



HAL
open science

On Adaptive Resampling Procedures for Sequential Monte Carlo Methods

Pierre del Moral, Arnaud Doucet, Ajay Jasra

► **To cite this version:**

Pierre del Moral, Arnaud Doucet, Ajay Jasra. On Adaptive Resampling Procedures for Sequential Monte Carlo Methods. [Research Report] 2008, pp.46. inria-00332436v1

HAL Id: inria-00332436

<https://inria.hal.science/inria-00332436v1>

Submitted on 21 Oct 2008 (v1), last revised 22 Oct 2008 (v4)

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

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



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

On Adaptive Sequential Monte Carlo Methods

Pierre Del Moral — Arnaud Doucet — Ajay Jasra —

N° 6438

May 2008

Thème NUM



*R*apport
de recherche

On Adaptive Sequential Monte Carlo Methods

Pierre Del Moral*, Arnaud Doucet†, Ajay Jasra ‡,

Thème NUM — Systèmes numériques
Équipes-Projets CQFD

Rapport de recherche n° 6438 — May 2008 — 49 pages

Abstract: This article is concerned with the modeling and convergence analysis of adaptive sequential Monte Carlo methods. These methods consist of an interaction (or resampling) of a simulated particle system with respect to some adaptive empirical criteria. We design a natural mean field particle interpretation of these Feynman-Kac type stochastic models in excursion spaces. The interaction times are computed on-line using empirical adaptive criteria, including effective sample size and Boltzmann entropy criteria; this is typically done in practice, however, there are very few convergence results in this instance. We combine a non-linear distribution flow analysis with an original coupling technique to obtain functional central limit theorems and uniform exponential concentration estimates for these algorithms, yielding what seems to be the first results of this type for this class of models. The original exponential concentration theorems presented in this study also improves recently developed concentration estimates for mean field interpretations of Feynman-Kac distribution flows models.

Key-words: Sequential Monte Carlo methods, Feynman-Kac formulae, mean field particle models, effective sample size, adaptive functional criteria.

* Centre INRIA Bordeaux et Sud-Ouest & Institut de Mathématiques de Bordeaux, Université de Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France, Pierre.DelMoral@inria.fr

† Department of Statistics, University of British Columbia, 333-6356 Agricultural Road, Vancouver, BC, V6T 1Z2, Canada, arnaud@stat.ubc.ca

‡ Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-Ku Tokyo 106-8569, Japan, ajay@ism.ac.jp

Méthodes de Monte Carlo séquentielles et adaptatives

Résumé : Cet article porte sur la modélisation et l'analyse de la convergence de méthodes de Monte Carlo séquentielles et adaptatives. Ces techniques de simulation stochastiques sont fondées sur l'évolution aléatoire de systèmes de particules en interaction et associées à des critères d'adaptation. Nous présentons des interprétations de ces algorithmes en termes d'approximation de type champ moyen de flots de mesures de Feynman-Kac dans des espaces d'excursions. Les instants d'interaction sont calculés en lignes selon des critères d'adaptation empiriques tels des critères de tailles de ré-échantillonnage effectives et des critères d'entropie relative de type Boltzmann. Ces stratégies d'adaptation empiriques sont d'un usage courant en pratique souvent présentées comme de simples heuristiques, sans aucun fondement mathématique rigoureux. Nous présentons de nouvelles techniques d'analyse de type semi-groupe fondées sur des couplage de processus à valeurs mesures. Cette analyse permet d'obtenir des théorèmes de la limite centrale fonctionnels et des résultats de concentration exponentielle uniformes par rapport au paramètre temporel. Notons enfin que ces travaux sont les premières études rigoureuses sur la concentration et les fluctuations de ces classes d'algorithmes de simulation adaptatifs. Par ailleurs, les théorèmes de concentration exponentielle présentés dans cette étude améliorent de récents travaux sur ce thème dans le cadre des interprétations particulières de flots de mesures de Feynman-Kac.

Mots-clés : Méthodes de Monte Carlo séquentielles et adaptatives, formules de Feynman-Kac, méthodes particulières de type champ moyen, taille d'échantillonnage effective, critères adaptatifs fonctionnels.

1 Introduction

Sequential Monte Carlo methods ([4, 9]) are one of the most important stochastic simulation algorithms and are extensively used in wide variety of applications in statistics, engineering, physics, biology and finance. The algorithm consists of the simulation, in parallel, of a collection of samples (or particles), sequentially in time, in order to approximate integrals with respect to a marginal of a Feynman-Kac formula. More details are given below.

One of the most critical aspects of the algorithm is termed resampling (or selection) which, essentially, ‘resets’ the particle system. In this interaction operation, new particles are sampled, with replacement from the current collection, according to some law which allows consistent estimation; some details are given in Section 2.1.2. This is required to enable to algorithm to accurately estimate integrals, but simultaneously, leads to a drawback in the algorithm: resampling too often yields many similar particles, and poor estimation of the integrals of interest. It is often the case, in many practical algorithms (e.g. [7]), that the time at which interacting resampling occurs, is not deterministic; that is, the operation occurs according to some judicious stochastic rule. Conversely, the theoretical analysis of these algorithms is studied in the deterministic case (see [8] for an exception) and it is the intention of this article to provide results in the stochastic case. We remark, that the results are not restricted to the case of resampling; many other elements of the algorithm can be adapted, and we discuss this when relevant.

The approach adopted here is to use a coupling argument. We start with an ‘ideal’ sequential Monte Carlo algorithm, that is, the algorithm resamples (or adapts) at deterministic times, which are not explicitly known, but which can be approximated. In effect, we show that the difference, in probability, between the ideal sequential Monte Carlo algorithm and the one simulated in practice (that is, approximating the adaptation times), is exponentially small. This allows us to appeal to a variety of existing convergence results in [4].

To our knowledge, the article [8] seems the only existing reference on the fluctuation analysis of adaptive sequential Monte Carlo models. Before we proceed, we provide a brief discussion on this work. In that article, the authors establish a convergence result (Theorem 6), for adaptive resampling schedules. However, their stochastic models applies in the context of state-space models (although, it is stated that the frame-work can also apply more generally, that is, for more general models), and only takes into account one of the adaptive criteria considered here. The arguments used by the authors are somehow misleading. Firstly, besides the fact that the resampling times are adaptive *random* times the authors propose a singular inductive proof with respect to deterministic time periods. Secondly, they ‘implicitly’ avoid the degenerate situation where the threshold parameters coincide with the limiting functional criteria. Last but not least, the fluctuations of the sum of the weighted and the resampled empirical approximations is deduced from the fluctuations of each separate components without the use of any multivariate type central limit theorem.

Our coupling approach, as well as the generality of the results (models, criteria), is an important development in the analysis of sequential Monte Carlo methods. That is to say, the analysis technique is explicitly designed for adaptive criteria and, as such, broadens the scope of the applications of the result, and the understanding of adaptive criteria. For example, our approach tells us

more about the finite sample (the algorithm simulated in practice) convergence properties, in terms of the rates of convergence. In addition, the concentration analysis developed in this paper allows us to deduce, without further work, multivariate and functional central limit theorems, yielding what seems to be the first results of this type for adaptive sequential Monte Carlo models. Lastly, and for sequential Monte Carlo methods in general, we present new exponential concentration estimates, including uniform estimates with respect to the time parameter. These results allow us to calibrate the number of samples so as to achieve a given precision with a given probability, uniformly with respect to the time parameter.

1.1 Structure of the Article

The article has the following structure.

In section 2, we give details of the models that we will study (2.1.2), along with the adaptive interpretation of these models (2.2) and some adaptive criteria (interaction or resampling times) (2.3) this allows us to give our main result (2.4).

In section 3, we review the basic construction of mean field particle models in adaptive excursion spaces. It begins with a description of abstract Feynman-Kac particle models (3.1). Excursion space models are discussed in section 3.4. In section 3.5, we present a general class of adaptive criteria and we discuss their regularity properties. In section 3.6, we design a Feynman-Kac functional representation in excursion spaces. Section 4 is concerned with the functional representation of the adaptive criteria introduced above.

In section 5, we design a mean field particle interpretation of the sequential Monte Carlo algorithm (that is as in [4]). In section 5.2, we provide a refined concentration analysis of the empirical adaptive criteria around their limiting values. In a short introduction, section 5.2.1, we briefly review some Feynman-Kac semigroup estimates derived in [4] and their connection with uniform convergence results for mean field particle models. Section 5.2.2 present some preliminary \mathbb{L}_m -mean error bounds estimates in terms of Feynman-Kac type semigroups. This section also provide an estimate of the bias of the particle empirical measures. In section 5.2.3, we present and prove an original exponential concentration theorem. This result improves the exponential estimates presented in earlier studies, such as those presented in section 7.4.3 in [4]. In section 5.3, we apply these results to estimate the concentration of particle adaptive criteria around their limiting functional values. The mean field type interpretation of the adaptive sequential Monte Carlo algorithm is provided in section 5.4.

The asymptotic analysis of this model as the number of samples tends to infinity is discussed in section 6. In section 6.1, we discuss the differences between the adaptive interaction times and their limiting values. In section 6.2, we present a randomized adaptive criteria that allows to compare these resampling periods, up to an event with an exponential small occurrence probability. In section 7, we analyze the fluctuations of adaptive sequential Monte Carlo methods.

2 The Main Result

2.1 Distribution models on path spaces

We now give some more substantive discussion, but in order to do this, we shall need to give our notations.

2.1.1 Some Notation

Let us briefly review some standard notation used in the present article. We denote respectively by $\mathcal{M}(E)$, $\mathcal{P}(E)$, and $\mathcal{B}_b(E)$, the set of bounded and signed measures, the subset of all probability measures on some measurable space (E, \mathcal{E}) , and the Banach space of all bounded and measurable functions f on E equipped with the uniform norm $\|f\| = \sup_{x \in E} |f(x)|$. We also denote by $\text{Osc}_1(E)$, the convex set of \mathcal{E} -measurable functions f with oscillations $\text{osc}(f) = \sup_{(x,y) \in E^2} \{|f(x) - f(y)|\} \leq 1$. We denote by $\mu(f) = \int \mu(dx) f(x)$, the Lebesgue integral of a function $f \in \mathcal{B}(E)$, with respect to a measure $\mu \in \mathcal{M}(E)$. We slightly abuse the notation, and sometimes denote by $\mu(A) = \mu(1_A)$ the measure of a measurable subset $A \in \mathcal{E}$. Recall that a bounded integral operator M from a measurable space E into an auxiliary measurable space F , is an operator $f \mapsto M(f)$ from $\mathcal{B}_b(F)$ into $\mathcal{B}_b(E)$ such that the functions $M(f)(x) = \int_F M(x, dy) f(y)$ are measurable and bounded, for any $f \in \mathcal{B}(E)$. A bounded integral operator M from a measurable space (E, \mathcal{E}) into an auxiliary measurable space (F, \mathcal{F}) also generates a dual operator $\mu \mapsto \mu M$ from $\mathcal{M}(E)$ into $\mathcal{M}(F)$ defined by $(\mu M)(f) := \mu(M(f))$. We also denote by $(p+q)_q := (p+q)!/p!$ the number of one-to-one mappings, from the set $\{1, \dots, p\}$ into $\{1, \dots, p+q\}$, and we use the conventions $(\sum_{\emptyset}, \prod_{\emptyset}) = (0, 1)$.

2.1.2 Distribution Models

Most of the research on Monte Carlo methods literature is centered around the computation of mean value type integrals

$$\gamma(f) := \int f(x) G(x) \mu(dx) = \mathbb{E}(f(X) G(X)) \quad \text{and} \quad \eta(f) := \gamma(f)/\gamma(1)$$

where X is a random variable with a prescribed distribution on some measurable state space (E, \mathcal{E}) , G is a given non negative potential type function on E , and f a given bounded measurable test function on E . If we can draw independent copies $(X^i)_{i \geq 1}$ of the reference random variable, an approximation of the pair of measures (γ, η) can be obtained as

$$\gamma^N := \frac{1}{N} \sum_{i=1}^N G(X^i) \delta_{X^i} \quad \text{and} \quad \eta^N := \sum_{i=1}^N \frac{G(X^i)}{\sum_{j=1}^N G(X^j)} \delta_{X^i}$$

In the above display, δ_a stands for the Dirac measure at the point $a \in E$. In practice, these rather crude approximations suffer from several problems. For instance, let us suppose that $G = 1_A$ is the indicator function of a measurable subset $A \subset E$ with a very small μ -probability mass $\mu(A) := \mathbb{P}(X \in A) \simeq 10^{-p}$, for some large parameter $p \geq 1$. In this situation, one may not be able to find a sample variable X^i within the set A , so that $\gamma^N = 0$ and the measure

η^N is not well defined. To overcome this difficulty, we often use importance sampling techniques. This strategy consists in finding an auxiliary ‘twisted probability measure’, say μ' , such that $\mu \ll \mu'$ with a well defined Radon-Nikodym derivative $\frac{d\mu}{d\mu'}$, and to set

$$\gamma(f) = \mathbb{E}(f(X') G'(X')) = \int f(x) G'(x) \mu'(dx) \quad \text{with} \quad G'(x) := G(x) \frac{d\mu}{d\mu'}(x)$$

In the above display, X' stands for a random variable with distribution μ' on E . This new model has exactly the same form as the one discussed above, so that we can use the same Monte Carlo strategy, replacing the pair of parameters (μ, G) by (μ', G') .

Even if we have found a good ‘twisted distributions’ another important problem often arises. Namely, the Radon-Nikodym weighted derivative reflects the relative differences between the desired reference measure and the importance sampling one. These weight functionals are generally degenerate (also known as *weight degeneracy*). For instance, let us suppose that the random variable X represents the random path of an auxiliary sequence of random variables X'_n , and the potential functions G is a weighted product function of the whole path sequence from the origin up to a given time horizon n :

$$X = (X'_0, \dots, X'_n) \quad \text{and} \quad G(X) = \prod_{0 < p < n} G'_p(X'_p)$$

In the above display, G'_p stands for an auxiliary potential function of the random states X'_p . In this situation, the pair of measures (γ, η) given below depend on the time parameter, and they are given for every test function f_n on the path space by the following Feynman-Kac integral representation formulae:

$$\eta_n(f_n) = \gamma_n(f_n) / \gamma_n(1) \quad \text{with} \quad \gamma_n(f_n) = \mathbb{E} \left(f_n(X'_0, \dots, X'_n) \prod_{0 < p < n} G'_p(X'_p) \right) \quad (2.1)$$

Once more, from the pure mathematical point of view, we emphasize that every importance sampling change of measure of the type discussed above reduce to the same type of Feynman-Kac model.

These probabilistic models and their mean field interacting particle interpretations are one of the most active contact points between probability, Bayesian statistics, theoretical chemistry, quantum physics, and engineering sciences, including rare event analysis and advanced signal processing. For instance, in signal processing, the chain X'_n represents the signal process and the potential functions stands for the likelihood functions associated with some partial and noisy observations. The corresponding measures η_n are versions of the conditional distribution of the random paths of the signal, given a series of observations. In particle physics, the law of a non absorbed random free evolution particle X'_n evolving in a absorbing medium E_n associated with an energy well V is given by a Feynman-Kac integral formulae with a Boltzmann-Gibbs potential function $G'_n(X'_n) = \exp(-\beta V(X'_n))$. For more details on these application model areas and their connection with molecular simulation and particle filtering technology, we refer the reader to chapter 12 in [4].

We emphasize that in most of the application areas discussed above, the free evolution process X'_n and the potential functions G'_n are dictated by the

physical or engineering problem at hand. In other instances, the choice of these two parameters is more flexible. For instance, let us suppose that the target measure is defined in terms of a Boltzmann-Gibbs mass transformation of a reference ‘static’ measure λ , with respect to some potential function g_n on some measurable state space E ; namely, these measures are defined by the following formulae:

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

For instance, in combinatorial counting problems λ often represents the uniform measure on a given finite set E and $g_n = 1_{A_n}$ the indicator function a subset $A_n \subset E$. In global optimization problems, λ may represent a reference measure on a given state space E and $g_n = \exp(-\beta_n V)$ a Boltzmann-Gibbs factor associated with an inverse temperature parameter β and an energy type non negative function V . The sampling of these Boltzmann-Gibbs measures is one of the central problems in the Markov chain Monte Carlo community. The literature is abounds with examples of Markov transitions M'_n with a prescribed invariant measure $\pi_n = \pi_n M'_n$. For instance, Metropolis-Hastings transitions often consist in sampling a local move, and to accept or reject the proposal with an appropriate acceptance ratio so that the target measure π_n is reversible with respect to these local transitions. In this connection, we mention that the standard and the rather well known simulated annealing algorithm consists in finding judicious increasing schedule $\beta_n \uparrow$ with sufficiently slow variations to ensure that successive Metropolis-Hasting moves between these changes are close to equilibrium. Another strategy, is to find a judicious series of multiplicative potential functions, so that

$$(g_{n+1} = g_n \times G'_n \quad \text{and} \quad \pi_n = \pi_n M'_n) \Rightarrow \pi_{n+1} = \Psi_{G'_n}(\pi_n) M'_{n+1} \quad (2.2)$$

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

$$G'_n = \exp(-(\beta_{n+1} - \beta_n)V), \quad \text{and respectively} \quad G'_n = 1_{A_{n+1}} \quad (2.3)$$

with an non decreasing inverse temperature parameter β_n and a non increasing sequence of subsets A_n .

Using the Markov property and the multiplicative structure of the Feynman-Kac measures (2.1), we can prove that the measures π_n given in (2.2) coincide with the n -th time marginals of the path measures defined in (2.1). That is, we have that

$$\pi_n(f) \propto \mathbb{E} \left(f(X'_n) \prod_{0 \leq p < n} G'_p(X'_p) \right) \quad \text{and} \quad \lambda(g_n) := \lambda(g_0) \times \prod_{0 \leq p < n} \pi_p(G'_{p+1}) \quad (2.4)$$

with the proportional sign \propto . As we mentioned above, we underline that these integral representation formulae are valid for any choice of the parameters (G'_n, M'_n) satisfying the pair of admissibility conditions presented in the l.h.s. of (2.2). As a result, we can choose a different series of parameters at every time step. As noticed by [14], in the context of molecular simulation, formula (2.4) is also known as Jarzynski’s non equilibrium dynamic model [10].

2.2 Adaptive sequential Monte Carlo

The literature on Monte Carlo computing is saturated with a variety of bio-inspired stochastic algorithms for sampling according to the Feynman-Kac path measures (2.1) and computing the normalizing constants. These stochastic algorithms are sometimes referred to as sequential Monte Carlo methods, population Monte Carlo methods, or branching and interacting particle systems, diffusion or quantum Monte Carlo methods. The central and natural idea is given below.

Firstly, we evolve a population of N individuals mimicking the free path evolution of the reference Markov chain X'_n . In a second stage, we kill individuals with poor G'_n -potential value, and we duplicate the ones with a relative larger potential value. In this sense, we define a branching particle model with a genealogical tree occupation measure that approximates the desired measures (2.1) as the precision parameter N tends to ∞ . The theoretical foundations of these classical models are rather well understood. From the pure mathematical point of view, they can be regarded as a mean field particle interpretation model associated with a non-linear Feynman-Kac distribution flow models. The interaction and branching rules are sampled at every time steps with an appropriate mean field selection transition that depends upon the current occupation measure of the system.

