



A Functional Central Limit Theorem for a Class of Interacting Markov Chain Monte Carlo Models

Bernard Bercu, Pierre del Moral, Arnaud Doucet

► To cite this version:

Bernard Bercu, Pierre del Moral, Arnaud Doucet. A Functional Central Limit Theorem for a Class of Interacting Markov Chain Monte Carlo Models. [Research Report] 2008, pp.25. inria-00227534v1

HAL Id: inria-00227534

<https://inria.hal.science/inria-00227534v1>

Submitted on 30 Jan 2008 (v1), last revised 18 Jun 2009 (v5)

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

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

A Functional Central Limit Theorem for a Class of Interacting Markov Chain Monte Carlo Models

B. Bercu*, P. Del Moral†, A. Doucet‡

January 30, 2008

Abstract

We present a functional central limit theorem for a new class of interaction Markov chain Monte Carlo interpretations of discrete generation measure valued equations. We provide an original stochastic analysis based on semigroup techniques on distribution spaces and fluctuation theorems for self-interaction random fields. Besides the fluctuation analysis of these models, we also present a series of sharp \mathbb{L}_m -mean error bounds in terms of the semigroup associated with the first order expansion of the limiting measure valued process, yielding what seems to be the first results of this type for this class of interacting processes. We illustrate these results in the context of Feynman-Kac integration semigroups arising in physics, biology and stochastic engineering science.

Keywords : Multivariate and functional central limit theorems, random fields, martingale limit theorems, self-interacting Markov chains, Markov chain Monte Carlo models, and Feynman-Kac integrals.

Mathematics Subject Classification :

Primary: 47H20, 60G35, 60J85, 62G09, ; Secondary: 47D08, 47G10, 62L20.

1 Introduction

1.1 Self interacting Markov chain models

Let $(S^{(l)})_{l \geq 0}$ be a sequence of measurable spaces equipped with some σ -fields $(\mathcal{S}^{(l)})_{l \geq 0}$, and for every $l \geq 0$ we denote by $\mathcal{P}(S^{(l)})$ the set of all probability measures on $S^{(l)}$. In all the sequel, we shall denote by $\pi^{(l)}$ a sequence of probability measures on $S^{(l)}$ that satisfies a nonlinear equation of the following form

$$\Phi_l(\pi^{(l-1)}) = \pi^{(l)} \quad (1.1)$$

for some mappings $\Phi_l : \mathcal{P}(S^{(l-1)}) \rightarrow \mathcal{P}(S^{(l)})$. These probabilistic models and their mean-field interacting particle interpretations are one of the most active contact points between probability, theoretical chemistry, quantum physics, and engineering sciences, including rare event analysis, and advanced signal processing. These stochastic particle models are sometimes referred as sequential Monte Carlo methods, population Monte Carlo models or branching and interacting particle systems. It is clearly out of the scope of this article

*Institut de Mathématiques de Bordeaux, Université de Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France, bercu@math.u-bordeaux1.fr

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

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

to enter into the details of these stochastic models, for a rather thorough discussion we recommend the interested reader to consult the pair of books [2, 6].

One drawback of these particle models is that it is impossible to increase iteratively the precision of the algorithm. To solve this important question, we have recently introduced in a pair of articles [1, 3] a new class of interacting Monte Carlo Markov chain models (*abbreviate i-MCMC*). This new class of interacting stochastic algorithms can be described as a dynamic and adaptive simulation algorithm which takes advantage of the information carried by the past history to increase the quality of the next series of samples. In contrast with other numerical method, the main advantage of the i-MCMC methodology is that it provides a natural adaptation and reinforced learning strategy of the physical or engineering evolution equation at hand. In this connection, we emphasize that this type of reinforcement with the past is observed frequently in nature and society, where "beneficial" interactions with the past history tend to be repeated. Another advantage of these probabilistic techniques is that they do not use any regularity information on the coefficients of the models and they apply to large scale models. Furthermore, in contrast to more traditional mean-field type particle models and related sequential Monte Carlo techniques, these stochastic algorithms increase iteratively the precision and the performance of the numerical approximation. Last but not least, another advantage of the i-MCMC methodology comes from the fact that it can be combined easily with the more traditional mean field particle model discussed above.

Let us give a rough description of the i-MCMC methodology. Firstly, we run a first Monte Carlo Markov chain type sampler, say $X_n^{(0)}$ with a prescribed simple target distribution $\pi^{(0)}$. Then, we use judiciously the occupation measure of the chain $X^{(0)}$ to design a second MCMC algorithm say $X_n^{(1)}$ with a more sophisticated and complex target limiting measure $\pi^{(1)}$. These two mechanisms are combined online so that the pair process interacts with the occupation measure of the system at each time. More formally, the elementary transition $X_n^{(1)} \rightsquigarrow X_{n+1}^{(1)}$ of the chain $X^{(1)}$ at a current time n depends on the occupation measure of the chain $X_p^{(0)}$, from the origin $p = 0$, up to the current time $p = n$. This strategy can be extended to design a series of MCMC samplers $(X^{(l)})_{l \geq 0}$ with a prescribed sequence of distributions $(\pi^{(l)})_{l \geq 0}$ with an increasing level of complexity. These i-MCMC samplers are sometimes called self interacting Markov chain in reference with the fact that the Markov chain $X_n^{[m]} := (X_n^{(l)})_{0 \leq l \leq m}$ associated with a fixed series of m levels evolves with elementary transitions which depend on the occupation measure of the whole system from the origin up to the current time.

As shown in [3], from the pure mathematical point of view, these i-MCMC models can be casted to the class of self-interacting processes. The convergence analysis of such stochastic models is essentially based on the study of the interacting sequences of conditionally independent random variables, whose distributions depend in a nonlinear way on the occupation measure of the chain. The theoretical analysis of these sophisticated models is much more involved than the one of traditional genetic type mean field interpretation models developed in [2]. In a pair of recent articles [1, 3] we initiated the theoretical analysis of i-MCMC models and we provided a variety of convergence results including exponential estimates and an uniform convergence theorem with respect to the time parameter. However, we left open the important questions related to the fluctuations and the large deviation principles. The main purpose of this paper is to provide the fluctuation analysis of the occupation measures of an i-MCMC algorithm around its limiting values $\pi^{(l)}$, as the time parameter n tends to infinity. We also make no restriction on the state spaces and we illustrate these models in the context of abstract Feynman-Kac distribution flows in path spaces.

An outline of the development of the paper is as follows : The i-MCMC models are described in section 1.2. We also connect these models with the more traditional mean field particle algorithms and self interacting processes. For the convenience of the reader, we have collected in a short preliminary section, section 1.3, some of the main notation and conven-

tions used in the article. We also recall some more or less well known regularity properties of integral operators of current use in the further development of the article. Section 2 is devoted to the two main results of this work. We provide a multivariate functional central limit theorem together with some sharp non asymptotic \mathbb{L}_m -estimates in terms of the semigroup associated with a first order expansion of the mappings Φ_l . To motivate our approach, in section 3 we illustrate these results in the context of Feynman-Kac semigroups. The last four sections are essentially concerned with the proof of the pair of theorems presented in section 2. Our analysis is based on two main ingredients; namely, a multivariate functional central limit theorem for local interaction fields and a multilevel expansion of the occupation measure of the i-MCMC model at a given level l in terms of the occupation measures at the levels with lower indexes. These two results are presented respectively in section 4 and section 5. The final two sections, sections 6 and section 7 are devoted respectively to the proofs of the functional central limit theorem and the sharp non asymptotic \mathbb{L}_m -estimates theorem. In an appendix, we have collected the proof of a series of technical lemmas.

1.2 Description of the i-MCMC methodology

The i-MCMC methodology is better introduced in terms of the traditional Metropolis-Hastings model. We recall that the Metropolis-Hastings model is an elementary Markov chain with a probability transition of the following form

$$M(x, dy) = (1 \wedge G(x, y)) P(x, dy) + \left(1 - \int_S (1 \wedge G(x, z)) P(x, dz)\right) \delta_x(dy) \quad (1.2)$$

In the above displayed formula, $P(x, dy)$ is an exploration Markov transition on some measurable space S , and $G : S^2 \rightarrow \mathbb{R}_+$. Let π be a probability measure on S , and let $(\pi \times P)_1$ and $(\pi \times P)_2$ be the pair of measures on $(S \times S)$ defined by

$$(\pi \times P)_1(dx, y) := \pi(dx) P(x, dy) =: (\pi \times P)_2(dy, x)$$

We further assume that $(\pi \times P)_1 \ll (\pi \times P)_2$ and we set $G := d(\pi \times P)_1 / d(\pi \times P)_2$ the corresponding Radon-Nykodim ratio. In this case, π is an invariant measure of the time homogeneous Markov chain with transition M . One well known difficulty with this stochastic algorithm comes from the fact that the acceptance rate potential function G may be very small in some regions when the proposition transition P is not judiciously chosen.

Next, we suppose that $\pi = \Phi(\eta)$ is related to a auxiliary probability measure η on some possibly different measurable space S' with some mapping $\Phi : \mathcal{P}(S') \rightarrow \mathcal{P}(S)$. We also suppose that there exists a running Markov chain Monte Carlo algorithm on S' whose occupation measures, say η_n approximate the measure η , as the time parameter n increases. At each current time, say n , we would like to choose a Metropolis-Hastings transition M_n associated with a pair (G_n, P_n) with invariant measure $\Phi(\eta_n)$. The rationale behind this is that the resulting chain will behave asymptotically as a Markov chain with invariant distribution $\Phi(\eta) = \pi$, as soon as η_n is a good approximation of η .

The choice of the transition M_n is far from being unique. In [3], we present a variety of solutions and one of them seems more appropriate. It is based on the following natural observations : If we simply choose $P_n(x, dy) = \Phi(\eta_n)(dy)$, then the corresponding acceptance-rejection ratio G_n is equal to one. In this sense, the best Metropolis-Hastings algorithm is to sample a series of independent random variables $X := (X_n)_{n \geq 0}$ with distribution $(\Phi(\eta_{n-1}))_{n \geq 0}$. Iterating this interacting stochastic modeling strategy, we design a series of i-MCMC models $X^{(k)} := (X_n^{(k)})_{n \geq 0}$ on every level set $S^{(k)}$ whose occupation measures $\eta_n^{(k)}$ approximate the solution $\pi^{(k)}$ of the system (1.1), for every $k \geq 0$.

To describe more precisely the resulting model, we need a few notation. We also fix a

series of initial probability measures ν_k on $S^{(k)}$, with $k \geq 0$, and we denote by

$$\eta_n^{(k)} := \frac{1}{n+1} \sum_{0 \leq p \leq n} \delta_{X_p^{(k)}}$$

the flow of occupation measures of the chain $X^{(k)}$ at level k from the origin up to time n , where δ_x stands for the Dirac measure at a given point $x \in S^{(k)}$.

