Branching and Interacting Particle Interpretations of Rare Event Probabilities

Pierre Del Moral[1] and Pascal Lezaud[2]

[1] delmoral@math.unice.fr. Université de Nice Sophia Antipolis-06108 Nice Cedex 02, France
[2] pascal.lezaud@recherche.enac.fr. Centre d'Etudes de la Navigation Aérienne-31055 Toulouse Cedex, France

Summary. This article focuses on branching particle interpretations of rare events. We connect importance sampling techniques with interacting particle algorithms, and multi-splitting branching models. These Monte Carlo methods are illustrated with a variety of examples arising in particle trapping analysis, as well as in ruin type estimation problems. We also provide a rather detailed presentation of the asymptotic theory of these particle algorithms, including exponential extinction probabilities, $\mathbb{L}_p$-mean error bounds, central limit theorem, and fluctuation variance comparaisons.

1 Introduction

The study of rare events is an important and very active area in a variety of scientific disciplines. In particle physics, rare event problems are often related to the estimation of non absorption probabilities of a particle evolving in a trapping medium. These quantities are also connected to the estimation of the Lyapunov exponent of Schrödinger type operators. In engineering sciences, these rare event problems arise in the analysis and prediction of major risks, as such earthquakes, floods, air collision risks, nuclear radiation dispersions. Studying major risks can undertaken utilising two main approaches, the statistical analysis of collected data and the probabilistic modelling of the processes leading to the incident. The statistical analysis of extreme values often needs an extended observation period, due to the very low occurrence probability of rare events. They are often based on the standard extreme value distributions, like the Gumbel, the Fréchet and the Weibull laws (see for instance [11, 7], and references therein). The probabilistic approach firstly consists in modelling the randomness of the underlying system, and secondly in using some mathematical, or simulation tools, to obtain an accurate estimate.

The use of analytical, and numerical approaches are often based on simplified, and ad hoc assumptions. On the other hand, the Monte Carlo simulation

is a practical alternative when the analysis calls for fewer simplifying assumptions. Nevertheless, obtaining accurate estimates of rare event probabilities, say about 10^{-9} to 10^{-12}, using traditional techniques really requires a huge amount of computing time.

Many techniques for reducing the number of trials in Monte Carlo simulation have been proposed, the more promising are based on importance sampling. Importance sampling consists in modifying the underlying probability distribution in such a way that rare events occur much more frequently. To use importance sampling, we need to have a deep knowledge of the system under study; and even in such a case, importance sampling may not reduce the number of trials. In addition, for large time scale problems, the importance weightings are also degenerate with respect to the time parameter. This degeneration reduces considerably the performance and the accuracy of the Monte Carlo approximation.

An alternative way to increase the relative number of visits to the rare event, is to use interactive evolution models and trajectory splitting techniques. These two approaches are based on the fact that there exist some physical potential functions reflecting the rare event regime, or there exist some intermediate levels that are visited much more often than the rare level. These intermediate events act as gateways to reach the desired rare level set. Particle methodologies were first introduced as heuristic algorithms in the beginning of the 1950's, in biology by M.N. Rosenbluth and A. W. Rosenbluth [12] for macromolecular simulations, as well as in physics with the article of T.E. Harris and H. Kahn [8] for particle transmission simulations. Since this period, the range of applications of these interactive particle ideas have increased, revealing unexpected connections between a variety of domains, including signal processing, financial mathematics, particle physics, biology and engineering sciences. The application in rare event simulation was firstly introduced in [1] and shortly after has been adapted to the hybrid systems [9]. For a detailed description of these applications models, and a precise mathematical analysis of these particle methods, the reader is referred to the research monograph [2], and references therein.

In the present review article, we focus on branching particle interpretations of rare events. A short description of the paper is as follows. Section 2 sets out a brief description of different types of branching particle methodologies. We connect importance sampling techniques with interacting particle algorithms, and multi-splitting branching models. We illustrate these Monte Carlo methods with a variety of examples arising in particle trapping analysis, as well as in ruin type estimation problems. Finally, we end this section with a description of a more refined interacting particle analysis of rare event probabilities based on multi-level decompositions of the state-space regions. The capture of the behavior of the Markov chain between each level needs the introduction of the random excursions models. This is provided by the Section 3, in which Markov chain models in abstract path are briefly intro-

duced. The law of the excursions in the rare regime are described by functional representations which belong to the class of Feynman-Kac formulas.

Section 4 describes the Feynman-Kac models which are at the corner of diverses disciplines. The Feynman-Kac models, in general nonhomogeneous state spaces, are built with two ingredients: A Markov chain associated with a reference probability measure, and a sequence of potential functions. From the pure mathematical point of view, they correspond to a change of probability on path space associated with a given sequence of potential functions. From the point of view of physics, they represent for instance the path distribution of a single particle evolving in absorbing and disordered media (see for instance [13]). In this interpretation, the potential function represents a "killing or creation" rate related to the absorbing nature of the medium. In engineering sciences the use of Feynman-Kac models is of course not restricted to rare event modelling and analysis. They are commonly used in non linear filtering to represent the conditional distribution of a given random signal with respect to a sequence of noisy observations delivered by some sensors. Different physical interpretations of the Feynman-Kac models are also provided in this section. These models have natural particle interpretations in terms of genealogical tree-based evolutions. The Section 5 is devoted to particle interpretations of Feynman-Kac models. These particle models can be sought in many different ways depending on the application we have in mind. For the analysis of rare events, we have chosen to describe these models as an abstract stochastic linearization technique for solving nonlinear and measure-valued equations. The basic idea is to associate to a given nonlinear dynamical structure, a sequence of N Markov processes, in such a way that the N-empirical measures of the configurations converge, as $N \to \infty$, to the desired distribution. The parameter N represents the precision parameter, as well as the size of the systems. In some sense, these particle models can be regarded as a new approximation simulation technique. All these particle models are built on the same paradigm: When exploring a state space with many particles, we duplicate better fitted individuals at the expense of having light particles with poor fitness die.

Finally, Section 6 is concerned with the asymptotic behavior of the particle methods when the size of the systems tends to infinity. We provide a rather detailed presentation of the asymptotic theory, including exponential extinction probabilities, $\mathbb{L}_p$-mean error bounds, central limit theorem, and fluctuation variance comparaisons between these particles algorithms.

2 Branching Particle Methodologies

This section sets out a brief description of four different types of branching particle methodologies and an interacting particle systems algorithm for estimating rare events. The first concerns an original genetic type interpretation of importance sampling representations of rare event probabilities. The second

one is concerned with a class of rare event problems arising in physics, and more particularly in nuclear engineering. We design an interacting particle interpretation for the evolution of a Markov chain in an absorbing medium. Special attention is paid to the study of the genealogical structure of these interacting jump particle models. We also connect these trapping problems with the estimation of the Lyapunov exponent of a Schrödinger type semigroup. The third and fourth are devoted respectively to an elementary Bernoulli branching method, and to a more sophisticated branching splitting variant.

2.1 Importance Sampling Branching Models

Let X_n be a Markov chain with transition probabilities $M_n(x_{n-1}, dx_n) = \mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1})$. Suppose we want to estimate the probability $\mathbb{P}(A)$, that X_n reaches some rare region A of the state space. To fix the ideas, we can think of a simple random walk starting at $X_0 = 0$ and evolving to the right with a small probability

$$\mathbb{P}(X_n = X_{n-1} + 1) = p = 1 - \mathbb{P}(X_n = X_{n-1} - 1) < 1/2 \,. \tag{1}$$

In this situation for large values of M, the probabilities $\mathbb{P}(X_n \geq M)$ are extremely small. As we mentioned in the introduction, the importance sampling methodology consists in changing the whole distribution of the chain X_n so that to deal with a new random process X'_n which is "attracted" by the rare event. If we let $M'_n(x_{n-1}, dx_n)$ be the Markov transitions of X'_n, and assuming that M_n and M'_n are mutually absolutely continuous, then we have the rare event probability representation.

$$\mathbb{P}(X_n \in A) = \mathbb{E}\left(1_A(X'_n) \prod_{k=1}^{n} G_k(X'_{k-1}, X'_k)\right) \,, \tag{2}$$

with

$$G_k(X'_{k-1}, X'_k) = \frac{\mathrm{d}M_k(X'_{k-1}, \cdot)}{\mathrm{d}M_k(X'_{k-1}, \cdot)}(X'_k) \,,$$

and for any bounded function f

$$\mathbb{E}(f(X_n)|X_n \in A) = \frac{\mathbb{E}\left(f(X'_n) 1_A(X'_n) \prod_{k=1}^{n} G_k(X'_{k-1}, X'_k)\right)}{\mathbb{E}\left(1_A(X'_n) \prod_{k=1}^{n} G_k(X'_{k-1}, X'_k)\right)} \,.$$

In the simple random walk example, we can exchange the role of p and $q = 1-p$. In this case, X'_n tends to move to the right, and the change of probability formula (2) holds true with $G_k(x, x+1) = p/q < 1$ and $G_k(x, x-1) = q/p > 1$.

The importance sampling method consists in evolving N independent copies X'^i of X', and taking the weighted Monte Carlo estimates

$$\frac{1}{N}\sum_{i=1}^{N} 1_A(X_n^{'i})\left[\prod_{k=1}^{n} G_k(X_{k-1}^{'i}, X_k^{'i})\right] \xrightarrow[N\to\infty]{} \mathbb{P}(X_n \in A)$$

$$\sum_{i=1}^{N} f(X_n^{'i}) \frac{1_A(X_n^{'i})\prod_{k=1}^{n} G_k(X_{k-1}^{'i}, X_k^{'i})}{\sum_{j=1}^{N} 1_A(X_n^{'j})\prod_{k=1}^{n} G_k(X_{k-1}^{'j}, X_k^{'j})} \xrightarrow[N\to\infty]{} \mathbb{E}(f(X_n)|X_n \in A)\ .$$

This Monte Carlo method works rather well when the so-called twisted process X'_n is well identified and the time parameter n is not too large, but it cannot be interpreted in any way as a simulation methodology of the process in the rare event regime. A complementary methodology is to interpret, at each stage, the local Radon-Nikodym potential functions G_n as birth rates. These favour the particle transitions $X'_{n-1} \to X'_n$ moving too slowly towards the rare level set. The corresponding algorithm consists in evolving N-particles according to a genetic type mutation/selection method:

$$(\widehat{X}_{n-2}^{'i})_{1\leq i\leq N} \xrightarrow{\text{Mutat.}} (X_{n-1}^{'i})_{1\leq i\leq N} \xrightarrow{\text{Select.}} (\widehat{X}_{n-1}^{'i})_{1\leq i\leq N} \xrightarrow{\text{Mutat.}} (X_n^{'i})_{1\leq i\leq N}\ .$$

- During the selection mechanism, we examine the potential value of each past transition $(\widehat{X}_{n-2}^{'i}, X_{n-1}^{'i})_{1\leq i\leq N}$ and we select randomly N states $\widehat{X}_{n-1}^{'i}$ according to the discrete distribution

$$\sum_{i=1}^{N} \frac{G_{n-1}(\widehat{X}_{n-2}^{'i}, X_{n-1}^{'i})}{\sum_{j=1}^{N} G_{n-1}(\widehat{X}_{n-2}^{'j}, X_{n-1}^{'j})} \delta_{X_{n-1}^{'i}}\ .$$

- During the mutation mechanism, we simply evolve each selected particle $\widehat{X}_{n-1}^{'i}$ with a random elementary transition $\widehat{X}_{n-1}^{'i} \rightsquigarrow X_n^{'i} \sim M'_n(\widehat{X}_{n-1}^{'i}, \cdot)$.

The particle approximation models are now given by the occupation measures:

$$1_{|I_n^N|>0} \times \frac{1}{|I_n^N|}\sum_{i\in I_n^N} f(\widehat{X}_n^{'i}) \xrightarrow[N\to\infty]{} \mathbb{E}(f(X_n)|X_n \in A)\ ,$$

and the product formula

$$\frac{|I_n^N|}{N}\left[\prod_{k=1}^{n}\frac{1}{N}\sum_{i=1}^{N} G_k(\widehat{X}_{k-1}^{'i}, X_k^{'i})\right] \xrightarrow[N\to\infty]{} \mathbb{P}(X_n \in A)\ ,$$

where $|I_n^N|$ represents the cardinality of the set of indices of the particles having succeeded to enter in A at time n. Furthermore, if we trace back the complete genealogy of the particles having succeeded to reach the level A at time n, then we have for any test function f_n on the path space

$$1_{|I_n^N|>0} \times \frac{1}{|I_n^N|}\sum_{i\in I_n^N} f_n(\widehat{X}_{0,n}^{'i}, \cdots, \widehat{X}_{n,n}^{'i}) \xrightarrow[N\to\infty]{} \mathbb{E}(f_n(X_0, \cdots, X_n)|X_n \in A)\ ,$$

where $(\widehat{X}'^i_{k,n})_{0\leq k\leq n}$ represents the ancestral line of the end-time particle $\widehat{X}'^i_{n,n} = \widehat{X}'^i_n$. Although, we can prove that $\mathbb{P}(I^N_n = \emptyset)$ decreases to 0 exponentially fast, as $N \to \infty$, in practice we still need to choose a sufficiently large number of particles to ensure that a reasonably large proportion arrives to the target set. The propagation of chaos properties of the interactive particle models ensure that the random variables $\widehat{X}'^i_n$ behaves asymptotically as independent copies of X'_n in the rare event regime.

2.2 Interacting Trapping Models

This section is concerned with rare event estimation problems arising in particle trapping analysis, and nuclear engineering. These probabilistic models also provide interesting physical interpretations of rare events in terms of interactive trapping particles, and the associated genealogical structure. We also connect these rare event estimations with the analysis of Lyapunov exponents of Schrödinger operators.

We consider a physical particle X_n evolving in an absorbing medium E, related to a given potential function $G : E \to [0, 1]$. In the state space regions, where $G = 1$, the particle evolves randomly, and freely, according to a given Markov transition kernel $M(x, dy)$. When it enters in other regions, where $G < 1$, its life time decreases, and it is instantly absorbed when it visits the subset of null potential values. For indicator potential function, $G = 1_A$, $A \subset E$, this model reduces to a particle evolution killed on the complementary set $A^c = E \setminus A$. To visualize these models, Fig. 1 shows a particle evolution on $E = \mathbb{Z}$ killed outside an interval A at a random time T, and Fig. 2 illustrates the evolution of an absorbed particle in a lattice.

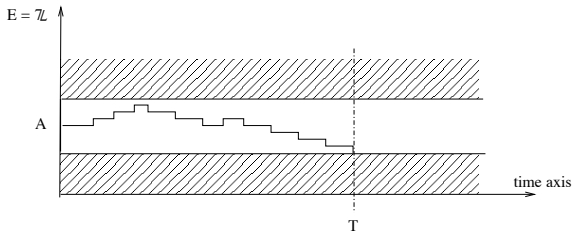

Fig. 1. Evolution of a particle in $E = \mathbb{Z}$ killed outside of A.

These probabilistic models arise in particle physics, such as in neutron collision/absorption analysis [8], as well as in nuclear engineering such as in the risk analysis of radiation containers shields. In this situation, the radiation source emits particles, which evolve in an absorbing shielding environment. In this context, the particle desintegrates when it visits the obstacles. The precise probabilistic model associated to these physical evolutions are discussed in Section 4.2.

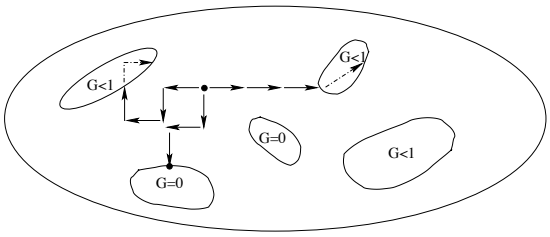

Fig. 2. Evolution of a particle in an absorbing medium.

If we let T be the first time the particle is absorbed, then we are interested in the rare event probabilities

$$\mathbb{P}(T > n) = \int_{E^{n+1}} \eta_0(dx_0)G(x_0)M(x_0, dx_1)G(x_1)\cdots M(x_{n-1}, dx_n)G(x_n) \ .$$

In the above formula, we integrate over all the particle paths $(x_0, \cdots, x_n) \in E^{n+1}$. The distribution η_0 represents the initial law of X_0 , $G(x_0)$ is the probability that the particle at x_0 is not killed, $M(x_0, dx_1)$ is the distribution of the transition from x_0 to x_1 , $G(x_1)$ is the probability that the particle at x_1 is not killed, and so on. For large values of the time parameter n , these probabilities are extremely small. In some sense, we have that

$$\mathbb{P}(T > n) = \mathbb{P}(T > 0)\prod_{i=1}^{n} \mathbb{P}(T > p | T > p-1) \approx e^{-n\lambda} \ , \tag{3}$$

for some constant $\lambda > 0$, which reflects the strength of the obstacles. This constant corresponds to the logarithmic Lyapunov exponent of the integral Schrödinger type semigroup, $G(x, dy) = G(x)M(x, dy)$. For more details, the reader is referred to [5].