In recent years, researchers have built new branching and interaction strategies. The idea is to evolve the individuals up to some judicious time periods that depend on given criteria that measure the quality of a series of path samples. These resampling type criteria often depend on the time behavior of the cumulated weights of the free path particles. These new generation of tuning parameters have several advantages:

1. In the context of the Feynman-Kac path integral measures (2.1), these criteria can be used to design several online adaptive control strategies, including potential weights variances, or entropy like distances between the free evolution distribution and the weighted one.
2. In the context of Boltzmann-Gibbs measures (2.2), they allow to define adaptive temperature schedules or adaptive subsets restrictions strategies. For instance, we can decide to change the temperature parameter at some appropriate time periods. These time parameters and the next cooling temperature may depend on the performance of the exploration of the individuals in the energy landscape.

In terms of the latter point, some related investigations can be found in [11].

2.2.1 Adaptive Framework

Roughly speaking, these new generations of adaptive stochastic sampling strategies can be cast into the following recursion:

As for traditional models, the first central idea is to propagate forward a series of N sample paths copies

$$\forall 1 \leq i \leq N \quad \mathcal{Y}_1^{(N,i)} := (Y_0^{(N,i)}, Y_1^{(N,i)}, \dots, Y_{t_N}^{(N,i)})$$

of the path valued random variables $(X'_0, \dots, X'_{t_1^N})$ up to the first time t_1^N the importance weight products

$$W_{0,t_1^N}(\mathcal{Y}_1^{(N,i)}) := \prod_{0 < p \leq t_1^N} G'_p(Y_p^{(N,i)})$$

of the N path samples become somehow degenerate; say, falling below a reasonable threshold (note that this is only an example of an interaction time; some additional examples are described in Section 2.3). Loosely speaking, at that time $t_1^N = t_1$ the weighted occupation measures of the system

$$\tilde{\eta}_1^N := \sum_{i=1}^N \frac{W_{0,t_1^N}(\mathcal{Y}_1^{(N,i)})}{\sum_{j=1}^N W_{0,t_1^N}(\mathcal{Y}_1^{(N,j)})} \delta_{\mathcal{Y}_1^{(N,i)}}$$

is an N -approximation measure of the updated target measure given by

$$\hat{\eta}_1(f) = \hat{\gamma}_1(f)/\hat{\gamma}_1(1) \quad \text{with} \quad \hat{\gamma}_1(f) = \mathbb{E} \left(f(X'_0, \dots, X'_{t_1}) \prod_{0 < p \leq t_1} G'_p(X'_p) \right)$$

The precise analysis of these approximations is one of the objectives of the article, and it will be discussed later.

To prevent the weighted measures $\tilde{\eta}_1^N$ to become degenerate after the selected time period $t_1^N = t_1$, we replace these measures by a bootstrap type empirical measure

$$\hat{\eta}_1^N := \frac{1}{N} \sum_{i=1}^N \delta_{\hat{\mathcal{Y}}_1^{(N,i)}}$$

associated with N path particles $\hat{\mathcal{Y}}_1^{(N,i)} := (\hat{Y}_{0,t_1^N}^{(N,i)}, \hat{Y}_{1,t_1^N}^{(N,i)}, \dots, \hat{Y}_{t_1^N,t_1^N}^{(N,i)})$ resampled for instance randomly with the current weighted measure $\tilde{\eta}_1^N$. More judicious resampling type strategies are discussed in section 3.3 dedicated to mean field particle interpretations of Feynman-Kac distribution flows.

In a second stage, we propagate forward from every end point of the selected paths $\hat{Y}_{t_1^N}^{(N,i)}$ a novel series of N sample paths copies

$$(Y_{t_1^N+1}^{(N,i)}, Y_{t_1^N+2}^{(N,i)}, \dots, Y_{t_2^N}^{(N,i)})$$

of the path valued random variables $(X'_{t_1^N+1}, X'_{t_1^N+2}, \dots, X'_{t_2^N})$ starting at $X'_{t_1^N} = \hat{Y}_{t_1^N}^{(N,i)}$, up to the first time t_2^N , the importance weight products of the N path samples become again degenerate

$$W_{t_1^N,t_2^N}(\mathcal{Y}_2^{(N,i)}) := \prod_{t_1^N < p \leq t_2^N} G'_p(Y_p^{(N,i)})$$

In the above display, $\mathcal{Y}_2^{(N,i)}$ stands for the resulting path sequences of length t_2^N given by the formula

$$\mathcal{Y}_2^{(N,i)} := \left(\hat{Y}_{t_1^N,t_1^N}^{(N,i)}, (Y_{t_1^N+1}^{(N,i)}, Y_{t_1^N+2}^{(N,i)}, \dots, Y_{t_2^N}^{(N,i)}) \right)$$

Arguing as above, the weighted occupation measures of the system

$$\tilde{\eta}_2^N := \sum_{i=1}^N \frac{W_{t_1^N, t_2^N}(\mathcal{Y}_2^{(N,i)})}{\sum_{j=1}^N W_{t_1^N, t_2^N}(\mathcal{Y}_2^{(N,j)})} \delta_{\mathcal{Y}_2^{(N,i)}}$$

is an N -approximation measure of the updated target measure at time $t_2^N = t_2$ given by

$$\hat{\eta}_2(f) = \hat{\gamma}_2(f)/\hat{\gamma}_2(1) \quad \text{with} \quad \hat{\gamma}_2(f) = \mathbb{E} \left(f(X'_0, \dots, X'_{t_2}) \prod_{0 < p \leq t_2} G'_p(X'_p) \right)$$

As before, to prevent the degeneracy of the weighted measures, after the selected time period t_2 , we replace the weighted measures $\tilde{\eta}_2^N$ by a bootstrap type empirical measure

$$\hat{\eta}_2^N := \frac{1}{N} \sum_{i=1}^N \delta_{\hat{\mathcal{Y}}_2^{(N,i)}}$$

associated with N path particles $\hat{\mathcal{Y}}_2^{(N,i)} := (\hat{Y}_{0, t_2^N}^{(N,i)}, \hat{Y}_{1, t_2^N}^{(N,i)}, \dots, \hat{Y}_{t_2^N, t_2^N}^{(N,i)})$ resampled with the current weighted measure $\tilde{\eta}_2^N$, and so on up to the desired final time horizon.

Remark 2.1 *This adaptive excursion model is rather flexible. For instance, in the context of the Boltzmann-Gibbs model (2.2), we can choose the same potential function $G'_{t_n^N+1} := G'_p$ between the interaction times $t_n^N < p \leq t_{n+1}^N$. In other words, we only change the functions at the random interaction times t_n^N*

$$\forall t_n^N < p \leq t_{n+1}^N \quad G'_{t_n^N+1} := G'_p \quad \rightsquigarrow \quad \forall t_{n+1}^N < p \leq t_{n+2}^N \quad G'_{t_{n+1}^N+1} := G'_p \quad \rightsquigarrow \dots$$

For instance, we can use a temperature schedule with the same decay rate between the interaction times $t_n^N < p \leq t_{n+1}^N$, and then change the decay rate parameters only at the interaction times. Using this strategy, we have used only n different potential functions at every time step t_{n+1}^N :

$$G'_{t_0^N+1} \rightsquigarrow G'_{t_1^N+1} \rightsquigarrow \dots \rightsquigarrow G'_{t_n^N+1} \rightsquigarrow$$

Without further work, the forthcoming analysis can be extended to situations where the potential functions $(G'_p)_{p > t_n^N}$ at a given level n depend on the series of parameters (t_0^N, \dots, t_n^N) . More formally, we can choose

$$\forall t_n^N < p \leq t_{n+1}^N \quad G'_p(x) := G''_{p, (t_0^N, \dots, t_n^N)}(x) \quad (2.5)$$

for some collection of mappings $G''_{p, (t_0, \dots, t_n)}$, with $0 = t_0 \leq \dots \leq t_n < p$. In much the same way, in the context of the Boltzmann-Gibbs model (2.2), we can use a series of Markov transitions $(M'_p)_{p > t_n^N}$ as well as an adaptive resampling criteria at a given level n that depend on the series of parameters (t_0^N, \dots, t_n^N) .

2.3 Some adaptive criteria

The literature on sequential Monte Carlo algorithms abounds with entropy like criteria to measure the degeneracy of the weighted samples; this can be used to design adaptive resampling strategies. In all the situations, the resampling times steps $t_n = t_n^N$ are random variables that depend on the N - sequential Monte Carlo algorithm.

To simplify the presentation, we suppose that the potential functions G'_n satisfy the following condition

$$(G') \quad \forall n \geq 0 \quad q'_n := \sup_{(x,y) \in (E'_n)^2} (G'_n(x)/G'_n(y)) < \infty \quad (2.6)$$

The analysis of more general models is technically more involved and it often requires some additional conditions on the free evolution process. We refer the interested reader to [4] for a more thorough discussion on these technicalities. Nevertheless, the forthcoming development can be easily combined with these models. The criteria are termed 1)-4), in the order that they appear. Because of their importance in practice, and to better connect these strategies with existing results in the literature, we develop a rather detailed discussion on these adaptive strategies.

2.3.1 Adaptive control of the weights

In this situation, the particle model explores the state-space with a series of local transitions

$$\widehat{Y}_{t_n^N, t_n^N}^{(N,i)} \longrightarrow Y_{t_{n+1}^N}^{(N,i)} \longrightarrow Y_{t_{n+2}^N}^{(N,i)} \longrightarrow \dots \longrightarrow Y_s^{(N,i)} \quad (2.7)$$

up to the first time $s = t_{n+1}^N$ ($\geq t_n^N$) the proportion of particles with a weight of at least size s_n is below a prescribed a_n -threshold

$$C_{t_n^N, s}^N := \frac{1}{N} \text{Card} \left\{ 1 \leq i \leq N : \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,i)}) \geq s_n \right\} \leq a_n \quad (2.8)$$

with a prescribed parameter $a_n \leq 1$. To our knowledge, this type of adaptive weight criteria was first discussed as an heuristic type resampling schedule in the end of the article [3].

Next, we provide some comments on the role and the choice of the adaptive parameters (s_n, a_n) . The forthcoming discussion applies to all the examples of adaptive criteria discussed in the article.

For $[0, 1]$ -valued potential functions G'_k , it is convenient to choose in (2.8) a weight size parameter $s_n \in [0, 1]$. In this situation, the random mappings

$$s \in [t_n^N, \infty[\mapsto C_{t_n^N, s}^N \in [0, 1]$$

are non increasing functions starting at $C_{t_n^N, t_n^N}^N = 1$ (where we have use the standard convention $\prod_{\emptyset} = 1$). We emphasize that when none of the sampled particles at time $s = t_n^N + 1$ have a weight greater than the prescribed threshold s_n , we simply have $t_{n+1}^N = t_n^N$ and we ‘restart’ the above adaptive sampling

path strategy. Using the extended adaptive models (2.5) discussed at the end of section 2.2, we can choose to restart the complete run with the same set of potentials and Markov transitions as we the one we used before and/or we can decide to decrease the weight size parameter s_n as well as the threshold level a_n .

In connection with the adaptive potential models discussed in (2.5), our framework allows to analyze without further work fully adaptive models associated with sequence of threshold parameters that may depend on the previous times steps

$$s_n = s'_n(t_0^N, \dots, t_n^N) \quad \text{and} \quad a_n = a'_n(t_0^N, \dots, t_n^N)$$

for some functions s' and a'_n on \mathbb{R}_+^{n+1} .

The choice of these threshold parameters depends on the problem at hand. From the previous discussion, if s_n is excessively large, the N -sampling path model may be stuck at a given time step t_n^N during several runs. In relation to this point, and in the further development of the article, we provide a set of regularity conditions under which these empirical criteria approximate a limiting functional integration criteria. In this case, the random time step t_n^N converges, as N tends to ∞ , to a series of deterministic sequences of time periods t_n expressed in terms of these functional criteria. In this situation, for a large population size N we are always ensure that $t_{n+1}^N > t_n^N$ as soon as $t_{n+1} > t_n$.

To illustrate this adaptive time step splitting, let us suppose that the reference stochastic process X'_k a Markov chain on $E'_k = \mathbb{Z}$ starting at the origin $X'_0 = 0$. For instance, X'_k can be taken as a simple random walk on the integer lattice. If we choose the indicator potential functions and the unit weight size parameters

$$G'_k = 1_{[k, \infty[} \quad \text{and} \quad s_n = 1 \quad (2.9)$$

then we find

$$\prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,i)}) \geq 1 \iff \left(\forall t_n^N < k \leq s \quad Y_k^{(N,i)} \geq k \right)$$

for every $1 \leq i \leq N$. Thus, we have

$$\begin{aligned} C_{t_n^N, s}^N &= \frac{1}{N} \text{Card} \left\{ 1 \leq i \leq N : \forall t_n^N < k \leq s \quad Y_k^{(N,i)} \geq k \right\} \\ &= \frac{1}{N} \sum_{i=1}^N \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,i)}) \end{aligned} \quad (2.10)$$

When $a_n = 1/2$, we find that t_{n+1}^N is such that

$$t_{n+1}^N = \inf \left\{ s \geq t_n^N : \right.$$

$$\left. \text{Card} \left\{ 1 \leq i \leq N : Y_{t_n^N+1}^{(N,i)} \geq t_n^N + 1, Y_{t_n^N+2}^{(N,i)} \geq t_n^N + 2, \dots, Y_s^{(N,i)} \geq s \right\} \leq N/2 \right\}$$

Roughly speaking, for the simple random walk model, it is ensured that about half of the N sampled paths are able to reach the next upper level set at every elementary time step. In this situation, the random times t_{n+1}^N are more likely to take the value $t_n^N + 1$, or to return the value t_n^N . If we take a lower threshold

parameter, say $a_n = 1/100$, these random times t_{n+1}^N are more likely to take large values. Indeed, in this case, about $N/2$ will succeed to reach the first level, thus about $N/4$ the second, $N/8$ the third one, and so on up to $N/64$ the sixth and $N/128$ the seventh. Therefore, for $a_n = 1/100$ we expect t_{n+1}^N to take the value $t_n^N + 6$.

The random walk model discussed above can easily be extended to multi-splitting branching excursion type models, absorbed at some lower level set, say $B \subset \mathbb{Z} \cap \mathbb{R}_-$. In this case, the random path processes $(Y_k^{(N,i)})_{k \geq t_n^N}$ are independent copies of a random excursion starting from a given level set t_n^N and ending at the absorbing set B . The next adaptive time step t_{n+1}^N coincides with the first upper level set that the N particles are able to reach with a proportion less than a_n (that is, before being absorbed by the set B). The resulting model for adaptive multi-level splitting coincides with the adaptive branching algorithm for rare event simulation discussed in [1].

2.3.2 Normalizing constants

In this situation, the particle model explores the state-space with a series of local transitions (2.7) up to the first time $s = t_{n+1}^N$ the empirical potential products goes below a prescribed a_n -threshold

$$C_{t_n^N, s}^N := \frac{1}{N} \sum_{i=1}^N \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,i)}) \leq a_n \quad (2.11)$$

For indicator potential functions, by (2.10) we see that this situation coincide with the one discussed above. Furthermore, for $]0, 1]$ -valued potential functions G'_k , it is convenient to choose in (2.11) a threshold parameter $a_n \in [0, 1]$. In this situation, the random mappings

$$s \in [t_n^N, \infty[\mapsto C_{t_n^N, s}^N \in [0, 1]$$

are non increasing functions starting at $C_{t_n^N, t_n^N}^N = 1$ (where we have use the standard convention $\prod_{\emptyset} = 1$). Notice that

$$a_n = 1 \implies t_{n+1}^N = t_n^N$$

For $]0, 1]$ -valued threshold parameters, we have the rather crude estimates

$$\begin{aligned} \underline{g}_k &:= \inf G'_k \leq \sup G'_k := \bar{g}_k \\ &\Downarrow \\ \prod_{t_n^N < k \leq s} \underline{g}_k &:= \underline{C}_{t_n^N, s}^N \leq C_{t_n^N, s}^N \leq \bar{C}_{t_n^N, s}^N := \prod_{t_n^N < k \leq s} \bar{g}_k \end{aligned} \quad (2.12)$$

This pair of estimates can be improved by considering the infimum and the supremum of the potential product weight functions along the M'_k -admissible paths sequences (2.7) starting at the selected states $\widehat{Y}_{t_n^N, t_n^N}^{(N,i)}$.

Notice that (2.12) implies that

$$\underline{t}_{n+1}^N \leq t_{n+1}^N \leq \bar{t}_{n+1}^N$$

with

$$t_{n+1}^N := \inf \left\{ s \geq t_n^N : \underline{C}_{t_n^N, s}^N \leq a_n \right\} \quad \text{and} \quad \bar{t}_{n+1}^N := \inf \left\{ s \geq t_n^N : \bar{C}_{t_n^N, s}^N \leq a_n \right\}$$

For instance, for the exponential Boltzmann-Gibbs potential given in (2.3), we have

$$\underline{V} := \inf V \leq \sup V := \bar{V} \implies \underline{g}_k = e^{-(\beta_{k+1} - \beta_k) \bar{V}} \quad \text{and} \quad \bar{g}_k = e^{-(\beta_{k+1} - \beta_k) \underline{V}}$$

This clearly implies that

$$\underline{C}_{t_n^N, s}^N = e^{-(\beta_{s+1} - \beta_{t_n^N+1}) \bar{V}} \quad \text{and} \quad \bar{C}_{t_n^N, s}^N = e^{-(\beta_{s+1} - \beta_{t_n^N+1}) \underline{V}}$$

We further assume the temperature schedule $(\beta_k)_{k > t_n^N}$ after time t_n^N is associated with a prescribed and fixed decay rate $\delta_n > 0$. In this situation, we have

$$\begin{aligned} \forall t_n^N < k \quad \Delta \beta_{k+1} = (\beta_{k+1} - \beta_k) = \delta_n &\implies (\beta_{s+1} - \beta_{t_n^N+1}) = (s - t_n^N) \delta_n \\ &\implies \begin{cases} \underline{C}_{t_n^N, s}^N = e^{-(s - t_n^N) \delta_n \bar{V}} \\ \bar{C}_{t_n^N, s}^N = e^{-(s - t_n^N) \delta_n \underline{V}} \end{cases} \end{aligned}$$

From the previous discussion, we conclude that

$$t_n^N + \frac{1}{\delta_n \bar{V}} \log(1/a_n) := t_{n+1}^N \leq t_{n+1}^N \leq \bar{t}_{n+1}^N = t_n^N + \frac{1}{\delta_n \underline{V}} \log(1/a_n)$$

For small values of the inverse temperature decay rates δ_n , we ensure that the adaptive criteria (2.11) is met during at least $\frac{1}{\delta_n \bar{V}} \log(1/a_n)$ time steps. In the reverse angle, to ensure that the adaptive criteria (2.11) is met during at least m_n time steps we need to choose $\delta_n \leq (m_n \bar{V})^{-1} \log(1/a_n)$; that is

$$\delta_n \leq (m_n \bar{V})^{-1} \log(1/a_n) \implies (t_{n+1}^N - t_n^N) \geq m_n$$

2.3.3 Weight variances criteria

In this situation, the particles explore the state-space with a series of local transitions (2.7), up to the first time $s = t_{n+1}^N$ the empirical particle variances is larger than some prescribed a_n -threshold

$$\sum_{i=1}^N \left(\frac{\prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,i)})}{\sum_{j=1}^N \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,j)})} \right)^2 \geq (1 + a_n)/N$$

or equivalently

$$C_{t_n^N, s}^N := \frac{1}{N} \sum_{i=1}^N \left(\frac{\prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,i)})}{\frac{1}{N} \sum_{j=1}^N \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,j)})} \right)^2 - 1 \geq a_n$$

In the statistical literature this criteria is called the effective sample size. To our knowledge, this approach was first discussed as a heuristic type resampling

criteria by [12]. For indicator potential functions, we have $(G'_k)^2 = G'_k$ from which we find that

$$C_{t_n^N, s}^N := \frac{1}{\frac{1}{N} \sum_{j=1}^N \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N,j)})} - 1$$

In this case, by (2.10), we see that this coincides with the pair of adaptive criteria discussed above.

