On parallel implementation of sequential Monte Carlo methods: the island particle model

Christelle Vergé · Cyrille Dubarry · Pierre Del Moral · Eric Moulines

Received: 14 June 2013 / Accepted: 16 October 2013 / Published online: 5 November 2013
© Springer Science+Business Media New York 2013

Abstract The approximation of the Feynman-Kac semigroups by systems of interacting particles is a very active research field, with applications in many different areas. In this paper, we study the parallelization of such approximations. The total population of particles is divided into sub-populations, referred to as islands. The particles within each island follow the usual selection/mutation dynamics. We show that the evolution of each island is also driven by a Feynman-Kac semigroup, whose transition and potential can be explicitly related to ones of the original problem. Therefore, the same genetic type approximation of the Feynman-Kac semi-group may be used at the island level; each island might undergo selection/mutation algorithm. We investigate the impact of the population size within each island and the number of islands, and study different type of interactions. We find conditions under which introducing interactions between islands is beneficial. The theoretical results are supported by some Monte Carlo experiments.

Keywords Particle approximation of Feynman-Kac flow · Island models · Parallel implementation

1 Introduction

Numerical approximation of Feynman-Kac semigroups by systems of interacting particles is a very active field of researches. Interacting particle systems are increasingly used to sample complex high dimensional distributions in a wide range of applications including nonlinear filtering, data assimilation problems, rare event sampling, hidden Markov chain parameter estimation, stochastic control problems, financial mathematics; see for example Doucet and Freitas (2001), Chopin (2002), Del Moral (2004), Cappé and Moulines (2005), Del Moral et al. (2012b) and the references therein.

Let \((E_n, \mathcal{E}_n)_{n \geq 0}\) be a sequence of measurable spaces. Denote by \(B_0(E_n)\) the Banach space of all bounded and measurable real valued functions \(f\) on \(E_n\), equipped with the uniform norm. Let \((g_n)_{n \in \mathbb{N}}\) be a sequence of measurable potential functions, \(g_n : E_n \to \mathbb{R}^+\). Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. In the sequel, all the processes are defined on this probability space. Let \((X_n)_{n \in \mathbb{N}}\) be a non-homogeneous Markov chain on the sequence of state-spaces \((E_n)_{n \in \mathbb{N}}\) with initial distribution \(\eta_0\) on \((E_0, \mathcal{E}_0)\) and Markov kernels \((M_n)_{n \in \mathbb{N}}\).\(^1\) We associate to the sequences of poten-
tial functions \((g_n)_{n \in \mathbb{N}}\) and Markov kernels \((M_n)_{n \in \mathbb{N}}\) the sequence of Feynman-Kac measures, defined for all \(n \geq 1\) and for any \(f_n \in \mathcal{B}_b(\mathbb{E}_n)\) by

\[
\begin{align*}
\eta_n(f_n) & \overset{\text{def}}{=} \gamma_n(f_n)/\gamma_n(1), \\
\gamma_n(f_n) & \overset{\text{def}}{=} \mathbb{E}\left[ f_n(X_n) \prod_{0 \leq p < n} g_p(X_p) \right] \\
& = \int \gamma_0(dx_0) \left[ \prod_{0 \leq p < n} g_p(x_p)M_{p+1}(x_p, dx_{p+1}) \right] \\
& \times f_n(x_n), \\
\end{align*}
\]

where we have set by convention \(\eta_0(f_0) = \gamma_0(f_0) \overset{\text{def}}{=} \mathbb{E}[f_0(X_0)].\)

The sequences of distributions \((\eta_n)_{n \geq 0}\) and their unnormalized versions \((\gamma_n)_{n \geq 0}\) are approximated sequentially using interacting particle systems (IPS). Such particle approximations are often referred to as sequential Monte Carlo (SMC) methods. The IPS consists in approximating for each \(n \in \mathbb{N}\) the probability \(\eta_n\) by a set of \(N_1\) particles \((X_n^i)_{i=1}^{N_1}\) which are generated recursively. Typically, the update of the particles may be decomposed into a mutation and a selection step. For example, the bootstrap algorithm proceeds as follows. In the selection step the particles are first sampled with weights proportional to the potential functions. In the mutation step, a new generation of particles \((X_{n+1}^i)_{i=1}^{N_1}\) is generated from the selected particles using the kernel \(M_{n+1}\). The asymptotic behavior of such particle approximation is now well understood (see Del Moral 2004 and Del Moral et al. 2012b).

Particle filtering is computationally an intensive method. Parallel computations provides an appealing solution to tackle this issue (see Durham and Geweke 2011 and the references therein for an in-depth description of parallelization of Bayesian computations). Moreover, parallel SMC has econometric applications. For instance, it is used in the Matlab package DeCo (Density Combination) to combine density forecasts with time-varying weights and different choices of scoring rule (see Casarin et al. 2013 for further details). The basic idea to implement interacting particle system in parallel goes as follows: instead of considering a single large batch of \(N = N_1N_2\) particles, the population is divided into \(N_2\) batches of \(N_1\) particles. These batches are referred in the sequel to as islands. The terminology island is borrowed from dynamic populations theory (like the genetic type interacting particle model). The particles within each island are selected and mutate, as described above. We might also introduce interactions among islands.

In this paper we introduce the island particle models. As we will see below, we may cast the island particle model in the Feynman-Kac framework, with appropriately defined potentials and transition kernels. The key observation is that the marginal distribution of the island Feynman-Kac model w.r.t. any individual coincide with (3). This interpretation allows to use the interacting particle model at the island level. The general idea of interacting SMC is already mentioned in Jasra et al. (2008), but in this contribution each island targets a different distribution.

The study of the island particle model gives rise to several challenging theoretical questions. In this paper, we investigate the impact of the number of particles in each island \(N_1\) compared to the number of islands \(N_2\) for a given total number of particles \(N \overset{\text{def}}{=} N_1N_2\), for an algorithm where the bootstrap mechanism is used both within and between the islands. We focus on the asymptotic bias and variance when both \(N_1\) and \(N_2\) goes to infinity. Fluctuation theorem and non-asymptotic results will be present in a forthcoming paper. We also investigate when and why introducing interactions at the island level improves the accuracy of the particle approximation. Intuitively, the trade-off might be understood as follows. When the \(N_2\) islands are run independently, the bias induced in each island only depends on their population size \(N_1\); when \(N_1\) is small compared to the total number \(N\), the bias will be large (and is of course not reduced by averaging across the islands). To reduce the bias, introducing an interaction between the islands is beneficial. However, this interaction increases the variance, due to the selection step. If we consider the mean squared error, the interaction between islands is beneficial when the improvement associated to the bias correction is not offset by the variance increase. When the number of particles \(N_1\) within each island is small and the number of islands \(N_2\) is large, then the interaction is typically beneficial. On the contrary, when \(N_2 \ll N_1\), the interaction between islands may increase the mean squared error. We then propose a method whose mean-squared error can be shown to be lower than the independent island model for any choice of \(N_1\) and \(N_2\). In this algorithm, the islands are only sampled when required, based on a generalization of the effective sample size at the island level.

The paper is organized as follows. In Sect. 2 the interacting particle approximation of the Feynman-Kac model is first reviewed. The island Feynman-Kac model is then introduced. We first investigate the double bootstrap algorithm, in which selection and mutation are applied at each iteration within and across the islands. The asymptotic bias and variance of this algorithm is presented in Sect. 3. The Feynman-Kac interpretation of the island model leads to several interacting island algorithms, based on different approximations of Feynman-Kac flows. Some of these are introduced and analyzed in Sect. 4. Some numerical experiments are
reported to support our findings and illustrate the impact of the numbers of islands and particles within each island in Sect. 5.

2 Algorithm derivation

In this section, we introduce the island particle model. We first briefly recall the bootstrap approximation of Feynman-Kac measures.

According to the definitions (1) and (3) of the sequences of the Feynman-Kac measures \((\eta_n)_{n \in \mathbb{N}}\) and \((\gamma_n)_{n \in \mathbb{N}}\), for all \(f_{n+1} \in \mathcal{B}_b(\mathbb{E}_{n+1})\) we get

\[
\gamma_{n+1}(f_{n+1}) = \eta_{n+1}(f_{n+1})\gamma_n(1),
\]

and since,

\[
\gamma_{n+1}(1) = \gamma_n(g_n) = \eta_n(g_n)\gamma_n(1),
\]

an easy induction shows that

\[
\gamma_{n+1}(1) = \prod_{0 \leq p < n+1} \eta_p(g_p)
\]

and then,

\[
\gamma_{n+1}(f_{n+1}) = \eta_{n+1}(f_{n+1}) \prod_{0 \leq p < n+1} \eta_p(g_p) .
\]

Moreover, the sequence \((\eta_n)_{n \in \mathbb{N}}\) satisfy a nonlinear recursive relation. Indeed,

\[
\eta_{n+1}(f_{n+1}) = \frac{\gamma_n(g_nM_{n+1}f_{n+1})}{\gamma_n(g_nM_{n+1})} = \frac{\eta_n(g_nM_{n+1}f_{n+1})}{\eta_n(g_n)} .
\]

Let \(\mathcal{P}(\mathbb{E}_n)\) be the set of probability measures on \(\mathbb{E}_n\). Using the Boltzmann-Gibbs transformation \(\Psi_n : \mathcal{P}(\mathbb{E}_n) \rightarrow \mathcal{P}(\mathbb{E}_n)\), defined for all \(\mu_n \in \mathcal{P}(\mathbb{E}_n)\) by

\[
\Psi_n(\mu_n)(dx_n) \overset{\text{def}}{=} \frac{g_n(x_n)\mu_n(dx_n)}{\mu_n(g_n)},
\]

the recursion (5) may be rewritten as

\[
\eta_{n+1} = \Psi_n(\eta_n)M_{n+1} .
\]

The sequence of probability \((\eta_n)_{n \in \mathbb{N}}\) can be approximated using the bootstrap algorithm. Other approximations can also be considered as well, but we only introduce the bootstrap for notational simplicity. Let \(N_1\) be a positive integer. For any nonnegative integer \(n\) we denote by

\[
(\mathbb{E}_n, \mathcal{E}_n) \overset{\text{def}}{=} \left(\mathbb{E}_n^{N_1} \times \mathbb{E}_n \otimes \cdots \otimes \mathbb{E}_n \otimes \mathbb{E}_n^{N_1}\right),
\]

the product space (the dependence of \(\mathbb{E}_n\) and \(\mathcal{E}_n\) in \(N_1\) is implicit). Thereafter, we omit to write the \(\sigma\)-field \(\mathcal{E}_n\) when there will be no confusion. We define the Markov kernel

\[
M_{n+1}(x_n, dx_{n+1}) \overset{\text{def}}{=} \prod_{1 \leq i \leq N_1} \sum_{j=1}^{N_1} \frac{g_n(x^j_n)}{\sum_{k=1}^{N_1} g_n(x^k_n)} M_{n+1}(x^j_n, dx_{n+1}) .
\]

In other words, this transition can be interpreted as follows:

- In the selection step, the components of the vector \(x_n\) are selected with probabilities proportional to their potential \(\{g_n(x^i_n)\}_{i=1}^{N_1}\);
- In the mutation step, the selected coordinates move conditionally independently to new positions using the Markov kernel \(M_{n+1}\).

