Electronic Journal of Probability

Vol. 14 (2009), Paper no. 73, pages 2130–2155.

Journal URL
http://www.math.washington.edu/~ejpecp/

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

Bernard Bercu*, Pierre Del Moral† and Arnaud Doucet‡

Abstract

We present a functional central limit theorem for a new class of interacting Markov chain Monte Carlo algorithms. These stochastic algorithms have been recently introduced to solve non-linear measure-valued equations. We provide an original theoretical analysis based on semigroup techniques on distribution spaces and fluctuation theorems for self-interacting random fields. Additionally we also present a series of sharp mean error bounds in terms of the semigroup associated with the first order expansion of the limiting measure-valued process. We illustrate our results in the context of Feynman-Kac semigroups.

Key words: Multivariate and functional central limit theorems, random fields, martingale limit theorems, self-interacting Markov chains, Markov chain Monte Carlo methods, Feynman-Kac semigroups.

AMS 2000 Subject Classification: Primary 60F05, 60J05, 60J20, 68U20, 80M31.

Submitted to EJP on February 20, 2008, final version accepted August 19, 2009.

*Centre INRIA Bordeaux et Sud-Ouest & Institut de Mathématiques de Bordeaux , Université Bordeaux, 351 cours de la Libération 33405 Talence cedex, France, Bernard.Bercu@math.u-bordeaux1.fr

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

‡Department of Statistics & Department of Computer Science, University of British Columbia, 333-6356 Agricultural Road, Vancouver, BC, V6T 1Z2, Canada and The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan, Arnaud@stat.ubc.ca

1 Introduction

1.1 Nonlinear Measure-Valued Equations and Self Interacting Markov chains

Let $(S^{(l)})_{l\geq 0}$ be a sequence of measurable spaces endowed with some σ-fields $(\mathscr{S}^{(l)})_{l\geq 0}$, and for every $l\geq 0$, denote by $\mathscr{P}(S^{(l)})$ the set of all probability measures on $S^{(l)}$. Let also $(\pi^{(l)})_{l\geq 0}$ be a sequence of probability measures on $S^{(l)}$ satisfying a nonlinear equation of the following form

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

for some mappings $\Phi_l : \mathscr{P}(S^{(l-1)}) \to \mathscr{P}(S^{(l)})$.

Being able to solve these equations numerically has numerous applications in nonlinear filtering, global optimization, Bayesian statistics and physics as it allow us to approximate any sequence of fixed "target" probability distributions $(\pi^{(l)})_{l\geq 0}$. For example, in a nonlinear filtering framework $\pi^{(l)}$ corresponds to the posterior distribution of the state of an unobserved dynamic model at time l given the observations collected from time 0 to time l. In an optimization framework, $\pi^{(l)}$ could correspond to a sequence of annealed versions of a distribution π that we are interested in maximizing. When Φ_l is a Feynman-Kac transformation [2], there has been much work on mean field interacting particle interpretations of measure-valued equations of the form (1.1); see for example [2; 6]. An alternative iterative approach to solve these equations named Interacting Markov chain Monte Carlo methods (i-MCMC) has recently been proposed in [1; 3].

Let us give a rough description of i-MCMC methods. At level $l = 0$, we run a standard MCMC chain say $X_n^{(0)}$ with a prescribed simple target distribution $\pi^{(0)}$. Then, we use the occupation measure of the chain $X^{(0)}$ to design a second MCMC chain say $X_n^{(1)}$ with a more 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 $(X^{(l)})_{l\geq 0}$ targeting a prescribed sequence of distributions $(\pi^{(l)})_{l\geq 0}$ of increasing complexity. These i-MCMC samplers are self interacting Markov chains as 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.

The theoretical analysis of these algorithms is more complex than the one of mean field interacting particle algorithms developed in [2]. In a pair of recent articles [1; 3], we initiated the theoretical analysis of i-MCMC algorithms and we provided a variety of convergence results including exponential estimates and an uniform convergence theorem with respect to the index l. However, we did not present any central limit theorem. The main purpose of this paper is to provide the fluctuation analysis of the occupation measures of a specific class of i-MCMC algorithms 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 Feynman-Kac semigroups.

The rest of the paper is organized as follows: the i-MCMC methodology is described in section 1.2. For the convenience of the reader, we have collected in the short preliminary section 1.3 some of the main notation and conventions used in the paper. We also recall some regularity properties of integral operators used in this 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 mean error bounds in terms of the semigroup associated with a first order expansion of the mappings Φ_l. To motivate our approach, we illustrate in section 3 these results in the context of Feynman-Kac semigroups. The last four sections are essentially concerned with the proof of the two 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 measures 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 section 6 and section 7 are devoted respectively to the proofs of the functional central limit theorem and the sharp non asymptotic mean error bounds theorem. In appendix, we have collected the proofs of a series of technical lemmas.

1.2 Description of the i-MCMC methodology

To introduce the i-MCMC methodology, we first recall how the traditional Metropolis-Hastings algorithm proceeds. This algorithm samples a Markov chain with the following transition kernel

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

where $P(x,dy)$ is a Markov transition on some measurable space S, and $G : S^2 \to \mathbb{R}_+$. 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

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

We further assume that $(\pi \times P)_2 \ll (\pi \times P)_1$ and the corresponding Radon-Nykodim ratio is given by

$$G = \frac{d(\pi \times P)_2}{d(\pi \times P)_1}.$$

In this case, π is an invariant measure of the time homogeneous Markov chain with transition M. One well-known difficulty with this algorithm is the choice of P.

