Quantitative Approximations of Evolving Probability Measures and Sequential Markov Chain Monte Carlo Methods

Andreas Eberle, Carlo Marinelli

no. 411

Diese Arbeit ist mit Unterstützung des von der Deutschen Forschungsgemeinschaft getragenen Sonderforschungsbereichs 611 an der Universität Bonn entstanden und als Manuskript vervielfältigt worden.

Bonn, Mai 2008

QUANTITATIVE APPROXIMATIONS OF EVOLVING PROBABILITY MEASURES AND SEQUENTIAL MARKOV CHAIN MONTE CARLO METHODS

ANDREAS EBERLE AND CARLO MARINELLI

Abstract. We study approximations of evolving probability measures by an interacting particle system. The particle system dynamics is a combination of independent Markov chain moves and importance sampling/resampling steps. Under global regularity conditions, we derive non-asymptotic error bounds for the particle system approximation. The main motivation are applications to sequential MCMC methods for Monte Carlo integral estimation.

1. Introduction and main results

1.1. **Setup.** Let μ_t $(t \geq 0)$ denote a family of mutually absolutely continuous probability measures on a set S. To keep the presentation as simple and non-technical as possible, we assume that S is finite, although most results of this paper extend to continuous state spaces under standard regularity assumptions. We will now explain how to obtain Fokker–Planck type evolution equations on the space of probability measures on S that are satisfied by μ_t, and how to approximate these equations by interacting particle systems. The main purpose of this paper is to study the error of the particle system approximations in an L^p sense.

We assume that the measures are represented in the form

$$\mu_t(x) \;=\; \frac{1}{Z_t}\,\exp\left(-\mathcal{H}_t(x)\,\right)\,\mu_0(x)\,, \qquad t \geq 0, \tag{1.1}$$

where Z_t is a normalization constant, and $(t,x) \mapsto \mathcal{H}_t(x)$ is a given function on $[0,\infty)\times S$ that is continuously differentiable in the first variable. If, for example, $\mathcal{H}_t(x) = t\cdot\mathcal{H}(x)$ for some function $\mathcal{H} : S \to \mathbb{R}$ then $(\mu_t)_{t\geq 0}$ is the exponential family corresponding to $\mathcal{H}$ and μ_0. Let

$$H_t(x) \;:=\; -\frac{\partial}{\partial t}\log\mu_t(x) \;=\; -\frac{\partial}{\partial t}\log\frac{\mu_t(x)}{\mu_0(x)}$$

denote the negative logarithmic time derivative of the measures μ_t. Note that

$$\mu_t(x) \;=\; \exp\left(-\int_0^t H_s(x)\,ds\right)\,\mu_0(x)\,, \tag{1.2}$$

and

$$\langle H_t, \mu_t\rangle \;=\; 0 \qquad \text{for all } t \geq 0, \tag{1.3}$$

where

$$\langle f,\nu\rangle := \int_S f\,d\nu \;=\sum_{x\in S} f(x)\,\nu(x)$$

Date: June 6, 2008.

2000 *Mathematics Subject Classification.* 65C05, 60J25, 60B10, 47H20, 47D08.

Key words and phrases. Markov Chain Monte Carlo, sequential Monte Carlo, importance sampling, spectral gap, Dirichlet forms, functional inequalities.

This work is supported by the Sonderforschungsbereich 611, Bonn. The second author also gratefully acknowledges the hospitality and financial support of the Centre de Recerca Matemàtica, Bellaterra, through an IPDE fellowship.

denotes the integral of a function $f : S \to \mathbb{R}$ w.r.t. a measure ν on S. In particular,

$$H_t \;=\; \frac{\partial}{\partial t}\mathcal{H}_t \;-\; \langle \frac{\partial}{\partial t}\mathcal{H}_t, \mu_t\rangle \,.$$

In applications we have in mind, the functions $\mathcal{H}_t$ are given explicitly. Hence H_t is known explicitly up to an additive time-dependent constant. The evaluation of this constant, however, would require integration w.r.t. μ_t.

If all the functions H_t, $t \geq 0$, vanish then $\mu_t = \mu_0$ for all $t \geq 0$. In this case the measures are invariant for a Markov transition semigroup $(p_t)_{t\geq 0}$, i.e.,

$$\mu_s p_{t-s} \;=\; \mu_t \qquad \text{for all } t \geq s \geq 0,$$

provided the generator $\mathcal{L}$ satisfies $\mathcal{L}^*\mu_0 = 0$, i.e.,

$$\langle \mathcal{L}f, \mu_0\rangle \;=\; 0 \qquad \text{for all } f : S \to \mathbb{R}.$$

This fact is exploited in Markov Chain Monte Carlo (MCMC) methods for approximating expectation values w.r.t. the measure μ_0. The particle systems studied below can be applied for the same purpose when the measures μ_t are time-dependent.

1.2. **Fokker-Planck equation and particle system approximation.** To obtain approximations of the measures μ_t, we consider generators (Q-matrices) $\mathcal{L}_t$, $t \geq 0$, of a time-inhomogeneous Markov process on S satisfying the detailed balance conditions

$$\mu_t(x)\mathcal{L}_t(x,y) = \mu_t(y)\mathcal{L}_t(y,x) \quad \forall\, t \geq 0,\; x,y \in S. \tag{1.4}$$

For example, $\mathcal{L}_t$ could be the generator of a Metropolis dynamics w.r.t. μ_t, i.e.,

$$\mathcal{L}_t(x,y) \;=\; K_t(x,y)\cdot \min\left(\frac{\mu_t(y)}{\mu_t(x)}, 1\right) \text{ for } x \neq y,$$

$\mathcal{L}_t(x,x) = -\sum_{y\neq x}\mathcal{L}_t(x,y)$, where the proposal matrix K_t is a given symmetric transition matrix on S. By (1.4), $\mathcal{L}_t^*\mu_t = 0$, i.e.,

$$\langle \mathcal{L}_t f, \mu_t\rangle \;=\; 0 \qquad \text{for all } f : S \to \mathbb{R} \text{ and } t \geq 0.$$

Therefore, for any choice of non-negative constants λ_t, $t \geq 0$, the measures μ_t are the unique solution of the evolution equation

$$\frac{\partial}{\partial t}\nu_t \;=\; \lambda_t\,\mathcal{L}_t^*\nu_t \;-\; H_t\nu_t \tag{1.5}$$

with initial condition $\nu_0 = \mu_0$. In general, the solutions of (1.5) are not probability measures. Therefore, we consider the equation

$$\frac{\partial}{\partial t}\eta_t \;=\; \lambda_t\,\mathcal{L}_t^*\eta_t \;-\; H_t\eta_t \;+\; \langle H_t, \eta_t\rangle\,\eta_t \tag{1.6}$$

satisfied by the normalized measures $\eta_t = \frac{\nu_t}{\nu_t(S)}$. Note that

$$\nu_t \;=\; \exp\left(-\int_0^t \langle H_s, \eta_s\rangle\right)\eta_t \;,$$

since the right hand side solves (1.5) with initial condition ν_0.

The Fokker Planck equation (1.6) is an evolution equation for probability measures which is satisfied by μ_t. In contrast to the unnormalized equation it is not modified by adding constants to the functions H_t. We now introduce interacting particle systems that discretize the evolution equations (1.6) and (1.5). Consider right continuous time-inhomogeneous Markov processes $(X_t^N, \mathbb{P})$, $N \in \mathbb{N}$, with state space S^N and generators at time t given by

$$\mathcal{L}_t^N\varphi(x_1,\dots,x_N) \;=\; \lambda_t\sum_{i=1}^N \mathcal{L}_t^{(i)}\varphi(x_1,\dots,x_N) + \frac{1}{N}\sum_{i,j=1}^N (H_t(x_i)-H_t(x_j))^+(\varphi(x^{i\to j})-\varphi(x)). \tag{1.7}$$

Here $x = (x_1, \ldots, x_N) \in S^N$ and

$$(x^{i \to j})_k \;=\; \begin{cases} x_k & \text{if } k \neq i, \\ x_j & \text{if } k = i. \end{cases}$$

Moreover, $\mathcal{L}_t^{(i)}$ stands for the operator $\mathcal{L}_t$ applied to the i-th component of x. Thus the components $X_{t,i}^N$, $i = 1, \ldots, N$, of the process X_t^N move like independent Markov processes with generator $\lambda_t \mathcal{L}_t$ and are occasionally replaced by components with a lower value of H_t. Note that to compute the generator (and hence to simulate the Markov process) it is enough to know the functions H_t up to an additive constant.

One can show that if the initial distributions of the Markov processes are the N-fold products π^N of a probability measure π on S, then almost surely, the empirical distributions

$$\eta_t^N \;=\; \frac{1}{N} \sum_{i=1}^{N} \delta_{X_{t,i}^N} \tag{1.8}$$

and the reweighted empirical distributions

$$\nu_t^N \;=\; \exp\left(- \int_0^t \langle H_s, \eta_s^N \rangle \right) \eta_t^N \tag{1.9}$$

converge weakly to the solutions of the equations (1.6) and (1.5) with initial conditions $\eta_0 = \nu_0 = \pi$, cf. Corollary 1.6 below. As a consequence, simulating the Markov process X_t^N with initial distribution μ_0^N yields a Monte Carlo method for approximating sequentially the probability measures μ_t, $t \geq 0$. This Monte Carlo method can be viewed as a combination of Markov chain Monte Carlo and importance sampling/resampling. It is a continuous-time analogue of a particular type of sequential Monte Carlo sampler. Sequential Monte Carlo samplers are used in various applications to estimate expectation values w.r.t. multimodal distributions, and have been introduced systematically by Del Moral, Doucet and Jasra [1]. They are related to several multi-level sampling methods including parallel tempering [6, 9, 12] and the equi-energy sampler [11].

1.3. **Quantitative convergence bounds.** Our major aim is to quantify the convergence properties of the approximating particle systems when the initial distribution is μ_0^N. A law of large numbers type convergence theorem and a corresponding central limit theorem have been established in [3, 2] for a related particle system approximation, cf. also [13]. A crucial question for algorithmic applications, however, are quantitative bounds on the approximation error

$$\langle f, \eta_t^N \rangle - \langle f, \mu_t \rangle \tag{1.10}$$

for a given function $f : S \to \mathbb{R}$ and fixed N. The central limit theorem in [3] in principle yields such bounds asymptotically as $N \to \infty$ (at least for a modified particle system). However, two important questions remain open:

- The expression for the asymptotic variance in the central limit theorem derived in [3] is not very explicit – it involves L^2 norms of an associated Feynman-Kac semigroup. Methods how to bound this expression *efficiently* in a general setup and in concrete models have to be developed.
- For applications it is crucial to derive *non-asymptotic bounds* (i.e., bounds for fixed N), because the asymptotic estimates could be far off when only a limited number of particles is available. The arguments in the proof of the CLT in [3] and the results in [13] in principle yield non-asymptotic bounds but the constants showing up are of order $\exp \int_0^t \operatorname{osc}(H_s)\, ds$. Even in many simple models with strong mixing properties, this quantity is extremely large. Hence in spite of the LLN and CLT, even in nice cases it remains unclear if the empirical distributions yield a reasonable approximation to μ_t for a realistic number N of particles (e.g. $N = 10,000$).