Let us introduce the particles and their evolution. Define by \((X_n)_{n \geq 0}\) the Markov chain where for each \(n \in \mathbb{N}\),

\[
X_n = (X_1^n, \ldots, X_{N_1}^n) \in \mathbb{E}_n, \quad n \geq 0
\]

with initial distribution \(\eta_0 \overset{\text{def}}{=} \eta_0 \otimes \cdots \otimes \eta_0\) and transition kernel \(M_{n+1}\). Denote by \(m_{N_1}\) the empirical measure on \(\mathbb{E}_n\), defined as the kernel on \(\mathbb{E}_n \times \mathbb{E}_n\) by

\[
m_{N_1}(x_n, dz_n) \overset{\text{def}}{=} \frac{1}{N_1} \sum_{i=1}^{N_1} \delta_{x^i_n}(dz_n),
\]

where \(\delta_{x_n}\) is the Dirac mass at \(x_n \in \mathbb{E}_n\). Equation (4) suggests the following \(N_1\)-particle approximations of the measures \(\eta_n\) and \(\gamma_n\) respectively defined for \(f_n \in B_b(\mathbb{E}_n)\) by

\[
\eta_{N_1}^n(f_n) \overset{\text{def}}{=} m_{N_1}^n f_n(X_n) = \frac{1}{N_1} \sum_{i=1}^{N_1} f_n(X^i_n) .
\]

\[
\gamma_{N_1}^n(f_n) \overset{\text{def}}{=} \eta_{N_1}^n(f_n) \prod_{0 \leq p < n+1} \eta_{N_1}^n(g_p) = \eta_{N_1}^n(f_n) \gamma_{N_1}^n(1) .
\]

For \(x_n = (x_1^n, \ldots, x_{N_1}^n) \in \mathbb{E}_n\), define the potential function

\[
g_n(x_n) \overset{\text{def}}{=} m_{N_1}^n g_n(x_n) = \frac{1}{N_1} \sum_{i=1}^{N_1} g_n(x_i^n) .
\]

The sequences of transition kernels \((M_n)_{n \in \mathbb{N}}\) and potential functions \((g_n)_{n \in \mathbb{N}}\) given by (9) and (13), respectively, define the Feynman-Kac process. The associated sequences of Feynman-Kac measures are defined, for each \(f_n \in B_b(\mathbb{E}_n)\), by the following recursions

\[
\eta_0(f_0) \overset{\text{def}}{=} \eta_0(f_0) = \mathbb{E}[f_0(X_0)] ,
\]

\[
\eta_n(f_n) \overset{\text{def}}{=} \gamma_{N_1}(f_n) / \gamma_n(1) , \quad \text{for all } n \geq 1 .
\]
\[ y_n(f_n) = \mathbb{E}\left[ f_n(X_n) \prod_{0 \leq p < n} g_p(X_p) \right], \]

for all \( n \geq 1. \) \( \tag{16} \]

where \( (X_n)_{n \geq 0} \) is a Markov chain with initial distribution \( \gamma_0 \) and transition kernel \( M_n. \) The key result, justifying the introduction of the island particle models, is the following theorem which links \((\eta_n, \gamma_n)_{n \geq 0}\) and \((\eta_n, \gamma_n)_{n \geq 0}. \)

**Theorem 1** For any \( f_n \in \mathcal{B}_0(\mathbb{E}_n) \) of the form \( f_n(x_n) = N_1^{-1} \sum_{i=1}^{N_1} f_n(x_n^i) \) where \( f_n \in \mathcal{B}_0(\mathbb{E}_n), \)

\[ y_n(f_n) = y_n(f_n) \quad \text{and} \quad \eta_n(f_n) = \eta_n(f_n). \] \( \tag{17} \]

**Proof** See Sect. 6.1.

The Feynman-Kac model \( (\eta_n, \gamma_n)_{n \geq 0} \) can be approximated by an interacting particle system at the island level. We first describe the double bootstrap algorithm where the bootstrap is also applied across the islands (this algorithm shares some similarities with Chopin et al. 2013). This is only one of the many possible algorithms that can be derived from this interpretation of the Feynman-Kac model at the island level; see Sect. 4 for other approximations.

Define by \( \mathcal{P}(\mathbb{E}_n) \) the set of probability measures on \( \mathbb{E}_n. \) One can easily check that the sequence of measures \((\eta_n)_{n \geq 0}\) satisfies the following recursion

\[ \eta_{n+1} = \Psi_n(\eta_n) M_{n+1}, \] \( \tag{18} \]

where \( \Psi_n : \mathcal{P}(\mathbb{E}_n) \to \mathcal{P}(\mathbb{E}_n) \) is the Boltzmann-Gibbs transformation defined for any \( \mu_n \in \mathcal{P}(\mathbb{E}_n) \) by

\[ \Psi_n(\mu_n)(dx_n) = \frac{\mathbb{E}_n(x_n^i) \mu_n(dx_n)}{\mu_n(g_n)}. \]

Let \( N_2 \) be a positive integer. We define the Markov kernel \( M_{n+1} \) from \((\mathbb{E}_n^{N_2}, \mathcal{E}_n^{\otimes N_2}) \) to \((\mathbb{E}_{n+1}^{N_2}, \mathcal{E}_{n+1}^{\otimes N_2}) \) as follows: for any \((x_1^1, \ldots, x_{N_2}^1) \in \mathbb{E}_n^{N_2} \) and \((x_{n+1}^1, \ldots, x_{n+1}^{N_2}) \in \mathbb{E}_{n+1}^{N_2}, \) we put

\[ M_{n+1}(x_n^{1, \ldots, x_n^{N_2}}, d(x_{n+1}^{1, \ldots, x_{n+1}^{N_2}})) = \prod_{1 \leq i \leq N_2} \sum_{j=1}^{N_2} \frac{g_n(x_j^i)}{\sum_{k=1}^{N_2} g_n(x_k^i)} M_n(x_n^i, dx_{n+1}^i). \] \( \tag{19} \]

For each \( n \in \mathbb{N}, \) \((X_1^1, \ldots, X_1^{N_2}) \in \mathbb{E}_n^{N_2} \) is a population of \( N_2 \) interacting islands each with \( N_1 \) individuals. The process \(((X_1^1, \ldots, X_1^{N_2}))_{n \geq 0} \) is a Markov chain with the transition kernel \( (M_{n+1})_{n \geq 0}. \)

In this interpretation, the \( N_2 \)-particle model defined above can be seen as an interacting particle approximation of the island Feynman-Kac measures \((\eta_n, y_n)_{n \geq 0}. \) The transition \( M_{n+1} \) can be interpreted as follows:

---

**Algorithm 1** Bootstrap within bootstrap island filter

1. Initialization:
2. for \( i \) from 1 to \( N_2 \) do
3. Sample \( N_1 \) independent random variables \( X_i^0 = (X_i^0)^{N_1}_{j=1} \) from \( \eta_0. \)
4. end for
5. for \( p \) from 0 to \( n - 1 \) do
6. Selection step between islands:
7. Sample \( I_p = (I_p^j)_{j=1}^{N_1} \) multinomially with probability proportional to \( (\frac{1}{N_1} \sum_{j=1}^{N_1} g_p(X_p^{i,j}))_{j=1}^{N_1}. \)
8. Island mutation step:
9. for \( i \) from 1 to \( N_2 \) do
10. Particle selection within each island:
11. Sample \( J_i^j = (J_i^j)_{j=1}^{N_1} \) multinomially with probability proportional to \( (g_p(X_p^{i,j}))_{j=1}^{N_1}. \)
12. Particle mutation:
13. For \( 1 \leq j \leq N_1 \), sample conditionally independently \( X_i^{j,1} \) from the Markov kernel \( M_{p+1}(X_i^{j,1}, \cdot). \)
14. end for
15. end for
16. Approximate \( \eta_n(f_n) \) by \( \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} f_n(X_i^{j,1}). \)

---

In the selection step, we sample randomly \( N_2 \) islands among the current islands \((X_i^0)_{1 \leq i \leq N_2} \in \mathbb{E}_n^{N_2} \) with probability proportional to the empirical mean of the potentials in each island \( g_n(X_i^n) = N_1^{-1} \sum_{j=1}^{N_1} g_n(X_i^{i,j}), \quad 1 \leq i \leq N_2. \)

In the mutation transition, the selected islands are independently updated using the Markov transition \( M_{n+1}. \)

Also observe that for \( N_1 = 1, \) every island has a single particle. In this situation, the island Feynman-Kac model coincides with the \( N_2 \)-particle model associated with the Feynman-Kac measures \( \eta_n. \)

Denote by \( m_{N_2} \) the empirical measure defined for any \( f_n \in \mathcal{B}_0(\mathbb{E}_n) \) and \((x_1^1, \ldots, x_{N_2}^1) \in \mathbb{E}_n^{N_2} \) by

\[ m_{N_2}(f_n) \equiv \frac{1}{N_2} \sum_{i=1}^{N_2} f_n(x_i^1). \]

The \( N_2 \)-particle approximations of the measures \( \eta_n \) and \( y_n \) are defined for any \( f_n \in \mathcal{B}_0(\mathbb{E}_n) \) by

\[ \eta^{N_2}_n(f_n) \equiv m_{N_2}(f_n(X_1^1, \ldots, X_{N_2}^1)), \] \( \tag{20} \]
\[ \gamma^{N_2}_n(f_n) \equiv \eta^{N_2}_n(f_n) \prod_{0 \leq p < n} \eta^{N_2}_p(g_p) = \eta^{N_2}_n(f_n) \gamma^{N_2}_n(1). \] \( \tag{21} \)
3 Asymptotic analysis of the double bootstrap algorithm

The bootstrap particle approximation of the Feynman-Kac semigroup can be studied using the techniques introduced in Del Moral (2004) and further developed in Del Moral et al. (2012b). For \( \ell \in \mathbb{N} \), consider the finite kernel \( Q_{\ell+1} \) from \((E, \mathcal{E})\) into \((E, \mathcal{E})\) for all \( x_\ell \in E \) by

\[
Q_{\ell+1}(x_\ell, dx_\ell+1) \overset{\text{def}}{=} g_\ell(x_\ell)M_{\ell+1}(x_\ell, dx_\ell+1).
\]

(22)

For \( p < n \), define by \( Q_{p,n} \) the finite kernel from \((E_p, \mathcal{E}_p)\) into \((E_n, \mathcal{E}_n)\) as the following product

\[
Q_{p,n} \overset{\text{def}}{=} Q_{p+1}Q_{p+2} \cdots Q_n,
\]

(23)

and set by convention \( Q_{n,n} \overset{\text{def}}{=} I_n \) where \( I_n \) is the identity kernel on \((E_n, \mathcal{E}_n)\). With this definition, the linear semigroup associated with the sequence of unnormalized Feynman-Kac measures \((\gamma_n)_{n \in \mathbb{N}}\) may be equivalently expressed as follows

\[
\gamma_n = \gamma_p Q_{p,n}.
\]

(24)

For any \( x_p \in E_p, A_n \in \mathcal{E}_n, Q_{p,n} \) may be written as the following conditional expectation,

\[
Q_{p,n}(x_p, A_n) = \mathbb{E} \left[ \mathbb{I}_{A_n}(X_n) \prod_{p \leq q < n} g_q(X_q) \bigg| X_p = x_p \right].
\]

where \((X_n)_{n \geq 0}\) is the non-homogeneous Markov chain on the sequence of state-spaces \((E_n, \mathcal{E}_n)_{n \geq 0}\) with initial distribution \( \eta_0 \) and Markov kernels \((M_n)_{n \geq 1}\).

