ESAIM: M2AN 44 (2010) 947–975
DOI: 10.1051/m2an/2010048
ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

A BACKWARD PARTICLE INTERPRETATION OF FEYNMAN-KAC FORMULAE

Pierre Del Moral[1], Arnaud Doucet[2,3] and Sumeetpal S. Singh[4]

Abstract. We design a particle interpretation of Feynman-Kac measures on path spaces based on a backward Markovian representation combined with a traditional mean field particle interpretation of the flow of their final time marginals. In contrast to traditional genealogical tree based models, these new particle algorithms can be used to compute normalized additive functionals "on-the-fly" as well as their limiting occupation measures with a given precision degree that does not depend on the final time horizon. We provide uniform convergence results w.r.t. the time horizon parameter as well as functional central limit theorems and exponential concentration estimates, yielding what seems to be the first results of this type for this class of models. We also illustrate these results in the context of filtering of hidden Markov models, as well as in computational physics and imaginary time Schroedinger type partial differential equations, with a special interest in the numerical approximation of the invariant measure associated to h-processes.

Mathematics Subject Classification. 65C05, 65C35, 60G35, 47D08.

Received July 19, 2009. Revised February 15, 2010.
Published online August 26, 2010.

1. Introduction

Let $(E_n)_{n\geq 0}$ be a sequence of measurable spaces equipped with some σ-fields $(\mathcal{E}_n)_{n\geq 0}$, and we let $\mathcal{P}(E_n)$ be the set of all probability measures over the set E_n. We let X_n be a Markov chain with Markov transition M_n on E_n, and we consider a sequence of $(0,1]$-valued potential functions G_n on the set E_n. The Feynman-Kac path measure associated with the pairs (M_n, G_n) is the probability measure $\mathbb{Q}_n$ on the product state space

Keywords and phrases. Feynman-Kac models, mean field particle algorithms, functional central limit theorems, exponential concentration, non asymptotic estimates.

[1] Centre INRIA Bordeaux et Sud-Ouest & Institut de Mathématiques de Bordeaux, Université de Bordeaux I, 351 cours de la Libération, 33405 Talence Cedex, France. Pierre.Del-Moral@inria.fr

[2] Department of Statistics & Department of Computer Science, University of British Columbia, 333-6356 Agricultural Road, Vancouver, BC, V6T 1Z2, Canada. arnaud@stat.ubc.ca

[3] The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan.

[4] Department of Engineering, University of Cambridge, Trumpington Street, CB2 1PZ, UK. sss40@cam.ac.uk

Article published by EDP Sciences

$E_{[0,n]} := \prod_{0\leq p\leq n} E_p$ defined by the following formula

$$\mathrm{d}\mathbb{Q}_n := \frac{1}{\mathcal{Z}_n} \left\{ \prod_{0\leq p<n} G_p(X_p) \right\} \mathrm{d}\mathbb{P}_n \tag{1.1}$$

where $\mathcal{Z}_n$ is a normalizing constant and $\mathbb{P}_n$ is the distribution of the paths $(X_p)_{0\leq p\leq n}$ of the Markov process X_p from the origin $p = 0$, up to the current time $p = n$. We also denote by $\Gamma_n = \mathcal{Z}_n\ \mathbb{Q}_n$ the unnormalized version of (1.1). In this article we design a particle interpretation of $\mathbb{Q}_n$ based on a backward Markovian representation combined with a traditional mean field particle interpretation of the flow of their final time marginals.

These distributions on path spaces arise in a variety of application areas, including nonlinear filtering, Bayesian inference, branching processes in biology, particle absorption problems in physics, and many other instances. We refer the reader to the books [7,16] and references therein. In filtering problems, we are given a Markov process (X_n, Y_n) taking values in some product space $(E_n \times E'_n)$, with transition probabilities of the following form

$$\mathbb{P}\left((X_n, Y_n) \in \mathrm{d}(x_n, y_n) | (X_{n-1}, Y_{n-1}) = (x_{n-1}, y_{n-1})\right) = M_n(x_{n-1}, \mathrm{d}x_n)\ g_n(x_n, y_n)\ \lambda_n(\mathrm{d}y_n)$$

with some density functions g_n w.r.t. some reference measure λ_n on E'_n. In this model $(X_n)_{n\geq 0}$ is unobserved, or the hidden Markov process, while only $(Y_n)_{n\geq 0}$ is observed. Given a series of observations $Y_0 = y_0, \ldots, Y_n = y_n$, and setting $G_n := g_n(\,.\,, y_n)$ in (1.1) we find that

$$\mathbb{Q}_n = \mathrm{Law}((X_0, \ldots, X_n) \mid \forall 0 \leq p < n,\quad Y_p = y_p).$$

Feynman-Kac models also play a central role in the numerical analysis of certain partial differential equations. It offers a natural way to solve these functional integral models by simulating random paths of stochastic processes. These Feynman-Kac models were originally presented by Kac in 1949 [21] for continuous time processes. Since then they have also been used in molecular chemistry and computational physics to calculate the ground state energy of some Hamiltonian operators associated with some potential function V describing the energy of a molecular configuration (see for instance [3,10,17,28], and references therein). To better connect these partial differential equation models with (1.1), let us assume that $M_n(x_{n-1}, \mathrm{d}x_n)$ is the Markov probability transition $X_n = x_n \rightsquigarrow X_{n+1} = x_{n+1}$ coming from a time discretization $X_n = X'_{t_n}$ of a continuous time E-valued Markov process X'_t. Let the time mesh be $(t_n)_{n\geq 0}$ with time step $(t_n - t_{n-1}) = \Delta t$. For potential functions of the form $G_n = \mathrm{e}^{-V\Delta t}$, the measures $\mathbb{Q}_n \simeq_{\Delta t \to 0} \mathbb{Q}_{t_n}$ represents the time discretization of the following distribution:

$$\mathrm{d}\mathbb{Q}_t = \frac{1}{\mathcal{Z}_t}\ \exp\left(-\int_0^t\ V(X'_s)\ \mathrm{d}s\right)\ \mathrm{d}\mathbb{P}_t^{X'} \tag{1.2}$$

where $\mathbb{P}_t^{X'}$ stands for the distribution of the random paths $(X'_s)_{0\leq s\leq t}$ with a given infinitesimal generator L. The errors introduced by the discretization of time discussed above are well understood for regular models, we refer the interested reader to [12,14,24,25] in the context of nonlinear filtering. The marginal distributions γ_t at time t of the unnormalized measures $\mathcal{Z}_t\ \mathrm{d}\mathbb{Q}_t$ are the solution of the so-called imaginary time Schroedinger equation, given in weak formulations[5] on every sufficiently regular function f by

$$\frac{\mathrm{d}}{\mathrm{d}t}\ \gamma_t(f) := \gamma_t(L^V(f)) \quad \text{with} \quad L^V = L - V. \tag{1.3}$$

[5] Consult the last paragraph of this section for a statement of the notation used in this article.

Another way to see how these imaginary time Schroedinger equations fits into the above abstract setting is to observe that the semigroup $(P_t^V)_{t\geq 0}$ of the above evolution equation has the following Feynman-Kac path-integral representation

$$\gamma_t(f) = \gamma_0(P_t^V(f)) \quad \text{with} \quad P_t^V(f)(x) = \mathbb{E}\left(f(X'_t)\ \exp\left\{-\int_0^t V(X'_s)\mathrm{d}s\right\}|X'_0 = x\right).$$

The above formulation is called the probabilistic solution to the evolution equation (1.3), and the above functional representation is customarily called the Feynman-Kac formula, in reference with the pioneering works by these two researchers on this subject in the beginning of the fifties. We let $E = \mathbb{R}^d$, and L be a second order elliptic differential operator written as

$$L(f) = \frac{1}{2}\sum_{i,j=1}^d a^{i,j}\partial_{x^i}\partial_{x^j}(f) + \sum_{i=1}^d b^i\partial_{x^i}(f)$$

for C^∞ functions a, b and the matrix $(a^{i,j}(x))$ is non-negative definite and invertible at each x. In this situation, the Markov process X'_t with infinitesimal operator L is given by the following Itô stochastic differential equations:

$$\mathrm{d}X'_t = b(X'_t)\mathrm{d}t + \sigma(X'_t)\mathrm{d}B_t \tag{1.4}$$

where B_t is a d-dimensional Brownian motion and $\sigma = (\sigma^{i,j})_{1\leq i,j\leq d}$ is chosen so that $\sigma\sigma^T = a$. We can also consider diffusion models on a d-dimensional smooth compact manifold E. In this situation, the elliptic differential operator is again written as above in a given chart, and ∂_{x^i} is the differential operator w.r.t. the ith coordinate of the underlying chart. In this context, we can also associate to the infinitesimal generator L a Markov process with continuous trajectories (see for instance the book of Émery [18], and Ikeda and Watanabe [20]).

Continuous time filtering problems are also connected in a natural way to stochastic partial differential equations. In this situation, the signal X'_t is given by a time homogeneous Markov process of the form (1.4), and the observation process is an $\mathbb{R}^{d'}$-valued process defined by

$$\mathrm{d}Y_t = H(X'_t)\,\mathrm{d}t + \sigma\,\mathrm{d}V_t \tag{1.5}$$

where V_t is a d'-vector standard Wiener process independent of the signal, and H is some regular function from $\mathbb{R}^d$ into $\mathbb{R}^{d'}$. We let t_n, $n \geq 0$, be a given time mesh with $t_0 = 0$ and $t_n \leq t_{n+1}$. Also let $X_n := X'_{[t_n,t_{n+1}]}$ be the Markov chain taking values at each time n in the space $E_n = C([t_n, t_{n+1}], \mathbb{R}^d)$ of continuous paths from $[t_n, t_{n+1}]$ into $\mathbb{R}^d$. Given the observation path Y_s, $0 \leq s \leq t_n$, we define the "random" potential functions G_n on E_n by setting for any $x_n = (x'_n(s))_{t_n\leq s\leq t_{n+1}} \in E_n$,

$$G_n(x_n) = \exp\left(\int_{t_n}^{t_{n+1}} H^\star(x'_n(s))\ \mathrm{d}Y_s - \frac{1}{2}\int_{t_n}^{t_n} H^\star(x'_n(s))H(x'_n(s))\ \mathrm{d}s\right).$$

In the above definition $(.)^\star$ stands for the transposition operator. The Kallianpur-Striebel formula (see for instance [22]) states that

$$\mathbb{Q}_n = \mathrm{Law}((X'_{[0,t_{n+1}]}) \mid \mathcal{Y}_{t_n}) \tag{1.6}$$

where $\mathcal{Y}_{t_n} = \sigma(Y_s,\ s \leq t_n)$ represents the sigma-field generated by the observation process. It is well known that the marginal $\eta_{t_n} := \mathrm{Law}(X'_{t_n} \mid \mathcal{Y}_{t_n})$ of $\mathbb{Q}_n$ w.r.t. time t_n is solution of the Kushner-Stratonovich stochastic partial differential equation defined below for sufficiently regular test functions f on $\mathbb{R}^d$

$$\eta_{t_n}(f) = \nu_0(f) + \int_0^{t_n} \eta_s(L(f))\mathrm{d}s + \int_0^{t_n} [\eta_s(H^\star f) - \eta_s(H^\star)\eta_s(f)]\ (\mathrm{d}Y_s - \eta_s(H)\mathrm{d}s)\,. \tag{1.7}$$

In the above displayed formula ν_0 stands for the initial distribution of the random state X'_0. To obtain a computationally feasible solution, a traditional discrete time approximation of the above conditional measures consists of selecting a sequence of meshes $(t_{n+1} - t_n) = \Delta > 0$ and an Euler type approximation $\left(X'^{,\Delta}_{t_n}, Y^{\Delta}_{t_n}\right)$ of the pair diffusion processes (X'_t, Y_t) given in (1.4) and (1.5). If we set in (1.1), $X_n = X'^{,\Delta}_{t_n}$ and for any $x_n \in \mathbb{R}^d$

$$G_n(x_n) := \exp\left(H^\star(x_n)\left(Y^{\Delta}_{t_{n+1}} - Y^{\Delta}_{t_n}\right) - \frac{1}{2}\,H^\star(x_n)H(x'_{t_n})\Delta\right)$$

then, we find that

$$\mathbb{Q}_n = \mathrm{Law}((X'^{,\Delta}_0, \ldots, X'^{,\Delta}_{t_{n+1}}) \mid Y^{\Delta}_0, \ldots, Y^{\Delta}_{t_n}).$$

For further details on these approximation models, we refer the reader to [6]. The robust version of the above continuous time filtering problem has the same form as (1.2), for some non homogeneous potential functions V_t that depends on the observation process. For further details on this model, we refer the reader to [9].

The mean field particle approximation models associated with the continuous time flow of measures (1.3), (1.6) and (1.7) are discussed in [6,9,10] (see also [3,17,28]). These interacting stochastic models can be interpreted in different ways depending on the application domain. In advanced signal processing, these stochastic models are called particle filters or sequential Monte Carlo methods. In molecular chemistry, these evolutionary type models are often interpreted as a quantum or diffusion Monte Carlo model. In this context, particles often are referred as walkers, to distinguish the virtual particle-like objects to physical particles, like electrons of atoms. The numerical approximation of *the full path-space measures* $\mathbb{Q}_t$, and their discrete time versions defined in (1.1), is clearly much more involved, mainly because it requires the use of stochastic analysis tools for Markov processes on path spaces. A standard strategy is to keep track of the particle ancestral information in the mean field particle interpretation of (1.3). The corresponding path-particle approximation model coincides with that of the traditional genealogical tree based particle approximation model (see for instance [7], and references therein).

In this article, we design an original numerical approximation scheme for the discrete time distributions $\mathbb{Q}_n$ based on the simulation of a sequence of mean field interacting particle systems. In contrast to the more traditional genealogical tree based approximations (see for instance [7]), the particle model presented in this article can approximate additive functionals of the form

$$\overline{F}_n(x_0, \ldots, x_n) = \frac{1}{(n+1)} \sum_{0\leq p\leq n} f_p(x_p), \tag{1.8}$$

where $\|f_p\| \leq 1$, with an asymptotic variance which decreases to zero with the time horizon n. Moreover this computation can be done "on-the-fly". We now summarize our main results.

Let $\mathbb{Q}^N_n$ denote the N-particle approximation of $\mathbb{Q}_n$. (The precision of the algorithm corresponds to the size N of the particle system.) Under some appropriate regularity properties, we can calibrate the performance of the model using the following uniform and non asymptotic Gaussian concentration estimate

$$\frac{1}{N}\log \sup_{n\geq 0} \mathbb{P}\left(\left|[\mathbb{Q}^N_n - \mathbb{Q}_n](\overline{F}_n)\right| \geq \frac{b}{\sqrt{N}} + \epsilon\right) \leq -\epsilon^2/(2b^2)$$

for any $\epsilon > 0$, and for some finite constant $b < \infty$. We also prove that the asymptotic variance of $\sqrt{N}[\mathbb{Q}^N_n - \mathbb{Q}_n](\overline{F}_n)$ decreases to zero at rate $1/n$. In newly obtained results [13], we have established the following non asymptotic bias and variance bounds

$$\sup_{n\geq 0}\left|\mathbb{E}\left(\mathbb{Q}^N_n(\overline{F}_n)\right) - \mathbb{Q}_n(\overline{F}_n)\right| \leq \frac{c}{N} \quad \text{and} \quad \forall n \geq 0 \quad \mathbb{E}\left(\left[\mathbb{Q}^N_n(\overline{F}_n) - \mathbb{Q}_n(\overline{F}_n)\right]^2\right) \leq \frac{c}{N}\,\left(\frac{1}{n+1} + \frac{1}{N}\right) \tag{1.9}$$

where c is some finite constant that does not depend on the time parameter n. Thus, for any large time horizon $n \geq N$, the above r.h.s. mean square error is of order $1/N^2$.

