The Annals of Applied Probability
2010, Vol. 20, No. 2, 593–639
DOI: 10.1214/09-AAP628

INTERACTING MARKOV CHAIN MONTE CARLO METHODS FOR SOLVING NONLINEAR MEASURE-VALUED EQUATIONS[1]

BY PIERRE DEL MORAL AND ARNAUD DOUCET

INRIA and Université Bordeaux and University of British Columbia and The Institute of Statistical Mathematics

We present a new class of interacting Markov chain Monte Carlo algorithms 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 evolutions depend on the occupation measure of their 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 behavior of these iterative algorithms which relies on measure-valued processes and semigroup techniques. We establish a variety of convergence results including exponential estimates and a uniform convergence theorem with respect to the number of target distributions. We also illustrate these algorithms in the context of Feynman–Kac distribution flows.

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

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

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 an analytic 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

Received February 2008; revised July 2009.

[1]Supported in part by the French national programs "Sécurité et Informatique" and "Systèmes Complexes et Modélisation Mathématique" under projects NEBBIANO ANR-06-SETI-009 and VIROSCOPY ANR-08-SYSC-016.

AMS 2000 subject classifications. Primary 47H20, 60G35, 60J85, 62G09; secondary 47D08, 47G10, 62L20.

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

$(\pi^{(l)})_{l\geq 0}$. 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 [5].

In recent years, there has been considerable interest in the development of interacting particle interpretations of measure-valued equations of the form (1.1) which we briefly review here.

1.2. *Interacting particle methods.* The central idea of interacting particle methods 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 measure $\pi_N^{(l)} := \frac{1}{N}\sum_{p=1}^N \delta_{X_p^{(l)}}$ approximates $\pi^{(l)}$ 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:

$$\mathbb{P}\big((X_1^{(l)}, \dots, X_N^{(l)}) \in dx | X^{(l-1)}\big) = \prod_{p=1}^{N} \Phi_l\bigg(\frac{1}{N}\sum_{1\leq q\leq N} \delta_{X_q^{(l-1)}}\bigg)(dx_p), \tag{1.2}$$

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

For Feynman–Kac transformations, these interacting particle models have been extensively studied and they are sometimes referred to as sequential Monte Carlo methods, particle filters and population Monte Carlo methods; see [5, 8] for a review of the literature. In this context, the convergence analysis of these particle algorithms is now well understood. A variety of theoretical results are available, including sharp propagations of chaos properties, fluctuations and large deviations theorems, as well as uniform convergence results with respect to the level index l.

These interacting particle methods suffer from two serious limitations. First, when the mapping Φ_l is complex, it may be impossible to generate independent draws from it. Second, it is typically impossible to determine beforehand the number of particles necessary to achieve a fixed precision for a given application and users usually have to perform multiple runs for an increasing number of particles until stabilization of the Monte Carlo estimates is observed. Markov chain Monte Carlo (MCMC) methods appear as a natural way to solve these two problems [12]. However, standard MCMC methods do not apply in this context as we have a sequence of target distributions defined on different spaces and the normalizing constants of these distributions are typically unknown.

1.3. *Self-interacting Markov chains*. We propose here a new class of interacting MCMC 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 sequence of samples. Moreover, in contrast to interacting particle methods, 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 [6, 7] presented by the first author in collaboration with Laurent Miclo. These studies are concerned with biology-inspired self-interacting Markov chain (SIMC) models with applications to genetic type algorithms involving a competition between a reinforcement mechanism and a potential function [6, 7]. These ideas 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 [4]. Related ideas have also appeared in computational chemistry [10] and statistics [9].

In the present article, we design a new general class of i-MCMC methods. Roughly speaking, these algorithms proceed as follows. At level $l = 0$ we run an MCMC algorithm to obtain a chain $X^{(0)} = (X_n^{(0)})_{n\geq 0}$ targeting $\pi^{(0)}$. Note that here the "time" index n corresponds to the number of iterations of the i-MCMC algorithm. We use the occupation measure of the chain $X^{(0)}$ at time n judiciously to design a second MCMC algorithm to generate $X^{(1)} = (X_n^{(1)})_{n\geq 0}$ at level 1 targeting $\pi^{(1)}$ which is typically more complex than $\pi^{(0)}$. More precisely, 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 $(X_0^{(0)}, X_1^{(0)}, \dots, X_n^{(0)})$. Similarly we use the empirical measure of $X^{(l-1)}$ at level $l-1$ to "feed" an MCMC algorithm generating $X^{(l)}$ targeting $\pi^{(l)}$ at level l. These i-MCMC samplers are 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 time 0 up to time n.

From the pure mathematical point of view, the convergence analysis of SIMC 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 SIMC is much more involved than the one of traditional Markov chains. It also differs significantly from interacting particle methods developed in [5]. 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 index n tends to infinity.

The rest of the paper is organized as follows:

The main notation used in this work are introduced in a brief preliminary Section 1.4. The i-MCMC methodology is detailed formally in Section 1.5. The main results of the article are presented in Section 1.6. Several examples of i-MCMC methods are provided in Section 2. This section also provides a discussion on how to combine interacting particle methods with i-MCMC methods. 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 sequence of time averaged semigroups on distribution flow state spaces. The asymptotic analysis of i-MCMC methods is 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 algorithms 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 sequence of target measures. We prove a uniform convergence theorem with respect to the level index l. 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 Sections 6.1 and in 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 used in the article. We also recall some 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 $\|f\| \le 1$, and by $\mathrm{Osc}_1(E)$, the convex set of $\mathcal{E}$-measurable functions f with oscillations less than one; that is,

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

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 $A \in \mathcal{E}$.

Let $M(x, dy)$ be a kernel from a measurable space $(E, \mathcal{E})$ into a measurable space $(F, \mathcal{F})$ of the bounded integral 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 $f \in \mathcal{B}(F)$. Such a kernel 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 $\|\mu\| = \sup_{f \in \mathcal{B}_1(E)} |\mu(f)|$. Using 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 kernel 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 kernel 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

$$\begin{aligned}
&\|\mu M\| = 2 \sup_{f \in \mathrm{Osc}_1(F)} |\mu M(f)| \leq \beta(M) 2 \sup_{f \in \mathrm{Osc}_1(E)} |\mu(f)| \\
&\quad \Longrightarrow \quad \|\mu M\| \leq \beta(M) \|\mu\|.
\end{aligned}$$

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 kernel

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

That is, we have

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

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

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

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 $\sum_\varnothing = 0$ and $\prod_\varnothing = 1$.

1.5. *Interacting Markov chain Monte Carlo methods*. We describe here the i-MCMC methodology to numerically solve (1.1). We consider a Markov transition $M^{(0)}$ from $S^{(0)}$ into itself and a collection of Markov transitions $M_\mu^{(l)}$ from $S^{(l)}$ into itself, indexed by the parameter $l \ge 0$ and the set of probability measures $\mu \in \mathcal{P}(S^{(l-1)})$. We further assume that the invariant measure of each operator $M_\mu^{(l)}$ is given by $\Phi_l(\mu)$; that is, we have

$$\forall l \ge 0, \forall \mu \in \mathcal{P}(S^{(l-1)}) \qquad \Phi_l(\mu) M_\mu^{(l)} = \Phi_l(\mu).$$

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

We let $X^{(0)} := (X_n^{(0)})_{n\ge 0}$ be a Markov chain on $S^{(0)}$ with initial distribution ν_0 and Markov transitions $M^{(0)}$. For every $k \ge 1$, given a realization of the chain $X^{(k-1)} := (X_n^{(k-1)})_{n\ge 0}$, the kth level chain $X_n^{(k)}$ is a Markov chain with initial distribution ν_k and with random Markov transitions $M^{(k)}_{\eta_n^{(k-1)}}$ depending on the current occupation measures $\eta_n^{(k-1)}$ of the chain at level $(k-1)$; that is, we have

$$\mathbb{P}(X_{n+1}^{(k)} \in dx | X^{(k-1)}, X_n^{(k)}) = M^{(k)}_{\eta_n^{(k-1)}}(X_n^k, dx) \tag{1.3}$$

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 kth level chain $X_n^{(k)}$ behaves asymptotically as a Markov chain with time homogeneous transitions $M^{(k)}_{\pi^{(k-1)}}$ as long as $\eta_n^{(k-1)}$ is a good approximation of $\pi^{(k-1)}$.

