Implicit equal-weights particle filter

Implicit Equal-Weights Particle Filter[†]

Mengbin Zhu[a,b*], Peter Jan van Leeuwen[b,c], Javier Amezcua[b,c]

[a]*Academy of Ocean Science and Engineering, NUDT, Changsha 410073, China*

[b]*Department of Meteorology, University of Reading, Reading RG6 6BB, UK*

[c]*National Centre for Earth Observation at the University of Reading, Reading RG6 6BB, UK*

[*]Correspondence to: Mengbin Zhu, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading, RG6 6BB, UK. E-mail: Mengbin.Zhu@pgr.reading.ac.uk

Filter degeneracy is the main obstacle for the implementation of particle filter in non-linear high-dimensional models. A new scheme, the implicit equal-weights particle filter (IEWPF), is introduced. In this scheme samples are drawn implicitly from proposal densities with a different covariance for each particle, such that all particle weights are equal by construction.

We test and explore the properties of the new scheme using a 1,000-dimensional simple linear model, and the 1,000-dimensional non-linear Lorenz96 model, and compare the performance of the scheme to a Local Ensemble Kalman Filter. The experiments show that the new scheme can easily be implemented in high-dimensional systems and is never degenerate, with good convergence properties in both systems.

Key Words: particle filter; non-degeneracy; implicit sampling; targeted weights

Received . . .

1. Introduction

Geophysical systems such as atmosphere or ocean systems are inherently nonlinear in nature. Numerical models which are used to simulate the true geophysical systems often have a state space of over one million variables, which results from discretising physical variables in a 3-D spatial grid. The dimension of state space keeps on growing due to the sustainable increase in model resolution and the computation capacity of the super computers. Numerical models describing atmosphere or ocean processes are discretisations of partial differential equations which need accurate initial and boundary conditions. The uncertainty in model equations and in initial and boundary conditions can be reduced by bringing in information from observations.

The prior knowledge of the state variables is described by the prior probability density function (pdf). The probability of each model state can be updated using Bayes' theorem by multiplying it with the probability of observations given that specific model state, the likelihood, resulting in the posterior probability density

[†]Please ensure that you use the most up to date class file, available from the QJRMS Home Page at
http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1477-870X

function, or posterior pdf of that model state. This update process is called data assimilation (DA). Present-day DA methods are tailored to specific statistics of the posterior pdf, e.g. mean, covariance, modes, etc. Variational methods such as 3DVar and 4DVar (**?**) search for the mode of the posterior pdf through the minimisation of a cost function. It cannot be guaranteed that the mode variational methods find is the global mode of the posterior pdf, which means that the search may stop at a local mode. Furthermore, it is hard to generate an uncertainty estimate of a variational solution due to its implicit use of the covariances involved. The Ensemble Kalman Filter (EnKF) (**??**) estimates the mean and covariance of the posterior pdf under the implicit assumptions of linearity and Gaussianity. Neither of these two popular methods can describe non-Gaussian posterior pdfs in an accurate manner, and it is still unclear what they estimate in a multimodal posterior pdf. Hybrids between the two methods like Ens4DVar, in which the ensemble from an EnKF is used to inform the background covariance of 3- or 4DVar about previous observations and flow structures, and 4DEnsVar, in which the space-time covariances in the 4DVar are explicitly generated from a forecast ensemble, like in an Ensemble Kalman Smoother, do not solve this issue.

The particle filter (PF) is a sequential Monte-Carlo method, which uses an ensemble of particles to represent the posterior pdf directly without linear or Gaussian assumptions, see e.g. (**?**). It has been successfully applied in systems with low dimensions, e.g. (**??**). But for geophysical systems with high dimensions, limited to the computation resources of modern super-computers, we cannot run enough model simulations to simulate the posterior pdf while avoiding the so-called "curse of dimensionality".

Different flavours of PFs exist, but all of them share two steps: forecast (also known as mutation) and weighting. When the numerical model equations contain errors (as they always do, of course, but these are often ignored), it is advantageous to slightly change the stochastic forecast model to stir the model closer to future observations. This is allowed in particle filtering as long as the weight of that model run is lowered accordingly in a well-specified way. Statistically this is known as drawing from a proposal density. When the model reaches the observation time these weights are multiplied by the likelihood of these observations assuming that they have been generated from that model state. The closer the model run to the original model, and the closer the model is to the observations the higher its weight will be. Most of the particle weights degenerate to a very small value as time evolves simply because it is hard to stay close to all observations. In high-dimensional situations with a large number of independent observations one particle obtains a weight close to one, while the others have weights very close to zero. The degeneracy of the particle weights leads to a loss of statistical information since the effective ensemble size reduces to 1. This is the main obstacle for PF to be applied operationally as an alternative in DA (**?**). **???** argue that the ensemble size must scale exponentially with respect to the "effective size" of the problem (proportional to the number of independent observations) for a particle filter to avoid degeneracy. They show that this is even the case for the proposal density that has the lowest variance in the weights, which they showed to be equal to the so-called Optimal Proposal Density, which is known to be optimal in a slightly different way. The analytical calculations were backed up by convincing experiments using a simple linear test case.

? and **?** introduce an implicit proposal density method that choose a map from the implicit sampling space to the original state space. Examples on 100-500 dimensional spaces show that the method is more robust than the original particle filter with resampling, but it is easy to show that the method reduces to the optimal proposal density when observations are present at every time step and when the model noise is state independent, having the same degeneracy issues.

The equivalent-weights particle filter (EWPF) of **??** and **??** explore a particle filter that uses a proposal density of a different class than studied before. It allows for a proposal density for each particle that depends not only on the position of the particle at previous time, but on all particles at previous time. Since all particles are involved it is straightforward to ensure that the final weights of part, or al, of the particles are equal. It has been shown to be non-degenerate in even high-dimensional spaces, e.g. the 65,000 dimensional barotropic vorticity model, and recently the over 2 million dimensional climate model HadCM3 (**?**). The scheme has a few tuning parameters that can be adjusted for optimal performance, measured by e.g. rank histograms. It can

be shown, however, that such a scheme is biased. It is well known that particle filters that explore resampling are always biased, but the bias is of a stronger nature in this filter. A bias is not an issue in itself as long as the bias is smaller than the statistical noise in the method. In the IWPF, in order to enforce weights that are close together, the particles are forced to be positioned close to a hyper-ellipsoidal shell, one for each particle. This means that the proposal density of all particles together does not explore the full state space. The equivalent-weights PF works extremely well for small ensemble sizes of order 10-100 for high-dimensional (order 1,000 or much more) systems, when the statistical noise is relatively large, but this scheme does perform less favourably when large ensemble sizes are used and the bias becomes apparent. Although we typically cannot afford more that 10-100 particles in geophysical systems this limits the usefulness of this scheme.

In this article, a new PF is proposed, which we label the implicit equal-weights particle filter (IEWPF). This scheme uses a proposal transition density in which each particle is drawn implicitly from a slightly different proposal density, the difference being a factor in front of the covariance of the proposal. . This factor depends on all other particles such that the equal-weight property is fulfilled. This scheme is applied during the last transition step before the observations. In between observation times a simple relaxation scheme is used, as in the equivalent-weights particle filter. One strong advantage of this new scheme is that the number of tuning parameters has been reduced drastically, and we will show that the bias is much smaller than for the EWPF.

This article is organised as follows. Section 2 describes the implicit equal-weights particle filter in detail, and its performance on the linear model used by **?** and 1,000 dimension Lorenz96 model are discussed in section 3, together with a comparison with the LETKF. A summary and conclusions are provided in section 4.

2. Implicit Equal-Weights Particle Filter

2.1. *The basic idea*

Bayes' theorem shows how the prior density $p(x)$ is changed when multiplying it with the density of observations y given a specific model state x, the likelihood. The posterior pdf of the model state given observations $p(x|y)$ is thus given by:

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)} \tag{1}$$

The posterior pdf of a filter is the probability of the state variable x^n at time-step n given the observations $y^{1:n}$ at time $1, \cdots, n$. For a Markovian system with observational errors that are independent from one time to another, the posterior pdf can be written as

$$p(x^n|y^{1:n}) = \frac{p(y^n|x^n)}{p(y^n)} \int p(x^n|x^{n-1})p(x^{n-1}|y^{1:n-1})dx^{n-1} \tag{2}$$

The transition density $p(x^n|x^{n-1})$ is related to the model equation via

$$x^n = \mathcal{M}(x^{n-1}) + \beta^n \tag{3}$$

in which $\mathcal{M}(.)$ is the nonlinear deterministic model equation, and β^n is a stochastic perturbation with mean zero that can, in principle, depend on x^{n-1}.

Let us assume for the moment that we run a particle filter and that the particle weights in the ensemble at previous time-step $n-1$ are equal:

$$p(x^{n-1}|y^{1:n-1}) = \frac{1}{N}\sum_{i=1}^{N} \delta(x^{n-1} - x_i^{n-1}) \tag{4}$$

When plugging equation (4) into equation (2), we find that:

$$p(x^n|y^{1:n}) = \frac{1}{N}\sum_{i=1}^{N} \frac{p(y^n|x^n)p(x^n|x_i^{n-1})}{p(y^n)} \tag{5}$$

One can now multiply the numerator and denominator of equation (5) by the same factor $q(x^n|\mathbf{x^{n-1}}, y^n)$, in which $\mathbf{x^{n-1}}$ is defined as the collection of all particles at time $n-1$.

$$p(x^n|y^{1:n}) = \frac{1}{N}\sum_{i=1}^{N} \frac{p(y^n|x^n)}{p(y^n)} \frac{p(x^n|x_i^{n-1})}{q(x^n|\mathbf{x^{n-1}}, y^n)} q(x^n|\mathbf{x^{n-1}}, y^n) \tag{6}$$

where the support of $q(x^n|\mathbf{x^{n-1}}, y^n)$ should be equal to or larger than that of $p(x^n|x_i^{n-1})$. $q(x^n|\mathbf{x^{n-1}}, y^n)$ is the so-called *proposal transition density*.

