



Heine, K., & Whiteley, N. P. (2016). Fluctuations, stability and instability of a distributed particle filter with local exchange. *Stochastic Processes and their Applications*. DOI: [10.1016/j.spa.2016.11.003](https://doi.org/10.1016/j.spa.2016.11.003)

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1016/j.spa.2016.11.003](https://doi.org/10.1016/j.spa.2016.11.003)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via Elsevier at <http://www.sciencedirect.com/science/article/pii/S0304414916302137>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/pure/about/ebr-terms.html>



# Fluctuations, stability and instability of a distributed particle filter with local exchange

Kari Heine\*, Nick Whiteley

*School of Mathematics, University Walk, Bristol, BS8 1TW, United Kingdom*

Received 12 May 2015; received in revised form 19 May 2016; accepted 27 November 2016

Available online xxxx

---

## Abstract

We study a distributed particle filter proposed by Bolić et al. (2005). This algorithm involves  $m$  groups of  $M$  particles, with interaction between groups occurring through a “local exchange” mechanism. We establish a central limit theorem in the regime where  $M$  is fixed and  $m \rightarrow \infty$ . A formula we obtain for the asymptotic variance can be interpreted in terms of colliding Markov chains, enabling analytic and numerical evaluations of how the asymptotic variance behaves over time, with comparison to a benchmark algorithm consisting of  $m$  independent particle filters. We prove that subject to regularity conditions, when  $m$  is fixed both algorithms converge time-uniformly at rate  $M^{-1/2}$ . Through use of our asymptotic variance formula we give counter-examples satisfying the same regularity conditions to show that when  $M$  is fixed neither algorithm, in general, converges time-uniformly at rate  $m^{-1/2}$ .

© 2016 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

MSC: 60F05; 60F99; 60G35

**Keywords:** Hidden Markov model; Particle filter; Central limit theorem; Asymptotic variance; Local exchange; Sequential Monte Carlo

---

\* Correspondence to: Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT, United Kingdom.

*E-mail addresses:* [k.heine@ucl.ac.uk](mailto:k.heine@ucl.ac.uk) (K. Heine), [Nick.Whiteley@bristol.ac.uk](mailto:Nick.Whiteley@bristol.ac.uk) (N. Whiteley).

<http://dx.doi.org/10.1016/j.spa.2016.11.003>

0304-4149/© 2016 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Since their introduction in [11], particle filters have become very popular tools in engineering, signal processing, econometrics and various other disciplines for approximate nonlinear filtering of hidden Markov models (HMM's). Investigations of particle filters have generated book-length studies, notably [6], demonstrating the well-developed state of knowledge about convergence rates, fluctuations, propagation of chaos, large deviations and various other properties, with more recent contributions to the literature focusing on specific algorithmic mechanisms, such as adaptive resampling [5,7].

Trends in the development of computers towards distributed and parallel architectures have influenced particle filtering methodology. One of the main bottlenecks for computational efficiency when implementing particle filters is the interaction between particles which occurs in the resampling step. This step is important because it ensures that the algorithm exhibits certain time-uniform convergence properties, but is difficult to parallelise.

A significant piece of work from the engineering literature which addresses this difficulty is [2], introducing an algorithm we refer to as the Local Exchange Particle Filter (LEPF), in which groups of particles are spread across computational units. What makes this algorithm unusual is that the  $m$  groups of  $M$  weighted particles interact through an “exchange” mechanism, which places it outside the frameworks of many existing studies, notably [6,5,7]. The practical rationale for the LEPF is to achieve a compromise between communication efficiency of the algorithm and the benefits brought about by resampling. In particular the interaction between particles in the LEPF occurs in a localised manner, making it suited to implementation on a network of computing devices without the need for global connections.

Despite substantial interest in [2] from practitioners – it has 250 citations according to Google scholar at the time of writing – relatively little is known about convergence properties of LEPF. Indeed the question of whether it truly exhibits the same time-uniform convergence properties as the original particle filter of [11] has not been fully answered. The few papers on analysis of the LEPF appear to be [16,17] and the recent technical report [18]. [16] concerns analysis over a single time-step, and [17,18] provide proofs of time-uniform convergence of the particle filtering approximation error, in  $L_1$  and  $L_p$  norms respectively, in the regime where  $M$  is fixed and  $m \rightarrow \infty$ , for an algorithm of which the LEPF as we present it is a special case. However, the proofs of [17,18] rely on key hypotheses on the particle weights which they do not rigorously verify, and which seem difficult to check in general. The results of [17,18] also do not establish a particular rate of convergence.

The structure of this paper and outline of our main contributions are as follows (precise statements are given later). In Section 2 we introduce the setup of the filtering problem, present the LEPF and describe the main result of [17,18]. We also introduce a standard algorithm consisting of  $m$  independent bootstrap particle filters (IBPF), each with  $M$  particles. The independence in the IBPF makes it very easy to parallelise, so from a computational point of view it is a natural alternative to the LEPF. In this paper the convergence properties of the IBPF, which are already well-understood, serve as benchmarks against which to compare the LEPF.

Section 3 introduces a general algorithm of which the LEPF and IBPF are special cases, and gives our main result, [Theorem 1](#), a central limit theorem (CLT) for the error in particle approximation of prediction filter distributions, in the regime where  $M$  is fixed and  $m \rightarrow \infty$ . We address time-uniform convergence in Section 4. Our first result here is a positive one: that under strong but standard regularity conditions, in  $L_p$  norm the error from the LEPF converges time-uniformly with rate  $M^{-1/2}$ , in the regime where  $m$  is fixed and  $M \rightarrow \infty$ . The same is true

of the IBPF. Our second result, [Proposition 5](#) in [Section 4.2](#), shows that growth without bound of the asymptotic variance in our CLT is sufficient to rule out time-uniform convergence at rate  $m^{-1/2}$  in the regime where  $M$  is fixed and  $m \rightarrow \infty$ .

[Section 5](#) investigates various properties of the asymptotic variance for the LEPF and compares them to those of the IBPF. In particular, we show by examples in [Sections 5.1](#) and [5.2](#) that under conditions which can be considered very favourable for performance, the asymptotic variance for the LEPF and IBPF can grow over time without bound. This can be considered a negative result for the LEPF, since the sequence of asymptotic variances (over time) for the original particle filter of [\[11\]](#) has been shown under weaker conditions to be bounded, or tight when the observations in the HMM are treated as random [\[4,15,9,22,8\]](#). Moreover, combined with [Proposition 5](#) in [Section 4.2](#), these examples serve as counter-examples to time-uniform convergence at rate  $m^{-1/2}$ . This does not contradict the time-uniform convergence results of [\[17,18\]](#), since the latter results do not pertain to a specific convergence rate, and they concern the updated filtering distributions. However, our [Proposition 5](#) allows us to confirm that a hypothesis slightly stronger than that of [\[18\]](#) does not hold in general, even under favourable conditions. [Section 6](#) contains further discussion and interpretation of our results. Our analysis allows us to explain qualitatively why the asymptotic variance for the LEPF may be lower or grow over time more slowly than that for the IBPF, and we illustrate this phenomenon with numerical results.

Some clarifications about originality are in order. To the knowledge of the authors, our CLT is the first result of its kind for the LEPF. Our starting point to prove this result consists of a martingale decomposition and error bounds, [Proposition 1](#) in [Section 3](#), which is an application of a result obtained by the authors in [\[23\]](#) for a class of algorithms which includes the LEPF. However, we emphasise that [Proposition 1](#) is only one of the first steps towards the CLT itself, leaving us with substantial work to do. In our study of time-uniform convergence, we also appeal to a result of [\[23\]](#) ([Proposition 4](#) in the present paper), but again we have some work to do in dealing with the specifics of the LEPF. We also point out that despite some superficial similarities, the details of LEPF and our analysis differ substantially from those of some resampling algorithms studied recently by the authors in [\[13\]](#).

### Notation

For any measurable space  $(\mathbb{X}, \mathcal{X})$  we use  $\mathcal{M}(\mathbb{X})$ ,  $\mathcal{P}(\mathbb{X})$  and  $\mathcal{B}(\mathbb{X})$  to denote the set of measures, probability measures and the set of bounded and measurable functions defined on  $\mathbb{X}$ , respectively.  $\mathbb{N}$  includes 0. For any  $\mathbb{N}$ -valued  $m \geq 1$  we write  $[m] := \{1, \dots, m\}$ . Whenever summation over a single variable appears without the summation set made explicit, the sum is taken over the set  $[N]$ , i.e.  $\sum_i \equiv \sum_{i=1}^N$  and for summations over multiple variables we write  $\sum_{(i_1, \dots, i_p)} \equiv \sum_{i_1} \cdots \sum_{i_p}$ . We use  $Id$  to denote the identity mapping for any domain of definition and  $\mathbf{1}$  to denote a constant function equal to 1 everywhere. For any function  $\varphi : A \rightarrow \mathbb{R}$ , we define  $\varphi^{\otimes 2}(x, y) := \varphi(x)\varphi(y)$  for all  $x, y \in A$ . For  $\varphi \in \mathcal{B}(\mathbb{X})$  we define  $\|\varphi\|_\infty := \sup_x |\varphi(x)|$  and  $\text{osc}(\varphi) := \sup_{x,y} |\varphi(x) - \varphi(y)|$ . For any  $\mu, \nu \in \mathcal{M}(\mathbb{X})$ ,  $\mu \otimes \nu$  denotes the product measure and  $\mu^{\otimes 2} := \mu \otimes \mu$ . We use  $\delta_x$  to denote the point mass located at  $x$ . We define  $\lfloor x \rfloor := \max(z \in \mathbb{Z} : z \leq x)$  and  $(y \bmod x) := y - \lfloor (y-1)/x \rfloor x$ . All random variables we encounter are considered to be defined on some underlying probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with expectation w.r.t.  $\mathbb{P}$  denoted by  $\mathbb{E}$ . Convergence in probability under  $\mathbb{P}$  is denoted by  $\xrightarrow{\mathbb{P}}$ .



**Algorithm 1** Local exchange particle filter

---

**for**  $i = 1, \dots, Mm$   
 Set  $W_0^i = 1$  and sample  $\zeta_0^i \sim \pi_0$   
 Set  $L^i = (i + \theta) \bmod Mm$   
**for**  $k = 1, \dots, m$   
 Set  $G_k = \{(k - 1)M + 1, \dots, (k - 1)M + M\}$   
**for**  $n = 1, 2, \dots$   
**for**  $k = 1, \dots, m$   
**for**  $i \in G_k$   
 Set  $W_n^i = (Mm)^{-1} \sum_{j \in G_k} W_{n-1}^{L^j} g_{n-1}(\zeta_{n-1}^{L^j})$   
 Sample  $\zeta_n^i \mid \zeta_0, \dots, \zeta_{n-1} \sim \frac{\sum_{j \in G_k} W_{n-1}^{L^j} g_{n-1}(\zeta_{n-1}^{L^j}) f(\zeta_{n-1}^{L^j}, \cdot)}{\sum_{j \in G_k} W_{n-1}^{L^j} g_{n-1}(\zeta_{n-1}^{L^j})}$

---

computations for the  $m$  groups are performed concurrently by a network of  $m$  computers, so the regime  $M$  fixed,  $m \rightarrow \infty$  can be thought of as corresponding to an increasingly large network, in which each computer handles  $M$  particles, see [2] for details.

[17,18] studied an algorithm of which the LEPF as we present it in Algorithm 1 is a special case. Our mapping  $i \mapsto L^i$  is a particular instance of the mapping denoted by  $\beta$  in [17,18] and if one sets their exchange period parameter  $n_0 = 1$ , one recovers Algorithm 1. The generality of  $\beta$  in [17,18] allows for other patterns of interaction between particles, beyond the ones considered in the present article. Whilst we focus on the prediction filter distributions  $\pi_n$ , [17,18] focus on particle approximations of the updated filtering distributions  $\hat{\pi}_n(A) := \pi_n(g_n \mathbb{I}_A) / \pi_n(g_n)$ ,  $A \in \mathcal{X}, n \geq 0$ . To allow us to state their result, for each  $n \geq 0$  let  $\{\hat{\zeta}_n^i; i \in [N]\}$  be random variables which are conditionally independent given  $\zeta_0, \dots, \zeta_n$ , with

$$\mathbb{P}(\hat{\zeta}_n^i \in \cdot \mid \zeta_0, \dots, \zeta_n) = \frac{\sum_{j \in G_k} W_n^{L^j} g_n(\zeta_n^{L^j}) \delta_{\zeta_n^{L^j}}(\cdot)}{\sum_{j \in G_k} W_n^{L^j} g_n(\zeta_n^{L^j})}, \quad \forall i \in G_k,$$

and

$$\hat{\pi}_n^N := \frac{\sum_i W_{n+1}^i \delta_{\hat{\zeta}_n^i}}{\sum_i W_{n+1}^i}.$$

The  $\{\hat{\zeta}_n^i, i \in [N]\}$  can be understood as “integrated out” in Algorithm 1.

In the notation of the present paper and with  $\tilde{W}_n^k := W_n^i = W_n^j$  for all  $i, j \in G_k$  and  $k \in [m]$ , the key hypothesis of [18, Assumption 3] can be equivalently written as follows: there exist  $\epsilon \in [0, 1)$  and  $q \geq 4$  such that,

$$\sup_{m \geq 1} \sup_{n \geq 0} m^{q-\epsilon} \mathbb{E} \left[ \left| \max_{k \in [m]} \frac{\tilde{W}_n^k}{\sum_{j \in [m]} \tilde{W}_n^j} \right|^q \right] < \infty. \tag{3}$$



special case of the so-called  $\alpha$ SMC algorithm, introduced by the authors in [23]. In turn this later allows us to leverage some results of [23] – in particular Proposition 1 – providing some building blocks for our CLT. The IBPF is also an instance of Algorithm 3 and this fact eases our presentation of comparisons between it and the LEPF in Section 5.

---

**Algorithm 3**

---

**for**  $i = 1, \dots, N$   
 Set  $W_0^i = 1$  and sample  $\zeta_0^i \sim \pi_0$   
**for**  $n = 1, 2, \dots$   
**for**  $i = 1, \dots, N$   
 Set  $W_n^i = \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$   
 Sample  $\zeta_n^i | \zeta_0, \dots, \zeta_{n-1} \sim (W_n^i)^{-1} \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)$

---

From henceforth, the integer  $M \geq 1$  is, unless stated otherwise, assumed to be fixed. In Algorithm 3,  $\alpha$  is a row-stochastic matrix, of size  $N \times N$ , with  $N = Mm$ . Assumption 2 introduces hypotheses on the matrix  $\alpha$  for each value  $N \in \{Mm : m \geq 1\}$ . To state these hypotheses precisely, we need to be clear about dependence of  $\alpha$  on  $N$  and hence write  $\alpha_N$  up until the end of Section 3.1, beyond which we revert to  $\alpha$  to reduce notational clutter.