Dobrushin contraction coefficients do not seem to be an appropriate tool to answer these more delicate questions. In [4] we propose an L^2 approach to quantify asymptotic stability properties of the Fokker-Planck equations for multimodal distributions where good global mixing properties fail. Our aim in the present article is to develop foundations of an L^p approach to non-asymptotic bounds for the particle system approximations.

To state our results in detail let us consider the Markov process $(X_t^N, \mathbb{N})$ with initial distribution μ_0^N. To derive error bounds for the particle system approximation it is convenient to consider at first the error for the Monte Carlo estimates based on the reweighted empirical distributions ν_t^N defined in (1.9). By a martingale argument it can be shown that $\langle f, \nu_t^N\rangle$ is an unbiased estimator of $\langle f, \mu_t\rangle$ for any function $f : S \to \mathbb{R}$ and $t \geq 0$, see Corollary 1.6 below. Elementary estimates show that the approximation error (1.10) can be controlled by variances of estimators based on ν_t^N:

Lemma 1.1. *For all functions $f : S \to \mathbb{R}$ and $t \geq 0$ we have*

$$\mathbb{E}\left[\left|\langle f, \eta_t^N\rangle - \langle f, \mu_t\rangle\right|^2\right] \;\leq\; 2\,\mathrm{Var}\left(\langle f, \nu_t^N\rangle\right) \,+\, 2\|f - \langle f, \mu_t\rangle\|_{\sup}^2\,\mathrm{Var}\left(\langle 1, \nu_t^N\rangle\right) \tag{1.11}$$

and

$$\begin{aligned}\mathbb{E}\left[\left|\langle f, \eta_t^N\rangle - \langle f, \mu_t\rangle\right|\right] \;\leq\;& \mathrm{Var}\left(\langle f, \nu_t^N\rangle\right)^{1/2} + \sqrt{2}\|f - \langle f, \mu_t\rangle\|_{\sup}\,\mathrm{Var}\left(\langle 1, \nu_t^N\rangle\right) \\ &+ \sqrt{2}\;\mathrm{Var}\left(\langle f, \nu_t^N\rangle\right)^{1/2}\;\mathrm{Var}\left(\langle 1, \nu_t^N\rangle\right)^{1/2}.\end{aligned} \tag{1.12}$$

The proof is given in Section 4 below. Because of the lemma, the errors can be quantified in terms of the variance bounds

$$\varepsilon_t^{N,p} := \sup\left\{\mathbb{E}\left[\left|\langle f, \nu_t^N\rangle - \langle f, \mu_t\rangle\right|^2\right] \,\middle|\, : \; f : S \to \mathbb{R},\; \|f\|_{L^p(\mu_t)} \leq 1\right\}, \qquad p \in [2, \infty). \tag{1.13}$$

To efficiently bound the quantities $\varepsilon_t^{N,p}$ we apply estimates of L^p-L^q operator norms of Feynman-Kac type transition operators $q_{s,t}$, which we now define. For $0 \leq s \leq t < \infty$ and a function $f : S \to \mathbb{R}$, let $q_{s,t}f(x)$ denote the unique solution of the backward equation

$$-\frac{\partial}{\partial s} q_{s,t}f = \lambda_s \mathcal{L}_s q_{s,t}f - H_s q_{s,t}f, \qquad s \in [0, t], \tag{1.14}$$

with terminal condition $q_{t,t}f = f$. It can be shown that $q_{s,t}f$ is also the unique solution of the corresponding forward equation

$$\frac{\partial}{\partial t} q_{s,t}f = q_{s,t}(\lambda_t \mathcal{L}_t f - H_t f), \qquad t \in [s, \infty), \tag{1.15}$$

with initial condition $q_{s,s}f = f$. As a consequence, a probabilistic representation of $q_{s,t}$ is given by Feynman-Kac formula

$$(q_{s,t}f)(x) = \mathbb{E}_{s,x}\left[e^{-\int_s^t H_r(X_r)\,dr} f(X_t)\right] \qquad \text{for all } x \in S, \tag{1.16}$$

where $(X_t)_{t \geq s}$ is a time-inhomogeneous Markov process w.r.t. $\mathbb{P}_{s,x}$ with generators $\mathcal{L}_t$ and initial condition $X_s = x$ $\mathbb{P}_{s,x}$-a.s., see e.g. [7], [8].

For p, $q \in [2, \infty]$ with $p \leq q$, let us consider the operator norms

$$C_{s,t}(p) = \sup_{f \neq 0} \frac{\|q_{s,t}f\|_{L^p(\mu_s)}}{\|f\|_{L^p(\mu_t)}} \qquad \text{and}$$

$$C_{s,t}(p,q) = \sup_{f \neq 0} \frac{\|q_{s,t}f\|_{L^{2r}(\mu_s)}}{\|f\|_{L^p(\mu_t)}} \vee \sup_{f \neq 0} \frac{\|q_{s,t}f\|_{L^p(\mu_s)}}{\|f\|_{L^{p/2}(\mu_t)}} \vee 1,$$

where $r \in [p, \infty]$ is chosen such that $p^{-1} = q^{-1} + r^{-1}$. Moreover, for $\delta > 0$ we set

$$\bar{C}_t(p, q, \delta) := \int_0^{(t-\delta)^+} \|H_s\|_{L^q(\mu_s)} C_{s,t}(p,q)^2\,ds.$$

Remark 1.2. Since we assume that the state space is finite, all the constants are finite, but their numerical values can be very large. It is a straightforward consequence of the forward equation (1.15) that

$$\mu_s q_{s,t} = \mu_t, \qquad 0 \le s \le t, \tag{1.17}$$

and hence $C_{s,t}(1) = 1$. On the other hand, in contrast to Markov transition operators which are contractions on L^∞, the constant $C_{s,t}(\infty)$ is extremely large in typical applications. Therefore bounds on $C_{s,t}(p)$ are very sensitive to the choice of p, see [5] for details. The constants $C_{s,t}(p,q)$ and $\bar{C}_t(p,q,\delta)$ are related to hypercontractivity properties and can only be expected to be bounded in a feasible way if $t-s$ and δ respectively are not too small.

For a function $f : S \to \mathbb{R}$ let

$$V_{s,t}(f) = -\langle H_s (q_{s,t} f)^2, \mu_s\rangle + \iint |H_s(x)| (q_{s,t} f(y) - q_{s,t} f(x))^2 \, \mu_s(dx)\, \mu_s(dy). \tag{1.18}$$

Our first main result shows that for $p > 4$ the asymptotic (as $N \to \infty$) variance of the estimator $\langle f, \nu_t^N \rangle$ is bounded from above by

$$N^{-1} \cdot \left(\mathrm{Var}_{\mu_t}(f) + \int_0^t V_{s,t}(f)\, ds + \|f\|^2_{L^p(\mu_t)} \right),$$

and it quantifies the deviation of the mean square error $\mathrm{Var}\left(\langle f, \nu_t^N\rangle\right)$ from the expression in the previous display non-asymptotically, i.e. for a fixed N:

Theorem 1.3. *Fix $t_0 \ge 0$, $q \in]6,\infty]$, and $p \in]\frac{4q}{q-2}, q[$. Let $N \in \mathbb{N}$ be such that*

$$N \ge 80 \cdot \max\left(1, \bar{C}_s(p,q,\delta), \bar{C}_s(\tilde{p},q,\delta)\right) \qquad \textit{for all } 0 \le s \le t_0,$$

where $\tilde{p}$ is defined by $\tilde{p}^{-1} = q^{-1} + (p/2)^{-1}$ and

$$\delta := \left(35 \sup_{s\in[0,t_0]} \operatorname{osc} H_s\right)^{-1}. \tag{1.19}$$

Then for $t \in [0,t_0]$ we have

$$\begin{aligned} N \cdot \mathbb{E}\left[|\langle f, \nu_t^N\rangle - \langle f, \mu_t\rangle|^2 \right] \;\le\; & \mathrm{Var}_{\mu_t}(f) + \int_0^t V_{s,t}(f)\, ds \\ & + \left[1 + \bar{C}_t(p,q,\delta) \cdot \sup_{s \le t} \left(19 (\varepsilon_s^{N,p})^{1/2} + 11 \varepsilon_s^{N,p} \right) \right] \cdot \|f\|^2_{L^p(\mu_t)}. \end{aligned} \tag{1.20}$$

In particular,

$$\sqrt{\varepsilon_t^{N,p}} \le (2 + v_t(p))^{1/2} N^{-1/2} + 23 \cdot \bar{C}_t(p,q,\delta) N^{-1} + 7 \bar{C}_t(p,q,\delta) (2 + v_t(p))^{1/2} N^{-3/2}, \tag{1.21}$$

where

$$v_t(p) \;:=\; \sup_{f \ne 0} \frac{\int_0^t V_{s,t}(f)\, ds}{\|f\|^2_{L^p(\mu_t)}}.$$

To apply Theorem 1.3 we need bounds for the constants $v_t(p)$ and $\bar{C}_t(p,q,\delta)$. We will now discuss how to derive such bounds from Poincaré and logarithmic Sobolev inequalities in the following particular cases:

a) The Markov processes with generators $\mathcal{L}_t$, $t \ge 0$, have "good" global mixing properties.
b) The state space S is decomposed into disjoint subsets S_i, $i \in I$, such that $\mathcal{L}_t(x,y) = 0$ for all $t \ge 0$, $x \in S_i$ and $y \in S_j$ with $i \ne j$, and "good" mixing properties hold on each of the subsets S_i.

