

COMBINED USE OF IMPORTANCE WEIGHTS AND RESAMPLING WEIGHTS IN SEQUENTIAL MONTE CARLO METHODS^{*,**}

FRANCOIS LE GLAND¹

Abstract. A particle approximation of Feynman–Kac distributions is presented here, that combines SIS and SIR algorithms in the sense that only a fraction of the importance weights is used for resampling, and two different approaches are proposed to analyze its performance. The first approach is based on a representation in terms of path–space distributions, and could be used to analyze the joint particle approximation of distributions for a reference model and for several alternate models at the same time. The second approach, which is of independent interest and seems very promising, is based on a representation in terms of a multiplicative functional, and could be used to analyze particle approximation with adaptive resampling schemes.

(Updated version from 18 December 2007)

INTRODUCTION

Consider the unnormalized and normalized Feynman–Kac distributions defined on the set E by

$$\langle \gamma_n, \phi \rangle = \mathbb{E}[\phi(X_n) \prod_{k=0}^n g_k(X_k)] \quad \text{and} \quad \langle \mu_n, \phi \rangle = \frac{\langle \gamma_n, \phi \rangle}{\langle \gamma_n, 1 \rangle},$$

respectively, where $\{X_k, k = 0, 1, \dots, n\}$ is a Markov chain taking values in E and characterized by

- its initial probability distribution $\eta_0(dx)$,
- and its transition probabilities $Q_k(x, dx')$, for any $k = 1, \dots, n$,

and where $g_k(x)$ is a bounded nonnegative function for any $k = 0, 1, \dots, n$, and more generally

$$\langle \gamma_n, \phi \rangle = \int_E \cdots \int_E \phi(x_n) \gamma_0(dx_0) \prod_{k=1}^n R_k(x_{k-1}, dx_k),$$

which includes the previous case, with

$$\gamma_0(dx) = g_0(x) \eta_0(dx) \quad \text{and} \quad R_k(x, dx') = Q_k(x, dx') g_k(x'),$$

^{*} *This work was partially supported by CNRS, under the MathSTIC project Chaînes de Markov Cachées et Filtrage Particulaire, and under the AS–STIC project Méthodes Particulières (AS 67), and by Electricité de France R&D.*

^{**} *This paper is dedicated to the memory of Natacha Caylus (1977–2006).*

¹ IRISA / INRIA, Campus de Beaulieu, 35042 RENNES Cédex, France

for any $k = 1, \dots, n$. Associated with this probabilistic (or more generally, integral) representation is the recurrence relation $\gamma_k = \gamma_{k-1} R_k$, for any $k = 1, \dots, n$. In full generality, it is always possible to decompose the nonnegative measure

$$\gamma_0(dx) = W_0(x) p_0(dx) , \quad (1)$$

in terms of a nonnegative function and a normalized probability distribution, and to decompose the nonnegative kernel

$$R_k(x, dx') = W_k(x, x') P_k(x, dx') , \quad (2)$$

in terms of a nonnegative function and a normalized Markov kernel, for any $k = 1, \dots, n$.

Assuming a further factorization of the nonnegative functions

$$W_0(x) = W_0^{\text{imp}}(x) W_0^{\text{red}}(x) \quad \text{and} \quad W_k(x, x') = W_k^{\text{imp}}(x, x') W_k^{\text{red}}(x, x') , \quad (3)$$

for any $k = 1, \dots, n$, the following decompositions hold

$$\gamma_0(dx) = W_0^{\text{imp}}(x) W_0^{\text{red}}(x) p_0(dx) \quad \text{and} \quad R_k(x, dx') = W_k^{\text{imp}}(x, x') W_k^{\text{red}}(x, x') P_k(x, dx') , \quad (4)$$

for any $k = 1, \dots, n$, and the main contribution of this paper is to exploit this decomposition to design and study particle approximations of the form

$$\mu_k^N = \sum_{i=1}^N u_k^i w_k^i \delta_{\xi_k^i} \quad \text{with} \quad \sum_{i=1}^N w_k^i = 1 \quad \text{and} \quad \sum_{i=1}^N u_k^i w_k^i = 1 ,$$

which combine SIS and SIR algorithms [10, Section 3.4.4], with two species of nonnegative weights

- the weights (w_k^1, \dots, w_k^N) are used for resampling,
- the importance weights (u_k^1, \dots, u_k^N) are used for weighting.

One possible motivation is the joint particle approximation of Feynman–Kac distributions associated with reference and alternate models, under the following absolute continuity assumptions

$$\gamma_0(dx) = r_0(x) \gamma_0^0(dx) \quad \text{and} \quad R_k(x, dx') = r_k(x, x') R_k^0(x, dx') , \quad (5)$$

for any $k = 1, \dots, n$. In this case indeed, assuming a decomposition of the reference nonnegative measure

$$\gamma_0^0(dx) = W_0^0(x) p_0^0(dx) ,$$

in terms of a nonnegative function and a normalized probability distribution, and assuming a decomposition of the reference nonnegative kernel

$$R_k^0(x, dx') = W_k^0(x, x') P_k^0(x, dx') ,$$

in terms of a nonnegative function and a normalized Markov kernel, for any $k = 1, \dots, n$, the following decompositions hold

$$\gamma_0(dx) = r_0(x) W_0^0(x) p_0^0(dx) \quad \text{and} \quad R_k(x, dx') = r_k(x, x') W_k^0(x, x') P_k^0(x, dx') , \quad (6)$$

for any $k = 1, \dots, n$, which clearly are of the form (4), and it is possible to design particle approximations of the form

$$\mu_k^N = \sum_{i=1}^N u_k^i w_k^{0,i} \delta_{\xi_k^{0,i}} \quad \text{with} \quad \sum_{i=1}^N w_k^{0,i} = 1 \quad \text{and} \quad \sum_{i=1}^N u_k^i w_k^{0,i} = 1$$

where

- the particle positions $(\xi_k^{0,1}, \dots, \xi_k^{0,N})$ and the resampling weights $(w_k^{0,1}, \dots, w_k^{0,N})$ depend on the reference model only,

- the importance weights (u_k^1, \dots, u_k^N) depend on both the reference and alternate models.

Clearly, the two point of views are mathematically equivalent, upon suitable substitution of the different nonnegative measures and nonnegative functions, and to decide which point of view to adopt usually depends on the application.

The paper is organized as follows : An original particle approximation is presented in Section 1, that combines importance weights and resampling weights. To analyze its performance, and in particular to derive a central limit theorem as the number N of particles goes to infinity, with an explicit expression for the asymptotic variance, two different approaches are proposed. A first representation is introduced in Section 2, in terms of path-space distributions and a path-particle approximation is presented in Section 3, with a central limit theorem stated in Theorem 3.2 and with an explicit expression for the asymptotic variance given in Proposition 3.5. A second representation is introduced in Section 4, in terms of the multiplicative functional of a Markov chain and an extended particle approximation is presented in Section 5, where importance weights are treated as particles, with a central limit theorem stated in Theorem 5.2 and with an explicit expression for the asymptotic variance given in Proposition 5.5. It is checked that all three particle approximations actually correspond to the same algorithm, and that the two expressions obtained for the asymptotic variance are actually equal. The respective merits of the two different approaches are discussed in the Conclusion.

The following abuse of notation $W_0(x, x') = W_0(x')$, $W_0^{\text{imp}}(x, x') = W_0^{\text{imp}}(x')$ and $W_0^{\text{red}}(x, x') = W_0^{\text{red}}(x')$ will be used throughout the paper.