**Assumption 2.** For all  $N \in \{Mm : m \geq 1\}$ ,

- (2.1)  $\alpha_N$  is doubly stochastic,
- (2.2) for all  $i, j \in [N]$  and  $z \in \mathbb{Z}$ ,

$$\alpha_N^{ij} = \alpha_N^{(i+zM) \bmod N, (j+zM) \bmod N}.$$

Additionally, for some integer  $\beta \geq 1$ ,

- (2.3)  $\alpha_N^{ij} = 0$  for  $N \geq 2\beta + 1$  and  $i, j \in [N]$  such that

$$\Delta(i, j) := \min_{\ell \in \mathbb{Z}} |i - j + \ell N| > \beta,$$

- (2.4) there exists  $\{\alpha_\infty^{ij} : i, j \in \mathbb{Z}\}$  such that for  $N \geq 2\beta + 1$ ,

$$\alpha_\infty^{ij} = \alpha_N^{i \bmod N, j \bmod N} \mathbb{I}[|i - j| \leq \beta], \quad i, j \in \mathbb{Z}. \tag{4}$$

Assumption (2.1) allows us to apply results from [23] to Algorithm 3. Assumption (2.2) asserts that the elements on each diagonal of  $\alpha_N$  are periodic with cycle length  $M$ . Intuitively, this captures the idea that the  $N$  particles in Algorithm 3 are in some sense organised into groups of size  $M$ . It is easily verified that the function  $\Delta$  appearing in Assumption (2.3) is a metric on  $[N]$ , in particular it is the graph distance on a cycle graph with vertex set  $[N]$  where there is an edge between each  $i \in [N]$  and  $(i + 1) \bmod N$ . Assumption (2.3) then asserts that  $\alpha_N$  is a band matrix in the sense that elements further than  $\beta$  away from the main diagonal in metric  $\Delta$  are equal to zero, in turn influencing the conditional independence structure of the particles in Algorithm 3. Finally Assumption (2.4) can be interpreted as meaning that there is some common structure to the matrices  $\alpha_N$  as  $N$  grows, and loosely speaking, this common structure is captured in the “limiting” doubly infinite matrix  $\alpha_\infty$ , which will show up later in our CLT.

---

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>



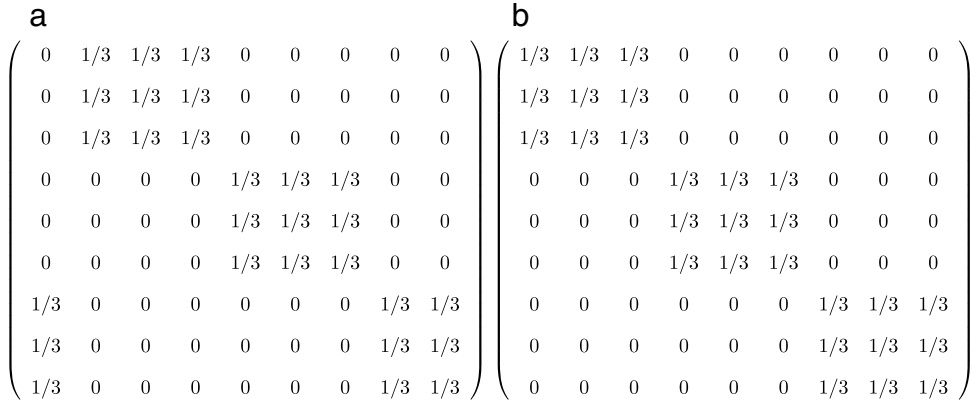


Fig. 1. Matrices in (a) and (b) correspond to the LEPF and IBPF, respectively.

Let us now state how the LEPF and IBPF fit in this framework. Consider

$$\begin{aligned} \alpha_N^{ij} &= M^{-1} \mathbb{I}[\lfloor (i-1)/M \rfloor = \lfloor ((j-\theta) \bmod N - 1)/M \rfloor] \quad \forall i, j \in [N] \\ \alpha_\infty^{ij} &= M^{-1} \mathbb{I}[\lfloor (i-1)/M \rfloor = \lfloor (j-\theta-1)/M \rfloor] \quad \forall i, j \in \mathbb{Z}. \end{aligned} \tag{5}$$

It is a matter of elementary but tedious manipulations to show that with  $\alpha = \alpha_N$  as in (5), Algorithm 3 reduces to the LEPF as in Algorithm 1, and to check that Assumptions (2.1)–(2.3) hold with  $\beta = M - 1 + \theta$ . Checking Assumption (2.4) involves some less trivial work and a proof is provided in the Appendix.

To recover the IBPF from Algorithm 3, we take

$$\begin{aligned} \alpha_N^{ij} &= M^{-1} \mathbb{I}[\lfloor (i-1)/M \rfloor = \lfloor (j-1)/M \rfloor] \quad \forall i, j \in [N] \\ \alpha_\infty^{ij} &= M^{-1} \mathbb{I}[\lfloor (i-1)/M \rfloor = \lfloor (j-1)/M \rfloor] \quad \forall i, j \in \mathbb{Z}. \end{aligned} \tag{6}$$

With  $\beta = M - 1$ , checking Assumptions (2.1)–(2.3) is again elementary, and in this case Assumption (2.4) is obviously satisfied.

Fig. 1(a) and (b) show the matrices defined in (5)–(6) in the case  $N = 9, M = 3$  and  $\theta = 1$ . It follows from Assumptions (2.3) and (2.4) that  $\alpha_\infty$  is, like each  $\alpha_N$ , a row-stochastic matrix, which can be thought of as specifying the transition probabilities of a  $\mathbb{Z}$ -valued Markov chain. It turns out that the asymptotic variance in our CLT is expressed in terms of two copies of this chain. To this end, denote by  $\mathbb{E}_{u,v}$ , where  $u, v \in \mathbb{Z}$ , the expectation w.r.t. the law of the bi-variate backward Markov chain  $(I_k, J_k)_{0 \leq k \leq n}$ , where

$$\begin{aligned} (I_n, J_n) &\sim \delta_u \otimes \delta_v, \\ \mathbb{P}(I_k = i_k, J_k = j_k \mid I_{k+1} = i_{k+1}, J_{k+1} = j_{k+1}) &= \alpha_\infty^{i_{k+1}i_k} \alpha_\infty^{j_{k+1}j_k}. \end{aligned} \tag{7}$$

Fig. 2 illustrates some segments of paths for  $I$  (or  $J$ ) which have strictly positive probability under the transitions  $\alpha_\infty$  for the LEPF and IBPF, with  $M = 3$  and  $\theta = 1$ .

Before stating our main result we introduce some more notation. For all  $n \geq 1$ , define non-negative kernels  $Q_n : \mathbb{X} \times \mathcal{X} \rightarrow \mathbb{R}_+$  as

$$Q_n(x, A) := g_{n-1}(x) f(x, A), \quad \forall x \in \mathbb{X}, A \in \mathcal{X}, \tag{8}$$

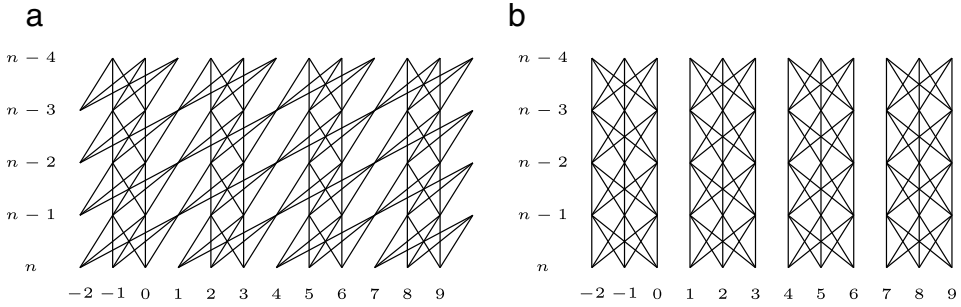


Fig. 2. Some of the paths assigned positive probability by  $\alpha_\infty$  for the (a) LEPF and (b) IBPF. In both cases  $M = 3$  and in (a)  $\theta = 1$ .

and the corresponding operators on functions and measures

$$Q_n(\varphi)(x) := \int Q_n(x, dx')\varphi(x'), \quad \forall x \in \mathbb{X}, \varphi \in \mathcal{B}(\mathbb{X}),$$

$$\mu Q_n(A) := \int Q_n(x, A)\mu(dx), \quad \forall \mu \in \mathcal{P}(\mathbb{X}), A \in \mathcal{X},$$

respectively. Moreover we define for  $n \geq 1$  and  $0 \leq p < n$

$$\frac{Q_{p,p}}{Q_{p,p}} := Id, \quad \frac{Q_{p,n}}{Q_{p,n}} := Q_{p+1} \cdots Q_n,$$

$$\frac{Q_{p,p}}{Q_{p,p}} := Id, \quad \frac{Q_{p,n}}{Q_{p,n}} := Q_{p+1} \cdots Q_n,$$

where  $\bar{Q}_n := Q_n/\pi_{n-1}(g_{n-1})$  for all  $n \geq 1$ . Also let

$$\gamma_n := \pi_0 Q_{0,n}, \quad \forall n \geq 0.$$

Define the tensor-product kernel  $Q_n^{\otimes 2}(x, y, d(x', y')) := Q_n(x, dx')Q_n(y, dy')$ , with the corresponding operators on functions and measures written similarly to those for  $Q_n$ , and finally define operators  $C_0$  and  $C_1$ , such that for any  $\varphi \in \mathcal{B}(\mathbb{X}^2)$ ,

$$C_0(\varphi)(x, y) = \varphi(x, y) \quad \text{and} \quad C_1(\varphi)(x, y) = \varphi(x, x), \quad \forall x, y \in \mathbb{X}.$$

We then have:

**Theorem 1.** Fix  $M > 1$  and  $\beta > 0$  and suppose that Assumption 2 holds. Then for any  $\varphi \in \mathcal{B}(\mathbb{X})$ , Algorithm 3 has the property

$$\sqrt{N}(\pi_n^N(\varphi) - \pi_n(\varphi)) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \sigma_n^2), \quad \forall n \in \mathbb{N},$$

where  $N$  goes to infinity along the sequence  $\{Mm : m = 1, 2, \dots\}$ , the following variances are assumed strictly positive,

$$\sigma_0^2 = \pi_0((\varphi - \pi_0(\varphi))^2)$$

$$\sigma_n^2 = \frac{1}{\gamma_n(\mathbf{1})^2 M} \sum_{\substack{0 \leq u < M \\ |v| \leq 2n\beta}} \mathbb{E}_{u, u+v} \left[ \pi_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \cdots Q_n^{\otimes 2} C_{\epsilon_n} (\bar{\varphi}^{\otimes 2}) \right], \quad n \geq 1,$$

with  $\epsilon_k = \mathbb{I}[I_k = J_k]$ , for all  $0 \leq k \leq n$ , and  $\bar{\varphi} := \varphi - \pi_n(\varphi)$ .

**Remark 1.** Since the LEPF and IBPF are special cases of Algorithm 3, Theorem 1 applies to them immediately. Note that the only distinction between the asymptotic variances for LEPF and

the IBPF arises from  $\alpha_\infty$ , as given for these two algorithms in (5) and (6). In Section 5 we shall examine  $\sigma_n^2$  for the LEPF and the IBPF in detail, which involves study of the  $I, J$  processes for these two algorithms.

3.2. Martingale array and the proof of the main result

Defining the random measures

$$\Gamma_n^N := \frac{1}{N} \sum_{i=1}^N W_n^i \delta_{\zeta_n^i}, \quad \bar{\Gamma}_n^N := \frac{\Gamma_n^N}{\gamma_n(\mathbf{1})}, \quad \forall n \in \mathbb{N}, \tag{9}$$

allows us to decompose the particle approximation error as

$$\sqrt{N}(\pi_n^N(\varphi) - \pi_n(\varphi)) = \sqrt{N} \bar{\Gamma}_n^N(\bar{\varphi}) - \pi_n^N(\bar{\varphi}) \frac{\sqrt{N}}{\gamma_n(\mathbf{1})} (\Gamma_n^N(\mathbf{1}) - \gamma_n(\mathbf{1})), \tag{10}$$

where  $\bar{\varphi} := \varphi - \pi_n(\varphi)$ .

Our overall strategy in proving Theorem 1 is to establish asymptotic normality of  $\sqrt{N} \bar{\Gamma}_n^N(\bar{\varphi})$  as  $N \rightarrow \infty$  using the CLT for martingale arrays [12], and to apply results from [23] to show that the second term on the r.h.s. of (10) converges to zero in probability. Our first step is to identify a martingale representation for  $\sqrt{N} \bar{\Gamma}_n^N(\bar{\varphi})$ , for which the setup is as follows.

Fix  $n \in \mathbb{N}$  and  $M \geq 1$ . For given  $m \geq 1$  and  $\varphi \in \mathcal{B}(\mathbb{X})$  define, for  $\varrho \in [Mm]$ ,

$$\xi_\varrho^m := \frac{1}{\sqrt{Mm}} \left( \bar{Q}_{0,n}(\varphi)(\zeta_\varrho^i) - \pi_0 \bar{Q}_{0,n}(\varphi) \right), \tag{11}$$

and for  $\varrho \in [(n+1)Mm] \setminus [Mm]$ ,

$$\xi_\varrho^m := \frac{W_p^i}{\sqrt{Mm} \gamma_p(\mathbf{1})} \left( \bar{Q}_{p,n}(\bar{\varphi})(\zeta_p^i) - \frac{\sum_j \alpha^{ij} W_{p-1}^j \bar{Q}_{p-1,n}(\bar{\varphi})(\zeta_{p-1}^j)}{\sum_j \alpha^{ij} W_{p-1}^j \bar{Q}_p(\mathbf{1})(\zeta_{p-1}^j)} \right), \tag{12}$$

where  $p = p_m(\varrho)$ ,  $i = i_m(\varrho)$  and

$$p_m(\varrho) := \left\lfloor \frac{\varrho - 1}{Mm} \right\rfloor \quad \text{and} \quad i_m(\varrho) := \overline{\varrho \bmod Mm}.$$

Writing out the expression for  $W_p^i$ ,  $p \geq 1$ , in Algorithm 3, using the fact that  $\alpha$  is row-stochastic and Assumption 1,

$$W_p^i = \sum_{(i_0, \dots, i_{p-1})} \prod_{q=0}^{p-1} \alpha^{i_{q+1} i_q} g_q(\zeta_q^{i_q}) \leq \prod_{q=0}^{p-1} \|g_q\|_\infty < \infty.$$

Combining this with (11), (12) and again using Assumption 1, we have

$$\sup_{\varrho \in [(n+1)Mm]} \left| \xi_\varrho^m \right| \leq \frac{1}{\sqrt{Mm}} \max_{p \in \{0, \dots, n\}} \frac{\prod_{q=0}^{p-1} \|g_q\|_\infty}{\gamma_p(\mathbf{1})} \text{osc}(\bar{Q}_{p,n}(\bar{\varphi})) < \infty, \tag{13}$$

with the convention  $\prod_{q=0}^{-1} \|g_q\|_\infty = 1$ .

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>



$Q \in [(n + 1)Mm]$ . If

$$\sum_{Q=1}^{(n+1)Mm} \mathbb{E} \left[ (\xi_Q^m)^2 \mathbb{I}[|\xi_Q^m| > \varepsilon] \mid \mathcal{F}_{Q-1}^m \right] \xrightarrow{m \rightarrow \infty} 0, \quad \forall \varepsilon > 0, \tag{19}$$

$$\sum_{Q=1}^{(n+1)Mm} \mathbb{E} \left[ (\xi_Q^m)^2 \mid \mathcal{F}_{Q-1}^m \right] \xrightarrow{m \rightarrow \infty} \sigma^2, \quad \sigma^2 > 0, \tag{20}$$

then

$$\sum_{Q=1}^{(n+1)Mm} \xi_Q^m \xrightarrow{m \rightarrow \infty} \mathcal{N}(0, \sigma^2). \tag{21}$$

We now present the main arguments in the proof of [Theorem 1](#).

