communications Mathematical Sciences*, 7(3).

Houtekamer, P. and H. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129(1), 123–137.

Houtekamer, P. L. and H. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126(3), 796–811.

Houtekamer, P. L. and H. Mitchell, 2006: Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society*, 131(613), 3269–3289.

Houtekamer, P. L., H. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Monthly Weather Review*, 133(3), 604–620.

Hunt, B., E. Kalnay, E. Kostelich, E. Ott, D. Patil, T. Sauer, I. Szunyogh, J. Yorke, and A. Zimin, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, 56A, 273–277.

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

Julier, S. and J. Uhlmann, 1997: A new extension of the Kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, 182–193.

Kalman, R., 1960: A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82, 35–45.

Kalman, R. and R. Bucy, 1961: New results in linear filtering and prediction theory. *Transactions of the ASME - Journal of Basic Engineering*, 83, 95–107.

Kivman, G. A., 2003: Sequential parameter estimation for stochastic systems. *Nonlinear Processes in Geophys.*, 10, 253–259.

Kong, A., J. Liu, and W. Wong, 1994: Sequential imputations and Bayesian missing data problems. *Journal of the American Statistical Association*, 89(425), 278–288.

Kuhl, D., I. Szunyogh, E. J. Kostelich, D. J. Patil, G. Gyarmati, M. Oczkowski, B. R. Hunt, E. Kalnay, E. Ott, and J. A. Yorke, 2007: Assessing predictability with a local ensemble Kalman filter. *Journal of the Atmospheric Sciences*, 64, 1116–1140.

Le Dimet, F.-X. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, 38A, 97–110.

Le Gland, F., V. Monbet, and V. Tran, 2009: Large sample asymptotics for the ensemble Kalman filter. Technical Report 7014, INRIA.

Le Gland, F., V. Monbet, and V. Tran, 2010: Large sample asymptotics for the ensemble Kalman filter. In *Handbook on Nonlinear Filtering*, Crisan, D. and Rozovskii, B., editors. Oxford University Press (in press).

Liu, J., E. Fertig, H. Li, E. Kalnay, B. Hunt, E. Kostelich, I. Szunyogh, and R. Todling, 2008: Comparison between local ensemble transform Kalman filter and psas in the nasa finite volume gcm-perfect model experiments. *Nonlinear Processes in Geophysics*, 15(4), 645–659.

Lorenc, A., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Monthly Weather Review*, 109(4), 701–721.

Miyoshi, T. and S. Yamane, 2007: Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Monthly Weather Review*, 135(11), 3841–3861.

Musso, C., N. Oudjane, and F. Le Gland, 2001: Improving regularized particle filters. In *Sequential Monte Carlo Methods in Practice*, Doucet, A., de Freitas, N., and Gordon, N., editors, Statistics for Engineering and Information Science. Springer–Verlag, New York, 247–271.

Ott, E., B. Hunt, I. Szunyogh, A. Zimin, M. C. E.J. Kostelich, E. Kalnay, D. Patil, and J. A. Yorke, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, 56A, 415–428.

Patil, D., B. Hunt, and J. Carton, 2001: Identifying low-dimensional nonlinear behavior in atmospheric data. *Monthly Weather Review*, 129(8), 2116–2125.

Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Monthly Weather Review*, 129, 1194–1207.

Rémy, S., O.Pannekoucke, T. Bergot, and C. Baehr, 2010: Adaptation of a particle filtering method for data assimilation in a 1D numerical model used for fog forecasting. *Quarterly Journal of the Royal Meteorological Society*, in press.

Shu, C.-W., 1998: *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations*, volume 1697 of *Lecture Notes in Mathematics*. Springer Berlin / Heidelberg, 325–432.

Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson, 2008: Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, 136(12), 4629–4640.

Szunyogh, I., E. J. . Kostelich, G. Gyarmati, E. Kalnay, B. Hunt, E. Ott, E. Satterfield, and J. Yorke, 2008: A local ensemble transform Kalman filter data assimilation system for the NCEP global model. *Tellus*, 60A, 113–130.

Tippett, M., J. Anderson, C. B. Craig, T. Hamill, and J. Whitaker, 2003: Ensemble square root filters. *Monthly Weather Review*, 131(7), 1485–1490.

van der Merwe, R., N. de Freitas, A. Doucet, and E. Wan, 2001: The unscented particle filter. In *Advances in Neural Information Processing Systems 13*, 584-590.

van Leeuwen, P. J., 2002: Ensemble Kalman filters, sequential importance resampling and beyond. In *ECMWF workshop on the role of the upper ocean in the medium and extended range forecasting*, ECMWF, Reading, UK, 46–56.

van Leeuwen, P. J., 2003: A variance-minimizing filter for large-

scale applications. *Monthly Weather Review*, 131(9), 2071–2084.

van Leeuwen, P. J., 2009: Particle filtering in geophysical systems. *Monthly Weather Review*, early online release.

Whitaker, J. and T. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Monthly Weather Review*.

Whitaker, J., T. Hamill, X. W. Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP global forecast system. *Monthly Weather Review*, 136(2), 463–482.