The i-MCMC model proposed in this article is defined inductively on the level parameter k . Firstly, to simplify the analysis and the presentation, we shall suppose that $\nu_0 = \pi^{(0)}$ and we let $X^{(0)} := (X_n^{(0)})_{n \geq 0}$ be a collection of independent random variables with common distribution $\nu_0 = \pi^{(0)}$. For every $k \geq 1$, and given a realization of the chain $X^{(k)}$, the k -th level chain $X^{(k+1)} := (X_n^{(k+1)})_{n \geq 0}$ is a collection of independent random variables with distributions given, for any $n \geq 0$, by the following formula

$$\mathbb{P}((X_0^{(k+1)}, \dots, X_n^{(k+1)}) \in d(x_0, \dots, x_n) \mid X^{(k)}) = \prod_{0 \leq p \leq n} \Phi_{k+1} \left(\eta_{p-1}^{(k)} \right) (dx_p) \quad (1.3)$$

In the above displayed formula, we have used the convention $\Phi_{k+1} \left(\eta_{-1}^{(k)} \right) = \nu_{k+1}$ for $p = 0$, and $d(x_0, \dots, x_n) = dx_0 \times \dots \times dx_n$ stands for an infinitesimal neighborhood of a generic path sequence $(x_0, \dots, x_n) \in (S^{(k+1)})^{(n+1)}$.

To better connect the i-MCMC methodology with the more traditional mean field particle models and self interacting processes, we end this section with a pair of remarks.

Suppose that we fix the time horizon n and we replace in formula (1.3) the flow of occupation measures $(\eta_p^{(k)})_{0 \leq p \leq n}$ by the occupation measure $\eta_n^{(k)}$ of the chain $X^{(k)}$. In this situation, we loose the recursive formulation of the algorithm but we end up with a traditional mean field interpretation model of the equation (1.1). Finally, we present a self interacting interpretation of the i-MCMC methodology. We fix a final time horizon m of the equation (1.1), and for every $0 \leq l \leq m$ we denote by $\eta^{(l)} \in \mathcal{P}(S^{(l)})$ the image measure of a measure $\eta \in \mathcal{P}(E_m)$ on the l -th section $S^{(l)}$ of the product space $E_m := \prod_{0 \leq l \leq m} S^{(l)}$. In this notation, it is readily checked that

$$X_n^{[m]} := (X_n^{(0)}, \dots, X_n^{(m)})$$

is an E_m -valued and self interacting Markov chain with elementary transitions defined by

$$\mathbb{P}(X_{n+1}^{[m]} \in dy \mid X_0^{[m]}, \dots, X_n^{[m]}) = \prod_{0 \leq l \leq m} \Phi_l \left(\eta_n^{(l-1)} \right) (dy^l) \quad \text{with} \quad \eta_n := \frac{1}{n+1} \sum_{p=0}^n \delta_{X_p^{[m]}} \quad (1.4)$$

In the above displayed formula we have use the convention $\Phi_0 \left(\eta_n^{(-1)} \right) = \pi^{(0)}$ for $l = 0$, and $dy := dy^0 \times \dots \times dy^m$ stands for an infinitesimal neighborhood of a generic point $y := (y^0, \dots, y^m) \in E_m$.

1.3 Notation and conventions

We denote respectively by $\mathcal{M}(E)$, $\mathcal{M}_0(E)$, $\mathcal{P}(E)$, and $\mathcal{B}(E)$, the set of all finite signed measures on some measurable space (E, \mathcal{E}) , the convex subset of measures with null mass, the set of all probability measures, and the Banach space of all bounded and measurable functions f on E . We equip $\mathcal{B}(E)$ with the uniform norm $\|f\| = \sup_{x \in E} |f(x)|$. We also denote by $\mathcal{B}_1(E) \subset \mathcal{B}(E)$ the unit ball of functions $f \in \mathcal{B}(E)$ with $\|f\| \leq 1$, and by $\text{Osc}_1(E)$, the convex set of \mathcal{E} -measurable functions f with oscillations less than one, which means that

$$\text{osc}(f) = \sup \{ |f(x) - f(y)| ; x, y \in E \} \leq 1$$

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 slight abuse the notation, and sometimes we denote by $\mu(A) = \mu(1_A)$ the measure of a measurable subset $A \in \mathcal{E}$.

We recall that a bounded integral operator M from a measurable space (E, \mathcal{E}) into an auxiliary measurable space (F, \mathcal{F}) into itself, is an operator $f \mapsto M(f)$ from $\mathcal{B}(F)$ into $\mathcal{B}(E)$ such that the functions

$$M(f)(x) = \int_F M(x, dy) f(y)$$

are \mathcal{E} -measurable and bounded, for any $f \in \mathcal{B}(F)$. 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 denote by $\|M\| := \sup_{f \in \mathcal{B}_1(F)} \|M(f)\|$ the norm of the operator $f \mapsto M(f)$ and we equip the Banach space $\mathcal{M}(E)$ with the corresponding total variation norm $\|\mu\| = \sup_{f \in \mathcal{B}_1(E)} |\mu(f)|$. We let $\beta(M)$ be the Dobrushin coefficient of a bounded integral operator M defined by the following formula

$$\beta(M) := \sup \{ \text{osc}(M(f)) ; f \in \text{Osc}_1(F) \}$$

When M has a constant mass $M(1)(x) = M(1)(y)$, for any $(x, y) \in E$, the operator $\mu \mapsto \mu M$ maps $\mathcal{M}_0(E)$ into $\mathcal{M}_0(F)$, and $\beta(M)$ coincides with the norm of this operator. We equip the sets of distribution flows $\mathcal{M}(E)^\mathbb{N}$ with the uniform total variation distance defined by

$$\forall \eta = (\eta_n)_{n \geq 0}, \mu = (\mu_n)_{n \geq 0} \in \mathcal{M}(E)^\mathbb{N} \quad \|\eta - \mu\| := \sup_{n \geq 0} \|\eta_n - \mu_n\|$$

We extend a given bounded integral operator $\mu \in \mathcal{M}(E) \mapsto \mu M \in \mathcal{M}(F)$ into an mapping

$$\eta = (\eta_n)_{n \geq 0} \in \mathcal{M}(E)^\mathbb{N} \mapsto \eta D = (\eta_n M)_{n \geq 0} \in \mathcal{M}(F)^\mathbb{N}$$

Sometimes, we slight abuse the notation and we denote by ν instead of $(\nu)_{n \geq 0}$ the constant distribution flows equal to a given measure $\nu \in \mathcal{P}(E)$.

For any \mathbb{R}^d -valued function $f = (f^i)_{1 \leq i \leq d} \in \mathcal{B}(F)^d$, any integral operator M from E into F , and any $\mu \in \mathcal{M}(F)$, we write $M(f)$ and $\mu(f)$ the \mathbb{R}^d -valued function and the point in \mathbb{R}^d given respectively by

$$M(f) := (M(f^1), \dots, M(f^d)) \quad \text{and} \quad \mu(f) := (\mu(f^1), \dots, \mu(f^d))$$

Unless otherwise is stated, we denote by $c(k)$, $k \in \mathbb{N}$, a constant whose values may vary from line to line, but they only depend on the parameter k . For any pair of integers $0 \leq m \leq n$, we denote by $(n)_m := n!/(n-m)!$ the number of one to one mappings from $\{1, \dots, m\}$ into $\{1, \dots, n\}$. Finally, we shall use the usual conventions $\sum_{\emptyset} = 0$ and $\prod_{\emptyset} = 1$.

2 Fluctuation theorems

This section is mainly concerned with the statement of the two main fluctuation theorems presented in this article. These results are based on a single first order weak regularity condition on the mappings Φ_l governing the measure valued equation (1.1). Namely, we further assume that for any $l \geq 1$, the mappings $\Phi_l : \mathcal{P}(S^{(l-1)}) \rightarrow \mathcal{P}(S^{(l)})$ satisfy the following first order decomposition

$$\Phi_l(\mu) - \Phi_l(\eta) = (\mu - \eta) \mathcal{D}_{l,\eta} + \Xi_l(\mu, \eta) \tag{2.1}$$

In the above display, $\mathcal{D}_{l,\eta}$ stands for some collection of bounded integral operators from $S^{(l-1)}$ into $S^{(l)}$, indexed by the set of probability measures $\eta \in \mathcal{P}(S^{(l-1)})$ and $\Xi_l(\mu, \eta)$

stands for some collection of remainder signed measures on $S^{(l)}$ and indexed by the set of probability measures $\mu, \eta \in \mathcal{P}(S^{(l-1)})$. We further require that

$$\sup_{\eta \in \mathcal{P}(S^{(l-1)})} \|\mathcal{D}_{l,\eta}\| < \infty \quad \text{and} \quad |\Xi_l(\mu, \eta)(f)| \leq \int |(\mu - \eta)^{\otimes 2}(g)| \Xi_l(f, dg) \quad (2.2)$$

for some integral operator Ξ_l from $\mathcal{B}(S^{(l)})$ into the set $\mathcal{T}_2(S^{(l-1)})$ of all tensor product functions $g = \sum_{i \in I} a_i (h_i^1 \otimes h_i^2)$, with $I \subset \mathbb{N}$, $(h_i^1, h_i^2)_{i \in I} \in (\mathcal{B}(S^{(l-1)})^2)^I$, and a sequence of numbers $(a_i)_{i \in I} \in \mathbb{R}^I$ such that

$$|g| := \sum_{i \in I} |a_i| \|h_i^1\| \|h_i^2\| < \infty \quad \text{and} \quad \chi_l := \sup_{f \in \mathcal{B}_1(S^{(l)})} \int |g| \Xi_l(f, dg) < \infty$$

This weak regularity condition is satisfied in a variety of models including Feynman-Kac integral semigroups discussed in the next section. We also mention that, under weaker assumptions on the mappings Φ_l , we already proved in [1, 3] that for every $l \geq 0$ and any function $f \in \mathcal{B}(S^{(l)})$, $\eta_n^{(l)}(f)$ converge almost surely to $\pi^{(l)}(f)$, as $n \rightarrow \infty$. The second referenced article also provide exponential inequalities and uniform estimates with respect to the time parameter l .

To describe precisely the fluctuations of the empirical measures $\eta_n^{(l)}$ around their limiting values $\pi^{(l)}$, we need to introduce some notation. We denote by D_l the first integral operator $\mathcal{D}_{l, \pi^{(l-1)}}$ associated with the target measure $\pi^{(l-1)}$, and we set $D_{k,l}$ with $0 \leq k \leq l$, for the corresponding semigroup. More formally, we have that

$$D_l := \mathcal{D}_{l, \pi^{(l-1)}} \quad \text{and} \quad \forall 1 \leq k \leq l \quad D_{k,l} = D_k D_{k+1} \dots D_l$$

For $k > l$, we use the convention $D_{k,l} = Id$, the identity operator. The reader may find in section 3 an explicit functional representation of these semigroups in the context of Feynman-Kac models.

The functional central limit theorem describing the fluctuations of the i-MCMC model around the solution of equation (1.1) is stated as follows.