**Proof of Theorem 1.** We first note that the case  $n = 0$  is trivial, since in [Algorithm 3](#),  $\{\xi_0^i : i \in [N]\}$  are i.i.d. samples from  $\pi_0$ . So it remains to consider  $n \geq 1$ . With the definitions [\(11\)](#), [\(12\)](#), and [\(14\)](#), [Proposition 1](#) establishes that

$$\left\{ \left( \sum_{s=1}^Q \xi_s^m, \mathcal{F}_Q^m \right) : Q \in [(n + 1)Mm] \right\}$$

constitutes the martingale array as in the statement of [Theorem 2](#), and our next task is to check conditions [\(19\)](#) and [\(20\)](#).

Condition [\(19\)](#) is easily seen to be satisfied due to [\(15\)](#). The majority of our work then goes into checking [\(20\)](#). Since, for given  $m \geq 1$ ,  $\{(\xi_Q^m, \mathcal{F}_Q^m) : Q \in [(n + 1)Mm]\}$  is a martingale difference sequence, we have

$$\begin{aligned} \sum_{Q=1}^{(n+1)Mm} \mathbb{E} \left[ (\xi_Q^m)^2 \mid \mathcal{F}_{Q-1}^m \right] &= \mathbb{E} \left[ \left( \sum_{Q=1}^{(n+1)Mm} \xi_Q^m \right)^2 \right] \\ &\quad + \sum_{Q=1}^{(n+1)Mm} \mathbb{E} \left[ (\xi_Q^m)^2 \mid \mathcal{F}_{Q-1}^m \right] - \mathbb{E} \left[ (\xi_Q^m)^2 \right]. \end{aligned}$$

[Proposition 2](#) in [Section 3.3](#) establishes convergence to zero of the residual, in the sense that

$$\sum_{Q=1}^{(n+1)Mm} \mathbb{E} \left[ (\xi_Q^m)^2 \mid \mathcal{F}_{Q-1}^m \right] - \mathbb{E} \left[ (\xi_Q^m)^2 \right] \xrightarrow{m \rightarrow \infty} 0.$$

[Proposition 3](#) in [Section 3.4](#) establishes convergence of the variance, in the sense that

$$\mathbb{E} \left[ \left( \sum_{Q=1}^{(n+1)Mm} \xi_Q^m \right)^2 \right] \xrightarrow{m \rightarrow \infty} \frac{1}{M\gamma_n(\mathbf{1})^2} \sum_{\substack{0 \leq u < M \\ |v| \leq 2n\beta}} \mathbb{E}_{u,v+u} \left[ \pi_0^{\otimes 2} \mathcal{C}_{\epsilon_0} \mathcal{Q}_1^{\otimes 2} \mathcal{C}_{\epsilon_1} \cdots \mathcal{Q}_n^{\otimes 2} \mathcal{C}_{\epsilon_n} (\bar{\varphi}^{\otimes 2}) \right],$$

where  $\epsilon_k = \mathbb{I}[I_k = J_k]$ . Thus condition [\(20\)](#) is satisfied and so by [\(16\)](#) in [Proposition 1](#),

$$\sqrt{Mm} \Gamma_n^{Mm}(\bar{\varphi}) \xrightarrow{m \rightarrow \infty} \mathcal{N}(0, \sigma_n^2). \tag{22}$$

By (2.1) we can use (17) and (18) of Proposition 1 and Hölder’s inequality to obtain

$$\begin{aligned} & \lim_{m \rightarrow \infty} \mathbb{E} \left[ \left| \pi_n^{Mm}(\bar{\varphi}) \sqrt{Mm} \left( \Gamma_n^{Mm}(\mathbf{1}) - \gamma_n(\mathbf{1}) \right) \right| \right] \\ & \leq \lim_{m \rightarrow \infty} \mathbb{E} \left[ \left| \pi_n^{Mm}(\bar{\varphi}) \right|^2 \right]^{\frac{1}{2}} \sup_{m \geq 1} \sqrt{Mm} \mathbb{E} \left[ \left| \Gamma_n^{Mm}(\mathbf{1}) - \gamma_n(\mathbf{1}) \right|^2 \right]^{\frac{1}{2}} = 0, \end{aligned}$$

implying

$$\pi_n^{Mm}(\bar{\varphi}) \sqrt{Mm} \left( \Gamma_n^{Mm}(\mathbf{1}) - \gamma_n(\mathbf{1}) \right) \xrightarrow[m \rightarrow \infty]{\mathbb{P}} 0. \tag{23}$$

The claim follows by Slutsky’s theorem from (10), (22) and (23). □

### 3.3. Convergence of the residual to zero

**Proposition 2.** *Under the assumptions of Theorem 1,*

$$\sum_{\varrho=1}^{(n+1)Mm} \mathbb{E} \left[ \left( \xi_{\varrho}^m \right)^2 \mid \mathcal{F}_{\varrho-1}^m \right] - \mathbb{E} \left[ \left( \xi_{\varrho}^m \right)^2 \right] \xrightarrow[m \rightarrow \infty]{\mathbb{P}} 0.$$

**Proof.** Define:

$$Z_{\varrho}^m := \mathbb{E} \left[ \left( \xi_{\varrho}^m \right)^2 \mid \mathcal{F}_{\varrho-1}^m \right] - \mathbb{E} \left[ \left( \xi_{\varrho}^m \right)^2 \right].$$

By Markov’s inequality we have for all  $\varepsilon > 0$  that

$$\mathbb{P} \left( \left| \sum_{\varrho=1}^{(n+1)Mm} Z_{\varrho}^m \right| \geq \varepsilon \right) \leq \frac{1}{\varepsilon^2} \sum_{\varrho=1}^{(n+1)Mm} \mathbb{E} \left[ \left( Z_{\varrho}^m \right)^2 \right] + \frac{1}{\varepsilon^2} \sum_{\varrho=1}^{(n+1)Mm} \sum_{\varrho' \neq \varrho} \mathbb{E} \left[ Z_{\varrho}^m Z_{\varrho'}^m \right]. \tag{24}$$

By (15),  $\left( Z_{\varrho}^m \right)^2 \leq 4C_n^4 / (Mm)^2$  and hence

$$\sum_{\varrho=1}^{(n+1)Mm} \mathbb{E} \left[ \left( Z_{\varrho}^m \right)^2 \right] \leq \frac{4(n+1)C_n^4}{Mm} \xrightarrow[m \rightarrow \infty]{} 0.$$

To establish convergence to zero of the second summation on the r.h.s. of (24), we shall show that suitably many pairs  $Z_{\varrho}^m, Z_{\varrho'}^m$  are independent, therefore making no contribution to the sum since  $\mathbb{E}[Z_{\varrho}^m] = 0$ , and use (15) to bound the remaining pairs. Introduce the notation, for  $i \in [Mm]$ ,

$$\text{pa}(\zeta_0^i) := \emptyset, \quad \text{pa}(\zeta_n^i) := \{ \zeta_q^j : 0 \leq q < n, j \in [Mm], (\alpha^{n-q})^{ij} > 0 \}, \quad n \geq 1, \tag{25}$$

and by convention let  $\sigma(\emptyset)$  be the trivial  $\sigma$ -algebra. Our strategy to obtain a lower bound for the number of independent pairs  $Z_{\varrho}^m, Z_{\varrho'}^m$  is as follows:

Lemma 1 shows that  $Z_{\varrho}^m$  is measurable w.r.t.  $\sigma(\text{pa}(\zeta_{p_m(\varrho)}^{i_m(\varrho)}))$ , and consequently

$$\sigma(\text{pa}(\zeta_{p_m(\varrho)}^{i_m(\varrho)})) \perp \sigma(\text{pa}(\zeta_{p_m(\varrho')}^{i_m(\varrho')})) \implies Z_{\varrho}^m \perp Z_{\varrho'}^m.$$

Lemma 2 shows that for any  $0 \leq p, q \leq n$  and  $i, j \in [Mm]$ ,

$$\text{pa}(\zeta_p^i) \cap \text{pa}(\zeta_q^j) = \emptyset \implies \sigma(\text{pa}(\zeta_p^i)) \perp \sigma(\text{pa}(\zeta_q^j)).$$

Lemma 3 shows that the number of pairs  $\varrho \neq \varrho'$  such that  $\text{pa}(\zeta_{p_m(\varrho)}^{i_m(\varrho)}) \cap \text{pa}(\zeta_{p_m(\varrho')}^{i_m(\varrho')}) = \emptyset$  is at least  $Mm(n + 1)^2(Mm - 4n\beta - 1)$ .

The total number of pairs  $(\varrho, \varrho')$  where  $\varrho \neq \varrho'$  is  $(n + 1)Mm((n + 1)Mm - 1)$  and hence by (15)

$$\sum_{\varrho=1}^{(n+1)Mm} \sum_{\varrho' \neq \varrho} \mathbb{E} \left[ Z_{\varrho}^m Z_{\varrho'}^m \right] \leq \frac{4C_n^4}{Mm} ((n + 1)((n + 1)Mm - 1) - (n + 1)^2(Mm - 4n\beta - 1))$$

which is easily seen to converge to 0 as  $m \rightarrow \infty$ , completing the proof of the proposition.  $\square$

Before presenting Lemmata 1–3 we point out the following useful consequence of (25). Note that  $\zeta_q^{i_q} \in \text{pa}(\zeta_p^{i_p})$  if and only if there exists a sequence  $(i_q, \dots, i_p)$ , such that  $\prod_{k=p-1}^q \alpha^{i_{k+1}i_k} > 0$ . Using this equivalence it follows that if  $\ell < q < p$  and  $\zeta_{\ell}^{i_{\ell}} \in \text{pa}(\zeta_q^{i_q})$  and  $\zeta_q^{i_q} \in \text{pa}(\zeta_p^{i_p})$ , then also  $\zeta_{\ell}^{i_{\ell}} \in \text{pa}(\zeta_p^{i_p})$ , and thus we have the implication

$$\zeta \in \text{pa}(\zeta_p^{i_p}) \implies \text{pa}(\zeta) \subset \text{pa}(\zeta_p^{i_p}). \tag{26}$$

**Lemma 1.** For any  $\varrho \in [(n + 1)Mm]$ ,  $Z_{\varrho}^m$  is measurable w.r.t.  $\sigma(\text{pa}(\zeta_{p_m(\varrho)}^{i_m(\varrho)}))$ .

**Proof.** The variables  $\{\zeta_0^i\}_{i \in [Mm]}$  are independent, so for  $\varrho \in [Mm]$ ,  $Z_{\varrho}^m = 0$ ,  $\mathbb{P}$ -a.s. and  $\sigma(\text{pa}(\zeta_{p_m(\varrho)}^{i_m(\varrho)})) = \{0, \Omega\}$ , hence the claimed measurability holds. For  $Mm < \varrho \leq (n + 1)Mm$  we need to show that

$$\mathbb{E} \left[ (\xi_{\varrho}^m)^2 \mid \mathcal{F}_{\varrho-1}^m \right] = \mathbb{E} \left[ (\xi_{\varrho}^m)^2 \mid \text{pa}(\zeta_{p_m(\varrho)}^{i_m(\varrho)}) \right], \quad \mathbb{P}\text{-a.s.} \tag{27}$$

According to Algorithm 3,

$$\mathbb{P}(\zeta_p^{i_p} \in A \mid \zeta_0, \dots, \zeta_{p-1}) = \frac{\sum_j \alpha^{i_p j} W_{p-1}^j g_{p-1}(\zeta_{p-1}^j) f(\zeta_{p-1}, A)}{W_p^{i_p}}, \quad \mathbb{P}\text{-a.s.} \tag{28}$$

Writing out the expression for  $W_p^{i_p}$  from Algorithm 3 gives

$$W_p^{i_p} = \sum_{(i_0, \dots, i_{p-1})} \prod_{q=0}^{p-1} g_q(\zeta_q^{i_q}) \alpha^{i_{q+1}i_q}, \tag{29}$$

which clearly is measurable w.r.t.  $\sigma(\text{pa}(\zeta_p^{i_p}))$ . Noting additionally (26), we find the r.h.s. of (28) also measurable w.r.t.  $\sigma(\text{pa}(\zeta_p^{i_p}))$ . The latter observation combined with the fact that in Algorithm 3 the variables  $\{\zeta_p^i\}_{i \in [Mm]}$  are conditionally independent given  $\zeta_0, \dots, \zeta_{p-1}$ , shows that

$$\mathbb{P}(\zeta_p^i \in A \mid \zeta_0, \dots, \zeta_{p-1}, \zeta_p^1, \dots, \zeta_p^{i-1}) = \mathbb{P}(\zeta_p^i \in A \mid \text{pa}(\zeta_p^i)), \quad \mathbb{P}\text{-a.s.} \tag{30}$$

Then using again the fact that  $W_p^{i_p}$  is measurable w.r.t.  $\sigma(\text{pa}(\zeta_p^{i_p}))$ , we have by (12) that (27) holds for  $Mm < \varrho \leq Mm(n + 1)$ , which completes the proof.  $\square$

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

**Lemma 2.** For any  $0 \leq p, q \leq n$  and  $i, j \in [Mm]$ ,

$$\text{pa}(\zeta_p^i) \cap \text{pa}(\zeta_q^j) = \emptyset \implies \sigma(\text{pa}(\zeta_p^i)) \perp \sigma(\text{pa}(\zeta_q^j)). \tag{31}$$