Xiong, X., I. Navon, and B. Uzunoglu, 2006: A note on the particle filter with posterior gaussian resampling. *Tellus*, 58A, 456–460.

Zhou, Y., D. McLaughlin, and D. Entekhabi, 2006: Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Monthly Weather Review*, 134(8), 2128–2142.

Zupanski, M., 2005: Maximum likelihood ensemble filter: Theoretical aspects. *Monthly Weather Review*, 133(6), 1710–1726.

LIST OF TABLES

Table 1. Experiments on 1D linear problems. Results obtained from WEnKF and EnKF with $N = 10$ particles are compared for different sets of parameters (σ_Q,σ_R,σ_B). Mean squared estimation errors with respect to the Kalman mean estimate are presented on the first table, errors with respect to the Kalman variance estimate on the second table, together with a comparison of variances estimated by both methods. The mean squared errors have been computed on 5000 independent trials. On both tables, we also indicate the percentage of cases among all time steps and trials where the WEnKF estimate was closer to the Kalman estimate.

Scenario			Results (mean estimation)		
			% WEnKF	Error	
σ_Q	σ_R	σ_B	closer	WEnKF	EnKF
0.5	1	0.5	47%	0.092	0.089
1	1	5	49%	0.119	0.117
1	1	1	56.1%	0.106	0.114
1	1	10	48.5%	0.179	0.172
1	0.5	1	61.9%	0.056	0.065
1	0.2	1	63.6%	0.020	0.024
0.2	1	0.2	38.2%	0.134	0.096
10	2	10	66.5%	0.227	0.285

Scenario			Results (variance estimation)				
			% WEnKF	Error		Estimate	
σ_Q	σ_R	σ_B	closer	WEnKF	EnKF	WEnKF	EnKF
0.5	1	0.5	31.7%	0.078	0.060	0.446	0.496
1	1	5	28.9%	0.114	0.087	0.547	0.604
1	1	1	31.8%	0.107	0.085	0.554	0.597
1	1	10	34.9%	0.112	0.092	0.564	0.609
1	0.5	1	33.3%	0.034	0.028	0.333	0.352
1	0.2	1	38%	0.007	0.006	0.158	0.167
0.2	1	0.2	38.4%	0.044	0.035	0.331	0.371
10	2	10	35.7%	0.726	0.619	1.589	1.655

Table 2. Experiments on 1D linear problems with an empirical propagation law (updating of weights with (55)). Results obtained from WEnKF and EnKF with $N = 10$ particles are compared for different sets of parameters (σ_Q,σ_R,σ_B). Mean squared estimation errors with respect to the Kalman mean estimate are presented on the first table, errors with respect to the Kalman variance estimate on the second table, together with a comparison of variances estimated by both methods. The mean squared errors have been computed on 5000 independent trials. On both tables, we also indicate the percentage of cases among all time steps and trials where the WEnKF estimate was closer to the Kalman estimate.

Scenario			Results (mean estimation)		
			% WEnKF	Error	
σ_Q	σ_R	σ_B	closer	WEnKF	EnKF
0.5	1	0.5	88%	0.046	0.089
1	1	5	94.5%	0.042	0.117
1	1	1	93.5%	0.052	0.114
1	1	10	91.7%	0.053	0.172
1	0.5	1	98.8%	0.028	0.065
1	0.2	1	96.9%	0.010	0.024
0.2	1	0.2	69.6%	0.055	0.096
10	2	10	96.8%	0.096	0.285

Scenario			Results (variance estimation)				
			% WEnKF	Error		Estimate	
σ_Q	σ_R	σ_B	closer	WEnKF	EnKF	WEnKF	EnKF
0.5	1	0.5	50.6%	0.056	0.060	0.412	0.496
1	1	5	63.2%	0.076	0.087	0.515	0.604
1	1	1	54.9%	0.077	0.085	0.502	0.597
1	1	10	60.4%	0.079	0.092	0.514	0.609
1	0.5	1	72.6%	0.023	0.028	0.304	0.352
1	0.2	1	85.4%	0.004	0.006	0.145	0.167
0.2	1	0.2	44.1%	0.041	0.035	0.306	0.371
10	2	10	88.5%	0.387	0.619	1.467	1.655

Table 3. Experiments on 1D non-linear problems without use of an inflation procedure. Results obtained from WEnKF and EnKF with $N = 10$ particles are compared for different sets of parameters (σ_Q,σ_R,σ_B). Mean squared estimation errors with respect to the ground truth are presented, together with a comparison of variances estimated by both methods. The mean squared errors have been computed on 5000 independent trials. We also indicate the percentage of cases among all time steps and trials where the WEnKF estimate was closer to the ground truth.

Scenario			Results				
			% WEnKF	Error		Variance estimate	
σ_Q	σ_R	σ_B	closer	WEnKF	EnKF	WEnKF	EnKF
0.2	0.2	0.2	54.9%	0.197	0.202	0.131	0.142
0.2	0.2	1	52.7%	0.241	0.244	0.124	0.139
0.2	1	0.2	46.2%	0.610	0.601	0.368	0.392
1	0.2	1	64.4%	0.220	0.235	0.156	0.171
1	1	1	61.3%	0.683	0.712	0.529	0.573

Table 4. Experiments on 1D non-linear problems with use of an inflation procedure. Results obtained from WEnKF and EnKF with $N = 10$ particles are compared for different sets of parameters (σ_Q,σ_R,σ_B). Mean squared estimation errors with respect to the ground truth are presented, together with a comparison of variances estimated by both methods. The mean squared errors have been computed on 5000 independent trials. We also indicate the percentage of cases among all time steps and trials where the WEnKF estimate was closer to the ground truth.