To estimate these constants, and these rare event probabilities, we evolve N interacting particles, $\xi_n = (\xi_n^i)_{1\leq i\leq N} \in E^N$, according to the following rules

$$\xi_n = (\xi_n^i)_{1\leq i\leq N} \xrightarrow{\text{trapping/selection}} \widehat{\xi}_n = (\widehat{\xi_n^i})_{1\leq i\leq N} \xrightarrow{\text{evolution}} \xi_{n+1} = (\xi_{n+1}^i) \ .$$

During the trapping transition, each particule ξ_n^i survives with a probability $G(\xi_n^i)$, and in this case we set $\widehat{\xi_n^i} = \xi_n^i$. Otherwise, with a probability $1 - G(\xi_n^i)$, the particle is absorbed, and instantly another randomly chosen particle in the current configuration duplicates. More precisely, when the particle ξ_n^i is absorbed, we chose randomly a new particle ξ_n^i according to the discrete Gibbs measure

$$\sum_{j=1}^{N} \frac{G(\xi_n^j)}{\sum_{k=1}^{N} G(\xi_n^k)} \delta_{\xi_n^j} \ .$$

During the evolution step, each selected particule $\widehat{\xi_n^i}$ evolves randomly according to the Markov transition M. The rare event probabilities are approxi-

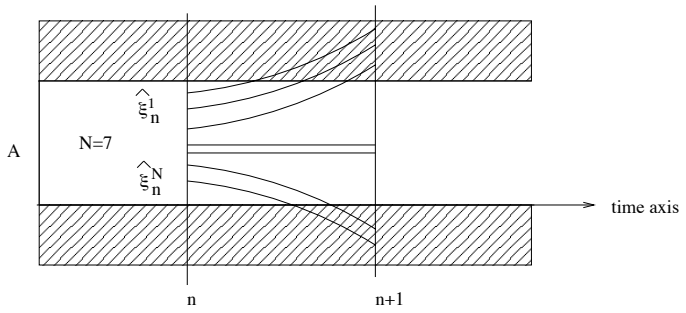

Fig. 3. Interacting particle with indicator potential function $G = 1_A$.

mated by the product formula

$$\mathbb{P}^N(T > n) = \prod_{p=0}^{n} \left(\frac{1}{N} \sum_{i=1}^{N} G(\xi_p^i) \right) \to \mathbb{P}(T > n)$$

$$= \mathbb{P}(T > 0) \prod_{i=1}^{n} \mathbb{P}(T > p | T > p - 1) \, .$$

In the case of indicator potential function $G = 1_A$, we notice that the empirical mean potentials corresponds to the population of evolving transitions which have not been absorbed. In Fig. 3, we illustrate an example with $N = 7$ and $N^{-1} \sum_{i=1}^{N} 1_A(\xi_{n+1}^i) = 2/7$.

For long time horizon, we also have a particle interpretation of the Lyapunov exponent λ, previously introduced in (3)

$$-\frac{1}{n+1} \sum_{p=0}^{n} \log \left(\frac{1}{N} \sum_{i=1}^{N} G(\xi_p^i) \right) \approx \lambda.$$

In the birth and death interpretation, we can trace back the complete genealogy of a given particle ξ_n^i . If we let

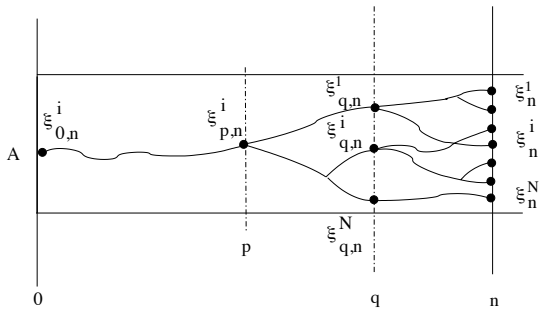

Fig. 4. Genealogical tree associated with the interactif trapping model.

$$\xi^i_{0,n} \leftarrow \xi^i_{1,n} \leftarrow \cdots \leftarrow \xi^i_{n-1,n} \leftarrow \xi^i_{n,n} = \xi^i_n$$

be the ancestral line of the particle with label i, at time n, then we have for any test function f_n on the state space E^{n+1},

$$\frac{1}{N}\sum_{i=1}^{N} f_n(\xi^i_{0,n}, \xi^i_{1,n}, \cdots, \xi^i_{n,n}) \xrightarrow[N\to\infty]{} \mathbb{E}\left(f(X_0, \cdots, X_n) | T \geq n\right) \ .$$

In some sense, the genealogical tree, associated with interaction trapping model, represents the path strategy used by the Markov particle to stay alive up to time n. Returning to the indicator potential function example, a model of a random tree is represented in Fig. 4.

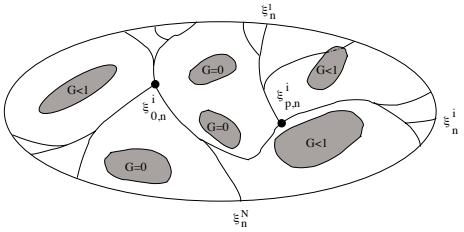

Fig. 5. Genealogical tree model in the lattice example.

In the lattice example, the genealogical tree models correspond to a spider web type strategy, as such illustrated in Fig. 5

2.3 A Bernoulli Splitting Technique

In contrast to importance sampling type algorithms, in the trajectory splitting methodology, the step-by-step evolution of the system follows the original probability measure. Entering the intermediate states, which is usually characterized by crossing a threshold by a control parameter, triggers the splitting of the trajectory. The current system state is held, and a number of independent subtrajectories are simulated from that state.

For example, let us consider $(m+1)$ sets B_i such that

$$B_{m+1} \subset B_m \subset \cdots \subset B_1 \, .$$

When the rare event A coincide with B_{m+1}, we have the product formula

$$\mathbb{P}(A) = \mathbb{P}(A|B_m)\mathbb{P}(B_m|B_{m-1}) \cdots \mathbb{P}(B_2|B_1)\mathbb{P}(B_1). \tag{4}$$

On the right hand side of (4, each conditioning event is "not rare". The branching splitting technique proceeds as follows. Make a $\{0,1\}$ Bernoulli trial to check whether or not the set event B_1 has occured. If B_1 has occured, then we

split this trial in R_1 Bernoulli additional trials, and for each of them we check again wether or not the event B_2 has occured. This procedure is repeated at each level. More precisely, each time the event B_i has occured, we sample R_i trials and we repeat this splitting technique each time B_{i+1} has occured. If an event level is not reached, neither is A, then we stop the current retrial. Using R_0 independent replications of this procedure, we have then considered $R_0 R_1 \cdots R_m$ trials, taking into account for example, that if we have failed to reach a level B_i at the i-th step, the $R_i \cdots R_m$ possible retrials have failed. An unbiased estimator of $\mathbb{P}(A)$ is clearly given by the quantity

$$\widehat{P} = \frac{N_A}{R_0 \prod_{i=1}^{m} R_i},$$

where N_A is the total number of trajectories having reached the set A. It can be proven [10] that in some sense the optimal simulation is obtained if

$$m = \lfloor -0.6275 \log P(A) - 1 \rfloor, \quad \mathbb{P}(B_i|B_{i-1}) \approx 1/5, \quad R_i = 5.$$

Nevertheless, in practice the trajectory splitting method may be difficult to apply. For example, the case of the estimation of the probability of a rare event in dynamical system is more complex, since the difficulty to find theoretically the optimal B_i, and R_i for each level i. Furthermore, the probability to reach B_i varies generally with the state of entrance in level B_{i-1}. Finally, but not the least, the conditional probabilities $\mathbb{P}(B_i|B_{i-1})$ are of course generally unknown! In this sense, this rather crude splitting strategy is of pure academic interest.

2.4 Branching Splitting Models

The branching strategy, we are about to describe, is rather close in spirit to the Bernoulli splitting method described above. The essential difference is that it enters the random evolution of the process in the rare event region. To be more precise, we consider a Markov process evolving in some state space, in such a way that a given region, or a particular site, say O, is visited infinitly often. Our objective is to estimate the probability $\mathbb{P}(A)$ to reach a rare level A before returning to O. For instance, if T_A represents the first hitting time of A, and T_O the first return time to O, then $\mathbb{P}(A) = \mathbb{P}(T_A < T_O)$. We proceed as shown in Fig. 6. If the first level B_1 is reached before going back to O, then we split the path into R_1 trials; otherwise, if we are back to O, we stop the exploration. At the next step we evolve each of these R_1 paths, starting from its entrance state in B_1. If a path hits B_2, before returning to O, then it is again splitted into R_2 trials; otherwise we stop its exploration. We repeat the branching transition at each level B_k. Finally, a path from level B_m that succeed to reach A (before returning to O) is considered as a success, and it is stopped. An implicit, and technical assumption is that the level $(i+1)$ cannot be reached from level $(i-1)$ without entering previously the intermediate level

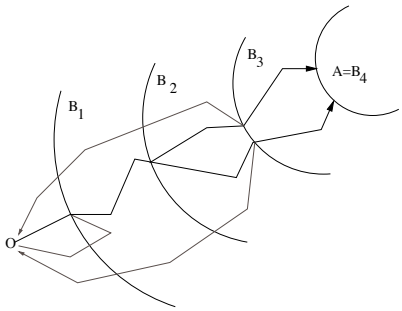

Fig. 6. Example with $R_i = 2$ at each level

i. This condition is clearly met, if we consider a decreasing sequence of level set $A = B_{m+1} \subset B_m \subset \cdots \subset B_1$.

In practice, the simulation time is limited to a given value T, so that we estimate

$$\mathbb{P}(T_A < \min(T_O, T)) .$$

Moreover, going back from B_i to O may take a long time, so we can reduce the computations by freezing the path exploration when it goes back to one, two or more levels down. Nevertheless, this rather crude strategy induces a bias, which is difficult to estimate. An alternative approach is the RESTART method, introduced in [15, 16].

2.5 Interacting Particle Systems Algorithm

In this section we design a genetic, and genealogical tree base model for estimating a rather general class of rare events, following [1]. The main idea behind this evolutionary type algorithm is again to decompose the state space into a judicious choice of threshold levels. This decomposition reflects the successive levels the stochastic process needs to cross before entering into the rare event. More precisely, we consider a strong Markov chain X_n, which is assumed to start in some set O with a given initial probability distribution. We associate to a given target set A, the first time the process hits the set A, namely

$$T_A = \inf\{n \geq 0 \ : \ X_n \in A\} .$$

We use the classical convention $\inf \emptyset = \infty$. We would like to estimate the quantities

$$\mathbb{P}(T_A \leq T) \quad \text{and} \quad \text{Law}((X_n), 0 \leq n \leq T_A | T_A \leq T) . \tag{5}$$

In the above formulas, T is either a deterministic finite horizon time, or the (finite) entrance time into a recurrent set R when $R \cap O = \emptyset$ (or the first return time to O, if $R = O$).

As previously, before visiting R, or entering into A, the random process passes through a decreasing sequence of level sets

$$A = B_{m+1} \subset B_m \subset \cdots \subset B_1 \ ,$$

with $B_1 \cap (O \cup R) = \emptyset$. To capture the precise behavior of X between the different levels, we consider the random excursions $\mathcal{X}_n$ of X between the successive random times T_{n-1} and T_n ,where $(T_n)_{n=1,\cdots,m+1}$ represent the entrance times of the level sets $(B_n)_{n=1,\cdots,m+1}$. A synthetic picture of these excursions is given in the Fig. 7. We observe that these excursions may have different

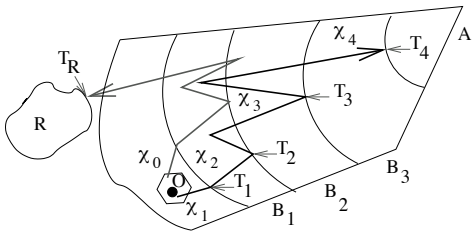

Fig. 7. Embedded Markov Chain

random lengths, and we have the decomposition formula

$$(T_A \leq T) = (T_{m+1} \leq T) = (T_1 \leq T, \cdots, T_{m+1} \leq T) \, .$$

To check wether or not a given path $(x_k)_{p \leq k \leq q}$, starting at $x_p \in B_{n-1}$ at time p, has succeeded to reach the level B_n at time q, it is convenient to introduce the indicator potential functions G_n defined on the excursion space by

$$G_n((x_k)_{p \leq k \leq q}) = \mathbf{1}_{\{x_q \in B_n\}}.$$

With this notation, and for each n we have

$$(T_n \leq T) = (G_1(\mathcal{X}_1) = 1, \cdots, G_n(\mathcal{X}_n) = 1) = \left(\prod_{p=1}^{n} G_p(\mathcal{X}_p) = 1 \right) \ .$$

Integrating over all the process excursions, we obtain the following formula

$$\mathbb{P}(T_n \leq T) = \mathbb{E}\left(\prod_{p=1}^{n} G_p(\mathcal{X}_p) \right) \ . \tag{6}$$

More generally, the law of the excursions in the rare event regime are described by the following formulas

$$\mathbb{E}(f(\mathcal{X}_0, \cdots, \mathcal{X}_n) | T_n \leq T) = \frac{\mathbb{E}\left(f(\mathcal{X}_0, \cdots, \mathcal{X}_n) \prod_{p=0}^{n} G_p(\mathcal{X}_p) \right)}{\mathbb{E}\left(\prod_{p=0}^{n} G_p(\mathcal{X}_p) \right)} \tag{7}$$

for any bounded test function f.

Functional representation of the form (6), and (7), belong to the class of Feynman-Kac formulas. A detailed account on these models can be found in [2], and the references therein. These models have natural particle interpretations, in terms of genealogical tree-based evolutions. An elementary genetic type approximation model is briefly described as follows: When a particle, starting at some level, does not succeed to enter into the next one, it is killed. Otherwise, each time it enters into a closer level of the rare set, it splits into several offsprings. Between the levels, these offsprings evolve as independent copies of the stochastic process $\mathcal{X}_n$, until they reach (or not) an even closer level, and so on.

To be more precise, we evolve N particles according to a two-steps, and genetic type mechanism:

$$(\xi_n^1, \cdots, \xi_n^N) \xrightarrow{\text{selection}} (\widehat{\xi}_n^1, \cdots, \widehat{\xi}_n^N) \xrightarrow{\text{mutation}} (\xi_{n+1}^1, \cdots, \xi_{n+1}^N) \tag{8}$$

- During the selection, each particle with label i having succeeded to reach the n-th level is held, and we set $\widehat{\xi}_n^i = \xi_n^i$. The others $\widehat{\xi}_n^j$ are chosen randomly (and uniformly) in the set of those having succeeded to reach the level B_n. If N_n denotes the number of particle which succeeded to reach the level B_n, then the estimate of the conditional probability $\mathbb{P}(B_n|B_{n-1})$ is simply given by the proportion ratio N_n/N.
- During the mutation, each particle $\widehat{\xi}_n^i$ evolves to a new location ξ_{n+1}^i, randomly chosen according to the transition probability of the chain $\mathcal{X}_n$ in the excursion space.

From previous consideration, a natural unbiased estimate of $\mathbb{P}(T_A \leq T)$ is simply given by the product $\prod_{p=1}^{m+1}(N_p/N)$. Using the propagation of chaos properties of the particle approximation models, one can prove that the above estimate converges to the true value, as $N \to \infty$, [1, 2]. More precisely, we have the almost sure convergence result

$$\mathbb{P}^N(T_A < T) = \prod_{p=1}^{m+1} \frac{N_p}{N} \xrightarrow[N\to\infty]{} \mathbb{P}(T_A < T) \, .$$

The genealogical tree based model associated with the above genetic-type algorithm represents the conditional distribution of the process evolving in the rare event regime. To be more precise, we let

$$\xi_{0,n}^i \leftarrow \xi_{1,n}^i \leftarrow \xi_{2,n}^i \leftarrow \cdots \leftarrow \xi_{n,n}^i$$

be the ancestral lines of the excursion-valued particles $(\xi_{n,n}^i)_{i\in I_n^N}$ having succeeded to reach the n-th level. For any bounded and measurable test function f defined on the excursion space, we have the almost sure result

$$1_{|I_n^N|>0} \times \frac{1}{|I_n^N|} \sum_{i\in I_n^N} f(\xi_{0,n}^i, \cdots, \xi_{n,n}^i) \xrightarrow[N\to\infty]{} \mathbb{E}\left(f(\mathcal{X}_0, \cdots, \mathcal{X}_n)|T_n < T\right) \, .$$

In the figure 8, we provide a schematic picture of a genealogical tree associated with $N = 4$ particles evolving on the lattice $\mathbb{Z}$ between a sequence of 3 upper-levels. Each particle starting at the origin, tends to move back to the set of non positive integer $O = -\mathbb{N}$. The prototype of this model is the simple random walk X_n on $\mathbb{Z}$ given by the transitions (1).

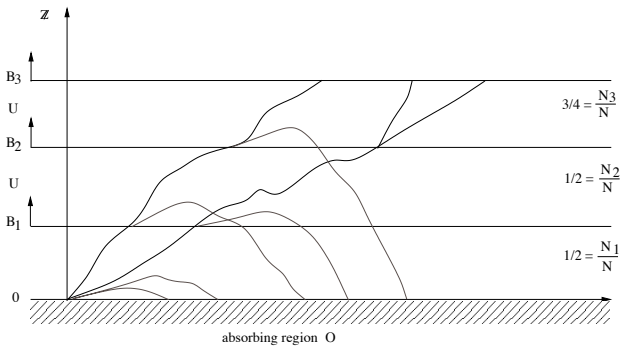

Fig. 8. Genealogical tree

3 Markov Chain and Random Excursion Models

Intuitively speaking, a sequence of random variables $(X_n)_{n\geq 0}$, taking values at each time n in some measurable state space $(E_n, \mathcal{E}_n)$, is said to be a Markov chain when its future and its past trajectories are independent, given the present state of the chain. This Markov property is an extension to a random phenomenon of the well-known property of a deterministic dynamical system, which basically says that the future position and velocity are uniquely defined as soon as they are known at a previous date.