1. PARTICLE APPROXIMATION WITH COMBINED WEIGHTING AND RESAMPLING

Recall that the unnormalized distributions satisfy the following recurrence relation

$$\gamma_k = \gamma_{k-1} R_k = \mu_{k-1} R_k \langle \gamma_{k-1}, 1 \rangle ,$$

for any $k = 1, \dots, n$. Recall also the decomposition (1), and introducing a particle approximation of the form

$$\mu_k \approx \mu_k^N = \sum_{i=1}^N u_k^i w_k^i \delta_{\xi_k^i} \quad \text{with} \quad \sum_{i=1}^N w_k^i = 1 \quad \text{and} \quad \sum_{i=1}^N u_k^i w_k^i = 1 ,$$

and using the decomposition (2), yields

$$\mu_{k-1}^N R_k(dx') = \sum_{i=1}^N \underbrace{u_{k-1}^i W_k(\xi_{k-1}^i, x')}_{w_k^i(x')} \underbrace{w_{k-1}^i P_k(\xi_k^i, dx')}_{m_k^i(dx')} ,$$

which can be interpreted as the marginal nonnegative measure on E associated with a nonnegative measure on the product set $\{1, \dots, N\} \times E$. Using the auxiliary variable approach [11], the resulting ASIR algorithm can be described by

$$\gamma_0^N = \frac{1}{N} \sum_{i=1}^N W_0(\xi_0^i) \delta_{\xi_0^i} \quad \text{and} \quad \gamma_k^N = \frac{1}{N} \sum_{i=1}^N u_{k-1}^{\tau_k^i} W_k(\xi_{k-1}^{\tau_k^i}, \xi_k^i) \delta_{\xi_k^i} \langle \gamma_{k-1}^N, 1 \rangle ,$$

where $(\xi_0^1, \dots, \xi_0^N)$ are i.i.d. random variables taking values in E and with common probability distribution $p_0(dx)$, and where $((\tau_k^1, \xi_k^1), \dots, (\tau_k^N, \xi_k^N))$ are i.i.d. random variables taking values in the product set $\{1, \dots, N\} \times E$ and with common probability distribution $(m_k^1(dx'), \dots, m_k^N(dx'))$, or equivalently

$$\tau_k^i \sim (w_{k-1}^1, \dots, w_{k-1}^N) \quad \text{and} \quad \xi_k^i \sim P_k(\xi_{k-1}^{\tau_k^i}, dx') ,$$

independently for any $i = 1, \dots, N$. Using the further factorization (3) results in the following approximations

$$\mu_0^N = \frac{\sum_{i=1}^N W_0^{\text{imp}}(\xi_0^i) W_0^{\text{red}}(\xi_0^i) \delta_{\xi_0^i}}{\sum_{j=1}^N W_0^{\text{imp}}(\xi_0^j) W_0^{\text{red}}(\xi_0^j)} = \sum_{i=1}^N u_0^i w_0^i \delta_{\xi_0^i},$$

and

$$\mu_k^N = \frac{\sum_{i=1}^N u_{k-1}^i W_k^{\text{imp}}(\xi_{k-1}^i, \xi_k^i) W_k^{\text{red}}(\xi_{k-1}^i, \xi_k^i) \delta_{\xi_k^i}}{\sum_{j=1}^N u_{k-1}^j W_k^{\text{imp}}(\xi_{k-1}^j, \xi_k^j) W_k^{\text{red}}(\xi_{k-1}^j, \xi_k^j)} = \sum_{i=1}^N u_k^i w_k^i \delta_{\xi_k^i},$$

for the normalized distributions, which defines implicitly the resampling weights

$$w_0^i = \frac{W_0^{\text{red}}(\xi_0^i)}{\sum_{j=1}^N W_0^{\text{red}}(\xi_0^j)} \quad \text{and} \quad w_k^i = \frac{W_k^{\text{red}}(\xi_{k-1}^i, \xi_k^i)}{\sum_{j=1}^N W_k^{\text{red}}(\xi_{k-1}^j, \xi_k^j)},$$

and the importance weights

$$u_0^i = \frac{W_0^{\text{imp}}(\xi_0^i)}{\sum_{j=1}^N w_0^j W_0^{\text{imp}}(\xi_0^j)} \quad \text{and} \quad u_k^i = \frac{u_{k-1}^i W_k^{\text{imp}}(\xi_{k-1}^i, \xi_k^i)}{\sum_{j=1}^N w_k^j u_{k-1}^j W_k^{\text{imp}}(\xi_{k-1}^j, \xi_k^j)},$$

for any $i = 1, \dots, N$.

Remark 1.1. If the substitution $W_0^{\text{imp}} \leftarrow r_0$, $W_0^{\text{red}} \leftarrow W_0^0$ and $p_0 \leftarrow p_0^0$ is made, and if the substitution $W_k^{\text{imp}} \leftarrow r_k$, $W_k^{\text{red}} \leftarrow W_k^0$ and $P_k \leftarrow P_k^0$ is made for any $k = 1, \dots, n$, with the notations of (6), then clearly the particle positions $(\xi_k^1, \dots, \xi_k^N)$ and the resampling weights (w_k^1, \dots, w_k^N) depend on the reference model only, whereas the importance weights (u_k^1, \dots, u_k^N) depend on both the reference and alternate models. If in addition the derivatives in (5) are continuous or differentiable w.r.t. some parameter of the model, then the importance weights will automatically inherit the same property, as suggested in [8]. This idea has been used in Monte Carlo maximum likelihood estimation [1, 7] or in smooth particle approximation of Feynman–Kac distributions [3, 4, 6, 9]. Using a single particle system for the reference value, with different importance weights corresponding to different values, makes the resulting approximation regular, but also poorly accurate for values too far from the reference. In opposition, using a different particle system for each different value, would make the resulting approximation very irregular, but also uniformly accurate for all values.

To derive a central limit theorem for the proposed particle approximation, no direct approach seems possible. However, as pointed out by Éric Moulines during the Oxford workshop, a particle approximation in path–space could be used for this purpose. Alternatively, as suggested by Pierre Del Moral on another occasion, an extended particle approximation, where importance weights are treated as particles, could also be used for the same purpose. These two different approaches are explored in Sections 2 and 3, and in Sections 4 and 5, respectively, and their merits are discussed in the Conclusion.

2. REPRESENTATION IN PATH-SPACE

Using the decompositions (4) yields

$$\begin{aligned} \langle \gamma_n, \phi \rangle &= \int_E \cdots \int_E \phi(x_n) \gamma_0(dx_0) \prod_{k=1}^n R_k(x_{k-1}, dx_k) \\ &= \int_E \cdots \int_E \phi(x_n) W_0^{\text{imp}}(x_0) W_0^{\text{red}}(x_0) p_0(dx_0) \prod_{k=1}^n W_k^{\text{imp}}(x_{k-1}, x_k) W_k^{\text{red}}(x_{k-1}, x_k) P_k(x_{k-1}, dx_k) \\ &= \mathbb{E}[\phi(X_n) \prod_{k=0}^n W_k^{\text{imp}}(X_{k-1}, X_k) \prod_{k=0}^n W_k^{\text{red}}(X_{k-1}, X_k)], \end{aligned}$$

where $\{X_k, k = 0, 1, \dots, n\}$ is a Markov chain taking values in E and characterized by

- its initial probability distribution $p_0(dx)$,
- and its transition probabilities $P_k(x, dx')$, for any $k = 1, \dots, n$.

Consider the unnormalized and normalized Feynman–Kac distributions defined on the path-space $E_{0:n} = E \times \cdots \times E$ by

$$\langle \gamma_n^\bullet, f_n \rangle = \mathbb{E}[f_n(X_n^\bullet) \prod_{k=0}^n g_k^\bullet(X_k^\bullet)] \quad \text{and} \quad \langle \mu_n^\bullet, f_n \rangle = \frac{\langle \gamma_n^\bullet, f_n \rangle}{\langle \gamma_n^\bullet, 1 \rangle},$$

respectively, where $\{X_k^\bullet, k = 0, 1, \dots, n\}$ is a path-space valued Markov chain defined by $X_k^\bullet = (X_0, \dots, X_k) = X_{0:k}$ for any $k = 0, 1, \dots, n$ and characterized by

