Linear Variance Bounds for Particle Approximations of Time-Homogeneous Feynman-Kac Formulae

NICK WHITELEY, NIKOLAS KANTAS & AJAY JASRA

Department of Mathematics, University of Bristol, Bristol, BS8 1TW, UK.

Department of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ, UK.

Department of Mathematics, Imperial College London, London, SW7 2AZ, UK.

August 22, 2011

Abstract

This article establishes sufficient conditions for a linear-in-time bound on the non-asymptotic variance of particle approximations of time-homogeneous Feynman-Kac formulae. These formulae appear in a wide variety of applications including option pricing in finance and risk sensitive control in engineering. In direct Monte Carlo approximation of these formulae, the non-asymptotic variance typically increases at an exponential rate in the time parameter. It is shown that a linear bound holds when a non-negative kernel, defined by the logarithmic potential function and Markov kernel which specify the Feynman-Kac model, satisfies a type of multiplicative drift condition and other regularity assumptions. Examples illustrate that these conditions are general and flexible enough to accommodate two rather extreme cases, which can occur in the context of a non-compact state space: 1) when the potential function is bounded above, not bounded below and the Markov kernel is not ergodic; and 2) when the potential function is not bounded above, but the Markov kernel itself satisfies a multiplicative drift condition.

Keywords: Feynman-Kac Formulae; Non-Asymptotic Variance; Multiplicative Drift Condition.

1 Introduction

On a state space X endowed with a σ-algebra $\mathcal{B}(\mathsf{X})$ let M be a Markov kernel and let $U : \mathsf{X} \to \mathbb{R}$ be logarithmic potential function. Then for $x \in \mathsf{X}$, consider the sequence of measures $\{\gamma_{n,x}; n \geq 1\}$ defined by

$$\gamma_{n,x}(\varphi) \quad := \quad \mathbb{E}_x \left[\exp\left(\sum_{k=0}^{n-1} U(X_k) \right) \varphi(X_n) \right], \tag{1.1}$$

for a suitable test function φ and where $\mathbb{E}_x$ denotes expectation with respect to the law of a Markov chain $\{X_n; n \geq 0\}$ with transition kernel M, initialised from $X_0 = x$.

Feynman-Kac formulae as in (1.1) arise in a variety of application domains. In the case that U is non-positive, the quantity $\gamma_{n,x}(1)$ can be interpreted as the probability of survival up to time step n of a

Markovian particle exploring an absorbing medium [Del Moral and Miclo, 2003, Del Moral and Doucet, 2004]; the particle evolves according to M and at time step k it is killed with probability $\exp(U(X_k))$. Another application is the calculation of expectations at a terminal time with respect to jump-diffusion processes which may or may not be partially observed (e.g. Jasra and Doucet [2009]). In particular, for option pricing in finance, there are a variety of options, (e.g. asian, barrier) which can be written in the form (1.1) where the potential function arises from the pay-off function/change of measure and the Markov kernel specifies finite dimensional marginals of some partially observed Lévy process (e.g. Jasra and Del Moral [2011]). It is remarked that in this latter example, the finite dimensional marginals can induce a time-homogeneous Markov chain that is not necessarily ergodic. Furthermore, functionals as in (1.1) arise in certain stochastic control problems, where one considers the bivariate process $\{X_n = (Y_n, A_n); n \geq 0\}$ with Y_n being a controlled Markov chain and $\{A_n; n \geq 0\}$ a control input process. In some cases the transition kernel M can be expressed as $M_1(y_n, da_n)M_2(y_n, a_n, dy_{n+1})$ with M_1 corresponding to the control law or policy and M_2 to the controlled process dynamics. In a risk-sensitive optimal control framework $\frac{1}{n}\log\gamma_{n,x}(1)$ arises as a cost function one aims to minimise with respect to an appropriate class of policies; see [Whittle, 1990, Di Masi and Stettner, 1999] for details. In such problems it is common to choose $U(y, a)$ to be unbounded above, e.g. U is usually chosen to be a quadratic for linear and Gaussian state space models [Whittle, 1990]. More generally (1.1) arises as a special case of a time-inhomogeneous Feynman-Kac formulae studied by Del Moral [2004].

The non-negative kernel $Q(x, dy) := \exp(U(x))M(x, dy)$, defines a linear operator on functions $Q(\varphi)(x) := \int Q(x, dy)\varphi(y)$ and (1.1) can be rewritten as $\gamma_{n,x}(\varphi) = Q_n(\varphi)(x)$, where Q_n denotes the n-fold convolution of Q. In the applications described above, Feynman-Kac formulae as in (1.1) typically cannot be evaluated analytically. However, they may be approximated using a system of interacting particles [Del Moral, 2004]. These particle systems, also known as sequential Monte Carlo methods in the computational statistics literature (e.g. Doucet et al. [2001]), have themselves become an object of intensive study, see amongst others [Crisan and Bain, 2008, Del Moral et al., 2009, van Handel, 2009, Chopin et al., 2011, Del Moral et al., 2011] and references therein for recent developments in a variety of settings.

The present work is concerned with second moment properties of errors associated with the particle approximations of $\{\gamma_{n,x}\}$. In order to obtain bounds on the relative variance, we control certain tensor-product functionals of these particle approximations, recently addressed by Cérou et al. [2011], using stability properties of the operators $\{Q_n; n \geq 1\}$. These stability properties are themselves derived from the multiplicative ergodic and spectral theories of linear operators on weighted ∞-norm spaces due to Kontoyiannis and Meyn [2003, 2005]; this is one of the main novelties of the paper. By doing so we obtain a linear-in-n relative variance bound under assumptions on Q which are weaker than those relied upon in the literature to date and which readily hold on non-compact spaces. Furthermore, to the knowledge of the authors, these are the first results which establish

- that a linear-in-n bound holds under conditions which can accommodate Q defined in terms of a non-ergodic Markov kernel M,
- that any form of non-asymptotic stability result for particle approximations of Feynman Kac formulae holds under conditions which can accommodate U not bounded above.

1.1 Interacting Particle Systems

Let $N \in \mathbb{N}$ be a population size parameter. For $n \in \mathbb{N}$, let $\zeta_n^{(N)} := \left\{\zeta_n^{(N,i)}; 1 \leq i \leq N\right\}$ be the n-th generation of the particle system, where each particle, $\zeta_n^{(N,i)}$, is a random variable valued in X. Denote $\eta_n^N := \frac{1}{N}\sum_{i=1}^N \delta_{\zeta_n^{(N,i)}}$. The generations of the particle system $\left\{\zeta_n^{(N)}; n \geq 0\right\}$ form a X^N-valued Markov chain: for $x \in \mathsf{X}$, the law of this chain is denoted by $\mathbb{P}_x^N$ and has transitions given in integral form by:

$$\begin{aligned}
\mathbb{P}_x^N\left(\zeta_0^{(N)} \in dy\right) &= \prod_{i=1}^N \delta_x\left(dy^i\right), \\
\mathbb{P}_x^N\left(\zeta_n^{(N)} \in dy \middle| \zeta_{n-1}^{(N)}\right) &= \prod_{i=1}^N \left(\frac{\eta_{n-1}^N Q(dy^i)}{\eta_{n-1}^N Q(1)}\right), \quad 1 \leq k \leq n,
\end{aligned} \tag{1.2}$$

where $dy = d\left(y^1, \ldots y^N\right)$, 1 is the unit function and for some test function φ, $\eta_n^N(\varphi) := \frac{1}{N}\sum_{i=1}^N \varphi\left(\zeta_n^{(N,i)}\right)$ (here the dependence of η_n^N on x is suppressed from the notation). These transition probabilities correspond to a simple selection-mutation operation: at each time step N particles are selected with replacement from the population, on the basis of "fitness" defined in terms of e^U, followed by them each being mutating in a conditionally-independent manner according to M.

The empirical measures $\left\{\gamma_{n,x}^N; n \geq 0\right\}$, defined by

$$\gamma_{n,x}^N(\varphi) := \prod_{k=0}^{n-1} \eta_k^N\left(e^U\right) \eta_n^N(\varphi), \quad n \geq 1,$$

and $\gamma_{0,x}^N := \delta_x$, are taken as approximations of $\{\gamma_{n,x}\}$. It is well known that $\mathbb{E}_x^N\left[\gamma_{n,x}^N(\varphi)\right] = \gamma_{n,x}(\varphi)$ [Del Moral, 2004, Chapter 9], where $\mathbb{E}_x^N$ denotes expectation with respect to the law of the N-particle system.

1.2 Standard Regularity Assumptions for Stability

Recent work on analysis of tensor product functionals associated with $\left\{\gamma_{n,x}^N; n \geq 0\right\}$, [Del Moral et al., 2009], has lead to important results regarding higher moments of the error associated with these particle approximations; in a possibly time-inhomogeneous context Cérou et al. [2011] have proved a remarkable linear-in-n bound on the relative variance of $\gamma_{n,x}^N(1)$. In the context of time-homogeneous Feynman-Kac models, the assumptions of Cérou et al. [2011] are that

$$\sup_{x \in \mathsf{X}} U(x) < \infty \tag{1.3}$$

and that for some $m_0 \geq 1$, there exists a finite constant c such that

$$Q_{m_0}(x, dy) \leq c Q_{m_0}(x', dy), \quad \forall (x, x') \in \mathsf{X}^2. \tag{1.4}$$

The result of Cérou et al. [2011] is then of the form:

$$N > c\,(n+1) \qquad \Longrightarrow \qquad \mathbb{E}_x^N\left[\left(\frac{\gamma_{n,x}^N(1)}{\gamma_{n,x}(1)}-1\right)^2\right] \leq c\frac{4}{N}\,(n+1)\,, \quad \forall x \in \mathsf{X}. \tag{1.5}$$

where c is as in (1.4). The efficiency of the particle approximation is therefore quite remarkable: a natural alternative scheme for estimation of $\gamma_{n,x}(1)$ is to simulate N independent copies of the Markov chain with transition M and approximate the expectation in (1.1) by simple averaging, but the relative variance in that case typically explodes *exponentially* in n. The restriction is that (1.4) rarely holds on non-compact spaces. The present work is concerned with proving a result of the same form as (1.5) under assumptions which are more readily verifiable when X is non-compact. The main result is summarized after the following discussion of (1.3)-(1.4) and how they relate to the assumptions we consider.

