Interacting Markov Chain Monte Carlo Methods For Solving Nonlinear Measure-Valued Equations

Pierre del Moral, Arnaud Doucet

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Interacting Markov Chain Monte Carlo Methods For Solving Nonlinear Measure-Valued Equations

Pierre Del Moral — Arnaud Doucet

N° 6435

January 2008

Thème NUM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6435--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
BORDEAUX - SUD OUEST

Interacting Markov Chain Monte Carlo Methods For Solving Nonlinear Measure-Valued Equations

Pierre Del Moral*, Arnaud Doucet†

Thème NUM — Systèmes numériques
Équipes-Projets CQFD

Rapport de recherche n° 6435 — January 2008 — 45 pages

Abstract: We present a new interacting Markov chain Monte Carlo methodology for solving numerically discrete-time measure-valued equations. The associated stochastic processes belong to the class of self-interacting Markov chains. In contrast to traditional Markov chains, their time evolution may depend on the occupation measure of the past values. This general methodology allows us to provide a natural way to sample from a sequence of target probability measures of increasing complexity. We develop an original theoretical analysis to analyze the behaviour of these algorithms as the time parameter tends to infinity. This analysis relies on measure-valued processes and semigroup techniques. We present a variety of convergence results including exponential estimates and a uniform convergence theorem with respect to the number of target distributions, yielding what seems to be the first results of this kind for this class of self-interacting models. We also illustrate these models in the context of Feynman-Kac distribution flows.

Key-words: Markov chain Monte Carlo methods, sequential Monte Carlo, self-interacting processes, time-inhomogeneous Markov chains, Metropolis-Hastings algorithm, Feynman-Kac formulae

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

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

Centre de recherche INRIA Bordeaux – Sud Ouest
Domaine Universitaire - 351, cours de la Libération 33405 Talence Cedex
Téléphone : +33 5 40 00 69 00

Méthodes de Monte Carlo par Chaînes de Markov en Interaction pour la Résolution de Processus à Valeurs Measures Non Linéaires

Résumé : Nous présentons de nouvelles méthodes de Monte Carlo par chaînes de Markov en interaction pour la résolution numérique de processus à valeurs mesures non linéaires et à temps discret. Ces algorithmes stochastiques appartiennent à la classe des modèles non linéaires de chaînes de Markov en auto-interaction. A la différence des chaînes de Markov traditionnelles, ces algorithmes explorent aléatoirement les espaces d'états en interaction avec leurs mesures d'occupation temporelle. Cette nouvelle méthodologie permet de simuler de façon naturelle un flot de mesures de probabilités cibles avec un degré de complexité croissant. Nous développons une analyse théorique originale du comportement en temps long de ces modèles. Cette analyse est fondée sur l'étude de processus à valeurs mesures et sur des techniques de semigroupes. Nous présentons une variété de résultats de convergence avec notamment des estimées exponentielles et un théorème de convergence uniforme par rapport au paramètre temporel. Ces résultats semblent être les premiers de ce type pour cette classe de processus en auto-interaction. Enfin, nous illustrons ces modèles dans le cadre des flots de mesures de Feynman-Kac.

Mots-clés : Méthodes de Monte Carlo par chaînes de Markov, méthodes de Monte Carlo séquentielles, processus en auto-interaction, chaînes de Markov non homogènes, Algorithmes de Metropolis-Hastings, formules de Feynman-Kac

1 Introduction

1.1 Nonlinear measure-valued processes

Let $(S^{(l)}, \mathcal{S}^{(l)})_{l\geq 0}$ be a sequence of measurable spaces. For every $l \geq 0$ we denote by $\mathcal{P}(S^{(l)})$ the set of probability measures on $S^{(l)}$. Suppose we have a sequence of probability measures $\pi^{(l)} \in \mathcal{P}(S^{(l)})$ where $\pi^{(0)}$ is known and we have for $l \geq 1$ the following nonlinear measure-valued equations

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

for some mappings $\Phi_l : \mathcal{P}(S^{(l-1)}) \to \mathcal{P}(S^{(l)})$. Except in some particular situations, these measure-valued equations do not admit a closed-form solution.

Being able to solve these equations numerically has numerous applications in nonlinear filtering, global optimization, Bayesian statistics and physics as it would 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. In both cases, Φ_l is a Feynman-Kac transformation [3].

In recent years, there has been considerable interest in the development of mean field interacting particle interpretations of measure-valued equations of the form (1) which we briefly review here.

1.2 Mean field particle interpretations

These mean field particle interpretations are sometimes referred to as sequential Monte Carlo methods, particle filters and population Monte Carlo methods when the mappings Φ_l are Feynman-Kac transformations [3]. The central idea is to construct a Markov chain $X^{(l)} = (X_p^{(l)})_{1\leq p\leq N}$ taking values in the product spaces $(S^{(l)})^N$ so that the empirical measures $\pi_N^{(l)} := \frac{1}{N}\sum_{p=1}^N \delta_{X_p^{(l)}}$ approximate the current distribution $\pi^{(l)}$ as $N \uparrow \infty$. In the simpler version, we construct inductively $X^{(l)} = (X_p^{(l)})_{1\leq p\leq N}$ by sampling N independent random variables with common law $\Phi_l(\pi_N^{(l-1)})$. The rationale behind this is that the resulting particle measure $\pi_N^{(l)}$ should be a good approximation of $\pi^{(l)}$ as long as $\pi_N^{(l-1)}$ is a good approximation of $\pi^{(l-1)}$. More formally, $X^{(l)}$ is an $(S^{(l)})^N$-valued Markov chain with elementary transitions given by the following formula

$$\mathbb{P}\left((X_1^{(l)}, \ldots, X_N^{(l)}) \in dx \mid X^{(l-1)}\right) = \prod_{p=1}^N \Phi_l\left(\frac{1}{N}\sum_{1\leq q\leq N} \delta_{X_q^{(l-1)}}\right)(dx_p) \tag{2}$$

In the above display, $dx = d(x_1, \ldots, x_N) = dx_1 \times \ldots \times dx_N$ stands for an infinitesimal neighborhood of a point in the product space $(x_1, \ldots, x_N) \in (S^{(l)})^N$.

For Feynman-Kac transformations, these interacting particle models have recently become an extensively studied subject in nonlinear filtering and in Bayesian statistics; for a review see the pair of books [3, 6]. In this framework, the convergence analysis of these stochastic particle algorithms is now well understood. A variety of theoretical results are available, including sharp propagations of chaos properties, fluctuation and large deviations theorems, as well as uniform convergence results with respect to the parameter l.

A severe limitation of these stochastic algorithms is that, it is impossible to iteratively increase the precision of the numerical approximation. An alternative class of stochastic simulations algorithms which allows us to iteratively increase the precision are MCMC methods; see [9] for a review. However, standard MCMC methods do not apply to this framework.

1.3 Self-interacting Markov chains

We propose here a new class of interacting Markov chain Monte Carlo methods (*i-MCMC*) to solve these nonlinear measure-valued equations numerically. These i-MCMC methods can be described as adaptive and dynamic simulation algorithms which take advantage of the information carried by the past history to increase the quality of the next series of samples. One critical aspect of i-MCMC as opposed to standard MCMC methods is that it provides a natural adaptation and reinforced learning strategy of the physical or engineering evolution equation at hand. 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. Moreover, in contrast to more traditional mean-field type particle models and related sequential Monte Carlo techniques, these stochastic algorithms can increase the precision and performance of the numerical approximations iteratively.

The origins of i-MCMC methods can be traced back to a pair of articles [4, 5] presented by the first author in collaboration with Laurent Miclo. These studies are concerned with biology-inspired self-interacting Markov chain models with applications to genetic type algorithms involving a competition between the natural reinforcement mechanisms and the potential attraction of a given exploration landscape [4, 5]. These lines of research have been extended to the MCMC methodology in the joint articles of the authors with Christophe Andrieu and Ajay Jasra [1], as well as in the more recent article of the authors with Anthony Brockwell [2]. Related ideas have also appeared in computational chemistry [7].

In the present article, we design a new general class of i-MCMC models. Roughly speaking, these stochastic models are defined as follows: Firstly, at level 0 we run an MCMC algorithm to obtain a chain $X_n^{(0)}$ with a prescribed target distribution $\pi^{(0)}$. Then, we use the occupation measure of the chain $X^{(0)}$ judiciously to design a second MCMC algorithm to generate $X_n^{(1)}$ at level 1 with a target measure $\pi^{(1)}$ which is typically more complex than $\pi^{(0)}$. These two mechanisms are combined online so that the pair process interacts with the occupation measure of the system at each iteration n. More formally, the elementary transition $X_n^{(1)} \rightsquigarrow X_{n+1}^{(1)}$ of the chain $X^{(1)}$ at 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 allows us to design a series of MCMC samplers $(X^{(l)})_{l\geq 0}$ with a prescribed sequence of target distributions $(\pi^{(l)})_{l\geq 0}$ of increasing complexity. These i-MCMC samplers are sometimes called self-interacting Markov chains (*SIMC*) in reference to the fact that the complete Markov chain $\overline{X}_n^m := (X_n^{(l)})_{0\leq l\leq m}$ associated with a fixed series of m levels evolves with elementary transitions $\overline{X}_n^m \rightsquigarrow \overline{X}_{n+1}^m$ that depend on the occupation measure of the whole system $\overline{X}_p^m$ from the origin $p=0$ up to the current time $p=n$. Note that here the time index n corresponds to the number of iterations of the i-MCMC algorithm.

From the pure mathematical point of view the convergence analysis of self-interacting chains is essentially based on the study of the stability properties of sophisticated Markov chains with elementary transitions depending in a nonlinear way on the occupation measure of the chains. Hence the theoretical analysis of these models is much more involved than the one of traditional Markov chains. It also differs from the more traditional mean field interacting interpretations models developed in [3]. Besides the introduction of a new methodology, our main contribution is a refined theoretical analysis based on measure-valued processes and semigroup methods to analyze their asymptotic behavior as the time parameter n tends to infinity. Additionally, we also present a variety of convergence results including exponential estimates and a uniform convergence theorem with respect to the level index l. We make no restriction on the state spaces and we illustrate these models in the context of abstract Feynman-Kac distribution flows. Finally, we present an avenue of new research problems.

The rest of the paper is organized as follows:

The main notation and convention used in this work are introduced in the brief preliminary section 1.4. The i-MCMC methodology is detailed in section 1.5. The main results of the article are presented in section 1.6. Several examples of i-MCMC models are provided in section 2, including Feynman-Kac path integral distributions and interacting Metropolis-Hasting algorithms. This section also provides a discussion on how to combine the mean field particle interpretation models with i-MCMC.

Section 3 is concerned with the asymptotic behavior of an abstract class of time inhomogeneous Markov chains. In section 3.2, we present a preliminary resolvent analysis to estimate the regularity properties of Poisson operator and invariant measure type mappings. In section 3.3, we apply these results to study the law of large numbers and the concentration properties of time inhomogeneous Markov chains.

In section 4 we discuss the regularity properties of a series of time averaged semigroups on distribution flow state spaces. The asymptotic analysis of i-MCMC models are discussed in section 5. The strong law of large numbers is presented in section 5.2. We also provide an $\mathbb{L}_r$-mean error bound for the occupation measures of the i-MCMC model at each level l. In section 5.3, we discuss the long time behavior of these stochastic models in terms of the exponential stability properties of a time averaged type semigroup associated with the flow of target measures. We prove a uniform convergence theorem with respect to the number of levels.

The asymptotic analysis of the occupation measures associated with the complete self-interacting model on a fixed series of levels is discussed in section 6. The $\mathbb{L}_r$-mean error bounds and the concentration analysis are presented respectively in section 6.1 and in section 6.2. The final section, section 7, is concerned with contraction properties of time averaged Feynman-Kac distribution flows.

1.4 Notation and conventions

For the convenience of the reader we have collected some of the main notation and conventions used in the article. We also recall some more or less well-known regularity properties of integral operators used further in the article.

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 $\mathrm{Osc}_1(E)$, the convex set of $\mathcal{E}$-measurable functions f with oscillations less than one; that is,

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

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

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

$$M(f)(x) = \int_F M(x, dy)\ f(y) \ \in \mathbb{R}$$

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 $\|\mu\| = \sup_{f\in\mathcal{B}_1(E)} |\mu(f)|$. In this slightly abusive notation, we have

$$\|M\| := \sup_{x\in E} \sup_{f\in\mathcal{B}_1(F)} |\delta_x M(f)| = \sup_{x\in E} \|\delta_x M\|$$

where δ_x stands for the Dirac measure at the point $x \in E$. We recall that the norm of any bounded integral operator M with null mass $M(1) = 0$ satisfies

$$\|M\| = \sup_{f\in\mathcal{B}_1(F)} \|M(f)\| = 2 \sup_{f\in\mathrm{Osc}_1(F)} \|M(f)\|$$

When M has a constant mass, that is $M(1)(x) = M(1)(y)$ for any $(x, y) \in E^2$, the operator $\mu \mapsto \mu M$ maps $\mathcal{M}_0(E)$ into $\mathcal{M}_0(F)$. In this situation, we let $\beta(M)$ be the Dobrushin coefficient of a bounded integral operator M defined by the following formula

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

By construction, we have $M(f)/\beta(M) \in \mathrm{Osc}_1(E)$ as soon as $\beta(M) \neq 0$, so that

$$\|\mu M\| = 2 \sup_{f\in\mathrm{Osc}_1(F)} |\mu M(f)| \leq \beta(M)\ 2 \sup_{f\in\mathrm{Osc}_1(E)} |\mu(f)| \Longrightarrow \|\mu M\| \leq \beta(M)\ \|\mu\|$$

Using the fact that $\|\delta_x - \delta_y\| = 2$ for $x \neq y$ and

$$\begin{aligned}\beta(M) &= \sup_{f\in \mathrm{Osc}_1(F)} \sup_{(x,y)\in E^2} |(\delta_x M - \delta_y M)(f)| = \sup_{(x,y)\in E^2} \frac{\|\delta_x M - \delta_y M\|}{\|\delta_x - \delta_y\|} \\ &\leq \sup_{\mu\in\mathcal{M}_0(E)} \frac{\|\mu M\|}{\|\mu\|}\end{aligned}$$

we prove that

$$\beta(M) = \sup_{\mu\in\mathcal{M}_0(E)} \frac{\|\mu M\|}{\|\mu\|} = \frac{1}{2} \sup_{(x,y)\in E^2} \|\delta_x M - \delta_y M\|$$

is also the norm of the operator

$$\mu \in \mathcal{M}_0(E) \mapsto \mu M \in \mathcal{M}_0(F)$$

That is, we have that

$$\beta(M) = \sup_{\mu\in\mathcal{M}_0(E)} (\|\mu M\|/\|\mu\|)$$

More generally, for every bounded integral operator K from an auxiliary measurable space $(E', \mathcal{E}')$ into an measurable space $(E, \mathcal{E})$, with null mass $K(1) = 0$, we have

$$\|KM\| = \sup_{x\in E'} \|(\delta_x K)M\| \leq \beta(M) \sup_{x\in E'} \|(\delta_x K)\| \Longrightarrow \|KM\| \leq \beta(M)\ \|K\|$$

Unless otherwise stated, we use the letter C to denote a universal constant whose value may vary from line to line. Finally, we shall use the conventions $\sum_\emptyset = 0$ and $\prod_\emptyset = 1$.

1.5 The i-MCMC methodology

We describe here the i-MCMC methodology to numerically solve (1). We consider a Markov transition $M^{(0)}$ from $S^{(0)}$ into itself and a collection of Markov transitions $M^{(l)}_\mu$ from $S^{(l)}$ into itself, indexed by the parameter $l \geq 0$ and the set of probability measures $\mu \in \mathcal{P}(S^{(l-1)})$. We further assume that the invariant measure of each operator $M^{(l)}_\mu$ is given by $\Phi_l(\mu)$; that is we have that

$$\forall l \geq 0 \quad \forall \mu \in \mathcal{P}(S^{(l)}) \qquad \Phi_l(\mu) = \Phi_l(\mu) M^{(l)}_\mu$$

For $l = 0$, we use the convention $\Phi_0(\pi^{(-1)}) = \pi^{(0)}$ and $M^{(0)}_\mu = M^{(0)}$. For every $l \leq m$, we denote by $\eta^{(l)} \in \mathcal{P}(S^{(l)})$ the image measure of a measure $\eta \in \mathcal{P}(\prod_{0\leq l\leq m} S^{(l)})$ on the l-th level set $S^{(l)}$. We also fix a series of probability measures ν_k on $S^{(k)}$, with $k \geq 0$.

We let $X^{(0)} := (X^{(0)}_n)_{n\geq 0}$ be a Markov chain on $S^{(0)}$ with initial distribution ν_0 and elementary Markov transitions $M^{(0)}$. For every $k \geq 1$, given a realization of the chain $X^{(k-1)} := (X^{(k-1)}_n)_{n\geq 0}$, the k-th level chain $X^{(k)}_n$ is a Markov chain with initial distribution ν_k and with random Markov transitions $M^{(k)}_{\eta^{(k-1)}_n}$

depending on the current occupation measures $\eta_n^{(k-1)}$ of the chain at level $(k-1)$; that is we have that

$$\mathbb{P}(X_{n+1}^{(k)} \in dx \mid X^{(k-1)},\ X_n^{(k)}) = M^{(k)}_{\eta_n^{(k-1)}}(X_n^k, dx) \tag{3}$$

with

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

The rationale behind this is that the k-th level chain $X_n^{(k)}$ behaves asymptotically as a Markov chain with time homogeneous elementary transition $M^{(k)}_{\pi^{(k-1)}}$ as long as $\eta_n^{(k-1)}$ is a good approximation of $\pi^{(k-1)}$.