In the context of filtering, $\mathbb{Q}_n^N$ corresponds to the sequential Monte Carlo approximation of the forward filtering backward smoothing recursion [19]. Recently, another theoretical study of this problem was undertaken by [15] – our results complement theirs. We provide new functional central limit theorems, non-asymptotic bias and variance bounds (see [13]) as well as uniform exponential concentration inequalities. Additionally, we show how the forward filtering backward smoothing estimates of additive functionals can be computed using a forward only recursion. This has applications to online parameter estimation for non-linear non-Gaussian state-space models [13,26].

For time homogeneous models $(M_n, f_n, G_n) = (M, f, G)$ with a lower bounded potential function $G > \delta$, and a M-reversible transition w.r.t. to some probability measure μ s.t. $M(x, .) \sim \mu$ and $(M(x, .)/\mathrm{d}\mu) \in \mathbb{L}_2(\mu)$, it can be established that $\mathbb{Q}_n(F_n)$ converges to $\mu_h(f)$, as $n \to \infty$, with the measure μ_h defined below

$$\mu_h(\mathrm{d}x) := \frac{1}{\mu(hM(h))}\ h(x)\ M(h)(x)\ \mu(\mathrm{d}x).$$

In the above display, h is a positive eigenmeasure associated with the largest eigenvalue of the integral operator $Q(x, \mathrm{d}y) = G(x)M(x, \mathrm{d}y)$ on $\mathbb{L}_2(\mu)$ (see for instance Sect. 12.4 in [7]). This measure μ_h is in fact the invariant measure of the h-process defined as the Markov chain with elementary Markov transitions $M_h(x, \mathrm{d}y) \propto M(x, \mathrm{d}y)h(y)$. As the initiated reader will have certainly noticed, the above convergence result is only valid under some appropriate mixing conditions on the h-process. The long time behavior of these h-processes and their connections to various applications areas of probability, analysis, geometry and partial differential equations, has been the subject of many papers for many years in applied probability. In our framework, using elementary manipulations, the Gaussian estimate given above can be used to calibrate the convergence of the particle estimate $\mathbb{Q}_n^N(F_n)$ towards $\mu_h(f)$, as the pair of parameters N and $n \to \infty$.

The rest of this article is organized as follows:

In Section 2, we describe the mean field particle models used to design the particle approximation measures $\mathbb{Q}_n^N$. In Section 3, we state the main results presented in this article, including a functional central limit theorem, and non asymptotic mean error bounds. Section 4 is dedicated to a key backward Markov chain representation of the measures $\mathbb{Q}_n$. The analysis of our particle approximations is provided in Section 5. Sections 6 and 7 detail the proof of the two main theorems presented in Section 3. In Section 8, we provide some comparisons between the backward particle model discussed in this article and the more traditional genealogical tree based particle model. In the final Section 9, we present a numerical example which is relevant for the problem of maximum likelihood estimation of the model parameters of a Hidden Markov Model (HMM). The gradient of the likelihood of the observed data can be computed using an Infinitesimal Perturbation Analysis (IPA) representation recently proposed in [5], and we compare the numerical implementations of this gradient using our backward particle model and the genealogical tree based particle model. The former is shown to have significantly less variance when both methods are implemented with the same number of particles.

Results contained in standard probability texts such as the book of Shiryaev [29], Billingsley [2] or the books of Stroock [30,31], are assumed and used without reference, as well are results from measure theory and elementary functional analysis. A few results on local sampling estimates and on the fluctuations of Feynman-Kac particle models are also needed and references for these are included. The study of any stochastic particle model often involves the analysis of the stability properties of the semigroup associated with some limiting measure valued equation. Instead of writing summations or integrals with respect to some density functions, these semigroups are preferably expressed in terms of integral operators acting on measures or on test functions. This integral operator framework is notationally lighter and of great practical value. Functional inequalities may be used and it also enters in a natural way contraction coefficients of semigroup operators, and other related quantities. For the convenience of the reader, we end this introduction with a summary of the notation used in the present article, and also present some basic definitions from probability and operator semigroup theory.

We denote respectively by $\mathcal{M}(E)$, and $\mathcal{B}(E)$, the set of all finite signed measures on some measurable space $(E, \mathcal{E})$, and the Banach space of all bounded and measurable functions f equipped with the uniform norm $\|f\|$. We let $\mu(f) = \int \mu(\mathrm{d}x)\, f(x)$, be the Lebesgue integral of a function $f \in \mathcal{B}(E)$, with respect to a measure

$\mu \in \mathcal{M}(E)$. We recall that a bounded integral kernel $M(x, \mathrm{d}y)$ from a measurable space $(E, \mathcal{E})$ into an auxiliary measurable space $(E', \mathcal{E}')$ is an operator $f \mapsto M(f)$ from $\mathcal{B}(E')$ into $\mathcal{B}(E)$ such that the functions

$$x \mapsto M(f)(x) := \int_{E'} M(x, \mathrm{d}y) f(y)$$

are $\mathcal{E}$-measurable and bounded, for any $f \in \mathcal{B}(E')$. In the above displayed formulae, $\mathrm{d}y$ stands for an infinitesimal neighborhood of a point y in E'. The kernel M also generates a dual operator $\mu \mapsto \mu M$ from $\mathcal{M}(E)$ into $\mathcal{M}(E')$ defined by $(\mu M)(f) := \mu(M(f))$. A Markov kernel is a positive and bounded integral operator M with $M(1) = 1$. Given a pair of bounded integral operators (M_1, M_2), we let $(M_1 M_2)$ the composition operator defined by $(M_1 M_2)(f) = M_1(M_2(f))$. For time homogeneous state spaces, we denote by $M^m = M^{m-1}M = MM^{m-1}$ the mth composition of a given bounded integral operator M, with $m \geq 1$. In the context of finite state spaces, these integral operations coincide with the traditional matrix operations on multidimensional state spaces. Given a positive function G on E, we let $\Psi_G : \eta \in \mathcal{P}(E) \mapsto \Psi_G(\eta) \in \mathcal{P}(E)$ be the Bayes-Boltzmann-Gibbs transformation defined by

$$\Psi_G(\eta)(\mathrm{d}x) := \frac{1}{\eta(G)}\, G(x)\, \eta(\mathrm{d}x).$$

2. Description of the models

The numerical approximation of the path-space distributions (1.1) requires extensive calculations. The mean field particle interpretation of these models are based on the fact that the flow of the nth time marginals η_n of the measures $\mathbb{Q}_n$ satisfy a nonlinear evolution equation of the following form

$$\eta_{n+1}(\mathrm{d}y) = \int \eta_n(\mathrm{d}x) K_{n+1,\eta_n}(x, \mathrm{d}y) \tag{2.1}$$

for some collection of Markov transitions $K_{n+1,\eta}$, indexed by the time parameter $n \geq 0$ and the set of probability measures $\mathcal{P}(E_n)$. The mean field particle interpretation of the nonlinear measure valued model (2.1) is the E_n^N-valued Markov chain

$$\xi_n = \left(\xi_n^1, \xi_n^2, \ldots, \xi_n^N\right) \in E_n^N$$

with elementary transitions defined as

$$\mathbb{P}\left(\xi_{n+1} \in \mathrm{d}x \mid \xi_n\right) = \prod_{i=1}^N K_{n+1,\eta_n^N}(\xi_n^i, \mathrm{d}x^i) \quad \text{with} \quad \eta_n^N := \frac{1}{N} \sum_{j=1}^N \delta_{\xi_n^j}. \tag{2.2}$$

In the above displayed formula, $\mathrm{d}x$ stands for an infinitesimal neighborhood of the point $x = (x^1, \ldots, x^N) \in E_{n+1}^N$. The initial system ξ_0 consists of N independent and identically distributed random variables with common law η_0. We let $\mathcal{F}_n^N := \sigma(\xi_0, \ldots, \xi_n)$ be the natural filtration associated with the N-particle approximation model defined above. The resulting particle model coincides with a genetic type stochastic algorithm $\xi_n \rightsquigarrow \widehat{\xi}_n \rightsquigarrow \xi_{n+1}$ with selection transitions $\xi_n \rightsquigarrow \widehat{\xi}_n$ and mutation transitions $\widehat{\xi}_n \rightsquigarrow \xi_{n+1}$ dictated by the potential (or fitness) functions G_n and the Markov transitions M_{n+1}.

During the selection stage $\xi_n \rightsquigarrow \widehat{\xi}_n$, for every index i, with a probability $\epsilon_n G_n(\xi_n^i)$, we set $\widehat{\xi}_n^i = \xi_n^i$, otherwise we replace ξ_n^i with a new individual $\widehat{\xi}_n^i = \xi_n^j$ randomly chosen from the whole population with a probability proportional to $G_n(\xi_n^j)$. The parameter $\epsilon_n \geq 0$ is a tuning parameter that must satisfy the constraint $\epsilon_n G_n(\xi_n^i) \leq 1$, for every $1 \leq i \leq N$. For $\epsilon_n = 0$, the resulting proportional selection transition corresponds to the so-called simple genetic model. During the mutation stage, the selected particles $\widehat{\xi}_n^i \rightsquigarrow \xi_{n+1}^i$ evolve independently according to the Markov transitions M_{n+1}.

If we interpret the selection transition as a birth and death process, then arises the important notion of the ancestral line of a current individual. More precisely, when a particle $\widehat{\xi}_{n-1}^i \longrightarrow \xi_n^i$ evolves to a new location ξ_n^i, we can interpret $\widehat{\xi}^i_{n-1}$ as the parent of ξ^i_n. Looking backwards in time and recalling that the particle $\widehat{\xi}^i_{n-1}$ has selected a site ξ^j_{n-1} in the configuration at time $(n-1)$, we can interpret this site ξ^j_{n-1} as the parent of $\widehat{\xi}^i_{n-1}$ and therefore as the ancestor denoted $\xi^i_{n-1,n}$ at level $(n-1)$ of ξ^i_n. Running backwards in time we may trace the whole ancestral line

$$\xi^i_{0,n} \longleftarrow \xi^i_{1,n} \longleftarrow \ldots \longleftarrow \xi^i_{n-1,n} \longleftarrow \xi^i_{n,n} = \xi^i_n. \tag{2.3}$$

More interestingly, the occupation measure of the corresponding N-genealogical tree model converges as $N \to \infty$ to the conditional distribution $\mathbb{Q}_n$. For any function F_n on the path space $E_{[0,n]}$, we have the following convergence (to be stated precisely later) as $N \to \infty$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N F_n(\xi^i_{0,n}, \xi^i_{1,n}, \ldots, \xi^i_{n,n}) = \int \mathbb{Q}_n(\mathrm{d}(x_0,\ldots,x_n))\ F_n(x_0,\ldots,x_n). \tag{2.4}$$

This convergence result can be refined in various directions. Nevertheless, even under very favourable mixing assumptions, the asymptotic variance $\sigma^2_n(F_n)$ of the above occupation measure around $\mathbb{Q}_n$ increases quadratically with the final time horizon n for additive functions of the form

$$F_n(x_0,\ldots,x_n) = \sum_{0\le p\le n} f_p(x_p) \Rightarrow \sigma^2_n(F_n) \simeq n^2 \tag{2.5}$$

comprised of some collection of non negative functions f_p on E_p. To be more precise, let us examine a time homogeneous model $(E_n, f_n, G_n, M_n) = (E, f, G, M)$ with constant potential functions $G_n = 1$ and mutation transitions M s.t. $\eta_0 M = \eta_0$. For the choice of the tuning parameter $\epsilon = 0$, using the asymptotic variance formulae in [7], equation (9.13), p. 304, for any function f s.t. $\eta_0(f) = 0$ and $\eta_0(f^2) = 1$ we prove that

$$\sigma^2_n(F_n) = \sum_{0\le p\le n} \mathbb{E}\left(\left[\sum_{0\le q\le n} M^{(q-p)_+}(f)(X_q)\right]^2\right)$$

with the positive part $a_+ = \max(a, 0)$ and the convention $M^0 = Id$, the identity transition. For $M(x, \mathrm{d}y) = \eta_0(\mathrm{d}y)$, we find that

$$\sigma^2_n(F_n) = \sum_{0\le p\le n} \mathbb{E}\left(\left[\sum_{0\le q\le p} f(X_q)\right]^2\right) = (n+1)(n+2)/2. \tag{2.6}$$

We further assume that the Markov transitions $M_n(x_{n-1}, \mathrm{d}x_n)$ are absolutely continuous with respect to some measures $\lambda_n(\mathrm{d}x_n)$ on E_n and we have

$$(H) \qquad \forall (x_{n-1}, x_n) \in (E_{n-1} \times E_n) \qquad H_n(x_{n-1}, x_n) = \frac{\mathrm{d}M_n(x_{n-1}, \centerdot)}{\mathrm{d}\lambda_n}(x_n) > 0.$$

In this situation, we have the backward decomposition formula

$$\mathbb{Q}_n(\mathrm{d}(x_0,\ldots,x_n)) = \eta_n(\mathrm{d}x_n)\ \mathcal{M}_n(x_n, \mathrm{d}(x_0,\ldots,x_{n-1})) \tag{2.7}$$

with the Markov transitions $\mathcal{M}_n$ defined below

$$\mathcal{M}_n(x_n, \mathrm{d}(x_0,\ldots,x_{n-1})) := \prod_{q=1}^n M_{q,\eta_{q-1}}(x_q, \mathrm{d}x_{q-1}). \tag{2.8}$$

In the above display, $M_{n+1,\eta}$ is the collection of Markov transitions defined for any $n \geq 0$ and $\eta \in \mathcal{P}(E_n)$ by

$$M_{n+1,\eta}(x, \mathrm{d}y) = \frac{1}{\eta\left(G_n H_{n+1}(.,x)\right)}\; G_n(y)\; H_{n+1}(y,x)\; \eta(\mathrm{d}y). \tag{2.9}$$

A detailed proof of this formula and its extended version is provided in Section 4. In a nonlinear filtering context, equations (2.7)–(2.8) correspond to the forward filtering backward smoothing decomposition [19].