The condition of (1.4) and its variants are very common in the filtering literature on exponential stability of the optimal filter and its particle approximations, see for example [Del Moral and Guionnet, 2001, Le Gland and Oudjane, 2004] and references therein. It can be interpreted as implying a uniform bound on the relative oscillations of the total mass of Q_{m_0}, i.e.,

$$\frac{Q_{m_0}(1)(x)}{Q_{m_0}(1)(x')} \leq c, \quad \forall\,(x,x') \in \mathsf{X}^2, \tag{1.6}$$

and this is very useful when controlling various functionals which arise when analysing the relative variance as in (1.5), (see Cérou et al. 2011, Proof of Theorem 5.1). However one may take the interpretation of (1.4) in another direction: it implies immediately that there exist finite measures, say β and ν, and $\epsilon > 0$ such that

$$Q_{m_0}(x,dy) \leq \beta(dy)\,, \quad Q_{m_0}(x,dy) \geq \epsilon\nu(dy)\,, \quad \forall x \in \mathsf{X}. \tag{1.7}$$

In the case that $U = 0$ (i.e $Q = M$ is a probabilistic kernel) and M is ψ-irreducible and aperiodic, this type of minorization over the entire state space X implies uniform ergodicity of Q, which is in turn equivalent to Q satisfying a Foster-Lyapunov drift condition with a bounded drift function [Meyn and Tweedie, 2009, Theorem 16.2.2]. In the scenario of present interest, where in general $U \neq 0$, one may take $V : \mathsf{X} \to [1,\infty)$ to be defined by $V(x) = 1$, for all x, and then when (1.3) holds, it is trivially true that there exists $\delta \in (0,1)$ and $b < \infty$ such that Q satisfies the *multiplicative* drift condition,

$$Q\left(e^V\right) \quad \leq \quad e^{V(1-\delta)+b\mathbb{I}_{\mathsf{X}}}, \tag{1.8}$$

where $\mathbb{I}_{\mathsf{X}}$ is the indicator function on X. Q may then also be viewed as a bounded linear operator on the space of real-valued and bounded functions on X endowed with the ∞-norm, which is norm-equivalent [in the sense of Meyn and Tweedie, 2009, p.393] to $\|\varphi\|_{e^V} := \sup_{x\in\mathsf{X}} \dfrac{|\varphi(x)|}{\exp V(x)}$, with V any bounded weighting function.

As explained in the next section, the interest in writing (1.7)-(1.8) is that conditions expressed

in this manner have natural generalisations in the context of weighted ∞-norm function spaces with possibly unbounded V.

1.3 Setting and Main Result

Del Moral [2004, (e.g. Chapter 4 and Section 12.4)] and Del Moral and Doucet [2004] address the setting in which $\{Q_n; n \geq 1\}$ is considered as a semigroup of bounded linear operators on the Banach space of real-valued and bounded functions on X, endowed with the ∞-norm, and Del Moral and Miclo [2003] address the L_2 setting, connecting stability properties of the measures$\{\gamma_{n,x}\}$ and their normalized counterparts to the spectral theory of bounded linear operators on Banach spaces.

Kontoyiannis and Meyn [2003, 2005] have developed multiplicative ergodic and spectral theories of operators of the form Q in the setting of *weighted* ∞-norm spaces; a function space setting which has already proved to be very fruitful for the study of general state-space Markov chains [Meyn and Tweedie, 2009, Chapter 16] without reversibility assumptions. The reader is referred to [Kontoyiannis and Meyn, 2003, 2005] for extensive historical perspective on this spectral theory and related topics, including (of particular relevance in the present context) the theory of non-negative operators due to Numellin [2004, Chapter 5]. The work of [Kontoyiannis and Meyn, 2003, 2005, Meyn, 2006] is geared towards large deviation theory for sample path ergodic averages $n^{-1}\sum_{k=0}^{n-1} U(X_k)$ under the transition M and in that context it is natural to state separately assumptions on M and U. By contrast, when studying the particle systems described above, we are not directly concerned with such sample paths, but rather the relationship between the properties of the particle approximations $\left\{\gamma_{n,x}^N\right\}$ and their exact counterparts $\{\gamma_{n,x}\}$. Some of the results of Kontoyiannis and Meyn [2003, 2005] will be applied to this effect, but starting from assumptions expressed directly in terms of Q which reflect the scenario of interest.

The core assumptions in the present work (see section 2.2 for precise statements) are that for some constants $m_0 \geq 1$, $\delta \in (0,1)$ and all $d \geq 1$ large enough,

$$Q_{m_0}(x, dy) \geq \epsilon_d \nu_d(dy), \quad \forall x \in C_d, \tag{1.9}$$

$$Q\left(e^V\right) \leq e^{(1-\delta)+b_d \mathbb{I}_{C_d}}, \tag{1.10}$$

with V unbounded and $C_d := \{x : V(x) \leq d\} \subset \mathsf{X}$ a sublevel set. It is noted that one recovers the minorization and drift of (1.7)-(1.8) in the case that V is bounded and $C_d = \mathsf{X}$. It will be illustrated through examples in section 4, that (1.9)-(1.10) can be satisfied in circumstances which allow M to be non-ergodic. Furthermore, it will also be demonstrated that, in contrast to (1.3), conditions (1.9)-(1.10) can be satisfied with U not bounded above, subject to strong enough assumptions on M and a restriction on the growth rate of the positive part of U.

The main result obtained in the present work (Theorem 3.2 in section 3) is a bound of the form:

$$N > c_1(n+1) \geq \phi(x) \qquad \Longrightarrow \qquad \mathbb{E}_x^N\left[\left(\frac{\gamma_{n,x}^N(1)}{\gamma_{n,x}(1)} - 1\right)^2\right] \leq c_2 \frac{4}{N}(n+1)\frac{v^{2+\epsilon}(x)}{h_0^2(x)},$$

with

$$\phi(x) := c_1\left(\left\lceil \frac{1}{B_1}\log\left[B_0^2 \frac{v(x)}{h_0(x)}\right]\right\rceil + 1\right),$$

and where $v(x) = e^{V(x)}$, and B_0, B_1, c_1, c_2 are constants which are independent of N, n and x. In this display h_0 is the eigenfunction associated with the principal eigenvalue of Q and the constant B_1 is directly related to the size of the spectral gap of Q. Verification of the existence of h_0 along with various other spectral quantities plays a central role in the proofs.

We note that Del Moral and Doucet [2004], Cérou et al. [2011] also consider the case in which $\exp U(x)$ may touch zero and the former are also directly concerned with approximation of the eigenvalue λ corresponding to h_0 via the empirical probability measures $\{\eta_n^N\}$. These issues are beyond the scope of the present article but the study by the authors of these and related issues in a more general time-inhomogeneous setting is underway. It is also remarked that Cérou et al. [2011] consider a more general type of particle system, which involves an accept/reject evolution mechanism. The approach taken here is also applicable in that context, but for simplicity of presentation we only consider the selection-mutation transition in (1.2).

The remainder of the paper is structured as follows. Section 2 is largely expository: it introduces various spectral definitions and the main assumptions of the present work and goes on to show how these assumptions validate the application of multiplicative ergodicity results of Kontoyiannis and Meyn [2005]. It is stressed that much of the content of this section is included in order to make clear the similarities and differences between the setting of interest and the main stated assumptions and results of Kontoyiannis and Meyn [2005]. Section 3 deals with the variance bounds for the particle approximations. Numerical examples are given in section 4. Many of the proofs of the results in Section 2 are in Appendix A. Some proofs and lemmas for the results in Section 3 can be found in Appendix B.

2 Multiplicative Ergodicity

2.1 Notations and Conventions

Let X be a state space and $\mathcal{B}(\mathsf{X})$ be an associated countably generated σ-algebra. We are typically interested in the case $\mathsf{X} = \mathbb{R}^{d_x}$, $d_x \geq 1$, but our results are readily applicable in the context of more general non-compact state-spaces. For a weighting function $v : \mathsf{X} \to [1, \infty)$, and φ a measurable real-valued function on X, define the norm $\|\varphi\|_v := \sup_{x \in \mathsf{X}} |\varphi(x)| / v(x)$ and let $\mathcal{L}_v := \{\varphi : \mathsf{X} \to \mathbb{R}; \|\varphi\|_v < \infty\}$ be the corresponding Banach space. Throughout, when dealing with weighting functions we employ an lower/upper-case convention for exponentiation and write interchangeably $v \equiv e^V$.

For K a kernel on $\mathsf{X} \times \mathcal{B}(\mathsf{X})$, a function φ and a measure μ denote $\mu(\varphi) := \int \varphi(x)\mu(dx)$, $K\varphi(x) := \int K(x, dy)\varphi(y)$ and $\mu K(\cdot) := \int \mu(dx)K(x, \cdot)$. Let $\mathcal{P}$ be the collection of probability measures on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$, and for a given weighting function $v : \mathsf{X} \to [1, \infty)$ let $\mathcal{P}_v$ denote the subset of such measures μ such that $\mu(v) < \infty$. For $0 \leq \ell < m$ the $(m - \ell)$-fold iterate of K is denoted:

$$K_{m-\ell} \quad = \quad \underbrace{K \dots K}_{m-\ell \text{ times}},$$

and when $\ell = 0$ it is suppressed from the notation and we write $K_m \equiv K_{m-0}$. By convention we set $K_0 = Id$.

The induced operator norm of a linear operator K acting $\mathcal{L}_v \to \mathcal{L}_v$ is

$$|||K|||_v := \sup \left\{ \frac{\|K\varphi\|_v}{\|\varphi\|_v}; \varphi \in \mathcal{L}_v, \|\varphi\|_v \neq 0 \right\} = \sup \left\{ \|K\varphi\|_v ; \varphi \in \mathcal{L}_v, |\varphi| \leq v \right\}.$$

The spectrum of K as an operator on $\mathcal{L}_v$, denoted by $\mathcal{S}_v(K)$, is the set of complex z such that $[Iz - K]^{-1}$ does not exist as a bounded linear operator on $\mathcal{L}_v$. The corresponding spectral radius of K, denoted by $\xi_v(K)$, is given by

$$\xi_v(K) := \sup \{|z| ; z \in \mathcal{S}_v(K)\} = \lim_{n\to\infty} |||K_n|||_v^{1/n},$$

where the limit always exists by subadditive arguments, but may be infinite. The following definitions are from Kontoyiannis and Meyn [2005].

