

# A non asymptotic variance theorem for unnormalized Feynman-Kac particle models

Frédéric Cérou, Pierre del Moral, Arnaud Guyader

► **To cite this version:**

Frédéric Cérou, Pierre del Moral, Arnaud Guyader. A non asymptotic variance theorem for unnormalized Feynman-Kac particle models. [Research Report] RR-6716, INRIA. 2008. inria-00337392

**HAL Id: inria-00337392**

**<https://hal.inria.fr/inria-00337392>**

Submitted on 6 Nov 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*A non asymptotic variance theorem for  
unnormalized Feynman-Kac particle models*

Frédéric Cérou — Pierre Del Moral — Arnaud Guyader

**N° 6716**

Novembre 2008

Thème NUM



*rapport  
de recherche*



## A non asymptotic variance theorem for unnormalized Feynman-Kac particle models

Frédéric Cérou\* , Pierre Del Moral† , Arnaud Guyader‡

Thème NUM — Systèmes numériques  
Équipes-Projets Aspi et Cqfd

Rapport de recherche n° 6716 — Novembre 2008 — 26 pages

**Abstract:** We present a non asymptotic theorem for interacting particle approximations of unnormalized Feynman-Kac models. We provide an original stochastic analysis based on Feynman-Kac semigroup techniques combined with recently developed coalescent tree-based functional representations of particle block distributions. We present some regularity conditions under which the  $\mathbb{L}_2$ -relative error of these weighted particle measures grows linearly with respect to the time horizon yielding what seems to be the first results of this type for this class of unnormalized models. We also illustrate these results in the context of particle simulation of static Boltzmann-Gibbs measures and restricted distributions, with a special interest in rare event analysis.

**Key-words:** Interacting particle systems, Feynman-Kac semigroups, non asymptotic estimates, genetic algorithms, Boltzmann-Gibbs measures, Monte Carlo models, rare events

The author names appear in alphabetical order.

This work is supported in part by the French national programme “Sécurité ET Informatique” under project NEBBIANO, ANR-06-SETIN-009.

\* Inria Rennes - Bretagne Atlantique

† Inria Bordeaux - Sud Ouest

‡ Université de Haute Bretagne - Rennes 2

# **Théorème de variance non asymptotique pour des modèles particuliers de Feynman-Kac non normalisés**

**Résumé :** Nous présentons un théorème non asymptotique pour les approximation par systèmes de particules en interaction des modèles de Feynman-Kac non normalisés. Nous introduisons une analyse stochastique originale basée sur des techniques de semi-groupes de Feynman-Kac, associées avec les représentation, récemment proposées, des distributions de blocks de particules, en terme de développement en arbre de coalescence. Nous présentons des conditions de régularité sous lesquelles l'erreur relative  $\mathbb{L}_2$  de ces mesures particulières pondérées croît linéairement par rapport à l'horizon temporel, conduisant à ce qui semble être le premier résultat de ce type pour cette classe de modèles non normalisés. Nous illustrons ces résultats dans le contexte des mesures statiques de Boltzmann-Gibbs et des distributions restreintes, avec un intérêt particulier pour les événements rares.

**Mots-clés :** Systèmes de particules en interaction, semi-groupes de Feynman-Kac, estimées non asymptotiques, mesures de Boltzmann-Gibbs, modèles de Monte-Carlo, événements rares

## 1 Introduction

The field of Feynman-Kac path integrals and their particle interpretations are one of the most active contact points between probability, theoretical chemistry, quantum physics, and engineering sciences, including rare event analysis, and advanced signal processing. For a rather thorough discussion, the interested reader is recommended to consult the pair of books [5, 9], and the references therein. During the last two decades, the asymptotic analysis of these interacting particle models has been developed in various directions, including propagation of chaos analysis,  $\mathbb{L}_p$ -mean error estimates, central limit type theorems, and large deviation principles. Nevertheless, we emphasize that most of the non asymptotic results developed in the literature are concerned with empirical particle measures and normalized Feynman-Kac probability distributions. Thus, they do not apply to engineering or physical problems involving the computation of unnormalized Feynman-Kac models including rare event particle simulation and partition functions estimation in statistical mechanics.

Loosely speaking, unnormalized Feynman-Kac measures represent the distribution of the paths of a Markov process, weighted by the product of a given collection of non negative potential functions. The total masses of these measures are also called the normalizing constants. For instance, for set indicator potential functions the total mass of these functional represents the probability that the reference Markov chain stays in that set for a given number of time steps. We already mention that the particle approximations of these unnormalized measures are defined in terms of weighted products of empirical potential functions. The length of these products is directly related to the time horizon. The refined analysis of these unnormalized particle approximations requires to control the degeneracy of these weighted products in terms of the time parameter. The main objective of this article is to present non asymptotic  $\mathbb{L}_2$ -estimates for these weighted particle measures. Our main result is a non asymptotic variance estimate of the relative error with a degeneracy degree that grows linearly with respect to the time parameter. As shown in [2, 12] in the context of rare events, this result is sharp in the sense that the asymptotic variance of the relative errors grows linearly with respect to the time horizon. We design an original stochastic analysis that combines refined Feynman-Kac semigroup techniques with the recently developed algebraic tree-based functional representations of particle block distributions obtained by the second author with F. Patras and S. Rubenthaler in [8].

The rest of this article is organized as follows:

In a preliminary section, section 1.1, we provide a mathematical description of the Feynman-Kac models and their probabilistic particle interpretations. The advantage of the general Feynman-Kac model presented here is that it unifies the theoretical analysis of a variety of genetic type algorithms currently used in Bayesian statistics, biology, particle physics, and engineering sciences. It is clearly out of the scope of this article to present a detailed review of these particle approximation models. We rather refer the reader to the pair of research books [5, 9], and references therein. The main results of this article are briefly presented in the end of this section.

In section 1.2, we illustrate these rather abstract models with three concrete applications. The first one is concerned with the stochastic particle simulation of static Boltzmann-Gibbs measures. The second one is concerned with restricted

measure models. The third one deals with rare event analysis. More details on these application areas can be found in the articles [6, 7, 2].

Section 2 is concerned with some key combinatorial properties of tensor product measures and their interpretations in terms of coalescent type transitions.

In section 3 we provide some preliminary coalescent tree based expansions for particle tensor product measures. These algebraic developments are pivotal in our analysis of non asymptotic  $\mathbb{L}_2$ -estimates for unnormalized Feynman-Kac measures. We already mention that these expansions are expressed in terms of pairwise coalescent Markov transitions and Feynman-Kac semigroups.

Section 4 is devoted to the analysis of the total mass of unnormalized Feynman-Kac semigroups. We provide a series of regularity conditions under which the relative variation of these quantities depends only linearly on the time horizon of these semigroups.

In section 5, we state and prove the main results of the present article. We examine non homogeneous models including degenerate potential functions that may vanish on some state space regions.

In the final section, section 6, we outline the preceding results in terms of efficiency for rare event probability estimation. Roughly speaking, we want to control the relative variance of our estimator when the event of interest is getting more and more rare. Our main result enables us to derive an efficiency result for rare event probability estimation, the first of its kind concerning the Interacting Particle System (IPS) approach applied to rare events.

We end this introduction with a brief review of some of the standard notation used in the present article. We denote respectively by  $\mathcal{M}(E)$ ,  $\mathcal{P}(E)$ , and  $\mathcal{B}_b(E)$ , the set of bounded and signed measures, the subset of all probability measures on some measurable space  $(E, \mathcal{E})$ , and the Banach space of all bounded and measurable functions  $f$  on  $E$  equipped with the uniform norm  $\|f\| = \sup_{x \in E} |f(x)|$ . We denote by  $\mu(f) = \int \mu(dx) f(x)$ , the Lebesgue integral of a function  $f \in \mathcal{B}_b(E)$ , with respect to a measure  $\mu \in \mathcal{M}(E)$ . We slightly abuse the notation, and sometimes denote by  $\mu(A) = \mu(1_A)$  the measure of a measurable subset  $A \in \mathcal{E}$ . Recall that a bounded integral operator  $M$  from a measurable space  $E$  into itself, is an operator  $f \mapsto M(f)$  from  $\mathcal{B}_b(E)$  into itself such that the functions  $M(f)(x) = \int_F M(x, dy) f(y)$  are measurable and bounded, for any  $f \in \mathcal{B}_b(E)$ . A bounded integral operator  $M$  from a measurable space  $(E, \mathcal{E})$  into itself also generates a dual operator  $\mu \mapsto \mu M$  from  $\mathcal{M}(E)$  into  $\mathcal{M}(E)$  defined by  $(\mu M)(f) := \mu(M(f))$ . Given a pair  $(M_1, M_2)$  of bounded integral operators we denote by  $M_1 M_2$  the composition of the operators given by the following formula  $(M_1 M_2)(x, dz) = \int M_1(x, dy) M_2(y, dz)$ . We also set  $M^m = M^{m-1} M = M M^{m-1}$  the  $m$  composition transition, with  $m \geq 1$ , and we use the conventions  $(\sum_{\emptyset}, \prod_{\emptyset}) = (0, 1)$ . Finally, the tensor product operator  $M^{\otimes 2}$  is the bounded integral operator defined for every function  $f \in \mathcal{B}_b(E \times E)$  by

$$M^{\otimes 2}(f)(x, x') = \int_{E \times E} M(x, dy) M(x', dy') f(y, y').$$

## 1.1 Description of the models and statement of some results

We consider a collection of bounded potential functions  $G_n$  on the state space  $E$ , a distribution  $\eta_0$  on  $E$ , and a collection of Markov transitions  $M_n(x, dy)$

from  $E$  into itself. We associate to these objects the Feynman-Kac measures, defined for any  $f \in \mathcal{B}_b(E)$  by the formulae

$$\eta_n(f) = \gamma_n(f)/\gamma_n(1) \quad \text{with} \quad \gamma_n(f) = \mathbb{E}[f(X_n) \prod_{0 \leq k < n} G_k(X_k)] \quad (1.1)$$