**Proof.** The implication in (31) holds immediately in the case that  $p, q \in \{0, 1\}$ , due to the convention that  $\sigma(\emptyset)$  is the trivial  $\sigma$ -algebra and the independence of the  $\zeta_0^i$ 's. So suppose w.l.o.g.  $p > 1$  and  $0 \leq q \leq p$ , fix any  $i, j \in [Mm]$  and assume that  $\text{pa}(\zeta_p^i) \cap \text{pa}(\zeta_q^j) = \emptyset$ . For  $0 \leq r < p$ , define the sets of random variables

$$\begin{aligned} \mathcal{Z}_r &:= \text{pa}(\zeta_p^i) \cap \{\zeta_s^k; 0 \leq s \leq r, k \in [Mm]\}, \\ \mathcal{Z}'_r &:= \text{pa}(\zeta_q^j) \cap \{\zeta_s^k; 0 \leq s \leq r, k \in [Mm]\}, \end{aligned}$$

notice that  $\mathcal{Z}_{p-1} = \text{pa}(\zeta_p^i)$  and similarly  $\mathcal{Z}'_{p-1} = \text{pa}(\zeta_q^j)$ , so our objective is to prove  $\sigma(\mathcal{Z}_{p-1}) \perp \sigma(\mathcal{Z}'_{p-1})$ . Notice also that  $\mathcal{Z}_r \cap \mathcal{Z}'_r = \emptyset$  for  $0 \leq r < p$  since we have assumed  $\text{pa}(\zeta_p^i) \cap \text{pa}(\zeta_q^j) = \emptyset$ . We proceed with an inductive argument, the induction hypothesis being that for some  $0 \leq r < p - 1$ ,

$$\sigma(\mathcal{Z}_r) \perp \sigma(\mathcal{Z}'_r). \tag{32}$$

To initialise, observe that (32) holds with  $r = 0$ , due when  $q = 0$  to the convention that  $\sigma(\emptyset)$  is trivial, and due when  $q > 0$  to the independence of the  $\zeta_0^i$ 's and  $\mathcal{Z}_0 \cap \mathcal{Z}'_0 = \emptyset$ . Now assume that (32) holds for some  $0 \leq r < p - 1$ , for each  $\zeta \in \mathcal{Z}_{r+1} \cup \mathcal{Z}'_{r+1}$  let  $B_\zeta$  be an arbitrary member of  $\mathcal{X}$  and let  $A_\zeta$  be the event  $\{\zeta \in B_\zeta\}$ . Then writing  $\mathcal{G}_r := \sigma(\zeta_0, \dots, \zeta_r)$ , and with the convention that products over the empty set are unity, we have

$$\begin{aligned} & \mathbb{P} \left( \bigcap_{\zeta \in \mathcal{Z}_{r+1} \cup \mathcal{Z}'_{r+1}} A_\zeta \right) \\ &= \mathbb{E} \left[ \mathbb{P} \left( \bigcap_{\zeta \in \mathcal{Z}_{r+1} \setminus \mathcal{Z}_r \cup \mathcal{Z}'_{r+1} \setminus \mathcal{Z}'_r} A_\zeta \mid \mathcal{G}_r \right) \prod_{\zeta \in \mathcal{Z}_r \cup \mathcal{Z}'_r} \mathbb{I}[A_\zeta] \right] \\ &= \mathbb{E} \left[ \mathbb{P} \left( \bigcap_{\zeta \in \mathcal{Z}_{r+1} \setminus \mathcal{Z}_r} A_\zeta \mid \mathcal{G}_r \right) \mathbb{P} \left( \bigcap_{\zeta \in \mathcal{Z}'_{r+1} \setminus \mathcal{Z}'_r} A_\zeta \mid \mathcal{G}_r \right) \prod_{\zeta \in \mathcal{Z}_r \cup \mathcal{Z}'_r} \mathbb{I}[A_\zeta] \right] \\ &= \mathbb{E} \left[ \mathbb{P} \left( \bigcap_{\zeta \in \mathcal{Z}_{r+1} \setminus \mathcal{Z}_r} A_\zeta \mid \sigma(\mathcal{Z}_r) \right) \mathbb{P} \left( \bigcap_{\zeta \in \mathcal{Z}'_{r+1} \setminus \mathcal{Z}'_r} A_\zeta \mid \sigma(\mathcal{Z}'_r) \right) \prod_{\zeta \in \mathcal{Z}_r \cup \mathcal{Z}'_r} \mathbb{I}[A_\zeta] \right] \\ &= \mathbb{P} \left( \bigcap_{\zeta \in \mathcal{Z}_{r+1}} A_\zeta \right) \mathbb{P} \left( \bigcap_{\zeta \in \mathcal{Z}'_{r+1}} A_\zeta \right). \end{aligned}$$

The first equality uses the tower property of conditional expectations and the fact that  $\sigma(\mathcal{Z}_r) \vee \sigma(\mathcal{Z}'_r) \subset \mathcal{G}_r$ . The second and third equalities use the following facts: in Algorithm 3,  $\zeta_{r+1} = \{\zeta_{r+1}^k : k \in [Mm]\}$  are conditionally independent given  $\mathcal{G}_r$ ; for any  $\zeta \in \mathcal{Z}_{r+1} \setminus \mathcal{Z}_r$  (resp.  $\zeta \in \mathcal{Z}'_{r+1} \setminus \mathcal{Z}'_r$ ),  $\mathbb{P}(A_\zeta \mid \mathcal{G}_r)$  is measurable w.r.t.  $\sigma(\text{pa}(\zeta))$  (see (28)–(30));  $\text{pa}(\zeta) \subset \{\zeta_s^k; 0 \leq s \leq r, k \in [Mm]\}$  and by (26)  $\text{pa}(\zeta) \subset \text{pa}(\zeta_p^i)$ , hence  $\sigma(\text{pa}(\zeta)) \subset \sigma(\mathcal{Z}_r)$  (resp.  $\sigma(\text{pa}(\zeta)) \subset \sigma(\mathcal{Z}'_r)$ ).

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>



The fourth equality holds by the induction hypothesis. By a monotone class argument, (32) then holds with  $r$  replaced by  $r + 1$ , which completes the induction and hence also the proof of (31). □

**Lemma 3.** *Under Assumption (2.3), the number of pairs  $\varrho \neq \varrho'$  such that  $\text{pa}(\zeta_{p_m(\varrho)}^{i_m(\varrho)}) \cap \text{pa}(\zeta_{p_m(\varrho')}^{i_m(\varrho')}) = \emptyset$  is at least  $Mm(n + 1)^2(Mm - 4n\beta - 1)$ .*

**Proof.** We start by proving the implication

$$\prod_{q=0}^{n-1} \alpha^{i_q+1i_q} > 0 \implies \Delta(i_p, i_n) \leq (n - p)\beta \leq n\beta \quad \forall 0 \leq p < n. \tag{33}$$

By (2.3),  $\prod_{q=0}^{n-1} \alpha^{i_q+1i_q} > 0$  implies  $\Delta(i_{p+1}, i_p) \leq \beta, \forall 0 \leq p < n$  and then since  $\Delta$  is a metric, (33) follows from the triangle inequality.

Note that by (25) and (33),  $\text{pa}(\zeta_r^i) \subset \{\zeta_r^k; 0 \leq r < p, \Delta(i, k) \leq p\beta\}$  and therefore when  $i = 1$ ,  $\text{pa}(\zeta_p^i) \cap \text{pa}(\zeta_q^j) = \emptyset$  for all  $0 \leq p, q \leq n$  and  $j \in \{2n\beta + 2, \dots, Mm - 2n\beta\}$ , the latter set being non-empty for all  $m$  large enough, since  $M, \beta$  and  $n$  are fixed. Hence for  $i = 1$  fixed, there are at least  $(n + 1)^2(Mm - 4n\beta - 1)$  pairs  $(\zeta_p^i, \zeta_q^j)$  such that  $\text{pa}(\zeta_p^i) \cap \text{pa}(\zeta_q^j) = \emptyset$ . Then allowing  $i$  to vary over the set  $[Mm]$  gives the lower bound as claimed. □

### 3.4. Convergence of the variance

The main result of Section 3.4 is:

**Proposition 3.** *Under the assumptions of Theorem 1, for all  $n > 0$*

$$\lim_{m \rightarrow \infty} \mathbb{E} \left[ \left( \sum_{\varrho=1}^{(n+1)Mm} \xi_{\varrho}^m \right)^2 \right] = \frac{1}{M\gamma_n(\mathbf{1})^2} \sum_{\substack{0 \leq u < M \\ |v| \leq 2n\beta}} \mathbb{E}_{u,v+u} \left[ \pi_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \cdots Q_n^{\otimes 2} C_{\epsilon_n} (\bar{\varphi}^{\otimes 2}) \right]. \tag{34}$$

where  $\epsilon_k = \mathbb{I}[I_k = J_k]$ , for all  $0 \leq k \leq n$ .

From (16) and (9) it follows that

$$\mathbb{E} \left[ \left( \sum_{\varrho=1}^{(n+1)Mm} \xi_{\varrho}^m \right)^2 \right] = \frac{Mm}{\gamma_n(\mathbf{1})^2} \mathbb{E} \left[ \left( \frac{1}{Mm} \sum_i W_n^i \bar{\varphi}(\zeta_n^i) \right)^2 \right]. \tag{35}$$

The first step towards proving Proposition 3 is to develop an expression for the expectation on the r.h.s. of (35) in the following lemma, which is inspired by tensor product analysis of [3].

**Lemma 4.** *Fix  $n \in \mathbb{N}, M \geq 1, m \geq 1$  and set  $N = Mm$ . For any  $\varphi \in \mathcal{B}(\mathbb{X})$ ,*

$$\begin{aligned} & \mathbb{E} \left[ \left( \frac{1}{N} \sum_i W_n^i \varphi(\zeta_n^i) \right)^2 \right] \\ &= \frac{1}{N^2} \sum_{(i_0:n, j_0:n)} \left( \prod_{q=0}^{n-1} \alpha^{i_q+1i_q} \alpha^{j_q+1j_q} \right) \pi_0^{\otimes 2} C_{\mathbb{I}[i_0=j_0]} Q_1^{\otimes 2} C_{\mathbb{I}[i_1=j_1]} \cdots Q_n^{\otimes 2} C_{\mathbb{I}[i_n=j_n]} (\varphi^{\otimes 2}). \end{aligned}$$

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>



$$\begin{aligned}
 &= \frac{1}{N^2} \sum_{(i_{k:n}, j_{k:n})} \left( \prod_{q=k}^{n-1} \alpha^{i_{q+1}i_q} \alpha^{j_{q+1}j_q} \right) \times \left( \left( \sum_{i_{k-1}} \alpha^{i_k i_{k-1}} W_{k-1}^{i_{k-1}} \delta_{\zeta_{k-1}^{i_{k-1}}} \right) \right. \\
 &\quad \left. \otimes \left( \sum_{j_{k-1}} \alpha^{j_k j_{k-1}} W_{k-1}^{j_{k-1}} \delta_{\zeta_{k-1}^{j_{k-1}}} \right) \right) \left( Q_k^{\otimes 2} C_{\epsilon_k} \cdots Q_n^{\otimes 2} C_{\epsilon_n}(\varphi) \right) \\
 &= \frac{1}{N^2} \sum_{(i_{k-1:n}, j_{k-1:n})} \left( \prod_{q=k-1}^{n-1} \alpha^{i_{q+1}i_q} \alpha^{j_{q+1}j_q} \right) \\
 &\quad \times \left( W_{k-1}^{i_{k-1}} \delta_{\zeta_{k-1}^{i_{k-1}}} \otimes W_{k-1}^{j_{k-1}} \delta_{\zeta_{k-1}^{j_{k-1}}} \right) \left( Q_k^{\otimes 2} C_{\epsilon_k} \cdots Q_n^{\otimes 2} C_{\epsilon_n}(\varphi) \right),
 \end{aligned}$$

proving that the induction hypothesis holds at rank  $k - 1$ . Thus

$$\begin{aligned}
 &\mathbb{E} \left[ \left( \left( \frac{1}{N} \sum_{i_n} W_n^{i_n} \delta_{\zeta_n^{i_n}} \right) \otimes \left( \frac{1}{N} \sum_{j_n} W_n^{j_n} \delta_{\zeta_n^{j_n}} \right) \right) (\varphi) \middle| \mathcal{G}_0 \right] \\
 &= \frac{1}{N^2} \sum_{(i_{0:n}, j_{0:n})} \left( \prod_{q=0}^{n-1} \alpha^{i_{q+1}i_q} \alpha^{j_{q+1}j_q} \right) \left( W_0^{i_0} \delta_{\zeta_0^{i_0}} \otimes W_0^{j_0} \delta_{\zeta_0^{j_0}} \right) \left( Q_1^{\otimes 2} C_{\epsilon_1} \cdots Q_n^{\otimes 2} C_{\epsilon_n}(\varphi) \right).
 \end{aligned}$$

Finally, since  $\{\zeta_0^i : i \in [N]\}$  are i.i.d. samples from  $\pi_0$  and  $W_0^i = 1$  for all  $i \in [N]$ , we have

$$\begin{aligned}
 &\mathbb{E} \left[ \left( \left( \frac{1}{N} \sum_{i_n} W_n^{i_n} \delta_{\zeta_n^{i_n}} \right) \otimes \left( \frac{1}{N} \sum_{j_n} W_n^{j_n} \delta_{\zeta_n^{j_n}} \right) \right) (\varphi) \right] \\
 &= \frac{1}{N^2} \sum_{(i_{0:n}, j_{0:n})} \left( \prod_{q=0}^{n-1} \alpha^{i_{q+1}i_q} \alpha^{j_{q+1}j_q} \right) \pi_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \cdots Q_n^{\otimes 2} C_{\epsilon_n}(\varphi),
 \end{aligned}$$

from which the claim follows by observing that

$$\left( \left( \frac{1}{N} \sum_{i_n} W_n^{i_n} \delta_{\zeta_n^{i_n}} \right) \otimes \left( \frac{1}{N} \sum_{j_n} W_n^{j_n} \delta_{\zeta_n^{j_n}} \right) \right) (\varphi^{\otimes 2}) = \left( \frac{1}{N} \sum_{i_n} W_n^{i_n} \varphi(\zeta_n^{i_n}) \right)^2. \quad \square$$

**Proof of Proposition 3.** Throughout the proof we use the shorthand notations  $i_{p:q} := (i_p, \dots, i_q)$ ,  $j_{p:q} := (j_p, \dots, j_q)$ ,  $i_{p:q} + u := (i_p + u, \dots, i_q + u)$  and  $j_{p:q} + u := (j_p + u, \dots, j_q + u)$  for any  $u \in \mathbb{Z}$  and  $p, q \in \mathbb{N}$  such that  $q < p$ . Also we define

$$\Xi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) := \pi_0^{\otimes 2} C_{\mathbb{I}[i_0=j_0]} Q_1^{\otimes 2} C_{\mathbb{I}[i_1=j_1]} \cdots Q_n^{\otimes 2} C_{\mathbb{I}[i_n=j_n]}(\varphi^{\otimes 2}), \tag{37}$$

and

$$\begin{aligned}
 \Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) &:= \prod_{q=0}^{n-1} \alpha^{i_{q+1}i_q} \alpha^{j_{q+1}j_q}, \\
 \Pi_{i_n, j_n}^\infty(i_{0:n-1}, j_{0:n-1}) &:= \prod_{q=0}^{n-1} \alpha_\infty^{i_{q+1}i_q} \alpha_\infty^{j_{q+1}j_q}.
 \end{aligned} \tag{38}$$

By Lemma 4, we have

$$(Mm)^2 \mathbb{E} \left[ \left( \frac{1}{Mm} \sum_i W_n^i \overline{\varphi}(z_n^i) \right)^2 \right] = \sum_{(i_{0:n}, j_{0:n})} \Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) \Xi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) = A_m + B_m \tag{39}$$

where  $A_m$  and  $B_m$  are obtained by partitioning the summation set:

$$A_m := \sum_{\substack{(i_{0:n}, j_{0:n}): \\ \Delta(i_n, j_n) > 2n\beta}} \Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) \Xi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}),$$

$$B_m := \sum_{i_n=1}^{Mm} \sum_{\substack{j_n: \\ \Delta(i_n, j_n) \leq 2n\beta}} \sum_{(i_{0:n-1}, j_{0:n-1})} \Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) \Xi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}). \tag{40}$$

Note that although not explicitly shown in the notation,  $\Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1})$  depends also on  $m$  through the size of matrix  $\alpha$ , whilst  $\Xi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1})$  does not.