Using the representation in (2.7), one natural way to approximate $\mathbb{Q}_n$ is to replace the measures η_n with their N-particle approximations η_n^N. The resulting particle approximation measure, $\mathbb{Q}_n^N$, is then

$$\mathbb{Q}_n^N(\mathrm{d}(x_0,\ldots,x_n)) := \eta_n^N(\mathrm{d}x_n)\; \mathcal{M}_n^N(x_n, \mathrm{d}(x_0,\ldots,x_{n-1})) \tag{2.10}$$

with the random transitions

$$\mathcal{M}_n^N(x_n, \mathrm{d}(x_0,\ldots,x_{n-1})) := \prod_{q=1}^{n} M_{q,\eta_{q-1}^N}(x_q, \mathrm{d}x_{q-1}). \tag{2.11}$$

At this point, it is convenient to recall that for any bounded measurable function f_n on E_n, the measures η_n can be written as follows

$$\eta_n(f_n) := \frac{\gamma_n(f_n)}{\gamma_n(1)} \quad \text{with} \quad \gamma_n(f_n) := \mathbb{E}\left(f_n(X_n) \prod_{0\leq p<n} G_p(X_p)\right) = \eta_n(f_n) \prod_{0\leq p<n} \eta_p(G_p). \tag{2.12}$$

The multiplicative formula in the r.h.s. of (2.12) is easily checked using the fact that $\gamma_{n+1}(1) = \gamma_n(G_n) = \eta_n(G_n)\;\gamma_n(1)$. Mimicking the above formulae, we set

$$\Gamma_n^N = \gamma_n^N(1) \times \mathbb{Q}_n^N \quad \text{with} \quad \gamma_n^N(1) := \prod_{0\leq p<n} \eta_p^N(G_p) \quad \text{and} \quad \gamma_n^N(\mathrm{d}x) = \gamma_n^N(1) \times \eta_n^N(\mathrm{d}x).$$

Notice that the N-particle approximation measures $\mathbb{Q}_n^N$ can be computed recursively with respect to the time parameter. For instance, for linear functionals of the form (2.5), we have

$$\mathbb{Q}_n^N(F_n) = \eta_n^N(F_n^N)$$

with a sequence of random functions F_n^N on E_n that can be computed "on-the-fly" according to the following recursion

$$F_n^N = \sum_{0\leq p\leq n} \left[M_{n,\eta_{n-1}^N} \ldots M_{p+1,\eta_p^N}\right](f_p) = f_n + M_{n,\eta_{n-1}^N}(F_{n-1}^N)$$

with the initial value $F_0^N = f_0$. In contrast to the genealogical tree based particle model (2.4), this new particle algorithm requires N^2 computations instead of N as

$$\forall 1 \leq j \leq N, \qquad F_n^N(\xi_n^j) = f_n(\xi_n^j) + \sum_{1\leq i\leq N} \frac{G_{n-1}(\xi_{n-1}^i) H_n(\xi_{n-1}^i, \xi_n^j)}{\sum_{1\leq i'\leq N} G_{n-1}(\xi_{n-1}^{i'}) H_n(\xi_{n-1}^{i'}, \xi_n^j)}\; F_{n-1}^N(\xi_{n-1}^i).$$

An important application of this recursion is to parameter estimation for non-linear non-Gaussian state-space models. For instance, it may be used to implement an on-line version of the Expectation-Maximization algorithm as detailed in [23], Section 3.2. In a different approach to recursive parameter estimation, an on-line particle algorithm is presented in [26] to compute the score for non-linear non-Gaussian state-space models. In fact, the algorithm of [26] is actually implementing a special case of the above recursion and may be reinterpreted as an "on-the-fly" computation of the forward filtering backward smoothing estimate of an additive functional derived from Fisher's identity.

The convergence analysis of the N-particle measures $\mathbb{Q}_n^N$ towards their limiting value $\mathbb{Q}_n$, as $N \to \infty$, is intimately related to the convergence of the flow of particle measures $(\eta_p^N)_{0\leq p\leq n}$ towards their limiting measures $(\eta_p)_{0\leq p\leq n}$. Several estimates can be easily derived more or less directly from the convergence analysis of the particle occupation measures η_n^N developed in [7], including $\mathbb{L}_p$-mean error bounds and exponential deviation estimates. It is clearly out of the scope of the present work to review all these consequences. One of the central objects in this analysis is the local sampling errors V_n^N induced by the mean field particle transitions and defined by the following stochastic perturbation formula

$$\eta_n^N = \eta_{n-1}^N K_{n,\eta_{n-1}^N} + \frac{1}{\sqrt{N}}\ V_n^N. \tag{2.13}$$

The fluctuations and deviations of these centered random measures V_n^N can be estimated using non asymptotic Kintchine's type $\mathbb{L}_r$-inequalities, as well as Hoeffding's or Bernstein's type exponential deviations [7,11]. We also proved in [9] that these random perturbations behave asymptotically as Gaussian random perturbations. More precisely, for any fixed time horizon $n \geq 0$, the sequence of random fields V_n^N converges in law, as the number of particles N tends to infinity, to a sequence of independent, Gaussian and centered random fields V_n such that, for any bounded function f on E_n, we have

$$\mathbb{E}(V_n(f)^2) = \int\ \eta_{n-1}(\mathrm{d}x)K_{n,\eta_{n-1}}(x,\mathrm{d}y)\left(f(y) - K_{n,\eta_{n-1}}(f)(x)\right)^2. \tag{2.14}$$

In Section 5, we provide some key decompositions expressing the deviation of the particle measures $(\Gamma_n^N, \mathbb{Q}_n^N)$ around their limiting values $(\Gamma_n, \mathbb{Q}_n)$ in terms of these local random fields models. These decompositions can be used to derive almost directly some exponential and $\mathbb{L}_p$-mean error bounds using the stochastic analysis developed in [7]. We shall use these functional central limit theorems and some of their variations in various places in the present article.

3. Statement of some results

In the present article, we have chosen to concentrate on functional central limit theorems, as well as on non asymptotic bias and variance theorems in terms of the time horizon. Unless otherwise is stated, we further assume that the potential functions G_n are lower bounded. To describe our results, it is necessary to introduce the following notation. Let $\beta(M)$ denote the Dobrushin coefficient of a Markov transition M from a measurable space E into another measurable space E' which defined by the following formula

$$\beta(M) := \sup\ \{\mathrm{osc}(M(f));\ \ f \in \mathrm{Osc}_1(E')\}$$

where $\mathrm{Osc}_1(E')$ stands the set of $\mathcal{E}'$-measurable functions f with oscillation $\mathrm{osc}(f)$ less than or equal to 1, with $\mathrm{osc}(f) = \sup\{[f(x) - f(y)];\ x,y \in E'\}$. Some stochastic models discussed in the present article are based on sequences of random Markov transitions M^N that depend on N random particles. In this case, $\beta(M^N)$ may fail to be measurable. For this type of models we shall use outer probability measures to integrate these quantities. For instance, the mean value $\mathbb{E}\left(\beta(M^N)\right)$ is to be understood as the infimum of the quantities $\mathbb{E}(B^N)$ where $B^N \geq \beta(M^N)$ are measurable dominating functions. We also recall that γ_n satisfies the linear recursive equation

$$\gamma_n = \gamma_p Q_{p,n} \quad \text{with} \quad Q_{p,n} = Q_{p+1}Q_{p+2}\ldots Q_n \quad \text{and} \quad Q_n(x,\mathrm{d}y) = G_{n-1}(x)\ M_n(x,\mathrm{d}y)$$

for any $0 \leq p \leq n$. Using elementary manipulations, we also check that

$$\Gamma_n(F_n) = \gamma_p D_{p,n}(F_n)$$

with the bounded integral operators $D_{p,n}$ from E_p into $E_{[0,n]}$ defined below

$$D_{p,n}(F_n)(x_p) := \int \mathcal{M}_p(x_p, \mathrm{d}(x_0,\dots,x_{p-1}))\mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1},\dots,x_n))\ F_n(x_0,\dots,x_n) \tag{3.1}$$

with

$$\mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1},\dots,x_n)) := \prod_{p\le q<n} Q_{q+1}(x_q, \mathrm{d}x_{q+1}).$$

We also let $(G_{p,n}, P_{p,n})$ be the pair of potential functions and Markov transitions defined below

$$G_{p,n} = Q_{p,n}(1)/\eta_p Q_{p,n}(1) \quad \text{and} \quad P_{p,n}(F_n) = D_{p,n}(F_n)/D_{p,n}(1). \tag{3.2}$$

Let the mapping $\Phi_{p,n} : \mathcal{P}(E_p) \to \mathcal{P}(E_n)$, $0 \le p \le n$, be defined as follows

$$\Phi_{p,n}(\mu_p) = \frac{\mu_p Q_{p,n}}{\mu_p Q_{p,n}(1)}.$$

Our first main result is a functional central limit theorem for the pair of random fields on $\mathcal{B}(E_{[0,n]})$ defined below

$$W_n^{\Gamma,N} := \sqrt{N}\left(\Gamma_n^N - \Gamma_n\right) \quad \text{and} \quad W_n^{\mathbb{Q},N} := \sqrt{N}\ [\mathbb{Q}_n^N - \mathbb{Q}_n] \tag{3.3}$$

$W_n^{\Gamma,N}$ is centered in the sense that $\mathbb{E}\left(W_n^{\Gamma,N}(F_n)\right) = 0$ for any $F_n \in \mathcal{B}(E_{[0,n]})$. The proof of this surprising unbiasedness property can be found in Corollary 5.3, in Section 5.

The first main result of this article is the following multivariate fluctuation theorem.

Theorem 3.1. *We suppose that the following regularity condition is met for any $n \ge 1$ and for any pair of states $(x,y) \in (E_{n-1}, E_n)$*

$$(H^+) \qquad h_n^-(y) \le H_n(x,y) \le h_n^+(y) \quad \textit{with} \quad (h_n^+/h_n^-) \in \mathbb{L}_4(\eta_n) \quad \textit{and} \quad h_n^+ \in \mathbb{L}_1(\lambda_n). \tag{3.4}$$

In this situation, the sequence of random fields $W_n^{\Gamma,N}$, resp. $W_n^{\mathbb{Q},N}$, converge in law, as $N \to \infty$, to the centered Gaussian fields W_n^Γ, resp. $W_n^{\mathbb{Q}}$, defined for any $F_n \in \mathcal{B}(E_{[0,n]})$ by

$$W_n^\Gamma(F_n) = \sum_{p=0}^n \gamma_p(1)\ V_p\left(D_{p,n}(F_n)\right) \quad \textit{and} \quad W_n^{\mathbb{Q}}(F_n) = \sum_{p=0}^n V_p\left(G_{p,n}\ P_{p,n}(F_n - \mathbb{Q}_n(F_n))\right).$$

An interpretation of the corresponding limiting variances in terms of conditional distributions of $\mathbb{Q}_n$ w.r.t. to the time marginal coordinates is provided in Section 8.1.

The second main result of the article is the following non asymptotic theorem.

Theorem 3.2. *For any $r \ge 1$, $n \ge 0$, $F_n \in \mathrm{Osc}_1(E_{[0,n]})$ we have the non asymptotic estimates*

$$\sqrt{N}\ \mathbb{E}\left(\left|[\mathbb{Q}_n^N - \mathbb{Q}_n](F_n)\right|^r\right)^{\frac{1}{r}} \le a_r \sum_{0\le p\le n} b_{p,n}^2\ c_{p,n}^N \tag{3.5}$$

for some finite constants $a_r < \infty$ whose values only depend on the parameter r, and a pair of constants $(b_{p,n}, c_{p,n}^N)$ such that

$$b_{p,n} \le \sup_{x,y}\left(Q_{p,n}(1)(x)/Q_{p,n}(1)(y)\right) \quad \textit{and} \quad c_{p,n}^N \le \mathbb{E}\left(\beta(P_{p,n}^N)\right).$$

In the above display, $P_{p,n}^N$ stands for the random Markov transitions defined as $P_{p,n}$ by replacing in (3.1) and (3.2) the transitions $\mathcal{M}_p$ by $\mathcal{M}_p^N$.

For linear functionals of the form (2.5), with $f_n \in \mathrm{Osc}_1(E_n)$, the $\mathbb{L}_r$-mean error estimate (3.5) is satisfied with a constant $c^N_{p,n}$ in (3.5) that can be chosen so that

$$c^N_{p,n} \leq \sum_{0\leq q<p} \mathbb{E}\left(\beta\left(M_{p,\eta^N_{p-1}} \ldots M_{q+1,\eta^N_q}\right)\right) + \sum_{p\leq q\leq n} b^2_{q,n}\ \beta(S_{p,q}) \tag{3.6}$$

with the Markov transitions $S_{p,q}$ from E_p into E_q defined for any function $f \in \mathcal{B}(E_q)$ by the following formula $S_{p,q}(f) = Q_{p,q}(f)/Q_{p,q}(1)$.

We emphasize that the $\mathbb{L}_r$-mean error bounds described in the above theorem enter the stability properties of the semigroups $S_{p,q}$ and the one associated with the backward Markov transitions M_{n+1,η^N_n}. In several instances, the term in the r.h.s. of (3.6) can be uniformly bounded with respect to the time horizon. For instance, in the toy example we discussed in (2.6), we have the variance formula (see Sect. 8 for more details)

$$\mathbb{E}\left(W^{\mathbb{Q}}_n(F_n)^2\right) = (n+1)$$

and the non asymptotic $\mathbb{L}_r$-estimates

$$b_{p,n} = 1 \quad \text{and} \quad c^N_{p,n} \leq 1 \implies \sqrt{N}\ \mathbb{E}\left(\left|[\mathbb{Q}^N_n - \mathbb{Q}_n](F_n)\right|^r\right)^{\frac{1}{r}} \leq a_r\ (n+1).$$

In more general situations, these estimates are related to the stability properties of the Feynman-Kac semigroup. To simplify the presentation, let us suppose that the state space E_n, and the pair of potential-transitions (G_n, M_n) are time homogeneous $(E_n, G_n, H_n, M_n) = (E, G, H, M)$, and chosen so that the following regularity condition is satisfied

$$(M)_m \qquad \forall (x,x') \quad G(x) \leq \delta\ G(x') \quad \text{and} \quad M^m(x, \mathrm{d}y) \leq \rho\ M^m(x', \mathrm{d}y) \tag{3.7}$$

for some $m \geq 1$ and some parameters $(\delta, \rho) \in [1,\infty)^2$. Under this rather strong condition, we have

$$b_{p,n} \leq \rho\delta^m \quad \text{and} \quad \beta(S_{p,q}) \leq \left(1 - \rho^{-2}\delta^{-m}\right)^{\lfloor (q-p)/m \rfloor}.$$

See for instance Corollary 4.3.3. in [7] and the more recent article [4]. On the other hand, let us suppose that

$$\inf_{x,y,y'} \left(H(x,y)/H(x,y')\right) = \alpha(h) > 0.$$

In this case, we have

$$M_{n,\eta}(x, \mathrm{d}y) \leq \alpha(h)^{-2}\ M_{n,\eta}(x', \mathrm{d}y) \Longrightarrow \beta\left(M_{p,\eta^N_{p-1}} \ldots M_{q+1,\eta^N_q}\right) \leq \left(1 - \alpha(h)^2\right)^{p-q}.$$

For linear functional models of the form (2.5) associated with functions $f_n \in \mathrm{Osc}_1(E_n)$, it is now readily checked that

$$\sqrt{N}\ \mathbb{E}\left(\left|[\mathbb{Q}^N_n - \mathbb{Q}_n](F_n)\right|^r\right)^{\frac{1}{r}} \leq a_r\ b\ (n+1) \tag{3.8}$$

for some finite constant $b < \infty$ whose values do not depend on the time parameter n. The above non asymptotic estimate is not sharp for $r = 2$. To obtain better bounds, we need to refine the analysis of the variance using first order decompositions to analyze separately the bias of the particle model. In a companion paper [13], we prove that

$$\sup_{n\geq 0} \left|\mathbb{E}\left(\mathbb{Q}^N_n(\overline{F}_n)\right) - \mathbb{Q}_n(\overline{F}_n)\right| \leq \frac{c}{N} \quad \text{and} \quad \mathbb{E}\left(W^{\mathbb{Q},N}_n(F_n)^2\right) \leq c\ (n+1)\left(1 + \frac{n+1}{N}\right)$$

where c is once again a finite time-independent constant.

With some information on the constants a_r, the above $\mathbb{L}_r$-mean error bounds can turned to uniform exponential estimates w.r.t. the time parameter for normalized additive functionals of the following form

$$\overline{F}_n(x_0,\ldots,x_n) := \frac{1}{n+1}\sum_{0\le p\le n} f_p(x_p).$$

To be more precise, by Lemma 7.3.3 in [7], the collection of constants a_r in (3.8) can be chosen so that

$$a_{2r}^{2r} \le (2r)!\; 2^{-r}/r! \quad \text{and} \quad a_{2r+1}^{2r+1} \le (2r+1)!\; 2^{-r}/r!. \tag{3.9}$$

In this situation, it is easily checked that for any $\epsilon > 0$, and $N \ge 1$, we have the following uniform Gaussian concentration estimates:

$$\frac{1}{N}\log\sup_{n\ge 0}\mathbb{P}\left(\left|[\mathbb{Q}_n^N - \mathbb{Q}_n](\overline{F}_n)\right| \ge \frac{b}{\sqrt{N}} + \epsilon\right) \le -\epsilon^2/(2b^2).$$

This result is a direct consequence of the fact that for any non negative random variable U

$$\left(\forall r\ge 1, \quad \mathbb{E}\left(U^r\right)^{\frac{1}{r}} \le a_r\; b\right) \Rightarrow \log\mathbb{P}\left(U \ge b+\epsilon\right) \le -\epsilon^2/(2b^2).$$

To check this claim, we develop the exponential to prove that

$$\log\mathbb{E}\left(\mathrm{e}^{tU}\right) \overset{\forall t\ge 0}{\le} bt + \frac{(bt)^2}{2} \Rightarrow \log\mathbb{P}\left(U\ge b+\epsilon\right) \le -\sup_{t\ge 0}\left(\epsilon t - \frac{(bt)^2}{2}\right)\cdot$$

We end this section with a brief discussion on the regularity condition $(M)_m$ introduced in (3.7).