In the special case where $M_\mu^{(k)}(x^k, .) = \Phi_k(\mu)$, the k-th level chain $(X_n^{(k)})_{n\geq 1}$ is a collection of independent random variables with distributions $(\Phi_k(\eta_{n-1}^{(k-1)}))_{n\geq 1}$; that is we have that

$$\mathbb{P}((X_1^{(k)}, \ldots, X_n^{(k)}) \in dx \mid X^{(k-1)}) = \prod_{p=1}^{n} \Phi_k\left(\frac{1}{p}\sum_{0\leq q<p} \delta_{X_q^{(k-1)}}\right)(dx_p) \tag{4}$$

In the above display $dx = d(x_1, \ldots, x_n) = dx_1 \times \ldots \times dx_n$ stands for for an infinitesimal neighborhood of a generic path sequence $(x_0, \ldots, x_n) \in (S^{(k)})^n$.

We end this section with a self-interacting Markov chain interpretation of the stochastic algorithm discussed above. We consider the product space

$$E_m := (S^{(0)} \times \ldots \times S^{(m)})$$

and we let $(K_\eta^{(m)})_{\eta\in\mathcal{P}(E_m)}$ be the collection of Markov transitions from E_m into itself given by

$$\forall x := (x^0, \ldots, x^m) \in E_m \qquad K_\eta^{(m)}(x, dy) = \prod_{0\leq l\leq m} M^{(l)}_{\eta^{(l-1)}}(x^l, dy^l) \tag{5}$$

In the above display $dy := dy^0 \times \ldots \times dy^m$ stands for for an infinitesimal neighborhood of a generic point $y := (y^0, \ldots, y^m) \in E_m$, and $\eta^{(l)} \in \mathcal{P}(S^{(l)})$ stands for the image measure of a measure $\eta \in \mathcal{P}(E_m)$ on the l-th level set $S^{(l)}$, with $m \geq l$. In this notation, it is readily checked that

$$\overline{X}_n^m := (X_n^{(0)}, \ldots, X_n^{(m)})$$

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

$$\mathbb{P}(\overline{X}_{n+1}^m \in dy \mid \overline{X}_0^m, \ldots, \overline{X}_n^m) = K^{(m)}_{\overline{\eta}_n^{[m]}}(\overline{X}_n^m, dy) \quad \text{with} \quad \overline{\eta}_n^{[m]} = \frac{1}{n+1}\sum_{p=0}^{n} \delta_{\overline{X}_n^m} \tag{6}$$

1.6 Statement of some results

We further assume that for any $l \geq 1$, the mappings $\Phi_l : \mathcal{P}(S^{(l-1)}) \to \mathcal{P}(S^{(l)})$ satisfy the following regularity condition for any pair of measures $(\mu, \nu) \in \mathcal{P}(S^{(l-1)})^2$

$$\forall l \geq 0 \quad \forall f \in \mathcal{B}(S^{(l)}) \qquad |[\Phi_l(\mu) - \Phi_l(\nu)]\,(f)| \leq \int\ |[\mu - \nu]\,(g)|\ \Gamma_l(f, dg) \tag{7}$$

for some bounded integral operator Γ_l from $\mathcal{B}(S^{(l)})$ into $\mathcal{B}(S^{(l-1)})$, with

$$\int_{\mathcal{B}(S^{(l-1)})} \Gamma_l(f, dg) \ \|g\| \leq \Lambda_l \ \|f\| \quad \text{and} \quad \Lambda_l < \infty$$

We also suppose that there exist some integer $n_l \geq 0$ and some constant c_l such that we have

$$\|M_\mu^{(l)} - M_\nu^{(l)}\| \ \leq \ c_l \ \|\mu - \nu\| \quad \text{and} \quad b_l(n_l) := \sup_{\mu \in \mathcal{P}(S^{(l-1)})} \beta((M_\mu^{(l)})^{n_l}) < 1 \quad (8)$$

In the special case where $M_\mu^{(l)}(x^l, .) = \Phi_l(\mu)$, the second condition is trivially met for $n_l = 1$ with $b_l(n_l) = 0$. The first condition is related to the Lipschitz property of the mapping $\Phi_l(\mu)$. In this particular situation, it takes the following form

$$\|\Phi_l(\mu) - \Phi_l(\nu)\| \ \leq \ c_l \ \|\mu - \nu\|$$

Under the conditions (8), for every $\eta \in \mathcal{P}(E_m)$, the invariant measure $\omega_{K_\eta^{(m)}}(\eta) \in \mathcal{P}(E_m)$ of $K_\eta^{(m)}$ defined in (5) is given the tensor product measure

$$\omega_{K_\eta^{(m)}}(\eta) = \pi^{(0)} \otimes \Phi_1(\eta^{(0)}) \otimes \ldots \otimes \Phi_m(\eta^{(m-1)}) \quad (9)$$

We observe that the tensor product measure

$$\overline{\pi}^{[m]} := \pi^{(0)} \otimes \ldots \otimes \pi^{(m)} \quad (10)$$

is a fixed point of the mapping $\omega_{K_\eta^{(m)}} : \eta \in \mathcal{P}(E_m) \to \omega_{K_\eta^{(m)}}(\eta) \in \mathcal{P}(E_m)$.

In this notation, our main results are basically stated as follows.

Theorem 1.1 *For any $r \geq 1$, $m \geq 1$, and any function $f \in \mathcal{B}(E_m)$ we have*

$$\sup_{n \geq 1} \sqrt{n} \ \mathbb{E}\left(\left|\overline{\eta}_n^{[m]}(f) - \overline{\pi}^{[m]}(f)\right|^r\right) < \infty$$

Under some addition regularity conditions, we have the exponential inequality

$$\forall t > 0 \qquad \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left(\left|\left[\overline{\eta}_n^{[m]} - \overline{\pi}^{[m]}\right](f)\right| > t\right) < -\frac{t^2}{2\overline{\sigma}_m^2}$$

for some finite constant $\overline{\sigma}_m < \infty$ as well as the following uniform convergence estimate

$$\sup_{k \geq 0} \sup_{n \geq 1} n^{\alpha/2} \ \mathbb{E}\left(\left|\eta_n^{(k)}(f_k) - \pi^{(k)}(f_k)\right|^r\right) < \infty$$

for some parameter $\alpha \in (0, 1]$, and for any collection of functions $(f_k)_{k \geq 0} \in \prod_{k \geq 0} \mathcal{B}_1(S^{(k)})$.

We end this introduction with a series of comments and open research questions.

Firstly, the mean error bounds and the exponential estimates presented above suggest the existence of Gaussian fluctuations of the occupation measures $\overline{\eta}_n^{[m]}$ around their limiting value $\overline{\pi}^{[m]}$, with a fluctuation rate $\sqrt{n}$. We will study these fluctuations and the associated large deviations principles in a forthcoming article.

Surprisingly enough, the decays to equilibrium presented in theorem 1.1 differ from the three types of decays exhibited in the pair of articles [4, 5] for evolutionary type self-interacting Markov chains. To understand the main differences between these classes of interacting processes, we recall that the decay rates to equilibrium often depends on the contraction coefficient of the invariant measure mapping associated with a given self-interacting model. In our context, these mappings are not necessarily contractive. Nevertheless, we shall see in section 6 that the semigroup associated with these mappings becomes constant after a sufficiently large number of iterations. In this sense, the self-interacting models discussed in the present article are more regular than the one analyzed in [4, 5].

Last but not least, the uniform convergence estimate with respect to the number of levels depends on the stability properties of a time average semigroup associated with the mappings Φ_l. The contraction properties of this new class of nonlinear semigroups on the flow of measures are studied in section 7 in the context of Feynman-Kac models. Roughly speaking, we show that the stability properties of the reference Feynman-Kac semigroups can be transferred to study the associated time averaged models. In more general situations this question remains open.

2 Motivating applications

2.1 Feynman-Kac models

The main example of mappings Φ_l considered here are the Feynman-Kac transformations given below

$$\forall l \geq 0 \quad \forall (\mu, f) \in (\mathcal{P}(S^{(l)}) \times \mathcal{B}(S^{(l+1)})) \qquad \Phi_{l+1}(\mu)(f) := \mu(G_l L_{l+1}(f))/\mu(G_l) \tag{11}$$

where 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) 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 the spectral analysis of Schrödinger type operators and directed polymer analysis [3]. Even if they look innocent, these Feynman-Kac distribution flows are 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; that is, we have that

$$Y_l := (Y'_0, \ldots, Y'_l) \in S^{(l)} := (E'_0 \times \ldots \times E'_l) \tag{12}$$

We also underline that the unnormalized measures $\gamma^{(l)}$ are expressed in terms of integrals on path spaces and we recall that $\gamma^{(l)}$ can be expressed in terms of the flow of measures $(\pi^{(k)})_{0\leq k<l}$ with the following formulae

$$\gamma^{(l)}(f) = \pi^{(l)}(f) \prod_{0\leq k<l} \pi^{(k)}(G_k)$$

Thus, the i-MCMC methodology allows us to estimate these unnormalized flows and thus the normalizing constants $\gamma^{(l)}(1)$ by replacing in the above displayed formula the measures $\pi^{(k)}$ by their approximations.

In this context, the right hand side product of the formula (4) associated to the i-MCMC algorithm takes the following form

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

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 randomly select 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 randomly evolve 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 tends to duplicate individuals with large potential values, at the expense of individuals with low potential values. The selected offspring randomly evolve from the state space $S^{(k-1)}$ to the state space $S^{(k)}$ at the next level. The same description for path space models (12) coincides with the evolution of genealogical tree based i-MCMC models.

For the Feynman-Kac transformations (11), we proved in [3] that the condition (8) ensuring convergence of the algorithm is satisfied with $c_l = \beta(L_l)/\epsilon_{l-1}(G)$ as soon as the potential functions satisfy the following condition

(G) *There exists a sequence* $(\epsilon_l(G))_{l\geq 0} \in (0,1)^{\mathbb{N}}$ *such that*

$$\forall l \geq 0 \quad \forall (x,y) \in (S^{(l)})^2 \qquad G_l(x) \geq \epsilon_l(G)\ G_l(y)$$

2.2 Interacting Metropolis-Hastings models

Suppose the reference state spaces $S^{(l)}$ are given for any $l \geq 1$ by

$$S^{(l)} = S^{(l-1)} \times E'_l$$

where E'_l is an auxiliary sequence of measurable spaces. For $l = 0$, we set $S^{(0)} = E'_0$. Returning to the Feynman-Kac model presented in (11), we can choose

$$\begin{aligned}\pi^{(l)}(d(x_0,\ldots,x_{l-1},x_l)) &\propto \left[\pi^{(0)}(dx_0)\prod_{k=1}^{l} L_k(x_{k-1},dx_k)\right] \times \left[\prod_{0\leq k<l} G_k(x_k)\right]\\ &\propto \pi^{(l-1)}(d(x_0,\ldots,x_{l-1}))\ L_l(x_{l-1},dx_l)\ G_{l-1}(x_{l-1}).\end{aligned}$$

We introduce the following integral operator from $S^{(l-1)}$ into E'_l

$$P_l((x_1,\ldots,x_{l-1}),dx_l) = L_l(x_{l-1},dx_l)\ G_{l-1}(x_{l-1}). \tag{13}$$

In this scenario, it is sensible to propose to use for $M_\mu^{(l)}$ in the i-MCMC algorithm the following Markov kernel on the product space $S^{(l)}$ indexed by the set of measures $\mu \in \mathcal{P}(S^{(l-1)})$

$$M_\mu^{(l)}(x,dy)$$

$$= (\mu\otimes Q_l)(dy)\ (1\wedge r_l(x,y)) + \left(1-\int_{S^{(l)}}(1\wedge r_l(x,z))\ (\mu\otimes Q_l)(dz)\right)\ \delta_x(dy)$$

where Q_l is a Markov transition from $S^{(l-1)}$ into E'_l and for every (u,v) and $(w,z)\in(S^{(l-1)}\times E'_l)$

$$r_l((u,v),(w,z)) := \frac{d\left(Q_l(u,.)\otimes P_l(w,.)\right)}{d\left(P_l(u,.)\otimes Q_l(w,.)\right)}(v,z) \tag{14}$$

where we assume that

$$Q_l(u,.)\otimes P_l(w,.) << P_l(u,.)\otimes Q_l(w,.).$$

It can be checked that the kernel $M_\mu^{(l)}$ is nothing but a Metropolis-Hastings kernel of proposal distribution $\mu\otimes Q_l$ and invariant distribution

$$\Phi_l(\mu)(dx) = \frac{1}{\mu P_l(1)}\ \left(\mu\otimes P_l\right)(dx)$$

and that we indeed have $\Phi_l(\pi^{(l-1)})=\pi^{(l)}$.

We can also easily establish that for any pair of measures $(\mu,\nu)\in\mathcal{P}(S^{(l-1)})^2$

$$\|M_\mu^{(l)}-M_\nu^{(l)}\|\ \leq\ 2\ \|\mu-\nu\|$$

so that the first condition on the left hand side of (8) is satisfied. Under the additional assumption that for any $(u,v)\in(S^{(l-1)}\times E'_l)$

$$\frac{dP_l(u,.)}{dQ_l(u,.)}(v)\leq C_l$$

it follows from [8, Theorem 2.1] that

$$\beta(M_\mu^{(l)})\leq(1-C_l^{-1})$$

from which we conclude that the second condition on the right hand side of (8) is met with $n_l=1$ and $b_l(n_l)=(1-C_l^{-1})$.

2.3 Mean field particle and i-MCMC models

As mentioned in the introduction, in contrast to mean field particle models presented in (1.2), we emphasize that the precision parameter n of i-MCMC models is not fixed but it increases at every time step. There exist several ways to combine a mean field particle model with an i-MCMC model.

For instance, suppose we are given a realization of a mean field algorithm $X^{(l)} = (X_p^{(l)})_{1\leq p\leq N}$ with a precision parameter N. One natural way to initialize the i-MCMC model is to start with a collection of initial random states $X_0^{(l)}$ sampled according to the N-particle approximation measures

$$\nu_l = \pi_N^{(l)} := \frac{1}{N}\sum_{i=1}^{N}\delta_{X_i^{(l)}}$$

Another strategy is to use the N-particle approximation measures $\pi_N^{(l)}$ in the evolution of the i-MCMC model. In other words we interpret the series of samples $X_i^{(l)}$, $1 \leq i \leq N$, as the first N iterations of the i-MCMC model at level l. More formally, this strategy simply substitutes the current occupation measure $\eta_n^{(k-1)}$ of the chain at level $(k-1)$ in (3) by the occupation measure $\eta_n^{(N,k-1)}$ of the whole sequence of random variables at level $(k-1)$ defined by

$$\eta_n^{(N,k-1)} = \frac{n+1}{N+n+1}\;\eta_n^{(k-1)} + \frac{N}{N+n+1}\;\pi_N^{(k-1)}$$

The convergence analysis of these two natural combinations of a mean field particle model with an i-MCMC algorithm can be conducted easily using the techniques developed in this article.

3 Time inhomogeneous Markov chains

3.1 Description of the models

We consider a collection of Markov transitions K_η on some measurable space $(E,\mathcal{E})$ indexed by the set of probability measures $\eta \in \mathcal{P}(F)$ on some possibly different measurable space $(F,\mathcal{F})$. We further assume that for any pair of measures $(\eta,\mu) \in \mathcal{P}(F)^2$ and some integer $n_0 \geq 0$ we have

$$\|K_\eta - K_\mu\| \;\leq\; c\;\|\eta-\mu\| \quad \text{and} \quad b(n_0) := \sup_{\eta\in\mathcal{P}(E)} \beta(K_\eta^{n_0}) < 1 \tag{15}$$

We associate with the collection of transitions K_η an E-valued inhomogeneous Markov chain X_n with elementary transitions defined by

$$\mathbb{P}(X_{n+1} \in dx \mid X_0,\ldots,X_n) = K_{\mu_n}(X_n,dx)$$

where μ_n is a sequence of possibly random distributions on F that only depends on the random sequence $(X_0,\ldots,X_n)$. More precisely, μ_n is a measurable random variable with respect to the σ-field generated by the random states X_p from the origin $p=0$, up to the current time horizon $p=n$. We further assume that the variations of the flow μ_n are controlled by some sequence of random variables $\epsilon(n)$ in the sense that

$$\forall n \geq 0 \qquad \|\mu_{n+1}-\mu_n\| \leq \epsilon(n)$$

We let $\overline{\epsilon}(n)$ be the mean variation of the distribution flow $(\mu_p)_{0\leq p\leq n}$; that is we have

$$\overline{\epsilon}(n) := \frac{1}{(n+1)}\sum_{p=0}^{n}\epsilon(p)$$

For self-interacting Markov chains, we have $F = E$ and the measures μ_p coincide with the occupation measures of the chain up to the current time. In this particular situation, we have

$$\mu_n = \eta_n := \frac{1}{(n+1)} \sum_{p=0}^{n} \delta_{X_p} \implies \epsilon(n) \leq \frac{2}{(n+2)} \tag{16}$$

This implies that

$$\overline{\epsilon}(n) \leq \frac{2}{(n+1)} \log(n+2)$$

Under assumption (15), every elementary transition $K_{\mu_n}(x, dy)$ admits an invariant measure

$$\omega(\mu_n) K_{\mu_n} = \omega(\mu_n) \in \mathcal{P}(E)$$

For sufficiently small variations $\epsilon(n)$ of the distribution flow μ_n, we expect that the occupation measures η_n have the same asymptotic behavior as the mean values $\overline{\omega}_n(\mu)$ of the instantaneous invariant measures $\omega(\mu_p)$ from the origin $p = 0$ up to the current time n. That is, in some sense, for large values of the time horizon we have

$$\eta_n \simeq \overline{\omega}_n(\mu) := \frac{1}{n+1} \sum_{p=0}^{n} \omega(\mu_p) \tag{17}$$

3.2 A resolvent analysis

We recall that assumption (15) ensures that K_η has a unique invariant measure for any $\eta \in \mathcal{P}(F)$

$$\omega(\eta) K_\eta = \omega(\eta) \in \mathcal{P}(E)$$

and the pair of sums given by

$$\alpha(\eta) := \sum_{n \geq 0} \beta(K_\eta^n) \in [1, \infty) \quad \text{and} \quad \sum_{n \geq 0} \left[K_\eta^n - \omega(\eta)\right](f) \tag{18}$$

are absolutely convergent for any $f \in \mathcal{B}(E)$. The main simplification of these conditions comes from the fact that the resolvent operator

$$P_\eta \; : \; f \in \mathcal{B}(E) \to P_\eta(f) := \sum_{n \geq 0} \left[K_\eta^n - \omega(\eta)\right](f) \in \mathcal{B}(E)$$

is a well-defined solution of the Poisson equation

$$\begin{cases} (K_\eta - Id) P_\eta &= (\omega(\eta) - Id) \\ \omega(\eta) P_\eta &= 0 \end{cases}$$