2.3.4 Boltzmann's entropy criteria

We further assume that the potential functions G'_k take values in $]0, \infty[$. In this case, the particles explore the state-space with a series of local transitions (2.7), up to the first time $s = t_{n+1}^N$ that the relative entropy of the empirical particle measures with respect to its weighted versions are larger than some a_n -threshold

$$C_{t_n^N, s}^N := -\frac{1}{N} \sum_{i=1}^N \sum_{t_n^N < k \leq s} \log G'_k(Y_k^{(N,i)}) \geq a_n \quad (2.13)$$

For the exponential Boltzmann-Gibbs potential given in (2.3), we have

$$C_{t_n^N, s}^N := \frac{1}{N} \sum_{i=1}^N \sum_{t_n^N < k \leq s} (\beta_{k+1} - \beta_k) V(Y_k^{(N,i)})$$

In this situation, the random mappings

$$s \in [t_n^N, \infty[\mapsto C_{t_n^N, s}^N \in [0, 1]$$

are non decreasing functions starting at $C_{t_n^N, t_n^N}^N = 0$ (where we have use the standard convention $\sum_{\emptyset} = 0$). Arguing as above, we have the rather crude estimates

$$\underline{V} := \inf V \leq \sup V := \bar{V}$$

↓

$$\underline{C}_{t_n^N, s}^N = (\beta_{s+1} - \beta_{t_n^N+1}) \underline{V} \leq C_{t_n^N, s}^N \leq \bar{C}_{t_n^N, s}^N := (\beta_{s+1} - \beta_{t_n^N+1}) \bar{V}$$

This implies that

$$\underline{t}_{n+1}^N \leq t_{n+1}^N \leq \bar{t}_{n+1}^N$$

with

$$\underline{t}_{n+1}^N := \inf \left\{ s \geq t_n^N : \bar{C}_{t_n^N, s}^N \geq a_n \right\} \quad \text{and} \quad \bar{t}_{n+1}^N := \inf \left\{ s \geq t_n^N : \underline{C}_{t_n^N, s}^N \geq a_n \right\}$$

When the temperature schedule $(\beta_k)_{k > t_n^N}$ after time t_n^N is associated with a prescribed and fixed decay rate $\delta_n > 0$, we find that

$$\underline{C}_{t_n^N, s}^N = (s - t_n^N) \delta_n \underline{V} \quad \text{and} \quad \bar{C}_{t_n^N, s}^N = (s - t_n^N) \delta_n \bar{V}$$

and therefore

$$\underline{t}_{n+1}^N := t_n^N + a_n / (\delta_n \bar{V}) \leq t_{n+1}^N \leq \bar{t}_{n+1}^N := t_n^N + a_n / (\delta_n \underline{V})$$

In this case, to ensure that the adaptive criteria (2.13) is met during at least m_n iterations, we need to choose an inverse temperature decay rate $\delta_n \leq a_n / (m_n \bar{V})$. More formally, we have that

$$\delta_n \leq a_n / (m_n \bar{V}) \implies (t_{n+1}^N - t_n^N) \geq m_n$$

2.4 Statement of some results

2.4.1 A limiting particle model

Roughly speaking, we let $(t_n)_{n \geq 0}$ be the deterministic sequence of time steps defined as $(t_n^N)_{n \geq 0}$ by replacing the empirical criteria $C_{t_n^N, s}^N$ by their limiting values $C_{t_n, s}$, as $N \uparrow \infty$; that is, in the examples 1) and 3) we have that

$$t_{n+1} := \inf \{t_n \leq s : C_{t_n, s} \leq a_n\}$$

In all the situations, we set $t_0^N = t_0 = 0$.

In the examples 2) and 4), the limiting times t_n are defined as above by replacing condition $C_{t_n, s} \leq a_n$ by the condition $C_{t_n, s} \geq a_n$. In the four situation examined above, these limiting criteria $C_{t_n, s}$ are respectively given by

$$1) \quad \mathbb{P}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \geq a_n \right) \quad 2) \quad \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)$$

and

$$3) \quad \frac{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k)^2 \right)}{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)^2} - 1 \quad 4) \quad - \sum_{t_n < k \leq s} \mathbb{E}_{n, \hat{\eta}_n} (\log G'_k(X'_k))$$

with the updated Feynman-Kac path measures $\hat{\eta}_n$ given below

$$\hat{\eta}_n(f) = \hat{\gamma}_n(f) / \hat{\gamma}_n(1) \quad \text{with} \quad \hat{\gamma}_n(f) = \mathbb{E} \left(f(X'_0, \dots, X'_{t_n}) \prod_{0 < p \leq t_n} G'_p(X'_p) \right)$$

In the above formulae, $\mathbb{E}_{n, \hat{\eta}_n}$ is the expectation with respect to the law of a free evolution X'_k starting at time t_n at the end point $X'_{t_n} = \hat{X}'_{t_n}$ of a random path $\hat{\mathcal{X}}_n := (\hat{X}'_{0, t_n}, \hat{X}'_{1, t_n}, \dots, \hat{X}'_{t_n, t_n})$ with distribution $\hat{\eta}_n$; that is, we have that

$$\hat{X}'_{t_n, t_n} \xrightarrow{M'_{t_n+1}} X'_{t_n+1} \xrightarrow{M'_{t_n+2}} X'_{t_n+2} \xrightarrow{M'_{t_n+3}} \dots \xrightarrow{M'_s} X'_s \quad (2.14)$$

Notice that $\hat{\eta}_n$ and $\hat{\gamma}_n$ are the updated Feynman-Kac measures associated with the pair of measures

$$\eta_n(f) = \gamma_n(f) / \gamma_n(1) \quad \text{with} \quad \gamma_n(f_n) = \mathbb{E} \left[f_n(X'_0, \dots, X'_{t_n}) \prod_{0 < k \leq t_n-1} G'_k(X'_k) \right]$$

Definition 2.2 We denote by $(\mathcal{X}_n, \hat{\mathcal{X}}_n)_{n \geq 0}$ the sequential Monte carlo algorithm defined as $(\mathcal{Y}_n, \hat{\mathcal{Y}}_n)_{n \geq 0}$ by replacing the empirical criteria $C_{t_n^N, s}^N$ by their deterministic limiting values $C_{t_n, s}$. We also suppose that both models coincide on every time interval $0 \leq n \leq m$, as soon as the random times $t_n^N = t_n$, for every $0 \leq n \leq m$. This condition simply amounts of coupling the pair of models with the same random exploration and selection variables on the events $\cap_{0 \leq n \leq m} \{t_n^N = t_n\}$.

2.4.2 An exponential coupling theorem

We give our main results, which holds under the regularity condition, given above.

The first result is the following exponential coupling theorem.

Theorem 2.3 *Assume that the threshold parameters $(a_n)_{n \geq 0}$ are sampled realizations of a collection of absolutely continuous random variables $A = (A_n)_{n \geq 0}$. In this situation, for almost every realization of these thresholds, there exists a canonical way to couple the stochastic algorithms $(\mathcal{X}_n, \hat{\mathcal{X}}_n)_{n \geq 0}$ and $(\mathcal{Y}_n, \hat{\mathcal{Y}}_n)_{n \geq 0}$ such that for every $m \geq 0$ and any $N \geq 1$, we have*

$$\mathbb{P} \left(\exists 0 \leq n \leq m \quad (\mathcal{Y}_n^{(N)}, \hat{\mathcal{Y}}_n) \neq (\mathcal{X}_n^{(N)}, \hat{\mathcal{X}}_n) \mid A \right) \leq c_1(m) e^{-N \epsilon^2(m)/c_2(m)}$$

for a pair of finite constants $c_1(m), c_2(m) < \infty$ whose values only depend on the time parameter, and an almost sure positive constant $\epsilon(m) > 0$ that only depends on $(A_n)_{0 \leq n \leq m}$.

The second result is the following non asymptotic exponential concentration result.

Theorem 2.4 *For any $n \geq 0$, any bounded test function f_n , with $\|f_n\| \leq 1$, any $N \geq 1$, and any $0 \leq \epsilon \leq 1/2$, we have the exponential concentration estimate*

$$\mathbb{P} \left(\left| [\eta_n^N - \eta_n] (f_n) \right| \geq \epsilon \right) \leq c_1 \exp(-N \epsilon^2/c_2(n))$$

with the empirical measures

$$\eta_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{(\hat{\mathcal{X}}_{n-1}^{(N,i)}, (X_{t_{n-1}+1}^{(N,i)}, X_{t_{n-1}+2}^{(N,i)}, \dots, X_{t_n}^{(N,i)}))}$$

In the above displayed formula c_1 stands for a finite universal constant c_1 , and $c_2(n)$ for some collection of finite constants. In addition, under appropriate regularity conditions on the potential transitions G'_k and on the Markov chain X'_k , the above estimates are valid for the marginal measures associated with the time parameters $t_{n-1} \leq p \leq t_n$, for some constant $c_2(n) = c_2$ whose values do not depend on the time parameter.

The proofs of theorem 2.3 and theorem 2.4 are respectively detailed in section 6.2 and in section 5.2.3. Let us mention some direct consequences of the concentration theorem, theorem 2.3. The sequential Monte carlo algorithm $(\mathcal{X}_n, \hat{\mathcal{X}}_n)_{n \geq 0}$ can be regarded as a mean field particle interpretation of Feynman-Kac integrals. A variety of convergence results have been developed for these particle approximation models (see for instance [4]), including non asymptotic exponential estimates and \mathbb{L}_p -mean error bounds. The above theorem allows us to transfer all of these estimates up to an event with an exponentially small occurrence probability. Furthermore, without further work these exponential estimates show that the fluctuation of the occupation measures $\frac{1}{N} \sum_{i=1}^N \delta_{\mathcal{Y}_n^{(N,i)}}$ around their limiting values η_n coincide with the ones $\frac{1}{N} \sum_{i=1}^N \delta_{\mathcal{X}_n^{(N,i)}}$ associated with the N -particle model \mathcal{X}_n . We shall return to this subject in section 7 dedicated to functional central limit theorems for weighted and resampled empirical approximations.

3 Description of the models

For the convenience of the reader, in the first two sections, (section 3.1 and section 3.3), we briefly present the definition of the mean field particle interpretation model associated with an abstract Feynman-Kac distribution flow model. These constructions are used in various places in the article. The excursion Feynman-Kac model associated with a prescribed sequence of time parameters are presented in section 3.4.

The definition of the functional adaptive criteria discussed in this work is provided in section 3.5, and the precise definition of "the limiting" time parameters $(t_n)_{n \geq 0}$ discussed in the introduction is provided in section 3.6.

3.1 Feynman-Kac distribution flows

We consider a sequence of measurable state spaces S_n , a probability measure η_0 on S_0 , and a sequence of Markov transitions $\mathcal{M}_n(x_{n-1}, dx_n)$ from S_{n-1} into S_n , with $n \geq 1$. We let \mathcal{X}_n be a Markov chain with initial distribution $\text{Law}(\mathcal{X}_0) = \eta_0$ on S_0 and elementary transitions

$$\mathbb{P}(\mathcal{X}_n \in dy \mid \mathcal{X}_{n-1} = x) = \mathcal{M}_n(x, dy)$$

We consider a sequence of non negative and bounded potential-likelihood function \mathcal{G}_n on S_n . To simplify the presentation, and to avoid unnecessary technicalities, we shall suppose that $\mathcal{G}_0 = 1$ and the potential functions \mathcal{G}_n , with $n \geq 1$, take values in $]0, 1[$ and they satisfy the following condition

$$(\mathcal{G}) \quad \forall n \geq 1 \quad q_n := \sup_{(x,y) \in S_n^2} (\mathcal{G}_n(x)/\mathcal{G}_n(y)) < \infty \quad (3.1)$$

This regularity condition can be relaxed in various ways. As before, we refer the reader to chapter 2 in [4], for a detailed discussion on Feynman-Kac models associated with more general potential functions, including indicator potential functions and unbounded potentials. The Boltzmann-Gibbs transformation Ψ_n associated with a given potential function \mathcal{G}_n is the mapping

$$\Psi_n : \mu \in \mathcal{P}(S_n) \mapsto \Psi_n(\mu) \in \mathcal{P}(S_n) \quad \text{with} \quad \Psi_n(\mu)(dx) := \frac{1}{\mu(\mathcal{G}_n)} \mathcal{G}_n(x) \mu(dx)$$

Notice that $\Psi_n(\mu)$ can be rewritten as a non-linear Markov transport equation

$$\Psi_n(\mu)(dy) = (\mu \mathcal{S}_{n,\mu})(dy) := \int_{S_n} \mu(dx) \mathcal{S}_{n,\eta_n}(x, dy)$$

with the acceptance-rejection type selection transitions:

$$\mathcal{S}_{n,\eta_n}(x, dy) = \mathcal{G}_n(x) \delta_{x_n}(dy) + (1 - \mathcal{G}_n(x)) \Psi_n(\eta_n)(dy)$$

We let $(\eta_n, \hat{\eta}_n)_{n \geq 0}$ be the flow of probability measures, both starting at $\eta_0 = \hat{\eta}_0$, and defined for any $n \geq 1$ by the following recursion

$$\forall n \geq 0 \quad \eta_{n+1} = \hat{\eta}_n \mathcal{M}_{n+1} \quad \text{with} \quad \hat{\eta}_n := \Psi_n(\eta_n) \quad (3.2)$$

Notice that $(\eta_n)_{n \geq 0}$ and $(\hat{\eta}_n)_{n \geq 0}$ satisfies the nonlinear recursion:

$$\eta_{n+1} = \Phi_{n+1}(\eta_n) := \Psi_n(\eta_n) \mathcal{M}_{n+1} \quad \text{and} \quad \hat{\eta}_{n+1} = \hat{\Phi}_{n+1}(\hat{\eta}_n) := \Psi_{n+1}(\hat{\eta}_n \mathcal{M}_{n+1})$$

It is worth mentioning that the one step mapping $\widehat{\Phi}_{n+1}$ can be rewritten as follows

$$\widehat{\Phi}_{n+1}(\widehat{\eta}_n) := \widehat{\Psi}_n(\widehat{\eta}_n) \widehat{\mathcal{M}}_{n+1}$$

with the Boltzmann-Gibbs transformation $\widehat{\Psi}_n$ associated with a given potential function $\widehat{\mathcal{G}}_n = \mathcal{M}_{n+1}(\mathcal{G}_n)$ and the Markov transition $\widehat{\mathcal{M}}_{n+1}$ given for any bounded measurable test function f on S_{n+1} , and any $x \in S_n$ by the formula:

$$\widehat{\mathcal{M}}_{n+1}(f)(x) = \frac{\mathcal{M}_{n+1}(\mathcal{G}_n f)(x)}{\mathcal{M}_{n+1}(\mathcal{G}_n)(x)}$$

From the pure mathematical point of view, this shows that the flow $(\widehat{\eta}_n)_{n \geq 0}$ has the same dynamical structure as the flow $(\eta_n)_{n \geq 0}$. Thus, without any loss of generality we shall concentrate our stochastic analysis on the flow of measures $(\eta_n)_{n \geq 0}$.

To get one step further, we readily check that the flow of measures $(\eta_n)_{n \geq 0}$ satisfies the non-linear transport equation

$$\eta_{n+1} = \eta_n \mathcal{K}_{n+1, \eta_n} \quad \text{with} \quad \mathcal{K}_{n+1, \eta_n} := \mathcal{S}_{n, \eta_n} \mathcal{M}_{n+1}$$

We also recall that the solution $(\eta_n, \widehat{\eta}_n)$ of these recursive updating-prediction equations have the following functional representations

$$\eta_n(f_n) = \gamma_n(f_n) / \gamma_n(1) \quad \text{and} \quad \widehat{\eta}_n(f_n) = \widehat{\gamma}_n(f_n) / \widehat{\gamma}_n(1) \quad (3.3)$$

with the unnormalized Feynman-Kac measures γ_n and $\widehat{\gamma}_n$ defined by the formulae

$$\gamma_n(f_n) = \mathbb{E} [f_n(\mathcal{X}_n) \prod_{0 < k < n} \mathcal{G}_k(\mathcal{X}_k)] \quad \text{and} \quad \widehat{\gamma}_n(f_n) = \gamma_n(f_n \mathcal{G}_n) \quad (3.4)$$

3.2 Feynman-Kac semigroups

For further use in the convergence analysis of adaptive particle models, we present in this section the definition of the Feynman-Kac semigroup associated with the flow of measures $(\gamma_n)_{n \geq 0}$ and $(\eta_n)_{n \geq 0}$.

Firstly, we denote by $\mathcal{Q}_{n+1}(x_n, dx_{n+1})$ the bounded integral operator from S_n into S_{n+1} defined by the elementary multiplicative formula:

$$\mathcal{Q}_{n+1}(x_n, dx_{n+1}) = \mathcal{G}_n(x_n) \mathcal{M}_{n+1}(x_n, dx_{n+1})$$

We let $(\mathcal{Q}_{p,n})_{0 \leq p \leq n}$ be the corresponding linear semi-group defined by

$$\mathcal{Q}_{p,n} = \mathcal{Q}_{p+1} \mathcal{Q}_{p+2} \dots \mathcal{Q}_n$$

with the convention $\mathcal{Q}_{n,n} = I$, the identity operator. Notice that $\mathcal{Q}_{p,n}$ is alternatively defined by the Feynman-Kac product formula:

$$\mathcal{Q}_{p,n}(f_n)(x_p) = \mathbb{E} \left[f_n(\mathcal{X}_n) \prod_{p \leq k < n} \mathcal{G}_k(\mathcal{X}_k) \mid \mathcal{X}_p = x_p \right] \quad (3.5)$$

Using the Markov property, we prove that

$$\gamma_n(f_n) = \mathbb{E} \left[\mathbb{E} \left[f_n(\mathcal{X}_n) \prod_{p \leq k < n} \mathcal{G}_k(\mathcal{X}_k) \mid \mathcal{X}_p \right] \prod_{0 \leq k < p} \mathcal{G}_k(\mathcal{X}_k) \right] = \gamma_p(\mathcal{Q}_{p,n}(f_n))$$

The last assertions shows that $(\mathcal{Q}_{p,n})_{0 \leq p \leq n}$ is the semigroup associated with the unnormalized measures $(\gamma_n)_{n \geq 0}$.