Next, assume $\pi = \Phi(\eta)$ is related to an auxiliary probability measure η on some possibly different measurable space S' with some mapping $\Phi : \mathscr{P}(S') \to \mathscr{P}(S)$. We also suppose that there exists an MCMC algorithm on S' whose occupation measure, say η_n, approximates 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 to 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 (G_n, P_n) is far from being unique and we present a few solutions in [3]. A natural and simple one consists of selecting $P_n(x,dy) = \Phi(\eta_n)(dy)$ which ensures that G_n is equal to one. This is the class of i-MCMC algorithms studied in this paper. Iterating this strategy, we design a sequence of i-MCMC chains $X^{(k)} = (X_n^{(k)})_{n\geq 0}$ 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 algorithm, we need a few notation. We introduce a sequence of initial probability measures ν_k on $S^{(k)}$, with $k \geq 0$, and we denote by

$$\eta_n^{(k)} := \frac{1}{n+1} \sum_{0 \leq p \leq n} \delta_{X_p^{(k)}} \tag{1.3}$$

the sequence 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 algorithm proposed here is defined inductively on the level parameter k. First, 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 $(k+1)$-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

$$\mathbb{P}((X_0^{(k+1)}, \ldots, X_n^{(k+1)}) \in d(x_0, \ldots, x_n) | X^{(k)}) = \prod_{0 \leq p \leq n} \Phi_{k+1}\left(\eta_{p-1}^{(k)}\right)(dx_p)\,, \tag{1.4}$$

where we use the convention that we have

$$\Phi_{k+1}\left(\eta_{-1}^{(k)}\right) = \nu_{k+1}.$$

Moreover, $d(x_0, \ldots, x_n) = dx_0 \times \ldots \times dx_n$ stands for an infinitesimal neighborhood of a generic path sequence $(x_0, \ldots, x_n) \in (S^{(k+1)})^{(n+1)}$.

1.3 Notation and conventions

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

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

We denote by

$$\mu(f) = \int_E \mu(dx)\, f(x)$$

the Lebesgue integral of a function $f \in \mathscr{B}(E)$ with respect to $\mu \in \mathscr{M}(E)$. We slightly abuse the notation, and we denote by $\mu(A) = \mu(1_A)$ the measure of a measurable subset $A \in \mathscr{E}$.

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

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

are $\mathscr{E}$-measurable and bounded for any $f \in \mathscr{B}(F)$. A bounded integral operator M from a measurable space $(E,\mathscr{E})$ into an auxiliary measurable space $(F,\mathscr{F})$ also generates a dual operator $\mu \mapsto \mu M$ from $\mathscr{M}(E)$ into $\mathscr{M}(F)$ defined by $(\mu M)(f) = \mu(M(f))$. We denote by

$$\|M\| := \sup_{f \in \mathscr{B}_1(E)} \|M(f)\|$$

the norm of the operator $f \mapsto M(f)$ and we equip the Banach space $\mathscr{M}(E)$ with the corresponding total variation norm

$$\|\mu\| = \sup_{f \in \mathscr{B}_1(E)} |\mu(f)|.$$

We also denote by $\beta(M)$ the Dobrushin coefficient of a bounded integral operator M defined as

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

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 $\mathscr{M}_0(E)$ into $\mathscr{M}_0(F)$, and $\beta(M)$ coincides with the norm of this operator. We equip the sets of sequence of distributions $\mathscr{M}(E)^{\mathbb{N}}$ with the uniform total variation distance defined for all $\eta = (\eta_n)_{n\geq 0}, \mu = (\mu_n)_{n\geq 0} \in \mathscr{M}(E)^{\mathbb{N}}$ by

$$\|\eta - \mu\| := \sup_{n\geq 0} \|\eta_n - \mu_n\|.$$

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

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

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

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

$$M(f) = \left(M(f^1), \ldots, M(f^d)\right) \quad \text{and} \quad \mu(f) = \left(\mu(f^1), \ldots, \mu(f^d)\right).$$

Finally we denote by $c(k)$ with $k \in \mathbb{N}$, a generic constant whose values may change 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 $\{1,\ldots,m\}$ into $\{1,\ldots,n\}$. Finally, we shall use the usual conventions $\sum_\emptyset = 0$ and $\prod_\emptyset = 1$.

2 Fluctuations theorems

This section is mainly concerned with the statement of the two main theorems presented in this article. These results are based on a first order weak regularity condition on the mappings (Φ_l) appearing in (1.1). We assume that the mappings $\Phi_l : \mathscr{P}(S^{(l-1)}) \to \mathscr{P}(S^{(l)})$ satisfy the following first order decomposition for any $l \geq 1$

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

In the formula above, $\mathscr{D}_{l,\eta}$ is a collection of bounded integral operators from $S^{(l-1)}$ into $S^{(l)}$, indexed by the set of probability measures $\eta \in \mathscr{P}(S^{(l-1)})$ and $\Xi_l(\mu,\eta)$ is a collection of remainder signed measures on $S^{(l)}$ indexed by the set of probability measures $\mu,\eta \in \mathscr{P}(S^{(l-1)})$. We also require that

$$\sup_{\eta\in\mathscr{P}(S^{(l-1)})} \|\mathscr{D}_{l,\eta}\| < \infty \tag{2.2}$$

together with

$$\left|\Xi_l(\mu,\eta)(f)\right| \leq \int \left|(\mu-\eta)^{\otimes 2}(g)\right| \Xi_l(f,dg) \tag{2.3}$$

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

$$|g| = \sum_{i\in I} |a_i|\ \|h_i^1\|\|h_i^2\| < \infty \quad \text{and} \quad \chi_l = \sup_{f\in\mathscr{B}_1(S^{(l)})} \int |g|\ \Xi_l(f,dg) < \infty. \tag{2.4}$$

This weak regularity condition is satisfied in a variety of models including Feynman-Kac semigroups discussed in the next section. We also mention that, under weaker assumptions on the mappings Φ_l, we already proved in [1, Theorem 1] and [3, Theorem 1.1.] that for every $l \geq 0$ and any function $f \in \mathscr{B}(S^{(l)})$, we have the almost sure convergence result

$$\lim_{n\to\infty} \eta_n^{(l)}(f) = \pi^{(l)}(f) \tag{2.5}$$

The article [3, Theorem 1.1.] also provides exponential inequalities and uniform estimates with respect to the index l.

In order 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 $\mathscr{D}_{l,\pi^{(l-1)}}$ 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

$$D_l = \mathscr{D}_{l,\pi^{(l-1)}}$$

and for all $1 \leq k \leq l$,

$$D_{k,l} = D_k D_{k+1} \ldots D_l.$$

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

We now present the functional central limit theorem describing the fluctuations of the i-MCMC process around the solution of equation (1.1). Recall that $\eta_n^{(k)}$ is the empirical measure (1.3) of $X_0^{(k)},\ldots,X_n^{(k)}$ which are i.i.d. samples from $\pi^{(0)}$ when $k = 0$ and generated according to (1.4) otherwise.