The assumption that observations appear at every time-step is made and we draw samples from the proposal transition density $q(x^n|\mathbf{x^{n-1}}, y^n)$, instead of the original transition density $p(x^n|x_i^{n-1})$. This leads the posterior pdf to be expressed as:

$$p(x^n|y^{1:n}) = \frac{1}{N}\sum_{i=1}^{N}\frac{p(y^n|x_i^n)}{p(y^n)}\frac{p(x_i^n|x_i^{n-1})}{q(x_i^n|\mathbf{x^{n-1}}, y^n)}\delta(x^n - x_i^n) \quad (7)$$

Consequently, the posterior pdf of model state at time-step n can be written as

$$p(x^n|y^{1:n}) = \frac{1}{N}\sum_{i=1}^{N} w_i\delta(x^n - x_i^n) \quad (8)$$

where w_i is the particle weights given by

$$w_i = \frac{p(y^n|x_i^n)}{p(y^n)}\frac{p(x_i^n|x_i^{n-1})}{q(x_i^n|\mathbf{x^{n-1}}, y^n)} \quad (9)$$

Now assuming that the model system is Markovian and using Bayes' theorem, the numerator in the expression for the weights can be expressed as

$$p(y^n|x^n)p(x^n|x_i^{n-1}) = p(x^n|x_i^{n-1}, y^n)p(y^n|x_i^{n-1}) \quad (10)$$

Therefore the particle weight of ensemble member with index i at observed time-step becomes

$$w_i = \frac{p(x_i^n|x_i^{n-1}, y^n)p(y^n|x_i^{n-1})}{p(y^n)q(x_i^n|\mathbf{x^{n-1}}, y^n)} \quad (11)$$

In the so-called optimal proposal density (**?**) one chooses $q(x_i^n|\mathbf{x^{n-1}}, y^n) = p(x_i^n|x_i^{n-1}, y^n)$, leading to weights $w_i \propto p(y^n|x_i^{n-1})$. For systems with a large number of independent observations these weights are degenerate, see e.g. (**?**).

The implicit part of our scheme follows from drawing samples implicitly from a standard Gaussian distributed proposal density $q(\xi)$ instead of the original one $q(x^n|\mathbf{x^{n-1}}, y^n)$ (**?**). These two pdfs are related by:

$$q(x^n|\mathbf{x^{n-1}}, y^n) = \frac{q(\xi)}{||\frac{dx}{d\xi}||} \quad (12)$$

where $||\frac{dx}{d\xi}||$ denotes the absolute value of the determinant of the Jacobian matrix of the $\mathcal{R}^{N_x} \to \mathcal{R}^{N_x}$ transformation $x_i = g(\xi_i)$. In the Implicit Equal-Weights Particle Filter this function $g(.)$ is defined via

$$x_i^n = x_i^a + \alpha_i^{1/2}P^{1/2}\xi_i^n \quad (13)$$

with x_i^a the mode of $q(x_i^n|\mathbf{x^{n-1}}, y^n)$, P is a measure of the width of that pdf, and α_i is a scalar. In the implicit particle filter of **?** α_i is determined by choosing the proposal density as the optimal proposal density, so again $q(x_i^n|\mathbf{x^{n-1}}, y^n) = p(x_i^n|x_i^{n-1}, y^n)$, and using the expression for x_i^n directly in

$$p(x_i^n|x_i^{n-1}, y^n) = \frac{q(\xi)}{||\frac{dx}{d\xi}||} \quad (14)$$

leading to a nonlinear scalar equation for α_i.

Our scheme is different in that we choose the α_i such that all particles get the same weight w_{target}, so we determine the scalar α_i for each particle from:

$$w_i = \frac{p(x_i^n|x_i^{n-1}, y^n)p(y^n|x_i^{n-1})}{Np(y^n)q(x_i^n|\mathbf{x^{n-1}}, y^n)} = w_{target} \quad (15)$$

This equation is at the heart of the IEWPF, showing the equal-weights part of the scheme. It ensures that the filter is not degenerate in systems with arbitrary large dimensions and with an arbitrary large number of independent observations.

We can expand this as follows. Sampling implicitly from $q(\xi)$ instead of $q(x_i^n|\mathbf{x^{n-1}}, y^n)$, the particle weights are now given by

$$w_i = \frac{p(x_i^n|x_i^{n-1}, y^n)p(y^n|x_i^{n-1})}{q(\xi)}\left|\left|\frac{dx}{d\xi}\right|\right| \cdot w_i^{prev} \quad (16)$$

where $q(\xi)$ is the standard Gaussian distribution and w_i^{prev} introduces the weight from previous time-steps. This equation demonstrates the implicit part of the scheme.

The determinant of the Jacobian depends only on the transformation from ξ to x, and is independent of the pdfs of the these variables. Hence, we can simply obtain it from (13) and get

$$\left|\left|\frac{dx}{d\xi}\right|\right| = \left|\left|\alpha_i^{1/2}P^{1/2} + P^{1/2}\xi_i^n\frac{\partial\alpha_i^{1/2}}{\partial\xi_i^n}\right|\right| \quad (17)$$

Factorising $\alpha_i^{1/2}P^{1/2}$ out from the right hand side leads to:

$$\left|\left|\frac{dx}{d\xi}\right|\right| = \alpha_i^{N_x/2}\left|\left|P^{1/2}\right|\right|\left|\left|I + \frac{\xi_i^n}{\alpha_i^{1/2}}\frac{\partial\alpha_i^{1/2}}{\partial\xi_i^n}\right|\right| \tag{18}$$

The last factor in this equation can be simplified to a scalar by using Sylvester's determinant lemma. Hence, the equation for $||\frac{dx}{d\xi}||$ reduces to:

$$\left|\left|\frac{dx}{d\xi}\right|\right| = \alpha_i^{N_x/2}\left|\left|P^{1/2}\right|\right|\left|1 + \frac{\partial\alpha_i^{1/2}}{\partial\xi_i^n}\frac{\xi_i^n}{\alpha_i^{1/2}}\right| \tag{19}$$

2.2. *Gaussian observation and model errors, and linear observation operator*

In this section the new scheme is explored for the case when observation errors and model errors are assumed to be Gaussian, and the observation operator $H \in \mathcal{R}^{N_y \times N_x}$ is assumed to be linear. With these assumptions we can write:

$$\begin{aligned}
&p(y^n|x^n)p(x^n|x_i^{n-1}) \\
&= \frac{1}{A}\exp\left[-\frac{1}{2}(y^n - Hx^n)^T R^{-1}(y^n - Hx^n)\right. \\
&\left. - \frac{1}{2}(x^n - f(x_i^{n-1}))^T Q^{-1}(x^n - f(x_i^{n-1}))\right] \\
&= \frac{1}{A}\exp\left(-\frac{1}{2}(x^n - \hat{x_i^n})^T P^{-1}(x^n - \hat{x_i^n})\right)\exp(-\frac{1}{2}\phi_i) \\
&= p(x^n|x_i^{n-1}, y^n)p(y^n|x_i^{n-1})
\end{aligned} \tag{20}$$

where

$$P = (Q^{-1} + H^T R^{-1} H)^{-1} \tag{21}$$

$$\hat{x_i^n} = f(x_i^{n-1}) + (Q^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1}(y^n - Hf(x_i^{n-1})) \tag{22}$$

$$\phi_i = (y^n - Hf(x_i^{n-1}))^T (HQH^T + R)^{-1}(y^n - Hf(x_i^{n-1})) \tag{23}$$

x_i^a in equation (13) is the mode of $p(x^n|x_i^{n-1}, y^n)$, given by

$$x_i^a = \hat{x_i^n} = f(x_i^{n-1}) + QH^T(HQH^T + R)^{-1}(y^n - Hf(x_i^{n-1})) \tag{24}$$

For ease of presentation we introduce:

$$\alpha_i = 1 + \varepsilon_i \tag{25}$$

Taking minus the logarithm of the expression for the weights derived in the previous section leads to:

$$\begin{aligned}
-2\log w_i = &-2\log w_i^{prev} \\
&+ \left\{-2\log\left(\frac{p(x_i^n|x_i^{n-1}, y^n)p(y^n|x_i^{n-1})}{q(\xi)}\left|\left|\frac{dx}{d\xi}\right|\right|\right)\right\}
\end{aligned} \tag{26}$$

Let J_i and J_i^{prev} stand for 2 times the logarithmic particle weights of analysis time and previous time-steps respectively, then the last equation can be rewritten as:

$$J_i = J_i^{prev} - 2\log\left(\frac{p(x_i^n|x_i^{n-1}, y^n)p(y^n|x_i^{n-1})}{q(\xi)}\left|\left|\frac{dx}{d\xi}\right|\right|\right) \tag{27}$$

Substituting the Jacobian factor obtained in equation (19) we find:

$$\begin{aligned}
J_i = &J_i^{prev} + (x_i^n - \hat{x_i^n})^T P^{-1}(x_i^n - \hat{x_i^n}) + \phi_i \\
&- \xi_i^{nT}\xi_i^n - 2\log\left(\alpha_i^{N_x/2}\left|\left|P^{1/2}\right|\right|\left|1 + \frac{\partial\alpha_i^{1/2}}{\partial\xi_i^n}\frac{\xi_i^n}{\alpha_i^{1/2}}\right|\right)
\end{aligned} \tag{28}$$

in which the constant term common to all particles is ignored as it plays no role in the following.