A Markov chain is characterized by its Markov kernels $M_n(x_{n-1}, dx_n)$, which describe the conditional probability of the transition from the point $X_{n-1} = x_{n-1} \in E_{n-1}$ to the infinitesimal neighborhood dx_n of the point $x_n \in E_n$. More formally we have that,

$$\mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1}) = M_n(x_{n-1}, dx_n).$$

3.1 Canonical Probability Space

Using the Markov dependence property, μ as the distribution of the initial random state X_0, we check that

$$\begin{aligned}\mathbb{P}_\mu((X_0,\cdots,X_n)\in d(x_0,\cdots,x_n))&\\ &= M_n(x_{n-1},dx_n)\mathbb{P}_\mu\left((X_0,\cdots,X_{n-1})\in d(x_0,\cdots,x_{n-1})\right)\\ &= \mu(dx_0)M_1(x_0,dx_1)\cdots M_n(x_{n-1},dx_n)\end{aligned}$$

where $d(x_0,\cdots,x_n)$ stands for an infinitesimal neighborhood of the path point

$$(x_0,\cdots,x_n)\in E_0\times\cdots\times E_n\ .$$

We define the distribution $\mathbb{P}_{\mu,n}$ of the canonical sequence $(X_0,\cdots,X_n)$ on $\Omega_n=\left(\prod_{p=0}^n E_p\right)$, equipped with the product σ-field $\mathcal{F}_n=\otimes_{p=0}^n\mathcal{E}_p$, by setting

$$\mathbb{P}_{\mu,n}(dx_{[0,n]})=\mu(dx_0)M_1(x_0,dx_1)\cdots M_n(x_{n-1},dx_n).$$

By the consistency property of the collection $\mathbb{P}_{\mu,n}$, $n\in\mathbb{N}$, the Ionescu Tulcea's theorem ensures the existence of an overall distribution $\mathbb{P}_\mu$, on the whole path space $\Omega=\left(\prod_{n\geq 0}E_n\right)$, with finite-dimensional distributions $\mathbb{P}_{\mu,n}$. If we denote by $X_n, n\in\mathbb{N}$, the canonical projection mappings

$$X_n:\omega=(\omega_n)_{n\geq 0}\in\Omega\longrightarrow X_n(\omega)=\omega_n\in E_n\ ,$$

then for any $A_p\in\mathcal{E}_p, p\geq 0$, we have that

$$\begin{aligned}&\mathbb{P}_\mu((X_0,\cdots,X_n)\in(A_0\times\cdots\times A_n))\\ &=\int_{A_0\times\cdots\times A_n}\mu(dx_0)M_1(x_0,dx_1)\cdots M_n(x_{n-1},dx_n).\end{aligned}$$

For obvious reasons, the probability model defined in this way

$$(\Omega,\mathcal{F}=(\mathcal{F}_n)_{n\geq 0},X=(X_n)_{n\geq 0},\mathbb{P}_\mu),\tag{9}$$

is called the canonical realisation of the Markov chain, with transitions M_n, and initial distribution μ.

3.2 Path-Space Markov Models

As we mentionned in the introduction, the path space modelling is dictated by the excursions analysis of the process in the rare event regime.

Let $(E_n,\mathcal{E}_n)$ be an auxilairy collection of mesurable spaces, and let X_n be a nonanticipative sequence of E_n-valued random variables in the sense that the distribution of X_{n+1} on E_{n+1} only depends on the random states $(X_0,\cdots,X_n)$. By direct inspection, we notice that the path sequence

$$\mathcal{X}_n=X_{[0,n]}=(X_0,\cdots,X_n),$$

forms a nonhomogeneous Markov chain taking values in the product space

$$E_{[0,n]} = (E_0 \times \cdots \times E_n) \,,$$

equipped with the product σ-field

$$\mathcal{F}_n = \mathcal{E}_0 \otimes \cdots \otimes \mathcal{E}_p \,.$$

In this situation, each point $x_{[0,n]} = (x_0, \cdots, x_n) \in E_{[0,n]}$ has to be thought of like a path from the origin up to time n.

When X_n is a Markov chain, with non necessarily time homogeneous transitions $M_n(x_{n-1}, dx_n)$ from E_{n-1} into E_n, the Markov chain $\mathcal{X}_n$ in path space is called the historical process, or the path process of the chain X_n.

The Markov transitions $\mathcal{M}_n$ of the chain $\mathcal{X}_n$ are connected to M_n by the formula

$$\mathcal{M}_{n+1}(x_{[0,n]}, dy_{[0,n+1]}) = \delta_{x_{[0,n]}}(dy_{[0,n]}) M_{n+1}(y_n, dy_{n+1}).$$

The motion of the path process $\mathcal{X}_n$ simply consists of extending each path of X_n with an elementary M_n-transition. In summary, we have the synthetic diagram

$$\mathcal{X}_{n-1} = X_{[0,n-1]} \longrightarrow \mathcal{X}_n = X_{[0,n]} = (X_{[0,n-1]}, X_n),$$

with the random state $X_n \sim M_{n+1}(X_{n-1}, \cdot)$.

4 Feynman-Kac Models

To describe precisely the Feynman-Kac models, we need to introduce some additional notation. Firstly, we denote by $\mathcal{B}_b(E)$ the set of bounded measurable functions on a given measurable space $(E, \mathcal{E})$. The expectation operators with respect to $\mathbb{P}_\mu$ and $\mathbb{P}_x$ are denoted by $\mathbb{E}_\mu$ and $\mathbb{E}_x$. For instance, for any $F_n \in \mathcal{B}_b(E_{[0,n]})$, we have

$$\mathbb{E}_\mu(F_n(X_{[0,n]})) = \int_{E_{[0,n]}} F_n(x_{[0,n]}) \mathbb{P}_{\mu,n}(dx_{[0,n]}).$$

We denote respectively by $\mathcal{M}(E_n)$, and $\mathcal{P}(E_n) \subset \mathcal{M}(E_n)$ the set of bounded and signed measures, and the subset of probability measures on the measurable space $(E_n, \mathcal{E}_n)$. We also recall that any Markov kernel M_n from E_{n-1} to E_n generates two operators: The first acting on $\mathcal{B}_b(E_n)$, taking value in $\mathcal{B}_b(E_{n-1})$, and defined by

$$\forall (x_{n-1}, f_n) \in (E_{n-1} \times \mathcal{B}_b(E_n)) \,,$$
$$M_n(f_n)(x_{n-1}) = \int_{E_n} M_n(x_{n-1}, dx_n) f_n(x_n) \,.$$

The other one acting on measures $\mu_{n-1} \in \mathcal{P}(E_{n-1})$, taking values in $\mathcal{P}(E_n)$, and defined by

$$\forall (\mu_{n-1}, A_n) \in (\mathcal{P}(E_{n-1}) \times \mathcal{E}_n) ,$$
$$(\mu_{n-1}M_n)(A_n) = \int_{E_{n-1}} \mu_{n-1}(dx_{n-1}) M_n(x_{n-1}, A_n) .$$

Finally, if $M_{n+1}(x_n, dx_{n+1})$ is a Markov transition from $(E_n, \mathcal{E}_n)$, to another measurable space $(E_{n+1}, \mathcal{E}_{n+1})$, then we denote by $M_n M_{n+1}$ the composition operator,

$$M_n M_{n+1}(x_{n-1}, dx_{n+1}) = \int_{E_n} M_n(x_{n-1}, dx_n) M_{n+1}(x_n, dx_{n+1}).$$

Finally, for a given integral operator M from E_0 into E_1, for any $x \in E_0$ and $\varphi^1, \varphi^2 \in \mathcal{B}_b(E_1)$, we simplify notation, and we write

$$M[(\varphi^1 - M\varphi^1)(\varphi^2 - M\varphi^2)](x)$$

instead of

$$M[(\varphi^1 - M(\varphi^1)(x))(\varphi^2 - M(\varphi^2)(x))](x) = M(\varphi^1\varphi^2)(x) - M(\varphi^1)(x)M(\varphi^2)(x). \tag{10}$$

4.1 Description of the Models

We consider a given collection of bounded and $\mathcal{E}_n$-measurable nonnegative functions $G_n : E_n \to [0, \infty)$ such that for any $n \in \mathbb{N}$, we have

$$\mathbb{E}_\mu\Big(\prod_{p=0}^{n} G_p(X_p)\Big) > 0. \tag{11}$$

Definition 1. *The Feynman-Kac prediction and updated path models, associated with the pair (G_n, M_n) (and the initial distribution μ), are the sequence of measures on path space defined respectively, for any $n \in \mathbb{N}$, by the formulas*

$$\mathbb{Q}_{\mu,n}(dx_{[0,n]}) = \frac{1}{\mathcal{Z}_n}\Big\{\prod_{p=0}^{n-1} G_p(x_p)\Big\}\mathbb{P}_{\mu,n}(dx_{[0,n]}) ,$$
$$\widehat{\mathbb{Q}}_{\mu,n}(dx_{[0,n]}) = \frac{1}{\widehat{\mathcal{Z}}_n}\Big\{\prod_{p=0}^{n} G_p(x_p)\Big\}\mathbb{P}_{\mu,n}(dx_{[0,n]}) .$$

The normalizing constants

$$\mathcal{Z}_n = \mathbb{E}_\mu\Big(\prod_{p=0}^{n-1} G_p(x_p)\Big) \quad \text{and} \quad \widehat{\mathcal{Z}}_n = \mathcal{Z}_{n+1} = \mathbb{E}_\mu\Big(\prod_{p=0}^{n} G_p(x_p)\Big)$$

are also often called the partition functions.

Note that for any test function $F_n \in \mathcal{B}_b(E_{[0,n]})$, we have

$$\mathbb{Q}_{\mu,n}(F_n) = \frac{1}{\mathcal{Z}_n}\mathbb{E}_\mu\Big(F_n(X_0,\cdots,X_n)\prod_{p=0}^{n-1}G_p(X_p)\Big)\ ,$$

$$\widehat{\mathbb{Q}}_{\mu,n}(F_n) = \frac{1}{\widehat{\mathcal{Z}}_n}\mathbb{E}_\mu\Big(F_n(X_0,\cdots,X_n)\prod_{p=0}^{n}G_p(X_p)\Big)\ .$$

The Feynman-Kac models have a particular dynamic structure. To describe precisely their evolution, it is convenient to introduce the flow of the time marginals.

Definition 2. *The sequence of bounded nonnegative measures γ_n, and $\widehat{\gamma}_n$, on E_n, and defined for any $f_n \in \mathcal{B}_b(E_n)$ by the formulas*

$$\gamma_n(f_n) = \mathbb{E}_\mu\Big(f_n(X_n)\prod_{p=0}^{n-1}G_p(X_p)\Big)\ ,$$

$$\widehat{\gamma}_n(f_n) = \mathbb{E}_\mu\Big(f_n(X_n)\prod_{p=0}^{n}G_p(X_p)\Big)\ ,$$

are called the unnormalized prediction, and updated, Feynman-Kac model associated with the pair (G_n, M_n). The sequence of distributions η_n, and $\widehat{\eta}_n$, on E_n, and defined for any $f_n \in \mathcal{B}_b(E_n)$ by

$$\eta_n(f_n) = \gamma_n(f_n)/\gamma_n(1) \quad \textit{and} \quad \widehat{\eta}_n(f_n) = \widehat{\gamma}_n(f_n)/\widehat{\gamma}_n(1)$$

are called the normalized prediction, and updated, Feynman-Kac model associated with the pair (G_n, M_n).

To get one step further, we notice that

$$\gamma_n(f_nG_n) = \widehat{\gamma}_n(f_n)\ , \quad \text{and} \quad \widehat{\eta}_n(f_n) = \frac{\gamma_n(f_nG_n)}{\gamma_n(G_n)} = \frac{\eta_n(f_nG_n)}{\eta_n(G_n)}. \tag{12}$$

An other key product formula that relates the "unnormalized models" $(\gamma_n, \widehat{\gamma}_n)$ with the Feynman-Kac distribution flow $(\eta_p)_{p\leq n}$, is given by

$$\gamma_n(f_n) = \eta_n(f_n)\prod_{p=0}^{n-1}\eta_p(G_p) \quad \text{and} \quad \widehat{\gamma}_n(f_n) = \widehat{\eta}_n(f_n)\prod_{p=0}^{n}\eta_p(G_p).$$

The identity (12) leads us to introduce the following transformation.

Definition 3. *The Boltzmann-Gibbs transformation associated with a potential function G_n on $(E_n, \mathcal{E}_n)$ is the mapping*

$$\Psi_n : \eta \in \mathcal{P}_n(E_n) \longrightarrow \Psi_n(\eta) \in \mathcal{P}_n(E_n)$$

from the subset $\mathcal{P}_n(E_n) = \{\eta \in \mathcal{P}(E_n) : \eta(G_n) > 0\}$ *into itself, and defined by*

$$\Psi_n(\eta)(dx_n) = \frac{1}{\eta(G_n)}\, G_n(x_n)\, dx_n \ .$$

In this notation, we see that

$$\widehat{\eta}_n = \Psi_n(\eta_n) \ , \quad \text{and} \quad \eta_n = \widehat{\eta}_{n-1} M_n \ . \tag{13}$$

The last identity comes from the following observation

$$\gamma_n(f_n) = \mathbb{E}_\mu\Big(M_n(f_n)(X_{n-1}) \prod_{p=0}^{n-1} G_p(X_p)\Big) = \widehat{\gamma}_{n-1}(M_n(f_n)) \ .$$

We conclude that, the Feynman-Kac flows $(\eta_n, \widehat{\eta}_n)$ are the solution of the nonlinear and measure-valued processes equations

$$\eta_n = \Phi_n(\eta_{n-1}) \ , \quad \text{and } \widehat{\eta}_n = \widehat{\Phi}_n(\widehat{\eta}_{n-1}) \ , \tag{14}$$

with the one step mappings Φ_n, and $\widehat{\Phi}_n$, defined by

$$\Phi_n(\eta) = \Psi_{n-1}(\eta) M_n \ , \qquad \widehat{\Phi}_n = \Psi_n(\eta M_n) \ .$$

We emphasize that the above evolution analysis strongly relies on the fact that the potential functions $(G_n)_{n\geq 0}$ satisfy the regularity condition stated in (11). For instance, the measure-valued equations (14) may not be defined for any initial distribution η_0 or $\widehat{\eta}_0$, since it may be happen that $\eta_0(G_0) = 0$, or $\widehat{\eta}_0(G_0) = 0$. On the other hand, when the potential functions G_n are unbounded, the Boltzmann-Gibbs transformation Ψ_n are only defined on the set $\{\eta \in \mathcal{P}(E_n), 0 < \eta(G_n) < \infty\}$.

To solve these problems, we further require that the pairs (G_n, M_n) satisfy for any $x_n \in E_n$ the following condition:

$$0 < \widehat{G}_n(x_n) = M_{n+1}(G_{n+1})(x_n) \quad \text{and} \quad \sup_{x_n} |\widehat{G}_n(x_n)| = \|\widehat{G}_n\| < \infty \, . \tag{15}$$

In this situation, the integral operators

$$\widehat{M}_n(x_{n-1}, dx_n) = \frac{M_n(x_{n-1}, dx_n) G_n(x_n)}{M_n(G_n)(x_{n-1})}$$

are well-defined Markov-kernels from E_{n-1} to E_n . With this notation, the mapping $\widehat{\Phi}_n$ can be expressed as follows

$$\widehat{\Phi}_n = \widehat{\Psi}_{n-1}(\eta) \widehat{M}_n \, ,$$

where $\widehat{\Psi}_n$ is the Boltzmann-Gibbs transformation associated with the pair potential/kernel $(\widehat{G}_n, \widehat{M}_n)$ and the initial measure $\widehat{\eta}_0$. Thus the updated

Feynman-Kac models associated with the pair (G_n, M_n) and initial measure η_0 coincide with the prediction Feynman-Kac models associated with the pairs $(\widehat{G}_n, \widehat{M}_n)$ starting at $\widehat{\eta}_0$. As we mentionned above, the interpretation of the updated flow as a prediction flow associated with the pair $(\widehat{G}_n, \widehat{M}_n)$ is often more judicious. To illustrate this observation, we examine the situation where the potential function G_n may take some null values, and we set

$$\widehat{E}_n = \{x_n \in E_n : \quad G_n(x_n) > 0\} .$$

It may happen that $\widehat{E}_n$ is not M_n-accessible from any point in E_{n-1}. In this case, we may have $M_n(x_{n-1}, \widehat{E}_n) = 0$, for some $x_{n-1} \in E_{n-1}$, and therefore $M_n(G_n)(x_{n-1}) = 0$. In this situation, the condition (15) is clearly not met. So, we weaken it by considering the following condition

$$(\mathcal{A}) \qquad \forall x_n \in \widehat{E}_n,\ M_{n+1}(x_n, \widehat{E}_{n+1}) > 0, \text{ and } \eta_0(\widehat{E}_0) > 0 , \qquad (16)$$

which says that the set $\widehat{E}_{n+1}$ is accessible from any point in $\widehat{E}_n$. This accessibility condition avoids some degenerate tunneling problems such as those represented in the figure 9.