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

$$\mathbb{P}((X_1^{(k)}, \dots, X_n^{(k)}) \in dx | X^{(k-1)}) = \prod_{p=1}^n \Phi_k\Big(\frac{1}{p} \sum_{0\le q<p} \delta_{X_q^{(k-1)}}\Big)(dx_p), \tag{1.4}$$

where $dx = d(x_1, \dots, x_n) = dx_1 \times \cdots \times dx_n$ stands for an infinitesimal neighborhood of a generic path sequence $(x_1, \dots, x_n) \in (S^{(k)})^n$.

We end this section with a SIMC interpretation of the stochastic algorithm discussed above. We consider the product space

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

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

$$\forall x := (x^0, \dots, x^m) \in E_m \qquad K_\eta^{(m)}(x, dy) = \prod_{0 \le l \le m} M_{\eta^{(l-1)}}^{(l)}(x^l, dy^l), \tag{1.5}$$

where $dy := dy^0 \times \cdots \times dy^m$ stands for for an infinitesimal neighborhood of a generic point $y := (y^0, \dots, y^m) \in E_m$, and $\eta^{(l)} \in \mathcal{P}(S^{(l)})$ stands for the image measure of a measure $\eta \in \mathcal{P}(E_m)$ on the lth level space $S^{(l)}$, with $m \ge l$. In other words, $\eta^{(l)}$ is the lth marginal of the measure η. In this notation, we can readily check that

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

is an E_m-valued SIMC with elementary transitions defined by

$$\mathbb{P}(\overline{X}_{n+1}^m \in dy | \mathcal{F}_n^{\overline{X}^m}) = K_{\overline{\eta}_n^{[m]}}^{(m)}(\overline{X}_n^m, dy) \qquad \text{with } \overline{\eta}_n^{[m]} = \frac{1}{n+1} \sum_{p=0}^n \delta_{\overline{X}_n^m}, \tag{1.6}$$

where $\mathcal{F}_n^{\overline{X}^m}$ stands for the filtration generated by $\overline{X}^m$.

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

$$\begin{aligned} &\forall l \ge 0, \forall f \in \mathcal{B}(S^{(l)}) \\ &|[\Phi_l(\mu) - \Phi_l(\nu)](f)| \le \int |[\mu - \nu](g)| \Gamma_l(f, dg) \end{aligned} \tag{1.7}$$

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

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

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

$$\|M_\mu^{(l)} - M_\nu^{(l)}\| \le c_l \|\mu - \nu\| \quad \text{and} \quad b_l(n_l) := \sup_{\mu \in \mathcal{P}(S^{(l-1)})} \beta\big((M_\mu^{(l)})^{n_l}\big) < 1. \tag{1.8}$$

This pair of abstract regularity conditions are rather standard. The first one (1.7) is a natural Lipschitz property on the weakly continuous integral mappings

$$\forall f \in \mathcal{B}(S^{(l)}) \qquad \mu \in \mathcal{P}(S^{(l-1)}) \mapsto \Phi_l(\mu)(f) \in \mathbb{R}.$$

Roughly speaking, this weak Lipschitz property simply expresses the fact that $\Phi_l(\mu)(f)$ only depends on integrals of functions with respect to the reference measure μ. This condition is clearly satisfied for linear Markov semigroups $\Phi_l(\mu) = \mu K_l$ associated with some Markov transition K_l. We shall discuss this condition in the context of nonlinear Feynman–Kac type semigroups (2.1) in Section 2.1.

In the special case where $M_\mu^{(l)}(x^l, \cdot) = \Phi_l(\mu)$, the second condition (1.8) is trivially met for $n_l = 1$ with $b_l(n_l) = 0$. In this particular situation, the first Lipschitz property of the mapping $\Phi_l(\mu)$ takes the following form:

$$\|\Phi_l(\mu) - \Phi_l(\nu)\| \le c_l \|\mu - \nu\|.$$

For more general models, condition (1.8) expresses the fact that the Markov transitions $M_\mu^{(l)}$ are strongly continuous and they satisfy Dobrushin's mixing condition, uniformly with respect to μ. We shall discuss this regularity condition in the context of Metropolis–Hastings type algorithms (2.7) in Section 2.2.

Under the conditions (1.8), 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 (1.5) is given by the tensor product measure

$$\omega_{K_\eta^{(m)}}(\eta) = \pi^{(0)} \otimes \Phi_1(\eta^{(0)}) \otimes \cdots \otimes \Phi_m(\eta^{(m-1)}). \tag{1.9}$$

We observe that the tensor product measure

$$\overline{\pi}^{[m]} := \pi^{(0)} \otimes \cdots \otimes \pi^{(m)} \tag{1.10}$$

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

Using this notation, our main results are basically as follows.

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

$$\sup_{n\ge 1} \sqrt{n}\, \mathbb{E}\big(\big|\overline{\eta}_n^{[m]}(f) - \overline{\pi}^{[m]}(f)\big|^r\big) < \infty.$$

Under some additional regularity conditions, we have the exponential inequality

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

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

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

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

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

First, 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 have recently studied these fluctuations in [2, 3].

It might be surprising that the decays to equilibrium presented in Theorem 1.1 differ from the three types of decays exhibited in [6, 7]. To understand the main differences between these classes of interacting processes, we recall that the decay rate 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 essentially constant after a sufficiently large number of iterations. In this respect, the self-interacting models discussed in the present article are more regular than the ones analyzed in [6, 7].

The uniform convergence estimate with respect to the number of levels depends on the stability properties of a time averaged semigroup associated with the mappings Φ_l. The contraction properties of this new class of nonlinear semigroups are studied in Section 7 in the context of Feynman–Kac models. 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:

$$\begin{aligned} &\forall l \geq 0, \forall (\mu, f) \in \big(\mathcal{P}\big(S^{(l)}\big) \times \mathcal{B}\big(S^{(l+1)}\big)\big) \\ &\Phi_{l+1}(\mu)(f) := \mu(G_l L_{l+1}(f))/\mu(G_l), \end{aligned} \tag{2.1}$$

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

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

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 very wide variety of applications including nonlinear filtering and rare event analysis as well the spectral analysis of Schroedinger type operators and directed polymer analysis [5]. 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 sequence of measures $(\pi^{(k)})_{0\leq k<l}$ with the following formulae:

$$\gamma^{(l)}(f)=\pi^{(l)}(f)\prod_{0\leq k<l}\pi^{(k)}(G_k). \tag{2.2}$$

To check this assertion, we simply observe that

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

and we have the key multiplicative formula

$$\begin{aligned}\gamma^{(l)}(1)=\gamma^{(l-1)}(G_{l-1})=\pi^{(l-1)}(G_{l-1})\times\gamma^{(l-1)}(1)\\ \Longrightarrow\quad \gamma^{(l)}(1)=\prod_{0\leq k<l}\pi^{(k)}(G_k).\end{aligned} \tag{2.3}$$

Thus the i-MCMC methodology allows us to estimate the normalizing constants $\gamma^{(l)}(1)$ by replacing the measures $\pi^{(k)}$ by their approximations in (2.3). These models are quite flexible. For instance, the reference Markov chain may represent the paths from the origin up to the current time l of an auxiliary chain Y'_l taking values in some state spaces E'_l with some Markov transitions $(\widetilde{L}_l)_{l\geq 1}$ and potentials $(\widetilde{G}_l)_{l\geq 1}$; that is, we have

$$Y_l:=(Y'_0,\ldots,Y'_l)\in S^{(l)}:=(E'_0\times\cdots\times E'_l) \tag{2.4}$$

and

$$\begin{aligned}L_l(y_{l-1},d\overline{y}_l)=\delta_{(y'_0,\ldots,y'_{l-1})}(d(\overline{y}'_0,\ldots,\overline{y}'_{l-1}))\widetilde{L}_l(y'_{l-1},d\overline{y}'_l),\\ G_l(y_l)=\widetilde{G}_l(y'_l).\end{aligned} \tag{2.5}$$

2.2. *Interacting Markov chain Monte Carlo methods for Feynman–Kac models*. In the Feynman–Kac context and assuming we are working on path spaces (2.4), we can propose the following two i-MCMC algorithms to approximate $\pi^{(l)}$. The first one simply consists of sampling directly $X_p^{(k)}=(X_p^{\prime(0)},X_p^{\prime(1)},\ldots,X_p^{\prime(k)})$ from the right-hand side product of the formula (1.4) which takes here the following form:

$$\Phi_k\biggl(\frac{1}{p}\sum_{0\leq q<p}\delta_{X_q^{(k-1)}}\biggr)(dx_p^{(k)})=\sum_{0\leq q<p}\frac{G_{k-1}(X_q^{(k-1)})}{\sum_{0\leq m<p}G_{k-1}(X_m^{(k-1)})}L_k\bigl(X_q^{(k-1)},dx_p^{(k)}\bigr),$$

where $dx_p^{(k)}=dx_p^{\prime(0)}\times\cdots\times dx_p^{\prime(k)}$. We see that $X_p^{(k)}$ is sampled according to two separate genetic type mechanisms. First, we randomly select one state $X_q^{(k-1)}$ at level $(k-1)$ with a probability proportional to its potential value $G_{k-1}(X_q^{(k-1)})$.