Theorem 2.1 *For every $k \geq 0$, the sequence of random fields $(U_n^{(k)})_{n \geq 0}$ on $\mathcal{B}(S^{(k)})$ defined by*

$$U_n^{(k)} := \sqrt{n} \left[\eta_n^{(k)} - \pi^{(k)} \right]$$

converges in law, as n tends to infinity and in the sense of finite dimensional distributions, to a sequence of Gaussian random fields $U^{(k)}$ on $\mathcal{B}(S^{(k)})$ given by the following formula

$$U^{(k)} := \sum_{0 \leq l \leq k} \frac{\sqrt{(2l)!}}{l!} V^{(k-l)} D_{(k-l)+1,k}$$

In the above display, $(V^{(l)})_{l \geq 0}$ stands for a collection of independent and centered Gaussian fields with a covariance function given for any $(f, g) \in \mathcal{B}(S^{(l)})^2$ by

$$\mathbb{E} \left(V^{(l)}(f) V^{(l)}(g) \right) = \pi^{(l)} \left[(f - \pi^{(l)}(f)) (g - \pi^{(l)}(g)) \right] \quad (2.3)$$

In view of these fluctuations, and recalling that for any centered Gaussian variable W , we have $\mathbb{E}(W^{2m}) = 2^{-m} (2m)_m$, the reader should be convinced that the estimates presented in next theorem are sharp with respect to the pair of parameters m and k .

Theorem 2.2 For any $k, n \geq 0$, and $f \in \text{Osc}_1(S^{(k)})$ and any integer $m \geq 1$ we have the non asymptotic estimates

$$\begin{aligned} & \sqrt{2(n+1)} \mathbb{E} \left(\left| [\eta^{(k)} - \pi^{(k)}](f) \right|^m \right)^{\frac{1}{m}} \\ & \leq a(m) \sum_{0 \leq l \leq k} \frac{\sqrt{(2l)!}}{l!} \beta(D_{(k-l)+1,k}) + b(m) c(k) \frac{(\log(n+1))^k}{\sqrt{(n+1)}} \end{aligned}$$

with the collection of constants $a(m)$ given by

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

and some constant $b(m)$ whose values only depend on the parameter m .

3 Feynman-Kac semigroups

In this section, we illustrate the fluctuation results presented in this article with the Feynman-Kac transformations Φ_l given below

$$\forall l \geq 0 \quad \forall (\mu, f) \in (\mathcal{P}(S^{(l)}) \times \mathcal{B}(S^{(l+1)})) \quad \Phi_{l+1}(\mu)(f) := \mu(G_l L_{l+1}(f)) / \mu(G_l) \quad (3.1)$$

In the above displayed formula, G_l is a positive potential function on $S^{(l)}$, and L_{l+1} stands for a collection of Markov transitions from $S^{(l)}$ into $S^{(l+1)}$. In this situation, the solution of the measure valued equation (1.1) is given by the normalized Feynman-Kac distribution flow described below

$$\pi^{(l)}(f) = \gamma^{(l)}(f) / \gamma^{(l)}(1) \quad \text{with} \quad \gamma^{(l)}(f) := \mathbb{E} \left(f(Y_l) \prod_{0 \leq k < l} G_k(Y_k) \right)$$

where $(Y_l)_{l \geq 0}$ stands for a Markov chain taking values in the state spaces $(S^{(l)})_{l \geq 0}$, with initial distribution $\pi^{(0)}$ and Markov transitions $(L_l)_{l \geq 1}$.

These probabilistic models arise in a variety of applications including nonlinear filtering and rare event analysis as well as spectral analysis of Schrödinger type operators and directed polymer analysis. Even if they look innocent, these Feynman-Kac distribution flows are very complex mathematical objects. For instance, the reference Markov chain may represent the paths from the origin up to the current time of an auxiliary sequence of random variables Y'_l taking values in some state spaces E'_l . To be more precise, we have

$$Y_l := (Y'_0, \dots, Y'_l) \in S^{(l)} := (E'_0 \times \dots \times E'_l) \quad (3.2)$$

To get an intuitive feel of the motion of these i-MCMC algorithm, we provide a brief description of the i-MCMC model associated with the Feynman-Kac distribution flows (3.1). In this situation, we first observe that each term in the right hand side product of the formula (1.3) takes the following form

$$\Phi_k \left(\frac{1}{p} \sum_{0 \leq q < p} \delta_{X_q^{(k-1)}} \right) = \sum_{0 \leq q < p} \frac{G_{k-1}(X_q^{(k-1)})}{\sum_{0 \leq q' < p} G_{k-1}(X_{q'}^{(k-1)})} L_k(X_q^{(k-1)}, \cdot)$$

From this observation, we see that each random state $X_p^{(k)}$ with $1 \leq p \leq n$, is sampled according to two separate genetic type mechanisms. Firstly, we select randomly one state $X_q^{(k-1)}$ at level $(k-1)$, with a probability proportional to its potential value $G_{k-1}(X_q^{(k-1)})$.

Then, we evolve randomly from this state according to the exploration transition L_k . This biology-inspired i-MCMC model can be interpreted as a spatial branching and interacting process. In this interpretation, the k -th chain duplicates the individual with large potential functions, at the expense of light individual with poor potential value die. The selected offsprings evolve randomly from the state spaces $S^{(k-1)}$ to the state spaces $S^{(k)}$ at the next level.

The same description for path space models (3.2) coincides with the evolution of genealogical tree based i-MCMC models.

For this class of Feynman-kac models, we observe that the decomposition (2.1) is satisfied with the first order integral operator $D_{l,\eta}$ defined by

$$\forall f \in \mathcal{B}(S^{(l)}) \quad D_{l,\eta}(f) := \frac{G_{l-1}}{\eta(G_{l-1})} L_l(f - \Phi_l(\eta)(f))$$

with the remainder measures $\Xi_l(\mu, \eta)$ given by

$$\forall f \in \mathcal{B}(S^{(l)}) \quad \Xi_l(\mu, \eta)(f) := -\frac{1}{\mu(G_{l-1})} (\mu - \eta)^{\otimes 2} (G_{l-1} \otimes D_{l,\eta}(f))$$

One can observe that the regularity condition (2.2) is satisfied as soon as we have

$$\forall l \geq 0 \quad 0 < \inf G_l \leq \sup G_l < \infty$$

Indeed, it is easily checked that for any function $f \in \mathcal{B}_1(S^{(l)})$, we have

$$|\Xi_l(\mu, \eta)(f)| \leq \frac{1}{(\inf G_{l-1})^2} [|(\mu - \eta)^{\otimes 2} (G_{l-1} \otimes L_l(f))| + |(\mu - \eta)^{\otimes 2} (G_{l-1}^{\otimes 2})|]$$

Finally, we mention that the semigroup $D_{k,l}$ introduced above can be explicitly described in terms of the semigroup

$$Q_{k,l} = Q_k Q_{k+1} \dots Q_l$$

associated with the Feynman-Kac integral operator $Q_l(f) := G_{l-1} L_l(f)$. More precisely, we have that

$$\forall 1 \leq k \leq l \quad D_{k,l}(f) = \frac{Q_{k,l}}{\pi^{(k-1)}(Q_{k,l}(1))} (f - \pi^{(l)}(f))$$

The Dobrushin coefficient $\beta(P_{k,l})$ can also be computed in terms of these semigroups. Firstly, we observe that

$$D_{k,l}(f)(x) = \int (P_{k,l}(f)(x) - P_{k,l}(f)(y)) \frac{Q_{k,l}(1)(y)}{\pi^{(k-1)} Q_{k,l}(1)} \frac{Q_{k,l}(1)(x)}{\pi^{(k-1)} Q_{k,l}(1)} \pi^{(k-1)}(dy)$$

with the Markov integral operator $P_{k,l}$ given below

$$\forall 0 \leq l \leq l \quad P_{k,l}(f)(x) := \frac{Q_{k,l}(f)(x)}{Q_{k,l}(1)(x)}$$

Thus, we find that

$$\|D_{k,l}(f)\| \leq \frac{\|Q_{k,l}(1)\|}{\pi^{(k-1)} Q_{k,l}(1)} \beta(P_{k,l}) \text{osc}(f)$$

from which we conclude that

$$\beta(D_{k,l}) \leq 2 \frac{\|Q_{k,l}(1)\|}{\pi^{(k-1)} Q_{k,l}(1)} \beta(P_{k,l})$$

4 Fluctuations of the local interaction fields

4.1 Introduction

As for mean field particle models (cf. section 9.2 in [2]), the local errors induced by the i-MCMC stochastic algorithm can be interpreted as small disturbances of the measure valued equation (1.1). To be more precise, we need a few definitions.

Definition 4.1 Denote by $\mathcal{F}_l^{(n)}$ the sigma field generated by the collection of random variables $(X_p^{(k)})_{0 \leq p \leq n}$, with $0 \leq k \leq l$, and let $\mathcal{F}^{(n)} := (\mathcal{F}_l^{(n)})_{l \geq 0}$ be the corresponding filtration. Consider the flow of centered random measures $\delta^{(l)} := (\delta_n^{(l)})_{n \in \mathbb{N}} \in \mathcal{M}(S^{(l)})^{\mathbb{N}}$ defined for any $n \geq 0$ by the following formula

$$\forall l \geq 0 \quad \forall n \geq 0 \quad \delta_n^{(l)} = \left[\delta_{X_n^{(l)}} - \Phi_l \left(\eta_{n-1}^{(l-1)} \right) \right]$$

For $n = 0$, we use the convention $\Phi_l \left(\eta_{-1}^{(l-1)} \right) = \nu^{(l)}$, so that $\delta_0^{(l)} = \left[\delta_{X_0^{(l)}} - \nu^{(l)} \right]$.

In our context the i-MCMC model at level l consists in a series of conditionally independent random variables $X_n^{(l)}$ with different distributions $\Phi_l(\eta_{n-1}^{(l-1)})$, at every time step n . Thus the centered local sampling errors $\Delta_n^{(l)}$ are defined by the following random fields models

$$\Delta_n^{(l)} := \frac{1}{\sqrt{n+1}} \sum_{0 \leq p \leq n} \delta_p^{(l)} \quad (4.1)$$

By definition of the i-MCMC models, we readily find that for any $f \in \mathcal{B}(S^{(l)})$, we have

$$\begin{aligned} \mathbb{E}(\Delta_n^{(l)}(f)) &= 0 \\ \mathbb{E}([\Delta_n^{(l)}(f)]^2) &= \frac{1}{n+1} \sum_{0 \leq p \leq n} \Phi_l \left(\eta_{p-1}^{(l-1)} \right) \left(\left[f - \Phi_l \left(\eta_{p-1}^{(l-1)} \right) (f) \right]^2 \right) \end{aligned} \quad (4.2)$$

$$\rightarrow \pi^{(l)} \left(\left[f - \pi^{(l)}(f) \right]^2 \right) \quad \text{as } n \rightarrow \infty \quad (4.3)$$

This shows that the random disturbances $\Delta_n^{(l)}(f)$ are unbiased with finite variance. Much more is true, applying the traditional central limit theorem for triangular arrays, see for instance theorem 4 on page 543 in the textbook of A. Shiryaev [10], we find that $\Delta_n^{(l)}(f)$ converges in law as $n \rightarrow \infty$ to a centered Gaussian random variable with variance given by (4.3).