- A pole $z_0 \in \mathcal{S}_v(K)$ is *of finite multiplicity* n if
  - for some $\epsilon_1 > 0$ we have $\{z \in \mathcal{S}_v(K); |z - z_0| \leq \epsilon_1\} = \{z_0\}$;
  - and the associated projection operator

$$J := \frac{1}{2\pi i} \int_{\partial\{z:|z-z_0|\leq\epsilon_1\}} [Iz - K]^{-1} \, \mathrm{d}z,$$

    can be expressed as a finite linear combination of some $\{s_i\} \subset \mathcal{L}_v$ and $\{\nu_i\} \subset \mathcal{P}_v$,

$$J = \sum_{i,j=0}^{n-1} m_{i,j} \left[s_i \otimes \nu_j \right],$$

    where $[s_i \otimes \nu_j](x, dy) = s_i(x)\nu_j(dy)$.

- K admits a *spectral gap* in $\mathcal{L}_v$ if there exists $\epsilon_0 > 0$ such that $\mathcal{S}_v(Q) \cap \{z : |z| \geq \xi_v(Q) - \epsilon_0\}$ is finite and contains only poles of finite multiplicity.
- K is *v-uniform* if it admits a spectral gap and there exists a unique pole $\lambda \in \mathcal{S}_v(Q)$ of multiplicity 1, satisfying $|\lambda| = \xi_v(Q)$.
- K has a *discrete spectrum* if for any compact set $B \subset \mathbb{C} \setminus \{0\}$, $\mathcal{S}_v(Q) \cap B$ is finite and contains only poles of finite multiplicity.
- K is *v-separable* if for any $\epsilon > 0$ there exists a finite rank operator $\hat{K}_\epsilon$ such that $|||K - \hat{K}_\epsilon|||_v \leq \epsilon$

2.2 Multiplicative Ergodic Theorem

In this section we present the main assumptions and state some results from Kontoyiannis and Meyn [2005] (see also Kontoyiannis and Meyn [2003]).

2.2.1 Assumptions

(**H1**) The semigroup $\{Q_n; n \geq 1\}$ is ψ-irreducible and aperiodic (see Meyn [2006, Section 2.1]).

(**H2**) There exists an unbounded $V : \mathsf{X} \to [1, \infty)$, constants $m_0 \geq 1$, $\delta \in (0, 1)$ and $\underline{d} \geq 1$ with the following properties.

For each $d \geq \underline{d}$, there exists $\epsilon_d > 0$ and $\nu_d \in \mathcal{P}_v$ such that $C_d := \{x : V(x) \leq d\}$ is (m_0, ϵ_d, ν_d)-small for Q, i.e.

$$Q_{m_0}(x, \cdot) \quad \geq \quad \mathbb{I}_{C_d}(x)\epsilon_d \nu_d(\cdot), \quad \forall x \in \mathsf{X}. \tag{2.1}$$

with $\nu_d(C_d) > 0$. Furthermore $Q_{m_0}(C_d)(x) > 0$ for all $x \in \mathsf{X}$. There also exists $b_d < \infty$ such that the following multiplicative drift condition holds:

$$Q\left(e^V\right) \quad \leq \quad e^{V(1-\delta)+b_d \mathbb{I}_{C_d}}. \tag{2.2}$$

(**H3**) $U : \mathsf{X} \to \mathbb{R}$ is such that

$$U^+ \quad := \quad \max(U, 0) \in \mathcal{L}_V.$$

(**H4**) There exists $t_0 \geq 1$ and for each $d \geq \underline{d}$ there exists a measure β_d, such that $\beta_d\left(e^V\right) < \infty$ and

$$\mathbb{P}_x\left(X_{t_0} \in A, \tau_{C_d^c} > t_0\right) \quad \leq \quad \beta_d(A), \quad x \in C_d, A \in \mathcal{B}(\mathsf{X}),$$

where $\mathbb{P}_x$ denotes the law of the Markov chain $\{X_n\}$ with transition M and $\tau_A := \inf\{n \geq 1 : X_n \in A\}$.

Remark 2.1. We take care to emphasize the following differences and similarities between the above assumptions and the setting of Kontoyiannis and Meyn [2005].

- Assumption (H2) equation (2.2) applies directly to the Q kernel, whereas Kontoyiannis and Meyn [2005] impose a multiplicative drift condition on M. The key issue is that the multiplicative drift condition for Q is the essential and implicit ingredient of Lemma B.4 of Kontoyiannis and Meyn [2005], and as we shall see in section 4, under the conditions that U is bounded above but not bounded below, assumption (H2) can hold without geometric drift assumptions on M. A related phenomenon is considered by Meyn [2006] in order to obtain "one-sided" large deviation principles for ergodic sample-path averages for the chain with transition M.

- Assumption (H2) requires the sublevel sets of V to be small for Q and this is exploited in Lemma A.1. The explicit m_0-step minorisation condition makes it easy to bound below the spectral radius of Q, see Lemma 2.1. In the setting of Kontoyiannis and Meyn [2003] the spectral radius of Q is bounded below by 1 as U is assumed centered with respect to the invariant probability distribution for M. In the present context, this centering assumption is unnatural, especially as we want to consider some situations where such an invariant probability does not exist.

- Assumption (H3) is weaker than the corresponding assumption in the statement of [Kontoyiannis and Meyn, 2005, Theorem 3.1]. However, (H3) coincides with the first part of [Kontoyiannis and Meyn, 2005, Equation 73], which combined with (H1), (H2) and (H4) in Lemma 2.2 below, is enough to prove that Q has a discrete spectrum in $\mathcal{L}_v$.

- As shown in [Kontoyiannis and Meyn, 2005, Theorem 3.4] and [Kontoyiannis and Meyn, 2003], a MET can be proved without (H4), but at the cost of restrictions on the class of functions to which U belongs which are a little unwieldy.

2.2.2 Results

We now give a collection of results which are used to prove the MET, Theorem 2.2. The proofs are given in Appendix A. It is remarked that the steps in the proof Theorem 2.2 are effectively the same as part of the proof of Theorem 3.1 of Kontoyiannis and Meyn [2005], our starting assumptions are stated differently.

The following preparatory lemma establishes that the Feynman-Kac formula (1.1) is well defined and presents bounds on the spectral radius of Q.

Lemma 2.1. *Assume (H2). Then for all $x \in \mathsf{X}$, $n \geq 1$, $\varphi \in \mathcal{L}_v$:*

$$0 < |\gamma_{n,x}(\varphi)| < \infty. \tag{2.3}$$

and for all $d \geq \underline{d}$:

$$\epsilon_d \nu_d(C_d) \leq \xi_v(Q) < \infty, \tag{2.4}$$

where $\underline{d}$ is as in (H2).

To clarify how assumptions (H1)-(H4) connect with the results of Kontoyiannis and Meyn [2005] we next present a lemma regarding the v-separability of Q which is a stepping stone to the MET. Observe that the multiplicative drift condition (H2) implies that Q can be approximated in norm to arbitrary precision by truncation to the sublevel sets of V, in the sense that for fixed d and any $r \geq 1$,

$$\mathbb{I}_{C_r^c} Q\left(e^V\right) \quad \leq \quad e^{V-\delta r+b_d}, \tag{2.5}$$

and then with $\hat{Q} := \mathbb{I}_{C_r} Q$, it follows immediately that $|||Q - \hat{Q}|||_v \leq e^{-\delta r+b_d}$. In the following lemma, which combines [Kontoyiannis and Meyn, 2005, Lemmata B.3-B.5] and is included here for completeness, the density assumption (H4) plays a key role in establishing that iterates of this truncation of Q can be approximated by a finite rank kernel.

Lemma 2.2. *Assume (H1)-(H4). Then Q_{2t_0+2} is v-separable, where t_0 is as in (H4).*

The following theorem makes a key connection between v-separability and a discrete spectrum.

Theorem 2.1. *[Kontoyiannis and Meyn, 2005, Theorem 3.5] If the linear operator $Q : \mathcal{L}_v \to \mathcal{L}_v$ is bounded and $Q_{t_0} : \mathcal{L}_v \to \mathcal{L}_v$ is v-separable for some $t_0 \geq 1$, then Q has a discrete spectrum in $\mathcal{L}_v$.*

Under (H2) Q is indeed bounded, so has a discrete spectrum in $\mathcal{L}_v$ and then by definition it also admits a spectral gap in $\mathcal{L}_v$. Fixing arbitrarily $\theta > \xi_v(Q)$ and $d \geq \underline{d}$, we may consider the resolvent

and corresponding potential operators, defined respectively by

$$R := [I\theta - Q]^{-1} = \sum_{k=0}^{\infty} \theta^{-k-1} Q_k, \tag{2.6}$$

$$H := \left[I\lambda_R - \left(R - \theta^{(-m_0-1)} \epsilon_d \mathbb{I}_{C_d} \otimes \nu_d\right)\right]^{-1} = \sum_{k=0}^{\infty} \lambda_R^{-k-1} \left(R - \theta^{(-m_0-1)} \epsilon_d \mathbb{I}_{C_d} \otimes \nu_d\right)^k, \tag{2.7}$$

where $\lambda_R := 1/(\theta - \xi_v(Q))$ and ϵ_d, C_d, ν_d, are as in (H2). The dependence of R and H on θ and d is suppressed from the notation. Also then consider the function h on X defined (with the same d as in H) by

$$h(x) := H(\mathbb{I}_{C_d})(x), \quad \forall x \in \mathsf{X}. \tag{2.8}$$

We can now state and prove the MET:

Theorem 2.2. *Assume (H1)-(H4). Then there exists a maximal and isolated eigenvalue λ, with $\lambda = \xi_v(Q)$, an eigenfunction $h_0 \in \mathcal{L}_v$ and an eigenmeasure $\mu_0 \in \mathcal{P}_v$ satisfying*

$$Qh_0 = \lambda h_0, \qquad \mu_0 Q = \lambda \mu_0,$$

and $\mu_0(h_0) = 1$. The eigenfunction is a normalisation of h given in (2.8), in the sense that for the same d taken in the definition of h, we have

$$h_0 := \frac{H(\mathbb{I}_{C_d})}{\mu_0 H(\mathbb{I}_{C_d})} \quad \text{with} \quad \mu_0 := \frac{\nu_d H}{\nu_d H(1)}. \tag{2.9}$$

Furthermore, there exist constants $B_0 < \infty$ and $B_1 > 0$ such for any $\varphi \in \mathcal{L}_v$, any $n \geq 1$ and any $x \in \mathsf{X}$,

$$\left|\lambda^{-n} \gamma_{n,x}(\varphi) - h_0(x)\mu_0(\varphi)\right| \leq \|\varphi\|_v B_0 e^{-nB_1} v(x). \tag{2.10}$$

Proof. We give only a sketch proof, as it is essentially that of Theorem 3.1 of Kontoyiannis and Meyn [2005]. As established in Lemma 2.1, under our assumptions $\xi_v(Q) < \infty$. In addition, as mentioned above Q is bounded on $\mathcal{L}_v$ has a discrete spectrum and admits a spectral gap in $\mathcal{L}_v$. By the minorization condition of (H2) one can obtain a minorization condition for R of (2.6):

$$R(x, dy) \geq \theta^{(-m_0-1)} \epsilon_d \mathbb{I}_{C_d}(x) \nu_d(dy)$$

and hence one can apply Proposition 2.8 of Kontoyiannis and Meyn [2005]. This proposition establishes that Q is v-uniform and that $h \in \mathcal{L}_v$ is an eigenfunction for Q with eigenvalue $\lambda = \xi_v(Q)$ (i.e. it is maximal and isolated, by the $v-$uniformity). The normalization is justified by the finiteness, under our assumptions, of the associated quantities. The eigenmeasure is easily obtained by similar arguments to the proof of Proposition 4.5 of Kontoyiannis and Meyn [2003].