- its initial probability distribution $\eta_0^\bullet(dx_0) = p_0(dx_0)$,
- and its transition probabilities $Q_k^\bullet(x_{0:k-1}, dx'_{0:k}) = \delta_{x_{0:k-1}}(dx'_{0:k-1}) P_k(x'_{k-1}, dx'_k)$, for any $k = 1, \dots, n$,

and where $g_k^\bullet(x_{0:k}) = W_k^{\text{red}}(x_{k-1}, x_k)$ for any $k = 0, 1, \dots, n$. Associated with this path-space probabilistic representation is the recurrence relation $\gamma_k^\bullet = \gamma_{k-1}^\bullet R_k^\bullet$, with

$$R_k^\bullet(x_{0:k-1}, dx'_{0:k}) = Q_k^\bullet(x_{0:k-1}, dx'_{0:k}) g_k^\bullet(x'_{0:k}) = \delta_{x_{0:k-1}}(dx'_{0:k-1}) P_k(x'_{k-1}, dx'_k) W_k^{\text{red}}(x'_{k-1}, x'_k),$$

for any $k = 1, \dots, n$. In particular for any function of the form

$$f_n(x_{0:n}) = F(x_n, \prod_{k=0}^n W_k^{\text{imp}}(x_{k-1}, x_k)),$$

defined on path-space, it holds

$$\langle \gamma_n^\bullet, f_n \rangle = \mathbb{E}[F(X_n, \prod_{k=0}^n W_k^{\text{imp}}(X_{k-1}, X_k)) \prod_{k=0}^n W_k^{\text{red}}(X_{k-1}, X_k)],$$

and for instance, for the function

$$T_n \phi(x_{0:n}) = \phi(x_n) \prod_{k=0}^n W_k^{\text{imp}}(x_{k-1}, x_k) \quad \text{then clearly} \quad \langle \gamma_n^\bullet, T_n \phi \rangle = \langle \gamma_n, \phi \rangle, \quad (7)$$

or in other words $\gamma_n = \gamma_n^\bullet T_n$ in terms of a transformed distribution, and for the function

$$T_n^\square \phi(x_{0:n}) = \phi_0(x_n) \left| \prod_{k=0}^n W_k^{\text{imp}}(x_{k-1}, x_k) \right|^2 \quad \text{it holds} \quad \langle \gamma_n^\bullet, T_n^\square \phi \rangle = \langle \gamma_n^\square, \phi \rangle,$$

with the unnormalized and normalized Feynman–Kac distributions defined on the set E by

$$\langle \gamma_n^\square, \phi \rangle = \mathbb{E}[\phi(X_n) | \prod_{k=0}^n W_k^{\text{imp}}(X_{k-1}, X_k)]^2 \prod_{k=0}^n W_k^{\text{red}}(X_{k-1}, X_k) \quad \text{and} \quad \langle \mu_n^\square, \phi \rangle = \frac{\langle \gamma_n^\square, \phi \rangle}{\langle \gamma_n^\square, 1 \rangle},$$

respectively.

3. PARTICLE APPROXIMATION IN PATH-SPACE

Recall that the unnormalized distributions satisfy the recurrence relation

$$\gamma_k^\bullet = \gamma_{k-1}^\bullet R_k^\bullet = g_k^\bullet (\mu_{k-1}^\bullet Q_k^\bullet) \langle \gamma_{k-1}^\bullet, 1 \rangle,$$

for any $k = 1, \dots, n$. Introducing a weighted particle approximation of the form

$$\mu_k^\bullet \approx \mu_k^{\bullet, N} = \sum_{i=1}^N w_k^i \delta_{\xi_k^{\bullet, i}} \quad \text{with} \quad \sum_{i=1}^N w_k^i = 1,$$

where $\xi_k^{\bullet, i} = (\xi_{0,k}^i, \dots, \xi_{k,k}^i)$ is a path-space particle with terminal position $\xi_k^i = \xi_{k,k}^i$ for any $i = 1, \dots, N$, yields

$$\mu_{k-1}^{\bullet, N} Q_k^\bullet(dx'_{0:k}) = \sum_{i=1}^N w_{k-1}^i \delta_{\xi_{k-1}^{\bullet, i}}(dx'_{0:k-1}) P_k(x'_{k-1}, dx'_k).$$

The corresponding SIR algorithm can be described by

$$\gamma_0^{\bullet, N} = \frac{1}{N} \sum_{i=1}^N g_0^\bullet(\xi_0^{\bullet, i}) \delta_{\xi_0^{\bullet, i}} \quad \text{and} \quad \gamma_k^{\bullet, N} = \frac{1}{N} \sum_{i=1}^N g_k^\bullet(\xi_k^{\bullet, i}) \delta_{\xi_k^{\bullet, i}} \langle \gamma_{k-1}^{\bullet, N}, 1 \rangle,$$

where $(\xi_0^{\bullet, 1}, \dots, \xi_0^{\bullet, N})$ are i.i.d. random variables on the set E with common probability distribution $p_0(dx_0)$, and where $g_0^\bullet(\xi_0^{\bullet, i}) = W_0^{\text{red}}(\xi_{0,0}^i)$ for any $i = 1, \dots, N$, and where $(\xi_k^{\bullet, 1}, \dots, \xi_k^{\bullet, N})$ are i.i.d. random variables on the path-space $E_{0:k} = E \times \dots \times E$ with common probability distribution $\mu_{k-1}^{\bullet, N} Q_k^\bullet(dx'_{0:k})$, or equivalently

$$\tau_k^i \sim (w_{k-1}^1, \dots, w_{k-1}^N) \quad \text{and} \quad (\xi_{0,k}^i, \dots, \xi_{k-1,k}^i) = (\xi_{0,k-1}^{\tau_k^i}, \dots, \xi_{k-1,k-1}^{\tau_k^i}) \quad \text{and} \quad \xi_{k,k}^i \sim P_k(\xi_{k-1,k}^i, dx'_k),$$

independently for any $i = 1, \dots, N$, and where $g_k^\bullet(\xi_k^{\bullet, i}) = W_k^{\text{red}}(\xi_{k-1,k}^i, \xi_{k,k}^i) = W_k^{\text{red}}(\xi_{k-1,k-1}^{\tau_k^i}, \xi_{k,k}^i)$ for any $i = 1, \dots, N$. This results in the following approximations

$$\mu_0^{\bullet, N} = \frac{\sum_{i=1}^N W_0^{\text{red}}(\xi_0^i) \delta_{\xi_0^{\bullet, i}}}{\sum_{j=1}^N W_0^{\text{red}}(\xi_0^j)} = \sum_{i=1}^N w_0^i \delta_{\xi_0^{\bullet, i}} \quad \text{and} \quad \mu_k^{\bullet, N} = \frac{\sum_{i=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i) \delta_{\xi_k^{\bullet, i}}}{\sum_{j=1}^N W_k^{\text{red}}(\xi_{k-1}^j, \xi_k^j)} = \sum_{i=1}^N w_k^i \delta_{\xi_k^{\bullet, i}},$$

for the normalized distributions, in terms of the terminal positions of the path-space particles, which defines implicitly the resampling weights as

$$w_0^i = \frac{W_0^{\text{red}}(\xi_0^i)}{\sum_{j=1}^N W_0^{\text{red}}(\xi_0^j)} \quad \text{and} \quad w_k^i = \frac{W_k^{\text{red}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i)}{\sum_{j=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^j}, \xi_k^j)},$$

for any $i = 1, \dots, N$. In particular for the function

$$T_k \phi(x_{0:k}) = \phi(x_k) \prod_{p=0}^k W_p^{\text{imp}}(x_{p-1}, x_p),$$

defined on path-space and already introduced in (7), it holds

$$T_k \phi(\xi_k^{\bullet, i}) = \phi(\xi_k^i) \prod_{p=0}^k W_p^{\text{imp}}(\xi_{p-1, k}^i, \xi_{p, k}^i),$$

hence

$$\langle \gamma_0^{\bullet, N}, T_0 \phi \rangle = \frac{1}{N} \sum_{i=1}^N g_0^{\bullet}(\xi_0^{\bullet, i}) T_0 \phi(\xi_0^{\bullet, i}) = \frac{1}{N} \sum_{i=1}^N W_0^{\text{red}}(\xi_0^i) v_0^i \phi(\xi_0^i),$$

where

$$v_0^i = W_0^{\text{imp}}(\xi_0^i),$$

for any $i = 1, \dots, N$, and

$$\langle \gamma_k^{\bullet, N}, T_k \phi \rangle = \frac{1}{N} \sum_{i=1}^N g_k^{\bullet}(\xi_k^{\bullet, i}) T_k \phi(\xi_k^{\bullet, i}) \langle \gamma_{k-1}^{\bullet, N}, 1 \rangle = \frac{1}{N} \sum_{i=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i) v_k^i \phi(\xi_k^i) \langle \gamma_{k-1}^{\bullet, N}, 1 \rangle,$$