Since $x_i^n = \hat{x_i^n} + \alpha_i^{1/2}P^{1/2}\xi_i^n$ and $\alpha_i = 1 + \varepsilon_i$, the equation of J_i can be simplified as

$$\begin{aligned}
J_i &= J_i^{prev} + \alpha_i\xi_i^{nT}P^{1/2}P^{-1}P^{1/2}\xi_i^n + \phi_i \\
&\quad - \xi_i^{nT}\xi_i^n - 2\log\left(\alpha_i^{N_x/2}\left|\left|P^{1/2}\right|\right|\left|1 + \frac{\partial\alpha_i^{1/2}}{\partial\xi_i^n}\frac{\xi_i^n}{\alpha_i^{1/2}}\right|\right) \\
&= J_i^{prev} + \varepsilon_i\xi_i^{nT}\xi_i^n + \phi_i \\
&\quad - 2\log\left(\alpha_i^{N_x/2}\left|\left|P^{1/2}\right|\right|\left|1 + \frac{\partial\alpha_i^{1/2}}{\partial\xi_i^n}\frac{\xi_i^n}{\alpha_i^{1/2}}\right|\right)
\end{aligned} \tag{29}$$

For ease of presentation we also introduce $\xi_i^{nT}\xi_i^n = \gamma_i$, such that J_i is given by

$$\begin{aligned}
J_i = &J_i^{prev} + \varepsilon_i\gamma_i + \phi_i \\
&- 2\log\left(\alpha_i^{N_x/2}\left|\left|P^{1/2}\right|\right|\left|1 + \frac{\partial\alpha_i^{1/2}}{\partial\xi_i^n}\frac{\xi_i^n}{\alpha_i^{1/2}}\right|\right)
\end{aligned} \tag{30}$$

Setting the weights of all particles equal to the target weight w_{target} is equal to putting all J_i equal to a constant number C,

leading to the following equation for ε_i:

$$\varepsilon_i\gamma_i - 2\log(\alpha_i^{N_x/2}) - 2\log(||P^{1/2}||) - 2\log\left(\left|1+\frac{\partial\alpha_i^{1/2}}{\partial\xi_i^n}\frac{\xi_i^n}{\alpha_i^{1/2}}\right|\right) + \phi_i + J_i^{prev} - C = 0 \quad (31)$$

in which $-2\log(||P^{1/2}||)$ is absorbed in C as it is also a constant.

Although this is a scalar equation, the derivative makes this implicit equation hard if not impossible to solve in general. Since we are interested in high-dimensional problems we consider this equation in the limit of large state dimension N_x. As detailed in Appendix A we can integrate this equation in this limit, leading to the much simpler equation:

$$\varepsilon_i\gamma_i - N_x\log(1+\varepsilon_i) + \phi_i + J_i^{prev} - C = 0 \quad (32)$$

If this equation could be solved and the real solutions of ε_i could be obtained, an absolute equal weights particle filter method that avoids filter degeneracy is discovered. The equation can be solved by iterative methods, such as Newton method, etc., but interestingly analytical solutions exist. The analytical solutions are based on the so-called Lambert W function, as detailed in the next section.

2.3. *Analytical Solutions*

2.3.1. *Lambert W Function*

The Lambert W function (**??**), also called the omega function or the product logarithm function, is the inverse function of

$$z = f(W) = W(z)e^{W(z)} \quad (33)$$

where $e^{W(z)}$ is the exponential function and z is any complex number.

The W function is multivalued (except at zero point) because $f(\cdot)$ is not injective. In this new scheme, our attention is restricted to real-valued W, the complex variable notation z is replaced by the real variable notation x, and $W(x)$ exists when $x \geq -1/e$, and is double-valued on $(-1/e, 0)$, see Figure 1. The branch satisfying $W(x) \geq -1$ is denoted by $W_0(x)$ and the branch satisfying $W(x) \leq -1$ is denoted by $W_{-1}(x)$ as indicated in Figure 1.

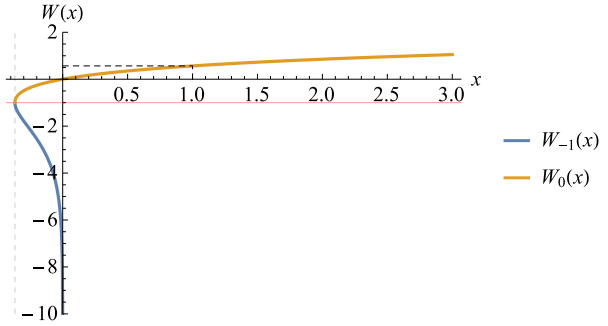


Figure 1. Lambert W function in real-valued $W(x)$

As can be seen in the figure, $W_0(0) = 0$ and $W_0(-1/e) = -1$. The Lambert W function decreases from $W_{-1}(-1/e) = -1$ to $W_{-1}(0^-) = -\infty$ in the branch $W_{-1}(x)$. Crucial identities of Lambert W function are its derivative

$$\frac{dW}{dz} = \frac{W(z)}{z(1+W(z))} \quad (34)$$

where $z \notin \{0, -1/e\}$ and the equation

$$W(x \cdot e^x) = x \quad (35)$$

which follows directly from its definition.

Its interest for our problem is that the Lambert W function gives the solution of the generalized problem:

$$\log(A + Bx) + C_w x = \log D \quad (36)$$

as

$$x = \frac{1}{C_w}W\left[\frac{C_w D}{B}\exp\left(\frac{AC_w}{B}\right)\right] - \frac{A}{B} \quad (37)$$

This allows us to solve equation (32) to obtain an analytical solution for ε_i.

2.3.2. *Solutions for α_i*

Equation (32) could be generalized as

$$ax - b\log(1+x) - c = 0 \quad (38)$$

in which $a = \gamma_i$, $b = N_x$, $c = C - \phi_i - J_i^{prev}$ and $x = \varepsilon_i$.

Following equation (36), the analytical solution of x is found as

$$x = -\frac{b}{a}W\left[-\frac{a}{b}\cdot e^{-\frac{a}{b}}\cdot e^{-\frac{c}{b}}\right] - 1 \quad (39)$$

so that

$$\varepsilon_i = -\frac{N_x}{\gamma_i} W\left[-\frac{\gamma_i}{N_x}\cdot e^{-\frac{\gamma_i}{N_x}}\cdot e^{-\frac{c}{N_x}}\right] - 1 \tag{40}$$

and

$$\alpha_i = 1+\varepsilon_i = -\frac{N_x}{\gamma_i} W\left[-\frac{\gamma_i}{N_x}\cdot e^{-\frac{\gamma_i}{N_x}}\cdot e^{-\frac{c}{N_x}}\right] \tag{41}$$

To ensure real-valued solutions c must satisfy

$$-\frac{\gamma_i}{N_x}\cdot e^{-\frac{\gamma_i}{N_x}}\cdot e^{-\frac{c}{N_x}} > -e^{-1} \tag{42}$$

so

$$c > N_x \log(\frac{\gamma_i}{N_x}) - \gamma_i + N_x \tag{43}$$

In accordance with the characersitics of Lambert W function, we find the following characteristics for α_i. First, there are two real solutions for $W(\cdot)$, and for ε_i and thus α_i. ε_i has a positive real solution give by W_{-1} branch and a negative real solution given by the negative x part of W_0 branch which is always larger than -1. Second, if the value of c is zero, the value of α_i becomes a single constant solution 1 because of identity equation (35):

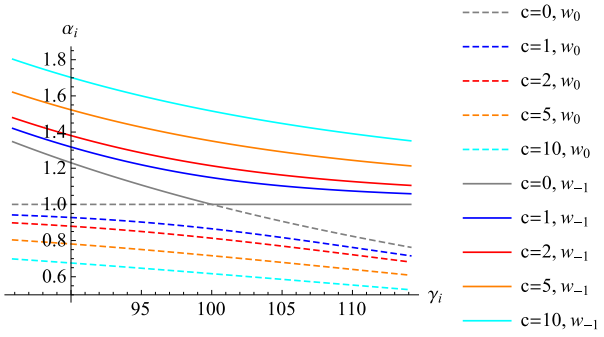
$$\begin{aligned}\alpha_i &= -\frac{N_x}{\gamma_i} W\left[-\frac{\gamma_i}{N_x}\cdot e^{-\frac{\gamma_i}{N_x}}\right] \\ &= -\frac{N_x}{\gamma_i}\cdot\left\{-\frac{\gamma_i}{N_x}\right\} = 1\end{aligned} \tag{44}$$

Practically, the solutions can be derived numerically via the Lambert W function, or by using numerical approximation methods for the original equation (32). The Lambert W function could be approximated using Newton's method or Halley's method (**?**).

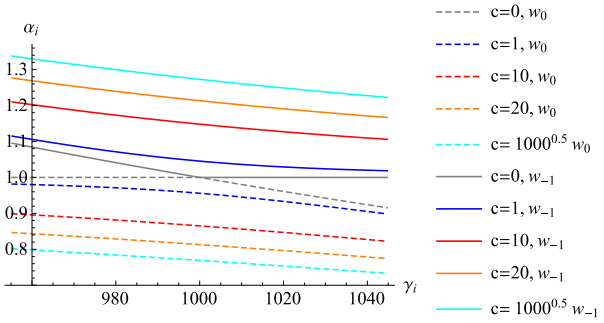
2.3.3. *Structure of solutions*

The analytical solution for α_i is a complicated form of Lambert W function of γ_i. Since γ_i is a χ^2 variable with N_x degrees of freedom a typical range for γ_i is $[N_x - \sqrt{2N_x}, N_x + \sqrt{2N_x}]$. The order of magnitude of c is $O(\sqrt{N_x})$. Figure 2 shows the plot of α_i as a function of γ_i with varying c values under different state space dimensions.

The solution for α_i has two branches related to Lambert W function. In Figure 2, the dashed line is the -1 branch of the α_i solutions and the full line is the 0 branch. Different line colours represent different c values.



(a) $\alpha_i(\gamma_i)$ with $N_x = 100$.



(b) $\alpha_i(\gamma_i)$ with $N_x = 1000$.

Figure 2. $\alpha_i(\gamma_i)$ with varying c under different N_x values.

The α_i values tend to be closer to 1 when N_x becomes larger, mainly because the fluctuation of $\frac{\gamma_i}{N_x}$ tends to be smaller when N_x increases.

When $c = 0$, the two branches of α_i meet in one point where $\gamma_i = N_x$. With increasing c, the gap between two branches becomes larger and larger and the values of α_i are further away from 1.