To obtain (2.10) one may define the *twisted* kernel:

$$\check{P}(x, dy) \quad := \quad \lambda^{-1} h_0^{-1}(x) Q(x, dy) h_0(y), \tag{2.11}$$

which can be seen to be well defined as a Markov kernel, as λ is strictly positive and finite and (H2) implies h_0 is everywhere finite and strictly positive. Furthermore one observes immediately that $\check{P}$ admits $\check{\pi}$, defined by $\check{\pi}(\varphi) = \mu_0(h_0\varphi)/\mu_0(h_0) = \mu_0(h_0\varphi)$, as an invariant probability distribution. By Lemma A.1 in Appendix A one can apply Theorem 3.4 of Kontoyiannis and Meyn [2005] to the Markov chain associated to the twisted kernel, (in the notation of of Theorem 3.4 of Kontoyiannis and Meyn [2005], take $g \equiv \varphi/h_0$, $F \equiv 0$). This results in the bound (2.10), which completes the proof. □

Remark 2.2. Upon dividing through by h_0, the equation (2.10) of the MET may be viewed as a probabilistic, geometric ergodic theorem for the twisted chain associated to the kernel (2.11) and the modified test function φ/h_0, with a naturally modified drift function $\check{v} = e^{\check{V}}$ proportional to v/h_0. See Lemma A.1 in Appendix A.

Remark 2.3. The constant B_1 in equation (2.10) is directly related to the size of the spectral gap of Q, see [Kontoyiannis and Meyn, 2003, Proof of Theorem 4.1].

3 Non-Asymptotic Variance

3.1 Tensor Product Functionals

The various tensor product functionals considered in the remainder of this paper require some additional notation. For a measurable function F on X^2 and a weighting function $v : \mathsf{X} \to [1, \infty)$, we define the norm $\|F\|_{v,2} := \sup_{x,y \in \mathsf{X}^2} |F(x,y)| / (v(x)v(y))$ and denote $\mathcal{L}_{v,2} := \left\{ F : \mathsf{X}^2 \to \mathbb{R}; \|F\|_{v,2} < \infty \right\}$ the corresponding function space. For two functions $\varphi_1, \varphi_2 \in \mathcal{L}_v$, we denote by $\varphi_1 \otimes \varphi_2 \in \mathcal{L}_{v,2}$ the tensor product function defined by $\varphi_1 \otimes \varphi_2(x, x') := \varphi_1(x)\varphi_2(x')$. Let $K : \mathsf{X} \times \mathcal{B}(\mathsf{X}) \to \mathbb{R}_+$ be a kernel on X. The two-fold tensor product operator corresponding to K is defined, for any $F \in \mathcal{L}_{v,2}$, by

$$K^{\otimes 2}(F)(x, x') \quad := \quad \int_{\mathsf{X}^2} K(x, dy) K(x', dy') F(y, y').$$

The iterated operator notation of the previous section is carried over so that for indices $0 \leq \ell < m$,

$$K_{\ell,m}^{\otimes 2}(F)(x, x') \quad = \quad \underbrace{K^{\otimes 2} \dots K^{\otimes 2}}_{m-\ell \text{ times}}(F)(x, x').$$

Corresponding to the particle empirical measures of section 1.1, for $n \geq 1$, we introduce the tensor product empirical measures (or 2-fold $V-$statistic):

$$\left(\eta_n^N\right)^{\otimes 2} := \frac{1}{N^2} \sum_{1 \leq i,j \leq N} \delta_{(\zeta_n^i, \zeta_n^j)}, \quad \left(\gamma_{n,x}^N\right)^{\otimes 2} := \gamma_{n,x}^N(1)^2 \left(\eta_n^N\right)^{\otimes 2}.$$

Following the definition of Cérou et al. [2011], the coalescent integral operator D, acting on func-

tions on X^2, is defined by

$$D(F)(x,x') = F(x,x), \quad (x,x') \in \mathsf{X}^2.$$

For any $0 \leq s \leq (n+1)$, we denote by $\mathcal{I}_{n,s} := \{(i_1,...,i_s) \in \mathbb{N}_0^s; 0 \leq i_1 < \ldots < i_s \leq n\}$ the set of coalescent time configurations over a horizon of length $n+1$ and for $(i_1,...,i_s) \in \mathcal{I}_{n,s}$ and $x \in \mathsf{X}$, the nonegative measure $\Gamma_{n,x}^{(i_1,...i_s)}$ on $\left(\mathsf{X}^2, \mathcal{B}\left(\mathsf{X}^2\right)\right)$, and its normalised counterpart $\bar{\Gamma}_{n,x}^{(i_1,...,i_s)}$, are defined by

$$\Gamma_{n,x}^{(i_1,...,i_s)} := \gamma_{i_1,x}^{\otimes 2} D Q_{i_1,i_2}^{\otimes 2} D \ldots Q_{i_{s-1},i_s}^{\otimes 2} D Q_{i_s,n}^{\otimes 2}, \qquad \bar{\Gamma}_{n,x}^{(i_1,...i_s)} := \frac{\Gamma_{n,x}^{(i_1,...,i_s)}}{\gamma_{n,x}(1)^2}, \tag{3.1}$$

for $s \geq 1$, and for $s = 0$, $\Gamma_{n,x}^{(\emptyset)}(F) := \gamma_{n,x}^{\otimes 2}(F)$ and $\bar{\Gamma}_{n,x}^{(\emptyset)}(F) := \eta_n^{\otimes 2}(F)$. We refer the reader to Cérou et al. [2011, Section 3.] for a helpful visual representation of the integrals in the transport equation (3.1). We have already checked in Lemma 2.1 that the Feynman-Kac formula (1.1) is well defined under our assumptions in the $\mathcal{L}_v$ setting, which validates the denominator of (3.1).

When Theorem 2.2 holds, we will denote by $\check{\mathbb{E}}_x$ expectation with respect to the law of the twisted Markov chain $\left\{\check{X}_k; k \geq 0\right\}$, i.e that with transition kernel $\check{P}$ as in equation (2.11) and initialised from $\check{X}_0 = x$.

3.2 Non-Asymptotic Variance

In this section we give our main result. The proof is detailed in section 3.3. The following additional assumption imposes some further restrictions on the function class considered, but this is not overly demanding, considering that we will be dealing with coalesced tensor product quantities.

(**H5**) Let V and $\bar{d}$ be as in assumption (H2). There exists $0 < \epsilon_0 < \epsilon$ and for all $d \geq \bar{d}$, there exists $b_d^* < \infty$ such that

$$Q\left(e^{(1+\epsilon)V}\right) \leq e^{(1+\epsilon)V-(1+\epsilon_0)V+b_d^* \mathbb{I}_{C_d}}.$$

The following theorem is due to Cérou et al. [2011] .

Theorem 3.1. *[Cérou et al., 2011, Proposition 3.4] For any $n \geq 1$, $x \in \mathsf{X}$ and $N \geq 1$ the following expansion holds:*

$$\begin{aligned}
&\mathbb{E}_x^N \left[\left(\frac{\gamma_{n,x}^N(1)}{\gamma_{n,x}(1)} - 1 \right)^2 \right] \\
&= \sum_{s=1}^{n+1} \left(1 - \frac{1}{N}\right)^{(n+1)-s} \frac{1}{N^s} \sum_{(i_1,\ldots,i_s) \in \mathcal{I}_{n,s}} \left[\bar{\Gamma}_{n,x}^{(i_1,\ldots i_s)}(1 \otimes 1) - 1 \right],
\end{aligned} \tag{3.2}$$

where $\mathbb{E}_x^N$ denotes expectation w.r.t. the law of the N-particle system.

It is remarked that there is a different error decomposition in [Chan and Lai, 2011], which can hold to any order under appropriate regularity conditions; one would conjecture that this decomposition can also be treated, but this is not considered here. The main result of this section is the following theorem, whose proof is postponed.

Theorem 3.2. *Assume (H1)-(H5). Then there exists $c_1 < \infty$ and $c_2 < \infty$ depending only on the quantities in (H1)-(H5) such that for all $x \in \mathsf{X}$,*

$$N > c_1 (n+1) \geq \phi(x) \qquad \Longrightarrow \qquad \mathbb{E}_x^N \left[\left(\frac{\gamma_{n,x}^N(1)}{\gamma_{n,x}(1)} - 1 \right)^2 \right] \leq c_2 \frac{4}{N} (n+1) \frac{v^{2+\epsilon}(x)}{h_0^2(x)},$$

with

$$\phi(x) := c_1 \left(\left\lceil \frac{1}{B_1} \log \left[B_0^2 \frac{v(x)}{h_0(x)} \right] \right\rceil + 1 \right),$$

and where B_0 and B_1 are as in Theorem 2.2.

3.3 Construction of the Proof

In the following Section, we detail the argument to prove Theorem 3.2. To that end, we present the essence of the argument with Proposition 3.1 and Lemma 3.1 below; the proofs of which are in Appendix B.1. In addition, these results will rely upon technical lemmas which can be found in Appendix B.2.