According to (1), \( \eta_n = \gamma_n/\gamma_n(1) \) implies that \( \eta_n = \gamma_p Q_{p,n}/\gamma_p Q_{p,n}(1) \). Denote by \( \Phi_{n+1} \) the mapping from \( \mathcal{P}(E_n) \) to \( \mathcal{P}(E_{n+1}) \) given, for any \( \mu_n \in \mathcal{P}(E_n) \) by

\[
\Phi_{n+1}(\mu_n) \overset{\text{def}}{=} \sum_{q \leq n} \mu_n Q_{n+1}(x_q, dx_q) = \frac{\mu_n Q_{n+1}}{\mu_n Q_{n+1}(1)}.
\]

(25)

Since \( \eta_p = \gamma_p/\gamma_p(1) \), these relations may be equivalently rewritten as

\[
\eta_n = \frac{\eta_p Q_{p,n}}{\eta_p Q_{p,n}(1)} = \Phi_{p,n}(\eta_p),
\]

(26)

where \( \Phi_{p,n} = \Phi_p \circ \Phi_{p-1} \circ \cdots \circ \Phi_1 \) is the nonlinear semigroup associated to the normalized Feynman-Kac measures \((\eta_n)_{n \geq 0}\). This nonlinear semigroup may be associated to the potential kernels

\[
P_{p,n} \overset{\text{def}}{=} \frac{Q_{p,n}}{\eta_p Q_{p,n}(1)} = \frac{\gamma_p(1)}{\gamma_n(1)} Q_{p,n},
\]

(27)

and therefore

\[
\eta_n = \eta_p P_{p,n}.
\]

(28)

For \( \ell \in \mathbb{N} \), consider the finite kernel \( Q_{\ell+1} \) from \((E, \mathcal{E})\) into \((E, \mathcal{E})\) for any \( x_\ell \in E \) by

\[
Q_{\ell+1}(x_\ell, dx_{\ell+1}) \overset{\text{def}}{=} g_\ell(x_\ell)M_{\ell+1}(x_\ell, dx_{\ell+1}),
\]

where \( M_\ell \) is defined in (9) and \( g_\ell \) in (13). For \( p \leq n \), define by \( Q_{p,n} \) the finite kernel from \((E_p, \mathcal{E}_p)\) into \((E_n, \mathcal{E}_n)\) by the equation \( Q_{p,n} \overset{\text{def}}{=} Q_{p+1}Q_{p+2} \cdots Q_n \). Note that, for any \( x_p \in E_p, A_n \in \mathcal{E}_n \),

\[
Q_{p,n}(x_p, A_n) = \mathbb{E} \left[ \mathbb{I}_{A_n}(X_n) \prod_{p \leq q < n} g_q(X_q) \bigg| X_p = x_p \right],
\]

where \((X_n)_{n \geq 0}\) is the island Markov chain defined in (10). With this notation, we may rewrite (14) as \( \gamma_n = \gamma_p Q_{p,n}/\gamma_p Q_{p,n}(1) \). According to (15), \( \eta_n = \gamma_n/\gamma_n(1) \) implies that \( \eta_n = \gamma_p Q_{p,n}/\gamma_p Q_{p,n}(1) \), and then

\[
\eta_n = \frac{\eta_p Q_{p,n}}{\eta_p Q_{p,n}(1)} = \eta_p P_{p,n},
\]

where \( P_{p,n} \) are given by

\[
P_{p,n} = \frac{Q_{p,n}}{\eta_p Q_{p,n}(1)} = \frac{\gamma_p(1)}{\gamma_n(1)} Q_{p,n}.
\]

(29)

To analyse the fluctuation of the interacting particle approximation \((\eta_nN)_{n \geq 0}\) around their limiting values \((\eta_n)_{n \geq 0}\), we introduced the local sampling errors. We first decompose the difference \( \gamma_nN - \gamma_n \) as follows

\[
\gamma_nN - \gamma_n = \sum_{p=1}^n \left[ \gamma_p N_1 Q_{p,n} - \gamma_p N_1 Q_{p-1,n} \right] + \gamma_0 N_1 Q_{0,n} - \gamma_n.
\]

(29)

For any \( p \geq 1 \), note that

\[
\gamma_p N_1 Q_p = \gamma_p N_1(1)\eta_p N_1 Q_p = \frac{\gamma_p N_1(1)\eta_p N_1(1)}{\gamma_n(1)} \Phi_p(\eta_p N_1(1)) = \frac{\gamma_p N_1(1)\gamma_p N_1(1)}{\gamma_p N_1(1)} \Phi_p(\eta_p N_1(1)) = \gamma_p N_1(1)\Phi_p(\eta_p N_1(1)),
\]

and therefore

\[
\eta_n = \eta_p P_{p,n}.
\]

Plugging in this relation in the local error yields to
\(\gamma_p^{N_1} Q_{p,n} - \gamma_p^{N_1} Q_{p-1,n}\)
\(= \gamma_p^{N_1} Q_{p,n} - \gamma_p^{N_1} Q_{p} Q_{p,n}\)
\(= (\gamma_p^{N_1} - \gamma_p^{N_1}(1)\Phi_p^{(N_1-1)}) Q_{p,n}\)
\(= \gamma_p^{N_1}(1)(\eta_p^{N_1} - \Phi_p^{(N_1-1)}) Q_{p,n},\)

which, together with (29), imply that,

\[W_n^{\gamma, N_1} \overset{\text{def}}{=} \sqrt{N_1} \left[ \gamma_n^{N_1} - \gamma_n \right] = \sum_{p=0}^{n} \gamma_p^{N_1}(1)W_p^{\gamma, N_1} Q_{p,n}. \tag{30}\]

where the local errors \((W_p^{\gamma, N_1})_{p \geq 0}\) are defined by

\[W_0^{N_1} = \sqrt{N_1}(\eta_0^{N_1} - \eta_0)\] and \[W_p^{N_1} = \sqrt{N_1}(\eta_p^{N_1} - \Phi_p^{(N_1-1)}), \quad \text{for all } p \geq 1. \tag{31}\]

The following results, adapted from Del Moral (2004, Corollary 9.3.1, pp. 295–298), establishes the convergence of \((W_p^{\gamma, N_1})_{1 \leq p \leq n}\) to centered Gaussian fields.

**Theorem 2** For the bootstrap filter, for any fixed time horizon \(n \geq 1\), the sequence \((W_p^{\gamma, N_1})_{1 \leq p \leq n}\) converges in law, as \(N_1 \to \infty\), to a sequence of \(n\) independent centered Gaussian random fields \((W_p)_{0 \leq p \leq n}\), converging in law, for any bounded function \(f_\rho \in \mathbb{B}_b(\mathbb{E}_\rho)\), and \(1 \leq p \leq n\), by

\[\mathbb{E}[W_p(f_\rho)^2] = \eta_\rho((f_\rho - \eta_\rho f_\rho)^2). \tag{32}\]

Now, consider the sequence of random fields \((W_n^{\eta, N_1})_{n \geq 0}\) defined for any function \(f_n \in \mathbb{B}_b(\mathbb{E}_n)\) by

\[W_n^{\eta, N_1}(f_n) \overset{\text{def}}{=} \sqrt{N_1} \left[ \eta_n^{N_1} - \eta_n \right](f_n) \]
\[= \sqrt{N_1} \eta_n^{N_1}(f_n - \eta_n(f_n)) \]
\[= \sqrt{N_1} \frac{\gamma_n^{N_1}(f_n - \eta_n(f_n))}{\gamma_n^{N_1}(1)}. \tag{33}\]

Using the fact that \(\gamma_n(f_n - \eta_n(f_n)) = 0\) and (30), we may write

\[W_n^{\eta, N_1}(f_n) = \sqrt{N_1} \frac{(\gamma_n^{N_1}(f_n - \eta_n(f_n)))}{\gamma_n^{N_1}(1)} \]
\[= W_n^{\gamma, N_1}(f_n - \eta_n(f_n)) \gamma_n^{N_1}(1). \tag{34}\]

The decomposition (30) and (33), combined with the Slutsky’s lemma, imply the following asymptotic decomposition (which remains valid for more general algorithms than the bootstrap algorithm)

\(\gamma_n^{N_1}(f_n - \eta_n(f_n)) = 0\) and (30), we may write

\[W_n^{\gamma, N_1}(f_n - \eta_n(f_n)) \gamma_n^{N_1}(1).

The asymptotic bias and variance for the single island interacting particle approximation of the sequence of Feynman-Kac measure formulated in the forthcoming theorem result almost immediately from Theorem 2.

**Theorem 3** Assume that the sequence of local errors \((W_p^{\gamma, N_1})_{1 \leq p \leq n}\) converges in law, as \(N_1 \to \infty\), to a sequence of \(n\) independent centered Gaussian random fields \((W_p)_{1 \leq p \leq n}\). Then, the sequence of random fields \((W_n^{\eta, N_1})_{N_1 \geq 0}\) converges in law, as \(N_1 \to \infty\), to the Gaussian random fields \(W_n\) defined for any bounded function \(f_n \in \mathbb{B}_b(\mathbb{E}_n)\) by

\[W_n(f_n) \overset{\text{def}}{=} \sum_{p=0}^{n} \gamma_p^{N_1} W_p(Q_{p,n} f_n) \]
\[= \gamma_n(1) \sum_{p=0}^{n} W_p(P_{p,n}(f_n - \eta_n(f_n))). \tag{36}\]

where \(P_{p,n}\) is defined in (27). The sequence of random fields \((W_n^{\eta, N_1})_{N_1 \geq 0}\) converges in law, as \(N_1 \to \infty\), to the Gaussian random fields \(W_n\) defined for any function \(f_n \in \mathbb{B}_b(\mathbb{E}_n)\) by

\[W_n(f_n) \overset{\text{def}}{=} \sum_{p=0}^{n} W_p(P_{p,n}(f_n - \eta_n(f_n))). \tag{37}\]

The asymptotic bias and variance for the single island interacting particle approximation of the sequence of Feynman-Kac measure formulated in the forthcoming theorem result almost immediately from Theorem 2.

**Theorem 4** Assume that the sequence of local errors \((W_p^{\gamma, N_1})_{1 \leq p \leq n}\) converges in law, as \(N_1 \to \infty\), to a sequence of \(n\) independent centered Gaussian random fields \((W_p)_{1 \leq p \leq n}\). Then, for any time horizon \(n \geq 0\) and any bounded function \(f_n \in \mathbb{B}_b(\mathbb{E}_n)\), we have

\[\lim_{N_1 \to \infty} N_1 \mathbb{E}[\eta_n^{N_1}(f_n) - \eta_n(f_n)] = B_n(f_n), \]
\[\lim_{N_1 \to \infty} N_1 \mathbb{V}[\eta_n^{N_1}(f_n)] = V_n(f_n), \]

with

\[B_n(f_n) \overset{\text{def}}{=} - \sum_{p=0}^{n} \mathbb{E}[W_p(P_{p,n}(f_n - \eta_n(f_n)))]. \tag{38}\]

\[V_n(f_n) \overset{\text{def}}{=} \sum_{p=0}^{n} \mathbb{E}[\{W_p(P_{p,n}(f_n - \eta_n(f_n)))\}^2]. \tag{39}\]

When the bootstrap algorithm is applied, we get the following expressions for \(B_n(f_n)\) and \(V_n(f_n)\) using Theorem 2:

\[B_n(f_n) = - \sum_{p=0}^{n} \eta_p(P_{p,n}(f_n - \eta_n(f_n)) \tag{40}\]
\[V_n(f_n) = \sum_{p=0}^{n} \eta_p(P_{p,n}(f_n - \eta_n(f_n))^2). \tag{41}\]
Proof See Sect. 6.2.

We now compute the bias and the variance for the double bootstrap algorithm. The asymptotic behavior of the bias and the variance is derived in the following theorem using techniques adapted from Del Moral (2004).

