

## Reviewer 1

We are grateful to Dr. Van Leeuwen for all the valuable comments and suggestions. We have revised the manuscript accordingly and provide our point-by-point responses below.

### **Comment:**

*Line 6: I would suggest a small change to emphasize that the KL divergence is minimized at each iteration between the pdf at that iteration and the posterior pdf, starting at the first iteration with the prior pdf.*

### **Response:**

In line 5, we have incorporated the reviewer's suggestion to clarify the relevant details.

### **Comment:**

*Line 32: This is still not correct as the likelihood can be non-Gaussian. Please be accurate. Note also the word 'also' in line 35. This is simply not correct.*

**Response:** The statements have been removed from lines 33 and 36.

### **Comment:**

*Introduction: I would strongly suggest to split the discussion of the particle filter and the particle flow filter. They are completely different methods, both in philosophy and in implementation. The authors keep mixing them in their discussion, which is not helpful to the community. My suggestion is to first discuss the need for nonlinearity, then discuss the particle filters, then their issues and potential solutions (and I notice proposal densities are not discussed at all, see e.g. the Particle filter review in Van Leeuwen et al. 2019 for a more complete discussion, but there are many other sources). Then discuss another solution, the particle flow filters, and the connection with MCMC methods for the stochastic variants.*

### **Response:**

We have restructured the introduction to clearly separate the discussion of particle filters and particle flow filters. The revised introduction now discusses proposal densities (lines 43-46), introduces particle flow filters as an alternative approach with their connection to MCMC methods (lines 52-56), and reorders the presentation to better distinguish these methodologies.

### **Comment:**

*Line 213: "with respect to the  $l$  variable."*

### **Response:**

We have made this grammatical correction in line 227.

### **Comment:**

*Line 510: The underlying reason that the Kalman filters show this behavior is that the data assimilation only scales the prior, it does not change its shape. Hence, if the prior has one or two outliers, these outliers will remain in the posterior, and can grow towards the next data assimilation time, etc. The PFF do change the shape of the prior due to the communication among the particles during the iterations. The authors might want to add something like this.*

### **Response:**

This suggestion has been addressed in line 521.

## Reviewer 2

We greatly appreciate Dr. Farchi's thoughtful follow-up comments and careful review of our revised manuscript. Our detailed responses are provided below.

**Comment:** *L 27-28: repetition “particularly” / “particular” in the first sentence.*

**Response:**

The repeated use of “particularly” and “particular” has been corrected by replacing “particularly” with “especially” in line 27.

**Comment:** *L31: “the challenge” → be more precise: which challenge?*

**Response:**

The reference to “the challenge”, in line 31, has been clarified as relating to the representation of enhanced non-linear and non-Gaussian behaviour.

**Comment:**

*L32: I think that this statement is still deceiving. The assumption of Gaussian errors in 4D-Var is not fundamental, and made for practical reasons, contrary to the Kalman filter which is by construction a Gaussian method. Perhaps you can simply remove 4D-Var from this sentence, without changing your message?*

**Response:**

The statement has been removed from line 33.

**Comment:**

*L 43-45: You could perhaps slightly reformulate the paragraph. In particular, avoid starting two consecutive sentences with “Other”.*

**Response:**

The paragraph has been slightly reformulated by removing the first “Other” and replacing the second one with “Further” to improve readability (line 48).

**Comment:**

*L 45: This is a detail, but “jittering” is actually the same technique as what other authors (including myself) call “regularisation” and is not only used after tempering steps, but also in more “standard” implementations of the PF after (or before) resampling.*

**Response:**

We modified the sentence to indicate that jittering (or regularisation) is not restricted to tempering steps, but can also be applied after resampling (line 50).

**Comment:**

*Eq. 5: Introducing a constant “C” to have consistent physical dimensions looks like a quick patch that could potentially hide a more serious issue. That being, reading the other reviewer comment, I have the feeling that fixing this issue seems more complex than one can think of. For the present work, I think that it would be perhaps more appropriate to recognise that there is some inconsistency, and that more research is needed to solve it. (But this is only my opinion)*

**Response:** We have now clarified that this is not an inconsistency but that this work uses nondimensionalized variables as in previous references, (line 153). If one wants to work with dimensionalized variables a diffusion coefficient must be included in the formulation. The diffusion coefficient value in that case is related to the convergence rate but it has no physical meaning because the convergence is in pseudo-time (at a fixed physical time).

**Comment:**

*L 175-180: perhaps you could give a little bit more detail and briefly explain to what each term in L. 177 and 179 correspond? Also, I strongly advise to number all equations.*

**Response:**

We have added clarifications explaining the computational complexity terms in line 190. Following the reviewer's advice, we have also numbered all equations throughout the manuscript.

**Comment:**

*L 417: “the LMPF’s” → “the LMPFs”?*

**Response:**

We corrected the plural of LMPF and applied the same change consistently throughout the manuscript.

## Editor

We thank the Editor, Dr. Talagrand, for the suggestions and careful evaluation of our manuscript. We have revised the paper in accordance with these remarks, and our responses are detailed below.

**Comment:** *Ll. 42-43, “... such as localization, first introduced in (Bengtsson et al., 2003) and independently in van Leeuwen (2003)”. You then mention (l. 75) earlier references to localization (in the context of EnKF, but that may be confusing). I suggest you change the text above to “... such as localization, first introduced for particle filters, and independently, in (Bengtsson et al., 2003) and van Leeuwen (2003)”.*

**Response:**

The sentence in lines 48–48 has been modified according to the editor's suggestion.

**Comment:** *Ll. 62-63, sentence starting “To address this lack ...” I suggest “To improve its convergence properties, Ba et al. (2022) proposed alternative formulations of SVGD”*

**Response:**

The text in line 71 has been revised as suggested by the editor.

**Comment:**

*L. 82, “While **that** work ...” is more appropriate here (compare with “This work ...” in l.47)*

**Response:**

The sentence in line 91 has been changed according to the editor's instructions.

**Comment:**

*L. 125, define superscripts  $f(j)$  and  $a(j)$*

**Response:**

This point is clarified in line 133.

**Comment:**

*Ll. 200-201, word missing ? “... in Hunt et al. (2007) **and** has some resemblance ...”*

**Response:**

The sentence in line 214 has been corrected according to the editor's instructions.

**Comment:**

*Ll. 205 and 214 (and maybe elsewhere). From what I understand, it is not a question of correlation, but of statistical dependence. I suggest to write on l. 205 “we assume that the variables located outside of a neighborhood  $C_l$  of  $x_l$  are statistically independent of  $x_l$ ”.*

**Response:**

We have addressed this suggestion, and the text has been modified in line 218.

**Comment:**

*L. 219,  $\Gamma_l \circ \Sigma$  , I understand  $\circ$  denotes the Schur product ?*

**Response:**

This point has been clarified in line 232.

**Comment:**

*L.219-220, "... around l with only one's and zeros, as in eq. (15) below"*

**Response:**

The text in line 234 has been revised as suggested by the editor.

**Comment:**

*L. 478, the number given there (16-17, 18-19) must be the number of positive Lyapunov exponents, not their value. Say it more precisely.*

**Response:**

This point has been clarified in line 492.