Also denote the normalized version

$$\mathcal{P}_{p,n}(f_n) := \mathcal{Q}_{p,n}(f_n) / \mathcal{Q}_{p,n}(1)$$

Finally, we denote by $(\Phi_{p,n})_{0 \leq p \leq n}$ the non linear semi-group associated to the flow of normalized measures $(\eta_n)_{n \geq 0}$:

$$\Phi_{p,n} = \Phi_n \circ \dots \circ \Phi_{p+2} \circ \Phi_{p+1}$$

with the convention $\Phi_{n,n} = I$, the identity operator. From the previous discussion, we notice that $(\Phi_{p,n})_{0 \leq p \leq n}$ is alternatively defined in terms of $(\mathcal{Q}_{p,n})_{0 \leq p \leq n}$ with the following formulae:

$$\Phi_{p,n}(\eta_p)(f_n) = \frac{\gamma_p \mathcal{Q}_{p,n}(f_n)}{\gamma_p \mathcal{Q}_{p,n}(1)} = \frac{\eta_p \mathcal{Q}_{p,n}(f_n)}{\eta_p \mathcal{Q}_{p,n}(1)} \quad (3.6)$$

3.3 Mean field interpretation models

The mean field particle interpretation of the evolution equation (3.2) is the Markov chain

$$\mathcal{X}_n^{(N)} = (\mathcal{X}_n^{(N,1)}, \mathcal{X}_n^{(N,2)}, \dots, \mathcal{X}_n^{(N,N)}) \in S_n^N$$

with elementary transitions

$$\mathbb{P}(\mathcal{X}_{n+1}^{(N)} \in dx_{n+1} \mid \mathcal{X}_n^{(N)}) = \prod_{i=1}^N \mathcal{K}_{n+1, \eta_n^N}(\mathcal{X}_n^{(N,i)}, dx_{n+1}^i) \quad (3.7)$$

with

$$\mathcal{K}_{n+1, \eta_n^N} = \mathcal{S}_{n, \eta_n^N} \mathcal{M}_{n+1} \quad \text{and} \quad \eta_n^N := \frac{1}{N} \sum_{j=1}^N \delta_{\mathcal{X}_n^{(N,j)}}$$

Loosely speaking this integral decomposition shows that this particle model has the same updating/prediction nature as the one of the ‘limiting’ Feynman-Kac model. More precisely the deterministic two steps updating/prediction transitions in distribution spaces

$$\eta_n \xrightarrow{\mathcal{S}_{n, \eta_n}} \widehat{\eta}_n = \eta_n \mathcal{S}_{n, \eta_n} = \Psi_n(\eta_n) \xrightarrow{\mathcal{M}_{n+1}} \eta_{n+1} = \widehat{\eta}_n \mathcal{M}_{n+1} \quad (3.8)$$

have been replaced by a two steps selection/mutation transition in product spaces

$$\mathcal{X}_n^{(N)} \in S_n^N \xrightarrow{\text{selection}} \widehat{\mathcal{X}}_n^{(N)} \in S_n^N \xrightarrow{\text{mutation}} \mathcal{X}_{n+1}^{(N)} \in S_{n+1}^N \quad (3.9)$$

3.4 Path space and excursion models

We let X'_n be a Markov chain taking values in some measurable state spaces E'_n with elementary transitions

$$\mathbb{P}(X'_n \in dx'_n \mid X'_{n-1} = x'_{n-1}) := M'_n(x'_{n-1}, dx'_n)$$

and the initial distribution $\eta'_0 = \text{Law}(X'_0)$. We also consider a sequence of potential functions G'_n on the state spaces E'_n . To simplify the presentation, we assume that $G'_0 = 1$ and the potential functions G'_n take values in $]0, 1[$.

We associate with a prescribed non decreasing sequence of time parameters $(t_n)_{n \geq 0}$, the excursion valued random variables

$$X_0 = X'_0 \quad \text{and} \quad \forall n \geq 1 \quad X_n := (X'_{t_{n-1}+1}, X'_{t_{n-1}+2}, \dots, X'_{t_n})$$

and the random path sequences

$$\mathcal{X}_n := (X_0, \dots, X_n) = (X'_0, X'_1, \dots, X'_{t_n}) \in E_{t_n} := E'_0 \times \dots \times E'_{t_n}$$

Notice that \mathcal{X}_n forms a Markov chain

$$\mathcal{X}_{n+1} := (\mathcal{X}_n, X_{n+1}) = \left(\mathcal{X}_n, \left(X'_{t_n+1}, \dots, X'_{t_{n+1}} \right) \right) \quad (3.10)$$

taking values in the excursion spaces $S_n := E_{t_n}$. Furthermore, if we choose, in (3.4), the potential functions

$$\forall n \geq 1 \quad \mathcal{G}_n(\mathcal{X}_n) = G_n(X_n) := \prod_{t_{n-1} < k \leq t_n} G'_k(X'_k) \quad (3.11)$$

then we readily find that

$$\begin{aligned} \gamma_n(f_n) &= \mathbb{E} \left[f_n(\mathcal{X}_n) \prod_{0 < k < n} \mathcal{G}_k(\mathcal{X}_k) \right] \\ &= \mathbb{E} \left[f_n(X_0, \dots, X_n) \prod_{0 < k < n} G_k(X_k) \right] = \mathbb{E} \left[f_n(X'_0, \dots, X'_{t_n}) \prod_{0 < k \leq t_{n-1}} G'_k(X'_k) \right] \end{aligned}$$

The derivation of the updated versions of these formulae follows the same arguments.

These Feynman-Kac path distribution models are rather flexible. For instance, by definition of the potential functions \mathcal{G}_n of the excursion Feynman-Kac model (3.11), we readily prove that the condition (\mathcal{G}) introduced in (3.1) is satisfied, as soon as the condition (G') introduced in (2.6) holds true. More precisely, we have that

$$(G') \implies (\mathcal{G}) \text{ with } q_n \leq M' - \sup \left\{ \prod_{t_{n-1} < k \leq t_n} (G'_k(x'_k)/G'_k(y'_k)) \right\} \left(\leq \prod_{t_{n-1} < k \leq t_n} q'_k \right)$$

In the above displayed formula, the essential supremum $M' - \sup \{ \cdot \}$ is taken over all M'_k -admissible paths $x'_{t_{n-1}:t_n} = (x'_{t_{n-1}+1}, x'_{t_{n-1}+2}, \dots, x'_{t_n})$ and $y'_{t_{n-1}:t_n} = (y'_{t_{n-1}+1}, y'_{t_{n-1}+2}, \dots, y'_{t_n})$ of the underlying Markov chain X'_k .

On the other hand, the marginal $\bar{\gamma}_n$ of γ_n on the state spaces $\bar{S}_n := (E'_{t_n+1} \times \dots \times E'_{t_{n+1}})$ is again given by for any bounded measurable function \bar{f}_n on \bar{S}_n by

$$\bar{\gamma}_n(\bar{f}_n) := \mathbb{E} \left[\bar{f}_n(X_n) \prod_{0 < k < n} G_k(X_k) \right]$$

The mean field particle interpretations of the flow of normalized measures $\bar{\eta}_n$ follows the same construction as the one associated with the flow η_n . Furthermore, the path-particle model associated to the particle approximation of η_n on path space coincides with the genealogical tree evolution model associated with the evolutionary interpretation of the particle approximation model of $\bar{\eta}_n$. For a more thorough discussion on these genealogical tree based models, we refer the reader to [4].

3.5 Functional criteria models

Some precise definitions on the laws of the excursions and criteria of the adaptive sequential Monte Carlo method are now given.

Empirical adaptive criteria are expressed in terms of integrals of with to the occupation measures of the mean field particle model on excursion spaces. These occupation measures converge in some sense to the law of an excursion of the reference Markov chain X'_k starting with some updated Feynman-Kac distribution at some appropriate interaction time. These limiting objects are described by the following definition.

Definition 3.1 *For any pair of integers $0 \leq p \leq n$ and every $\eta \in \mathcal{P}(E_p)$, we denote by $\mathbb{P}_{\eta, (p, n)}$ the law of a random path sequence of length n , starting with an initial distribution η on the path space E_p and performing $(n - p)$ elementary moves with the local probability transitions M'_k , with $p < k \leq n$. More precisely, $\mathbb{P}_{\eta, (p, n)}$ is the measure defined by*

$$\begin{aligned} \mathbb{P}_{\eta, (p, n)}(d(x_p, x'_{p+1}, \dots, x'_n)) &:= \eta(dx_p) M'_{p+1}(x'_p, dx'_{p+1}) \dots M'_n(x'_{n-1}, dx'_n) \\ &\in \mathcal{P}(E_p \times (E'_{p+1} \times \dots \times E'_n)) = \mathcal{P}(E_n) \end{aligned}$$

In the above displayed formula, we have used the notation $dx_p = d(x'_0, \dots, x'_p)$ to denote an infinitesimal neighborhood of a path sequence

$$x_p = (x'_0, \dots, x'_p) \in E_p = (E'_0 \times \dots \times E'_p)$$

To connect the empirical adaptive criteria with their limiting functional versions, these criteria need to satisfy some weak regularity condition. Most of the entropy like criteria encountered in the literature are expressed in terms of a finite number of integrals with respect to the occupation measures of the samplers. Our analysis applies without further work to general models with countable or integral type criteria models. This class of stochastic models is described in the following definition.

Definition 3.2 *We consider a sequence of adaptive criteria type functionals*

$$\forall n \geq 0 \quad \forall p \leq q \quad \mathcal{H}_{p, q}^{(n)} : \mu \in \mathcal{P}(E_q) \mapsto \mathcal{H}_{p, q}^{(n)}(\mu) \in \mathbb{R}_+$$

satisfying the following Lipschitz type regularity condition

$$\left| \mathcal{H}_{p, q}^{(n)}(\mu_1) - \mathcal{H}_{p, q}^{(n)}(\mu_2) \right| \leq \int |\mu_1 - \mu_2|(h) |H_{p, q}^{(n)}(dh)| \quad (3.12)$$

for some collection of bounded measures $H_{p, q}^{(n)}$ on $\mathcal{B}(E_q)$ such that

$$\delta(H_{p, q}^{(n)}) := \sup_{h \in \text{Osc}_1(E_q)} \int \text{osc}(h) H_{p, q}^{(n)}(dh) < \infty$$

To illustrate this rather abstract construction, we provide four examples of adaptive criteria type functionals. All of them are defined in terms of the updating weight product potential function

$$W_{p,q} : (x'_0, \dots, x'_q) \in E_q \mapsto W_{p,q}(x'_0, \dots, x'_q) := \prod_{p < k \leq q} G'_k(x'_k) \in]0, 1[$$

In all the situations described below, the measures $H_{t_n, s}^{(n)}$, in (3.12), have a finite support with one or two elements.

1. **Weights thresholds control:** The functional defined below

$$\mathcal{H}_{p,q}^{(n)}(\mu) = \mu \left(A_{p,q}^{(n)} \right) \quad \text{with} \quad A_{p,q}^{(n)} := \{x_q \in E_q : W_{p,q}(x_q) \geq s_n\}$$

measures the proportion of the path sequences with a weight product larger than some parameter s_n .

2. **Partition functions:** In some sense, the functional defined below

$$\mathcal{H}_{p,q}^{(n)}(\mu) = \mu(W_{p,q})$$

measures the degeneracy of the partition functions for large values of the parameter $(q - p)$.

3. **Weights variances w.r.t. free motions:** The functional defined below

$$\mathcal{H}_{p,q}^{(n)}(\mu) = \mu \left(\left[\frac{W_{p,q}}{\mu(W_{p,q})} - 1 \right]^2 \right)$$

coincide with the variance of the weights with respect the reference measure μ .

4. **Boltzmann's entropy criteria:** The functional defined below

$$\mathcal{H}_{p,q}^{(n)}(\mu) = \text{Ent}(d\mu | W_{p,q} d\mu) := -\mu(\log W_{p,q})$$

measures the relative entropy distance between the reference measure μ and the updated weighted measure.

3.6 Adaptive time steps construction

Given a sequence of adaptive criteria type functionals $\mathcal{H}_{p,q}^{(n)}$, with $n \geq 0$ and $0 \leq p \leq q$, we define a non decreasing sequence of deterministic time steps $(t_n)_{n \geq 0}$ and a flow of Feynman-Kac measures $(\eta_n, \hat{\eta}_n)$ as follows:

- Set $n = 0$ and $t_n = 0$ and $\hat{\eta}_0 = \eta_0 = \eta'_0$.
- Suppose that we have defined the time step t_n and the pair of measures $(\eta_n, \hat{\eta}_n) \in \mathcal{P}(E_{t_n})^2$ at a given rank n . To define the time step t_{n+1} and the pair of measures $(\eta_{n+1}, \hat{\eta}_{n+1}) \in \mathcal{P}(E_{t_{n+1}})^2$, we consider a sequence of

adaptive criteria $\mathcal{H}_{t_n, s}^{(n)}$, with $s \geq t_n$ and a given interval $I_n \subset \mathbb{R}_+$ of the form $I_n = [a_n, \infty[$ or $I_n =]-\infty, a_n]$ such that

$$\mathcal{H}_{t_n, t_n}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, t_n)}) \notin I_n$$

We define the time step t_{n+1} as the first time $s \geq t_n$ the quantity $\mathcal{H}_{t_n, s}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)})$ hits the set I_n ; that is, we have that

$$t_{n+1} := \inf \left\{ t_n \leq s : \mathcal{H}_{t_n, s}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)}) \in I_n \right\}$$

Given that time step $t_{n+1} (> t_n)$, we set

$$\eta_{n+1} = \mathbb{P}_{\hat{\eta}_n, (t_n, t_{n+1})} \quad \text{and} \quad \hat{\eta}_{n+1} = \Psi_{n+1} (\eta_{n+1}) \quad (3.13)$$

with the Boltzmann-Gibbs transformation Ψ_{n+1} associated with the potential function $\mathcal{G}_{n+1} = W_{t_n, t_{n+1}}$.

We end this construction with an important observation. By definition of the Markov transition \mathcal{M}_{n+1} of the excursion model \mathcal{X}_n defined in section 3.4, we readily check that

$$\eta_{n+1} = \mathbb{P}_{\hat{\eta}_n, (t_n, t_{n+1})} = \hat{\eta}_n \mathcal{M}_{n+1} \quad (3.14)$$

This yields the recursion

$$(3.13) \quad \text{and} \quad (3.14) \implies \eta_{n+1} = \Psi_n (\eta_n) \mathcal{M}_{n+1}$$

From the above discussion, we find that the distributions η_n defined by the above inductive construction coincide with the Feynman-Kac distribution models presented in section 3.1. More precisely, the flow of measures η_n and $\hat{\eta}_n$ we have constructed inductively coincide with the Feynman-Kac distribution flow defined in (3.3), with the pair of Markov chain and potential function $(\mathcal{X}_n, \mathcal{G}_n)$ on excursion spaces defined in (3.10) and (3.11). The only difference between these models and the ones discussed in section 3.4 comes from the fact that the sequence of time parameters $(t_n)_{n \geq 0}$ is inductively defined in terms of a prescribed sequence of functional adaptive criteria. The mean field particle interpretation of these flows is discussed in section 5.

4 Some applications

In this section, we examine in some details the inductive construction of the deterministic time steps t_n introduced in section 3.6 in the four situations discussed in section 2.3.

4.1 Adaptive control of the weights

The time steps are constructed to control the degeneracy of the Feynman-Kac weights. To be more precise, we set

$$\mathcal{H}_{t_n, s}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)}) = \mathbb{E}_{n, \hat{\eta}_n} \left[1_{A_{t_n, s}^{(n)}} (X'_{t_n+1}, \dots, X'_s) \right] = \mathbb{P}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k (X'_k) \geq s_n \right)$$

with a ‘size’ parameter $s_n < 1$ and the sets

$$A_{t_n, s}^{(n)} := \left\{ (x'_{t_n+1}, \dots, x'_s) \in E'_{]t_n, s]} : \prod_{t_n < k \leq s} G'_k(x'_k) \geq s_n \right\}$$

In the above displayed formulae, $\mathbb{E}_{n, \hat{\eta}_n}$ is the expectation with respect to the law $\mathbb{P}_{n, \hat{\eta}_n}$ of a free evolution X'_k starting at time t_n at the end point X'_{t_n} of a random path (X_0, \dots, X_n) with distribution $\hat{\eta}_n$ on the space E_n .

For $[0, 1[$ -valued potential functions, the mappings

$$t_n \leq s \mapsto \mathbb{P}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \geq s_n \right)$$

are clearly non-increasing functions starting at 1, for $s = t_n$. If we choose an interval

$$I_n :=]-\infty, a_n] \quad \text{for some } a_n < 1$$

then the time step t_{n+1} coincides with the first time s the flow of probabilities

$$s \mapsto \mathbb{P}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \geq s_n \right)$$

go below the level a_n ; that is, we have that

$$t_{n+1} := \inf \left\{ t_n \leq s : \mathbb{P}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \geq s_n \right) \leq a_n \right\}$$

This adaptive time step strategy allows us to decompose the trajectories of the reference free motion X'_k into a sequence of excursion like processes

$$X_0 = X'_0 \quad \text{and} \quad \forall n \geq 1 \quad X_n := (X'_{t_{n-1}+1}, X'_{t_{n-1}+2}, \dots, X'_{t_n}) \quad (4.1)$$

with

$$\forall s < t_{n+1} \quad \mathbb{P}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \geq s_n \right) > a_n$$

and

$$s = t_{n+1} \implies \mathbb{P}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq t_{n+1}} G'_k(X'_k) \geq s_n \right) \leq a_n$$

We illustrate these adaptive time step splitting with the random walk model discussed in (2.9) (section 2.3.1). In this situation, we recall that X'_k is a Markov chain on $E'_k = \mathbb{Z}$ starting at the origin $X'_0 = 0$. If we choose the indicator potential functions

$$G'_k = 1_{[k, \infty[}$$

then we find that

$$\mathbb{E}_{0, \hat{\eta}_0} \left(\prod_{0 < k \leq t_1} G'_k(X'_k) \right) = \mathbb{P}(X'_1 \geq 1, \dots, X'_k \geq k)$$

When $a_0 = 1/2$, we find that t_1 coincides with the first level k the chain X'_k can reach with a probability less than $1/2$ with $X'_0 = 0$; more formally, we have

$$t_1 = \inf \{k \geq 0 : \mathbb{P}(X'_0 \geq 0, X'_1 \geq 1, \dots, X'_k \geq k) \leq 1/2\}$$

In much the same way, when $a_n = 1/2$, we find that t_{n+1} coincides with the first level k the chain X'_k can reach with a probability less than $1/2$ with $X'_{t_n} = x'_{t_n}$; more formally, we have

$$t_{n+1} = \inf \{k \geq t_n : \mathbb{P}_{n, \hat{\eta}_n}(X'_{t_n+1} \geq t_n + 1, \dots, X'_k \geq k) \leq 1/2\}$$

4.2 Normalizing constants

The time steps are built to control the degeneracy of the normalizing constants of a Feynman-Kac model. To be more precise, we set

$$\mathcal{H}_{t_n, s}^{(n)}(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}) = \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right) := \int \hat{\eta}_n(dx_{t_n}) \mathbb{E}_{t_n, x'_{t_n}} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)$$

Notice that $\mathcal{H}_{t_n, s}^{(n)}(\mathbb{P}_{\hat{\eta}_n, (t_n, s)})$ can be reformulated in terms of an integral with respect to the measure $\mathbb{P}_{\eta_n, (t_n, s)}$; that is we have that

$$\mathcal{H}_{t_n, s}^{(n)}(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}) = \int \hat{\eta}_n(dx_{t_n}) M'_{t_n+1}(x'_{t_n}, dx'_{t_n+1}) \dots M'_s(x'_{s-1}, dx'_s) \prod_{t_n < k \leq s} G'_k(x'_k)$$

Recalling the convention $\prod_{\emptyset} = 1$, we have

$$s = t_n \implies \mathcal{H}_{t_n, t_n}^{(n)}(\mathbb{P}_{\hat{\eta}_n, (t_n, t_n)}) = 1$$

