

Answers to the reviews of "Backward importance sampling for online estimation of state space models" (JCGS-21-171)

August 2022

We would like to thank the Referee and the Associate Editor for their comments which led us to improve the original version of our work. We address below the concerns raised by the reviewer and the associate editor.

Referee's comment *My main concern is about the real data application which is quite short and presented together with the applications to simulated data, as a subsection of Section 5. It would be interesting to have (and would substantially improve the paper) two separate sections, one concerning the simulation experiments with emphasis to the comparison of the proposed approach to existing methods, and another one containing some applications to real problems. This section would contain some real data application(s) of the proposed algorithm, to a couple of different models and data sets, say, and the comparison of the respective results to the results from existing methods, listing possible advantages of the new approach.*

Answer We agree that the real data setting was not sufficiently highlighted and that it required to be developed. In the revised version of our work the experiments are split in two sections.

- Section 5 contains all simulations with synthetic data. We first compare the performance of our approach with the standard and widespread Poor man's (or path space) smoother for RNN. Then, we provide a comparison of our method with the Acceptance-Rejection (AR) method used in the PaRIS procedure for the Sine model and show that we outperform the AR method. Finally, we illustrate that our algorithm can be used for recursive maximum likelihood using the Sine model. As stated by the Referee, this section now emphasizes the different settings in which our method can be applied and how it compares to the most common alternatives in our opinion. The results are also commented in the new conclusion section.
- Section 6 is devoted to real data applications to highlight that our method can be used in such settings. We would like to add that the hares-lynx example illustrates a complex challenge in the case of multidimensional diffusion processes where the exact approaches proposed since Beskos et al. (2008) could not be applied. It is now clearly explained that Exact Algorithms cannot be used in this setting and we therefore compare our method to another SMC algorithm. In this case, we discuss now more extensively the results provided in Fig.7: our method proposes estimators with a much smaller variance for a comparable computational cost. We now highlight that this feature is of great practical interest in abundance estimation problems.

In addition to this section, we now apply our method in a real unsupervised problem in finance, by estimating the log-return of the classical S & P 500 using a stochastic volatility model, using daily log returns from 2005 to 2018 (consisting in $n \simeq 3300$ observations). In this simple setting where state of the art competitors can be applied, the estimated parameters were similar as the one obtained with a FFBSi or a path-space smoother. We used these estimates to perform smoothing and estimate the volatility during this period. Our estimates highlight periods with high uncertainty in finance such as

the aftermath of the 2008 crisis. However, as a) this data set example is for simple comparison for a well known problem, and b) the associate editor asked that *"the revision should not increase the length of the manuscript"*, we postponed this study in the appendix for the interested reader.

Referee's comment *The paper seems to have ended rather abruptly, without either an application section nor a concluding section, which is also needed to summarize the contributions and findings of the work.*

The application to real data is now a separate section which is clearly identified. We agree that a conclusion was missing and we summarise all our contributions.

- A new SMC smoother that can be applied in the setting of Gloaguen et al (2022) and computationally outperforms their algorithm.
- A whole framework using our algorithm and exact algorithms for SDEs to perform online MLE in the context of partially observed diffusion processes. This algorithm relies on a novel unbiased estimator of the gradient of the logarithm of the intractable transition density.
- An example of estimation for a multidimensional partially observed diffusion process which is not in the classes of exact algorithms of Beskos et al (2008). To the best of our knowledge, this is the first case of SMC estimation using the exact parametrix algorithms, together with the Wald's trick for positive weights.
- An illustration of how our framework can be used for novel machine learning models like recurrent neural networks.

Miscellaneous typos We thank the Referee for the thorough review of the paper, all typos were corrected excepted for the term "parametrix" which is the dedicated name of the algorithm proposed by Anderson and Kohatsu-Higa (2017).

Editorial issues. Please find below our updates based on the editorial issued mentioned by the Associate Editor.

- The tower property is now detailed in equation Eq. (4).
- The choice of a small \tilde{N} is motivated by the seminal paper of Olsson et al. (2017) where the authors show (empirically and theoretically) that it is enough to choose $\tilde{N} \geq 2$. This is now stated just after Eq. (12).
- Colors and legends have been modified with colors from the viridis palette wich aims to be both colorblind friendly and suited for black and white printing.
- The code have been reread and runs after installing 2 packages available on our gihub. We used as much as possible the tidyverse style, and parallelization to reduce time cost.