Second, we randomly evolve from this state according to the mutation transition L_k. This i-MCMC model can be interpreted as a spatial branching and interacting process. In this interpretation, the kth 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.

For the Feynman–Kac transformations (2.1), we proved in [5] that the condition (1.8) ensuring convergence of the algorithm is satisfied with $c_l = \beta(\widetilde{L}_l)/\varepsilon_{l-1}(G)$ as soon as the potential functions satisfy the following condition:

(G) *For any $l \geq 0$, the potential functions G_l are bounded above and bounded away from zero, so that*

$$\varepsilon_l(G) := \inf_{x,y} \frac{G_l(x)}{G_l(y)} \in (0, 1).$$

We can also propose the following alternative i-MCMC algorithm to approximate $\pi^{(l)}$ which relies on using a transition kernel $M_\mu^{(l)}$ different from $\Phi_l(\mu)$. We introduce the following kernel from $S^{(l-1)}$ into E'_l:

$$R_l((x'_0, \dots, x'_{l-1}), dx'_l) = \widetilde{L}_l(x'_{l-1}, dx'_l)\widetilde{G}_{l-1}(x'_{l-1}). \tag{2.6}$$

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 $\mu \in \mathcal{P}(S^{(l-1)})$

$$\begin{aligned} M_\mu^{(l)}(x, dy) &= (\mu \otimes K_l)(dy)\big(1 \wedge r_l(x, y)\big) \\ &\quad + \left(1 - \int_{S^{(l)}} \big(1 \wedge r_l(x, z)\big)(\mu \otimes K_l)(dz)\right)\delta_x(dy), \end{aligned} \tag{2.7}$$

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

$$r_l((u, v), (w, z)) := \frac{d(K_l(u, \cdot) \otimes R_l(w, \cdot))}{d(R_l(u, \cdot) \otimes K_l(w, \cdot))}(v, z), \tag{2.8}$$

where we assume that

$$K_l(u, \cdot) \otimes R_l(w, \cdot) \ll R_l(u, \cdot) \otimes K_l(w, \cdot).$$

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

We can also easily establish that for any 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 (1.8) is satisfied. Under the additional assumption that for any $(u, v) \in (S^{(l-1)} \times E'_l)$

$$\frac{dP_l(u, \cdot)}{dK_l(u, \cdot)}(v) \leq C_l$$

it follows from [11], Theorem 2.1, that

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

from which we conclude that the second condition on the right-hand side of (1.8) is met with $n_l = 1$ and $b_l(n_l) = (1 - C_l^{-1})$.

2.3. *Interacting particle and Markov chain Monte Carlo methods.* As mentioned in the Introduction, in contrast to interacting particle methods presented in Section 1.2, we emphasize that the precision parameter n of i-MCMC models is not fixed but increases at every time step. There exist several ways to combine an interacting particle method with an i-MCMC method.

For instance, suppose we are given a realization of an interacting particle 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 (1.3) 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 an interacting particle method and i-MCMC method 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

$$\|K_\eta - K_\mu\| \leq c\|\eta - \mu\| \quad \text{and} \quad b(n_0) := \sup_{\eta \in \mathcal{P}(E)} \beta(K_\eta^{n_0}) < 1. \tag{3.1}$$

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

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

where μ_n is a sequence of possibly random distributions on F that only depends on the random sequence $(X_0, \dots, X_n)$. 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 $\varepsilon(n)$ in the sense that

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

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

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

For SIMC, we have $F = E$ and the measure μ_n coincides with the occupation measures of the chain up to the current time n. In this particular situation, we have

$$\mu_n = \eta_n := \frac{1}{n+1} \sum_{p=0}^{n} \delta_{X_p} \quad \Longrightarrow \quad \varepsilon(n) \leq \frac{2}{n+2}. \tag{3.2}$$

This implies that

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

Under assumption (3.1), 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 $\varepsilon(n)$ 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 time $p = 0$ up to the current time $p = n$. That is, for large values of the time horizon n, we have in some sense

$$\eta_n \simeq \overline{\omega}_n(\mu) := \frac{1}{n+1} \sum_{p=0}^{n} \omega(\mu_p). \tag{3.3}$$

3.2. *A resolvent analysis.* We recall that assumption (3.1) 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

$$\alpha(\eta) := \sum_{n\geq 0} \beta(K_\eta^n) \in [1,\infty) \quad \text{and} \quad \sum_{n\geq 0}[K_\eta^n - \omega(\eta)](f) \tag{3.4}$$

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

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

is a well-defined solution of the Poisson equation

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

The reader should not be misled by the notation P_η. In this context, P_η is not a Markov transition kernel. We have used the letter P in reference to the solution of the Poisson equation.

PROPOSITION 3.1. *For any $\eta \in \mathcal{P}(F)$, P_η is a bounded integral operator on $\mathcal{B}(E)$ 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 $\beta(P_\eta) \leq \alpha(\eta)$ is readily deduced from the following decomposition:

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

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

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

In much the same way, we use the fact that

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

to check that

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

and

$$\|P_\eta(f)\| \le \Big[\sum_{n\ge 0}\beta(K_\eta^n)\Big]\operatorname{osc}(f) \quad\Rightarrow\quad \|P_\eta\| \le 2\sum_{n\ge 0}\beta(K_\eta^n).$$

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

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

Since we have

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

we conclude that $\alpha(\eta) \le n_0\sum_{p\ge 0}\beta(K_\eta^{n_0})^p = \frac{n_0}{1-\beta(K_\eta^{n_0})}$. The end of the proof of the proposition is now complete. □

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

$$\|\omega(\eta)-\omega(\mu)\| \le \delta_{n_0}(\eta,\mu)\|\eta-\mu\| \tag{3.5}$$

and

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

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

$$\delta_{n_0}(\eta,\mu) \le \frac{cn_0}{1-(\beta(K_\eta^{n_0})\wedge\beta(K_\mu^{n_0}))}. \tag{3.6}$$

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

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

we find that

$$\|\omega(\eta)-\omega(\mu)\| \le \frac{1}{1-(\beta(K_\mu^{n_0})\wedge\beta(K_\eta^{n_0}))}\|\omega(\eta)(K_\eta^{n_0}-K_\mu^{n_0})\|. \tag{3.7}$$

On the other hand, we have

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

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

$$\|K_\eta^{n_0} - K_\mu^{n_0}\| \leq cn_0\|\eta - \mu\| \quad \Longrightarrow \quad \|\omega(\eta)(K_\eta^{n_0} - K_\mu^{n_0})\| \leq cn_0\|\eta - \mu\|.$$

The proof of (3.5) is now a direct consequence of (3.7).

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

$$P_\mu(K_\mu - \mathrm{Id}) = (K_\mu - \mathrm{Id})P_\mu = (\omega(\mu) - \mathrm{Id}).$$

This yields

$$P_\mu(K_\mu - \mathrm{Id})P_\eta = (\omega(\mu) - \mathrm{Id})P_\eta.$$

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

$$P_\mu(K_\eta - \mathrm{Id})P_\eta = P_\mu(\omega(\eta) - \mathrm{Id}) = -P_\mu.$$

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. $\square$

3.3. $\mathbb{L}_r$*-inequalities and concentration analysis.* First, we examine some of the consequences of the pair of regularity conditions presented in (3.1). The second condition ensures that the functions $\alpha(\eta)$ and $\delta_{n_0}(\eta,\mu)$ introduced in (3.4) and (3.6) are uniformly bounded; that is, we have

$$1 \le a(n_0) := \sup_{\eta\in\mathcal{P}(F)} \alpha(\eta) \le \frac{n_0}{1-b(n_0)} \tag{3.8}$$

and

$$d(n_0) := \sup_{(\eta,\mu)\in\mathcal{P}(F)^2} \delta_{n_0}(\eta,\mu) \le \frac{cn_0}{1-b(n_0)} < \infty. \tag{3.9}$$

We recall that $\overline{\omega}_n(\mu)$ is defined in (3.3). We are now in a position to state and prove the main result of this section.