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

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

where $\left(V^{(l)}\right)$ *is a collection of independent and centered Gaussian fields with a covariance function given for any* $(f,g) \in \mathscr{B}(S^{(l)})^2$ *by*

$$\mathbb{E}\left(V^{(l)}(f)V^{(l)}(g)\right) = \pi^{(l)}\left[\left(f - \pi^{(l)}(f)\right)\left(g - \pi^{(l)}(g)\right)\right]. \tag{2.6}$$

We now recall that if Z is a Gaussian $\mathcal{N}(0,1)$ random variable, then for any $m \geq 1$,

$$\mathbb{E}(|Z|^m) = 2^{m/2}\frac{1}{\sqrt{\pi}}\Gamma\left(\frac{m+1}{2}\right),$$

where Γ stands for the standard Gamma function. Consequently, in view of Theorem 2.1, the reader should be convinced that the estimates presented in the next theorem are sharp with respect to the parameters m and k.

Theorem 2.2. *For any* $k,n \geq 0$*, any* $f \in \mathrm{Osc}_1(S^{(k)})$*, and any integer* $m \geq 1$*, we have the non asymptotic mean error bounds*

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

with the collection of constants $a(m)$ *given by* $a(2m)^{2m} = (2m)_m$ *and*

$$a(2m+1)^{2m+1} = \frac{1}{\sqrt{m+1/2}}\,(2m+1)_{(m+1)}$$

and for some constants $b(m)$ *(respectively* $c(k)$*) whose values only depend on the parameter* m *(respectively* k*).*

3 Feynman-Kac semigroups

In this section, we shall illustrate the fluctuation results presented in this paper with the Feynman-Kac mappings (Φ_l) given for all $l \geq 0$ and all $(\mu, f) \in (\mathscr{P}(S^{(l)}) \times \mathscr{B}(S^{(l+1)}))$ by

$$\Phi_{l+1}(\mu)(f) := \frac{\mu(G_l L_{l+1}(f))}{\mu(G_l)}, \tag{3.1}$$

where G_l is a positive potential function on $S^{(l)}$ and L_{l+1} is a Markov transition from $S^{(l)}$ to $S^{(l+1)}$. In this scenario, the solution of the measure-valued equation (1.1) is given by

$$\pi^{(l)}(f) = \frac{\gamma^{(l)}(f)}{\gamma^{(l)}(1)}$$

with

$$\gamma^{(l)}(f) = \mathbb{E}\left(f(Y_l)\prod_{0\leq k<l} G_k(Y_k)\right)$$

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. These Feynman-Kac distributions are quite 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

$$Y_l = (Y'_0, \ldots, Y'_l) \in S^{(l)} = (E'_0 \times \ldots \times E'_l). \tag{3.2}$$

To get an intuitive understanding of i-MCMC algorithms in this context, we note that for the Feynman-Kac mappings (3.1) we have for all $f \in \mathscr{B}(S^{(k)})$

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

It follows that each random state $X_p^{(k)}$ is sampled according to two separate genetic type mechanisms. First, we select randomly one state $X_q^{(k-1)}$ at level $(k-1)$, with a probability proportional to its potential value $G_{k-1}(X_q^{(k-1)})$. Second, we evolve randomly from this state according to the exploration transition L_k. This i-MCMC algorithm model can be interpreted as a spatial branching and interacting process. In this interpretation, the k-th chain duplicates samples with large potential values, at the expense of samples with small potential values. The selected offspring evolve randomly from the state space $S^{(k-1)}$ to the state space $S^{(k)}$ at the next level. The same description for path space models (3.2) coincides with the evolution of genealogical tree based i-MCMC algorithms.

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 for all $f \in \mathscr{B}(S^{(l)})$ by

$$D_{l,\eta}(f) := \frac{G_{l-1}}{\eta(G_{l-1})}\, L_l\left(f - \Phi_l(\eta)(f)\right)$$

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

$$\Xi_l(\mu,\eta)(f) := -\frac{1}{\mu(G_{l-1})}\,(\mu-\eta)^{\otimes 2}\left(G_{l-1}\otimes D_{l,\eta}(f)\right).$$

We can observe that the regularity condition (2.3) is satisfied if, for all $l \geq 0$,

$$0 < \inf G_l \leq \sup G_l < \infty.$$

Indeed, it is easy to check that for any function $f \in \mathscr{B}_1(S^{(l)})$,

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

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

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

associated with the Feynman-Kac operator $Q_l(f) = G_{l-1}L_l(f)$. More precisely, for all $1 \le k \le l$, we have

$$D_{k,l}(f) = \frac{Q_{k,l}}{\pi^{(k-1)}(Q_{k,l}(1))}\left(f - \pi^{(l)}(f)\right).$$

The Dobrushin coefficient $\beta(D_{k,l})$ can also be computed in terms of these semigroups. We have

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

with the Markov integral operator $P_{k,l}$ given, for all $1 \le k \le l$, by

$$P_{k,l}(f)(x) = \frac{Q_{k,l}(f)(x)}{Q_{k,l}(1)(x)}.$$

Therefore, it follows that

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

which leads to

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

4 Fluctuations of the local interaction fields

4.1 Introduction

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

Definition 4.1. *Denote by $\mathscr{F}_l^{(n)}$ the sigma field generated by the collection of random variables $(X_p^{(k)})_{0 \le p \le n}$, with $0 \le k \le l$, and let $\mathscr{F}^{(n)} = (\mathscr{F}_l^{(n)})_{l \ge 0}$ be the corresponding filtration. Consider the sequence of centered random measures $\delta^{(l)} = \left(\delta_n^{(l)}\right) \in \mathscr{M}(S^{(l)})^{\mathbb{N}}$ defined for all $n \ge 0$ and all $l \ge 0$ by*

$$\delta_n^{(l)} := \left[\delta_{X_n^{(l)}} - \Phi_l\left(\eta_{n-1}^{(l-1)}\right)\right].$$

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 process 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

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

By definition of the i-MCMC algorithm, we have $\mathbb{E}(\Delta_n^{(l)}(f)) = 0$ for any $f \in \mathscr{B}(S^{(l)})$. In addition, we have

$$\mathbb{E}([\Delta_n^{(l)}(f)]^2) = \frac{1}{n+1} \sum_{0 \le p \le n} \Phi_l\left(\eta_{p-1}^{(l-1)}\right)\left(\left[f - \Phi_l\left(\eta_{p-1}^{(l-1)}\right)(f)\right]^2\right).$$