In (1.1),  $(X_n)_{n \geq 0}$  represents a Markov chain with initial distribution  $\eta_0$ , and elementary transitions  $(M_n)_{n > 0}$ . To simplify the presentation, and avoid unnecessary technical discussion, we shall suppose that the potential functions take values in  $[0, 1]$  and for any  $n \geq 0$  we have  $\eta_n(G_n) > 0$ . By the Markov property and the multiplicative structure of (1.1), it is easily checked that the flow  $(\eta_n)_{n \geq 0}$  satisfies the following equation

$$\eta_{n+1} = \Phi_n(\eta_n) := \Psi_{G_n}(\eta_n)M_{n+1} \quad (1.2)$$

with the Boltzmann-Gibbs transformation  $\Psi_{G_n}$  defined below :

$$\Psi_{G_n}(\eta_n)(dx) := \frac{1}{\eta_n(G_n)} G_n(x) \eta_n(dx)$$

We also readily check the following multiplicative formula

$$\gamma_n(1) = \prod_{0 \leq p < n} \eta_p(G_p) \quad (1.3)$$

The particle approximation of the flow (1.2) depends on the choice of the McKean interpretation model. These probabilistic interpretations consist of a chosen collection of Markov transitions  $K_{n+1, \eta_n}$ , indexed by the set of probability measures  $\eta_n$  on  $E$ , and satisfying the compatibility condition  $\Phi_n(\eta_n) = \eta_n K_{n+1, \eta_n}$ . The choice of these Markov transitions is far from being unique. By (1.2), we find that

$$\forall n \geq 0 \quad \forall \alpha \in [0, 1] \quad \eta_{n+1} = \eta_n K_{n+1, \eta_n}^{(\alpha)} \quad (1.4)$$

with the McKean transition  $K_{n+1, \eta_n}^{(\alpha)} = S_{\alpha G_n, \eta_n} M_{n+1}$  and the selection type transition

$$S_{\alpha G_n, \eta_n}(x, dy) = \alpha G_n(x) \delta_x(dy) + (1 - \alpha G_n(x)) \Psi_{G_n}(\eta_n)(dy)$$

**Definition 1.1** *The mean field particle interpretation of the evolution equation (1.4) is the  $E^N$ -valued Markov chain  $X_n^{(N)} = \left( X_n^{(N, i)} \right)_{1 \leq i \leq N}$  with elementary transitions*

$$\mathbb{P} \left( X_{n+1}^{(N)} \in dx_{n+1} \mid X_n^{(N)} \right) = \prod_{i=1}^N K_{n+1, \eta_n^N}^{(\alpha)}(X_n^{(N, i)}, dx_{n+1}^i), \quad (1.5)$$

where  $\eta_n^N$  stands for the occupation measure  $\eta_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_n^{(N, i)}}$  of the  $N$ -uple  $X_n^{(N)}$  at time  $n$ . The initial configuration  $X_0^{(N)} = \left( X_0^{(N, i)} \right)_{1 \leq i \leq N}$  consists of  $N$  independent and identically distributed random variables with distribution  $\eta_0$ .



In our context, it is worth mentioning that the elementary transitions of the chain  $X_n^{(N)} \rightsquigarrow X_{n+1}^{(N)}$  are decomposed into two separate mechanisms: Firstly, the current state  $X_n^{(N,i)}$  of each individual with label  $i \in \{1, \dots, N\}$  performs an acceptance-rejection type transition  $X_n^{(N,i)} \rightsquigarrow \widehat{X}_n^{(N,i)}$  according to Markov transition:  $S_{\alpha G_n, \eta_n^N}$ . In other words with a probability  $\alpha G_n(X_n^{(N,i)})$  the particle remains in the same site and we set  $\widehat{X}_n^{(N,i)} = X_n^{(N,i)}$ . Otherwise it jumps to a new location randomly chosen according to the Boltzmann-Gibbs distribution

$$\Psi_{G_n}(\eta_n^N) = \sum_{j=1}^N \frac{G_n(X_n^{(N,j)})}{\sum_{k=1}^N G_n(X_n^{(N,k)})} \delta_{X_n^{(N,j)}}$$

After the acceptance-rejection stage, the selected individuals  $\widehat{X}_n^{(N,i)}$  evolve independently to a new site  $X_{n+1}^{(N,i)}$  randomly chosen with distribution given by  $M_{n+1}(\widehat{X}_n^{(N,i)}, dx)$ .

The above model can be extended in various ways. For instance we can consider acceptance parameters  $\alpha_n(\eta_n)$  that depend on the time parameter as well as on the current measure  $\eta_n$ . All the results presented in this article remain valid for these extended models. We also emphasize that for  $\alpha = 0$ , we find that  $K_{n+1, \eta_n}^{(0)}(x, dy) = \Phi_{n+1}(\eta_n)(dy)$ . In this situation,  $X_n^{(N)}$  evolves as a simple genetic algorithm with mutation transitions  $M_n$ , and selection fitness functions  $G_n$ .

Besides the fact that  $\eta_n(G_n) > 0$ , it is important to mention that the empirical quantities  $\eta_n^N(G_n)$  may vanish at a given random time. The formal definition of this time is given below.

**Definition 1.2** We let  $\tau^N$  be the first time  $n$  we have  $\eta_n^N(G_n) = 0$ .

$$\tau^N = \inf \{n \geq 0 : \eta_n^N(G_n) = 0\}$$

At time  $\tau^N$ , the particle algorithm stops and from that time the particle approximation measures are defined as the null measures (see for instance chapter 7, section 7.2.2. in [5]) :

$$\forall n > \tau^N \quad \eta_n^N = 0$$

Mimicking the multiplicative formula (1.3), we also consider the following  $N$ -particle approximation of the unnormalized Feynman-Kac measures.

**Definition 1.3** The  $N$ -particle approximation measures  $\gamma_n^N$  associated with the unnormalized Feynman-Kac models  $\gamma_n$  introduced in (1.1) are defined for any  $f \in \mathcal{B}_b(E)$  by the following formulae:

$$\gamma_n^N(f) = \gamma_n^N(1) \times \eta_n^N(f) \quad \text{with} \quad \gamma_n^N(1) = \prod_{0 \leq p < n} \eta_p^N(G_p)$$

As an aside, we observe that  $\gamma_n^N = 0$ , for any time  $n > \tau^N$ . It is well known that the particle measures  $\gamma_n^N$  are unbiased estimates of the unnormalized Feynman-Kac measures  $\gamma_n$  (see for instance [5], theorem 7.4.2. p.239). That is we have that

$$\forall f \in \mathcal{B}_b(E) \quad \mathbb{E}(\gamma_n^N(f)) = \gamma_n(f)$$

It is obviously out of the scope of this article to present a full asymptotic analysis of these particle models. We refer the interested reader to the book [5] and the series of articles [6, 7, 11] and the references therein. For instance, it is well known that the particle occupation measures converge to the desired Feynman-Kac measures as the size of the population tends to infinity. That is, we have with various precision estimates, and as  $N$  tends to infinity, the weak convergence results  $\lim_{N \rightarrow \infty} \eta_n^N = \eta_n$  and  $\lim_{N \rightarrow \infty} \gamma_n^N = \gamma_n$ .

To give a flavor of our main results, we discuss non asymptotic variance estimates only for time homogeneous models  $(G_n, M_n) = (G, M)$ .

**Theorem 1.4** *For the simple genetic algorithm corresponding to the choice  $\alpha = 0$ , we have an explicit decomposition formula of the following form*

$$\forall N > 1 \quad \mathbb{E} \left( \left[ \frac{\gamma_n^N(1)}{\gamma_n(1)} - 1 \right]^2 \right) = \left( 1 - \frac{1}{N} \right)^{(n+1)} \sum_{s=1}^{n+1} \frac{1}{(N-1)^s} v_n(s) \quad (1.6)$$

for some finite constants  $v_n(s)$  explicitly described in terms of Feynman-Kac type coalescent tree based expansions and whose values do not depend on the precision parameter  $N$ . A similar upper bound holds true for any choice of the acceptance parameter.

In addition, suppose that the pair of potential-transitions  $(G, M)$  are chosen so that

$$\forall (x, x') \in E^2 \quad G(x) \leq \delta G(x') \quad \text{and} \quad M^m(x, dy) \leq \beta M^m(x', dy) \quad (1.7)$$

for some  $m \geq 1$  and some parameters  $(\delta, \beta) \in [1, \infty]^2$ . In this situation, for any  $n \geq 0$  and any  $N > (n+1)\beta\delta^m$  we have

$$\mathbb{E} \left( \left[ \frac{\gamma_n^N(1)}{\gamma_n(1)} - 1 \right]^2 \right) \leq \frac{4}{N} (n+1) \beta \delta^m$$

A similar assertion is also true if we replace in condition (1.7) the triplet  $(E, G, M)$  by the triplet  $(\widehat{E}, \widehat{G}, \widehat{M})$  given below

$$\widehat{E} := G^{-1}([0, 1]) \quad \widehat{G}(x) := M(G)(x)$$

and

$$\widehat{M}(x, dy) := M(x, dy)G(y)/M(G)(x)$$

The proof of the first assertion of the theorem is housed in section 3. The second part of the theorem is proved in section 5.

We end this section with some comments on the impact of these non asymptotic estimates. Firstly, we mention that the convergence analysis of the occupation measures  $\eta_n^N$  around their limiting values is rather well understood (see for instance [5], theorem 7.4.4. p.246). For instance, under the regularity conditions (1.7) it is well known that for every bounded Borel function  $f$ , with  $\|f\| \leq 1$ , we have the uniform estimates

$$\sup_{n \geq 0} \mathbb{E} \left[ \left[ \eta_n^N(f) - \eta_n(f) \right]^2 \right]^{\frac{1}{2}} \leq \frac{c(\beta, \delta)}{\sqrt{N}}$$

for some finite constant whose values only depend on the pair of parameters  $(\beta, \delta)$ . Using the following decomposition

$$\frac{\gamma_n^N(f)}{\gamma_n(1)} - \eta_n(f) = \left[ \frac{\gamma_n^N(1)}{\gamma_n(1)} - 1 \right] \eta_n^N(f) + [\eta_n^N(f) - \eta_n(f)]$$

we readily find that for any  $N > (n+1)\beta\delta^m$

$$\sqrt{N} \mathbb{E} \left[ \left[ \frac{\gamma_n^N(f)}{\gamma_n(1)} - \eta_n(f) \right]^2 \right]^{\frac{1}{2}} \leq 2 \sqrt{(n+1) \beta \delta^m} + c(\beta, \delta)$$

We slightly abuse the notation and we consider the following flow of Feynman-Kac measures on path spaces

$$\eta_n(f_n) = \gamma_n(f_n)/\gamma_n(1) \quad \text{with} \quad \gamma_n(f_n) = \mathbb{E}[f_n(X_0, \dots, X_n) \prod_{0 \leq k < n} G_k(X_k)]$$

for any bounded function  $f_n \in \mathcal{B}_b(E^{n+1})$ . It is rather well known that these models can also be expressed as the ones presented in (1.1) through a state-space enlargement. In this context, under the regularity conditions (1.7) it is more or less well known that for every function  $f_n$ , with  $\|f_n\| \leq 1$ , we have the uniform estimates

$$\sup_{n \geq 0} \mathbb{E} \left[ [\eta_n^N(f_n) - \eta_n(f_n)]^2 \right]^{\frac{1}{2}} \leq c(\beta, \delta) \sqrt{\frac{n+1}{N}}$$

