



# On the Particle Approximation of Lagged Feynman-Kac Formulae

Elsiddig Awadelkarim, Michel Caffarel, Pierre Del Moral, Ajay Jasra

## ► To cite this version:

Elsiddig Awadelkarim, Michel Caffarel, Pierre Del Moral, Ajay Jasra. On the Particle Approximation of Lagged Feynman-Kac Formulae. 2024. hal-04705748

**HAL Id: hal-04705748**

**<https://hal.science/hal-04705748v1>**

Preprint submitted on 23 Sep 2024

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

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

# On the Particle Approximation of Lagged Feynman-Kac Formulae

BY ELSIDDIG AWADELKARIM<sup>1</sup>, MICHEL CAFFAREL<sup>2</sup>, PIERRE DEL MORAL<sup>3</sup> & AJAY JASRA<sup>4</sup>

<sup>1</sup>Applied Mathematics and Computational Science Program, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, 23955-6900, KSA.

<sup>2</sup>University of Toulouse, CNRS-Lab. de Chimie et Physique Quantiques, Toulouse, 31062, FR.

<sup>3</sup>Centre de Recherche Inria Bordeaux Sud-Ouest, Talence, 33405, FR.

<sup>4</sup>School of Data Science, The Chinese University Hong Kong, Shenzhen, CN.

E-Mail: [elsiddigawadelkarim.elsiddig@kaust.edu.sa](mailto:elsiddigawadelkarim.elsiddig@kaust.edu.sa); [caffarel@irsamc.ups-tlse.fr](mailto:caffarel@irsamc.ups-tlse.fr);

[pierre.del-moral@inria.fr](mailto:pierre.del-moral@inria.fr); [ajayjasra@cuhk.edu.cn](mailto:ajayjasra@cuhk.edu.cn)

## Abstract

In this paper we examine the numerical approximation of the limiting invariant measure associated with Feynman-Kac formulae. These are expressed in a discrete time formulation and are associated with a Markov chain and a potential function. The typical application considered here is the computation of eigenvalues associated with non-negative operators as found, for example, in physics or particle simulation of rare-events. We focus on a novel *lagged* approximation of this invariant measure, based upon the introduction of a ratio of time-averaged Feynman-Kac marginals associated with a positive operator iterated  $l \in \mathbb{N}$  times; a lagged Feynman-Kac formula. This estimator and its approximation using Diffusion Monte Carlo (DMC) have been extensively employed in the physics literature. In short, DMC is an iterative algorithm involving  $N \in \mathbb{N}$  particles or walkers simulated in parallel, that undergo sampling and resampling operations. In this work, it is shown that for the DMC approximation of the lagged Feynman-Kac formula, one has an almost sure characterization of the  $\mathbb{L}_1$ -error as the time parameter (iteration) goes to infinity and this is at most of  $\mathcal{O}(\exp\{-\kappa l\}/N)$ , for  $\kappa > 0$ . In addition a non-asymptotic in time, and time uniform  $\mathbb{L}_1$ -bound is proved which is  $\mathcal{O}(l/\sqrt{N})$ . We also prove a novel central limit theorem to give a characterization of the exact asymptotic in time variance. This analysis demonstrates that the strategy used in physics, namely, to run DMC with  $N$  and  $l$  small and, for long time enough, is mathematically justified. Our results also suggest how one should choose  $N$  and  $l$  in practice. We emphasize that these results are not restricted to physical applications; they have broad relevance to the general problem of particle simulation of the Feynman-Kac formula, which is utilized in a great variety of scientific and engineering fields.

**Keywords:** Feynman-Kac Formula, Diffusion Monte Carlo, Eigenvalue approximation.

## 1 Introduction

The Feynman-Kac (FK) formula plays a central role in many scientific fields including mathematics [5], statistics [17], physics [7, 29], control theory [15], quantitative finance [24], to name some of the main ones, see for instance [11, 12] for book-length introductions to the subject. In the context of this work, the Feynman-Kac formula can be formulated in the following way. Consider a time-homogeneous Markov chain with initial distribution  $\eta_0$  and transition kernel  $M$  on a measurable state-space  $\mathbf{E}$  and let  $G$  be a bounded, measurable and strictly positive real-valued function on  $\mathbf{E}$ . The Feynman-Kac expresses the evolution of the initial distribution as follows. Let  $n \in \mathbb{N}$  and any  $\varphi : \mathbf{E} \rightarrow \mathbb{R}$  bounded and measurable, the Feynman-Kac measure  $\eta_n(\varphi)$  is given by

$$\eta_n(\varphi) := \frac{\mathbb{E} \left[ \varphi(X_n) \prod_{p=0}^{n-1} G(X_p) \right]}{\mathbb{E} \left[ \prod_{p=0}^{n-1} G(X_p) \right]} \quad (1.1)$$

where  $\mathbb{E}[\cdot]$  is the expectation operator w.r.t. the probability law of the afore-mentioned Markov chain. In this work, we will be in particular interested in evaluating  $\eta_\infty(\varphi) := \lim_{n \rightarrow +\infty} \eta_n(\varphi)$  which exists under certain mathematical assumptions. However, except for trivial cases, the quantity  $\eta_n(\varphi)$  cannot be evaluated analytically and one must rely upon numerical methods. These approaches are referred to as sequential Monte Carlo or particle filters in mathematics/statistics and under the general denomination of

Diffusion Monte Carlo (DMC) in physics. Several variants of DMC have been introduced such as Green's function Monte Carlo [6, 25], Fixed-Node Diffusion Monte Carlo [27], Pure Diffusion Monte Carlo [7, 8], Stochastic Reconfiguration Monte Carlo [2, 22, 30, 31] and Reptation Monte Carlo [3], to cite the main ones.

In practice, calculations are done by introducing a McKean interacting particle system involving a number of particles to which stochastic rules built from  $M$  (sampling using the Markov kernel) and  $G$  (resampling using the weight  $G$ ) are applied iteratively. In the mathematical literature, these methods are by now rather well understood, with a plethora of results; see for instance [4, 11, 12, 18, 33]. In the following where the estimator discussed is borrowed to physics, we will adopt the name DMC for the preceding numerical algorithm. Note that in physics, the  $N$  samples/particles of the evolving population are called walkers, so the two terms "particles" and "walkers" will be used interchangeably in this article. Now, a key point is that a great freedom exists in choosing either the stochastic rules and/or the estimator employed for  $\eta_n(\varphi)$ . The problem of making some optimal choice leading to smaller variances and reduced finite-population biases is thus a central issue in numerical implementations.

In this work we analyze an original estimator for the Feynman-Kac measure introduced in physics [7, 8] and which appears, to the best of our knowledge, not to have been considered in the mathematical literature. In physics, the problem considered consists in solving the imaginary-time dependent Schrödinger equation using the Feynman-Kac formula.

In its simplest formulation, the Feynman-Kac formula is a particular case of Eq.(1.1) with the following setting: i)  $Q(x, y) = G(x)M(x, dy)$  is the kernel of the operator  $e^{-\tau H}$  where  $H$  is the Hamiltonian operator describing the system and  $\tau$  a small positive quantity playing the role of a time step, ii)  $M(x, dy)$  the Markov kernel corresponding to the brownian process, and iii)  $G(x) = e^{-\tau V(\mathbf{x})}$  where  $V(\mathbf{x})$  is the potential function. In practice, for non-trivial quantum systems, calculations are possible (small enough statistical errors) only if importance sampling is introduced. We are then led to a so-called importance sampled Feynman-Kac formula, which is still a particular case of Eq.(1.1), but with some more general operator  $Q$ . Details are given in Appendix A. A general presentation of this setting and the way it is derived in physics can be found for example in [7, 29]. As known, the quantity  $\eta_\infty(G)$  obtained from Eq.(1.1) turns out to be the eigenvalue of the operator  $Q$ , that is the  $\lambda \in \mathbb{R}^+$  such that for some  $h : \mathbb{E} \rightarrow \mathbb{R}^+$ ,  $\int_{\mathbb{E}} h(y)Q(x, dy) = \lambda h(x)$ ; the computation of which has found several applications in mathematics, for example, in rare-events estimation [13, 34]. In physics, getting  $\lambda$  gives the ground-state energy,  $E_0$ , via the relation  $\lambda = e^{-\tau E_0}$  and  $h$  is the ground-state eigenstate from which physical properties can be computed [28].

Based upon the Feynman-Kac formula representation, different FK-like formulae can be proposed to evaluate  $\eta_\infty(\varphi)$ . For instance, one can consider the following quantity

$$\bar{\eta}_n(\varphi) := \frac{1}{n} \sum_{k=0}^{n-1} \eta_k(\varphi), \quad (1.2)$$

which, as  $n$  grows large, will converge, under assumptions, to  $\eta_\infty(\varphi)$ . In this work, we introduce the alternative quantity built from  $\eta_k$  and defined as follows. Let  $l \in \mathbb{N}$  be a lag, which can be chosen to be any fixed value and define the function for  $x_0 \in \mathbb{E}$ :

$$Q^l(\varphi)(x_0) := \int_{\mathbb{E}^l} \varphi(x_l) \prod_{k=1}^l Q(x_{k-1}, dx_k).$$

We define the quantity

$$\frac{\bar{\eta}_n(Q^l(\varphi))}{\bar{\eta}_n(Q^l(1))} \quad (1.3)$$

which leads to the following FK-like formula for the limiting invariant measure

$$\eta_\infty(\varphi) = \lim_{n \rightarrow \infty} \frac{\bar{\eta}_n(Q^l(\varphi))}{\bar{\eta}_n(Q^l(1))}. \quad (1.4)$$

In the following, we will refer to (1.3) as the *lagged* Feynman-Kac formula.

The general idea of why one may prefer to apply the numerical DMC algorithm to (1.3) instead of (1.2) is the flexibility resulting from the introduction of the lag parameter  $l$ . For example, and without entering into details at this stage, using DMC for approximating (1.2) necessitates in general a large number  $N$  of walkers to extrapolate to zero the population bias of order  $\frac{1}{N}$  as  $N$  goes to infinity. In sharp contrast, this problem becomes much less severe using (1.3) since, as we shall show, the  $N$ -dependence of the results decreases rapidly as the lag  $l$  increases. From that, one might obtain a substantial computational saving in the DMC simulation. Finally, as already mentioned and to the best of our knowledge, a mathematical study of the DMC approximation of (1.3) has not been undertaken; this is the topic of the present study.

## 1.1 Summary of Main Results and Article Structure

The analysis of the DMC approximation of (1.3) is rather challenging. As we will show later on, this consists of a normalized estimator of walkers over the lag  $l$ , time-averaged in as in (1.3). The standard theory associated to DMC (e.g. [11]) is not trivially applicable to the estimator that is used in practice, hence requiring an original approach. In this article, under assumptions, we prove the following interesting results:

- For a fixed number of particles  $N$  we prove an almost sure (as  $n$  grows) limit for the  $\mathbb{L}_1$ -error of the DMC approximation of (1.3) and show that this is at most  $\mathcal{O}(\exp\{-\kappa l\}/N)$ ,  $\kappa > 0$ . This is Theorem 4.1 and Corollary 4.1.
- For a fixed number of particles  $N$  and time horizon  $n$  we prove that the appropriately centered  $\mathbb{L}_1$ -error DMC approximation of (1.3) is at most  $\mathcal{O}(l/\sqrt{N})$ . The bound is uniform in time ( $n$ ). This is Theorem 4.2.
- For a fixed number of particles  $N$  we prove an asymptotic in  $n$  central limit theorem (CLT) for the estimator. This is in Theorem 4.3.