**Theorem 5** For the double bootstrap algorithm, for any time horizon \( n \geq 0 \) and any \( f_n \in B_b(\mathbb{E}_n) \), we have

\[
\lim_{N_1 \to \infty} \lim_{N_2 \to \infty} N_1 N_2 \mathbb{E} \left[ \eta_{n}^{N_2} (m_{N_1, f_n}) - \eta_n (f_n) \right] = B_n(f_n) + \tilde{B}_n(f_n),
\]

\[
\lim_{N_1 \to \infty} \lim_{N_2 \to \infty} N_1 N_2 \text{Var}(\eta_{n}^{N_2} (m_{N_1, f_n})) = V_n(f_n) + \tilde{V}_n(f_n),
\]

where \( B_n(f_n) \) and \( V_n(f_n) \) are defined respectively in (38) and in (39), and where \( \tilde{B}_n(f_n) \) and \( \tilde{V}_n(f_n) \) are given by:

\[
\tilde{B}_n(f_n) \overset{\text{def}}{=} - \sum_{\ell=0}^{n} (n-\ell) \mathbb{E} \left[ W_\ell(P_{\ell,n}(1)) \right. \\
\times W_\ell\left( P_{\ell,n}(f_n - \eta_n(f_n)) \right) \\
+ \sum_{\ell=0}^{n} \mathbb{E} \left[ W_\ell \left( \sum_{p=\ell}^{n} P_{\ell,p}(1) \right) \times W_\ell(P_{\ell,n}(f_n - \eta_n(f_n))) \right],
\]

\[
\tilde{V}_n(f_n) \overset{\text{def}}{=} \sum_{\ell=0}^{n} (n-\ell) \mathbb{E} \left[ W_\ell(P_{\ell,n}(f_n - \eta_n(f_n)))^2 \right].
\]

When the bootstrap algorithm is applied, we get the following expressions for \( \tilde{B}_n(f_n) \) and \( \tilde{V}_n(f_n) \) using Theorem 2:

\[
\tilde{B}_n(f_n) = - \sum_{\ell=0}^{n} (n-\ell) \eta_\ell \left( (P_{\ell,n}(1) - \eta_n(1)) \right.
\times P_{\ell,n}(f_n - \eta_n(f_n)) \\
+ \sum_{\ell=0}^{n} \eta_\ell \left( \sum_{p=\ell}^{n} (P_{\ell,p}(1) - \eta_p(1)) \right)
\times P_{\ell,n}(f_n - \eta_n(f_n))),
\]

\[
\tilde{V}_n(f_n) = \sum_{\ell=0}^{n} (n-\ell) \eta_\ell (P_{\ell,n}(f_n - \eta_n(f_n)))^2.
\]

**Proof** See Sect. 6.3.

We can also consider the case where the \( N_2 \) islands are kept independent (a bootstrap filter is still applied within each island, but there is no interaction between islands). To that purpose, denote by \( \tilde{X}_n^{i,j} \) \( N_2 \) independent islands of size \( N_1 \), each evolving using the bootstrap filter and, define the estimator of \( \eta_n^i (f_n) \) for any \( f_n \in B_b(\mathbb{E}_n) \), given by the empirical mean across islands

\[
\tilde{\eta}_n^{N_2} (f_n) \overset{\text{def}}{=} \frac{1}{N_2} \sum_{i=1}^{N_2} f_n(\tilde{X}_n^{i,j}).
\]

For functions \( f_n \) on \( \mathbb{E}_n \) of the form \( f_n(x_n) = m_{N_1} f_n(x_n) \), with \( f_n \in B_b(\mathbb{E}_n) \), we have

\[
\tilde{\eta}_n^{N_2} (f_n) = \frac{1}{N_2} \sum_{i=1}^{N_2} m_{N_1} f_n(\tilde{X}_n^{i,j}) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_2} \sum_{j=1}^{N_1} f_n(\tilde{X}_n^{i,j}).
\]

The asymptotic behavior of the bias and variance of \( m_{N_1} f_n(\tilde{X}_n^{i,j}) \) may be easily deduced from the one of \( \eta_n^{N_1} (f_n) \); Theorem 4 implies that

**Theorem 6** For any time horizon \( n \geq 0 \) and any \( f_n \in B_b(\mathbb{E}_n) \), we have

\[
\lim_{N_1 \to \infty} N_1 \mathbb{E} \left[ \tilde{\eta}_n^{N_2} (m_{N_1} f_n) - \eta_n (f_n) \right] = B_n(f_n),
\]

\[
\lim_{N_1 \to \infty} N_1 N_2 \text{Var}(\tilde{\eta}_n^{N_2} (m_{N_1} f_n)) = V_n(f_n),
\]

where \( B_n(f_n) \) and \( V_n(f_n) \) are defined respectively in (38) and (39).

The variance of the particle approximation is inversely proportional to \( N_1 N_2 \), but because the islands do not interact, the bias is independent of \( N_2 \) and is inversely proportional to \( N_1 \).

As shown by Theorems 5 and 6, a trade-off has to be made between the bias and the variance to decide which of the two estimators \( \eta_n^{N_1} \) and \( \eta_n^{N_2} \) is the best. We can compare the mean squared error (MSE) when the islands interact or when they are kept independent. The MSE for independent islands is given by \( V_n(f_n) + \tilde{V}_n(f_n) \) whereas the MSE for the double bootstrap is given by \( \frac{V_n(f_n) + \tilde{V}_n(f_n)}{N_1 N_2} \). Therefore,

\[
\frac{V_n(f_n) + \tilde{V}_n(f_n)}{N_1 N_2} < \frac{V_n(f_n)}{N_1 N_2} + \frac{B_n(f_n)^2}{N_1^2} \iff N_1 < \frac{B_n(f_n)^2}{\tilde{V}_n(f_n) N_2}.
\]

Consequently, the double bootstrap algorithm outperforms the independent islands when the number of particles \( N_1 \) within each island is small compared to the number of islands \( N_2 \); the interaction improves the bias (which is independent of \( N_2 \) when the islands are kept independent).
the contrary, when \( N_1 \) is larger than \( N_2 \), the variance increase introduced by the interaction (because of the selection step) may be larger than the bias reduction.

4 Extensions

In Sect. 3 we have described and analyzed an interacting island model where the bootstrap algorithm is used both within and across the islands. Of course, other IPS approximations may be considered within and across islands. We will describe how the results of the previous sections may be adapted. The IPS approximation of each individual island may be cast in the Feynman-Kac framework. This section is devoted to check these conditions for various IPS approximations.

4.1 Epsilon-bootstrap interaction

\( \epsilon \)-bootstrap interaction is a variant of the bootstrap, in which the selection step is slightly modified: only a fraction of the particles are resampled. Let \( \epsilon_n \) be a nonnegative constant such that \( \epsilon_n \| g_n \|_\infty \in [0, 1] \), where \( \| g_n \|_\infty = \sup_{x \in \mathbb{E}} |g_n(x_n)| \). For any measure \( \mu_n \in \mathcal{P}(\mathbb{E}_n) \), define \( S_{n,\mu_n} \) the Markov kernel on \( (\mathbb{E}_n, \mathcal{E}_n) \) given for \( x_n \in \mathbb{E}_n \) and \( A_n \in \mathcal{E}_n \) by

\[
S_{n,\mu_n}(x_n, A_n) \defeq \epsilon_n g_n(x_n) \delta_{x_n}(A_n) + (1 - \epsilon_n g_n(x_n)) \Psi_n(\mu_n)(A_n),
\]

where \( \Psi_n \) is defined in (6). \( \epsilon \)-bootstrap interaction algorithm proceeds as follows. At iteration \( n \), a particle \( X_n^i \) is kept with a probability equal to \( \epsilon_n g_n(X_n^i) \) or resampled with a probability \( 1 - \epsilon_n g_n(X_n^i) \). Resampling a particle consists in replacing it by a particle selected at random in the current population with weights proportional to their potential \( \Psi_n(X_n^i) \). Then, each selected particle is independently updated according to the Markov kernel \( M_{n+1} \).

When \( \epsilon_n = 0 \), all the particles are resampled, which correspond to the bootstrap filter. Define the Markov kernel \( M_{n+1}(x_n, dx_{n+1}) \) from \( \mathbb{E}_n \) into \( \mathbb{E}_{n+1} \) by

\[
M_{n+1}(x_n, dx_{n+1}) \defeq \prod_{1 \leq i \leq N_1} S_{n,\gamma_n} M_{n+1}(x_n^i, dx_{n+1}^i).
\]

Consider a Markov chain \( (X_n)_{n \geq 0} \) where for each \( n \in \mathbb{N} \), \( X_n = (X_n^1, \ldots, X_n^{N_1}) \in \mathbb{E}_n \), with initial distribution \( \eta_0 \) and transition kernel \( M_{n+1} \). Define the same approximations of the measures \( \eta_n \) and \( \gamma_n \) as in (11) and (12). Then, consider the island Feynman-Kac model associated to the Markov chain (14) and the potential function (13). The associated sequence \( \{ (\eta_n, \gamma_n) \}_{n \geq 0} \) of Feynman-Kac measures is given in (14), (15) and (16). We may establish the following extension of Theorem 1.

**Theorem 7** Let \( \{ (X^i_0, \ldots, X^i_N) \}_{0 \leq i \leq N} \) be a population of particles generated by the \( \epsilon \)-bootstrap interaction algorithm specified by (47). Then, for any \( f_n \in \mathcal{B}_b(\mathbb{E}_n) \) of the form \( f_n(x_n) = N_1^{-1} \sum_{i=1}^{N_1} f_n(x_n^i) \) with \( f_n \in \mathcal{B}_b(\mathbb{E}_n) \), we get

\[
\eta_n(f_n) = \gamma_n(f_n) = \eta_n(f_n).
\]

**Proof** See Sect. 6.4.

For each \( n \in \mathbb{N} \), let \( (X^1_1, \ldots, X^{N_2}_n) \in \mathbb{E}^{N_2}_n \) be a population of \( N_2 \) islands each of \( N_1 \) individuals. The process \( (X^1_1, \ldots, X^{N_2}_n) \) is a Markov chain evolving according to selection and mutation steps, defined as follows:

- **Selection step**: each island \( X^i_n \) is kept with a probability equal to \( \epsilon_n g_n(X^i_n) \) or resampled with a probability \( 1 - \epsilon_n g_n(X^i_n) \). Resampling an island consists in replacing it by an island selected at random in the current population with weights proportional to their potential \( (g_n(X^1_n), \ldots, g_n(X^{N_2}_n)) \).

- **Mutation step**: each selected island is updated independently according to the Markov transition \( M_{n+1} \).

These islands particles allow to build the \( N_2 \)-particle approximation of the measures \( \eta_n \) and \( \gamma_n \), for any \( f_n \in \mathcal{B}_b(\mathbb{E}_n) \), as

\[
\eta_n^{N_2}(f_n) \defeq \frac{1}{N_2} \sum_{i=1}^{N_2} f_n(X^i_n),
\]

\[
\gamma_n^{N_2}(f_n) \defeq \prod_{0 \leq p < n} \eta_n^{N_2}(f_n) = \eta_n^{N_2}(f_n) \gamma_n^{N_2}(1).
\]

For this selection scheme, the following results, adapted from Del Moral (2004, Corollary 9.3.1, pp. 295–298), establishes the convergence of \( (W_p^{N_1})_{1 \leq p \leq n} \) to centered Gaussian fields:

**Theorem 8** For the \( \epsilon \)-bootstrap filter, for any fixed time horizon \( n \geq 1 \), the sequence \( (W_p^{N_1})_{1 \leq p \leq n} \) defined in (31) converges in law, as \( N_1 \) goes to infinity, to a sequence of \( n \) independent centered Gaussian random fields \( (W_p)_{0 \leq p \leq n} \) with variance given, for any bounded function \( f_p \in \mathcal{B}_b(\mathbb{E}_p) \), and \( 1 \leq p \leq n \), by

\[
\mathbb{E}[W_p(f_p)^2] = \eta_{p-1}[S_{p-1,\eta_{p-1}} M_p f_p^2 - (S_{p-1,\eta_{p-1}} M_p f_p)^2].
\]

This variance is smaller than the variance of the bootstrap algorithm.