1.4. **Non-asymptotic bounds from global Poincaré and log Sobolev inequalities.** For $t \geq 0$ and $q \in [1, \infty]$ let us define

$$K_t(q) = \int_0^t \|H_s\|_{L^q(\mu_s)}\, ds.$$

Since $H_s = -\frac{\partial}{\partial s} \log \mu_s$, the quantities $K_t(q)$ are a way to control how much the measures μ_s change for $s \in [0, t]$. A rough estimate yields

$$v_t(p) \;\leq\; 5 \cdot K_t(2) \cdot \sup_{s \in [0,t]} C_{s,t}(4)^2 \qquad \text{for all } p \geq 4, \qquad \text{and} \tag{1.22}$$

$$\bar{C}_t(p, q, \delta) \;\leq\; K_t(q) \sup_{s \in [0, t-\delta]} C_{s,t}(p, q)^2 \qquad \text{for all } q \geq p \geq 1. \tag{1.23}$$

Hence estimates for $v_t(p)$ and $\bar{C}_t(p, q, \delta)$ follow from appropriate L^p-L^q bounds for the Feynman-Kac propagators $q_{s,t}$. In [5], we derive such bounds systematically from Poincaré and logarithmic Sobolev inequalities. To apply these results let us define the weighted Poincaré and log Sobolev constants

$$A_t := \sup_{f \in \mathcal{S}_0} \frac{-\int H_t f^2\, d\mu_t}{\mathcal{E}_t(f)},$$

$$B_t := \sup_{f \in \mathcal{S}_0} \frac{\left| \int H_t f\, d\mu_t \right|^2}{\mathcal{E}_t(f)},$$

$$\gamma_t := \sup_{f \in \mathcal{S}_1} \frac{\int f^2 \log |f|\, d\mu_t}{\mathcal{E}_t(f)},$$

where $\mathcal{S}_0 = \{f : S \to \mathbb{R} |\ \langle f, \mu_t \rangle = 0,\ f \not\equiv 0\}$, $\mathcal{S}_1 = \{f : S \to \mathbb{R} |\ \langle f^2, \mu_t \rangle = 1\}$, and

$$\mathcal{E}_t(f) \;=\; -\int f \mathcal{L}_t f\, d\mu_t \;=\; \frac{1}{2} \sum_{x,y \in S} (f(y) - f(x))^2 \mathcal{L}_t(x, y) \mu_t(x)$$

denotes the Dirichlet form of the self-adjoint operator $\mathcal{L}_t$ on $L^2(S, \mu_t)$. We refer to [14] for background on Poincaré and logarithmic Sobolev inequalities and their applications to estimate L^p contractivity properties of transition semigroups and mixing times of reversible time-homogeneous Markov chains. In [5] we apply similar techniques to derive L^p-L^q bounds for Feynman-Kac propagators. We show that $C_{s,t}(p)$ and $C_{s,t}(p, q)$ are small (in particular less than 2) if the intensities λ_s, $0 \leq s \leq t$, of MCMC moves are sufficiently large in terms of the constants A_s, B_s and γ_s, respectively. By combining these results with Theorem 1.3 we obtain:

Theorem 1.4. *Fix $t_0 \geq 0$, $q \in]6, \infty[$ and $p \in]\frac{4q}{q-2}, q[$. Suppose that*

$$N \geq 120 \cdot \max(K_{t_0}(q), 1), \qquad \text{and}$$

$$\lambda_s \;\geq\; \max\left(\frac{pA_s}{4} + \frac{p(p+3)}{4} t_0 B_s\ ,\ \frac{\gamma_s}{4\delta} \log a(p, q) \right) \qquad \text{for all } s \in [0, t_0], \tag{1.24}$$

where δ is defined by (1.19) and

$$a(p, q) = \max\left(\frac{2r-1}{p-1}, \frac{2p-2}{p-2}, \frac{p-1}{\tilde{p}-1}, \frac{2\tilde{p}-2}{\tilde{p}-2} \right)$$

with $\tilde{p}$ and r determined by $\tilde{p}^{-1} = q^{-1} + 2p^{-1}$ and $p^{-1} = q^{-1} + r^{-1}$. Then for $t \in [0, t_0]$,

$$\sqrt{\varepsilon_t^{N,p}} \;\leq\; (2 + 9\, K_t(2))^{1/2}\, N^{-1/2} + 35\, K_t(q)\, N^{-1} + 11\, K_t(q) \cdot (2 + 9\, K_t(2))^{1/2}\, N^{-3/2}. \tag{1.25}$$

Remark 1.5. (i) The assumptions on p and q guarantee that $\tilde{p} > 2$, so that $a(p, q)$ is finite.
(ii) In particular, if (1.24) holds, then

$$\sqrt{\varepsilon_t^{N,p}} \;\le\; \left(\frac{N}{2+9K_t(q)}\right)^{-1/2} + 4\cdot\left(\frac{N}{2+9K_t(q)}\right)^{-1} + 2\cdot\left(\frac{N}{2+9K_t(q)}\right)^{-3/2}.$$

Therefore a number N of particles of order $O(K_t(q)/\alpha^2)$ is sufficient to ensure that $\sqrt{\varepsilon_t^{N,p}} < \alpha$ for α small enough.

Theorems 1.3 and 1.4 provide non-asymptotic bounds on the variance of the Monte Carlo estimator $\langle f, \nu_t^N\rangle$ that hold uniformly over all functions $f \in L^p(\mu_t)$. One can also combine these bounds with (1.20) and (1.12) to obtain more precise non-asymptotic bounds for the Monte Carlo approximation errors $\langle f, \nu_t^N\rangle$ and $\langle f, \eta_t^N\rangle$ for a fixed function f.

Corollary 1.6. *Suppose that the assumptions of Theorem 1.4 hold, and let $f \in L^p(\mu_t)$. Then*

$$N\cdot\mathbb{E}\left[\left|\langle f,\nu_t^N\rangle - \langle f,\mu_t\rangle\right|^2\right] \;\le\; \mathrm{Var}_{\mu_t}(f) + \int_0^t V_{s,t}(f)\,ds + \|f\|^2_{L^p(\mu_t)} + R(t,N)\cdot\|f\|^2_{L^p(\mu_t)},$$

$$N^{1/2}\cdot\mathbb{E}\left[\left|\langle f,\eta_t^N\rangle - \langle f,\mu_t\rangle\right|\right] \;\le\; \left(\mathrm{Var}_{\mu_t}(f) + \int_0^t V_{s,t}(f)\,ds + \|f-\langle f,\mu_t\rangle\|^2_{L^p(\mu_t)}\right)^{1/2} + \tilde{R}(t,N)\cdot\|f-\langle f,\mu_t\rangle\|_{\sup}$$

with explicit constants $R(t, N)$ and $\tilde{R}(t, N)$ of order $O(N^{1/2})$.

1.5. **Non-asymptotic bounds from local estimates.** It is not very surprising that the empirical distributions are good approximations of the measures μ_t if strong global mixing properties hold. On the other hand, with suitable modifications the above analysis can be also be applied to derive bounds when good mixing properties hold only locally. As an illustration, we consider another extreme case in which the state space is decomposed into several components that are not connected by the underlying Markovian dynamics.

Suppose that

$$S = \bigcup_{i\in I} S_i,$$

is a decomposition of S into disjoint non-empty subsets $\{S_i\}$, $i \in I$, such that

$$\mathcal{L}_t(x,y) \;=\; 0 \qquad \text{for all } t\ge 0,\ x\in S_i \text{ and } y\in S_j \text{ with } i\ne j.$$

Let $\mu_t^i := \mu_t(\cdot|S_i)$ denote the measure μ_t conditioned by S_i. Then we can apply the arguments above with the L^p norm replaced by the stronger norm

$$\|f\|^{\sim}_{L^p(\mu_t)} \;:=\; \max_{i\in I}\|f\|_{L^p(S_i,\mu_t^i)}.$$

Since Hölder's inequality and related estimates hold for these modified L^p norms as well, the assertion of Theorem 1.3 still remains true if $\varepsilon_t^{N,p}$ is replaced by

$$\tilde{\varepsilon}_t^{N,p} \;:=\; \sup\left\{\mathbb{E}\left[\left|\langle f,\nu_t^N\rangle - \langle f,\mu_t\rangle\right|^2\right] \;:\; f: S\to\mathbb{R} \text{ with } \|f\|^{\sim}_{L^p(\mu_t)}\le 1\right\},$$

and the constants $C_{s,t}(p, q)$ and $\bar{C}_t(p, q, \delta)$ are defined w.r.t. the modified L^p and L^q norms as well. Moreover, the representation (1.2) and (1.3) hold for μ_t^i instead of μ_t if H_t is replaced by

$$H_t^i \;:=\; H_t - \langle H_t, \mu_t^i\rangle\,.$$

Let A_t^i, B_t^i and γ_t^i denote the Poincaré and logarithmic Sobolev constants defined as above but with S, μ_t and H_t replaced by S_i, μ_t^i and H_t^i respectively. We set

$$\tilde{A}_t := \max_{i\in I} A_t^i, \qquad \tilde{B}_t := \max_{i\in I} B_t^i, \qquad \tilde{\gamma}_t := \max_{i\in I}\gamma_t^i,$$

$$\tilde{K}_t(q) \;:=\; \int_0^t \|H_s\|_{L^q(\mu_s)}^{\sim}\, ds, \qquad \text{and}$$
$$\tilde{M}_t \;:=\; \max_{i\in I} \sup_{0\le r\le s\le t} \frac{\mu_s(S_i)}{\mu_r(S_i)}\,.$$

Then, by estimating L^p norms separately on each component, we can prove the following extension of Theorem 1.4:

Theorem 1.7. *Fix $t_0 \ge 0$, $q \in]6,\infty[$ and $p \in]\frac{4q}{q-2}, q[$. Suppose that*

$$N \ge 120 \cdot \max(\tilde{K}_{t_0}(q), 1), \qquad \textit{and}$$
$$\lambda_s \;\ge\; \max\left(\frac{p\tilde{A}_s}{4} + \frac{p(p+3)}{4} t_0 \tilde{B}_s \;,\; \frac{\tilde{\gamma}_s}{4\delta} \log a(p,q) \right) \qquad \textit{for all } s \in [0,t_0]. \tag{1.26}$$

Then for $t \in [0,t_0]$,

$$\sqrt{\tilde{\varepsilon}_t^{N,p}} \;\le\; (2+9K_t(2))^{1/2}\, \tilde{M}_t\, N^{-1/2} + 35\, \tilde{K}_t(q)\, \tilde{M}_t^2\, N^{-1} + 11\, \tilde{K}_t(q)\, (2+9K_t(2))^{1/2}\, N^{-3/2}.$$

Remark 1.8. (i) If there is only one component, the assertion of Theorem 1.7 reduces to that of Theorem 1.4.
(ii) Error bounds for the estimators $\langle f, \nu_t^N\rangle$ and $\langle f, \eta_t^N\rangle$ for a fixed function f hold analogously to Corollary 1.6.

1.6. **Open problems.** The cases discussed in Sections 1.4 and 1.5 are extreme cases. In many typical applications, one would expect the state space to split up as time evolves into more and more components that get almost disconnected by the dynamics (local modes, metastable states). The study of such more complicated situations is an important topic for future research.

1.7. **Outline.** The rest of this paper contains the proofs of the results above. We will first in Section 2 derive by martingale arguments an explicit formula for the variances of the estimators $\langle f, \nu_t^N\rangle$, cf. Theorem 2.1 below. In Section 3 we apply this formula to prove Theorem 1.3. Finally, Section 4 contains the proofs of Theorem 1.4 and Theorem 1.7.