Scenario			Results				
			% WEnKF	Error		Variance estimate	
σ_Q	σ_R	σ_B	closer	WEnKF	EnKF	WEnKF	EnKF
0.2	0.2	0.2	59.7%	0.135	0.146	0.138	0.148
0.2	0.2	1	49.5%	0.093	0.093	0.135	0.148
0.2	1	0.2	53.6%	0.568	0.570	0.391	0.398
1	0.2	1	65.9%	0.224	0.241	0.158	0.172
1	1	1	64%	0.455	0.487	0.531	0.573

Table 5. Experiments on a high-dimensional non-linear problem. Results obtained from WEnKF and EnKF are compared for different sets of parameters (σ_B,σ_R,σ_Q) and different measurement latencies Δk (associated to different numbers of particles $N = 20 \times \Delta k$). Mean squared estimation errors with respect to the ground truth are presented, together with a comparison of the dispersion criteria computed for both methods. The mean squared errors have been computed on 100 independent trials.

Scenario			Results			
			Error		Dispersion	
$\sigma_B = \sigma_R$	σ_Q	Δk	WEnKF	EnKF	WEnKF	EnKF
0.01	0.01	1	0.006	0.007	0.015	0.016
		2	0.013	0.042	0.039	0.137
		3	0.009	0.026	0.027	0.072
		4	0.011	0.031	0.037	0.149
		5	0.013	0.047	0.046	0.531
0.01	0.05	1	0.016	0.017	0.025	0.027
		2	0.022	0.052	0.049	0.150
		3	0.018	0.035	0.039	0.083
		4	0.021	0.042	0.050	0.168
		5	0.023	0.059	0.060	0.625
0.01	0.1	1	0.048	0.048	0.057	0.057
		2	0.055	0.083	0.085	0.182
		3	0.051	0.065	0.078	0.122
		4	0.052	0.071	0.091	0.205
		5	0.052	0.095	0.101	0.712
0.05	0.01	1	0.007	0.008	0.016	0.017
		2	0.013	0.043	0.038	0.138
		3	0.009	0.026	0.027	0.072
		4	0.011	0.034	0.037	0.154
		5	0.013	0.048	0.046	0.559
0.05	0.05	1	0.018	0.019	0.027	0.029
		2	0.025	0.056	0.052	0.153
		3	0.020	0.037	0.041	0.086
		4	0.022	0.046	0.052	0.171
		5	0.024	0.066	0.062	0.666
0.05	0.1	1	0.053	0.051	0.062	0.060
		2	0.056	0.085	0.087	0.188
		3	0.052	0.068	0.080	0.127
		4	0.053	0.074	0.093	0.214
		5	0.055	0.095	0.105	0.762
0.1	0.01	1	0.009	0.010	0.018	0.019
		2	0.015	0.048	0.041	0.148
		3	0.011	0.027	0.030	0.075
		4	0.014	0.036	0.040	0.164
		5	0.015	0.058	0.048	0.704
0.1	0.05	1	0.021	0.020	0.030	0.030
		2	0.025	0.056	0.052	0.160
		3	0.022	0.036	0.043	0.087
		4	0.023	0.042	0.053	0.172
		5	0.024	0.061	0.061	0.630
0.1	0.1	1	0.051	0.050	0.061	0.060
		2	0.056	0.084	0.088	0.188
		3	0.052	0.067	0.081	0.127
		4	0.054	0.073	0.094	0.218
		5	0.054	0.090	0.104	0.712

LIST OF FIGURES

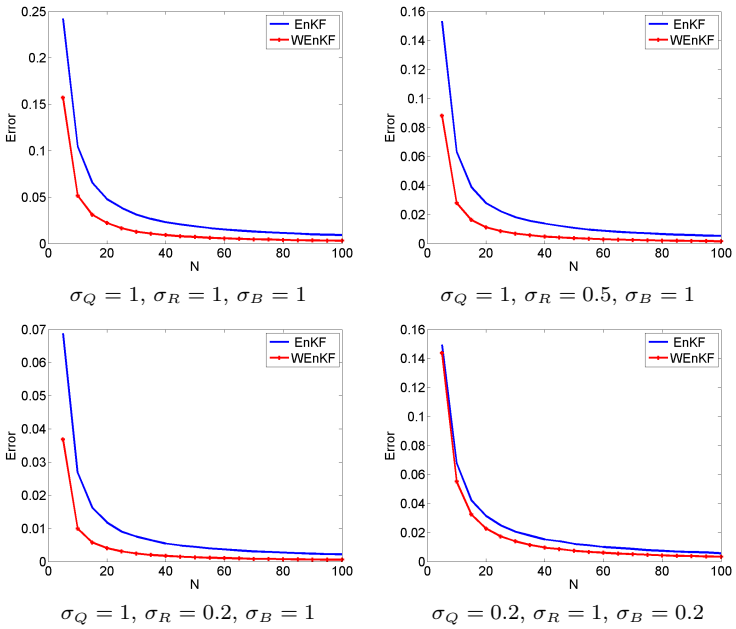


Fig. 1. Experiments on 1D linear problems. Results obtained from WEnKF (line with crosses) and EnKF (continuous line) are compared for four different sets of parameters (σ_Q,σ_R,σ_B) and increasing numbers N of particles. The mean squared estimation error with respect to the mean of the optimal Kalman filter distribution is plotted versus the number of particles. The mean squared errors have been computed on 5000 independent trials.