Recalling (2.5), as $\eta_n^{(l)}(f)$ converges almost surely to $\pi^{(l)}(f)$, as $n \to \infty$, for every $l \ge 0$ and any function $f \in \mathscr{B}(S^{(l)})$, we deduce that $\Phi_l(\eta_n^{(l-1)})(f)$ converges almost surely to $\pi^{(l)}(f)$, as $n \to \infty$, which implies, via Cesaro mean convergence that, as $n \to \infty$,

$$\mathbb{E}([\Delta_n^{(l)}(f)]^2) \to \pi^{(l)}\left(\left[f - \pi^{(l)}(f)\right]^2\right). \tag{4.2}$$

This shows that the random disturbances $\Delta_n^{(l)}(f)$ are unbiased with finite variance. It is possible to obtain a much stronger result by applying the traditional central limit theorem for triangular arrays; see for instance Theorem 4, page 543 of [11]. More precisely, we find that $\Delta_n^{(l)}(f)$ converges in law as $n \to \infty$ to a centered Gaussian random variable with variance given by (4.2). If we rewrite $\Delta_n^{(l)}$ in a different way, we have the following decomposition

$$\begin{aligned}\eta_n^{(l)} &= \frac{1}{n+1} \sum_{0 \le p \le n} \delta_{X_p^{(l)}} \\ &= \frac{1}{n+1} \sum_{0 \le p \le n} \Phi_l(\eta_{p-1}^{(l-1)}) + \frac{1}{\sqrt{n+1}}\, \Delta_n^{(l)}.\end{aligned} \tag{4.3}$$

Equation (4.3) shows that the evolution equation of the sequence of occupation measures $\eta_n^{(l)}$ can be interpreted as a stochastic perturbation of the following measure-valued equation

$$\mu_n^{(l)} = \frac{1}{n+1} \sum_{0 \le p \le n} \Phi_l(\mu_{p-1}^{(l-1)}) \tag{4.4}$$

initialized using conditions $\mu_n^{(0)} = \eta_n^{(0)}$. Note that the constant distribution flow $\mu_n^{(l)} = \pi^{(l)}$ associated with the solution of (1.1) also satisfies (4.4).

Although the above discussion gives some insight on the asymptotic normal behavior of the local errors accumulated by the i-MCMC algorithm, it does not give directly the precise fluctuations of the sequence of measures $(\eta_n^{(l)})_{n \ge 0}$ around their limiting values $\pi^{(l)}$. The analysis of these fluctuations is based on the way the semigroup associated with the evolution (4.4) propagates these local perturbation random fields. We can observe that the one step transformation of the equation (4.4) is an averaging of a nonlinear transformation Φ_l of the complete sequence of measures $\mu_p^{(l-1)}$ from the origin $p = 0$ up to the current time n. Iterating these transformations, we end up with a complex semigroup on the sequence 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 developments 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 algorithm. We present an admissible class of array type weight functions for which the corresponding weighted perturbation random fields behave as a collection of independent and centered Gaussian fields.

4.2 A martingale convergence theorem

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

Definition 4.2. *Denote by $\mathscr{W}$ the set of non negative weight array functions $(w_n(p))_{n\geq 0, 0\leq p\leq n}$ satisfying, for some $m\geq 1$,*

$$\sum_{n\geq 0} w_n^m(0) < \infty$$

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

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

for some scaling function ϖ such that

$$\lim_{\varepsilon\to 0+} \varpi(\varepsilon) = 0 \qquad \text{and} \qquad \lim_{\varepsilon\to 1-} \varpi(\varepsilon) = 1.$$

We observe that the standard sequence $w_n(p) = 1/\sqrt{n}$ belongs to $\mathscr{W}$, with the identity function $\varpi(\varepsilon) = \varepsilon$.

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

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

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

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

In addition, for any $l\geq 0$ and any $n\geq 0$, we also denote $V_n^{(l)} = W_n^{(l)}(\delta^{(l)})$.

Denote by $f = (f_l)_{l\geq 0}\in\prod_{l\geq 0}\mathscr{B}(S^{(l)})^d$ 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 $f^i = (f_l^i)_{l\geq 0}\in\prod_{l\geq 0}\mathscr{B}(S^{(l)})$. We consider the $\mathbb{R}^d$-valued and $\mathscr{F}^{(n)}$-martingale $\mathscr{M}^{(n)}(f) = (\mathscr{M}^{(n)}(f^i))_{1\leq i\leq d}$ defined for any $l\geq 0$ and any $1\leq i\leq d$, by

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

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

$$\mathscr{M}_t^{(n)}(f) := \mathscr{M}_{\lfloor t\rfloor}^{(n)}(f) + \sum_{p=0}^{\lfloor\{t\}(n+1)\rfloor-1} w_n^{(\lfloor t\rfloor+1)}(p)\ \delta^{(\lfloor t\rfloor+1)}\left(f_{(\lfloor t\rfloor+1)}\right). \tag{4.5}$$

We can observe that $(\mathscr{M}_t^{(n)}(f))$ is a càdlàg martingale with respect to the filtration associated with the σ-fields $\mathscr{F}_t^{(n)}$ 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 formula (4.5) converges in law, as $n \to \infty$, to an $\mathbb{R}^d$-valued 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 $1 \leq i, j \leq d$*

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

with

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

Before establishing the proof of this convergence result, we illustrate some direct consequences of this theorem.

On the one hand, it follows from this convergence that the discrete generation martingales $(\mathcal{M}_l^{(n)}(f))$, where $l \in \mathbb{N}$, converge in law as $n \to \infty$ to an $\mathbb{R}^d$-valued and discrete 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

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

On the other hand, we can fix a parameter l and introduce a collection of functions $g_k^j = (g_k^{j,i})_{1\leq i\leq d_j} \in \mathscr{B}(S^{(k)})^{d_j}$, with $0 \leq j \leq l$ and $0 \leq k \leq l$, where $d_j \geq 1$ is some fixed parameter. For every $0 \leq k \leq l$, we take $f = (f_k)_{0\leq k\leq l}$ with $f_k = (g_k^j)_{0\leq j\leq l} \in \mathscr{B}(S^{(k)})^d$ 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

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

We can deduce from the above martingale convergence theorem that the random fields $(V_n^{(j)})$ with $0 \leq j \leq l$, converge in law, as n tends to infinity and in the sense of finite dimensional distributions, to a sequence of independent and centered Gaussian fields $\left(V^{(j}\right)_{0\leq j\leq l}$ such that, for any $1 \leq i, i' \leq d_j$,

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

We can achieve this discussion by the following corollary of Theorem 4.4.