2. Variances of weighted empirical averages

In this section we will prove the following theorem which shows that $\langle f, \nu_t^N\rangle$ is an unbiased estimator for $\langle f, \mu_t\rangle$ and gives an explicit formula for the variance:

Theorem 2.1. *For all $f : S \to \mathbb{R}$,*

$$\mathbb{E}\left[\langle f, \nu_t^N\rangle\right] \;=\; \langle f, \mu_t\rangle, \qquad \textit{and}$$
$$\mathbb{E}\left[\left|\langle f, \nu_t^N\rangle - \langle f, \mu_t\rangle\right|^2\right] \;=\; \frac{1}{N}\operatorname{Var}_{\mu_t}(f) + \frac{1}{N}\int_0^t \mathbb{E}\left[V_{s,t}^N(f)\right]\, ds\,,$$

where

$$\begin{aligned} V_{s,t}^N(f) = &- \langle H_s(q_{s,t}f)^2, \nu_s^N\rangle\langle 1, \nu_s^N\rangle - \langle H_s, \nu_s^N\rangle\langle q_{s,t}f^2 - (q_{s,t}f)^2, \nu_s^N\rangle \\ &+ \frac{1}{2}\iint |H_s(z) - H_s(y)|(q_{s,t}f(z) - q_{s,t}f(y))^2\, \nu_s^N(dy)\, \nu_s^N(dz). \end{aligned} \tag{2.1}$$

The proof of Theorem 2.1 relies on the identification of appropriate martingales. Recall that the Carré du champ (square field) operator Γ_t^N associated to $\mathcal{L}_t^N$ is defined for functions $\varphi : S^N \to \mathbb{R}$ by

$$\Gamma_t^N(\varphi) \;=\; \mathcal{L}_t^N\varphi^2 - 2\varphi\mathcal{L}_t^N\varphi,$$

i.e.,

$$\Gamma_t^N(\varphi)(x) \;=\; \sum_{y\in S} \mathcal{L}_t^N(x,y)\cdot(\varphi(y)-\varphi(x))^2 \qquad \text{for all } x \in S^N. \tag{2.2}$$

It is well-known that the processes

$$M_t^\varphi = \varphi(t, X_t^N) - \varphi(0, X_0^N) - \int_0^t (\frac{\partial}{\partial s} + \mathcal{L}_s^N)\varphi(s, X_s^N)\, ds, \qquad \text{and} \tag{2.3}$$

$$N_t^\varphi = (M_t^\varphi)^2 - \int_0^t \Gamma_s^N(\varphi(s,\,\cdot\,))(X_s^N)\, ds \tag{2.4}$$

are martingales w.r.t. the filtration induced by the process X_t^N for any function $\varphi : \mathbb{R}^+ \times S^N \to \mathbb{R}$ that is twice continuously differentiable in the first variable, cf. e.g. Kipnis and Landim [10], Appendix 1, Lemma 5.1. For $x \in S^N$ let

$$\eta(x) = \frac{1}{N}\sum_{i=1}^N \delta_{X_i^N}$$

denote the corresponding empirical average. In the next lemma we derive expressions for $\mathcal{L}_t^N$ and Γ_t^N acting on linear functions on S^N of the form

$$\varphi_f(x) = \langle f, \eta(x)\rangle = N^{-1}\sum_{i=1}^N f(x_i).$$

Lemma 2.2. *For any function $f : S \to \mathbb{R}$ and $t \geq 0$, one has*

$$\mathcal{L}_t^N \langle f, \eta\rangle = \lambda_t \langle \mathcal{L}_t f, \eta\rangle + \langle H_t, \eta\rangle\langle f, \eta\rangle - \langle H_t f, \eta\rangle$$

and

$$\Gamma_t^N(\langle f, \eta\rangle) = \frac{\lambda_t}{N}\langle \Gamma_t(f), \eta\rangle + \frac{1}{N}\iint \big(H_t(y) - H_t(z)\big)^+ \big(f(z) - f(y)\big)^2\, \eta(dy)\,\eta(dz),$$

where Γ_t denotes the Carré du champ operator w.r.t. $\mathcal{L}_t$.

Proof. The definition of $\mathcal{L}_t^N$ immediately yields

$$\mathcal{L}_t^N \langle f, \eta\rangle(x) = \frac{\lambda_t}{N}\sum_{i=1}^N \mathcal{L}_t f(x_i) + \frac{1}{N^2}\sum_{i,j=1}^N (H_t(x_i) - H_t(x_j))^+ (f(x_j) - f(x_i)). \tag{2.5}$$

Moreover,

$$\begin{aligned}
\sum_{i,j=1}^N (H_t(x_i) - H_t(x_j))^+ (f(x_j) - f(x_i)) &= \sum_{i,j: H_t(x_i) > H_t(x_j)} (H_t(x_i) - H_t(x_j))(f(x_j) - f(x_i)) \\
&= \sum_{i,j: H_t(x_j) > H_t(x_i)} (H_t(x_j) - H_t(x_i))(f(x_i) - f(x_j)) \\
&= \sum_{i,j: H_t(x_j) > H_t(x_i)} (H_t(x_i) - H_t(x_j))(f(x_j) - f(x_i)) \\
&= -\sum_{i,j=1}^N (H_t(x_i) - H_t(x_j))^- (f(x_j) - f(x_i)),
\end{aligned}$$

and hence

$$\sum_{i,j=1}^N (H_t(x_i) - H_t(x_j))(f(x_j) - f(x_i)) = 2\sum_{i,j=1}^N (H_t(x_i) - H_t(x_j))^+ (f(x_j) - f(x_i)).$$

Therefore the second term on the right hand side of (2.5) is equal to

$$\begin{aligned}
\frac{1}{2N^2}\sum_{i,j=1}^N (H_t(x_i) - H_t(x_j))(f(x_j) - f(x_i)) &= \Big(\frac{1}{N}\sum_{i=1}^N H_t(x_i)\Big)\Big(\frac{1}{N}\sum_{j=1}^N f(x_j)\Big) - \frac{1}{N}\sum_{i=1}^N H_t(x_i) f(x_i) \\
&= \langle H_t, \eta(x)\rangle\langle f, \eta(x)\rangle - \langle H_t f, \eta(x)\rangle,
\end{aligned}$$

from which the first claim follows.

Furthermore, since

$$\langle f,\eta(x^{i\to j})\rangle - \langle f,\eta(x)\rangle \;=\; N^{-1}\cdot\left(f(x_j)-f(x_i)\right),$$

(2.2) and (1.7) imply

$$\begin{aligned}\Gamma_t^N\langle f,\eta\rangle(x) \;&=\; \frac{\lambda_t}{N^2}\sum_{i=1}^N\sum_{y\in S}\mathcal{L}_t(x_i,y)(f(y)-f(x_i))^2\\ &\quad+\frac{1}{N^3}\sum_{i,j=1}^N\big(H_t(x_i)-H_t(x_j)\big)^+\big(f(x_j)-f(x_i)\big)^2,\end{aligned}$$

from which the second claim follows noting that the first term on the right hand side of the previous expression is equal to

$$\frac{\lambda_t}{N^2}\sum_{i=1}^N\Gamma_t(f)(x_i) \;=\; \frac{\lambda_t}{N}\langle\Gamma_t(f),\eta(x)\rangle. \qquad\square$$

Now let us define

$$\bar{A}^f_{s,t} \;=\; \langle q_{s,t}f,\eta^N_s\rangle \;=\; \frac{1}{N}\sum_{i=1}^N(q_{st}f)(X^N_{s,i})\,.$$

As a consequence of Lemma 2.2 we obtain:

Proposition 2.3. *The processes $\bar{M}^f_u$ and $\bar{N}^f_u$, $u\in[0,t]$, defined by*

$$\begin{aligned}\bar{M}^f_u \;&=\; \bar{A}^f_{u,t}-\bar{A}^f_{0,t}-\int_0^u\langle H_s,\eta^N_s\rangle\langle q_{s,t}f,\eta^N_s\rangle\,ds,\\ \bar{N}^f_u \;&=\; (\bar{M}^f_u)^2-\frac{1}{N}\int_0^u\lambda_s\,\langle\Gamma_s(q_{s,t}f),\eta^N_s\rangle\,ds\\ &\quad-\frac{1}{N}\int_0^u\iint\big(H_s(y)-H_s(z)\big)^+\big(q_{s,t}f(z)-q_{s,t}f(y)\big)^2\,\eta^N_s(dy)\,\eta^N_s(dz)\,ds\end{aligned}$$

are martingales w.r.t. the filtration $\mathcal{F}_t=\sigma(X^N_s\mid s\in[0,t])$.

Proof. Note that $\bar{A}^f_s=\varphi(s,X^N_s)$, where

$$\varphi(s,x) \;=\; N^{-1}\sum_{i=1}^N q_{st}f(x_i).$$

By the backward equation (1.14),

$$\begin{aligned}\frac{\partial}{\partial s}\varphi(s,x) \;&=\; -\frac{\lambda_s}{N}\sum_{i=1}^N\mathcal{L}_s q_{st}f(x_i)+\frac{1}{N}\sum_{i=1}^N H_s q_{st}f(x_i)\\ &=\; -\lambda_s\langle\mathcal{L}_s q_{st}f,\eta(x)\rangle+\langle H_s q_{st}f,\eta(x)\rangle,\end{aligned}$$

and by lemma 2.2,

$$\left(\mathcal{L}^N_s\varphi\right)(s,x) \;=\; \lambda_s\langle\mathcal{L}_s q_{s,t}f,\eta(x)\rangle+\langle H_s,\eta(x)\rangle\langle q_{s,t}f,\eta(x)\rangle-\langle H_s q_{s,t}f,\eta(x)\rangle$$

Hence

$$\left(\frac{\partial}{\partial s}+\mathcal{L}^N_s\right)\varphi(s,x) \;=\; \langle H_s,\eta(x)\rangle\langle q_{s,t}f,\eta(x)\rangle,$$

which proves that $\bar{M}^f = M^\varphi$ is a martingale, cf. (2.3). Similarly, by Lemma 2.2,

$$\begin{aligned}\Gamma_s^N(\varphi)(s,x) &= \frac{\lambda_s}{N}\langle \Gamma_s(q_{s,t}f), \eta(x)\rangle \\ &+\frac{1}{N}\iint \big(H_s(y)-H_s(z)\big)^+\big(q_{s,t}f(z)-q_{s,t}f(y)\big)^2\,\eta_s^N(dy)\,\eta_s^N(dz),\end{aligned}$$

which proves that $\bar{N}^f = N^\varphi$ is a martingale, cf. (2.4). □

Since in general, $\bar{A}^f_{s,t}$ is not a martingale, $\langle f, \eta_t^N\rangle$ is not an unbiased estimator for $\langle f, \mu_t\rangle$. This motivates considering $\langle f, \nu_t^N\rangle$ instead. Let

$$A^f_{s,t} \;=\; \langle q_{s,t}f, \nu_s^N\rangle \;=\; e^{-\int_0^s \langle H_r, \eta_r^N\rangle\,dr}\,\bar{A}^f_{s,t}. \tag{2.6}$$