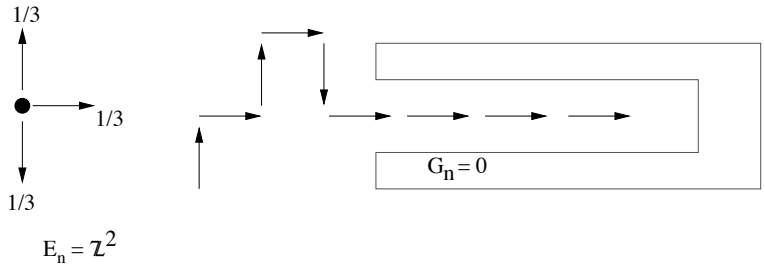

Fig. 9. Tunneling problem

Assuming the condition $(\mathcal{A})$, the condition (15) is only met for any $x_n \in \widehat{E}_n$, and the operators $\widehat{M}_n$ (defined for any $x_{n-1} \in \widehat{E}_{n-1}$) are well-defined Markov kernels from $\widehat{E}_{n-1}$ into $\widehat{E}_n$. Finally, we note that for any $\eta_0 \in \mathcal{P}(E_0)$, with $\eta_0(\widehat{E}_0) > 0$, the updated measure $\widehat{\eta}_0 = \Psi_0(\eta_0)$ is such that $\widehat{\eta}_0(\widehat{E}_0) = 1$.

Summarizing the discussion above, the updated Feynman-Kac measures $\widehat{\eta}_n \in \mathcal{P}(\widehat{E}_n)$ can be interpreted as the prediction models associated with the pair potential/kernel $(\widehat{G}_n, \widehat{M}_n)$ on the restricted state space $(\widehat{E}_n, \widehat{\mathcal{E}}_n)$, as soon as the accessibility condition $\mathcal{A}$ is met. We can also check that

$$\mathbb{E}_{\eta_0}\left(f_n(X_n)\prod_{p=0}^{n} G_p(X_p)\right) = \eta_0(G_0)\,\widehat{\mathbb{E}}_{\widehat{\eta}_0}\left(f_n(X_n)\prod_{p=0}^{n-1}\widehat{G}_p(X_p)\right) > 0 .$$

In particular, this shows that for any $n \in \mathbb{N}$, we have

$$\eta_n \in \mathcal{P}_n(E_n) = \{\eta \in \mathcal{P}(E_n) : \ \eta(G_n) > 0\} \ .$$

Therefore, the Feynman-Kac flow is a well-defined two-step updating/prediction model

$$\eta_n \in \mathcal{P}_n(E_n) \xrightarrow{\text{updating}} \widehat{\eta}_n \in \mathcal{P}_n(\widehat{E}_n) \xrightarrow{\text{prediction}} \eta_{n+1} \in \mathcal{P}_{n+1}(E_{n+1}) \ .$$

Finally, when the accessibility condition $(\mathcal{A})$ is not met, it may happen that

$$\widehat{\eta}_n M_{n+1}(G_{n+1}) = \eta_{n+1}(G_{n+1}) = 0 \ .$$

In this situation, the Feynman-Kac flow η_n is well-defined, up to the first time τ we have $\eta_\tau(G_\tau) = 0$. At time τ, the measure η_τ cannot be updated anymore. Recalling that $\eta_\tau(G_\tau) = \gamma_{\tau+1}(1)/\gamma_\tau(1)$, we also see that τ coincides with the first time that

$$\widehat{\gamma}_\tau(1) = \gamma_{\tau+1}(1) = \mathbb{E}_{\eta_0}\Big(\prod_{p=0}^{\tau} G_p(X_p)\Big) = 0 \ .$$

4.2 Physical Interpretations of the Feynman-Kac Models

We now provide different physical interpretations of the Feynman-Kac models. The first one is the traditional trapping interpretation, the second one is based on measure-valued, and interacting processes ideas, such as those arising in mathematical biology.

In the first part, we design a Feynman-Kac representation of distribution flows of a Markov particle evolving in an absorbing medium. As we mentionned in the introduction, these probabilistic models provide a physical interpretation of rare event probabilities in terms of absorption time distributions. In the second part, we set out an alternative representation in terms of nonlinear and measure valued processes, the so-called McKean interpretation. The cornerstone of the particle interpretations, developped in this section, is the interpretation of the Feynman-Kac model as such the distribution of a non absorbed particle.

To clarify the presentation, we assume that the potential functions G_n are strictly positive. On the other hand, since the potential functions G_n are assumed to be bounded, we can replace in the definition of the normalized measures η_n, $\widehat{\eta}_n$, the functions G_n by $G_n/\|G_n\|$, without altering their nature. So, there is no loss of generality to assume that $0 < G_n(x_n) \leq 1$.

Killing Interpretation

Now, we identify the potential functions G_n with the multiplicative operator $\mathcal{G}_n$, acting on $\mathcal{B}_b(E_n)$, and defined by the formula

$$\mathcal{G}_n(f_n)(x_n) = G_n(x_n) \ f_n(x_n) \ .$$

We can alternatively see $\mathcal{G}_n$ as the integral operator on E_n defined by

$$\mathcal{G}_n(x_n, dy_n) = G_n(x_n)\delta_{x_n}(dy_n) \; .$$

In this connection, we note that $\mathcal{G}_n$ is a sub-Markovian kernel

$$\mathcal{G}_n(x_n, E_n) = G_n(x_n) \leq 1 \; .$$

The first way to turn the sub-Markovian kernels $\mathcal{G}_n$ into the Markov case consists in adding a cemetery point c to the state space E_n , and then extending the various quantities on the space $E_n^c = E_n \cup \{c\}$ as follows:

- The test functions f_n and the potential functions G_n are extended by setting $f_n(c) = 0 = G_n(c)$.
- The Markov transitions M_n are extended into transitions from E_{n-1}^c to E_n^c by setting $M_n^c(c, \cdot) = \delta_c$, and for each $x_{n-1} \in E_{n-1}$,

$$M_n^c(x_{n-1}, dx_n) = M_n(x_{n-1}, dx_n) \; .$$

- Finally, the Markov extension $\mathcal{G}_n^c$ of $\mathcal{G}_n$ is given by

$$\mathcal{G}_n^c(x_n, dy_n) = G_n(x_n)\delta_{x_n}(dy_n) + (1 - G_n(x_n))\delta_c(dy_n) \; .$$

The corresponding Markov chain

$$\left(\Omega^c = \prod_n E_n^c, \mathcal{F}^c = (\mathcal{F}_n^c)_{n\geq 0}, X = (X_n)_{n\geq 0}, \mathbb{P}_\mu^c\right) ,$$

with initial distribution $\mu \in \mathcal{P}(E_0)$ and elementary transitions

$$Q_{n+1}^c = \mathcal{G}_n^c M_{n+1}^c \; , \tag{17}$$

can be regarded as a Markov particle evolving in an environment, with absorbing obstacles related to potential functions G_n . In view of (17), we see that the motion is decomposed into two separate killing/exploration transitions,

$$X_n \xrightarrow{\text{killing}} \widehat{X}_n \xrightarrow{\text{exploration}} X_{n+1}$$

which are defined as follows:

- **Killing:** If $X_n = c$, then we set $\widehat{X}_n = c$. Otherwise the particle X_n is still alive. In this case, we perform the following random choice: With a probability $G(X_n)$, it remains in the same site and we set $\widehat{X}_n = X_n$; and with probability $1 - G_n(X_n)$, it is killed, and we set $\widehat{X}_n = c$.
- **Exploration:** Firstly, when the particle has been killed, we hace $\widehat{X}_n = c$, and we set $X_p = \widehat{X}_p = c$ for any $p > n$. Otherwise, the particle $\widehat{X}_n \in E_n$ evolves to a new location X_{n+1} in E_{n+1} , randomly chosen according to the distribution $M_{n+1}(\widehat{X}_n, \cdot)$.

In this physical interpretation, the Feynman-Kac flows $(\widehat{\eta}_n, \eta_n)$ represent the conditional distributions of a nonabsorbed Markov particle. To see this claim, we denote by T the time at which the particle has been killed

$$T = \inf\{n \geq 0 : \ \widehat{X}_n = c\} \ .$$

By construction, we have

$$\mathbb{P}^c_\mu(T > n) = \mathbb{P}^c_\mu(\widehat{X}_0 \in E_0, \cdots, \widehat{X}_n \in E_n) = \mathbb{E}_\mu\Big(\prod_{p=0}^{n} G_p(X_p)\Big) \ .$$

This shows that the normalized constants of $\widehat{\eta}_n$, and η_n , represent respectively the probability for the particle to be killed at a time strictly greater than or at least equal to n . That is, we have that

$$\widehat{\gamma}_n(1) = \mathbb{P}^c_\mu(T > n) \ \text{ and } \ \gamma_n(1) = \mathbb{P}^c_\mu(T \geq n) \ .$$

Similar arguments yield that

$$\widehat{\gamma}_n(f_n) = \mathbb{E}^c_\mu\left(f_n(X_n)1_{\{T>n\}}\right) \ \text{ and } \ \gamma_n(f_n) = \mathbb{E}^c_\mu\left(f_n(X_n)1_{\{T\geq n\}}\right) \ .$$

Finally, we conlude that

$$\widehat{\eta}_n(f_n) = \mathbb{E}^c_\mu(f_n(X_n)|T > n) \ \text{ and } \ \eta_n(f_n) = \mathbb{E}^c_\mu(f_n(X_n)|T \geq n) \ .$$

The subsets $G_n^{-1}((0,1))$ and $G_n^{-1}(0)$ are called respectively, the sets of soft and hard obstacles (at time n). A particle entering into a hard obstacle is instantly killed; whereas if it enters into a soft obstacle, its lifetime decreases. When the accessibility condition $(\mathcal{A})$ is met, we can replace the mathematical objects (η_0, E_n, G_n, M_n) by $(\widehat{\eta}_0, \widehat{E}_n, \widehat{G}_n, \widehat{M}_n)$. We define in this way a particle motion in an absorbing medium, with no hard obstacles. Loosely speaking, the hard obstacles have been replaced by repulsive obstacles. For instance, in the situation where $G_n = 1_{\widehat{E}_n}$, the Feynman-Kac model associated with (η_0, G_n, M_n) corresponds to a particle motion in an absorbing medium, with pure hard obstacle sets $\widehat{E}_n$; while the Feynman-Kac associated with $(\widehat{\eta}_0, \widehat{G}_n, \widehat{M}_n)$, corresponds to a particle motion in an absorbing medium, with only soft obstacles related to the potential functions $\widehat{G}_n$.

Interacting Process Interpretation

In interacting process literature, Feynman-Kac flows are alternatively interpreted as nonlinear measure-valued process. For instance, the distribution η_n in (14) is regarded as a solution of nonlinear recursive equations. This equation can be rewritten in the following form

$$\eta_{n+1} = \eta_n K_{n+1,\eta_n} \ , \tag{18}$$

where K_{n+1,η_n} is the collection of Markov kernels given by

$$K_{n+1,\eta_n}(x,dz) = S_{n,\eta_n} M_{n+1}(x,dz) = \int_{E_n} S_{n,\eta_n}(x,dy) M_{n+1}(y,dz) \ ,$$

with the selection type transitions

$$S_{n,\eta_n}(x,dy) = G_n(x)\delta_x(dy) + (1 - G_n(x))\Psi_n(\eta_n)(dy) \ .$$

Note that the corresponding evolution equation is now decomposed into two separate transitions

$$\eta_n \xrightarrow{S_{n,\eta_n}} \widehat{\eta}_n = \eta_n S_{n,\eta_n} \xrightarrow{M_{n+1}} \eta_{n+1} = \widehat{\eta}_n M_{n+1} \ , \tag{19}$$

In constrast with the killing interpretation, we have turned the sub-Markovian kernel $\mathcal{G}_n$ into the Markov case in a nonlinear way, by replacing the Dirac measure δ_c , by the Boltzmann-Gibbs jump distribution $\Psi_n(\eta_n)$.

The choice of $K_{n,\eta}$ is not unique. A collection of Markov kernels $K_{n,\eta}$, $\eta \in \mathcal{P}(E_n)$ satisfying the compatibility condition

$$\Phi_n(\eta) = \eta K_{n,\eta}$$

for any $\eta \in \mathcal{P}(E_n)$ is called a McKean interpretation of the flow η_n . In comparaison with (17), the motion of the canonical model $X_n \to X_{n+1}$ associated with the Markov kernels $(K_{n,\eta})_{\eta \in \mathcal{P}(E_n)}$ is the overlapping of an interacting jump, and an exploration transition

$$X_n \xrightarrow{\text{interacting jump}} \widehat{X}_n \xrightarrow{\text{exploration}} X_{n+1} \ .$$

These two mechanisms are defined as follows:

- **Interacting jump:** Given the position, and the distribution η_n at time n of the particle X_n , a jump is performed to a new site $\widehat{X}_n$, randomly chosen according to the distribution

  $$S_{n,\eta_n}(X_n,\cdot) = G_n(X_n)\delta_{X_n} + (1 - G_n(X_n))\Psi_n(\eta_n) \ .$$

  In other words, with a probability $G_n(X_n)$ the particle remains in the same site, and we set $\widehat{X}_n = X_n$. Otherwise, it jumps to a new location, randomly chosen according to the Boltzmann-Gibbs distribution $\Psi_n(\eta_n)$. Notice that particles are attracted by regions with high potential values.
- **Exploration:** The exploration transition coincides with that of the killed particle model. During this stage, the particle evolves to a new site X_{n+1} , randomly chosen according to $M_{n+1}(\widehat{X}_n,\cdot)$.

5 Interacting Particle Systems

The basic idea behind the interacting particle systems is to associate to a given nonlinear dynamical structure, a sequence of E_n^N-valued Markov processes, in such a way that the configuration occupation measures converge, as $N \to \infty$, to the desired distribution. The parameter N represents the precision parameter, as well as the size of the systems. The state components of the E_n^N-valued Markov process are called particles.

5.1 Interacting Particle Interpretations

Hereafter, we suppose the potential functions G_n are bounded and strictly positive (the situation where G_n may take null values can be reduced to this situation, under appropriate accessibility conditions, by replacing η_n by $\widehat{\eta}_n$).

We recall that η_n satisfy the nonlinear recursive equation (18) where the kernels $K_{n,\eta}$ are a combination of a selection and mutation transition

$$K_{n+1,\eta} = S_{n,\eta} M_{n+1} \,. \tag{20}$$

The selection transition $S_{n,\eta}$ on E_n is given by

$$S_{n,\eta_n}(x, dy) = \varepsilon_n G_n(x)\delta_x(dy) + (1 - \varepsilon_n G_n(x))\Psi_n(\eta_n)(dy) \,, \tag{21}$$

where ε_n stands for non negative number such that $\varepsilon_n G_n \leq 1$.

Definition 4. *The interacting particle model associated with a collection of Markov transitions $K_{n,\eta}, \eta \in \mathcal{P}(E_n), n \geq 1$, and with initial distribution η_0, is a sequence of nonhomogeneous Markov chains*

$$\left(\Omega^{(N)} = \prod_{n\geq 0} E_n^N,\ \mathcal{F}^N = (\mathcal{F}_n^N)_{n\geq 0},\ \xi = (\xi_n)_{n\geq 0},\ \mathbb{P}_{\eta_0}^N\right) ,$$

taking values at each time n in the product space E_n^N. That is, we have

$$\xi_n = (\xi_n^1, \cdots, \xi_n^N) \in E_n^N = \underbrace{E_n \times \cdots \times E_n}_{N times} .$$

The initial configuration ξ_0 consists of N independent, and identically distributed random variables, with common law η_0. Its elementary transitions from E_{n-1}^N into E_n^N are given by

$$\mathbb{P}_{\eta_0}^N\Big(\xi_n \in dx_n | \xi_{n-1}\Big) = \prod_{p=1}^{N} K_{n,m(\xi_{n-1})}(\xi_{n-1}^p, dx_n^p) \,,$$

where

$$m(\xi_{n-1}) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_{n-1}^i}$$

is the empirical measure of the configuration ξ_{n-1} of the system, and $dx_n = dx_n^1 \times \cdots \times dx_n^N$ is an infinitesimal neighborhood of a point $x_n = (x_n^1, \cdots, x_n^N) \in E_n^N$.