Corollary 4.5. *For every collection of weight functions $(w^{(l)}) \in \mathscr{W}^{\mathbb{N}}$, the sequence of random fields $(V_n^{(l)})_{l\geq 0}$ 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 $\left(V^{(l)}\right)_{l\geq 0}$ with covariance function given for any $l \geq 0$ and any $(f, g) \in \mathscr{B}(S^{(l)})^2$ by*

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

We now come back to the proof of Theorem 4.4.

Proof of Theorem 4.4: For every $1 \le i \le d$ and any $0 \le k \le l$ such that $f_k^i \in \mathscr{B}(S^{(k)})$, the martingale increments are given by

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

Our goal is to make use of the central limit theorem for triangular arrays of $\mathbb{R}^d$-valued martingales given in [8] page 437. We first rewrite the martingale $\mathscr{M}_l^{(n)}(f)$ in a suitable form

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

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

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

We further denote by $\mathscr{G}_i^{(n)}$ the σ-field generated by the random variables $X_p^{(k)}$ for any pair of parameters (k,p) such that $k(n+1)+p \le i$. By construction, for any sequence of functions $f = (f_l)_{l \ge 0}$ and $g = (g_l)_{l \ge 0} \in \prod_{l \ge 0} \mathscr{B}(S^{(l)})$ and for every $0 \le i \le l(n+1)+n$, with $i = k(n+1)+p$ for some $0 \le k \le l$, and $0 \le p \le n$, we find that

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

with the local covariance function

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

This implies that

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

Recalling (2.5), we have that $\eta_n^{(k)}(h)$ converges almost surely to $\pi^{(k)}(h)$, as $n \to \infty$, for every $k \ge 0$ and any function $h \in \mathscr{B}(S^{(k)})$. Under our regularity conditions on the sequence of mappings (Φ_k), we find that $\Phi_k(\eta_n^{(k-1)})(h)$ converges almost surely to $\pi^{(k)}(h)$, as $n \to \infty$, for any function $h \in \mathscr{B}(S^{(k)})$. It immediately implies the almost sure convergence

$$\lim_{n\to\infty} C_n^{(k)}(f,g) = C^{(k)}(f,g) = \pi^{(k)}\Big((f_k - \pi^{(k)}(f_k))\,(g_k - \pi^{(k)}(g_k))\Big).$$

Under our assumptions on the weight array functions $(w_n^{(k)}(p))$, we can deduce that almost surely

$$\lim_{n\to\infty} \sum_{p=0}^{n} w_n^{(k)}(p)^2\, C_p^{(k)}(f,g) = C^{(k)}(f,g).$$

Consequently, it leads to the almost sure convergence

$$\lim_{n\to\infty} \langle \mathscr{M}^{(n)}(f), \mathscr{M}^{(n)}(g)\rangle_l = \sum_{0\le k\le l} C^{(k)}(f,g) = \mathscr{C}_l(f,g).$$

In order to go one step further, we introduce the sequence of $\mathbb{R}^d$-valued and continuous time càdlàg random processes given, for any $f = (f_l)_{l\ge 0} \in \prod_{l\ge 0} \mathscr{B}(S^{(l)})^d$ and $t \in \mathbb{R}_+$, by

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

It is not hard to see that for any $l \in \mathbb{N}$, and any $0 \le k \le n$,

$$l + \frac{k}{n+1} \le t < l + \frac{k+1}{n+1} \Longrightarrow \lfloor t \rfloor = l$$

and $\lfloor t(n+1)\rfloor - \lfloor t\rfloor(n+1) = \lfloor \{t\}(n+1)\rfloor = k$ where $\{t\} = t - \lfloor t \rfloor$ stands for the fractional part of $t \in \mathbb{R}_+$. 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

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

This readily implies that

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

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

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

Under our assumptions on the weight array functions $(w_n^{(k)}(p))$, we find that

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

Hereafter, we have for every $0 \le i \le l(n+1)+n$

$$\|\mathscr{V}_i^{(n)}(f)\| \le \sup_{0\le k\le l} \|f_k\| \sup_{0\le k\le l} \sup_{0\le p\le n} w_n^{(k)}(p).$$

In addition, we also have

$$\lim_{n\to\infty} \sup_{0\le k\le l} \sup_{0\le p\le n} w_n^{(k)}(p) = 0.$$

Consequently, we conclude that the conditional Lindeberg condition is satisfied. Hence, the $\mathbb{R}^d$-valued martingale $(\mathscr{M}_t^{(n)}(f))$ converge in law, as $n\to\infty$, to a continuous martingale $\mathscr{M}_t(f)$ with predictable bracket given, for any air of indexes $1 \le j, j' \le d$, by

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

which completes the proof of Theorem 4.4. ∎

5 A multilevel expansion formula

We present in this section a multilevel decomposition of the sequence of occupation measures $(\eta_n^{(k)})$ around their limiting value $\pi^{(k)}$. In section 4.1, we have shown that the sequence $(\eta_n^{(k)})$ satisfies a stochastic perturbation equation (4.3) of the measure-valued equation (4.4). In this interpretation, the forthcoming developments provide a first order expansion of the semigroup associated to equation (4.3). This result will be used in various places in the rest of the paper. In section 6, we combine these first order developments with the local fluctuation analysis presented in section 4 to derive a "two line" proof of the functional central limit theorem 2.1. In section 7, we use these decompositions to prove the non asymptotic mean error bounds presented in Theorem 2.2. The forthcoming multilevel expansion is expressed in terms of the following time averaging operators.

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

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

defined for any sequence of measures $\eta = (\eta_n) \in \mathscr{M}(S^{(l)})^{\mathbb{N}}$, and any $n \ge 0$, by

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

We also denote by $\mathbb{S}^k = \mathbb{S} \circ \mathbb{S}^{k-1}$ the k-th iterate of the mapping $\mathbb{S}$.

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

Proposition 5.2. *For every $k \ge 0$, we have the following multilevel expansion*

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

with a sequence of signed random measures $\Xi^{(k)} = (\Xi_n^{(k)})$ such that, for any $m \geq 1$,

$$\sup_{f \in \mathscr{B}_1(S^{(k)})} \mathbb{E}\left(|\Xi_n^{(k)}(f)|^m\right)^{1/m} \leq b(m)\, c(k)\, \frac{(\log(n+1))^k}{(n+1)} \tag{5.3}$$

for some constants $b(m)$ (respectively $c(k)$) whose values only depend on the parameter m (respectively k).