**Proposition 1** The asymptotic variance of \( \eta_n^{N_1} \) is smaller with respect to a non-zero sequence \( \{ \epsilon_p \}_{0 \leq p \leq n-1} \) introduced in (46) than in the bootstrap algorithm.
Algorithm 2 ESS within $\epsilon_p$-bootstrap interaction for $\epsilon_p = (\text{essup}_{n_p} \eta_p(g_p))^{-1}$

1: Initialization:
2: for $i$ from 1 to $N_2$ do
3: \hspace{1em} Set $\omega^i_0 = (\omega^{i,j}_0)_{j=1}^{N_1} = (1, \ldots, 1)$.
4: \hspace{1em} Sample $X^i_0 = (X^{i,j}_0)_{j=1}^{N_1}$ independently distributed according to $\eta_0$.
5: end for
6: for $p$ from 0 to $n - 1$ do
7: \hspace{1em} Island selection step:
8: \hspace{2em} for $i$ from 1 to $N_2$ do
9: \hspace{3em} With probability $g_p(X^i_p)/\max_{1 \leq k \leq N_2} g_p(X^k_p)$, set $I^i_p = i$.
10: \hspace{3em} With probability $1 - g_p(X^i_p)/\max_{1 \leq k \leq N_2} g_p(X^k_p)$, sample $I^i_p$ multinomially with probability proportional to $(g_p(X^i_p)/\max_{1 \leq k \leq N_2} g_p(X^k_p))_{i=1}^{N_2}$.
11: end for
12: \hspace{1em} Island mutation step:
13: \hspace{2em} for $i$ from 1 to $N_2$ do
14: \hspace{3em} Particle selection and weight updating within each island:
15: \hspace{4em} Set $N^\text{eff} = (\sum_{j=1}^{N_1} \omega^{i,j}_p g_p(X^j_p)^2)/(\sum_{j=1}^{N_1} \omega^{i,j}_p g_p(X^j_p))^2$.
16: \hspace{4em} if $N^\text{eff}_i \geq \alpha_{\text{Particles}} N_1$ then
17: \hspace{5em} For $1 \leq j \leq N_1$, set $\omega^{i,j+1}_p = \omega^{i,j}_p g_p(X^{i,j}_p)$.
18: \hspace{5em} Set $J^i_p = (J^{i,j}_p)_{j=1}^{N_1} = (1, 2, \ldots, N_1)$.
19: \hspace{4em} else
20: \hspace{5em} Set $\omega^{i,j+1}_p = (\omega^{i,j+1}_p)_{j=1}^{N_1}$.
21: \hspace{5em} Sample $J^i_p = (J^{i,j}_p)_{j=1}^{N_1}$ multinomially with probability proportional to $(\omega^{i,j}_p g_p(X^{i,j}_p))_{j=1}^{N_1}$.
22: end if
23: \hspace{4em} Particle mutation:
24: \hspace{5em} For $1 \leq j \leq N_1$, sample independently $X^{i,j}_{p+1}$ according to $M_{p+1}(X^{i,j}_{p+1}, \cdot)$, where $I^{i,j}_p = J^{i,j}_p$.
25: end for
26: end for
27: Approximate $\eta_n(f_n)$ by $\frac{1}{N_2 \sum_{j=1}^{N_1} \omega^{i,j}_n} \sum_{i=1}^{N_2} \sum_{j=1}^{N_1} \omega^{i,j}_n f_n(X^{i,j}_n)$.

Proof The proof is given in Sect. 6.5. □

For example, for $\epsilon_p = (\text{essup}_{n_p} \eta_p(g_p))^{-1}$, $0 \leq p \leq n$ the asymptotic variance of $\eta^{N_1}_n(f_n), 0 \leq p \leq n$ is lower than for the bootstrap. We can also adapt it at the island level. For instance, Algorithm 2 describes the $\epsilon_p = (\max_{1 \leq j \leq N_2} g_p(X^j_p))^{-1}$-bootstrap islands interaction with ESS filter within the islands.

4.2 Effective sample size interaction

We describe the particle approximation of the probabilities $(\eta_n)_{n \geq 0}$ using the effective sample size (ESS) method introduced in Liu and Chen (1995); see also Liu (2001), Del Moral et al. (2012a), Douc and Moulines (2008) and Arnaud and Le Gland (2009). The difference with the bootstrap filter stems from the selection step of the current particles which is not performed at each step, but only when the importance weights do not satisfy some appropriately defined criterion. Contrary to the bootstrap filter, we now keep both the particles and the weights. Denote by $x^i_n$ a particle and $w^i_n$ its associated weight, assumed to be nonnegative. For a weighted sample $((w^i_n, x^i_n))^{N_1}_{i=1}$, the criterion

$$\left(\sum_{i=1}^{N_1} w^i_n g_n(x^i_n)\right)^2 \left/ \sum_{i=1}^{N_1} (w^i_n g_n(x^i_n))^2 \right.$$ 

is the effective sample size (ESS). The algorithm goes as follows. When the ESS is less than $\alpha N_1$, for some $\alpha \in (0, 1)$, the particles are multinomially resampled with probabilities proportional to their weights times their potential functions and the weights are all reset to 1. When the ESS is greater
than αN₁, then the weights are simply multiplied by the
potential function. The selected particles are then updated us-
ing the transition kernel Mₙ₊₁. For any nonnegative integer
p we set (Eₚ, Eₚ) def = ((Eₚ × ω⁺)ᴺ₁, (Eₚ ⊗ B(ω⁺))⊗ᴺ₁).
Introduce the following set

\[ \Theta_{n,α} = \left\{ x_n = \left( (x_n^1, w_n^1), \ldots, (x_n^{N_1}, w_n^{N_1}) \right) \in E_n \mid \right. \\ \left. \frac{\left( \sum_{i=1}^{N_1} w_n^i g_n(x_n^i) \right)^2}{\sum_{i=1}^{N_1} (w_n^i g_n(x_n^i))^2} \geq \alpha N_1 \right\}. \]

Define the Markov kernel Mₙ₊₁ from Eₙ into Eₙ₊₁ by

\[ M_{n+1}(x_n, dx_{n+1}) \defeq \mathbb{1}_{\Theta_{n,α}}(x_n) \left[ \prod_{1 \leq i \leq N_1} \delta_{w_n^i g_n(x_n^i)}(dx_{n+1}^i) M_{n+1}(x_{n}^i, dx_{n+1}^i) \right] \\
+ \prod_{1 \leq i \leq N_1} \delta_{w_n^i g_n(x_n^i)}(dx_{n+1}^i) \times \sum_{j=1}^{N_1} \sum_{k=1}^{N_1} \frac{w_n^j g_n(x_n^j)}{\sum_{i=1}^{N_1} w_n^i g_n(x_n^i)} M_{n+1}(x_n^j, dx_{n+1}^j). \quad (49) \]

where \( x_n = [(x_n^1, w_n^1), \ldots, (x_n^{N_1}, w_n^{N_1})] \in E_n \) and \( \Theta_{n,α} \) is the complement of \( \Theta_{n,α} \). We define a Markov chain \( (X_n)_{n \geq 0} \) where for each \( n \in \mathbb{N} \),

\[ X_n = \left( (X_n^1, \omega_n^1), \ldots, (X_n^{N_1}, \omega_n^{N_1}) \right) \in E_n, \quad (50) \]

with initial distribution \( \eta_0 \defeq (\eta_0 \otimes \delta_1) \otimes N_1 \) and transition kernel \( M_{n+1} \). Equation (4) suggests the following \( N_1 \)-particle approximations of the measures \( \eta_n \) and \( \gamma_n \) defined for \( f_n \in B_b(E_n) \) by

\[ \eta_{N_1}^n(f_n) \defeq \frac{1}{\sum_{j=1}^{N_1} \omega_n^j} \sum_{j=1}^{N_1} \omega_n^j f_n(X_n^j) = m_{N_1} f_n(X_n), \quad (51) \]

\[ \gamma_{N_1}^n(f_n) \defeq \eta_{N_1}^n(f_n) \prod_{0 \leq p < n} \eta_{N_1}^p(g_p) = \eta_{N_1}^n(f_n) \gamma_{N_1}^n(1), \quad (52) \]

where \( m_{N_1} \) stands for the operator given for any \( f_n \in B_b(E_n) \) by

\[ m_{N_1} f_n(x_n) \defeq \frac{1}{\sum_{i=1}^{N_1} \omega_n^i} \sum_{i=1}^{N_1} \omega_n^i f_n(x_n^i). \]

For \( x_n = [(x_n^1, w_n^1), \ldots, (x_n^{N_1}, w_n^{N_1})] \in E_n \), define the potential function

\[ g_n(x_n) \defeq m_{N_1} g_n(x_n) = \frac{1}{\sum_{i=1}^{N_1} \omega_n^i} \sum_{i=1}^{N_1} \omega_n^i g_n(x_n^i). \quad (53) \]

We can defined the island Feynman-Kac model associated to the Markov chain (49) and the potential function (53), as in (14), (15) and (16). \( \square \)

**Theorem 9** For a particle system \( x_n = \left( (x_n^1, w_n^1), \ldots, (x_n^{N_1}, w_n^{N_1}) \right) \in E_n \) generated by the ESS algorithm and for any \( f_n \in B_b(E_n) \) of the form

\[ f_n(x_n) = \left( \sum_{i=1}^{N_1} w_n^i \right)^{-1} \sum_{i=1}^{N_1} w_n^i f_n(x_n^i) \]

where \( f_n \in B_b(E_n) \),

\[ \gamma_n(f_n) = \gamma_n(f_n) \quad \text{and} \quad \eta_n(f_n) = \eta_n(f_n). \quad (54) \]

**Proof** See Sect. 6.6. \( \square \)

For each \( n \in \mathbb{N} \), let \( (X_n^1, \ldots, X_n^{N_2}) \in E_n^{N_2} \) be a population of \( N_2 \) islands each of \( N_1 \) individuals. We associate to each island, a weight \( \Omega_n^i \), for \( i \in \{1, \ldots, N_2\} \). We can also make the islands interact using an ESS criterion.

The process \( \left( (X_n^1, \Omega_n^1), \ldots, (X_n^{N_2}, \Omega_n^{N_2}) \right) \) is a Markov chain which evolves according to selection and mutation steps, defined as follows

- **Selection step:** if the ESS criterion \( \sum_{i=1}^{N_2} \Omega_n^i g_n(X_n^i)^2 / \sum_{i=1}^{N_2} (\Omega_n^i g_n(X_n^i))^2 \) is larger than \( \beta N_2 \) for some \( \beta \in (0, 1) \), we do not resample the islands and we update the weights thanks to the potential function \( \Omega_{n+1}^i = \Omega_n^i g_n(X_n^i) \); otherwise, we resample the islands with probability proportional to \( (\Omega_n^i g_n(X_n^i))_{i=1}^{N_2} \) and the weights are all reset to 1.

- **Mutation step:** each selected island is updated independently according to the Markov transition \( M_{n+1} \).

These islands particles allow to define the \( N_2 \)-particle approximation of the measures \( \eta_n \) and \( \gamma_n \), for any \( f_n \in B_b(E_n) \), as

\[ \eta_{N_2}^n(f_n) \defeq \frac{1}{\sum_{i=1}^{N_2} \Omega_n^i} \sum_{i=1}^{N_2} \Omega_n^i f_n(X_n^i), \]

\[ \gamma_{N_2}^n(f_n) \defeq \eta_{N_2}^n(f_n) \prod_{0 \leq p < n} \eta_{N_2}^p(g_p) = \eta_{N_2}^n(f_n) \gamma_{N_2}^n(1). \]

Algorithm 3 describes the ESS within ESS island filter.