Proposition 3.1 *For any $\eta \in \mathcal{P}(F)$, P_η is a bounded integral operator on E and we have*

$$(\|P_\eta\|/2) \vee \beta(P_\eta) \leq \alpha(\eta) \leq \frac{n_0}{1 - \beta(K_\eta^{n_0})}$$

Proof:
The fact that $\beta\left(P_\eta\right) \leq \alpha(\eta)$ is readily deduced from the following decomposition

$$P_\eta(f)(x) - P_\eta(f)(y) := \sum_{n\geq 0}\left[K_\eta^n(f)(x) - K_\eta^n(f)(y)\right]$$

Indeed, using this decomposition we find that $\mathrm{osc}(P_\eta(f)) \leq \sum_{n\geq 0}\mathrm{osc}(K_\eta^n(f))$. Recalling that $\mathrm{osc}(K_\eta^n(f)) \leq \beta(K_\eta^n)\ \mathrm{osc}(f)$, we conclude that

$$\mathrm{osc}(P_\eta(f)) \leq \left[\sum_{n\geq 0}\beta(K_\eta^n)\right]\ \mathrm{osc}(f) \Rightarrow \beta(P_\eta) \leq \sum_{n\geq 0}\beta(K_\eta^n)$$

In much the same way, we use the fact that

$$P_\eta(f)(x) = \sum_{n\geq 0}\int\ \left[K_\eta^n(f)(x) - K_\eta^n(f)(y)\right]\omega(\eta)(dy)$$

to check that

$$\|P_\eta(f)\| \leq \sum_{n\geq 0}\mathrm{osc}(K_\eta^n(f))$$

and

$$\|P_\eta(f)\| \leq \left[\sum_{n\geq 0}\beta(K_\eta^n)\right]\ \mathrm{osc}(f) \Rightarrow \|P_\eta\| \leq 2\ \sum_{n\geq 0}\beta(K_\eta^n)$$

To prove that $\alpha(\eta) \leq \frac{n_0}{1-\beta(K_\eta^{n_0})}$, we use the decomposition

$$\alpha(\eta) := \sum_{n\geq 0}\beta(K_\eta^n) = \sum_{p\geq 1}\ \sum_{n=(p-1)n_0}^{pn_0-1}\beta(K_\eta^n) = \sum_{p\geq 1}\ \sum_{r=0}^{n_0-1}\beta(K_\eta^{(p-1)n_0+r})$$

Since we have

$$\beta(K_\eta^{(p-1)n_0+r}) \leq \beta(K_\eta^{(p-1)n_0})\ \beta(K_\eta^r) \leq \beta(K_\eta^{n_0})^{(p-1)}\ \beta(K_\eta^r) \leq \beta(K_\eta^{n_0})^{(p-1)}$$

we conclude that $\alpha(\eta)\ \leq\ n_0\ \sum_{p\geq 0}\ \beta(K_\eta^{n_0})^p = \frac{n_0}{1-\beta(K_\eta^{n_0})}$. The end of the proof of the proposition is now completed. ■

Proposition 3.2 *For any pair of measures $(\eta,\mu) \in \mathcal{P}(F)^2$, we have*

$$\|\omega(\eta) - \omega(\mu)\| \leq\ \delta_{n_0}(\eta,\mu)\ \|\eta - \mu\| \tag{19}$$

and

$$\|P_\mu - P_\eta\| \leq \alpha(\eta)\ \left[2c\ \alpha(\mu) + \delta_{n_0}(\eta,\mu)\right]\ \|\eta - \mu\|$$

for some finite constant $\delta_{n_0}(\eta,\mu)$ such that

$$\delta_{n_0}(\eta,\mu) \leq \frac{cn_0}{1-(\beta(K_\eta^{n_0})\wedge\beta(K_\mu^{n_0}))} \tag{20}$$

Proof:
The proof of the first assertion is based on the following decomposition

$$\omega(\eta) - \omega(\mu) = \omega(\eta)(K_\eta^{n_0} - K_\mu^{n_0}) + [\omega(\eta) - \omega(\mu)]\, K_\mu^{n_0}$$

Using the fact that

$$\|\, [\omega(\eta) - \omega(\mu)]\, K_\mu^{n_0}\| \leq \beta(K_\mu^{n_0})\ \|\omega(\eta) - \omega(\mu)\|$$

we find that

$$\|\omega(\eta) - \omega(\mu)\| \leq \frac{1}{1 - (\beta(K_\mu^{n_0}) \wedge \beta(K_\eta^{n_0}))}\ \|\omega(\eta)(K_\eta^{n_0} - K_\mu^{n_0})\| \tag{21}$$

On the other hand, we have

$$\|\omega(\eta)(K_\eta^{n_0} - K_\mu^{n_0})\| \leq \|K_\eta^{n_0} - K_\mu^{n_0}\|\ \|\omega(\eta)\| = \|K_\eta^{n_0} - K_\mu^{n_0}\|$$

Using the decomposition

$$K_\eta^{n_0} - K_\mu^{n_0} = \sum_{p=0}^{n_0-1}\ K_\mu^p (K_\eta - K_\mu) K_\eta^{n_0-(p+1)}$$

we find that

$$\|K_\eta^{n_0} - K_\mu^{n_0}\| \leq \sum_{p=0}^{n_0-1}\ \|K_\mu^p (K_\eta - K_\mu) K_\eta^{n_0-(p+1)}\|$$

For any $0 \leq p \leq n_0$ we have

$$\begin{aligned}\|K_\mu^p (K_\eta - K_\mu) K_\eta^{n_0-(p+1)}\| &\leq \|K_\mu^p\|\ \|K_\eta - K_\mu\|\ \|K_\eta^{n_0-(p+1)}\| \\ &\leq \|K_\eta - K_\mu\| \leq c\ \|\eta - \mu\|\end{aligned}$$

from which we conclude that

$$\|K_\eta^{n_0} - K_\mu^{n_0}\| \leq c n_0\ \|\eta - \mu\| \implies \|\omega(\eta)(K_\eta^{n_0} - K_\mu^{n_0})\| \leq c n_0\ \|\eta - \mu\|$$

The proof of (19) is now a direct consequence of (21)

The proof of the second assertion is based on the following decomposition

$$P_\eta - P_\mu = P_\mu (K_\eta - K_\mu) P_\eta + [\omega(\mu) - \omega(\eta)]\, P_\eta$$

To check this formula, we first use the fact that $K_\mu P_\mu = P_\mu K_\mu$ to prove that

$$P_\mu (K_\mu - Id) = (K_\mu - Id) P_\mu = (\omega(\mu) - Id)$$

This yields that

$$P_\mu (K_\mu - Id) P_\eta = (\omega(\mu) - Id) P_\eta$$

Using the Poisson equation and using the fact that $P_\mu(1) = 0$ we also have the decomposition

$$P_\mu (K_\eta - Id) P_\eta = P_\mu (\omega(\eta) - Id) = -P_\mu$$

Combining these two formulae, we conclude that

$$P_\mu(K_\eta - K_\mu)P_\eta = [P_\eta - P_\mu] - [\omega(\mu) - \omega(\eta)]\, P_\eta$$

It follows that

$$\|P_\eta - P_\mu\| \leq \|P_\mu(K_\eta - K_\mu)P_\eta\| + \|\, [\omega(\mu) - \omega(\eta)]\, P_\eta\|$$

The term on the right hand side is easily estimated. Indeed, under our assumptions we readily find that

$$\begin{aligned}\|\, [\omega(\mu) - \omega(\eta)]\, P_\eta\| &\leq \beta(P_\eta)\ \|\omega(\eta) - \omega(\mu)\| \\ &\leq \alpha(\eta)\ \|\omega(\eta) - \omega(\mu)\| \leq \alpha(\eta)\ \delta_{n_0}(\eta,\mu)\ \|\eta - \mu\|\end{aligned}$$

On the other hand, we have

$$\|P_\mu(K_\eta - K_\mu)P_\eta\| \ \leq\ \beta(P_\eta)\ \|P_\mu(K_\eta - K_\mu)\| \leq\ \beta(P_\eta)\ \|P_\mu\|\|K_\eta - K_\mu\|$$

from which we conclude that

$$\|P_\mu(K_\eta - K_\mu)P_\eta\| \ \leq\ 2c\ \alpha(\mu)\alpha(\eta)\ \|\eta - \mu\|$$

The end of the proof is now clear. ∎

3.3 $\mathbb{L}_r$-inequalities and concentration analysis

Firstly, we examine some of the consequences of the pair of regularity conditions presented in (15). The second condition ensures that the functions $\alpha(\eta)$ and $\delta_{n_0}(\eta,\mu)$ introduced in (18) and (20) are uniformly bounded; that is we have

$$1 \leq a(n_0) := \sup_{\eta\in\mathcal{P}(F)} \alpha(\eta) \leq \frac{n_0}{1-b(n_0)} \tag{22}$$

and

$$d(n_0) := \sup_{(\eta,\mu)\in\mathcal{P}(F)^2} \delta_{n_0}(\eta,\mu) \leq \frac{cn_0}{1-b(n_0)} < \infty \tag{23}$$

We recall that $\overline{\omega}_n(\mu)$ is defined in (17). We are now in a position to state and prove the main result of this section.

Theorem 3.3 *For any $n \geq 0$, $f \in \mathcal{B}_1(E)$ and $r \geq 1$ we have the estimate*

$$\mathbb{E}\left(|[\eta_n - \overline{\omega}_n(\mu)]\,(f)|^r\right)^{\frac{1}{r}} \leq e(r)\ \left(\frac{n_0}{1-b(n_0)}\right)^2\ \left[\frac{1}{\sqrt{n+1}} + c\ \mathbb{E}(\overline{\epsilon}(n)^r)^{\frac{1}{r}}\right]$$

for some finite constant $e(r) < \infty$ whose value only depends on the parameter r. In addition, for any $\delta \in (0,1)$ and any time horizon $n \geq 1$, the probability that

$$\begin{aligned}&|[\eta_n - \overline{\omega}_n(\mu)]\,(f)| \\ &\leq \tfrac{n_0}{1-b(n_0)}\ \left[\sqrt{\tfrac{2\log(2/\delta)}{n+1}} + (1+c)\ \left(\tfrac{4n_0}{1-b(n_0)}\right)\left[\overline{\epsilon}(n) \vee \tfrac{1}{n+1}\right]\right]\end{aligned}$$

is greater than $(1-\delta)$ (where c is the constant introduced in (15)).

Corollary 3.4 *For the self-interacting Markov chain associated with the occupation measure distribution flow (16), we have for any $n \geq 0$, $f \in \mathcal{B}_1(E)$ and any $r \geq 1$*

$$\sqrt{n+1}\ \mathbb{E}\left(\left|\left[\eta_n - \overline{\omega}_n(\mu)\right](f)\right|^r\right)^{\frac{1}{r}} \leq e(r)\ (1+c)\left(\frac{n_0}{1-b(n_0)}\right)^2$$

for some finite constant $e(r) < \infty$ whose value only depends on the parameter r. In addition, for any $\delta \in (0,1)$ and any time horizon $n \geq 1$, the probability that

$$\left|\left[\eta_n - \overline{\omega}_n(\mu)\right](f)\right| \leq \left(\frac{2n_0}{1-b(n_0)}\right)^2\ \sqrt{\frac{2}{n+1}}\ \left[\sqrt{\log\left(2/\delta\right)} + 2(1+c)\right]$$

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

Proof of theorem 3.3: Firstly, we examine some consequences of the regularity conditions presented in (15) on the resolvent function P_η introduced in (18). Using proposition 3.1 and proposition 3.2 we find the following uniform estimates

$$\sup_{\eta \in \mathcal{P}(F)} \left(\left(\|P_\eta\|/2\right) \vee \beta(P_\eta)\right) \leq \frac{n_0}{1-b(n_0)}$$

and

$$\|P_\mu - P_\eta\| \leq 3c\left(\frac{n_0}{1-b(n_0)}\right)^2\ \|\mu - \eta\| \tag{24}$$

In addition, using proposition 3.2 again we find that the invariant measure mapping ω is uniform Lipschitz in the sense that

$$\|\omega(\eta) - \omega(\mu)\| \leq\ \frac{cn_0}{1-b(n_0)}\ \|\eta - \mu\|$$

For any $n \geq 0$ and any function $f \in \mathcal{B}_1(E)$, we set

$$I_n(f) := (n+1)\left[\eta_n - \overline{\omega}_n(\mu)\right](f) = \sum_{p=0}^{n}\ \left[f(X_p) - \omega(\mu_p)(f)\right]$$

Using the Poisson equation, we have

$$\left[Id - \omega(\mu_p)\right] = (Id - K_{\mu_p})P_{\mu_p}$$

From this formula, we find the decomposition

$$\begin{aligned}\left[f(X_p) - \omega(\mu_p)(f)\right] &= P_{\mu_p}(f)(X_p) - K_{\mu_p}(P_{\mu_p}(f))(X_p) \\ &= \left[P_{\mu_p}(f)(X_p) - P_{\mu_p}(f)(X_{p+1})\right] + \Delta M_{p+1}(f)\end{aligned} \tag{25}$$

with the increments

$$\Delta M_{p+1}(f) := \left[P_{\mu_p}(f)(X_{p+1}) - K_{\mu_p}(P_{\mu_p}(f))(X_p)\right]$$

of the martingale $M_{n+1}(f)$ defined by

$$M_{n+1}(f) := \sum_{p=1}^{n+1} \Delta M_p(f) = \sum_{p=1}^{n+1}\left[P_{\mu_{p-1}}(f)(X_p) - K_{\mu_{p-1}}(P_{\mu_{p-1}}(f))(X_{p-1})\right]$$

For $n = 0$, we set $M_0(f) = 0$. The first term in the right hand side of (25) can also be rewritten in the following form

$$\begin{aligned}&P_{\mu_p}(f)(X_p) - P_{\mu_p}(f)(X_{p+1})\\&= \left[P_{\mu_p}(f)(X_p) - P_{\mu_{p+1}}(f)(X_{p+1})\right] + \left[P_{\mu_{p+1}}(f)(X_{p+1}) - P_{\mu_p}(f)(X_{p+1})\right]\end{aligned}$$

This yields the decomposition

$$\sum_{p=0}^{n} \left[P_{\mu_p}(f)(X_p) - P_{\mu_p}(f)(X_{p+1})\right] = \left[P_{\mu_0}(f)(X_0) - P_{\mu_{n+1}}(f)(X_{n+1})\right] + L_{n+1}(f)$$

with the random sequence

$$L_{n+1}(f) := \sum_{p=0}^{n} \left[P_{\mu_{p+1}} - P_{\mu_p}\right](f)(X_{p+1})$$

In summary, we have established the following decomposition

$$I_n(f) = M_{n+1}(f) + L_{n+1}(f) + \left[P_{\mu_0}(f)(X_0) - P_{\mu_{n+1}}(f)(X_{n+1})\right]$$

We estimate each term separately. Firstly, using (24) we prove that

$$\left|P_{\mu_0}(f)(X_0) - P_{\mu_{n+1}}(f)(X_{n+1})\right| \leq \|P_{\mu_0}\| + \|P_{\mu_{n+1}}\| \leq \frac{4n_0}{1-b(n_0)}$$

In much the same way, using (24) we find that

$$\begin{aligned}\|L_{n+1}\| \leq \sum_{p=0}^{n} \|P_{\mu_{p+1}} - P_{\mu_p}\| &\leq 3c \left(\frac{n_0}{1-b(n_0)}\right)^2 \sum_{p=0}^{n} \|\mu_{p+1} - \mu_p\|\\&= 3c\ (n+1)\ \left(\frac{n_0}{1-b(n_0)}\right)^2\ \overline{\epsilon}(n)\end{aligned}$$

From these two estimates, we conclude that

$$|I_n(f)| \leq |M_{n+1}(f)| + 3c\ (n+1)\ \left(\frac{n_0}{1-b(n_0)}\right)^2\ \overline{\epsilon}(n) + \frac{4n_0}{1-b(n_0)} \tag{26}$$

To estimate the martingale term, we recall that the unpredictable quadratic variation process $[M(f), M(f)]_n$ of the martingale $M_n(f)$ is the cumulated sum of the square of its increments from the origin up to the current time; that is, we have that

$$[M(f), M(f)]_n := \sum_{p=1}^{n}\ (\Delta M_p(f))^2$$

The main simplification of our regularity conditions comes from the fact that the increments $|\Delta M_p(f)|$ are uniformly bounded. More precisely, we have the almost sure estimates

$$\begin{aligned}|\Delta M_{p+1}(f)| &= \left|P_{\mu_p}(f)(X_{p+1}) - K_{\mu_p}(P_{\mu_p}(f))(X_p)\right|\\&= \left|\int\ \left[P_{\mu_p}(f)(X_{p+1}) - P_{\mu_p}(f)(x)\right]\ K_{\mu_p}(X_p, dx)\right|\\&\leq \int\ \left|P_{\mu_p}(f)(X_{p+1}) - P_{\mu_p}(f)(x)\right|\ K_{\mu_p}(X_p, dx)\end{aligned}$$

from which we conclude that

$$|\Delta M_{p+1}(f)| \leq \mathrm{osc}(P_{\mu_p}(f)) \leq \beta(P_{\mu_p}) \leq \frac{n_0}{1-b(n_0)}$$

By definition of the quadratic variation process $[M(f), M(f)]_n$, this implies that

$$[M(f), M(f)]_n \leq \left(\frac{n_0}{1-b(n_0)}\right)^2 \; n$$

The end of the proof is now a direct consequence of the Burkholder-Davis-Gundy inequality for discrete generation martingales. For any $r \geq 1$, there exists some finite constant $e(r)$ whose value only depends on r, and such that for any n

$$\mathbb{E}\left(\max_{1\leq p\leq n} |M_p(f)|^r\right)^{\frac{1}{r}} \;\leq\; e(r)\; \mathbb{E}\left([M(f), M(f)]_n^{\frac{r}{2}}\right)^{\frac{1}{r}} \leq\; e(r)\; \frac{n_0}{1-b(n_0)}\; \sqrt{n}$$