Firstly, we recall the mixing condition stated in the r.h.s. of (3.7) is satisfied for any aperiodic and irreducible Markov chains on finite state spaces, as well as for bi-Laplace exponential transitions on the real line associated with a bounded drift function, and for Gaussian transitions with a mean drift function that is constant outside some compact domain. We also mention that this condition is met with $m = 1$ for the Δ-time step Markov transition M of a continuous time regular diffusion models on some compact manifold. In this context, the r.h.s. of (3.7) is satisfied with $m = 1$ and $\rho = a\; \exp(b/\Delta)$, for some $a, b < \infty$ (see for instance [1]). Of course, when applied to a particular situation the uniform estimates presented above are not sharp. They only reflect the worst case error in a rather large class of models.

Moreover these two regularity conditions are sometimes not satisfied in some important applications. In particular, this is the case for Coulomb interactions in molecular Hamiltonian models as well as indicator type potential functions in hard obstacle particle absorption models. In these singular settings a more refined analysis is required. Several strategies can be underlined:

The first classical idea is to use cutoff techniques, such as replacing the original process X_n by its restriction to a given bounded subset of the state space, or by replacing the original potential G_n by the cutoff function $(G_n \vee \epsilon)$, for some parameter $\epsilon > 0$. These cutoff techniques can be thought as another approximation level. For singular potential models, using stochastic perturbation type arguments the expected performance of our particle technology applied to cutoff approximation models depends on the stability properties of the corresponding limiting flow of Feynman-Kac measures.

Another strategy, which is more probabilistic in nature, is to change the reference probability measure $\mathbb{P}_n$ so as to work with a new Markov chain model X_n whose elementary transitions depend on the potential function G_n. For instance, for indicator type potential functions $G_n := 1_A$ associated with some measurable subset $A \subset E_n$, we can choose a reference Markov chain $\widehat{X}_n$ with elementary transitions restricted to the set A, and a new potential function $\widehat{G}_n$ that reflects the probability to stay in the desired set A starting from some given state. In the context of particle absorption models, this rather well known technique is sometimes used to turn a hard obstacle model into a soft obstacle one (see for instance [8]). From the pure mathematical point of view, these change of probability models have the same form as the one given in (1.1), but their mean field particle model differs. For more details on this change of measure technique and the corresponding particle approximations, we refer the reader to Section 3 in the book [7], as well as [8].

When the potential functions G_n vanish in some regions of the state space, we also mention that the particle model is only defined up to the first time $\tau^N = k$ we have $\eta_k^N(G_k) = 0$. When $\gamma_n(1) > 0$, we can prove that the event $\{\tau^N \leq n\}$ has an exponentially small probability to occur, and the estimates presented in the above theorems can be extended to these singular situations by replacing $\mathbb{Q}_n^N(F_n)$ by the particle estimates $\mathbb{Q}_n^N(F_n)\ 1_{\tau^N \geq n}$. The stochastic analysis of these singular models are quite technical, for further details we refer the reader to Sections 7.2.2 and 7.4 in the book [7].

4. A backward Markov chain formulation

This section is mainly concerned with the proof of the backward decomposition formula (2.7). Before proceeding, we recall that the measures (γ_n, η_n) satisfy the nonlinear equations

$$\gamma_n = \gamma_{n-1} Q_n \quad \text{and} \quad \eta_{n+1} := \Phi_{n+1}(\eta_n) := \Psi_{G_n}(\eta_n) M_{n+1}$$

and their semigroups are given by

$$\gamma_n = \gamma_p Q_{p,n} \quad \text{and} \quad \eta_n(f_n) := \eta_p Q_{p,n}(f_n)/\eta_p Q_{p,n}(1)$$

for any function $f_n \in \mathcal{B}(E_n)$. In this connection, we also mention that the semigroup of the pair of measures $(\Gamma_n, \mathbb{Q}_n)$ defined in (1.1) for any $0 \leq p \leq n$ and any $F_n \in \mathcal{B}(E_{[0,n]})$, we have

$$\Gamma_n(F_n) = \gamma_p D_{p,n}(F_n) \quad \text{and} \quad \mathbb{Q}_n(F_n) = \eta_p D_{p,n}(F_n)/\eta_p D_{p,n}(1). \tag{4.1}$$

These formulae are a direct consequence of the following observation

$$\eta_p D_{p,n}(F_n) = \int \mathbb{Q}_p(\mathrm{d}(x_0, \ldots, x_p))\ \mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_n)) F_n(x_0, \ldots, x_n).$$

Lemma 4.1. *For any $0 \leq p < n$, we have*

$$\gamma_p(\mathrm{d}x_p)\ \mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_n)) = \gamma_n(\mathrm{d}x_n)\ \mathcal{M}_{n,p}(x_n, \mathrm{d}(x_p, \ldots, x_{n-1})) \tag{4.2}$$

with

$$\mathcal{M}_{n,p}(x_n, \mathrm{d}(x_p, \ldots, x_{n-1})) := \prod_{p \leq q < n} M_{q+1,\eta_q}(x_{q+1}, \mathrm{d}x_q).$$

In particular, for any time $n \geq 0$, the Feynman-Kac path measures $\mathbb{Q}_n$ defined in (1.1) can be expressed in terms of the sequence of marginal measures $(\eta_p)_{0 \leq p \leq n}$, with the following backward Markov chain formulation

$$\mathbb{Q}_n(\mathrm{d}(x_0, \ldots, x_n)) = \eta_n(\mathrm{d}x_n)\ \mathcal{M}_{n,0}(x_n, \mathrm{d}(x_0, \ldots, x_{n-1})). \tag{4.3}$$

Before entering into the details of the proof of this lemma, we mention that (4.3) holds true for any well defined Markov transition $M_{n+1,\eta_n}(y, \mathrm{d}x)$ from E_n into E_{n+1} satisfying the local backward equation

$$\Psi_{G_n}(\eta_n)(\mathrm{d}x)\ M_{n+1}(x, \mathrm{d}y) = \Phi_{n+1}(\eta_n)(\mathrm{d}y)\ M_{n+1,\eta_n}(y, \mathrm{d}x)$$

or equivalently

$$\eta_n(\mathrm{d}x)\ Q_{n+1}(x, \mathrm{d}y) = (\eta_n Q_{n+1})(\mathrm{d}y)\ M_{n+1,\eta_n}(y, \mathrm{d}x). \tag{4.4}$$

In other words, we have the duality formula

$$\Psi_{G_n}(\eta_n)\left(f\ M_{n+1}(g)\right) = \Phi_{n+1}(\eta_n)\left(g\ M_{n+1,\eta_n}(f)\right). \tag{4.5}$$

Also notice that for any pair of measures μ, ν on E_n s.t. $\mu \ll \nu$, we have $\mu M_{n+1} \ll \nu M_{n+1}$. Indeed, if we have $\nu M_{n+1}(A) = 0$, the function $M_{n+1}(1_A)$ is null ν-almost everywhere, and therefore μ-almost everywhere from which we conclude that $\mu M_{n+1}(A) = 0$. For any bounded measurable function g on E_n we set

$$\Psi^g_{G_n}(\eta_n)(\mathrm{d}x) = \Psi_{G_n}(\eta_n)(\mathrm{d}x)\ g(x) \ll \Psi_{G_n}(\eta_n)(\mathrm{d}x).$$

From the previous discussion, we have $\Psi^g_{G_n}(\eta_n)M_{n+1} \ll \Psi_{G_n}(\eta_n)M_{n+1}$ and it is easily checked that

$$M_{n+1,\eta_n}(g)(y) = \frac{\mathrm{d}\Psi^g_{G_n}(\eta_n)M_{n+1}}{\mathrm{d}\Psi_{G_n}(\eta_n)M_{n+1}}(y)$$

is a well defined Markov transition from E_{n+1} into E_n satisfying the desired backward equation. These manipulations are rather classical in the literature on Markov chains (see for instance [27], and references therein). Under the regularity condition (H) the above transition is explicitly given by the formula (2.9).

Now, we come to the proof of Lemma 4.1.

Proof of Lemma 4.1*.* We prove (4.2) using a backward induction on the parameter p. By (4.4), the formula is clearly true for $p = (n-1)$. Suppose the result has been proved at rank p. Since we have

$$\gamma_{p-1}(\mathrm{d}x_{p-1})\ \mathcal{Q}_{p-1,n}(x_{p-1}, \mathrm{d}(x_p, \ldots, x_n)) = \gamma_{p-1}(\mathrm{d}x_{p-1})\ Q_p(x_{p-1}, \mathrm{d}x_p)\ \mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_n))$$

and

$$\gamma_{p-1}(\mathrm{d}x_{p-1})\ Q_p(x_{p-1}, \mathrm{d}x_p) = \gamma_p(\mathrm{d}x_p)\ M_{p,\eta_{p-1}}(x_p, \mathrm{d}x_{p-1}).$$

Using the backward induction we conclude that the desired formula is also met at rank $(p-1)$. The second assertion is a direct consequence of (4.2). The end of the proof of the lemma is now completed. □

We end this section with some properties of backward Markov transitions associated with a given initial probability measure that may differ from the one associated with the Feynman-Kac measures. These mathematical objects appear in a natural way in the analysis of the N-particle approximation transitions $\mathcal{M}^N_n$ introduced in (2.11).

Definition 4.2. For any $0 \le p \le n$ and any probability measure $\eta \in \mathcal{P}(E_p)$, we denote by $\mathcal{M}_{n+1,p,\eta}$ the Markov transition from E_{n+1} into $E_{[p,n]} = (E_p \times \ldots \times E_n)$ defined by

$$\mathcal{M}_{n+1,p,\eta}\left(x_{n+1}, \mathrm{d}(x_p, \ldots, x_n)\right) = \prod_{p \le q \le n} M_{q+1,\Phi_{p,q}(\eta)}(x_{q+1}, \mathrm{d}x_q).$$

Notice that this definition is consistent with the definition of the Markov transitions $\mathcal{M}_{p,n}$ introduced in Lemma 4.1:

$$\mathcal{M}_{n+1,p,\eta_p}\left(x_{n+1}, \mathrm{d}(x_p, \ldots, x_n)\right) = \mathcal{M}_{n+1,p}\left(x_{n+1}, \mathrm{d}(x_p, \ldots, x_n)\right).$$

Also observe that $\mathcal{M}_{n+1,p,\eta}$ can alternatively be defined by the pair of recursions

$$\begin{aligned}\mathcal{M}_{n+1,p,\eta}\left(x_{n+1}, \mathrm{d}(x_p, \ldots, x_n)\right) &= \mathcal{M}_{n+1,p+1,\Phi_{p+1}(\eta)}\left(x_{n+1}, \mathrm{d}(x_{p+1}, \ldots, x_n)\right) \times M_{p+1,\eta}(x_{p+1}, \mathrm{d}x_p)\\ &= M_{n+1,\Phi_{p,n}(\eta)}(x_{n+1}, \mathrm{d}x_n)\ \mathcal{M}_{n,p,\eta}\left(x_n, \mathrm{d}(x_p, \ldots, x_{n-1})\right).\end{aligned} \tag{4.6}$$

The proof of the following lemma follows the same lines of arguments as the ones used in proof of Lemma 4.1. For the convenience of the reader, the details of this proof are postponed to the appendix.

Lemma 4.3. *For any $0 \le p < n$ and any probability measure $\eta \in \mathcal{P}(E_p)$, we have*

$$\eta Q_{p,n}(\mathrm{d}x_n)\ \mathcal{M}_{n,p,\eta}(x_n, \mathrm{d}(x_p, \ldots, x_{n-1})) = \eta(\mathrm{d}x_p)\ \mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_n)).$$

In other words, we have

$$\mathcal{M}_{n,p,\eta}(x_n, \mathrm{d}(x_p, \ldots, x_{n-1})) = \frac{(\eta \times \mathcal{Q}_{p,n-1})(\mathrm{d}(x_p, \ldots, x_{n-1}))G_{n-1}(x_{n-1})\ H_n(x_{n-1}, x_n)}{(\eta Q_{p,n-1})\,(G_{n-1}\ H_n(\,.\,, x_n))} \tag{4.7}$$

with the measure $(\eta \times \mathcal{Q}_{p,n-1})$ defined below

$$(\eta \times \mathcal{Q}_{p,n-1})(\mathrm{d}(x_p, \ldots, x_{n-1})) := \eta(\mathrm{d}x_p)\ \mathcal{Q}_{p,n-1}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_{n-1})).$$

5. Particle approximation models

We provide in this section some preliminary results on the convergence of the N-particle measures $(\Gamma_n^N, \mathbb{Q}_n^N)$ to their limiting values $(\Gamma_n, \mathbb{Q}_n)$, as $N \to \infty$. Most of the forthcoming analysis is developed in terms of the following integral operators.

Definition 5.1. For any $0 \le p \le n$, we let $D_{p,n}^N$ be the $\mathcal{F}_{p-1}^N$-measurable integral operators from $\mathcal{B}(E_{[0,n]})$ into $\mathcal{B}(E_p)$ defined below

$$D_{p,n}^N(F_n)(x_p) := \int \mathcal{M}_p^N(x_p, \mathrm{d}(x_0, \ldots, x_{p-1}))\mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_n))F_n(x_0, \ldots, x_n)$$

with the conventions $D_{0,n}^N = \mathcal{Q}_{0,n}$, and resp. $D_{n,n}^N = \mathcal{M}_n^N$, for $p = 0$, and resp. $p = n$

The main result of this section is the following theorem.