We shall prove that  $A_m = 0$  and that for all  $m$  large enough  $B_m$  is equal to the r.h.s. of (34). First consider  $A_m$ . We can use the implication (33), given in the proof of Lemma 3, and observe that if  $\Delta(i_n, j_n) > 2n\beta$  and  $\Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) > 0$ , then by two applications of the triangle inequality

$$\Delta(i_p, j_p) \geq \Delta(i_p, j_n) - \Delta(j_p, j_n) \geq \Delta(i_n, j_n) - \Delta(i_p, i_n) - \Delta(j_p, j_n) > 0,$$

and hence  $\mathbb{I}[i_p = j_p] = 0$ , for all  $0 \leq p \leq n$ . Consequently, by using the fact that

$$\pi_0^{\otimes 2} Q_1^{\otimes 2} \dots Q_n^{\otimes 2} (\overline{\varphi}^{\otimes 2}) = (\pi_0 Q_1 \dots Q_n (\overline{\varphi}))^2 = \gamma_n(\mathbf{1})^2 \pi_n(\overline{\varphi})^2 = 0,$$

we have

$$A_m = \sum_{\substack{(i_{0:n}, j_{0:n}): \\ \Delta(i_n, j_n) > 2n\beta}} \Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) \pi_0^{\otimes 2} Q_1^{\otimes 2} \dots Q_n^{\otimes 2} (\overline{\varphi}^{\otimes 2}) = 0. \tag{41}$$

Next we consider  $B_m$ . Let us start by writing

$$B_{i_n, j_n}^m := \sum_{(i_{0:n-1}, j_{0:n-1})} \Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) \Xi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}), \tag{42}$$

and

$$\phi(a_1, \dots, a_p) := ((a_1 + kM) \bmod N, \dots, (a_p + kM) \bmod N), \quad \forall (a_1, \dots, a_p) \in \mathbb{Z}^p,$$

for some fixed  $k \in \mathbb{Z}$  and any  $p > 0$ . First we prove that  $B_{i_n, j_n}^m$  satisfies

$$B_{i_n, j_n}^m = B_{\phi(i_n, j_n)}^m. \tag{43}$$

By (2.2) we have immediately, for all  $i_{0:n}, j_{0:n} \in [Mm]^{n+1}$ ,

$$\Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) = \Pi_{\phi(i_n, j_n)}(\phi(i_{0:n-1}, j_{0:n-1})), \tag{44}$$

and also

$$\Xi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1}) = \Xi_{\phi(i_n, j_n)}(\phi(i_{0:n-1}, j_{0:n-1})). \tag{45}$$

Combining (42), (44), (45) and using the fact that  $\phi : [Mm]^n \times [Mm]^n \rightarrow [Mm]^n \times [Mm]^n$  is a bijection to perform a change of variable, we can write

$$\begin{aligned} B_{i_n, j_n}^m &= \sum_{(i_{0:n-1}, j_{0:n-1})} \Pi_{\phi(i_n, j_n)}(\phi(i_{0:n-1}, j_{0:n-1})) \Xi_{\phi(i_n, j_n)}(\phi(i_{0:n-1}, j_{0:n-1})) \\ &= \sum_{(i_{0:n-1}, j_{0:n-1})} \Pi_{\phi(i_n, j_n)}(i_{0:n-1}, j_{0:n-1}) \Xi_{\phi(i_n, j_n)}(i_{0:n-1}, j_{0:n-1}) \\ &= B_{\phi(i_n, j_n)}^m, \end{aligned}$$

establishing (43).

Since for any  $u, v \in [N]$ , and  $0 \leq c \leq N/2$ ,  $\Delta(u, v) = c$  if and only if  $u = (v \pm c) \overline{\text{mod} N}$ , we can re-parametrise the summations in (40), and by using (43) we have

$$\begin{aligned} B_m &= \sum_{k=0}^{m-1} \sum_{\ell=1}^M \sum_{|c| \leq 2n\beta} B_{\ell+kM, (\ell+kM+c) \overline{\text{mod} N}}^m \\ &= m \sum_{\ell=1}^M \sum_{|c| \leq 2n\beta} B_{u_0+\ell, (u_0+\ell+c) \overline{\text{mod} N}}^m \end{aligned} \tag{46}$$

for any  $0 \leq u_0 \leq (m - 1)M$ .

Recall (33) from the proof of Lemma 3. An analogous implication

$$\prod_{q=0}^{n-1} \alpha_{\infty}^{i_{q+1}i_q} > 0 \implies |i_p - i_n| \leq (n - p)\beta \leq n\beta \quad \forall 0 \leq p < n, \tag{47}$$

can be established for  $\alpha_{\infty}$  by using the absolute difference instead of the metric  $\Delta$ .

Let us set  $u_0 = 3n\beta$  and assume that  $m > (u_0 + M + 3n\beta)/M$ , which is legitimate since our aim is to find the limit of  $B_m$  as  $m \rightarrow \infty$ . We then have

$$(u_0 + \ell + c) \overline{\text{mod} N} = u_0 + \ell + c, \quad \forall \ell \in [M], |c| \leq 2n\beta, \tag{48}$$

and by using (33) and (47) one can check that when  $i_n = u_0 + \ell$  and  $j_n = u_0 + \ell + c$ , then  $\Pi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1})$  and  $\Xi_{i_n, j_n}(i_{0:n-1}, j_{0:n-1})$  are greater than zero only if  $\beta < i_{q+1}$ ,  $j_{q+1} \leq Mm - \beta$ , for all  $0 \leq q < n$ . But by (2.4),  $\alpha^{i_{q+1}i_q} = \alpha_{\infty}^{i_{q+1}i_q}$  for all  $\beta < i_{q+1} \leq Mm - \beta$  and  $i_q \in [Mm]$ . Thus we have by (42) and (48)

$$\begin{aligned} &B_{u_0+\ell, (u_0+\ell+c) \overline{\text{mod} N}}^m \\ &= \sum_{\substack{i_{0:n-1} \in \mathbb{Z}^n \\ j_{0:n-1} \in \mathbb{Z}^n}} \Pi_{u_0+\ell, u_0+\ell+c}^{\infty}(i_{0:n-1}, j_{0:n-1}) \Xi_{u_0+\ell, u_0+\ell+c}(i_{0:n-1}, j_{0:n-1}). \end{aligned} \tag{49}$$

Finally we use the fact that by (2.2) and (2.4),  $\alpha_{\infty}^{i+kM, j+kM} = \alpha_{\infty}^{ij}$  for all  $i, j, k \in \mathbb{Z}$  and hence by (46), (49) and the fact that  $\Xi_{u_0+\ell, u_0+\ell+c}(i_{0:n-1}, j_{0:n-1}) = \Xi_{\ell, \ell+c}(i_{0:n-1}, j_{0:n-1})$ ,

$$\begin{aligned} \frac{B_m}{m} &= \sum_{\ell=0}^{M-1} \sum_{|c| \leq 2n\beta} \sum_{\substack{i_{0:n-1} \in \mathbb{Z}^n \\ j_{0:n-1} \in \mathbb{Z}^n}} \Pi_{\ell, \ell+c}^{\infty}(i_{0:n-1}, j_{0:n-1}) \Xi_{\ell, \ell+c}(i_{0:n-1}, j_{0:n-1}) \\ &= \sum_{\ell=0}^{M-1} \sum_{|c| \leq 2n\beta} \mathbb{E}_{\ell, \ell+c} \left[ \pi_0^{\otimes 2} \mathcal{C}_{\epsilon_0} Q_1^{\otimes 2} \mathcal{C}_{\epsilon_1} \cdots Q_n^{\otimes 2} \mathcal{C}_{\epsilon_n} (\overline{\varphi}^{\otimes 2}) \right], \end{aligned} \tag{50}$$

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

where the last form is independent of  $m$ . The claim then follows by combining (35), (39), (41) and (50). □

**4. Time-uniform convergence**

Recall from Proposition 1 that for Algorithm 3, if Assumption (2.1) holds, then for each  $n \in \mathbb{N}$  and  $p \geq 1$ ,

$$\sup_{M, m \geq 1} \sqrt{Mm} \mathbb{E}[|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)|^p]^{1/p} < \infty. \tag{51}$$

In this section we establish conditions under which the LEPF and IBPF satisfy, for all  $p \geq 1$ :

$$\sup_{M \geq 1} \sup_{n \geq 0} \sqrt{M} \mathbb{E}[|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)|^p]^{1/p} < \infty, \tag{52}$$

and do not satisfy, for any  $p \geq 1$ :

$$\sup_{m \geq 1} \sup_{n \geq 0} \sqrt{m} \mathbb{E}[|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)|^p]^{1/p} < \infty, \tag{53}$$

where in (52),  $m$  is fixed and in (53),  $M$  is fixed. We note that (52) and (53) are equivalent to corresponding inequalities with  $\sup_{M \geq 1}$  and  $\sup_{m \geq 1}$  replaced by  $\limsup_{M \rightarrow \infty}$  and  $\limsup_{m \rightarrow \infty}$  respectively, since for  $\varphi \in \mathcal{B}(\mathbb{X})$ ,  $|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)| \leq \text{osc}(\varphi) < \infty$ .

We shall again leverage the fact that the LEPF and IBPF are instances of Algorithm 3, which is itself an instance of  $\alpha$ SMC from [23], where it was shown that

$$\mathcal{E}_n^N := \frac{\left(\frac{1}{N} \sum_i W_n^i\right)^2}{\frac{1}{N} \sum_i (W_n^i)^2}, \quad N_n^{\text{eff}}(M, m) := Mm\mathcal{E}_n^{Mm},$$

play a central role in time-uniform convergence. The quantity  $N_n^{\text{eff}}$  is commonly called the *effective sample size*. Note that by Jensen’s inequality we always have  $\mathcal{E}_n^N \leq 1$ , or equivalently  $N_n^{\text{eff}}(M, m) \leq Mm$ . We shall appeal to the following result, which is a special case of [23, Proposition 3] (in particular see the last displayed equation in [23, Proof of Theorem 2]).

**Proposition 4.** *Suppose that Assumption (2.1) holds and additionally,*

$$\exists(\delta, \epsilon) \in [1, \infty)^2 \quad \text{s.t.} \quad \sup_{n \geq 0} \sup_{x, y} \frac{g_n(x)}{g_n(y)} \leq \delta, \quad \text{and} \quad f(x, \cdot) \leq \epsilon f(y, \cdot), \quad \forall x, y \in \mathbb{X}. \tag{54}$$

*Then there exists  $\rho < 1$  and for each  $p \geq 1$  a finite constant  $c_p$  such that for any  $n \geq 0$ ,  $M \geq 1$ ,  $m \geq 1$  and  $\varphi \in \mathcal{B}(\mathbb{X})$ , Algorithm 3 has the property:*

$$\mathbb{E}[|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)|^p]^{1/p} \leq \|\varphi\|_\infty \frac{c_p}{\sqrt{Mm}} \sum_{q=0}^n \rho^{n-q} \mathbb{E} \left[ |\mathcal{E}_q^{Mm}|^{-p/2} \right]^{1/p}. \tag{55}$$

**4.1. The regime  $m$  fixed and  $M \rightarrow \infty$**

We shall now show that under the assumptions of Proposition 4, both the LEPF and IBPF satisfy (52). For the LEPF, this is a new result. For the IBPF, the result is not very surprising, since

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

it is well known that under the strong but standard hypothesis (54), a single bootstrap particle filter is time-uniformly convergent (see [6, Section 7.4.3.] and references therein). However, perhaps more surprising is the simplicity of the following argument, which applies to both the IBPF and LEPF.

It was noted in Section 2 Eq. (2) that for both the LEPF and IBPF, for any  $k \in [m]$ ,

$$W_n^i = W_n^j =: \tilde{W}_n^k, \quad \forall i, j \in G_k. \tag{56}$$

Consequently, for any  $M, m \geq 1$  and  $n \geq 0$ ,

$$\begin{aligned} \mathcal{E}_n^{Mm} &= \frac{\left(\frac{1}{N} \sum_{k=1}^m M \tilde{W}_n^k\right)^2}{\frac{1}{N} \sum_{k=1}^m M (\tilde{W}_n^k)^2} \\ &= \frac{1}{m} \frac{\left(\frac{1}{\sqrt{m}} \sum_{k=1}^m \tilde{W}_n^k\right)^2}{\frac{1}{m} \sum_{k=1}^m (\tilde{W}_n^k)^2} \\ &= \frac{1}{m} \left(1 + \frac{\sum_{k=1}^m \sum_{\ell \neq k} \tilde{W}_n^k \tilde{W}_n^\ell}{\sum_{k=1}^m (\tilde{W}_n^k)^2}\right) \geq \frac{1}{m}, \end{aligned} \tag{57}$$

or alternatively  $N_n^{\text{eff}} \geq M$ . Substituting the lower bound (57) into (55) gives (52) as claimed.

4.2. The regime  $M$  fixed and  $m \rightarrow \infty$

The following proposition establishes that  $\limsup_{n \rightarrow \infty} \sigma_n^2 = \infty$  is a sufficient condition for failure of (53). In Sections 5.1 and 5.2 we present examples for the IBPF and LEPF such that  $\lim_{n \rightarrow \infty} \sigma_n^2 = \infty$  and (54) holds.

**Proposition 5.** Consider Algorithm 3. Assume that the hypotheses of Theorem 1 hold, fix  $\varphi \in \mathcal{B}(\mathbb{X})$  and  $M > 1$ . Then for any  $n \geq 0$  and  $p \geq 1$ ,

$$\lim_{m \rightarrow \infty} \sqrt{m} \mathbb{E}[|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)|^p]^{1/p} = \frac{\sigma_n}{\sqrt{M}} \sqrt{2} \left(\frac{\Gamma((p+1)/2)}{\sqrt{\pi}}\right)^{1/p}. \tag{58}$$

If  $\limsup_{n \rightarrow \infty} \sigma_n^2 = \infty$ , then (53) does not hold for any  $p \geq 1$ . If additionally (54) holds, then for the LEPF and IBPF, for any  $p \geq 1$

$$\limsup_{m \rightarrow \infty} \sup_{n \geq 0} m^{p/2} \mathbb{E} \left[ \left| \sqrt{\sum_{k \in [m]} \left( \frac{\tilde{W}_n^k}{\sum_{j \in [m]} \tilde{W}_n^j} \right)^2} \right|^p \right] = \infty. \tag{59}$$

**Remark 4.** The condition in (59) clearly rules out:

$$\sup_{m \geq 1} \sup_{n \geq 0} m^p \mathbb{E} \left[ \max_{k \in [m]} \frac{\tilde{W}_n^k}{\sum_{j \in [m]} \tilde{W}_n^j} \right]^p < \infty,$$

which is exactly the key hypothesis of [18] as written in (3) in the case  $\epsilon = 0$ . Note however, that whilst proposition establishes that (53) does not hold, i.e. time-uniform convergence at rate  $m^{-1/2}$  does not occur, we have not ruled out the possibility that time-uniform convergence occurs at some slower rate. Moreover, our negative result is of course valid only for the specific local exchange mechanism appearing in Algorithm 1, which is only a special case of the more general framework of [18]. In Section 6.3 we shall comment on some possible algorithmic modifications to ensure time-uniform convergence.

**Proof.** To prove (58), we follow arguments used in the proof of [8, Theorem 12], who established a limit of the same form for a standard particle filter. We first recall the fact that for a sequence of random variables  $(A_m)_{m \geq 1}$ , if  $A_m \xrightarrow{d} A$  for some  $A$ , and for some  $p > 0$ ,  $(|A_m|^p)_{m \geq 1}$  is uniformly integrable, then  $\lim_{m \rightarrow \infty} \mathbb{E}[|A_m|^p] = \mathbb{E}[|A|^p]$ , see [19, p.14, Theorem A]. As in the statement, fix  $\varphi \in \mathcal{B}(\mathbb{X})$ ,  $M > 1$  and  $n \geq 0$ . Then set  $A_m = \sqrt{m} \pi_n^{Mm}(\varphi)$ . By Theorem 1,  $A_m$  converges in distribution to a zero-mean Gaussian variable with variance  $\sigma_n^2/M$ . For any given  $p \geq 1$  and  $\delta > 0$ , (51) implies  $\sup_{m \geq 1} \mathbb{E}[|A_m|^{p+\delta}] < \infty$ , so by [20, Lemma II.6.3],  $(|A_m|^p)_{m \geq 1}$  is uniformly integrable. Therefore (58) holds.