Combining this estimate with (26), we find that

$$\mathbb{E}\left(|I_n(f)|^r\right)^{\frac{1}{r}} \;\leq\; e(r)\; \left(\frac{n_0}{1-b(n_0)}\right)^2 \; \left[\sqrt{(n+1)} + c\; (n+1)\; \mathbb{E}(\overline{\epsilon}(n)^r)^{\frac{1}{r}}\right]$$

with again some finite constant $e(r)$ whose values may vary from line to line, but only depends on r. Recalling the definition of $I_n(f)$, we conclude that

$$\mathbb{E}\left(|[\eta_n - \overline{\omega}_n(\mu)]\,(f)|^r\right)^{\frac{1}{r}} \leq e(r)\; \left(\frac{n_0}{1-b(n_0)}\right)^2 \; \left[\frac{1}{\sqrt{(n+1)}} + c\; \mathbb{E}(\overline{\epsilon}(n)^r)^{\frac{1}{r}}\right]$$

This ends the proof of the first assertion. To prove the concentration estimates, we use the fact that

$$|[\eta_n - \overline{\omega}_n(\mu)]\,(f)| \leq \frac{|M_{n+1}(f)|}{n+1} + \frac{n_0}{1-b(n_0)}\left[\frac{3c\; n_0}{1-b(n_0)}\;\overline{\epsilon}(n) + \frac{4}{n+1}\right]$$

from which we deduce the rather crude upper bound

$$|[\eta_n - \overline{\omega}_n(\mu)]\,(f)| \leq \frac{|M_{n+1}(f)|}{n+1} + (1+c)\; \left(\frac{2n_0}{1-b(n_0)}\right)^2 \left[\overline{\epsilon}(n) \vee \frac{1}{n+1}\right] \tag{27}$$

The Chernov-Hoeffding exponential inequality states that for every martingale M_n with $M_0 = 0$ and uniformly bounded increments $\sup_n |\Delta M_n| \leq a$, we have

$$\mathbb{P}(|M_n| \geq tn) \leq 2\; e^{-nt^2/2a^2}$$

In our context, we have proved that $\sup_n |\Delta M_n(f)| \leq n_0/(1-b(n_0))$, from which we conclude that

$$\mathbb{P}\left(|[\eta_n - \overline{\omega}_n(\mu)]\,(f)| > t + (1+c)\; \left(\tfrac{2n_0}{1-b(n_0)}\right)^2 \left[\overline{\epsilon}(n) \vee \tfrac{1}{n+1}\right]\right)$$

$$\leq 2\; \exp\left(-(n+1)\; \tfrac{t^2}{2}\; \left(\tfrac{1-b(n_0)}{n_0}\right)^2\right)$$

We conclude the proof of the theorem by choosing $t = \frac{n_0}{1-b(n_0)}\sqrt{\frac{2\log(2/\delta)}{n+1}}$. ∎

4 Distribution flows models

In this section, we have collected the definition of a series of semigroups on distribution flow state spaces. We also take the opportunity to describe some of their regularity properties we shall use in the further developments of the article.

We equip the sets of distribution flows $\mathcal{P}(S^{(l)})^{\mathbb{N}}$ with the uniform total variation distance defined by

$$\forall(\eta,\mu)\in\left(\mathcal{P}(S^{(l)})^{\mathbb{N}}\right)^2 \qquad \|\eta-\mu\|:=\sup_{n\geq 0}\|\eta_n-\mu_n\|$$

We extend a given integral operator $\mu\in\mathcal{P}(S^{(l)})\mapsto\mu L\in\mathcal{P}(S^{(l+1)})$ into a mapping

$$\eta=(\eta_n)_{n\geq 0}\ \in\mathcal{P}(S^{(l)})^{\mathbb{N}}\mapsto\eta L=(\eta_n L)_{n\geq 0}\mathcal{P}(S^{(l+1)})^{\mathbb{N}}$$

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

4.1 Time averaged semigroups

We associate with the mappings Φ_l introduced in (1) the mappings

$$\Phi^{(l)}\ :\ \eta\in\mathcal{P}(S^{(l-1)})^{\mathbb{N}}\ \mapsto\ \Phi^{(l)}(\eta)=\left(\Phi^{(l)}_n(\eta)\right)_{n\geq 0}\in\mathcal{P}(S^{(l)})^{\mathbb{N}}$$

defined by the coordinate mappings

$$\forall\eta\in\mathcal{P}(S^{(l-1)})^{\mathbb{N}}\quad\forall n\geq 0\qquad \Phi^{(l)}_n(\eta):=\Phi_l(\eta_n)$$

We denote by

$$\Phi^{(k,l)}=\Phi^{(k)}\circ\Phi^{(k-1,l)}$$

with $0\leq l\leq k$, the semigroup associated with the mappings $\Phi^{(l)}$. We also consider the time averaged transformations

$$\overline{\Phi}^{(l)}\ :\ \eta\in\mathcal{P}(S^{(l-1)})^{\mathbb{N}}\ \mapsto\ \overline{\Phi}^{(l)}(\eta)=\left(\overline{\Phi}^{(l)}_n(\eta)\right)_{n\geq 0}\in\mathcal{P}(S^{(l)})^{\mathbb{N}}$$

defined by the coordinate mappings

$$\begin{aligned}\forall\eta\in\mathcal{P}(S^{(l-1)})^{\mathbb{N}}\quad\forall n\geq 0\qquad \overline{\Phi}^{(l)}_n(\eta) &:= \frac{1}{n+1}\sum_{p=0}^{n}\Phi^{(l)}_p(\eta)\\ &= \frac{1}{n+1}\sum_{p=0}^{n}\Phi_l(\eta_p)\in\mathcal{P}(S^{(l)})\end{aligned}$$

For $l=0$, we use the convention $\Phi_0(\eta_p)=\pi^{(0)}$ for any $0\leq p\leq n$, so that with some abusive but obvious notation $\overline{\Phi}^{(0)}(\eta)=\pi^{(0)}$ represents the constant sequence $(\pi^{(0)})_{n\geq 0}$ such that $\pi^{(0)}_n=\pi^{(0)}$.

We also denote $\overline{\Phi}^{(k,l)}\ :\ \mathcal{P}(S^{(l-1)})^{\mathbb{N}}\to\mathcal{P}(S^{(k)})^{\mathbb{N}}$ with $0\leq l\leq k$, the semigroup associated with the mappings $\overline{\Phi}^{(l)}$ and defined by

$$\overline{\Phi}^{(k,l)}:=\overline{\Phi}^{(k)}\circ\overline{\Phi}^{(k-1)}\circ\ldots\circ\overline{\Phi}^{(l)}$$

We use the convention $\overline{\Phi}^{(k,l)}=Id$, the identity operator, for $l>k$.

4.2 Integral operators

We associate with the integral operator Γ_k from $\mathcal{B}(S^{(k)})$ into $\mathcal{B}(S^{(k-1)})$ introduced in (7) the integral operator $\overline{\Gamma}^{(k)}$ from $(\mathbb{N}\times\mathcal{B}(S^{(k)}))$ into the set $(\mathbb{N}\times\mathcal{B}(S^{(k-1)}))$ defined by

$$\overline{\Gamma}^{(k)}((n,f),d(p,g)) := \Sigma(n,dp)\times\Gamma_k(f,dg) \quad \text{with} \quad \Sigma(n,dp) := \frac{1}{n+1}\sum_{q=0}^{n}\delta_q(dp)$$

The semigroup $\overline{\Gamma}^{(l_2,l_1)}$ $(0\leq l_1\leq l_2)$ associated with the integral operators $\overline{\Gamma}^{(l)}$ ise defined by

$$\overline{\Gamma}^{(l_2,l_1)} := \overline{\Gamma}^{(l_2)}\overline{\Gamma}^{(l_2-1)}\ldots\overline{\Gamma}^{(l_1)}$$

For $l_1=l_2=0$, we use the convention $\overline{\Gamma}^{(0,0)}=\overline{\Gamma}^{(0)}=0$ for the null measure on $(\mathbb{N}\times\mathcal{B}(S^{(0)}))$. Also observe that

$$\overline{\Gamma}^{(l_2,l_1)} = \Sigma^{l_2-l_1+1}\times\Gamma_{l_2,l_1}$$

where the semigroups Σ^{l_1} and Γ_{l_2,l_1}, $0\leq l_1\leq l_2$ associated with the pair of integral operators Σ and Γ_l are

$$\Sigma^{l_1}=\Sigma\Sigma^{l_1-1}=\Sigma^{l_1-1}\Sigma \quad \text{and} \quad \Gamma_{l_2,l_1}:=\Gamma_{l_2}\Gamma_{l_2-1}\ldots\Gamma_{l_1}$$

We use the convention $\Sigma^0=Id$.

We end this section with a technical lemma relating the regularity properties (7) of the mappings Φ_k to the regularity properties of the semigroups $\overline{\Phi}^{(k,l)}$.

Lemma 4.1 *For any $0\leq l_1\leq l_2$, $n\geq 0$, any flow of measures $\eta,\mu\in\mathcal{P}(S^{(l_1-1)})^{\mathbb{N}}$ and any function $f\in\mathcal{B}(S^{(l_2)})$ we have*

$$\begin{aligned}&\left|\left[\overline{\Phi}_n^{(l_2,l_1)}(\eta)-\overline{\Phi}_n^{(l_2,l_1)}(\mu)\right](f)\right|\\ &\leq\int_{(\mathbb{N}\times\mathcal{B}(S^{(l_1-1)}))}\left|[\eta_p-\mu_p](g)\right|\;\overline{\Gamma}^{(l_2,l_1)}((n,f),d(p,g))\end{aligned}$$

Proof:
Notice that we have $\overline{\Gamma}^{(l,l)}=\overline{\Gamma}^{(l)}$. We also observe that $\overline{\Gamma}^{(l_2,l_1)}$ is a bounded integral operator from $(\mathbb{N}\times\mathcal{B}(S^{(l_2)}))$ into $(\mathbb{N}\times\mathcal{B}_n(S^{(l_1-1)}))$. We prove the lemma by induction on the parameter $k=l_2-l_1$. The result is clearly true for $k=0$. Indeed, by (7) we find that for any $l\geq 0$

$$\begin{aligned}\left|\left[\overline{\Phi}_n^{(l)}(\eta)-\overline{\Phi}_n^{(l)}(\mu)\right](f)\right| &\leq \frac{1}{n+1}\sum_{p=0}^{n}\left|[\Phi_l(\eta_p)-\Phi_l(\mu_p)](f)\right|\\ &\leq\frac{1}{n+1}\sum_{p=0}^{n}\int_{\mathcal{B}(S^{(l-1)})}\left|[\eta_p-\mu_p](g)\right|\;\Gamma(f,dg)\end{aligned}$$

Rewritten in terms of $\overline{\Gamma}^{(l)}$, we have proved that

$$\left|\left[\overline{\Phi}_n^{(l)}(\eta)-\overline{\Phi}_n^{(l)}(\mu)\right](f)\right|\leq\int_{(\mathbb{N}\times\mathcal{B}(S^{(l-1)}))}\left|[\eta_p-\mu_p](g)\right|\;\overline{\Gamma}^{(l)}((n,f),d(p,g))$$

This ends the proof of the result for $k = 0$. Now, suppose we have proved that

$$\left|\left[\overline{\Phi}_p^{(l_2,l_1)}(\eta) - \overline{\Phi}_p^{(l_2,l_1)}(\mu)\right](g)\right| \leq \int \ |[\eta_q - \mu_q]\,(h)|\ \overline{\Gamma}^{(l_2,l_1)}((p,g), d(q,h))$$

for any pair of integers $l_1 < l_2$ with $l_2 - l_1 = k$ for some $k \geq 1$. In this case, for any $l < k$ and any function $f \in \mathcal{B}(S^{(l+1)})$, we have

$$\begin{aligned}
&\left|\left[\overline{\Phi}_n^{(l+1,l-k)}(\eta) - \overline{\Phi}_n^{(l+1,l-k)}(\mu)\right](f)\right| \\
&= \left|\left[\overline{\Phi}_n^{(l+1)}(\overline{\Phi}^{(l,l-k)}(\eta)) - \overline{\Phi}_n^{(l+1)}(\overline{\Phi}^{(l,l-k)}(\mu))\right](f)\right|
\end{aligned}$$

and therefore

$$\begin{aligned}
&\left|\left[\overline{\Phi}_n^{(l+1,l-k)}(\eta) - \overline{\Phi}_n^{(l+1,l-k)}(\mu)\right](f)\right| \\
&\leq \int \ \left|\left[\overline{\Phi}_p^{(l,l-k)}(\eta) - \overline{\Phi}_p^{(l,l-k)}(\mu)\right](g)\right|\ \overline{\Gamma}^{(l+1)}((n,f), d(p,g))
\end{aligned}$$

Under our induction hypothesis, this implies that

$$\begin{aligned}
&\left|\left[\overline{\Phi}_n^{(l+1,l-k)}(\eta) - \overline{\Phi}_n^{(l+1,l-k)}(\mu)\right](f)\right| \\
&\leq \int \ |[\eta_q - \mu_q]\,(h)|\ \int \ \overline{\Gamma}^{(l+1)}((n,f), d(p,g))\ \overline{\Gamma}^{(l,l-k)}((p,g), d(q,h)) \\
&= \int \ |[\eta_q - \mu_q]\,(h)|\ \overline{\Gamma}^{(l+1,l-k)}((n,f), d(q,h))
\end{aligned}$$

Letting $l_1 = (l-k)$ and $l_2 = (l+1)$, we have proved that for any $l_1 < l_2$ with $l_2 - l_1 = (k+1)$

$$\left|\left[\overline{\Phi}_n^{(l_2,l_1)}(\eta) - \overline{\Phi}_n^{(l_2,l_1)}(\mu)\right](f)\right| \leq \int \ |[\eta_p - \mu_p]\,(g)|\ \overline{\Gamma}^{(l_2,l_1)}((n,f), d(p,g))$$

This ends the proof of the lemma. ■

4.3 Path space semigroups

To simplify the presentation, we fix a time horizon $m \geq 1$ and write ω instead of $\omega_{K_\eta^{(m)}}$, the invariant measure mapping defined in (9). We also write E instead of E_m.

We extend the mapping ω on $\mathcal{P}(E)$ to $\mathcal{P}(E)^{\mathbb{N}}$ by setting

$$\omega \ :\ \eta = (\eta_n)_{n\geq 0} \in \mathcal{P}(E)^{\mathbb{N}} \ \mapsto\ \omega(\eta) = (\omega_n(\eta))_{n\geq 0} \in \mathcal{P}(E)^{\mathbb{N}}$$

with the coordinate mappings ω_n defined by

$$\omega_n(\eta) := \omega(\eta_n) = \pi^{(0)} \otimes \Phi_1(\eta_n^{(0)}) \otimes \ldots \otimes \Phi_m(\eta_n^{(m-1)})$$

For every $l \leq m$, we recall that $\eta_n^{(l)}$ stands for the image measure on $S^{(l)}$ of a given measure $\eta_n \in \mathcal{P}(E_m)$. We also consider the mappings

$$\overline{\omega} \; : \; \eta \in \mathcal{P}(E)^{\mathbb{N}} \; \mapsto \; \overline{\omega}(\eta) = (\overline{\omega}_n(\eta))_{n\geq 0} \in \mathcal{P}(E)^{\mathbb{N}}$$

defined by the coordinate mappings

$$\forall \eta = (\eta_n)_{n\geq 0} \in \mathcal{P}(E)^{\mathbb{N}} \quad \forall n \geq 0 \qquad \overline{\omega}_n(\eta) := \frac{1}{n+1}\sum_{p=0}^{n} \omega_p(\eta) = \frac{1}{n+1}\sum_{p=0}^{n} \omega(\eta_p)$$

Lemma 4.2 *For any $1 \leq k \leq m$ and any flow of measures $\eta \in \mathcal{P}(E)^{\mathbb{N}}$, we have*

$$\omega^k(\eta) = \overline{\pi}^{[k-1]} \otimes \bigotimes_{i=0}^{m-k} \Phi^{(i+k,i+1)}(\eta^{(i)})$$

For $k = m+1$, we have

$$\forall \eta \in \mathcal{P}(E)^{\mathbb{N}} \qquad \omega^{m+1}(\eta) = \pi^{[m]}$$

Proof:
We use a simple induction on the parameter k. The result is clearly true for $k = 1$. Suppose we have proved the result at some rank k. In this case we have

$$\begin{aligned}
\omega^k(\omega(\eta)) &= \overline{\pi}^{[k-1]} \otimes \Phi_{k,1}(\omega(\eta)^{(0)}) \otimes \bigotimes_{i=1}^{m-k} \Phi_{i+k,i+1}(\omega(\eta)^{(i)}) \\
&= \overline{\pi}^{[k-1]} \otimes \pi^{(k)} \otimes \bigotimes_{i=1}^{m-k} \Phi_{i+k,i}(\eta^{(i-1)}) \\
&= \overline{\pi}^{[k]} \otimes \bigotimes_{i=0}^{m-(k+1)} \Phi_{i+(k+1),i+1}(\eta^{(i)})
\end{aligned}$$

This ends the proof of the lemma. ∎

Lemma 4.3 *For any $1 \leq k \leq m$ and any $\eta = (\eta_n)_{n\geq 0} \in \mathcal{P}(E)^{\mathbb{N}}$, we have*

$$\overline{\omega}_n^k(\eta) = \frac{1}{n+1}\sum_{p=0}^{n}\left[\overline{\pi}^{[k-1]} \otimes \bigotimes_{i=0}^{m-k} \Phi_p^{(i+k)}\left(\overline{\Phi}^{(i+(k-1),i+1)}(\eta^{(i)})\right)\right]$$

For $k = m+1$, we have

$$\forall \eta \in \mathcal{P}(E)^{\mathbb{N}} \qquad \overline{\omega}^{m+1}(\eta) = \pi^{[m]}$$

Proof:
We use a simple induction on the parameter k. The result is clearly true for $k = 1$. Indeed, we have in this case

$$\overline{\omega}_n(\eta) = \frac{1}{n+1}\sum_{p=0}^{n}\left[\overline{\pi}^{[k-1]} \otimes \bigotimes_{i=0}^{m-1} \Phi_p^{(i+1)}\left(\eta^{(i)}\right)\right]$$

We also observe that

$$\overline{\omega}_n(\eta)^{(i)} = \frac{1}{n+1}\sum_{p=0}^{n} \Phi_p^{(i)}\left(\eta^{(i-1)}\right) = \overline{\Phi}_n^{(i)}\left(\eta^{(i-1)}\right) \Rightarrow \overline{\omega}(\eta)^{(i)} = \overline{\Phi}^{(i)}\left(\eta^{(i-1)}\right)$$

Suppose we have proved the result at some rank k. In this case, we have

$$\overline{\omega}^k(\overline{\omega}(\eta)) = \frac{1}{n+1}\sum_{p=0}^{n}\left[\overline{\pi}^{[k]} \otimes \bigotimes_{i=1}^{m-k} \Phi_p^{(i+k)}\left(\overline{\Phi}^{(i+(k-1),i)}(\eta^{(i-1)})\right)\right]$$

from which we conclude that

$$\overline{\omega}^{k+1}(\eta) = \frac{1}{n+1}\sum_{p=0}^{n}\left[\overline{\pi}^{[k]} \otimes \bigotimes_{i=0}^{m-(k+1)} \Phi_p^{(i+(k+1))}\left(\overline{\Phi}^{(i+k,i+1)}(\eta^{(i)})\right)\right]$$

This ends the proof of the lemma. ■

5 Asymptotic analysis

5.1 Introduction

This section is concerned with the asymptotic behavior of i-MCMC models as the time index n tends to infinity.