If we rewrite $\Delta_n^{(l)}$ in a different way, we have the following decomposition

$$\begin{aligned} \eta_n^{(l)} &= \frac{1}{n+1} \sum_{0 \leq p \leq n} \delta_{X_p^{(l)}} \\ &= \frac{1}{n+1} \sum_{0 \leq p \leq n} \Phi_l(\eta_{p-1}^{(l-1)}) + \frac{1}{\sqrt{n+1}} \Delta_n^{(l)} \end{aligned} \quad (4.4)$$

The above formula shows that the evolution equation of the flow of occupation measures $\eta_n^{(l)}$ can be interpreted as a solution of the following measure valued equation

$$\forall n \geq 0 \quad \mu_n^{(l)} = \frac{1}{n+1} \sum_{0 \leq p \leq n} \Phi_l(\mu_{p-1}^{(l-1)}) \quad (4.5)$$

with a sequence of initial conditions $\mu_n^{(0)} = \eta_n^{(0)}$. We end this discussion with the observation that the constant distribution flow $\mu_n^{(l)} = \pi^{(l)}$ associated with the solution of (1.1) also

satisfies (4.5). In this sense, we can interpret the evolution equation (4.4) of the flow of occupation measures $\eta_n^{(l)}$ as a stochastic perturbation of the flow (4.5).

Although the above discussion gives some insight on the asymptotic normal behavior of the local errors accumulated by the i-MCMC model, it do not give directly the precise fluctuations of the sequence of measures $(\eta_n^{(l)})_{n \geq 0}$ around their limiting values $\pi^{(l)}$. The stochastic analysis of these fluctuations is based on the way the semigroup associated with the evolution (4.5) propagates these local perturbation random fields. One can observe that the one step transformation of the equation (4.5) is an averaging of a non linear transformation Φ_l of the complete flow of measures $\mu_p^{(l-1)}$ from the origin $p = 0$ up to the current time n . Iterating these transformation, we end up with a complex semigroup on the flow of measures between successive levels. We refer the interested reader to [3] for an "explicit" calculation of these semigroups in the context of Feynman-Kac models. In the further development of section 5, we shall derive a first order expansion of these semigroups. These multilevel decompositions express the propagations of the local perturbation errors in terms of weighted random fields. The next section is devoted to the fluctuations analysis of a general class of weighted random fields associated with the i-MCMC model. We present an admissible class of array type weight functions for which the corresponding weighted perturbation random fields behaves as a collection of independent and centered Gaussian fields.

4.2 A martingale convergence theorem

This section is concerned with the fluctuation analysis of weighted random fields associated with the local perturbation random fields (4.1) discussed in section 4.1. Our fluctuation theorem is basically stated in terms of the convergence of a sequence of martingales. To describe these processes, we need a pair of definitions.

Definition 4.2 *We let \mathcal{W} be the set of non negative weight array functions $(w_n(p))_{0 \leq p \leq n, 0 \leq n}$, satisfying the following conditions*

$$\lim_{n \rightarrow \infty} \vee_{0 \leq p \leq n} w_n(p) = 0$$

and, for all $\epsilon \in [0, 1]$

$$\varpi(\epsilon) := \lim_{n \rightarrow \infty} \sum_{0 \leq p \leq \lfloor \epsilon n \rfloor} w_n^2(p) < \infty$$

for some scaling function ϖ such that $\lim_{(\epsilon_0, \epsilon_1) \rightarrow (0-, 1+)} (\varpi(\epsilon_0), \varpi(\epsilon_1)) = (0, 1)$. We observe that the traditional and constant fluctuation rates sequence $w_n(p) = 1/\sqrt{n}$ belongs to \mathcal{W} , with the identity function $\varpi(\epsilon) = \epsilon$.

Definition 4.3 *For any $l \geq 0$, we associate to a given weight sequence $w^{(l)} \in \mathcal{W}$ with scaling functions $\varpi^{(l)}$ the mapping*

$$W^{(l)} : \eta \in \mathcal{M}(S^{(l)})^{\mathbb{N}} \mapsto W^{(l)}(\eta) = (W_n^{(l)}(\eta))_{n \geq 0} \in \mathcal{M}(S^{(l)})^{\mathbb{N}}$$

defined for any flow of measures $\eta = (\eta_n)_{n \geq 0} \in \mathcal{M}(S^{(l)})^{\mathbb{N}}$, and any $n \geq 0$, by the weighted measures

$$W_n^{(l)}(\eta) = \sum_{0 \leq p \leq n} w_n^{(l)}(p) \eta_p$$

Denote by $f = (f_l)_{l \geq 0} \in \prod_{l \geq 0} \mathcal{B}(S^{(l)})^d$ a collection of d -valued functions, and for every $1 \leq i \leq d$, let f^i be the i -th coordinate collection of functions $f^i = (f_l^i)_{l \geq 0} \in \prod_{l \geq 0} \mathcal{B}(S^{(l)})$.

We consider the \mathbb{R}^d -valued and $\mathcal{F}^{(n)}$ -martingale $\mathcal{M}^{(n)}(f) = (\mathcal{M}^{(n)}(f^i))_{1 \leq i \leq d}$ defined for any $l \geq 0$ and any $1 \leq i \leq d$, by

$$\mathcal{M}_l^{(n)}(f^i) := \sum_{0 \leq k \leq l} W_n^{(k)}(\delta^{(k)})(f_k^i)$$

We extend this discrete generation process to the half line $\mathbb{R}_+ = [0, \infty[$ by setting

$$\forall t \in \mathbb{R}_+ \quad \mathcal{M}_t^{(n)}(f) := \mathcal{M}_{[t]}^{(n)}(f) + \sum_{p=0}^{\lfloor \{t\}(n+1) \rfloor - 1} w_n^{(\lfloor t \rfloor + 1)}(p) \delta^{(\lfloor t \rfloor + 1)}(f_{(\lfloor t \rfloor + 1)}) \quad (4.6)$$

Notice that $\mathcal{M}_t^{(n)}(f)$ is a càdlàg martingale with respect to the filtration associated with the σ -fields $\mathcal{F}_t^{(n)}$ generated by the random variables $(X_p^{(k)})_{0 \leq p \leq n}$, with $k \leq \lfloor t \rfloor$, and $X_p^{(\lfloor t \rfloor + 1)}$ with $0 \leq p < \lfloor \{t\}(n+1) \rfloor$.

Theorem 4.4 *The sequence of martingales $\mathcal{M}_t^{(n)}(f)$ defined in (4.6) converges in law, as $n \rightarrow \infty$ to an \mathbb{R}^d -valued and Gaussian martingale $\mathcal{M}_t(f) = (\mathcal{M}_t(f^i))_{1 \leq i \leq d}$ such that for any $l \geq 0$ and any pair of indexes $1 \leq i, j \leq d$*

$$\langle \mathcal{M}(f^i), \mathcal{M}(f^j) \rangle_t = \mathcal{C}_{[t]}(f^i, f^j) + \varpi^{(\lfloor t \rfloor + 1)}(\{t\}) [\mathcal{C}_{[t] + 1}(f^i, f^j) - \mathcal{C}_{[t]}(f^i, f^j)]$$

with

$$\mathcal{C}_t(f^i, f^j) := \sum_{0 \leq k \leq \lfloor t \rfloor} \pi^{(k)} \left[(f_k^i - \pi^{(k)}(f_k^i)) (f_k^j - \pi^{(k)}(f_k^j)) \right]$$

Before getting into the proof of this convergence, we illustrate some direct consequences of this theorem.

Firstly, we observe that the discrete generation martingales $\mathcal{M}_l^{(n)}(f)$, with $l \in \mathbb{N}$, converges in law, as $n \rightarrow \infty$ to an \mathbb{R}^d -valued and discrete generation Gaussian martingale $\mathcal{M}_l(f) = (\mathcal{M}_l(f^i))_{1 \leq i \leq d}$ with bracket given for any $l \geq 0$ and any $1 \leq i, j \leq d$ by

$$\langle \mathcal{M}(f^i), \mathcal{M}(f^j) \rangle_l = \mathcal{C}_l(f^i, f^j)$$

Next, we fix a parameter l and we introduce a collection of functions $g_k^j = (g_k^{j,i})_{1 \leq i \leq d_j} \in \mathcal{B}(S^{(k)})^{d_j}$, with some fixed dimension parameters $d_j \geq 1$, with $0 \leq j \leq l$ and $0 \leq k \leq l$.

For every $0 \leq k \leq l$, we take $f = (f_k)_{0 \leq k \leq l}$ with $f_k = (g_k^j)_{0 \leq j \leq l} \in \mathcal{B}(S^{(k)})^d$ and $d = \sum_{i=0}^k d_k$. We further assume that $g_k^j = 1_{j=k} g_k^k$. In other words, all the functions g_k^j are null except on the diagonal $k = j$. By construction, for every $0 \leq k \leq l$, we have

$$\mathcal{M}_k^{(n)}(f) = (\mathcal{M}_k^{(n)}(g^j))_{0 \leq j \leq l} \quad \text{and} \quad \mathcal{M}_k^{(n)}(g^j) = 1_{k \geq j} V_n^{(j)}(g^j)$$

Assuming that the martingale convergence theorem given above holds true, we can conclude that the random fields $V_n^{(j)}$, with $0 \leq j \leq l$, converges in law, as n tends to infinity and in the sense of finite dimensional distributions, to a sequence of independent and centered Gaussian fields $(V^{(j)})_{0 \leq j \leq l}$, with for any $1 \leq i, i' \leq d_j$

$$\begin{aligned} \mathbb{E} \left[V^{(j)}(g_j^{j,i}) V^{(j')}(g_j^{j',i'}) \right] &= \langle \mathcal{M}(g_j^{j,i}) \mathcal{M}(g_j^{j',i'}) \rangle_l \\ &= 1_{j=j'} \pi^{(j)} \left[(g_j^{j,i} - \pi^{(j)}(g_j^{j,i})) (g_j^{j,i'} - \pi^{(j)}(g_j^{j,i'})) \right] \end{aligned}$$

We summarize this discussion with the following corollary of theorem 4.4.

Corollary 4.5 *For every collection of weight functions $(w^{(l)})_{l \geq 0} \in \mathcal{W}^{\mathbb{N}}$, the sequence of random fields $(V_n^{(l)})_{l \geq 0}$ defined by $V_n^{(l)} := W_n^{(l)}(\delta^{(l)})$, converges in law, as n tends to infinity and in the sense of finite dimensional distributions, to a sequence of independent and centered Gaussian fields $(V^{(l)})_{l \geq 0}$ with covariance function given by*

$$\forall l \geq 0 \quad \forall (f, g) \in \mathcal{B}(S^{(l)})^2 \quad \mathbb{E} \left(V^{(l)}(f) V^{(l)}(g) \right) = \pi^{(l)} \left[(f - \pi^{(l)}(f)) (g - \pi^{(l)}(g)) \right]$$

Now, we come to the proof of theorem 4.4.