The proof of Theorem 3.2 is constructed in the following manner. By Theorem 3.1) we have the decomposition (3.2) in terms of the operators $\left\{ \bar{\Gamma}_{n,x}^{(i_1,...i_s)} \right\}$. The proof in Cérou et al. [2011] focuses upon controlling these expressions via the regularity conditions mentioned in section 1.2; our proof will do the same, except under (H1)-(H5).

Throughout the remainder of this paper, let $V^* : \mathsf{X} \to [1, \infty)$ is defined by

$$V^*(x) \quad := \quad V(x)(1+\epsilon) - \log h_0(x) + \log \|h_0\|_{v^{(1+\epsilon)}}, \tag{3.3}$$

where ϵ as in (H5). We proceed with the following key proposition:

Proposition 3.1. *Assume (H1)-(H5). Then there exists $c < \infty$ depending only on the quantities in (H1)-(H5) such that for all $n \geq 1$, $0 \leq s \leq n+1$, $(i_1, ...i_s) \in \mathcal{I}_{n,s}$, $F \in \mathcal{L}_{v^{1/2},2}$ and $x \in \mathsf{X}$,*

$$\bar{\Gamma}_{n,x}^{(i_1,...i_s)}(F) \quad \leq \quad \|F\|_{v^{1/2},2} \, c^{s+1} \frac{v(x)}{h_0(x)} \frac{\check{\mathbb{E}}_x \left[\prod_{k \in \{i_1,...,i_{s-1}\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s}\right) \right]}{\check{\mathbb{E}}_x \left[1/h_0\left(\check{X}_n\right) \right]^2}, \tag{3.4}$$

with the conventions that the product in the numerator is unity when $s \leq 1$, and in the case of $s = 0$, $i_s = 0$. In the above displays, v is as in (H2), $h_0 \in \mathcal{L}_v$ is eigenfunction as in Theorem 2.2 and $v^* = e^{V^*}$ is as in (3.3).

This result connects the operators $\left\{ \bar{\Gamma}_{n,x}^{(i_1,...i_s)} \right\}$ with expectations of the Lyapunov functions v and v^* and the eigenfunction, w.r.t. the twisted chain. Given this Proposition, one needs to control the numerator and denominator. The latter can be achieved by the MET of Theorem 2.2 and the former via the following:

Lemma 3.1. *Assume (H1)-(H5). Then there exists $c < \infty$ depending only on the quantities in (H1)-*

(H5) such that for any $n \geq 1$, $1 \leq s \leq n+1$, $(i_1, ..., i_s) \in \mathcal{I}_{n,s}$,

$$\check{\mathbb{E}}_x \left[\prod_{k \in \{i_1,\ldots,i_s\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{n+1}\right) \right] \quad \leq \quad c^{s+1} v^*(x), \quad \forall x \in \mathsf{X}, \tag{3.5}$$

where v^* *is as in* (3.3).

We now proceed with the proof of Theorem 3.2.

Proof. [Proof of Theorem 3.2] By Proposition 3.1 and Lemma 3.1 we have that there exists a finite constant c depending only on the quantities in (H1)-(H5) such that

$$\bar{\Gamma}_{n,x}^{(i_1,\ldots i_s)}(1 \otimes 1) \quad \leq \quad c^{s+1} \frac{v(x)}{h_0(x)} v^*(x) \frac{1}{\check{\mathbb{E}}_x\left[1/h_0\left(\check{X}_n\right)\right]^2}. \tag{3.6}$$

Using the fact that $\check{\mathbb{E}}_x\left[1/h_0\left(\check{X}_n\right)\right] = \gamma_{n,x}(1)/[\lambda^n h_0(x)]$ we appeal to (2.10) of the MET of Theorem 2.2 as follows. Without loss of generality, it can be assumed that $B_0 > 1$. Then for all $x \in \mathsf{X}$

$$n \geq \left\lceil \frac{1}{B_1} \log\left[B_0^2 \frac{v(x)}{h_0(x)} \right] \right\rceil \quad \Rightarrow \quad 1 - B_0 e^{-B_1 n} \frac{v(x)}{h_0(x)} \geq \frac{B_0 - 1}{B_0} \quad \Rightarrow \quad \check{\mathbb{E}}_x\left[1/h_0\left(\check{X}_n\right)\right] \geq \frac{B_0 - 1}{B_0}. \tag{3.7}$$

Throughout the remainder of the proof the left-most inequality in (3.7) is assumed to hold. Then combining (3.7) with (3.6) and recalling the definition of v^* we then have that there exists $c_0 < \infty$ such that

$$\bar{\Gamma}_{n,x}^{(i_1,\ldots i_s)}(1 \otimes 1) \quad \leq \quad c_0 c^{s+1} \frac{v^{2+\epsilon}(x)}{h_0^2(x)}.$$

Proceeding by the essentially the same argument as in [Cérou et al., 2011, Proof of Theorem 5.1], we use the identity:

$$\sum_{s=1}^{n+1} \sum_{(i_1,\ldots,i_s) \in \mathcal{I}_{n,s}} \prod_{j \in \{i_1,\ldots,i_s\}} a_j = \left[\prod_{s=0}^{n} (1 + a_s) \right] - 1,$$

which holds for any $n \geq 1$ and $\{a_s; s \geq 0\}$, to establish via Theorem 3.1 that

$$\begin{aligned}
\mathbb{E}_x^N \left[\left(\frac{\gamma_{n,x}^N(1)}{\gamma_{n,x}(1)} - 1 \right)^2 \right] \quad &\leq \quad c_0 c \frac{v^{2+\epsilon}(x)}{h_0^2(x)} \sum_{s=1}^{n+1} \left(1 - \frac{1}{N}\right)^{(n+1)-s} \frac{1}{N^s} \sum_{(i_1,\ldots,i_s) \in \mathcal{I}_{n,s}} c^s \\
&= \quad c_0 c \frac{v^{2+\epsilon}(x)}{h_0^2(x)} \left(1 - \frac{1}{N}\right)^{n+1} \left[\left(1 + \frac{c}{N-1}\right)^{n+1} - 1 \right] \\
&\leq \quad c_0 c \frac{v^{2+\epsilon}(x)}{h_0^2(x)} \left[\left(1 + \frac{c}{N-1}\right)^{n+1} - 1 \right].
\end{aligned}$$

Then exactly as in [Cérou et al., 2011, Proof of Corollary 5.2],

$$N > 1 + c\,(n+1) \qquad \Rightarrow \qquad \left(1 + \frac{c}{N-1}\right)^{n+1} - 1 \leq \frac{2}{N-1} c\,(n+1) \leq \frac{4}{N} c\,(n+1)\,.$$

This completes the proof. □

4 Examples

This section gives some discussion and examples of circumstances in which the assumptions can be satisfied. In particular we focus on the drift assumption of (H2). It seems natural to consider two general cases: those in which it is not assumed, or it is assumed, that the Markov kernel M itself satisfies a multiplicative drift condtion.

4.1 Cases without a multiplicative drift assumption on M

In this situation, the decay of the potential function plays a key role in establishing the multiplicative drift condition, illustrated as follows.

Lemma 4.1. *Assume that there exists $V : \mathsf{X} \to [1, \infty)$ unbounded such that $|||M|||_v < \infty$ and the sublevels of V are small for M. If for any finite d, $\inf_{x \in C_d} U(x) > -\infty$, and there exists d_1 such that $\sup_{x \in C_{d_1}} U(x) < \infty$ and for some $\delta_1 \in (0,1)$, $\sup_{x \in C_{d_1}^c} U(x)/V(x) \leq -\delta_1$, assumption (H2) is satisfied.*

Proof. We have

$$Q\left(e^V\right)(x) \leq \exp\left(V(x) + U(x) + \log |||M|||_v\right), \quad \forall x \in \mathsf{X}.$$

As V is unbounded, for any $\delta \in (0, \delta_1)$ there exists $\underline{d}$ large enough such that for all $x \in \mathsf{X}$ and $d \geq \underline{d}$,

$$\mathbb{I}_{C_d^c}(x) Q\left(e^V\right)(x) \leq \exp\left(V(x)(1-\delta)\right), \quad \mathbb{I}_{C_d}(x) Q\left(e^V\right)(x) \leq \exp\left(d + \sup_{y \in C_d} U(y) + \log |||M|||_v\right),$$

which is enough to verify the drift part of (A2). The minorization part is direct. □

In the extensive literature on drift for Markov kernels there are several conditions which guarantee the existence of v such that $|||M|||_v < \infty$. For example, any M satisfying the polynomial drift condition of **?** automatically satisfies $|||M|||_v < \infty$ for the same v upto a factor of e. However, ergodicity of M is not necessary, as illustrated in the following simple example.

4.1.1 Gaussian Random Walk

Let $\mathsf{X} := \mathbb{R}$ and U and M be defined by

$$U(x) := -x^2, \quad M(x, dy) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-x)^2}{2}\right) dy,$$

where dy denotes Lebesgue measure. With ψ Lebesgue measure, the ψ-irreducibility and aperiodicity of $\{Q_n; n \geq 1\}$ is immediate. For the drift and minorization conditions of (H2), elementary manipulations show that equation (2.2) holds with $V(x) = x^2/\left(2\left(1+\delta_0\right)\right) + 1$ for suitable $\delta_0 > 0$ and solutions of the minorization condition (2.1) are also easily obtained. Condition (H3) is trivially satisfied because U is