Arguing as above, we readily find that for any  $N > (n+1)\beta\delta^m$

$$\sqrt{\frac{N}{n+1}} \mathbb{E} \left[ \left[ \frac{\gamma_n^N(f)}{\gamma_n(1)} - \eta_n(f) \right]^2 \right]^{\frac{1}{2}} \leq 2 \sqrt{\beta \delta^m} + c(\beta, \delta)$$

## 1.2 Some model application areas

We consider a flow of Boltzmann-Gibbs measures  $\eta_n$  associated with some reference measure  $\lambda$  and some  $[0, 1]$ -valued potential function  $g_n$  on some measurable state space  $E$  and given by the following formulae:

$$\eta_n(dx) = \Psi_{g_n}(\lambda)(dx) := \frac{1}{\mathcal{Z}_n} g_n(x) \lambda(dx) \quad \text{with} \quad \mathcal{Z}_n := \lambda(g_n) > 0$$

For instance, in combinatorial counting problems  $\lambda$  often represents the uniform measure on a given finite set  $E$  and  $g_n = 1_{A_n}$  is the indicator function of a subset  $A_n \subset E$ . In global optimization problems,  $\lambda$  may represent a reference measure on a given state space  $E$  and  $g_n = \exp(-\beta_n V)$  a Boltzmann-Gibbs factor associated with an inverse temperature parameter  $\beta_n$  and an energy type non negative function  $V$ . We further assume that there exists a sequence of potential functions  $G_n$  satisfying the multiplicative product condition

$$g_{n+1} = g_n \times G_n$$

To simplify the presentation, we assume that  $g_0 = 1$ , so that  $\eta_0 = \lambda$  and  $g_1 = G_0$ .

Returning to the pair of examples discussed above, we can choose the exponential Boltzmann-Gibbs potential and respectively the indicator potential

functions defined by  $G_n = \exp(-(\beta_{n+1} - \beta_n)V)$  and respectively  $G_n = 1_{A_{n+1}}$ , with a non decreasing inverse temperature parameter  $\beta_n$  and a non increasing sequence of subsets  $A_n$ . By construction, we readily check that the sequence of target measures  $(\eta_n)_{n \geq 0}$  satisfies the following recursion  $\eta_{n+1} = \Psi_{G_n}(\eta_n)$ . To go one step further, we consider a sequence of Markov transitions  $M_n(x, dy)$  from  $E$  into itself, with a prescribed invariant measure  $\eta_n = \eta_n M_n$ . We let  $X_n$  be a Markov chain with initial distribution  $\text{Law}(X_0) = \lambda$  and elementary transitions  $M_n$ . For instance, Metropolis-Hastings transitions often consist in sampling a local move, and to accept or reject the proposal with an appropriate acceptance ratio so that the target measure  $\pi_n$  is reversible with respect to these local transitions. In this connection, we mention that the standard and rather well known simulated annealing algorithm consists in finding judicious increasing schedule  $\beta_n \uparrow$  with sufficiently slow variations to ensure that successive Metropolis-Hasting moves between these changes are close to equilibrium. For product state space models, another strategy is to consider a Gibbs sampling transition. These random transitions consist in sampling local moves by updating each coordinate of the state according to appropriate conditional distributions. By construction, we readily check that the flow of measures  $\eta_n$  satisfies the recursion (1.2). Using the Markov property and the multiplicative structure of our models, we can also prove that the measures  $\eta_n$  given above coincide with the flow of Feynman-Kac measures defined in (1.1). Finally, we observe that  $\eta_p(g_p) = \lambda(g_p G_p) / \lambda(g_p) = \lambda(g_{p+1}) / \lambda(g_p)$ , from which we conclude that  $\gamma_n(1) = \lambda(g_n)$ . As noticed by [13], in the context of molecular simulation, the Feynman-Kac representation formula of static Boltzmann-Gibbs measures is also known as Jarzynski's nonequilibrium dynamic model [10].

Another important practical application concerns rare events, as discussed for instance in [2]. By rare event we mean an event whose probability is too small to be accurately estimated by a simple Monte Carlo procedure in a reasonable time. Practically, this is the case if this probability is less than, say  $10^{-9}$ . In this case, the normalizing constant  $\gamma_n(1)$  is the probability, to be estimated, of the rare event under consideration.

One of the most used model for rare event is the following. Let  $Z = \{Z_t, t \geq 0\}$  be a continuous-time strong Markov process taking values in some Polish state space  $S$ . For a given target Borel set  $A \subset S$  we define the hitting time

$$T_B = \inf\{t \geq 0 : Z_t \in A\},$$

as the first time when the process  $Z$  hits  $A$ . In many applications, the set  $A$  is the (super) level set of a scalar measurable function  $\phi$  defined on  $S$ , i.e.

$$A = \{z \in S : \phi(z) \geq \lambda_A\}.$$

It may happen that most of the realizations of  $X$  never reach the set  $A$ . The corresponding rare event probabilities are extremely difficult to analyze. In particular one would like to estimate the quantities

$$\mathbb{P}(T_B \leq T) \quad \text{and} \quad \text{Law}(Z_t, 0 \leq t \leq T_B \mid T_B \leq T),$$

where  $T$  is a  $\mathbb{P}$ -almost surely finite stopping time, for instance the hitting time of a recurrent Borel set  $R \subset S$ , i.e.  $T = T_R$  with

$$T_R = \inf\{t \geq 0 : Z_t \in R\} \quad \text{and} \quad \mathbb{P}(T_R < \infty) = 1.$$

In practice the process  $Z$ , before visiting  $R$  or entering into the desired set  $A$ , passes through a decreasing sequence of closed sets

$$A = A_{n^*} \subset A_{n^*-1} \subset \cdots \subset A_2 \subset A_1 \subset A_0 . \quad (1.8)$$

The parameter  $n^*$  and the sequence of level sets depend on the problem at hand. We can easily fit this problem in the Feynman-Kac model presented in section 1.1 simply by setting

$$\forall 1 \leq n \leq n^* \quad X_n := (Z_t, T_{n-1} \wedge T \leq t \leq T_n \wedge T) \in E$$

with

$$E := \bigcup_{t' \leq t''} \mathbb{D}([t', t''], S)$$

where, with a slight abuse of notation,  $T_n$  stands for the first time  $T_{A_n}$  the process  $Z$  reaches  $A_n$ , that is

$$T_n = \inf\{t \geq 0 : Z_t \in A_n\}$$

with the convention  $\inf \emptyset = \infty$ , and with  $\mathbb{D}(I, S)$  being the set of the trajectories in  $S$  indexed by the interval  $I$ . The potential functions  $G_n$  on  $E$  defined for each  $x = (x_t, t' \leq t \leq t'') \in \mathbb{D}([t', t''], S)$  with  $t' \leq t''$  by

$$G_n(x) = 1_{(x_{t''} \in A_n)} .$$

In this notation, we have  $T_A = T_{n^*}$  and for every  $n \leq n^*$

$$\gamma_n(1) = \mathbb{P}(T_n \leq T) \quad \text{and} \quad \eta_n = \text{Law}(X_n \mid T_n \leq T) , \quad (1.9)$$

For more details on these excursion valued Feynman-Kac models, we refer the reader to [2]. As we will show in the last section, our main result enables us to derive an efficiency result for rare event probability estimation, the first of its kind concerning the Interacting Particle System (IPS) approach applied to rare events.

## 2 Particle tensor product measures

In this short section, we provide some key combinatorial properties of empirical measures. We also define the unnormalized versions of the  $N$ -tensor product measures associated with the particle models introduced in section 1.1.

Firstly, we denote by  $m(x) = \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{x^i}$ , the empirical measure associated with some  $N$ -uple  $x = (x^1, \dots, x^N) \in E^N$ . We also introduce the tensor product measures

$$m(x)^{\otimes 2} = \frac{1}{N^2} \sum_{1 \leq i, j \leq N} \delta_{(x^i, x^j)} \quad \text{and} \quad m(x)^{\odot 2} = \frac{1}{N(N-1)} \sum_{1 \leq i \neq j \leq N} \delta_{(x^i, x^j)}$$

**Definition 2.1** We denote by  $C$  the coalescent type integral operator defined by

$$\forall F \in \mathcal{B}_b(E^2) \quad \forall (x, y) \in E^2 \quad C(F)(x, y) = F(x, x)$$

**Lemma 2.2** For any  $F \in \mathcal{B}_b(E^2)$ , we have

$$m(x)^{\otimes 2}(F) = \frac{1}{N} m(x)^{\odot 2}(C(F)) + \left(1 - \frac{1}{N}\right) m(x)^{\odot 2}(F) \quad (2.1)$$

**Proof:**

We use the fact that

$$\begin{aligned} & m(x)^{\otimes 2}(F) \\ &= \frac{1}{N^2} \sum_{1 \leq i, j \leq N} F(x^i, x^j) \\ &= \frac{1}{N^2} \sum_{1 \leq i \leq N} F(x^i, x^i) + \frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} F(x^i, x^j) \\ &= \frac{1}{N} \left( \frac{1}{N(N-1)} \sum_{1 \leq i \neq j \leq N} F(x^i, x^i) \right) + \frac{N(N-1)}{N^2} \left( \frac{1}{N(N-1)} \sum_{1 \leq i \neq j \leq N} F(x^i, x^j) \right) \end{aligned}$$

to check that

$$m(x)^{\otimes 2}(F) = \frac{1}{N} m(x)^{\odot 2}(C(F)) + \left(1 - \frac{1}{N}\right) m(x)^{\odot 2}(F)$$

■

We end this section with the definition of the tensor product measures associated with the mean field particle model presented in section 1.1.

**Definition 2.3** For any population size  $N \geq 1$  and any time parameter  $n \geq 0$ , we set

$$(\gamma_n^N)^{\otimes 2}(F) = \gamma_n^N(1)^2 \times (\eta_n^N)^{\otimes 2}(F) \quad \text{and} \quad (\gamma_n^N)^{\odot 2}(F) := \gamma_n^N(1)^2 \times (\eta_n^N)^{\odot 2}(F)$$

with the tensor product measures given by  $(\eta_n^N)^{\otimes 2} := m(X_n^{(N)})^{\otimes 2}$  and  $(\eta_n^N)^{\odot 2} := m(X_n^{(N)})^{\odot 2}$ .