The implication of these results are rather important. Theorem 4.1 and Corollary 4.1 establish that, as is done in practice, DMC can be run with  $N$  the number of walkers small and  $l$  small, for long-times and the error of the approximation can be vanishingly (exponentially) small. Theorem 4.2 also shows that one should not necessarily grow  $l$  arbitrarily large as there is a price to pay. We also use our theorems to compare with the DMC approximation of (1.2). Theorem 4.3 provides an exact characterization of the asymptotic in time variance. Our results are also confirmed in numerical simulations. The proofs of the afore mentioned results require some non-standard arguments based upon novel path-wise solutions of the Poisson equation for Markov chains. Our proofs are not confined to problems in physics and indeed, apply to any application where the particle approximation of (1.3) is of practical interest. This constitutes many applications in rare-events estimation and stochastic control problems; see [13, 34] for instance.

This paper is structured as follows. In Section 2 we detail the formulae of interest and the Diffusion Monte Carlo algorithm. In Section 3 we give the numerical estimator under study. Section 4 presents our mathematical results with discussion. The appendix contains the technical details associated to the proofs of our main results as well as some details of the application in physics.

## 2 Feynman-Kac Formulae and Diffusion Monte Carlo

### 2.1 Notation

Let  $(E, \mathcal{E})$  be a measurable space. Denote by  $\mathcal{P}(E)$  and  $\mathcal{M}(E)$  the collection of probability measures and non-negative measures on  $(E, \mathcal{E})$ . For  $\varphi : E \rightarrow \mathbb{R}$  we write  $\mathcal{B}_b(E)$  as the collection of bounded and measurable functions on  $(E, \mathcal{E})$ . For  $\varphi \in \mathcal{B}_b(E)$ , the sup-norm is written  $\|\varphi\| \sup_{x \in E} |\varphi(x)|$ . For  $(\mu, \varphi) \in \mathcal{M}(E) \times \mathcal{B}_b(E)$ ,  $\mu(\varphi) = \int_E \varphi(x) \mu(dx)$ . Let  $P : E \rightarrow \mathcal{M}(E)$ ,  $(\mu, \varphi) \in \mathcal{M}(E) \times \mathcal{B}_b(E)$  we use the

short-hand notation for  $x \in \mathbf{E}$ :

$$\mu P(\varphi) = \int_{\mathbf{E}} \mu(dx) P(\varphi)(x) \quad \text{and} \quad P(\varphi)(x) = \int_{\mathbf{E}} P(x, dy) \varphi(y).$$

Note that one can use the notation  $\mu P(\varphi)$  and  $\mu(P(\varphi))$  exchangeably, but we generally use the former.  $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ . For  $P : \mathbf{E} \rightarrow \mathcal{M}(\mathbf{E})$ ,  $\varphi \in \mathcal{B}_b(\mathbf{E})$  and  $n \in \mathbb{N}$ ,  $P^n(\varphi)(x_0) = \int_{\mathbf{E}^n} \varphi(x_n) \prod_{k=1}^n P(x_{k-1}, dx_k)$  with the convention that  $P^0(\varphi)(x_0) = \varphi(x_0)$ . For  $A \in \mathcal{E}$  the Dirac measure is written  $\delta_A(dx)$ , with the convention that if  $A = \{x\}$ ,  $x \in \mathbf{E}$ , we write  $\delta_x(dy)$ . For  $(\mu, \nu) \in \mathcal{P}(\mathbf{E})^2$ , we denote by  $\|\mu - \nu\|_{\text{tv}} = \sup_{A \in \mathcal{E}} |\mu(A) - \nu(A)|$  as the total variation distance.  $\mathcal{N}(0, \sigma^2)$  denotes the one-dimensional normal distribution of mean 0 and variance  $\sigma^2 > 0$ .

## 2.2 Feynman-Kac Semigroups

Consider the Feynman-Kac measures on a measurable space  $(\mathbf{E}, \mathcal{E})$ , for  $(n, \varphi) \in \mathbb{N}_0 \times \mathcal{B}_b(\mathbf{E})$ :

$$\eta_n(\varphi) = \gamma_n(\varphi) / \gamma_n(1) \quad \text{with} \quad \gamma_n(\varphi) := \mathbb{E} \left( \varphi(X_n) \prod_{p=0}^{n-1} G(X_p) \right)$$

where  $G : \mathbf{E} \rightarrow \mathbb{R}^+$  is a strictly positive and bounded potential and the expectation  $\mathbb{E}(\cdot)$  is w.r.t. the law of a time-homogenous Markov chain with initial distribution  $\eta_0$  and transition kernel  $M : \mathbf{E} \rightarrow \mathcal{P}(\mathbf{E})$ . We use the convention  $\gamma_0 = \eta_0 = \text{Law}(X_0)$ , for  $n = 0$ . The unnormalized measures  $\gamma_n$  have a linear evolution given for any  $n \geq 1$  by the recursion

$$\gamma_n(\varphi) = \gamma_{n-1} Q(\varphi) \quad \text{and we have} \quad \gamma_n(\varphi) = \eta_n(\varphi) \prod_{0 \leq p < n} \eta_p(G) \quad (2.1)$$

where  $Q(x, dy) = G(x)M(x, dy)$ . The normalized measures have a nonlinear evolution

$$\eta_{n+1}(\varphi) = \Phi(\eta_n)(\varphi) = \Psi_G(\eta_n)M(\varphi) \quad \text{with} \quad \Psi_G(\eta_n)(\varphi) := \frac{\eta_n(G\varphi)}{\eta_n(G)}. \quad (2.2)$$

The evolution semigroup associated with the flow  $\eta_n$  is given for any  $p \leq n$  by

$$\eta_n(\varphi) = \Phi^{n-p}(\eta_p)(\varphi) \quad \text{with} \quad \Phi^{n-p}(\eta_p)(\varphi) = \Phi \circ \Phi \circ \dots \circ \Phi(\eta_p)$$

where the composition is  $(n-p)$ -times. Note that

$$\Phi^{n-p}(\eta_p)(\varphi) = \frac{\eta_p Q^{n-p}(\varphi)}{\eta_p Q^{n-p}(1)} \quad \text{and} \quad \Phi^l(\bar{\eta}_n)(\varphi) = \frac{\bar{\eta}_n Q^l(\varphi)}{\bar{\eta}_n Q^l(1)}.$$

For any given  $l \geq 1$  we set

$$\mu_n^l(\varphi) := \frac{1}{n} \sum_{0 \leq k < n} \gamma_{l+k}(\varphi) / \gamma_k(1) = \frac{1}{n} \sum_{0 \leq k < n} \eta_k(Q^l(\varphi)) = \bar{\eta}_n Q^l(\varphi)$$

and their normalized versions

$$\bar{\mu}_n^l(\varphi) := \mu_n^l(\varphi) / \mu_n^l(1) \implies \bar{\mu}_n^l = \Phi^l(\bar{\eta}_n).$$

In our context, under some stability conditions (see e.g. [11, 34] for instance), there exists an invariant measure  $\eta_\infty$ , a parameter  $\lambda > 0$  and a function  $h \geq 0$  such that for any  $(x, \varphi) \in \mathbf{E} \times \mathcal{B}_b(\mathbf{E})$

$$Q(h)(x) = \lambda h(x) \quad \text{and} \quad \eta_\infty(\varphi) = \Phi(\eta_\infty)(\varphi) \quad (\implies \quad Q^n(h) = \lambda^n h).$$

Applying the above fixed point equation to  $\varphi = h$  we check that

$$\Phi^n(\eta_\infty)(h) = \frac{\eta_\infty Q^n(h)}{\eta_\infty Q^n(1)} = \eta_\infty(h) = \lambda^n \frac{\eta_\infty(h)}{\eta_\infty Q^n(1)} \iff \eta_\infty Q^n(1) = \lambda^n.$$

Also note that

$$G(x) = Q(1)(x) \implies \eta_\infty(G) = \lambda.$$

More generally, for any  $\varphi \in \mathcal{B}_b(\mathbf{E})$  we have

$$\Phi^n(\eta_\infty)(\varphi) = \eta_\infty(\varphi) \iff \eta_\infty Q^n(\varphi) = \lambda^n \eta_\infty(\varphi).$$

### 2.3 Diffusion Monte Carlo

DMC is a discrete-time system of  $N$  walkers  $\xi_n = (\xi_n^i)_{1 \leq i \leq N}$ . The system starts with  $N$  independent copies of a random variable with distribution  $\eta_0$ . Given the system  $\xi_n$  at some time  $n \geq 0$ , we sample  $N$  conditionally independent walkers  $\xi_{n+1}^i$  with their respective distribution  $\Phi(\eta_n^N)$ , with

$$\eta_n^N := m(\xi_n) := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}.$$

In other words, the DMC method consists of approximating the measure  $\eta_n$  by using the occupation measure  $\eta_n^N$  associated with a system of  $N$  walkers. The initial positions of the walkers are randomly chosen from the distribution  $\eta_0$ . The evolution of each walker follows then the following selection/mutation steps:

- Selection: We evaluate the current position  $\xi_n^i$  of all of the walkers at its potential value  $G(\xi_n^i)$ . For each walker, we select a walker amongst the current collection with probability  $G(\xi_n^i)/\eta_n^N(G)$ , giving the new position  $\widehat{\xi}_n^i$ .
- Mutation: We move the selected walker  $\widehat{\xi}_n^i = x$  to a new location  $\xi_{n+1}^i = y$  using the transition kernel  $M(x, dy)$ .