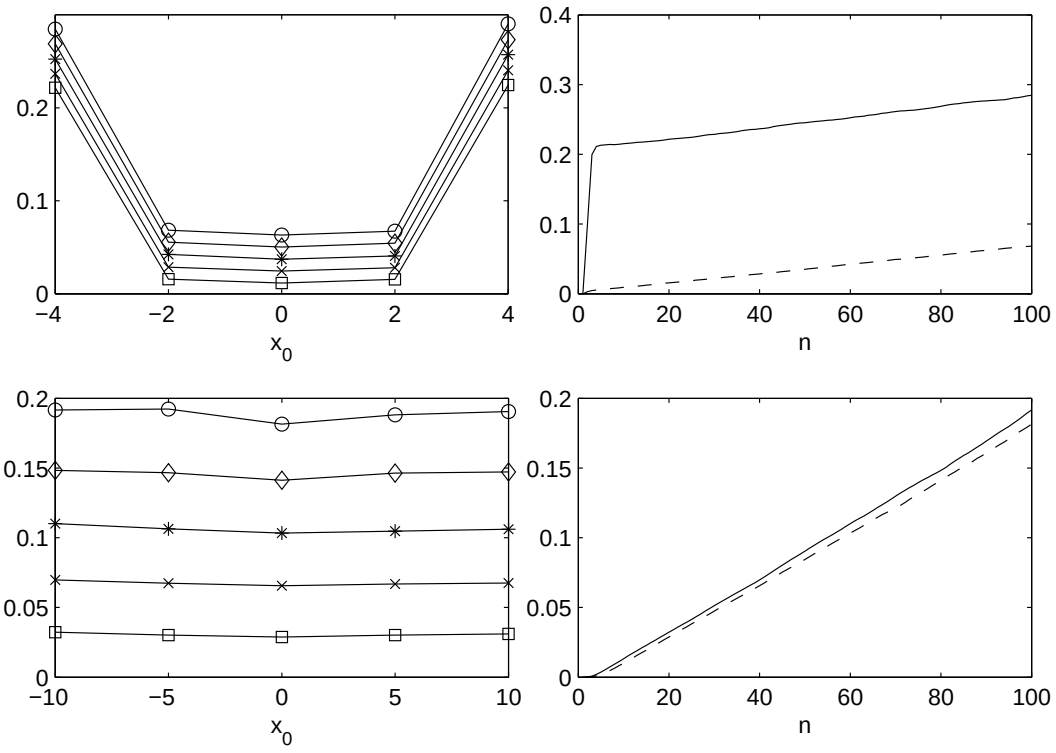

Figure 4.1: Top: Gaussian random walk model. Bottom: ergodic autoregression model. Left: Relative variance vs. initial condition x_0, at times $\square, n = 20$; $\times, n = 40$; $*, n = 60$; $\diamond, n = 80$; $\circ, n = 100$. Right: Relative variance vs. n, from initial conditions (dashed) $x_0 = 0$, (solid - top) $x_0 = 4$, (solid - bottom) $x_0 = 10$.

non-positive. The density assumption (H4) is satisfied with β_d the restriction of Lebesgue measure to C_d. Assumption (H5) holds for ϵ small enough and $\epsilon_0 = \epsilon/2$.

It is generally not easy to obtain or estimate values for the constants in Theorem 3.2. In the numerical examples below, we consider a fixed value of N and consider the relative variance as a function of the n and the initial condition x.

The numerical results of Figure 4.1 show estimates of $\mathbb{E}_x^N\left[\left(\frac{\gamma_{n,x}^N(1)}{\gamma_{n,x}(1)} - 1\right)^2\right]$ with fixed $N = 2000$, for various x and n, with in each case the expectation approximated by averaging over 2×10^4 independent simulations of the particle system. For this model $\gamma_{n,x}(1)$ can be computed analytically, and this exact value was used in the estimates. The linear growth of the relative variance and its dependence on the initial point x is apparent from the figure.

4.2 Cases with a multiplicative drift assumption on M

The following Lemma shows that condition (H2) holds for suitable U when M itself satisfies a multiplicative drift condition.

Lemma 4.2. *Assume that there exists $V : \mathsf{X} \to [1, \infty)$ unbounded, $\delta_1 > 0$, $d_1 \geq 1$ and for each $d \geq d_1$*

there exists $b_d < \infty$ such that

$$M\left(e^V\right) \quad \leq \quad e^{V(1-\delta_1)+b_d \mathbb{I}_{C_d}}, \tag{4.1}$$

and the set $C_d = \{x; V(x) \leq d\}$ is small for M. Then if $U^+ \in \mathcal{L}_V$, $\lim_{r\to\infty} \left\|\mathbb{I}_{C_r^c} U^+\right\|_V = 0$ and for all finite d, $\inf_{x\in C_d} U(x) > -\infty$ assumption (H2) holds.

Proof. Due to the drift condition (4.1), for any $\delta \in (0, \delta_1)$,

$$Q\left(e^V\right) \quad \leq \quad \exp\left(V(1-\delta) - (\delta_1 - \delta)V + U^+ + b_d \mathbb{I}_{C_d}\right),$$

and due to $\lim_{r\to\infty} \left\|\mathbb{I}_{C_r^c} U^+\right\|_V = 0$, there exists $\underline{d}$ such that for all $d \geq \underline{d}$,

$$Q\left(e^V\right) \leq \exp\left(V(1-\delta) + \bar{b}_d \mathbb{I}_{C_d}\right),$$

where $\bar{b}_d := b_d + d\left\|U^+\right\|_V$, which verifies the drift part of (H2). The minorization part is immediate from the assumptions that C_d is small for M and $U(x)$ is bounded below on C_d. □

4.2.1 Ergodic Autoregression

For this example we consider the lag-1 autoregressive Markov kernel,

$$M(x, dy) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-\alpha x)^2}{2}\right) dy,$$

for fixed $|\alpha| < 1$. Elementary manipulations then show that, for $\delta_0 > 0$ and $\underline{d}$ large enough, M satisfies (4.1) with $V(x) = x^2/\left(2\left(1+\delta_0\right)\right) + 1$. Also M readily admits minorization on the sublevel sets C_d.

To complete the specification of Q consider the potential function $U(x) = |x|$, which clearly satisfies (H3). Lemma 4.2 shows that (H2) is satisfied. The density assumption (H4) is satisfied for β_d proportional to Lebesgue measure restricted to C_d. Again it is straightforward to check that (H5) is satisfied for $\epsilon > 0$ small enough and $\epsilon_0 = \epsilon/2$.

Figure 4.1 also shows estimates of the relative variance obtained by simulation for this model with $\alpha = 0.4$ and using $N = 10^4$ particles, averaged over 10^4 independent realizations. Again the linear growth of the variance is apparent, but there appears to be less variation with respect to the initial condition than in the random walk example.

4.2.2 Cox-Ingersoll-Ross Process

The Cox-Ingersoll-Ross (CIR) process, [Cox et al., 1985], is a diffusion process that is typically used in financial applications to capture mean-reverting behaviour and state-dependent volatility, which is thought to be typical in many real scenarios. The process is defined via the stochastic differential equation:

$$dX_t = \theta\left(\mu - X_t\right) dt + \sigma\sqrt{X_t} dW_t$$

where $\{W_t\}$ is standard Brownian motion, $\theta > 0$ is the mean-reversion rate, $\mu > 0$ is the level of mean-reversion and $\sigma > 0$ is the volatility. We assume that $\frac{2\theta\mu}{\sigma^2} > 1$ so that the process is stationary an never touches zero.

Throughout the remainder of section 4.2.2, for $\Delta > 0$ we denote by M^Δ the transition probability from any time t to $t + \Delta$ of the CIR process with parameters θ, μ, σ. The following lemma identifies a drift function for M^Δ, exhibiting a trade-off between growth rate of the drift function specified by a parameter s, the parameters of the CIR process and the time step size Δ.

Lemma 4.3. *For $s > 0$ and $\Delta > 0$, consider the candidate drift function $V : \mathbb{R}_+ \to [1, \infty)$, defined by*

$$V(x) := 1 + \frac{4\theta s x}{\sigma^2 \left(1 - e^{-\theta\Delta}\right)}. \tag{4.2}$$

Then subject to the conditions:

$$s \in \left(0, \frac{1 - e^{-\theta\Delta}}{2}\right), \quad \delta \in \left(0, \frac{e^{-\theta\Delta}}{1 - 2s}\right), \quad d \geq \frac{1 - 2\theta\mu \log\left(1 - 2s\right)/\sigma^2}{e^{-\theta\Delta}/\left(1 - 2s\right) - \delta} =: \underline{d}, \tag{4.3}$$

the following multiplicative drift condition is satisfied:

$$M^\Delta\left(e^V\right) \quad \leq \quad e^{V(1-\delta) + b_d \mathbb{I}_{C_d}},$$

with V as in (4.2) and $b_d := \frac{d e^{-\theta\Delta}}{1 - 2s} - \frac{2\theta\mu}{\sigma^2} \log\left(1 - 2s\right) + 1$.

Proof. For $t \geq 0$ define

$$c_t := \frac{2\theta}{\sigma^2\left(1 - e^{-\theta t}\right)}, \quad \kappa := \frac{4\theta\mu}{\sigma^2},$$

and the scaled process $Z_t := 2c_t X_t$. Conditional on $X_0 = x$, Z_t has a non-central chi-square distribution with degree of freedom κ and non-centrality parameter taking the value $2c_t x e^{-\theta t}$ [Cox et al., 1985]. We then have for any $x \in \mathsf{X}$,

$$\begin{aligned} M^\Delta\left(e^V\right)(x) &= \mathbb{E}_x\left[\exp\left(sZ_\Delta\right)\right]\exp(1) \\ &= \exp\left[2c_\Delta x s\left(\frac{e^{-\theta\Delta}}{1 - 2s}\right) - \frac{\kappa}{2}\log\left(1 - 2s\right) + 1\right] \\ &\leq \exp\left[V(x)\left(\frac{e^{-\theta\Delta}}{1 - 2s}\right) - \frac{\kappa}{2}\log\left(1 - 2s\right) + 1\right]. \end{aligned}$$

where the equalities hold due to the existence of the moment generating function $\mathbb{E}_x\left[\exp\left(sZ_t\right)\right]$, for $s < 1/2$, which is satisfied under the conditions of (4.3). Under these conditions we also then have for $d \geq \underline{d}$ and $x \notin C_d$,

$$\begin{aligned} M^\Delta\left(e^V\right)(x) &\leq \exp\left[V(x)\left(1 - \delta\right) - d\left(\frac{e^{-\theta\Delta}}{1 - 2s} - \delta\right) - \frac{\kappa}{2}\log\left(1 - 2s\right) + 1\right] \\ &\leq \exp\left[V(x)\left(1 - \delta\right)\right], \end{aligned}$$