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

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

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 $n\ge 1$, the probability that

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

is greater than $(1-\delta)$ [where c is the constant introduced in (3.1)].

COROLLARY 3.4. *For the SIMC associated with the occupation measure distribution flow (3.2), we have for any $n\ge 0$, $f\in\mathcal{B}_1(E)$ and any $r\ge 1$*

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

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 $n\ge 1$, the probability that

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

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

PROOF OF THEOREM 3.3. First, we examine some consequences of the regularity conditions presented in (3.1) on the resolvent function P_η introduced in (3.4). Using Propositions 3.1 and 3.2 we find the following uniform estimates:

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

and

$$\|P_\mu - P_\eta\| \leq 3c\left(\frac{n_0}{1-b(n_0)}\right)^2 \|\mu - \eta\|. \tag{3.10}$$

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

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

Using the Poisson equation, we have

$$[\mathrm{Id} - \omega(\mu_p)] = (\mathrm{Id} - K_{\mu_p})P_{\mu_p}.$$

From this formula, we find the decomposition

$$\begin{aligned}
&[f(X_p) - \omega(\mu_p)(f)] \\
&\qquad = P_{\mu_p}(f)(X_p) - K_{\mu_p}(P_{\mu_p}(f))(X_p) \\
&\qquad = [P_{\mu_p}(f)(X_p) - P_{\mu_p}(f)(X_{p+1})] + \Delta M_{p+1}(f)
\end{aligned} \tag{3.11}$$

with the increments

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

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

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

For $n = 0$, we set $M_0(f) = 0$. The first term in the right-hand side of (3.11) can also be rewritten in the following form:

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

This yields the decomposition

$$\begin{aligned}
&\sum_{p=0}^{n}[P_{\mu_p}(f)(X_p) - P_{\mu_p}(f)(X_{p+1})] \\
&\qquad = [P_{\mu_0}(f)(X_0) - P_{\mu_{n+1}}(f)(X_{n+1})] + L_{n+1}(f)
\end{aligned}$$

with the random sequence

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

In summary, we have established the following decomposition:

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

We estimate each term separately. First, using (3.10) we prove that

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

In much the same way, using (3.10) we obtain

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

From these two estimates, we conclude that

$$|I_n(f)| \leq |M_{n+1}(f)| + 3c(n+1)\left(\frac{n_0}{1-b(n_0)}\right)^2 \overline{\varepsilon}(n) + \frac{4n_0}{1-b(n_0)}. \tag{3.12}$$

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

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

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

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

from which we conclude that

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

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 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

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

Combining this estimate with (3.12), we find that

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

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

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

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

$$|[\eta_n - \overline{\omega}_n(\mu)](f)| \leq \frac{|M_{n+1}(f)|}{n+1} + \frac{n_0}{1-b(n_0)}\left[\frac{3cn_0}{1-b(n_0)}\overline{\varepsilon}(n) + \frac{4}{n+1}\right]$$

from which we deduce the rather crude upper bound

$$\begin{aligned} &|[\eta_n - \overline{\omega}_n(\mu)](f)| \\ &\qquad \leq \frac{|M_{n+1}(f)|}{n+1} + (1+c)\left(\frac{2n_0}{1-b(n_0)}\right)^2 \left[\overline{\varepsilon}(n) \vee \frac{1}{n+1}\right]. \end{aligned} \tag{3.13}$$

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

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

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

$$\begin{aligned} &\mathbb{P}\left(|[\eta_n - \overline{\omega}_n(\mu)](f)| > t + (1+c)\left(\frac{2n_0}{1-b(n_0)}\right)^2 \left[\overline{\varepsilon}(n) \vee \frac{1}{n+1}\right]\right) \\ &\qquad \leq 2\exp\left(-(n+1)\frac{t^2}{2}\left(\frac{1-b(n_0)}{n_0}\right)^2\right). \end{aligned}$$

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

4. Distribution flows models. In this section, we have collected the definition of a series of semigroups on distribution flow 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

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

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

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

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

4.1. *Time averaged semigroups.* We associate with the mappings Φ_l introduced in (1.1) the mappings

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

defined by the coordinate mappings

$$\forall\eta\in\mathcal{P}\big(S^{(l-1)}\big)^{\mathbb{N}},\forall n\geq 0 \qquad \Phi_n^{(l)}(\eta):=\Phi_l(\eta_n).$$

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

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

defined by the coordinate mappings

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

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 $\pi_n^{(0)}=\pi^{(0)}$.

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

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

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

4.2. *Integral operators.* We associate with the kernel Γ_k from $\mathcal{B}(S^{(k)})$ into $\mathcal{B}(S^{(k-1)})$ introduced in (1.7) the kernel $\overline{\Gamma}^{(k)}$ from $(\mathbb{N} \times \mathcal{B}(S^{(k)}))$ into the set $(\mathbb{N} \times \mathcal{B}(S^{(k-1)}))$ defined by

$$\begin{aligned} \overline{\Gamma}^{(k)}((n,f),d(p,g)) &:= \Sigma(n,dp) \times \Gamma_k(f,dg) \\ &\text{with } \Sigma(n,dp) := \frac{1}{n+1}\sum_{q=0}^{n}\delta_q(dp). \end{aligned} \tag{4.1}$$

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

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

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}, $0 \le l_1 \le l_2$ associated with the pair of integral operators Σ and Γ_l are

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

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

We end this section with a technical lemma relating the regularity properties (1.7) of the mappings Φ_k to the regularity properties of the semigroups $\overline{\Phi}^{(k,l)}$.

LEMMA 4.1. *For any $0 \le l_1 \le l_2$, $n \ge 0$, 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*

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

PROOF. Notice that we have $\overline{\Gamma}^{(l,l)} = \overline{\Gamma}^{(l)}$. We also observe that $\overline{\Gamma}^{(l_2,l_1)}$ is a kernel 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

(1.7) we find that for any $l \geq 0$

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

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

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

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

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

and therefore

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

Under our induction hypothesis, this implies that

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

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 (1.9). 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

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

with the coordinate mappings ω_n defined by

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

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

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

defined by the coordinate mappings

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

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

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

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

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

This ends the proof of the lemma. $\square$

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

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

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

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

We also observe that

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

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

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

from which we conclude that

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

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 nonasymptotic $\mathbb{L}_r$-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 toward the solution $\pi^{(k)}$ of the measure-valued equation (1.1).

Section 5.3 is concerned with uniform convergence results with respect to the level index k. We examine this important question in terms of the stability properties of the time averaged semigroups introduced in Section 4.1. We present nonasymptotic $\mathbb{L}_r$-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 (1.1).

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*

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

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 $\frac{1}{p'} + \frac{1}{p} = 1$. First, we notice that the functions

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

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

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

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

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

By construction, we have

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

from which we conclude that

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

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.

THEOREM 5.2. *Under the regularity conditions* (1.7) *and* (1.8), *we have for any* $k \geq 0$, *any function* $f \in \mathcal{B}_1(S^{(k)})$ *and any* $n \geq 0$ *and* $r \geq 1$

$$\begin{aligned}&\sqrt{(n+1)}\mathbb{E}(|[\eta_n^{(k)} - \pi^{(k)}](f)|^r)^{1/r}\\ &\quad\leq e(r)\sum_{l=0}^{k}(1+c_l)\left(\frac{n_l}{1-b_l(n_l)}\right)^2 \prod_{l+1\leq i\leq k} 2\Lambda_i.\end{aligned}\tag{5.1}$$

PROOF. We prove the theorem by induction on the parameter k. First, we observe that the estimate (5.1) is true for $k = 0$. Indeed, by Corollary 3.4 we have that

$$\sqrt{(n+1)}\mathbb{E}(|[\eta_n^{(0)} - \pi^{(0)}](f)|^r)^{1/r} \leq e(r)(1+c_0)\left(\frac{n_0}{1-b_0(n_0)}\right)^2$$

for some finite constant $e(r) < \infty$ whose value only depends on the parameter r. We further suppose that the estimate (5.1) is true at rank $(k-1)$. To prove that it is also true at rank k, we use the decomposition

$$[\eta_n^{(k)} - \pi^{(k)}] = [\eta_n^{(k)} - \overline{\Phi}_n^{(k)}(\eta^{(k-1)})] + [\overline{\Phi}_n^{(k)}(\eta^{(k-1)}) - \overline{\Phi}_n^{(k)}(\pi^{(k-1)})].\tag{5.2}$$