Proof of Proposition 5.2: Recall that we use the convention $D_{1,0} = I_d$. We prove the mean error bounds by induction on the integer parameter k. For $k = 0$, we readily find that

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

Hence, (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. We shall make use of the following decomposition

$$\begin{aligned} \eta_n^{(k+1)} - \pi^{(k+1)} &= \mathbb{S}_n(\delta^{(k+1)}) + \frac{1}{n+1} \sum_{0 \leq p \leq n} \left[\Phi_{k+1}(\eta_{p-1}^{(k)}) - \Phi_{k+1}(\eta_p^{(k)})\right] \\ &\quad + \frac{1}{n+1} \sum_{0 \leq p \leq n} \left[\Phi_{k+1}(\eta_p^{(k)}) - \Phi_{k+1}(\pi^{(k)})\right]. \end{aligned} \tag{5.4}$$

It follows from our assumptions on the mappings (Φ_k) that for all $\mu, \eta \in \mathscr{P}(S^{(k)})$,

$$\|\Phi_{k+1}(\mu) - \Phi_{k+1}(\eta)\| \leq c(k)\, \|\mu - \eta\|$$

where

$$c(k) \leq \sup_{\gamma \in \mathscr{P}(S^{(k)})} \|D_{k+1,\gamma}\| + \chi_{k+1}.$$

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

$$\left\|\frac{1}{n+1} \sum_{0 \leq p \leq n} \left[\Phi_{k+1}(\eta_p^{(k)}) - \Phi_{k+1}(\eta_{p-1}^{(k)})\right]\right\| \leq c(k)\, \frac{\log(n+1)}{(n+1)}. \tag{5.5}$$

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

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

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

$$\mathbb{E}\left(\left|(\eta_p^{(k)} - \pi^{(k)})^{\otimes 2}(g)\right|^m\right)^{1/m}$$

$$\leq \sum_{i\in I} |a_i| \, \mathbb{E}\left(\left([\eta_p^{(k)} - \pi^{(k)}](h_i^1) \right)^{2m} \right)^{1/2m} \, \mathbb{E}\left(\left([\eta_p^{(k)} - \pi^{(k)}](h_i^2) \right)^{2m} \right)^{1/2m}.$$

Moreover, via the mean error bounds established in [3, Theorem 1.1.], it follows that for any $i \in I$ and $j = 1, 2$

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

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

$$\mathbb{E}\left(\left| (\eta_p^{(k)} - \pi^{(k)})^{\otimes 2}(g) \right|^m \right)^{1/m} \leq b(m) \, c(k) \, \left| g \right| / (p+1) \tag{5.6}$$

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

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

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

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

This yields the decomposition

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

In the last assertion, we have used the convention that $D_{k+2,k+1} = I_d$ stands for the identity operator. We also note that

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

The end of the proof of the result at rank $(k+1)$ is now a consequence of decomposition (5.4) in conjunction with the estimates (5.5) and (5.6). This completes the proof of Proposition 5.2. ■

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. We first need to express the time averaging semigroup $\mathbb{S}^k$ introduced in definition 5.1 in terms of the following weighted summations

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

with the weight array functions $s_n^{(k)} = (s_n^{(k)}(p))_{0 \le p \le n}$ defined by the initial condition $s_n^{(1)}(p) = 1$ and, for any $k \ge 1$ and, for all $n \ge 0$ and $0 \le p \le n$, by the recursive equation

$$s_n^{(k+1)}(p) = \sum_{p \le q \le n} \frac{1}{(q+1)}\ s_q^{(k)}(p).$$

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 $k \ge 0$, we have*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{0 \le p \le n} s_n^{(k+1)}(p)^2 = \frac{(2k)!}{(k!)^2}. \tag{6.1}$$

In addition, for any $k \ge 1$, the weight array functions $(w^{(k)})$ defined, for all $n \ge 0$ and $0 \le p \le n$, by

$$w_n^{(k)}(p) := \frac{s_n^{(k)}(p)}{\sqrt{\sum_{0 \le q \le n} s_n^{(k)}(q)^2}}$$

belongs to the set $\mathscr{W}$ introduced in Definition 4.2, with $(\varpi^{(k)})$ given, for all $\varepsilon \in [0,1]$, by

$$\varpi^{(k)}(\varepsilon) = \sum_{l=0}^{2(k-1)} \frac{(-1)^l}{l!}\ \varepsilon\ (\log(\varepsilon))^l.$$

We shall now proceed to the "two line" proof of Theorem 2.1.

Proof of Theorem 2.1: Fix a parameter $k \ge 0$ and for any $0 \le l \le k$, 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

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

where

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

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

$$(n+1)\ \mathbb{S}_n^{l+1}(\eta) = \sum_{0 \le p \le n} s_n^{(l+1)}(p)\ \eta_p = \Big(\sum_{0 \le q \le n} s_n^{(l+1)}(q)^2 \Big)^{1/2}\ W_n^{(k-l)}(\eta).$$

Using the multilevel expansion presented in Proposition 5.2, we obtain that

$$\sqrt{(n+1)}\left[\eta_n^{(k)} - \pi^{(k)}\right] = \sqrt{n+1}\Big(\sum_{0\le l\le k} \mathbb{S}_n^{l+1}(\delta^{(k-l)})\, D_{(k-l)+1,k} + \Xi_n^{(k)}\Big),$$

$$= \sum_{0\le l\le k}\Big(\frac{1}{n+1}\sum_{0\le q\le n} s_n^{(l+1)}(q)^2\Big)^{1/2}\, V_n^{(k-l)}\, D_{(k-l)+1,k} + \sqrt{n+1}\Big(\Xi_n^{(k)}\Big)$$

where we recall that, for any $k\geq 0$, $V_n^{(k)} = W_n^{(k)}(\delta^{(k)})$. Finally, Theorem 2.1 is a direct consequence of Corollary 4.5 together with the mean error bound (5.3) with $m=1$ and (6.1). ∎

The proof of Proposition 6.1 relies on two pivotal lemmas. Their proofs follow elementary techniques but require tedious calculations which are given in appendix.