One can then easily check that for every test function  $\phi \in \mathcal{B}_b(E \times E)$  :

$$\mathbb{E} \left( (\eta_n^N)^{\odot 2}(\phi) \mid X_{n-1}^{(N)} \right) = \mathbb{E} \left( (\eta_{n-1}^N)^{\odot 2} \left[ K_{n, \eta_{n-1}^N}^{\otimes 2}(\phi) \right] \mid X_{n-1}^{(N)} \right) \quad (2.2)$$

### 3 Coalescent tree based expansions

The functional coalescent tree based expansions developed in this section are described in terms of the Feynman-Kac semigroups defined below.

**Definition 3.1** We let  $Q_{p,n}$ , with  $0 \leq p \leq n$ , be the Feynman-Kac semi-group associated with the flow  $\gamma_n = \gamma_p Q_{p,n}$ . For  $p = n$ , we use the convention that  $Q_{n,n} = Id$ .

Using the Markov property, it is not difficult to check that  $Q_{p,n}$  has the following functional representation

$$Q_{p,n}(f_n)(x_p) = \mathbb{E} \left[ f_n(X_n) \prod_{p \leq k < n} G_k(X_k) \mid X_p = x_p \right] \quad (3.1)$$

for any test function  $f_n \in \mathcal{B}_b(E)$ , and any state  $x_p \in E$ .

To simplify the presentation, we notice that formula (2.1) stated in lemma 2.2 can be rewritten as follows

$$m(x)^{\otimes 2}(F) = \mathbb{E} \left( m(x)^{\odot 2}(C_\epsilon(F)) \right) \quad (3.2)$$

where  $(C_0, C_1) = (Id, C)$  and  $\epsilon$  stands for a  $\{0, 1\}$ -valued random variable with distribution

$$\mathbb{P}(\epsilon = 1) = 1 - \mathbb{P}(\epsilon = 0) = \frac{1}{N}$$

The following technical lemma is pivotal.

**Lemma 3.2** *For any non negative function  $F \in \mathcal{B}_b(E^2)$ , any acceptance parameter  $\alpha \in [0, 1]$ , and any  $n \geq 0$ , we have the upper bound*

$$\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \right) \leq \left( \frac{N}{N-1} \right)^{n+1} \mathbb{E} \left( \eta_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \dots Q_n^{\otimes 2} C_{\epsilon_n}(F) \right) \quad (3.3)$$

In the particular case where  $\alpha = 0$ , we have the formula

$$\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \right) = \mathbb{E} \left( \eta_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \dots Q_n^{\otimes 2} C_{\epsilon_n}(F) \right) \quad (3.4)$$

where  $(\epsilon_n)_{n \geq 0}$  stands for a sequence of independent and identically distributed random variables with common law:

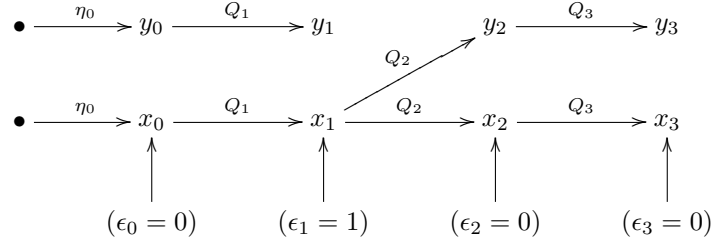
$$\mathbb{P}(\epsilon_1 = 1) = 1 - \mathbb{P}(\epsilon_1 = 0) = \frac{1}{N}$$

Before getting into the details of the proof of the lemma, let us pause for a while and give some comments on the interpretations of these results. Firstly we observe that the functional representation formulae stated in lemma 3.2 are expressed in terms of coalescent operators and Feynman-Kac tensor product semigroups. We emphasize that the second assertion, formula (3.4), is a particular case of the general functional tree based representations presented by the second author with F. Patras and S. Rubenthaler in [8]. This result is only met for the simple genetic model associated with a null acceptance parameter  $\alpha = 0$ . Up to our knowledge there still does not exist any explicit functional representation formula for more general models associated with an acceptance parameter  $\alpha \in ]0, 1]$ .

Next, we turn our attention to a coalescent tree based formulation of the integral expansion stated in the r.h.s. of (3.4). Let us start with an elementary example. Suppose that  $n = 3$  and  $(\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3) = (0, 1, 0, 0)$ . In this situation, we have that

$$\begin{aligned} & \eta_0^{\otimes 2} C_0 Q_1^{\otimes 2} C_1 Q_2^{\otimes 2} C_0 Q_3^{\otimes 2} C_0(F) \\ &= \eta_0^{\otimes 2} Q_1^{\otimes 2} C Q_2^{\otimes 2} Q_3^{\otimes 2}(F) \\ &= \int [\eta_0(dx_0) \eta_0(dy_0)] [Q_1(x_0, dx_1) Q_1(y_0, dy_1)] \\ & \quad \times [Q_2(x_1, dx_2) Q_2(x_1, dy_2)] [Q_3(x_2, dx_3) Q_3(y_2, dy_3)] F(x_3, y_3) \end{aligned}$$

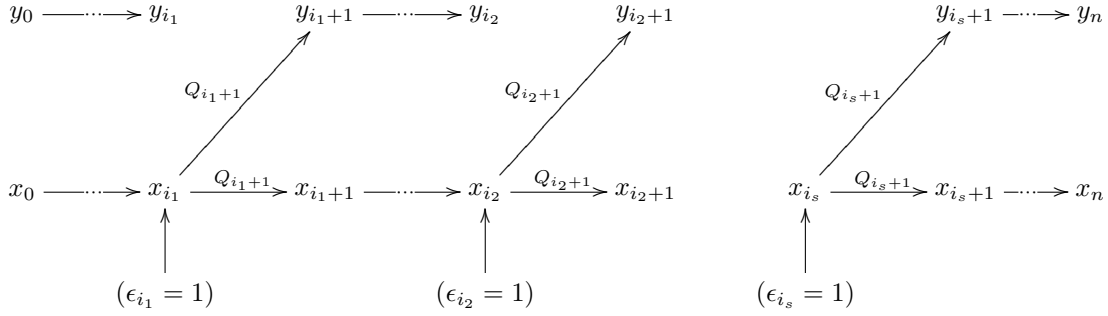
The integration coordinates  $(x_p, y_p)$  from the origin  $p = 0$  up to the third and last level  $p = 3$  can be associated in a canonical way to the following coalescent tree:



More generally, suppose that

$$\forall j \in \{i_1, \dots, i_s\} \quad \epsilon_j = 1 \quad \text{and} \quad \forall k \notin \{i_1, \dots, i_s\} \quad \epsilon_k = 0$$

for some collection of coalescence time indexes  $0 \leq i_1 < \dots < i_s \leq n$ , with  $0 \leq s \leq n$ . The corresponding coalescent tree picture is given below:



This formulation is associated with the following integral expansion:

$$\eta_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \dots Q_n^{\otimes 2} C_{\epsilon_n} = \eta_0^{\otimes 2} Q_{0,i_1}^{\otimes 2} C Q_{i_1,i_2}^{\otimes 2} C \dots Q_{i_{s-1},i_s}^{\otimes 2} C Q_{i_s,n}^{\otimes 2}$$

The coalescent tree based expansions associated with the mean tensor product measures (3.4) will be described in terms of the bounded positive measures defined below.

**Definition 3.3** We associate with any  $0 \leq s \leq (n+1)$  and any coalescence time indexes  $0 \leq i_1 < \dots < i_s \leq n$  the non negative measure  $\Gamma_n^{(i_1, \dots, i_s)} \in \mathcal{M}(E^2)$  defined by the transport equation

$$\Gamma_n^{(i_1, \dots, i_s)} := \gamma_{i_1}^{\otimes 2} C Q_{i_1, i_2}^{\otimes 2} C Q_{i_2, i_3}^{\otimes 2} \dots C Q_{i_{s-1}, i_s}^{\otimes 2} C Q_{i_s, n}^{\otimes 2}$$

We also denote by  $\bar{\Gamma}_n^{(i_1, \dots, i_s)}$  its normalized version given for any  $F \in \mathcal{B}_b(E^2)$  by the following formula:

$$\bar{\Gamma}_n^{(i_1, \dots, i_s)}(F) := \frac{1}{\gamma_n(1)^2} \Gamma_n^{(i_1, \dots, i_s)}(F)$$

We also use the conventions

$$\Gamma_n^{(\emptyset)}(F) = \gamma_n^{\otimes 2}(F) \quad \text{and} \quad \bar{\Gamma}_n^{(\emptyset)}(F) = \eta_n^{\otimes 2}(F) \quad \text{for } s = 0$$



Using lemma 3.2, we readily prove the following decomposition:

$$\begin{aligned} & \mathbb{E} (\eta_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \dots Q_n^{\otimes 2} C_{\epsilon_n} (F)) \\ &= \left(1 - \frac{1}{N}\right)^{(n+1)} \gamma_n^{\otimes 2}(F) \\ &+ \sum_{s=1}^{n+1} \left(1 - \frac{1}{N}\right)^{(n+1)-s} \frac{1}{N^s} \\ &\quad \sum_{0 \leq i_1 < \dots < i_s \leq n} \mathbb{E} (\eta_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \dots Q_n^{\otimes 2} C_{\epsilon_n} (F) \mid \Omega_n(i_1, \dots, i_s)) \end{aligned}$$

with the sets of events

$$\Omega_n(i_1, \dots, i_s) := \{\forall j \in \{i_1, \dots, i_s\} \quad \epsilon_j = 1 \text{ and } \forall k \notin \{i_1, \dots, i_s\} \quad \epsilon_k = 0\}$$

On these sets we have that

$$\eta_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \dots Q_n^{\otimes 2} C_{\epsilon_n} = \eta_0^{\otimes 2} Q_{0,i_1}^{\otimes 2} C Q_{i_1,i_2}^{\otimes 2} C \dots Q_{i_{s-1},i_s}^{\otimes 2} C Q_{i_s,n}^{\otimes 2} = \Gamma_n^{(i_1, \dots, i_s)}$$

Then, for the simple genetic particle model associated with the case where  $\alpha = 0$ , we find the following functional representation formula:

$$\begin{aligned} & \mathbb{E} \left( \frac{(\gamma_n^N)^{\otimes 2}(F)}{\gamma_n(1)^2} \right) - \eta_n^{\otimes 2}(F) \\ &= \sum_{s=1}^{n+1} \left(1 - \frac{1}{N}\right)^{(n+1)-s} \frac{1}{N^s} \sum_{0 \leq i_1 < \dots < i_s \leq n} \left[ \bar{\Gamma}_n^{(i_1, \dots, i_s)}(F) - \eta_n^{\otimes 2}(F) \right] \end{aligned}$$

In particular, if we choose the constant unit function  $F = 1$ , we obtain the first assertion (1.6) of theorem 1.4.