.

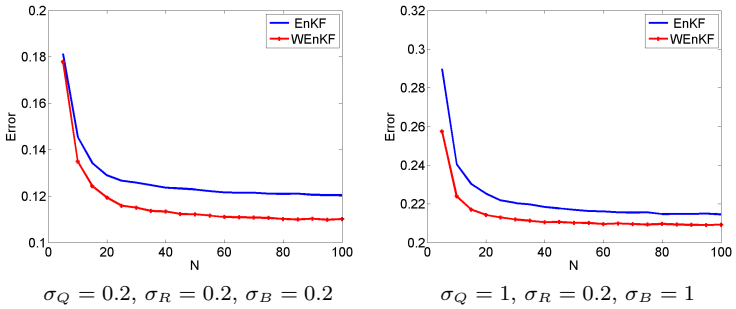
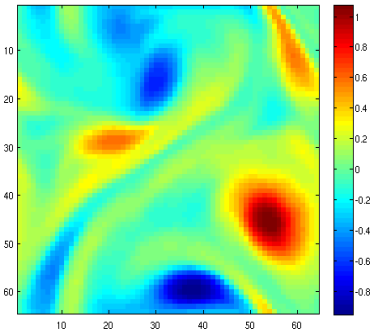
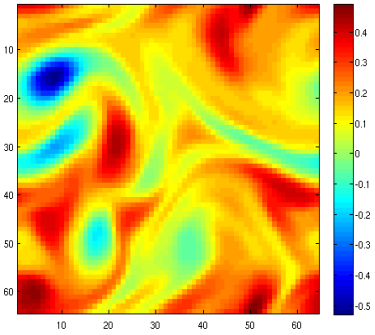


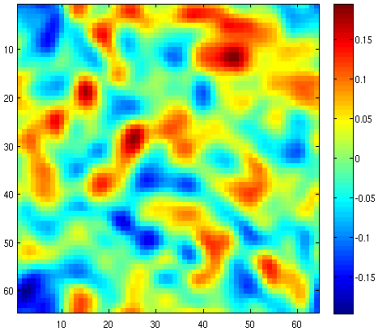
Fig. 2. Experiments on 1D non-linear problems. Results obtained from WEnKF (line with crosses) and EnKF (continuous line) are compared for two different sets of parameters (σ_Q,σ_R,σ_B) and increasing numbers N of particles. The mean squared estimation error with respect to the ground truth is plotted versus the number of particles. The mean squared errors have been computed on 5000 independent trials.