The N-particle model, associated with the Markov transition $K_{n,\eta}$ given by (20), is the Markov chain ξ_n with elementary transitions

$$\mathbb{P}_{\eta_0}^N\Big(\xi_{n+1} \in dx_{n+1} | \xi_n\Big) = \int_{E_n^N} \mathcal{S}_n(\xi_n, dx_n) \mathcal{M}_{n+1}(x_n, dx_{n+1}) \; .$$

The Boltzmann-Gibbs transition $\mathcal{S}_n$, from E_n^N into itself, and the mutation transition $\mathcal{M}_{n+1}$, from E_n^N into E_{n+1}^N , are defined by the product formulas

$$\mathcal{S}_n(\xi_n, dx_n) = \prod_{p=1}^N S_{n,m(\xi_n)}(\xi_n^p, dx_n^p) \; ,$$

$$\mathcal{M}_{n+1}(x_n, dx_{n+1}) = \prod_{p=1}^N M_{n+1}(x_n^p, dx_{n+1}^p) \; .$$

This integral decomposition shows that (the deterministic) two-step updating/prediction transitions in (19) have been replaced by a two-step selection/mutation transitions (8)

$$\xi_n \in E_n^N \xrightarrow{\text{selection}} \widehat{\xi}_n \in E_n^N \xrightarrow{\text{mutation}} \xi_{n+1} \in E_{n+1}^N \; .$$

In more details, the motion of the particles is defined as follows:

- **Selection:** Given the configuration $\xi_n \in E_n^N$ of the system at time n , the selection transition consists in selecting randomly N particles $\widehat{\xi}_n^i$ with respective distribution $S_{n,m(\xi_n)}(\xi_n^i, \cdot)$. In other words, with a probability $\varepsilon_n G_n(\xi_n^i)$, we set $\widehat{\xi}_n^i = \xi_n^i$; otherwise, we select randomly a particle $\widetilde{\xi}_n^i$ with distribution

$$\Psi_n(m(\xi_n)) = \sum_{i=1}^N \frac{G_n(\xi_n^i)}{\sum_{j=1}^N G_n(\xi_n^j)} \delta_{\xi_n^i}, \text{ and we set } \widehat{\xi}_n^i = \widetilde{\xi}_n^i \; .$$

- **Mutation:** Given the selected configuration $\widehat{\xi}_n \in E_n^N$, the mutation transition consists in sampling randomly N independent random particles ξ_{n+1}^i with respective distributions $M_{n+1}(\widehat{\xi}_n^i, \cdot)$.

5.2 Particle Models with Degenerate Potential

We now discuss the situation where G_n is not necessarily strictly positive. To avoid some complications, we suppose the accessibility condition $(\mathcal{A})$ is met.

Two strategies can be underlined. In view of the discussion given in Sect. 4.1, the first idea is to consider the N-particle approximation model associated with some McKean interpretation of the updated model $\widehat{\eta}_n = \Psi_n(\eta_n)$ which can be regarded as a sequence of measures on $\widehat{E}_n = G_n^{-1}(0,\infty)$. Furthermore, $\widehat{\eta}_n$ coincide with the prediction model starting at $\widehat{\eta}_0$ and associated with the pair of potentials/kernels $(\widehat{G}_n, \widehat{M}_n)$ on the state spaces $\widehat{E}_n$.

The potential function $\widehat{G}_n$ is now a strictly positive function on $\widehat{E}_n$ and the updated model $\widehat{\eta}_n$ satisfies the recursive equation

$$\widehat{\eta}_{n+1} = \widehat{\eta}_n \widehat{K}_{n+1,\widehat{\eta}_n} \quad \text{with} \quad \widehat{K}_{n+1,\eta} = \widehat{S}_{n,\eta}\widehat{M}_{n+1} \ .$$

The selection transitions are now Markov kernels, from $\widehat{E}_n$ into itself, and they are defined for any $x_n \in \widehat{E}_n$ by the formula

$$\widehat{S}_{n,\eta}(x_n, dy_n) = \varepsilon_n \widehat{G}_n(x_n)\delta_{x_n}(dy_n) + (1 - \varepsilon_n \widehat{G}_n(x_n))\widehat{\Psi}_n(\eta)(dy_n) \ .$$

The Boltzmann-Gibbs transformation $\widehat{\Psi}_n$ is given by

$$\widehat{\Psi}_n(\eta)(dx_n) = \frac{1}{\eta(\widehat{G}_n)}\, \widehat{G}_n(x_n)\, \eta(dx_n) \ .$$

In this interpretation, the model $\widehat{\eta}_n$ satisfies the deterministic evolution equation

$$\widehat{\eta}_n \xrightarrow{\text{updating}} \widetilde{\eta}_n = \widehat{\eta}_n \widehat{S}_{n,\widehat{\eta}_n} \xrightarrow{\text{prediction}} \widehat{\eta}_{n+1} = \widetilde{\eta}_n \widehat{M}_{n+1} \ .$$

The N-particle associated with this McKean interpretation is defined as before.

The second strategy consists in still working with the McKean interpretation of the prediction flow associated with the collection of transitions $K_{n+1,\eta} = S_{n,\eta}M_{n+1}$ with $\eta \in \mathcal{P}_n(E_n)$. In this case the particle interpretation given in Definition 4 is not well-defined. Indeed, it may happen that the whole configuration ξ_n moves out of the set $\widehat{E}_n$. To describe rigorously the particle model we proceed as in Sect. 4.2. We add a cemetery point Δ to the product space E_n^N and we extend the test functions and the mutation/selection transitions $(\mathcal{S}_n, \mathcal{M}_n)$ on E_n^N to $E_n^N \cup \{\Delta\}$ as follows:

- The test functions $\varphi_n \in \mathcal{B}_b(E_n^N)$ are extended by setting $\varphi_n(\Delta) = 0$.
- The selection transitions $\mathcal{S}_n$, from E_n^N into itself, are extended into transitions on $E_n^N \cup \{\Delta\}$ by setting $\mathcal{S}_n(x,\cdot) = \delta_\Delta$, as soon as the empirical measure $m(x) \notin \mathcal{P}_n(E_n)$.
- The mutation transitions $\mathcal{M}_{n+1}$ are extended into transitions from $E_n^N \cup \{\Delta\}$ to $E_{n+1}^N \cup \{\Delta\}$ by setting $\mathcal{M}_{n+1}(\Delta,\cdot) = \delta_\Delta$.

The corresponding interacting particle model is a sequence of nonhomogeneous Markov chains, taking values at each time n in $E_n^N \cup \{\Delta\}$. It is defined by a two-step selection/mutation transition of the same nature as before:

$$\xi_n \in E_n^N \cup \{\Delta\} \xrightarrow{\text{selection}} \widehat{\xi}_n \in E_n^N \cup \{\Delta\} \xrightarrow{\text{mutation}} \xi_{n+1} \in E_{n+1}^N \cup \{\Delta\} .$$

The only difference is that the chain is killed at the first time n, we have $m(\xi_n) \notin \mathcal{P}_n(E_n)$. Let τ^N and τ be the dates at which respectively the chain and the Feynman-Kac model are killed:

$$\tau^N = \inf\{n \in \mathbb{N}; m(\xi_n)(G_n) = 0\}, \quad \text{and} \quad \tau = \inf\{n \in \mathbb{N}; \eta_n(G_n) = 0\} .$$

Then it is intuitively clear that $\tau^N \leq \tau$, and in Sect. 6.3 it will be proved that for any $n \leq \tau$ and $N \geq 1$ we have exponential estimate

$$\mathbb{P}_{\eta_0}^N(\tau^N \leq n) \leq a(n) \exp(-N/b(n)) .$$

In particular, this shows that $\lim_{N\to\infty} \mathbb{P}_{\eta_0}^N(\tau^N = \tau) = 1$.

5.3 Application to Particle Analysis of Rare Events

We use the notations and conventions as were introduced in Sects. 2.5 and 3. We recall that $X = (X_n)_{n\in\mathbb{N}}$ is a strong Markov chain taking values in some metric state space (S, d). The process X starts in some Borel set $O \subset S$ with a given probability distribution ν_0. We also consider a pair of Borel subsets (A, R), such that $A_0 \cap R = \emptyset = A \cap R$.

We associate with this pair, the first time T the process hits $A \cup R$, and we let T_R be the hitting time of the set R. We also assume that for any initial $x_0 \in O$, we have $\mathbb{P}_x(T < \infty) = 1$. One would like to estimate the quantities

$$\begin{aligned} \mathbb{P}(T < T_R) &= \mathbb{P}(X_T \in A) , \qquad (22) \\ \text{Law}(X_n; 0 \leq n \leq T | T < T_R) &= \text{Law}(X_n; 0 \leq n \leq T | X_T \in A) . \end{aligned}$$

It often happens that most of the realizations of X never reach the target set A, but are attracted, and absorbed by some non empty set R. These rare events are difficult to analyze numerically. One strategy to estimate these events is to consider the sequence of level-crossing excursions $\mathcal{X}_n$ associated with a splitting of the state space, namely

$$\mathcal{X}_0 = (0, X_0), \quad \text{and} \quad \mathcal{X}_n = (T_n, X_{[T_{n-1}, T_n]}) ,$$

with the entrance times $T_n = \inf\{n \geq 0 : X_n \in B_n \cup R\}$. This sequence forms a Markov chain taking value in the set of excursions $E = \cup_{p\geq 0}(\{p\} \times S^p)$.

One way to check whether or not a random path has succeeded to reach the desired n-th level is to consider the indicator potential functions $\mathcal{G}_n(q, x_{[p,q]}) = 1_{B_n}(x_q)$, with the convention $B_0 = O$. Using elementary calculations, we obtain the following Feynman-Kac representation of the desired quantities (22).

Proposition 1. *For any n and any $f_n \in \mathcal{B}_b(E)$, we have that*

$$\mathbb{E}\left(f_n(\mathcal{X}_0, \cdots, \mathcal{X}_n) \; ; \; T_n < T_R\right) = \mathbb{E}\left(f_n(\mathcal{X}_0, \cdots, \mathcal{X}_n) \prod_{p=0}^{n} \mathcal{G}_p(\mathcal{X}_p)\right) .$$

The prediction Feynman-Kac model $\eta_n \in \mathcal{P}(E)$, defined by

$$\eta_n(f) = \gamma_n(f)/\gamma_n(1) \quad \text{with} \quad \gamma_n(f) = \mathbb{E}\left(f(\mathcal{X}_n)\prod_{p=0}^{n-1}\mathcal{G}_p(\mathcal{X}_p)\right) ,$$

satisfies the measure-valued dynamical system

$$\eta_{n+1} = \Phi_{n+1}(\eta_n) \quad \text{with} \quad \eta_0 = \delta_0 \otimes \nu_0 .$$

The mappings Φ_{n+1}, from $\mathcal{P}_n(E)$ into $\mathcal{P}(E)$, are defined by $\Phi_{n+1}(\eta) = \Psi_n(\eta)\mathcal{M}_{n+1}$, where the Markov kernels $\mathcal{M}_n(u, dv)$ represent the Markov transitions of the chain excursions $\mathcal{X}_n$. We have the following lemma

Lemma 1. *For any $n \geq 0$, we have*

$$\mathcal{P}(T_n < T_R) = \widehat{\gamma}_n(1) = \gamma_{n+1}(1) .$$

In addition, we have $\mathcal{P}(T_n < T_R | T_{n-1} < T_R) = \eta_n(G_n)$, and for any $f \in \mathcal{B}_b(E)$

$$\eta_n(f) = \mathbb{E}\left(f(T_n, X_{[T_{n-1},T_n]})|T_{n-1} < T_R\right) ,$$
$$\widehat{\eta}_n(f) = \mathbb{E}\left(f(T_n, X_{[T_{n-1},T_n]})|T_n < T_R\right) .$$

This lemma gives a Feynman-Kac interpretation of rare events probabilities. Since the potentials are indicator functions, it is more judicious to rewrite the Boltzmann-Gibbs transformations $\Psi_n(\eta) = \eta S_{n,\eta}$ in terms of the selection Markov transitions

$$S_{n,\eta}(u, dv) = (1 - 1_{\{\mathcal{G}_n^{-1}(1)\}}(u))\Psi_n(\eta)(dv) + 1_{\{\mathcal{G}_n^{-1}(1)\}}(u)\delta_u(dv) .$$

Note that $\mathcal{G}_n^{-1}(1)$ represents the collection of excursions in S entering the nth level B_n; that is, we have that

$$\mathcal{G}_n^{-1}(1) = \{u = (q, x_{[p,q]}) \in E;\ x_q \in B_n\} .$$

The particle interpretation of these discrete Feynman-Kac model is simply derived from Sect. 5.2. In this context, the particle model consists in evolving a collection of N-excursion valued particles

$$\xi_n^i = (T_n^i, X^i_{[T^i_{n-1},T^i_n]}) \in E \cup \{\Delta\} ,$$
$$\widehat{\xi}_n^i = (\widehat{T}_n^i, \widehat{X}^i_{[\widehat{T}^i_{n-1},\widehat{T}^i_n]}) \in E \cup \{\Delta\} .$$

The auxiliary point Δ stands for a cemetery point, the random time pairs (T^i_{n-1}, T^i_n) and $(\widehat{T}^i_{n-1}, \widehat{T}^i_n)$ represent the length of the corresponding excursions. At the time $n = 0$, the initial system consists of N independent, and identically distributed, S-valued random variables $\xi_0^i = (0, X_0^i)$, with common

law $\eta_0 = \delta_0 \otimes \nu_0$. Since we have $\mathcal{G}_0(0,u) = 1$, there is no updating transition at time $n = 0$, and we set $\widehat{\xi}_0^i = \xi_0^i$, for each $1 \leq i \leq N$.

Mutation: The mutation stage $\widehat{\xi}_n \to \xi_{n+1}$ at time $n+1$ is defined as follows. If $\widehat{\xi}_n = \Delta$, we set $\xi_{n+1} = \Delta$. Otherwise, during the mutation, each selected excursion $\widehat{\xi}_n^i$ evolves randomly, and independently of each other, according to the Markov transition $\mathcal{M}_{n+1}$ of the chain $\mathcal{X}_n$. Thus, ξ_{n+1}^i is a random variable with distribution $\mathcal{M}_{n+1}(\widehat{\xi}_n^i, \cdot)$. More precisely, we set $T_n^i = \widehat{T}_n^i$, and the particle $\widehat{X}^i_{[\widehat{T}^i_{n-1}, \widehat{T}^i_n]}$ evolves randomly as a copy of the excursion process $(X_s)_{s \geq T_n^i}$ starting at $X_{T_n^i}$, and up to the first time T_{n+1}^i it visits B_{n+1}, or returns to R. The stopping time T_{n+1}^i represents the first time $t \geq T_n^i$ the ith excursion hits the set $B_{n+1} \cup R$.

Selection: The selection mechanism $\xi_{n+1} \to \widehat{\xi}_{n+1}$ is defined as follows. In the mutation stage, we have sampled N excursions ξ_{n+1}^i. Some of these particles have succeeded to reach the desired set B_{n+1}, and the other ones have entered into R. We denote by $I^N(n+1)$ the set of the labels of the particles having reached the $(n+1)$-th level, and we set $m(\xi_{n+1}) = N^{-1} \sum_{i=1}^N \delta_{(\xi_{n+1}^i)}$. Two situations may occur. If $I^N(n+1) = \emptyset$ then none of the particles have succeeded to hit the desired level. In this situation, we have $m(\xi_{n+1}) \notin \mathcal{P}_{n+1}(E)$, and the algorithm has to be stopped. In this case, we set $\widehat{\xi}_{n+1} = \Delta$. Otherwise, the selection transition is defined as follows. Each particle $\widehat{\xi}_{n+1}$ is sampled according to the selection distribution

$$
\begin{aligned}
&S_{n,m(\xi_{n+1})}(\xi_{n+1}^i, dv) \\
&= 1_{B_{n+1}}(X^i_{T^i_{n+1}}) \delta_{\xi^i_{n+1}}(dv) + 1_{B^C_{n+1}}(X^i_{T^i_{n+1}}) \Psi_n(m(\xi_{n+1}))(dv) \; .
\end{aligned}
$$

More precisely, if the i-th excursion has reached the desired level, then we set $\widehat{\xi}_{n+1}^i = \xi_{n+1}^i$. In the opposite case, the particle has not reached the $(n+1)$-th level, but it has visited the set R. In this case, $\widehat{\xi}_{n+1}^i$ is chosen randomly and uniformly in the set $\{\xi_{n+1}^j ; j \in I^N(n+1)\}$ of all excursions having entered into B_{n+1}. In other words, each particle that doesn't enter into the $(n+1)$-th level is killed, and instantly a different particle in the B_{n+1} level splits into two offsprings.

For each time $n < \tau^N = \inf\{n \geq 0 \; : \; X^i_{T^i_n} \in R, 1 \leq i \leq n\}$, the N-particle approximation measures $(\gamma_n^N, \eta_n^N, \widehat{\eta}_n^N)$ associated with $(\gamma_n, \eta_n, \widehat{\eta}_n)$ are defined by

$$\widehat{\gamma}_n^N(1) = \gamma_n^N(\mathcal{G}_n) = N^{-n} \prod_{p=1}^{n} \mathrm{Card}(I^N(p)) \ ,$$

$$\eta_n^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\xi_n^i} \ ,$$

$$\widehat{\eta}_n^N = \Psi_n(\eta_n^N) = \frac{1}{\mathrm{Card}(I^N(n))} \sum_{i \in I^N(n)} \delta_{(T_n^i, X^i_{[T^i_{n-1}, T^i_n]})} \ .$$

Thus, $\widehat{\gamma}_n^N(1)$ is the proportion product of excursions having entered levels $B_1, \cdots, B_n$. Also notice that $\widehat{\eta}_n^N$ is the occupation measure of the excursions entering the nth level.