The strong law of large numbers is discussed in section 5.2. We present non-asymptotic $\mathbb{L}_r$-inequalities that allow us to quantify the convergence of the occupation measures $\eta_n^{(k)} = \frac{1}{n+1}\sum_{p=0}^{n}\delta_{X_p^{(k)}}$ of i-MCMC models towards the solution $\pi^{(k)}$ of the measure-valued equation (1).

Section 5.3 is concerned with uniform convergence results with respect to the number of levels k. We examine this important question in terms of the stability properties of the time averaged semigroups introduced in section 4.1. We present non-asymptotic $\mathbb{L}_r$-inequalities for a series of i-MCMC models that do not depend on the number of levels. These estimates are probably the most important in practice since they allow us to quantify the running time of a i-MCMC to achieve a given precision independently of the time horizon of the limiting measure-valued equation (1).

Our approach is based on an original combination of nonlinear semigroup techniques with the asymptotic analysis of time inhomogeneous Markov chains developed in section 3. The following technical lemma presents a more or less well known generalized Minkowski integral inequality which will be used in our proofs.

Lemma 5.1 (generalized Minkowski integral inequality) *For any pair of bounded positive measures μ_1 and μ_2 on some measurable spaces $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$, any bounded measurable function φ on the product space $(E_1 \times E_2)$ any $p \geq 1$, we have*

$$\left[\int_{E_1} \mu_1(dx_1) \left|\int_{E_2} \varphi(x_1, x_2)\ \mu_2(dx_2)\right|^p\ \right]^{\frac{1}{p}}$$

$$\leq \int_{E_2} \left(\int_{E_1} |\varphi(x_1, x_2)|^p\ \mu_1(dx_1)\right)^{\frac{1}{p}}\ \mu_2(dx_2)$$

Proof:
Without loss of generality, we suppose that φ is a nonnegative function. For $p = 1$, the lemma is a direct consequence of Fubini's theorem. Let us assume that $p > 1$, and let p' be such that $\frac{1}{p'} + \frac{1}{p} = 1$. Firstly, we notice that the functions

$$\varphi_1(x_1) := \int_{E_2} \varphi(x_1, x_2)\, \mu_2(dx_2) \quad \text{and} \quad \phi_p(x_2) := \left(\int_{E_1} |\varphi(x_1, x_2)|^p\ \mu_1(dx_1)\right)^{\frac{1}{p}}$$

are measurable for every $p \geq 1$. In this notation, we need to prove that $\mu_1(\varphi_1^p)^{\frac{1}{p}} \leq \mu_2(\phi_p)$. It is also convenient to consider the function

$$\psi(x_1, x_2) := \varphi(x_1, x_2)/\phi_p(x_2)^{\frac{1}{p'}}$$

We use the convention $\psi(x_1, x_2) = 0$, for every $x_1 \in E_1$ as long as $\phi_p(x_2) = 0$. We observe that

$$\left(\int_{E_1} \psi(x_1, x_2)^p\ \mu_1(dx_1)\right)^{\frac{1}{p}} = \phi_p(x_2)/\phi_p(x_2)^{\frac{1}{p'}} = \phi_p(x_2)^{\frac{1}{p}}$$

By construction, we have

$$\begin{aligned}\varphi_1(x_1) &= \int_{E_2} \psi(x_1, x_2)\ \phi_p(x_2)^{\frac{1}{p'}}\ \mu_2(dx_2) \\ &\leq \left[\int_{E_2} \psi(x_1, x_2)^p\ \mu_2(dx_2)\right]^{\frac{1}{p}} \times \mu_2(\phi_p)^{\frac{1}{p'}}\end{aligned}$$

from which we conclude that

$$\begin{aligned}\mu_1(\varphi_1^p) &\leq \mu_2(\phi_p)^{\frac{p}{p'}} \times \left[\int_{E_2} \psi(x_1, x_2)^p\ \mu_1(dx_1)\mu_2(dx_2)\right] \\ &= \mu_2(\phi_p)^{\frac{p}{p'}} \times \mu_2(\phi_p) = \mu_2(\phi_p)^p\end{aligned}$$

The end of the proof is now clear. ■

5.2 Strong law of large numbers

This section is mainly concerned with the proof of the following $\mathbb{L}_r$-inequalities for the occupation measure of an i-MCMC model at a given level set.

Theorem 5.2 *Under the regularity conditions (7) and (8), we have for any $k \geq 0$, any function $f \in \mathcal{B}_1(S^{(k)})$ and any $n \geq 0$ and $r \geq 1$*

$$\sqrt{(n+1)}\,\mathbb{E}\left(\left|\left[\eta_n^{(k)} - \pi^{(k)}\right](f)\right|^r\right)^{\frac{1}{r}} \leq e(r)\,\sum_{l=0}^{k}(1+c_l)\,\left(\frac{n_l}{1-b_l(n_l)}\right)^2 \prod_{l+1\leq i\leq k} 2\Lambda_i \tag{28}$$

Proof:
We prove the theorem by induction on the parameter k. Firstly, we observe that the estimate (28) is true for $k = 0$. Indeed, by corollary 3.4 we have that

$$\sqrt{(n+1)}\,\mathbb{E}\left(\left|\left[\eta_n^{(0)} - \pi^{(0)}\right](f)\right|^r\right)^{\frac{1}{r}} \leq e(r)\,\left(\tfrac{n_0}{1-b_0(n_0)}\right)^2$$

for some finite constant $e(r) < \infty$ whose value only depends on the parameter r. We further suppose that the estimate (28) is true at rank $(k-1)$. To prove that it is also true at rank k, we use the decomposition

$$\left[\eta_n^{(k)} - \pi^{(k)}\right] = \left[\eta_n^{(k)} - \overline{\Phi}_n^{(k)}(\eta^{(k-1)})\right] + \left[\overline{\Phi}_n^{(k)}(\eta^{(k-1)}) - \overline{\Phi}_n^{(k)}(\pi^{(k-1)})\right] \tag{29}$$

For every $k \geq 0$, given a realization of the chain $X^{(k-1)} := (X_p^{(k-1)})_{p\geq 0}$ the k-th level chain $X_n^{(k)}$ behaves as a Markov chain with random Markov transitions $M^{(k)}_{\eta_n^{(k-1)}}$ dependent on the current occupation measure of the chain at level $(k-1)$. Therefore, using corollary 3.4 again we notice that

$$\sqrt{(n+1)}\,\mathbb{E}\left(\left|\left[\eta_n^{(k)} - \overline{\Phi}_n^{(k)}(\eta^{(k-1)})\right](f)\right|^r\right)^{\frac{1}{r}} \leq e(r)\,(1+c_k)\,\left(\tfrac{n_k}{1-b_k(n_k)}\right)^2$$

for some finite constant $e(r) < \infty$ whose values only depends on the parameter r.

Using the decomposition (29) and lemma 4.1, we obtain

$$\begin{aligned}
&\left|[\left[\eta_n^{(k)} - \pi^{(k)}\right](f)\right| \\
&\leq \left|\left[\eta_n^{(k)} - \overline{\Phi}_n^{(k)}(\eta^{(k-1)})\right](f)\right| + \int\ \left|\left[\eta_p^{(k-1)} - \pi^{(k-1)}\right](g)\right|\ \overline{\Gamma}^{(k)}((n,f),d(p,g))
\end{aligned}$$

For every function $f \in \mathcal{B}_1(S^{(l)})$, and any $n \geq 0$, $k \geq 0$, $r \geq 1$, we set

$$J_n^{(k)}(f) := \sqrt{n+1}\,\mathbb{E}\left(\left|\left[\eta_n^{(k)} - \pi^{(k)}\right](f)\right|^r\right)^{\frac{1}{r}} \quad \text{and} \quad j^{(k)} := \sup_{n\geq 1}\ \sup_{f:\|f\|\leq 1} J_n^{(k)}(f)$$

By the generalized Minkowski integral inequality presented in lemma 5.1, we find that

$$\begin{aligned}
J_n^{(k)}(f) \quad &\leq \quad e(r)(1+c_k)\,\left(\frac{n_k}{1-b_l(n_k)}\right)^2 \\
&\qquad + \sqrt{n+1}\int\ J_p^{(k-1)}(g)\,\frac{1}{\sqrt{p+1}}\,\overline{\Gamma}^{(k)}((n,f),d(p,g))
\end{aligned}$$

Since we have

$$\int_{\mathbb{N}} \frac{1}{\sqrt{q+1}}\ \Sigma(n,dq) = \frac{1}{n+1}\sum_{q=0}^{n}\frac{1}{\sqrt{q+1}} \leq \frac{2}{\sqrt{n+1}} \tag{30}$$

we conclude that

$$J_n^{(k)}(f) \leq e(r)(1+c_k)\ \left(\frac{n_k}{1-b_l(n_k)}\right)^2 + 2\ j^{(k-1)}\ \sup_f \int\ \|g\|\ \Gamma_k(f,dg)$$

and therefore

$$j^{(k)} \leq e(r)(1+c_k)\ \left(\frac{n_k}{1-b_k(n_k)}\right)^2 + j^{(k-1)}\ 2\Lambda_k$$

Under the induction hypothesis, we have

$$j^{(k-1)}\ 2\Lambda_k\ \leq e(r)\ \sum_{l=0}^{k-1}(1+c_l)\ \left(\frac{n_l}{1-b_l(n_l)}\right)^2 \prod_{l+1\leq i\leq k} 2\Lambda_i$$

and therefore

$$\begin{aligned} j^{(k)} \quad &\leq \quad e(r)\left[(1+c_k)\ \left(\frac{n_k}{1-b_k(n_k)}\right)^2\right. \\ &\qquad\qquad \left. + \sum_{l=0}^{k-1}(1+c_l)\ \left(\frac{n_l}{1-b_l(n_l)}\right)^2 \prod_{l+1\leq i\leq k} 2\Lambda_i\right] \\ &= \quad \sum_{l=0}^{k}(1+c_l)\ \left(\frac{n_l}{1-b_l(n_l)}\right)^2 \prod_{l+1\leq i\leq k} 2\Lambda_i \end{aligned}$$

This ends the proof of the theorem. ∎

5.3 A uniform convergence theorem

This section focuses on the behavior of an i-MCMC model associated with a large number of levels. We establish an uniform convergence theorem under the assumption that the time averaged semigroup $\overline{\Phi}^{(k,l)}$ introduced in section 4.1 is exponentially stable; that is there exist a pair of positive constants $\lambda_1, \lambda_2 > 0$ and an integer k_0 such that for every $l \geq 0$, $\eta, \mu \in \mathcal{P}(S^{(l)})^{\mathbb{N}}$ and any $k \geq k_0$ we have

$$\|\overline{\Phi}^{(l+k,l+1)}(\eta) - \overline{\Phi}^{(l+k,l+1)}(\mu)\| \leq \lambda_1\ e^{-\lambda_2 k} \tag{31}$$

We also assume that the parameters $(b_k, c_k, n_k, \Lambda_k)$ are chosen so that

$$A = \sup_{k\geq 0}\left[(1+c_k)\ \left(\frac{n_k}{1-b_k(n_k)}\right)^2\right] < \infty \quad \text{and} \quad B := 2\sup_{k\geq 1}\Lambda_k < \infty \tag{32}$$

For the Feynman-Kac transformations (11), we give in section 7 sufficient conditions on G_l and L_{l+1} ensuring (31) is satisfied. If (31) and (32) are both satisfied, we have the following uniform convergence result.

Theorem 5.3 *If $B=1$, then we have for any $r\geq 1$, any parameter n such that $(n+1)\geq e^{2\lambda_2(k_0+1)}$, and for any $(f_l)_{l\geq 0}\in\prod_{l\geq 0}\mathrm{Osc}_1(S^{(l)})$*

$$\sup_{l\geq 0}\mathbb{E}\left(\left|\left[\eta_n^{(l)}-\pi^{(l)}\right](f_l)\right|^r\right)^{\frac{1}{r}}\leq\frac{e(r)}{\sqrt{n+1}}\left(A\ \left(1+\frac{\log(n+1)}{2\lambda_2}\right)+\lambda_2\ e^{\lambda_2}\right)$$

If $B>1$, then we have for any $r\geq 1$, any n such that $(n+1)\geq e^{2(\lambda_2+\log B)(k_0+1)}$, and for any $(f_l)_{l\geq 0}\in\prod_{l\geq 0}\mathrm{Osc}_1(S^{(l)})$

$$\sup_{l\geq 0}\mathbb{E}\left(\left|\left[\eta_n^{(l)}-\pi^{(l)}\right](f_l)\right|^r\right)^{\frac{1}{r}}\leq e(r)\ \left[\frac{AB}{B-1}+\lambda_1\right]\ \frac{e^{\lambda_2}}{(n+1)^{\alpha/2}}$$

with $\alpha:=\frac{\lambda_2}{(\lambda_2+\log B)}$.

Proof:

Firstly, we notice that for any $k\geq 0$, we have the following estimate from (28) and (32)

$$\sqrt{(n+1)}\ \mathbb{E}\left(\left|\left[\eta_n^{(k)}-\pi^{(k)}\right](f_k)\right|^r\right)^{\frac{1}{r}}\leq e(r)\ A\ \frac{B^{k+1}-1}{B-1}\tag{33}$$

For $B=1$, we use the convention $\frac{B^k-1}{B-1}=k$.

We have the following decomposition

$$\begin{aligned}\eta_n^{(l+k)}-\pi^{(l+k)}&=\left[\eta_n^{(l+k)}-\overline{\Phi}_n^{(l+k,l+1)}(\eta^{(l)})\right]+\left[\overline{\Phi}_n^{(l+k,l+1)}(\eta^{(l)})-\overline{\Phi}_n^{(l+k,l+1)}(\pi^{(l)})\right]\\&=\sum_{i=l+1}^{l+k}\left[\overline{\Phi}_n^{(l+k,i+1)}(\eta^{(i)})-\overline{\Phi}_n^{(l+k,i+1)}(\overline{\Phi}^{(i)}(\eta^{(i-1)}))\right]\\&\qquad+\left[\overline{\Phi}_n^{(l+k,l+1)}(\eta^{(l)})-\overline{\Phi}_n^{(l+k,l+1)}(\pi^{(l)})\right]\end{aligned}\tag{34}$$

Recall that we use the convention $\overline{\Phi}^{(l_1,l_2)}=Id$ for $l_1<l_2$, so that

$$i=l+k\Longrightarrow\overline{\Phi}_n^{(l+k,i+1)}(\eta^{(i)})=\overline{\Phi}_n^{(l+k,l+k+2)}(\eta^{(l+k)})=\eta_n^{(l+k)}$$

Using lemma 4.1, we find that

$$\begin{aligned}&\left|\left[\overline{\Phi}_n^{(l_2,l_1+1)}(\eta^{(l_1)})-\overline{\Phi}_n^{(l_2,l_1)}(\overline{\Phi}^{(l_1)}(\eta^{(l_1-1)}))\right](f_{l_2})\right|\\&\leq\int\ \left|\left[\eta_p^{(l_1)}-\overline{\Phi}_p^{(l_1)}(\eta^{(l_1-1)})\right](g)\right|\ \overline{\Gamma}^{(l_2,l_1+1)}((n,f_{l_2}),d(p,g))\end{aligned}$$

By the generalized Minkowski integral inequality, this implies that

$$\begin{aligned}&\mathbb{E}\left(\left|\left[\overline{\Phi}_n^{(l_2,l_1+1)}(\eta^{(l_1)})-\overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\Phi}^{(l_1)}(\eta^{(l_1-1)}))\right](f_{l_2})\right|^r\right)^{\frac{1}{r}}\\&\leq\int\ \mathbb{E}\left(\left|\left[\eta_p^{(l_1)}-\overline{\Phi}_p^{(l_1)}(\eta^{(l_1-1)})\right](g)\right|^r\right)^{\frac{1}{r}}\ \overline{\Gamma}^{(l_2,l_1+1)}((n,f_{l_2}),d(p,g))\end{aligned}$$