2.3.4. *Discussion*

Figure 2 illustrates how α_i performs with varying c. There is a gap when $c \neq 0$, restricting the state space that $\alpha_i^{1/2} P^{1/2} \xi_i^n$ explores. Since $P^{1/2}$ is a constant matrix, we ignore it in this section. We define function $f(\xi_i^n)$ as

$$f(\xi_i^n) = \alpha_i^{1/2}\xi_i^n \tag{45}$$

The full expression is given by

$$f(\xi_i^n) = \sqrt{-\frac{N_x}{\xi_i^{nT}\xi_i^n} W\left[-\frac{\xi_i^{nT}\xi_i^n}{N_x}\cdot e^{-\frac{\xi_i^{nT}\xi_i^n}{N_x}}\cdot e^{-\frac{c}{N_x}}\right]}\,\xi_i^n \tag{46}$$

We choose N_x to be 1 as the simplest case, and c/N_x has three values, 0, 1/2 and 1. Figure 3 shows the state space that $f(\xi_i^n)$ explores with varying c/N_x values when $N_x = 1$.

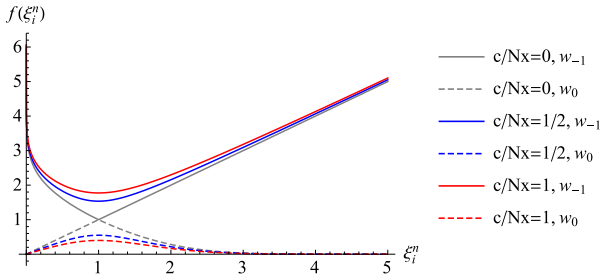


Figure 3. Function $f(\xi_i^n)$ with different c/N_x values

With varying c/N_x, the changing behaviour of the solution can be seen clearly in Figure 3. The gap exists for other particles except the one with $c/N_x = 0$ and it becomes larger when c/N_x is becoming larger in this case. The gap becomes a hyper-sphere for high dimensional systems.

The importance of the gap lies in the fact that the proposal density does not explore the full state space for those particles that have a gap, so all particles except one. This means that the new scheme will be biased, although it is unclear what form this bias takes. The gap position will be different for each particle, so the space missed out by several particles will be much smaller than the gap of an individual particle. And, because one particle has no gap, the ensemble as a whole will explore full state space.

The scheme is tailored to high-dimensional systems, so we studied the importance of the gap when N_x increases. For each particle there are two high probability hyper-spheres surrounding the gap, and we show in the appendix B that the ratio of the gap volume and the volumes of the two high probability hyper-spheres will become smaller when the state space dimension increases, suggestion that the bias decreases when N_x increases.

2.4. *Multi Time-steps Between Observations*

In typical geophysical systems several model time-steps exist between observations times. In principle one can extend the formulation above for a number of time steps, as e.g. the implicit particle filter does. In that case x_i^a becomes a model trajectory over time, and can be found as weak-constraint 4DVar solution with fixed initial condition. The random vector ξ_i^n will now extend over space and time, and so will P. This will again result in a highly nonlinear equation for α_i, which can be solved numerically.

In this paper we use a simpler approach and use the relaxation proposal density also explored in e.g. **?**. If it is assumed that the original model error is Gaussian with known covariance matrix Q, then the model transition density is expressed as

$$p(x^j|x_i^{j-1}) \sim N(\mathcal{M}(x_i^{j-1}), Q) \tag{47}$$

The relaxation proposal transition density of the time-steps before the last time step towards the observations is chosen as

$$q(x_i^j|x_i^{j-1}, y^n) \sim N(\mathcal{M}(x_i^{j-1}) + B(\tau)[y^n - h\mathcal{M}(x_i^{j-1})], \widehat{Q}) \tag{48}$$

In this equation $B(\tau)[y^n - h\mathcal{M}(x_i^{j-1})]$ is the relaxation term forcing the model state towards the observations at time-step n. $\widehat{Q}$ is the covariance of the random forcing in the modified model, which we choose equal to Q in our experiments. Sampling from the proposal transition density instead of the original model equation leads to:

$$\frac{p(x_i^j|x_i^{j-1})}{q(x_i^j|x_i^{j-1}, y^n)} \propto \exp\left[-\frac{1}{2}v^T Q^{-1} v + \frac{1}{2}(\widehat{d\beta_i^j})^T Q^{-1}(\widehat{d\beta_i^j})\right] \tag{49}$$

where we introduced the short-hand notation v for $B(\tau)[y^n - h\mathcal{M}(x_i^{j-1})] + \widehat{d\beta_i^j}$.

The relaxation strength $B(\tau)$ is given by

$$B(\tau) = b\tau Q H^T R^{-1} \tag{50}$$

where τ increases linearly from zero to one at the *previous* time-steps and b is a constant. $B(\tau)$ controls the strength of relaxation, but, via its dependence on Q, also spreads the information from the observed grid points to unobserved grid points.

This expression for p/q allows to generate w_i^{prev} simply by multiplying the particle weight by the p/q factor for each time-step, see **?** for more details.

3. Experiments

3.1. *Linear Model Experiments*

Although geophysical models tend to be high-dimensional non-linear systems, linear models are still a simple benchmark for testing new DA schemes. Furthermore, analytical solutions are usually available in these cases. Consider the model equation and

the observations (**?**):

$$x^n = x^{n-1} + \eta^{n-1} \tag{51}$$

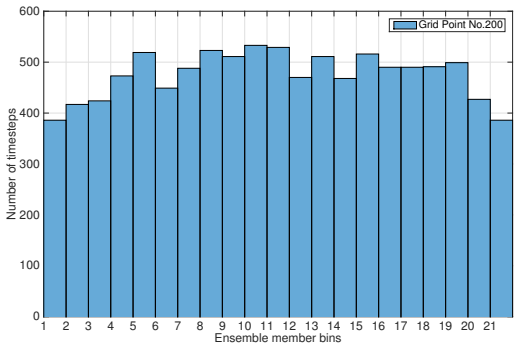
$$y^n = x^n_{truth} + \epsilon^n \tag{52}$$

where x^n is the state variable at time-step n and y^n is the observation vector at time-step n. Random model perturbations η are drawn from the model error pdf $N(0, Q)$, observation errors ϵ are drawn from observation eror pdf $N(0, R)$. We sample the ensemble members x^0_i from the background errors $N(0, B)$. Observation and background errors are mutually uncorrelated, so the three matrices are diagonal. This means that effectively we are running N_x independent data assimilation systems at the same time, in which N_x is the dimension of the state space. We choose $N_x = 1000$ and observations are taken at every grid point and every time-step. The diagonal elements of Q, B and R are 0.04, 1, and 0.12 respectively.

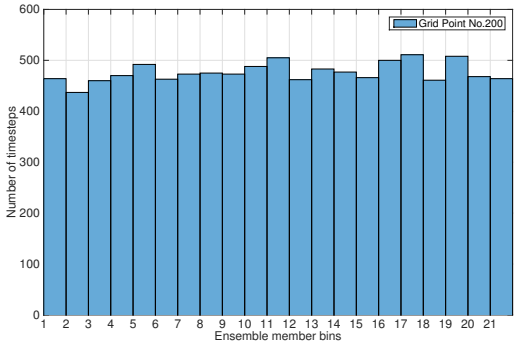
Rank histograms are used to evaluate the spread and quality of the ensemble over all the observation time steps. They are generated by ranking the observations in the set of perturbed forecast state variable ensemble members. In general, a flat histogram means that the observations are indistinguishable for any of the perturbed ensemble members and a humped rank histogram reveals too much spread for the ensemble. A U-shaped rank histogram is the evidence of being too little spread for the ensemble, see **?**.

Different percentages of the positive ε_i are chosen to test the rank histograms of this new scheme, 100%, 50% and 0%. In the 50% case ε_i is chosen positive or negative randomly with equal probability. After running the linear model for 10,000 time-steps with 20 ensemble members, the rank histograms of the particles are shown in Figure 4.

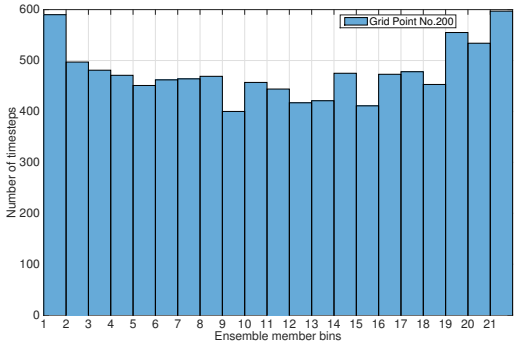
The figure shows that a different percentage of positive ε_i results in a different shape of the rank histogram. The humped histogram seen in Figure 4 indicates that all positive ε_i brings too much spread in the ensemble, while a U-shaped histogram seen in the third all negative ε_i histogram in Figure 4 is the evidence of too little spread. Fifty percent of randomly chosen positive ε_i generates a flat rank histogram that indicates a good quality ensemble.



(a) 100% positive ε_i



(b) 50% randomly chosen positive ε_i



(c) 0% positive ε_i

Figure 4. Rank histograms of grid point 200 using different percentage of positive ε_i in linear model experiment.

Comparing the ratio of RMSE and the spread of the analysis ensemble for the first 200 time-steps in Figure 5, fifty percentage of positive ε_i shows a stable ratio almost equal to one after some spin-up time-steps. Increasing or decreasing the percentage of positive ε_i causes a degradation in spread of the ensemble, which can be seen clearly from Figure 5.

We can also look at the shape and structure of the posterior pdf. Since we know the true posterior pdf is a Gaussian we can test how good our ensemble is. However, since the ensemble size is small a direct calculating of the posterior pdf is not very useful.

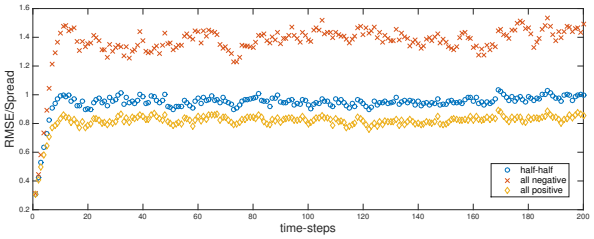


Figure 5. Ratio of RMSE and spread of the ensemble with different percentage of positive ε_i.