The asymptotic analysis of these particles measures will be discussed in the following sections. We will prove the following results (see notation (10)):

Theorem 1. *For any $n \geq 0$ and $N \geq 1$ we have*

$$\mathbb{P}(\tau^N \leq n) \leq a(n) \exp(-N/b(n)) \ .$$

The particle estimates are unbiased, $\mathbb{E}(\widehat{\gamma}_n^N(1) 1_{\{n<\tau^N\}}) = \mathbb{P}(T_n < T_R)$, and for any $p \geq 1$, and any $n \geq 0$, we have

$$\sqrt{N}\ \mathbb{E}\left(|\widehat{\gamma}_n^N(1) 1_{\{n<\tau^N\}} - \mathbb{P}(T_n < T_R)|^p\right)^{1/p} \leq a(p) b(n) \ ,$$

for some finite constants $a(p)$, $b(n) < \infty$ whose values only depend respectively on the parameters p and n.

In addition, for any $0 \leq n \leq m+1$, the sequence of random variables

$$W_{n+1}^N = \sqrt{N}(\gamma_n^N(1) 1_{\{\tau^N > n\}} - \mathbb{P}(T_n < T_R))$$

converges in law (as N tends to ∞) to a centered Gaussian random variable W_{n+1} with variance

$$\sigma_n^2 = \sum_{q=0}^{n+1} (\gamma_q(1))^2 \eta_{q-1}(K_{q,\eta_{q-1}}[Q_{q,n}(1) - K_{q,\eta_{q-1}} Q_{q,n}(1)]^2) \ .$$

The collection of functions $Q_{q,n+1}(1)$ on the excursion space E are defined for any $x = (x_n)_{s \leq n \leq t}$ by

$$Q_{q,n+1}(1)(t,x) = 1_{B_q}(x_t) \mathbb{P}(T_n < T_R | T_q = t, X_{T_q} = x_t) \ .$$

Example 1. When the set $S = \mathbb{R}^d$ is the Euclidean space, we can think of a sequence of centered decreasing balls with radius $1/(n+1)$

$$B_n = \mathcal{B}(0, \frac{1}{n+1}) \quad \text{and} \quad R = S \setminus \mathcal{B}(0, 1+\varepsilon)$$

for some $\varepsilon > 0$. Further assume that the process X exits the ball of radius $1+\varepsilon$ in finite time. In this situation, $\mathbb{P}(T < T_R)$ is the probability that X hits the smallest ball B_m, starting with $1/2 < |X_0| \leq 1$, and before exiting the ball of radius $1+\varepsilon$. The distribution (22) represents the conditional distribution of the process X in this ballistic regime (see Fig. 10).

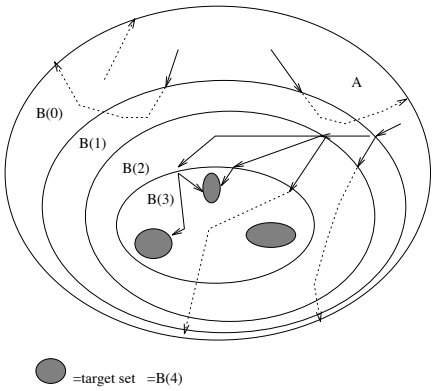

Fig. 10. Ballistic regime, target $B(4)$ with $N = 4$

6 Asymptotic Behavior

This section is concerned with the asymptotic behavior of particle approximation models, as the size of the systems tends to infinity. The principal convergence results are the following. Firstly, γ_n^N is an unbiaised estimator; that is, we have for any $f_n \in \mathcal{B}_b(E_n)$

$$\mathbb{E}_{\eta_0}^N(\gamma_n^N(f_n)1_{\{\tau_n^N \geq n\}}) = \gamma_n(f_n) \ .$$

Furthermore, we have the $\mathbb{L}_p$-estimates

$$\sqrt{N}\ \mathbb{E}_{\eta_0}^N[|\eta_n^N(f_n) - \eta_n(f_n)|^p]^{1/p} \leq a(p)b(n)\|f\| \ ,$$

which can be extended to a countable collection of uniformly bounded functions $\mathcal{F}_n \subset \mathcal{B}_b(E_n)$,

$$\sqrt{N}\ \mathbb{E}_{\eta_0}^N\left[\sup_{f_n \in \mathcal{F}_n} |\eta_n^N(f_n) - \eta_n(f_n)|^p\right]^{1/p} \leq a(p)b(n)I(\mathcal{F}_n) \ ,$$

for some finite constant $I(\mathcal{F}_n) < \infty$ that only depends on the class $\mathcal{F}_n$. Similar but exponential type estimates will be also covered. By instance, we have for any $\varepsilon > 0$ and N sufficiently large

$$\mathbb{P}_{\eta_0}^N\left[\sup_{f_n \in \mathcal{F}_n} |\eta_n^N(f_n) - \eta_n(f_n)| > \varepsilon\right] \leq d_n(\varepsilon, \mathcal{F}_n)\mathrm{e}^{-N\varepsilon^2/b(n)} \ ,$$

with a finite constant $d(\varepsilon, \mathcal{F}_n)$ depending on ε and the class $\mathcal{F}_n$. From these estimates and using the Borel-Cantelli lemma, we conclude the almost sure convergence result

$$\lim_{N\to\infty} \sup_{f_n \in \mathcal{F}_n} |\eta_n^N(f_n) - \eta_n(f_n)| = 0 \ .$$

The corresponding fluctuations and Central Limits Theorems will also be discussed in Sect. 6.5, in which the following result will be proved: For any $n \geq 0$, and $f \in \mathcal{B}_b(E_n)$, the sequence of random variables

$$W_n^N(f) = \sqrt{N}(\gamma_n^N(f_n)1_{\{\tau^N \geq n\}} - \gamma_n(f_n))$$

converges in law (as N tends to ∞) to a centered Gaussian random variable $W_n(f)$ with variance

$$\sigma_n^2(f) = \sum_{q=0}^{n} \gamma_q(1)^2 \eta_{q-1} \left(K_{q,\eta_{q-1}}[Q_{q,n}(f) - K_{q,\eta_{q-1}} Q_{q,n}(f)]^2 \right) ,$$

where $Q_{p,n}(f)$ are some functions defined hereafter. We use the convention $\eta_{-1} = \eta_0 = K_{0,\eta_{-1}}$. Rephrasing these asymptotic results in the context of analysis of rare events leads to the Theorem 1.

6.1 Preliminaries

Feynman-Kac Semigroups

In this short section, we introduce the Feynman-Kac semigroups, $Q_{p,n}$ and $\Phi_{p,n}$, associated respectively with γ_n and η_n. They are defined by the formulas

$$Q_{p,n} = Q_{p+1} \cdots Q_{n-1} Q_n , \quad \text{and} \quad \Phi_{p,n} = \Phi_n \circ \Phi_{n-1} \circ \ldots \circ \Phi_{p+1} ,$$

with $Q_n(x_{n-1}, dx_n) = G_{n-1}(x_{n-1}) M_n(x_{n-1}, dx_n)$. We use the convention $Q_{n,n} = Id$ and $\Phi_{n,n} = Id$. These semigroups are alternatively defined by

$$Q_{p,n}(f_n)(x_p) = \mathbb{E}_{p,x_p} \left(f_n(X_n) \prod_{q=p}^{n-1} G_q(X_p) \right), \quad \Phi_{p,n}(\mu_p)(f_n) = \frac{\mu_p(Q_{p,n}(f_n))}{\mu_p(Q_{p,n}(1))} ,$$

where $\mathbb{E}_{p,x_p}$ is the expectation with respect the law of the shifted chain $(X_{p+n})_{n \geq 0}$. By definition of η_n and $Q_{p,n}$, we observe that

$$\eta_n(f_n) = \frac{\eta_p(Q_{p,n}(f_n))}{\eta_p(Q_{p,n}(1))} , \qquad \gamma_p(Q_{p,n}(1)) = \gamma_n(1) . \tag{23}$$

Now, introducing the pair potential/transition $(G_{p,n}, P_{p,n})$ defined by

$$G_{p,n} = Q_{p,n}(1) \quad \text{and} \quad P_{p,n}(f_n) = \frac{Q_{p,n}(f_n)}{Q_{p,n}(1)} ,$$

we deduce the following formula for the semigroup $\Phi_{p,n}$

$$\Phi_{p,n}(\mu_p) = \Psi_{p,n}(\mu_p) P_{p,n} ,$$

with the Boltzmann-Gibbs transformation, $\Psi_{p,n}$ from $\mathcal{E}_p$ into itself, defined by

$$\Psi_{p,n}(\mu_p)(f_p) = \mu_p(G_{p,n}(f_n))/\mu_p(G_{p,n}(1)) .$$

Some Inequalities for Independent Random Variables

In this section, we discuss some general inequalities for sequences of independent variables. These inequalities will be used in the following sections.

Let $(\mu_i)_{i\geq 1}$ be a sequence of probability measures on a given measurable state space $(E,\mathcal{E})$. We also consider a sequence of $\mathcal{E}$-measurable functions $(h_i)_{i\geq 1}$ such that $\mu_i(h_i)=0$, for all $i\geq 1$. During the further development of this section we fix an integer $N\geq 1$. To clarify the presentation we slight abuse the notation and we denote respectively by

$$m(X)=\frac{1}{N}\sum_{i=1}^{N}\delta_{X^i} \quad \text{and} \quad \mu=\frac{1}{N}\sum_{i=1}^{N}\mu_i\ ,$$

the N-empirical measure associated to a collection of independent random variables $X=(X^i)_{i\geq 1}$, with respective distributions $(\mu_i)_{i\geq 1}$ and the N-averaged measure associated to the sequence of measures $(\mu_i)_{i\geq 1}$. When we are given N-sequences of points $x=(x^i)_{1\leq i\leq N}\in E^N$ and functions $(h_i)_{1\leq i\leq N}\in\mathcal{B}_b(E)^N$ we shall also use the following notations

$$m(x)(h)=\frac{1}{N}\sum_{i=1}^{N}h_i(x^i) \quad \text{and} \quad \sigma^2(h)=\frac{1}{N}\sum_{i=1}^{N}\mathrm{osc}^2(h_i)\ ,$$

where $\mathrm{osc}(h)=\sup\{|h(x)-h(y)|\}$ is the oscillation of the function h.

For any pair of integers (p,n), with $1\leq p\leq n$, we denote by $(n)_p$ the quantity

$$(n)_p=\frac{n!}{(n-p)!}\ .$$

We have the following lemmas [2][§7.3]:

Lemma 2 (Chernov-Hoeffding).

$$\mathbb{P}\left(|m(X)(h)|\geq\varepsilon\right)\leq 2\mathrm{e}^{-2N\varepsilon^2/\sigma^2(h)}\ .$$

Lemma 3. *For any sequence of $\mathcal{E}$-measurable functions $(h_i)_{i\geq 1}$ such that $\mu_i(h_i)=0$ and $\sigma(h)<\infty$ we have for any $p\geq 1$*

$$\sqrt{N}\ \mathbb{E}(|m(X)(h)|^p)^{\frac{1}{p}}\leq d(p)^{\frac{1}{p}}\ \sigma(h)\ , \tag{24}$$

with the sequence of finite constants $(d(n))_{n\geq 0}$ defined, for any $n\geq 1$, by the formulas

$$d(2n)=(2n)_n\ 2^{-n} \quad \text{and} \quad d(2n-1)=\frac{(2n-1)_n}{\sqrt{n-1/2}}\ 2^{-(n-1/2)}\ . \tag{25}$$

In addition we have for any $\varepsilon>0$

$$\mathbb{E}(\exp\left(\varepsilon\sqrt{N}|m(X)(h)|\right))\leq(1+\varepsilon\sigma(h)/\sqrt{2})\ \exp\left(\varepsilon^2\sigma^2(h)/2\right)\ .$$

We now extend the previous results to the convergence of empirical processes with respect to some Zolotarev seminorm. Let $\mathcal{F}$ be a given collection of measurable functions $f : E \to \mathbb{R}$ such that $\|f\| = \sup_{x\in E} |f(x)| \leq 1$. We associate with $\mathcal{F}$ the Zolotarev seminorm on $\mathcal{P}(E)$ defined by

$$\|\mu - \nu\|_{\mathcal{F}} = \sup\{|\mu(f) - \nu(f)| \ : \ f \in \mathcal{F}\} \ .$$

No generality is lost and much convenience is gained by supposing that the unit constant function $f = 1 \in \mathcal{F}$. Furthermore, we shall suppose that $\mathcal{F}$ contains a countable and dense subset.

To measure the size of a given class $\mathcal{F}$, one considers the covering numbers $\mathcal{N}(\varepsilon, \mathcal{F}, L_p(\mu))$ defined as the minimal number of $L_p(\mu)$-balls of radius $\varepsilon > 0$ needed to cover $\mathcal{F}$. By $\mathcal{N}(\varepsilon, \mathcal{F})$ and by $I(\mathcal{F})$ we denote the uniform covering numbers and entropy integral given by

$$\mathcal{N}(\varepsilon, \mathcal{F}) = \sup\{\mathcal{N}(\varepsilon, \mathcal{F}, L_2(\eta)); \eta \in \mathcal{P}(E)\} \ ,$$
$$I(\mathcal{F}) = \int_0^1 \sqrt{\log(1 + \mathcal{N}(\varepsilon, \mathcal{F}))} d\varepsilon \ .$$

For more details and various examples the reader is invited to consult [14]. We have the following lemma [2][§7.3]:

Lemma 4. *For any $p \geq 1$, we have*

$$\sqrt{N}\, \mathbb{E}\left(\|m(X) - \mu\|_{\mathcal{F}}^p\right)^{1/p} \leq c\lfloor p/2 \rfloor !\, I(\mathcal{F}) \ ,$$

where c is a universal constant.

For any $\varepsilon > 0$ and $\sqrt{N} \geq 4\varepsilon^{-1}$, we have that

$$\mathbb{P}\left(\|m(X) - \mu\|_{\mathcal{F}} > 8\varepsilon\right) \leq 8\mathcal{N}(\varepsilon, \mathcal{F}) \mathrm{e}^{-N\varepsilon^2/2} \ .$$

6.2 Strong Law of Large Numbers

In the following picture, we have illustrated the random evolution of the N-particle approximation model:

$$\begin{array}{ccccccccc}
\eta_0 & \to \eta_1 = \Phi_1(\eta_0) & \to \eta_2 = \Phi_{0,2}(\eta_0) & \to & \cdots & \to \eta_n = \Phi_{0,n}(\eta_0) \\
\Downarrow & & & & & \\
\eta_0^N \to & \Phi_1(\eta_0^N) & \to \Phi_{0,2}(\eta_0^N) & \to & \cdots & \to \Phi_{0,n}(\eta_0^N) \\
& \Downarrow & & & & \\
& \eta_1^N & \to \Phi_2(\eta_0^N) & \to & \cdots & \to \Phi_{1,n}(\eta_1^N) \\
& & \Downarrow & & & \\
& & \eta_2^N & \to & \cdots & \to \Phi_{2,n}(\eta_2^N) \\
& & & \Downarrow & & \vdots \\
& & & \eta_{n-1}^N & \to \Phi_{n-1,n}(\eta_{n-1}^N) \\
& & & & & \Downarrow \\
& & & & & \eta_n^N
\end{array}$$

In this picture, the sampling errors are represented by the implication sign "⇓". Using the identity $\Phi_{q-1,n}(\eta_{q-1}^N) = \Phi_{q,n}(\Phi_q(\eta_{q-1}^N))$, we observe that

$$\eta_n^N - \eta_n = \sum_{q=0}^{n} \left[\Phi_{q,n}(\eta_q^N) - \Phi_{q,n}(\Phi_q(\eta_{q-1}^N))\right] , \tag{26}$$

with the convention $\Phi_0(\eta_{-1}^N) = \eta_0$. Note that each term on the r.h.s. represents the propagation of the pth sampling local error $\Phi_q(\eta_{q-1}^N) \Rightarrow \eta_q^N$. This pivotal formula will be of important use in the following. In addition, we have for each $\eta_1, \eta_2 \in \mathcal{P}(E_q)$ and $f \in \mathcal{B}_b(E_n)$

$$\Phi_{q,n}(\eta_1)(f) - \Phi_{q,n}(\eta_2)(f) = \frac{1}{\eta_2(G_{q,n})}[(\eta_1(Q_{q,n}(f)) - \eta_2(Q_{q,n}(f))) + \Phi_{q,n}(\eta_1)(f)(\eta_2(G_{q,n}) - \eta_1(G_{q,n}))] .$$

We deduce the following formula which highlights the sampling errors:

$$\eta_n^N(f) - \eta_n(f) = \sum_{q=0}^{n} \frac{1}{\eta_{q-1}^N(G_{q,n})}[(\eta_q^N(Q_{q,n}(f)) - \Phi_q(\eta_{q-1}^N)(Q_{q,n}(f))) + \Phi_{q,n}(\eta_q^N)(f)(\Phi_q(\eta_{q-1}^N)(G_{q,n}) - \eta_q^N(G_{q,n}))] . \tag{27}$$

6.3 Extinction Probabilities

The objective of this short section is to estimate the probability of extinction of a class of particle models, associated with bounded (by one) potential functions that may take null values. Let us recall that the limiting flow η_n is well-defined, only up to the first time τ we have $\eta_\tau(G_\tau) = 0$; that is

$$\tau = \inf\{n \in \mathbb{N} \,:\, \eta_n(G_n) = 0\} = \inf\{n \in \mathbb{N} \,:\, \gamma_{n+1} = 0\} .$$

In the same way, the N-interacting particle systems are only defined up to the time τ^N the whole configuration $\xi_n \in E_n^N$ first hits the hard obstacle set $(E_n \setminus \widehat{E}_n)^N$:

$$\tau^N = \inf\{n \in \mathbb{N} \,:\, \eta_n^N(G_n) = 0\}.$$

It follows the equivalence $(\tau^N \geq n) \Leftrightarrow (\xi_0 \in \widehat{E}_0, \cdots, \xi_{n-1} \in \widehat{E}_{n-1})$, which indicates that τ^N is a predictable Markov time with respect to the filtration $(\mathcal{F}_n^N)$, in the sense that $\{\tau^N \geq n\} \in \mathcal{F}_{n-1}^N$. We have the following rather crude but reassuring result [2][Theorem 7.4.1]

Theorem 2. *Suppose we have $\gamma_n(1) > 0$ for any $n \geq 0$. Then, for any $N \geq 1$ and $n \geq 0$, we have the estimate*

$$\mathbb{P}(\tau^N \leq n) \leq a(n)\mathrm{e}^{-N/b(n)} ,$$

for some constants $a(n)$ and $b(n)$ which depend only on n and $\gamma_{n+1}(1)$.

