Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2017-44-AC1, 2017

Nonlinear Processes in Geophysics Discussions

Interactive comment on "The Onsager–Machlup functional for data assimilation" *by* Nozomi Sugiura

N. Sugiura

nsugiura@jamstec.go.jp

Received and published: 13 September 2017

The author really appreciates the 1st reviewer for careful reading of the manuscript including the Appendix. Although the revision according to the reviewer's general remark will take some time, I provide quick and tentative replies to the specific comments in order to enhance the discussion.

1. *The first sentence in the abstract is confusing or inaccurate, because one does not have to resort to the more advanced and more difficult OM functional, as long as the stochastic differential equation (SDE) is to be solved in a time-discretized form (rather than time-continuous form) for data assimilation application. Nonlinear SDEs can rarely be solved analytically in time-continuous form. Although the OM functional is*

useful and important for rigorous theoretical considerations when the time-continuous limit is applied to a time-discretized form of quadratic cost function, the time-continuous limit can be derived formally (or intuitively) without considering the OM functional, as shown in (5.36)-(5.40) on pages 155-156 of Jazwinski (1970: Stochastic Processes and Filtering Theory). Since the SDEs are actually solved in time-continuous forms in this paper (as well as in most data assimilation studies), the importance and utility of the OM functional for real data assimilation appear to be overstated in this paper.

All of the SDEs are actually solved in time-discrete form in this paper, and thus I am sure that the importance of OM functional in data assimilation is properly illustrated. The derivation in Jazwinski (1970) is valid for the assignment of each path probability, but we should be careful when we consider an optimization problem. If we solve the optimization problem with Jazwinski's strategy, that should lead to curves like "E" or "T" in Fig.3 and Fig.5, which is clearly less meaningful than "ED" or "TD".

Even in a discrete-time setting, a path drawn with model error term is generally not differentiable in time direction since the random term on each time slice adds an independent noise, and thus a smoother, whose object is smooth functions, cannot optimize the paths itself. What we can do is to draw smooth curves and compare the densities of paths in their ϵ-neighborhoods, which is shown in the manuscript.

2. *The noise intensity σ in (1) is a constant rather than a function of x_t . In this case, as explained at the end of section 4.6 (pages 119-120) of Jazwinski (1970), the associated Ito integral and Stratonovich integral are identical. Thus, as a stochastic differential equation, (1) can be viewed either as an Ito equation or its equivalent Stratonovich equation. The authors may need to clarify this point.*

I agree with the reviewer's comment that there is no need to distinguish the Ito integral from the Stratonovich integral with regard to the discretization of the stochastic differential equation (SDE). Note that the distinction is applied in the manuscript only to the discretization of the OM functional, not to the SDE itself, because the quadratic term in the former contains the product of the noise dw_t and the process dependent term $f(x_t)$.

3. *Eq. (13) is derived from (A2) by the scaling for a smooth curve, but the scaling $x_n - x_{n-1} = O(\delta_t)$ is not explained in Appendix A.*

In the case of a smooth curve, there is no stochastic term, and thus $x_n - x_{n-1}$ is the product of a bounded function $f(x_{n-1})$ and δ_t, which results in a value with $O(\delta_t)$. I will mention that in the revised manuscript.

4. *It appears that (21) is derived from (13) and (19) with ϕ changed into x, but it is not clear why ϕ can be changed into x.*

The symbol x in (21) represents either ϕ or x. Since the notation was confusing, I will change the symbol x to χ (or something) in (21) and the related expressions.

5. *It is not shown and unclear how (22) is derived.*

The four cases are the conceivable combinations of the timing of the drift term $f(x_t)$ and the presence or absence of the divergence term. Equation (22) is just one of them. I will mention that in the revised manuscript.

6. *Due to the questions in above comments 3-4, it is not clear whether the four schemes considered in section 2 all converge to the same time-continuous limit. If the answer is yes, then the differences between the numerical results obtained from the four schemes for each example in section 3 are caused the differences in discretization, and these differences should diminish as δ_t approaches to zero. To verify this numerically, the authors need to show for each example that the differences between the numerical results obtained from the four schemes become increasingly small as δ_t decreases toward zero.*

Not all of them converge to the same limit. Rather, the schemes that I judged to be applicable (see table 1) are expected to converge to the common answers, which you can see clearly in Figs. 2 to 5. I will mention that in the revised manuscript.

7. *Section 6.3.2 of Law et al. (2015) is cited for the derivation of divergence term in Appendix B. I checked but found that there are only 5 chapters in Law et al. (2015).*

I am afraid you refer to the preprint version of Law et al. (2015) in arxiv. Please refer to the commercially published version, which has 9 chapters.

8. *The formulation on the line above (B7) appears to be for $\delta\mu_0/\delta x$ or $d\mu_0/dx$ [that is, the variation of μ_0 with respect to variation of $x(t)$ for $0 \le t \le T$] rather than for the Wiener measure μ_0 itself. Similarly, μ should be $\delta\mu/\delta x$ or $d\mu/dx$ on the left-hand side of (B7) and (B8). Correct?*

As the reviewer pointed out, the expression for the measure was inaccurate. I will change μ_0 to $\mu_0(d\Omega)$, μ to $\mu(d\Omega)$ as well, and add the explanation as follows.

We divide a time interval $[0,T]$ to N equal segments $[(n-1)\delta_t, n\delta_t],\ n = 1, 2, \cdots, N$ with $T = N\delta_t$, and consider a cylinder set $\Omega \equiv \Omega_1 \times \Omega_2 \times \cdots \times \Omega_N$ where $\Omega_n \subset \mathbb{R}^D$ is on time slice $t = n\delta_t$. We can define a measure for the cylinder set Ω as

$$\mu_0(\Omega) = \int_{\Omega_1} p(\delta_t, x_1 - x_0)dx_1 \int_{\Omega_2} p(\delta_t, x_2 - x_1)dx_2 \cdots \int_{\Omega_N} p(\delta_t, x_N - x_{N-1})dx_N,$$

$$p(\delta_t, x' - x) = \frac{1}{\sqrt{2\pi\delta_t}} \mathrm{e}^{-\frac{|x'-x|^2}{2\delta_t}}.$$

The Wiener measure is the extension of the above μ_0 to an infinite number of segments ($\delta_t \to 0$). We write it symbolically as

$$\mu_0(\Omega) = \int_\Omega \mu_0(d\Omega),$$

$$\mu_0(d\Omega) \propto \exp\left[-\frac{1}{2}\int_0^T \left|\frac{dx}{dt}\right|^2 dt\right] d\Omega.$$

9. *As a reader, I would like to see the detailed step-by-step derivations (with adequate interpretations) of (B11)-(B14).*

Thank you for the interest. I will append explanations to the derivation.

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-44, 2017.