### 5 Numerical simulations

**Example 1 (Linear Gaussian model)** In order to assess numerically the previous results, we now consider the Linear Gaussian Model (LGM) defined by:

\[ X_{n+1} = \phi X_n + \sigma u X_n, \quad Y_n = X_n + \sigma V_n, \]

\( \odot \) Springer
Algorithm 3 ESS within ESS island filter

1: Initialization:
2: Set \( \mathbf{\Omega}_0 = (\mathbf{\Omega}_0^i)_{i=1}^{N_2} = (1, \ldots, 1) \).
3: for \( i \) from 1 to \( N_2 \) do
4:   Set \( \omega_0^i = (\omega_0^{i,j})_{j=1}^{N_1} = (1, \ldots, 1) \).
5: Sample \( \mathbf{X}_0^i = (X_{0,j}^i)_{j=1}^{N_1} \) independently distributed according to \( \eta_0 \).
6: end for
7: for \( p \) from 0 to \( n - 1 \) do
8:   Island selection step and weight updating:
9:   Set \( N_2^p = (\sum_{i=1}^{N_2} \mathbf{\Omega}_p^i g_p(X_p^i))^2 / \sum_{i=1}^{N_2} (\mathbf{\Omega}_p^i g_p(X_p^i))^2 \).
10: if \( N_2^p \geq \alpha_{\text{islands}} N_2 \) then
11:   For \( 1 \leq i \leq N_2 \), set \( \mathbf{\Omega}_p^i = \mathbf{\Omega}_p^i g_p(X_p^i) \).
12:   Set \( I_p = (I_p^i)_{i=1}^{N_2} = (1, 2, \ldots, N_2) \).
13: else
14:   Set \( \mathbf{\Omega}_p^i = (\mathbf{\Omega}_{p+1}^i)_{i=1}^{N_2} = (1, \ldots, 1) \).
15: Sample \( I_p = (I_p^i)_{i=1}^{N_2} \) multinomially with probability proportional to \((\mathbf{\Omega}_p^i g_p(X_p^i, \omega^i_p))^2\).
16: end if
17: Island mutation step:
18: for \( i \) from 1 to \( N_2 \) do
19:   Particle selection and weight updating within each island:
20:   Set \( N_1^p = (\sum_{j=1}^{N_1} \omega_{p,j}^{i,j} g_p(X_p^{i,j}))^2 / \sum_{j=1}^{N_1} (\omega_{p,j}^{i,j} g_p(X_p^{i,j}))^2 \).
21: if \( N_1^p \geq \alpha_{\text{particles}} N_1 \) then
22:   For \( 1 \leq j \leq N_1 \), set \( \omega_{p,j}^{i,j} = \omega_{p+1,j}^{i,j} g_p(X_p^{i,j}) \).
23:   Set \( J_p^i = (J_p^{i,j})_{j=1}^{N_1} = (1, 2, \ldots, N_1) \).
24: else
25:   Set \( \omega_{p+1,j} = (\omega_{p,j}^{i,j})_{i=1}^{N_2} = (1, \ldots, 1) \).
26: Sample \( J_p^i = (J_p^{i,j})_{j=1}^{N_1} \) multinomially with probability proportional to \((\omega_{p,j}^{i,j} g_p(X_p^{i,j}))^2\).
27: end if
28: Particle mutation:
29: For \( 1 \leq j \leq N_1 \), sample independently \( X_p^{i,j} \) according to \( M_{p+1}(X_p^{i,j}, \lambda_p^{i,j}, \cdot, \cdot) \), where \( \lambda_p^{i,j} = J_p^{i,j} \).
30: end for
31: end for
32: Approximate \( \eta_n(f_n) \) by \( \frac{1}{\sum_{i=1}^{N_2} \mathbf{\Omega}_n} \sum_{i=1}^{N_2} \sum_{j=1}^{N_1} \omega_n^{i,j} f_n(X_n^{i,j}) \).

where \( X_0 \sim N(0, \sigma_n^2/(1 - \phi^2)) \). \( \{U_p\}_{p \geq 1} \) and \( \{V_p\}_{p \geq 1} \) are independent sequences of i.i.d. standard Gaussian random variables, independent of \( X_0 \). In the simulations, we have used \( n = 150 \) observations, generated using the model with \( \phi = 0.9 \), \( \sigma_u = 0.6 \) and \( \sigma_v = 1 \). We focus on the filtering problem, consisting in computing the distribution of the state \( X_n \) given \( Y_0, \ldots, Y_n \). This problem can be cast in the Feynman-Kac setting by framing it as

\[
M_{p+1}(x_p, dx_{p+1}) = \frac{1}{\sqrt{2\pi\sigma_u}} \exp \left[ -\frac{(x_{p+1} - \phi x_p)^2}{2\sigma_u^2} \right] dx_{p+1},
\]

\[
g_p(x_p) = \frac{1}{\sqrt{2\pi\sigma_v}} \exp \left[ -\frac{(y_p - x_p)^2}{2\sigma_v^2} \right].
\]

We estimate the mean of the latent state \( \mathbb{E}[X_n \mid Y_0, \ldots, Y_n] \). We compare the results obtained for different interactions across the islands and for different values of \( N_1 \) and \( N_2 \); in all the simulations, the bootstrap filter is used within the islands. We have run the simulations independently 250 times and we have compared these estimators with the value computed using the Kalman filter. Figure 1 displays the boxplots of the 250 values of these estimators. As expected, for small values of \( N_1 \) compared to \( N_2 \), the bias of independent islands is large compared to cases where islands interact; on the contrary, the variance is smaller for independent islands than for bootstrap island interaction. In this example, the type of interaction between islands does not have a signifi-
Fig. 1 Comparison of different interactions across the islands with bootstrap within each island for the LGM
(1) Bootstrap/independent; (2) Bootstrap/ESS; (3) Bootstrap/Bootstrap; (4) Bootstrap/(1/\|g_n\|_\infty)-bootstrap; (5) Bootstrap/essup_{\eta_p}(g_n)-bootstrap

Table 1 Island interaction number using bootstrap within $\epsilon_p$-bootstrap and double bootstrap for the LGM

| $N_1$ | $N_2$ | 1   | 10  | 100 | 1000 |
|-------|-------|-----|-----|-----|------|
| 1     | 0     | 150 | 631 | 1500| 15000|
| 10    | 0     | 150 | 387 | 1500| 5244 |
| 100   | 0     | 150 | 162 | 2473| 30262|
| 1000  | 0     | 150 | 57  | 915 | 15000|

Table 2 Island interaction number using bootstrap within ESS and double bootstrap for the LGM

| $N_1$ | $N_2$ | 1   | 10  | 100 | 1000 |
|-------|-------|-----|-----|-----|------|
| 1     | 0     | 150 | 718 | 1500| 8052 |
| 10    | 0     | 150 | 190 | 2473| 30262|
| 100   | 0     | 150 | 21  | 247 | 2472 |
| 1000  | 0     | 150 | 0   | 0   | 0    |

Table 3 Percentage of the variance gain using bootstrap within $\epsilon_p$-bootstrap on the left side and ESS within bootstrap on the right side, compared to the double bootstrap, in the LGM example

| $N_1$ | $N_2$ | 10  | 100 | 1000 |
|-------|-------|-----|-----|------|
| 10    | 24.7  | 27.4| 27  | 14.7 |
| 100   | 24.7  | 29.7| 16.1| 12   |
| 1000  | 34.1  | 29.2| 26.4| 41.4 |

An important aspect for the efficiency of the algorithms is the number of interactions between islands. The smaller this number is, the quicker the algorithm will be. The number of interactions in the bootstrap case is $nN_2$. We have compared the island interaction number for the $\epsilon_p$-bootstrap and the ESS interactions w.r.t. the bootstrap one, when we apply the bootstrap filter within the islands. We have computed the empirical number of interactions over the 250 simulations; the results are respectively given in Tables 1 and 2. For a given number of islands, the island interaction number for the ESS and the $\epsilon_p$-bootstrap decrease when the island size grows, whereas it is constant for the bootstrap. The island interaction number is always much smaller using the ESS or the $\epsilon_p$-bootstrap than the bootstrap, across the islands. Moreover, as soon as the number of particles in each island is large enough, the ESS is no longer resampling the islands.

Theorem 8 assures that the variance is smaller using the $\epsilon_p$-bootstrap than the bootstrap interaction. The variance gain using $\epsilon_p$-bootstrap or ESS instead of bootstrap across the islands is given in Table 3. The bootstrap interaction is applied within the islands. The variance is significantly reduced using the $\epsilon_p$-bootstrap or the ESS interaction across the islands, instead of the bootstrap, up to 41 percent variance reduction.

Example 2 (Stochastic volatility model) We consider the stochastic volatility model:

$$X_{p+1} = \alpha X_p + \sigma U_{p+1}, \quad Y_p = \beta e^{\frac{X_p}{2}} V_p,$$

where $X_0 \sim \mathcal{N}(0, \sigma^2/(1-\alpha^2))$, $\{U_p\}_{p\geq 0}$ and $\{V_p\}_{p\geq 0}$
are independent sequences of standard Gaussian random variables independent of $X_0$. In the simulations, we have used $n = 100$ observations, generated using the model with $\alpha = 0.98$, $\sigma = 0.5$ and $\beta = 1$. We estimate the mean of the latent state $X_n$ given the observations $Y_0, \ldots, Y_n$. This problem can be cast in the Feynman-Kac framework by setting for all $p \geq 0$

$$M_{p+1}(x_p, dx_{p+1}) = \frac{1}{\sqrt{2\pi \sigma}} \exp\left[-\frac{(x_{p+1} - \alpha x_p)^2}{2\sigma^2}\right] dx_{p+1},$$

$$g_p(x_p) = \frac{1}{\sqrt{2\pi \beta \sigma}} \exp\left[-\frac{x_p/2 - \gamma_p}{2\beta^2}\right].$$

We have computed this quantity using a single run of bootstrap filter with $10^6$ particles. In the following results, we always consider bootstrap interaction within each island, and we compare different interactions across the islands, for several values of $N_1$ and $N_2$. We have run the simulations independently 250 times. Figure 2 displays the boxplots of the 250 values of these estimators. The behavior of the different methods is similar to the one observed for the Linear Gaussian Model example.

We have compared the island interaction number for the $\epsilon_p$-bootstrap and the ESS interaction w.r.t. the bootstrap one, when we apply the bootstrap filter within the islands. We have computed the empirical number of interactions over the 250 simulations; the results are respectively given in Tables 4 and 5. The number of interactions in the bootstrap case is $nN_2$. The same phenomena are observed as for the Linear Gaussian Model example.

The variance gain using the $\epsilon_p$-bootstrap or the ESS instead of the bootstrap across the islands is given in Table 6. The bootstrap interaction is applied within the islands. The variance is significantly reduced using the $\epsilon_p$-bootstrap or the ESS interaction across the islands, instead of the bootstrap, up to 66 percent variance reduction.
6 Proofs

6.1 Proof of Theorem 1

Using (11) and (13), \( g_n(X_n) \) may be expressed as \( g_n(X_n) = \eta_{N_n}^{N_1}(g_n) \) where \( X_n \) and \( \eta_{N_n}^{N_1} \) are defined in (11) and (10), respectively. Similarly, for any \( f_n \in B_0(E_n) \) of the form \( f_n(x_n) = N_n^{-1} \sum_{i=1}^{N_n} f_n(x_i^n) \) where \( f_n \in B_0(E_n) \), \( f_n(X_n) \) is given by \( f_n(x_n) = \eta_{N_n}^{N_1}(f_n) \). Note that

\[
\gamma_n^{N_1}(f_n) \equiv \mathbb{E} \left[ f_n(X_n) \prod_{0 \leq p < n} g_p(X_p) \right] = \mathbb{E} \left[ \eta_{N_n}^{N_1}(f_n) \prod_{0 \leq p < n} \eta_{N_n}^{N_1}(g_p) \right],
\]

and since by (12), it suffices to prove that \( \gamma_n^{N_1}(f_n) \) is an unbiased estimator of \( \gamma_n(f_n) \), i.e.

\[
\mathbb{E}[\gamma_n^{N_1}(f_n)] = \gamma_n(f_n).
\] (56)

Define the filtration \( F_{n+1}^{N_1} \equiv \sigma(X_p, 0 \leq p \leq n) \). Note that

\[
\mathbb{E}[\eta_{N_n}^{N_1}(f_p) \mid F_{n-1}^{N_1}] = \mathbb{E} \left[ f_p(X_p) \mid F_{n-1}^{N_1} \right] = \frac{1}{\eta_{n-1}^{N_1}(g_{n-1})} \sum_{i=1}^{N_n} \eta_{n-1}^{N_1}(Q_n f_n) \prod_{0 \leq p < n-1} \eta_{n-p}^{N_1}(g_p),
\]

where \( Q_p \) is defined in (22).