For a detailed proof, the reader is referred to [2][§7.4]. Its key idea is based on the following observation. Using formula (23), we obtain for any $p \leq n$,

$$\eta_n(G_n) = \frac{\eta_p(G_{p,n+1})}{\eta_p(G_{p,n})} = \frac{\gamma_{n+1}(1)}{\gamma_n(1)} .$$

Now, referring to the setting of Theorem 2, we obtain that $\eta_q(G_q) > 0$ for any $1 \leq q \leq n$, and therefore that $\tau > n$. In fact, assuming the condition $\gamma_n(1) > 0$ for all n, avoids the tunneling problems with probability one, so an exponential decrease of the extinction probabilities.

6.4 Convergence of Empirical Processes

This section provides precise estimates on the convergence of the particle density profiles when the size of the system tends to infinity. We start with the analysis of the unnormalized particles models and we show that this approximation particle has no bias. The central idea consists in expressing the difference between the particle measures and the limiting Feynman-Kac ones as such end values of martingale sequence.

We recall that a square integrable and $\mathcal{F}^N$-martingale $M^N = (M_n^N)_{n\geq 0}$ is an $\mathcal{F}^N$-adapted sequence such that $\mathbb{E}(M_n^N)^2 < \infty$ for all $n \geq 0$ and

$$\mathbb{E}(M_{n+1}^N | \mathcal{F}_n^N) = M_n^N \quad (\mathbb{P}^N - \text{a.s.}) .$$

The predictable quadratic characteristic of M^N is the sequence of random variables $\langle M^N \rangle = (\langle M^N \rangle_n)_{n\geq 0}$ defined by

$$\langle M^N \rangle_n = \sum_{p=0}^{n} \mathbb{E}((M_p^N - M_{p-1}^N)^2 | \mathcal{F}_{p-1}^N) ,$$

with the convention $\mathbb{E}((M_0^N - M_{-1}^N)^2 | \mathcal{F}_{-1}^N) = \mathbb{E}(M_0^N)^2$. The stochastic process $\langle M^N \rangle$ is also called the angle bracket of M^N and is the unique predictable increasing process such that the sequence $((M_n^N)^2 - \langle M^N \rangle_n)_{n\geq 0}$ is an $\mathcal{F}^N$-martingale.

In the following, we will use the simplified notation (10). For instance, if we consider the McKean model

$$K_{n,\eta}(x, \cdot) = G_{n-1}(x) M_n(x, \cdot) + (1 - G_{n-1}(x)) \Phi_n(\eta) , \tag{28}$$

we first observe that

$$K_{q,\eta}(\varphi - \Phi_q(\varphi)) = K_{q,\eta}(\varphi) - \Phi_q(\eta)(\varphi) = G_{q-1}(M_q(\varphi) - \Phi_q(\eta)(\varphi)) .$$

So, let $\tilde{\varphi}_q$ be the function defined by $\tilde{\varphi}_q = \varphi - \Phi_q(\eta)(\varphi)$. We obtain

$$\begin{aligned} K_{q,\eta}[\varphi - K_{q,\eta}(\varphi)]^2 &= K_{q,\eta}[\tilde{\varphi}_q - K_{q,\eta}(\tilde{\varphi}_q)]^2 \\ &= K_{q,\eta}(\tilde{\varphi}_q)^2 - (K_{q,\eta}(\tilde{\varphi}_q))^2 \\ &= K_{q,\eta}[\varphi - \Phi_q(\eta)(\varphi)]^2 - G_{q-1}^2[M_q(\varphi) - \Phi_q(\eta)(\varphi)]^2 \ . \end{aligned} \tag{29}$$

Furthermore, if we consider the McKean model

$$K_{n,\eta}(x,\cdot) = \Phi_n(\eta)(\cdot) \ , \tag{30}$$

we obtain

$$K_{q,\eta}[\varphi - K_{q,\eta}(\varphi)]^2 = \Phi_q(\eta)[\varphi - \Phi_q(\eta)(\varphi)]^2 \ . \tag{31}$$

These two formulas indicate that the particle model in the first case is more accurate than the other one.

Proposition 2. *For each $n \geq 0$ and $f_n \in \mathcal{B}_b(E_n)$, we let $\Gamma_{\cdot,n}^N(f_n)$ be the $\mathbb{R}$-valued process defined for any $p \in \{0, \cdots, n\}$ by*

$$\Gamma_{p,n}^N(f_n) = \gamma_p^N(Q_{p,n}f_n)1_{\{\tau^N \geq p\}} - \gamma_p(Q_{p,n}f_n) \ . \tag{32}$$

For any $p \leq n$, $\Gamma_{\cdot,n}^N(f_n)$ has the $\mathcal{F}^N$-martingale decomposition

$$\Gamma_{p,n}^N(f_n) = \sum_{q=0}^{p} \gamma_q^N(1)1_{\{\tau^N \geq p\}} \left[\eta_q^N(Q_{q,n}f_n) - \eta_{q-1}^N K_{q,\eta_{q-1}^N}(Q_{q,n}f_n)\right] \ , \tag{33}$$

and its bracket is given by

$$\langle \Gamma_{\cdot,n}^N(f_n)\rangle_p = \frac{1}{N}\sum_{q=0}^{p} (\gamma_q^N(1))^2 \ 1_{\{\tau^N \geq p\}} \eta_{q-1}^N \left(K_{q,\eta_{q-1}^N}\left[Q_{q,n}f_n - K_{q,\eta_{q-1}^N}Q_{q,n}f_n\right]^2\right) \ ,$$

with the convention $\Phi_0(\eta_{-1}^N) = \eta_0 = K_{0,\eta_{-1}^N}$.

The first consequence of Proposition 2 is that γ_n^N is unbiased. More precisely, using the martingale decomposition (33) with $p = n$, we obtain for any $f \in \mathcal{F}_n$ the following identity

$$\mathbb{E}(\gamma_n^N(f)1_{\{\tau^N \geq p\}}) = \gamma_n(f) \ .$$

In fact, we have the more precise result [2][Theorem 7.4.2]

Theorem 3. *For each $p \geq 1, n \in \mathbb{N}$, and for any (separable) collection $\mathcal{F}_n$ of measurable functions $f : E_n \to \mathbb{R}$ such that $\|f\| \leq 1$ (and $1 \in \mathcal{F}_n$), we have for any $f \in \mathcal{F}_n$*

$$\mathbb{E}(\gamma_n^N(f)1_{\{\tau^N \geq p\}}) = \gamma_n(f) \ ,$$

and for any $r \leq n$

$$\sqrt{N}\,\mathbb{E}(\|1_{\{\tau^N\geq r\}}\gamma_r^N Q_{r,n}-\gamma_r Q_{r,n}\|_{\mathcal{F}_n}^p)^{1/p}\leq c(n+1)\lfloor p/2\rfloor! I(\mathcal{F}_n)\;.$$

In addition, for any $\varepsilon\geq 4/\sqrt{N}$, *we have the exponential estimate*

$$\mathbb{P}\left(\|1_{\{\tau^N\geq r\}}\gamma_r^N Q_{r,n}-\gamma_r Q_{r,n}\|_{\mathcal{F}_n}>\varepsilon\right)\leq 8(n+1)\mathcal{N}(\varepsilon_n,\mathcal{F}_n)\mathrm{e}^{-N\varepsilon_n^2/2}\;, \quad (34)$$

with $\varepsilon_n=\varepsilon/(n+1)$.

Applying the exponential estimate (34) with $r=n$ and $\varepsilon=\gamma_n(1)/2$, we obtain, for any pair (n,N) such that $\sqrt{N}\geq 8/\gamma_n(1)$, the following inequality

$$\mathbb{P}\left(1_{\{\tau^N\geq r\}}\gamma_n^N(1)\geq\gamma_n(1)/2\right)\geq 1-8(n+1)\mathcal{N}(\varepsilon_n,\mathcal{F}_n)\mathrm{e}^{-N\varepsilon_n^2/2}\;,$$

with $\varepsilon_n=\gamma_n(1)/(2(n+1))$. Now, to obtain some exponential estimate for the measure η_n^N, we use the following decomposition

$$(\eta_n^N(f)-\eta_n(f))1_{\{\tau^N\geq n\}}=\frac{\gamma_n(1)}{\gamma_n^N(1)}\,\gamma_n^N\left(\frac{1}{\gamma_n(1)}(f-\eta_n(f))\right)1_{\{\tau^N\geq n\}}\;. \quad (35)$$

If we set $f_n=\frac{1}{\gamma_n(1)}(f-\eta_n(f))$, then since $\gamma_n(f_n)=0$, (35) also reads

$$\begin{aligned}(\eta_n^N(f)-\eta_n(f))1_{\{\tau^N\geq n\}}&=\frac{\gamma_n(1)}{\gamma_n^N(1)}(\gamma_n^N(f_n)1_{\{\tau^N\geq n\}}-\gamma_n(f_n))\\&=\frac{\gamma_n(1)}{\gamma_n^N(1)}\Gamma_{n,n}^N(f_n)\;. \quad (36)\end{aligned}$$

Let Ω_n^N be the set of events

$$\Omega_n^N=\{\gamma_n^N(1)1_{\{\tau^N\geq n\}}\geq\gamma_n(1)/2\}\subset\{\tau^N\geq n\}\;.$$

Using Theorem 3, we have

$$\mathbb{P}(\Omega_n^N)\geq 1-\frac{b(n)^2}{N}\;,$$

where $b(n)$ is a constant which depends on n only. If we combine this estimate with Theorem 3 and (36), we find that for any $f\in\mathcal{B}_b(E_n)$, with $\|f\|\leq 1$

$$\begin{aligned}|\mathbb{E}\left((\eta_n^N(f)-\eta_n(f))1_{\{\tau^N\geq n\}}\right)|&\leq|\mathbb{E}\left((\eta_n^N(f)-\eta_n(f))1_{\Omega_n^N}\right)|+2\mathbb{P}((\Omega_n^N)^2)\\&\leq\frac{b(n)^2}{N}\;,\end{aligned}$$

where $b(n)$ is a new constant which depends on n only. Finally by Theorem 2, we conclude that

$$|\mathbb{E}\left((\eta_n^N(f)1_{\{\tau^N\geq n\}}-\eta_n(f))\right)|\leq\frac{b(n)^2}{N}+a(n)\mathrm{e}^{-N/b(n)}\;.$$

A consequence of this result is the following extension of the Glivenko-Cantelli theorem to particle models.

Corollary 1. *Let $\mathcal{F}_n$ be a countable collection of functions f such that $\|f\| \leq 1$ and $\mathcal{N}(\varepsilon, \mathcal{F}_n) < \infty$ for any $\varepsilon > 0$. Then, for any time $n \geq 0$, $\|\eta_n^N(f)1_{\{\tau^N \geq n\}} - \eta_n(f)\|_{\mathcal{F}_n}$ converges almost surely to 0 as $N \to \infty$.*

Some time-uniform estimates can also be obtained when the pair (G_n, M_n) satisfies some regularity conditions. When these conditions are met the nonlinear Feynman-Kac semigroup $\Phi_{p,n}$ has asymptotic stability properties which ensure that in some sense for each elementary term

$$[\Phi_{q,n}(\eta_n^N) - \Phi_{q,n}(\Phi_q(\eta_{q_1}^N))] \to 0 \quad \text{as} \quad (n-q) \to \infty \, .$$

Consequently, according to (26), a uniform estimate of the sum of the "small errors" can be proved. The reader is invited to consult [2][§7.4] for more details about this subject.

6.5 Central Limit Theorems

Let us consider the particle approximation model $\xi_n = (\xi_n^i)_{1 \leq i \leq N}$ associated with a nonlinear measure-valued equation of the form

$$\eta_n = \eta_{n-1} K_{n,\eta_{n-1}} \, . \tag{37}$$

We will assume that $\gamma_n(1) > 0$ for all n. The n-th sampling error is the measure-valued random variable V_n^N defined by the formula

$$\eta_n^N = \eta_{n-1}^N K_{n,\eta_{n-1}}^N + V_n^N/\sqrt{N} \, . \tag{38}$$

Notice that V_n^N is itself the sum of the local errors induced by the random elementary transitions $\xi_{n-1}^i \rightsquigarrow \xi_n^i$ of the N particles; that is, we have

$$V_n^N = \sum_{i=1}^N \Delta_i V_n^N \, ,$$

with the "local" terms given for any $\varphi_n \in \mathcal{B}_b(E_n)$ by

$$\Delta_i V_n^N(\varphi_n) = \frac{1}{\sqrt{N}} [\varphi_n(\xi_n^i) - K_n, \eta_{n-1}^N(\varphi_n)(\xi_{n-1}^i)] \, .$$

By definition of the particle model, η_n^N is the empirical measure associated with a collection of conditionnaly independent random variables ξ_n^i with distributions $K_{n,\eta_{n-1}^N}(\xi_{n-1}^i, \cdot)$. From this we obtain that

$$\mathbb{E}_{\eta_0}^N[\eta_n^N(f_n)|\mathcal{F}_n^N] = \Phi_n(\eta_{n-1}^N)(f_n) = \eta_{n-1}^N K_{n,\eta_{n-1}^N} \, ,$$

where $\mathcal{F}_n^N = \sigma(\xi_0, \cdots, \xi_{n-1})$ is the σ-field asociated with the $\xi_0, \cdots, \xi_{n-1}$.

So we readily find that $\mathbb{E}(V_n^N(\varphi_n)) = 0$ and

$$\mathbb{E}(V_n^N(\varphi_n)^2) = \mathbb{E}(\eta_{n-1}^N(K_{n,\eta_{n-1}^N}[\varphi_n - K_{n,\eta_{n-1}^N}(\varphi_n)]^2)) \; .$$

In addition, for sufficiently regular McKean interpretation models, we have the asymptotic result

$$\lim_{N\to\infty} \mathbb{E}(V_n^N(\varphi_n)^2) = \eta_{n-1}(K_{n,\eta_{n-1}}[\varphi_n - K_{n,\eta_{n-1}}(\varphi_n)]^2) \; .$$

The formula (38) shows that the particle density η_n^N satisfy almost the same equation (37) as the limiting measures η_n. In fact [2][§9.3], $\mathcal{V}_n^N(\varphi_n)$ converges in law to a Gaussian random variable $V_n(\varphi_n)$ such that

$$\mathbb{E}(V_n(\varphi_n)) = 0 \quad \text{and} \quad \mathbb{E}(V_n(\varphi_n)^2) = \eta_{n-1}(K_{n,\eta_{n-1}}[\varphi_n - K_{n,\eta_{n-1}}(\varphi_n)]^2) \; .$$

These elementary fluctuations give some insight on the asymptotic normal behavior of the local errors accumulated by the sampling scheme. Nevertheless, they do not give directly CLT result for the difference between the particle measures η_n^N or γ_n^N and the corresponding limiting measures η_n and γ_n.

Preliminaries

The key idea is to consider the one-dimensional $\mathcal{F}^N$-martingale

$$M_n^N(f) = \sqrt{N} \sum_{p=0}^{n} 1_{\{\tau^N \geq p\}}[\eta_p^N(f_p) - \Phi_p(\eta_{p-1}^N)(f_p)] \; ,$$

where f_p stands for some collection of measurable and bounded functions defined on E_p. The angle bracket of this martingale is given by the formula

$$\langle M^N(f)\rangle_n = \sum_{p=0}^{n} \eta_{p-1}^N[K_{p,\eta_{p-1}^N}((f_p - K_{p,\eta_{p-1}^N} f_p)^2)] \; .$$

Then [2][Theorem 9.3.1], for any sequence of bounded measurable functions f_p and $p \geq 0$, the $\mathcal{F}^N$-martingale $M_n^N(f)$ converges in law to a Gaussian martingale $M_n(f)$ such that for any $n \geq 0$

$$\langle M(f)\rangle_n = \sum_{p=0}^{n} \eta_{p-1}[K_{p,\eta_{p-1}}((f_p - K_{p,\eta_{p-1}} f_p)^2)] \; .$$

A first consequence of this result is the next corollary which expresses the fact that the local errors associated with the particle approximation sampling steps behave asymptotically as a sequence of independent and centered Gaussian random variables.