## 3 Numerical Estimator

### 3.1 Unnormalized Particle Measures

We consider some well-known facts about particle approximation of unnormalized measures; this discussion will prove useful as we give lagged estimators that are used in the physics literature. Mimicking the r.h.s. formula in (2.1), the unbiased estimate of  $\gamma_n(\varphi)$  (see [11, Chapter 7]) is defined by

$$\gamma_n^N(\varphi) := \eta_n^N(\varphi) \prod_{0 \leq p < n} \eta_p^N(G).$$

The unbiasedness property ensures that

$$\mathbb{E}(\gamma_n^N(\varphi)) = \gamma_n(\varphi) = \eta_0 Q^n(\varphi).$$

To check this claim, note that

$$\mathbb{E}(\eta_n^N(\varphi) \mid \xi_{n-1}) = \Phi(\eta_{n-1}^N)(\varphi) = \frac{\eta_{n-1}^N Q(\varphi)}{\eta_{n-1}^N Q(1)} = \frac{\eta_{n-1}^N Q(\varphi)}{\eta_{n-1}^N(G)}.$$

This implies that

$$\mathbb{E}(\gamma_n^N(\varphi) \mid \xi_{n-1}) = \eta_{n-1}^N Q(\varphi) \prod_{0 \leq p < (n-1)} \eta_p^N(G) = \gamma_{n-1}^N Q(\varphi).$$

Iterating the argument, for any  $k < n$  we check that

$$\mathbb{E}(\gamma_n^N(\varphi) \mid \xi_k) = \gamma_k^N Q^{n-k}(\varphi). \quad (3.1)$$

Applying the above to  $k = 0$  we recover the unbiasedness property

$$\mathbb{E}(\gamma_n^N(\varphi) \mid \xi_0) = \eta_0^N Q^n(\varphi) \quad \text{and} \quad \mathbb{E}(\gamma_n^N(\varphi)) = \eta_0 Q^n(\varphi) = \gamma_n(\varphi).$$

Note that

$$\eta_0 = \eta_\infty \implies \mathbb{E}(\gamma_n^N(\varphi)) = \eta_\infty(Q^n(\varphi)) = \lambda^n \eta_\infty(\varphi).$$

### 3.2 Fix Lagged Estimator

For any given  $(l, k, \varphi) \in \mathbb{N}_0^2 \times \mathcal{B}_b(\mathbf{E})$  we have

$$\gamma_{l+k}^N(\varphi)/\gamma_k^N(1) = \eta_{l+k}^N(\varphi) \prod_{k \leq p < k+l} \eta_p^N(G) = F_\varphi(\xi_k, \xi_{k+1}, \dots, \xi_{k+l})$$

with

$$F_\varphi(\xi_k, \xi_{k+1}, \dots, \xi_{k+l}) = m(\xi_{l+k})(\varphi) \prod_{k \leq p < k+l} m(\xi_p)(G).$$

The estimate we want to analyze is based upon:

$$\mu_n^{l,N}(\varphi) := \frac{1}{n} \sum_{0 \leq k < n} \gamma_{l+k}^N(\varphi)/\gamma_k^N(1) = \frac{1}{n} \sum_{0 \leq k < n} F_\varphi(\xi_k, \xi_{k+1}, \dots, \xi_{k+l}).$$

The estimator is:

$$\bar{\mu}_n^{l,N}(\varphi) := \frac{\mu_n^{l,N}(\varphi)}{\mu_n^{l,N}(1)} = \frac{\frac{1}{n} \sum_{0 \leq k < n} F_\varphi(\xi_k, \xi_{k+1}, \dots, \xi_{k+l})}{\frac{1}{n} \sum_{0 \leq k < n} F_1(\xi_k, \xi_{k+1}, \dots, \xi_{k+l})}. \quad (3.2)$$

To understand the intuition for this estimator, observe that

$$l = 0 \implies \gamma_k^N(\varphi)/\gamma_k^N(1) = \eta_k^N(\varphi)$$

and more generally for any  $l \geq 0$  we have, noting (3.1):

$$\frac{\mathbb{E}(\gamma_{k+l}^N(\varphi) \mid \xi_k)}{\gamma_k^N(1)} = \frac{\gamma_k^N Q^l(\varphi)}{\gamma_k^N(1)} = \eta_k^N Q^l(\varphi)$$

and hence

$$\mathbb{E}(\mu_n^{l,N}(\varphi)) = \frac{1}{n} \sum_{0 \leq k < n} \mathbb{E}(\gamma_{l+k}^N(\varphi)/\gamma_k^N(1)) = \frac{1}{n} \sum_{0 \leq k < n} \mathbb{E}(\eta_k^N Q^l(\varphi)).$$

The limiting object is given by

$$\frac{1}{n} \sum_{0 \leq k < n} \eta_k(Q^l(\varphi)) \simeq_{n \uparrow \infty} \eta_\infty(Q^l(\varphi)) = \lambda^l \eta_\infty(\varphi).$$

As a result, we expect that for long-time intervals  $n$  that  $\bar{\mu}_n^{l,N}(G)$  should give a good approximation of  $\lambda$  the eigenvalue; or indeed,  $\bar{\mu}_n^{l,N}(\varphi)$  an approximation of  $\eta_\infty(\varphi)$ . What is important, is to understand the bias and variance of this estimator, precisely in the long-time regime, which is the topic of the next section. Particularly, what the effect of  $l \in \mathbb{N}$  is.

## 4 Main Results

### 4.1 Assumption

(A1) There exists a  $C < +\infty$  such that

$$\sup_{(x,y) \in \mathbf{E}^2} \frac{G(x)}{G(y)} \leq C.$$

There exists a  $(\theta, \beta) \in (0, 1) \times \mathcal{P}(\mathbf{E})$  such that for any  $(x, A) \in \mathbf{E} \times \mathcal{E}$

$$\theta \int_A \beta(dy) \leq \int_A M(x, dy) \leq \theta^{-1} \int_A \beta(dy).$$

This assumption is well-studied in the context of particle approximations of Feynman-Kac formulae; see [11] for example. The assumption will typically hold in scenarios where the state-space  $\mathbf{E}$  is compact and seldom holds otherwise. For instance, it is satisfied for Markov transitions of elliptic diffusions on compact manifolds  $\mathbf{E}$ , see for instance the pioneering work of [1, 26, 32] on Gaussian estimates for heat kernels on manifolds. The assumption can be relaxed using the methods in [4, 9, 18, 19, 33, 36] at the cost of more technical proofs that one would expect to lead to the same qualitative conclusions. Note that (A1) assures that for  $\varphi \in \mathcal{B}_b(\mathbf{E})$

$$\eta_\infty(\varphi) := \lim_{n \rightarrow \infty} \frac{\gamma_n(\varphi)}{\gamma_n(1)}.$$

is well-defined. Indeed in [11, 12, 13, 16, 34] rates of convergence of  $\eta_n(\varphi)$  to  $\eta_\infty(\varphi)$  have been established.

### 4.2 Particle System as a Markov Chain

We can consider the particle system  $\xi_0, \xi_1, \dots$  as a time-homogeneous Markov chain on  $(\mathbf{E}^N, \mathcal{E}^{\otimes N})$  with initial distribution  $\prod_{i=1}^N \eta_0(d\xi_0^i)$  and transition  $R^N(\xi_{p-1}, d\xi_p) = \prod_{i=1}^N \Phi(m(\xi_{p-1}))(d\xi_p^i)$ . Moreover, under (A1) the Markov chain is uniformly ergodic and there exist a unique invariant measure, which we shall call  $\Pi^N$ . Note that it is simple to see that  $\Pi^N$  is deterministic, in that it does not depend on the randomness of the simulated particle system and, indeed, an exact expression is available; see [23, Corollary 1] for example. Now as the particle system is exchangeable, the marginal of  $\Pi^N$  in any co-ordinate is the same and we will write it as  $\pi^N \in \mathcal{P}(\mathbf{E})$ . We have the following result which shall prove useful in the subsequent discussion.

**Proposition 4.1.** *Assume (A1). Then there exists a  $C \in (0, \infty)$  such that for any  $N \in \mathbb{N}$ :*

$$\|\pi^N - \eta_\infty\|_{\text{tv}} \leq \frac{C}{N}.$$

*Proof.* Recall from [14, 15] the time-uniform estimate, for  $\varphi \in \mathcal{B}_b(\mathbf{E})$

$$\sup_{n \geq 0} |\mathbb{E} [\eta_n^N(\varphi) - \eta_n(\varphi)]| \leq \frac{C \|\varphi\|}{N}. \quad (4.1)$$

Now we have that

$$\begin{aligned} |\mathbb{E} [\varphi(\xi_n^1) - \eta_\infty(\varphi)]| &= |\mathbb{E} [\eta_n^N(\varphi) - \eta_\infty(\varphi)]| \\ &\leq |\mathbb{E} [\eta_n^N(\varphi) - \eta_n(\varphi)]| + |\eta_n(\varphi) - \eta_\infty(\varphi)|. \end{aligned}$$

Then on noting that

$$|\pi^N(\varphi) - \eta_\infty(\varphi)| = \lim_{n \rightarrow +\infty} |\mathbb{E} [\varphi(\xi_n^1) - \eta_\infty(\varphi)]|$$

the result is concluded by using (4.1),  $\lim_{n \rightarrow +\infty} |\eta_n(\varphi) - \eta_\infty(\varphi)| = 0$  and basic properties of the total variation distance.  $\square$



### 4.3 The Effect of the Lag

**Theorem 4.1.** *Assume (A1). Then for any  $(N, l, \varphi) \in \mathbb{N}^2 \times \mathcal{B}_b(\mathbb{E})$  we have almost surely:*

$$\lim_{n \rightarrow +\infty} \bar{\mu}_n^{l,N}(\varphi) - \eta_\infty(\varphi) = \Phi^l(\pi^N)(\varphi) - \Phi^l(\eta_\infty)(\varphi)$$

*Proof.* Using the exposition in Section 4.2 we can use Proposition B.1 in the Appendix along with a simple first Borel-Cantelli argument, to determine that

$$\mu_n^{l,N}(\varphi) \rightarrow_{a.s.} \int_{\mathbb{E}^{(l+1)N}} m(\xi_{l+1})(\varphi) \left\{ \prod_{p=1}^{l-1} m(\xi_p)(G) \right\} \Pi^N(d\xi_1) \prod_{p=1}^{l+1} R^N(\xi_{p-1}, d\xi_p). \quad (4.2)$$

where  $\rightarrow_{a.s.}$  denotes almost sure convergence as  $n \rightarrow \infty$ . Moreover, by using standard properties of particle filters, (see e.g. [11, Chapter 7]) the R.H.S. of (4.2) is equal to

$$\Pi^N(m(Q^l(\varphi)))$$

where  $m$  is the  $N$ -equally weighted empirical measure. Therefore we one has that

$$\bar{\mu}_n^{l,N}(\varphi) \rightarrow_{a.s.} \frac{\Pi^N(m(Q^l(\varphi)))}{\Pi^N(m(Q^l(1)))}. \quad (4.3)$$

As the particle system is exchangeable, the marginal of  $\Pi^N$  in any co-ordinate is the same, we have that

$$\frac{\Pi^N(m(Q^l(\varphi)))}{\Pi^N(m(Q^l(1)))} = \frac{\pi^N(Q^l(\varphi))}{\pi^N(Q^l(1))} = \Phi^l(\pi^N)(\varphi).$$

As

$$\eta_\infty(\varphi) = \Phi^l(\eta_\infty)(\varphi)$$

the proof is completed.  $\square$

**Corollary 4.1.** *Assume (A1). Then there exists a  $(C, \kappa) \in (0, \infty)^2$  such that for any  $(N, l, \varphi) \in \mathbb{N}^2 \times \mathcal{B}_b(\mathbb{E})$  we have almost surely:*

$$\lim_{n \rightarrow +\infty} |\bar{\mu}_n^{l,N}(\varphi) - \eta_\infty(\varphi)| \leq \frac{C e^{-\kappa l} \|\varphi\|}{N}.$$

*Proof.* Using Theorem 4.1 the proof can be completed by using the exponential stability property of the semigroup  $\Phi(\cdot)$  which holds under (A1); see [11, Chapter 4] for example. More precisely, we have the exponential decay

$$\|\Phi^l(\pi^N) - \Phi^l(\eta_\infty)\|_{\text{tv}} \leq C e^{-\kappa l} \|\pi^N - \eta_\infty\|_{\text{tv}}$$

for some constants  $(C, \kappa) \in (0, \infty)^2$  that do not depend upon  $(N, l, \varphi)$ . Application of Proposition 4.1 allows one to conclude.  $\square$

**Remark 4.1.** *The approach here relies on a novel path-based Markov chain  $\mathbb{L}_p$ -bound using a martingale plus remainder structure. It is not the first theoretical analysis for particle methods that keeps  $N$  fixed and allows another parameter to grow; see [4, 35].*