Proposition 2.4. *The process $A^f_{u,t}$, $u\in[0,t]$, is a martingale with increasing process given by*

$$\begin{aligned}\langle A^f_{\bullet,t}\rangle_u &= \frac{1}{N}\int_0^u \lambda_s\,\langle 1,\nu_s^N\rangle\langle \Gamma_s(q_{s,t}f),\nu_s^N\rangle\,ds \\ &+\frac{1}{N}\int_0^u\iint \big(H(x)-H(y)\big)^+\big(q_{s,t}f(y)-q_{s,t}f(x)\big)^2\,\nu_s^N(dx)\,\nu_s^N(dy)\,ds.\end{aligned}$$

Proof. By the integration by parts formula for Stieltjes integrals and Proposition 2.3, we get

$$\begin{aligned}A^f_{u,t}-A^f_{0,t} &= \int_0^u e^{-\int_0^s\langle H_r,\eta_r^N\rangle\,dr}d\bar{A}^f_{s,t} - \int_0^u \langle H_s,\eta_s^N\rangle e^{-\int_0^s\langle H_r,\eta_r^N\rangle\,dr}\bar{A}^f_{s,t}\,ds \\ &= \int_0^u e^{-\int_0^s\langle H_r,\eta_r^N\rangle\,dr}d\bar{M}^f_s + \langle H_s,\eta_s^N\rangle A^f_s\,ds - \langle H_s,\eta_s^N\rangle A^f_s\,ds.\end{aligned}$$

Hence $A^f_{s,t}$, $s\in[0,t]$, is a martingale, whose increasing process can be written as

$$\langle A^f_{\bullet,t}\rangle_u \;=\; \int_0^u e^{-2\int_0^s\langle H_r,\eta_r^N\rangle\,dr}\,d\langle \bar{M}^f\rangle_s.$$

The result now follows by Proposition 2.3 and Equation (1.9). □

The purpose of the next lemma is to obtain an alternative representation (mudulo martingale terms) of the term involving the Carré du champ operator in the expression for $\langle A^f_{\bullet,t}\rangle$.

Lemma 2.5. *The following decomposition holds:*

$$\begin{aligned}&\int_0^u \lambda_s\langle 1,\nu_s^N\rangle\langle\Gamma_s(q_{st}f),\nu_s^N\rangle\,ds \\ &\quad= \tilde{M}_u + \langle 1,\nu_u^N\rangle\,\big\langle (q_{ut}f)^2,\nu_u^N\big\rangle + \int_0^u\langle H_s,\nu_s^N\rangle\langle (q_{st}f)^2,\nu_s^N\rangle\,ds \\ &\qquad-\int_0^u\langle 1,\nu_s^N\rangle\langle H_s(q_{st}f)^2,\nu_s^N\rangle\,ds,\end{aligned}$$

where $\tilde{M}$ is a martingale.

Proof. Let

$$Y_u := e^{-2\int_0^u\langle H_r,\eta_r^N\rangle\,dr}\,\big\langle (q_{ut}f)^2,\eta_u^N\big\rangle = \langle 1,\nu_u^N\rangle\langle (q_{ut}f)^2,\nu_u^N\rangle.$$

By applying the martingale problem to the functions $\varphi(s,x) = \big\langle (q_{st}f)^2,\eta(x)\big\rangle$, we obtain that the following processes differ only by a martingale term:

$$\begin{aligned}e^{-2\int_0^u\langle H_r,\eta_r^N\rangle\,dr}\,\big\langle (q_{ut}f)^2,\eta_u^N\big\rangle &\sim -2\int_0^u e^{-2\int_0^s\langle H_r,\eta_r^N\rangle\,dr}\langle H_s,\eta_s^N\rangle\,\big\langle (q_{st}f)^2,\eta_s^N\big\rangle\,ds \\ &+\int_0^u e^{-2\int_0^s\langle H_r,\eta_r^N\rangle\,dr}\left(\frac{\partial}{\partial s}+\mathcal{L}_s^N\right)\varphi(s,X_s^N)\,ds.\end{aligned}$$

Proceeding as in the proof of proposition 2.3, we get that

$$\frac{\partial}{\partial s}\varphi(s,X_s^N) \;=\; 2\langle q_{st}f\frac{\partial}{\partial s}q_{st}f,\eta_s^N\rangle \;=\; -2\lambda_s\langle q_{st}f\mathcal{L}_s q_{st}f,\eta_s^N\rangle + 2\langle H_s(q_{st}f)^2,\eta_s^N\rangle,$$

and

$$\mathcal{L}_s^N\varphi(s,X_s^N) \;=\; \lambda_s\langle\mathcal{L}_s(q_{st}f)^2,\eta_s^N\rangle + \langle H_s,\eta_s^N\rangle\langle(q_{st}f)^2,\eta_s^N\rangle - \langle H_s(q_{st}f)^2,\eta_s^N\rangle.$$

Recalling that $\mathcal{L}_s(q_{st}f)^2 - 2q_{st}f\mathcal{L}_s q_{st}f = \Gamma_s(q_{s,t}f)$ and $\nu_s^N = \exp(-\int_0^s\langle H_r,\nu_r^N\rangle dr)\,\eta_s^N$, we obtain

$$\begin{aligned}\langle 1,\nu_u^N\rangle\langle(q_{ut}f)^2,\nu_u^N\rangle \;\sim\; & -\int_0^u\langle H_s,\nu_s^N\rangle\langle(q_{st}f)^2,\nu_s^N\rangle\,ds\\ & +\int_0^u\langle 1,\nu_s^N\rangle\langle H_s(q_{st}f)^2,\nu_s^N\rangle\,ds + \int_0^u\lambda_s\,\langle 1,\nu_s^N\rangle\langle\Gamma_s(q_{st}f),\nu_s^N\rangle\,ds,\end{aligned}$$

which proves the assertion. □

Lemma 2.6. *For all $t\geq 0$,*

$$\mathbb{E}\Big[\langle 1,\nu_t^N\rangle\langle f^2,\nu_t^N\rangle\Big] \;=\; \langle f^2,\mu_t\rangle - \mathbb{E}\left[\int_0^t\langle H_s,\nu_s^N\rangle\langle q_{st}f^2,\nu_s^N\rangle\,ds\right].$$

Proof. By the product rule for Stieltjes integrals,

$$\begin{aligned}\langle 1,\nu_s^N\rangle\langle q_{st}f^2,\nu_s^N\rangle &= e^{-\int_0^s\langle H_r,\eta_r^N\rangle\,dr}A_{s,t}^{f^2}\\ &= \int_0^s e^{-\int_0^u\langle H_r,\eta_r^N\rangle\,dr}dA_{u,t}^{f^2} \;-\; \int_0^s\langle H_u,\nu_u^N\rangle A_{u,t}^{f^2}\,du.\end{aligned}$$

Since $s\mapsto A_{s,t}^{f^2}$ is a martingale,

$$\mathbb{E}\left[\langle 1,\nu_t^N\rangle\langle f^2,\nu_t^N\rangle\right] \;=\; \langle q_{0,t}f^2,\mu_0\rangle - \mathbb{E}\left[\int_0^t\langle H_u,\nu_u^N\rangle A_{u,t}^{f^2}\,du\right].$$

The proof is completed by noting that $\langle q_{0,t}f^2,\mu_0\rangle = \langle f^2,\mu_t\rangle$. □

Proof of Theorem 2.1. Fix a function $f:S\to\mathbb{R}$ and $t\geq 0$. Recalling that by (1.17), $\langle f,\mu_t\rangle = \langle q_{0,t}f,\mu_0\rangle$, we have

$$\begin{aligned}\langle f,\nu_t^N\rangle - \langle f,\mu_t\rangle &= \langle q_{t,t}f,\nu_t^N\rangle - \langle q_{0,t}f,\nu_0^N\rangle + \langle q_{0,t}f,\nu_0^N\rangle - \langle q_{0,t}f,\mu_0\rangle\\ &= A_{t,t}^f - A_{0,t}^f + \langle q_{0,t}f,\nu_0^N\rangle - \langle q_{0,t}f,\mu_0\rangle.\end{aligned}$$

Taking expectations on both sides, we immediately obtain

$$\mathbb{E}\left[\langle f,\nu_t^N\rangle\right] \;=\; \langle f,\mu_t\rangle,$$

because $s\mapsto A_{s,t}f$ is a martingale by Proposition 2.4, and ν_0^N is the empirical distribution of N i.i.d. random variables with distribution μ_0. Moreover, by Proposition 2.4 and Lemma 2.5,

$$\begin{aligned}N\cdot\mathbb{E}\left[\left|\langle f,\nu_t^N\rangle - \langle f,\mu_t\rangle\right|^2\right] \;&=\; N\cdot\mathbb{E}\left[(A_{t,t}^f - A_{0,t}^f)^2\right] \;+\; N\cdot\mathbb{E}\left[\left(\langle q_{0,t}f,\nu_0^N\rangle - \langle q_{0,t}f,\mu_0\rangle\right)^2\right]\\ =\;& N\cdot\mathbb{E}\left[\langle A_{\bullet,t}^f\rangle_t\right] + \mathrm{Var}_{\mu_0}(q_{0,t}f)\\ =\;& \mathbb{E}\Big[\langle 1,\nu_t^N\rangle\langle f^2,\nu_t^N\rangle - \langle(q_{0,t}f)^2,\nu_0^N\rangle\Big] \;+\; \mathrm{Var}_{\mu_0}(q_{0,t}f)\\ & +\mathbb{E}\int_0^t\langle H_s,\nu_s^N\rangle\langle(q_{st}f)^2,\nu_s^N\rangle\,ds \;-\; \mathbb{E}\int_0^t\langle 1,\nu_s^N\rangle\langle H_s(q_{st}f)^2,\nu_s^N\rangle\,ds\\ & +\mathbb{E}\int_0^t\iint\big(H(x)-H(y)\big)^+\big(q_{s,t}f(y)-q_{s,t}f(x)\big)^2\,\nu_s^N(dx)\,\nu_s^N(dy)\,ds.\end{aligned}$$

The assertion now follows from Lemma 2.6 by noting that

$$\begin{aligned} -\mathbb{E}\Big[\langle (q_{0,t}f)^2, \nu_0^N\rangle\Big] \;+\; \mathrm{Var}_{\mu_0}(q_{0,t}f) &= -\langle (q_{0,t}f)^2, \mu_0\rangle \;+\; \mathrm{Var}_{\mu_0}(q_{0,t}f) \\ &= \langle q_{0,t}f, \mu_0\rangle^2 \;=\; \langle f, \mu_t\rangle^2. \end{aligned}$$