Proof of theorem 4.4:

Firstly, we notice that the martingale increments $V_n^{(l)}(f_l^i)$ are given by

$$V_n^{(k)}(f_k^i) := W_n^{(k)}(\delta^{(k)})(f_k^i) = \sum_{0 \leq p \leq n} w_n^{(k)}(p) \left[f_k^i(X_p^{(k)}) - \Phi_k \left(\eta_{p-1}^{(k-1)} \right) (f_k^i) \right]$$

To use the central limit theorem for triangular arrays of \mathbb{R}^d -valued random variables (Theorem 3.33, p. 437 in [7]) we first rewrite the martingale $\mathcal{M}_l^{(n)}(f)$ in the following form

$$\mathcal{M}_l^{(n)}(f) = \sum_{i=0}^{l(n+1)+n} \mathcal{V}_i^{(n)}(f)$$

where for every $0 \leq i \leq l(n+1)+n$, with $i = k(n+1)+p$ for some $0 \leq k \leq l$, and $0 \leq p \leq n$

$$\mathcal{V}_i^{(n)}(f) := w_n^{(k)}(p) \left[f_k(X_p^{(k)}) - \Phi_k \left(\eta_{p-1}^{(k-1)} \right) (f_k) \right]$$

We further denote by $\mathcal{G}_i^{(n)}$ the σ -field generated by the random variables $X_p^{(k)}$ for any pair of parameters (k, p) such that $k(n+1)+p \leq i$.

By construction, for any flow of functions $f = (f_l)_{l \geq 0}$ and $g = (g_l)_{l \geq 0} \in \prod_{l \geq 0} \mathcal{B}(S^{(l)})$ and for every $0 \leq i \leq l(n+1)+n$, with $i = k(n+1)+p$ for some $0 \leq k \leq l$, and $0 \leq p \leq n$, we find that

$$\begin{aligned} \mathbb{E} \left[\mathcal{V}_i^{(n)}(f) \mid \mathcal{G}_{i-1}^{(n)} \right] &= 0 \\ \mathbb{E} \left[\mathcal{V}_i^{(n)}(f) \mathcal{V}_i^{(n)}(g) \mid \mathcal{G}_{i-1}^{(n)} \right] &= w_n^{(k)}(p)^2 C_p^{(k)}(f, g) \end{aligned}$$

with the local covariance function

$$C_p^{(k)}(f, g) := \Phi_k(\eta_{p-1}^{(k-1)}) \left([f_k - \Phi_k(\eta_{p-1}^{(k-1)})(f_k)] [g_k - \Phi_k(\eta_{p-1}^{(k-1)})(g_k)] \right)$$

This implies that

$$\sum_{i=k(n+1)}^{k(n+1)+n} \mathbb{E} \left[\mathcal{V}_i^{(n)}(f) \mathcal{V}_i^{(n)}(g) \mid \mathcal{G}_{i-1}^{(n)} \right] = \sum_{p=0}^n w_n^{(k)}(p)^2 C_p^{(k)}(f, g)$$

We recall that the quantities $\eta_n^{(k)}(h)$ converge almost surely to $\pi^{(k)}(h)$, as $n \rightarrow \infty$, for every $k \geq 0$ and any function $h \in \mathcal{B}(S^{(k)})$. Under our regularity conditions on the mappings Φ_k , we find that $\Phi_k(\eta_n^{(k-1)})(h)$ converge almost surely to $\pi^{(k)}(h)$, as $n \rightarrow \infty$, for any function $h \in \mathcal{B}(S^{(k)})$. It readily follows the following convergence result

$$\lim_{n \rightarrow \infty} C_n^{(k)}(f, g) = C^{(k)}(f, g) := \pi^{(k)}((f_k - \pi^{(k)}(f_k)) (g_k - \pi^{(k)}(g_k))) \quad \mathbb{P} - \text{a.e.}$$

Under our assumptions on the weight functions $w^{(k)}$, we deduce that

$$\lim_{n \rightarrow \infty} \sum_{p=0}^n w_n^{(k)}(p)^2 C_p^{(k)}(f, g) = C^{(k)}(f, g) \quad \mathbb{P} - \text{a.e.}$$

from which we conclude that

$$\lim_{n \rightarrow \infty} \langle \mathcal{M}^{(n)}(f), \mathcal{M}^{(n)}(g) \rangle_l = \sum_{0 \leq k \leq l} C^{(k)}(f, g) = \mathcal{C}_l(f, g) \quad \mathbb{P} - \text{a.e.}$$

To get one step further, we introduce the sequence of \mathbb{R}^d -valued and continuous time càdlàg random process defined for any $f = (f_l)_{l \geq 0} \in \prod_{l \geq 0} \mathcal{B}(S^{(l)})^d$ by

$$t \in \mathbb{R}_+ \mapsto \sum_{i=0}^{\lfloor t(n+1) \rfloor + n} \mathcal{V}_i^{(n)}(f)$$

Notice that for any $l \in \mathbb{N}$, and any $0 \leq k \leq n$, we have

$$l + \frac{k}{n+1} \leq t < l + \frac{k+1}{n+1} \implies \lfloor t \rfloor = l \quad \text{and} \quad \lfloor t(n+1) \rfloor - \lfloor t \rfloor(n+1) = \lfloor \{t\}(n+1) \rfloor = k$$

where $\{t\} := t - \lfloor t \rfloor$ stands for the fractional part of a real number $t \in \mathbb{R}_+$. Using the fact that

$$\lfloor t(n+1) \rfloor + n = (\lfloor t \rfloor(n+1) + n) + \lfloor \{t\}(n+1) \rfloor$$

we obtain the decomposition

$$\begin{aligned} & \sum_{i=(\lfloor t \rfloor + 1)(n+1)}^{\lfloor t(n+1) \rfloor + n} \mathcal{V}_i^{(n)}(f) \\ &= \sum_{p=0}^{\lfloor \{t\}(n+1) \rfloor - 1} \mathcal{V}_{(\lfloor t \rfloor + 1)(n+1) + p}^{(n)}(f) \\ &= \sum_{p=0}^{\lfloor \{t\}(n+1) \rfloor - 1} w_n^{(\lfloor t \rfloor + 1)}(p) \left(f_{(\lfloor t \rfloor + 1)}(X_p^{(\lfloor t \rfloor + 1)}) - \Phi_{(\lfloor t \rfloor + 1)}(\eta_{p-1}^{\lfloor t \rfloor})(f_{(\lfloor t \rfloor + 1)}) \right) \end{aligned}$$

This readily implies that

$$\sum_{i=0}^{\lfloor t(n+1) \rfloor + n} \mathcal{V}_i^{(n)}(f) = \mathcal{M}_t^{(n)}(f)$$

Furthermore, using the same computations as above, for any pair of integers $1 \leq j, j' \leq d$, we find that

$$\begin{aligned} \sum_{i=0}^{\lfloor t(n+1) \rfloor + n} \mathbb{E}(\mathcal{V}_i^{(n)}(f^j) \mathcal{V}_i^{(n)}(f^{j'}) \mid \mathcal{G}_{i-1}^{(n)}) &= \sum_{k=0}^{\lfloor t \rfloor} \sum_{p=0}^n w_n^{(k)}(p)^2 C_p^{(k)}(f^j, f^{j'}) \\ &\quad + \sum_{p=0}^{\lfloor \{t\}(n+1) \rfloor - 1} w_n^{(\lfloor t \rfloor + 1)}(p)^2 C_p^{(\lfloor t \rfloor + 1)}(f^j, f^{j'}) \end{aligned}$$

Under our assumptions on the weight functions, we find that

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{i=0}^{\lfloor t(n+1) \rfloor + n} \mathbb{E} \left[\mathcal{V}_i^{(n)}(f^j) \mathcal{V}_i^{(n)}(f^{j'}) \mid \mathcal{G}_{i-1}^{(n)} \right] &= \mathcal{C}_{\lfloor t \rfloor}(f^j, f^{j'}) \\ &\quad + \varpi^{(\lfloor t \rfloor + 1)}(\{t\}) \left[\mathcal{C}_{\lfloor t \rfloor + 1}(f^j, f^{j'}) - \mathcal{C}_{\lfloor t \rfloor}(f^j, f^{j'}) \right] \end{aligned}$$

Since we also have

$$\|\mathcal{V}_i^{(n)}(f)\| \leq \left[\sup_{0 \leq k \leq l} \|f_k\| \right] \left[\sup_{0 \leq k \leq l} \sup_{0 \leq p \leq n} w_n^{(k)}(p) \right]$$

for every $0 \leq i \leq l(n+1) + n$, and

$$\lim_{n \rightarrow \infty} \bigvee_{0 \leq k \leq l} \bigvee_{0 \leq p \leq n} w_n^{(k)}(p) = 0$$

we conclude that the conditional Lindeberg condition is satisfied. Thus, the \mathbb{R}^d -valued martingale $\mathcal{M}_t^{(n)}(f)$ converge in law, as n tends to infinity, to a continuous martingale $\mathcal{M}_t(f)$ with predictable bracket given, for any pair of indexes $1 \leq j, j' \leq d$, by

$$\langle \mathcal{M}_t(f^j), \mathcal{M}_t(f^{j'}) \rangle_t = \mathcal{C}_{[t]}(f^j, f^{j'}) + \varpi^{(\lfloor t \rfloor + 1)}(\{t\}) \left[\mathcal{C}_{[t]+1}(f^j, f^{j'}) - \mathcal{C}_{[t]}(f^j, f^{j'}) \right]$$

which completes the proof of the theorem. \blacksquare

5 A multilevel expansion formula

We present in this section a multilevel decomposition of the flow of occupation measures $\eta_n^{(k)}$ around their limiting value $\pi^{(k)}$. In section 4.1, we have shown that $\eta_n^{(k)}$ satisfies a stochastic perturbation equation (4.4) of a measure valued equation (4.5) involving complex time averaging one step mappings. In this interpretation, the forthcoming developments provides a first order expansion of the semigroup associated with the equation (4.4).

This result will be used in various places in the further development of the article. Firstly, in section 6 we combine these first order developments with the local fluctuation analysis presented in section 4 to derive a "two lines proof" of the functional central limit theorem 2.1. In section 7, we use these decompositions to prove the non asymptotic \mathbb{L}_m -estimates presented in theorem 2.2.

The forthcoming multilevel expansion is expressed in terms of the following time averaging operators.

Definition 5.1 For any $l \geq 0$, denote by \mathbb{S} the mapping

$$\mathbb{S} : \eta \in \mathcal{M}(S^{(l)})^{\mathbb{N}} \mapsto \mathbb{S}(\eta) = (\mathbb{S}_n(\eta))_{n \geq 0} \in \mathcal{M}(S^{(l)})^{\mathbb{N}}$$

defined for any flow of measures $\eta = (\eta_n)_{n \geq 0} \in \mathcal{M}(S^{(l)})^{\mathbb{N}}$, and any $n \geq 0$, by the measures

$$\mathbb{S}_n(\eta) = \frac{1}{n+1} \sum_{0 \leq p \leq n} \eta_p \quad (5.1)$$

We let $\mathbb{S}^k := \mathbb{S} \circ \mathbb{S}^{k-1}$ be k -th iterate of the mapping \mathbb{S} .

We are now in position to state and to prove the following pivotal multilevel expansion.