Theorem 5.2. *For any $0 \le p \le n$, and any function F_n on the path space $E_{[0,n]}$, we have*

$$\mathbb{E}\left(\Gamma_n^N(F_n) \,\middle|\, \mathcal{F}_p^N\right) = \gamma_p^N\left(D_{p,n}^N(F_n)\right) \quad \textit{and} \quad W_n^{\Gamma,N}(F_n) = \sum_{p=0}^{n} \gamma_p^N(1)\ V_p^N\left(D_{p,n}^N(F_n)\right)$$

Proof of Theorem 5.2. To prove the first assertion, we use a backward induction on the parameter p. For $p = n$, the result is immediate since we have

$$\Gamma_n^N(F_n) = \gamma_n^N(1)\ \eta_n^N\left(D_{n,n}^N(F_n)\right).$$

We suppose that the formula is valid at a given rank $p \le n$. In this situation, we have

$$\begin{aligned}
\mathbb{E}\left(\Gamma_n^N(F_n) \,\middle|\, \mathcal{F}_{p-1}^N\right) &= \gamma_p^N(1)\ \mathbb{E}\left(\eta_p^N\left(D_{p,n}^N(F_n)\right) \,\middle|\, \mathcal{F}_{p-1}^N\right) \\
&= \gamma_{p-1}^N(1) \int \eta_{p-1}^N(G_{p-1}H_p(\,.\,, x_p))\ \lambda_p(\mathrm{d}x_p)\ D_{p,n}^N(F_n)(x_p).
\end{aligned} \tag{5.1}$$

Using the fact that

$$\gamma_{p-1}^N(1)\ \eta_{p-1}^N(G_{p-1}H_p(\,.\,, x_p))\ \lambda_p(\mathrm{d}x_p)\ M_{p,\eta_{p-1}^N}(x_p, \mathrm{d}x_{p-1}) = \gamma_{p-1}^N(\mathrm{d}x_{p-1})Q_p(x_{p-1}, \mathrm{d}x_p)$$

we conclude that the r.h.s. term in (5.1) takes the form

$$\int \gamma_{p-1}^N(\mathrm{d}x_{p-1})\mathcal{M}_{p-1}^N(x_{p-1}, \mathrm{d}(x_0, \ldots, x_{p-2}))\mathcal{Q}_{p-1,n}(x_{p-1}, \mathrm{d}(x_p, \ldots, x_n))F_n(x_0, \ldots, x_n) = \gamma_{p-1}^N\left(D_{p-1,n}^N(F_n)\right).$$

This ends the proof of the first assertion. The proof of the second assertion is based on the following decomposition

$$\begin{aligned}\left(\Gamma_n^N-\Gamma_n\right)(F_n) &= \sum_{p=0}^{n}\left[\mathbb{E}\left(\Gamma_n^N(F_n)\,\middle|\,\mathcal{F}_p^N\right)-\mathbb{E}\left(\Gamma_n^N(F_n)\,\middle|\,\mathcal{F}_{p-1}^N\right)\right]\\ &= \sum_{p=0}^{n}\gamma_p^N(1)\ \left(\eta_p^N\left(D_{p,n}^N(F_n)\right)-\frac{1}{\eta_{p-1}^N(G_{p-1})}\ \eta_{p-1}^N\left(D_{p-1,n}^N(F_n)\right)\right)\end{aligned}$$

where $\mathcal{F}_{-1}^N$ is the trivial sigma field. By definition of the random fields V_p^N, it remains to prove that

$$\eta_{p-1}^N\left(D_{p-1,n}^N(F_n)\right)=\left(\eta_{p-1}^N Q_p\right)\left(D_{p,n}^N(F_n)\right).$$

To check this formula, we use the decomposition

$$\begin{aligned}\eta_{p-1}^N(\mathrm{d}x_{p-1})\ \mathcal{M}_{p-1}^N(x_{p-1},\mathrm{d}(x_0,\ldots,x_{p-2}))\ \mathcal{Q}_{p-1,n}(x_{p-1},\mathrm{d}(x_p,\ldots,x_n)) &= \eta_{p-1}^N(\mathrm{d}x_{p-1})Q_p(x_{p-1},\mathrm{d}x_p)\\ \times\mathcal{M}_{p-1}^N(x_{p-1},\mathrm{d}(x_0,\ldots,x_{p-2}))\ \mathcal{Q}_{p,n}(x_p,\mathrm{d}(x_{p+1},\ldots,x_n)).\end{aligned}\tag{5.2}$$

Using the fact that

$$\eta_{p-1}^N(\mathrm{d}x_{p-1})Q_p(x_{p-1},\mathrm{d}x_p)\quad=\quad(\eta_{p-1}^N Q_p)(\mathrm{d}x_p)\ M_{p,\eta_{p-1}^N}(x_p,\mathrm{d}x_{p-1})$$

we conclude that the term in the r.h.s. of (5.2) is equal to

$$(\eta_{p-1}^N Q_p)(\mathrm{d}x_p)\ \mathcal{M}_p^N(x_p,\mathrm{d}(x_0,\ldots,x_{p-1}))\ \mathcal{Q}_{p,n}(x_p,\mathrm{d}(x_{p+1},\ldots,x_n)).$$

This ends the proof of the theorem. □

Several consequences of Theorem 5.2 are now emphasized. On the one hand, using the fact that the random fields V_n^N are centered given $\mathcal{F}_{n-1}^N$, we find that

$$\mathbb{E}\left(\Gamma_n^N(F_n)\right)=\Gamma_n(F_n).$$

On the other hand, using the fact that

$$\frac{\gamma_p(1)}{\gamma_n(1)}=\frac{\gamma_p(1)}{\gamma_p Q_{p,n}(1)}=\frac{1}{\eta_p Q_{p,n}(1)}$$

we prove the following decomposition

$$\overline{W}_n^{\Gamma,N}(F_n)=\sqrt{N}\left(\overline{\gamma}_n^N(1)\ \mathbb{Q}_n^N-\mathbb{Q}_n\right)(F_n)=\sum_{p=0}^{n}\overline{\gamma}_p^N(1)\ V_p^N\left(\overline{D}_{p,n}^N(F_n)\right)\tag{5.3}$$

with the pair of parameters $\left(\overline{\gamma}_n^N(1),\overline{D}_{p,n}^N\right)$ defined below

$$\overline{\gamma}_n^N(1):=\frac{\gamma_n^N(1)}{\gamma_n(1)}\quad\text{and}\quad\overline{D}_{p,n}^N(F_n)=\frac{D_{p,n}^N(F_n)}{\eta_p Q_{p,n}(1)}\cdot\tag{5.4}$$

Using again the fact that the random fields V_n^N are centered given $\mathcal{F}_{n-1}^N$, we have

$$\mathbb{E}\left(\overline{W}_n^{\Gamma,N}(F_n)^2\right) = \sum_{p=0}^{n} \mathbb{E}\left(\overline{\gamma}_p^N(1)^2 \; \mathbb{E}\left[V_p^N\left(\overline{D}_{p,n}^N(F_n)\right)^2 \;\mid \mathcal{F}_{p-1}^N\right]\right).$$

Using the estimates

$$\begin{aligned} \|D_{p,n}^N(F_n)\| &\leq \|Q_{p,n}(1)\| \; \|F_n\| \\ \|\overline{D}_{p,n}^N(F_n)\| &\leq \|\overline{Q}_{p,n}(1)\| \; \|F_n\| \quad \text{with} \quad \overline{Q}_{p,n}(1) = \frac{Q_{p,n}(1)}{\eta_p Q_{p,n}(1)} \end{aligned} \tag{5.5}$$

we prove the non asymptotic variance estimate

$$\mathbb{E}\left(\overline{W}_n^{\Gamma,N}(F_n)^2\right) \leq \sum_{p=0}^{n} \mathbb{E}\left(\overline{\gamma}_p^N(1)^2\right) \; \|\overline{Q}_{p,n}(1)\|^2 = \sum_{p=0}^{n}\left[1 + \mathbb{E}\left([\overline{\gamma}_p^N(1) - 1]^2\right)\right] \; \|\overline{Q}_{p,n}(1)\|^2$$

for any function F_n such that $\|F_n\| \leq 1$. On the other hand, using the decomposition

$$\left(\overline{\gamma}_n^N(1) \; \mathbb{Q}_n^N - \mathbb{Q}_n\right) = \left[\overline{\gamma}_n^N(1) - 1\right] \; \mathbb{Q}_n^N + \left(\mathbb{Q}_n^N - \mathbb{Q}_n\right)$$

we prove that

$$\mathbb{E}\left(\left[\mathbb{Q}_n^N(F_n) - \mathbb{Q}_n(F_n)\right]^2\right)^{1/2} \leq \frac{1}{\sqrt{N}} \; \mathbb{E}\left(W_n^{\Gamma}(F_n)^2\right)^{1/2} + \mathbb{E}\left(\left[\overline{\gamma}_n^N(1) - 1\right]^2\right)^{1/2}.$$

Some interesting bias estimates can also be obtained using the fact that

$$\mathbb{E}\left(\mathbb{Q}_n^N(F_n)\right) - \mathbb{Q}_n(F_n) = \mathbb{E}\left(\left[1 - \overline{\gamma}_n^N(1)\right] \; \left[\mathbb{Q}_n^N(F_n) - \mathbb{Q}_n(F_n)\right]\right)$$

and the following easily proved upper bound

$$\left|\mathbb{E}\left(\mathbb{Q}_n^N(F_n)\right) - \mathbb{Q}_n(F_n)\right| \leq \mathbb{E}\left(\left[1 - \overline{\gamma}_n^N(1)\right]^2\right)^{1/2} \mathbb{E}\left(\left[\mathbb{Q}_n^N(F_n) - \mathbb{Q}_n(F_n)\right]^2\right)^{1/2}.$$

Under the regularity condition $(M)_m$ stated in (3.7), we proved in a recent article [4], that for any $n \geq p \geq 0$, and any $N > (n+1)\rho\delta^m$ we have

$$\|\overline{Q}_{p,n}(1)\| \leq \delta^m \rho \quad \text{and} \quad N \; \mathbb{E}\left[\left(\overline{\gamma}_n^N(1) - 1\right)^2\right] \leq 4 \; (n+1) \; \rho \; \delta^m.$$

From these estimates, we readily prove the following corollary.

Corollary 5.3. *Assume that condition $(M)_m$ is satisfied for some parameters (m, δ, ρ). In this situation, for any $n \geq p \geq 0$, any F_n such that $\|F_n\| \leq 1$, and any $N > (n+1)\rho\delta^m$ we have*

$$\mathbb{E}\left(\overline{W}_n^{\Gamma,N}(F_n)\right) = 0 \quad \textit{and} \quad \mathbb{E}\left(\overline{W}_n^{\Gamma,N}(F_n)^2\right) \leq (\delta^m \rho)^2 (n+1) \left(1 + \frac{2}{N} \; \rho\delta^m (n+2)\right).$$

In addition, we have

$$N \; \mathbb{E}\left(\left[\mathbb{Q}_n^N(F_n) - \mathbb{Q}_n(F_n)\right]^2\right) \quad \leq \quad 2(n+1)\rho\delta^m \left(4 + \rho\delta^m \left[1 + \frac{2}{N}(n+2)\right]\right)$$

and the bias estimate

$$N\left|\mathbb{E}\left(\mathbb{Q}_n^N(F_n)\right)-\mathbb{Q}_n(F_n)\right|\leq 2\sqrt{2}\ (n+1)\rho\delta^m\left(4+\rho\delta^m\left[1+\frac{2}{N}(n+2)\right]\right)^{1/2}.$$

6. Fluctuation properties

This section is mainly concerned with the proof of Theorem 3.1. Unless otherwise is stated, in the further developments of this section, we assume that the regularity condition (H^+) presented in (3.4) is satisfied for some collection of functions (h_n^-, h_n^+). Our first step to establish Theorem 3.1 is the fluctuation analysis of the N-particle measures $(\Gamma_n^N, \mathbb{Q}_n^N)$ given in Proposition 6.2 whose proof relies on the following technical lemma.

Lemma 6.1.

$$\begin{aligned}\mathcal{M}_n^N(x_n,\mathrm{d}(x_0,\ldots,x_{n-1}))-\mathcal{M}_n(x_n,\mathrm{d}(x_0,\ldots,x_{n-1}))=\sum_{0\leq p\leq n}\left[\mathcal{M}_{n,p,\eta_p^N}-\mathcal{M}_{n,p,\Phi_p\left(\eta_{p-1}^N\right)}\right]\\ \times(x_n,\mathrm{d}(x_p,\ldots,x_{n-1}))\ \ \mathcal{M}_p^N(x_p,\mathrm{d}(x_0,\ldots,x_{p-1})).\end{aligned}$$

The proof of this lemma follows elementary but rather tedious calculations; thus it is postponed to the appendix. We now state Proposition 6.2.

Proposition 6.2. *For any $N\geq 1$, $0\leq p\leq n$, $x_p\in E_p$, $m\geq 1$, and $F_n\in\mathcal{B}(E_{[0,n]})$ such that $\|F_n\|\leq 1$, we have*

$$\sqrt{N}\ \mathbb{E}\left(\left|D_{p,n}^N(F_n)-D_{p,n}(F_n)(x_p)\right|^m\right)^{\frac{1}{m}}\leq a(m)\ b(n)\ \left(\frac{h_p^+}{h_p^-}(x_p)\right)^2 \tag{6.1}$$

for some finite constants $a(m)<\infty$, resp. $b(n)<\infty$, whose values only depend on the parameters m, resp. on the time horizon n.

Proof. Using Lemma 6.1, we find that

$$D_{p,n}^N(F_n)-D_{p,n}(F_n)=\sum_{0\leq q\leq p}\left[\mathcal{M}_{p,q,\eta_q^N}-\mathcal{M}_{p,q,\Phi_q\left(\eta_{q-1}^N\right)}\right]\left(T_{p,q,n}^N(F_n)\right)$$

with the random function $T_{p,q,n}^N(F_n)$ defined below

$$T_{p,q,n}^N(F_n)(x_q,\ldots,x_p):=\int\mathcal{Q}_{p,n}(x_p,\mathrm{d}(x_{p+1},\ldots,x_n))\ \mathcal{M}_q^N(x_q,\mathrm{d}(x_0,\ldots,x_{q-1}))\ F_n(x_0,\ldots,x_n).$$

Using formula (4.7), we prove that for any $m\geq 1$ and any function F on $E_{[q,p]}$

$$\sqrt{N}\ \mathbb{E}\left(\left|\left[\mathcal{M}_{p,q,\eta_q^N}-\mathcal{M}_{p,q,\Phi_q\left(\eta_{q-1}^N\right)}\right](F)\,(x_p)\right|^m\ \mid\mathcal{F}_{q-1}^N\right)^{\frac{1}{m}}\leq a(m)\ b(n)\ \|F\|\ \left(\frac{h_p^+}{h_p^-}(x_p)\right)^2$$

for some finite constants $a(m)<\infty$ and $b(n)<\infty$ whose values only depend on the parameters m and n. Using these almost sure estimates, we easily prove (6.1). This ends the proof of the proposition. □

Now, we come to the proof of Theorem 3.1.