If (53) were to hold, the r.h.s. of (58) would be upper-bounded by a finite constant possibly depending on  $p$  and  $M$ , but independent of  $n$ . The latter would contradict  $\limsup_{n \rightarrow \infty} \sigma_n^2 = \infty$ . Hence (53) does not hold when  $\limsup_{n \rightarrow \infty} \sigma_n^2 = \infty$ .

Now assume (54) holds in addition to  $\limsup_{n \rightarrow \infty} \sigma_n^2 = \infty$ . In order to establish (59) by a contradiction, assume that for some  $p \geq 1$  there is a constant  $d_p$  such that

$$\limsup_{m \rightarrow \infty} \sup_{n \geq 0} m^{p/2} \mathbb{E} \left[ \left| \sum_{k \in [m]} \left( \frac{\tilde{W}_n^k}{\sum_{j \in [m]} \tilde{W}_n^j} \right)^2 \right|^p \right] = d_p < \infty. \tag{60}$$

Since for the IBPF and LEPF,  $W_n^i = W_n^j = \tilde{W}_n^k$  for all  $i, j \in G_k$ , we have

$$m \sum_{k \in [m]} \left( \frac{\tilde{W}_n^k}{\sum_{j \in [m]} \tilde{W}_n^j} \right)^2 = 1/\mathcal{E}_n^{Mm}.$$

Combining this and (60) into the bound (55) of Proposition 4 gives

$$\limsup_{m \rightarrow \infty} \sup_{n \geq 0} \sqrt{m} \mathbb{E}[|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)|^p]^{1/p} \leq \|\varphi\|_\infty \frac{c_p}{\sqrt{M}} \frac{d_p}{1 - \rho} < \infty,$$

in turn implying (53), since  $|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)| \leq \text{osc}(\varphi) < \infty$ . But we have already proved that (53) does not hold for any  $p \geq 1$  when  $\limsup_{n \rightarrow \infty} \sigma_n^2 = \infty$ , hence the inequality in (60) does not hold for any  $p \geq 1$ . This completes the proof.  $\square$



5. A closer look at the asymptotic variance

Our objective in this section is to develop more insight into the asymptotic variance in Theorem 1,

$$\sigma_n^2 = \frac{1}{M\gamma_n(\mathbf{1})^2} \sum_{\substack{0 \leq u < M \\ |v| \leq 2n\beta}} \mathbb{E}_{u,v+u} \left[ \pi_0^{\otimes 2} \mathcal{C}_{\mathbb{I}[I_0=J_0]} \mathcal{Q}_1^{\otimes 2} \mathcal{C}_{\mathbb{I}[I_1=J_1]} \cdots \mathcal{Q}_n^{\otimes 2} \mathcal{C}_{\mathbb{I}[I_n=J_n]} (\bar{\varphi}^{\otimes 2}) \right] \quad (61)$$

for the LEPF and IBPF, especially regarding its behaviour as  $n \rightarrow \infty$ .

For the convenience of the reader we recall that in (61),  $\mathbb{E}_{u,v}$  denotes expectation w.r.t. to the law of the bi-variate Markov chain:

$$\begin{aligned} (I_n, J_n) &\sim \delta_u \otimes \delta_v, \\ \mathbb{P}(I_k = i_k, J_k = j_k \mid I_{k+1} = i_{k+1}, J_{k+1} = j_{k+1}) &= \alpha_\infty^{i_{k+1}i_k} \alpha_\infty^{j_{k+1}j_k}, \end{aligned} \quad (62)$$

and thus the only distinction between the asymptotic variances for the LEPF and IBPF is through  $\alpha_\infty$ , as given in (5) and (6).

To help develop insight, we consider a much simplified HMM:

$$f(x, \cdot) = \pi_0(\cdot) \quad \text{and} \quad g_n = g \in \mathcal{B}(\mathbb{X}), \quad \forall n \in \mathbb{N}. \quad (63)$$

This is obviously quite unrealistic, so let us be clear about our motives:

Firstly, (63) can be understood as being a favourable assumption for the performance of the LEPF and IBPF:  $f(x, \cdot) = \pi_0(\cdot)$  implies that  $\pi_n = \pi_0$  and that the particles  $\{\zeta_n^i : i \in [N]\}$  in both Algorithms 1 and 2 are i.i.d. samples from  $\pi_0$  for all  $n \in \mathbb{N}$ . Never-the-less, we shall see in Section 5.2 in conjunction with Section 4.2 that under this favourable assumption certain *negative* results can hold for the IBPF and LEPF, namely  $\lim_{n \rightarrow \infty} \sigma_n^2 = \infty$  and lack of time-uniform convergence.

Secondly, we shall see that (63) makes the expression in (61) considerably more tractable, allowing us to make precise comparisons between the LEPF and IBPF. We shall see in Section 6.2 that our conclusions for this simplified HMM are consistent with results obtained by simulation for a more realistic stochastic volatility model.

Under (63), we have  $\pi_n = \pi_0$ ,  $\gamma_n(\mathbf{1}) = \pi_0 \mathcal{Q}_{0,n}(\mathbf{1}) = \pi_0(g)^n$ , and for all  $\Phi \in \mathcal{B}(\mathbb{X}^2)$  and  $1 \leq p \leq n$ ,

$$\mathcal{C}_{\mathbb{I}[I_{p-1}=J_{p-1}]} \mathcal{Q}_p^{\otimes 2}(\Phi) = (\mathbb{I}[I_{p-1} = J_{p-1}]g^2 + \mathbb{I}[I_{p-1} \neq J_{p-1}]g^{\otimes 2})\pi_0^{\otimes 2}(\Phi)$$

so

$$\begin{aligned} &\frac{1}{\gamma_n(\mathbf{1})^2} \pi_0^{\otimes 2} \mathcal{C}_{\mathbb{I}[I_0=J_0]} \mathcal{Q}_1^{\otimes 2} \mathcal{C}_{\mathbb{I}[I_1=J_1]} \cdots \mathcal{Q}_n^{\otimes 2} \mathcal{C}_{\mathbb{I}[I_n=J_n]} (\bar{\varphi}^{\otimes 2}) \\ &= \mathbb{I}[I_n = J_n] \pi_0(\bar{\varphi}^2) \prod_{p=0}^{n-1} \left( 1 + \mathbb{I}[I_p = J_p] \left( \frac{\pi_0(g^2)}{\pi_0(g)^2} - 1 \right) \right) \\ &= \mathbb{I}[I_n = J_n] \pi_0(\bar{\varphi}^2) (1 + c)^{Z_n}, \end{aligned} \quad (64)$$

where

$$Z_n = \sum_{p=0}^{n-1} \mathbb{I}[I_p = J_p] \quad (65)$$

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

and  $c = \pi_0(g^2)/\pi_0(g)^2 - 1$ . By (61) and (64), we thus have

$$\frac{\sigma_n^2}{\pi_0(\bar{\varphi}^2)} = \frac{1}{M} \sum_{\substack{0 \leq u < M \\ |v| \leq 2n\beta}} \mathbb{E}_{u,v+u} \left[ \mathbb{I}[I_n = J_n](1+c)^{Z_n} \right] = \frac{1}{M} \sum_{u=0}^{M-1} \mathbb{E}_{u,u} [e^{tZ_n}] \tag{66}$$

where  $t = \log(1+c)$  and the second equality follows from the initial condition part of (62).

We thus observe the key role in the asymptotic variance played by the moment generating function of the random variable  $Z_n$ , whose interpretation is clear by (65):  $Z_n$  is the number of times the Markov chains  $I$  and  $J$  collide in  $n$  steps. Intuitively, the more frequent these collisions tend to be, the faster the growth of the asymptotic variance.

To help formalise this intuition, our next step is to characterise the law of  $Z_n$  under (62) with  $u = v$ , for the IBPF and the LEPF, in order to understand how  $\sigma_n^2$  behaves as  $n \rightarrow \infty$ . We stress that this law is a consequence only of (62) and does not depend on (63).

### 5.1. Law of $Z_n$ for the IBPF

In the case of the IBPF we see immediately by inspecting  $\alpha_\infty$  in (6) (see also Fig. 2(b)) that when  $u = v$  for any  $u \in \mathbb{Z}$  in (62),  $I$  and  $J$  are sequences of i.i.d. random variables, each uniformly distributed on the set  $\{\lfloor (u-1)/M \rfloor M + 1, \dots, \lfloor (u-1)/M \rfloor M + M\}$ . Hence the random variables  $(\mathbb{I}[I_k - J_k])_{0 \leq k < n}$  constitute a sequence of Bernoulli variables with success probability  $M^{-1}$  and consequently

$$Z_n \sim \text{Binomial}(n, 1/M), \tag{67}$$

whatever the value of  $u$  (we note that this conclusion can also be deduced from [3, Lemma 3.2], which provides a non-asymptotic variance formula for a single bootstrap particle filter, i.e.  $N = M$ ). Hence (66) can be further simplified to

$$\frac{\sigma_n^2}{\pi_0(\bar{\varphi}^2)} = \frac{1}{M} \sum_{u=0}^{M-1} \mathbb{E}_{u,u} [e^{tZ_n}] = \mathbb{E}_{0,0} [e^{tZ_n}], \tag{68}$$

where  $t = \log(1+c)$ ,  $c = \pi_0(g^2)/\pi_0(g)^2 - 1$ .

By (67),  $\mathbb{E}_{0,0} [e^{tZ_n}]$  is the moment generating function of a binomial distribution, so readily,

$$\frac{\sigma_n^2}{\pi_0(\bar{\varphi}^2)} = \left(1 + \frac{c}{M}\right)^n. \tag{69}$$

Thus when (63) holds, and assuming that  $\pi_0(\bar{\varphi}^2) > 0$  and  $c > 0$ , for the IBPF  $\sigma_n^2$  grows exponentially fast as  $n \rightarrow \infty$ . This can be considered a negative result for the IBPF compared to the standard bootstrap particle filter, for which it has been shown that under a variety of more realistic conditions the sequence  $(\sigma_n^2)_{n \in \mathbb{N}}$  may be bounded by a finite constant, or is tight when the observation sequence is treated as random [4,15,9,22,8]. When (63) holds one can easily construct  $\pi_0$  and  $g$  such that (54) holds and  $c > 0$ .

### 5.2. Example of $\sigma_n^2 \rightarrow \infty$ for the LEPF

Let us point out an example which satisfies (63), (54) and for which  $\sigma_n^2 \rightarrow \infty$ . Notice that for the LEPF, it follows easily from (62) and (5) that for any  $u \in \mathbb{Z}$  and whatever the values of  $M$

and  $\theta$ ,

$$\mathbb{E}_{u,u}[\mathbb{I}[Z_n = n]] = \mathbb{E}_{u,u} \left[ \prod_{p=0}^{n-1} \mathbb{I}[I_p = J_p] \right] = \frac{1}{M^n},$$

hence we have the crude lower bound,

$$\mathbb{E}_{u,u}[e^{tZ_n}] \geq \left( \frac{\pi_0(g^2)}{\pi_0(g)^2} \frac{1}{M} \right)^n.$$

As we shall now demonstrate, one can readily construct examples for which  $\frac{\pi_0(g^2)}{\pi_0(g)^2} \frac{1}{M} > 1$  and hence such that  $\sigma_n^2 \rightarrow \infty$  exponentially fast for any  $\varphi$  with  $\pi_0(\bar{\varphi}^2) > 0$ . Let  $\mathbb{X} = \{0, 1\}$ ,  $p \in (0, 1/M)$ ,  $\delta \in (0, 1)$  and

$$\pi_0(0) = p, \quad \pi_0(1) = 1 - p, \quad g(0) = 1 - \delta, \quad g(1) = \delta.$$

Then, since

$$\frac{\pi_0(g^2)}{\pi_0(g)^2} = \frac{p(1 - \delta)^2 + (1 - p)\delta^2}{(p(1 - \delta) + (1 - p)\delta)^2} \xrightarrow{\delta \rightarrow 0} 1/p > M,$$

we can choose  $\delta$  small enough that  $\frac{\pi_0(g^2)}{\pi_0(g)^2} \frac{1}{M} > 1$ , whilst satisfying  $g \in \mathcal{B}(\mathbb{X})$  and  $g(x) > 0$ , as required for Assumption 1 and (54).

### 5.3. Law of $Z_n$ for the LEPF

The interaction pattern illustrated in Fig. 2(a) makes study of the law of  $Z_n$  more difficult for the LEPF than for the IBPF, but never-the-less we shall below derive an exact characterisation of the distribution of  $Z_n$ . Observe that  $Z_n$  depends on  $I$  and  $J$  only through the sequence of indicator variables  $(\mathbb{I}[I_k = J_k])_{0 \leq k < n}$ , but this sequence is unfortunately non-Markov and difficult to analyse directly. However the bi-variate process  $(D, E)$ , with  $D := (D_k)_{0 \leq k \leq n}$ ,  $E := (E_k)_{0 \leq k \leq n}$  and

$$D_k := \left\lfloor \frac{I_{n-k} - 1}{M} \right\rfloor - \left\lfloor \frac{J_{n-k} - 1}{M} \right\rfloor, \quad E_k := \mathbb{I}[I_{n-k} = J_{n-k}], \quad \forall 0 \leq k \leq n$$

is easier to deal with.

It follows from  $\alpha_\infty$  in (5), (62) and some elementary manipulations (omitted for brevity) that the bi-variate sequence  $(D_k, E_k)_{0 \leq k \leq n}$  is Markov and for any  $u \in \mathbb{Z}$  the initial condition  $(I_n, J_n) \sim \delta_u \otimes \delta_u$  implies  $(D_0, E_0) \sim \delta_0 \otimes \delta_1$ . Thus all statements about the law of functionals of  $(D, E)$  in the remainder of Section 5.3 hold irrespective of the particular value of  $u = v$  in (62).

By similarly elementary but lengthy manipulations it can be checked that  $D$  is also Markov, with for all  $1 \leq k \leq n$ ,  $d \in \mathbb{Z}$ ,

$$D_k \mid \{D_{k-1} = d\} \sim \frac{\theta(M - \theta)}{M^2} (\delta_{d+1} + \delta_{d-1}) + \left( \frac{(M - \theta)^2}{M^2} + \frac{\theta^2}{M^2} \right) \delta_d, \tag{70}$$

and if  $d_{k-1}, d_k \in \mathbb{Z}$  and  $x \in \{0, 1\}$  such that  $d_{k-1} = d_k = 0$ , i.e. that the integer parts of  $(I_{n-k} - 1)/M$  and  $(J_{n-k} - 1)/M$ , as well as  $(I_{n-k+1} - 1)/M$  and  $(J_{n-k+1} - 1)/M$ , coincide,

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

then

$$E_k \mid \{D_{k-1} = d_{k-1}, D_k = d_k, E_{k-1} = x\} \sim \text{Bernoulli}\left(\frac{M}{(M - \theta)^2 + \theta^2}\right) \tag{71}$$

and otherwise

$$E_k \mid \{D_{k-1} = d_{k-1}, D_k = d_k, E_{k-1} = x\} \sim \delta_0. \tag{72}$$

By (71) and (72), for all  $b \in [n]$ ,

$$Z_n \mid \{B = b\} \sim \text{Binomial}\left(b, \frac{M}{(M - \theta)^2 + \theta^2}\right) \tag{73}$$

where

$$B := \sum_{k=1}^n \mathbb{I}[D_{k-1} = 0] \mathbb{I}[D_k = 0]. \tag{74}$$

Therefore it remains to derive the distribution of  $B$ , the distribution of  $Z_n$  is then available by marginalisation.