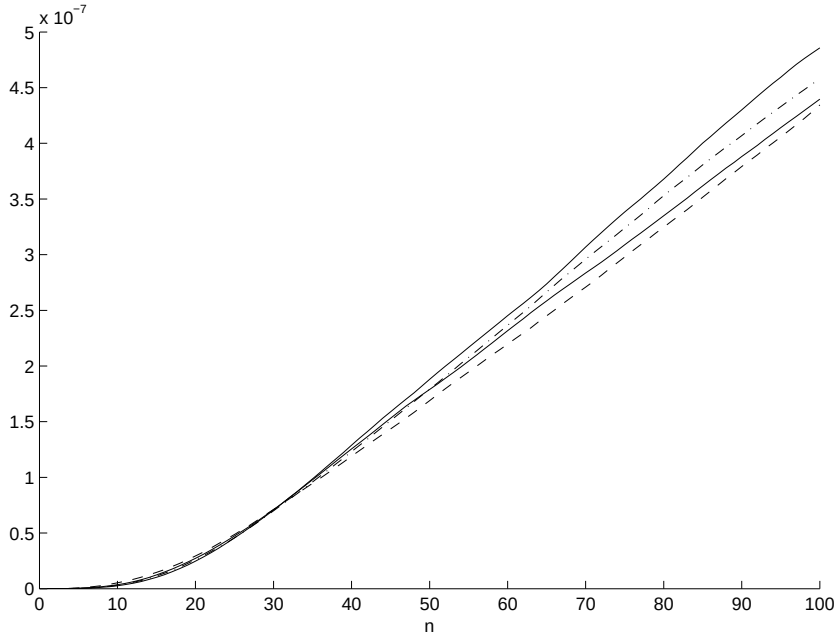

Figure 4.2: Cox-Ingersoll-Ross process. Relative variance vs. n, from initial conditions $x_0 = 0.1$ (dashed), $x_0 = 1$ (solid - bottom), $x_0 = 3$ (dot-dashed), $x_0 = 10$ (solid top).

and for $x \in C_d$,

$$M\left(e^V\right)(x) \quad \leq \quad \exp\left[d\left(\frac{e^{-\theta\Delta}}{1-2s}\right) - \frac{\kappa}{2}\log\left(1-2s\right) + 1\right] = \exp\left(b_d\right).$$

□

We will consider as an example the case where the Markov chain $\{X_n\}$ the skeleton of the CIR process over a discrete time grid of spacing Δ and $U(x) := \alpha \log x$ for some fixed α. Lemmata 4.2 and 4.3 establish that (H2)-(H3) are satisfied and one can check (H4)-(H5) are satisfied similarly to the previous example.

Figure 4.2 displays estimates of the relative variance for this model, computed via simulation, when $\Delta = 0.01$, (i.e. $M \equiv M^{0.01}$), $\alpha = 0.01$, $\theta = 10$, $\mu = 1$, and $\sigma = 0.1$. This was obtained using $N = 10^3$ particles, averaged over 3×10^3 independent realizations. Again the linear growth of the relative variance is present for different initial conditions. Note one may interpret $\gamma_{100,x}(1)$ as the geometric mean $\mathbb{E}_x[\prod_{k=0}^{99} X_k^{1/100}]$, which can be used for prediction in a variety of financial application.

5 Summary

In this paper we have established a linear-in-n bound on the non-asymptotic variance associated with particle approximations of time-homogeneous Feynman-Kac formulae, under assumptions that can be verified on non-compact state-spaces.

There are several possible extensions to this work. Firstly, to consider non-homogeneous Feynman-

Kac formula, which occur routinely in applications such as filtering and Bayesian statistics. Secondly, an important developing area in the analysis of sequential Monte Carlo methods is the case when the dimension of the state-space can be very large [Beskos et al., 2011]. Such analysis has relied on classical geometric drift conditions and it would be interesting to consider the role of multiplicative drift conditions in this context.

Acknowledgements

The first and third authors acknowledge the assistance of the London Mathematical Society for their funding, via a research in pairs grant. The second author was supported by the EPSRC programme grant on Control For Energy and Sustainability EP/G066477/1.

A Proofs and Auxiliary Results for Section 2

Proof. [Proof of Lemma 2.1] Fix any $d \geq \underline{d}$. The lower bound of (2.3) is due to $e^{U(x)} > 0$ for all x. The upper bound of (2.3) is an immediate consequence of the inequality $Q\left(e^V\right)/e^V \leq e^{b_d}$, implied by (2.2).

For the upper bound of (2.4), use the standard inequality $\xi_v(Q) \leq |||Q|||_v$ and then also due to the drift condition in (2.2), $|||Q|||_v < \infty$. Now consider the lower bound. It is claimed that for any $k \geq 3$ and $1 \leq j \leq k-1$,

$$Q_{km_0}\left(e^V\right)(x) \quad \geq \quad Q_{(k-j)m_0}\left(\mathbb{I}_{C_d}\right)(x)\epsilon_d^j \nu_d\left(C_d\right)^{j-1}\nu_d\left(e^V\right), \quad \forall x \in \mathsf{X}, \tag{A.1}$$

where m_0 is as in (H2). For each k, the claim is verified by induction in j; fix $k \geq 3$ arbitrarily. For $j = 1$,

$$Q_{km_0}\left(e^V\right)(x) \geq Q_{(k-1)m_0}\left(\mathbb{I}_{C_d} Q\left(e^V\right)\right)(x) \geq Q_{(k-1)m_0}\left(\mathbb{I}_{C_d}\right)(x)\epsilon_d \nu_d\left(e^V\right)$$

which initializes the induction. Now assume that (A.1) holds at rank $1 \leq j < k-1$. Then at rank $j+1$, applying the induction hypothesis

$$\begin{aligned} Q_{km_0}\left(e^V\right)(x) &\geq Q_{(k-j-1)m_0}\left(\mathbb{I}_{C_d} Q\left(\mathbb{I}_{C_d}\right)\right)(x)\epsilon_d^j \nu_d\left(C_d\right)^{j-1}\nu_d\left(e^V\right) \\ &\geq Q_{(k-j-1)m_0}\left(\mathbb{I}_{C_d}\right)(x)\epsilon_d^{j+1}\nu_d\left(C_d\right)^{j}\nu_d\left(e^V\right), \quad \forall x \in \mathsf{X}, \end{aligned}$$

where (2.1) has been applied, thus the claim is verified.

Now applying (A.1) with $j = k-1$ gives,

$$\frac{Q_{km_0}\left(e^V\right)(x)}{e^{V(x)}} \quad \geq \quad \frac{Q_{m_0}\left(\mathbb{I}_{C_d}\right)(x)}{e^{V(x)}}\epsilon_d^{k-1}\nu_d\left(C_d\right)^{k-2}\nu_d\left(e^V\right) > 0, \quad \forall x \in \mathsf{X},$$

which implies that

$$|||Q_{km_0}|||_v^{1/(km_0)} \geq \epsilon_d^{1-1/(km_0)}\nu_d\left(C_d\right)^{1-2/(km_0)}\nu_d\left(e^V\right)^{1/(km_0)}\left[\sup_{x\in\mathsf{X}}\frac{Q\left(\mathbb{I}_{C_d}\right)(x)}{e^{V(x)}}\right]^{1/(km_0)}.$$

Taking $k \to \infty$ is enough to verify (2.4), as $\lim_{n\to\infty}|||Q^n|||_v^{1/n}$ always exists by subadditivity. □

Proof. [Proof of Lemma 2.2] Set $r \geq \underline{d}$ arbitrarily and let $\hat{Q} := \mathbb{I}_{C_r} Q$, where the dependence of $\hat{Q}$ on r is suppressed for notational convenience. Then under (H3),

$$\begin{aligned}\hat{Q}_{t_0+1}(x,A) &= \mathbb{E}_x\left[\prod_{n=0}^{t_0} \mathbb{I}_{C_r}(X_n)\exp\left(U(X_n)\right)\mathbb{I}_A(X_{t_0+1})\right] \\ &\leq \exp\left(rt_0\left\|U^+\right\|_V\right)\mathbb{E}_x\left[\prod_{n=0}^{t_0} \mathbb{I}_{C_r}(X_n)\exp\left(U(X_{t_0})\right)\mathbb{I}_A(X_{t_0+1})\right], \quad \forall x \in \mathsf{X}, A \in \mathcal{B}(\mathsf{X}),\end{aligned}$$

and therefore under (H4),

$$\hat{Q}_{t_0+1}(x,A) \leq \beta_r^*(A) := \exp\left(rt_0\left\|U^+\right\|_V\right)\int_{C_r} \beta_r(dy)Q(y,A), \quad \forall x \in \mathsf{X}, A \in \mathcal{B}(\mathsf{X}). \tag{A.2}$$

Lemma B3 of [Kontoyiannis and Meyn, 2005] then implies that $\hat{Q}_{2t_0+2}$ is v-separable. Upon considering the difference $Q_{2t_0+2} - \hat{Q}_{2t_0+2}$

$$\left[(Q_{2t_0+2} - \hat{Q}_{2t_0+2})(\varphi)\right] = \sum_{n=0}^{2t_0+1}\left[Q_{2t_0+2-n}\hat{Q}_n - Q_{2t_0+n+1}\hat{Q}_{n+1}\right](\varphi)$$

applying the sub-multiplicative property of the operator norm and using (2.5) which implies that $|||Q - \hat{Q}|||_v \to 0$ as $r \to \infty$ one can conclude by the above arguments. □

The following lemma considers the twisted kernel $\check{P}$ defined in (2.11).

Lemma A.1. *Assume (H1)-(H4). Then there exists $\delta_0 \in (0,\delta)$, $d_0 \geq 1$ and for any $d \geq d_0$, there exists $\check{b}_d < \infty$ such that*

$$\check{P}\left(e^{\check{V}}\right) \leq e^{\check{V} - \delta_0 V + \check{b}_d \mathbb{I}_{C_d}}, \tag{A.3}$$

$$\sup_{x \in C_d} e^{\check{V}(x)} < \infty, \tag{A.4}$$

where $\check{V} : \mathsf{X} \to [1,\infty)$ is defined by $\check{V}(x) := V(x) - \log h_0(x) + c$ with $c =: \log \|h_0\|_v$. Furthermore, there exists $\rho < 1$, depending only on d_0 and δ_0, and for any $d \geq d_0$ there exists $\check{b}'_d < \infty$ such that

$$\check{P}\left(e^{\check{V}}\right) \leq \rho e^{\check{V}} + \check{b}'_d \mathbb{I}_{C_d}. \tag{A.5}$$

Proof. Under the assumptions of the lemma, we have already seen via [Kontoyiannis and Meyn, 2005, Proposition 2.8] that the twisted kernel is well defined. First consider, (A.3); under (H2), setting $\delta_0 \in (0,\delta)$, for any $d \geq \underline{d}$,

$$\begin{aligned}\check{P}\left(\frac{e^V}{h_0}\right) &= \lambda^{-1}h_0^{-1}Q\left(e^V\right) \\ &\leq \exp\left(V - \log h_0 - \delta_0 V - (\delta - \delta_0)V - \log\lambda + b_d\mathbb{I}_{C_d}\right).\end{aligned}$$

As V is unbounded, there exists d_0 such that for all $d \geq d_0$, equation (A.3) holds with $\check{b}_d := b_d - \log\lambda$.