$\sigma_B = 0.1$, $\sigma_Q = 0.01$

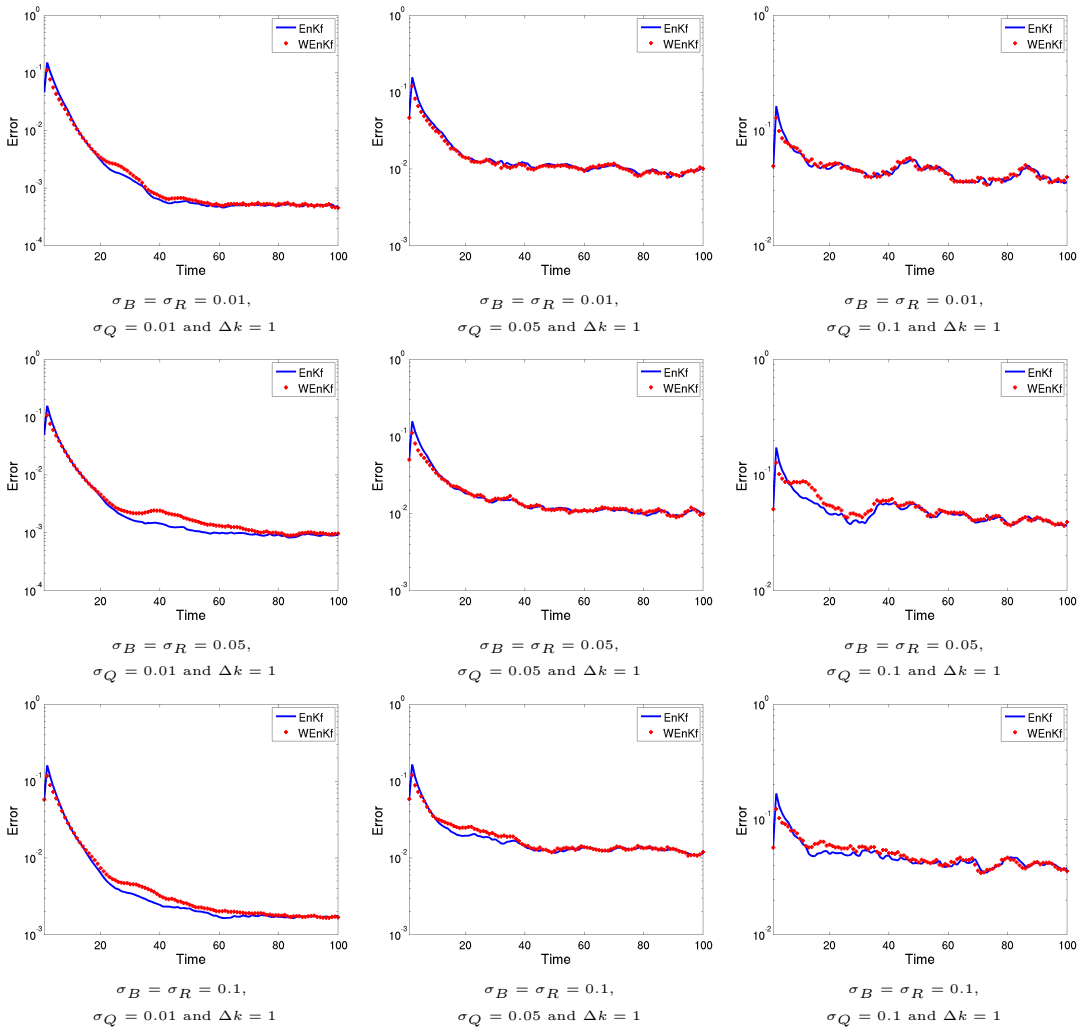


$\sigma_B = 0.05$, $\sigma_Q = 0.01$



$\sigma_B = 0.01$, $\sigma_Q = 0.01$

Fig. 3. Experiments on a high-dimensional non-linear problem. Examples of ground truth vorticity maps are presented for three different sets of parameters (σ_B,σ_Q). It can be noted that the amplitude and the size of vorticity structures vary a lot with respect to the scenarios.



$\sigma_B = \sigma_R = 0.01$, $\sigma_Q = 0.01$ and $\Delta k = 1$ | $\sigma_B = \sigma_R = 0.01$, $\sigma_Q = 0.05$ and $\Delta k = 1$ | $\sigma_B = \sigma_R = 0.01$, $\sigma_Q = 0.1$ and $\Delta k = 1$

$\sigma_B = \sigma_R = 0.05$, $\sigma_Q = 0.01$ and $\Delta k = 1$ | $\sigma_B = \sigma_R = 0.05$, $\sigma_Q = 0.05$ and $\Delta k = 1$ | $\sigma_B = \sigma_R = 0.05$, $\sigma_Q = 0.1$ and $\Delta k = 1$

$\sigma_B = \sigma_R = 0.1$, $\sigma_Q = 0.01$ and $\Delta k = 1$ | $\sigma_B = \sigma_R = 0.1$, $\sigma_Q = 0.05$ and $\Delta k = 1$ | $\sigma_B = \sigma_R = 0.1$, $\sigma_Q = 0.1$ and $\Delta k = 1$

Fig. 4. Experiments on a high-dimensional non-linear problem, with a measurement latency $\Delta k = 1$. The time evolution of the mean squared estimation error with respect to the ground truth is plotted for the WEnKF (dotted line) and the EnKF (continuous line), for different sets of parameters (σ_B,σ_R,σ_Q). The mean squared errors have been computed on 100 independent trials.