For every $k \geq 0$, given a realization of the chain $X^{(k-1)} := (X_p^{(k-1)})_{p\geq 0}$ the kth 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

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

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

Using the decomposition (5.2) and Lemma 4.1, we obtain

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

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

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

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

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

Since we have

$$\int_{\mathbb{N}} \frac{1}{\sqrt{q+1}}\Sigma(n, dq) = \frac{1}{n+1}\sum_{q=0}^{n}\frac{1}{\sqrt{q+1}} \leq \frac{2}{\sqrt{n+1}} \tag{5.3}$$

we conclude that

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

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

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

This ends the proof of the theorem. $\square$

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 some 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

$$\|\overline{\Phi}^{(l+k,l+1)}(\eta) - \overline{\Phi}^{(l+k,l+1)}(\mu)\| \leq \lambda_1 e^{-\lambda_2 k}. \tag{5.4}$$

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

$$A = \sup_{k\geq 0}\left[(1+c_k)\left(\frac{n_k}{1-b_k(n_k)}\right)^2\right] < \infty \quad \text{and} \quad B := 2\sup_{k\geq 1}\Lambda_k < \infty. \tag{5.5}$$

For the Feynman–Kac transformations (2.1), we give in Section 7 sufficient conditions on G_l and L_{l+1} ensuring (5.4) is satisfied. If (5.4) and (5.5) 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 $(f_l)_{l\geq 0} \in \prod_{l\geq 0}\mathrm{Osc}_1(S^{(l)})$*

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

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 $(f_l)_{l\geq 0} \in \prod_{l\geq 0}\mathrm{Osc}_1(S^{(l)})$.

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

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

PROOF. First, we notice that we have the following estimate from (5.1) and (5.5) for any $k \geq 0$:

$$\sqrt{(n+1)}\mathbb{E}(|[\eta_n^{(k)} - \pi^{(k)}](f_k)|^r)^{1/r} \leq e(r)A\frac{B^{k+1}-1}{B-1}. \tag{5.6}$$

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

We have the following decomposition:

$$\begin{aligned}
\eta_n^{(l+k)} - \pi^{(l+k)} &= [\eta_n^{(l+k)} - \overline{\Phi}^{(l+k,l+1)_n}(\eta^{(l)})] \\
&\quad + [\overline{\Phi}_n^{(l+k,l+1)}(\eta^{(l)}) - \overline{\Phi}_n^{(l+k,l+1)}(\pi^{(l)})] \\
&= \sum_{i=l+1}^{l+k}[\overline{\Phi}_n^{(l+k,i+1)}(\eta^{(i)}) - \overline{\Phi}_n^{(l+k,i+1)}(\overline{\Phi}^{(i)}(\eta^{(i-1)}))] \\
&\quad + [\overline{\Phi}_n^{(l+k,l+1)}(\eta^{(l)}) - \overline{\Phi}_n^{(l+k,l+1)}(\pi^{(l)})].
\end{aligned} \tag{5.7}$$

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

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

Using Lemma 4.1, we find that

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

By the generalized Minkowski integral inequality, this implies that

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

Using Corollary 3.4, we find that

$$\begin{aligned}&\mathbb{E}\big(\big|\big[\overline{\Phi}_n^{(l_2,l_1+1)}(\eta^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\Phi}^{(l_1)}(\eta^{(l_1-1)}))\big](f_{l_2})\big|^r\big)^{1/r}\\ &\qquad \le e(r)(1+c_{l_1})\left(\frac{n_{l_1}}{1-b_{l_1}(n_{l_1})}\right)^2\\ &\qquad\qquad \times \int_{\{0,\dots,n\}} \frac{1}{\sqrt{(p+1)}}\Sigma^{(l_2-l_1)}(n,dp) \times \int \|g\|\Gamma_{l_2,l_1+1}(f_{l_2},dg).\end{aligned}$$

By (5.3) and

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

we conclude that

$$\begin{aligned}&\sqrt{(n+1)}\mathbb{E}\big(\big|\big[\overline{\Phi}_n^{(l_2,l_1+1)}(\eta^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\Phi}^{(l_1)}(\eta^{(l_1-1)}))\big](f_{l_2})\big|^r\big)^{1/r}\\ &\qquad \le e(r)AB^{l_2-l_1}\|f_{l_2}\|.\end{aligned} \tag{5.8}$$

Using the decomposition (5.7), we prove that for every $f_{l+k} \in \mathcal{B}_1(S^{(l+k)})$ and any $k \ge k_0$

$$\sup_{l\ge 0}\mathbb{E}\big(\big|\big[\eta_n^{(l+k)} - \pi^{(l+k)}\big](f_{l+k})\big|^r\big)^{1/r} \le e(r)\frac{A}{\sqrt{n+1}}\frac{B^k-1}{B-1} + \lambda_1 e^{-\lambda_2 k}.$$

Finally, by (5.6), we conclude that for every $k \ge k_0$

$$\sup_{l\ge 0}\mathbb{E}\big(\big|\big[\eta_n^{(l)} - \pi^{(l)}\big](f_l)\big|^r\big)^{1/r} \le e(r)\frac{A}{\sqrt{n+1}}\frac{B^{k+1}-1}{B-1} + \lambda_1 e^{-\lambda_2 k}.$$

For $B = 1$, we have

$$\sup_{l\ge 0}\mathbb{E}\big(\big|\big[\eta_n^{(l)} - \pi^{(l)}\big](f_l)\big|^r\big)^{1/r} \le e(r)A\frac{(k+1)}{\sqrt{n+1}} + \lambda_1 e^{-\lambda_2 k}.$$

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

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

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

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

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

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

from which we conclude that

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

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

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

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

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

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

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

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. $\square$

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 $\eta_n^{(k)}$ toward the solutions $\pi_n^{(k)}$ of the measure-valued equation (1.1). We show here that it is also possible to establish such results to quantify the convergence of the path-space occupation measures $\overline{\eta}_n^{[m]}$ introduced in (1.6) toward the tensor product measure $\overline{\pi}^{(m)}$ defined in (1.10).

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*

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

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 (1.9). 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 sequence of 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 lth level space $S^{(l)}$.

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

$$\eta - \overline{\pi}^{[m]} = \sum_{k=0}^{m}\big[\overline{\omega}^k(\eta) - \overline{\omega}^{k+1}(\eta)\big]. \tag{6.1}$$

In the above-displayed formula, $\overline{\pi}^{[m]} = (\overline{\pi}_n^{[m]})_{n\in\mathbb{N}} \in \mathcal{P}(E)^{\mathbb{N}}$ stands for the constant sequence of measures $\overline{\pi}_n^{[m]} = \overline{\pi}^{[m]}$, for any $n \in \mathbb{N}$.

Using Proposition 4.3, the kth 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:

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

Here the mappings

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

are defined for any $n \geq 0$ by

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

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 $0 \leq i \leq m-k$

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

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

and

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

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

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

and therefore

$$
\begin{aligned}
&\overline{\omega}_n^{k}(\eta) - \overline{\omega}_n^{k+1}(\eta)\\
&\quad= \frac{1}{n+1}\sum_{p=0}^{n}\big[\overline{\pi}^{[k-1]}\otimes\big\{\Pi_p^{(k,m)}\big((\eta^{(l)})_l\big) - \Pi_p^{(k,m)}\big(\big(\overline{\Phi}^{(l)}\big(\eta^{(l-1)}\big)\big)_l\big)\big\}\big].
\end{aligned}
\tag{6.2}
$$

Moving one step further, we introduce the decomposition

$$
\begin{aligned}
&\Pi^{(k,m)}(\mu) - \Pi^{(k,m)}(\nu)\\
&\quad= \sum_{j=0}^{m-k}\left\{\left(\bigotimes_{i=0}^{j-1}\Pi^{(k,m),(i)}(\nu)\right)\right.\\
&\qquad\qquad\otimes\big[\Pi^{(k,m),(j)}(\mu) - \Pi^{(k,m),(j)}(\nu)\big]\\
&\qquad\qquad\qquad\left.\otimes\left(\bigotimes_{i=j+1}^{m-k}\Pi^{(k,m),(i)}(\mu)\right)\right\}
\end{aligned}
\tag{6.3}
$$