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

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

where the remainder sequence $(r_n^{(k)}(p))$ is such that $r_n^{(0)}(p) = r_n^{(1)}(p) = 0$ and, for all $k\geq 2$ and $0\leq p\leq n$, we have

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

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

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

with $|\rho^{(k)}(n)| \leq c(k)\ \big(\log(n+1)\big)^k$. In addition, for any increasing sequence of integers (κ_n) such that $\kappa_n\leq n$ and $\kappa_n/n\to\kappa>0$, as $n\to\infty$, we have

$$\lim_{n\to\infty}\frac{1}{k!}\,\frac{1}{n}\sum_{0\le p\le \kappa_n}\left(\sum_{p\le q\le n}\frac{1}{q+1}\right)^k = \sum_{0\le l\le k}\frac{(-1)^l}{l!}\kappa\,(\log(\kappa))^l.$$

We are now in position to prove Proposition 6.1.

Proof of Proposition 6.1: First, it clearly follows from Lemma 6.2 that

$$\sup_{0\le p\le n} s_n^{(k)}(p) \le c(k)\,(\log(n+1))^{k-1}. \tag{6.2}$$

In addition, we claim that

$$\sum_{0\le p\le n} s_n^{(k)}(p)^2 = \frac{(2(k-1))!}{((k-1)!)^2}\,(n+1) + \widetilde{r}_n^{(k)} \tag{6.3}$$

with $|\widehat{r}_n^{(k)}| \leq c(k)\ \left(\log(n+1)\right)^{2(k-1)}$. We can actually observe that

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

with

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

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 a similar way, we find that

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

In summary, we have shown that $|\overline{r}_n^{(k)}| \leq c(k)\ \log(n+1)^{(2k-3)}$. Consequently, we obtain from (6.2) and (6.3) that

$$\lim_{n\to\infty} \sup_{0\leq p\leq n} w_n^{(k)}(p) = 0.$$

We achieve the proof of (6.3) by a direct application of Lemma 6.3. From the above discussion, we also find that

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

with $|\widehat{r}_n^{(k)}| \leq |\overline{r}_n^{(k)}|$. We deduce from Lemma 6.3 that for any increasing sequence of integers (κ_n) such that $\kappa_n \leq n$ and $\kappa_n/n \to \kappa > 0$,

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

leading to

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

which completes the proof of Proposition 6.1. ∎

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 $f \in \mathscr{B}(S^{(l)})$, and any integer $m \geq 1$,*

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

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: Via expansion (5.2), we find from Lemma 7.1 that

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

In addition, by the elementary inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ together with (6.3), we obtain that

$$\left(\frac{1}{(n+1)} \sum_{0\leq p\leq n} s_n^{(l+1)}(p)^2\right)^{1/2} \leq \frac{\sqrt{(2l)!}}{l!} + c(l)\, \frac{\left(\log(n+1)\right)^l}{\sqrt{n+1}}$$

which completes the proof of Theorem 2.2. ■

The proof of Lemma 7.1 relies on the classical Khintchine inequality for Rademacher sequences established by Haagerup [7].

Lemma 7.2. *Let (ε_n) be a symmetric Rademacher sequence which means that (ε_n) is a sequence of independent symmetric Bernoulli random variables. Then, for any real number $m \geq 1$, there exist two positive constant A_m and B_m such that, for any arbitrary sequence (c_n) of real numbers, we have*

$$A_m \left(\sum_{0\leq p\leq n} c_p\right)^{1/2} \leq \mathbb{E}\left(\left|\sum_{0\leq p\leq n} c_p \varepsilon_p\right|^m\right)^{1/m} \leq B_m \left(\sum_{0\leq p\leq n} c_p\right)^{1/2}. \tag{7.1}$$

Moreover, the best constants A_m and B_m may be explicitly calculated. For example, $B_m = 1$ if $1 \leq m \leq 2$ while if $m > 2$,

$$B_m = \sqrt{2}\left(\frac{\Gamma\left((m+1)/2\right)}{\sqrt{\pi}}\right)^{1/m}.$$

Proof of Lemma 7.1: We shall make use of a classical symmetrization argument. For all $l \geq 0$, we consider a independent copy $(Y_n^{(l)})_{n\geq 0}$ of the reference i-MCMC process $(X_n^{(l)})_{n\geq 0}$. We also assume that these two processes are defined on the same probability space, and we denote by $\mathscr{H}_l^{(n)}$ 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 $f_l \in \mathscr{B}(S^{(l)})$, and for any $0 \leq p \leq n$

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

This readily implies that

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

where $(\Delta_n^{(k)}(p, f_l))_{0\leq p\leq n}$ is a sequence of symmetric and conditionally independent random variables given by

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

We finish the proof by a direct application of Lemma 7.2. ∎

Appendix

Proof of Lemma 6.2

We prove Lemma 6.2 by induction on the integer parameter k. For $k = 1$, as $s_n^{(1)}(p) = 1$, the result is clearly valid with the null remainder $r_n^{(0)}(p) = 0$. This is also true for $k = 2$ with $r_n^{(2)}(p) = 0$. From now on, we assume $k \geq 3$ and we shall make use of the induction hypothesis. In what follows $c(k)$ stands for some finite constant whose values may change from line to line but they only depend on the parameter k. For any $0 \leq p \leq n$, we denote

$$I_q(p) := \sum_{q\leq r\leq p} \frac{1}{r+1}$$

and $\Delta I_q(p) = I_q(p) - I_q(p-1)$. By definition of $s_n^{(k)}$ and the induction hypothesis, we find that

$$\begin{aligned} s_n^{(k+1)}(p) &= \sum_{p\leq q\leq n} \frac{1}{(q+1)}\, s_q^{(k)}(p) \\ &= \frac{1}{(k-1)!} \sum_{p\leq q\leq n} I_p(q)^{(k-1)}\ \Delta I_p(q) + r_{1,n}^{(k)}(p) \end{aligned}$$

with the remainder sequence

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

Next, it follows that

$$\begin{aligned} (I_p(q))^{k-1} &= \left(I_p(q-1)+\Delta I_p(q)\right)^{k-1}, \\ &= I_p(q-1)^{k-1} + \sum_{0\leq l\leq k-2} \binom{k-1}{l} I_p(q-1)^l \,(\Delta I_p(q))^{k-1-l} \end{aligned}$$

where

$$s_n^{(k+1)}(p) = \frac{1}{(k-1)!} \sum_{p\leq q\leq n} \Delta I_p(q)\, I_p(q-1)^{k-1} + r_{1,n}^{(k)}(p) + r_{2,n}^{(k)}(p)$$

and

$$\begin{aligned} r_{2,n}^{(k)}(p) &= \frac{1}{(k-1)!} \sum_{p\leq q\leq n} \sum_{0\leq l\leq k-2} \binom{k-1}{l}\; I_p(q-1)^l \,(\Delta I_p(q))^{k-l}, \\ &\leq c(k)\, I_p(n)^{k-2} \sum_{p\leq q\leq n} (\Delta I_q(p))^2, \\ &\leq c(k)\, \left(\log(n+1)\right)^{k-2} \sum_{p\leq r\leq n} \frac{1}{(r+1)^2}. \end{aligned}$$