Using corollary 3.4, we find that

$$\mathbb{E}\left(\left|\left[\overline{\Phi}_n^{(l_2,l_1+1)}(\eta^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\Phi}^{(l_1)}(\eta^{(l_1-1)}))\right](f_{l_2})\right|^r\right)^{\frac{1}{r}}$$

$$\leq e(r)\ (1+c_{l_1})\ \left(\tfrac{n_{l_1}}{1-b_{l_1}(n_{l_1})}\right)^2$$

$$\times \int_{\{0,\ldots,n\}} \frac{1}{\sqrt{(p+1)}}\ \Sigma^{(l_2-l_1)}(n,dp) \times \int\ \|g\|\ \Gamma_{l_2,l_1+1}(f_{l_2},dg)$$

By (30) and

$$\int\ \Gamma_{k,l}(f_{l_2},dg)\ \|g\| \leq \Lambda_{k,l}\ \|f_{l_2}\| \quad \text{with} \quad \Lambda_{k,l} \leq \prod_{l\leq i\leq k} \Lambda_i \leq B^{k-l+1} < \infty$$

we conclude that

$$\sqrt{(n+1)}\mathbb{E}\left(\left|\left[\overline{\Phi}_n^{(l_2,l_1+1)}(\eta^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\Phi}^{(l_1)}(\eta^{(l_1-1)}))\right](f_{l_2})\right|^r\right)^{\frac{1}{r}} \leq e(r)\ A\ B^{l_2-l_1}\ \|f_{l_2}\| \tag{35}$$

Using the decomposition (34), we prove that for every $f_{l+k} \in \mathcal{B}_1(S^{(l+k)})$ and any $k \geq k_0$

$$\sup_{l\geq 0}\mathbb{E}\left(\left|\left[\eta_n^{(l+k)} - \pi^{(l+k)}\right](f_{l+k})\right|^r\right)^{\frac{1}{r}} \leq e(r)\ \frac{A}{\sqrt{n+1}}\ \frac{B^k-1}{B-1} + \lambda_1\ e^{-\lambda_2\ k}$$

Finally, by (33), we conclude that for every $k \geq k_0$

$$\sup_{l\geq 0}\mathbb{E}\left(\left|\left[\eta_n^{(l)} - \pi^{(l)}\right](f_l)\right|^r\right)^{\frac{1}{r}} \leq e(r)\ \frac{A}{\sqrt{n+1}}\ \frac{B^{k+1}-1}{B-1} + \lambda_1\ e^{-\lambda_2\ k}$$

For $B=1$, we have

$$\sup_{l\geq 0}\mathbb{E}\left(\left|\left[\eta_n^{(l)} - \pi^{(l)}\right](f_l)\right|^r\right)^{\frac{1}{r}} \leq e(r)\ A\ \frac{(k+1)}{\sqrt{n+1}} + \lambda_1\ e^{-\lambda_2\ k}$$

In this situation, we choose the parameters k, n such that

$$k = k(n) := \lfloor \frac{\log(n+1)}{2\lambda_2} \rfloor \geq k_0$$

Notice that $k(n)$ is the largest integer k satisfying

$$k \leq \frac{\log(n+1)}{2\lambda_2} \quad \left(\Leftrightarrow \frac{1}{\sqrt{n+1}} \leq e^{-\lambda_2\ k}\right)$$

Since $(k(n)+1) \geq \frac{\log(n+1)}{2\lambda_2}$, we have

$$e^{-\lambda_2\ k(n)} \leq e^{\lambda_2}\ e^{-\lambda_2\ \frac{\log(n+1)}{2\lambda_2}} = \frac{e^{\lambda_2}}{\sqrt{n+1}}$$

from which we conclude that

$$A\ \frac{(k(n)+1)}{\sqrt{n+1}} + \lambda_1\ e^{-\lambda_2\ k(n)} \leq \frac{1}{\sqrt{n+1}}\left(A\ \left(1+\frac{\log{(n+1)}}{2\lambda_2}\right)+\lambda_2\ e^{\lambda_2}\right)$$

For $B>1$, we choose the parameters k,n such that

$$k = k(n) := \lfloor \frac{\log{(n+1)}}{2(\lambda_2+\log B)} \rfloor \geq k_0$$

Notice that $k(n)$ is the largest integer k such that

$$k \leq \frac{\log{(n+1)}}{2(\lambda_2+\log B)} \quad \left(\Leftrightarrow \frac{B^k}{\sqrt{n+1}} \leq e^{-\lambda_2\ k}\right)$$

Since $(k(n)+1) \geq \frac{\log{(n+1)}}{2(\lambda_2+\log B)}$, we have

$$\frac{B^{k(n)}}{\sqrt{n+1}} \leq e^{-\lambda_2\ k(n)} \leq e^{\lambda_2}\ e^{-\lambda_2\ \frac{\log{(n+1)}}{2(\lambda_2+\log B)}} = \frac{e^{\lambda_2}}{(n+1)^{\alpha/2}}$$

with $\alpha := \frac{\lambda_2}{(\lambda_2+\log B)}$, from which we conclude that

$$\frac{A}{\sqrt{n+1}}\ \frac{B^{k(n)+1}-1}{B-1} + \lambda_1\ e^{-\lambda_2\ k(n)} \leq \left[\frac{AB}{B-1}+\lambda_1\right]\ \frac{e^{\lambda_2}}{(n+1)^{\alpha/2}} - \frac{AB}{B-1}\ \frac{1}{\sqrt{n+1}}$$

This ends the proof of the theorem. ■

6 Path space models

In the previous section, we have established $\mathbb{L}_r$-mean error bounds and exponential estimates quantifying the convergence of the occupation measures $\eta_n^{(k)}$ towards the solutions $\pi_n^{(k)}$ of the measure-valued equation (1). We show here that it also possible to establish such results to quantify the convergence of the path-space occupation measures $\overline{\eta}_n^{[m]}$ introduced in (6) towards the tensor product measure $\overline{\pi}^{(m)}$ defined in (10).

6.1 $\mathbb{L}_r$-mean error bounds

Our main result is the following theorem.

Theorem 6.1 *For every $f \in \mathcal{B}(E_m)$, we have*

$$\sup_{n\geq 1} \sqrt{n}\ \mathbb{E}\left(\left|\left[\overline{\eta}_n^{[m]} - \overline{\pi}^{(m)}\right](f)\right|^r\right)^{\frac{1}{r}} < \infty$$

Proof:
To simplify the presentation, we fix a time horizon $m\geq 1$ and write ω instead of $\omega_{K_\eta^{(m)}}$, the invariant measure mapping defined in (9). We also write E instead of E_m, and $\overline{\eta}_n$ instead of $\overline{\eta}_n^{[m]}$. In this notation, $(\overline{\eta}^{(l)})$ represents the flow of the

occupation measures $\overline{\eta}_n^{(l)} := \frac{1}{n+1}\sum_{p=0}^n \delta_{X_p^{(l)}} \in \mathcal{P}(S^{(l)})$ of the i-MCMC model on the l-th level set $S^{(l)}$.

Using the fact that $\overline{\omega}^{m+1}(\eta) = \overline{\pi}^{(m)}$, we obtain the following decomposition for any $\eta \in \mathcal{P}(E)^{\mathbb{N}}$

$$\eta - \overline{\pi}^{[m]} = \sum_{k=0}^{m} \left[\overline{\omega}^k(\eta) - \overline{\omega}^{k+1}(\eta)\right] \tag{36}$$

Using proposition 4.3, the k-th iterate $\overline{\omega}^k$ of the mapping $\overline{\omega}$ can be rewritten for any $\eta \in \mathcal{P}(E)^{\mathbb{N}}$ in the following form

$$\overline{\omega}_n^k(\eta) = \frac{1}{n+1}\sum_{p=0}^{n} \left[\overline{\pi}^{[k-1]} \otimes \Pi_p^{(k,m)}((\eta^{(l)})_{0\leq l\leq m})\right]$$

Here the mappings

$$\Pi^{(k,m)} \ : \ \mu \in \prod_{0\leq i\leq m} \mathcal{P}(S^{(i)})^{\mathbb{N}} \mapsto \Pi^{(k,m)}(\mu) = \left(\Pi_n^{(k,m)}(\mu)\right)_{n\geq 0} \in \left(\bigotimes_{i=k}^{m} \mathcal{P}(S^{(i)})\right)^{\mathbb{N}}$$

are defined for any $n \geq 0$ by

$$\Pi_n^{(k,m)}(\mu) := \bigotimes_{i=0}^{m-k} \Pi_n^{(k,m),(i)}(\mu) \in \bigotimes_{i=0}^{m-k} \mathcal{P}(S^{(i+k)})$$

with for any $(\mu^{(l)})_{0\leq l\leq m} \in \prod_{0\leq i\leq m} \mathcal{P}(S^{(i)})^{\mathbb{N}}$ and any $0 \leq i \leq m-k$

$$\Pi_n^{(k,m),(i)}((\mu^{(l)})_l) := \Phi_{i+k}\left(\overline{\Phi}_n^{(i+(k-1),i+1)}(\mu^{(i)})\right) \in \mathcal{P}(S^{(i+k)})$$

We emphasize that $\Pi_n^{(k,m)}(\mu)$ only depends on the flow of measures $(\mu^{(l)})_{0\leq l\leq m-k}$, and

$$\begin{aligned}
\overline{\omega}_n^{k+1}(\eta) &= \frac{1}{n+1}\sum_{p=0}^{n} \left[\overline{\pi}^{[k]} \otimes \Pi_p^{(k+1,m)}((\eta^{(l)})_l)\right] \\
&= \frac{1}{n+1}\sum_{p=0}^{n} \left[\overline{\pi}^{[k-1]} \otimes \pi^{(k)} \otimes \bigotimes_{i=0}^{m-(k+1)} \Phi_{i+k+1}\left(\overline{\Phi}_p^{(i+k,i+2)}(\overline{\Phi}^{(i+1)}(\eta^{(i)}))\right)\right] \\
&= \frac{1}{n+1}\sum_{p=0}^{n} \left[\overline{\pi}^{[k-1]} \otimes \bigotimes_{i=0}^{m-k} \Phi_{i+k}\left(\overline{\Phi}_p^{(i+(k-1),i+1)}(\overline{\Phi}^{(i)}(\eta^{(i-1)}))\right)\right]
\end{aligned}$$

with the convention $\overline{\Phi}^{(0)}(\eta^{(-1)})) = \pi^{(0)}$, for $i = 0$. This implies that for any $0 \leq k \leq m$

$$\overline{\omega}_n^{k+1}(\eta) = \frac{1}{n+1}\sum_{p=0}^{n} \left[\overline{\pi}^{[k-1]} \otimes \Pi_p^{(k,m)}((\overline{\Phi}^{(l)}(\eta^{(l-1)}))_l)\right]$$

and therefore

$$\begin{aligned}
&\overline{\omega}_n^{k}(\eta) - \overline{\omega}_n^{k+1}(\eta) \\
&= \tfrac{1}{n+1}\textstyle\sum_{p=0}^{n} \left[\overline{\pi}^{[k-1]} \otimes \left\{\Pi_p^{(k,m)}((\eta^{(l)})_l) - \Pi_p^{(k,m)}\left(\left(\overline{\Phi}^{(l)}(\eta^{(l-1)})\right)_l\right)\right\}\right]
\end{aligned} \tag{37}$$

Moving one step further, we introduce the decomposition

$$\begin{aligned}
&\Pi^{(k,m)}(\mu) - \Pi^{(k,m)}(\nu) \\
&= \sum_{j=0}^{m-k} \left\{ \left(\bigotimes_{i=0}^{j-1} \Pi^{(k,m),(i)}(\nu) \right) \otimes \left[\Pi^{(k,m),(j)}(\mu) - \Pi^{(k,m),(j)}(\nu) \right] \right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad \left. \otimes \left(\bigotimes_{i=j+1}^{m-k} \Pi^{(k,m),(i)}(\mu) \right) \right\}
\end{aligned} \tag{38}$$

for any $\mu = (\mu^{(l)})_{0\leq l\leq m}$ and $\nu = (\nu^{(l)})_{0\leq l\leq m} \in \prod_{0\leq i\leq m} \mathcal{P}(S^{(i)})^{\mathbb{N}}$, with the flow of signed measures

$$\begin{aligned}
&\Pi_n^{(k,m),(j)}(\mu) - \Pi_n^{(k,m),(j)}(\nu) \\
&= \left[\Phi_{j+k}\left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\mu^{(j)}) \right) - \Phi_{j+k}\left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\nu^{(j)}) \right) \right]
\end{aligned}$$

For every $f \in \mathcal{B}(S^{(j+k)})$, we find that

$$\begin{aligned}
&\left| \left[\Pi_n^{(k,m),(j)}(\mu) - \Pi_n^{(k,m),(j)}(\nu) \right](f) \right| \\
&\leq \int \left| \left[\left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\mu^{(j)}) \right) - \left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\nu^{(j)}) \right) \right](g) \right| \; \Gamma_{j+k}(f,dg)
\end{aligned} \tag{39}$$

We let $\mathcal{F}_n^{m,j}$ be the sigma field given by

$$\mathcal{F}_n^{m,j} = \sigma\left(X_p^{(l)} \;:\; 0\leq p\leq n,\; 0\leq l\leq m,\; l\neq j \right)$$

Combining the generalized Minkowski integral inequality presented in lemma 5.1 with the inequality (35), we prove that

$$\begin{aligned}
&\mathbb{E}\left(\left| \left[\Pi_n^{(k,m),(j)}((\overline{\eta}^{(l)})_l) - \Pi_n^{(k,m),(j)}\left(\left(\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)}) \right)_l \right) \right](f) \right|^r \;\mid \mathcal{F}_n^{m,j} \right)^{\frac{1}{r}} \\
&\leq \int \mathbb{E}\left(\left| \left[\left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\overline{\eta}^{(j)}) \right) - \left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\overline{\Phi}^{(j)}(\overline{\eta}^{(j-1)})) \right) \right](g) \right|^r \;\mid \mathcal{F}_n^{m,j} \right)^{\frac{1}{r}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times \; \Gamma_{j+k}(f,dg) \\
&\leq \frac{e(r)}{\sqrt{n+1}} \; A \; B^k \; \|f\|
\end{aligned}$$

Notice that the decomposition (38) can be rewritten for any $f \in \mathcal{B}\left(\prod_{l=k}^m S^{(l)} \right)$ in the following form

$$\begin{aligned}
&\left[\Pi_n^{(k,m)}(\mu) - \Pi_n^{(k,m)}(\nu) \right](f) \\
&= \sum_{j=0}^{m-k} \left[\Pi_n^{(k,m),(j)}(\mu) - \Pi_n^{(k,m),(j)}(\nu) \right] \left(R_n^{(k,m),(j)}(\mu,\nu)(f) \right)
\end{aligned} \tag{40}$$

with the integral operators $R_n^{(k,m),(j)}(\mu,\nu) \; : \; \mathcal{B}\left(\prod_{l=k}^m S^{(l)}\right) \mapsto \mathcal{B}(S^{(j+k)})$ given below

$$R_n^{(k,m),(j)}(\mu,\nu)(f)(x_{k+j}) = \int \; f(x_k,\ldots,x_{k+(j-1)},x_{k+j},x_{k+j+1},\ldots,x_m)$$

$$\times \left(\prod_{i=0}^{j-1} \Pi_n^{(k,m),(i)}(\nu)\right)(dx_{i+k}) \times \left(\prod_{i=j+1}^{m-k} \Pi_n^{(k,m),(i)}(\mu)(dx_{i+k})\right)$$

Using the fact that the pair of measures

$$\bigotimes_{i=0}^{j-1} \Pi_n^{(k,m),(i)}\left(\left(\overline{\Phi}^{(l)}(\eta^{(l-1)})\right)_l\right) \quad \text{and} \quad \bigotimes_{i=j+1}^{m-k} \Pi_n^{(k,m),(i)}((\eta^{(l)})_l)$$

only depend on the distribution flow $\left(\overline{\Phi}^{(i)}(\eta^{(i-1)})\right)_{0\leq i\leq j-1}$ and $(\eta^{(i)})_{j+1\leq i\leq m-k}$, we find that the random functions

$$f_n^{(k,m),(j)} := R_n^{(k,m),(j)}\left((\overline{\eta}^{(l)})_l, \left(\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)})\right)_l\right)(f) \in \mathcal{B}(S^{(j+k)})$$

do not depend on the distribution flows $\eta^{(j)}$ and $\eta^{(j-1)}$. This shows that $f_n^{(k,m),(j)}$ are measurable with respect to $\mathcal{F}_n^{m,j}$. From previous calculations (and again using the generalized Minkowski integral inequality presented in lemma 5.1) we find that

$$\mathbb{E}\left(\left|\left[\Pi_n^{(k,m),(j)}((\overline{\eta}^{(l)})_l) - \Pi_n^{(k,m),(j)}\left(\left(\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)})\right)_l\right)\right](f_n^{(k,m),(j)})\right|^r \;\mid \mathcal{F}_n^{m,j}\right)^{\frac{1}{r}}$$

$$\leq \int \; \Gamma_{j+k}(f_n^{(k,m),(j)},dg)$$

$$\times \; \mathbb{E}\left(\left|\left[\left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\overline{\eta}^{(j)})\right) - \left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\overline{\Phi}^{(j)}(\overline{\eta}^{(j-1)}))\right)\right](g)\right|^r \;\mid \mathcal{F}_n^{m,j}\right)^{\frac{1}{r}}$$

$$\leq \frac{e(r)}{\sqrt{n+1}} \; A \; B^k \; \|f\|$$

We conclude that for any $f \in \mathcal{B}(\prod_{k\leq j\leq m} S^{(j)})$

$$\mathbb{E}\left(\left|\left[\Pi_n^{(k,m)}((\overline{\eta}^{(l)})_l) - \Pi_n^{(k,m)}\left(\left(\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)})\right)_l\right)\right](f)\right|^r\right)^{\frac{1}{r}}$$

$$\leq (m-k+1)\; \frac{e(r)}{\sqrt{n+1}} \; A \; B^k \; \|f\|$$

Using (40), it is now easily checked that for every $f \in \mathcal{B}(E)$

$$\mathbb{E}\left(\left|\left[\overline{\omega}_n^k(\overline{\eta}) - \overline{\omega}_n^{k+1}(\overline{\eta})\right](f)\right|^r\right)^{\frac{1}{r}} \leq (m-k+1)\; \frac{e(r)}{\sqrt{n+1}} \; A \; B^k \; \|f\|$$

Finally, by (36) we conclude that

$$\mathbb{E}\left(\left|\left[\overline{\eta}_n - \overline{\pi}^{[m]}\right](f)\right|^r\right)^{\frac{1}{r}} \;\leq \frac{e(r)}{\sqrt{n+1}} \; A \; \|f\| \; \sum_{k=0}^m (m-k+1) \; B^k$$

This ends the proof of the theorem. ∎

6.2 Concentration analysis

This section is mainly concerned with exponential bounds for the deviations of the occupation measures $\overline{\eta}_n^{[m]}$ around the limiting tensor product measure $\overline{\pi}^{[m]}$. We restrict our attention to models satisfying the Lipschitz type condition (7) for some integral operators Γ_k with uniformly finite support

$$\sup_{f\in\mathcal{B}(S^{(k)})} \mathrm{Card}\left(\mathrm{Supp}(\Gamma_k(f,.))\right) < \infty$$

To simplify the presentation, we fix a parameter $m \geq 1$, and sometimes we write $\overline{\eta}_n$ instead of $\overline{\eta}_n^{[m]}$. We shall also use the letters c_i, $i \geq 1$ to denote some finite constants whose values may vary from line to line but do not depend on the time parameter n.