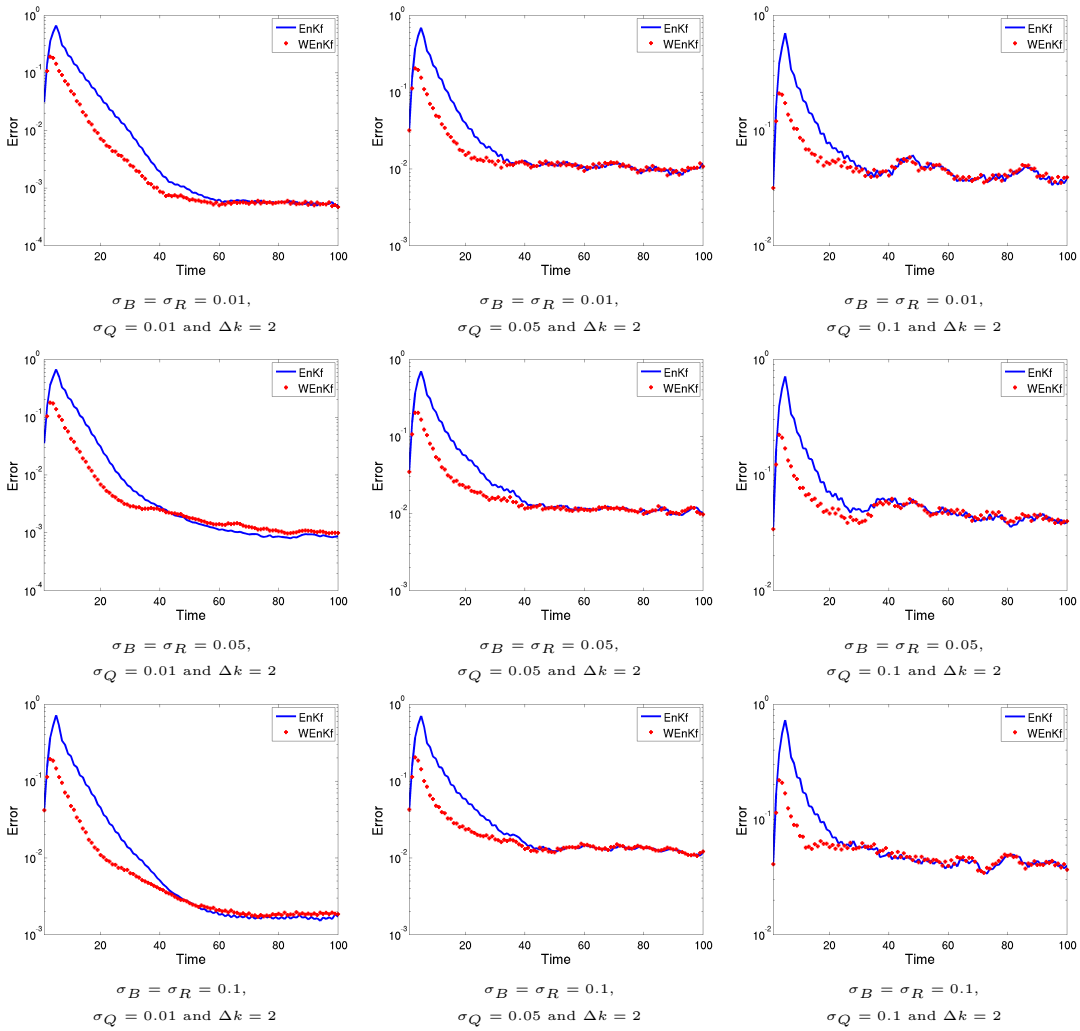
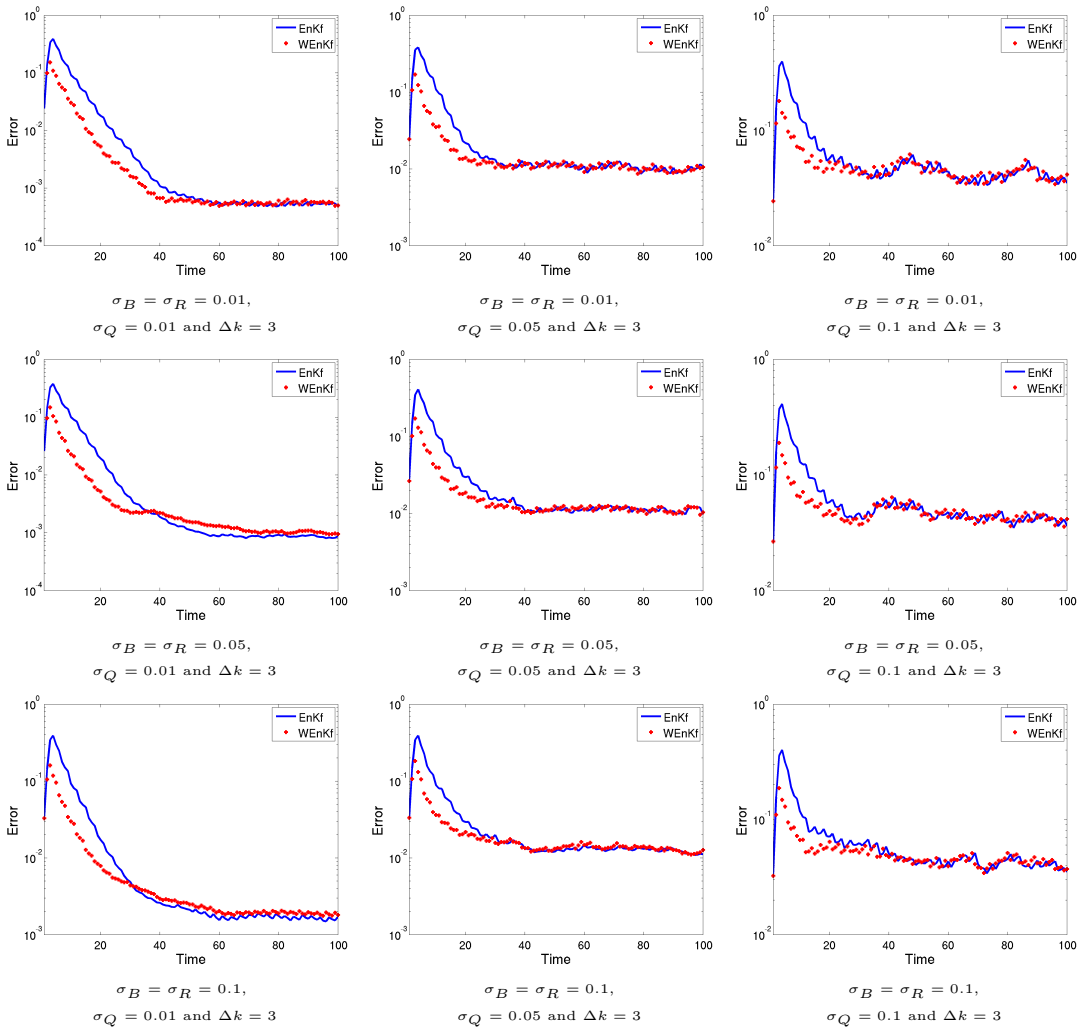