where

$$\begin{aligned} v_k^i &= \prod_{p=0}^k W_p^{\text{imp}}(\xi_{p-1, k}^i, \xi_{p, k}^i) \\ &= W_k^{\text{imp}}(\xi_{k-1, k}^i, \xi_{k, k}^i) \prod_{p=0}^{k-1} W_p^{\text{imp}}(\xi_{p-1, k}^i, \xi_{p, k}^i) \\ &= W_k^{\text{imp}}(\xi_{k-1, k-1}^{\tau_k^i}, \xi_{k, k}^i) \prod_{p=0}^{k-1} W_p^{\text{imp}}(\xi_{p-1, k-1}^{\tau_k^i}, \xi_{p, k-1}^{\tau_k^i}) \\ &= W_k^{\text{imp}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i) v_{k-1}^{\tau_k^i}, \end{aligned}$$

for any $i = 1, \dots, N$. Notice the underlying recursive structure, in the form of a multiplicative functional of a Markov chain up to resampling. In view of the interpretation $\gamma_n = \gamma_n^{\bullet} T_n$ in terms of a transformed distribution, this results in the following approximations

$$\gamma_0^N = \frac{1}{N} \sum_{i=1}^N W_0^{\text{red}}(\xi_0^i) v_0^i \delta_{\xi_0^i} \quad \text{and} \quad \gamma_k^N = \frac{1}{N} \sum_{i=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i) v_k^i \delta_{\xi_k^i} \langle \gamma_{k-1}^{\bullet, N}, 1 \rangle,$$

for the unnormalized distributions, hence

$$\mu_0^N = \frac{\sum_{i=1}^N W_0^{\text{red}}(\xi_0^i) v_0^i \delta_{\xi_0^i}}{\sum_{j=1}^N W_0^{\text{red}}(\xi_0^j) v_0^j} = \sum_{i=1}^N u_0^i w_0^i \delta_{\xi_0^i} \quad \text{and} \quad \mu_k^N = \frac{\sum_{i=1}^N W_k^{\text{red}}(\xi_{k-1}^{T_k^i}, \xi_k^i) v_k^i \delta_{\xi_k^i}}{\sum_{j=1}^N W_k^{\text{red}}(\xi_{k-1}^{T_k^j}, \xi_k^j) v_k^j} = \sum_{i=1}^N u_k^i w_k^i \delta_{\xi_k^i},$$

for the normalized distributions, which defines implicitly the importance weights as

$$u_k^i = \frac{v_k^i}{\sum_{j=1}^N v_k^j w_k^j} \quad \text{with} \quad v_0^i = W_0^{\text{imp}}(\xi_0^i) \quad \text{and} \quad v_k^i = W_k^{\text{imp}}(\xi_{k-1}^{T_k^i}, \xi_k^i) v_{k-1}^i,$$

for any $i = 1, \dots, N$.

Remark 3.1. It is clear that the particle positions (defined as the terminal positions of the path-space particles) and resampling weights defined here are the same as the particle positions and resampling weights defined in Section 1, and it is easy to check by induction that the importance weights defined here are the same as the importance weights defined in Section 1.

For this algorithm, the following central limit theorem holds [5, Chapter 9], [2, Chapter 9].

Theorem 3.2.

$$\sqrt{N} \frac{\langle \gamma_n^{\bullet, N} - \gamma_n^{\bullet}, f_n \rangle}{\langle \gamma_n^{\bullet}, 1 \rangle} \Longrightarrow \mathcal{N}(0, V_n^{\bullet}(f_n)),$$

in distribution as $N \uparrow \infty$, with asymptotic variance

$$V_n^{\bullet}(f_n) = \frac{\text{var}(g_0^{\bullet} R_{1:n}^{\bullet} f_n, \eta_0^{\bullet})}{\langle \eta_0^{\bullet}, g_0^{\bullet} R_{1:n}^{\bullet} 1 \rangle^2} + \sum_{k=1}^n \frac{\text{var}(g_k^{\bullet} R_{k+1:n}^{\bullet} f_n, \eta_k^{\bullet})}{\langle \eta_k^{\bullet}, g_k^{\bullet} R_{k:n}^{\bullet} 1 \rangle^2},$$

where $\eta_k^{\bullet} = \mu_{k-1}^{\bullet} Q_k^{\bullet}$ for any $k = 1, \dots, n$, and where

$$R_{k+1:n}^{\bullet} f_n(x_{0:k}) = R_{k+1}^{\bullet} \cdots R_n^{\bullet} f_n(x_{0:k}) = \mathbb{E}[f_n(X_n^{\bullet}) \prod_{p=k+1}^n g_p^{\bullet}(X_p^{\bullet}) \mid X_k^{\bullet} = x_{0:k}],$$

for any $k = 0, 1, \dots, n$, with $R_{n+1:n}^{\bullet} f_n(x_{0:n}) = f_n(x_{0:n})$ by convention.

Remark 3.3 (CLT for normalizing constants). It follows from the identity

$$\frac{\langle \gamma_n^N - \gamma_n, 1 \rangle}{\langle \gamma_n, 1 \rangle} = \frac{\langle \gamma_n^{\bullet}, 1 \rangle}{\langle \gamma_n^{\bullet}, 1 \rangle} \frac{\langle \gamma_n^{\bullet, N} - \gamma_n^{\bullet}, T_n 1 \rangle}{\langle \gamma_n^{\bullet}, 1 \rangle},$$

and from Theorem 3.2 that

$$\sqrt{N} \frac{\langle \gamma_n^N - \gamma_n, 1 \rangle}{\langle \gamma_n, 1 \rangle} \Longrightarrow \mathcal{N}\left(0, V_n\left(\frac{\langle \gamma_n^{\bullet}, 1 \rangle}{\langle \gamma_n^{\bullet}, 1 \rangle} T_n 1\right)\right),$$

in distribution as $N \uparrow \infty$.

Remark 3.4 (CLT for normalized distributions). It follows from the decomposition

$$\langle \mu_n^N - \mu_n, \phi \rangle = \frac{\langle \gamma_n^{\bullet}, 1 \rangle}{\langle \gamma_n, 1 \rangle} \frac{\langle \gamma_n^{\bullet}, T_n 1 \rangle}{\langle \gamma_n^{\bullet, N}, T_n 1 \rangle} \left\langle \frac{\gamma_n^{\bullet, N} - \gamma_n^{\bullet}}{\langle \gamma_n^{\bullet}, 1 \rangle}, T_n(\phi - \langle \mu_n, \phi \rangle) \right\rangle,$$

from Theorem 3.2 and from the Slutsky lemma, that

$$\sqrt{N} \langle \mu_n^N - \mu_n, \phi \rangle \Longrightarrow \mathcal{N}(0, V_n(\frac{\langle \gamma_n^\bullet, 1 \rangle}{\langle \gamma_n, 1 \rangle} T_n(\phi - \langle \mu_n, \phi \rangle))) ,$$

in distribution as $N \uparrow \infty$.

As a consequence, it is worth finding an explicit expression, in terms of Feynman–Kac distributions defined on the set E , of the asymptotic variance $V_n(f_n)$ for functions of the form $f_n = T_n \phi$ defined on the path-space $E_{0:n} = E \times \cdots \times E$. Let

$$R_{k+1:n} \phi(x) = R_{k+1} \cdots R_n \phi(x) = \mathbb{E}[\phi(X_n) \prod_{p=k+1}^n W_p(X_{p-1}, X_p) \mid X_k = x] ,$$

for any $k = 0, 1, \dots, n$, with $R_{n+1:n} \phi(x) = \phi(x)$ by convention. Then

$$\frac{\text{var}(g_0^\bullet R_{1:n}^\bullet T_n(\frac{\langle \gamma_n^\bullet, 1 \rangle}{\langle \gamma_n, 1 \rangle} \phi), \eta_0^\bullet)}{\langle \eta_0^\bullet, g_0^\bullet R_{1:n}^\bullet 1 \rangle^2} = \frac{\text{var}(W_0 R_{1:n} \phi, p_0)}{\langle p_0, W_0 R_{1:n} 1 \rangle^2} ,$$

and

$$\begin{aligned} \frac{\text{var}(g_k^\bullet R_{k+1:n}^\bullet T_n(\frac{\langle \gamma_n^\bullet, 1 \rangle}{\langle \gamma_n, 1 \rangle} \phi), \eta_k^\bullet)}{\langle \eta_k^\bullet, g_k^\bullet R_{k+1:n}^\bullet 1 \rangle^2} &= \frac{\langle \gamma_{k-1}^\square, 1 \rangle \langle \gamma_{k-1}^\bullet, 1 \rangle}{\langle \gamma_{k-1}, 1 \rangle^2} \frac{\text{var}(W_k (R_{k+1:n} \phi) \circ \pi, \mu_{k-1}^\square \otimes P_k)}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} \\ &+ [\frac{\langle \gamma_{k-1}^\square, 1 \rangle \langle \gamma_{k-1}^\bullet, 1 \rangle}{\langle \gamma_{k-1}, 1 \rangle^2} \frac{\langle \mu_{k-1}^\square, R_{k:n} \phi \rangle^2}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} - \frac{\langle \mu_{k-1}, R_{k:n} \phi \rangle^2}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2}] , \end{aligned}$$

for any $k = 1, \dots, n$.