Corollary 2. *The sequence of random fields $\mathcal{V}_n^N = (V_p^N)_{0\leq p\leq n}$ converges in law, as $N \to \infty$, to a sequence $\mathcal{V}_n = (V_p)_{0\leq p\leq n}$ of $(n+1)$ independent and Gaussian random fields V_p with, for any $\varphi_p^1, \varphi_p^2 \in \mathcal{B}_p(E_p)$, $\mathbb{E}(V_p(\varphi_p^1)) = 0$ and*

$$\mathbb{E}(V_p(\varphi_p^1)V_p(\varphi_p^2)) = \eta_{p-1}(K_{p,\eta_{p-1}}[\varphi_p^1 - K_{p,\eta_{p-1}}(\varphi_p^1)][\varphi_p^2 - K_{p,\eta_{p-1}}(\varphi_p^2)]) \; .$$

We now are concerned with the fluctuations of the particle approximation measures γ_n^N nd η_n^N. Nevertheless, before we start, we recall some tools to transfer CLT such as the Slutsky's technique and the δ-method. Firstly, the Slutsky's theorem states that for any sequences of random variables $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$, taking value in some separable metric space (E,d), which are such that X_n converges in law, as $n\to\infty$, to some random variable X, and $d(X_n,Y_n)$ converges to 0 in probability, then Y_n converges in law, as $N\to\infty$, to X. We deduce of this theorem, that if X_n converges in law to some finite constant c (which implies the convergence in probability) and Y_n converges in law to some variable Y, then X_nY_n converges in law to cY.

The other tool, also known as the δ-method [2][§9.3], is the following lemma.

Lemma 5. *Let $(U_0^N,\cdots,U_n^N)_{N\geq 1}$ be a sequence of $\mathbb{R}^{n+1}$-valued random variables defined on some probability space and $(u_p)_{0\leq p\leq n}$ be a given point in $\mathbb{R}^{n+1}$. Suppose that*

$$\sqrt{N}\,(U_0^N-u_0,\cdots,U_n^N-u_n)$$

converges in law, as $N\to\infty$, to some random vector $(U_0,\cdots,U_n)$. Then, for any differentiable function $F_n:\mathbb{R}^{n+1}\to\mathbb{R}$ at the point $(u_p)_{0\leq p\leq n}$, the sequence

$$\sqrt{N}\,[F_n(U_0^N(\omega),\cdots,U_n^N(\omega))-F_n(u_0,\cdots,u_n)]$$

converges in law as $N\to\infty$ to the random variable $\sum_{p=0}^n \frac{\partial F_n}{\partial u_i}(u_0,\cdots,u_n)U_p$.

Unnormalized Measures

We consider the $\mathbb{R}$ valued process $\Gamma_{\cdot,n}^N(f_n)$ introduced in Proposition 2. As the reader may have certainly noticed, the martingale decomposition of $\Gamma_{\cdot,n}^N$, exhibited in Proposition, 2 is expressed in terms of the sequence of local errors V_n^N.

Let $\overline{\Gamma}_{\cdot,n}^N(f_n)$ be the random sequence defined as in (33) by replacing, in the summation, the terms $\gamma_q^N(1)1_{\{\tau^N\geq q\}}$ by their limiting values $\gamma_q(1)$. In order to combine the CLT stated in Corollary 2 with the δ-method, we rewrite the resulting random sequence as

$$\begin{aligned}\sqrt{N}\,\overline{\Gamma}_{n,n}^N(f_n)&=\sqrt{N}\,\sum_{q=0}^p\gamma_q(1)\left[\eta_q^N-\eta_{q-1}^NK_{q,\eta_{q-1}^N}\right](Q_{q,n}f_n)\\&=\sqrt{N}\,F_n(U_{0,n}^N,\cdots,U_{n,n}^N)\,,\end{aligned}$$

with the random sequence $(U_{p,n}^N)_{0\leq p\leq n}$, and the function F_n given by

$$U_{p,n}^N=V_p^N(Q_{p,n}f_n)/\sqrt{N}\quad\text{and}\quad F_n(v_0,\cdots,v_n)=\sum_{q=0}^n\gamma_q(1)v_q\,.$$

Since for any $n \geq 0$ we have $\lim_{N\to\infty} \gamma_q^N(1)\ 1_{\{\tau^N \geq q\}} = \gamma_q(1)$ in probability, we easily deduce from Corollary 2, the Slutsky's theorem and the δ-method that the real-valued random variable $\sqrt{N}\ (\gamma_n^N(f_n)1_{\{\tau^N \geq n\}} - \gamma_n(f_n))$ converges in law to the centered Gaussian random variable $W_n^\gamma(f_n) = \sum_{q=0}^n \gamma_q(1)V_p(Q_{p,n}f_n)$ with variance

$$\sigma_n^2(f) = \sum_{q=0}^{n} (\gamma_q(1))^2 \eta_{q-1}\left(K_{q,\eta_{q-1}}[Q_{p,n}f_n - K_{q,\eta_{q-1}}Q_{p,n}f_n]^2\right)\ .$$

With the McKean model (28), the formula (29) gives the following new expression for the variance

$$\sigma_n^2(f) = \sum_{q=0}^{n} (\gamma_q(1))^2 \eta_q((Q_{q,n}f - \eta_q(Q_{q,n}f))^2)$$
$$- \sum_{q=1}^{n} (\gamma_q(1))^2 \eta_{q-1}\left(G_{q-1}^2(M_qQ_{q,n}f - \eta_q(Q_{q,n}f))^2\right). \quad (39)$$

Normalized Measures

Using formula (35) and the Slutsky's theorem, we obtain that the sequence of real-valued random variables

$$W_n^{\eta,N}(f) = \sqrt{N}\ (\eta_n^N(f) - \eta_n(f))1_{\{\tau^N \geq n\}}$$

converges to the Gaussian random variable W_n^η given by

$$W_n^\eta(f) = W_n^\gamma\left(\frac{1}{\gamma_n(1)}(f - \eta_n(f))\right)\ .$$

Now, let the semigroups $\overline{Q}_{p,n}$ and the functions $f_{p,n}$ be respectively defined by

$$\overline{Q}_{p,n} = \frac{\gamma_p(1)}{\gamma_n(1)}Q_{p,n}\ , \quad \text{and} \quad f_{p,n} = \overline{Q}_{p,n}(f - \eta_n f)\ . \quad (40)$$

Then, the variance of the Gaussian random variable $W_n^\eta(f)$ is given by the formula

$$\mathbb{E}(W_n^\eta(f)^2) = \sum_{p=0}^{n} \eta_{p-1}\left(K_{p,\eta_{p-1}}[f_{p,n} - K_{p,\eta_{p-1}}f_{p,n}]^2\right)\ . \quad (41)$$

Killing Interpretations and Related Comparisons

One of the best ways to interpret the fluctuations variances developed previously is to use the Feynman-Kac killing interpretations provided in Sect. 4.2.

In this context, X_n is regarded as a Markov particle evolving in an absorbing medium with obstacles related to $[0,1]$-valued potentials. Using the same notation and terminology as was used in Sect. 4.2, the Feynman-Kac semigroup $Q_{p,n}$ has the following interpretation

$$\begin{aligned} Q_{p,n}(x_p, dx_n) &= \int \left\{ \prod_{q=p}^{n-1} G_q(x_q) \right\} M_{p+1}(x_p, dx_{p+1} \cdots M_n(x_{n-1}, dx_n) \\ &= \mathbb{P}^c_{p,x_p}(X_n \in dx_n, T \geq n) \, , \end{aligned}$$

where $\mathbb{P}^c_{p,x_p}$ represents the distribution of the absorbed particle evolution model starting at $X_p = x_p$ at time p. In this context, the variance of the fluctuation variable $W^\gamma_n(1)$, associated with the McKean interpretation model (30), is given by

$$\begin{aligned} \mathbb{E}(W^\gamma_n(1)^2) &= \gamma_n(1)^2 \sum_{p=0}^{n} \eta_p \left([1 - G_{p,n}/\eta_p(G_{p,n})]^2 \right) \\ &= \mathbb{P}^c(T \geq n)^2 \sum_{p=0}^{n} \int_{E_p} \mathbb{P}^c(X_p \in dx_p | T \geq p) \left[\frac{\mathbb{P}^c_{p,x_p}(T \geq n)}{\mathbb{P}^c(T \geq n | T \geq p)} - 1 \right]^2 . \end{aligned}$$

We further assume that for any $n \geq p$ and η_p-a.e. $x_p, y_p \in \widehat{E}_p$, we have

$$\mathbb{P}^c_{p,x_p}(T \geq n) \geq \delta \mathbb{P}^c_{p,y_p}(T \geq n) \, , \tag{42}$$

for some $\delta > 0$ (see [2][Proposition 4.3.3] for sufficient conditions to obtain the condition (42)). In this case we have

$$\mathbb{E}(W^\gamma_n(1)^2) \leq b(\delta)(n+1)\mathbb{P}^c(T \geq n)^2,$$

for some finite constant $b(\delta)$.

The killing interpretation also suggests another evolution model based on N independent and identically distributed copies X^i of the absorbed particle evolution model. The Monte Carlo approximation is now given by $N^{-1} \sum_{i=1}^N 1_{\{T^i \geq n\}}$, where T^i represents the absorption time of the i-th particle. It is well known that the fluctuation variance $\sigma^{MC}_n(1)^2$ of this scheme is given by

$$\sigma^{MC}_n(1)^2 = \mathbb{P}^c(T \geq n)(1 - \mathbb{P}^c(T \geq n)) \, .$$

From previous considerations we find that

$$\frac{\sigma^{MC}_n(1)^2}{\mathbb{E}(W^\gamma_n(1)^2)} \geq \frac{1}{b(\delta)(n+1)} \frac{1 - \mathbb{P}^c(T \geq n)}{\mathbb{P}^c(T \geq n)} \to \infty \, ,$$

as soon as $\mathbb{P}^c(T \geq n) = o(1/n)$.

In addition, according to the formulas (41) and (31), and the observation that $\eta_q(f_{q,n}) = 0$, the variance of the random field W^η_n can also be described for any $f \in \mathcal{B}_b(E_n)$ as

$$\mathbb{E}(W_n^\eta(f)^2) = \sum_{p=0}^{n} \eta_p(f_{p,n}^2) \; .$$

If we choose the McKean model (28) then, according to the formula (29), we conclude that the variance of the random field W_n^η is defined for any $f \in \mathcal{B}_b(E_n)$ by the formula

$$\mathbb{E}(W_n^\eta(f)^2) = \sum_{p=0}^{n} \eta_p(f_{p,n}^2) - \sum_{p=1}^{n} \eta_{p-1}[(G_{p-1}M_p(f_{p,n}))^2] \; .$$

Then, we readily see that the variance of the corresponding CLT is strictly smaller than the one associated with the McKean interpretation $K_{n,\eta}(x_{n-1}, \cdot) = \Phi_n(\eta)$.

Application to Rare Event Analysis

We use the same notation and conventions as introduced in Sect. 5.3. Using the fluctuation analysis stated in the Sect. 6.5, we have the following theorem

Theorem 4. *For any $0 \leq n \leq m+1$, the sequence of random variables*

$$W_{n+1}^N = \sqrt{N} \; (1_{\{\tau^N > n\}} \gamma_{n+1}^N(1) - \mathbb{P}(T_n < T_R))$$

converges in law (as N tends to ∞) to a Gaussian random variable W_{n+1} with mean 0 and variance

$$\sigma_n^2 = \sum_{q=0}^{n+1} (\gamma_q(1))^2 \eta_{q-1} \left(K_{q,\eta_{q-1}}[Q_{q,n+1}(1) - K_{q,\eta_{q-1}} Q_{q,n+1}(1)]^2 \right) \; .$$

The collection of functions $Q_{q,n+1}(1)$ on the excursion space E are defined for any $x = (x_n)_{s \leq n \leq t}$ by

$$Q_{q,n+1}(1)(t,x) = 1_{B_q}(x_t) \mathbb{P}(T_n < T_R | T_q = t, X_{T_q} = x_t) \; .$$

Explicit calculations of σ_n are in general difficult to obtain since they rely on an explicit knowledge of the semigroup $Q_{q,n}$. Nevertheless, in the context of rare event analysis, an alternative can be provided. Firstly, according to the formula (39), the variance σ_n^2 takes the form

$$\sigma_n^2 = \mathbb{P}(T_n < T_R)^2 (a_n - b_n) \; ,$$

with

$$a_n = \frac{1}{\gamma_{n+1}(1)^2} \sum_{q=0}^{n+1} (\gamma_q(1))^2 \eta_q((Q_{q,n+1}(1) - \eta_q(Q_{q,n+1}(1)))^2)$$

$$b_n = \frac{1}{\gamma_{n+1}(1)^2} \sum_{q=1}^{n+1} (\gamma_q(1))^2 \eta_{q-1} \left(G_{q-1}^2 (M_q Q_{q,n+1}(1) - \eta_q(Q_{q,n+1}(1)))^2 \right) \; .$$

Then we observe that $\gamma_p(1) = \mathbb{P}(T_{p-1} < T_R)$ and

$$\eta_q Q_{q,n+1}(1) = \gamma_{n+1}(1)/\gamma_p(1) = \mathbb{P}(T_n < T_R | T_{p-1} < T_R) \ ,$$

from which we conclude that

$$a_n = \sum_{q=0}^{n+1} \mathbb{E}\left([\Delta^n_{q-1,q}(T_q, X_{T_q}) 1_{\{T_q < T_R\}} - 1]^2 | T_{q-1} < T_R \right) \ ,$$

where

$$\Delta^n_{p,q}(t,x) = \mathbb{P}(T_n < T_R | T_q = t, X_{T_q} = x) / \mathbb{P}(T_n < T_R | T_p < T_R) \ .$$

In much the same way, we find

$$\begin{aligned} b_n &= \sum_{q=0}^{n} \mathbb{E}(1_{\{T_q < T_R\}} [\Delta^n_{q,q}(T_q, X_{T_q}) - 1]^2 | T_{q-1} < T_R) \\ &= \sum_{q=0}^{n} \mathbb{P}(T_q < T_R | T_{q-1} < T_R) \mathbb{E}[\Delta^n_{q,q}(T_q, X_{T_q}) - 1]^2 | T_q < T_R) \ . \end{aligned}$$

References

1. F. Cérou, P. Del Moral, F. LeGland, and P. Lezaud. Genetic Genealogical Models in Rare Events Analysis. *Publication du Laboratoire de Statistique et Probabilités, Toulouse III*, 2002.
2. P. Del Moral. *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. Springer-Verlag, 2004.
3. P. Del Moral and J. Jacod. Interacting Particle Filtering With Discrete Observations. In A. Doucet, N. de Freitas, and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*. Springer Verlag, 2001.
4. P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering. In J. Azma, M. Émery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXIV*, volume 1729 of *Lecture Notes in Mathematics*, pages 1–14. Springer, 2000.
5. P. Del Moral and L. Miclo. Particle Approximations of Lyapunov Exponents Connected to Schrödinger Operators and Feynman-Kac Semigroups. *ESAIM: Probability and Statistics*, 7:171–208, 2003.
6. A. Doucet, N. de Freitas, and N. Gordon. An Introduction to Sequential Monte Carlo Methods. In A. Doucet, N. de Freitas, and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*. Springer Verlag, 2001.
7. P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer, 2003.
8. T.E. Harris and H. Kahn. Estimation of particle transmission by random sampling. *Natl. Bur. Stand. Appl. Math. Ser.*, 12:27–30, 1951.

9. J. Krystul and H. Blom. Sequential Monte Carlo Simulation of Rare Event Probability. *Hybridge report D9.3, http://www.nlr.nl/public/hosted-sites/hybridge/*, 2004.
10. A. Lagnoux. Rare Events Simulation. *Publication du Laboratoire de Statistique et Probabilités, Toulouse III*, 2003.
11. R. D. Reiss and M. Thomas. *Statistical Analysis of Extreme Values*. Birkhäuser, 1997.
12. M.N. Rosenbluth and A.W. Rosenbluth. Monte-Carlo calculations of the average extension of macromolecular chains. *J. Chem. Phys.*, 23:356–359, 1995.
13. A. Sznitman. *Brownian Motion Obstacles and Random Media*. Springer-Verlag, 1998.
14. A.N. Van der Vaart and J.A. Wellner. *Weak Convergence and Empirical Processes with Applications to Statistics*. Series in Statistics. Springer, 1996.
15. M. Villén-Altamirano and J. Villén-Altamirano. A method for accelerating rare event simulations. In *13th Int. Teletraffic Congress,ITC 13 (Queueing, Performance and Control in ATM)*, pages 71–76, Copenhagen, Denmark, 1991.
16. M. Villén-Altamirano and J. Villén-Altamirano. RESTART: An efficient and general method for fast simulation of rare events. Technical Report 7, Departamento de Mætmatica Aplicada, E.U. Informática, Universidad Politécnica de Madrid, 1997.

Index