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

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

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

$$\begin{aligned} &\bigl|\bigl[\Pi_n^{(k,m),(j)}(\mu) - \Pi_n^{(k,m),(j)}(\nu)\bigr](f)\bigr| \\ &\quad \le \int \bigl|\bigl[\bigl(\overline{\Phi}_n^{(j+(k-1),j+1)}\bigl(\mu^{(j)}\bigr)\bigr) \\ &\qquad - \bigl(\overline{\Phi}_n^{(j+(k-1),j+1)}\bigl(\nu^{(j)}\bigr)\bigr)\bigr](g)\bigr|\Gamma_{j+k}(f,dg). \end{aligned} \tag{6.4}$$

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

$$\mathcal{F}_n^{m,j} = \sigma\bigl(X_p^{(l)} : 0 \le p \le n, 0 \le l \le m, l \ne j\bigr).$$

Combining the generalized Minkowski integral inequality presented in Lemma 5.1 with the inequality (5.8), we prove that

$$\begin{aligned} &\mathbb{E}\bigl(\bigl|\bigl[\Pi_n^{(k,m),(j)}\bigl(\bigl(\overline{\eta}^{(l)}\bigr)_l\bigr) - \Pi_n^{(k,m),(j)}\bigl(\bigl(\overline{\Phi}^{(l)}\bigl(\overline{\eta}^{(l-1)}\bigr)\bigr)_l\bigr)\bigr](f)\bigr|^r \big| \mathcal{F}_n^{m,j}\bigr)^{1/r} \\ &\quad \le \int \mathbb{E}\bigl(\bigl|\bigl[\bigl(\overline{\Phi}_n^{(j+(k-1),j+1)}\bigl(\overline{\eta}^{(j)}\bigr)\bigr) \\ &\qquad\qquad - \bigl(\overline{\Phi}_n^{(j+(k-1),j+1)}\bigl(\overline{\Phi}^{(j)}\bigl(\overline{\eta}^{(j-1)}\bigr)\bigr)\bigr)\bigr](g)\bigr|^r \big| \mathcal{F}_n^{m,j}\bigr)^{1/r} \times \Gamma_{j+k}(f,dg) \\ &\quad \le \frac{e(r)}{\sqrt{n+1}} A B^k \|f\|. \end{aligned}$$

Notice that the decomposition (6.3) can be rewritten for any $f \in \mathcal{B}(\prod_{l=k}^m S^{(l)})$ in the following form:

$$\begin{aligned} &\bigl[\Pi_n^{(k,m)}(\mu) - \Pi_n^{(k,m)}(\nu)\bigr](f) \\ &\qquad = \sum_{j=0}^{m-k} \bigl[\Pi_n^{(k,m),(j)}(\mu) - \Pi_n^{(k,m),(j)}(\nu)\bigr]\bigl(R_n^{(k,m),(j)}(\mu,\nu)(f)\bigr) \end{aligned} \tag{6.5}$$

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

$$\begin{aligned} &R_n^{(k,m),(j)}(\mu,\nu)(f)(x_{k+j}) \\ &\quad = \int f\bigl(x_k, \ldots, x_{k+(j-1)}, x_{k+j}, x_{k+j+1}, \ldots, x_m\bigr) \\ &\qquad\qquad \times \Biggl(\prod_{i=0}^{j-1} \Pi_n^{(k,m),(i)}(\nu)\Biggr)(dx_{i+k}) \times \Biggl(\prod_{i=j+1}^{m-k} \Pi_n^{(k,m),(i)}(\mu)(dx_{i+k})\Biggr). \end{aligned}$$

Using the fact that the pair of measures

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

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

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

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

$$\begin{aligned}
&\mathbb{E}\big(\big|\big[\Pi_n^{(k,m),(j)}\big((\overline{\eta}^{(l)})_l\big) - \Pi_n^{(k,m),(j)}\big((\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)}))_l\big)\big]\big(f_n^{(k,m),(j)}\big)\big|^r \,|\mathcal{F}_n^{m,j}\big)^{1/r} \\
&\quad\le \int \Gamma_{j+k}\big(f_n^{(k,m),(j)}, dg\big) \\
&\qquad \times \mathbb{E}\big(\big|\big[\big(\overline{\Phi}_n^{(j+(k-1),j+1)}(\overline{\eta}^{(j)})\big) \\
&\qquad\qquad - \big(\overline{\Phi}_n^{(j+(k-1),j+1)}\big(\overline{\Phi}^{(j)}(\overline{\eta}^{(j-1)})\big)\big)\big](g)\big|^r \,|\mathcal{F}_n^{m,j}\big)^{1/r} \\
&\quad\le \frac{e(r)}{\sqrt{n+1}} AB^k \|f\|.
\end{aligned}$$

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

$$\begin{aligned}
&\mathbb{E}\big(\big|\big[\Pi_n^{(k,m)}\big((\overline{\eta}^{(l)})_l\big) - \Pi_n^{(k,m)}\big((\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)}))_l\big)\big](f)\big|^r\big)^{1/r} \\
&\quad\le (m-k+1)\frac{e(r)}{\sqrt{n+1}} AB^k \|f\|.
\end{aligned}$$

Using (6.5), it is now easily checked that for every $f \in \mathcal{B}(E)$

$$\mathbb{E}\big(|[\overline{\omega}_n^k(\overline{\eta}) - \overline{\omega}_n^{k+1}(\overline{\eta})](f)|^r\big)^{1/r} \le (m-k+1)\frac{e(r)}{\sqrt{n+1}} AB^k \|f\|.$$

Finally, by (6.1) we conclude that

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

This ends the proof of the theorem. $\square$

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 (1.7) for some kernel Γ_k with uniformly finite support

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

To simplify the presentation, we fix a parameter $m \ge 1$, and sometimes we write $\overline{\eta}_n$ instead of $\overline{\eta}_n^{[m]}$. We shall also use the letters c_i, $i \ge 1$ 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 a finite constant* $\overline{\sigma}_m < \infty$ *such that for any* $f \in \mathcal{B}_1(E_m)$ *and* $t > 0$

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

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* $n \geq 1$

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

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*:

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

where Σ^k *is the semigroup associated to the operator* Σ *defined in* (4.1). *For every* $k \geq 0$, $n \geq 1$, *and* $t \geq 0$ *we have the exponential inequalities*:

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

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

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

and

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

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

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

This implies that

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

from which we conclude that

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

This ends the proof of the lemma. $\square$

LEMMA 6.4. *For every $l_1 < l_2$, there exists some nonincreasing function*

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

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

$$\mathbb{P}(\sqrt{n+1}|[\overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\eta}^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1)}(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))](f)| > t)$$
$$\leq (c_1(n+1))^{(l_2-l_1)} \exp(-c_2 t^2/c_3^{l_2-l_1}).$$

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

$$\mathbb{P}(|[\overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\eta}^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1)}(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))](f)| > t)$$
$$\leq (c_1(n+1))^{(l_2-l_1)} \exp(-c_2(n+1)t^2/c_3^{l_2-l_1}).$$

Using the decomposition

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

we prove the following inclusion of events:

$$
\begin{aligned}
&\{|[\overline{\eta}_n^{(k)} - \pi^{(k)}](f)| > t\} \\
&\qquad \subset \{\exists 0 \le l \le k : |[\overline{\Phi}_n^{(k,l+1)}(\eta^{(l)}) - \overline{\Phi}_n{}^{(k,l+1)}(\overline{\Phi}^{(l)}(\eta^{(l-1)}))](f)| > t/(k+1)\}.
\end{aligned}
$$

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 $n \ge N(t)$

$$
\begin{aligned}
&\mathbb{P}(|[\overline{\eta}_n^{(k)} - \pi^{(k)}](f)| > t) \\
&\quad \le \sum_{0 \le l \le k} \mathbb{P}\left(|[\overline{\Phi}_n^{(k,l+1)}(\overline{\eta}^{(l)}) - \overline{\Phi}_n^{(k,l)}(\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)}))](f)| > \frac{t}{k+1}\right) \\
&\quad \le (k+1)\bigl(c_1(n+1)\bigr)^k e^{-(n+1)t^2 c_2/((k+1)^2 c_3^k)}.
\end{aligned}
$$