Proposition 3.5. *In particular for the normalizing constant*

$$\sqrt{N} \frac{\langle \gamma_n^N - \gamma_n, 1 \rangle}{\langle \gamma_n, 1 \rangle} \Longrightarrow \mathcal{N}(0, V_n) ,$$

in distribution as $N \uparrow \infty$, with asymptotic variance

$$\begin{aligned} V_n &= \frac{\text{var}(W_0 R_{1:n} \phi, p_0)}{\langle p_0, W_0 R_{1:n} 1 \rangle^2} + \sum_{k=1}^n \frac{\langle \gamma_{k-1}^\square, 1 \rangle \langle \gamma_{k-1}^\bullet, 1 \rangle}{\langle \gamma_{k-1}, 1 \rangle^2} \frac{\text{var}(W_k (R_{k+1:n} 1) \circ \pi, \mu_{k-1}^\square \otimes P_k)}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} \\ &+ \sum_{k=1}^n [\frac{\langle \gamma_{k-1}^\square, 1 \rangle \langle \gamma_{k-1}^\bullet, 1 \rangle}{\langle \gamma_{k-1}, 1 \rangle^2} \frac{\langle \mu_{k-1}^\square, R_{k:n} 1 \rangle^2}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} - 1] . \end{aligned}$$

Remark 3.6. In the extreme case where only resampling weights are used, i.e. if $W_k^{\text{red}} = W_k$ and $W_k^{\text{imp}} \equiv 1$ for any $k = 0, 1, \dots, n$, then the last sum cancels out, and

$$V_n = \frac{\text{var}(W_0 R_{1:n} \phi, p_0)}{\langle p_0, W_0 R_{1:n} 1 \rangle^2} + \sum_{k=1}^n \frac{\text{var}(W_k (R_{k+1:n} 1) \circ \pi, \mu_{k-1} \otimes P_k)}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} ,$$

as expected.

4. REPRESENTATION IN TERMS OF A MULTIPLICATIVE FUNCTIONAL

Starting rather from the absolute continuity assumptions

$$\gamma_0(dx) = W_0^{\text{imp}}(x) \gamma_0^0(dx) \quad \text{and} \quad R_k(x, dx') = W_k^{\text{imp}}(x, x') R_k^0(x, dx'), \quad (8)$$

which define implicitly, in view of the decompositions (4)

$$\gamma_0^0(dx) = W_0^{\text{red}}(x) p_0(dx) \quad \text{and} \quad R_k^0(x, dx') = W_k^{\text{red}}(x, x') P_k(x, dx'), \quad (9)$$

for any $k = 1, \dots, n$, consider the unnormalized and normalized Feynman–Kac distributions defined on the product set $E \times [0, \infty)$ by

$$\langle \gamma_n^e, F \rangle = \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\text{imp}}(x_{k-1}, x_k)) \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \quad \text{and} \quad \langle \mu_n^e, F \rangle = \frac{\langle \gamma_n^e, F \rangle}{\langle \gamma_n^e, 1 \rangle},$$

respectively. In particular

$$\langle \gamma_k^e, \phi \otimes e_0 \rangle = \int_E \cdots \int_E \phi(x_n) \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) = \langle \gamma_k^0, \phi \rangle,$$

where $e_0(v) \equiv 1$ by definition, and using the absolute continuity assumptions (8)

$$\begin{aligned} \langle \gamma_k^e, \phi \otimes e \rangle &= \int_E \cdots \int_E \phi(x_n) W_0^{\text{imp}}(x_0) \prod_{k=1}^n W_k^{\text{imp}}(x_{k-1}, x_k) \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \\ &= \int_E \cdots \int_E \phi(x_n) \gamma_0(dx_0) \prod_{k=1}^n R_k(x_{k-1}, dx_k) = \langle \gamma_k, \phi \rangle, \end{aligned}$$

where $e(v) \equiv v$ by definition. Associated with these two integral representations are the two recurrence relations $\gamma_k^0 = \gamma_{k-1}^0 R_k^0$ and $\gamma_k = \gamma_{k-1} R_k$ respectively, for any $k = 1, \dots, n$. Finally

$$\begin{aligned} \langle \gamma_k^e, \phi \otimes e^2 \rangle &= \int_E \cdots \int_E \phi(x_n) |W_0^{\text{imp}}(x_0)|^2 \prod_{k=1}^n |W_k^{\text{imp}}(x_{k-1}, x_k)|^2 \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \\ &= \int_E \cdots \int_E \phi(x_n) \gamma_0^\square(dx_0) \prod_{k=1}^n R_k^\square(x_{k-1}, dx_k) = \langle \gamma_k^\square, \phi \rangle, \end{aligned}$$

where

$$\gamma_0^\square(dx) = |W_0^{\text{imp}}(x)|^2 \gamma_0^0(dx) \quad \text{and} \quad R_k^\square(x, dx') = |W_k^{\text{imp}}(x, x')|^2 R_k^0(x, dx'),$$

for any $k = 1, \dots, n$. In particular for $\phi(x) \equiv 1$, it holds

$$\langle \gamma_k^e, 1 \rangle = \langle \gamma_k^e, 1 \otimes e_0 \rangle = \langle \gamma_k^0, 1 \rangle \quad \text{and} \quad \langle \gamma_k^e, 1 \otimes e \rangle = \langle \gamma_k, 1 \rangle,$$

for the normalizing constants, and

$$\langle \mu_k^e, \phi \otimes e_0 \rangle = \frac{\langle \gamma_k^e, \phi \otimes e_0 \rangle}{\langle \gamma_k^e, 1 \rangle} = \frac{\langle \gamma_k^0, \phi \rangle}{\langle \gamma_k^0, 1 \rangle} = \langle \mu_k^0, \phi \rangle, \quad (10)$$

$$\langle \mu_k^e, \phi \otimes e \rangle = \frac{\langle \gamma_k^e, \phi \otimes e \rangle}{\langle \gamma_k^e, 1 \rangle} = \frac{\langle \gamma_k, \phi \rangle}{\langle \gamma_k^0, 1 \rangle} = \frac{\langle \gamma_k, 1 \rangle}{\langle \gamma_k^0, 1 \rangle} \langle \mu_k, \phi \rangle, \quad (11)$$

and

$$\langle \mu_k^e, \phi \otimes e^2 \rangle = \frac{\langle \gamma_k^e, \phi \otimes e^2 \rangle}{\langle \gamma_k^e, 1 \rangle} = \frac{\langle \gamma_k^\square, \phi \rangle}{\langle \gamma_k^0, 1 \rangle} = \frac{\langle \gamma_k^\square, 1 \rangle}{\langle \gamma_k^0, 1 \rangle} \langle \mu_k^\square, \phi \rangle,$$

for the normalized Feynman–Kac distributions. In other words, the extended unnormalized Feynman–Kac distribution encodes all the different Feynman–Kac distributions, normalized or unnormalized, for the reference and alternate models, and in particular

$$\langle \gamma_n^e, \phi_0 \otimes e_0 + \phi \otimes e \rangle = \langle \gamma_n^0, \phi_0 \rangle + \langle \gamma_n, \phi \rangle.$$