By the definition of \( \gamma_n^{N_1} \) given in (12), we have

\[
\mathbb{E}[\gamma_n^{N_1}(f_n)] = \mathbb{E} \left[ \mathbb{E}[\eta_{N_n}^{N_1}(f_n) \mid F_{n-1}^{N_1}] \prod_{0 \leq p < n} \eta_{N_n}^{N_1}(g_p) \right] = \mathbb{E} \left[ \mathbb{E} \left[ \frac{\eta_{n-1}^{N_1}(Q_n f_n)}{\eta_{n-1}^{N_1}(g_{n-1})} \prod_{0 \leq p < n-1} \eta_{n-p}^{N_1}(g_p) \right] \right] = \mathbb{E} \left[ \eta_{n-1}^{N_1}(Q_n f_n) \prod_{0 \leq p < n-1} \eta_{n-p}^{N_1}(g_p) \right].
\]

By iterating this step we get

\[
\mathbb{E}[\gamma_n^{N_1}(f_n)] = \mathbb{E}[\eta_0^{N_1}(Q_1 \ldots Q_n f_n)] = \mathbb{E}[Q_1 \ldots Q_n f_n(X_1^n)] = \gamma_0 Q_1 \ldots Q_n f_n = \gamma_n(f_n).
\]

6.2 Proof of Theorem 4

We preface the proof by the following Lemma.

**Lemma 1** For any \( f_n^1, f_n^2 \in B_0(E_n) \), the pair \((W_n^{\gamma_n^{N_1}}, f_n^1), (W_n^{\gamma_n^{N_1}}, f_n^2)\) converges in law, as \( N_1 \) tends to infinity, to \((W_n^\gamma, f_n^1), (W_n^\gamma, f_n^2)\). In addition, for any polynomial function \( \Phi : \mathbb{R}^2 \rightarrow \mathbb{R} \), we have:

\[
\lim_{N_1 \rightarrow \infty} \mathbb{E}[\Phi(W_n^{\gamma_n^{N_1}}, f_n^1), W_n^{\gamma_n^{N_1}}, f_n^2)] = \mathbb{E}[\Phi(W_n^{\gamma}, f_n^1), W_n^{\gamma}, f_n^2)]).
\]

**Proof** For any \((\alpha, \beta) \in \mathbb{R}^2\) by the definitions (30) of \( W_n^{\gamma_n^{N_1}} \) and (33) of \( W_n^{\gamma_n^{N_1}} \), we have

\[
\alpha W_n^{\gamma_n^{N_1}}(f_n^1) + \beta W_n^{\gamma_n^{N_1}}(f_n^2)
\]

\[
= \sum_{p=0}^{N_1} \left[ \alpha \gamma_p^{N_1}(1) W_n^{\gamma_n^{N_1}}(Q_{p+n} f_n^1) + \beta \gamma_p^{N_1}(1) W_n^{\gamma_n^{N_1}}(Q_{p+n} (f_n^2 - \eta_n(f_n^2))) \right].
\]

As in the proof of Theorem 2, a simple application of Slutsky’s Lemma allows to show that \( \alpha W_n^{\gamma_n^{N_1}}(f_n^1) + \beta W_n^{\gamma_n^{N_1}}(f_n^2) \) converges in law to \( \alpha W_n^\gamma(f_n^1) + \beta W_n^\gamma(f_n^2) \).

The proof follows from Del Moral (2004, Theorem 7.4.4), using that for any \( p \geq 1 \),

\[
\sup_{N_1 \geq 1} \mathbb{E}[W_n^{\gamma_n^{N_1}}(f_n^1) | f_n^2]^{1/p} \leq c_p(n) \| f_n^1 \|, \quad (58)
\]

\[
\sup_{N_1 \geq 1} \mathbb{E}[W_n^{\gamma_n^{N_1}}(f_n^2) | f_n^1]^{1/p} \leq c_p(n) \| f_n^2 \|, \quad (59)
\]

for some finite constant \( c_p(n) \) depending only on \( p \) and \( n \). \( \square \)

**Proof of Theorem 4** Consider first the bias term. We decompose the error as follows using (33):

\[
N_1[\eta_n^{N_1}(f_n) - \eta_n(f_n)]
\]

\[
= \sqrt{N_1} W_n^{\gamma_n^{N_1}}(f_n) - \sqrt{N_1} \frac{\gamma_n(1)}{\gamma_n(1)} W_n^{\gamma_n^{N_1}} \left( \frac{f_n - \eta_n(f_n)}{\gamma_n(1)} \right)
\]

\[
= \sqrt{N_1} \left[ \frac{\gamma_n(1)}{\gamma_n(1)} - 1 \right] W_n^{\gamma_n^{N_1}} \left( \frac{f_n - \eta_n(f_n)}{\gamma_n(1)} \right) + \sqrt{N_1} W_n^{\gamma_n^{N_1}} \left( \frac{f_n - \eta_n(f_n)}{\gamma_n(1)} \right).
\]

Since \( W_n^{\gamma_n^{N_1}} = \sqrt{N_1}[\gamma_n^{N_1} - \gamma_n] \), Theorem 1 shows that, the expectation of the second term on the RHS of the previous equation, is zero. By noting that
From the definition (33) of $W^{\gamma,N_i}$, we have
\[
{N_1}E[W^{\gamma,N_i}(f_n) - \eta_n(f_n)] = - \frac{1}{\gamma_n(N_1)} W^{\gamma,N_i}(1) W^{\gamma,N_i}(f_n - \eta_n(f_n)),
\]
where $W^{\gamma,N_i}$ is defined in (30), we get
\[
{N_1}E[W^{\gamma,N_i}(f_n) - \eta_n(f_n)] = - \frac{1}{\gamma_n(N_1)} W^{\gamma,N_i}(1) W^{\gamma,N_i}(f_n - \eta_n(f_n)),
\]
where $W^{\gamma,N_i}$ is given in (33). According to Lemma 1:
\[
\lim_{N_i \to \infty} \frac{1}{N_1} E[W^{\gamma,N_i}(f_n) - \eta_n(f_n)] = B_n(f_n),
\]
by the definitions of $W^{\gamma}$ and $W^{\eta}$. Consider now the variance. We use the decomposition
\[
\text{Var}(\eta_n(f_n)) = \text{E}[(\eta_n(f_n) - \eta_n(f_n))^2] - [\text{E}(\eta_n(f_n) - \eta_n(f_n))]^2.
\]
Using (60), we get $[\text{E}[\eta_n(N_1) - \eta_n(f_n))]^2 = O(N_1^{-2})$. From the definition (33) of $W^{\eta,N_i}$, it follows $\text{E}(\eta_n(f_n) - \eta_n(f_n))^2 = N_1^{-1} \text{E}[W^{\eta,N_i}(f_n)^2]$, implying that
\[
\lim_{N_i \to \infty} \frac{1}{N_1} \text{Var}(\eta_n(f_n)) = \text{E}[W^{\eta}(f_n)^2] = V_n(f_n),
\]
by the definition of $W^{\eta}$ and using again Lemma 1. □

6.3 Proof of Theorem 5

We preface the proof of Theorem 5 by the following result on the usual Feynman-Kac model.

Lemma 2 For any time horizon $n \geq 1$ and any functions $f_n^1, f_n^2 \in B_b(E_n)$ such that $\eta_n(f_n^1) = 0$, we have
\[
\lim_{N_i \to \infty} \frac{1}{N_1} \text{E}[\eta_n^{N_i}(f_n^1) \eta_n^{N_i}(f_n^2)] = \gamma_n(1) \sum_{p=0}^{n} \text{E}[W_p(P_{p,n}(f_n^1)) W_p(P_{p,n}(f_n^2 - \eta_n(f_n^2)))].
\]

Proof By the definition (12) of $\gamma_n$ we have $\eta_n^{N_i}(f_n^1) \prod_{p=0}^{n} \eta_n^{N_i}(g_p) = \gamma_n^{N_i}(f_n^1)$, and, according to (56), $\text{E}[\gamma_n^{N_i}(f_n^1)] = \gamma_n(f_n^1) = \gamma_n(1) \eta_n(f_n^1) = 0$, so that we get
\[
\text{E}[\eta_n^{N_i}(f_n^1) \eta_n^{N_i}(f_n^2)] = \frac{1}{\gamma_n(N_1)} [\gamma_n^{N_i}(f_n^1) \eta_n^{N_i}(f_n^2)] = \frac{1}{\gamma_n(N_1)} \text{E}[\eta_n^{N_i}(f_n^1) \eta_n^{N_i}(f_n^2)],
\]
which implies
\[
Q_n f_n(x_{n-1}) = m^{N_1} Q_n f_n(x_{n-1}),
\]
\[
Q_{p,n} f_n(x_p) = m^{N_1} Q_{p,n} f_n(x_p),
\]
for any $p \leq n$. (63)

Proof We have from (57)
\[
M_n f_n(x_{n-1}) = \text{E} \left[ f_n(X_n) | X_{n-1} = x_{n-1} \right] = \frac{m^{N_1} Q_n f_n(x_{n-1})}{m^{N_1} g_{n-1}(x_{n-1})},
\]
which implies
\[
Q_n f_n(x_{n-1}) = g_{n-1}(x_{n-1}) M_n f_n(x_{n-1}) = m^{N_1} g_{n-1}(x_{n-1}) \times \frac{m^{N_1} Q_n f_n(x_{n-1})}{m^{N_1} g_{n-1}(x_{n-1})},
\]
showing (61). The proof of (62) follows by an induction since
\[
Q_{p,n} f_n(x_p) = Q_{p,n-1} Q_{p,n} f_n(x_p).
\]

Proof of Theorem 5 Asymptotic bias behavior: For any fixed $N_1$, the asymptotic bias behavior of $\eta_n^{N_1}(f_n)$ is given
for any $f_n \in B_b(\mathbb{E}_n)$ by applying Theorem 4 to the island particle model in the bootstrap case:

$$
\lim_{N_2 \to \infty} N_2 \mathbb{E}\left[ \eta_n^{N_2}(f_n) - \eta_n(f_n) \right] = - \sum_{p=0}^{n} \eta_p \left[ P_{p,n}(1) P_{p,n}(f_n - \eta_n(f_n)) \right].
$$

For linear functions $f_n$ of the form $f_n = m N_1 f_n$ where $f_n \in B_b(\mathbb{E}_n)$, Lemma 3 states that

$$
P_{p,n}(f_n - \eta_n(f_n)) = m N_1 P_{p,n}(f_n - \eta_n(f_n)),
$$

Therefore, we get

$$
P_{p,n}(1)(X_p) = m N_1 P_{p,n}(1)(X_p) = \eta_p N_1 \left( P_{p,n}(1) \right). \tag{65}
$$

Then we get

$$
\eta_p \left[ P_{p,n}(1) P_{p,n}(f_n - \eta_n(f_n)) \right]
\begin{align*}
(1) & \quad \gamma_p \left[ P_{p,n}(1) P_{p,n}(f_n - \eta_n(f_n)) \right] \\
(2) & \quad \mathbb{E}[P_{p,n}(1)(X_p) P_{p,n}(f_n - \eta_n(f_n))(X_p)] \prod_{\ell=0}^{p-1} W_{p}(X_{\ell}) \\
(3) & \quad \mathbb{E}[\eta_p N_1 \left( P_{p,n}(f_n - \eta_n(f_n)) \right) \eta_p N_1 \left( P_{p,n}(1) \right) \prod_{\ell=0}^{p-1} \eta_p N_1 (g_{\ell})] \\
& \quad \gamma_p(1), \tag{66}
\end{align*}
$$

where (1) is simply the definition (15) of $\eta_p$, (2) stems from the definition (16) of $\gamma_p$, and (3) follows from Theorem 1, the definition (13) of $(g_{\ell})_{\ell \geq 0}$ and equations (64) and (65). As $\eta_p \left( P_{p,n}(f_n - \eta_n(f_n)) \right) = 0$ we can apply Lemma 2 and

$$
\lim_{N_1 \to \infty} \frac{N_1 \eta_p \left[ P_{p,n}(1) P_{p,n}(f_n - \eta_n(f_n)) \right]}{\gamma_p(1)}
\begin{align*}
(1) & \quad \sum_{\ell=0}^{p} \mathbb{E}[W_{\ell} \left( P_{\ell,p} \left( P_{p,n}(1) - \eta_p P_{p,n}(1) \right) \right) ] \\
(2) & \quad \sum_{\ell=0}^{p} \mathbb{E}[W_{\ell} \left( P_{\ell,n}(1) - P_{\ell,p}(1) \right) W_{\ell} \left( P_{\ell,n}(f_n - \eta_n(f_n)) \right)] \\
& \quad \sum_{\ell=0}^{p} \mathbb{E}[W_{\ell} \left( P_{\ell,n}(f_n - \eta_n(f_n)) \right)^2],
\end{align*}
$$

from which we conclude that

$$
\lim_{N_1 \to \infty} \lim_{N_2 \to \infty} \frac{N_1 N_2 \mathbb{E}\left[ \eta_n^{N_2}(f_n) - \eta_n(f_n) \right]}{\gamma_p(1)}
\begin{align*}
(1) & \quad \sum_{\ell=0}^{p} \sum_{p=0}^{n} \mathbb{E}[W_{\ell} \left( P_{\ell,n}(f_n - \eta_n(f_n)) \right)^2] \\
(2) & \quad \sum_{\ell=0}^{p} \sum_{p=0}^{n} \mathbb{E}[W_{\ell} \left( P_{\ell,n}(f_n - \eta_n(f_n)) \right)^2] \\
& \quad \sum_{\ell=0}^{p} \sum_{p=0}^{n} \mathbb{E}[W_{\ell} \left( P_{\ell,n}(f_n - \eta_n(f_n)) \right)^2],
\end{align*}
$$

where $V_n(f_n)$ is defined in (39) and $\tilde{V}_n(f_n)$ is given in (43).
6.4 Proof of Theorem 7

Lemma 4 Let $\epsilon_n$ be a nonnegative constant such that $\epsilon_n g_n \in [0, 1]$. Then
\[
\Psi_n(\mu_n) = \mu_n S_{n, \mu_n},
\]
where $S_{n, \mu_n}$ is defined in (46).