Furthermore, we observe that

$$I_p(q)^k - I_p(q-1)^k = k\, I_p(q-1)^{k-1}\, \Delta I_p(q) + \sum_{0\leq l\leq k-2} \binom{k}{l} I_p(q-1)^l (\Delta I_p(q))^{k-l}$$

which implies that

$$\begin{aligned} \frac{1}{(k-1)!}\, I_p(q-1)^{(k-1)}\, \Delta I_p(q) &= \frac{1}{k!}\, k\, I_p(q-1)^{(k-1)}\, \Delta I_p(q), \\ &= \frac{1}{k!}\, \left(I_p(q)^k - I_p(q-1)^k\right) - \frac{1}{k!} \sum_{0\leq l\leq k-2} \binom{k}{l} I_p(q-1)^l (\Delta I_p(q))^{k-l}. \end{aligned}$$

Consequently, as $I_p(p-1)=0$, it leads to

$$\begin{aligned} s_n^{(k+1)}(p) &= \frac{1}{k!} \sum_{p\leq q\leq n} \left(I_p(q)^k - I_p(q-1)^k\right) + r_{1,n}^{(k)}(p) + r_{2,n}^{(k)}(p) + r_{3,n}^{(k)}(p), \\ &= \frac{1}{k!}\, I_p(n)^k + r_{1,n}^{(k)}(p) + r_{2,n}^{(k)}(p) + r_{3,n}^{(k)}(p) \end{aligned}$$

with

$$r_{3,n}^{(k)}(p) = -\frac{1}{k!} \sum_{p \le q \le n} \sum_{0 \le l \le k-2} \begin{pmatrix} k \\ l \end{pmatrix} I_p(q-1)^l (\Delta I_p(q))^{k-l}.$$

As before, we observe that

$$\begin{aligned} \left| r_{3,n}^{(k)}(q) \right| &\le c(k)\, I_p(n)^{k-2} \sum_{p \le q \le n} (\Delta I_p(q))^2, \\ &\le c(k)\, (\log(n+1))^{k-2} \sum_{p \le q \le n} \frac{1}{(q+1)^2}. \end{aligned}$$

In summary, we have proved that

$$s_n^{(k+1)}(p) = \frac{1}{k!}\, I_p(n)^k + r_n^{(k)}(p)$$

with some remainder sequence satisfying

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

which completes the proof of Lemma 6.2. ■

Proof of Lemma 6.3

We have for any $1 \le p \le n$

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

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

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

we find that

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

In a similar way, we can establish that

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

By a simple change of variables, we obtain

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

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

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

It implies that

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

from which we conclude that

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

and therefore

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

and

$$\frac{1}{k!} \sum_{0\le p\le n} \left(\sum_{p\le q\le n} \frac{1}{q+1}\right)^k \le c(k)\,(\log(n+1))^k + (n+1)$$

which achieves the proof of the first assertion of Lemma 6.3. To prove the second one, we observe that

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

and

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

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 find that

$$\begin{aligned}\frac{1}{k!}\int_{\frac{1}{n+1}}^{\frac{\kappa_n+1}{n+1}} \left(\log\left(\frac{1}{t}\right)\right)^k dt &= \frac{\kappa_n+1}{n+1} \sum_{0\le l\le k} \frac{1}{l!} \left(\log\left(\frac{n+1}{\kappa_n+1}\right)\right)^l \\ &\quad - \frac{1}{n+1} \sum_{0\le l\le k} \frac{1}{l!} \left(\log(n+1)\right)^l.\end{aligned}$$

It leads to

$$\lim_{n\to\infty}\frac{1}{k!}\int_{\frac{1}{n+1}}^{\frac{\kappa_n+1}{n+1}}\left(\log\left(\frac{1}{t}\right)\right)^k dt=\sum_{0\le l\le k}\frac{(-1)^l}{l!}\,\kappa(\log(\kappa))^l.$$

Therefore, we conclude that

$$\lim_{n\to\infty}\frac{1}{k!}\frac{1}{n}\int_{1}^{\kappa_n+1}\left(\log\left(\frac{(n+1)}{t}\right)\right)^k dt=\sum_{0\le l\le k}\frac{(-1)^l}{l!}\,\kappa(\log(\kappa))^l.$$

Finally, the second assertion of Lemma 6.3 follows from (7.2) and (7.3). ■

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). MR2044973

[3] P. Del Moral, and A. Doucet. *Interacting Markov Chain Monte Carlo Methods For Solving Nonlinear Measure-Valued Equations*. INRIA Research Report, RR-6435, 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). MR2220075

[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). MR2052266

[6] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Pratice*. Statistics for Engineering and Information Science. Springer New York, 2001. MR1847783

[7] U. Haagerup. *The best constants in the Khintchine inequality.* Studia Math., vol. 70, pp. 231–283 (1982). MR0654838

[8] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, A Series of Comprehensive Studies in Mathematics 288, Springer-Verlag (1987). MR0959133

[9] A. Jasra, C. Andrieu, A. Doucet and P. Del Moral. *Non-Linear Markov chain Monte Carlo via Self Interacting Approximations.* Preprint (2007). MR2405652

[10] A. Jasra, C. Andrieu, A. Doucet and P. Del Moral. *A Note on Convergence of the Equi-energy Sampler.* Stochastic Analysis and its Applications, vol. 26, pp. 298-312 (2008). MR2399737

[11] A. Shiryaev, *Probability*, Second edition, Vol 95 in Graduate Text in Mathematics, Springer Verlag, New York (1996). MR1368405