We will write  $\text{Beta-Binomial}(n, a, b)$  for the so-called beta binomial distribution [21] specified for any  $a, b > 0$  by the probability mass function

$$p(k) = \binom{n}{k} \frac{B(k + a, n - k + b)}{B(a, b)}, \quad \forall 0 \leq k \leq n, \tag{75}$$

where  $B(a, b)$  denotes the beta-function. The case  $\text{Beta-Binomial}(n, 1, 0)$  is understood as the point mass  $\delta_n$ . Moreover, we write  $\text{RWZ}(n)$  for the distribution specified by the probability mass function:

$$p(x) = \frac{2^x}{2^{2\lfloor n/2 \rfloor}} \binom{2\lfloor n/2 \rfloor - x}{\lfloor n/2 \rfloor} \quad \forall 0 \leq x \leq \lfloor n/2 \rfloor, \tag{76}$$

with convention  $\binom{n}{0} = 1$  for all  $n \geq 0$ . As shown in [10, Theorem 2], (76) is the distribution of the number of times a symmetric simple random walk on  $\mathbb{Z}$  starting from zero returns to zero in  $n$  time-steps. The following result, in conjunction with (73), characterises the distribution of  $Z_n$ . The proof is in the [Appendix](#).

**Lemma 5.** Fix  $M \geq 1$  and  $\theta \in \{1, \dots, M - 1\}$ , let  $B$  be as defined in (74) and let  $\bar{B}$ ,  $S$  and  $V$  be random variables such that

$$\bar{B} \mid \{V = v, S = s\} \sim \text{Beta-Binomial}(v, s + 1, n - v - s), \tag{77}$$

and

$$S \mid \{V = v\} \sim \text{RWZ}(n - v), \quad V \sim \text{Binomial}\left(n, \frac{(M - \theta)^2}{M^2} + \frac{\theta^2}{M^2}\right). \tag{78}$$

Then  $B$  has the same distribution as  $\bar{B}$ .

## 6. Interpretation of results and discussion

One of the main conclusions which can be drawn from our results thus far is quite negative: we have seen in Section 5, that for the IBPF and LEPF, the asymptotic variance can increase over

time at an exponential rate. However, taken in isolation, this fact does not convey information about the relative performance of the two algorithms. The aim of Section 6 is to address this matter, qualitatively and numerically.

In Section 6.1, we continue with a toy model for which we are able to numerically evaluate asymptotic variances without simulation and explain the behaviour we see in terms of the collision count  $Z_n$ . We also examine dependence on the parameters  $M$  and  $\theta$ , compare asymptotic variance values with nonasymptotic values obtained by simulation, and explore the behaviour of the effective sample size. Section 6.2 considers a more realistic stochastic volatility model, and Section 6.3 provides some concluding perspectives and describes avenues for future investigation.

6.1. Evaluation of asymptotic variances

Recall that for the toy model of Section 5, the asymptotic variances for the IBPF and LEPF are proportional to  $\mathbb{E}_{0,0}[e^{tZ_n}]$ , where  $Z_n$  counts collisions of the Markov chains with transition probabilities given by  $\alpha_\infty$ . Due to the graph in Fig. 2(a) having only one connected component, versus several in Fig. 2(b), it seems natural to suppose that  $Z_n$  is “typically” lower for the LEPF than for the IBPF, and thus the LEPF will exhibit lower asymptotic variance.

To explore this idea, we now use (73) and Lemma 5 to make numerical evaluations of  $\mathbb{E}_{0,0}[e^{tZ_n}]$ . We do so for the specific instance of the model (63) where

$$X_0 \sim \mathcal{N}(0, 1) \quad \text{and} \quad g(x) = e^{-(x+1/2)^2/2} / \sqrt{2\pi}, \tag{79}$$

and define  $t_0 := \log(\pi_0(g^2)/\pi_0(g)^2) \approx .1855077$ .

Fig. 3(a) shows  $\mathbb{E}_{0,0}[e^{tZ_n}]$  vs.  $n$  for the LEPF. Noting the logarithmic scale, the plot suggests that  $\mathbb{E}_{0,0}[e^{tZ_n}]$  grows without bound as  $n \rightarrow \infty$ . In Fig. 3(b),  $R_n$  denotes the ratio of  $\mathbb{E}_{0,0}[e^{tZ_n}]$  for the IBPF to that for the LEPF. It is apparent that  $R_n$  is growing exponentially fast with  $n$ , suggesting the interaction structure of the LEPF has significant benefits in terms of asymptotic variance.

Fig. 3(c) compares  $R_n$  to the ratio of non-asymptotic mean square errors estimated by:

$$R_n^N := \frac{\sum_{i=1}^{N_{MC}} (\pi_{n,IBPF}^{N,i}(\varphi) - \pi_n(\varphi))^2}{\sum_{i=1}^{N_{MC}} (\pi_{n,LEPF}^{N,i}(\varphi) - \pi_n(\varphi))^2}, \tag{80}$$

where  $\pi_{n,IBPF}^{N,i}(\varphi)$  and  $\pi_{n,LEPF}^{N,i}(\varphi)$ ,  $i = 1, \dots, N_{MC}$ ,  $N_{MC} = 2000$  are independent approximations of  $\pi_n(\varphi)$ , with  $\varphi = Id$ , obtained from the IBPF and LEPF. It is apparent that as  $N$  grows,  $R_n^N$  approaches  $R_n$  and that the benefit of the LEPF over the IBPF becomes more substantial.

The main algorithmic difference between the LEPF and the IBPF is the number of particles exchanged between groups. For the IBPF, this number is 0, for the LEPF, is specified by the parameter  $\theta$ . Fig. 3(d) shows the behaviour of  $R_n$  for different values of  $\theta$ . The results suggest that highest value of  $R_n$  is obtained when  $\theta = M/2$ , i.e. half of the particles in each group are exchanged.

By (68), the behaviour of  $R_n$  is explained entirely by the distribution of  $Z_n$ . Fig. 4(a) shows a comparison of these distributions in the case that  $M = 3$  and  $\theta = 1$ , i.e. the same settings as

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

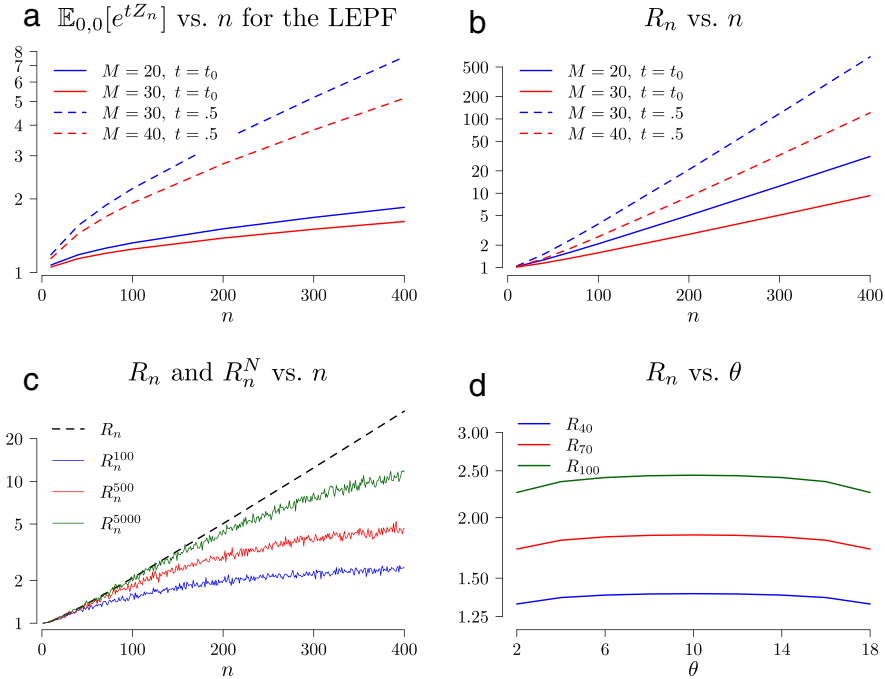


Fig. 3. (a)  $\mathbb{E}_{0,0}[e^{tZ_n}]$  vs.  $n$  for the LEPF with  $\theta = 1$ . (b)  $R_n$  vs.  $n$  for  $\theta = 1$ . (c)  $R_n$  and  $R_n^N$  vs.  $n$  for  $\theta = 1$ ,  $M = 20$  and  $t = t_0$ . (d)  $R_n$  vs.  $\theta$  for  $M = 20$  and  $t = t_0$ .

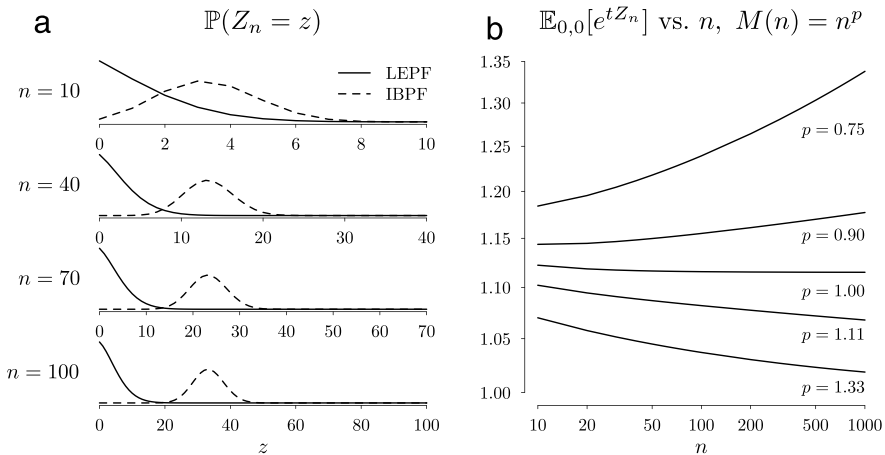


Fig. 4. (a) Probability mass functions of  $Z_n$  for IBPF and LEPF with  $M = 3$ ,  $\theta = 1$ . (b)  $\mathbb{E}_{0,0}[e^{tZ_n}]$  vs.  $n$  for  $M(n) = n^p$ .

in Fig. 2. By (67) the distribution of  $Z_n$  for the IBPF is centred at  $n/M$ , while the corresponding distribution in the case of LEPF remains concentrated near 0 and, in particular, we observe that the distributions become increasingly distinct for large  $n$ .

To help illustrate the connection to the convergence results of Section 4.2, Fig. 5 shows a simulation of  $\mathcal{E}_n^{Mm}$  over 70 000 time steps for the LEPF using the same model as before with

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

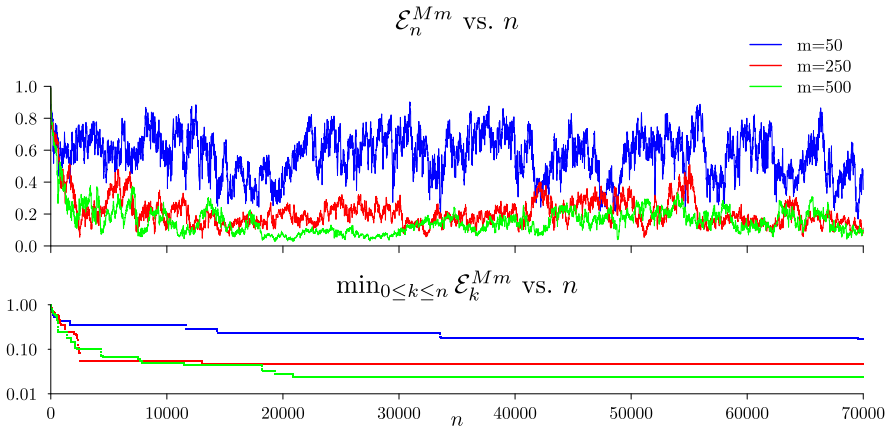


Fig. 5.  $\mathcal{E}_n^{Mm}$  (top) and its running minimum on log-scale (bottom) for  $M = 20$  and  $m = 50, 250, 500$ .

$M = 20$  and  $m = 50, 250, 500$ . For each fixed value of  $m$ ,  $\mathcal{E}_n^{Mm}$  does not crash to zero and stick there, but rather it fluctuates to some extent, eventually as  $n$  grows reaching values which are closer to 0 for larger values of  $m$ . It is relevant here to recall the possibly quite loose lower bound  $\mathcal{E}_n^{Mm} \geq 1/m$  derived in (57). Informally, some connection between this phenomenon and the convergence rate can be observed in Eq. (55), where  $(\mathcal{E}_n^{Mm})^{-1/2}$  appears in the  $L_p$  error bound.

Lastly we consider the question of how  $M = M(n)$  should be scaled with  $n$  in order to prevent explosion of  $\sigma_n^2$  as  $n \rightarrow \infty$ . So let  $\sigma_n^2$  be as given in (61) but with  $M$  replaced by  $M(n)$ . For the IBPF we see straightforwardly that if  $\limsup_{n \rightarrow \infty} n/M(n) < \infty$ , then by (69),

$$\limsup_{n \rightarrow \infty} \frac{\sigma_n^2}{\pi_0(\bar{\varphi}^2)} = \limsup_{n \rightarrow \infty} \left(1 + \frac{c}{n} \frac{n}{M(n)}\right)^n < \infty.$$

We address the same issue for the LEPF through numerical evaluations again using the formulae of Section 5.3. Fig. 4(b) shows the behaviour of  $\mathbb{E}_{0,0}[e^{tZ_n}] = \sigma_n^2/\pi_0(\bar{\varphi}^2)$  for the LEPF with  $M(n) = n^p$ ,  $p = 0.75, 0.90, 1.00, 1.11, 1.33$ . The results suggest that the “right” scaling may be  $M(n) = n$ , as for IBPF, in the sense that for  $p > 1$ ,  $\sigma_n^2$  tends towards  $\pi_0(\bar{\varphi}^2)$ , and for  $p < 1$ ,  $\limsup_{n \rightarrow \infty} \sigma_n^2 = \infty$ . We also note that for  $M(n) = n$ , we have from (79) for the IBPF that  $\lim_{n \rightarrow \infty} \mathbb{E}_{0,0}[e^{tZ_n}] = e^c \approx 1.23$ , where as for the LEPF in Fig. 4(b), with  $M(n) = n$ , it appears that  $\limsup_{n \rightarrow \infty} \mathbb{E}_{0,0}[e^{tZ_n}] \approx 1.12$ .

6.2. Simulations

We now see if some of the phenomena observed for the simplified model carry over to the case of a more realistic stochastic volatility model:

$$\begin{aligned} X_0 &\sim \mathcal{N}(0, 1), \quad X_{k+1} = aX_k + V_k, \quad V_k \sim \mathcal{N}(0, \sigma_V^2), \quad \forall k \geq 0, \\ Y_k &= b \exp(X_k/2)\varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0, 1), \quad \forall k \geq 0. \end{aligned} \tag{81}$$

For the parameter values in the model we took  $a = 0.9$ ,  $b = 0.1$ ,  $\sigma_V = 0.5$ , and simulated a sequence of observations from the model. For the parameters of IBPF and LEPF we took  $M = 20$  and  $\theta = 1$ .