Fig. 5. Experiments on a high-dimensional non-linear problem, with a measurement latency $\Delta k = 2$. The time evolution of the mean squared estimation error with respect to the ground truth is plotted for the WEnKF (dotted line) and the EnKF (continuous line), for different sets of parameters (σ_B,σ_R,σ_Q). The mean squared errors have been computed on 100 independent trials.



$\sigma_B = \sigma_R = 0.01$, $\sigma_Q = 0.01$ and $\Delta k = 3$ | $\sigma_B = \sigma_R = 0.01$, $\sigma_Q = 0.05$ and $\Delta k = 3$ | $\sigma_B = \sigma_R = 0.01$, $\sigma_Q = 0.1$ and $\Delta k = 3$

$\sigma_B = \sigma_R = 0.05$, $\sigma_Q = 0.01$ and $\Delta k = 3$ | $\sigma_B = \sigma_R = 0.05$, $\sigma_Q = 0.05$ and $\Delta k = 3$ | $\sigma_B = \sigma_R = 0.05$, $\sigma_Q = 0.1$ and $\Delta k = 3$

$\sigma_B = \sigma_R = 0.1$, $\sigma_Q = 0.01$ and $\Delta k = 3$ | $\sigma_B = \sigma_R = 0.1$, $\sigma_Q = 0.05$ and $\Delta k = 3$ | $\sigma_B = \sigma_R = 0.1$, $\sigma_Q = 0.1$ and $\Delta k = 3$

Fig. 6. Experiments on a high-dimensional non-linear problem, with a measurement latency $\Delta k = 3$. The time evolution of the mean squared estimation error with respect to the ground truth is plotted for the WEnKF (dotted line) and the EnKF (continuous line), for different sets of parameters (σ_B,σ_R,σ_Q). The mean squared errors have been computed on 100 independent trials.