The main result of this section is the following concentration theorem.

Theorem 6.2 *There exists some finite constant $\overline{\sigma}_m < \infty$ such that for any $f \in \mathcal{B}_1(E_m)$ and $t > 0$*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left|\left[\overline{\eta}_n^{[m]} - \overline{\pi}^{[m]}\right](f)\right| > t\right) < -\frac{t^2}{2\overline{\sigma}_m^2}$$

The proof of this theorem is based on two technical lemmas.

Lemma 6.3 *We let $M = (M_n)_{n\geq 1}$ be a random process such that the following exponential inequality is satisfied for some positive constants $a, b > 0$ and for any $t \geq 0$ and $n \geq 1$*

$$\mathbb{P}(|M_n| \geq t\ \sqrt{n}) \leq a\ e^{-bt^2}$$

We consider the collection of random processes $\overline{M}^{(k)} = (\overline{M}_n^{(k)})_{n\geq 1}$ defined for any $n \geq 0$ and $k \geq 0$ by the following formula

$$\overline{M}_{n+1}^{(k)} := (n+1) \int \Sigma^k(n,dp)\ \frac{1}{p+1}\ M_{p+1}$$

For every $k \geq 0$, $n \geq 1$, and $t \geq 0$ we have the exponential inequalities

$$\mathbb{P}\left(\left|\overline{M}_n^{(k)}\right| \geq t\ \sqrt{n}\right) \leq a\ n^k\ e^{-bt^2/2^{2k}}$$

Proof:
We prove the lemma by induction on the parameter k. For $k = 0$, we have $\overline{M}_{n+1}^{(0)} := M_{n+1}$ so that the exponential estimate holds true with $a(0) = a$ and $b(0) = b$. Suppose we have proved the result at rank k. Using the fact that

$$\begin{aligned}
\overline{M}_{n+1}^{(k+1)} &= (n+1) \int \Sigma^{k+1}(n,dp)\ \frac{1}{p+1}\ M_{p+1} \\
&= (n+1) \int \Sigma(n,dp)\ \ \frac{1}{p+1} \left((p+1)\int\ \Sigma^k(p,dq)\ \frac{1}{q+1}\ M_{q+1}\right)
\end{aligned}$$

we prove the recursion formula

$$\overline{M}_{n+1}^{(k+1)} = (n+1) \int \Sigma(n,dp)\ \frac{1}{p+1}\ \overline{M}_{p+1}^{(k)}$$

On the other hand, we have

$$\frac{1}{2}\,\frac{\overline{M}^{(k+1)}_{n+1}}{\sqrt{n+1}} = \frac{1}{2}\,\sqrt{n+1}\,\int\,\Sigma(n,dp)\,\frac{1}{\sqrt{p+1}}\,\frac{\overline{M}^{(k)}_{p+1}}{\sqrt{p+1}}$$

and

$$\begin{aligned}\frac{1}{2}\sqrt{n+1}\,\int\,\Sigma(n,dp)\,\frac{1}{\sqrt{p+1}} &= \frac{1}{2\sqrt{n+1}}\sum_{p=0}^{n}\,\frac{1}{\sqrt{p+1}}\\ &\leq \frac{1}{2\sqrt{n+1}}\sum_{p=0}^{n}\,\int_{p}^{p+1}\frac{1}{\sqrt{t}}\,dt = 1\end{aligned}$$

Under the induction hypothesis, we have for any $0 \leq p \leq n$

$$\mathbb{P}\left(\left|\overline{M}^{(k)}_{p+1}\right| \geq t\,\sqrt{p+1}\right) \leq a\,(n+1)^k\,e^{-bt^2/2^{2k}}$$

This implies that

$$\begin{aligned}\mathbb{P}\left(\frac{1}{2}\,\frac{\overline{M}^{(k+1)}_{n+1}}{\sqrt{n+1}} > t\right) &\leq \mathbb{P}\left(\exists 0 \leq p \leq n\ :\ \overline{M}^{(k)}_{p+1} > t\sqrt{p+1}\right)\\ &\leq a\,(n+1)\,(n+1)^k\,e^{-bt^2/2^{2k}}\end{aligned}$$

from which we conclude that

$$\mathbb{P}\left(\overline{M}^{(k+1)}_{n+1} > t\,\sqrt{n+1}\right) \leq a\,(n+1)^{k+1}\,e^{-bt^2/2^{2(k+1)}}$$

This ends the proof of the lemma. ■

Lemma 6.4 *For every $l_1 < l_2$, there exists some non increasing function*

$$N\ :\ t \in [0,\infty) \mapsto N(t) \in [0,\infty)$$

such that for every $n \geq N(t)$ and any function $f \in \mathcal{B}_1(S^{(l_2)})$ we have

$$\mathbb{P}\left(\sqrt{n+1}\,\left|\left[\overline{\Phi}^{(l_2,l_1+1)}_n(\overline{\eta}^{(l_1)}) - \overline{\Phi}^{(l_2,l_1)}_n(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))\right](f)\right| > t\right)$$

$$\leq (c_1(n+1))^{(l_2-l_1)}\,\exp\left(-c_2t^2/c_3^{l_2-l_1}\right)$$

Before getting into the details of the proof of this lemma, it is interesting to mention a direct consequence of the above exponential estimates. Firstly, we observe that $N(t\sqrt{n+1}) \leq N(t)$ so that for any $t > 0$ and $n \geq N(t)$ we have

$$\mathbb{P}\left(\left|\left[\overline{\Phi}^{(l_2,l_1+1)}_n(\overline{\eta}^{(l_1)}) - \overline{\Phi}^{(l_2,l_1)}_n(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))\right](f)\right| > t\right)$$

$$\leq (c_1(n+1))^{(l_2-l_1)}\,\exp\left(-c_2(n+1)t^2/c_3^{l_2-l_1}\right)$$

Using the decomposition

$$\eta_n^{(k)} - \pi^{(k)} = \sum_{l=0}^{k} \left[\overline{\Phi}_n^{(k,l+1)}(\eta^{(l)}) - \overline{\Phi}_n^{(k,l+1)}(\overline{\Phi}^{(l)}(\eta^{(l-1)})) \right]$$

we prove the following inclusion of events

$$\left\{ \left| [\overline{\eta}_n^{(k)} - \pi^{(k)}](f) \right| > t \right\}$$

$$\subset \left\{ \exists 0 \leq l \leq k \ : \ \left| \left[\overline{\Phi}_n^{(k,l+1)}(\eta^{(l)}) - \overline{\Phi}_n^{(k,l+1)}(\overline{\Phi}^{(l)}(\eta^{(l-1)})) \right] (f) \right| > t/(k+1) \right\}$$

By lemma 6.4 we can find a sufficiently large integer $N(t)$ that may depend on the parameter k and such that for every $n \geq N(t)$

$$\mathbb{P}\left(\left| [\overline{\eta}_n^{(k)} - \pi^{(k)}](f) \right| > t \right)$$

$$\leq \textstyle\sum_{0 \leq l \leq k} \ \mathbb{P}\left(\left| \left[\overline{\Phi}_n^{(k,l+1)}(\overline{\eta}^{(l)}) - \overline{\Phi}_n^{(k,l)}(\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)})) \right] (f) \right| > \frac{t}{k+1} \right)$$

$$\leq (k+1) \ (c_1(n+1))^k \ e^{-(n+1)t^2 c_2/((k+1)^2 c_3^k)}$$

This clearly implies the existence of some finite constant $\sigma_k < \infty$ such that

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left(\left| [\overline{\eta}_n^{(k)} - \pi^{(k)}](f) \right| > t \right) < -\frac{t^2}{2\sigma_k^2}$$

Proof of lemma 6.4 :
Using lemma 4.1, we find that

$$\left| \left[\overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\eta}^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1)}(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)})) \right] (f) \right|$$

$$\leq \int \ \left| \left[\overline{\eta}_p^{(l_1)} - \overline{\Phi}_p^{(l_1)}(\overline{\eta}^{(l_1-1)}) \right] (g) \right| \ \overline{\Gamma}^{(l_2,l_1+1)}((n,f), d(p,g))$$

Arguing as in (27), we find that for any $g \in \mathcal{B}(S^{(l_1)})$, we have

$$\left| \left[\overline{\eta}_p^{(l_1)} - \overline{\Phi}_p^{(l_1)}(\overline{\eta}^{(l_1-1)}) \right] (g) \right| \leq \frac{\left| M_{p+1}^{(l_1)}(g) \right|}{p+1} + c_1 \ \frac{\log{(p+2)}}{p+2} \ \|g\| \qquad (41)$$

with a sub-Gaussian process $M_n^{(l_1)}(g)$ satisfying the following exponential inequality for any $t > 0$ and any time parameter $n \geq 1$

$$\mathbb{P}(\left| M_n^{(l_1)}(g) \right| \geq t\sqrt{n}) \leq 2 \ \exp\left(-c_2 t^2/\|g\|^2 \right)$$

We notice that

$$\frac{1}{n+2} \sum_{p=0}^{n} \frac{(\log{(p+2)})^k}{p+2} \leq \frac{(\log{(n+2)})^k}{n+2} \sum_{p=0}^{n} \frac{1}{p+2}$$

$$\leq \frac{(\log{(n+2)})^k}{n+2} \sum_{p=0}^{n} \int_{p+1}^{p+2} \frac{1}{t} \ dt = \frac{(\log{(n+2)})^{k+1}}{n+2}$$

This implies that

$$\int \Sigma(n,dp)\ \frac{\log(p+2)}{p+2} \leq 2\ \frac{(\log(p+2))^2}{p+2}$$

More generally for any $k \geq 0$, we have that

$$\int \Sigma^k(n,dp)\ \frac{\log(p+2)}{p+2} \leq 2^k\ \frac{(\log(n+2))^{k+1}}{n+2}$$

from which we prove that

$$\begin{aligned}
&\int \frac{\log(p+2)}{p+2}\ \|g\|\ \overline{\Gamma}^{(l_2,l_1+1)}((n,f),d(p,g)) \\
&\leq 2^{(l_2-l_1)}\ \frac{(\log(n+2))^{(l_2-l_1)+1}}{n+2}\ \int \|g\|\ \Gamma_{l_2,l_1+1}(f,dg) \\
&\leq 2^{(l_2-l_1)}\ \frac{(\log(n+2))^{(l_2-l_1)+1}}{n+2}\ \left(\prod_{l_1<i\leq l_2} \Lambda_i\right) \leq c_3^{(l_2-l_1)}\ \frac{(\log(n+2))^{(l_2-l_1)+1}}{n+2}
\end{aligned} \tag{42}$$

For any $g \in \mathcal{B}(S^{(l_1)})$ we set

$$\overline{\mathcal{M}}_{n+1}^{(l_1,l_2)}(g) := \int \Sigma^{(l_2-l_1)}(n,dp)\ \frac{\left|M_{p+1}^{(l_1)}(g)\right|}{p+1}$$

Using lemma 6.3, we prove that

$$\mathbb{P}\left(\overline{\mathcal{M}}_{n+1}^{(l_1,l_2)}(g) > t\right) \leq 2\ (n+1)^{(l_2-l_1)}\ \exp\left(-c_2(n+1)t^2/[2^{2(l_2-l_1)}\|g\|^2]\right)$$

We observe that

$$\int \frac{1}{p+1}\ \left|M_{p+1}^{(l_1)}(g)\right|\ \overline{\Gamma}^{(l_2,l_1+1)}((n,f),d(p,g)) = \int \overline{\mathcal{M}}_{n+1}^{(l_1,l_2)}(g)\ \Gamma_{l_2,l_1+1}(f,dg)$$

In addition, using (41) and (42) we find that

$$\begin{aligned}
&\left|\left[\overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\eta}^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1)}(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))\right](f)\right| \\
&\leq \int \overline{\mathcal{M}}_{n+1}^{(l_1,l_2)}(g)\ \Gamma_{l_2,l_1+1}(f,dg) + \epsilon_{l_1,l_2}(n)
\end{aligned} \tag{43}$$

with

$$\epsilon_{l_1,l_2}(n) := c_1 c_3^{(l_2-l_1)}\ \frac{(\log(n+2))^{(l_2-l_1)+1}}{n+2}$$

Using the inclusion of events

$$\begin{aligned}
&\left\{\int \overline{\mathcal{M}}_{n+1}^{(l_1,l_2)}(g)\ \Gamma_{l_2,l_1+1}(f,dg) > t\right\} \\
&\subset \left\{\exists g \in \mathrm{Supp}(\Gamma_{l_2,l_1+1}(f,.)) \quad \text{such that} \quad \overline{\mathcal{M}}_{n+1}^{(l_1,l_2)}(g) > t\|g\|\ /\ (\Lambda_{l_2,l_1+1})\right\}
\end{aligned}$$

we find that

$$\mathbb{P}\left(\int\ \overline{\mathcal{M}}^{(l_1,l_2)}_{n+1}(g)\ \Gamma_{l_2,l_1+1}(f,dg)>t\right)$$

$$\leq S_{l_2,l_1+1}(f)\ \mathbb{P}\left(\overline{\mathcal{M}}^{(l_1,l_2)}_{n+1}(g)>t\|g\|\ /\ (\Lambda_{l_2,l_1+1})\right)$$

Finally, under our assumptions we have

$$\begin{aligned} S_{l_2,l_1+1}(f) &= \mathrm{Card}\left(\mathrm{Supp}(\Gamma_{l_2,l_1+1}(f,\,.))\right) \\ &\leq \prod_{l_1+1\leq k\leq l_2}\ \sup_{f\in\mathcal{B}(S^{(k)})}\ \mathrm{Card}\left(\mathrm{Supp}(\Gamma_k(f,\,.))\right)\leq c_4^{(l_2-l_1)} \end{aligned}$$

from which we check that

$$\mathbb{P}\left(\int\ \overline{\mathcal{M}}^{(l_1,l_2)}_{n+1}(g)\ \Gamma_{l_2,l_1+1}(f,dg)>t\right)$$

$$\leq (c_5(n+1))^{(l_2-l_1)}\ \exp\left(-c_6(n+1)t^2/c_7^{(l_2-l_1)}\right)$$

Using (43), we conclude that

$$\mathbb{P}\left(\left|\left[\overline{\Phi}^{(l_2,l_1+1)}_n(\overline{\eta}^{(l_1)})-\overline{\Phi}^{(l_2,l_1)}_n(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))\right](f)\right|>t+\epsilon_{l_1,l_2}(n)\right)$$

$$\leq (c_5(n+1))^{(l_2-l_1)}\ \exp\left(-c_6(n+1)t^2/c_7^{(l_2-l_1)}\right)$$

To take the final step, we observe that

$$\mathbb{P}\left(\sqrt{n+1}\ \left|\left[\overline{\Phi}^{(l_2,l_1+1)}_n(\overline{\eta}^{(l_1)})-\overline{\Phi}^{(l_2,l_1)}_n(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))\right](f)\right|\right.$$

$$\left.>t+\sqrt{n+1}\ \epsilon_{l_1,l_2}(n)\right)$$

$$\leq \mathbb{P}\left(\left|\left[\overline{\Phi}^{(l_2,l_1+1)}_n(\overline{\eta}^{(l_1)})-\overline{\Phi}^{(l_2,l_1)}_n(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))\right](f)\right|>\tfrac{t}{\sqrt{n+1}}+\epsilon_{l_1,l_2}(n)\right)$$

We also notice that for any $t>0$ we can find some non increasing function $N(t)$ such that

$$\forall n\geq N(t)\qquad \sqrt{n+1}\ \epsilon_{l_1,l_2}(n)<t$$

This implies that for any $n\geq N(t)$ we have

$$\begin{aligned} &\mathbb{P}\left(\sqrt{n+1}\ \left|\left[\overline{\Phi}^{(l_2,l_1+1)}_n(\overline{\eta}^{(l_1)})-\overline{\Phi}^{(l_2,l_1)}_n(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))\right](f)\right|>2t\right) \\ &\leq (c_5(n+1))^{(l_2-l_1)}\ \exp\left(-c_6t^2/c_7^{(l_2-l_1)}\right) \end{aligned}$$

The end of the proof is now straightforward. ■

We are now in position to prove theorem 6.2.