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

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

PROOF OF LEMMA 6.4. Using Lemma 4.1, we find that

$$
\begin{aligned}
&|[\overline{\Phi}_n^{(l_2,l_1+1)}(\overline{\eta}^{(l_1)}) - \overline{\Phi}_n^{(l_2,l_1)}(\overline{\Phi}^{(l_1)}(\overline{\eta}^{(l_1-1)}))](f)| \\
&\qquad \le \int |[\overline{\eta}_p^{(l_1)} - \overline{\Phi}_p^{(l_1)}(\overline{\eta}^{(l_1-1)})](g)| \overline{\Gamma}^{(l_2,l_1+1)}((n,f), d(p,g)).
\end{aligned}
$$

Arguing as in (3.13), we find that for any $g \in \mathcal{B}(S^{(l_1)})$, we have

$$
|[\overline{\eta}_p^{(l_1)} - \overline{\Phi}_p^{(l_1)}(\overline{\eta}^{(l_1-1)})](g)| \le \frac{|M_{p+1}^{(l_1)}(g)|}{p+1} + c_1 \frac{\log(p+2)}{p+2} \|g\| \tag{6.6}
$$

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

$$
\mathbb{P}(|M_n^{(l_1)}(g)| \ge t\sqrt{n}) \le 2\exp(-c_2 t^2/\|g\|^2).
$$

We notice that

$$
\begin{aligned}
\frac{1}{n+2} \sum_{p=0}^{n} \frac{(\log(p+2))^k}{p+2} &\le \frac{(\log(n+2))^k}{n+2} \sum_{p=0}^{n} \frac{1}{p+2} \\
&\le \frac{(\log(n+2))^k}{n+2} \sum_{p=0}^{n} \int_{p+1}^{p+2} \frac{1}{t}\, dt \\
&= \frac{(\log(n+2))^{k+1}}{n+2}.
\end{aligned}
$$

This implies that

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

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

$$\begin{aligned}
&\int \frac{\log(p+2)}{p+2} \|g\| \overline{\Gamma}^{(l_2,l_1+1)}((n, f), d(p, g)) \\
&\quad \leq 2^{(l_2-l_1)} \frac{(\log(n+2))^{(l_2-l_1)+1}}{n+2} \int \|g\| \Gamma_{l_2,l_1+1}(f, dg) \\
&\quad \leq 2^{(l_2-l_1)} \frac{(\log(n+2))^{(l_2-l_1)+1}}{n+2} \biggl(\prod_{l_1<i\leq l_2} \Lambda_i\biggr) \\
&\quad \leq c_3^{(l_2-l_1)} \frac{(\log(n+2))^{(l_2-l_1)+1}}{n+2}.
\end{aligned} \tag{6.7}$$

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

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

Using Lemma 6.3, we prove that

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

We observe that

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

In addition, using (6.6) and (6.7) we find that

$$\begin{aligned}
&\bigl|\bigl[\overline{\Phi}_n^{(l_2,l_1+1)}\bigl(\overline{\eta}^{(l_1)}\bigr) - \overline{\Phi}_n^{(l_2,l_1)}\bigl(\overline{\Phi}^{(l_1)}\bigl(\overline{\eta}^{(l_1-1)}\bigr)\bigr)\bigr](f)\bigr| \\
&\qquad \leq \int \overline{\mathcal{M}}_{n+1}^{(l_1,l_2)}(g) \Gamma_{l_2,l_1+1}(f, dg) + \varepsilon_{l_1,l_2}(n)
\end{aligned} \tag{6.8}$$

with

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

Using the inclusion of events

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

we find that

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

Finally, under our assumptions we have

$$S_{l_2,l_1+1}(f) = \mathrm{Card}(\mathrm{Supp}(\Gamma_{l_2,l_1+1}(f,\cdot)))$$
$$\leq \prod_{l_1+1\leq k\leq l_2} \sup_{f\in\mathcal{B}(S^{(k)})} \mathrm{Card}(\mathrm{Supp}(\Gamma_k(f,\cdot))) \leq c_4^{(l_2-l_1)}$$

from which we check that

$$\mathbb{P}\left(\int \overline{\mathcal{M}}_{n+1}^{(l_1,l_2)}(g)\Gamma_{l_2,l_1+1}(f,dg) > t\right)$$
$$\leq \big(c_5(n+1)\big)^{(l_2-l_1)} \exp\big(-c_6(n+1)t^2/c_7^{(l_2-l_1)}\big).$$

Using (6.8), we conclude that

$$\mathbb{P}\big(\big|\big[\overline{\Phi}_n^{(l_2,l_1+1)}\big(\overline{\eta}^{(l_1)}\big) - \overline{\Phi}_n^{(l_2,l_1)}\big(\overline{\Phi}^{(l_1)}\big(\overline{\eta}^{(l_1-1)}\big)\big)\big](f)\big| > t + \varepsilon_{l_1,l_2}(n)\big)$$
$$\leq \big(c_5(n+1)\big)^{(l_2-l_1)} \exp\big(-c_6(n+1)t^2/c_7^{(l_2-l_1)}\big).$$

To take the final step, we observe that

$$\mathbb{P}\big(\sqrt{n+1}\big|\big[\overline{\Phi}_n^{(l_2,l_1+1)}\big(\overline{\eta}^{(l_1)}\big) - \overline{\Phi}_n^{(l_2,l_1)}\big(\overline{\Phi}^{(l_1)}\big(\overline{\eta}^{(l_1-1)}\big)\big)\big](f)\big| > t + \sqrt{n+1}\varepsilon_{l_1,l_2}(n)\big)$$
$$\leq \mathbb{P}\Bigg(\big|\big[\overline{\Phi}_n^{(l_2,l_1+1)}\big(\overline{\eta}^{(l_1)}\big) - \overline{\Phi}_n^{(l_2,l_1)}\big(\overline{\Phi}^{(l_1)}\big(\overline{\eta}^{(l_1-1)}\big)\big)\big](f)\big|$$
$$> \frac{t}{\sqrt{n+1}} + \varepsilon_{l_1,l_2}(n)\Bigg).$$

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

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

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

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

The end of the proof is now straightforward. $\square$

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 (6.4) we find that

$$\begin{aligned}
&\big|\big[\Pi_n^{(k,m),(j)}(\mu) - \Pi_n^{(k,m),(j)}(\nu)\big](f)\big| > t \\
&\quad\Longrightarrow\quad \exists g \in \mathrm{Supp}(\Gamma_{j+k}(f,\cdot)) : \big|\big[\big(\overline{\Phi}_n^{(j+(k-1),j+1)}(\mu^{(j)})\big) \\
&\qquad\qquad - \big(\overline{\Phi}_n^{(j+(k-1),j+1)}(\nu^{(j)})\big)\big](g)\big| > t\|g\|/\Lambda_{j+k}.
\end{aligned}$$

Therefore, using Lemma 6.4 we can find a nonincreasing 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

$$\begin{aligned}
&\mathbb{P}\big(\sqrt{n+1}\big|\big[\Pi_n^{(k,m),(j)}(\mu) - \Pi_n^{(k,m),(j)}(\nu)\big](f)\big| > t\big) \\
&\qquad \leq \big(c_1(n+1)\big)^{(k-1)} \exp\big(-c_2 t^2/c_3^{(k-1)}\big).
\end{aligned}$$

In much the same way, by the decomposition (6.5) we find the following assertion:

$$\begin{aligned}
&\big|\big[\Pi_n^{(k,m)}(\mu) - \Pi_n^{(k,m)}(\nu)\big](f)\big| > t \\
&\quad\Longrightarrow\quad \exists 0 \leq j \leq (m-k) : \big|\big[\Pi_n^{(k,m),(j)}(\mu) - \Pi_n^{(k,m),(j)}(\nu)\big] \\
&\qquad\qquad \times \big(R_n^{(k,m),(j)}(\mu,\nu)(f)\big)\big| > t/(m-k+1).
\end{aligned}$$

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 $n \geq N(t)$

$$\begin{aligned}
&\mathbb{P}\big(\sqrt{n+1}\big|\big[\Pi_n^{(k,m)}\big((\overline{\eta}^{(l)})_l\big) - \Pi_n^{(k,m)}\big((\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)}))_l\big)\big](f)\big| > t\big) \\
&\qquad \leq (m-k+1)\big(c_1(n+1)\big)^{k-1} \exp\big(-c_2 t^2/\big((m-k+1)^2 c_3^{k-1}\big)\big).
\end{aligned}$$