□

3. Proof of Theorem 1.3

Proposition 3.1. *Let p, q, $r \in [1,\infty]$ be such that $p^{-1} = q^{-1} + r^{-1}$. Then for $0 \le s \le t$,*

$$\begin{aligned} \mathbb{E}\left[V_{s,t}^N(f)\right] \;&\le V_{s,t}(f) + 19\|H_s\|_{L^q(\mu_s)} \max\left(\|q_{s,t}f\|^2_{L^{2r}(\mu_s)}, \|f\|^2_{L^2(\mu_t)}\right)\sqrt{\varepsilon_s^{N,p}} \\ &\quad + \left(10\|H_s\|_{L^q(\mu_s)}\|q_{s,t}f\|^2_{L^{2r}(\mu_s)} + \|H_s\|_{L^p(\mu_s)}\|q_{s,t}f^2\|_{L^p(\mu_s)}\right)\varepsilon_s^{N,p}. \end{aligned}$$

Proof. Let us first observe that by the Cauchy-Schwarz inequality,

$$\begin{aligned} &\mathbb{E}\left[\left|\langle f, \nu_s^N\rangle\langle g, \nu_s^N\rangle - \langle f, \mu_s\rangle\langle g, \mu_s\rangle\right|\right] \\ &\le\;\; (\varepsilon_t^{N,p})^{1/2}\left(|\langle f, \mu_s\rangle|\cdot\|g\|_{L^p(\mu_s)} + |\langle g, \mu_s\rangle|\cdot\|f\|_{L^p(\mu_s)}\right) \;+\; \varepsilon_t^{N,p}\|f\|_{L^r(\mu_s)}\|g\|_{L^q(\mu_s)} \end{aligned} \tag{3.1}$$

for all $0 \le s \le t$ and all functions f, $g : S \to \mathbb{R}$. Since the last term on the right-hand side of (2.1) can be estimated by

$$\iint |H_s(y)|(q_{s,t}f(z) - q_{s,t}f(y))^2\,\nu_s^N(dz)\,\nu_s^N(dy),$$

an application of (3.1) yields

$$\begin{aligned} \mathbb{E}\left[V_{s,t}^N(f)\right] \;\le\; &-\langle H_s(q_{s,t}f)^2, \mu_s\rangle\langle 1, \mu_s\rangle - \langle H_s, \mu_s\rangle\langle q_{s,t}f^2 - (q_{s,t}f)^2, \mu_s\rangle \\ &+ \iint |H_s(y)|(q_{s,t}f(z) - q_{s,t}f(y))^2\,\mu_s(dz)\,\mu_s(dy) \;+\; \sqrt{\varepsilon_s^{N,p}}R_{s,t}^N(f) + \varepsilon_s^{N,p}\bar{R}_{s,t}^N(f), \end{aligned}$$

where

$$\begin{aligned} R_{s,t}^N(f) \;=\;& \|H_s(q_{s,t}f)^2\|_{L^p(\mu_s)} + \langle H_s(q_{s,t}f)^2, \mu_s\rangle + \|H_s\|_{L^p(\mu_s)}|\langle f^2, \mu_t\rangle - \langle (q_{s,t}f)^2, \mu_s\rangle| \\ &+ 4\|H_s(q_{s,t}f)^2\|_{L^p(\mu_s)} + 4\langle |H_s|(q_{s,t}f)^2, \mu_s\rangle + 4\|H_s\|_{L^p(\mu_s)}\langle (q_{s,t}f)^2, \mu_s\rangle \\ &+ 4\langle |H_s|, \mu_s\rangle\|(q_{s,t}f)^2\|_{L^p(\mu_s)} \\ \le\;& 19\|H_s\|_{L^q(\mu_s)}\max\left(\|q_{s,t}f\|^2_{L^{2r}(\mu_s)}, \|f\|^2_{L^2(\mu_t)}\right), \end{aligned}$$

and

$$\begin{aligned} \bar{R}_{s,t}^N(f) \;=\;& \|H_s(q_{s,t}f)^2\|_{L^p(\mu_s)} + \|H_s\|_{L^p(\mu_s)}\|q_{s,t}f^2 - (q_{s,t}f)^2\|_{L^p(\mu_s)} \\ &+ 4\|H_s(q_{s,t}f)^2\|_{L^p(\mu_s)} + 4\|H_s\|_{L^p(\mu_s)}\|(q_{s,t}f)^2\|_{L^p(\mu_s)} \\ \le\;& \|H_s\|_{L^p(\mu_s)}\|q_{s,t}f^2\|_{L^p(\mu_s)} + 10\|H_s\|_{L^q(\mu_s)}\|q_{s,t}f\|^2_{L^{2r}(\mu_s)}. \end{aligned}$$

Here we have used the elementary inequality

$$\left(q_{s,t}f(y) - q_{s,t}f(x)\right)^2 \le 2\left(q_{s,t}f(y)^2 + q_{s,t}f(x)^2\right)$$

and the fact that $\langle q_{s,t}f, \mu_s\rangle = \langle f, \mu_t\rangle$. □

In order to bound $V_{s,t}^N(f)$ uniformly over $f \in L^p(\mu_t)$ with $\|f\|_{L^p(\mu_t)} \le 1$, one needs to be able to control $\|q_{s,t}f\|_{L^{2r}(\mu_t)}$ in terms of $\|f\|_{L^p(\mu_t)}$. This is possible if hypercontractivity holds and $t - s$ is sufficiently large. Over short time intervals $[s,t]$ we apply in a first step another rough estmate instead:

Lemma 3.2. *Let $p \geq 2$ and $N \in \mathbb{N}$. Then for $0 \leq s \leq t$,*

$$\frac{1}{N} \cdot \mathbb{E}\left[V_{s,t}^N(f)\right] \;\leq\; 4\,\mathrm{osc}(H_s)\big(1 + (\varepsilon_s^{N,p})^{1/2} \exp \int_s^t \mathrm{osc}(H_r)\, dr\big)^2 \|f\|^2_{L^p(\mu_t)}.$$

Proof. Setting

$$A_t^f \;:=\; \langle f, \nu_t^N \rangle \;=\; \langle f, \eta_t^N \rangle \cdot \exp\Big(- \int_0^t \langle H_s, \eta_s^N \rangle \, ds \Big),$$

we have $A_t^f = \langle f, \eta_t^N \rangle \cdot A_t^1$ for all $f : S \to \mathbb{R}$. Since

$$\langle f^2, \eta_t^N \rangle = \frac{1}{N} \sum_{i=1}^N f(X_i)^2 \leq \frac{1}{N} \Big(\sum_{i=1}^N |f(X_i)| \Big)^2 = N \langle |f|, \eta_t^N \rangle^2,$$

we obtain, recalling that η_t^N is a probability measure,

$$\begin{aligned} V_{s,t}^N(f) \;&\leq\; N \cdot (A_s^1)^2 \cdot \Big((\max H_s^- + \max H_s^+) \langle |q_{s,t} f|, \eta_s^N \rangle^2 + \max H_s^- \langle (q_{s,t} f^2)^{1/2}, \eta_s^N \rangle^2 \\ &\qquad + 2\,\mathrm{osc}(H_s) \langle |q_{s,t} f|, \eta_s^N \rangle^2 \Big) \\ &\leq\; N\, \mathrm{osc}(H_s) \Big(3 \langle q_{s,t} |f|, \nu_s^N \rangle^2 + \langle (q_{s,t} f^2)^{1/2}, \nu_s^N \rangle^2 \Big). \end{aligned} \tag{3.2}$$

Moreover, observing that inequality (3.1) implies

$$\mathbb{E}\left[\langle f, \nu_t^N \rangle^2 \right] \;\leq\; \Big(|\langle f, \mu_t \rangle| + \sqrt{\varepsilon_t^{N,p}} \|f\|_{L^p(\mu_t)} \Big)^2,$$

we obtain, taking expectations on both sides of (3.2),

$$\begin{aligned} \frac{1}{N} \cdot \mathbb{E}\left[V_{s,t}^N(f)\right] \;&\leq\; 3\,\mathrm{osc}(H_s) \big[\langle q_{s,t}|f|, \mu_s \rangle + (\varepsilon_s^{N,p})^{1/2} \big\| q_{s,t}|f| \big\|_{L^p(\mu_s)} \big]^2 \\ &\quad + \mathrm{osc}(H_s) \big[\langle q_{s,t} f^2, \mu_s \rangle^{1/2} + (\varepsilon_s^{N,p})^{1/2} \big\| q_{s,t} f^2 \big\|_{L^{p/2}(\mu_t)} \big]^2 \\ &\leq\; 4\,\mathrm{osc}(H_s) \cdot \left[\langle f^2, \mu_t \rangle (\varepsilon_s^{N,p})^{1/2} \exp\left(\int_s^t \mathrm{osc}(H_r)\, dr \right) \cdot \|f\|_{L^p(\mu_t)} \right]^2, \end{aligned}$$

where we have used the fact that $\langle q_{s,t} f, \mu_s \rangle = \langle f, \mu_t \rangle$, and the estimate

$$\big\| q_{s,t} f \big\|_{L^p(\mu_t)} \;\leq\; \exp\left(\int_s^t \mathrm{osc}(H_r)\, dr \right) \cdot \|f\|_{L^p(\mu_s)}. \tag{3.3}$$

The proof of (3.3) is elementary, cf. [5]. □

Remark 3.3. Note that, in particular, by (1.19),

$$\frac{1}{N} \int_{(t-\delta)^+}^t \mathbb{E}\left[V_{s,t}^N(f)\right] ds \;\leq\; \frac{4}{35} \left(1 + e^{1/35} \right)^2 \cdot \|f\|^2_{L^p(\mu_t)} \sup_{s<t} \varepsilon_s^{N,p} \vee 1 \;\leq\; \frac{1}{2} \|f\|^2_{L^p(\mu_t)} \sup_{s<t} \varepsilon_s^{N,p} \vee 1. \tag{3.4}$$

Combining Proposition 3.1 and Lemma 3.2 we obtain the following (rough) a priori estimate:

Lemma 3.4. *Let p, q, $r \in [2, \infty]$ be such that $p^{-1} = q^{-1} + r^{-1}$. If*

$$N \geq 80 \cdot \max\big(1, \bar{C}_t(p, q, \delta)\big) \qquad \textit{for all } t \in [0, t_0],$$

then

$$\varepsilon_t^{N,p} < 1 \qquad \textit{for all } t \in [0, t_0].$$

Proof. Proposition 3.1 implies

$$\mathbb{E}\left[V_{s,t}^N(f)\right] \;\le\; \|H_s\|_{L^q(\mu_s)} C_{s,t}(p,q)^2 \|f\|^2_{L^p(\mu_t)} \big(9 + 19(\varepsilon_s^{N,p})^{1/2} + 11\varepsilon_s^{N,p}\big)$$

for all $f : S \to \mathbb{R}$ and $0 \le s \le t$. Choosing N as stated we get, for $t \le t_0$,

$$\frac{1}{N}\int_0^{(t-\delta)^+} \mathbb{E}\left[V_{s,t}^N(f)\right]\, ds \;\le\; \frac{39}{80}\|f\|^2_{L^p(\mu_t)} \sup_{s<t} \varepsilon_s^{N,p} \vee 1.$$