The implication of Theorem 4.1 and Corollary 4.1 is rather interesting. The results say that for the asymptotic in  $n$  regime, that the almost sure  $\mathbb{L}_1$ -error is exponentially small as a function of  $l$  *irregardless of the number of walkers*, that is one can take  $N = \mathcal{O}(1)$ . From a practical point of view it suggests that a small number of walkers can be run for a long time and so long as  $l$  is moderate, the error of the estimator should be small. This shows that the parameter  $l$  provides quite a substantial freedom. For instance, one could estimate  $\lambda$  using the estimator:

$$\bar{\eta}_n^N(G) := \frac{1}{n} \sum_{k=0}^{n-1} \eta_k^N(G) \quad (4.4)$$

for which, using the approaches in [11] and under mathematical assumptions would have a bias

$$\left| \mathbb{E} \left[ \frac{1}{n} \sum_{k=0}^{n-1} \{ \eta_k^N(G) - \eta_k(G) \} \right] \right|$$

of  $\mathcal{O}(N^{-1})$  for any fixed  $n$ . Then it is simple to establish that the asymptotic in  $n$  bias is upper-bounded by a term that is  $\mathcal{O}(N^{-1})$ . What this means is that the only way in which this conventional estimate can indeed recover  $\lambda$ , for long-time periods, is by increasing the number of walkers, which could be substantially more expensive than using the estimator (3.2). More formally, when considering the estimator (3.2), for  $\epsilon \in (0, 1)$  given, asymptotically in  $n$ , to obtain an error of  $\mathcal{O}(\epsilon)$ , one can choose  $l = \mathcal{O}(-\log(\epsilon))$ ,  $N = \mathcal{O}(1)$ . When considering the estimate (4.4) then  $N = \mathcal{O}(\epsilon^{-1})$  to have an asymptotic bias of  $\mathcal{O}(\epsilon)$ . The cost of this computation could be substantially more than considering (3.2) depending on the relative cost of computing (3.2) over (4.4).

#### 4.4 Non-Asymptotic Bounds

We now investigate the non-asymptotic  $\mathbb{L}_1$ -error, where  $n, N, l$  are all fixed. Below  $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ . We have the following, uniform in time, non-asymptotic error bound.

**Theorem 4.2.** *Assume (A1). Then there exists a  $C \in (0, \infty)$  such that for any  $(n, N, l, \varphi) \in \mathbb{N}_0 \times \mathbb{N}^2 \times \mathcal{B}_b(\mathbb{E})$ :*

$$\mathbb{E} [ |\bar{\mu}_n^{l,N}(\varphi) - \bar{\mu}_n^l(\varphi)| ] \leq \frac{C \|\varphi\| l}{\sqrt{N}}.$$

*Proof.* We have:

$$\bar{\mu}_n^{l,N}(\varphi) - \bar{\mu}_n^l(\varphi) = T_1 + T_2$$

where

$$\begin{aligned} T_1 &= \frac{\mu_n^{l,N}(\varphi)}{\mu_n^{l,N}(1) \mu_n^l(1)} (\mu_n^l(1) - \mu_n^{l,N}(1)) \\ T_2 &= \frac{1}{\mu_n^l(1)} (\mu_n^{l,N}(\varphi) - \mu_n^l(\varphi)). \end{aligned}$$

The proof is easily completed by using Minkowski, the fact that

$$\frac{\mu_n^{l,N}(\varphi)}{\mu_n^{l,N}(1)} \leq \|\varphi\|$$

and Lemma C.2. □

**Remark 4.2.** *If one wants to consider  $\mathbb{E} [ |\bar{\mu}_n^{l,N}(\varphi) - \eta_\infty(\varphi)| ]$  then by using results on the rates of the bias (see [13, 34]) one would, under (A1), have an upper-bound of the type:*

$$\mathcal{O} \left( \frac{l}{\sqrt{N}} + \kappa^n \right)$$

for  $\kappa \in (0, 1)$ .

Theorem 4.2 shows that one cannot naively increase  $l$  so as to obtain a small error (as in Corollary 4.1) and there is at most a linear increase in the non-asymptotic  $\mathbb{L}_1$ -error. On the basis of Corollary 4.1 and Theorem 4.2 one expects that the asymptotic (in  $n$ )  $\mathbb{L}_1$ -error is at most

$$\mathcal{O} \left( \min \left\{ \frac{e^{-\kappa l}}{N}, \frac{l}{\sqrt{N}} \right\} \right)$$

so that again one can choose  $l = \mathcal{O}(-\log(\epsilon))$ ,  $N = \mathcal{O}(1)$  to obtain the error as  $\mathcal{O}(\epsilon)$  ( $\epsilon \in (0, 1)$ ). That is, there is not a contradiction with our previous discussion. None-the-less, the afore-mentioned points are based upon asymptotics in  $n$ . For instance, combining Theorem 4.2 and Remark 4.2 choose  $n = \mathcal{O}(-\log(\epsilon))$ ,  $l = f(\epsilon)$ , where  $f$  is a non-decreasing function which explodes at zero and  $N = \mathcal{O}(\epsilon^{-2}f(\epsilon)^2)$  would achieve a non-asymptotic  $\mathbb{L}_1$ -error of  $\mathcal{O}(\epsilon)$ . This again establishes the flexibility of the estimator where one can of course control the error with  $N$  if needed.

## 4.5 Central Limit Theorem

Whilst the previous results give a characterization of the bias and variance for large  $n$ , one can give exact expressions of this asymptotic error. This is done in the following central limit theorem.

We require several notations, which are given now. For  $(\xi_0, \dots, \xi_l, \varphi) \in \mathbb{E}^{N(l+1)} \times \mathcal{B}_b(\mathbb{E})$  define:

$$\widehat{F}_\varphi(\xi_0, \dots, \xi_l) := F_\varphi(\xi_0, \dots, \xi_l) - \pi^N(Q^l(\varphi)) + \sum_{q=1}^{\infty} \left\{ \mathbb{E} \left[ \Phi(\eta_{(q-1)l}^N)(Q^l(\varphi)) | \Xi_0 = \xi_l \right] - \pi^N(Q^l(\varphi)) \right\} \quad (4.5)$$

where  $\mathbb{E}[\cdot | \Xi_0 = \xi_l]$  is the expectation w.r.t. the law associated to the DMC algorithm, with the initial particles at time zero equal to  $\xi_l$ . Using (A1) one can show that the function  $\widehat{F}_\varphi$  is upper-bounded by a constant (which would explode as  $N$  grows). Now define for  $(q, r) \in \{1, \dots, n\} \times \{0, \dots, l\}$

$$(R^N)^{\otimes(l-r)}(\widehat{F}_\varphi)(\xi_q, \dots, \xi_{q+r}) := \int_{\mathbb{E}^{N(l-r)}} \widehat{F}_\varphi(\xi_q, \dots, \xi_{q+l}) \prod_{j=q+r+1}^{q+l} R^N(\xi_{j-1}, d\xi_j) \quad (4.6)$$

$$(R^N)^{\otimes(l+1-r)}(\widehat{\varphi})(\xi_q, \dots, \xi_{q+r-1}) := \int_{\mathbb{E}^{N(l+1-r)}} \widehat{F}_\varphi(\xi_q, \dots, \xi_{q+l}) \prod_{j=q+r}^{q+l} R^N(\xi_{j-1}, d\xi_j) \quad (4.7)$$

with the convention that  $(\xi_q, \xi_{q-1}) = \xi_{q-1}$ . We will also write

$$(R^N)^{\otimes(0)}(\widehat{\varphi})(\xi_q, \dots, \xi_{q+l}) = \widehat{F}_\varphi(\xi_q, \dots, \xi_{q+l}). \quad (4.8)$$

Next, for  $(l, k, \varphi) \in \mathbb{N} \times \{l, l+1, \dots\} \times \mathcal{B}_b(\mathbb{E})$

$$\sigma_{F_\varphi}^2 := \int_{\mathbb{E}^{Nl}} \tilde{F}_\varphi(\xi_{k-l}, \dots, \xi_{k-1}) \Pi^N(d\xi_{k-l}) \prod_{s=k-l+1}^{k-1} R^N(\xi_{s-1}, d\xi_s) \quad (4.9)$$

$$\begin{aligned} \tilde{F}_\varphi(\xi_{k-l}, \dots, \xi_{k-1}) &= \int_{\mathbb{E}^N} \left( \sum_{j=0}^l (R^N)^{\otimes(l-j)}(\widehat{F}_\varphi)(\xi_{k-j}, \dots, \xi_k) \right)^2 R^N(\xi_{k-1}, d\xi_k) - \\ &\quad \left( \sum_{j=0}^l (R^N)^{\otimes(l+1-j)}(\widehat{F}_\varphi)(\xi_{k-j}, \dots, \xi_{k-1}) \right)^2. \end{aligned} \quad (4.10)$$

Finally set

$$\begin{aligned} \overline{F}_\varphi(\xi_0, \dots, \xi_l) &:= \frac{1}{\pi^N(Q^l(1))} F_\varphi(\xi_0, \dots, \xi_l) - \frac{\pi^N(Q^l(\varphi))}{\pi^N(Q^l(1))^2} F_1(\xi_0, \dots, \xi_l) \\ \widehat{\overline{F}}_\varphi(\xi_0, \dots, \xi_l) &:= \frac{1}{\pi^N(Q^l(1))} \widehat{F}_\varphi(\xi_0, \dots, \xi_l) - \frac{\pi^N(Q^l(\varphi))}{\pi^N(Q^l(1))^2} \widehat{F}_1(\xi_0, \dots, \xi_l). \end{aligned}$$

Below we use  $\xrightarrow{d}$  to denote convergence in distribution as  $n$  increases.

**Theorem 4.3.** Assume (A1). For any  $(N, l, \varphi) \in \mathbb{N}^2 \times \mathcal{B}_b(\mathbb{E})$  we have

$$\sqrt{n}(\bar{\mu}_n^{l,N}(\varphi) - \Phi^l(\pi^N)(\varphi)) \xrightarrow{d} \mathcal{N}(0, \sigma_{F_\varphi}^2).$$

*Proof.* This is Proposition B.3 in the appendix, amended to the notation used in the main text. For instance, one can compare (4.6)-(4.8) with (B.4)-(B.6) and (4.9)-(4.10) with (B.8)-(B.9).  $\square$

**Remark 4.3.** We note that

$$\sqrt{n}|\Phi^l(\pi^N)(\varphi) - \eta_\infty(\varphi)| \leq \frac{C\sqrt{n}\exp\{-\kappa l\}}{N}$$

which means one must choose  $l$  and  $N$  appropriately to control  $\sqrt{n}|\Phi^l(\pi^N)(\varphi) - \eta_\infty(\varphi)|$ . For instance one could have  $n = \mathcal{O}(\epsilon^{-2})$ ,  $l = \mathcal{O}(-\log(\epsilon))$  and  $N = \mathcal{O}(1)$  to make this latter term small for large  $n$ .