Proposition 5.2 For every $k \geq 0$, we have the following multilevel expansion

$$\eta^{(k)} - \pi^{(k)} = \sum_{0 \leq l \leq k} \mathbb{S}^{l+1}(\delta^{(k-l)}) D_{(k-l)+1,k} + \Xi^{(k)} \quad (5.2)$$

with a flow of signed and random measures $\Xi^{(k)} = (\Xi_n^{(k)})_{n \geq 0}$ such that

$$\forall m \geq 1 \quad \sup_{f \in \mathcal{B}_1(S^{(k)})} \mathbb{E}(|\Xi_n^{(k)}(f)|^m)^{\frac{1}{m}} \leq b(m) c(k) (\log(n+1))^k / (n+1)$$

In the above display, $b(m)$ stands for some constant whose values only depend on m .

Proof:

We prove the expansion by induction on the parameter k . For $k = 0$, we readily find that

$$\eta_n^{(0)} - \pi^{(0)} = \frac{1}{n+1} \sum_{0 \leq p \leq n} \left[\delta_{X_p^{(0)}} - \pi^{(0)} \right] = \mathbb{S}_n(\delta^{(0)})$$

Notice that in this case (5.2) is satisfied with the sequence of null measures $\Xi_n^{(0)} = 0$. We further assume that the expansion is satisfied at rank k . The proof of the result at rank $(k+1)$ is based on the following decomposition

$$\begin{aligned} \eta_n^{(k+1)} - \pi^{(k+1)} &= \mathbb{S}_n(\delta^{(k+1)}) + \frac{1}{n+1} \sum_{0 \leq p \leq n} \left[\Phi_{k+1}(\eta_{p-1}^{(k)}) - \Phi_{k+1}(\eta_p^{(k)}) \right] \\ &\quad + \frac{1}{n+1} \sum_{0 \leq p \leq n} \left[\Phi_{k+1}(\eta_p^{(k)}) - \Phi_{k+1}(\pi^{(k)}) \right] \end{aligned} \quad (5.3)$$

It follows from our assumptions on the mappings Φ_k , that we have

$$\forall \mu, \eta \in \mathcal{P}(S^{(k)}) \quad \|\Phi_{k+1}(\mu) - \Phi_{k+1}(\eta)\| \leq c(k) \|\mu - \eta\|$$

with $c(k) \leq \left(\sup_{\gamma \in \mathcal{P}(S^{(k)})} \|D_{k+1, \gamma}\| + \chi_{k+1} \right)$. Using the fact that

$$\eta_p^{(k)} - \eta_{p-1}^{(k)} = \frac{1}{p+1} \left(\delta_{X_p^{(k)}} - \eta_{p-1}^{(k)} \right)$$

we can deduce that

$$\|\Phi_{k+1}(\eta_p^{(k)}) - \Phi_{k+1}(\eta_{p-1}^{(k)})\| \leq c(k)/(p+1)$$

which leads to

$$\left\| \frac{1}{n+1} \sum_{0 \leq p \leq n} \left[\Phi_{k+1}(\eta_p^{(k)}) - \Phi_{k+1}(\eta_{p-1}^{(k)}) \right] \right\| \leq c(k) \log(n+1)/(n+1) \quad (5.4)$$

On the other hand, using the first order expansion of the mapping Φ_{k+1} we have the decomposition

$$\Phi_{k+1}(\eta_p^{(k)}) - \Phi_{k+1}(\pi^{(k)}) = (\eta_p^{(k)} - \pi^{(k)})D_{k+1} + \Xi_{k+1}(\eta_p^{(k)}, \pi^{(k)})$$

We let $g = \sum_{i \in I} a_i (h_i^1 \otimes h_i^2) \in \mathcal{T}_2(S^{(k)})$ be a tensor product function associated with a subset $I \subset \mathbb{N}$, a pair of functions $(h_i^1, h_i^2)_{i \in I} \in (\mathcal{B}(S^{(k)})^2)^I$, and a sequence of numbers $(a_i)_{i \in I} \in \mathbb{R}^I$. Using the generalized Minkowski integral inequality we have

$$\begin{aligned} &\mathbb{E} \left[\left| (\eta_p^{(k)} - \pi^{(k)})^{\otimes 2}(g) \right|^m \right]^{\frac{1}{m}} \\ &\leq \sum_{i \in I} |a_i| \mathbb{E} \left[\left([\eta_p^{(k)} - \pi^{(k)}](h_i^1) \right)^{2m} \right]^{\frac{1}{2m}} \mathbb{E} \left[\left([\eta_p^{(k)} - \pi^{(k)}](h_i^2) \right)^{2m} \right]^{\frac{1}{2m}} \end{aligned}$$

Hence, via the \mathbb{L}_r -estimates presented in [3], we infer that for any $j = 1, 2$ and $i \in I$

$$\mathbb{E} \left[\left([\eta_p^{(k)} - \pi^{(k)}](h_i^j) \right)^{2m} \right]^{\frac{1}{2m}} \leq a(m) \|h_i^j\| / \sqrt{(p+1)}$$

for some finite constant $a(m)$ whose values only depend on the parameter m . These estimates readily implies that

$$\mathbb{E} \left[\left| (\eta_p^{(k)} - \pi^{(k)})^{\otimes 2}(g) \right|^m \right]^{\frac{1}{m}} \leq b(m) c(k) |g|/(p+1) \quad (5.5)$$

for some finite constant $b(m)$ whose values only depend on the parameter m . This implies that for any $f \in \mathcal{B}_1(S^{(k)})$

$$\mathbb{E} \left(\left| \frac{1}{n+1} \sum_{0 \leq p \leq n} \Xi_{k+1}(\eta_p^{(k)}, \pi^{(k)})(f) \right|^m \right)^{\frac{1}{m}} \leq b(m) c(k) \log(n+1)/(n+1)$$

To take the final step, we use the induction hypothesis to check that

$$(\eta^{(k)} - \pi^{(k)})D_{k+1} = \sum_{0 \leq l \leq k} \mathbb{S}^{(l+1)}(\delta^{(k-l)}) D_{(k-l)+1, k+1} + \Xi^{(k)} D_{k+1}$$

and

$$\sup_{f \in \mathcal{B}_1(S^{(k)})} \mathbb{E}(|\Xi_p^{(k)}(D_{k+1}f)|^m)^{\frac{1}{m}} \leq b(m) c(k) (\log(p+1))^k/(p+1)$$

This yields that

$$\begin{aligned} & \mathbb{S}_n(\delta^{(k+1)}) + \frac{1}{n+1} \sum_{0 \leq p \leq n} (\eta_p^{(k)} - \pi^{(k)})D_{k+1} \\ &= \mathbb{S}_n(\delta^{(k+1)}) + \sum_{0 \leq l \leq k} \mathbb{S}_n^{(l+2)}(\delta^{(k-l)}) D_{(k-l)+1, k+1} + \frac{1}{n+1} \sum_{0 \leq p \leq n} \Xi_p^{(k)} D_{k+1} \\ &= \sum_{0 \leq l \leq k+1} \mathbb{S}_n^{(l+1)}(\delta^{((k+1)-l)}) D_{((k+1)-l)+1, k+1} + \frac{1}{n+1} \sum_{0 \leq p \leq n} \Xi_p^{(k)} D_{k+1} \end{aligned}$$

In the last assertion, we have used the convention $D_{k+2, k+1} = Id$, the identity operator. Also notice that

$$\sup_{f \in \mathcal{B}_1(S^{(k)})} \mathbb{E} \left(\left| \frac{1}{n+1} \sum_{0 \leq p \leq n} \Xi_p^{(k)}(D_{k+1}(f)) \right|^m \right)^{\frac{1}{m}} \leq b(m) c(k) (\log(n+1))^{k+1}/(n+1)$$

The end of the proof of the result at rank $(k+1)$ is now a consequence of the decomposition (5.3) in conjunction with the estimates (5.4) and (5.5). The end of the proof of the proposition is now completed. \blacksquare

6 Proof of the central limit theorem

This section is mainly concerned with the proof of the functional central limit theorem presented in section 2.1. The first ingredient is to express the time averaging semigroup \mathbb{S}^k introduced in definition 5.1 in terms of the following weighted summations

$$\mathbb{S}_n^k(\eta) = \frac{1}{n+1} \sum_{0 \leq p \leq n} s_n^{(k)}(p) \eta_p$$

with the weight array functions $s_n^{(k)} := (s_n^{(k)}(p))_{0 \leq p \leq n}$ defined by the induction

$$\forall k \geq 1 \quad \forall 0 \leq p \leq n \quad s_n^{(k+1)}(p) = \sum_{p \leq q \leq n} \frac{1}{(q+1)} s_n^{(k)}(q) \quad \text{with} \quad s_n^{(1)}(p) := 1$$

The proof of theorem 2.1 is a direct consequence of the following proposition whose proof is postponed to the end of the section.

Proposition 6.1 *For any $k \geq 0$, we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{0 \leq q \leq n} s_n^{(k+1)}(q)^2 = (2k)!/k!^2$$

In addition, for any $k \geq 1$, the weight array functions $w^{(k)}$ defined by

$$\forall n \geq 0 \quad \forall 0 \leq p \leq n \quad w_n^{(k)}(p) := s_n^{(k)}(p) / \sqrt{\sum_{0 \leq q \leq n} s_n^{(k)}(q)^2}$$

belongs the set \mathcal{W} introduced in definition 4.2, with the functions $\varpi^{(k)}$ defined by

$$\varpi^{(k)}(\epsilon) = \sum_{l=0}^{2(k-1)} \frac{1}{l!} \epsilon \left(\log \left(\frac{1}{\epsilon} \right) \right)^l$$

As promised, we are now in position to derive a "two lines" proof of theorem 2.1.

Proof of theorem 2.1:

We fix a parameter k and for any $0 \leq l \leq k$, we let $W^{(l)}$ be the distribution flow mappings associated with the weight functions $w^{((k-l)+1)}$ defined in proposition 6.1 and given by

$$W^{(l)} : \eta \in \mathcal{M}(S^{(l)})^{\mathbb{N}} \mapsto W^{(l)}(\eta) = (W_n^{(l)}(\eta))_{n \geq 0} \in \mathcal{M}(S^{(l)})^{\mathbb{N}}$$

with for any $n \geq 0$ and $\eta = (\eta_n)_{n \geq 0} \in \mathcal{M}(S^{(l)})^{\mathbb{N}}$

$$W_n^{(l)}(\eta) = \sum_{0 \leq p \leq n} w_n^{((k-l)+1)}(p) \eta_p$$

By construction, we have for any $0 \leq l \leq k$

$$(n+1) \mathbb{S}_n^{l+1}(\eta) = \sqrt{\sum_{0 \leq q \leq n} s_n^{(l+1)}(q)^2} W_n^{(k-l)}(\eta)$$

Using the multilevel expansion presented in proposition 5.2, we prove that

$$\sqrt{(n+1)} \left[\eta_n^{(k)} - \pi^{(k)} \right] = \sum_{0 \leq l \leq k} \sqrt{\frac{1}{n+1} \sum_{0 \leq q \leq n} s_n^{(l+1)}(q)^2} W_n^{(k-l)}(\delta^{(k-l)}) D_{(k-l)+1,k} + \bar{\Xi}_n^{(k)}$$

with the remainder signed measure $\bar{\Xi}_n^{(k)}$ satisfying

$$\sup_{f \in \mathcal{B}_1(S^{(k)})} \mathbb{E}(|\bar{\Xi}_n^{(k)}(f)|) \leq c(k) (\log(n+1))^k / \sqrt{(n+1)}$$

The proof of theorem 2.1 is now a direct consequence of corollary 4.5. ■

The proof of proposition 6.1 is in turn based on a pair of pivotal lemmas. Their proof follows elementary techniques but tedious calculations and they are housed in the appendix.