Proof By (46) and (6) we have for any $A \in \mathcal{E}_n$
\[
\mu_n S_{n, \mu_n}(A_n) = \int \mu_n(dx_n)S_{n, \mu_n}(x_n, A_n)
\]
\[
= \int \mu_n(dx_n)\left[\epsilon_n g_n(x_n)\delta(x_n)(A_n) + (1 - \epsilon_n g_n(x_n))\Psi_n(\mu_n)(A_n)\right]
\]
\[
= \epsilon_n \int A_n \mu_n(dx_n)g_n(x_n)
\]
\[
+ (1 - \epsilon_n \mu_n(g_n))\Psi_n(\mu_n)(A_n)
\]
\[
= \Psi_n(\mu_n)(A_n).
\]
\[\square\]

Let $\mathcal{F}_{n}^{N_i}$ be the increasing filtration associated to the particle evolution $\mathcal{F}_{n}^{N_i} \equiv \sigma(X_p, 0 \leq p \leq n)$. As in the proofs of Theorems 1 and 9, the only point is to prove that
\[
\mathbb{E}[\eta_{n}^{N_i}(f_p) \mid \mathcal{F}_{n}^{N_i}] = \frac{\eta_{n-1}^{N_i}(Q_p f_p)}{\eta_{n-1}^{N_i}(g_p)},
\]
where $Q_p$ is defined in (22). Or,
\[
\mathbb{E}[\eta_{n}^{N_i}(f_p) \mid \mathcal{F}_{n}^{N_i}] = \frac{1}{N_i} \sum_{i=1}^{N_i} \mathbb{E}[f_p(X_i^p) \mid \mathcal{F}_{n-1}^{N_i}] = \frac{1}{N_i} \sum_{i=1}^{N_i} M_p(f_p)(X_i^{p-1})
\]
\[
= \eta_{n-1}^{N_i} M_p(f_p) = \eta_{n-1}^{N_i} S_{p-1, n-1}^{N_i} M_p(f_p)
\]
\[
= \Psi_{p-1}(\eta_{n-1}^{N_i}) M_p(f_p) = \frac{\eta_{n-1}^{N_i}(Q_p f_p)}{\eta_{n-1}^{N_i}(g_p)},
\]
using respectively (11), (47), Lemma 4 and (6).

6.5 Proof of Proposition 1

For the $\epsilon$-interaction bootstrap, the sequence $(W_{p}^{N_i})_{0 \leq p \leq n}$ converges in law, as $N_i$ tends to infinity, to a sequence of $n$ independent centered Gaussian random fields $(W_p)_{0 \leq p \leq n}$ with variance given by
\[
\mathbb{E}[W_p(f_p)^2]
\]
\[
= \eta_{p-1} S_{p-1, n-1} M_p f_p^2 - \eta_{p-1}[(S_{p-1, n-1} M_p f_p)^2]
\]
\[
= \Psi_{p-1}(\eta_{p-1}) M_p f_p^2 - \eta_{p-1}[(S_{p-1, n-1} M_p f_p)^2],
\]
thanks to Lemma 4.

In the special case $\epsilon_p = 0$ (the bootstrap case), the function $S_{p, n}^g$ is constant and equal to $\Psi_p(\eta_p)(g_p)$ and the variance for the bootstrap is just
\[
\Psi_{p-1}(\eta_{p-1}) (M_p f_p^2) - (\Psi_{p-1}(\eta_{p-1}) M_p f_p)^2
\]

Therefore, the variance of the $\epsilon$-interaction bootstrap may be decomposed as follows
\[
\mathbb{E}[W_p(f_p)^2]
\]
\[
= (\Psi_{p-1}(\eta_{p-1}) (M_p f_p^2) - (\Psi_{p-1}(\eta_{p-1}) M_p f_p)^2)
\]
\[
- (\eta_{p-1}[(S_{p-1, n-1} M_p f_p)^2]
\]
\[
- (\Psi_{p-1}(\eta_{p-1}) M_p f_p)^2).
\]

Observing,
\[
\eta_{p-1}[(S_{p-1, n-1} M_p f_p)^2] - (\Psi_{p-1}(\eta_{p-1}) M_p f_p)^2
\]
\[
= \eta_{p-1}[(S_{p-1, n-1} M_p f_p - \Psi_{p-1}(\eta_{p-1}) M_p f_p)]^2 \geq 0,
\]
allows to conclude.

6.6 Proof of Theorem 9

Using (51), (53), (16), and for $f_n$ such that $f_n(X_n) = (\sum_{i=1}^{N_i} u^i_{n-1} \sum_{i=1}^{N_i} u^i_n f_n(X_n)) = m^{N_i} f_n(X_n) = \eta_{n}^{N_i}(f_n)$, we get
\[
\gamma_n(f_n) \equiv \mathbb{E}\left[\prod_{0 \leq p < n} g_p(X_p)\right]
\]
\[
= \mathbb{E}\left[\eta_{n}^{N_i}(f_n) \prod_{0 \leq p < n} \eta_{p}^{N_i}(g_p)\right].
\]

By (52), it suffices to prove that $\mathbb{E}[\eta_{n}^{N_i}(f_n)] = \gamma_n(f_n)$. We define by $\mathcal{F}_{n}^{N_i}$ the increasing filtration associated to the particle evolution $\mathcal{F}_{n}^{N_i} \equiv \sigma(X_p, 0 \leq p \leq n)$. We will show that for any $p > 0$ and $f_p \in B_\mathcal{B}(\mathbb{E})$, we have $\mathbb{E}[\eta_{n}^{N_i}(f_p) \mid \mathcal{F}_{n}^{N_i}] = \eta_{p}^{N_i}(Q_p f_p) / \eta_{p-1}^{N_i}(g_p)$, where $Q_p$ is defined in (22). Indeed, by the definitions (49) of $M_p$ and (51) of $\eta_{n}^{N_i}$, Springer
\[
E[\eta_{p}^{N_{1}}(f_{p}) | F_{p-1}^{N_{1}}] = \sum_{i=1}^{N_{1}} \frac{\omega_{i}^{p}}{\sum_{j=1}^{N_{1}} \omega_{j}^{p}} E[f_{p}(X_{p}^{i}) | X_{p-1}]
\]
\[
= 1_{\theta_{p-1,\alpha}}(X_{p-1}) \times \left[ \sum_{i=1}^{N_{1}} \frac{\omega_{i}^{p-1} g_{p-1}(X_{p-1}^{i}) M_{p} f_{p}(X_{p-1}^{i})}{\sum_{i=1}^{N_{1}} \omega_{i}^{p-1} g_{p-1}(X_{p-1}^{i})} \right]
\]
\[
+ 1_{\theta_{p-1,\alpha}}(X_{p-1}) \times \left[ \frac{1}{N_{1}} \sum_{i=1}^{N_{1}} \frac{\omega_{i}^{p-1} g_{p-1}(X_{p-1}^{i}) M_{p} f_{p}(X_{p-1}^{i})}{\sum_{i=1}^{N_{1}} \omega_{i}^{p-1} g_{p-1}(X_{p-1}^{i})} \right]
\]
\[
= \frac{\eta_{p}^{N_{1}}(Q_{p} f_{p})}{\eta_{p}^{N_{1}}(g_{p-1})}.
\]

The proof follows exactly along the same lines as Theorem 1. By iterating this step we get
\[
E[\gamma_{n}^{N_{1}}(f_{n})] = E[\eta_{0}^{N_{1}}(Q_{1} \ldots Q_{n} f_{n})]
\]
\[
= E[Q_{1} \ldots Q_{n} f_{n}(X_{1}^{i})]
\]
\[
= \gamma_{0} Q_{1} \ldots Q_{n} f_{n} = \gamma_{n}(f_{n}).
\]

As the reader may have noticed, this unbias property doesn’t depend on the definition of the sets \(\theta_{p,\alpha}\) defining the resampling times. From this observation, we underline that Theorem 9 is also true for more general classes of resampling time criterion.

**Acknowledgements** This work is supported by the Agence Nationale de la Recherche through the 2009-2012 project Big MC. The work of Christelle Vergé is financially supported by CNES (Centre National d’Études Spatiales) and Onera, The French Aerospace Lab.

**References**

Arnaud, E., Le Gland, F.: SMC with Adaptive Resampling: Large Sample Asymptotics, pp. 481–484 (2009)

Cappé, O., Moulines, E.: On the use of particle filtering for maximum likelihood parameter estimation. In: European Signal Processing Conference (EUSIPCO), Antalya, Turkey (2005)

Casarin, R., Grassi, S., Ravazzolo, F., Van Dijk, H.K.: Parallel sequential Monte Carlo for efficient density combination: the deco Matlab toolbox. Tinbergen Institute Discussion Paper, 055/III (2013)

Chopin, N.: A sequential particle filter method for static models. Biometrika 89, 539–552 (2002)

Chopin, N., Jacob, P., Papaspiliopoulos, O.: SMC²; a sequential Monte Carlo algorithm with particle Markov chain Monte Carlo updates. J. R. Stat. Soc. B. 75(3), 397–426 (2013). doi:10.1111/j.1467-9868.2012.01046.x

Del Moral, P.: Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications. Springer, Berlin (2004)

Del Moral, P., Doucet, A., Jasra, A.: On adaptive resampling strategies for sequential Monte Carlo methods. Bernoulli 18(1), 252–278 (2012a)

Del Moral, P., Hu, P., Wu, L.: On the concentration properties of interacting particle processes. Found. Trends Mach. Learn. 3(3–4), 225–289 (2012b)

Douc, R., Moulines, E.: Limit theorems for weighted samples with applications to sequential Monte Carlo methods. Ann. Stat. 36(5), 2344–2376 (2008)

Doucet, A., Freitas, N.D. (eds.): Sequential Monte Carlo Methods in Practice. Springer, New York (2001)

Durham, G., Geweke, J.: Massively parallel sequential Monte Carlo for Bayesian inference. Manuscript, URL http://www.censoc.uts.edu.au/pdfs/geweke_papers/gp_working_9.pdf (2011)

Jasra, A., Doucet, A., Stephens, D.A., Holmes, C.C.: Interacting sequential Monte Carlo samplers for trans-dimensional simulation. Comput. Stat. Data Anal. 52, 1765–1791 (2008)

Liu, J.S.: Monte Carlo Strategies in Scientific Computing. Springer, New York (2001)

Liu, J., Chen, R.: Blind deconvolution via sequential imputations. J. Am. Stat. Assoc. 90(420), 567–576 (1995)