When the potential functions take values in $[0, 1[$, one concludes that the mapping

$$s \geq t_n \mapsto \mathcal{H}_{t_n, s}^{(n)}(\mathbb{P}_{\hat{\eta}_n, (t_n, s)})$$

is clearly non-increasing starting at 1 at time $s = t_n$. If we choose an interval

$$I_n :=]-\infty, a_n] \quad \text{for some } a_n < 1$$

then the time step t_{n+1} coincide with the first time s the quantities

$$s \mapsto \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)$$

go below the level a_n ; that is, we have that

$$t_{n+1} := \inf \left\{ t_n \leq s : \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right) \leq a_n \right\}$$

The adaptive time step strategy discussed above permits us to decompose the trajectories of the reference free motion X'_k into a sequence of excursion like processes

$$X_0 = X'_0 \quad \text{and} \quad \forall n \geq 1 \quad X_n := (X'_{t_{n-1}+1}, X'_{t_{n-1}+2}, \dots, X'_{t_n}) \quad (4.2)$$

with

$$\forall s < t_{n+1} \quad \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right) > a_n \quad \text{and} \quad \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq t_{n+1}} G'_k(X'_k) \right) \leq a_n$$

4.3 Weights variances criteria w.r.t. free motions

The time steps are built to control the weights variances with respect to the law of the free motions trajectories. To be more precise, we set

$$\begin{aligned} \mathcal{H}_{t_n, s}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)}) &= \mathbb{E}_{n, \hat{\eta}_n} \left(\left(\frac{\prod_{t_n < k \leq s} G'_k(X'_k)}{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)} - 1 \right)^2 \right) \\ &= \frac{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k)^2 \right)}{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)^2} - 1 \end{aligned}$$

Recalling the convention $\prod_{\emptyset} = 1$, we already notice that

$$s = t_n \implies \mathcal{H}_{t_n, t_n}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, t_n)}) = 0$$

The variance mappings $s \mapsto \mathcal{H}_{t_n, s}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)})$ are generally increasing. For instance, in the case of indicator potential functions $G'_k = 1_{A'_k}$ we have

$$\forall t_n \leq s \leq r \quad \mathcal{H}_{t_n, s}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)}) = \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)^{-1} - 1 \leq \mathcal{H}_{t_n, r}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, r)})$$

For another example in filtering problems, we refer the reader to [5]. One natural way to control these variances is to choose an interval

$$I_n := [a_n, \infty[\quad \text{for some } a_n > 0$$

then the time step t_{n+1} coincide with the first time s the functions

$$s \mapsto \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k)^2 \right) / \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)^2$$

go above the level $1 + a_n$; that is, we have that

$$t_{n+1} := \inf \left\{ t_n \leq s : \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k)^2 \right) \geq [1 + a_n] \mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)^2 \right\}$$

The adaptive time step strategy discussed above allows us to decompose the trajectories of the reference free motion X'_k into a sequence of excursion like processes

$$X_0 = X'_0 \quad \text{and} \quad \forall n \geq 1 \quad X_n := (X'_{t_{n-1}+1}, X'_{t_{n-1}+2}, \dots, X'_{t_n}) \quad (4.3)$$

with controlled Feynman-Kac weights variances

$$\forall t_n \leq s < t_{n+1} \quad \frac{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k)^2 \right)}{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq s} G'_k(X'_k) \right)^2} < 1 + a_n$$

and for $s = t_{n+1}$

$$\frac{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq t_{n+1}} G'_k(X'_k)^2 \right)}{\mathbb{E}_{n, \hat{\eta}_n} \left(\prod_{t_n < k \leq t_{n+1}} G'_k(X'_k) \right)^2} = \frac{\mathbb{E}_{n, \hat{\eta}_n} (G_{n+1}(X_{n+1})^2)}{\mathbb{E}_{n, \hat{\eta}_n} (G_{n+1}(X_{n+1}))^2} \geq 1 + a_n$$

4.4 Boltzmann's entropy criteria

The final adaptive time step strategy presented in this section allows us to control the entropy like distance between the free motion trajectories and the weighted Feynman-Kac measures. To be more precise, we set

$$\mathcal{H}_{t_n, s}^{(n)} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)} | \mathbb{Q}_{\hat{\eta}_n, (t_n, s)}) = - \sum_{t_n < k \leq s} \mathbb{E}_{n, \hat{\eta}_n} (\log G'_k(X'_k))$$

with the weighted measures $\mathbb{Q}_{\eta, (p, n)}$ defined by

$$\mathbb{Q}_{\eta, (p, n)}(d(x_p, x'_{p+1}, \dots, x'_n)) = \mathbb{P}_{\eta, (p, n)}(d(x_p, x'_{p+1}, \dots, x'_n)) \times \prod_{p < k \leq n} G'_k(x'_k)$$

For $]0, 1[$ -valued potential functions, the mappings

$$t_n \leq s \mapsto \text{Ent} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)} | \mathbb{Q}_{\hat{\eta}_n, (t_n, s)})$$

are clearly non decreasing functions starting at 0, for $s = t_n$. If we choose an interval

$$I_n := [a_n, \infty[\quad \text{for some } a_n > 0$$

then the time step t_{n+1} coincide with the first time s the entropy distances go above the level a_n ; that is, we have that

$$t_{n+1} := \inf \{ t_n \leq s : \text{Ent} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)} | \mathbb{Q}_{\hat{\eta}_n, (t_n, s)}) \geq a_n \}$$

This adaptive time step strategy allows to decompose the trajectories of the reference free motion X'_k into a sequence of excursion like processes

$$X_0 = X'_0 \quad \text{and} \quad \forall n \geq 1 \quad X_n := (X'_{t_{n-1}+1}, X'_{t_{n-1}+2}, \dots, X'_{t_n}) \quad (4.4)$$

with

$$\forall s < t_{n+1} \quad \text{Ent} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)} | \mathbb{Q}_{\hat{\eta}_n, (t_n, s)}) < a_n \quad \text{and} \quad \text{Ent} (\mathbb{P}_{\hat{\eta}_n, (t_n, s)} | \mathbb{Q}_{\hat{\eta}_n, (t_n, s)}) \geq a_n$$

5 Adaptive particle excursion models

This section is concerned with the mean field Feynman-Kac particle models associated with adaptive excursion spaces presented in section 3.6. We further assume that the adaptive time steps $(t_n)_{n \geq 0}$ defined in section 3.6 are explicitly known. Of course, this hypothesis is generally never satisfied. One objective of this section is provide a set of conditions under which this hypothesis can be relaxed. More precisely, our ultimate objective is to design a feasible mean field particle model with online adaptive time steps. This model is described in section 5.4.

As we mentioned in the introduction, the reference mean field particle model associated with the adaptive time steps $(t_n)_{n \geq 0}$ is an ‘ideal’ sequential Monte Carlo model that updates the samples at adaptive and deterministic times steps that are generally unknown but which can be approximated. The concentration analysis of these reference models is developed in section 5.2. In section 5.2.1, we provide a brief discussion on the set of constants arising in our estimates. Special mention is made to uniform control of these parameters with respect to the time horizon. In section 5.2.2, we provide several estimates on the bias and on the \mathbb{L}_m -mean errors. The section 5.2.3 provides rather sharp and non asymptotic exponential concentration estimates. In section 5.3, we use these concentration results to analyze the convergence of empirical adaptive criteria towards their limiting functional values.

5.1 A reference particle model

We recall that the mean particle interpretations (3.9) of the Feynman-Kac path measures (see section 3.4) is a mean field particle model with a two step updating-selection and mutation-exploration transition:

$$\mathcal{X}_n^{(N)} \in E_{t_n}^N \xrightarrow{\text{updating}} \widehat{\mathcal{X}}_n^{(N)} \in E_{t_n}^N \xrightarrow{\text{exploration}} \mathcal{X}_{n+1}^{(N)} \in E_{t_{n+1}}^N \quad (5.1)$$

In our context, the stochastic algorithm keeps track of all the paths of the sampled particles, so that the mean field particle algorithm can be interpreted as a genealogical tree based sampler. The corresponding ancestral lines are denoted as follows.

$$\begin{aligned} \widehat{\mathcal{X}}_n^{(N,i)} &= \left(\widehat{X}_{0,t_n}^{(N,i)}, \widehat{X}_{1,t_n}^{(N,i)}, \dots, \widehat{X}_{t_n,t_n}^{(N,i)} \right) \\ \mathcal{X}_n^{(N,i)} &= \left(X_{0,t_n}^{(N,i)}, X_{1,t_n}^{(N,i)}, \dots, X_{t_n,t_n}^{(N,i)} \right) \in E_{t_n} := (E'_0 \times \dots \times E'_{t_n}) \end{aligned}$$

Notice that the state space at the $(n+1)$ -th iteration is the path space given by

$$E_{t_{n+1}} := \left(E'_0 \times \dots \times E'_{t_{n+1}} \right) = E_{t_n} \times \left(E'_{t_n+1} \times \dots \times E'_{t_{n+1}} \right)$$

By definition of the reference Markov model \mathcal{X}_n given in (3.10), every path particle $\mathcal{X}_{n+1}^{(N,i)} \in E_{t_{n+1}}$ keeps track of the selected excursion $\widehat{\mathcal{X}}_n^{(N,i)} \in E_{t_n}$ and it evolves its terminal state $\widehat{X}_{t_n,t_n}^{(N,i)}$ with $(t_{n+1} - t_n)$ elementary random moves

with the Markov transitions M'_k . More formally, we have that

$$\begin{aligned} \mathcal{X}_{n+1}^{(N,i)} &= \left(\underbrace{\left(X_{0,t_{n+1}}^{(N,i)}, X_{1,t_{n+1}}^{(N,i)}, \dots, X_{t_n,t_{n+1}}^{(N,i)} \right)}_{\widehat{\mathcal{X}}_n^{(N,i)}}, \underbrace{\left(X_{t_{n+1},t_{n+1}}^{(N,i)}, \dots, X_{t_{n+1},t_{n+1}}^{(N,i)} \right)}_{\left(X_{t_{n+1}}^{(N,i)}, \dots, X_{t_{n+1}}^{(N,i)} \right)} \right) \\ &= \left(\widehat{\mathcal{X}}_n^{(N,i)}, \left(X_{t_{n+1}}^{(N,i)}, \dots, X_{t_{n+1}}^{(N,i)} \right) \right) \end{aligned}$$

with the $(t_{n+1} - t_n)$ M'_k -exploration transitions given by the following synthetic diagram:

$$\widehat{X}_{t_n,t_n}^{(N,i)} \xrightarrow{M'_{t_n+1}} X_{t_{n+1}}^{(N,i)} \xrightarrow{M'_{t_n+2}} X_{t_{n+2}}^{(N,i)} \xrightarrow{M'_{t_n+3}} \dots \xrightarrow{M'_{t_n+1}} X_{t_{n+1}}^{(N,i)}$$

From this discussion, it is worth mentioning a further convention that the particle empirical measures

$$\eta_{n+1}^N = \frac{1}{N} \sum_{i=1}^N \delta_{\left(\widehat{\mathcal{X}}_n^{(N,i)}, \left(X_{t_{n+1}}^{(N,i)}, X_{t_{n+2}}^{(N,i)}, \dots, X_{t_{n+1}}^{(N,i)} \right) \right)}$$

are the terminal values at time $s = t_{n+1}$ of the flow of random measures

$$t_n \leq s \leq t_{n+1} \mapsto \mathbb{P}_{\widehat{\eta}_n^N, (t_n, s)}^N = \frac{1}{N} \sum_{i=1}^N \delta_{\left(\widehat{\mathcal{X}}_n^{(N,i)}, \left(X_{t_{n+1}}^{(N,i)}, X_{t_{n+2}}^{(N,i)}, \dots, X_s^{(N,i)} \right) \right)} \quad (5.2)$$

5.2 Concentration analysis

5.2.1 Introduction

This section is concerned with the concentration analysis of the empirical measures η_n^N associated with an abstract mean-field particle approximation model (3.7) around their limiting values η_n defined in (3.3). Our concentration estimates are expressed in terms of the following constants:

$$q_{p,n} = \sup_{(x,y) \in S_p^2} \frac{\mathcal{Q}_{p,n}(1)(x)}{\mathcal{Q}_{p,n}(1)(y)} \quad \text{and} \quad \beta(\mathcal{P}_{p,n}) := \sup_{f \in \text{Osc}_1(S_n)} \text{osc}(\mathcal{P}_{p,n}(f))$$

We mention that these parameters can be expressed in terms of the mixing properties of the reference free motion associated with the Markov transitions \mathcal{M}_n . It is clearly out of the scope of the present article to derive these estimates, we refer the interested reader to chapter 4 of [4]. Nevertheless, under appropriate mixing type properties one can prove that the series $\sum_{p=0}^n q_{p,n}^\alpha \beta(\mathcal{P}_{p,n})$ is uniformly bounded with respect to the final time horizon n , for any parameter $\alpha \geq 0$. Most of the results presented in this section are expressed in terms of these series, including bias estimates, \mathbb{L}_p -mean error bounds, as well as exponential concentration estimates. As a result, these non-asymptotic results can be converted into uniform in time convergence results. To get a flavor of these uniform estimates, let us suppose that the Markov transitions \mathcal{M}_k satisfy the following regularity property

$(\mathcal{M})_m$ There exists an integer $m \geq 1$ and a sequence $(\delta_p)_{p \geq 0} \in]0, 1[^\mathbb{N}$ such that

$$\forall p \geq 0 \quad \forall (x, y) \in S_p^2 \quad \mathcal{M}_{p,p+m}(x, \cdot) \geq \delta_p \mathcal{M}_{p,p+m}(y, \cdot)$$

To get one step further in our discussion, we introduce the following quantities:

$$\forall k \leq l \quad r_{k,l} := \sup \prod_{k \leq p < l} \frac{\mathcal{G}_p(x_p)}{\mathcal{G}_p(y_p)} \left(\leq \prod_{k \leq p < l} q_p \right)$$

with the collection of constants $(q_n)_{n \geq 1}$ introduced in (3.1). In the above displayed formula, the supremum stands for the \mathcal{M} -essential supremum, in the sense that it is taken over all admissible pair of paths $(x_p)_{p \geq 0}$ and $(y_p)_{p \geq 0}$ with elementary transitions \mathcal{M}_p .

Under the condition $(\mathcal{M})_m$ we have for any $n \geq m \geq 1$, and $p \geq 1$,

$$q_{p,p+n} \leq \delta_p^{-1} r_{p,p+m} \quad \text{and} \quad \beta(\mathcal{P}_{p,p+n}) \leq \prod_{k=0}^{\lfloor n/m \rfloor - 1} \left(1 - \delta_p^2 r_{p+km+1, p+(k+1)m}^{-1} \right) \quad (5.3)$$

The proof of these crude estimates relies on technical semigroup techniques; see chapter 4 of [4] for details. Several contraction inequalities can be deduced from these estimates. To give a flavour of these results, we further assume that $(\mathcal{M})_m$ is satisfied with $m = 1$, and we have $\delta = \wedge_n \delta_n > 0$ and $q = \vee_{n \geq 1} q_n$. In this case, we can check that

$$q_{p,p+n} \leq \delta^{-1} q \quad \text{and} \quad \beta(\mathcal{P}_{p,p+n}) \leq (1 - \delta^2)^n \implies \forall \alpha \geq 0 \quad \sum_{p=0}^n q_{p,n}^\alpha \beta(\mathcal{P}_{p,n}) \leq q^\alpha / \delta^{(2+\alpha)}$$

More generally, suppose that $(\mathcal{M})_m$ is satisfied for some $m \geq 1$. Also suppose that the collection parameters δ_p and $r_{k,l}$ are such that

$$\wedge_p \delta_p := \delta > 0 \quad \vee_p r_{p,p+m} := \bar{r} < \infty \quad \text{and} \quad \vee_p r_{p+1, p+m} := \underline{r} < \infty \quad (5.4)$$

In this situation, we have

$$q_{p,p+n} \leq \delta^{-1} \bar{r} \quad \text{and} \quad \beta(\mathcal{P}_{p,p+n}) \leq (1 - \delta^2 \underline{r}^{-1})^{\lfloor n/m \rfloor}$$

and therefore

$$\forall \alpha \geq 0 \quad \sum_{p=0}^n q_{p,n}^\alpha \beta(\mathcal{P}_{p,n}) \leq m \underline{r} \bar{r}^\alpha / \delta^{(2+\alpha)} \quad (5.5)$$

It is well known that the mixing type condition $(\mathcal{M})_m$ is satisfied for any aperiodic and irreducible Markov chains on finite spaces, as well as for bi-Laplace exponential transitions associated with a bounded drift function. We also mention that this mixing condition is of course never met for the complete historical process $\mathcal{X}_n = (X'_p)_{0 \leq p \leq t_n}$ on $S_n = \prod_{0 \leq p \leq t_n} E'_p$ discussed in section 3.4.

Nevertheless, under appropriate conditions on the Markov transitions M'_k , it is satisfied for the time marginal model associated with the excursion valued Markov chain model given by

$$\mathcal{X}_n := (X'_p)_{t_{n-1} < p \leq t_n} \in S_n := \prod_{t_{n-1} < p \leq t_n} E'_p$$

For instance, if we have

$$\forall k \geq 1 \quad \forall (x, y) \in (E'_p)^2 \quad M'_k(x, \cdot) \geq \delta' M'_k(y, \cdot)$$

for some $\delta' > 0$, then condition $(\mathcal{M})_m$ is met with $m = 1$ and $\delta_p = \delta'$.

5.2.2 Some \mathbb{L}_m -mean error bounds

At this point, it is convenient to observe that the local sampling errors induced by the mean field particle model are expressed in terms of the collection of local random fields models defined below.

Definition 5.1 *For any $n \geq 0$ and any $N \geq 1$, we let V_n^N be the collection of random fields defined by the following stochastic perturbation formulae*

$$\eta_n^N = \eta_{n-1}^N \mathcal{K}_{n, \eta_{n-1}^N} + \frac{1}{\sqrt{N}} V_n^N \quad \left(\iff V_n^N := \sqrt{N} \left[\eta_n^N - \eta_{n-1}^N \mathcal{K}_{n, \eta_{n-1}^N} \right] \right) \quad (5.6)$$

For $n = 0$, we use the conventions $K_{0, \eta_{-1}^N}(x, dy) = \eta_0(dy)$ and $\eta_{-1}^N \mathcal{K}_{0, \eta_{-1}^N} = \eta_0$.

In order to quantify high order \mathbb{L}_m -mean errors we need the following Kintchine type inequality for martingales with symmetric and independent increments. This result is more or less well known. For the convenience of the reader, a proof is provided in the appendix.