Theorem 4.3 gives an *exact* description of the errors of the estimator in  $N$  and  $l$  through the expression of the asymptotic variance  $\sigma_{F_\varphi}^2$ . As is typical in Markov chain CLTs, this latter variance is written in terms of a solution to the Poisson equation (see e.g. [20]) as in (4.5).  $\sigma_{F_\varphi}^2$  is further complicated as one has the form of a ratio of lagged Markov chain estimators. In general one would need numerical methods to approximate the asymptotic variance and this is considered below. Understanding how  $\sigma_{F_\varphi}^2$  behaves as a function of  $N, l$  is an important mathematical question, which requires further investigation, but one might expect that it is related to Theorem 4.2 and this is then  $\mathcal{O}(l^2/N)$ .

To conclude this section on CLTs one can also give a result related to an approach that is considered in the literature. In some cases, authors have proposed running two different (independent) DMC algorithms, one that gives the estimate  $\mu_n^{l,N}(\varphi)$  and the other which gives the estimate  $\mu_n^{l,N}(1)$ . In this case defining

$$\sigma_\varphi^{2,\text{ind}} = \frac{1}{\pi^N(Q^l(1))} \sigma_{F_\varphi}^2 + \frac{\pi^N(Q^l(\varphi))^2}{\pi^N(Q^l(1))^4} \sigma_{F_1}^2$$

one can apply Proposition B.4 in the appendix to show that the asymptotic variance for such estimators (under (A1)) is exactly  $\sigma_\varphi^{2,\text{ind}}$ . In general, one expects that  $\sigma_{F_\varphi}^2 < \sigma_\varphi^{2,\text{ind}}$  but proving this is rather difficult. We consider a numerical comparison below.

## 4.6 Numerical Study

Consider the quantum harmonic oscillator Hamiltonian

$$H = -\frac{1}{2m}\nabla^2 + \frac{1}{2}m\omega^2 x^2.$$

The ground state energy (smallest eigenvalue) of the operator  $H$  is known to be  $E_0 = (n + \frac{1}{2})\omega$ . Following the discussion in Appendix A, we set  $\tau = 1/16$  (time discretization),  $\omega = 1$ , and  $m = 1$ . Define  $G(x) = e^{-\tau x^2/2}$  and  $M(\cdot, dx)$  to be the Brownian motion transition kernel (over time unit  $\tau$ ). The operator  $Q$  is given by  $Q = e^{-\tau H}$  and the largest eigenvalue of the operator  $Q$  is given by  $\lambda = e^{-\tau E_0}$ .

We ran the DMC algorithm and calculated the fixed lag estimator(3.2) with  $\varphi = G$ . We varied the lag values from 0 to 50, with  $N = 10$  and  $n = 50000$ . In order to calculate the bias and the variance we conducted 128 independent runs. For each lag value, we obtain an estimator of  $\lambda$ . Lag 0 case corresponds to the standard method the DMC particles are used to estimate  $\lambda$ . Non-zero lag values correspond to the estimators studied in this paper. Corollary 4.1 predicts that for a fixed number of particles  $N$  and a large  $n$ , the bias will decay at an exponential rate. Figure 1 presents two plots, one for the absolute bias and one for the log absolute bias each plotted against the different lag values. The figure exhibits a linear trend in the log absolute bias plot which affirms the statement about the exponential decaying bias stated in Corollary 4.1.

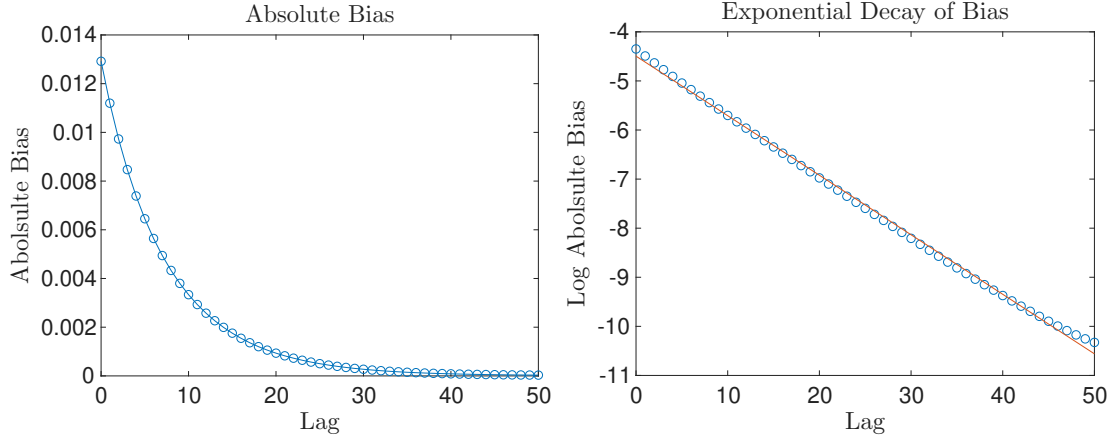


Figure 1: Bias Comparison. Left: Absolute Bias vs Lag. Right: Log Absolute Bias vs Lag.

Let  $(\xi_i)_{i \geq 0}$  and  $(\tilde{\xi}_i)_{i \geq 0}$  be identically distributed and independent DMC particles. Consider the following two estimators:

$$\bar{\mu}_n^{l,N} = \frac{\frac{1}{n} \sum_{0 \leq k \leq n} F_G(\xi_k, \dots, \xi_{k+l})}{\frac{1}{n} \sum_{0 \leq k \leq n} F_1(\xi_k, \dots, \xi_{k+l})}, \quad \tilde{\mu}_n^{l,N} = \frac{\frac{1}{n} \sum_{0 \leq k \leq n} F_G(\xi_k, \dots, \xi_{k+l})}{\frac{1}{n} \sum_{0 \leq k \leq n} F_1(\tilde{\xi}_k, \dots, \tilde{\xi}_{k+l})}.$$

Figure 2 shows the variance of both estimators under the same setting used for the bias simulation above. The simulation results indicate that the variance of the estimator  $\bar{\mu}_n^{l,N}$  is much smaller than the variance of  $\tilde{\mu}_n^{l,N}$ . This is because we expect that the numerator and denominator of  $\bar{\mu}_n^{l,N}$  are highly positively correlated which reduces the overall variance of the estimator  $\bar{\mu}_n^{l,N}$ . Moreover, comparing Figures 1 and 2 reveals an opposing trend between the bias and the variance as the lag value increases. The bias exponentially decreases, while the variance increases.

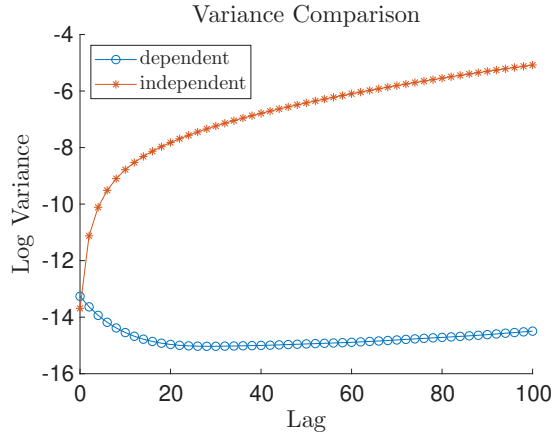


Figure 2: Variance Comparison.

## Acknowledgements

AJ was supported by CUHK-SZ UDF01003537. MC thanks the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 863481) for financial support.

## A Feynman-Kac formula in physics

In physics we are interested in solving the imaginary-time dependent Schrödinger equation

$$\frac{\partial \psi(\mathbf{x}, t)}{\partial t} = -H\psi(\mathbf{x}, t) \quad (\text{A.1})$$

with some function  $\psi(\mathbf{x}, 0)$  as initial condition and  $\mathbf{x} \in \mathbf{E} = \mathbb{R}^d$  with  $d \geq 1$ . Here,  $H$  is the Schrödinger Hamiltonian given by

$$H = -\frac{1}{2}\nabla^2 + V(\mathbf{x}).$$

The solution of Eq.(A.1) is expressed using the Feynman-Kac formula as follows

$$\psi(\mathbf{x}, t) = \mathbb{E}_{\mathbf{x}}[e^{-\int_0^t ds V(\mathbf{X}(s))}]$$

where the expectation  $\mathbb{E}_{\mathbf{x}}$  is with respect to the law of the Brownian process with initial distribution  $\psi(\mathbf{x}, 0)$  and final condition  $\mathbf{X}(t) = \mathbf{x}$ . Using the notations of the present article, this FK formula can be put into the form

$$\gamma_t(\phi) \equiv \int d\mathbf{x} \phi(\mathbf{x}) \psi(\mathbf{x}, t) = \mathbb{E}[\phi(\mathbf{X}(t)) e^{-\int_0^t ds V(\mathbf{X}(s))}].$$

In DMC the continuous time variable needs to be discretized. Writing  $t = n\tau$  where  $\tau$  is a time-step (a "small" positive quantity),  $\gamma_t(\phi)$  is approximated by

$$\gamma_n(\phi) = \mathbb{E}[\phi(\mathbf{X}_n) \prod_{p=0}^{n-1} G(\mathbf{X}_p)] \quad (\text{A.2})$$

with  $G(\mathbf{x}) = e^{-\tau V(\mathbf{x})}$  and initial distribution  $\eta_0(\mathbf{x}) = \psi(\mathbf{x}, 0)$ . As seen,  $\gamma_n(\phi)$  is the unrenormalized Feynman-Kac measure introduced in this work, see section 2.2, with operator  $Q$  given by  $e^{-tH}$ .

In physics where the number of degrees of freedom  $d$  of the quantum systems studied is large (typically,  $d$  is proportional to the number of *physical* particles), DMC calculations using Eq.(A.2) are just unfeasible because of the uncontrolled fluctuations of  $G(\mathbf{x})$ . In practice, this problem is solved by introducing importance sampling. Let  $\psi_G : \mathbf{E} \rightarrow \mathbb{R}^+$  be a so-called guiding wavefunction (in practice, a good approximation of the ground-state wavefunction). Let us introduce the distribution  $f(\mathbf{x}, t) = \psi_G(\mathbf{x})\psi(\mathbf{x}, t)$ . The partial differential equation obeyed by  $f$  is easily shown to be

$$\frac{\partial f(\mathbf{x}, t)}{\partial t} = Lf(\mathbf{x}, t) - E_L(\mathbf{x})f(\mathbf{x}, t)$$

where  $L$  is a Fokker-Planck operator written as

$$L = \frac{1}{2}\nabla^2 - \nabla[\mathbf{b}.] \quad (\text{A.3})$$

where the drift vector is  $\mathbf{b}(\mathbf{x}) = \frac{\nabla \psi_G}{\psi_G}$  and  $E_L(\mathbf{x})$  is a new potential function, called the local energy, given by

$$E_L(\mathbf{x}) = \frac{H\psi_G}{\psi_G}. \quad (\text{A.4})$$

The FK formula, Eq.(A.2), becomes

$$\gamma_n(\psi_G \phi) = \mathbb{E}[\phi(\mathbf{X}_n) \prod_{p=0}^{n-1} G(\mathbf{X}_p)]$$

where the expectation,  $\mathbb{E}$ , now refers to the law of the *drifted* Brownian process associated with the Fokker-Planck operator, Eq.(A.3), the weight is  $G(x) = e^{-\tau E_L(\mathbf{x})}$ , and the initial condition is  $\eta_0(\mathbf{x}) = \psi_G(\mathbf{x})\psi(\mathbf{x}, 0)$ . Note that the operator  $Q$  is now given by  $Q = e^{-\tau \psi_G H \frac{1}{\psi_G}}$ .