Instead we run the model for 1000 time steps and choose one grid-point, 200 in this case. The first 7 time steps are abandoned as the initial ensemble is quite wide, leading to statistical noise in the estimates. Then we move the mean of the ensemble at every time-step to zero and scale the ensemble values with $\frac{\sigma^{starttime}_{sample}}{\sigma^{j}_{sample}}$. The resulting ensembles at every time-step are then taken as one big ensemble, and a pdf is created, see Figure 6.

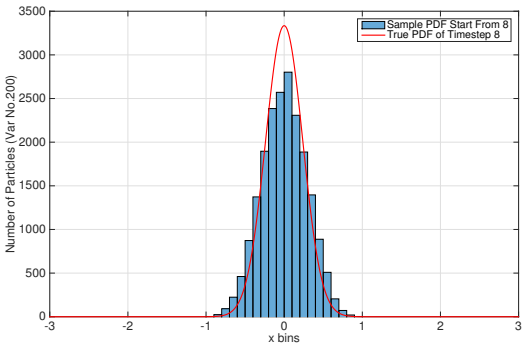
The red line is the true posterior pdf which is perfectly consistent with the 50% positive sample posterior pdf case. The sample posterior pdf becomes narrower when the percentage of positive ε_i decreases, consistent with the spread of the ensemble found in the rank histograms. Meanwhile, the sample posterior pdf becomes wider as the percentage of positive ε_i increases, again consistent with the rank histograms.

To check whether this scheme feasible for large number of ensemble members case, the number of ensemble members is increased to 1000, which leads to bias issue in the equivalent weights particle filter. Figure 7 shows that a choice of 50% positive ε_i leads to a pdf that is slightly too wide. Decreasing the percentage to 35 gives a better result. This result suggests that we might be able to choose ε_i in a better way, e.g. according to the probability mass on each side of the gap. This will be left for future research.

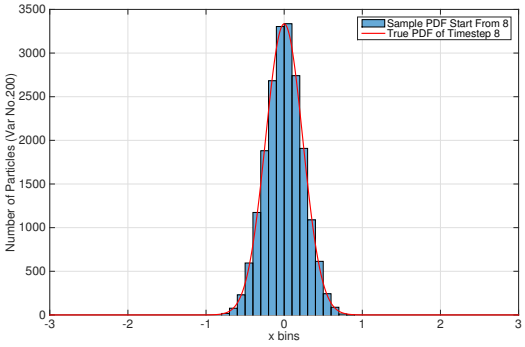
3.2. *High-Dimensional Lorenz96 Model Experiments*

In this section the new scheme is compared to the LETKF in a moderately high-dimensional setting of the Lorenz 1996 model. The Lorenz 96 model (**?**) is a dynamical model often used as a test model for new DA methods. It is defined as

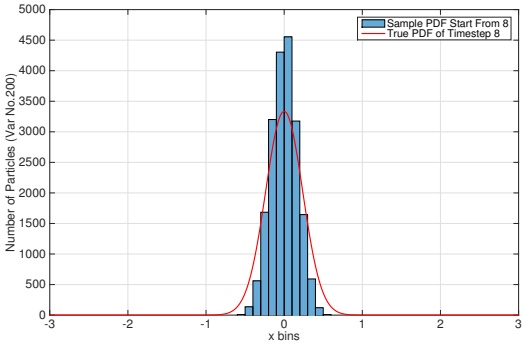
$$\frac{dx_i}{dt} = -x_{i-2}x_{i-1} + x_{i-1}x_{i+1} - x_i + F \qquad (53)$$



(a) 100% positive ε_i



(b) 50% randomly chosen positive ε_i

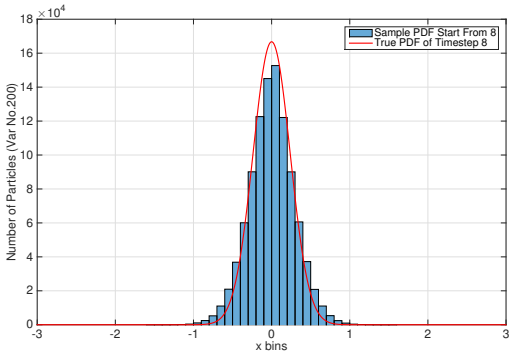


(c) 0% positive ε_i

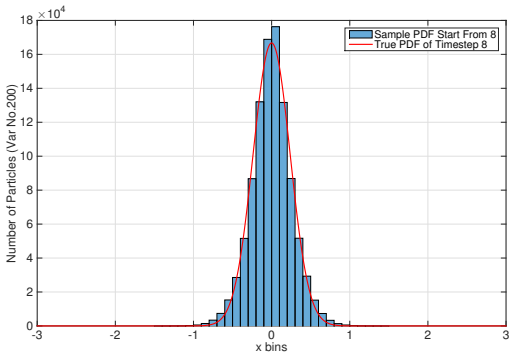
Figure 6. Posterior pdf represented by the particles using different percentage of positive ε_i in linear experiment.

where x_i is the state variable of the model at position i and F is a forcing constant, which is typically chosen as 8 for chaotic behaviour. The dimension of the Lorenz 96 model can be easily extended from 40 to 1,000, or more. In this section, 1,000 dimension Lorenz 96 model with 20 ensemble members is chosen for all the experiments.

As described in the previous section, the ε_i has two different real solutions in this new scheme, one is positive and the other is negative. We will explore the sensitivity to different choices of ε_i. To mimic realistic geophysical situations the grid points are observed every five time steps and three scenarios of spatial observation densities will be explored. The first one is observations at every grid point, the second one is observations



(a) 50% randomly chosen positive ε_i



(b) 35% randomly chosen positive ε_i

Figure 7. Posterior PDF in 1000 ensemble members with different percentage of positive ε_i.

at every other grid point and the last one is observations at the first half of the domain. We choose the model error covariance matrix Q and background error covariance matrix B as tridiagonal matrices. We used a time step of $\Delta t = 0.05$ with an RK4 scheme for the deterministic and an Euler scheme for the stochastic part of the model.

The LETKF uses same background error covariance matrix, model error covariance matrix, observation error covariance matrix and observation operator, and the initial ensemble of the two methods is the same. After some experimentation the localization radius of the LETKF is set to one grid point for best performance on RMSE, which is the standard measure of performance of the LETKF.

3.2.1. Parameter setting for the IEWPF

In this section we explore the parameter values of the IEWPF to determine the optimal setting for the Lorenz96 model, in which all variables are observed directly, every 5th timestep.

3.2.1.1. IEWPF without Relaxation Term In this experiment the relaxation term in the IEWPF is not used. The background covariance matrix B is a tridiagonal matrix with main diagonal value 1 and sub-diagonal value and super-diagonal value 0.25. The main diagonal value of Q is 0.10 and both sub- and super- diagonal values are 0.025. R is a diagonal matrix with main diagonal value 0.16. We observe every grid point every $5th$ time step. We use 50% positive ε_i.

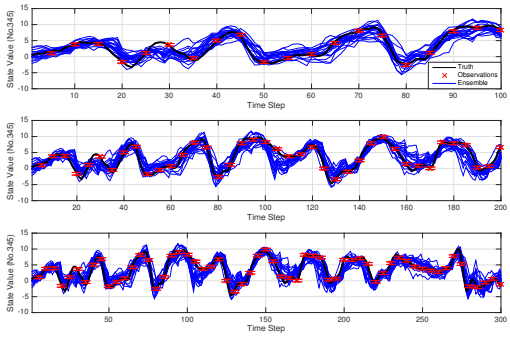


Figure 8. Trajectories of grid point 345 in this new scheme without relaxation term. The black line is the truth and the blue lines depict the evolution of the particles. Note the different time resolutions in the different plots.

The ensemble needs some spin-up time to reach a more-or-less steady spread, as depicted in Figure 8.

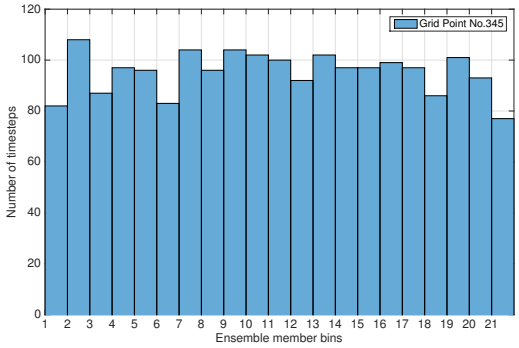


Figure 9. Rank histogram of grid point 345 after a 10,000 time steps model running.

Figure 9 illustrates rank histogram of observations with respect to the perturbed forecast ensemble members after a 10,000 time-steps model running for grid point 345. The flatness of the histogram shows that the ensemble has a good spread and quality.

3.2.1.2. IEWPF with Relaxation Term The relaxation makes particles move towards the high probability area at the time-steps between two adjacent observations. b and τ are the control parameters for the relaxation term in this new scheme, which decide the relaxation strength. b and τ are selected to be 0.25 and 0.50 as the best values in this scheme. The matrix parameters in this experiment are the same as those in the previous experiment

without relaxation term, and the observational scheme is also the same. ε_i is randomly chosen to be 50% positive.

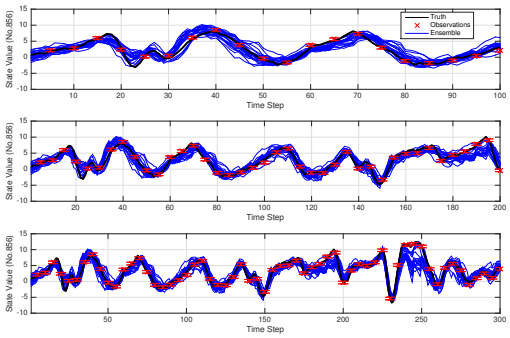


Figure 10. Trajectories of grid point 856 in this new scheme with relaxation term. The black line is the truth and the blue lines depict the evolution of the particles.

Figure 10 shows that only 20 ensemble trajectories can follow the truth properly under implicit equal-weights particle filter. Furthermore, the ensemble members are closer to the truth when the relaxation term is present. The spin-up time of ensemble model runs could be seen to be around 90 time-steps in Figure 10.

