

Revisions to “Simple conditions for convergence of sequential Monte Carlo genealogies with applications”

S. Brown, P. A. Jenkins, A. M. Johansen, J. Koskela

November 17, 2020

We thank the editor and the reviewer for the constructive comments which we believe have been fully addressed in the current revision. Responses to the reviewer’s comments are presented below, and we are also attaching a diff for your convenience.

Major comments

1. *It could be better indicated what is the applied relevance of the obtained results. My understanding is that the practical importance is mostly the time rescaling by τ_N , but it is not clear to me what practical questions this answers as this is not articulated in the paper.*

We agree that this could have been more clearly articulated. The timescale on which the trajectories of particles within these algorithms coalesce is of interest for a number of reasons: establishing the timescale over which smoothing distributions can be reasonably characterised by simple particle filters and the time intervals over which conditional SMC algorithms can provide effective “particle Gibbs” update steps, and determining storage requirements for online algorithms. Other properties of the genealogies are also interesting, for example in the context of recently developed techniques to characterize the variance of these algorithms. We have attempted to make this more explicit in the manuscript with the following changes:

p.2 l.11 We have extended the paragraph with more discussion about the practical consequences of our results.

2. *It is suggested that Theorem 1 is a true generalization of the result in Koskela et al. (2018), but this is not made explicit. If this is the case it is worth mentioning explicitly as a remark or so and including a proof or motivation of this statement. If not, the phrase ‘weaker and more tractable conditions’ (proof of Theorem 1) deserves some modification. It is not clear at first sight why this would be a generalization since many of the conditions of Koskela et al. (2018) are in fact implied by the assumption (2) of this paper.*

As you point out, we do not actually show that our result is a true generalisation of Koskela et al. (2018). We do show that one of the four conditions of the earlier result (equation (6) in the current manuscript) is unnecessary, and there are examples, one of which is given in the paper, of systems exhibiting this type of coalescent behaviour which are excluded by that assumption. For the other three conditions we do not have the converse result required to prove that our result is more general; however we note that in the neutral case condition (2) is minimal, so we expect that (2) is in general not significantly stronger than (3)–(5). On the other hand, we believe our result to be a considerable improvement upon Koskela et al. (2018) because the single, simpler condition, without mixed moments, is much easier to verify in practice. We have endeavoured to clarify these points and have made the following changes to the manuscript:

p.2 l.25 We have reworded the paragraph to make these considerations explicit.

p.5 l.-11 We have changed “weaker and more tractable” to just “more tractable”.

p.5 l.-7 We have added a paragraph contrasting condition (2) with conditions (3)–(6).

Typos and suggestions

- p.2 1.-6** *‘with respective densities’: I understand that (q_t) are the transition densities of the transition kernels (as opposed to stationary densities, which caused me some confusion)?* We have clarified that these are *transition* densities.
- p.2 1.-5** *‘sequence of potential functions’: please specify, what I understand from Algorithm 1, that $g_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ and $g_{t+1} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$.* We have added the domain/range of g_t when the states spaces are common subspaces of \mathbb{R}^d and an additional sentence explaining why we allow potential functions to depend upon the previous as well as current states.
- p.3 1.4** *Algorithm 1: What is the algorithm supposed to compute?* We have added a discussion and some examples of the possible outputs of Algorithm 1, but this depends to a significant degree on the particular application of the algorithm and so we have avoided making this explicit in the statement of the algorithm (every sampled variable may be important in some settings).
- p.4 1.-9** *‘we exclude the case where $\mathbb{P}\{\tau_N(t) = \infty\} > 0$ for finite t .’ Is it possible to give (tractable) sufficient conditions to be able to exclude this case?* It is difficult to give general, tractable conditions which are not very much stronger than necessary, which is why we have formulated things in this way. We have added more discussion of the finite time-scale condition at this point in the manuscript, noting that we do provide tractable conditions in the examples given and that in practice this is an extremely weak condition which is required only to exclude algorithms with such good properties that they can never be realised in practice.
- p.6 proof of Lemma 1** *the summation $\sum_{j \neq i}^N$ would be better to formulate as... or just briefly $\sum_{j \neq i}$* We have clarified the limits of all sums previously written in this way.

Other

- p.1 1.-4** We have added a citation to the recently-published book by Chopin & Papaspiliopoulos.