Lemma 6.2 *For any $0 \leq p \leq n$, and any $k \geq 1$, we have the formula*

$$s_n^{(k)}(p) = \frac{1}{(k-1)!} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^{(k-1)} + r_n^{(k)}(p)$$

with a remainder sequence or arrays $r_n^{(k)} = (r_n^{(k)}(p))_{0 \leq p \leq n}$ such that

$$\forall 0 \leq p \leq n \quad |r_n^{(k)}(p)| \leq c(k) (\log(n+1))^{k-3} \sum_{p \leq q \leq n} \frac{1}{(q+1)^2}$$

Lemma 6.3 For any $n \geq 0$, and any $k \geq 0$, we have the estimate

$$\frac{1}{k!} \sum_{0 \leq p \leq n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k = (n+1) + l^k(n)$$

with $|l^k(n)| \leq c(k) (\log(n+1))^k$. In addition, for any increasing sequence $\kappa_n \leq n$ such that $\lim_{n \rightarrow \infty} (\kappa_n/n) = \kappa > 0$, we have

$$\lim_{n \rightarrow \infty} \frac{1}{k!} \frac{1}{n} \sum_{0 \leq p \leq \kappa_n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k = \sum_{0 \leq l \leq k} \frac{1}{l!} \kappa \left(\log \left(\frac{1}{\kappa} \right) \right)^l$$

Now, we are in position to prove proposition 6.1.

Proof of proposition 6.1:

The fact that $\lim_{n \rightarrow \infty} \sup_{0 \leq p \leq n} w_n^{(k)}(p) = 0$ is a direct consequence of the pair of estimates given below

$$\sup_{0 \leq p \leq n} s_n^{(k)}(p) \leq c(k) (\log(n+1))^{k-1}$$

and

$$\sum_{0 \leq p \leq n} s_n^{(k)}(p)^2 = \frac{(2(k-1))!}{(k-1)!^2} (n+1) + \tilde{r}_n^{(k)} \quad (6.1)$$

with $|\tilde{r}_n^{(k)}| \leq c(k) (\log(n+1))^{2(k-1)}$. The first assertion is a direct consequence of lemma 6.2. To prove the second one, we observe that

$$\sum_{0 \leq p \leq n} s_n^{(k)}(p)^2 = \frac{1}{(k-1)!^2} \sum_{0 \leq p \leq n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^{2(k-1)} + \bar{r}_n^{(k)}$$

with

$$\bar{r}_n^{(k)} := \frac{2}{(k-1)!} \sum_{0 \leq p \leq n} \left(\left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^{(k-1)} r_n^{(k)}(p) \right) + \sum_{0 \leq p \leq n} r_n^{(k)}(p)^2$$

Furthermore, we also have

$$\sum_{0 \leq p \leq n} r_n^{(k)}(p)^2 \leq c(k) (\log(n+1))^{2(k-3)} \sum_{0 \leq p \leq n} \left(\sum_{p \leq q \leq n} \frac{1}{(q+1)^2} \right)^2$$

from which we prove the rather crude estimate

$$\begin{aligned} \sum_{0 \leq p \leq n} r_n^{(k)}(p)^2 &\leq c(k) (\log(n+1))^{2(k-3)} \sum_{0 \leq q \leq n} \sum_{0 \leq p \leq q} \frac{1}{(q+1)^2} \\ &= c(k) (\log(n+1))^{2(k-3)} \sum_{0 \leq q \leq n} \frac{1}{(q+1)} \leq c(k) (\log(n+1))^{2(k-3)+1} \end{aligned}$$

In much the same way, we find that

$$\begin{aligned} \sum_{0 \leq p \leq n} \left(\left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^{(k-1)} |r_n^{(k)}(p)| \right) &\leq \left(\sum_{0 \leq q \leq n} \frac{1}{q+1} \right)^{(k-1)} \sum_{0 \leq p \leq n} |r_n^{(k)}(p)| \\ &\leq c(k) \log(n+1)^{(k-1)+(k-3)+1} \end{aligned}$$

In summary, we have proved that $|\bar{r}_n^{(k)}| \leq c(k) \log(n+1)^{(2k-3)}$. We end the proof of (6.1) by a direct application of lemma 6.3. From the above discussion, we also find that

$$\sum_{0 \leq p \leq \kappa_n} s_n^{(k)}(p)^2 = \frac{1}{(k-1)!^2} \sum_{0 \leq p \leq \kappa_n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^{2(k-1)} + \bar{\bar{r}}_n^{(k)} \quad \text{with} \quad |\bar{\bar{r}}_n^{(k)}| \leq |\bar{r}_n^{(k)}|$$

By lemma 6.3, for any increasing sequence $\kappa_n \leq n$ such that $\lim_{n \rightarrow \infty} (\kappa_n/n) = \kappa > 0$ we find that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{0 \leq p \leq \kappa_n} s_n^{(k)}(p)^2 = \frac{(2(k-1))!}{(k-1)!^2} \sum_{0 \leq l \leq 2(k-1)} \frac{1}{l!} \kappa \left(\log \left(\frac{1}{\kappa} \right) \right)^l$$

from which we conclude that

$$\lim_{n \rightarrow \infty} \sum_{0 \leq p \leq \kappa_n} w_n^{(k)}(p)^2 = \sum_{0 \leq l \leq 2(k-1)} \frac{1}{l!} \kappa \left(\log \left(\frac{1}{\kappa} \right) \right)^l$$

This ends the proof of the proposition. ■

7 Proof of the sharp \mathbb{L}_m -estimates

This short section is concerned with the proof of non asymptotic estimates presented in theorem 2.2. We start with a pivotal technical lemma.

Lemma 7.1 *For any $k, l, n \geq 0$, any function $f \in \mathcal{B}(S^{(l)})$, and any integer $m \geq 1$, we have*

$$\mathbb{E} \left[\left| \mathbb{S}_n^{(k)}(\delta^{(l)})(f) \right|^m \right]^{\frac{1}{m}} \leq \frac{a(m)}{\sqrt{2(n+1)}} \left(\frac{1}{(n+1)} \sum_{0 \leq p \leq n} s_n^{(k)}(p)^2 \right)^{\frac{1}{2}} \text{osc}(f)$$

with the collection of constants $a(m)$ defined in theorem 2.2.

Using this lemma, the proof of theorem 2.2 is a routine consequence of the multilevel expansion presented in proposition 5.2, so we give it first.

Proof of theorem 2.2:

Using the expansion (5.2), we find that

$$\begin{aligned} & \sqrt{(n+1)} \mathbb{E} \left[\left| [\eta^{(k)} - \pi^{(k)}](f) \right|^m \right]^{\frac{1}{m}} \\ & \leq a(m) \sum_{0 \leq l \leq k} \left(\frac{1}{2(n+1)} \sum_{0 \leq p \leq n} s_n^{(l+1)}(p)^2 \right)^{\frac{1}{2}} \beta(D_{(k-l)+1,k}) + b(m) c(k) \frac{(\log(n+1))^k}{\sqrt{(n+1)}} \end{aligned}$$

Since $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, by (6.1) we conclude that

$$\left(\frac{1}{(n+1)} \sum_{0 \leq p \leq n} s_n^{(l+1)}(p)^2 \right)^{\frac{1}{2}} \leq \frac{\sqrt{(2l)!}}{l!} + \epsilon^{(k)}(n) \quad \text{with} \quad \epsilon^{(k)}(n) \leq c(l) \frac{(\log(n+1))^l}{\sqrt{n+1}}$$

This clearly ends the proof of the theorem. ■

The proof of lemma 7.1 is in turn based on the following Burkholder-Davis-Gundy type inequality for martingales with symmetric and independent increments.

Lemma 7.2 *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*

$$\mathbb{E}(|M_n|^m)^{\frac{1}{m}} \leq \frac{1}{\sqrt{2}} a(m) \mathbb{E}([M]_n^{m'/2})^{\frac{1}{m'}} \quad (7.1)$$

with the bracket process $[M]_n := \sum_{0 \leq p \leq n} \Delta_p^2$, the smallest even integer $m' \geq m$ and the collection of constants $a(m)$ introduced in proposition 7.1.

This lemma is more or less well known, for completeness an elementary proof is housed in the end of the appendix.

Now, we come to the proof of the lemma.

Proof of lemma 7.1:

We use a symmetry argument. We consider a independent copy $(Y_n^{(l)})_{n \geq 0}$, with $l \geq 0$, of the reference i-MCMC model $(X_n^{(l)})_{n \geq 0}$, with $l \geq 0$. We also assume that these two processes are defined on the same probability space, and we let $\mathcal{H}_l^{(n)}$ be the filtration generated by the process $(Y_p^{(k)})_{0 \leq p \leq n}$ with $k \leq l$. By construction, we have for any $l \geq 0$, any $f_l \in \mathcal{B}(S^{(l)})$ and any $0 \leq p \leq n$

$$\delta_p^{(l)}(f_l) = f_l(X_p^{(l)}) - \Phi_l(\eta_{p-1}^{(l-1)})(f_l) = \mathbb{E} \left(\left[f_l(X_p^{(l)}) - f_l(Y_p^{(l)}) \right] \middle| \mathcal{F}_l^{(n)} \vee \mathcal{H}_{l-1}^{(n)} \right)$$

This readily implies that

$$\sum_{0 \leq p \leq n} s_n^{(k)}(p) \delta_p^{(l)}(f_l) = \mathbb{E} \left(\sum_{0 \leq p \leq n} \Delta_n^{(k)}(p, f_l) \middle| \mathcal{F}_l^{(n)} \vee \mathcal{H}_{l-1}^{(n)} \right)$$

with the sequence of symmetric, and $(\mathcal{F}_{l-1}^{(n)} \vee \mathcal{H}_{l-1}^{(n)})$ -conditionally independent and symmetric random variables $(\Delta_n^{(k)}(p, f_l))_{0 \leq p \leq n}$ given below

$$\Delta_n^{(k)}(p, f_l) := s_n^{(k)}(p) \left[f_l(X_p^{(l)}) - f_l(Y_p^{(l)}) \right]$$

The end of the proof is a direct consequence of lemma 7.2. This ends the proof of the lemma. ■