In summary, we have proved that there exists some nonincreasing 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

$$\begin{aligned}
&\mathbb{P}\big(\sqrt{n+1}\big|\big[\overline{\pi}^{[k-1]} \otimes \big\{\Pi_n^{(k,m)}\big((\overline{\eta}^{(l)})_l\big) - \Pi_n^{(k,m)}\big((\overline{\Phi}^{(l)}(\overline{\eta}^{(l-1)}))_l\big)\big\}\big](f)\big| > t\big) \\
&\qquad \leq \big(c_4(n+1)\big)^m \exp(-c_5 t^2/c_6^m).
\end{aligned}$$

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

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

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

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

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

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

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

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

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

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

This implies that $N'(s) \leq N'(t)$. Thus, we have constructed a nonincreasing 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 (6.2), we can find a nonincreasing function $N(t)$ such that for any $n \geq N(t)$ and any $0 \leq k \leq m$

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

The end of the proof of the theorem is now a direct consequence of the decomposition (6.1). □

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 (5.4). In this section, we study the mappings $\overline{\Phi}^{(k,l)}$ associated with the Feynman–Kac transformations discussed in (7.2). We provide necessary conditions ensuring that (5.4) 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 $\mathcal{B}(S^{(l)})$ into $\mathcal{B}(S^{(l-1)})$ given by

$$\forall f \in \mathcal{B}(S^{(l)}) \qquad Q_l(f) := G_{l-1} \times L_l(f) \in \mathcal{B}(S^{(l-1)}). \tag{7.1}$$

By definition of the mappings Φ_l given in (2.1), it is easy to check that

$$\begin{aligned} &\overline{\Phi}^{(l)}(\eta) = \overline{\Psi}^{(l),Q_l(1)}(\eta) L_l \\ &\text{with } \forall n \geq 0 \overline{\Psi}_n^{(l),Q_l(1)}(\eta) = \frac{1}{n+1} \sum_{p=0}^{n} \Psi_l^{Q_l(1)}(\eta_p). \end{aligned} \tag{7.2}$$

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

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

the semigroup associated with the integral operator Q_l introduced in (7.1).

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

$$\overline{\Phi}^{(k,l)}(\eta) = \overline{\Psi}^{(k,l)}(\eta) P_{l,k} \qquad \textit{with } P_{l,k}(f) = \frac{Q_{l,k}(f)}{Q_{l,k}(1)}, \tag{7.3}$$

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

$$\begin{aligned} \overline{\Psi}^{(k,l)} &= \overline{\Psi}^{(l),H_{l,k}} \circ \overline{\Psi}^{(k-1,l)} \\ &= \overline{\Psi}^{(l),H_{l,k}} \circ \overline{\Psi}^{(l),H_{l,k-1}} \circ \cdots \circ \overline{\Psi}^{(l),H_{l,l}} \qquad \textit{with } H_{l,k} := \frac{Q_{l,k}(1)}{Q_{l,k-1}(1)}. \end{aligned}$$

For $l = k$, we use the conventions $\overline{\Psi}^{(k-1,l)} = \overline{\Psi}^{(l-1,l)} = \mathrm{Id}$ 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

$$P_{l,l}(f) = \frac{Q_l(f)}{Q_l(1)} = L_l(f)$$

and

$$\overline{\Psi}^{(l,l)} = \overline{\Psi}^{(l),Q_l(1)} \quad \Longrightarrow \quad \overline{\Phi}^{(l)}(\eta) = \overline{\Psi}^{(l,l)}(\eta) P_{l,l}.$$

Suppose we have proved formula (7.3) 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

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

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

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

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

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

and therefore

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

In summary, we have proved that

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

This ends the proof of the proposition. $\square$

7.2. *Contraction inequalities.*

PROPOSITION 7.4. *For any $l \le k$ 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

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

This implies that

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

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

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

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

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

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 $L_{l,k} = L_l L_{l-1} \cdots L_k$ associated with the Markov operators L_l. We introduce the following regularity conditions.

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

$$\forall l \ge 0, \forall (x, y) \in (S^{(l-1)})^2 \qquad L_{l+1,l+m}(x, \cdot) \ge \varepsilon_l(L) L_{l+1,l+m}(y, \cdot).$$

It is well known that the above condition is satisfied for any aperiodic and irreducible Markov chain on a finite space. Loosely speaking, for noncompact spaces this condition is related to the tails of the transition distributions on the boundaries of the state space. For instance, let us assume 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 $\operatorname{osc}(A_l) < \infty$. In this case, it is readily checked that condition $(L)_m$ holds true for $m = 1$ with the parameter

$$\varepsilon_{l-1}(L) = \exp\left(-c(l)\operatorname{osc}(A_l)\right).$$

Under the condition (G) presented on page 11 and the mixing condition $(L)_m$ stated above, we proved in [5] (see Corollary 4.3.3 on page 141) that we have for any $k \geq m \geq 1$, and $l \geq 1$

$$\beta(P_{l+1,l+k}) \leq \prod_{i=0}^{\lfloor k/m \rfloor - 1} \left(1 - \varepsilon_{l+im}^{(m)}\right) \qquad \text{with } \varepsilon_l^{(m)} := \varepsilon_l^2(L) \prod_{l+1 \leq k < l+m} \varepsilon_k(G).$$

Several contraction inequalities can be deduced from these estimates, we refer to Chapter 4 of the book [5]. To give a flavor of these results, we further assume that $(M)_m$ is satisfied with $m = 1$ and $\varepsilon(L) = \inf_l \varepsilon_l(L) > 0$. In this case, we can check that

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

Acknowledgments. We would like to thanks the anonymous referee for helpful comments and valuable suggestions that have improved the presentation of the article.

REFERENCES

[1] ANDRIEU, C., JASRA, A., DOUCET, A. and DEL MORAL, P. (2007). Non-linear Markov chain Monte Carlo via self-interacting approximations. Technical report, Dept. Mathematics, Bristol Univ.

[2] BERCU, B., DEL MORAL, P. and DOUCET, A. (2009). A functional central limit theorem for a class of interacting Markov chain Monte Carlo methods. *Electron. J. Probab.* **73** 2130–2155.

[3] BERCU, B., DEL MORAL, P. and DOUCET, A. (2008). Fluctuations of interacting Markov chain Monte Carlo models. Technical Report INRIA 6438.

[4] BROCKWELL, A., DEL MORAL, P. and DOUCET, A. (2010). Sequentially interacting Markov chain Monte Carlo. *Ann. Statist.* To appear.

[5] DEL MORAL, P. (2004). *Feynman–Kac Formulae: Genealogical and Interacting Particle Systems With Applications*. Springer, New York. MR2044973

[6] DEL MORAL, P. and MICLO, L. (2006). Self-interacting Markov chains. *Stoch. Anal. Appl.* **24** 615–660. MR2220075

[7] DEL MORAL, P. and MICLO, L. (2003). On convergence of chains with time empirical self-interactions. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460** 325–346. MR2052266

[8] DOUCET, A., DE FREITAS, N. and GORDON, N., eds. (2001). *Sequential Monte Carlo Methods in Practice*. Springer, New York. MR1847783

[9] KOU, S. C., ZHOU, Q. and WONG, W. H. (2006). Equi-energy sampler with applications in statistical inference and statistical mechanics. *Ann. Statist.* **34** 1581–1652. MR2283711

[10] LYMAN, E. and ZUCKERMAN, D. M. (2006). Resolution exchange simulation with incremental coarsening. *J. Chem. Theory Comp.* **2** 656–666.

[11] MENGERSEN, K. L. and TWEEDIE, R. L. (1996). Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.* **24** 101–121. MR1389882

[12] ROBERT, C. P. and CASELLA, G. (2004). *Monte Carlo Statistical Methods*, 2nd ed. Springer, New York. MR2080278

CENTRE INRIA BORDEAUX ET SUD-OUEST
AND INSTITUT DE MATHÉMATIQUES DE BORDEAUX
UNIVERSITÉ BORDEAUX
351 COURS DE LA LIBÉRATION
33405 TALENCE CEDEX
FRANCE
E-MAIL: Pierre.Del-Moral@inria.fr

DEPARTMENT OF STATISTICS
AND DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF BRITISH COLUMBIA
333-6356 AGRICULTURAL ROAD
VANCOUVER, BC, V6T 1Z2
CANADA
AND
THE INSTITUTE OF STATISTICAL MATHEMATICS
4-6-7 MINAMI-AZABU
MINATO-KU, TOKYO 106-8569
JAPAN
E-MAIL: Arnaud@stat.ubc.ca