Recalling now (3.4) and applying theorem 2.1 we get

$$\begin{aligned}\varepsilon_t^{N,p} \;&=\; \sup\left\{\frac{1}{N}\operatorname{Var}_{\mu_t}(f) + \frac{1}{N}\int_0^t \mathbb{E}\left[V_{s,t}^N(f)\right]\, ds \;:\; f : S \to \mathbb{R} \text{ with } \|f\|_{L^p(\mu_t)} \le 1\right\}\\ &<\; \frac{1}{80} + \left(\frac{39}{80} + \frac{1}{2}\right)\cdot \sup_{s<t} \varepsilon_s^{N,p} \vee 1\end{aligned}$$

for all $t \in [0, t_0]$. The claim follows since S is finite, and therefore the functions $t \mapsto \varepsilon_t^{N,p}$ are right continuous. □

The a priori estimate just obtained can be used instead of Lemma 3.2 to estimate $\mathbb{E}\left[V_{s,t}^N(f)\right]$ when $t - s$ is small:

Lemma 3.5. *Let $q \in]6,\infty]$ and $p \in]4q/(q-2),\infty[$. Suppose that*

$$N \;\ge\; 80\max\big(1, \bar{C}_t(\tilde{p}, q, \delta)\big) \qquad \textit{for all } t \in [0, t_0],$$

where $\tilde{p}$ is defined by $\tilde{p}^{-1} = q^{-1} + (p/2)^{-1}$. Then for $0 \le s \le t \le t_0$,

$$\mathbb{E}\left[V_{s,t}^N(f)\right] \;\le\; V_{s,t}(f) + 30\,\exp\left(2\int_s^t \operatorname{osc} H_r\, dr\right)\|H_s\|_{L^q(\mu_s)}\|f\|^2_{L^p(\mu_t)}.$$

Proof. Noting that $\tilde{p}^{-1} = q^{-1} + (p/2)^{-1} < 1/2$ by the assumptions on p and q, Proposition 3.1 yields

$$\begin{aligned}\mathbb{E}\left[V_{s,t}^N(f)\right] \;&\le V_{s,t}(f) + 19\,\|H_s\|_{L^q(\mu_s)} \max\big(\|q_{s,t}f\|^2_{L^p(\mu_s)}, \|f\|^2_{L^2(\mu_t)}\big)(\varepsilon_s^{N,p})^{1/2}\\ &\quad+\; \big[10\,\|H_s\|_{L^q(\mu_s)}\|q_{s,t}f\|^2_{L^p(\mu_s)} + \|H_s\|_{L^{\tilde{p}}(\mu_s)}\|q_{s,t}f^2\|_{L^{\tilde{p}}(\mu_s)}\big]\varepsilon_s^{N,p}.\end{aligned}$$

Since $\tilde{p} < \min(q, p/2)$, the claim follows by Lemma 3.4 and the estimate (3.3). □

We are now ready to prove the theorem:

Proof of Theorem 1.3. For $t \in [0, t_0]$ let

$$\bar{\varepsilon}_t^{N,p} \;:=\; \sup_{s \le t} \varepsilon_s^{N,p}.$$

By Proposition 3.1 we have

$$\mathbb{E}\left[V_{s,t}^N(f)\right] \;\le\; V_{s,t}(f) + \|H_s\|_{L^q(\mu_s)} C_{s,t}(p,q)^2 \|f\|^2_{L^p(\mu_t)}\big(19(\bar{\varepsilon}_t^{N,p})^{1/2} + 11\bar{\varepsilon}_t^{N,p}\big)$$

for all $f : S \to \mathbb{R}$ and $0 \le s \le t$. Therefore by Theorem 2.1, Lemma 3.5, and (1.19),

$$\begin{aligned}N\cdot\mathbb{E}\left[\left|\langle f, \nu_t^N\rangle - \langle f, \mu_t\rangle\right|^2\right] \;&=\; \operatorname{Var}_{\mu_t}(f) + \int_0^{(t-\delta)^+} \mathbb{E}\left[V_{s,t}^N(f)\right]\, ds \;+\; \int_{(t-\delta)^+}^t \mathbb{E}\left[V_{s,t}^N(f)\right]\, ds\\ \le\;& \operatorname{Var}_{\mu_t}(f) + \int_0^t V_{s,t}(f)\, ds\\ +\;& \left[\bar{C}_t(p,q,\delta)\big(19(\bar{\varepsilon}_t^{N,p})^{1/2} + 11\bar{\varepsilon}_t^{N,p}\big) + 30\,e^{2/35}\int_{(t-\delta)^+}^t \|H_s\|_{L^q(\mu_s)}\, ds\right]\cdot\|f\|^2_{L^p(\mu_t)}.\end{aligned}$$

Observing that $\|H_s\|_{L^q(\mu_s)} \le \operatorname{osc}(H_s)$ and that $30\cdot e^{2/35}/35 < 1$, we obtain (1.20).

Furthermore, by maximizing (1.20) over all $f : S \to \mathbb{R}$ with $\|f\|_{L^p(\mu_t)} \le 1$ and over t, we get

$$N\bar{\varepsilon}_t^{N,p} \;\le\; 2 + v_t^p + \bar{C}_t(p,q,\delta) \cdot \Big(19(\bar{\varepsilon}_t^{N,p})^{1/2} + 11\bar{\varepsilon}_t^{N,p}\Big)$$

for all $t \in [0, t_0]$. Hence, recalling that $N > 11\,\bar{C}_t(p,q,\delta)$ by assumption,

$$N\bar{\varepsilon}_t^{N,p} \;\le\; \alpha + \beta(N\bar{\varepsilon}_t^{N,p})^{1/2},$$

with

$$\begin{aligned}\alpha &= (2 + v_t^p) \cdot N/(N - 11 \cdot \bar{C}_t(p,q,\delta)),\\ \beta &= 19 \cdot \bar{C}_t(p,q,\delta) \cdot N^{1/2}/(N - 11 \cdot \bar{C}_t(p,q,\delta)).\end{aligned}$$

Therefore

$$(N\bar{\varepsilon}_t^{N,p})^{1/2} \le \frac{\beta}{2} + \big((\beta/2)^2 + \alpha\big)^{1/2} \le \alpha^{1/2} + \beta. \tag{3.5}$$

Since $\bar{C}_t(p,q,\delta) \le N/80$, we have $N - 11 \cdot \bar{C}_t(p,q,\delta) \ge \frac{69}{80}N$ and

$$\frac{N}{N - 11 \cdot \bar{C}_t(p,q,\delta)} \;=\; 1 + \frac{11 \cdot \bar{C}_t(p,q,\delta)}{N - 11 \cdot \bar{C}_t(p,q,\delta)} \;\le\; 1 + \frac{11 \cdot 80}{69}\bar{C}_t(p,q,\delta)\,N^{-1}.$$

Therefore, by (3.5),

$$(N\bar{\varepsilon}_t^{N,p})^{1/2} \;\le\; (2 + v_t^p)^{1/2} \cdot \left(1 + \frac{11 \cdot 80}{2 \cdot 69}\,\bar{C}_t(p,q,\delta)\,N^{-1}\right) \;+\; \frac{80 \cdot 19}{69}\,\bar{C}_t(p,q,\delta)\,N^{-1/2},$$

which implies (1.21). □

4. Proofs of Theorems 1.4 and 1.7

Proof of Theorem 1.4. By the estimates in [5] we have, for $0 \le s \le t \le t_0$,

$$\|q_{s,t}f\|_{L^p(\mu_s)} \le 2^{1/4}\|f\|_{L^p(\mu_s)}$$

for all $f : S \to \mathbb{R}$, provided

$$\lambda_s \;\ge\; \frac{p}{4}\,A_s \;+\; \frac{p(p+3)}{4}\,t_0\,B_s \qquad \text{for all } s \in [0, t_0]. \tag{4.1}$$

Hence, under this condition, we get $C_{s,t}(p) \le 2^{1/4}$. Moreover, by [5],

$$\|q_{t-\delta,t}f\|_{L^q(\mu_{t-\delta})} \;\le\; \exp\big(\int_{t-\delta}^{t} \max H_r^-\,dr\big) \cdot \|f\|_{L^p(\mu_t)}$$

for all $f : S \to \mathbb{R}$ and $0 \le \delta \le t \le t_0$, provided

$$\lambda_s \;\ge\; \frac{\gamma_s}{4\delta}\log\frac{q-1}{p-1} \qquad \text{for all } s \in [0, t_0]. \tag{4.2}$$

Since $\delta = \big(35\sup_{s\in[0,t]} \operatorname{osc} H_s\big)^{-1}$, we obtain that, for $s \le t - \delta$,

$$\|q_{s,t}f\|_{L^p(\mu_s)} \;=\; \|q_{s,t-\delta}q_{t-\delta,t}f\|_{L^p(\mu_s)} \;\le\; 2^{1/4}e^{1/35}\|f\|_{L^q(\mu_t)},$$

if both (4.1) and (4.2) hold. Hence

$$C_{s,t}(p,q) \;\le\; 2^{1/4}e^{1/35}$$

provided (4.1) holds and

$$\lambda_s \;\ge\; \frac{\gamma_s}{4\delta}\log\max\big(\frac{2r-1}{p-1}, \frac{2p-2}{p-2}\big) \qquad \text{for all } s \in [0, t_0]. \tag{4.3}$$

Since $2 < \tilde{p} < p$ and $\tilde{p}^{-1} = q^{-1} + (p/2)^{-1}$, we obtain similarly that $C_{s,t}(\tilde{p}, q) \le 2^{1/4}e^{1/35}$ provided (4.1) holds and

$$\lambda_s \;\ge\; \frac{\gamma_s}{4\delta}\log\max\big(\frac{p-1}{\tilde{p}-1}, \frac{2\tilde{p}-2}{\tilde{p}-2}\big) \qquad \text{for all } s \in [0, t_0]. \tag{4.4}$$

The assertion now follows from Theorem 1.3 and the estimates (1.22) and (1.23). □