Lemma 5.2 (Kintchine's inequality) *Let $M_n := \sum_{0 \leq p \leq n} \Delta_p$ be a real-valued martingale with symmetric and independent increments $(\Delta_n)_{n \geq 0}$. For any integer $m \geq 1$, and any $n \geq 0$, we have*

$$\mathbb{E}(|M_n|^m)^{\frac{1}{m}} \leq b(m) \mathbb{E} \left([M]_n^{m'/2} \right)^{1/m'} \quad \text{with} \quad [M]_n := \sum_{0 \leq p \leq n} \Delta_p^2 \quad (5.7)$$

In the above display, m' stands for the smallest even integer $m' \geq m$ and $(b(m))_{m \geq 1}$ is the collection of constants given below:

$$b(2m)^{2m} := (2m)_m 2^{-m} \quad \text{and} \quad b(2m+1)^{2m+1} := \frac{(2m+1)_{(m+1)}}{\sqrt{m+1/2}} 2^{-(m+1/2)} \quad (5.8)$$

Proposition 5.3 *For any $N \geq 1$, $m \geq 1$, $n \geq 0$ and any test function $f_n \in \mathcal{B}_b(S_n)$ we have the almost sure estimate*

$$\mathbb{E} \left(|V_n^N(f_n)|^m \middle| \mathcal{F}_{n-1}^{(N)} \right)^{\frac{1}{m}} \leq b(m) \text{osc}(f_n) \quad (5.9)$$

Proof:

By construction, we have

$$V_n^N(f_n) = \sum_{i=1}^N \Delta_{n,i}^{(N)}(f_n) \quad \text{with} \quad \Delta_{n,i}^{(N)}(f_n) := \frac{1}{\sqrt{N}} \left[f_n(\mathcal{X}_n^{(N,i)}) - \mathcal{K}_{n, \eta_{n-1}^N}(f_n)(\mathcal{X}_{n-1}^{(N,i)}) \right]$$

Given $\mathcal{X}_{n-1}^{(N)}$, we let $(\mathcal{Y}_n^{(N,i)})_{1 \leq i \leq N}$ be an independent copy of $(\mathcal{X}_n^{(N,i)})_{1 \leq i \leq N}$. We readily check that

$$\Delta_{n,i}^{(N)}(f_n) := \mathbb{E} \left(\frac{1}{\sqrt{N}} \left[f_n(\mathcal{X}_n^{(N,i)}) - f_n(\mathcal{Y}_n^{(N,i)}) \right] \mid \mathcal{X}_{n-1}^{(N)} \right)$$

This yields the formula

$$V_n^N(f_n) = \mathbb{E} \left(M_{n,N}^{(N)}(f_n) \mid \mathcal{X}_{n-1}^{(N)} \right)$$

with the terminal value $M_{n,N}^{(N)}(f_n)$ of the martingale sequence defined below

$$i \in \{1, \dots, N\} \mapsto M_{n,i}^{(N)}(f_n) := \sum_{j=1}^i \frac{1}{\sqrt{N}} \left[f_n(\mathcal{X}_n^{(N,j)}) - f_n(\mathcal{Y}_n^{(N,j)}) \right]$$

Using Kintchine's inequality the end of the proof of (5.9) is clear. \blacksquare

The proof of the following lemma is rather technical. Thus, it is provided in the appendix.

Lemma 5.4 *For any $0 \leq p \leq n$, any $\eta, \mu \in \mathcal{P}(S_p)$ and any $f_n \in \text{Osc}_1(S_n)$, we have the first order decomposition:*

$$[\Phi_{p,n}(\mu) - \Phi_{p,n}(\eta)](f_n) = 2 q_{p,n} \beta(\mathcal{P}_{p,n}) [\mu - \eta](\mathcal{U}_{p,n,\eta}(f_n)) + \mathcal{R}_{p,n}(\mu, \eta)(f_n)$$

with a remainder term

$$|\mathcal{R}_{p,n}(\mu, \eta)(f_n)| \leq 4 q_{p,n}^3 \beta(\mathcal{P}_{p,n}) \left[|\mu - \eta](\mathcal{V}_{p,n,\eta}(f))| \times |\mu - \eta](\mathcal{W}_{p,n,\eta}(f_n))| \right]$$

and a collection of functions $\mathcal{U}_{p,n,\eta}(f)$, $\mathcal{V}_{p,n,\eta}(f)$, $\mathcal{W}_{p,n,\eta}(f)$ in $\text{Osc}_1(S_p)$, whose values only depends on the parameters (p, n, η) .

We now present a bias estimate and some \mathbb{L}_m bounds of independent interest.

Theorem 5.5 *For any $n \geq 0$, $f_n \in \text{Osc}_1(S_n)$, and any $N \geq 1$, we have*

$$N \left| \mathbb{E}(\eta_n^N(f_n)) - \eta_n(f_n) \right| \leq \sigma_{1,n} \quad \text{with} \quad \sigma_{1,n} := 4 \sum_{p=0}^n q_{p,n}^3 \beta(\mathcal{P}_{p,n})$$

In addition, for any $m \geq 1$ we have

$$\sqrt{N} \mathbb{E} \left(\left| [\eta_n^N - \eta_n](f_n) \right|^m \right)^{\frac{1}{m}} \leq \frac{1}{\sqrt{N}} b(2m)^2 \sigma_{1,n} + b(m) \sigma_{2,n}$$

with

$$\sigma_{2,n} := 2 \sum_{p=0}^n q_{p,n} \beta(\mathcal{P}_{p,n})$$

Proof:

Using lemma 5.4, we have the telescoping sum decomposition

$$\begin{aligned} W_n^N &:= \sqrt{N} [\eta_n^N - \eta_n] \\ &= \sqrt{N} \sum_{p=0}^n [\Phi_{p,n}(\eta_p^N) - \Phi_{p,n}(\Phi_p(\eta_{p-1}^N))] = \mathcal{I}_n^N + \mathcal{J}_n^N \end{aligned}$$

with the pair of random measures $(\mathcal{I}_n^N, \mathcal{J}_n^N)$ given for any $f_n \in \text{Osc}_1(S_n)$ by

$$\begin{aligned} \mathcal{I}_n^N(f_n) &:= 2 \sum_{p=0}^n q_{p,n} \beta(\mathcal{P}_{p,n}) V_p^N(\mathcal{U}_{p,n,\Phi_p(\eta_{p-1}^N)}(f_n)) \\ \mathcal{J}_n^N(f_n) &:= \sqrt{N} \sum_{p=0}^n \mathcal{R}_{p,n}(\eta_p^N, \Phi_p(\eta_{p-1}^N))(f_n) \end{aligned}$$

At this point, we observe that

$$\mathbb{E}(W_n^N(f_n)) = \mathbb{E}(\mathcal{J}_n^N(f_n))$$

Using proposition 5.3, we readily check that for any $f_n \in \text{Osc}_1(S_n)$

$$\mathbb{E}\left(|\mathcal{I}_n^N(f_n)|^m\right)^{\frac{1}{m}} \leq b(m) \sigma_{2,n}$$

In much the same way, we find that

$$\sqrt{N} \mathbb{E}\left(|\mathcal{J}_n^N(f_n)|^m\right)^{\frac{1}{m}} \leq b(2m)^2 \sigma_{1,n} \quad (5.10)$$

The end of the proof is now clear. \blacksquare

5.2.3 A concentration theorem

The following concentration theorem is the main result of this section. It improves the exponential estimates presented in section 7.4.3 in [4], by removing a multiplicative correction factor of order \sqrt{N} in front of the exponential decay.

Theorem 5.6 *For any $n \geq 0$, $f_n \in \text{Osc}_1(S_n)$, $N \geq 1$, and any $0 \leq \epsilon \leq 1/2$, we have*

$$\mathbb{P}\left(|[\eta_n^N - \eta_n](f_n)| \geq \epsilon\right) \leq 6 \exp\left(-\frac{N \epsilon^2}{8 \sigma_{1,n}}\right) \quad (5.11)$$

and the constant $\sigma_{1,n}$ defined in theorem 5.5.

In addition, suppose that $(\mathcal{M})_m$ is satisfied for some $m \geq 1$ and condition (5.4) holds true for some $\delta > 0$ and some finite constants (\underline{r}, \bar{r}) . In this situation, for any value of the time parameter n , for any $f_n \in \text{Osc}_1(S_n)$, $N \geq 1$ and for any $\rho \in]0, 1[$ the probability that

$$|[\eta_n^N - \eta_n](f_n)| \leq \frac{4\bar{r}}{\delta^2} \sqrt{\frac{2m\bar{r}}{N\delta} \log\left(\frac{6}{\rho}\right)}$$

is greater than $(1 - \rho)$.

Proof:

We use the same notation as in the proof of theorem 5.5. We recall that $b(2m)^{2m} = \mathbb{E}(X^{2m})$ for every Gaussian centered random variable with $\mathbb{E}(X^2) = 1$ and

$$\forall s \in [0, 1/2[\quad \mathbb{E}(\exp\{sX^2\}) = \sum_{m \geq 0} \frac{s^m}{m!} b(2m)^{2m} = \frac{1}{\sqrt{1-2s}}$$

Using (5.10), for any $f_n \in \text{Osc}_1(S_n)$ and $0 \leq s < 1/(2\sigma_{1,n})$ we find that

$$\mathbb{E}\left(\exp\left\{s\sqrt{N}\mathcal{J}_n^N(f_n)\right\}\right) \leq \sum_{m \geq 0} \frac{(s\sigma_{1,n})^m}{m!} b(2m)^{2m} = \frac{1}{\sqrt{1-2s\sigma_{1,n}}} \quad (5.12)$$

To simplify the presentation, we set

$$f_{p,n}^N := \mathcal{U}_{p,n,\Phi_p(\eta_{p-1}^N)}(f_n) \quad \text{and} \quad \alpha_{p,n} := 2 q_{p,n} \beta(\mathcal{P}_{p,n})$$

In this notation, we find that

$$V_p^N(f_{p,n}^N) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \left(f_{p,n}^N(\mathcal{X}_p^{(N,i)}) - \mathcal{K}_{p,\eta_{p-1}^N}(f_{p,n}^N)(\mathcal{X}_{p-1}^{(N,i)}) \right)$$

We recall that $\mathbb{E}(e^{tX}) \leq e^{t^2 c^2/2}$, for every real-valued and centered random variable X with $|X| \leq c$. To prove this assertion, we use the convexity property of the exponential function, and the fact that $\cosh(t) \leq e^{t^2/2}$ to check that for any $|x| \leq 1$

$$e^{tx} \leq \frac{x+1}{2} e^t + \frac{1-x}{2} e^{-t} = x \sinh(t) + \cosh(t) \Rightarrow \mathbb{E}(e^{tX/c}) \leq e^{t^2/2}$$

The desired inequality is proved by replacing t by ct . From this elementary inequality, we prove that

$$\begin{aligned} & \mathbb{E}\left(\exp\left\{t \alpha_{p,n} V_p^N(f_{p,n}^N)\right\} \middle| \mathcal{X}_{p-1}^{(N)}\right) \\ &= \prod_{i=1}^N \int_{S_p} \mathcal{K}_{p,\eta_{p-1}^N}(\mathcal{X}_{p-1}^{(N,i)}, dx) e^{\frac{t\alpha_{p,n}}{\sqrt{N}} \left(f_{p,n}^N(x) - \mathcal{K}_{p,\eta_{p-1}^N}(f_{p,n}^N)(\mathcal{X}_{p-1}^{(N,i)}) \right)} \leq \exp\left(\frac{t^2 \alpha_{p,n}^2}{2}\right) \end{aligned}$$

Iterating the argument, we find that

$$\mathbb{E}\left(e^{t \mathcal{I}_n^N(f_n)}\right) = \mathbb{E}\left(\exp\left\{t \sum_{p=0}^n \alpha_{p,n} V_p^N(f_{p,n}^N)\right\}\right) \leq \exp\left(\frac{t^2 \sigma_n^2}{2}\right) \quad (5.13)$$

with

$$\sigma_n^2 := 4 \sum_{p=0}^n q_{p,n}^2 \beta(\mathcal{P}_{p,n})^2$$

From these Laplace upper bounds, the proof of the exponential estimates now follows standard arguments. Indeed, for any $0 \leq s < 1/(2\sigma_{1,n})$ and any $\epsilon > 0$, by (5.12) we have

$$\mathbb{P}\left(\sqrt{N} \mathcal{J}_n^N(f_n) \geq \epsilon\right) \leq \frac{1}{\sqrt{1-2s\sigma_{1,n}}} \exp\{-\epsilon s\}$$

Replacing ϵ by ϵN we prove

$$\mathbb{P}\left(\mathcal{J}_n^N(f_n)/\sqrt{N} \geq \epsilon\right) \leq \frac{1}{\sqrt{1-2s\sigma_{1,n}}} \exp\{-\epsilon s N\}$$

Now, if we choose $s = 3/(8\sigma_{1,n})$, then we find that

$$\mathbb{P}\left(\mathcal{J}_n^N(f_n)/\sqrt{N} \geq \epsilon\right) \leq 2 \exp\{-\epsilon N/(3\sigma_{1,n})\}$$

To estimate the probability tails of $\mathcal{I}_n^N(f_n)$, we use (5.13) and the fact that $\epsilon > 0$ and $t \geq 0$

$$\mathbb{P}\left(\mathcal{I}_n^N(f_n) \geq \epsilon\right) \leq \exp\left\{-\left(\epsilon t - \frac{t^2}{2} \sigma_n^2\right)\right\}$$

If we choose $t = \epsilon/\sigma_n^2$ and then replace ϵ by $\sqrt{N}\epsilon$ then we obtain

$$\forall \epsilon > 0 \quad \mathbb{P}\left(\mathcal{I}_n^N(f)/\sqrt{N} \geq \epsilon\right) \leq \exp\left(-\frac{N\epsilon^2}{2\sigma_n^2}\right)$$

Using the decomposition

$$[\eta_n^N - \eta_n] = \mathcal{I}_n^N/\sqrt{N} + \mathcal{J}_n^N/\sqrt{N}$$

we find that for any parameter $\alpha \in [0, 1]$ we have

$$\mathbb{P}\left([\eta_n^N - \eta_n](f_n) \geq \epsilon\right) \leq \mathbb{P}\left(\mathcal{I}_n^N(f_n)/\sqrt{N} \geq \alpha\epsilon\right) + \mathbb{P}\left(\mathcal{J}_n^N(f_n)/\sqrt{N} \geq (1-\alpha)\epsilon\right)$$

From previous calculations, we find that

$$\mathbb{P}\left([\eta_n^N - \eta_n](f_n) \geq \epsilon\right) \leq \exp\left(-\frac{N\epsilon^2\alpha^2}{2\sigma_n^2}\right) + 2 \exp\left(-\frac{N\epsilon(1-\alpha)}{3\sigma_{1,n}}\right) \quad (5.14)$$

If we choose $\alpha = (1-\epsilon)(\geq 1/2)$, we have $\alpha^2 \geq 1/4$ and

$$\mathbb{P}\left([\eta_n^N - \eta_n](f_n) \geq \epsilon\right) \leq \exp\left(-\frac{N\epsilon^2}{8\sigma_n^2}\right) + 2 \exp\left(-\frac{N\epsilon^2}{3\sigma_{1,n}}\right) \leq 3 \exp\left(-\frac{N\epsilon^2}{8(\sigma_{1,n} \vee \sigma_n^2)}\right)$$

It remains to observe that

$$q_{p,n} \geq 1 \quad \text{and} \quad \beta(\mathcal{P}_{p,n}) \leq 1 \implies \sigma_n^2 \leq \sigma_{1,n}$$

and

$$\begin{aligned} |[\eta_n^N - \eta_n](f_n)| \geq \epsilon &\iff [\eta_n^N - \eta_n](f_n) \geq \epsilon \quad \text{or} \quad [\eta_n^N - \eta_n](f_n) \leq -\epsilon \\ &\iff [\eta_n^N - \eta_n](f_n) \geq \epsilon \quad \text{or} \quad [\eta_n^N - \eta_n](-f_n) \geq \epsilon \end{aligned}$$

so that

$$\mathbb{P}\left(|[\eta_n^N - \eta_n](f_n)| \geq \epsilon\right) \leq \mathbb{P}\left([\eta_n^N - \eta_n](f_n) \geq \epsilon\right) + \mathbb{P}\left([\eta_n^N - \eta_n](-f_n) \geq \epsilon\right)$$

The end of the proof of (5.11) is now easily completed. We further assume that the mixing condition $(\mathcal{M})_m$ is satisfied for some $m \geq 1$ and condition (5.4) holds

true for some $\delta > 0$ and some finite constants (r, \bar{r}) . In this situation, by (5.5) we have the uniform concentration estimate

$$\sup_{n \geq 0} \mathbb{P} (|[\eta_n^N - \eta_n](f_n)| \geq \epsilon) \leq 6 \exp \left(-\frac{N \epsilon^2 \delta^5}{32 m r \bar{r}^3} \right)$$

We conclude the proof of the theorem by choosing $\epsilon := \frac{1}{\sqrt{N}} \frac{4\bar{r}}{\delta^2} \sqrt{\frac{2m\bar{r}}{\delta} \log \frac{6}{\rho}}$. ■

Remark 5.7 *Returning to the end of the proof of theorem 5.6, we see the exponential concentration estimates can be improved, marginally, by choosing in (5.14) the parameter $\alpha = \alpha_n(\epsilon) \in [0, 1]$ such that $a_n(\epsilon) \alpha^2 = b_n (1 - \alpha)$, with $a_n(\epsilon) := \frac{\epsilon}{2\sigma_n^2}$ and $b_n = \frac{1}{3\sigma_{1,n}}$. Elementary manipulations yield*

$$\alpha_n(\epsilon) = \frac{b_n}{2a_n(\epsilon)} \left(\sqrt{1 + \frac{4a_n(\epsilon)}{b_n}} - 1 \right) = \frac{\sigma_n^2}{3\sigma_{1,n}} \frac{1}{\epsilon} \left(\sqrt{1 + \frac{6\sigma_{1,n}}{\sigma_n^2} \epsilon} - 1 \right) (\xrightarrow{\epsilon \downarrow 0} 1)$$

and therefore

$$\forall \epsilon \geq 0 \quad \mathbb{P} ([\eta_n^N - \eta_n](f) \geq \epsilon) \leq 3 \exp \left(-N \frac{\epsilon^2}{2\sigma_n^2} \alpha_n^2(\epsilon) \right)$$

5.3 Approximating the adaptive criteria

By construction, the particle occupation measures $\mathbb{P}_{\hat{\eta}_n, (t_n, s)}^N$ approximate the measures $\mathbb{P}_{\hat{\eta}_n, (t_n, s)}$ introduced in definition 3.1; that is, we have in some sense

$$\mathbb{P}_{\hat{\eta}_n, (t_n, s)}^N \simeq_{N \uparrow \infty} \mathbb{P}_{\hat{\eta}_n, (t_n, s)}$$

Conversely, we observe that $\mathbb{P}_{\hat{\eta}_n, (t_n, s)}^N$, and respectively $\mathbb{P}_{\hat{\eta}_n, (t_n, s)}$ are the marginals of the measures η_{n+1}^N , and respectively η_{n+1} with respect to the $(s - t_n) + 1$ first coordinates. In other words, the measures $\mathbb{P}_{\hat{\eta}_n, (t_n, s)}^N$, and respectively $\mathbb{P}_{\hat{\eta}_n, (t_n, s)}$ are the projections of the measures η_{n+1}^N , and respectively η_{n+1} on the state space

$$E_s := E_{t_n} \times (E'_{t_{n+1}} \times \dots \times E'_s)$$

For instance, the following proposition is more or less direct consequence of theorem 5.6.

Proposition 5.8 *For any $N \geq 1$, $n \geq 0$, $t_n \leq s \leq t_{n+1}$, and any $\epsilon > 0$, we have the concentration inequality:*

$$\mathbb{P} \left(\left| \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}^N \right) - \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) \right| \geq \epsilon \right) \leq (1 + \epsilon \sqrt{N/2}) \exp \left(-\frac{N\epsilon^2}{c(n)} \right)$$

for some finite constant $c(n) < \infty$ whose values only depend on the time parameter. In addition, when the measures $H_{t_n, s}^{(n)}$ have a finite support, we have the concentration inequality:

$$\mathbb{P} \left(\left| \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}^N \right) - \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) \right| \geq \epsilon \right) \leq c_1(n) \exp \left(-\frac{N\epsilon^2}{c_2(n)} \right)$$

with a pair of finite constants $c_1(n), c_2(n) < \infty$ whose values only depend on the time parameter.