**Proposition 3.4** *For the simple genetic particle model associated with the case where  $\alpha = 0$  we have*

$$\mathbb{E} \left[ \left( \frac{\gamma_n^N(1)}{\gamma_n(1)} - 1 \right)^2 \right] = \sum_{s=1}^{n+1} \left(1 - \frac{1}{N}\right)^{(n+1)-s} \frac{1}{N^s} v_n(s)$$

with the collection of constants  $v_n(s)$  defined below

$$v_n(s) := \sum_{0 \leq i_1 < \dots < i_s \leq n} \left[ \bar{\Gamma}_n^{(i_1, \dots, i_s)}(1) - 1 \right]$$

The above functional representation shows that the order of precision is directly related to the coalescence degree of the trees discussed above. The order  $\frac{1}{N}$  corresponds to a coalescent tree with a single coalescence, the order  $\frac{1}{N^2}$  corresponds to a coalescent tree with a pair coalescence, and so on. The main difficulty in estimating these variances comes from the fact that the total mass of the coalescent type measures  $\bar{\Gamma}_n^{(i_1, \dots, i_s)}$  are generally unknown and we need to resort to a more refined analysis. We shall return to these questions in section 5 dedicated to non asymptotic  $\mathbb{L}_2$ -estimates.

We end this section with the proof of lemma 3.2.

**Proof of lemma 3.2:**

By construction, we have

$$\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \mid X_{n-1}^{(N)} \right) = \gamma_n^N(1)^2 \times \mathbb{E} \left( (\eta_n^N)^{\otimes 2}(F) \mid X_{n-1}^{(N)} \right)$$

By (3.2), we find that

$$\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \mid X_{n-1}^{(N)} \right) = \gamma_n^N(1)^2 \times \mathbb{E} \left( (\eta_n^N)^{\otimes 2}(C_{\epsilon_n} F) \mid X_{n-1}^{(N)} \right)$$

where  $\epsilon_n$  is a  $\{0, 1\}$ -valued random variable with distribution

$$\mathbb{P}(\epsilon_n = 1) = 1 - \mathbb{P}(\epsilon_n = 0) = \frac{1}{N}$$

As noticed in (2.2), we have that

$$\mathbb{E} \left( (\eta_n^N)^{\otimes 2}(C_{\epsilon_n} F) \mid X_{n-1}^{(N)} \right) = \mathbb{E} \left( (\eta_{n-1}^N)^{\otimes 2} \left[ K_{n, \eta_{n-1}^N}^{\otimes 2}(C_{\epsilon_n}(F)) \right] \mid X_{n-1}^{(N)} \right) \quad (3.5)$$

We use lemma 2.2 to check that

$$(\eta_{n-1}^N)^{\otimes 2} = \frac{N}{N-1} \left[ (\eta_{n-1}^N)^{\otimes 2} - \frac{1}{N} (\eta_{n-1}^N)^{\otimes 2} C \right]$$

This implies that for any non negative function  $F$ , we have the upper bound

$$\mathbb{E} \left( (\eta_n^N)^{\otimes 2}(C_{\epsilon_n} F) \mid X_{n-1}^{(N)} \right) \leq \frac{N}{N-1} \mathbb{E} \left( (\eta_{n-1}^N)^{\otimes 2} \left[ K_{n, \eta_{n-1}^N}^{\otimes 2}(C_{\epsilon_n}(F)) \right] \mid X_{n-1}^{(N)} \right)$$

Using the fact that

$$\begin{aligned} \eta_{n-1}^N K_{n, \eta_{n-1}^N} &= \Phi_n(\eta_{n-1}^N) \\ \implies (\eta_{n-1}^N)^{\otimes 2} K_{n, \eta_{n-1}^N}^{\otimes 2} &= \left( \eta_{n-1}^N K_{n, \eta_{n-1}^N} \right)^{\otimes 2} = \Phi_n(\eta_{n-1}^N)^{\otimes 2} \end{aligned}$$

we obtain

$$\mathbb{E} \left( (\eta_n^N)^{\otimes 2}(C_{\epsilon_n} F) \mid X_{n-1}^{(N)} \right) \leq \frac{N}{N-1} \mathbb{E} \left( \Phi_n(\eta_{n-1}^N)^{\otimes 2}(C_{\epsilon_n}(F)) \mid X_{n-1}^{(N)} \right) \quad (3.6)$$

In summary, we have proved that

$$\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \mid X_{n-1}^{(N)} \right) \leq \frac{N}{N-1} \gamma_n^N(1)^2 \times \mathbb{E} \left( \Phi_n(\eta_{n-1}^N)^{\otimes 2}(C_{\epsilon_n}(F)) \mid X_{n-1}^{(N)} \right)$$

To take the final step, we use the fact that

$$\Phi_n(\eta_{n-1}^N)(f) = \frac{\eta_{n-1}^N Q_n(f)}{\eta_{n-1}^N Q_n(1)} \implies \Phi_n(\eta_{n-1}^N)^{\otimes 2}(F) = \frac{(\eta_{n-1}^N)^{\otimes 2} Q_n^{\otimes 2}(F)}{(\eta_{n-1}^N)^{\otimes 2} Q_n^{\otimes 2}(1)}$$

and

$$Q_n(1)(x) = G_{n-1}(x) M_n(1)(x) = G_{n-1}(x) \implies (\eta_{n-1}^N)^{\otimes 2} Q_n^{\otimes 2}(1) = \eta_{n-1}^N(G_{n-1})^2$$

This yields that

$$\begin{aligned} \gamma_n^N(1)^2 \times \frac{1}{(\eta_{n-1}^N)^{\otimes 2} Q_n^{\otimes 2}(1)} &= \left\{ \prod_{0 \leq p < n} \eta_p^N(G_p) \right\}^2 \times \frac{1}{\eta_{n-1}^N(G_{n-1})^2} \\ &= \left\{ \prod_{0 \leq p < (n-1)} \eta_p^N(G_p) \right\}^2 \\ &= \gamma_{n-1}^N(1)^2 \end{aligned}$$

This implies that

$$\begin{aligned} &\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \mid X_{n-1}^{(N)} \right) \\ &\leq \frac{N}{N-1} \gamma_{n-1}^N(1)^2 \times \mathbb{E} \left( (\eta_{n-1}^N)^{\otimes 2} Q_n^{\otimes 2}(C_{\epsilon_n}(F)) \mid X_{n-1}^{(N)} \right) \end{aligned}$$

This readily implies that

$$\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \mid X_{n-1}^{(N)} \right) \leq \frac{N}{N-1} \mathbb{E} \left( (\gamma_{n-1}^N)^{\otimes 2} Q_n^{\otimes 2}(C_{\epsilon_n}(F)) \mid X_{n-1}^{(N)} \right)$$

from which we find that

$$\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \right) \leq \frac{N}{N-1} \mathbb{E} \left( (\gamma_{n-1}^N)^{\otimes 2} Q_n^{\otimes 2}(C_{\epsilon_n}(F)) \right)$$

This ends the proof of (3.3). The proof of the second assertion follows the same lines of arguments. Thus, it is only sketched. Indeed, in particular case where  $\alpha = 0$ , we have that

$$\begin{aligned} K_{n, \eta_{n-1}^N}(x, dy) &= \Phi_n(\eta_{n-1}^N)(dy) \\ \implies K_{n, \eta_{n-1}^N}^{\otimes 2}((x, x'), d(y, y')) &= \Phi_n(\eta_{n-1}^N)^{\otimes 2}(d(y, y')) \end{aligned}$$

In this situation, the formula (3.5) takes the following form

$$\mathbb{E} \left( (\eta_n^N)^{\otimes 2}(C_{\epsilon_n} F) \mid X_{n-1}^{(N)} \right) = \mathbb{E} \left( \Phi_n(\eta_{n-1}^N)^{\otimes 2}(C_{\epsilon_n}(F)) \mid X_{n-1}^{(N)} \right) \quad (3.7)$$

In other words, loosely speaking, the upper bound recursion (3.6) is replaced by the equation (3.7). The remainder of the proof follows exactly the same line of arguments, thus it is omitted. This ends the proof of the lemma.  $\blacksquare$

## 4 Regularity properties of Feynman-Kac semigroups

This section is concerned with some regularity properties of the Feynman-Kac semigroups involved in the coalescent tree based functional expansions presented in section 3. We start with a rather strong condition on the pair  $(G_n, M_n)$ .

*Condition  $(H)_m$ :*

- (G) The potential functions  $G_n$  satisfy the following conditions

$$\forall n \geq 0 \quad \delta_n := \sup_{(x,y) \in E^2} \frac{G_n(x)}{G_n(y)} < \infty$$

- (M)<sub>m</sub> There exists some integer  $m \geq 1$  and some sequence of numbers  $\beta_p^{(m)} \in [1, \infty[$  such that for any  $p \geq 0$  and any  $(x, x') \in E^2$  we have

$$M_{p,p+m}(x, dy) \leq \beta_p^{(m)} M_{p,p+m}(x', dy)$$

with

$$M_{p,p+m} = M_{p+1} M_{p+2} \dots M_{p+m}$$

The main simplification due to this regularity condition is that the total mass mapping  $x \mapsto Q_{p,n}(1)(x)$  of the Feynman-Kac semigroup  $Q_{p,n}$  introduced in (3.1) has uniformly bounded relative oscillations.

**Lemma 4.1** *We suppose that the condition (H)<sub>m</sub> is met for some parameters  $(m, \delta_n, \beta_p^{(m)})$ . Then, we have for any  $p \geq 0$*

$$\sup_{(x,y) \in E^2} \frac{Q_{p,n}(1)(x)}{Q_{p,n}(1)(y)} \leq \delta_p^{(m)} \beta_p^{(m)} \quad \text{with} \quad \delta_p^{(m)} = \prod_{p \leq q < p+m} \delta_q \quad (4.1)$$

**Proof:**

For any nonnegative function  $f \in \mathcal{B}_b(E)$ , any pair of points  $(x, y) \in E^2$  and for any  $p \leq n$  with  $|n - p| \geq m$  we have

$$\frac{Q_{p,n}(f)(x)}{Q_{p,n}(f)(y)} = \frac{G_p(x) M_{p+1} Q_{p+1,n}(f)(x)}{G_p(y) M_{p+1} Q_{p+1,n}(f)(y)} \leq \delta_p \frac{M_{p+1} Q_{p+1,n}(f)(x)}{M_{p+1} Q_{p+1,n}(f)(y)}$$

Using a simple induction, we find that

$$\frac{Q_{p,n}(f)(x)}{Q_{p,n}(f)(y)} \leq \delta_p^{(m)} \frac{M_{p,p+m} [Q_{p+m,n}(f)](x)}{M_{p,p+m} [Q_{p+m,n}(f)](y)}$$

Under condition (M)<sub>m</sub> we conclude that

$$\sup_{(x,y) \in E^2} \frac{Q_{p,n}(f)(x)}{Q_{p,n}(f)(y)} \leq \delta_p^{(m)} \beta_p^{(m)}$$

We further assume that  $|n - p| < m$ . In this case, we readily find that

$$\frac{Q_{p,n}(1)(x)}{Q_{p,n}(1)(y)} \leq \delta_p^{(n-p)} \leq \delta_p^{(m)} \leq \delta_p^{(m)} \beta_p^{(m)}$$

This ends the proof of the lemma. ■

Condition (H)<sub>m</sub> is clearly not met for indicator potential functions. Our next objective is to relax this condition in order to analyze these models. To describe precisely these new conditions, we need to introduce another round of notations.