Proof of Theorem 3.1. Using Theorem 5.2, we have the decomposition

$$W_n^{\Gamma,N}(F_n)=\sum_{p=0}^n\gamma_p^N(1)\ V_p^N\left(D_{p,n}(F_n)\right)+R_n^{\Gamma,N}(F_n)$$

with the second order remainder term

$$R_n^{\Gamma,N}(F_n) := \sum_{p=0}^{n} \gamma_p^N(1)\ V_p^N\left(F_{p,n}^N\right) \quad \text{and the function} \quad F_{p,n}^N := [D_{p,n}^N - D_{p,n}](F_n).$$

By Slutsky's lemma and by the continuous mapping theorem it clearly suffices to check that $R_n^{\Gamma,N}(F_n)$ converge to 0, in probability, as $N \to \infty$. To prove this claim, we notice that

$$\mathbb{E}\left(V_p^N\left(F_{p,n}^N\right)^2 \ \middle|\ \mathcal{F}_{p-1}^N\right) \leq \Phi_p\left(\eta_{p-1}^N\right)\left(\left(F_{p,n}^N\right)^2\right).$$

On the other hand, we have

$$\begin{aligned}\Phi_p\left(\eta_{p-1}^N\right)\left(\left(F_{p,n}^N\right)^2\right) &= \int \lambda_p(\mathrm{d}x_p)\ \Psi_{G_{p-1}}\left(\eta_{p-1}^N\right)(H_p(.,x_p))\ \ F_{p,n}^N(x_p)^2 \\ &\leq \eta_p\left(\left(F_{p,n}^N\right)^2\right) + \int \lambda_p(\mathrm{d}x_p)\ \left|\left[\Psi_{G_{p-1}}\left(\eta_{p-1}^N\right) - \Psi_{G_{p-1}}(\eta_{p-1})\right](H_p(.,x_p))\right|\ F_{p,n}^N(x_p)^2.\end{aligned}$$

This yields the rather crude estimate

$$\begin{aligned}\Phi_p\left(\eta_{p-1}^N\right)\left(\left(F_{p,n}^N\right)^2\right) &= \int \lambda_p(\mathrm{d}x_p)\ \Psi_{G_{p-1}}\left(\eta_{p-1}^N\right)(H_p(.,x_p))\ \ F_{p,n}^N(x_p)^2 \\ &\leq \eta_p\left(\left(F_{p,n}^N\right)^2\right) + 4\|Q_{p,n}(1)\|^2 \int \lambda_p(\mathrm{d}x_p)\ \left|\left[\Psi_{G_{p-1}}\left(\eta_{p-1}^N\right) - \Psi_{G_{p-1}}(\eta_{p-1})\right](H_p(.,x_p))\right|\end{aligned}$$

from which we conclude that

$$\begin{aligned}\mathbb{E}\left(V_p^N\left(F_{p,n}^N\right)^2\right) \leq &\int \eta_p(\mathrm{d}x_p)\ \mathbb{E}\left[\left(F_{p,n}^N(x_p)\right)^2\right] + 4\|Q_{p,n}(1)\|^2 \\ &\times \int \lambda_p(\mathrm{d}x_p)\ \mathbb{E}\left(\left|\left[\Psi_{G_{p-1}}\left(\eta_{p-1}^N\right) - \Psi_{G_{p-1}}(\eta_{p-1})\right](H_p(.,x_p))\right|\right).\end{aligned}$$

We can establish that

$$\sqrt{N}\ \mathbb{E}\left(\left|\left[\Psi_{G_{p-1}}\left(\eta_{p-1}^N\right) - \Psi_{G_{p-1}}(\eta_{p-1})\right](H_p(.,x_p))\right|\right) \leq b(n)\ h_p^+(x_p).$$

See for instance Section 7.4.3, Theorem 7.4.4, in [7]. Using Proposition 6.2,

$$\sqrt{N}\ \mathbb{E}\left(V_p^N\left(F_{p,n}^N\right)^2\right) \leq c(n)\left(\frac{1}{\sqrt{N}}\ \eta_p\left(\left(\frac{h_p^+}{h_p^-}\right)^4\right) + \lambda_p(h_p^+)\right)$$

for some finite constant $c(n) < \infty$. The end of the proof of the first assertion now follows standard computations. To prove the second assertion, we use the following decomposition

$$\sqrt{N}\ [\mathbb{Q}_n^N - \mathbb{Q}_n](F_n) = \frac{1}{\overline{\gamma}_n^N(1)}\ \overline{W}_n^{\Gamma,N}(F_n - \mathbb{Q}_n(F_n))$$

with the random fields $\overline{W}_n^{\Gamma,N}$ defined in (5.3). We complete the proof using the fact that $\overline{\gamma}_n^N(1)$ tends to 1, almost surely, as $N \to \infty$. This ends the proof of the theorem. □

We end this section with some comments on the asymptotic variance associated to the Gaussian fields $W_n^{\mathbb{Q}}$. Using (4.1), we prove that

$$\mathbb{Q}_n = \Psi_{\overline{D}_{p,n}(1)}(\eta_p)P_{p,n}$$

with the pair of integral operators $(\overline{D}_{p,n}, P_{p,n})$ from $\mathcal{B}(E_{[0,n]})$ into $\mathcal{B}(E_p)$

$$\overline{D}_{p,n}(F_n) := \frac{D_{p,n}(F_n)}{\eta_p Q_{p,n}(1)} = \frac{D_{p,n}(1)}{\eta_p Q_{p,n}(1)}\ P_{p,n}(F_n) \quad \text{and} \quad P_{p,n}(F_n) := \frac{D_{p,n}(F_n)}{D_{p,n}(1)}$$

from which we deduce the following formula

$$\overline{D}_{p,n}(F_n - \mathbb{Q}_n(F_n))(x_p) = \overline{D}_{p,n}(1)(x_p) \int [P_{p,n}(F_n)(x_p) - P_{p,n}(F_n)(y_p)]\ \Psi_{\overline{D}_{p,n}(1)}(\eta_p)(\mathrm{d}y_p). \tag{6.2}$$

Under condition $(M)_m$, for any function F_n of the form (2.5) with $\mathrm{osc}(f_p) \leq 1$, we have the following estimate

$$\|\overline{D}_{p,n}(1)\| \leq \delta^m \rho \Longrightarrow \mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right) \leq c(n+1).$$

This follows since $\mathrm{osc}(P_{p,n}(F_n))$ is uniformly (in p and n) bounded by virtue of the contraction property of the single-time marginals of $P_{p,n}$. (See Sect. 8 for more details.)

7. Non asymptotic $\mathbb{L}_r$-mean error estimates

This section is mainly concerned with the proof of Theorem 3.2. We follow the same semigroup techniques as the ones we used in Section 7.4.3 in [7] to derive uniform estimates w.r.t. the time parameter for the N-particle measures η_n^N. We use the decomposition

$$[\mathbb{Q}_n^N - \mathbb{Q}_n](F_n) = \sum_{0\leq p\leq n} \left(\frac{\eta_p^N D_{p,n}^N(F_n)}{\eta_p^N D_{p,n}^N(1)} - \frac{\eta_{p-1}^N D_{p-1,n}^N(F_n)}{\eta_{p-1}^N D_{p-1,n}^N(1)} \right)$$

with the conventions $\eta_{-1}^N D_{-1,n}^N = \eta_0 \mathcal{Q}_{0,n}$, for $p = 0$. Next, we observe that

$$\begin{aligned} \eta_{p-1}^N D_{p-1,n}^N(F_n) &= \int \eta_{p-1}^N(\mathrm{d}x_{p-1}) \mathcal{M}_{p-1}^N(x_{p-1}, \mathrm{d}(x_0,\ldots,x_{p-2})) \mathcal{Q}_{p-1,n}(x_{p-1}, \mathrm{d}(x_p,\ldots,x_n)) F_n(x_0,\ldots,x_n) \\ &= \int \eta_{p-1}^N(\mathrm{d}x_{p-1}) Q_p(x_{p-1},\mathrm{d}x_p) \mathcal{M}_{p-1}^N(x_{p-1}, \mathrm{d}(x_0,\ldots,x_{p-2})) \mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1},\ldots,x_n)) F_n(x_0,\ldots,x_n). \end{aligned}$$

On the other hand, we have

$$\eta_{p-1}^N(\mathrm{d}x_{p-1}) Q_p(x_{p-1},\mathrm{d}x_p) = \eta_{p-1}^N Q_p(\mathrm{d}x_p)\ M_{p,\eta_{p-1}^N}(x_p, \mathrm{d}x_{p-1})$$

from which we conclude that

$$\eta_{p-1}^N D_{p-1,n}^N(F_n) = (\eta_{p-1}^N Q_p)(D_{p,n}^N(F_n)).$$

This yields the decomposition

$$[\mathbb{Q}_n^N - \mathbb{Q}_n](F_n) = \sum_{0\leq p\leq n} \left(\frac{\eta_p^N D_{p,n}^N(F_n)}{\eta_p^N D_{p,n}^N(1)} - \frac{\Phi_p(\eta_{p-1}^N)(D_{p,n}^N(F_n))}{\Phi_p(\eta_{p-1}^N)(D_{p,n}^N(1))} \right) \tag{7.1}$$

with the convention $\Phi_0(\eta_{-1}^N) = \eta_0$, for $p = 0$. If we set

$$\widetilde{F}_{p,n}^N = F_n - \frac{\Phi_p(\eta_{p-1}^N)(D_{p,n}^N(F_n))}{\Phi_p(\eta_{p-1}^N)(D_{p,n}^N(1))}$$

then every term in the r.h.s. of (7.1) takes the following form

$$\frac{\eta_p^N D_{p,n}^N(\widetilde{F}_{p,n}^N)}{\eta_p^N D_{p,n}^N(1)} = \frac{\eta_p Q_{p,n}(1)}{\eta_p^N Q_{p,n}(1)} \times \left[\eta_p^N \overline{D}_{p,n}^N(\widetilde{F}_{p,n}^N) - \Phi_p(\eta_{p-1}^N)\overline{D}_{p,n}^N(\widetilde{F}_{p,n}^N)\right]$$

with the integral operators $\overline{D}_{p,n}^N$ defined in (5.4). Next, we observe that $D_{p,n}^N(1) = Q_{p,n}(1)$, and $\overline{D}_{p,n}^N(1) = \overline{D}_{p,n}(1)$. Thus, in terms of the local sampling random fields V_p^N, we have proved that

$$\frac{\eta_p^N D_{p,n}^N(\widetilde{F}_{p,n}^N)}{\eta_p^N D_{p,n}^N(1)} = \frac{1}{\sqrt{N}} \times \frac{1}{\eta_p^N \overline{D}_{p,n}(1)} \times V_p^N \overline{D}_{p,n}^N(\widetilde{F}_{p,n}^N) \tag{7.2}$$

and

$$\overline{D}_{p,n}^N(F_n) = \overline{D}_{p,n}(1) \times P_{p,n}^N(F_n) \quad \text{with} \quad P_{p,n}^N(F_n) := \frac{D_{p,n}^N(F_n)}{D_{p,n}^N(1)}\cdot \tag{7.3}$$

From these observations, we prove that

$$\frac{\Phi_p(\eta_{p-1}^N)(D_{p,n}^N(F_n))}{\Phi_p(\eta_{p-1}^N)(D_{p,n}^N(1))} = \frac{\Phi_p(\eta_{p-1}^N)(Q_{p,n}(1)\ P_{p,n}^N(F_n))}{\Phi_p(\eta_{p-1}^N)(Q_{p,n}(1))} = \Psi_{Q_{p,n}(1)}\left(\Phi_p(\eta_{p-1}^N)\right) P_{p,n}^N(F_n).$$

Arguing as in (6.2) we obtain the following decomposition

$$\overline{D}_{p,n}^N(\widetilde{F}_{p,n}^N)(x_p) = \overline{D}_{p,n}(1)(x_p) \times \int \left[P_{p,n}^N(F_n)(x_p) - P_{p,n}^N(F_n)(y_p)\right]\ \Psi_{Q_{p,n}(1)}(\Phi_p(\eta_{p-1}^N))(\mathrm{d}y_p)$$

and therefore

$$\left\|\overline{D}_{p,n}^N(\widetilde{F}_{p,n}^N)\right\| \leq b_{p,n}\operatorname{osc}(P_{p,n}^N(F_n)) \leq b_{p,n}\,\beta(P_{p,n}^N)\operatorname{osc}(F_n) \quad \text{with} \quad b_{p,n} \leq \sup_{x_p,y_p}\frac{Q_{p,n}(1)(x_p)}{Q_{p,n}(1)(y_p)}\cdot$$

We end the proof of (3.5) using the fact that for any $r \geq 1$, $p \geq 0$, $f \in \mathcal{B}(E_p)$ s.t. $\operatorname{osc}(f) \leq 1$ we have the almost sure Kintchine type inequality

$$\mathbb{E}\left(\left|V_p^N(f)\right|^r\ \mid \mathcal{F}_{p-1}^N\right)^{\frac{1}{r}} \leq a_r$$

for some finite (non random) constants $a_r < \infty$ whose values only depend on r. Indeed, using the fact that each term in the sum of (7.1) takes the form (7.2) we prove that

$$\sqrt{N}\ \mathbb{E}\left(\left|[\mathbb{Q}_n^N - \mathbb{Q}_n](F_n)\right|^r\right)^{\frac{1}{r}} \leq a_r \sum_{0\leq p\leq n} b_{p,n}^2\ \mathbb{E}\left(\operatorname{osc}(P_{p,n}^N(F_n))\right). \tag{7.4}$$

This ends the proof of the first assertion (3.5) of Theorem 3.2. For linear functionals of the form (2.5), it is easily checked that

$$D_{p,n}^N(F_n) = Q_{p,n}(1) \sum_{0\leq q\leq p}\left[M_{p,\eta_{p-1}^N}\dots M_{q+1,\eta_q^N}\right](f_q) + \sum_{p<q\leq n} Q_{p,q}(f_q\ Q_{q,n}(1))$$

with the convention $M_{p,\eta_{p-1}^N}\dots M_{p+1,\eta_p^N} = Id$, the identity operator, for $q = p$. Recalling that $D_{p,n}^N(1) = Q_{p,n}(1)$, we conclude that

$$P_{p,n}^N(F_n) = f_p + \sum_{0\leq q<p}\left[M_{p,\eta_{p-1}^N}\dots M_{q+1,\eta_q^N}\right](f_q) + \sum_{p<q\leq n}\frac{Q_{p,q}(Q_{q,n}(1)\ f_q)}{Q_{p,q}(Q_{q,n}(1))}$$

and therefore

$$P^N_{p,n}(F_n) = \sum_{0\leq q<p} \left[M_{p,\eta^N_{p-1}} \dots M_{q+1,\eta^N_q}\right](f_q) + \sum_{p\leq q\leq n} \frac{Q_{p,q}(Q_{q,n}(1)\ f_q)}{Q_{p,q}(Q_{q,n}(1))}$$

$$\frac{Q_{p,q}(Q_{q,n}(1)\ f_q)}{Q_{p,q}(Q_{q,n}(1))} = \frac{S_{p,q}(\overline{Q}_{q,n}(1)\ f_q)}{S_{p,q}(\overline{Q}_{q,n}(1))} \quad \text{with} \quad S_{p,q}(g) = \frac{Q_{p,q}(g)}{Q_{p,q}(1)}$$

with the potential functions $\overline{Q}_{q,n}(1)$ defined in (5.5). After some elementary computations, we obtain the following estimates

$$\operatorname{osc}(P^N_{p,n}(F_n)) \leq \sum_{0\leq q<p} \beta\left(M_{p,\eta^N_{p-1}} \dots M_{q+1,\eta^N_q}\right) \operatorname{osc}(f_q) + \sum_{p\leq q\leq n} b^2_{q,n}\ \beta(S_{p,q})\ \operatorname{osc}(f_q).$$

This ends the proof of the second assertion (3.6) of Theorem 3.2.

8. Comparisons with genealogical tree particle models

In this section, we provide with a brief comparison between these particle models and the genealogical tree particle interpretations of the measures $\mathbb{Q}_n$ discussed in (2.4).