Introducing importance sampling has two important consequences which make in practice the DMC simulations feasible. First, the drift contribution of the Markov kernel "pushes" the particles toward the regions where  $\Psi_G$  is important (large), thus avoiding to accumulate statistics in regions where the product  $\prod_{p=0}^{n-1} G(\mathbf{X}_p)$  is (very) small. Second, the statistical fluctuations of the weights are much reduced since the variations of the local energy, Eq.(A.4), are directly related to the quality of the approximate guiding wavefunction (no fluctuations when  $\psi_G$  is the exact ground-state).

## B Proofs

### B.1 Structure of Appendix

This appendix is split into two sections. The first is Section B.2 which considers some technical results for Markov chains, which are applied particularly for the proof of Theorem 4.1 and also Section B.2.1 used for the proof of Theorem 4.3. Section C houses all the proofs for Theorem 4.2. Throughout this appendix  $C$  is generic finite constant whose value may change from line-to-line of the computations and any dependencies in terms of the parameters of DMC or of FK models will be made clear in each context.

### B.2 Technical Results for Markov Chains

The notations in this section, in particular the symbols used, should be taken as independent of the main text. The reason for this is that these results are of independent interest and do not need the notation of DMC to be written. We consider a time-homogenous Markov chain on general state-space  $(\mathbf{X}, \mathcal{X})$  of initial distribution  $\mu$  and kernel  $K$ , with the latter having invariant measure  $\pi$ . Let  $l \in \mathbb{N}$  be fixed and consider for  $\varphi \in \mathcal{B}_b(\mathbf{X})$

$$\frac{1}{n} \sum_{q=1}^n \varphi(X_{q:q+l-1}).$$

We shall prove, under assumptions, convergence of the above quantity to

$$\pi \otimes K^{\otimes(l-1)}(\varphi) = \int_{\mathbf{X}^l} \varphi(x_{1:l}) \pi(dx_1) \prod_{j=2}^l K(x_{j-1}, dx_j)$$

where we use the notation  $x_{1:l} = (x_1, \dots, x_l)$ .

**(H1)** There exists a  $(\epsilon, \nu) \in (0, 1) \times \mathcal{P}(\mathbf{X})$  such that for any  $(x, A) \in \mathbf{X} \times \mathcal{X}$  we have

$$\int_A K(x, dy) \geq \epsilon \int_A \nu(dy).$$

**Proposition B.1.** *Assume (H1). Then for any  $p \geq 1$  there exist a  $C < +\infty$  such that for any  $(n, l, \varphi) \in \mathbb{N}^2 \times \mathcal{B}_b(\mathbf{X}^l)$ :*

$$\mathbb{E} \left[ \left| \frac{1}{n} \sum_{q=1}^n \varphi(X_{q:q+l-1}) - \pi \otimes K^{\otimes(l-1)}(\varphi) \right|^p \right]^{1/p} \leq \frac{C \|\varphi\| l \epsilon^{-1}}{\sqrt{n}}$$

where  $\epsilon$  is as (H1) and  $\|\varphi\| = \sup_{x \in \mathbf{X}^l} |\varphi(x)|$ .

*Proof.* Define the function

$$\widehat{\varphi}(x_{1:l}) = \sum_{q \geq 0} \left\{ (K^{\otimes l})^q(\varphi)(x_{1:l}) - \pi \otimes K^{\otimes(l-1)}(\varphi) \right\} \quad (\text{B.1})$$

where for  $q \in \mathbb{N}$

$$\begin{aligned} (K^{\otimes l})^q(\varphi)(x_{1:l}) &= \int_{\mathbf{X}^{lq}} \varphi(x_{1:l}^q) \prod_{j=1}^q K^{\otimes l}(x_{1:l}^{j-1}, dx_{1:l}^j) \\ &= \int_{\mathbf{X}^{lq}} \varphi(x_{1:l}^q) \prod_{j=1}^q \left\{ \prod_{k=1}^l K(x_{k-1}^j, dx_k^j) \right\} \end{aligned}$$

with  $x_{1:l}^0 = x_{1:l}$  and  $x_0^j = x_l^{j-1}$  (notice that there is 0 subscript of  $x$  in the first line). In the summand in (B.1) if  $q = 0$  we set  $(K^{\otimes l})^q(\varphi)(x_{1:l}) = \varphi(x_{1:l})$ . Note that because of (H1) it easily follows that

$$\sup_{x_{1:l} \in \mathbf{X}^l} |\widehat{\varphi}(x_{1:l})| \leq 2\epsilon^{-1} \|\varphi\|.$$

Then it is easy to check that for any  $x_{1:l} \in \mathbf{X}^l$

$$\begin{aligned} \varphi(x_{1:l}) - \pi \otimes K^{\otimes(l-1)}(\varphi) &= \widehat{\varphi}(x_{1:l}) - K^{\otimes l}(\widehat{\varphi})(x_{1:l}) \\ &= \widehat{\varphi}(x_{1:l}) - K^{\otimes l}(\widehat{\varphi})(x_l). \end{aligned}$$

Now, denoting the natural filtration of the Markov chain  $(X_k)_{k \geq 0}$  as  $(\mathcal{F}_k)_{k \geq 0}$ , we remark that

$$\mathbb{E}[\varphi(X_{q:q+l-1}) | \mathcal{F}_{q-1}] = K^{\otimes l}(\widehat{\varphi})(x_{q-1}).$$

Therefore it easily follows that

$$\begin{aligned} \frac{1}{n} \sum_{q=1}^n \left\{ \varphi(X_{q:q+l-1}) - \pi \otimes K^{\otimes(l-1)}(\varphi) \right\} &= \frac{1}{n} \sum_{q=1}^n \left\{ \widehat{\varphi}(x_{q:q+l-1}) - K^{\otimes l}(\widehat{\varphi})(x_q) \right\} \\ &= \frac{1}{n} \{M_n + R_n\} \end{aligned} \tag{B.2}$$

where

$$\begin{aligned} M_n &= \sum_{q=1}^n \left\{ \widehat{\varphi}(x_{q:q+l-1}) - K^{\otimes l}(\widehat{\varphi})(x_{q-1}) \right\} \\ R_n &= \left\{ K^{\otimes l}(\widehat{\varphi})(x_0) - K^{\otimes l}(\widehat{\varphi})(x_n) \right\}. \end{aligned}$$

Then combining (B.2) with the Minkowski inequality we have that

$$\mathbb{E} \left[ \left| \frac{1}{n} \sum_{q=1}^n \varphi(X_{q:q+l-1}) - \pi \otimes K^{\otimes(l-1)}(\varphi) \right|^p \right]^{1/p} \leq \mathbb{E} \left[ \left| \frac{1}{n} M_n \right|^p \right]^{1/p} + \frac{C \|\varphi\| \epsilon}{n} \tag{B.3}$$

Now we adopt the notation for any  $(q, r) \in \{1, \dots, n\} \times \{0, \dots, l-1\}$

$$K^{\otimes(l-r-1)}(\widehat{\varphi})(x_{q:q+r}) := \int_{\mathbf{X}^{l-r-1}} \widehat{\varphi}(x_{q:q+l-1}) \prod_{j=q+r+1}^{q+l-1} K(x_{j-1}, dx_j) \tag{B.4}$$

$$K^{\otimes(l-r)}(\widehat{\varphi})(x_{q:q+r-1}) := \int_{\mathbf{X}^{l-r}} \widehat{\varphi}(x_{q:q+l-1}) \prod_{j=q+r}^{q+l-1} K(x_{j-1}, dx_j) \tag{B.5}$$

with the convention that  $x_{q:q-1} = x_{q-1}$ . We will also write

$$K^{\otimes(0)}(\widehat{\varphi})(x_{q:q+l-1}) = \widehat{\varphi}(x_{q:q+l-1}). \tag{B.6}$$



Then we have that

$$M_n = \sum_{q=1}^n \sum_{r=0}^{l-1} \xi_{q,r} \quad (\text{B.7})$$

where

$$\xi_{q,r} := K^{\otimes(l-1-r)}(\hat{\varphi})(x_{q;q+r}) - K^{\otimes(l-r)}(\hat{\varphi})(x_{q;q+r-1}).$$

Now set

$$M_n(r) := \sum_{q=1}^n \xi_{q,r}$$

and note we clearly have  $M_n = \sum_{r=0}^{l-1} M_n(r)$ . For  $r \in \{0, \dots, l-1\}$  fixed define the filtration  $\mathcal{F}_q^r := \sigma(X_0, \dots, X_{q+r})$ ,  $q \geq 1$  with  $\mathcal{F}_0^r := \sigma(X_0)$ . Then it is clear that  $M_n(r)$  is a  $\mathcal{F}_n^r$ -martingale. Then applying the Minkowski inequality  $l$ -times followed by the Burkholder-Gundy-Davis inequality  $l$ -times, we obtain the upper-bound:

$$\mathbb{E} \left[ \left| \frac{1}{n} M_n \right|^p \right]^{1/p} \leq \frac{C \|\varphi\| l \epsilon^{-1}}{\sqrt{n}}$$

and then combining the above upper-bound with (B.3) allows us to conclude.  $\square$

### B.2.1 Central Limit Theorems

Recall the definitions in (B.1) and (B.4)-(B.6) and set for  $(l, k, \varphi) \in \{2, 3, \dots\} \times \{l, l+1, \dots\} \times \mathcal{B}_b(\mathcal{X}^l)$

$$\sigma_\varphi^2 := \int_{\mathcal{X}^{l-1}} \tilde{\varphi}(x_{k-l+1:k-1}) \pi(dx_{k-l+1}) \prod_{s=k-l+2}^{k-1} K(x_{s-1}, dx_s) \quad (\text{B.8})$$

$$\tilde{\varphi}(x_{k-l+1:k-1}) = \int_{\mathcal{X}} \left( \sum_{j=0}^{l-1} K^{\otimes(l-1-j)}(\hat{\varphi})(x_{k-j:k}) \right)^2 K(x_{k-1}, dx_k) - \left( \sum_{j=0}^{l-1} K^{\otimes(l-j)}(\hat{\varphi})(x_{k-j:k-1}) \right)^2. \quad (\text{B.9})$$

Below we use  $\xrightarrow{d}$  to denote convergence in distribution as  $n$  increases.

**Proposition B.2.** Assume (H1). For any  $(l, \varphi) \in \{2, 3, \dots\} \times \mathcal{B}_b(\mathcal{X}^l)$  we have

$$\frac{1}{\sqrt{n}} \sum_{q=1}^n (\varphi(X_{q;q+l-1}) - \pi \otimes K^{\otimes(l-1)}(\varphi)) \xrightarrow{d} \mathcal{N}(0, \sigma_\varphi^2)$$

where  $\sigma_\varphi^2$  is defined in (B.8)-(B.9).