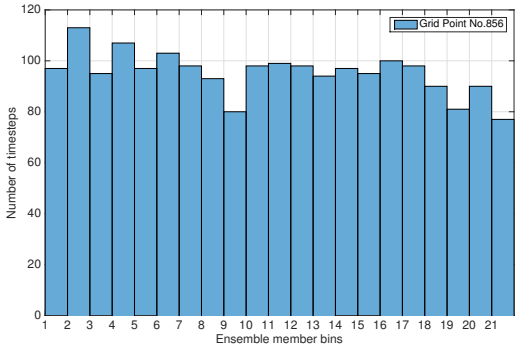
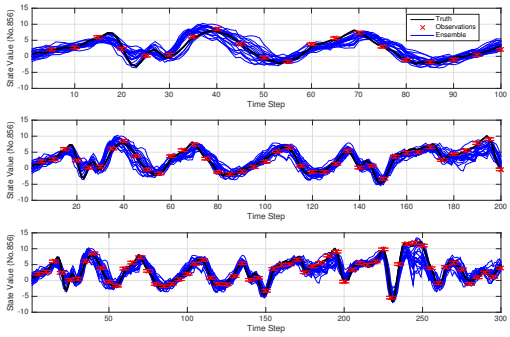


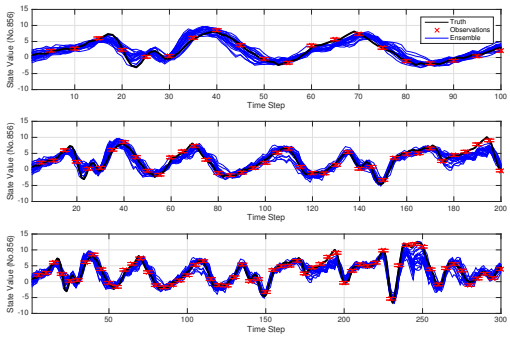
Figure 11. Rank histogram of grid point 856 after a 10,000 time-steps model running.

Figure 11 shows the rank histogram of observations with respect to the perturbed forecast particle ensemble members after a 10,000 time steps model running of grid point 856. The flatness of the histogram elucidates that the observations are indistinguishable from any perturbed ensemble member in the situation of observations at every grid point. The distribution histogram of the state variable is not Gaussian (not shown).

To investigate the sensitivity of IEWPF to the sign of ε_i, Figure 12 and Figure 13 illustrate the trajectories and rank histograms of the ensemble members for all positive (left) and all negative (right) ε_i. The system is still 1,000 dimensions and 20 ensemble members, and every grid point is observed every 5 time-steps.

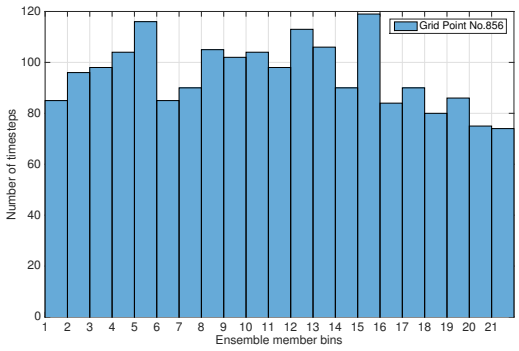


(a) all positive ε_i

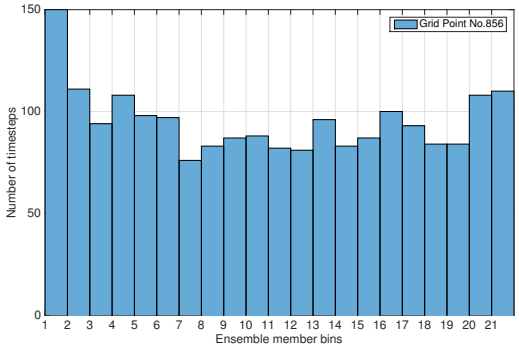


(b) all negative ε_i

Figure 12. Trajectories of grid point 856 in this new scheme with different percentage of positive ε_i. The black line is the truth and the blue lines depict the evolution of the particles.



(a) all positive ε_i



(b) all negative ε_i

Figure 13. Rank histogram of grid point 856 after a 10,000 time steps model running.

Figure 13 shows the rank histogram of the observations with respect to the perturbed forecast particle ensemble members of

grid point 856 accumulated over 10,000 time steps. The hump of the left histogram shows that ensemble has too much spread for all-positive ε_i. The right figure is a U-shaped rank histogram for the all-negative situation. There is no clear metric in measuring the sensitivity of rank histogram to the percentage of positive ε_i in a more quantitative way. For Figure 13, the sensitivity is moderate and not steep.

Comparing these two experiments, the IEWPF with relaxation term performs better than that without relaxation term, but the relaxation parameters need to be tuned. It is similar to the ensemble Kalman filters for systems that are not too nonlinear: the raw schemes are consistent, and by tuning the inflation factor and the localisation area a better performance can be achieved. This relaxation strength is used for all further experiments on this model, with 50% of the ε_i positive.

3.2.2. *Comparison of IEWPF and LETKF*

In this section the IEWPF will be compared to the LETKF. The localisation radius of the LETKF is set to 2 grid points, and the covariance inflation factor is 1.05, found as giving the lowest root-mean-square error (RMSE) after extensive experimentation.

Three experiments are compared, observing the whole state (exp1), every other grid point (exp2), and half of the state (exp3), all with observations every 5th time step. For the IEWPF we use $b = 0.25$ and $\tau = 0.5$ for exp1, $b = 1.2$ and $\tau = 0.6$ for exp2, and $b = 0.25$ and $\tau = 0.5$ for exp3.

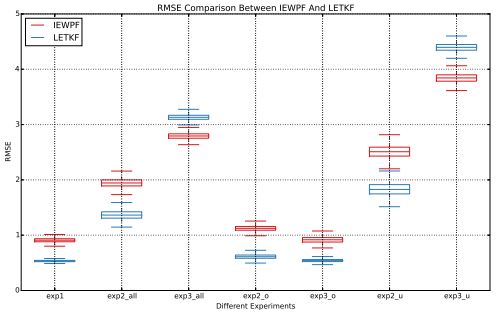


Figure 14. Root-mean-square error (RMSE) of the two methods for the different experiments, red for the IEWPF and blue for the LETKF. The subscript 'all' means RMSE over all gridpoint, 'o' denotes RMSE over only the observed grid points, and 'u' denotes RMES over only the unobserved grid points.

Figure 14 shows that the RMSE of the LETKF is systematically lower than that of the IEWPF apart from at the unobserved points. There it is outperformed by the IEWPF. Arguably importantly than the actual RMSE is the ratio of the RMSE to the spread in the ensemble. Figure 15 shows that both methods perform well on this measure for the first two experiments, and that the spread in the LETKF is way too low in experiment 3, while the IEWPF still performs well.

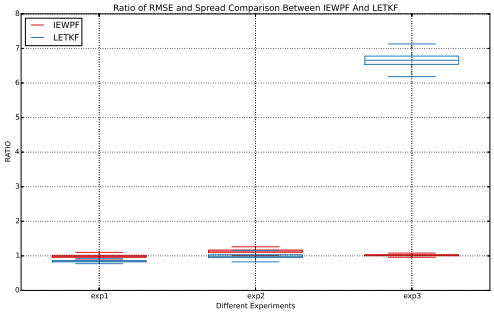


Figure 15. Ratio of the RMSE to the ensemble spread for the different experiments, red for the IEWPF and blue for the LETKF.

As a further comparison we look at the histograms for all experiments in figure 16. The LETKF is slightly over dispersive for the observed grid points in experiment 3, and strongly under dispersive on the unobserved grid points in that experiment. In contrast, the IEWPF performs well in all settings.

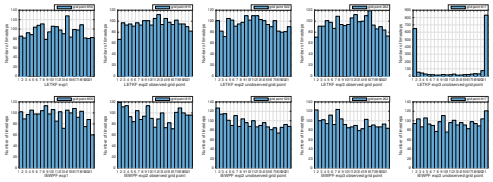


Figure 16. Rank histograms for the LETKF and the IEWPF for the different experiments.

4. Conclusions

A new DA method, the implicit equal-weights particle filter, has been presented in this paper. A flexible proposal density with a covariance that varies with the performance of each sample is used to make the particle weights all equal. It is essential that a model error term is included for this new method to work. This is not a serious drawback as it is well recognised that model errors are present and significant, but in practical applications model errors tend to be ignored. A relaxation term is included in the time-steps between two adjacent observations to make this new scheme more efficient. The addition of the relaxation term is included in the proposal density weights. The equal-weights conditions leads to a complicated matrix determinant ordinary differential equation that is hard to solve in general. We concentrated on the

high-dimensional systems, which allows for approximations that make the problem tractable. Interestingly, one particle can explore the full state space while all others experience a gap with zero proposal probability. However, this gap diminishes with the size of the system, making the new filter ideal for high-dimensional geophysical applications.

The IEWPF can be easily implemented in high-dimensional chaotic models. Two series of high-dimensional model experiments have been conducted, using 1,000 dimensional linear model and 1,000 dimensional non-linear Lorenz 96 model.

The linear model experiment using only 20 ensemble members shows that the particle weights does not degenerate as the dimension of the model state increases, without the exponential growth of the ensemble members reported by **?**. The new scheme preserves the posterior Gaussian pdf in linear model experiment. Increasing the ensemble size to 1,000, the simulated posterior pdf shows a Gaussian distribution which is slightly too wide. After decreasing the percentage of positive ε_i for 1,000 ensemble members, the simulated posterior pdf does resemble the true posterior. This bias is subject to further study and is likely related to the fact that we should choose the percentage of positive ε_i equal to the percentage of probability mass that has positive ε_i.

The performance of implicit equal-weights particle filter is also examined in 1000-dimensional non-linear Lorenz 96 model. Again the experiments show that this new scheme has very good consistency and convergence properties without filter degeneracy. A comparison with a tuned LETKF reveals that the RMSE errors of the latter tend to be smaller than those of the IEWPF, but the ensemble spread in the LETKF is too small when the observation density decreases. The ensemble spread is always equal to the RMSE in the IEWPF. This is also reflected in the rank histograms, which are too narrow for the LETKF when large portions of the system are unobserved. A lesson to learn from this is that concentrating only on the RMSE is not good practise in nonlinear data-assimilation systems.