It follows from the definition that

$$\gamma_0^e(dx, dv) = \gamma_0^0(dx) \delta_{W_0^{\text{imp}}(x)}(dv),$$

and introducing the extended nonnegative kernel

$$R_k^e(x, v, dx', dv') = R_k^0(x, dx') \delta_{v W_k^{\text{imp}}(x, x')}(dv'),$$

defined on the product set $E \times [0, \infty)$, for any $k = 1, \dots, n$, it is easily seen that

$$\begin{aligned} \langle \gamma_n^e, F \rangle &= \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\text{imp}}(x_{k-1}, x_k)) \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \\ &= \int_E \cdots \int_E \left\{ \int_0^\infty \cdots \int_0^\infty F(x_n, v_n) \delta_{W_0^{\text{imp}}(x_0)}(dv_0) \prod_{k=1}^n \delta_{v_{k-1} W_k^{\text{imp}}(x_{k-1}, x_k)}(dv_k) \right\} \\ &\quad \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \\ &= \int_E \int_0^\infty \cdots \int_E \int_0^\infty F(x_n, v_n) \gamma_0^e(dx_0, dv_0) \prod_{k=1}^n R_k^e(x_{k-1}, v_{k-1}, dx_k, dv_k), \end{aligned}$$

which justifies the interpretation of this nonnegative measure as an unnormalized Feynman–Kac distribution. Associated with this integral representation is the recurrence relation $\gamma_k^e = \gamma_{k-1}^e R_k^e$, for any $k = 1, \dots, n$. Furthermore, using the decompositions (9) yields

$$\begin{aligned} \langle \gamma_n^e, F \rangle &= \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\text{imp}}(x_{k-1}, x_k)) \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \\ &= \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\text{imp}}(x_{k-1}, x_k)) \prod_{k=0}^n W_k^{\text{red}}(x_{k-1}, x_k) p_0(dx_0) \prod_{k=1}^n P_k(x_{k-1}, dx_k) \\ &= \mathbb{E}[F(X_n, \prod_{k=0}^n W_k^{\text{imp}}(X_{k-1}, X_k)) \prod_{k=0}^n W_k^{\text{red}}(X_{k-1}, X_k)], \end{aligned}$$

where $\{X_k, k = 0, 1, \dots, n\}$ is a Markov chain taking values in E and characterized by

- its initial probability distribution $p_0(dx)$,

- and its transition probabilities $P_k(x, dx')$, for any $k = 1, \dots, n$.

Introducing

$$M_0 = W_0^{\text{imp}}(X_0) \quad \text{and} \quad M_k = W_k^{\text{imp}}(X_{k-1}, X_k) M_{k-1},$$

for any $k = 1, \dots, n$, defines a multiplicative functional $\{M_k, k = 0, 1, \dots, n\}$ associated with the Markov chain $\{X_k, k = 0, 1, \dots, n\}$, and yields yet another representation as

$$\langle \gamma_n^e, F \rangle = \mathbb{E}[F(X_n, M_n) \prod_{k=0}^n W_k^{\text{red}}(X_{k-1}, X_k)] \quad \text{with} \quad M_n = \prod_{k=0}^n W_k^{\text{imp}}(X_{k-1}, X_k),$$

where jointly $\{(X_k, M_k), k = 0, 1, \dots, n\}$ form another Markov chain taking values in the product set $E \times [0, \infty)$, characterized by

- its initial probability distribution $p_0^e(dx, dv) = p_0(dx) \delta_{W_0^{\text{imp}}(x)}(dv)$,
- and its transition probabilities $P_k^e(x, v, dx', dv') = P_k(x, dx') \delta_{v W_k^{\text{imp}}(x, x')}(dv')$, for any $k = 1, \dots, n$.

Finally, let $W_0^e(x, v) = W_0^{\text{red}}(x)$ and $W_k^e(x, v, x', v') = W_k^{\text{red}}(x, x')$ for any $k = 1, \dots, n$.

5. PARTICLE APPROXIMATION USING A MULTIPLICATIVE FUNCTIONAL

Recall that the unnormalized distributions satisfy the recurrence relation

$$\gamma_k^e = \gamma_{k-1}^e R_k^e = \mu_{k-1}^e R_k^e \langle \gamma_{k-1}^e, 1 \rangle,$$

for any $k = 1, \dots, n$. Using the decompositions (9), notice that

$$\gamma_0^e(dx, dv) = W_0^{\text{red}}(x) p_0(dx) \delta_{W_0^{\text{imp}}(x)}(dv) = W_0^e(x, v) p_0^e(dx, dv),$$

and

$$R_k^e(x, v, dx', dv') = W_k^{\text{red}}(x, x') P_k(x, dx') \delta_{v W_k^{\text{imp}}(x, x')}(dv') = W_k^e(x, v, x', v') P_k^e(x, v, dx', dv'),$$

and introducing a weighted particle approximation of the form

$$\mu_k^e \approx \mu_k^{e,N} = \sum_{i=1}^N w_k^i \delta_{(\xi_k^i, v_k^i)} \quad \text{with} \quad \sum_{i=1}^N w_k^i = 1,$$

yields

$$\mu_{k-1}^{e,N} R_k^e(dx', dv') = \sum_{i=1}^N W_k^{\text{red}}(\xi_{k-1}^i, x') \underbrace{w_{k-1}^i P_k(\xi_{k-1}^i, dx') \delta_{v_{k-1}^i W_k^{\text{imp}}(\xi_{k-1}^i, x')}(dv')}_{m_k^i(dx', dv')},$$

which can be interpreted as the marginal nonnegative measure on the product set $E \times [0, \infty)$ associated with a nonnegative measure defined on the product set $\{1, \dots, N\} \times E \times [0, \infty)$. Using the auxiliary variable approach [11], the resulting ASIR algorithm can be described by

$$\gamma_0^{e,N} = \frac{1}{N} \sum_{i=1}^N W_0^{\text{red}}(\xi_0^i) \delta_{(\xi_0^i, v_0^i)} \quad \text{and} \quad \gamma_k^{e,N} = \frac{1}{N} \sum_{i=1}^N W_k^{\text{red}}(\xi_{k-1}^i, \xi_k^i) \delta_{(\xi_k^i, v_k^i)} \langle \gamma_{k-1}^{e,N}, 1 \rangle,$$

where $((\xi_0^1, v_0^1), \dots, (\xi_0^N, v_0^N))$ are i.i.d. random variables taking values in the product set $E \times [0, \infty)$ and with common probability distribution $p_0^e(dx, dv)$, or equivalently

$$\xi_0^i \sim p_0(dx) \quad \text{and} \quad v_0^i = W_0^{\text{imp}}(\xi_0^i),$$

independently for any $i = 1, \dots, N$, and where $((\tau_k^1, \xi_k^1, v_k^1), \dots, (\tau_k^N, \xi_k^N, v_k^N))$ are i.i.d. random variables taking values in the product set $\{1, \dots, N\} \times E \times [0, \infty)$ and with common probability distribution $(m_k^1(dx', dv'), \dots, m_k^N(dx', dv'))$, or equivalently

$$\tau_k^i \sim (w_{k-1}^1, \dots, w_{k-1}^N) \quad \text{and} \quad \xi_k^i \sim P_k(\xi_{k-1}^{\tau_k^i}, dx') \quad \text{and} \quad v_k^i = v_{k-1}^{\tau_k^i} W_k^{\text{imp}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i),$$

independently for any $i = 1, \dots, N$. In view of the relation (10), this results in the following approximations

$$\mu_0^{0,N} = \frac{\sum_{i=1}^N W_0^{\text{red}}(\xi_0^i) \delta_{\xi_0^i}}{\sum_{j=1}^N W_0^{\text{red}}(\xi_0^j)} = \sum_{i=1}^N w_0^i \delta_{\xi_0^i} \quad \text{and} \quad \mu_k^{0,N} = \frac{\sum_{i=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i) \delta_{\xi_k^i}}{\sum_{j=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^j}, \xi_k^j)} = \sum_{i=1}^N w_k^i \delta_{\xi_k^i},$$

for the normalized distributions under the reference model, which defines implicitly the resampling weights as

$$w_0^i = \frac{W_0^{\text{red}}(\xi_0^i)}{\sum_{j=1}^N W_0^{\text{red}}(\xi_0^j)} \quad \text{and} \quad w_k^i = \frac{W_k^{\text{red}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i)}{\sum_{j=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^j}, \xi_k^j)},$$