*Proof.* We follow the notation of Proposition B.1. We have  $|\frac{1}{\sqrt{n}} R_n| \leq C \|\varphi\| \epsilon / \sqrt{n} \rightarrow 0$ . For  $\frac{1}{\sqrt{n}} M_n$  we have the representation

$$M_n = \frac{1}{\sqrt{n}} \sum_{k=l}^n \sum_{j=0}^{l-1} \xi_{k-j,j} + \frac{1}{\sqrt{n}} \sum_{k=1}^{l-1} \sum_{j=0}^{l-k} \xi_{k,j} + \frac{1}{\sqrt{n}} \sum_{k=n-l+1}^n \sum_{j=n-k+1}^{l-1} \xi_{k,j}.$$

The number of terms in the second and third terms on the R.H.S. is  $\mathcal{O}(l^2)$  and all the terms are bounded thus they approach zero as  $n \rightarrow \infty$ . For the first term, we use the Martingale array CLT ([21, Corollary 3.1]). Notice that  $\sum_{k=l}^n \sum_{j=0}^{l-1} \xi_{k-j,j}$  is a  $\mathcal{F}_n^0$ -martingale. Let  $\delta > 0$ , we have

$$\begin{aligned} \sum_{k=l}^n \mathbb{E} \left[ \left( \frac{1}{\sqrt{n}} \sum_{j=0}^{l-1} \xi_{k-j,j} \right)^2 \mathbb{1}_{\{|x|>\delta\}} \left( \frac{1}{\sqrt{n}} \sum_{j=0}^{l-1} \xi_{k-j,j} \right) \middle| \mathcal{F}_{k-1}^0 \right] &\leq \frac{1}{n^{3/2}\delta} \sum_{k=l}^n \mathbb{E} \left[ \left( \sum_{j=0}^{l-1} \xi_{k-j,j} \right)^3 \middle| \mathcal{F}_{k-1}^0 \right] \\ &\leq \frac{8l^3(n-l+1)\|\varphi\|\epsilon^{-1}}{n^{3/2}\delta} \rightarrow 0. \end{aligned}$$

For the asymptotic variance, noting (B.9), we use the definition of  $\xi_{k-j,j}$  and write

$$\sum_{k=l}^n \mathbb{E} \left[ \left( \frac{1}{\sqrt{n}} \sum_{j=0}^{l-1} \xi_{k-j,j} \right)^2 \middle| \mathcal{F}_{k-1}^0 \right] = \frac{1}{n} \sum_{k=l}^n \tilde{\varphi}(X_{k-l+1:k-1}) \xrightarrow{\mathbb{P}} \sigma_{\varphi}^2,$$

where  $\rightarrow_{\mathbb{P}}$  denotes convergence in probability as  $n$  grows and this follows from Proposition B.1.  $\square$

**Proposition B.3.** Assume (H1). For any  $(l, \varphi, \psi) \in \{2, 3, \dots\} \times \mathcal{B}_b(\mathbf{X}^l)^2$  we have

$$\sqrt{n} \left( \frac{\frac{1}{n} \sum_{q=1}^n \varphi(X_{q:q+l-1})}{\frac{1}{n} \sum_{q=1}^n \psi(X_{q:q+l-1})} - \frac{\pi \otimes K^{\otimes(l-1)}(\varphi)}{\pi \otimes K^{\otimes(l-1)}(\psi)} \right) \xrightarrow{d} \mathcal{N}(0, \sigma_{\tau}^2),$$

with  $\tau = \frac{1}{\pi \otimes K^{\otimes(l-1)}(\psi)} \varphi - \frac{\pi \otimes K^{\otimes(l-1)}(\varphi)}{(\pi \otimes K^{\otimes(l-1)}(\psi))^2} \psi$ .

*Proof.* Using the delta method, the limit is the same as

$$\begin{aligned} & \frac{1}{\pi \otimes K^{\otimes(l-1)}(\psi)} \frac{1}{\sqrt{n}} \left( \sum_{q=1}^n \varphi(X_{q:q+l-1}) - \pi \otimes K^{\otimes(l-1)}(\varphi) \right) \\ & - \frac{\pi \otimes K^{\otimes(l-1)}(\varphi)}{(\pi \otimes K^{\otimes(l-1)}(\psi))^2} \frac{1}{\sqrt{n}} \left( \sum_{q=1}^n \psi(X_{q:q+l-1}) - \pi \otimes K^{\otimes(l-1)}(\psi) \right). \end{aligned}$$

Similar to Proposition B.2 we use the decomposition

$$\begin{aligned} \xi_{q,r} &= \frac{1}{\pi \otimes K^{\otimes(l-1)}(\psi)} \left( K^{\otimes(l-1-r)}(\hat{\varphi})(x_{q:q+r}) - K^{\otimes(l-r)}(\hat{\varphi})(x_{q:q+r-1}) \right) \\ & - \frac{\pi \otimes K^{\otimes(l-1)}(\varphi)}{(\pi \otimes K^{\otimes(l-1)}(\psi))^2} \left( K^{\otimes(l-1-r)}(\hat{\psi})(x_{q:q+r}) - K^{\otimes(l-r)}(\hat{\psi})(x_{q:q+r-1}) \right) \\ & = K^{\otimes(l-1-r)}(\hat{\tau})(x_{q:q+r}) - K^{\otimes(l-r)}(\hat{\tau})(x_{q:q+r-1}) \end{aligned} \quad (\text{B.10})$$

and doing similar calculations we get the desired result.  $\square$

**Proposition B.4.** Assume (H1). Let  $(\tilde{X}_q)_{q \geq 0}$  be an independent copy of  $(X_q)_{q \geq 0}$ . For any  $(l, \varphi, \psi) \in \{2, 3, \dots\} \times \mathcal{B}_b(\mathbf{X}^l)^2$  we have

$$\sqrt{n} \left( \frac{\frac{1}{n} \sum_{q=1}^n \varphi(X_{q:q+l-1})}{\frac{1}{n} \sum_{q=1}^n \psi(\tilde{X}_{q:q+l-1})} - \frac{\pi \otimes K^{\otimes(l-1)}(\varphi)}{\pi \otimes K^{\otimes(l-1)}(\psi)} \right) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

with

$$\sigma^2 = \frac{1}{(\pi \otimes K^{\otimes(l-1)}(\psi))^2} \sigma_{\varphi}^2 + \frac{(\pi \otimes K^{\otimes(l-1)}(\varphi))^2}{(\pi \otimes K^{\otimes(l-1)}(\psi))^4} \sigma_{\psi}^2,$$

where  $\sigma_{\varphi}^2$  and  $\sigma_{\psi}^2$  are defined via (B.8).

*Proof.* By the delta method the distributional limit of

$$\sqrt{n} \left( \frac{\frac{1}{n} \sum_{q=1}^n \varphi(X_{q:q+l-1})}{\frac{1}{n} \sum_{q=1}^n \psi(\tilde{X}_{q:q+l-1})} - \frac{\pi \otimes K^{\otimes(l-1)}(\varphi)}{\pi \otimes K^{\otimes(l-1)}(\psi)} \right)$$

is the same as the distributional limit of

$$\begin{aligned} & \frac{1}{\pi \otimes K^{\otimes(l-1)}(\psi)} \frac{1}{\sqrt{n}} \left( \sum_{q=1}^n \varphi(X_{q:q+l-1}) - \pi \otimes K^{\otimes(l-1)}(\varphi) \right) \\ & - \frac{\pi \otimes K^{\otimes(l-1)}(\varphi)}{(\pi \otimes K^{\otimes(l-1)}(\psi))^2} \frac{1}{\sqrt{n}} \left( \sum_{q=1}^n \psi(\tilde{X}_{q:q+l-1}) - \pi \otimes K^{\otimes(l-1)}(\psi) \right) \end{aligned}$$

The result follows because of the independence of  $(X_q)_{q \geq 0}$  and  $(\tilde{X}_q)_{q \geq 0}$  and the CLT for each term.  $\square$

## C $\mathbb{L}_1$ -Proofs

**Lemma C.1.** *We have the following decomposition for any  $(n, l, N, \varphi) \in \mathbb{N}_0 \times \mathbb{N}^2 \times \mathcal{B}_b(\mathbb{E})$  whenever the formulae exist:*

$$\mu_n^{l,N}(\varphi) = \frac{1}{n} \sum_{p=0}^{n-1} \eta_p^N Q^l(\varphi) + \frac{1}{\sqrt{N}} V_n^{l,N}(\varphi)$$

where

$$\begin{aligned} V_n^{l,N}(\varphi) &:= \frac{1}{n} \sum_{p=0}^{n-1} \sum_{k=1}^l \frac{\gamma_{p+k}^N(1)}{\gamma_p^N(1)} V_{p+k}^N(Q^{l-k}(\varphi)) \\ V_n^N(\varphi) &:= \sqrt{N} (\eta_n^N(\varphi) - \Phi(\eta_{n-1}^N)(\varphi)). \end{aligned}$$

*Proof.* We start by noting that

$$\gamma_{p+l}^N - \gamma_p^N Q^l = \sum_{k=1}^l \left( \gamma_{p+k}^N Q^{l-k} - \gamma_{p+(k-1)}^N Q^{l-(k-1)} \right).$$

Conversely, we have

$$\eta_{p+(k-1)}^N Q^{l-(k-1)} = \eta_{p+(k-1)}^N Q Q^{l-k} = \eta_{p+(k-1)}^N(G) \frac{\eta_{p+(k-1)}^N Q Q^{l-k}}{\eta_{p+(k-1)}^N Q(1)} = \eta_{p+(k-1)}^N(G) \Phi(\eta_{p+(k-1)}^N) Q^{l-k}.$$

This yields the decomposition

$$\gamma_{p+l}^N - \gamma_p^N Q^l = \sum_{k=1}^l \gamma_{p+k}^N(1) \left( \eta_{p+k}^N - \Phi(\eta_{p+(k-1)}^N) \right) Q^{l-k}$$

from which we check that

$$\frac{\gamma_{p+l}^N(\varphi)}{\gamma_p^N(1)} - \eta_p^N Q^l(\varphi) = \frac{1}{\sqrt{N}} \sum_{k=1}^l \frac{\gamma_{p+k}^N(1)}{\gamma_p^N(1)} V_{p+k}^N(Q^{l-k}(\varphi))$$

with the centered local perturbation random fields

$$V_n^N = \sqrt{N} (\eta_n^N - \Phi(\eta_{n-1}^N)).$$

This implies that

$$\mu_n^{l,N}(\varphi) = \frac{1}{n} \sum_{p=0}^{n-1} \eta_p^N Q^l(\varphi) + \frac{1}{\sqrt{N}} V_n^{l,N}(\varphi)$$

with

$$V_n^{l,N}(\varphi) := \frac{1}{n} \sum_{p=0}^{n-1} \sum_{k=1}^l \frac{\gamma_{p+k}^N(1)}{\gamma_p^N(1)} V_{p+k}^N(Q^{l-k}(\varphi))$$

which completes the proof.  $\square$

**Lemma C.2.** *Assume (A1). Then there exists a  $C \in (0, \infty)$  such that for any  $(n, N, l, \varphi) \in \mathbb{N}_0 \times \mathbb{N}^2 \times \mathcal{B}_b(\mathbb{E})$ :*

$$\mathbb{E} \left[ \left\| \frac{\mu_n^{l,N}(\varphi) - \mu_n^l(\varphi)}{\mu_n^l(1)} \right\| \right] \leq \frac{C \|\varphi\| l}{\sqrt{N}}.$$

*Proof.* Using Lemma C.1 and the fact that

$$\mu_n^l(\varphi) = \frac{1}{n} \sum_{p=0}^{n-1} \eta_p(Q^l(\varphi))$$

we have that by the Minkowski inequality

$$\mathbb{E} \left[ \left| \frac{\mu_n^{l,N}(\varphi) - \mu_n^l(\varphi)}{\mu_n^l(1)} \right| \right] \leq T_1 + T_2$$

where

$$\begin{aligned} T_1 &= \frac{1}{n\mu_n^l(1)} \sum_{p=0}^{n-1} \mathbb{E} [|\eta_p^N(Q^l(\varphi)) - \eta_p(Q^l(\varphi))|] \\ T_2 &= \frac{1}{n\mu_n^l(1)} \sum_{p=0}^{n-1} \sum_{k=1}^l \mathbb{E} \left[ \left| \frac{\gamma_{p+k}^N(1)}{\gamma_p^N(1)} \{ \eta_{p+k}^N(Q^{l-k}(\varphi)) - \Phi(\eta_{p+k-1}^N(Q^{l-k}(\varphi))) \} \right| \right]. \end{aligned}$$

We deal with  $T_1$  and  $T_2$  individually to conclude.