The IEWPF was implemented with a weak relaxation term between observations to control the spread and to achieve a better converging trajectories of the ensemble members. This is a weaker part of the scheme, also it needs tuning. More sophisticated proposal densities can be used to improve performance further, and increase robustness of the scheme. For instance, one could extend the implicit equal weights staep over the whole trajectory between observations, as the Implicit Particle Filter does. The drawback of such a proposal is that an adjoint of the model is needed to make this efficient, although ensemble schemes like 4DEnsVar might also be explored.

The new scheme has been implemented into the EMPIRE data-assimilation software system (**?**), and experiments on high-dimensional geophysical systems are being planned.

Appendices

A. The high-dimensional limit

We need to solve the equation:

$$(\alpha_i - 1)\gamma_i - 2N_x \log \alpha_i^{1/2} - 2\log\left(\left|1 + \frac{\partial \alpha_i^{1/2}}{\partial \xi_i^n}\frac{\xi_i^n}{\alpha_i^{1/2}}\right|\right) - 2\log|P^{1/2}| + \phi_i + J_i^{prev} = C \tag{A1}$$

in which we have absorbed the constant factor $||P^{1/2}||$ in the constant C.

Let us now introduce the notation $a = \alpha_i^{1/2}$, $g = \gamma_i$, $n = N_x$, $\xi = \xi_i^n$ and $c = C + 2\log|P^{1/2}| - \phi_i - J_i^{prev}$. So each particle will have a different g, ξ and c and we need to solve for a. The equation to solve now becomes:

$$(a^2 - 1)g - 2n\log a - 2\log\left(\left|1 + \frac{\partial a}{\partial \xi}\frac{\xi}{a}\right|\right) = c \tag{A2}$$

To proceed we look for solutions in which α_i, so a, is only a function of ξ_i via $\gamma_i = \xi_i^T \xi_i$. The derivative now becomes:

$$\frac{\partial a}{\partial \xi} = \frac{da}{dg}\frac{\partial g}{\partial \xi} = 2\frac{da}{dg}\xi^T \tag{A3}$$

so that we end up with:

$$(a^2 - 1)g - 2n\log a - 2\log\left(\left|1 + 2\frac{da}{dg}\frac{g}{a}\right|\right) = c \tag{A4}$$

We proceed by pulling all log terms together as follows:

$$(a^2-1)g - 2\log\left[a^n\left(\left|1+2\frac{da}{dg}\frac{g}{a}\right|\right)\right] = c \tag{A5}$$

The argument of the log can be evaluated as:

$$a^n\left(1+2\frac{da}{dg}\frac{g}{a}\right) = a^{n-2}\left(a^2+2ga\frac{da}{dg}\right) = a^{n-2}\frac{da^2g}{dg} \tag{A6}$$

To proceed we introduce a new variable $b = a^2 g$ so that $a^2 = b/g$, leading to:

$$b - g + 2\log g^{n/2-1} - 2\log\left(b^{n/2-1}\left|\frac{db}{dg}\right|\right) = c \tag{A7}$$

Now pull all terms with b and g together to find:

$$\log\left(e^{-b/2}b^{n/2-1}\left|\frac{db}{dg}\right|\right) = \log(e^{-g/2}g^{n/2-1}) - \frac{c}{2} \tag{A8}$$

We take the exponential of both sides, leading to:

$$e^{-b/2}b^{n/2-1}\left|\frac{db}{dg}\right| = e^{-g/2}g^{n/2-1}e^{-c/2} \tag{A9}$$

This equation can be integrated over g to find:

$$\int e^{-b/2}b^{n/2-1}db = \int e^{-g/2}g^{n/2-1}dge^{-c/2} \tag{A10}$$

Now use

$$\int x^m e^{-\beta x}dx = -\frac{\Gamma(m+1,\beta x)}{\beta^{m+1}} \tag{A11}$$

to find:

$$\pm\,\Gamma(n/2, a^2 g/2) = \Gamma(n/2, g/2)e^{-c/2} \tag{A12}$$

The usefulness of this expansion comes from the fact that we can expand the $\Gamma(m,x)$ function for large arguments m and x in the following way (**?**). Write

$$Q(m,x) = \frac{\Gamma(m,x)}{\Gamma(m)} \tag{A13}$$

Now define $y = x/m$ and

$$z = y - 1 - \log y \tag{A14}$$

According to **?** we do not need to worry about the sign of z and we can write in general:

$$Q(m,x) = \frac{1}{2}erfc(\sqrt{mz}) + \frac{e^{-mz}}{\sqrt{2\pi m}}\left[\frac{1}{y-1} - \frac{1}{\sqrt{2z}} + O(\frac{1}{m})\right] \tag{A15}$$

Furthermore, the error function $erfc$ can be approximated for large arguments as

$$erfc(\sqrt{mz}) = \frac{e^{-mz}}{\sqrt{\pi mz}}\left[1 - \frac{3}{2mz} + O(\frac{1}{(mz)^2})\right] \tag{A16}$$

where we note that

$$mz = x - m - m\log y = O(N_x) \tag{A17}$$

Combining the two expansions we find for large x and m:

$$Q(m,x) = \frac{e^{-mz}}{\sqrt{2\pi m}}\left[\frac{1}{y-1} + O(\frac{1}{m}, \frac{1}{mz})\right] \tag{A18}$$

We thus find for our equation for a:

$$\pm\,\frac{e^{-w}}{\sqrt{2\pi}}\left[\frac{1}{a^2 g/n - 1}\right] = \frac{e^{-v}}{\sqrt{2\pi}}\left[\frac{1}{g/n-1}\right]e^{-c/2} \tag{A19}$$

in which

$$w = mz = \frac{1}{2}(a^2 g - n - n\log a^2 g + n\log n) \tag{A20}$$

and

$$v = mz = \frac{1}{2}(g - n - n\log g + n\log n) \tag{A21}$$

This can be evaluated with the absolute value as:

$$e^{-1/2[(a^2-1)g - n\log a^2]}\left|\frac{g-n}{a^2 g - n}\right| = e^{-c/2} \tag{A22}$$

which, taking the logarithm, results in:

$$(a^2-1)g - n\log a^2 - 2\log\left|\frac{g-n}{a^2 g - n}\right| = c \tag{A23}$$

In original variables we find:

$$\begin{aligned}\gamma_i\alpha_i - N_x\log\alpha_i + 2\log(|\alpha_i\gamma_i - N_x|)\\ = c + \gamma_i + 2\log(|\gamma_i - N_x|)\end{aligned} \tag{A24}$$

We can extract N_x from the logarithms on right and left hand side of the equation, leading to:

$$\begin{aligned}\gamma_i\alpha_i - N_x\log\alpha_i + 2\log\left(\left|\alpha_i\frac{\gamma_i}{N_x} - 1\right|\right)\\ = c + \gamma_i + 2\log\left(\left|\frac{\gamma_i}{N_x} - 1\right|\right)\end{aligned} \tag{A25}$$

We now note that the third term on the left-hand side is much smaller than the second term, and similarly on the right-hand side, leading to equation (32):

$$(\alpha_i - 1)\gamma_i - N_x\log\alpha_i = C - \phi_i - J_i^{prev} \tag{A26}$$

Now that we have found a solution for α_i we need to check if γ_i and $\alpha_i\gamma_i$ are much larger than zero. We know that for large N_x that γ_i is distributed according to $\chi^2_{N_x}$, so γ_i is large. For $\alpha_i\gamma_i$ we use the solution we generated equation (41), adopting the short-hand notation:

$$\alpha_i = a^2 = -\frac{n}{g}W\left[-\frac{g}{n}e^{-g/n}e^{-c/n}\right] \tag{A27}$$

Two solutions exist, $\alpha_i > 1$ and $\alpha_i < 1$. The former fulfils our requirement because if $\alpha_i > 1$ then $\alpha_i\gamma_i \gg 1$. So we have to check if $\alpha_i\gamma_i \gg 1$ for the W_0 solution, for any γ_i. We find

$$a^2 g = -nW_0\left[-\frac{g}{n}e^{-g/n}e^{-c/n}\right] \tag{A28}$$

We know $g \sim \chi^2_n$, so it has mean n and standard deviation $\sqrt{2n}$. Hence $g/n = O(1)$. Hence the smallness of the argument of W_0 comes from $e^{-c/n}$. For small arguments we can approximate:

$$\lim_{z\to 0} W_0(z) = z + O(z^2) \tag{A29}$$

so that, for c large:

$$a^2 g = n\frac{g}{n}e^{-g/n}e^{-c/n} = ge^{-g/n}e^{-c/n} \approx ne^{c/n} \tag{A30}$$

To understand the size of this term we can estimate c as:

$$\begin{aligned}c = C - \phi_i - J_i^{prev}\\ = \max_i(\phi + J^{prev}) - \phi_i - J_i^{prev} \approx \underset{i}{max}(\phi) - \phi_i\end{aligned} \tag{A31}$$

The standard deviation in $\phi \approx \delta\sqrt{N_y}$ in which δ a constant of order 1, and N_y is the number of independent observations. This suggests that $c = O(\sqrt{N_y})$, such that $e^{-c/n} \approx e^{-\sqrt{N_y}/N_x} = O(1)$. This suggests that $a^2 g \gg 1$ always, and the full proof is given.

B. The size of the gap

As discussed in the main text there is one particle that explores state space fully, but all others experience a part of state space that cannot be reached. We show in this appendix that this gap decreases compared to the high-probability area that can be reached when the system size increases. First we calculate the width of the hypersphere in the high probability region of state space. We concentrate on the f_0 branch as the width of the f_{-1} branch will be larger. This will be followed by an estimate of the width of the gap.

The high-probability region is defined as the area in state space resulting from varying $|\xi|$ within its standard deviation (or a multiple of that, but that factor won't matter in the order of magnitude calculation).

