Reply to rev. 2

We would like to thank the referee for his/her review on our paper and for giving us the opportunity to improve our paper.

We have improve the description of the numerical experiments with some details on the implementation used: finite difference for the spacial discretization, a fourth order Runge-Kutta for the time scheme and we have specified the numerical setting (time step, numerical value of the diffusion coefficient). The ensemble size has been increased to 6400 in order to limit the sampling noise, and a single ensemble of normalized error has been generated then used with appropriate initial error magnitude – this reduces the sampling fluctuations when comparing the numerical results from a method to another.

In order to investigate the limitation of the tangent-linear covariance dynamics the manuscript incorporates new results (even if further research are still needed to investigation more completely the limitations of the parametric formulation as highlighted in the manuscript):

A study of the mean predicted by the parametric model and estimated from the ensemble has been introduced in order to illustrate the ability of the PKF to provide an estimation of the true mean state when small non-linearities are present: see Fig. 4 and the new section 4.3.1. A long term experiment has been introduced to determine if there is an exponential growth of the error that could be a side effect of the tangent-linear approximation: see Fig. 8 and the new section 4.3.3. The discussion of the results has been put in a new section 4.3.4.

We copied your commentary in italics below, we reply in normal blue font

1) *page 1 Abstract: "Abstract: I am not sure if you can say that "this study extends the PKF to nonlinear dynamics". Rather, It is a required preliminary step."*
Since the parametric model is designed to the nearly nonlinear dynamics we think this terminology is appropriate. This is supported by the introduction of the mean predicted by the ensemble versus the parametric model (Fig. 4) where the mean is different from the reference state at T.

2) *page 2 line 20 "numerical tests"*
The typos is now corrected

3) *page 3 line 2 "Background on the"*
The typos is now corrected

4) *page 4 line 15 I am not really keen to use a semigroup in order to define a multivariate matrix. What stops you writing the matrix explicitly ?*
The matrix results from numerical integration of the diffusion equation, it is not easy to write the matrix explicitly.

5) *page 5 line 10 "meaningful"*
The typos is now corrected.

6) *page 5 line 13 The dimension of the space plays no role in here.*
We agree with the referee comment but we think it is interesting to extrapolate the 1D situation to the 2D/3D case much important for further applications, where Eq. (11) offers a systematic derivation of the dynamics, whatever the dimension is.

7) *page 6 line 5 Can I suggest that you avoid using the term "nonlinear Kalman filter".*

The Kalman filter is, by definition, linear. Yes, there are extensions of the Kalman filter methodology to nonlinear frameworks (the extended Kalman Filter, the ensemble Kalman filter, etc), but I believe that here you are referring to the framework which is either nonlinear or linear (Kalman).

The terminology has been employed by Cohn (1993) and corresponds to the next order of the extended Kalman filter where the magnitude of the fluctuation influences the mean state. The sentence has been rephrase as follows:

" Note that the fluctuation-mean flow interaction leads to the
Gaussian second-order filter \citep[sec. 9.3]{Jazwinski1970book},
and is important in nonlinear Kalman-like filters \citep{Cohn1993MWR}."

8) *page 6 line 12 "deduced from the difference"*

The typos is now corrected.

9) *page 7 line 21 As I understand it, \overline{∂ x ε 2} is the expectation of ∂ x ε 2 , not the approximation obtained by the ensemble average."*

Yes this is correct: in this expression there is no approximation, and you are right this corresponds to the expectation operator of the derivative of $ε^2$. *At a theoretical level, the "ensemble average" considering an infinite ensemble size is equivalent to the expectation operator. And this infinite ensemble average should not be confused with the finite ensemble average as used with an EnKF: the theoretical derivation does not rely on any finite ensemble as encountered in the EnKF. The deterministic equations Eq.(29) of the parametric model does not need any ensemble.*

10) *page 9 line 13 There is a comma missing after the exponential.*

The typos is now corrected.

11) *page 10 line 8 "expressed" .*

The typos is now corrected.

12) *page 10 line 23 How large is the ensemble for the numerical experiment ? Why do you call it nonlinear ?*

The ensemble size was mentioned in p11, line 12: 1600 members. The sentence has been rephrased as: *"Then, the PKF is assessed using a large ensemble of nonlinear forecasts (6400 members)".*

13) page 11 Caption for Figure 1: the figure has 6 curves. I assume you plot the solution at 0.8T too (same for Figure 4 page 15, Figure 5 page 16) :

The time 0.8T has been added in the caption.

14) *page 13 line 17 What does the 17.6% figure signify ?*

This refers to the result at time t=T. The sentence has been rephrase to clarify: *"with a low relative error $||K - K^{GC}||/||K||$ of 9.4% (respectively 17.6%) at time 0 (respectively T), "*

15) *page 14 Figure 3 What is the maximum difference between the two graphs ?*

We have precise the maximum difference of the two kurtosis normalized by K_G: 0.05

16) *page 15 line 10 In the case of Burgers' equation, can you provide an estimate of the length of the time interval throughout which the method presented in the paper gives meaningful results.*

In order to improve this point, additional results have been incorporated in the manuscript:

-1- Prediction of the mean:

For the Burgers equation, in the parametric model Eq.(29), the dynamics of the mean flow is

the true dynamics of the ensemble mean. We have insisted on the exactness of Eq.(14a) to predict the ensemble mean with the sentence:
"Moreover, as pointed out in Ménard (1994),
Eq. (14a) is the exact the ensemble mean for the Burgers dy-
namics, while Eq. (14b) is an approximated dynamics. As a
consequence, if the variance field is the true one, then the
mean predicted by Eq. (14a) is the true ensemble mean (Mé-
nard, 1994, sec. 5.5.2)." (sec. 3.1)

The new section 4.3.1 illustrates the case of the mean (see also the new Fig. 4) .

-2- Introduction of a long term experiment:

The new Fig. 8 and new section 4.3.4 have been introduce to tackle (but only in part) the time where the parametric model is valid.

17) *page 16 Figure 16. Again, it would help to know how large is the ensemble used in the numerical experiments.*
In the new version of the manuscript, the ensemble size has been increased to 6400 with a single ensemble of normalized error to facilitate the comparison of the numerical experiments.

18) *page 17 line 22 "The aim of this section"*
The typos is now corrected.

19) *page 18 line 2,4 Something went wrong with the display of the formulae (A3) and (A4)*
This is due to the "manuscript version" configuration of the manuscript (single column while "article version" is two columns). The two column version of the manuscript does not present this wrong display.