For  $T_1$  one has

$$T_1 = \frac{1}{n\mu_n^l(1)} \sum_{p=0}^{n-1} \eta_p(Q^l(1)) \mathbb{E} \left[ \left| \eta_p^N \left( \frac{Q^l(\varphi)}{\eta_p(Q^l(1))} \right) - \eta_p \left( \frac{Q^l(\varphi)}{\eta_p(Q^l(1))} \right) \right| \right].$$

By using standard  $\mathbb{L}_1$ -bounds for Feynman-Kac formula (e.g. [11, Theorem 7.4.4.]) it follows that

$$T_1 \leq C \frac{1}{n\sqrt{N}\mu_n^l(1)} \sum_{p=0}^{n-1} \eta_p(Q^l(1)) \left\| \frac{Q^l(\varphi)}{\eta_p(Q^l(1))} \right\|.$$

By (A1) one can show that for any  $s \in \mathbb{N}_0$

$$\sup_{p \geq 0} \left\| \frac{Q^s(\varphi)}{\eta_p(Q^s(1))} \right\| \leq C \|\varphi\| \quad (\text{C.1})$$

where  $C$  does not depend upon  $s$ ; see [10, Lemma 4.1] for example. Therefore we have shown that

$$T_1 \leq \frac{C \|\varphi\|}{\sqrt{N}}.$$

For  $T_2$  using the conditional i.i.d. property of the particle system one has

$$T_2 \leq \frac{C}{n\sqrt{N}\mu_n^l(1)} \sum_{p=0}^{n-1} \sum_{k=1}^l \eta_p(Q^{l-k}(1)) \left\| \frac{Q^{l-k}(\varphi)}{\eta_p(Q^{l-k}(1))} \right\| \mathbb{E} \left[ \frac{\gamma_{p+k}^N(1)}{\gamma_p^N(1)} \right]$$

and then using (C.1)

$$T_2 \leq \frac{C \|\varphi\|}{n\sqrt{N}\mu_n^l(1)} \sum_{p=0}^{n-1} \sum_{k=1}^l \eta_p(Q^{l-k}(1)) \mathbb{E} \left[ \frac{\gamma_{p+k}^N(1)}{\gamma_p^N(1)} \right].$$

Now

$$\mathbb{E} \left[ \frac{\gamma_{p+k}^N(1)}{\gamma_p^N(1)} \right] = \eta_p(Q^k(1)) \mathbb{E} \left[ \eta_p^N \left( \frac{Q^k(1)}{\eta_p(Q^k(1))} \right) \right]$$

so combining with (C.1) yields

$$T_2 \leq \frac{C}{n\sqrt{N}\mu_n^l(1)} \sum_{p=0}^{n-1} \sum_{k=1}^l \eta_p(Q^{l-k}(1))\eta_p(Q^k(1)).$$

Now due to (A1) there exist finite and positive constants  $C, \bar{C}$  so that for any  $(x, y) \in \mathbb{E}^2$  and any  $s \in \mathbb{N}_0$

$$\bar{C}Q^s(1)(y) \leq Q^s(1)(x) \leq CQ^s(1)(y)$$

so that it follows that there exists a finite  $C$  that does not depend on  $p, k, l$  or  $\eta_p$  so that

$$\frac{\eta_p(Q^{l-k}(1))\eta_p(Q^k(1))}{\eta_p(Q^l(1))} \leq C.$$

Hence

$$T_2 \leq \frac{C\|\varphi\|l\mu_n^l(1)}{\sqrt{N}\mu_n^l(1)} = \frac{C\|\varphi\|l}{\sqrt{N}}$$

and this concludes the proof.  $\square$

**Remark C.1.** Extending Lemma C.2 and hence Theorem 4.2 to  $\mathbb{L}_q$ -is challenging as one has to control the moment

$$\mathbb{E} \left[ \left| \frac{\gamma_{p+k}^N(1)}{\gamma_p^N(1)} \right|^q \right]^{1/q}.$$

This is particularly simple when  $q = 1$ , but even the case  $q = 2$  is not trivial; see [10] for instance.

## References

- [1] ARONSON, D.G. (1967). Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, **73**, 890–896.
- [2] ASSARAF, R., CAFFAREL, M., & KHELIF, A. (2000). Diffusion Monte Carlo methods with a fixed number of walkers. *Phys. Rev. E*, **61**, 4566.
- [3] BARONI, S., & MORONI, S. (1999). Reptation Quantum Monte Carlo: A Method for Unbiased Ground-State Averages and Imaginary-Time Correlations. *Phys. Rev. Lett.*, **82**, 4745.
- [4] BESKOS, A., CRISAN, D. & JASRA, A. (2014). On the stability of sequential Monte Carlo samplers. *Ann. Appl. Probab.*, **24**, 1396–1445.
- [5] BESKOS, A., JASRA, A., LAW, K., TEMPONE, R. & ZHOU, Y. (2017). Multilevel sequential Monte Carlo samplers. *Stoch. Proc. Appl.*, **127**, 1417–1440.
- [6] BINDER, K. & CEPERLEY, D.M. (1986). *Monte Carlo Methods in Statistical Physics*. Springer-Verlag Berlin.
- [7] CAFFAREL, M., & CLAVERIE, P. (1988). Development of a pure diffusion quantum Monte Carlo method using a full generalized Feynman-Kac formula. *J. Chem. Phys.*, **88**, 1088.
- [8] CAFFAREL, M., & CLAVERIE, P. (1986). Treatment of the Schrödinger equation through a Monte Carlo method based upon the generalized Feynman-Kac formula. *J. Stat. Phys.*, **43**, 797–801.
- [9] CAFFAREL, M., DEL MORAL, P., & DE MONTELLA, L. (2024). On the Mathematical foundations of Diffusion Monte Carlo. ArXiv preprint arXiv:2402.04642.

- [10] CEROU, F., DEL MORAL, P., & GUYADER, A. (2011). A non-asymptotic theorem for unnormalized Feynman-Kac particle models. *Ann. Inst. Henri Poincaré, Prob. Stat.*, **47**, 629–649.
- [11] DEL MORAL, P., (2004). *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. Springer: New York.
- [12] DEL MORAL, P., (2013). *Mean Field Simulation for Monte Carlo Integration*. Chapman & Hall: New York.
- [13] DEL MORAL, P. & DOUCET, A. (2004). Particle motions in absorbing medium with hard and soft obstacles. *Stoch. Anal. Appl.*, **22**, 1175–1207.
- [14] DEL MORAL, P. & GUIONNET, A. (2001). On the stability of interacting processes with applications to filtering and genetic algorithms. *Ann. Inst. Henri Poincaré, Prob. Stat.*, **37**, 155–194.
- [15] DEL MORAL, P., & MICLO, L. (2000). Branching and interacting particle systems. Approximations of Feynman-Kac formulae with applications to non-linear filtering. *Séminaire de probabilités de Strasbourg* **34**, 1-145. Springer: Lecture Notes in Mathematics.
- [16] DEL MORAL, P., & MICLO, L. (2003). Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman-Kac semigroups. *ESAIM: Probability and Statistics* **7**: 171–208.
- [17] DEL MORAL, P., DOUCET, A., & JASRA, A. (2006). Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B*, **68**, 411–436.
- [18] DEL MORAL, P., HORTON, E. & JASRA, A. (2023). On the stability of positive semigroups. *Ann. Appl. Probab.*, **33**, 4424-4490.
- [19] DOUC, R., MOULINES, E. & OLSSON, J. (2014). Long-term stability of sequential Monte Carlo methods under verifiable conditions. *Ann. Appl. Probab.*, **24**, 1767–1802.
- [20] GLYNN, P. & MEYN, S. (1996). A Lyapunov bound for solutions to the Poisson equation. *Ann. Probab.*, **24**, 916-931.
- [21] HALL, P. & HEYDE, C. C. (1980). *Martingale Limit Theory and its Application*. Academic Press: New York.
- [22] HETHERINGTON, J. H. (1984). Observations on the statistical iteration of matrices. *Phys. Rev. A*, **30**, 2713.
- [23] HOBERT, J. & ROBERT, C. (2004). A mixture representation of  $\pi$  with applications in Markov chain Monte Carlo and perfect sampling. *Ann. Appl. Probab.*, **14**, 1295–1305.
- [24] HUYÈN, P. (2009). *Continuous-time stochastic control and optimisation with financial applications*. Springer-Verlag.
- [25] KALOS, M.H. (1962). Monte Carlo calculations of the ground state of three-and four-body nuclei. *Physical Review*, **128**, 1791.
- [26] NASH, J. (1958). Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, **80**, 931–954.
- [27] REYNOLDS, P.J. & CEPERLEY, D.M. & ALDER, B.J. & LESTER, W.A. Jr. (1982). Fixed-node quantum Monte Carlo for molecules *J. Chem. Phys.*, **77**, 5593.
- [28] ROUSSET, M. (2006). On the control of an interacting particle estimation of Schrodinger ground states. *SIAM J. Math. Anal.*, **38**, 824–844.
- [29] SIMON, B. (2005). *Functional Integration and quantum physics*. AMS Chelsea Publishing.

- [30] SORELLA, S. (1998). Green function Monte Carlo with stochastic reconfiguration. *Phys. Rev. Lett.*, **80**, 4558.
- [31] SORELLA, S., & CAPRIOTTI, L. (2000). Green function Monte Carlo with stochastic reconfiguration: An effective remedy for the sign problem. *Phys. Rev. B*, **61**, 2599.
- [32] VAROPOULOS, N.T. (1990). Small time gaussian estimates of heat diffusion kernels. II. The theory of large deviations. *J. Func. Anal.*, **93**, 1-33.
- [33] WHITELEY, N. (2013). Stability properties of some particle filters. *Ann. Appl. Probab.*, **23**, 2500–2537.
- [34] WHITELEY, N. & KANTAS, N. (2017). Calculating principal eigen-functions of non-negative integral kernels: Particle Approximations and Applications. *Math. Oper. Res.* **42**, 1007–1034.
- [35] WHITELEY, N. & LEE, A. (2014). Twisting particle filters. *Ann. Stat.*, **42**, 115–141.
- [36] WHITELEY, N. KANTAS, N. & JASRA, A. (2012). Linear variance bounds for particle approximations of time-homogeneous Feynman-Kac formulae. *Stoch. Proc. Appl.* **122**, 1840–1865.