for any $i = 1, \dots, N$. Similarly, in view of the relation (11), this results in the following approximations

$$\mu_0^N = \frac{\sum_{i=1}^N W_0^{\text{red}}(\xi_0^i) v_0^i \delta_{\xi_0^i}}{\sum_{j=1}^N W_0^{\text{red}}(\xi_0^j) v_0^j} = \sum_{i=1}^N u_0^i w_0^i \delta_{\xi_0^i} \quad \text{and} \quad \mu_k^N = \frac{\sum_{i=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i) v_k^i \delta_{\xi_k^i}}{\sum_{j=1}^N W_k^{\text{red}}(\xi_{k-1}^{\tau_k^j}, \xi_k^j) v_k^j} = \sum_{i=1}^N u_k^i w_k^i \delta_{\xi_k^i},$$

for the normalized distributions under the alternate model, which defines implicitly the importance weights as

$$u_0^i = \frac{W_0^{\text{imp}}(\xi_0^i)}{\sum_{j=1}^N w_0^j W_0^{\text{imp}}(\xi_0^j)} \quad \text{and} \quad u_k^i = \frac{v_{k-1}^{\tau_k^i} W_k^{\text{imp}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i)}{\sum_{j=1}^N w_k^j v_{k-1}^{\tau_k^j} W_k^{\text{imp}}(\xi_{k-1}^{\tau_k^j}, \xi_k^j)},$$

for any $i = 1, \dots, N$.

Remark 5.1. It is clear that the particle positions, resampling weights and importance weights defined here are the same as the particle positions, resampling weights and importance weights defined in Sections 1 and 3.

For this algorithm, the following central limit theorem holds [5, Chapter 9], [2, Chapter 9].

Theorem 5.2.

$$\sqrt{N} \frac{\langle \gamma_n^{e,N} - \gamma_n^e, F \rangle}{\langle \gamma_n^e, 1 \rangle} \implies \mathcal{N}(0, V_n(F)),$$

in distribution as $N \uparrow \infty$, with asymptotic variance

$$V_n(F) = \frac{\text{var}(W_0^e R_{1:n}^e F, p_0^e)}{\langle p_0^e, W_0^e R_{1:n}^e 1 \rangle^2} + \sum_{k=1}^n \frac{\text{var}(W_k^e (R_{k+1:n}^e F) \circ \pi, p_k^e)}{\langle p_k^e, W_k^e (R_{k+1:n}^e 1) \circ \pi \rangle^2},$$

where $p_k^e = \mu_{k-1}^e \otimes P_k^e$ for any $k = 1, \dots, n$, and where

$$R_{k+1:n}^e F(x, v) = R_{k+1}^e \cdots R_n^e F(x, v) = \mathbb{E}[F(X_n, M_n) \prod_{p=k+1}^n W_p^{\text{red}}(X_{p-1}, X_p) \mid X_k = x, M_k = v],$$

for any $k = 0, 1, \dots, n$, with $R_{n+1:n}^e F(x, v) = F(x, v)$ by convention.

Remark 5.3 (CLT for normalizing constants). It follows from the identity

$$\lambda_0 \frac{\langle \gamma_n^{0,N} - \gamma_n^0, 1 \rangle}{\langle \gamma_n^0, 1 \rangle} + \lambda \frac{\langle \gamma_n^N - \gamma_n, 1 \rangle}{\langle \gamma_n, 1 \rangle} = \left\langle \frac{\gamma_n^{e,N} - \gamma_n^e}{\langle \gamma_n^e, 1 \rangle}, \lambda_0 + \lambda \frac{\langle \gamma_n^0, 1 \rangle}{\langle \gamma_n, 1 \rangle} (1 \otimes e) \right\rangle,$$

and from Theorem 5.2 that

$$\sqrt{N} \left[\lambda_0 \frac{\langle \gamma_n^{0,N} - \gamma_n^0, 1 \rangle}{\langle \gamma_n^0, 1 \rangle} + \lambda \frac{\langle \gamma_n^N - \gamma_n, 1 \rangle}{\langle \gamma_n, 1 \rangle} \right] \Longrightarrow \mathcal{N}(0, V_n(\lambda_0 + \lambda \frac{\langle \gamma_n^0, 1 \rangle}{\langle \gamma_n, 1 \rangle} (1 \otimes e))),$$

in distribution as $N \uparrow \infty$.

Remark 5.4 (CLT for normalized distributions). It follows from the decomposition

$$\begin{aligned} & \langle \mu_n^{0,N} - \mu_n^0, \phi_0 \rangle + \langle \mu_n^N - \mu_n, \phi \rangle \\ &= \left\langle \frac{\gamma_n^{e,N} - \gamma_n^e}{\langle \gamma_n^e, 1 \rangle}, \frac{\langle \gamma_n^e, 1 \rangle}{\langle \gamma_n^{e,N}, 1 \rangle} (\phi_0 - \langle \mu_n^0, \phi_0 \rangle) \otimes e_0 + \frac{\langle \gamma_n^e, 1 \otimes e \rangle}{\langle \gamma_n^{e,N}, 1 \otimes e \rangle} \frac{\langle \gamma_n^0, 1 \rangle}{\langle \gamma_n, 1 \rangle} (\phi - \langle \mu_n, \phi \rangle) \otimes e \right\rangle, \end{aligned}$$

from Theorem 5.2 and from the Slutsky lemma, that

$$\sqrt{N} \left[\langle \mu_n^{0,N} - \mu_n^0, \phi_0 \rangle + \langle \mu_n^N - \mu_n, \phi \rangle \right] \Longrightarrow \mathcal{N}(0, V_n((\phi_0 - \langle \mu_n^0, \phi_0 \rangle) \otimes e_0 + \frac{\langle \gamma_n^0, 1 \rangle}{\langle \gamma_n, 1 \rangle} (\phi - \langle \mu_n, \phi \rangle) \otimes e)),$$

in distribution as $N \uparrow \infty$.

As a consequence, it is worth finding an explicit expression, in terms of Feynman–Kac distributions defined on the set E , of the asymptotic variance $V_n(F)$ for functions of the form $F = \phi_0 \otimes e_0 + \phi \otimes e$ defined on the product set $E \times [0, \infty)$. Let

$$R_{k+1:n}^0 \phi_0(x) = R_{k+1}^0 \cdots R_n^0 \phi_0(x) \quad \text{and} \quad R_{k+1:n} \phi(x) = R_{k+1} \cdots R_n \phi(x),$$

for any $k = 0, 1, \dots, n$, with $R_{n+1:n}^0 \phi_0(x) = \phi_0(x)$ and $R_{n+1:n} \phi(x) = \phi(x)$ by convention. Then

$$\begin{aligned} & \frac{\text{var}(W_0^e R_{1:n}^e (\phi_0 \otimes e_0 + \frac{\langle \gamma_n^0, 1 \rangle}{\langle \gamma_n, 1 \rangle} \phi \otimes e), p_0^e)}{\langle p_0^e, W_0^e R_{1:n}^e 1 \rangle^2} \\ &= \frac{\text{var}(W_0^{\text{red}} R_{1:n}^0 \phi_0, p_0)}{\langle p_0, W_0^{\text{red}} R_{1:n}^0 1 \rangle^2} + \frac{\text{cov}(W_0^{\text{red}} R_{1:n}^0 \phi_0, W_0 R_{1:n} \phi, p_0)}{\langle p_0, W_0^{\text{red}} R_{1:n}^0 1 \rangle \langle p_0, W_0 R_{1:n} \phi \rangle} + \frac{\text{var}(W_0 R_{1:n} \phi, p_0)}{\langle p_0, W_0 R_{1:n} \phi \rangle^2}, \end{aligned}$$