**Definition 4.2** We denote by  $A_n$  the support of the potential functions  $G_n$ , that is

$$A_n := \{x \in E : G_n(x) > 0\}$$

We let  $(\hat{\gamma}_n, \hat{\eta}_n)$  be the updated Feynman-Kac measures on the set  $A_n$  given by

$$\forall n \geq 0 \quad \hat{\gamma}_n(dx) = \gamma_n(dx) G_n(x)$$

and

$$\hat{\eta}_n(dx) := \frac{1}{\eta_n(G_n)} G_n(x) \eta_n(dx)$$

We let  $(\hat{G}_n, \hat{M}_n)$  be the pair of potential functions and Markov transitions given by :

$$\forall x \in A_n \quad \hat{G}_n(x) := M_{n+1}(G_{n+1})(x)$$

and

$$\forall x \in A_{n-1} \quad \hat{M}_n(x, dy) := \frac{M_n(x, dy) G_n(y)}{M_n(G_n)(x)}$$

Notice that the updated Feynman-Kac measures  $(\hat{\gamma}_n, \hat{\eta}_n)$  can be rewritten in terms of  $(\hat{G}_n, \hat{M}_n)$  with the following change of reference measure formula

$$\hat{\eta}_n(f) := \frac{\hat{\gamma}_n(f)}{\hat{\gamma}_n(1)} \quad \text{with} \quad \hat{\gamma}_n(f) = \eta_0(G_0) \mathbb{E} \left( f(\hat{X}_n) \prod_{0 \leq p < n} \hat{G}_p(\hat{X}_p) \right) \quad (4.2)$$

In the above display,  $\hat{X}_n$  stands for a non homogeneous Markov chain with initial distribution  $\hat{\eta}_0$  and elementary Markov transitions  $\hat{M}_n$  from  $A_{n-1}$  into  $A_n$ . We are now in position to describe these new conditions.

*Condition  $(\hat{H})_m$ :*

- $(\hat{G})$  The potential functions  $\hat{G}_n$  satisfy the following conditions

$$\forall n \geq 0 \quad \hat{\delta}_n := \sup_{(x,y) \in A_n^2} \frac{\hat{G}_n(x)}{\hat{G}_n(y)} < \infty$$

- $(\hat{M})_m$  There exists some integer  $m \geq 1$  and some sequence of numbers  $\hat{\beta}_p^{(m)} \in [1, \infty[$  such that for any  $p \geq 0$  and any  $(x, x') \in A_p^2$  we have

$$\hat{M}_{p,p+m}(x, dy) \leq \hat{\beta}_p^{(m)} \hat{M}_{p,p+m}(x', dy)$$

with

$$\hat{M}_{p,p+m} = \hat{M}_{p+1} \hat{M}_{p+2} \dots \hat{M}_{p+m}$$

Using elementary but rather tedious calculations we readily prove that

$$(H)_m \implies (\hat{H})_m \quad \text{with} \quad \hat{\delta}_n \leq \delta_n \quad \text{and} \quad \hat{\beta}_p^{(m)} \leq \delta_{p+1}^{(m)} \beta_p^{(m)}$$

For  $m = 1$ , we also observe that

$$(M)_1 \implies (\hat{H})_1 \quad \text{with} \quad \hat{\delta}_n \leq \beta_n^{(1)} \quad \text{and} \quad \hat{\beta}_p^{(1)} \leq (\beta_p^{(1)})^2$$

Using the change of measure formula (4.2) we observe that the semigroup of the updated measures  $\widehat{\gamma}_n$  is given by

$$\widehat{Q}_{p,n} = \widehat{Q}_{p+1}\widehat{Q}_{p+2}\dots\widehat{Q}_n \quad \text{with} \quad \widehat{Q}_n(x, dy) = \widehat{G}_{n-1}(x) \widehat{M}_n(x, dy)$$

In other words,  $\widehat{Q}_{p,n}$  is defined as the semigroup  $Q_{p,n}$  by replacing the pair of objects  $(G_n, M_n)$  by the quantities  $(\widehat{G}_n, \widehat{M}_n)$ . From this simple observation, applying lemma 4.1 to the semigroup  $\widehat{Q}_{p,n}$ , without further work we readily find that for any  $p \geq 0$

$$\sup_{(x,y) \in A_p^2} \frac{\widehat{Q}_{p,n}(1)(x)}{\widehat{Q}_{p,n}(1)(y)} \leq \widehat{\delta}_p^{(m)} \widehat{\beta}_p^{(m)} \quad \text{with} \quad \widehat{\delta}_p^{(m)} = \prod_{p \leq q < p+m} \widehat{\delta}_q \quad (4.3)$$

as soon as the regularity condition  $(\widehat{H})_m$  is met for some parameters  $(m, \widehat{\delta}_n, \widehat{\beta}_p^{(m)})$ .

Using the easily checked formula

$$Q_{p,n}(f)(x) = G_p(x) \left[ \widehat{Q}_{p,n-1} M_n \right] (f)(x) \quad (4.4)$$

we readily prove the following lemma.

**Lemma 4.3** *We suppose that the condition  $(\widehat{H})_m$  is met for some parameters  $(m, \widehat{\delta}_n, \widehat{\beta}_p^{(m)})$ . In addition, we assume that the potential functions  $G_n$  satisfy the following conditions*

$$\forall n \geq 0 \quad \widetilde{\delta}_n := \sup_{(x,y) \in A_n^2} \frac{G_n(x)}{G_n(y)} < +\infty$$

Then, for any  $p \geq 0$ , we have the estimates

$$\sup_{(x,y) \in A_p^2} \frac{Q_{p,n}(1)(x)}{Q_{p,n}(1)(y)} \leq \widetilde{\delta}_p \widehat{\delta}_p^{(m)} \widehat{\beta}_p^{(m)} \quad (4.5)$$

## 5 Non asymptotic $\mathbb{L}_2$ -estimates

This section is concerned with the statement and the proof of the main results of this article.

**Theorem 5.1** *We suppose that the condition  $(\widehat{H})_m$  is met for some parameters  $(m, \widehat{\delta}_n, \widehat{\beta}_p^{(m)})$ . In addition, we assume that the potential functions  $G_n$  satisfy the following conditions*

$$\forall n \geq 0 \quad \widetilde{\delta}_n := \sup_{(x,y) \in A_n^2} \frac{G_n(x)}{G_n(y)} < +\infty \quad (5.1)$$

Then, for any non negative function  $F \in \mathcal{B}_b(E^2)$  with  $\|F\| \leq 1$ , and any  $N > 1$  we have

$$\mathbb{E} \left( (\gamma_n^N)^{\otimes 2}(F) \right) \leq \gamma_n(1)^2 \times \prod_{s=0}^n \left( 1 + \frac{1}{N-1} \frac{\widetilde{\delta}_s \widehat{\delta}_s^{(m)} \widehat{\beta}_s^{(m)}}{\eta_s(A_s)} \right) \quad (5.2)$$

Furthermore, if condition  $(H)_m$  is met for some  $(m, \delta_n, \beta_p^{(m)})$  then we have the estimate

$$\mathbb{E}((\gamma_n^N)^{\otimes 2}(F)) \leq \gamma_n(1)^2 \times \prod_{s=0}^n \left(1 + \frac{1}{N-1} \delta_s^{(m)} \beta_s^{(m)}\right) \quad (5.3)$$

Before getting into the proof of this theorem, we already present a simple consequence of the above estimates. Notice that theorem 1.4 stated in the introduction is a direct consequence of the following corollary.

**Corollary 5.2** *When conditions (5.1) and  $(\widehat{H})_m$  are met for some  $(m, \widehat{\delta}_n, \widehat{\beta}_p^{(m)})$ , we have the non asymptotic estimates*

$$N > \sum_{s=0}^n \frac{\widetilde{\delta}_s \widehat{\delta}_s^{(m)} \widehat{\beta}_s^{(m)}}{\eta_s(A_s)} \implies \mathbb{E} \left( \left[ \frac{\gamma_n^N(1)}{\gamma_n(1)} - 1 \right]^2 \right) \leq \frac{4}{N} \sum_{s=0}^n \frac{\widetilde{\delta}_s \widehat{\delta}_s^{(m)} \widehat{\beta}_s^{(m)}}{\eta_s(A_s)}$$

In addition, if condition  $(H)_m$  is met for some  $(m, \delta_n, \beta_p^{(m)})$ , then we have

$$N > \sum_{s=0}^n \left[ \delta_s^{(m)} \beta_s^{(m)} \right] \implies \mathbb{E} \left( \left[ \frac{\gamma_n^N(1)}{\gamma_n(1)} - 1 \right]^2 \right) \leq \frac{4}{N} \sum_{s=0}^n \left[ \delta_s^{(m)} \beta_s^{(m)} \right]$$

The proof of the corollary is elementary, thus we give it first.

**Proof of Corollary 5.2:**

Using the fact that  $\log(1+x) \leq x$  for any  $x \geq 0$ , and  $e^x \leq 1+2x$  for any  $x \in [0, 1]$ , we conclude that

$$\prod_{s=0}^n \left(1 + \frac{a_s}{N-1}\right) - 1 = e^{\sum_{s=0}^n \log\left(1 + \frac{a_s}{N-1}\right)} - 1 \leq e^{\frac{1}{N-1} \sum_{s=0}^n a_s} - 1 \leq \frac{2}{N-1} \sum_{s=0}^n a_s$$

for every  $(a_s)_{s \geq 0} \in [1, \infty]^{\mathbb{N}}$  and any  $N \geq 1 + \sum_{s=0}^n a_s$ . Also observe that  $\frac{1}{N-1} \leq \frac{2}{N}$  for any  $N \geq 1 + \sum_{s=0}^n a_s (\geq 2)$ . This yields that

$$\prod_{s=0}^n \left(1 + \frac{a_s}{N-1}\right) - 1 \leq \frac{4}{N} \sum_{s=0}^n a_s$$

Using these estimates the proof of the corollary is a direct consequence of theorem 5.1. This ends the proof of the corollary.  $\blacksquare$

Now, we come to the proof of theorem 5.1.