Appendix

Proof of lemma 6.2

We prove the lemma by induction on the parameter k . For $k = 2$, the result is clearly met with a null remainder sequence $r_n^{(2)} = 0$. We further assume that the result is satisfied at rank k . For any $0 \leq p \leq n$, we denote

$$I_n(p-1) := \sum_{0 \leq q \leq n} \frac{1}{q+1} - \sum_{0 \leq q \leq (p-1)} \frac{1}{q+1} = \sum_{p \leq q \leq n} \frac{1}{q+1}$$

and

$$\Delta I_n(p) = I_n(p) - I_n(p-1) = -\frac{1}{(p+1)}$$

In this notation and under our induction hypothesis, we have for all $p \leq q \leq n$

$$\frac{1}{q+1} s_n^{(k)}(q) = -\frac{1}{(k-1)!} I_n^{k-1}(q-1) \Delta I_n(q) + \frac{1}{(q+1)} r_n^{(k)}(q)$$

By definition of $s_n^{(k+1)}$, we find that

$$s_n^{(k+1)}(p) = -\frac{1}{(k-1)!} \sum_{p \leq q \leq n} I_n^{k-1}(q-1) \Delta I_n(q) + \sum_{p \leq q \leq n} \frac{1}{(q+1)} r_n^{(k)}(q)$$

On the other hand, by a direct application of the binomial formula, we find that

$$\begin{aligned} \Delta I_n^k(q) &:= I_n^k(q) - I_n^k(q-1) \\ &= [I_n(q-1) + \Delta I_n(q)]^k - I_n^k(q-1) \\ &= k I_n(q-1)^{k-1} \Delta I_n(q) + \sum_{2 \leq l \leq k} \binom{k}{l} [\Delta I_n(q)]^l I_n^{k-l}(q-1) \end{aligned}$$

from which we conclude that

$$-\frac{1}{(k-1)!} I_n^{k-1}(q-1) \Delta I_n(q) = -\frac{1}{k!} \Delta I_n^k(q) + \frac{1}{k!} \sum_{2 \leq l \leq k} \binom{k}{l} [\Delta I_n(q)]^l I_n^{k-l}(q-1)$$

Using the fact that $I_n(n) = 0$, we clearly have

$$\sum_{p \leq q \leq n} \Delta I_n^k(q) = I_n^k(n) - I_n^k(p-1) = -I_n^k(p-1)$$

This implies that

$$-\frac{1}{k!} \sum_{p \leq q \leq n} \Delta I_n^k(q) = \frac{1}{k!} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k$$

leading to the following decomposition

$$s_n^{(k+1)}(p) = \frac{1}{k!} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k + r_n^{(k+1)}(p)$$

with a remainder sequence or arrays $r_n^{(k+1)}$ given by

$$r_n^{(k+1)}(p) := \frac{1}{k!} \sum_{2 \leq l \leq k} \sum_{p \leq q \leq n} \binom{k}{l} [\Delta I_n(q)]^l I_n^{k-l}(q-1) + \sum_{p \leq q \leq n} \frac{1}{(q+1)} r_n^{(k)}(q)$$

It is easy to see that

$$\left| \frac{1}{k!} \sum_{2 \leq l \leq k} \binom{k}{l} [\Delta I_n(q)]^l I_n^{k-l}(q-1) \right| \leq c(k) |\Delta I_n(q)|^2 I_n^{k-2}(-1)$$

Consequently

$$\left| \frac{1}{k!} \sum_{2 \leq l \leq k} \sum_{p \leq q \leq n} \binom{k}{l} [\Delta I_n(q)]^l I_n^{k-l}(q-1) \right| \leq c(k) [\log(n+1)]^{k-2} \sum_{p \leq q \leq n} \frac{1}{(q+1)^2}$$

In addition, we also have

$$\begin{aligned} \sum_{p \leq q \leq n} \frac{1}{(q+1)} \left| r_n^{(k)}(q) \right| &\leq c(k) (\log(n+1))^{k-3} \sum_{p \leq q \leq n} \frac{1}{(q+1)} \sum_{q \leq r \leq n} \frac{1}{(r+1)^2} \\ &= c(k) (\log(n+1))^{k-3} \sum_{p \leq r \leq n} \frac{1}{(r+1)^2} \sum_{p \leq q \leq r} \frac{1}{(q+1)} \\ &\leq c(k) (\log(n+1))^{k-3} \left(\sum_{p \leq r \leq n} \frac{1}{(r+1)^2} \right) \left(\sum_{p \leq q \leq n} \frac{1}{(q+1)} \right) \end{aligned}$$

and therefore

$$\sum_{p \leq q \leq n} \frac{1}{(q+1)} \left| r_n^{(k)}(q) \right| \leq c(k) (\log(n+1))^{k-2} \sum_{p \leq q \leq n} \frac{1}{(q+1)^2}$$

Finally, we obtain that

$$|r_n^{(k+1)}(p)| \leq 2c(k) (\log(n+1))^{k-2} \sum_{p \leq q \leq n} \frac{1}{(q+1)^2}$$

which achieves the proof of the lemma. ■

Proof of lemma 6.3

We obviously have for any $1 \leq p \leq n$

$$\log\left(\frac{n+2}{p+1}\right) \leq \sum_{p \leq q \leq n} \frac{1}{(q+1)} \leq \log\left(\frac{n+1}{p}\right)$$

One can observe that the first inequality in the left hand side also holds true for $p = 0$. Moreover, using the fact that for any $p \geq 2$, we have

$$\left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k \leq \left(\log\left(\frac{n+1}{p}\right) \right)^k \leq \int_{p-1}^p \left(\log\left(\frac{n+1}{t}\right) \right)^k dt$$

we find that

$$\sum_{0 \leq p \leq n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k \leq c(k) (\log(n+1))^k + \int_1^{n+1} \left(\log\left(\frac{n+1}{t}\right) \right)^k dt$$

In much the same way, we find that

$$\sum_{0 \leq p \leq n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k \geq \int_1^{n+2} \left(\log\left(\frac{n+2}{t}\right) \right)^k dt$$

By a simple change of variable, we have

$$\int_1^{n+1} \left(\log\left(\frac{n+1}{t}\right) \right)^k dt = (n+1) \int_{\frac{1}{n+1}}^1 \left(\log\left(\frac{1}{t}\right) \right)^k dt$$

One easily check that a primitive of the power of a logarithm is given by the formula

$$\frac{1}{k!} \int \left(\log\left(\frac{1}{t}\right) \right)^k dt = t \sum_{0 \leq l \leq k} \frac{1}{l!} \left(\log\left(\frac{1}{t}\right) \right)^l$$

This yields that

$$\frac{1}{k!} \int_{\frac{1}{n+1}}^1 \left(\log\left(\frac{1}{t}\right) \right)^k dt = 1 - \frac{1}{n+1} \sum_{0 \leq l \leq k} \frac{1}{l!} (\log(n+1))^l$$

from which we conclude that

$$\frac{1}{k!} \int_1^{n+1} \left(\log\left(\frac{n+1}{t}\right) \right)^k dt = (n+1) - \sum_{0 \leq l \leq k} \frac{1}{l!} (\log(n+1))^l$$

and therefore

$$(n+2) - \sum_{0 \leq l \leq k} \frac{1}{l!} (\log(n+2))^l \leq \frac{1}{k!} \sum_{0 \leq p \leq n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k \leq c(k) (\log(n+1))^k + (n+1)$$

The end of the proof of the first assertion is now easily completed. To prove the second one, we observe that

$$\sum_{0 \leq p \leq \kappa_n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k \geq \int_1^{\kappa_n+2} \left(\log \left(\frac{(n+2)}{t} \right) \right)^k dt \quad (7.2)$$

and

$$\sum_{0 \leq p \leq \kappa_n} \left(\sum_{p \leq q \leq n} \frac{1}{q+1} \right)^k \leq c(k) (\log(n+1))^k + \int_1^{\kappa_n+1} \left(\log \left(\frac{(n+1)}{t} \right) \right)^k dt \quad (7.3)$$

Using the fact that

$$\frac{1}{k!} \int_1^{\kappa_n+1} \left(\log \left(\frac{(n+1)}{t} \right) \right)^k dt = (n+1) \int_{\frac{1}{n+1}}^{\frac{\kappa_n+1}{n+1}} \left(\log \left(\frac{1}{t} \right) \right)^k dt$$

we prove that

$$\frac{1}{k!} \int_{\frac{1}{n+1}}^{\frac{\kappa_n+1}{n+1}} \left(\log \left(\frac{1}{t} \right) \right)^k dt = \frac{\kappa_n+1}{n+1} \sum_{0 \leq l \leq k} \frac{1}{l!} \left(\log \left(\frac{n+1}{\kappa_n+1} \right) \right)^l - \frac{1}{n+1} \sum_{0 \leq l \leq k} \frac{1}{l!} (\log(n+1))^l$$

from which we conclude that

$$\lim_{n \rightarrow \infty} \frac{1}{k!} \int_{\frac{1}{n+1}}^{\frac{\kappa_n+1}{n+1}} \left(\log \left(\frac{1}{t} \right) \right)^k dt = \sum_{0 \leq l \leq k} \frac{1}{l!} \kappa \left(\log \left(\frac{1}{\kappa} \right) \right)^l$$

Thus, we have proved that

$$\lim_{n \rightarrow \infty} \frac{1}{k!} \frac{1}{n} \int_1^{\kappa_n+1} \left(\log \left(\frac{(n+1)}{t} \right) \right)^k dt = \sum_{0 \leq l \leq k} \frac{1}{l!} \kappa \left(\log \left(\frac{1}{\kappa} \right) \right)^l$$

The end of the proof is now a direct consequence of the estimates (7.2) and (7.3). ■

Proof of lemma 7.2

We prove the lemma by induction on the parameter n . The result is clearly satisfied for $n = 0$. Suppose the estimate (7.1) 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

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

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 twice the Cauchy-Schwarz inequality 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. ■

References

- [1] A. Brockwell, P. Del Moral, and A. Doucet. *Sequentially interacting Markov chain Monte Carlo*. Technical report, Department of Statistics, University of British Columbia (2007).
- [2] P. Del Moral, *Feynman-Kac formulae. Genealogical and interacting particle systems, with applications*, Springer Verlag New York, Series: Probability and its Applications (2004).
- [3] P. Del Moral, and A. Doucet. *Interacting Markov Chain Monte Carlo Methods For Solving Nonlinear Measure-Valued Equations*. INRIA Research Report, Jan. 2008.
- [4] P. Del Moral and L. Miclo. *Self Interacting Markov chains*. Stochastic Analysis and Applications, vol. 24, no. 3, pp. 615–660 (2006).
- [5] P. Del Moral and L. Miclo. *On convergence of chains with time empirical self-interactions*. Proc. Royal Soc. Lond. A., vol. 460, pp. 325–346 (2003).

- [6] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science. Springer New York, 2001.
- [7] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, A Series of Comprehensive Studies in Mathematics 288, Springer-Verlag (1987).
- [8] A. Jasra, C. Andrieu, A. Doucet and P. Del Moral. *Non-Linear Markov chain Monte Carlo via Self Interacting Approximations*. Preprint (2007).
- [9] A. Jasra, C. Andrieu, A. Doucet and P. Del Moral. *A Note on Convergence of the Equi-energy Sampler*. to appear in the Journal Stochastic Analysis and its Applications (2008).
- [10] A. Shiryaev, *Probability*, Second edition, Vol 95 in Graduate Text in Mathematics, Springer Verlag, New York (1996).