and

$$\begin{aligned}
 & \frac{\text{var}(W_k^e (R_{k+1:n}^e (\phi_0 \otimes e_0 + \frac{\langle \gamma_n^0, 1 \rangle}{\langle \gamma_n, 1 \rangle} \phi \otimes e)) \circ \pi, \mu_{k-1}^e \otimes P_k^e)}{\langle \mu_{k-1}^e \otimes P_k^e, W_k^e (R_{k+1:n}^e 1) \circ \pi \rangle^2} \\
 &= \frac{\text{var}(W_k^{\text{red}} (R_{k+1:n}^0 \phi_0) \circ \pi, \mu_{k-1}^0 \otimes P_k)}{\langle \mu_{k-1}^0, R_{k:n}^0 1 \rangle^2} \\
 &+ 2 \frac{\text{cov}(W_k^{\text{red}} (R_{k+1:n}^0 \phi_0) \circ \pi, W_k (R_{k+1:n} \phi) \circ \pi, \mu_{k-1} \otimes P_k)}{\langle \mu_{k-1}^0, R_{k:n}^0 1 \rangle \langle \mu_{k-1}, R_{k:n} 1 \rangle} \\
 &+ 2 \frac{\langle \mu_{k-1}, R_{k:n}^0 \phi_0 \rangle - \langle \mu_{k-1}^0, R_{k:n}^0 \phi_0 \rangle}{\langle \mu_{k-1}^0, R_{k:n}^0 1 \rangle} \\
 &+ \frac{\langle \gamma_{k-1}^\square, 1 \rangle \langle \gamma_{k-1}^0, 1 \rangle}{\langle \gamma_{k-1}, 1 \rangle^2} \frac{\text{var}(W_k (R_{k+1:n} \phi) \circ \pi, \mu_{k-1}^\square \otimes P_k)}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} \\
 &+ \left[\frac{\langle \gamma_{k-1}^\square, 1 \rangle \langle \gamma_{k-1}^0, 1 \rangle}{\langle \gamma_{k-1}, 1 \rangle^2} \frac{\langle \mu_{k-1}^\square, R_{k:n} \phi \rangle^2}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} - \frac{\langle \mu_{k-1}, R_{k:n} \phi \rangle^2}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} \right],
 \end{aligned}$$

for any $k = 1, \dots, n$.

Proposition 5.5. *In particular for the normalizing constant*

$$\sqrt{N} \frac{\langle \gamma_n^N - \gamma_n, 1 \rangle}{\langle \gamma_n, 1 \rangle} \implies \mathcal{N}(0, V_n),$$

in distribution as $N \uparrow \infty$, with asymptotic variance

$$\begin{aligned}
 V_n &= \frac{\text{var}(W_0 R_{1:n} \phi, p_0)}{\langle p_0, W_0 R_{1:n} 1 \rangle^2} + \sum_{k=1}^n \frac{\langle \gamma_{k-1}^\square, 1 \rangle \langle \gamma_{k-1}^0, 1 \rangle}{\langle \gamma_{k-1}, 1 \rangle^2} \frac{\text{var}(W_k (R_{k+1:n} 1) \circ \pi, \mu_{k-1}^\square \otimes P_k)}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} \\
 &+ \sum_{k=1}^n \left[\frac{\langle \gamma_{k-1}^\square, 1 \rangle \langle \gamma_{k-1}^0, 1 \rangle}{\langle \gamma_{k-1}, 1 \rangle^2} \frac{\langle \mu_{k-1}^\square, R_{k:n} 1 \rangle^2}{\langle \mu_{k-1}, R_{k:n} 1 \rangle^2} - 1 \right].
 \end{aligned}$$

Remark 5.6. It is clear that the expression for the asymptotic variance given here is the same as the expression given in Proposition 3.5.

6. CONCLUSION

Even though the two different approaches used here to obtain a central limit theorem for the proposed particle approximation do actually provide the same explicit expression for the asymptotic variance, they differ in the following aspects.

In the first approach, based on a representation in path-space, the importance weight functions appear only in the test function, and it is therefore easy to analyze the joint particle approximation of unnormalized distributions (and normalizing constants and normalized distributions, as a by-product) for the reference model and for several alternate models at the same time, just by choosing the appropriate test function in the central limit theorem. In view of (7), the idea is simply to use one test function per alternate model, i.e.

$$\frac{\langle \gamma_n^{a,N} - \gamma_n^a, \phi_a \rangle}{\langle \gamma_n^a, 1 \rangle} = \frac{\langle \gamma_n^\bullet, 1 \rangle}{\langle \gamma_n^a, 1 \rangle} \frac{\langle \gamma_n^{\bullet,N} - \gamma_n^\bullet, T_n^a \phi_a \rangle}{\langle \gamma_n^\bullet, 1 \rangle},$$

hence

$$\sqrt{N} \sum_{a \in A} \frac{\langle \gamma_n^{a,N} - \gamma_n^a, \phi_a \rangle}{\langle \gamma_n^a, 1 \rangle} \implies \mathcal{N}(0, V_n(\sum_{a \in A} \frac{\langle \gamma_n^\bullet, 1 \rangle}{\langle \gamma_n^a, 1 \rangle} T_n^a \phi_a)),$$

in distribution as $N \uparrow \infty$, with a correlation structure reflected in the asymptotic covariance matrix, easily obtained by polarization. In other words, this first approach seems appropriate to analyze particle approximations in statistical models depending on some parameter, in sensitivity analysis, etc.

In the second approach, based on a representation in terms of a multiplicative functional, the importance weight functions appear in the extended Markov model, and it would be necessary to change the Markov model to analyze the joint particle approximation of unnormalized distributions for the reference model and several alternate models at the same time. On the other hand, it would be easy with this Markov interpretation to analyze particle approximation with adaptive resampling schemes, where the decision to use resampling weights only vs. importance weights only is made dependent on an empirical criterion (effective number of particles, entropy of the sample, etc.) evaluated using the current particle approximation. The idea would be simply, as a generalization of (3), to introduce a factorization depending on the current normalized or unnormalized distribution on the product set $E \times [0, \infty)$: this would result in a representation in terms of a McKean model, and the associated particle approximation could be easily analyzed [5].

REFERENCES

- [1] Olivier Cappé, Randal Douc, Éric Moulines, and Christian P. Robert. On the convergence of the Monte Carlo maximum likelihood method for latent variable models. *Scandinavian Journal of Statistics*, 29(4):615–635, December 2002.
- [2] Olivier Cappé, Éric Moulines, and Tobias Rydén. *Inference in Hidden Markov Models*. Springer Series in Statistics. Springer–Verlag, New York, 2005.
- [3] Natacha Caylus, Arnaud Guyader, François Le Gland, and Nadia Oudjane. Application du filtrage particulaire à l’inférence statistique des HMM. In *Actes des 36èmes Journées de Statistique, Montpellier*. SFdS, May 2004.
- [4] Frédéric Cérou, François Le Gland, and Nigel J. Newton. Stochastic particle methods for linear tangent filtering equations. In José-Luis Menaldi, Edmundo Rofman, and Agnès Sulem, editors, *Optimal Control and Partial Differential Equations. In honour of professor Alain Bensoussan’s 60th birthday*, pages 231–240. IOS Press, Amsterdam, 2001.
- [5] Pierre Del Moral. *Feynman–Kac Formulae. Genealogical and Interacting Particle Systems with Applications*. Probability and its Applications. Springer–Verlag, New York, 2004.
- [6] Arnaud Doucet and Vladislav B. Tadić. Parameter estimation in general state–space models using particle methods. *Annals of the Institute of Statistical Mathematics*, 55(2):409–422, June 2003.
- [7] Charles J. Geyer. On the convergence of Monte Carlo maximum likelihood calculations. *Journal of the Royal Statistical Society, Series B*, 56(1):261–274, 1994.
- [8] Charles J. Geyer. Estimation and optimization of functions. In Walter R. Gilks, Sylvia Richardson, and David J. Spiegelhalter, editors, *Markov Chain Monte Carlo in Practice*, chapter 14, pages 241–258. Chapman & Hall, London, 1996.
- [9] Arnaud Guyader, François Le Gland, and Nadia Oudjane. A particle implementation of the recursive MLE for partially observed diffusions. In *Proceedings of the 13th Symposium on System Identification (SYSID), Rotterdam*, pages 1305–1310. IFAC / IFORS, August 2003.
- [10] Jun S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer Series in Statistics. Springer–Verlag, New York, 2001.
- [11] Michael K. Pitt and Neil Shephard. Filtering via simulations : auxiliary particle filter. *Journal of the American Statistical Association*, 94(446):590–599, June 1999.