Fig. 6(a) shows the ratio (80) for  $N_{MC} = 10\,000$ ,  $\varphi(x) = x$ . The true value of  $\pi_n(\varphi)$  was estimated with standard BPF using  $10^6$  particles. Roughly similar behaviour to that in Fig. 3(d)

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

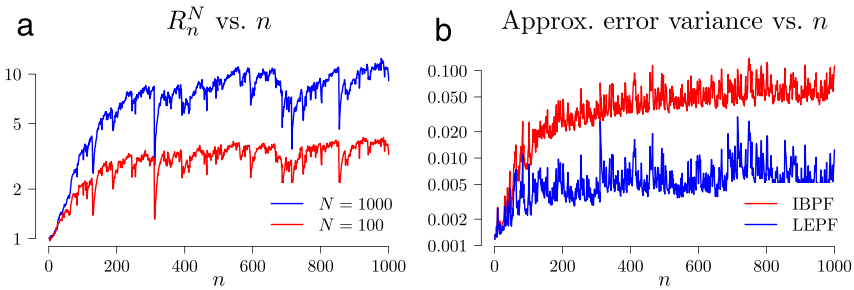


Fig. 6. (a) Approximate relative variance. (b) Approximate error variance for  $N = 1000$ .

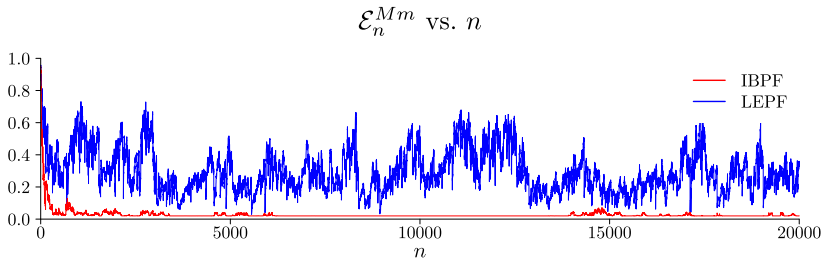


Fig. 7.  $\mathcal{E}_n^{Mm}$  vs.  $n$  with  $M = 20$  and  $m = 50$ .

can be observed, although of course for the stochastic volatility model we are not able to evaluate  $R_n$ . Fig. 6(b) shows estimated mean square errors for IBPF and LEPF, proportional to the numerator and denominator in (80), respectively.

Fig. 7 shows  $\mathcal{E}_n^{Mm}$  against  $n$  with  $M = 20$  and  $m = 50$  for a single run of each algorithm over  $2 \times 10^4$  time steps, for the stochastic volatility model (81). For both the LEPF and the IBPF,  $\mathcal{E}_n^{Mm}$  never goes below  $1/m = 0.02$ , in accordance with (57), but it is notable that for the IBPF,  $\mathcal{E}_n^{Mm}$  stays quite close to  $1/m = 0.02$ , whereas for LEPF,  $\mathcal{E}_n^{Mm}$  fluctuates around higher values.

### 6.3. Concluding remarks

Although our results establish that the asymptotic variance for the LEPF can grow over time without bound, so that time-uniform convergence at rate  $m^{-1/2}$  does not hold, our numerical experiments indicate that the errors from the LEPF may be substantially smaller than those from the IBPF, and that this difference can become more substantial as the time-horizon grows. The interaction structure of the LEPF therefore has clear benefits. The question of how to maximise these benefits, by considering variants of the LEPF arising from different  $\alpha$  matrices, seems challenging. Outside of the toy model scenario, the formula for the asymptotic variance (61) is rather complicated. However, it can be written in terms of a composition of a sequence of non-negative integral operators. If the observations  $(y_n)$  are treated as a random and stationary sequence, then the sequence of integral operators becomes also random and stationary. In light of this, Oseledec’s theorem or similar results for non-negative integral operators may provide some tools to describe the rate of growth of the asymptotic variance over time.

More extreme modifications to the LEPF and IBPF may allow time-uniform convergence at rate  $m^{-1/2}$  to be achieved. For instance, choosing  $\alpha$  adaptively in a time-varying manner so as to control the effective sample size can provably help to control errors [23]. The price to pay is that



doing so may compromise the communication efficiency of the algorithm on a distributed computing architecture. Another possible approach is to stabilise the performance of the algorithm by artificially regulating the values taken by the weights  $W_n^i$  and thus introduce some bias, but avoid degeneracy and prevent low values of effective sample size. A drawback of this approach is that it would compromise the lack-of-bias properties which validate the use of particle filters within particle MCMC. Rigorous treatment of these ideas is a potential topic for future research.

**Acknowledgements**

This work was supported by the EPSRC through First Grant EP/KO23330/1 and SuStaIn (EP/D063485/1).

**Appendix. Auxiliary proofs**

**Lemma 6.** *The matrices defined in (5) satisfy Assumption (2.4).*

**Proof.** For  $i, j \in \mathbb{Z}$  such that  $|i - j| > \beta := M - 1 + \theta$ , by (5),  $\alpha_\infty^{ij} = 0$  and (4) clearly holds. For  $|i - j| \leq \beta$ , we observe that by (5)

$$\alpha_\infty^{ij} = \alpha_\infty^{i+kM, j+kM}, \quad \forall k \in \mathbb{Z}, \tag{A.1}$$

and provided that  $i + kM, j + kM \in [N]$ , then by (2.2) we also have

$$\alpha_N^{i \bmod N, j \bmod N} = \alpha_N^{(i+kM) \bmod N, (j+kM) \bmod N} = \alpha_N^{i+kM, j+kM}. \tag{A.2}$$

So, to complete the verification of (4) we shall, for each  $i, j \in \mathbb{Z}$  such that  $|i - j| \leq \beta$ , find  $k$  such that  $i + kM, j + kM \in [N]$  and check that  $\alpha_N^{i+kM, j+kM} = \alpha_\infty^{i+kM, j+kM}$ .

First consider the case  $j \geq i$ , and set  $k = -\lfloor (i - 1)/M \rfloor$ . In this case, by using (5) together with the assumptions that  $Mm \geq 2\beta + 1$ , and  $j - i \leq \beta$  we have

$$\begin{aligned} \alpha_\infty^{i+kM, j+kM} &= M^{-1} \mathbb{I}[\lfloor (j + kM - \theta - 1)/M \rfloor = 0] \\ &= M^{-1} \mathbb{I}[j + kM - \theta \in [M]] \\ &= M^{-1} \mathbb{I}[\lfloor ((j + kM - \theta) \bmod N - 1)/M \rfloor = 0] \\ &= \alpha_N^{i+kM, j+kM}, \end{aligned}$$

and moreover  $i + kM, j + kM \in [N]$  holds and thus by (A.1), (A.2) we have (4) for  $j \geq i$ .

For the case  $i > j$ , we can take  $k = -\lfloor (i - 1)/M \rfloor + m - 1$ , for which  $i + kM, j + kM \in [N]$ , and similarly as above

$$\alpha_\infty^{i+kM, j+kM} = M^{-1} \mathbb{I}[(m - 1)M \leq j + kM - \theta - 1 \leq mM - 1] = \alpha_N^{i+kM, j+kM},$$

from which we conclude, by (A.1), (A.2) that (2.4) holds for all  $i, j \in \mathbb{Z}$ .  $\square$

**Proof of Lemma 5.** Let  $V$  be distributed as in (78), thus by (70)  $V$  has the same distribution as the number of zero increments in  $D$ . Our strategy is to construct a collection of sequences  $\{D^p\}_{0 \leq p \leq V}$  and random variables  $\{B^p\}_{0 \leq p \leq V}$  such that  $D^V$  and  $B^V$  have the same distributions as  $D$  and  $B$ , respectively. The construction is done in a manner that allows us to identify explicitly the distribution of  $B^V$  and hence the distribution of  $B$ .

To start, take a sequence  $D^0 := (D_k^0)_{0 \leq k \leq n-V}$  where  $D_0^0 = 0$  and the increments  $(D_k^0 - D_{k-1}^0)_{1 \leq k \leq n-V}$  are i.i.d. with common distribution  $\delta_{-1}/2 + \delta_1/2$ . We then define sequences  $D^p := (D_k^p)_{0 \leq k \leq n-V+p}$  for  $1 \leq p \leq V$  recursively

$$D^p := (D_0^{p-1}, \dots, D_{K_p}^{p-1}, D_{K_p}^{p-1}, \dots, D_{n-V+p-1}^{p-1}), \tag{A.3}$$

where  $K_p$  is a uniform random variable on the set  $\{0, \dots, n - V + p - 1\}$ , for all  $1 \leq p \leq V$ .

By this construction,  $D^0$  is of length  $n - V + 1$  and has only non-zero increments.  $D^1$  is of length  $n - V + 2$  and has exactly one zero increment at a uniformly random location. Finally,  $D^V$  is of length  $n + 1$  and has exactly  $V$  zero increments at uniformly random locations and hence can be checked to have the same distribution as  $D$ .

The random variables  $\{B^p\}_{0 \leq p \leq V}$  are defined as

$$B^p := \sum_{k=1}^{n-V+p} \mathbb{I}[D_{k-1}^p = 0] \mathbb{I}[D_k^p = 0], \quad \forall 0 \leq p \leq V, \tag{A.4}$$

for which we have, by (A.3), the recursive expression

$$B^p = B^{p-1} + \mathbb{I}[D_{K_p}^{p-1} = 0], \quad \forall 0 < p \leq V. \tag{A.5}$$

By the definition of  $K_p$ ,  $\mathbb{I}[D_{K_p}^{p-1} = 0]$  in (A.5) is a Bernoulli random variable with success probability

$$\frac{\sum_{k=0}^{n-V+p-1} \mathbb{I}[D_k^{p-1} = 0]}{n - V + p} = \frac{1 + B^{p-1} + \sum_{k=1}^{n-V+p-1} \mathbb{I}[D_{k-1}^{p-1} \neq 0] \mathbb{I}[D_k^{p-1} = 0]}{n - V + p}, \tag{A.6}$$

and if we define  $S := \sum_{k=1}^{n-V} \mathbb{I}[D_k^0 = 0]$  when  $V < n$ , and  $S := 0$  when  $V = n$ . This means that  $S$  is the number of times zero occurs in the sequence  $D^0$ , excluding the first element, then by induction, one can check that for all  $0 \leq p \leq V$ ,

$$S = \sum_{k=1}^{n-V+p} \mathbb{I}[D_{k-1}^p \neq 0] \mathbb{I}[D_k^p = 0],$$

and hence by (A.6)

$$\mathbb{I}[D_{K_p}^{p-1} = 0] \sim \text{Bernoulli}\left(\frac{1 + S + B^{p-1}}{n - V + p}\right). \tag{A.7}$$

The key observation is that by (A.5) and (A.7), the sequence  $(\mathbb{I}[D_{K_p}^{p-1} = 0])_{0 < p \leq V}$ , is distributed according to a Pólya’s urn model, for which we have readily (see, e.g. [14])

$$\sum_{p=1}^V \mathbb{I}[D_{K_p}^{p-1} = 0] \sim \text{Beta-Binomial}(V, S + 1, n - V - S), \tag{A.8}$$

with the convention that  $\text{Beta-Binomial}(k, 1, 0)$  corresponds to point mass  $\delta_k$  for any  $k > 0$ .

To conclude the proof we observe that because the increments of  $D^0$  are non-zero, we have  $B^0 = 0$  and hence, by (A.5),  $B^V = \sum_{p=1}^V \mathbb{I}[D_{K_p}^{p-1} = 0]$  and therefore, by (A.8), if we set  $\bar{B} = B^V$ , it remains to point out that because  $D^0$  is a simple random walk, we know by

Please cite this article in press as: K. Heine, N. Whiteley, Fluctuations, stability and instability of a distributed particle filter with local exchange, Stochastic Processes and their Applications (2016), <http://dx.doi.org/10.1016/j.spa.2016.11.003>

[10, Theorem 2], that  $S$  is distributed as described in (78). Finally, since  $D^V$  has the same distribution as  $D$ , then by (A.4) and (74),  $B^V$  must have the same distribution as  $B$ , concluding the proof.  $\square$

## References

- [1] C. Andrieu, A. Doucet, H. Holenstein, Particle Markov chain Monte Carlo methods, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 72 (2010) 269–342.
- [2] M. Bolić, P.M. Djurić, S. Hong, Resampling algorithms and architectures for distributed particle filters, *IEEE Trans. Signal Process.* 53 (2005) 2442–2450.
- [3] F. Cérou, P. Del Moral, A. Guyader, A nonasymptotic theorem for unnormalized Feynman–Kac particle models, *Ann. Inst. Henri Poincaré (B) Probab. Stat.* 47 (2011) 629–649.
- [4] N. Chopin, Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference, *Ann. Statist.* 32 (2004) 2385–2411.
- [5] D. Crisan, O. Obanubi, Particle filters with random resampling times, *Stochastic Process. Appl.* 122 (2012) 1332–1368.
- [6] P. Del Moral, Feynman–Kac Formulae, in: *Genealogical and Interacting Particle Systems with Applications*, in: *Probab. and its Appl.*, Springer Verlag, New York, 2004.
- [7] P. Del Moral, A. Doucet, A. Jasra, On adaptive resampling strategies for sequential Monte Carlo methods, *Bernoulli* 18 (2012) 252–278.
- [8] R. Douc, E. Moulines, J. Olsson, Long-term stability of sequential Monte Carlo methods under verifiable conditions, *Ann. Appl. Probab.* 24 (2014) 1767–1802.
- [9] B. Favetto, On the asymptotic variance in the central limit theorem for particle filters, *ESAIM: Probab. Stat.* 16 (2012) 151–164.
- [10] W. Feller, The numbers of zeros and of changes of sign in a symmetric random walk, *L'Enseign. Math.* 3 (1957) 229–235.
- [11] N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proc. F* 140 (1993) 107–113.
- [12] P. Hall, C.C. Heyde, *Martingale Limit Theory and Its Applications*, in: *Probab. and Math. Stat.*, Academic Press, 1980.
- [13] K. Heine, N. Whiteley, A.T. Cemgil, H. Gültaş, Butterfly resampling: Asymptotics for particle filters with constrained interactions, 2014. [arXiv:1411.5876](https://arxiv.org/abs/1411.5876).
- [14] N.L. Johnson, S. Kotz, *Urn Models and their Application*, John Wiley & Sons, 1977.
- [15] H.R. Künsch, Recursive Monte Carlo filters: algorithms and theoretical analysis, *Ann. Statist.* 33 (2005) 1983–2021.
- [16] J. Míguez, Analysis of parallelizable resampling algorithms for particle filtering, *Signal Process.* 87 (2007) 3155–3174.
- [17] J. Míguez, On the uniform asymptotic convergence of a distributed particle filter, in: *Proc. of the 8th Sens. Array and Multichannel Signal Process. Workshop, SAM 2014*, 2014, pp. 241–244.
- [18] J. Míguez, M.A. Vázquez, A proof of uniform convergence over time for a distributed particle filter, *Signal Process.* 122 (2016) 152–163.
- [19] R.J. Serfling, *Approximation Theorems of Mathematical Statistics*, in: *Wiley Ser. in Probab. and Math. Stat.*, John Wiley & Sons, 1980.
- [20] A.N. Shiryaev, *Probability*, in: *Grad. Texts in Math*, vol. 95, Springer-Verlag, 1984.
- [21] J.G. Skellam, A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 10 (1948) 257–261.
- [22] N. Whiteley, Stability properties of some particle filters, *Ann. Appl. Probab.* 23 (2013) 2500–2537.
- [23] N. Whiteley, A. Lee, K. Heine, On the role of interaction in sequential Monte Carlo algorithms, *Bernoulli* 22 (1) (2016) 494–529.