We introduce the short-hand notation $x = |\xi|$. The distribution of x can be found as follows. We know that $y = x^2 \sim N(N_x, 2N_x)$. Hence we find:

$$p_x(x) = p_y(y)\frac{dy}{dx} = 2xp_y(x^2) = \frac{2x}{\sqrt{4\pi N_x}}\exp\left[-\frac{(x^2 - N_x)^2}{4N_x}\right] \tag{B1}$$

Typical variations in x are given by its standard deviation, so we calculate its variance:

$$\delta_x^2 = \int_0^\infty \frac{x^3}{\sqrt{\pi N_x}}\exp\left[-\frac{(x^2 - N_x)^2}{4N_x}\right]dx \tag{B2}$$

The transformation $z = (x^2 - N_x)/\sqrt{4N_x}$, so $dz = 2xdx/\sqrt{4N_x}$ leads to:

$$\delta_x^2 = 2\sqrt{\frac{N_x}{\pi}}\int_{-\sqrt{N_x}/2}^\infty\left(z + \frac{\sqrt{N_x}}{2}\right)\exp\left[-z^2\right]dz \tag{B3}$$

This can be evaluated as:

$$\delta_x^2 = -\sqrt{\frac{N_x}{\pi}}\exp\Big[-z^2\Big]_{z=-\sqrt{N_x}/2}^{z=\infty} + \frac{N_x}{\sqrt{\pi}}\int_{-\sqrt{N_x}/2}^{\infty}\exp\Big[-z^2\Big]dz \tag{B4}$$

which, for N_x large, becomes:

$$\delta_x^2 = \frac{N_x}{\sqrt{\pi}}\frac{\sqrt{\pi}}{2}erfc\Big[-\sqrt{N_x}/2\Big] = \frac{N_x}{2}\left(1 - erf\Big[-\sqrt{N_x}/2\Big]\right) \tag{B5}$$

For N_x large the erf approaches -1, so we find:

$$\delta_x^2 \approx N_x \tag{B6}$$

leading to a standard deviation of

$$\delta \approx \sqrt{N_x} \tag{B7}$$

We now estimate the size of the width as:

$$\begin{aligned} dR_0 &= f\Big(\sqrt{N_x}+\delta_x\Big) - f\Big(\sqrt{N_x}\Big) \\ &= 2\sqrt{N_x}\sqrt{-\frac{1}{4}W(-4\exp(-4-c/N_x))} \\ &\qquad - \sqrt{N_x}\sqrt{-W(-\exp(-1-c/N_x))} \\ &= \sqrt{N_x}\Big[\sqrt{-W(-4\exp(-4-c/N_x))} \\ &\qquad - \sqrt{-W(-\exp(-1-c/N_x))}\Big] \\ &= \sqrt{N_x}\Big[\sqrt{-W(-4\exp(-4-c/N_x))} - 1\Big] \approx \sqrt{N_x} \end{aligned} \tag{B8}$$

Our next step is to calculate the size of the gap. The gap is situated around the transition from the f_0 solution to the f_{-1} solution. The asymptotic expansion of Lambert W function around $z = -e^{-1}$ is given by

$$W_0(z) = -1 + p_0 - O(p_0^2) \tag{B9}$$

and

$$W_{-1}(z) = -1 + p_{-1} - O(p_{-1}^2) \tag{B10}$$

in which $p_0 = \sqrt{2(ez+1)}$ and $p_{-1} = -p_0$. As detailed in Appendix C , the gap is the difference between the two solutions from the two branches situated at $\gamma_i = N_x$. Therefore, the analytical expression for $f_{-1} - f_0$ at the gap with $z = -e^{-x-1}$ and $x = c/N_x$ is given by:

$$\begin{aligned} dR_{gap} &= f_{-1} - f_0 \\ &= \sqrt{N_x}[\sqrt{(-W_{-1}(z))} - \sqrt{(-W_0(z))}] \\ &= \sqrt{N_x}[\sqrt{(1+\sqrt{2(ez+1)})} - \sqrt{(1-\sqrt{2(ez+1)})}] \\ &\approx \sqrt{N_x}\sqrt{2(ez+1)} = \sqrt{N_x}\sqrt{2(1-e^{-x})} \\ &\approx \sqrt{N_x}\sqrt{(2x)} = \sqrt{(2c)}. \end{aligned} \tag{B11}$$

We now find that the ratio of the gap width, so the forbidden area, to the width of the allowed area with high probability is given by:

$$\frac{dR_{gap}}{dR_0} \approx \frac{\sqrt{2c}}{\sqrt{N_x}} \tag{B12}$$

Since the typical value of c is $c \approx \sqrt{N_y}$ which can be maximised by $\sqrt{N_x}$, we find that the relative area of the forbidden part decreases faster than $N_x^{-1/4}$, proving our point.

C. The position of the gap

We have found in the main text that there are two branches of solutions that do not connect for all but one particle. In this appendix we determine where the branches are closest, which allows us to find the width of the gap, explored in Appendix B. To this end we need the position in ξ space where the maximum of the f_0 branch and the minimum of the f_{-1} branch are. Since there is no directional preference for random ξ we can evaluate the derivative along one arbitrary direction. We choose for ξ the vector $\xi_i = (s, 0, 0, \cdots)^T$ in which $s = \sqrt{\gamma_i}$. This means that the functions are now scalar functions of scalar s, with derivative:

$$\frac{\partial f(s)}{\partial s} = \alpha_i^{1/2} + s\frac{\partial \alpha_i^{1/2}}{\partial s} \tag{C1}$$

To evaluate this we go step by step. $\frac{\partial \alpha_i^{1/2}}{\partial s}$ is calculated firstly. We use

$$\frac{\partial \alpha_i^{1/2}}{\partial s} = \frac{1}{2}\alpha_i^{-1/2}\frac{\partial \alpha_i}{\partial s} \tag{C2}$$

and calculate $\frac{\partial \alpha_i}{\partial s}$ using the analytical solution of α_i to find

$$\begin{aligned}\frac{\partial \alpha_i}{\partial s} &= -N_x \frac{\partial \frac{1}{s^2}}{\partial s} W[g(s)] - \frac{N_x}{s^2} \cdot \frac{\partial W[g(s)]}{\partial \xi_i^n} \\ &= -N_x \frac{-2}{s^3} W[g(s)] - \frac{N_x}{s^2} \cdot \frac{\partial W[g(s)]}{\partial g(s)} \cdot \frac{\partial g(s)}{\partial s} \end{aligned} \tag{C3}$$

where $g(s) = -\frac{\gamma_i}{N_x} \cdot e^{-\frac{\gamma_i}{N_x}} \cdot e^{-\frac{c}{N_x}}$.

Progressively do the calculation of $\frac{\partial W[g(s)]}{\partial g(s)} \frac{\partial g(s)}{\partial s}$,

$$\begin{aligned}&\frac{\partial W[g(s)]}{\partial g(s)} \cdot \frac{\partial g(s)}{\partial s} \\ &= \frac{W(\cdot)}{g(s)(1+W(\cdot))} \left\{ \frac{-2s}{N_x} \cdot e^{-\frac{\gamma_i}{N_x}} \cdot e^{-\frac{c}{N_x}} \right. \\ &\qquad \left. - \frac{s^2}{N_x} \cdot \frac{-2s}{N_x} \cdot e^{-\frac{\gamma_i}{N_x}} \cdot e^{-\frac{c}{N_x}} \right\} \\ &= \frac{W(\cdot)}{g(s)(1+W(\cdot))} \cdot e^{-\frac{\gamma_i}{N_x}} \cdot e^{-\frac{c}{N_x}} \left\{ -\frac{2s}{N_x} + \frac{2s^3}{N_x^2} \right\} \\ &= \frac{W(\cdot)}{g(s)(1+W(\cdot))} (-\frac{2s}{N_x}) e^{-\frac{\gamma_i}{N_x}} \cdot e^{-\frac{c}{N_x}} \left\{ 1 - \frac{s^2}{N_x} \right\} \\ &= \frac{W(\cdot)}{1+W(\cdot)} \cdot \frac{2}{s} \left(1 - \frac{s^2}{N_x} \right) \end{aligned} \tag{C4}$$

where $W(\cdot)$ is the representation of $W[g(\xi)]$.

Now we go backward to achieve the full expression of the original equation. Plug equation (C4) into (C3), it is easy to see that

$$\begin{aligned}\frac{\partial \alpha_i}{\partial s} &= -N_x \frac{-2}{s^3} W[g(s)] - \frac{N_x}{s^2} \cdot \frac{W(\cdot)}{1+W(\cdot)} \cdot \frac{2}{s} \left(1 - \frac{s^2}{N_x} \right) \\ &= \frac{2N_x}{s^3} W[g(\xi)] \left\{ 1 - \left(\frac{1}{1+W[g(s)]} \right) \left(1 - \frac{s^2}{N_x} \right) \right\} \\ &= -\frac{2}{s} \alpha_i \left\{ 1 - \left(\frac{1}{1+W[g(s)]} \right) \left(1 - \frac{s^2}{N_x} \right) \right\} \end{aligned} \tag{C5}$$

Turn back to equation (C2), to find:

$$\frac{\partial \alpha_i^{1/2}}{\partial s} = \alpha_i^{1/2} \left(-\frac{1}{s} \right) \left\{ 1 - \left(\frac{1}{1+W[g(s)]} \right) \left(1 - \frac{s^2}{N_x} \right) \right\} \tag{C6}$$

Therefore, we get that

$$\begin{aligned}\frac{\partial f(s)}{\partial s} &= \alpha_i^{1/2} - \alpha_i^{1/2} \frac{W(\cdot) + \frac{s^2}{N_x}}{W(\cdot)+1} \\ &= \alpha_i^{1/2} \left(1 - \frac{W(\cdot) + \frac{\gamma_i}{N_x}}{W(\cdot)+1} \right) \end{aligned} \tag{C7}$$

from which we immediately see that the gap appears where $\gamma_i = N_x$.

Acknowledgement

The authors thank the National Centre for Earth Observation (NCEO), funded by the Natural Environment Research Council (NERC), National Natural Science Foundation of China (NSFC, Grant Nos. 41375113, 41475094), for supporting this collaboration.