**Proof of theorem 5.1:** To simplify the presentation, firstly we suppose that  $(H)_m$  is met for some parameters  $(m, \delta_n, \beta_p^{(m)})$ . We observe that

$$\begin{aligned} \Gamma_n^{(i_1, \dots, i_s)}(F) &= \gamma_{i_1}(1) \int \gamma_{i_1}(dx) \left[ Q_{i_1, i_2}^{\otimes 2} C \dots Q_{i_{s-1}, i_s}^{\otimes 2} C Q_{i_s, n}^{\otimes 2}(F) \right](x, x) \\ &= \Gamma_{i_s}^{(i_1, \dots, i_{s-1})}(C Q_{i_s, n}^{\otimes 2}(F)) \end{aligned}$$

On the other hand, we have

$$\frac{\gamma_{i_1}(1)}{\gamma_n(1)} = \frac{\gamma_{i_1}(1)}{\gamma_{i_1} Q_{i_1, n}(1)} = \frac{1}{\eta_{i_1} Q_{i_1, n}(1)}$$

and

$$\frac{\gamma_{i_1}(dx)}{\gamma_n(1)} = \frac{\gamma_{i_1}(1)}{\gamma_n(1)} \times \eta_{i_1}(dx) = \frac{\eta_{i_1}(dx)}{\eta_{i_1}Q_{i_1,n}(1)}$$

Now, we prove (5.3) using an induction on the parameter  $s$ . For  $s = 1$  we observe that

$$\begin{aligned} \bar{\Gamma}_n^{(i_1)}(F) &= \frac{\gamma_{i_1}(1)}{\gamma_n(1)} \int \frac{\gamma_{i_1}(dx)}{\gamma_n(1)} Q_{i_1,n}^{\otimes 2}(F)(x, x) \\ &= \int \frac{\eta_{i_1}(dx)Q_{i_1,n}(x, dy)}{\eta_{i_1}Q_{i_1,n}(1)} \times \frac{Q_{i_1,n}(x, dy')}{\eta_{i_1}Q_{i_1,n}(1)} F(y, y') \end{aligned}$$

Since  $\|F\| \leq 1$  this yields that

$$\bar{\Gamma}_n^{(i_1)}(F) \leq \int \frac{\eta_{i_1}(dx)Q_{i_1,n}(x, dy)}{\eta_{i_1}Q_{i_1,n}(1)} \times \frac{Q_{i_1,n}(1)(x)}{\eta_{i_1}Q_{i_1,n}(1)}$$

Using the estimate (4.1) we find that

$$\bar{\Gamma}_n^{(i_1)}(F) \leq \delta_{i_1}^{(m)} \beta_{i_1}^{(m)}$$

Next, we suppose that the desired upper bound is valid at rank  $(s - 1)$ , that is we have that for any  $\|F\| \leq 1$

$$\bar{\Gamma}_n^{(i_1, \dots, i_{s-1})}(F) \leq \prod_{k \in \{i_1, \dots, i_{s-1}\}} \left( \delta_k^{(m)} \beta_k^{(m)} \right)$$

To check that the result is also true at rank  $s$ , we use the decompositions

$$\begin{aligned} &\bar{\Gamma}_n^{(i_1, \dots, i_s)}(F) \\ &= \frac{1}{(\eta_{i_1}Q_{i_1,n}(1))^2} \int \eta_{i_1}(dx) \left[ \left\{ Q_{i_1, i_2}^{\otimes 2} CQ_{i_2, i_3}^{\otimes 2} \dots CQ_{i_{s-1}, i_s}^{\otimes 2} \right\} CQ_{i_s, n}^{\otimes 2} \right] (F)(x, x) \\ &= \frac{1}{(\eta_{i_1}Q_{i_1,n}(1))^2} \\ &\quad \int \eta_{i_1}(dx) \left[ Q_{i_1, i_2}^{\otimes 2} CQ_{i_2, i_3}^{\otimes 2} \dots CQ_{i_{s-1}, i_s}^{\otimes 2} \right] ((x, x), d(u, v)) Q_{i_s, n}^{\otimes 2}(F)(u, u) \end{aligned}$$

We observe that

$$\begin{aligned} Q_{i_s, n}^{\otimes 2}(1 \otimes 1)(u, u) &= Q_{i_s, n}(1)(u)^2 \leq \left( \delta_{i_s}^{(m)} \beta_{i_s}^{(m)} \right) Q_{i_s, n}(1)(u) Q_{i_s, n}(1)(u') \\ &= \left( \delta_{i_s}^{(m)} \beta_{i_s}^{(m)} \right) Q_{i_s, n}^{\otimes 2}(1 \otimes 1)(u, u') \end{aligned}$$



from which we conclude that

$$\begin{aligned}
& \bar{\Gamma}_n^{(i_1, \dots, i_s)}(F) \\
& \leq \frac{\left(\delta_{i_s}^{(m)} \beta_{i_s}^{(m)}\right)}{\left(\eta_{i_1} Q_{i_1, n}(1)\right)^2} \\
& \quad \int \eta_{i_1}(dx) \left[ Q_{i_1, i_2}^{\otimes 2} C Q_{i_2, i_3}^{\otimes 2} \dots C Q_{i_{s-1}, i_s}^{\otimes 2} \right] ((x, x), d(u, u')) Q_{i_s, n}^{\otimes 2}(1 \otimes 1)(u, u') \\
& = \left(\delta_{i_s}^{(m)} \beta_{i_s}^{(m)}\right) \frac{1}{\left(\eta_{i_1} Q_{i_1, n}(1)\right)^2} \\
& \quad \int \eta_{i_1}(dx) \left[ Q_{i_1, i_2}^{\otimes 2} C Q_{i_2, i_3}^{\otimes 2} \dots C Q_{i_{s-1}, n}^{\otimes 2} \right] (1 \otimes 1)(x, x)
\end{aligned}$$

and therefore

$$\bar{\Gamma}_n^{(i_1, \dots, i_s)}(F) \leq \left(\delta_{i_s}^{(m)} \beta_{i_s}^{(m)}\right) \bar{\Gamma}_n^{(i_1, \dots, i_{s-1})}(1 \otimes 1)$$

Using the induction hypothesis, we conclude that

$$\bar{\Gamma}_n^{(i_1, \dots, i_s)}(F) \leq \left\{ \prod_{k \in \{i_1, \dots, i_{s-1}\}} \left(\delta_k^{(m)} \beta_k^{(m)}\right) \right\} \times \left(\delta_{i_s}^{(m)} \beta_{i_s}^{(m)}\right)$$

This ends the proof of the inductive proof. Using the above estimates, we find that

$$\begin{aligned}
& \frac{1}{\gamma_n(1)^2} \mathbb{E}(\eta_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \dots Q_n^{\otimes 2} C_{\epsilon_n}(F)) \\
& \leq \left(1 - \frac{1}{N}\right)^{(n+1)} \\
& \quad + \sum_{s=1}^{n+1} \left(1 - \frac{1}{N}\right)^{(n+1)-s} \frac{1}{N^s} \sum_{0 \leq i_1 < \dots < i_s \leq n} \prod_{k \in \{i_1, \dots, i_s\}} \left(\delta_k^{(m)} \beta_k^{(m)}\right)
\end{aligned}$$

Using the decomposition

$$\prod_{s=0}^n (1 + a_s) = 1 + \sum_{1 \leq s \leq n+1} \sum_{0 \leq i_1 < \dots < i_s \leq n} \left( \prod_{j \in \{i_1, \dots, i_s\}} a_j \right)$$

which is valid for any  $n \geq 0$  and any collection of numbers  $(a_p)_{p \geq 0}$ , we prove that

$$\begin{aligned}
& \frac{1}{\gamma_n(1)^2} \mathbb{E}(\eta_0^{\otimes 2} C_{\epsilon_0} Q_1^{\otimes 2} C_{\epsilon_1} \dots Q_n^{\otimes 2} C_{\epsilon_n}(F)) \\
& \leq \left(1 - \frac{1}{N}\right)^{(n+1)} \left(1 + \sum_{s=1}^{n+1} \sum_{0 \leq i_1 < \dots < i_s \leq n} \prod_{k \in \{i_1, \dots, i_s\}} \frac{\delta_k^{(m)} \beta_k^{(m)}}{N-1}\right) \\
& \leq \left(1 - \frac{1}{N}\right)^{(n+1)} \prod_{s=0}^n \left(1 + \frac{\delta_s^{(m)} \beta_s^{(m)}}{N-1}\right)
\end{aligned}$$

The end of the inductive proof of (5.3) is now easily completed. Indeed, using (3.3) we conclude that for any  $\alpha \in [0, 1]$

$$\mathbb{E} \left( \frac{(\gamma_n^N)^{\otimes 2}(F)}{\gamma_n(1)^2} \right) \leq \prod_{s=0}^n \left( 1 + \frac{\delta_s^{(m)} \beta_s^{(m)}}{N-1} \right)$$

The proof of (5.2) follows the same lines of arguments, thus it is only sketched. We suppose condition  $(\widehat{H})_m$  is met for some parameters  $(m, \widehat{\delta}_n, \widehat{\beta}_p^{(m)})$ . In this situation, for every  $1 \leq i \leq n$  we have

$$\overline{\Gamma}_n^{(i)}(F) \leq \|F\| \int \frac{\eta_i(dx) Q_{i,n}(x, dy)}{\eta_i Q_{i,n}(1)} \times \frac{Q_{i,n}(1)(x)}{\eta_i Q_{i,n}(1)}$$

Using lemma 4.3 and recalling that

$$Q_{i,n}(1)(x) = G_i(x) M_{i+1}(Q_{i+1,n}(1))(x) = 1_{A_i}(x) Q_{i,n}(1)(x)$$

we find that

$$\begin{aligned} \frac{Q_{i,n}(1)(x)}{\eta_i Q_{i,n}(1)} &= \frac{Q_{i,n}(1)(x)}{\int \eta_i(dy) Q_{i,n}(1)(y)} \\ &= 1_{A_i}(x) \frac{Q_{i,n}(1)(x)}{\int \eta_i(dy) 1_{A_i}(y) Q_{i,n}(1)(y)} \leq \frac{\widetilde{\delta}_i \times \widehat{\delta}_i^{(m)} \widehat{\beta}_i^{(m)}}{\eta_i(A_i)} \end{aligned}$$

The inductive proof of (5.2) now follows exactly the one of (5.3) and it is omitted. This ends the proof of the theorem.  $\blacksquare$

## 6 Application to rare events

In this section, we want to outline the use of our main result in terms of efficiency for rare event probability estimation. Basically, it is about asymptotics when the rare event probability goes to 0: we want to control the relative variance of our estimator when the event of interest is getting more and more unlikely. In the context of *importance sampling*, a discussion about various efficiency (or robustness) properties may be found in [1]. Among all those, we will focus here on logarithmic efficiency.