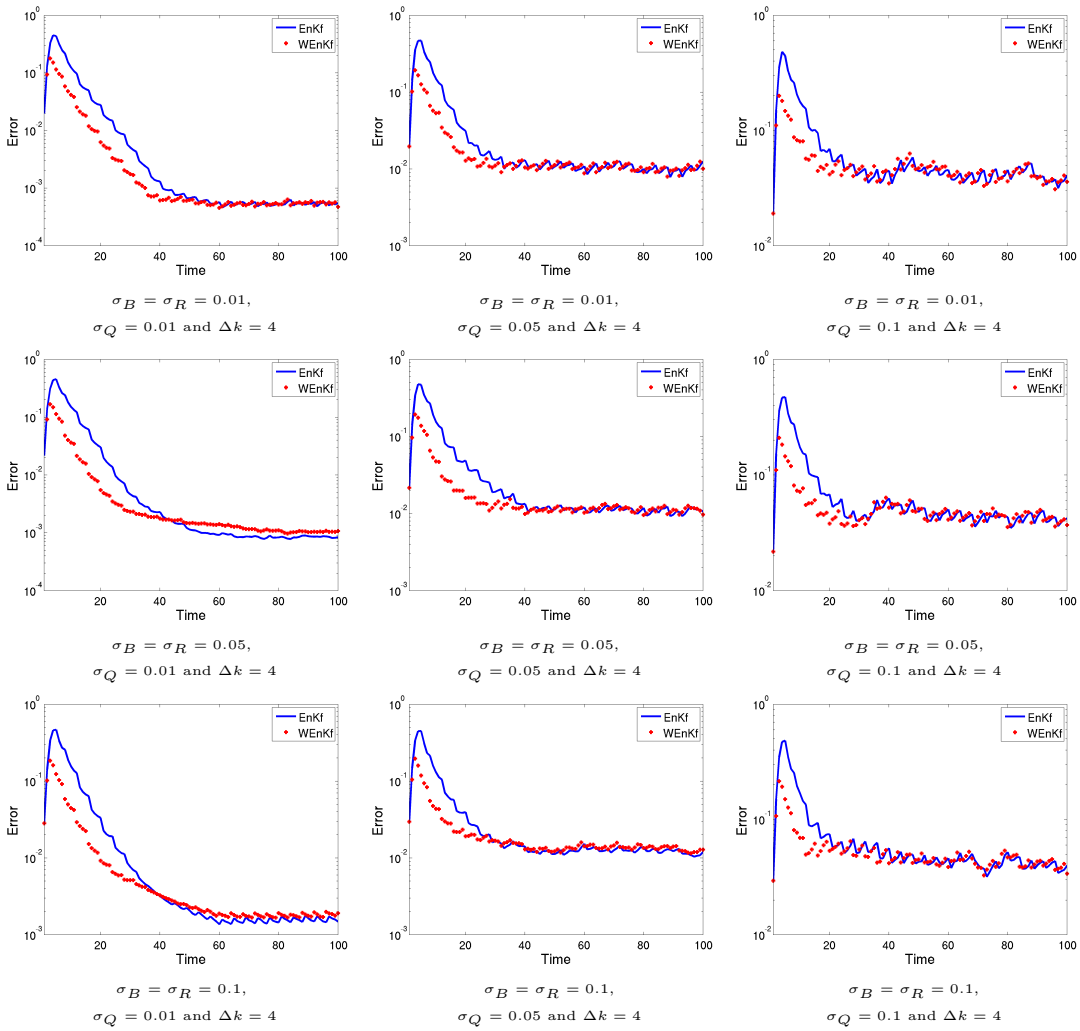


Fig. 7. Experiments on a high-dimensional non-linear problem, with a measurement latency $\Delta k = 4$. The time evolution of the mean squared estimation error with respect to the ground truth is plotted for the WEnKF (dotted line) and the EnKF (continuous line), for different sets of parameters (σ_B,σ_R,σ_Q). The mean squared errors have been computed on 100 independent trials.

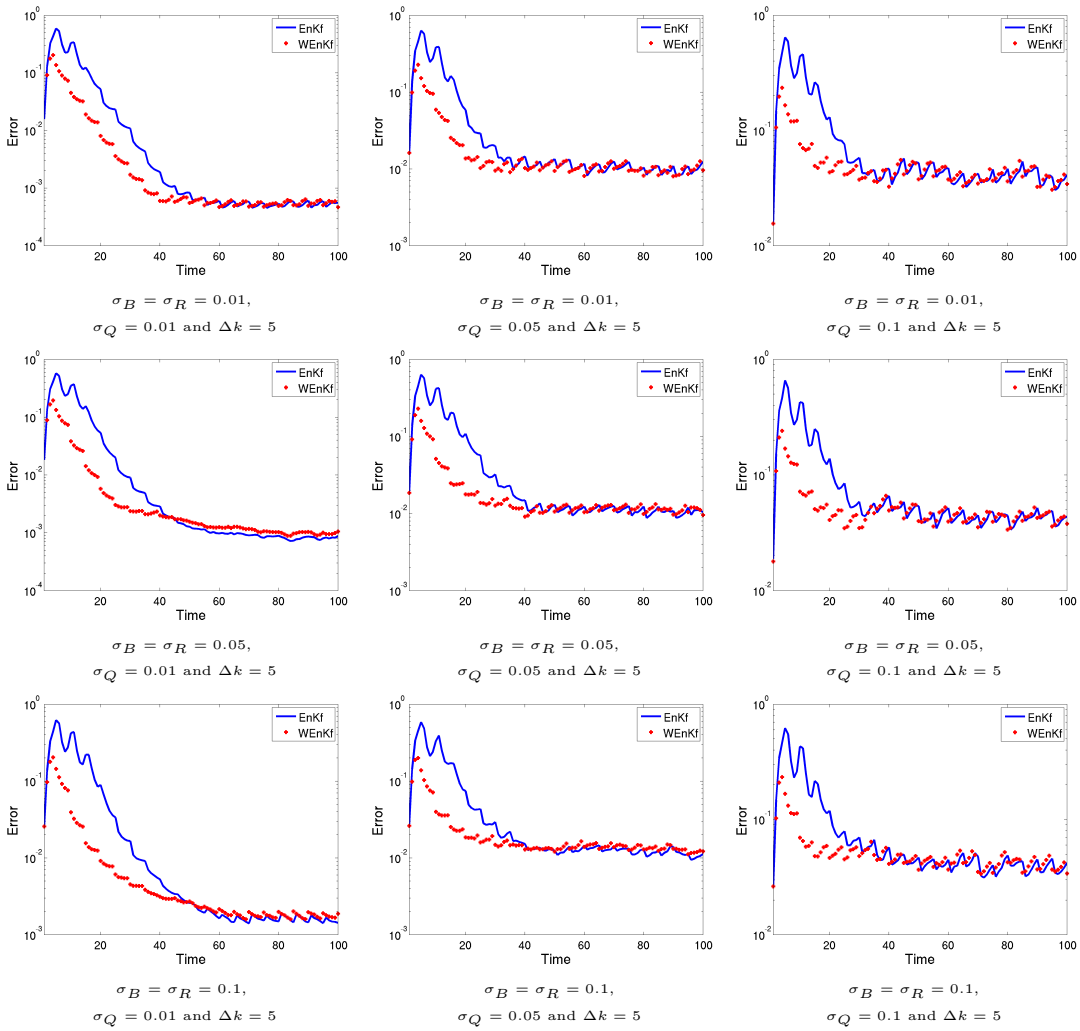


Fig. 8. Experiments on a high-dimensional non-linear problem, with a measurement latency $\Delta k = 5$. The time evolution of the mean squared estimation error with respect to the ground truth is plotted for the WEnKF (dotted line) and the EnKF (continuous line), for different sets of parameters (σ_B,σ_R,σ_Q). The mean squared errors have been computed on 100 independent trials.