Proof of theorem 6.2:

We use the same notation as we used in the proof of theorem 6.1. Using (39) we find that

$$\left|\left[\Pi_n^{(k,m),(j)}(\mu)-\Pi_n^{(k,m),(j)}(\nu)\right](f)\right|>t$$

$$\Longrightarrow \exists g\in \mathrm{Supp}(\Gamma_{j+k}(f,\boldsymbol{.}))\ :$$

$$\left|\left[\left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\mu^{(j)})\right)-\left(\overline{\Phi}_n^{(j+(k-1),j+1)}(\nu^{(j)})\right)\right](g)\right|>t\|g\|/\Lambda_{j+k}$$

Therefore, using lemma 6.4 we can find a non-increasing function $N(t)$ (that may depend on the parameter k), such that for every $n\geq N(t)$ and any $f\in\mathcal{B}_1(S^{(j+k)})$ we have

$$\mathbb{P}\left(\sqrt{n+1}\ \left|\left[\Pi_n^{(k,m),(j)}(\mu)-\Pi_n^{(k,m),(j)}(\nu)\right](f)\right|>t\right)$$

$$\leq (c_1(n+1))^{(k-1)}\ \exp\left(-c_2t^2/c_3^{(k-1)}\right)$$

In much the same way, by the decomposition (40) we find the following assertion

$$\left|\left[\Pi_n^{(k,m)}(\mu)-\Pi_n^{(k,m)}(\nu)\right](f)\right|>t$$

$$\Longrightarrow \exists 0\leq j\leq (m-k)\ :$$

$$\left|\left[\Pi_n^{(k,m),(j)}(\mu)-\Pi_n^{(k,m),(j)}(\nu)\right]\left(R_n^{(k,m),(j)}(\mu,\nu)(f)\right)\right|>t/(m-k+1)$$

Since $R_n^{(k,m),(j)}(\mu,\nu)$ maps $\mathcal{B}_1(\prod_{l=k}^m S^{(l)})$ into $\mathcal{B}_1(S^{(j+k)})$ we have for every parameter $n\geq N(t)$

$$\mathbb{P}\left(\sqrt{n+1}\ \left|\left[\Pi_n^{(k,m)}((\overline{\eta}^{(l)})_l)-\Pi_n^{(k,m)}((\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)}))_l)\right](f)\right|>t\right)$$

$$\leq (m-k+1)\ (c_1(n+1))^{k-1}\ \exp\left(-c_2t^2/((m-k+1)^2c_3^{k-1})\right)$$

In summary, we have proved that there exists some non-increasing function $N(t)$ that may depend on the parameter m such that for any $0\leq k\leq m$, any $f\in\mathcal{B}_1(E)$, and any $n\geq N(t)$ we have

$$\mathbb{P}\left(\sqrt{n+1}\ \left|\left[\overline{\pi}^{[k-1]}\otimes\left\{\Pi_n^{(k,m)}((\overline{\eta}^{(l)})_l)-\Pi_n^{(k,m)}((\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)}))_l)\right\}\right](f)\right|>t\right)$$

$$\leq (c_4(n+1))^m\ \exp\left(-c_5t^2/c_6^m\right)$$

Let $(U_n)_{n\geq 1}$ be a collection of $[0,1]$-valued random variables such that for any t there exists some non-increasing function $N(t)$, so that for $n\geq N(t)$

$$\mathbb{P}(\sqrt{n}\ U_n\geq t)\leq a\ n^{\alpha}\ e^{-t^2b}$$

for some integer $\alpha\geq 1$ and some pair of positive constants (a,b). In this situation, we can find a non-increasing function $N'(t)$ and a pair of positive constants (a',b') such that

$$\forall n\geq N'(t)\qquad \mathbb{P}\left(\sum_{p=1}^n U_p>\sqrt{n}\ t\right)\leq a'\ n^{\alpha+1}\ e^{-t^2b'}$$

To prove this claim, we simply use the fact that for any $n \geq N(t)$ we have

$$\frac{1}{\sqrt{n}} \sum_{p=1}^{n} U_p \leq \frac{N(t)}{\sqrt{n}} + \frac{1}{\sqrt{n}} \sum_{p=N(t)}^{n} \frac{1}{\sqrt{p}} \left(\sqrt{p} U_p\right) \quad \text{and} \quad \frac{1}{2\sqrt{n}} \sum_{p=1}^{n} \frac{1}{\sqrt{p}} \leq 1$$

This yields that for any $n \geq N(t)$

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{p=1}^{n} U_p > t + \frac{N(t)}{\sqrt{n}}\right) \leq \sum_{p=N(t)}^{n} \mathbb{P}\left(\sqrt{p} U_p > t/2\right)$$

We let $N'(t)$ be the smallest integer n such that $N(t)/\sqrt{n} \leq t$. Recalling that $N(t)$ is a non-decreasing function, we find that for any $s \geq t$

$$N(t)/\sqrt{n} \leq t \Longrightarrow N(s)/\sqrt{n} \leq N(t)/\sqrt{n} \leq t \leq s \Longrightarrow N(s)/\sqrt{n} \leq s$$

This implies that $N'(s) \leq N'(t)$. Thus, we have constructed a non-increasing function $N'(t)$ such that for any $n \geq N'(t)$

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{p=1}^{n} U_p > 2t\right) \leq a\ n^{\alpha+1}\ e^{-t^2 b/4}$$

This ends the proof of the assertion with $(a', b') = (a, b/2^4)$. Applying this property to the decomposition (37), we can find a non-increasing function $N(t)$ such that for any $n \geq N(t)$ and any $0 \leq k \leq m$

$$\mathbb{P}\left(\sqrt{n+1}\ \left|\left[\overline{\omega}_n^k(\eta) - \overline{\omega}_n^{k+1}(\eta)\right](f)\right| > t\right) \leq (c_7(n+1))^{m+1}\ \exp\left(-c_8 t^2/c_9^m\right)$$

The end of the proof of the theorem is now a direct consequence of the decomposition (36). ■

7 Feynman-Kac semigroups

In section 5.3, we established a uniform convergence theorem under the assumption that the time averaged semigroup $\overline{\Phi}^{(k,l)}$ introduced in section 4.1 is exponentially stable; that is it satisfies (31). In this section, we study the mappings $\overline{\Phi}^{(k,l)}$ associated with the Feynman-Kac transformations discussed in (45). In particular, we provide necessary conditions ensuring that (31) is satisfied in this case.

7.1 Description of the models

To precisely describe these mappings we need a few definitions.

Definition 7.1 *We denote by Ψ_l^G the Boltzman-Gibbs transformation associated with a positive potential function G on $S^{(l)}$, and defined for any $f \in \mathcal{B}(S^{(l)})$ by the following formula*

$$\Psi_l^G(\eta_p)(f) = \eta_p(Gf)/\eta_p(G)$$

We let Q_l be the integral operator from $S^{(l-1)}$ into $S^{(l)}$ given by

$$\forall f \in \mathcal{B}(S^{(l)}) \qquad Q_l(f) := G_{l-1} \times L_l(f) \in \mathcal{B}(S^{(l-1)}) \tag{44}$$

By definition of the mappings Φ_l given in (11), it is easy to check that

$$\overline{\Phi}^{(l)}(\eta) = \overline{\Psi}^{(l),Q_l(1)}(\eta)L_l \quad \text{with} \quad \forall n \geq 0 \quad \overline{\Psi}_n^{(l),Q_l(1)}(\eta) = \frac{1}{n+1}\sum_{p=0}^{n} \Psi_l^{Q_l(1)}(\eta_p) \tag{45}$$

Definition 7.2 *We let $\overline{\Phi}^{(k,l)}$ be the semigroup associated with the Feynman-Kac transformations Φ_l discussed in (45), and we denote by*

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

the semigroup associated with the integral operator Q_l introduced in (44).

Proposition 7.3 *For any $l \leq k$ we have that*

$$\overline{\Phi}^{(k,l)}(\eta) = \overline{\Psi}^{(k,l)}(\eta)P_{l,k} \quad \text{with} \quad P_{l,k}(f) = \frac{Q_{l,k}(f)}{Q_{l,k}(1)} \tag{46}$$

and the mapping $\overline{\Psi}^{(k,l)}$ from $\mathcal{P}(S^{(l-1)})^{\mathbb{N}}$ into itself given below

$$\begin{aligned}\overline{\Psi}^{(k,l)} &= \overline{\Psi}^{(l),H_{l,k}} \circ \overline{\Psi}^{(k-1,l)} \\ &= \overline{\Psi}^{(l),H_{l,k}} \circ \overline{\Psi}^{(l),H_{l,k-1}} \circ \ldots \circ \overline{\Psi}^{(l),H_{l,l}} \quad \text{with} \quad H_{l,k} := \frac{Q_{l,k}(1)}{Q_{l,k-1}(1)}\end{aligned}$$

For $l = k$, we use the conventions $\overline{\Psi}^{(k-1,l)} = \overline{\Psi}^{(l-1,l)} = Id$ and $Q_{l,k-1}(1) = Q_{l,l-1}(1) = 1$, so that $H_{l,l} = Q_{l,l}(1) = Q_l(1)$ and $\overline{\Psi}^{(l,l)} = \overline{\Psi}^{(l),Q_l(1)}$.

Proof:
We prove the proposition by induction on the parameter $m = (k-l)$. For $k = l$, we clearly have

$$P_{l,l}(f) = \frac{Q_l(f)}{Q_l(1)} = L_l(f) \quad \text{and} \quad \overline{\Psi}^{(l,l)} = \overline{\Psi}^{(l),Q_l(1)} \Longrightarrow \overline{\Phi}^{(l)}(\eta) = \overline{\Psi}^{(l,l)}(\eta)P_{l,l}$$

Suppose we have proved formula (46) for some $m = (k-l) \geq 0$. To check the result at level $m+1 = (k-l)+1 = ((k+1)-l)$, we first observe that

$$\overline{\Phi}^{(k+1)}\left(\overline{\Phi}^{(k,l)}(\eta)\right) = \overline{\Psi}^{(k+1),Q_{k+1}(1)}(\overline{\Phi}^{(k,l)}(\eta))P_{k+1,k+1}$$

For any $\mu \in \mathcal{P}(S^{(k)})$, we also have that

$$\overline{\Psi}_n^{(k+1),Q_{k+1}(1)}(\mu)(P_{k+1}(f)) = \frac{1}{n+1}\sum_{p=0}^{n} \frac{\mu_p(Q_{k+1}(f))}{\mu_p(Q_{k+1}(1))}$$

so that

$$\overline{\Psi}_n^{(k+1),Q_{k+1}(1)}(\overline{\Phi}^{(k,l)}(\eta))P_{k+1,k+1} = \frac{1}{n+1}\sum_{p=0}^{n} \frac{\overline{\Phi}_p^{(k,l)}(\eta)(Q_{k+1}(f))}{\overline{\Phi}_p^{(k,l)}(\eta)(Q_{k+1}(1))}$$

Using the induction hypothesis, we find that

$$\overline{\Phi}_p^{(k,l)}(\eta)(Q_{k+1}(f)) = \overline{\Psi}^{(k,l)}(\eta)[P_{l,k}(Q_{k+1}(f))]$$

We also have

$$P_{l,k}(Q_{k+1}(f)) = \frac{Q_{l,k+1}(1)}{Q_{l,k}(1)}\ P_{l,k+1}(f) = H_{l,k+1}\ P_{l,k+1}(f)$$

from which we prove that

$$\overline{\Psi}^{(k,l)}(\eta)[P_{l,k}(Q_{k+1}(f))] = \overline{\Psi}^{(k,l)}(\eta)[H_{l,k+1}\ P_{l,k+1}(f)]$$

This clearly yields that

$$\begin{aligned}\frac{\overline{\Phi}_p^{(k,l)}(\eta)(Q_{k+1}(f))}{\overline{\Phi}_p^{(k,l)}(\eta)(Q_{k+1}(1))} &= \frac{\overline{\Psi}_p^{(k,l)}(\eta)[H_{l,k+1}\ P_{l,k+1}(f)]}{\overline{\Psi}_p^{(k,l)}(\eta)[H_{l,k+1}]} \\ &= \Psi_l^{H_{l,k+1}}\left(\overline{\Psi}_p^{(k,l)}(\eta)\right) P_{l,k+1}(f)\end{aligned}$$

and therefore

$$\begin{aligned}\overline{\Psi}_n^{(k+1),Q_{k+1}(1)}(\overline{\Phi}^{(k,l)}(\eta))P_{k+1,k+1} &= \frac{1}{n+1}\sum_{p=0}^{n}\Psi_l^{H_{l,k+1}}\left(\overline{\Psi}_p^{(k,l)}(\eta)\right) P_{l,k+1}(f) \\ &= \overline{\Psi}_n^{(l),H_{l,k+1}}\left(\overline{\Psi}^{(k,l)}(\eta)\right) P_{l,k+1}(f)\end{aligned}$$

In summary, we have proved that

$$\overline{\Phi}^{(k+1,l)}(\eta) = \overline{\Psi}^{(k+1,l)}(\eta)P_{l,k+1}(f) \quad \text{with} \quad \overline{\Psi}^{(k+1,l)}(\eta) = \overline{\Psi}_n^{(l),H_{l,k+1}}\left(\overline{\Psi}^{(k,l)}(\eta)\right)$$

This ends the proof of the proposition. ∎

7.2 Contraction inequalities

Proposition 7.4 *For any $l \leq k$ we have*

$$\beta(P_{l,k}) = \frac{1}{2}\ \sup_{\eta,\mu}\|\overline{\Phi}^{(k,l)}(\eta) - \overline{\Phi}^{(k,l)}(\mu)\|$$

Proof:
Using proposition 7.3, we find that

$$\begin{aligned}\|\overline{\Phi}^{(k,l)}(\eta) - \overline{\Phi}^{(k,l)}(\mu)\| &= \|\left[\overline{\Psi}^{(k,l)}(\eta) - \overline{\Psi}^{(k,l)}(\mu)\right] P_{l,k}\| \\ &\leq \beta(P_{l,k})\ \|\overline{\Psi}^{(k,l)}(\eta) - \overline{\Psi}^{(k,l)}(\mu)\|\end{aligned}$$

This implies that

$$\sup_{\eta,\mu}\|\overline{\Phi}^{(k,l)}(\eta) - \overline{\Phi}^{(k,l)}(\mu)\| \leq 2\ \beta(P_{l,k})$$

On the other hand, if we chose the constant Dirac distribution flows $\eta = (\eta_n)_{n\geq 0}$ and $\mu = (\mu_n)_{n\geq 0}$ given by

$$\forall n \geq 0 \quad \eta_n = \delta_x \quad \text{and} \quad \mu_n = \delta_y$$

for some $x, y \in S^{(l-1)}$, we also have that

$$\overline{\Phi}^{(k,l)}(\delta_x) - \overline{\Phi}^{(k,l)}(\delta_y) = \delta_x P_{l,k} - \delta_y P_{l,k}$$

This implies that

$$\sup_{\eta,\mu} \|\overline{\Phi}^{(k,l)}(\eta) - \overline{\Phi}^{(k,l)}(\mu)\| \geq \sup_{x,y} \|\delta_x P_{l,k} - \delta_y P_{l,k}\| = 2\ \beta(P_{l,k})$$

This ends the proof of the proposition. ■

Our next objective is to estimate the contraction coefficient $\beta(P_{l,k})$ in terms of the mixing type properties of the semigroup $L_{l,k} = L_l L_{l-1} \ldots L_k$ associated with the Markov operators L_l. We introduce the following regularity conditions.

$(L)_m$ *There exists an integer $m \geq 1$ and a sequence $(\epsilon_l(L))_{l\geq 0} \in (0,1)^{\mathbb{N}}$ such that*

$$\forall l \geq 0 \quad \forall (x,y) \in (S^{(l-1)})^2 \qquad L_{l+1,l+m}(x, \centerdot) \geq \epsilon_l(L) \quad L_{l+1,l+m}(y, \centerdot)$$

It is well known that the above condition is satisfied for any aperiodic and irreducible Markov chain on a finite space. Loosely speaking, for non-compact spaces this condition is related to the tails of the transition distributions on the boundaries of the state space. For instance, let us suppose that $S^{(l)} = \mathbb{R}$ and L_l is the bi-Laplace transition given by

$$L_l(x, dy) = \frac{c(l)}{2}\ e^{-c(l)\,|y-A_l(x)|}\ dy$$

for some $c(l) > 0$ and some drift function A_n with bounded oscillations $\mathrm{osc}(A_l) < \infty$. In this case, it is readily checked that condition $(L)_m$ holds true for $m = 1$ with the parameter

$$\epsilon_{l-1}(L) = \exp\left(-c(l)\ \mathrm{osc}(A_l)\right)$$

Under the condition (G) presented on page 11 and the mixing condition $(L)_m$ stated above, we proved in [3] (see corollary 4.3.3 on page 141) that we have for any $k \geq m \geq 1$, and $l \geq 1$

$$\beta(P_{l+1,l+k}) \leq \prod_{i=0}^{\lfloor k/m \rfloor - 1} \left(1 - \epsilon_{l+im}^{(m)}\right) \quad \text{with} \quad \epsilon_l^{(m)} := \epsilon_l^2(L) \prod_{l+1\leq k<l+m} \epsilon_k(G)$$

Several contraction inequalities can be deduced from these estimates, we refer to chapter 4 of the book [3]. To give a flavor of these results, we further assume that $(M)_m$ is satisfied with $m = 1$, and we have $\epsilon(L) = \inf_l \epsilon_l(L) > 0$. In this case, we can check that

$$\beta(P_{l+1,l+k}) \leq \left(1 - \epsilon(L)^2\right)^k$$

References

[1] C. Andrieu, A. Jasra, A. Doucet and P. Del Moral. *Non-Linear Markov chain Monte Carlo via self-interacting approximations.* Technical report, Department of Mathematics, Bristol University (2007).

[2] A. Brockwell, P. Del Moral and A. Doucet. *Sequentially interacting Markov chain Monte Carlo.* Technical report, Department of Statistics, University of British Columbia (2007).

[3] P. Del Moral, *Feynman-Kac formulae. Genealogical and interacting particle systems, with applications*, Springer Verlag New York, Series: Probability and its Applications (2004).

[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 (eds.). *Sequential Monte Carlo Methods in Practice*, Springer Verlag New York, Series: Statistics for Engineering and Information Science (2001).

[7] E. Lyman and D.M. Zuckerman. *Resolution exchange simulation with incremental coarsening.* J. Chem. Theory Comp., vol. 2, pp. 656-666 (2006).

[8] K.L. Mengersen and R. Tweedie. *Rates of convergence of the Hastings and Metropolis algorithms.* Annals of Statistics, vol. 24, pp. 101-121 (1996).

[9] C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*, Springer Verlag New York (2004).

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

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

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