Proof:

By theorem 7.4.4 in [4], for any $N \geq 1$, $p \geq 1$, $n \geq 0$ and any test function $f_n \in \text{Osc}_1(E_{t_n})$, we have

$$\sup_{N \geq 1} \sqrt{N} \mathbb{E} \left(\left| \eta_n^N(f_n) - \eta_n(f_n) \right|^p \right)^{\frac{1}{p}} \leq b(p) c(n)$$

with some finite constant $c(n) < \infty$ and with the collection of constants $b(p)$ defined in (5.8). These estimates clearly imply that for any $t_n \leq s \leq t_{n+1}$, and any test function $h_n \in \text{Osc}_1(E_s)$

$$\sup_{N \geq 1} \sqrt{N} \mathbb{E} \left(\left| \mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N(h_n) - \mathbb{P}_{\hat{\eta}_n, (t_n, s)}(h_n) \right|^p \right)^{\frac{1}{p}} \leq b(p) c(n)$$

Under our Lipschitz regularity condition (3.12) on the criteria type functionals $\mathcal{H}_{t_n, s}^{(n)}$, using the generalized integral Minkowski inequality we conclude that

$$\sup_{N \geq 1} \sqrt{N} \mathbb{E} \left(\left| \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N \right) - \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) \right|^p \right)^{\frac{1}{p}} \leq b(p) c(n) \delta(H_{t_n, s}^{(n)})$$

The proof of the exponential estimate follows exactly the same lines of arguments as the ones used in the proof of corollary 7.4.3 in [4]; thus, it is omitted. The last assertion is a direct consequence of theorem 5.6. This ends the proof of the proposition \blacksquare

5.4 An online adaptive particle model

The above proposition shows that the functional criteria $\mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right)$ can be approximated by $\mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N \right)$, up to an exponentially small probability. Therefore, it is tempting to define a particle model with an online particle computation of the adaptive criteria and hence the corresponding time steps. The natural idea is to replace the deterministic sequence $(t_n)_{0 \leq n \leq m+1}$ the measures $\mathbb{P}_{\hat{\eta}_n, (t_n, s)}$ by their current N -particle approximation measures.

Definition 5.9 *We use the letters \mathcal{Y} and Y to define the corresponding particle model. In other words, the particle systems $\mathcal{Y}^{(N)} = (\mathcal{Y}^{(N, i)})$, $\hat{\mathcal{Y}}^{(N)} = (\hat{\mathcal{Y}}^{(N, i)})$, and $Y_{s, t}^{(N, i)}$ and $\hat{Y}_{s, t}^{(N, i)}$ are defined as $\mathcal{X}^{(N)} = (\mathcal{X}^{(N, i)})$, $\hat{\mathcal{X}}^{(N)} = (\hat{\mathcal{X}}^{(N, i)})$, and $X_{s, t}^{(N, i)}$ and $\hat{X}_{s, t}^{(N, i)}$ by replacing in the inductive construction of the deterministic sequence $(t_n)_{n \geq 0}$ the measures $\mathbb{P}_{\hat{\eta}_n, (t_n, s)}$ by their current N -particle approximation measures*

$$\mathbb{P}_{\hat{\eta}_n^N, (t_n^N, s)}^N := \frac{1}{N} \sum_{i=1}^N \delta \left(\hat{\mathcal{Y}}_n^{(N, i)}, \left(Y_{t_n^N+1}^{(N, i)}, Y_{t_n^N+2}^{(N, i)}, \dots, Y_s^{(N, i)} \right) \right)$$

In the above displayed formula, with a slight abuse of notation $\hat{\eta}_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{\hat{\mathcal{Y}}_n^{(N, i)}}$ stands for the updated occupation measure of the particle model $\hat{\mathcal{Y}}_n^{(N)}$. We also suppose that both models are constructed in such a way that they coincide on every time interval $0 \leq n \leq m$, as soon as the random times $t_n^N = t_n$, for every $0 \leq n \leq m$.

We emphasize that the measures $\overline{\mathbb{P}}_{\hat{\eta}_n^N, (t_n^N, s)}^N$ generally differ from the reference empirical measures $\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N$ introduced in (5.2). Indeed, the reference measures $\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N$ are built using the deterministic times t_n based on the functional criteria $\mathcal{H}_{t_{n-1}, s}^{(n-1)}(\mathbb{P}_{\hat{\eta}_{n-1}, (t_{n-1}, s)})$, whilst the empirical measures $\overline{\mathbb{P}}_{\hat{\eta}_n^N, (t_n^N, s)}^N$ are inductively constructed using random times t_n^N based on the current empirical criteria $\mathcal{H}_{t_{n-1}^N, s}^{(n-1)}(\overline{\mathbb{P}}_{\hat{\eta}_{n-1}^N, (t_{n-1}^N, s)})$.

To illustrate the resulting particle adaptive scheme, we give the particle approximation of the four functional criteria discussed in section 4.

1. **Adaptive control of the weights** The N -particle approximation of the criteria discussed in section 4.1 are given below

$$\mathcal{H}_{t_n^N, s}^{(n)}(\overline{\mathbb{P}}_{\hat{\eta}_n^N, (t_n^N, s)}^N) = \frac{1}{N} \text{Card} \left\{ 1 \leq i \leq N : \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N, i)}) \geq a_n \right\}$$

In this situation, the particle model explores the state-space with a series of local transitions (5.15), up to the first time t_{n+1}^N the proportion of particles with a weight of at least size s_n is below a prescribed a_n level

$$\frac{1}{N} \text{Card} \left\{ 1 \leq i \leq N : \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N, i)}) \geq s_n \right\} \leq a_n$$

2. **Normalizing constants and survival probabilities:** The N -particle approximation of the criteria discussed in section 4.2 are given below

$$\mathcal{H}_{t_n^N, s}^{(n)}(\overline{\mathbb{P}}_{\hat{\eta}_n^N, (t_n^N, s)}^N) = \frac{1}{N} \sum_{i=1}^N \prod_{t_n^N < k \leq s} G'_k(Y_k^{(N, i)})$$

In this situation, the particle model explore the state space with a series of local transitions

$$\widehat{Y}_{t_n^N, t_n^N}^{(N, i)} \xrightarrow{M'_{t_n^N+1}} Y_{t_n^N+1}^{(N, i)} \xrightarrow{M'_{t_n^N+2}} Y_{t_n^N+2}^{(N, i)} \xrightarrow{M'_{t_n^N+3}} \dots \xrightarrow{M'_{t_n^N+1}} Y_{t_{n+1}^N}^{(N, i)} \quad (5.15)$$

up to the first time t_{n+1}^N the mean particle potential products goes below a prescribed a_n level

$$\frac{1}{N} \sum_{i=1}^N \prod_{t_n^N < k \leq t_{n+1}^N} G'_k(Y_k^{(N, i)}) \leq a_n$$

3. **Weights variances criteria w.r.t. free motions:** The N -particle approximation of the criteria discussed in section 4.3 are given below

$$\mathcal{H}_{t_n^N, s}^{(n)}(\overline{\mathbb{P}}_{\hat{\eta}_n^N, (t_n^N, s)}^N) = N \sum_{i=1}^N \left(\frac{\prod_{t_n < k \leq s} G'_k(Y_k^{(N, i)})}{\sum_{j=1}^N \prod_{t_n < k \leq s} G'_k(Y_k^{(N, j)})} \right)^2 - 1$$

In this situation, the particle model explores the state-space with a series of local transitions (5.15), up to the first time t_{n+1}^N the empirical particle variances is larger than some prescribed a_n -threshold

$$\sum_{i=1}^N \left(\frac{\prod_{t_n < k \leq s} G'_k(Y_k^{(N,i)})}{\sum_{j=1}^N \prod_{t_n < k \leq s} G'_k(Y_k^{(N,j)})} \right)^2 \geq (1 + a_n)/N$$

4. **Boltzmann's entropy criteria:** The N -particle approximation of the criteria discussed in section 4.4 are given below

$$\mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N \right) = -\frac{1}{N} \sum_{i=1}^N \sum_{t_n^N < k \leq s} \log G'_k(Y_k^{(N,i)})$$

In this situation, the particle model explores the state-space with a series of local transitions (5.15), up to the first time t_{n+1}^N the relative entropy of the empirical particle measures with respect to its weighted versions are larger than some a_n -threshold

$$-\frac{1}{N} \sum_{i=1}^N \sum_{t_n^N < k \leq s} \log G'_k(Y_k^{(N,i)}) \geq a_n$$

6 Asymptotic analysis

6.1 A key approximation lemma

To go one step further in our discussion, it is convenient to introduce the following collection of events.

Definition 6.1 For any $\delta \in]0, 1[$, $m \geq 0$, $a_n \in \mathbb{R}$, and $N \geq 1$, we denote by $\Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})$, the collection of events defined by:

$$\Omega_m^N(\delta, (a_n)_{0 \leq n \leq m}) := \{ \forall 0 \leq n \leq m, \forall t_n \leq s \leq t_{n+1} \}$$

$$\left| \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N \right) - \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) \right| \leq \delta \left| \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) - a_n \right|$$

Lemma 6.2 On the event $\Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})$, for any $n \leq m$ and any $t_n \leq s \leq t_{n+1}$, we have

$$\mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) > a_n \implies \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N \right) > a_n$$

$$\mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) < a_n \implies \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N \right) < a_n$$

Proof:

We observe that

$$\left| \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n^N, (t_n, s)}^N \right) - \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) \right| \leq \delta \left| \mathcal{H}_{t_n, s}^{(n)} \left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)} \right) - a_n \right|$$

is equivalent to that of saying that

$$\mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) - \delta \left| \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) - a_n \right| \leq \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n^N,(t_n,s)}^N)$$

and

$$\mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n^N,(t_n,s)}^N) \leq \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) + \delta \left| \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) - a_n \right|$$

Thus, if we have $\mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) < a_n$ we also have that

$$\begin{aligned} \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n^N,(t_n,s)}^N) &\leq \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) + \delta \left(a_n - \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) \right) \\ &\leq \delta a_n + (1 - \delta) \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) < a_n \end{aligned}$$

In much the same way, if we have $\mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) > a_n$ we also have that

$$\begin{aligned} \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n^N,(t_n,s)}^N) &\geq \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) - \delta \left(\mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) - a_n \right) \\ &\geq \delta a_n + (1 - \delta) \mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) > a_n \end{aligned}$$

This ends the proof of the lemma. \blacksquare

Proposition 6.3 *We assume that the thresholds parameters a_n are chosen so that we have $\mathcal{H}_{t_n,s}^{(n)}(\mathbb{P}_{\hat{\eta}_n,(t_n,s)}^N) \neq a_n$, for any $n \geq 0$. In this situation, for any $\delta \in]0, 1[$, $m \geq 0$ and $N \geq 1$, we have*

$$\cap_{0 \leq n \leq m} \{ t_n^N = t_n \} \supset \Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})$$

Proof:

We prove the result by induction on the parameter $m \geq 0$. Under our assumptions, for $m = 0$ we have $t_0^N = t_0 = 0$. Thus, by our coupling construction the pair of particle models coincide up to the time $(t_1^N \wedge t_1)$. Therefore, we have

$$\forall s < (t_1^N \wedge t_1) \quad \mathbb{P}_{\hat{\eta}_0^N,(t_0,s)}^N = \bar{\mathbb{P}}_{\hat{\eta}_0^N,(t_0^N,s)}^N$$

By lemma 6.2, on the event $\Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})$ we have $t_1^N = t_1$. This proves the inclusion for $m = 0$ and $m = 1$. Suppose the result is true at rank m . Thus, on the event $\Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})$ we have $t_n^N = t_n$, for any $0 \leq n \leq m$. By our coupling construction, the pair of particle models coincides up to $(t_{m+1}^N \wedge t_{m+1})$; that is, we have that

$$t_m^N = t_m \quad \text{and} \quad \forall s < (t_{m+1}^N \wedge t_{m+1}) \quad \mathbb{P}_{\hat{\eta}_m^N,(t_m,s)}^N = \bar{\mathbb{P}}_{\hat{\eta}_m^N,(t_m^N,s)}^N$$

Once again, by lemma 6.2, on the event $\Omega_{m+1}^N(\delta, (a_n)_{0 \leq n \leq m+1})$ we also have that $t_{m+1}^N = t_{m+1}$. This ends the proof of the proposition. \blacksquare

6.2 Randomized adaptive criteria

The situation where the threshold parameters coincide with the adaptive criteria values $\mathcal{H}_{t_n, s}^{(n)}(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}) = a_n$ cannot be handled using our analysis. Indeed, in this rather degenerate situation it is impossible (at least with our approach) to compare the random times t_n^N with the reference deterministic ones t_n . The analysis of this situation is clearly more involved since it requires us to control both the empirical approximating criteria and the mean field approximation model.

To avoid this technical problem, one natural strategy is to introduce randomized criteria thresholds. We further assume that the parameters $(a_n)_{n \geq 0}$ are sampled realizations of a collection of absolutely continuous random variables $(A_n)_{n \geq 0}$ (in practice, this element can be incorporated into the simulation, with no real change in the performance). The main simplification of these randomized criteria comes from the fact that the parameters

$$\epsilon_m := \inf_{0 \leq n \leq m} \inf_{t_n \leq s \leq t_{n+1}} \left| \mathcal{H}_{t_n, s}^{(n)}(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}) - a_n \right|$$

are strictly positive for almost every realizations $A_n = a_n$ of the threshold parameters.

Theorem 6.4 *For almost every realization of the random threshold parameters, and for any $\delta \in]0, 1[$, we have the following exponential estimates :*

$$\mathbb{P}(\exists 0 \leq n \leq m \quad t_n^N \neq t_n \mid (A_n)_{0 \leq n \leq m}) \leq c_1(m) (1 + \delta \epsilon_m \sqrt{N/2}) \exp(-N\delta^2 \epsilon_m^2 / c_2(m))$$

for some constants $c_1(m), c_2(m) < \infty$ whose values only depend on the time parameter m . In addition, when the measures $H_{t_n, s}^{(n)}$ have a finite support, for any $\delta \in]0, 1/(2\epsilon_m)[$ we have

$$\mathbb{P}(\exists 0 \leq n \leq m \quad t_n^N \neq t_n \mid (A_n)_{0 \leq n \leq m}) \leq c_1(m) \exp(-N\delta^2 \epsilon_m^2 / c_2(m))$$

for a possibly different pair of finite constants $c_1(m), c_2(m) < \infty$.

Proof:

Using proposition 5.8, we obtain the rather crude estimate

$$\begin{aligned} & \mathbb{P}(\Omega - \Omega_m^N(\delta, (A_n)_{0 \leq n \leq m}) \mid (A_n)_{0 \leq n \leq m} = (a_n)_{0 \leq n \leq m}) \\ & \leq \sum_{n=0}^m \sum_{s=t_n}^{t_{n+1}} \mathbb{P}\left(\left|\mathcal{H}_{t_n, s}^{(n)}\left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}^N\right) - \mathcal{H}_{t_n, s}^{(n)}\left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}\right)\right| \geq \delta \left|\mathcal{H}_{t_n, s}^{(n)}\left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}\right) - a_n\right|\right) \\ & \leq \sum_{n=0}^m \sum_{s=t_n}^{t_{n+1}} \mathbb{P}\left(\left|\mathcal{H}_{t_n, s}^{(n)}\left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}^N\right) - \mathcal{H}_{t_n, s}^{(n)}\left(\mathbb{P}_{\hat{\eta}_n, (t_n, s)}\right)\right| \geq \delta \epsilon_m\right) \\ & \leq c_1(m) (1 + \delta \epsilon_m \sqrt{N/2}) \exp(-N\delta^2 \epsilon_m^2 / c_2(m)) \end{aligned}$$

for a pair of finite constants $c_1(m), c_2(m) < \infty$ whose values only depend on the time parameter. The final line is a direct consequence of proposition 5.8 and application of proposition 6.3 completes the proof. \blacksquare

One concludes that for almost every realization $(A_n)_{0 \leq n \leq m} = (a_n)_{0 \leq n \leq m}$ the pair particle models $(\mathcal{X}_n^{(N)}, \widehat{\mathcal{X}}_n)_{0 \leq n \leq m}$ and $(\mathcal{Y}_n^{(N)}, \widehat{\mathcal{Y}}_n)_{0 \leq n \leq m}$ only differ on events $\Omega - \Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})$ with exponentially small probabilities:

$$\begin{aligned} & \mathbb{P} \left(\exists 0 \leq n \leq m \quad (\mathcal{Y}_n^{(N)}, \widehat{\mathcal{Y}}_n) \neq (\mathcal{X}_n^{(N)}, \widehat{\mathcal{X}}_n) \mid (A_n)_{0 \leq n \leq m} = (a_n)_{0 \leq n \leq m} \right) \\ & \leq c_1(m) (1 + \delta \epsilon_m \sqrt{N/2}) \exp(-N\delta^2 \epsilon_m^2 / c_2(m)) \end{aligned}$$

7 A functional central limit theorem

7.1 A direct approach

In this section, we discuss some direct consequences of the exponential coupling estimates presented in section 6.2. For almost every realization $(A_n)_{0 \leq n \leq m} = (a_n)_{0 \leq n \leq m}$ and for any test function $f_n \in \text{Osc}_1(E_{t_n})$ we have the decomposition (writing $\overline{\eta}_n^N$ for the online adaptive approximation introduced in definition 5.9)

$$\sqrt{N} [\overline{\eta}_n^N - \eta_n] = \sqrt{N} [\eta_n^N - \eta_n] + \sqrt{N} [\overline{\eta}_n^N - \eta_n^N] \mathbf{1}_{\Omega - \Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})}$$

with

$$\mathbb{E} \left(\sqrt{N} [\overline{\eta}_n^N - \eta_n^N] (f_n) \mathbf{1}_{\Omega - \Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})} \right) \leq \sqrt{N} \mathbb{P}(\Omega - \Omega_m^N(\delta, (a_n)_{0 \leq n \leq m})) \rightarrow_{N \uparrow \infty} 0$$

One concludes that, for almost every realization $(A_n)_{0 \leq n \leq m} = (a_n)_{0 \leq n \leq m}$, the random fields

$$\overline{W}_n^N := \sqrt{N} [\overline{\eta}_n^N - \eta_n] = \quad \text{and} \quad W_n^N := \sqrt{N} [\eta_n^N - \eta_n]$$

converges in law, as $N \uparrow \infty$, to the same Gaussian and centered random field W_n .

7.2 Functional central limit theorems

To demonstrate the impact of this functional fluctuation result and to connect this work with related studies on adaptive sequential Monte Carlo models, we provide a brief discussion on the proof of this multivariate central limit theorem.

Firstly, we recall a rather well known functional fluctuation theorem of the local errors associated with the mean field particle approximation sampling transitions introduced in (5.6). This result was initially presented in [6] in the case of elementary genetic type models. It has been extended to general mean field particle models in section 9.3 in [4].