Returning to the framework presented in the end of section 1.2, we further assume that we have a family of rare sets  $A^\varepsilon$  indexed by  $\varepsilon \geq 0$ , of the form

$$A^\varepsilon = \{z \in S \text{ s.t. } \phi(z) > \lambda_\varepsilon\},$$

for some real valued function  $\phi$ . Denote as usual

$$T_{A^\varepsilon} = \inf\{t \geq 0, Z_t \in A^\varepsilon\} \quad \text{and} \quad T_R = \inf\{t \geq 0, Z_t \in R\}.$$

for some recurrent Borel subset  $R \subset S$ . Assume further that we have for some  $\theta > 0$

$$\mathbb{P}(T_{A^\varepsilon} < T_R) = e^{-\theta/\varepsilon},$$

which is typical of behavior driven by a large deviation principle. We further assume that we are given a non increasing sequence of level sets

$$A^\varepsilon = A_{n_\varepsilon} \subset A_{n_\varepsilon-1} \subset \cdots \subset A_2 \subset A_1 \subset A_0$$

with a real valued function  $\psi$  so that

$$A_n = \{x \in S, \quad \psi(x) > L_n\}.$$

In the above displayed formula  $(L_n)_{1 \leq n \leq n_\varepsilon}$  stands for a non decreasing sequence of real numbers, with some fixed time horizon  $n_\varepsilon$  that may depend on the parameter  $\varepsilon$ , and so that  $A_{n_\varepsilon} = A^\varepsilon$ . In the rare event literature, such a function  $\psi$  is called an *importance function*. In this notation, by (1.9) the rare event probability of interest is given by

$$\gamma_{n_\varepsilon}(1) = \mathbb{P}(T_{n_\varepsilon} \leq T_R) = e^{-\theta/\varepsilon}$$

Then we say that our estimator  $\gamma_{n_\varepsilon}^N(1)$  has the logarithmic efficiency property if we have

$$\lim_{\varepsilon \rightarrow 0} \frac{\log \mathbb{E} \left[ (\gamma_{n_\varepsilon}^N(1))^2 \right]}{2 \log \gamma_{n_\varepsilon}(1)} = 1.$$

Next, we discuss the regularity conditions  $(\widehat{G})$  and  $(\widehat{M})_m$  introduced on page 18:

Before to proceed, we observe that the parameters  $\widetilde{\delta}_n$  introduced in (4.3) are simply given by  $\widetilde{\delta}_n = 1$ . We check this claim using the fact that the potential functions  $G_n$  are the indicator functions on excursion subsets ending at the level sets  $A_n$ .

The assumption  $(\widehat{M})_m$  is clearly a mixing type property. In this context  $\widehat{M}_n(x_{n-1}, dx_n)$  is the elementary transition probability of an excursion  $\widehat{X}_n$  starting at  $A_{n-1}$  (at the terminal state of an excursion  $x_{n-1}$  ending at  $A_{n-1}$ ) and ending at the next level set  $A_n$ . We illustrate condition  $(\widehat{M})_m$  for the simple random walk on the one dimensional lattice  $S = \mathbb{Z}$  starting at the origin, with the increasing sequence of level sets  $A_n = [n, \infty[$ . In this context, we readily find that  $(\widehat{M})_m$  is satisfied with  $m = 1$  and  $\widehat{\beta}_n^{(1)} = 1$ , for every  $n \geq 1$ . More generally, in the simple setting of one dimension (i.e. the random process  $Z$  lives in  $\mathbb{R}$ ), we always have  $\widehat{\beta}_n^{(m)} = 1$  for all  $n$ .

Now we discuss to the regularity condition  $(\widehat{G})$ . Firstly, we observe that

$$\widehat{G}_n(x_n) = M_{n+1}(G_{n+1})(x_n)$$

is the probability of reaching the set  $A_{n+1}$ , starting from the terminal value of a random excursion  $x_n$  ending at  $A_n$ . The less this quantity depends on  $x_n$ , the lower is the variance, as it is already well known for the asymptotic variance (as seen in [2]). So a good choice of the sets  $A_n$  is such that they are close to level sets for the probability of reaching the rare event.

From corollary 5.2, we see that  $\eta_n(A_n)$  is another quantity of interest. In this situation, we recall that  $A_n$  is the set of all random excursions ending at the level  $A_n$  and  $\eta_n$  is the distribution of the  $n$ -th excursion  $X_n$  of the process  $Z_t$  given the fact that it has reached the level  $A_{n-1}$  at time  $T_{n-1}$ . Thus,  $\eta_n(A_n)$

is the probability of reaching level  $A_n$ , knowing that the trajectory has reached  $A_{n-1}$ . It is well known already (see [2, 12]) that we need to have these quantities  $\eta_n(A_n)$  as close to each other as possible (the best would be equal). So not only do we need to have an importance function close to the optimal one, but also to have the sets  $A_n$  evenly spaced in terms of hitting probabilities.

The issue of constructing a good importance function is far from trivial, and has been nicely addressed in [4] in the case of *importance splitting* techniques (which are close to the IPS approach). Their choice of importance function allows them to prove the asymptotically optimal efficiency of the importance splitting with their choice of the importance function.

From now on, we assume that we know how to construct a good importance function, in such a way that for all  $n$ ,  $\widehat{\delta}_n^{(m)} < \delta$  for some  $\delta$ , and we know how to construct the level sets  $A_n$  so that

$$\mathbb{P}(T_n < T_R \mid T_{n-1} < T_R) = \eta_n(A_n) \approx p > 0$$

for some  $p \in [0, 1]$ . A practical way for doing this has been proposed by two of the authors in [3]. We also suppose that the Markov process  $\widehat{X}_n$  is sufficiently mixing, so that  $\widehat{\beta}_n^{(m)} < \beta$ , for some  $\beta$ . In this situation, using the fact that  $\eta_n(A_n) \approx p > 0$ , we get that the number  $n_\varepsilon$  of steps needed to get to the rare event is of order  $-\frac{\theta}{\varepsilon \log p}$ . Using our theorem 5.1, we see that

$$\mathbb{E}[(\gamma_{n_\varepsilon}^N(1))^2] \leq \gamma_{n_\varepsilon}(1)^2 \left( 1 + \frac{\delta\beta}{(N-1)p} \right)^{-\frac{\theta}{\varepsilon \log p}},$$

Using Jensen's inequality and the fact that the estimator  $\gamma_{n_\varepsilon}^N(1)$  is unbiased, we conclude that

$$1 + \frac{1}{2 \log p} \log \left( 1 + \frac{1}{N-1} \frac{\delta\beta}{p} \right) \leq \frac{\log \mathbb{E}[(\gamma_{n_\varepsilon}^N(1))^2]}{2 \log \gamma_{n_\varepsilon}(1)} \leq 1,$$

which means that we get the asymptotic logarithmic efficiency at any (slow) rate, in the sense that

$$\lim_{N \uparrow \infty} \frac{\log \mathbb{E}[(\gamma_{n_\varepsilon}^N(1))^2]}{2 \log \gamma_{n_\varepsilon}(1)} = 1$$

## References

- [1] J. H. BLANCHET, P. W. GLYNN, P. L'ECUYER, W. SANDMANN, B. TUFFIN (2007) Asymptotic Robustness of Estimators in Rare-Event Simulation. Proc. of the 2007 INFORMS workshop.
- [2] F. CEROU, P. DEL MORAL, F. LE GLAND, P. LEZAUD (2006). Genealogical Models in Entrance Times Rare Event Analysis. Alea, Latin American Journal of Probability And Mathematical Statistics.
- [3] F. CEROU AND A. GUYADER (2007). Adaptive Multilevel Splitting for Rare Event Analysis, Stochastic Analysis and Applications, 25, 417-433.
- [4] T. DEAN AND P. DUPUIS (2008). Splitting for rare event simulation: A large deviations approach to design and analysis, to appear in Stochastic Processes and their Applications.

- 
- [5] P. DEL MORAL (2004). Feynman-Kac formulae. Genealogical and interacting particle systems. Springer-Verlag: New York.
  - [6] P. DEL MORAL, A. DOUCET, A. JASRA (2006). Sequential Monte Carlo samplers. *J. R. Statist. Soc. B*, **68**, 411–436.
  - [7] P. DEL MORAL, A. DOUCET, G. W. PETERS (2006). Sharp Propagations of Chaos Estimates for Feynman-Kac particle Models *Probability Theory and its Applications SIAM*, vol. **51**, no. **3**.
  - [8] P. DEL MORAL, F. PATRAS, S. RUBENTHALER (2006). Coalescent Tree Based Functional Representations for some Feynman-Kac Particle Models, to appear in the Annals of Applied Probability (2009), *also available at <https://hal.ccsd.cnrs.fr/ccsd-00086532>* .
  - [9] A. DOUCET, N. DE FREITAS, N. GORDON, EDITORS (2001). *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science. Springer New York.
  - [10] C. JARZYNSKI (1997). Equilibrium free energy differences from nonequilibrium measurements: a master equation approach, *Phys. Rev. E* 56(5), pp. 5018-5035.
  - [11] A. M. JOHANSEN, P. DEL MORAL, A. DOUCET (2006). Sequential Monte Carlo Samplers for Rare Events. *Proceedings of 6th International Workshop on Rare Event Simulation*, Bamberg, Germany.
  - [12] A. LAGNOUX (2006). Rare event simulation. PEIS 20 (1), pp. 45-66.
  - [13] M. ROUSSET & G. STOLTZ (2006). Equilibrium sampling from non-equilibrium dynamics. *J. Stat. Phys.*, **123**, 1251-1272.



---

Centre de recherche INRIA Rennes – Bretagne Atlantique  
IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex  
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier  
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq  
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique  
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex  
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex  
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex  
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

---

Éditeur  
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)  
<http://www.inria.fr>  
ISSN 0249-6399