8.1. Limiting variance interpretation models

Our first objective is to present a new interpretation of the pair of potential-transitions $(G_{p,n}, P_{p,n})$ defined in (3.2). We fix the time horizon n and we denote by $\mathbb{E}_{\mathbb{Q}_n}$ the expectation operator of a canonical random path $(X_0, \dots, X_n)$ under the measure $\mathbb{Q}_n$. For any function $F \in \mathcal{B}(E_{[p,n]})$, $p \leq n$, using (2.7) we check that

$$\mathbb{E}_{\mathbb{Q}_n}\left(F(X_p, \dots, X_n)\right) = \int \eta_n(\mathrm{d}x_n) \prod_{p<q\leq n} M_{q,\eta_{q-1}}(x_q, \mathrm{d}x_{q-1})\ F(x_p, \dots, x_n).$$

This implies that for any $F \in \mathcal{B}(E_{[0,p]})$, we have the $\mathbb{Q}_n$-almost sure formula

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}_n}\left(F(X_0, \dots, X_p) \mid (X_p, \dots, X_n)\right) &= \int \mathcal{M}_p(X_p, \mathrm{d}(x_0, \dots, x_{p-1}))F((x_0, \dots, x_{p-1}), X_p) \\
&= \mathbb{E}_{\mathbb{Q}_n}\left(F(X_0, \dots, X_p) \mid X_p\right).
\end{aligned}$$

Using elementary calculations, it is also easily checked that for any function $F \in \mathcal{B}(E_{[0,n]})$ we have the $\mathbb{Q}_n$-almost sure formula

$$\mathbb{E}_{\mathbb{Q}_n}\left(F(X_0, \dots, X_n) \mid (X_0, \dots, X_p)\right) = \frac{1}{\mathcal{Q}_{p,n}(1)(X_p)} \int \mathcal{Q}_{p,n}(X_p, \mathrm{d}(x_{p+1}, \dots, x_n))F((X_0, \dots, X_p), (x_{p+1}, \dots, x_n))$$

and therefore, for any function $F_n \in \mathcal{B}(E_{[0,n]})$, we prove that

$$\mathbb{E}_{\mathbb{Q}_n}\left(F_n(X_0, \dots, X_n) \mid X_p\right) = P_{p,n}(F_n)(X_p).$$

In much the same way, if we denote by $\mathbb{Q}^{(p)}_n$ the time marginal of the measure $\mathbb{Q}_n$ with respect to the pth coordinate, we have

$$\mathbb{Q}^{(p)}_n \ll \eta_p \quad \text{with} \quad \frac{\mathrm{d}\mathbb{Q}^{(p)}_n}{\mathrm{d}\eta_p} = G_{p,n}.$$

For centered functions F_n s.t. $\mathbb{Q}_n(F_n) = 0$, by the functional central limit Theorem 3.1, the limiting variance of the measures $\mathbb{Q}_n^N$ associated with the genetic model (2.2) with acceptance parameters $\epsilon_n = 0$ has the following interpretation:

$$\mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right) = \sum_{p=0}^{n} \eta_p \left[G_{p,n}^2 \; P_{p,n}(F_n)^2\right] = \sum_{p=0}^{n} \mathbb{E}_{\mathbb{Q}_n}\left(\frac{\mathrm{d}\mathbb{Q}_n^{(p)}}{\mathrm{d}\eta_p}(X_p)\; \mathbb{E}_{\mathbb{Q}_n}\left(F_n(X_0,\ldots,X_n)\,|X_p\right)^2\right).$$

We end this section with some estimates of these limiting variances. Arguing as in (6.2), for any $F_n \in \mathcal{B}(E_{[0,n]})$, we readily prove the estimate

$$\mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right) \leq \sum_{0\leq p\leq n} b_{p,n}^2 \; \mathrm{osc}(P_{p,n}(F_n))^2.$$

For linear functionals of the form (2.5), with functions $f_n \in \mathrm{Osc}_1(E_n)$, using the same lines of arguments as those we used at the end of Section 7, it is easily checked that

$$\mathrm{osc}(P_{p,n}(F_n)) \leq \sum_{0\leq q<p} \beta\left(M_{p,\eta_{p-1}}\ldots M_{q+1,\eta_q}\right) + \sum_{p\leq q\leq n} b_{q,n}^2 \; \beta(S_{p,q}).$$

Under the regularity condition $(M)_m$ stated in (3.7), the r.h.s. term in the above display is uniformly bounded with respect to the time parameters $0 \leq p \leq n$, from which we conclude that

$$\mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right) \leq c\;(n+1) \tag{8.1}$$

for some finite constant $c < \infty$, whose values do not depend on the time parameter.

8.2. **Variance comparisons**

We recall that the genealogical tree evolution models associated with the genetic type particle systems discussed in this article can be seen as the mean field particle interpretation of the Feynman-Kac measures η_n defined as in (2.12), by replacing the pair (X_n, G_n) by the historical process $\mathcal{X}_n$ and the potential function $\mathcal{G}_n$ defined below:

$$\mathcal{X}_n := (X_0,\ldots,X_n) \quad \text{and} \quad \mathcal{G}_n(\mathcal{X}_n) := G_n(X_n).$$

We also have a nonlinear transport equation defined as in (2.1) by replacing $K_{n,\eta_{n-1}}$ by some Markov transition $\mathcal{K}_{n,\eta_{n-1}}$ from $E_{[0,n-1]}$ into $E_{[0,n]}$. In this notation, the genealogical tree model coincides with the mean field particle model defined as in (2.2) by replacing K_{n,η_{n-1}^N} by $\mathcal{K}_{n,\eta_{n-1}^N}$, where η_{n-1}^N stands for the occupation measure of the genealogical tree model at time $(n-1)$. The local sampling errors are described by a sequence of random field model $\mathcal{V}_n^N, \mathcal{V}_n$ on $\mathcal{B}(E_{[0,n]})$ defined as in (2.13) and (2.14), by replacing $K_{n,\eta}$ by $\mathcal{K}_{n,\eta}$. More details on the path space technique can be found in Chapter 3 of the book [7].

The fluctuations of the genealogical tree occupation measures

$$\eta_n^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{(\xi_{0,n}^i,\xi_{1,n}^i,\ldots,\xi_{n,n}^i)} \quad \text{and} \quad \gamma_n^N := \left(\prod_{0\leq p<n} \eta_p^N(\mathcal{G}_p)\right) \times \eta_n^N \tag{8.2}$$

around their limiting values η_n and γ_n are described by the pair of empirical random fields defined below

$$\mathcal{W}_n^{\gamma,N} := \sqrt{N}\left(\gamma_n^N - \gamma_n\right) \quad \text{and} \quad \mathcal{W}_n^{\eta,N} := \sqrt{N}\;[\eta_n^N - \eta_n].$$

To describe the limiting Gaussian random fields $\mathcal{W}_n^{\gamma}$ and $\mathcal{W}_n^{\eta}$, we need another round of notation. Firstly, we observe that the pair of measures (γ_n, η_n) on the path space $E_{[0,n]}$ coincide with the measures $(\Gamma_n, \mathbb{Q}_n)$ we defined in the introduction of the present article. For these path space models, it is easily checked that

$$\gamma_n = \gamma_p \mathcal{D}_{p,n}$$

with the integral operator from $\mathcal{B}(E_{[0,n]})$ into $\mathcal{B}(E_{[0,p]})$ defined below

$$\mathcal{D}_{p,n}(F_n)(x_0,\dots,x_p) := \int \mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1},\dots,x_n))\ F_n((x_0,\dots,x_p),(x_{p+1},\dots,x_n)).$$

In the above display $\mathcal{Q}_{p,n}$ is the integral operator defined in (3.1). Notice that

$$\mathcal{D}_{p,n}(1)(x_0,\dots,x_p) = \mathcal{Q}_{p,n}(1)(x_p) = D_{p,n}(1)(x_p) = Q_{p,n}(1)(x_p).$$

As in (3.2), we consider be the pair of potential functions and Markov transitions $(\mathcal{G}_{p,n}, \mathcal{P}_{p,n})$ defined below

$$\mathcal{G}_{p,n}(x_0,\dots,x_p) = G_{p,n}(x_p) \quad \text{and} \quad \mathcal{P}_{p,n}(F_n) = \mathcal{D}_{p,n}(F_n)/\mathcal{D}_{p,n}(1). \tag{8.3}$$

In terms of conditional expectations, we readily prove that

$$\mathbb{E}_{\mathbb{Q}_n}\left(F_n(X_0,\dots,X_n) \mid (X_0,\dots,X_p)\right) = \mathcal{P}_{p,n}(F_n)(X_0,\dots,X_p) \tag{8.4}$$

for any function $F_n \in \mathcal{B}(E_{[0,n]})$.

It is more or less well known that the sequence of random fields $\mathcal{W}_n^{\gamma,N}$, resp. $\mathcal{W}_n^{\eta_n,N}$, converge in law, as $N \to \infty$, to the centered Gaussian fields $\mathcal{W}_n^{\gamma}$, resp. $\mathcal{W}_n^{\eta}$, defined as W_n^{Γ}, resp. $W_n^{\mathbb{Q}}$, by replacing the quantities $(V_p, G_{p,n}, D_{p,n}, P_{p,n}, \mathbb{Q}_n)$ by the path space models $(\mathcal{V}_p, \mathcal{G}_{p,n}, \mathcal{D}_{p,n}, \mathcal{P}_{p,n}, \eta_n)$; that is we have that

$$\begin{aligned}
\mathcal{W}_n^{\gamma}(F_n) &= \sum_{p=0}^{n} \gamma_p(1)\ \mathcal{V}_p\left(\mathcal{D}_{p,n}(F_n)\right) \\
\mathcal{W}_n^{\eta}(F_n) &= \sum_{p=0}^{n} \mathcal{V}_p\left(\mathcal{G}_{p,n}\ \mathcal{P}_{p,n}(F_n - \eta_n(F_n))\right).
\end{aligned}$$

A detailed discussion on these functional fluctuation theorems can be found in Chapter 9 in [7]. Arguing as before, for centered functions F_n s.t. $\mathbb{Q}_n(F_n) = 0$, the limiting variance of the genealogical tree occupation measures η_n^N associated with the genetic model (2.2) with acceptance parameters $\epsilon_n = 0$ has the following interpretation:

$$\begin{aligned}
\mathbb{E}\left(\mathcal{W}_n^{\eta}(F_n)^2\right) &= \sum_{p=0}^{n} \mathbb{E}_{\mathbb{Q}_n}\left(\frac{\mathrm{d}\mathbb{Q}_n^{(p)}}{\mathrm{d}\eta_p}(X_0,\dots,X_p)\ \mathbb{E}_{\mathbb{Q}_n}\left(F_n(X_0,\dots,X_n) \mid (X_0,\dots,X_p)\right)^2\right) \\
&= \mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right) + \sum_{p=0}^{n} \mathbb{E}_{\mathbb{Q}_n}\left(\frac{\mathrm{d}\mathbb{Q}_n^{(p)}}{\mathrm{d}\eta_p}(X_p)\ \mathrm{Var}_{\mathbb{Q}_n}\left(\mathcal{P}_{p,n}(F_n) \mid X_p\right)\right)
\end{aligned}$$

with the $\mathbb{Q}_n$-conditional variance of the conditional expectations (8.4) with respect to X_p given by

$$\mathrm{Var}_{\mathbb{Q}_n}\left(\mathcal{P}_{p,n}(F_n) \mid X_p\right) = \mathbb{E}_{\mathbb{Q}_n}\left(\left[\mathbb{E}_{\mathbb{Q}_n}\left(F_n(X_0,\dots,X_n) \mid (X_0,\dots,X_p)\right) - \mathbb{E}_{\mathbb{Q}_n}\left(F_n(X_0,\dots,X_n) \mid X_p\right)\right]^2 \mid X_p\right).$$

For sufficiently regular models, and for linear functionals of the form (2.5), with local functions $f_n \in \mathrm{Osc}_1(E_n)$, we have proved in (8.1) that $\mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right) \leq c\ (n+1)$, for some finite constant $c < \infty$, whose values do not depend on the time parameter. In this context, we also have that

$$\mathrm{Var}_{\mathbb{Q}_n}\left(\mathcal{P}_{p,n}(F_n)\,|X_p\right) = \mathbb{E}_{\mathbb{Q}_n}\left(\left[\sum_{0\leq q<p}\left(f_q(X_q) - \mathbb{E}_{\mathbb{Q}_n}\left(f_q(X_q)|X_p\right)\right)\right]^2 |X_p\right).$$

These local variance quantities may grow dramatically with the parameter p, so that the resulting variance $\mathbb{E}\left(\mathcal{W}_n^{\eta}(F_n)^2\right)$ will be much larger than $\mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right)$. For instance, in the toy model discussed in (2.6), we clearly have $\mathbb{Q}_n^{(p)} = \eta_p = \eta_0$ and

$$\mathbb{E}_{\mathbb{Q}_n}\left(F_n(X_0,\ldots,X_n)\,|(X_0,\ldots,X_p)\right) = \sum_{0\leq q\leq p} f(X_q)\cdot$$

from which we conclude that

$$\mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right) = (n+1) \quad \text{and} \quad \mathbb{E}\left(\mathcal{W}_n^{\eta}(F_n)^2\right) = \mathbb{E}\left(W_n^{\mathbb{Q}}(F_n)^2\right) + \frac{n(n+1)}{2}\cdot \tag{8.5}$$

9. Example

Consider the following hidden Markov model which is comprised of a $\mathbb{R}^d$-valued Markov chain

$$X_n = h_\theta(X_{n-1}, V_n), \tag{9.1}$$

where $(V_n)_{n\geq1}$ is a sequence of independent identically distributed random variables and a $\mathbb{R}^{d'}$-valued observed process $(Y_n)_{n\geq0}$

$$\mathbb{P}\left(Y_n \in \mathrm{d}y \mid X_n = x_n\right) = g_\theta(x_n, y_n)\ \lambda_n(\mathrm{d}y_n). \tag{9.2}$$

In the above displayed formulae, $h_\theta : \mathbb{R}^d \times \mathbb{R}^{n_v} \to \mathbb{R}^d$, $g_\theta(x_n, y_n)$ is the conditional density, with respect to the reference measure λ_n, of Y_n given $X_n = x_n$. The vector $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$ parameterizes the model. Given $(X_n)_{n\geq0} = (x_n)_{n\geq0}$, $(Y_n)_{n\geq0}$ is independent with laws $g_\theta(x_n, y_n)\lambda_n(\mathrm{d}y_n)$. We also assume that the law of $(V_n)_{n\geq1}$ and X_0 are independent of θ. By construction, the distribution of the observation path from the origin, up to a given time T is given by

$$\mathbb{P}_\theta\left((Y_0,\ldots,Y_T) \in \mathrm{d}(y_0,\ldots,y_T)\right) = \mathbb{E}_\theta\left\{\prod_{n=0}^{T} g_\theta(X_n, y_n)\right\}\ \lambda_0(\mathrm{d}y_0)\ldots\lambda_n(\mathrm{d}y_n)$$

where the expectation is computed with respect to the law of $(X_n)_{0\leq n\leq T}$ with parameter θ.

An important problem in statistics is that of estimating the model parameters that most aptly describes an observed real time series. There are many methods to solve this problem (see [23] for a recent review) and here we concentrate on the Maximum Likelihood method. Given a finite observation history $(y_n)_{0\leq n\leq T}$, the model parameter that best describes the data can be taken to be the maximizer of the density function given below

$$\theta \in \Theta \longrightarrow p_\theta(y_0,\ldots,y_T) = \mathbb{E}_\theta\left\{\prod_{n=0}^{T} g_\theta(X_n, y_n)\right\}.$$

An iterative procedure such as a gradient ascent method can be used to solve for this maximizer or for any one of them if there are local maxima. To implement gradient ascent, an estimate of the gradient of the likelihood is needed. Recently, an IPA gradient estimate has been proposed and implemented using the standard genealogical tree based particle system [5]. This approach is compared here with our backward particle system which is demonstrated below to have a much lower variance.

The IPA gradient of the log-likelihood $\log p_\theta(y_0,\dots,y_T)$ is given by

$$\nabla_\theta \log p_\theta(y_0,\dots,y_T) = \frac{\mathbb{E}_\theta\left\{R_n \prod_{n=0}^{T} g_\theta(X_n,y_n)\right\}}{\mathbb{E}_\theta\left\{\prod_{n=0}^{T} g_\theta(X_n,y_n)\right\}} \tag{9.3}$$

where

$$\begin{aligned} Z_n &= \nabla_\theta h_\theta(X_{n-1},V_n) + \nabla_x h_\theta(X_{n-1},V_n) Z_{n-1},\\ R_n &= R_{n-1} + \frac{Z_n^{\mathrm{T}} \nabla_x g_\theta(X_n,y_n) + \nabla_\theta g_\theta(X_n,y_n)}{g_\theta(X_n,y_n)}, \end{aligned}$$

with X_n evolving as in (9.1). (See [5] for the regularity conditions the model must satisfy for this to be true.) Here ∇_x indicates differentiation with respect to the state vector while ∇_θ differentiation with respect to the model vector. The quantity $Z_n^{\mathrm{T}} \in \mathbb{R}^{n_\theta} \times \mathbb{R}^d$ can be interpreted as the gradient of X_n with respect to θ, which indeed depends on all past state values. Z_0 and R_0 are initialized to the zero matrix and vector respectively.

The gradient $\nabla_\theta \log p_\theta(y_0,\dots,y_T)$ may be estimated using the standard genealogical tree based particle system for the Markov process $(X_n, Z_n, R_n)_{n\geq 0}$ with potentials $G_n(X_n,Z_n,R_n) = g_\theta(X_n,y_n)$ as was proposed by [5]. It is also straightforward to implement an on-line version of the backward particle model to estimate this same expectation. (Note though that this involves a function F_n in (2.4) which is slightly more complicated than the additive form in (2.5).) Details are omitted.