Theorem 7.1 *For any fixed time horizon $n \geq 0$, the sequence $(V_p^N)_{0 \leq p \leq n}$ converges in law, as N tends to infinity, to a sequence of n independent, Gaussian and centered random fields $(V_p)_{0 \leq p \leq n}$; with, for any $f_p, g_p \in \mathcal{B}(E_p)$, and $1 \leq p \leq n$,*

$$\mathbb{E}(V_p(f_p)V_p(g_p)) = \eta_{p-1}K_{p,\eta_{p-1}}([f_p - K_{p,\eta_{p-1}}(f_p)][g_p - K_{p,\eta_{p-1}}(g_p)]) \quad (7.1)$$

On the other hand, using the same lines of arguments as in the proof of lemma 5.4, we find the first order decomposition formula:

$$[\Phi_n(\mu) - \Phi_n(\eta)](f) = (\mu - \eta)\mathcal{D}_{n,\eta} + \mathcal{R}_n(\mu, \eta)$$

with the signed measure $\mathcal{R}_n(\mu, \eta)$ given by

$$\mathcal{R}_n(\mu, \eta)(f) := -\frac{1}{\mu(\mathcal{G}_{n,\eta})} [\mu - \eta]^{\otimes 2}(\mathcal{G}_{n,\eta} \otimes \mathcal{D}_{n,\eta}(f)) \quad \text{with} \quad \mathcal{G}_{n,\eta} := \mathcal{G}_{n-1}/\eta(\mathcal{G}_{n-1})$$

and the bounded integral operator

$$\mathcal{D}_{n,\eta}(f)(x) := \mathcal{G}_{n,\eta}(x) \times \mathcal{M}_n(f - \Phi_n(\eta)(f))(x)$$

Definition 7.2 We denote by $D_{p,n}$ the semigroup associated with the integral operators $D_n := \mathcal{D}_{n,\eta_{n-1}}$; that is, we have that

$$D_{p,n} := D_{p+1} \dots D_{n-1} D_n$$

For $p = n$, we use the convention $D_{n,n} = Id$, the identity operator.

We mention that the semigroup $D_{p,n}$ introduced above can be explicitly described in terms of the semigroup $\mathcal{Q}_{p,n}$, with the following formula

$$D_{p,n}(f) = \frac{\mathcal{Q}_{p,n}}{\eta_p(\mathcal{Q}_{p,n}(1))} (f - \eta_n(f))$$

The next technical lemma provides a key first order decomposition of the random fields W_n^N in terms of the local fluctuation errors. Its proof follows an elementary induction on the time parameter n , and is in the appendix. Note that R_p is a remainder term that can be understood in the proof.

Lemma 7.3 For any $N \geq 1$ and any $0 \leq p \leq n$, we have

$$W_n^N = \sum_{p=0}^n V_p^N D_{p,n} + \mathcal{R}_n^N \quad \text{with} \quad \mathcal{R}_n^N := \sqrt{N} \sum_{p=0}^{n-1} R_{p+1}(\eta_p^N, \eta_p) D_{p+1,n} \quad (7.2)$$

Using the \mathbb{L}_m -mean error estimates presented in section 5.2.2, we easily prove that the sequence of remainder random fields \mathcal{R}_n^N introduced in (7.2) converge in law, in the sense of finite distributions, to the null random field as N tends to ∞ . Therefore the fluctuations of W_n^N follows from theorem 7.1.

Corollary 7.4 For any fixed time horizon $n \geq 0$, the sequence of random fields $(W_n^N)_{n \geq 0}$ converges in law, as N tends to infinity, to the sequence of Gaussian and centered random fields $(W_n)_{n \geq 0}$ given below

$$\forall n \geq 0 \quad W_n = \sum_{p=0}^n V_p D_{p,n}$$

7.3 On the fluctuations of weighted occupation measures

We end this article with some comments on the fluctuations of weighted occupation measures on path spaces. Returning to the online adaptive particle model, given a pair of time steps $(t_n^N, t_{n+1}^N) = (t_n, t_{n+1})$ the N -particle measures

$$\bar{\eta}_{n+1}^N = \frac{1}{N} \sum_{i=1}^N \delta \left(\hat{\mathcal{Y}}_n^{(N,i)}, \left(Y_{t_n^N+1}^{(N,i)}, Y_{t_n^N+2}^{(N,i)}, \dots, Y_{t_{n+1}^N}^{(N,i)} \right) \right)$$

can be used to approximate the flow of updated Feynman-Kac path distributions $(\hat{\eta}_{n+1,s})_{t_n \leq s \leq t_{n+1}}$ given for any bounded test function $f_{n+1} \in \mathcal{B}_b(S_{n+1})$ by

$$s \in [t_n, t_{n+1}] \mapsto \hat{\eta}_{n+1,s}(f_{n+1}) \propto \mathbb{E} \left[f_{n+1}(X'_0, \dots, X'_{t_{n+1}}) \prod_{0 < k \leq s} G'_k(X'_k) \right]$$

Indeed, if we choose the weighted functions

$$T_{n+1}^{(1)}(f_{n+1})(x'_0, \dots, x'_{t_{n+1}}) := f_{n+1}(x'_0, \dots, x'_{t_{n+1}}) \prod_{t_n < k \leq s} G'_k(x'_k)$$

in some sense we have

$$\hat{\eta}_{n+1,s}^N(f_{n+1}) := \frac{\bar{\eta}_{n+1}^N \left(T_{n+1}^{(1)}(f_{n+1}) \right)}{\bar{\eta}_{n+1}^N \left(T_{n+1}^{(1)}(1) \right)} \simeq_{N \uparrow \infty} \hat{\eta}_{n+1,s}(f_{n+1}) := \frac{\eta_{n+1} \left(T_{n+1}^{(1)}(f_{n+1}) \right)}{\eta_{n+1} \left(T_{n+1}^{(1)}(1) \right)}$$

with the flow of Feynman-Kac measures η_{n+1} on path spaces introduced in section 3.4.

Since the adaptive interaction time is taken such that $t_{n+1}^N = t_{n+1}$, we also have

$$\frac{1}{N} \sum_{i=1}^N \delta_{\hat{\mathcal{Y}}_{n+1}^{(N,i)}} \simeq_{N \uparrow \infty} \hat{\eta}_{n+1,t_{n+1}} = \hat{\eta}_{n+1}$$

In other words, if we choose the marginal type functions

$$T_{n+1}^{(0)}(f_{n+1})(x'_0, \dots, x'_{t_{n+1}}) := f_{n+1}(x'_0, \dots, x'_{t_{n+1}})$$

we have

$$\eta_{n+2} \left(T_{n+1}^{(0)}(f_{n+1}) \right) = \hat{\eta}_{n+1}(f_{n+1}) \propto \mathbb{E} \left[f_{n+1}(X'_0, \dots, X'_{t_{n+1}}) \prod_{0 < k \leq t_{n+1}} G'_k(X'_k) \right]$$

and therefore

$$\bar{\eta}_{n+2}^N \left(T_{n+1}^{(0)}(f_{n+1}) \right) = \frac{1}{N} \sum_{i=1}^N f_{n+1} \left(\hat{\mathcal{Y}}_{n+1}^{(N,i)} \right) \simeq_{N \uparrow \infty} \eta_{n+2} \left(T_{n+1}^{(0)}(f_{n+1}) \right)$$

From the previous discussion, using the continuous mapping theorem or alternatively the standard δ -method, for almost every realization $(A_n)_{0 \leq n \leq m} =$

$(a_n)_{0 \leq n \leq m}$ a central limit theorem is easily derived for the collection of random fields given below

$$\widehat{W}_{n+1}^{N,(0)}(f_{n+1}) := \sqrt{N} \left[\bar{\eta}_{n+2}^N \left(T_{n+1}^{(0)}(f_{n+1}) \right) - \eta_{n+2} \left(T_{n+1}^{(0)}(f_{n+1}) \right) \right]$$

and

$$\widehat{W}_{n+1,s}^{N,(1)}(f_{n+1}) := \sqrt{N} \left[\widehat{\eta}_{n+1,s}^N(f_{n+1}) - \widehat{\eta}_{n+1,s}(f_{n+1}) \right]$$

as well as for the mixture of random field sequences

$$\widehat{W}_{n+1,s}^N = 1_{t_n^N \leq s < t_{n+1}^N} \widehat{W}_{n+1,s}^{N,(1)} + 1_{s=t_{n+1}^N} \widehat{W}_{n+1}^{N,(0)} \quad (7.3)$$

The fluctuation analysis of these random fields strongly relies on the functional central limit theorem stated in corollary 7.4. In particular, the fluctuations of the random fields (7.3) depend on the fluctuations of a pair of random fields. From elementary stochastic calculus, the corresponding fluctuations cannot be deduced directly from the separate fluctuations of each components.

In this connection, we mention that these functional central limit theorems extend theorem 6 in [8] in the context of multinomial resampling schemes. In our context, these simple genetic type models correspond to the McKean particle interpretation model associated with the elementary selection transitions $\mathcal{S}_{n,\eta}(x, dy) = \Psi_n(\eta)(dy)$ (cf. section 3.1, in the present article).

In [8], the authors use a standard one dimensional central limit theorem on every test function to deduce the fluctuations of the sum (7.3) of the weighted and the resampled occupation measures. As we mentioned in the introduction, these arguments are combined with an inductive, and misleading, proof on the time parameter s . We also remark that [8] do not consider the degenerate situation where the limiting adaptive criteria coincide at a given time period with the threshold parameter.

Appendix

Proof of lemma 5.2

We prove the lemma by induction on the parameter n . The result is clearly satisfied for $n = 0$. Suppose the estimate (5.7) has been check up to rank $(n - 1)$. To prove the result at rank n , we use the binomial decomposition

$$(M_{n-1} + \Delta_n)^{2m} = \sum_{p=0}^{2m} \binom{2m}{p} M_{n-1}^{2m-p} (\Delta_n)^p$$

Using the symmetry condition, all the odd moments of Δ_n are null. Consequently, we find that

$$\mathbb{E} \left((M_{n-1} + \Delta_n)^{2m} \right) = \sum_{p=0}^m \binom{2m}{2p} \mathbb{E} \left(M_{n-1}^{2(m-p)} \right) \mathbb{E} \left(\Delta_n^{2p} \right)$$

Using the induction hypothesis, we prove that the above expression is upper bounded by the quantity

$$\sum_{p=0}^m \binom{2m}{2p} 2^{-(m-p)} (2(m-p))_{(m-p)} \mathbb{E} \left([M]_{n-1}^{m-p} \right) \mathbb{E} \left(\Delta_n^{2p} \right)$$

To take the final step, we use the fact that

$$\binom{2m}{2p} 2^{-(m-p)} (2(m-p))_{(m-p)} = \frac{2^{-m} (2m)_m}{2^{-p} (2p)_p} \binom{m}{p} \quad \text{and} \quad (2p)_p \geq 2^p$$

to conclude that

$$\begin{aligned} \mathbb{E} \left((M_{n-1} + \Delta_n)^{2m} \right) &\leq 2^{-m} (2m)_m \sum_{p=0}^m \binom{m}{p} \mathbb{E} \left([M]_{n-1}^{m-p} \right) \mathbb{E} \left(\Delta_n^{2p} \right) \\ &= 2^{-m} (2m)_m \mathbb{E} \left([M]_n^m \right) \end{aligned}$$

For odd integers we use the Cauchy-Schwarz inequality twice to deduce that

$$\begin{aligned} \mathbb{E}(|M_n|^{2m+1})^2 &\leq \mathbb{E}(M_n^{2m}) \mathbb{E}(M_n^{2(m+1)}) \\ &\leq 2^{-(2m+1)} (2m)_m (2(m+1))_{(m+1)} \mathbb{E} \left([M]_n^{m+1} \right)^{\frac{2m+1}{m+1}} \end{aligned}$$

Since we also have that

$$(2(m+1))_{(m+1)} = \frac{(2(m+1))!}{(m+1)!} = 2 \frac{(2m+1)!}{m!} = 2 (2m+1)_{(m+1)}$$

and

$$(2m)_m = \frac{1}{2m+1} \frac{(2m+1)!}{m!} = \frac{1}{2m+1} (2m+1)_{(m+1)}$$

we conclude that

$$\mathbb{E}(|M_n|^{2m+1}) \leq 2^{-(m+1/2)} \frac{(2m+1)_{(m+1)}}{\sqrt{m+1/2}} \mathbb{E} \left([M]_n^{m+1} \right)^{1 - \frac{1}{2(m+1)}}$$

This ends the proof of the lemma. ■

Proof of lemma 5.4

Using (3.6), for any $f \in \mathcal{B}_b(S_{n+1})$ we find that

$$[\Phi_{p,n}(\mu) - \Phi_{p,n}(\eta)](f) = \frac{1}{\mu(\mathcal{G}_{p,n,\eta})} (\mu - \eta) \mathcal{D}_{p,n,\eta}(f)$$

with the bounded integral operator

$$\mathcal{D}_{p,n,\eta}(f)(x) := \mathcal{G}_{p,n,\eta}(x) \times \mathcal{P}_{p,n}(f - \Phi_{p,n}(\eta)(f))(x)$$

In the above display $\mathcal{G}_{p,n,\eta}$ and $\mathcal{P}_{p,n}$ stand for the potential function and the Markov operator given by

$$\mathcal{G}_{p,n,\eta} := \mathcal{Q}_{p,n}(1)/\eta(\mathcal{Q}_{p,n}(1)) \quad \text{and} \quad \mathcal{P}_{p,n}(f) = \mathcal{Q}_{p,n}(f)/\mathcal{Q}_{p,n}(1)$$

Using the fact that $\eta(\mathcal{G}_{p,n,\eta}) = 1$, it is now immediate to check that

$$[\Phi_{p,n}(\mu) - \Phi_{p,n}(\eta)] = (\mu - \eta) \mathcal{D}_{p,n,\eta} + \mathcal{R}_{p,n}(\mu, \eta)$$

with the measure $\mathcal{R}_{p,n}(\mu, \eta)$ given by

$$\mathcal{R}_{p,n}(\mu, \eta)(f) := -\frac{1}{\mu(\mathcal{G}_{p,n,\eta})} [\mu - \eta]^{\otimes 2}(\mathcal{G}_{p,n,\eta} \otimes \mathcal{D}_{p,n,\eta}(f))$$

Using the fact that

$$\mathcal{D}_{p,n,\eta}(f)(x) = \mathcal{G}_{p,n,\eta}(x) \int [\mathcal{P}_{p,n}(f)(x) - \mathcal{P}_{p,n}(f)(y)] \mathcal{G}_{p,n,\eta}(y) \eta(dy)$$

we find that

$$\forall f \in \text{Osc}_1(E_n) \quad \|\mathcal{D}_{p,n,\eta}(f)\| \leq q_{p,n} \beta(\mathcal{P}_{p,n})$$

Finally, for any $f \in \text{Osc}_1(S_n)$ we observe that

$$|\mathcal{R}_{p,n}(\mu, \eta)(f)| \leq (4q_{p,n}^3 \beta(\mathcal{P}_{p,n})) \left| [\mu - \eta]^{\otimes 2}(\bar{\mathcal{G}}_{p,n,\eta} \otimes \bar{\mathcal{D}}_{p,n,\eta}(f)) \right|$$

with the pair of functions

$$\bar{\mathcal{G}}_{p,n,\eta} := \frac{\mathcal{G}_{p,n,\eta}}{2q_{p,n}} \quad \text{and} \quad \bar{\mathcal{D}}_{p,n,\eta}(f) := \frac{D_{p,n,\eta}(f)}{2q_{p,n}\beta(\mathcal{P}_{p,n})} \in \text{Osc}_1(S_p)$$

This ends the proof of the lemma. \blacksquare

7.4 Proof of lemma 7.3

We prove the lemma by induction on the time parameter n . For $n = 0$, we readily find that $W_n^N = V_0^N = \sqrt{N} [\eta_0^N - \Phi_0(\eta_{-1}^N)]$, with the convention $\Phi_0(\eta_{-1}^N) = \eta_0$. Assuming that the decomposition is satisfied at rank n we prove that

$$\begin{aligned} W_{n+1}^N &= V_{n+1}^N + \sqrt{N} [\Phi_{n+1}(\eta_n^N) - \Phi_{n+1}(\eta_n)] = V_{n+1}^N + W_n^N D_{n+1} + \sqrt{N} R_{n+1}(\eta_n^N, \eta_n) \\ &= V_{n+1}^N + \sum_{p=0}^n V_p^N D_{p,n+1} + \sqrt{N} \sum_{p=0}^{n-1} R_{p+1}(\eta_p^N, \eta_p) D_{p+1,n+1} + \sqrt{N} R_{n+1}(\eta_n^N, \eta_n) \end{aligned}$$

Using the convention $D_{n+1,n+1} = Id$ we find that

$$\begin{aligned} V_{n+1}^N + \sum_{p=0}^n V_p^N D_{p,n+1} &= \sum_{p=0}^{n+1} V_p^N D_{p,n+1} \\ \sum_{p=0}^{n-1} R_{p+1}(\eta_p^N, \eta_p) D_{p+1,n+1} + R_{n+1}(\eta_n^N, \eta_n) &= \sum_{p=0}^n R_{p+1}(\eta_p^N, \eta_p) D_{p+1,n+1} \end{aligned}$$

from which we conclude that (7.2) is satisfied at rank $(n + 1)$. This ends the proof of the lemma. \blacksquare

References

- [1] F. CEROU & A. GUYADER (2006). Adaptive multilevel splitting for rare event analysis, *Rapport de Recherche INRIA*, no. 5710 (2005), *Stochastic Analysis and its Applications (2007)*.

-
- [2] F. CEROU, P. DEL MORAL, A. GUYADER, F. LEGLAND, P. LEZAUD, & H. TOPART. *Some recent improvements to importance splitting*. Proceedings of 6th International Workshop on Rare Event Simulation, Bamberg, Germany (2006).
- [3] P. DEL MORAL (1998). Measure valued processes and interacting particle systems. Application to non linear filtering problems. *Annals of Applied Probability*, vol. 8, no. 2, 438–495.
- [4] P. DEL MORAL (2004). *Feynman-Kac formulae. Genealogical and interacting particle systems*, Springer-Verlag: New York.
- [5] P. DEL MORAL, & J. JACOD (2000). Interacting particle filtering with discrete-time observations: asymptotic behaviour in the Gaussian case, in *Stochastics in finite and infinite dimensions Trends in mathematics*, Birkhäuser, pp.101-123.
- [6] P. DEL MORAL, & L. MICLO (2000). Branching and Interacting Particle Systems Approximations of Feynman-Kac Formulae. *Séminaire de Probabilités XXXIV, Lecture Notes in Mathematics, Springer-Verlag Berlin*, Vol. 1729, 1-145.
- [7] P. DEL MORAL,, A. DOUCET, & A. JASRA (2006). Sequential Monte Carlo samplers. *J. R. Statist. Soc. B*, **68**, 411–436.
- [8] R. DOUC, & É. MOULINES (2008). Limit theorems for weighted samples with applications to sequential Monte Carlo. *Ann. Statist.*, (to appear).
- [9] A. DOUCET, N. DE FREITAS, & N. GORDON (2001). *Sequential Monte Carlo methods in practice*, Springer Verlag New York, Series: Statistics for Engineering and Information Science.
- [10] C. JARZYNSKI (1997). Equilibrium free energy differences from nonequilibrium measurements: a master equation approach, *Phys. Rev. E* 56(5), pp. 5018-5035.
- [11] A. JASRA, D.A. STEPHENS, A. DOUCET & T. TSGARIS, (2008). Inference for Lévy driven stochastic volatility models via adaptive sequential Monte Carlo, Technical Report, McGill University.
- [12] J. LIU & R. CHEN (1998). Sequential Monte Carlo for dynamic systems. *J. Amer. Statist. Assoc.*, 93, 1032-1044.
- [13] C.P. ROBERT & G. CASELLA (2004). *Monte Carlo statistical methods*, 2nd edition, Springer Verlag New York.
- [14] M. ROUSSET & G. STOLTZ (2006). Equilibrium sampling from non-equilibrium dynamics. *J. Stat. Phys.*, **123**, 1251-1272.



Centre de recherche INRIA Bordeaux – Sud Ouest
Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex (France)

Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

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