Proof of Lemma 1.1 and Corollary 1.6. For a function $f : S \to \mathbb{R}$ and $t \geq 0$ let $f_t := f - \langle f, \mu_t \rangle$. Then

$$\begin{aligned} \langle f_t, \eta_t^N \rangle &= \langle f, \eta_t^N \rangle - \langle f, \mu_t \rangle, \qquad \text{and} \\ \mathbb{E}\left[\langle f_t, \eta_t^N \rangle^2\right] &\leq 2 \cdot \mathbb{E}\left[\left(\langle f_t, \eta_t^N \rangle - \langle f_t, \nu_t^N \rangle\right)^2\right] + 2 \cdot \mathbb{E}\left[\langle f_t, \nu_t^N \rangle^2\right] \\ &\leq 2 \cdot \|f_t\|_{\sup}^2 \cdot \mathbb{E}\left[\left(\langle 1, \nu_t^N \rangle - 1\right)^2\right] + 2 \cdot \mathbb{E}\left[\langle f_t, \nu_t^N \rangle^2\right] . \end{aligned}$$

Applying this bound, we obtain an improved L^1 estimate:

$$\begin{aligned} \mathbb{E}\left[|\langle f_t, \eta_t^N \rangle|\right] &\leq \mathbb{E}\left[\langle f_t, \eta_t^N \rangle^2\right]^{1/2} \mathbb{E}\left[\left(\langle 1, \nu_t^N \rangle - 1\right)^2\right]^{1/2} + \mathbb{E}\left[\langle f_t, \nu_t^N \rangle^2\right]^{1/2} \\ &\leq \mathbb{E}\left[\langle f_t, \nu_t^N \rangle^2\right]^{1/2} + \sqrt{2}\, \|f_t\|_{\sup} \mathbb{E}\left[\left(\langle 1, \nu_t^N \rangle - 1\right)^2\right] + \sqrt{2}\, \mathbb{E}\left[\langle f_t, \nu_t^N \rangle^2\right]^{1/2} \mathbb{E}\left[\left(\langle 1, \nu_t^N \rangle - 1\right)^2\right]^{1/2} . \end{aligned}$$

This proves Lemma 1.1. The assertion of Corollary 1.6 now follows from the lemma and Theorem 1.4. □

Proof of Theorem 1.7. Fix $i \in I$ and define

$$h_t(i) := \langle H_t, \mu_t^i \rangle = \int_{S_i} H_t \, d\mu_t / \mu_t(S_i).$$

Note that

$$h_t(i) = -\frac{d}{dt} \log \mu_t(S_i).$$

Since (1.2) and (1.3) hold, $H_t^i = H_t - h_t(i)$ is the negative logarithmic time derivative of μ_t^i. If we define $q_{s,t}^i f$ for functions $f : S_i \to \mathbb{R}$ in the same way as $q_{s,t} f$ with H_t replaced by H_t^i, then

$$q_{s,t} f(x) = \exp\big(-\int_s^t h_r(i)\, dr\big)\, q_{s,t}^i f(x) = \frac{\mu_t(S_i)}{\mu_s(S_i)}\, q_{s,t}^i f(x).$$

In particular, for $p \in [1, \infty]$, we have

$$\|q_{s,t} f\|_{\widetilde{L^p(\mu_s)}} = \max_{i \in I} \|q_{s,t} f\|_{L^p(\mu_s^i)} \leq \max_{i \in I} \frac{\mu_t(S_i)}{\mu_s(S_i)} \|q_{s,t}^i f\|_{L^p(\mu_s^i)}. \tag{4.5}$$

Assuming Poincaré and log Sobolev inequalities with respect to the measures μ_t^i and the functions H_t^i, we obtain the same type of L^p-L^q bounds for the operators $q_{s,t}^i$ as we did for the operators $q_{s,t}$ in the proof of Theorem 1.4. Because of (4.5) the assertion then follows similarly as above. □

References

1. P. Del Moral, A. Doucet, and A. Jasra, *Sequential Monte Carlo samplers*, J. R. Statist. Soc. B **68** (2006), no. 3, 411–436. MR MR1819122 (2002k:60013)
2. P. Del Moral and A. Guionnet, *On the stability of interacting processes with applications to filtering and genetic algorithms*, Ann. Inst. H. Poincaré Probab. Statist. **37** (2001), no. 2, 155–194. MR MR1819122 (2002k:60013)
3. P. Del Moral and L. Miclo, *Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering*, Séminaire de Probabilités, XXXIV, Lecture Notes in Math., vol. 1729, Springer, Berlin, 2000, pp. 1–145. MR MR1768060 (2001g:60091)
4. A. Eberle and C. Marinelli, *Convergence of sequential Markov chain Monte Carlo methods: I. Nonlinear flow of probability measures*, 2006, arXiv:math/0612074v1 [math.PR].
5. ———, *L^p estimates for Feynman-Kac propagators*, (2008), Preprint.
6. C. J. Geyer, *Markov Chain Monte Carlo maximum likelihood*, Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface (1991), 156–163.

7. Ĭ. Ī. Gīhman and A. V. Skorohod, *The theory of stochastic processes. II*, Springer-Verlag, New York, 1975.
8. A. Gulisashvili and J. A. van Casteren, *Non-autonomous Kato classes and Feynman-Kac propagators*, World Scientific Publishing Co., 2006. MR MR2253111 (2008b:60161)
9. K. Hukushima and K. Nemoto, *Exchange Monte Carlo method and application to spin glass simulations*, Journal of the Physical Society of Japan **65** (1996), no. 6, 1604–1608.
10. C. Kipnis and C. Landim, *Scaling limits of interacting particle systems*, Springer-Verlag, Berlin, 1999. MR MR1707314 (2000i:60001)
11. S. Kou, Q. Zhou, and W. Wong, *Equi-energy sampler with applications in statistical inference and statistical mechanics*, Annals of Statistics **34** (2006), no. 4, 1581–1619.
12. N. Madras and Z. Zheng, *On the swapping algorithm*, Random Structures Algorithms **22** (2003), no. 1, 66–97. MR MR1943860 (2004c:82117)
13. M. Rousset, *On the control of an interacting particle approximation of Schrödinger ground states*, SIAM Journal on Mathematical Analysis **38** (2006), no. 3, 824–844.
14. L. Saloff-Coste, *Lectures on finite Markov chains*, Lectures on probability theory and statistics (Saint-Flour, 1996), Lecture Notes in Math., vol. 1665, Springer, Berlin, 1997, pp. 301–413. MR MR1490046 (99b:60119)

Institut für Angewandte Mathematik, Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany
E-mail address: eberle@uni-bonn.de
URL: http://wiener.iam.uni-bonn.de/∼eberle

Institut für Angewandte Mathematik, Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany
E-mail address: marinelli@wiener.iam.uni-bonn.de
URL: http://www.uni-bonn.de/∼cm788

Bestellungen nimmt entgegen:

Institut für Angewandte Mathematik
der Universität Bonn
Sonderforschungsbereich 611
Wegelerstr. 6
D - 53115 Bonn

Telefon: 0228/73 4882
Telefax: 0228/73 7864
E-mail: astrid.link@iam.uni-bonn.de http://www.sfb611.iam.uni-bonn.de/

Verzeichnis der erschienenen Preprints ab No. 385

385. Albeverio, Sergio; Hryniv, Rostyslav; Mykytyuk, Yaroslav: Factorisation of Non-Negative Fredholm Operators and Inverse Spectral Problems for Bessel Operators

386. Otto, Felix: Optimal Bounds on the Kuramoto-Sivashinsky Equation

387. Gottschalk, Hanno; Thaler, Horst: A Comment on the Infra-Red Problem in the AdS/CFT Correspondence

388. Dickopf, Thomas; Krause, Rolf: Efficient Simulation of Multi-Body Contact Problems on Complex Geometries: A Flexible Decomposition Approach Using Constrained Minimization

389. Groß, Christian; Krause, Rolf: On the Convergence of Recursive Trust-Region Methods for Multiscale Non-Linear Optimization and Applications to Non-Linear Mechanics

390. Nemitz, Oliver; Nielsen, Michael Bang; Rumpf, Martin; Whitaker, Ross: Finite Element Methods on Very Large, Dynamic Tubular Grid Encoded Implicit Surfaces

391. Barbu, Viorel; Marinelli, Carlo: Strong Solutions for Stochastic Porous Media Equations with Jumps

392. Albeverio, Sergio; Cebulla, Christof: Synchronizability of Stochastic Network Ensembles in a Model of Interacting Dynamical Units; erscheint in: Physica A: Statistical Mechanics and its Applications

393. Albeverio, Sergio; Cebulla, Christof: Synchronizability of a Stochastic Version of FitzHugh-Nagumo Type Neural Oscillator Networks

394. Bacher, Kathrin: On Borell-Brascamp-Lieb Inequalities on Metric Measure Spaces

395. Croce, Roberto; Griebel, Michael; Schweitzer, Marc Alexander: Numerical Simulation of Droplet-Deformation by a Level Set Approach with Surface Tension

396. Schweizer, Nikolaus: Spectral Gaps of Simulated Tempering for Bimodal Distributions

397. Koivisto, Juha; Rost, Martin; Alt, Wolfgang: A Phase Field Model for Phospholipid Surfactant

398. Reiter, Philipp; Felix, Dieter; von der Mosel, Heiko; Alt, Wolfgang: Energetics and Dynamics of Global Integrals Modeling Interaction between Stiff Filaments

399. Bogachev, Leonid; Daletskii, Alexei: Poisson Cluster Measures: Quasi-Invariance, Integration by Parts and Equilibrium Stochastic Dynamics

400. Lapid, Erez; Müller, Werner: Spectral Asymptotics for Arithmetic Quotients of SL(n,R)/SO(n)

401. Finis, Tobias; Lapid, Erez; Müller, Werner: On The Spectral Side of Arthur's Trace Formula II

402. Griebel, Michael; Knapek, Stephan: Optimized General Sparse Grid Approximation Spaces for Operator Equations

403. Griebel, Michael; Jager, Lukas: The BGY3dM Model for the Approximation of Solvent Densities

404. Griebel, Michael; Klitz, Margit: Homogenization and Numerical Simulation of Flow in Geometries with Textile Microstructures

405. Dickopf, Thomas; Krause, Rolf: Weak Information Transfer between Non-Matching Warped Interfaces

406. Engel, Martin; Griebel, Michael: A Multigrid Method for Constrained Optimal Control Problems

407. Nepomnyaschikh, Sergey V.; Scherer, Karl: A Domain Decomposition Method for the Helmholtz-Problem

408. Harbrecht, Helmut: A Finite Element Method for Elliptic Problems with Stochastic Input Data

409. Groß, Christian; Krause, Rolf: A Recursive Trust–Region Method for Non-convex Constrained Minimization

410. Eberle, Andreas; Marinelli, Carlo: Lp Estimates for Feynman-Kac Propagators

411. Eberle, Andreas; Marinelli, Carlo: Quantitative Approximations of Evolving Probability Measures and Sequential Markov Chain Monte Carlo Methods