Results for the following linear Gaussian model are presented in Figure 1:

$$X_n = \phi X_{n-1} + \sigma_V V_n, \tag{9.4}$$
$$Y_n = c X_n + \sigma_W W_n, \tag{9.5}$$

where $(V_n)_{n\geq 1}$ and $(W_n)_{n\geq 1}$ are independent sequences of independent Gaussian random variables with zero mean and unit variance. The model vector is $\theta = (\phi, \sigma_V, c, \sigma_W)^{\mathrm{T}} \in \mathbb{R}\times(0,\infty)\times\mathbb{R}\times(0,\infty)$. Data was generated from this model with $(\phi,\sigma_V,c,\sigma_W) = (0.8, 0.1, 1, 1)$. Such a simple model was assumed since the gradient of the log-likelihood can be computed exactly with the Kalman filter, which serves as a benchmark. As can be seen from the boxplots, our estimator outperforms the genealogical tree based particle system. Note though that the computational cost is significantly higher for our method, *i.e.* $\mathcal{O}(N^2)$ compared to $\mathcal{O}(N)$ for the genealogical approach. However, from (8.5), we see that the new method with N particles may still outperform the genealogical tree based particle system with N^2 particles (in variance performance) as the variance of the latter can grow quadratically with time compared to linear growth of the former. (Admittedly, this growth was established in Section 8 under stringent mixing assumptions.) This has indeed been demonstrated in similar estimation problems involving models for which mixing has not been established [26].

Appendix

Proof of Lemma 4.3. We prove the lemma by induction on the parameter $n(>p)$. For $n = p+1$, we have

$$\mathcal{M}_{p+1,p,\eta}(x_{p+1}, \mathrm{d}x_p) = M_{p+1,\eta}(x_{p+1}, \mathrm{d}x_p) \quad \text{and} \quad \mathcal{Q}_{p,p+1}(x_p, \mathrm{d}x_{p+1}) = Q_{p+1}(x_p, \mathrm{d}x_{p+1}).$$

By definition of the transitions $M_{p+1,\eta}$, we have

$$\eta Q_{p+1}(\mathrm{d}x_{p+1})\ \mathcal{M}_{p+1,p,\eta}(x_{p+1}, \mathrm{d}x_p) = \eta(\mathrm{d}x_p)\ \mathcal{Q}_{p,p+1}(x_p, \mathrm{d}x_{p+1}).$$

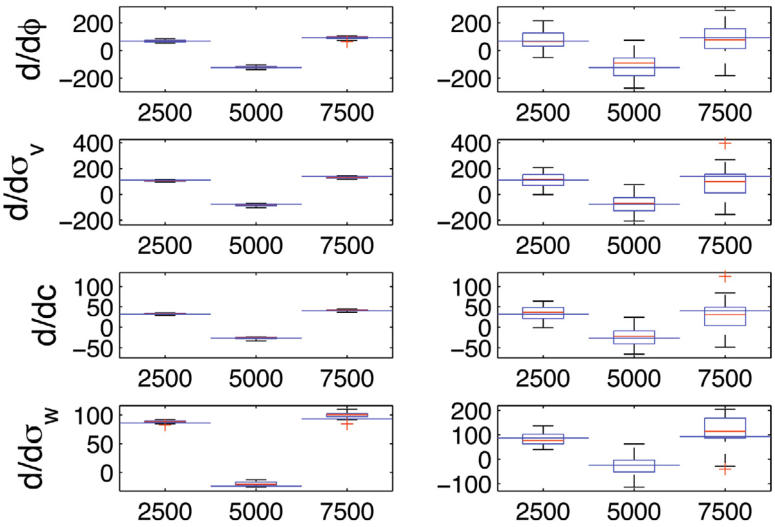

FIGURE 1. Comparison of the new backward particle system (left column) and genealogical tree based particle system (right column) implementation of the IPA derivative. Each column, from top to bottom, displays the estimate of the gradient of the log-likelihood computed with respect to $(\phi, \sigma_V, c, \sigma_W) = (0.8, 0.1, 1, 1)$. The gradient of the log-likehood was computed after 2500, 5000 and 7500 observations had been gathered. The long horizontal line dissecting the box plots are the true values computed with the Kalman filter.

We suppose that the result has been proved at rank n. In this situation, we notice that

$$\begin{aligned}\eta(\mathrm{d}x_p)\ \mathcal{Q}_{p,n+1}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_{n+1})) &= \eta(\mathrm{d}x_p)\ \mathcal{Q}_{p,n}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_n))Q_{n+1}(x_n, \mathrm{d}x_{n+1})\\ &= \eta Q_{p,n}(\mathrm{d}x_n)\ Q_{n+1}(x_n, \mathrm{d}x_{n+1})\ \mathcal{M}_{n,p,\eta}(x_n, \mathrm{d}(x_p, \ldots, x_{n-1}))\\ &= \eta Q_{p,n}(1)\ \Phi_{p,n}(\eta)(\mathrm{d}x_n)\ Q_{n+1}(x_n, \mathrm{d}x_{n+1})\ \mathcal{M}_{n,p,\eta}(x_n, \mathrm{d}(x_p, \ldots, x_{n-1})).\end{aligned}$$

Using the fact that

$$\Phi_{p,n}(\eta)(\mathrm{d}x_n)\ Q_{n+1}(x_n, \mathrm{d}x_{n+1}) = \Phi_{p,n}(\eta)Q_{n+1}(\mathrm{d}x_{n+1})\ M_{n+1,\Phi_{p,n}(\eta)}(x_{n+1}, \mathrm{d}x_n)$$

and

$$\eta Q_{p,n}(1)\ \Phi_{p,n}(\eta)Q_{n+1}(\mathrm{d}x_{n+1})\ = \eta Q_{p,n+1}(\mathrm{d}x_{n+1})$$

we conclude that

$$\begin{aligned}\eta(\mathrm{d}x_p)\ \mathcal{Q}_{p,n+1}(x_p, \mathrm{d}(x_{p+1}, \ldots, x_{n+1})) &= \eta Q_{p,n+1}(\mathrm{d}x_{n+1})\ M_{n+1,\Phi_{p,n}(\eta)}(x_{n+1}, \mathrm{d}x_n)\mathcal{M}_{n,p,\eta}(x_n, \mathrm{d}(x_p, \ldots, x_{n-1}))\\ &= \eta Q_{p,n+1}(\mathrm{d}x_{n+1})\ \mathcal{M}_{n+1,p,\eta}\left(x_{n+1}, \mathrm{d}(x_p, \ldots, x_n)\right).\end{aligned}$$

This ends the proof of the lemma. □

Proof of Lemma 6.1

Using the recursions (4.6), we prove that

$$\mathcal{M}_{n+1,p,\eta_p^N}\left(x_{n+1}, \mathrm{d}(x_p, \ldots, x_n)\right) = \mathcal{M}_{n+1,p+1,\Phi_{p+1}\left(\eta_p^N\right)}\left(x_{n+1}, \mathrm{d}(x_{p+1}, \ldots, x_n)\right) \times M_{p+1,\eta_p^N}(x_{p+1}, \mathrm{d}x_p).$$

On the other hand, we also have

$$\mathcal{M}^N_{p+1}(x_{p+1}, \mathrm{d}(x_0, \ldots, x_p)) = M_{p+1,\eta^N_p}(x_{p+1}, \mathrm{d}x_p)\mathcal{M}^N_p(x_p, \mathrm{d}(x_0, \ldots, x_{p-1}))$$

from which we conclude that

$$\mathcal{M}_{n+1,p+1,\Phi_{p+1}(\eta^N_p)}(x_{n+1}, \mathrm{d}(x_{p+1}, \ldots, x_n))\, \mathcal{M}^N_{p+1}(x_{p+1}, \mathrm{d}(x_0, \ldots, x_p)) = \mathcal{M}_{n+1,p,\eta^N_p}(x_{n+1}, \mathrm{d}(x_p, \ldots, x_n)) \times \mathcal{M}^N_p(x_p, \mathrm{d}(x_0, \ldots, x_{p-1})).$$

The end of the proof is now a direct consequence of the following decomposition

$$\begin{aligned}\mathcal{M}^N_n(x_n, \mathrm{d}(x_0, \ldots, x_{n-1})) - \mathcal{M}_n(x_n, \mathrm{d}(x_0, \ldots, x_{n-1})) = \sum_{1\le p\le n} \Big[\mathcal{M}_{n,p,\eta^N_p}(x_n, \mathrm{d}(x_p, \ldots, x_{n-1})) \\ \times \mathcal{M}^N_p(x_p, \mathrm{d}(x_0, \ldots, x_{p-1})) - \mathcal{M}_{n,p-1,\eta^N_{p-1}}(x_n, \mathrm{d}(x_{p-1}, \ldots, x_{n-1}))\, \mathcal{M}^N_{p-1}(x_{p-1}, \mathrm{d}(x_0, \ldots, x_{p-2}))\Big] \\ + \mathcal{M}_{n,0,\eta^N_0}(x_n, \mathrm{d}(x_0, \ldots, x_{n-1})) - \mathcal{M}_{n,0,\eta_0}(x_n, \mathrm{d}(x_0, \ldots, x_{n-1}))\end{aligned}$$

with the conventions

$$\mathcal{M}_{n,0,\eta^N_0}(x_n, \mathrm{d}(x_0, \ldots, x_{n-1}))\, \mathcal{M}^N_0(x_0, \mathrm{d}(x_0, \ldots, x_1)) = \mathcal{M}_{n,0,\eta^N_0}(x_n, \mathrm{d}(x_0, \ldots, x_{n-1}))$$

for $p = 0$, and for $p = n$

$$\mathcal{M}_{n,n,\eta^N_n}(x_n, \mathrm{d}(x_n, \ldots, x_{n-1}))\, \mathcal{M}^N_n(x_n, \mathrm{d}(x_0, \ldots, x_{n-1})) = \mathcal{M}^N_n(x_n, \mathrm{d}(x_0, \ldots, x_{n-1})).$$

Acknowledgements. S.S. Singh's research is funded by the Engineering and Physical Sciences Research Council under the First Grant Scheme (EP/G037590/1) and is gratefully acknowledged.

References

[1] D. Bakry, L'hypercontractivitée et son utilisation en théorie des semigroupes, in *Lecture Notes in Math.* **1581**, *École d'été de St. Flour XXII*, P. Bernard Ed. (1992).

[2] P. Billingsley, *Probability and Measure*. Third edition, Wiley series in probability and mathematical statistics (1995).

[3] E. Cancès, B. Jourdain and T. Lelièvre, Quantum Monte Carlo simulations of fermions. A mathematical analysis of the fixed-node approximation. *ESAIM: M2AN* **16** (2006) 1403–1449.

[4] F. Cerou, P. Del Moral and A. Guyader, A non asymptotic variance theorem for unnormalized Feynman-Kac particle models. *Ann. Inst. Henri Poincaré* (to appear).

[5] P.-A. Coquelin, R. Deguest and R. Munos, *Numerical methods for sensitivity analysis of Feynman-Kac models*. Available at http://hal.inria.fr/inria-00336203/en/, HAL-INRIA Research Report 6710 (2008).

[6] D. Crisan, P. Del Moral and T. Lyons, Interacting Particle Systems. Approaximations of the Kushner-Stratonovitch Equation. *Adv. Appl. Probab.* **31** (1999) 819–838.

[7] P. Del Moral, *Feynman-Kac formulae. Genealogical and interacting particle systems with applications*. Probability and its Applications, Springer Verlag, New York (2004).

[8] P. Del Moral and A. Doucet, Particle motions in absorbing medium with hard and soft obstacles. *Stoch. Anal. Appl.* **22** (2004) 1175–1207.

[9] P. Del Moral and L. Miclo, Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering, in *Séminaire de Probabilités XXXIV*, *Lecture Notes in Math.* **1729**, Springer, Berlin (2000) 1–145.

[10] P. Del Moral and L. Miclo, Particle approximations of Lyapunov exponents connected to Schroedinger operators and Feynman-Kac semigroups. *ESAIM: PS* **7** (2003) 171–208.

[11] P. Del Moral and E. Rio, *Concentration inequalities for mean field particle models*. Available at http://hal.inria.fr/inria-00375134/fr/, HAL-INRIA Research Report 6901 (2009).

[12] P. Del Moral, J. Jacod and P. Protter, The Monte Carlo Method for filtering with discrete-time observations. *Probab. Theory Relat. Fields* **120** (2001) 346–368.

[13] P. Del Moral, A. Doucet and S.S. Singh, *Forward smoothing using sequential Monte Carlo*. Cambridge University Engineering Department, Technical Report CUED/F-INFENG/TR 638 (2009).

[14] G.B. Di Masi, M. Pratelli and W.G. Runggaldier, An approximation for the nonlinear filtering problem with error bounds. *Stochastics* **14** (1985) 247–271.

[15] R. Douc, A. Garivier, E. Moulines and J. Olsson, *On the forward filtering backward smoothing particle approximations of the smoothing distribution in general state spaces models*. Technical report, available at `arXiv:0904.0316`.

[16] A. Doucet, N. De Freitas and N. Gordon Eds., *Sequential Monte Carlo Methods in Pratice*. Statistics for engineering and Information Science, Springer, New York (2001).

[17] M. El Makrini, B. Jourdain and T. Lelièvre, Diffusion Monte Carlo method: Numerical analysis in a simple case. *ESAIM: M2AN* **41** (2007) 189–213.

[18] M. Émery, *Stochastic calculus in manifolds*. Universitext, Springer-Verlag, Berlin (1989).

[19] S.J. Godsill, A. Doucet and M. West, Monte Carlo smoothing for nonlinear time series. *J. Am. Stat. Assoc.* **99** (2004) 156–168.

[20] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes* **24**. Second edition, North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam (1989).

[21] M. Kac, On distributions of certain Wiener functionals. *Trans. Am. Math. Soc.* **65** (1949) 1–13.

[22] G. Kallianpur and C. Striebel, *Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes*. Tech. Rep. 103, Department of Statistics, University of Minnesota, Minneapolis (1967).

[23] N. Kantas, A. Doucet, S.S. Singh and J.M. Maciejowski, An overview of sequential Monte Carlo methods for parameter estimation in general state-space models, in *Proceedings IFAC System Identification (SySid) Meeting*, available at `http://publications.eng.cam.ac.uk/16156/` (2009).

[24] H. Korezlioglu and W.J. Runggaldier, Filtering for nonlinear systems driven by nonwhite noises: an approximating scheme. *Stoch. Stoch. Rep.* **44** (1983) 65–102.

[25] J. Picard, Approximation of the nonlinear filtering problems and order of convergence, in *Filtering and control of random processes*, *Lecture Notes in Control and Inf. Sci.* **61**, Springer (1984) 219–236.

[26] G. Poyiadjis, A. Doucet and S.S. Singh, *Sequential Monte Carlo computation of the score and observed information matrix in state-space models with application to parameter estimation*. Technical Report CUED/F-INFENG/TR 628, Cambridge University Engineering Department (2009).

[27] D. Revuz, *Markov chains*. North-Holland (1975).

[28] M. Rousset, On the control of an interacting particle approximation of Schroedinger ground states. *SIAM J. Math. Anal.* **38** (2006) 824–844.

[29] A.N. Shiryaev, *Probability*, *Graduate Texts in Mathematics* **95**. Second edition, Springer (1986).

[30] D.W. Stroock, *Probability Theory: an Analytic View*. Cambridge University Press, Cambridge (1994).

[31] D.W. Stroock, *An Introduction to Markov Processes*, *Graduate Texts in Mathematics* **230**. Springer (2005).