For (A.4) by iteration of the eigenfunction equation, we have that for any $d \geq d_0$,

$$h_0(x) = \lambda^{-m_0} Q_{m_0}(h_0)(x) \geq \epsilon_d \nu_d(h_0), \quad \forall x \in C_d$$

where we apply the minorization part of (H2) to obtain the inequality.

To obtain (A.5), first when $x \notin C_d$, one can use (A.3), which implies that $\check{P}\left(e^{\check{V}}\right)(x) \leq e^{\check{V}(x)-\delta_0 V(x)} \leq e^{\check{V}(x)-\delta_0 d}$ so that (A.5) holds with $\rho := e^{-\delta_0 d_0}$. Equation (A.3) then shows that holds for $x \in C_d$ with $\check{b}'_d := \exp(d - \log\epsilon_d - \log\nu_d(h_0) + \check{b}_d + \log\|h_0\|_v)$. □

B Proofs and Auxiliary Results for Section 3

B.1 Proofs

Proof. [Proof of Proposition 3.1] The starting point of the proof is to write, using the definition of the twisted kernel,

$$\bar{\Gamma}^{(i_1,\ldots i_s)}_{n,x}(F) = \frac{\lambda^{-2n}\Gamma^{(i_1,\ldots,i_s)}_{n,x}(F)}{\lambda^{-2n}\gamma_{n,x}(1)^2} = \frac{\lambda^{-2n}\Gamma^{(i_1,\ldots,i_s)}_{n,x}(F)}{h_0^2(x)\check{\mathbb{E}}_x\left[1/h_0\left(\check{X}_n\right)\right]^2}.$$

Thus in order prove (3.4), we need to prove

$$\lambda^{-2n}h_0^{-2}(x)\Gamma^{(i_1,\ldots,i_s)}_{n,x}(F) \leq \|F\|_{v^{1/2},2}\, c^{s+1}\frac{v(x)}{h_0(x)}\check{\mathbb{E}}_x\left[\prod_{k\in\{i_1,\ldots,i_{s-1}\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s}\right)\right],$$

for each $n \geq 1$, $0 \leq s \leq n+1$ and each possible configuration of the coalescent time indices $(i_1,...,i_s) \in \mathcal{I}_{n,s}$. Throughout the remainder of the proof, c denotes a finite and positive constant, whose value may change on each appearance but depends only the constants in (H1)-(H5).

We first address the case $s = 0$. We then have $\Gamma^{(\emptyset)}_{n,x}(F) = \gamma^{\otimes 2}_{n,x}(F) = Q^{\otimes 2}_{0,n}(F)(x,x) \leq \|F\|_{v^{1/2},2} Q^{\otimes 2}_{0,n}(v\otimes v)(x,x)$ and therefore (recall $\check{v}$ from lemma A.1)

$$\begin{aligned}
\lambda^{-2n}h_0^{-2}(x)\Gamma^{(\emptyset)}_{n,x}(F) &\leq \|F\|_{v^{1/2},2}\,\lambda^{-2n}h_0^{-2}(x)Q^{\otimes 2}_{0,n}(v\otimes v)(x,x) \\
&\leq c\,\|F\|_{v^{1/2},2}\,\check{P}^{\otimes 2}_{0,n}(\check{v}\otimes\check{v})(x,x) \\
&\leq c\,\|F\|_{v^{1/2},2}\,\frac{v(x)}{h_0(x)}v^*(x). \qquad \text{(B.1)}
\end{aligned}$$

where the final inequality follows by iteration of the geometric drift condition (A.5) and the definition of v^*. For the case $s = 1, i_1 = 0$, note that $\Gamma^{(\emptyset)}_{n,x}(F) = \Gamma^{(0)}_{n,x}(F)$ and re-use (B.1).

For the case $s = 1, i_1 > 0$, Lemma B.3 applied with $k = 0, \ell = i_1, m = n$ implies

$$\begin{aligned}
\lambda^{-2n}Q^{\otimes 2}_{0,i_1}DQ^{\otimes 2}_{i_1,n}\left(v^{1/2}\otimes v^{1/2}\right)(x,x) &= \lambda^{-2n}DQ^{\otimes 2}_{0,i_1}DQ^{\otimes 2}_{i_1,n}\left(v^{1/2}\otimes v^{1/2}\right)(x,x') \\
&\leq cv(x)h_0(x)\check{\mathbb{E}}_x\left[v^*\left(\check{X}_{i_1}\right)\right].
\end{aligned}$$

Noting that $\Gamma_{n,x}^{(i_1)}(F) \leq \|F\|_{v^{1/2},2} Q_{0,i_1}^{\otimes 2} DQ_{i_1,n}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2}\right)(x,x)$, this shows that

$$\lambda^{-2n} h_0^{-2}(x) \Gamma_{n,x}^{(i_1)}(F) \quad \leq \quad c \|F\|_{v^{1/2},2} \frac{v(x)}{h_0(x)} \check{\mathbb{E}}_x \left[v^* \left(\check{X}_{i_1}\right)\right], \tag{B.2}$$

which concludes the treatment of the case $s = 1, i_1 > 0$.

We now address the cases $s > 1$. It is claimed that there exists a finite constant c such that for any $n \geq 1$, $(x,x') \in \mathsf{X}^2$, $F \in \mathcal{L}_{v^{1/2},2}$, $1 < s \leq n+1$, and any $(i_1, \ldots, i_s) \in \mathcal{I}_{n,s}$, (one can remove F by using $F \in \mathcal{L}_{v^{1/2},2}$ and hence replace it with $v^{1/2} \otimes v^{1/2}$)

$$\begin{aligned} &\lambda^{-2(n-i_1)} DQ_{i_1,i_2}^{\otimes 2} \ldots DQ_{i_{s-1},i_s}^{\otimes 2} DQ_{i_s,n}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2}\right)(x,x') \\ &\leq c^{s+1} v(x) h_0(x) \check{\mathbb{E}}_x \left[\prod_{k \in \{i_2 - i_1, \ldots, i_{s-1} - i_1\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s - i_1}\right) \right], \end{aligned} \tag{B.3}$$

with the convention that the product is equal to unity when $s = 2$. For a given n, the claim is proved by backward induction through the coalescent time indices. The inductive hypothesis is that at rank $1 \leq j \leq s-1$,

$$\begin{aligned} &\lambda^{-2(n-i_j)} DQ_{i_j,i_{j+1}}^{\otimes 2} \ldots DQ_{i_{s-1},i_s}^{\otimes 2} DQ_{i_s,n}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2}\right)(x,x') \\ &\leq c^{s-j+1} v(x) h_0(x) \check{\mathbb{E}}_x \left[\prod_{k \in \{i_{j+1} - i_j, \ldots, i_{s-1} - i_j\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s - i_j}\right) \right], \end{aligned} \tag{B.4}$$

with the convention that the product equals unity when $j + 1 = s$. When this hypothesis holds, we have at rank $j-1$:

$$\begin{aligned} &\lambda^{-2(n-i_{j-1})} DQ_{i_{j-1},i_j}^{\otimes 2} \ldots DQ_{i_{s-1},i_s}^{\otimes 2} DQ_{i_s,n}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2}\right)(x,x') \\ &\leq c^{s-j+2} v(x) h_0(x) \left(\int \check{P}_{i_{j-1},i_j}(x,dy)\, v(y) \check{\mathbb{E}}_y \left[\prod_{k \in \{i_{j+1} - i_j, \ldots, i_{s-1} - i_j\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s - i_j}\right) \right] \right) \\ &= c^{s-j+2} v(x) h_0(x) \check{\mathbb{E}}_x \left[\prod_{k \in \{i_j - i_{j-1}, \ldots, i_{s-1} - i_{j-1}\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s - i_{j-1}}\right) \right], \end{aligned}$$

where the inequality follows from applying the induction hypothesis and then Lemma B.2, noting that $v \geq 1$. Now consider (B.4) at rank $j = s-1$: the left hand side is

$$\lambda^{-2(n-i_{s-1})} DQ_{i_{s-1},i_s}^{\otimes 2} DQ_{i_s,n}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2}\right)(x,x'),$$

and Lemma B.3 implies that (B.4) does indeed hold at rank $s-1$. This concludes the inductive proof of (B.3).

In the case $i_1 = 0$, (B.3) implies that

$$\lambda^{-2n} h_0^{-2}(x) \Gamma_{n,x}^{(i_1,\ldots,i_s)}(F) \quad \leq \quad \|F\|_{v^{1/2},2}\, c^{s+1} \frac{v(x)}{h_0(x)} \check{\mathbb{E}}_x \left[\prod_{k \in \{i_1,\ldots,i_{s-1}\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s}\right) \right].$$

In the case $i_1 > 0$, we apply Lemma B.2 once more to (B.3) to yield

$$\begin{aligned} &\lambda^{-2n} DQ_{0,i_1}^{\otimes 2} DQ_{i_1,i_2}^{\otimes 2} \ldots DQ_{i_{s-1},i_s}^{\otimes 2} DQ_{i_s,n}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2}\right)(x,x') \\ &\leq c^{s+2} v(x) h_0(x) \check{\mathbb{E}}_x \left[\prod_{k \in \{i_1,\ldots,i_{s-1}\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s}\right) \right] \end{aligned}$$

from which we conclude that

$$\lambda^{-2n} h_0^{-2}(x) \Gamma_{n,x}^{(i_1,\ldots,i_s)}(F) \quad \leq \quad \|F\|_{v^{1/2},2}\, c^{s+2} \frac{v(x)}{h_0(x)} \check{\mathbb{E}}_x \left[\prod_{k \in \{i_1,\ldots,i_{s-1}\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{i_s}\right) \right].$$

This completes the proof of the proposition. □

Proof. [Proof of Lemma 3.1] We first consider some bounds on iterates of the twisted kernel. Standard iteration of the geometric drift condition in equation (B.9) shows that there exists a finite constant c_1 such that

$$\sup_{0 \leq \ell \leq m} \check{P}_{\ell,m}(v^*)(x) \leq c_1 v^*(x), \quad x \in \mathsf{X}, \tag{B.5}$$

and then due to the multiplicative drift condition in equation (B.8),

$$\sup_{0 \leq \ell < m} v(x) \check{P}_{\ell,m}(v^*)(x) = \sup_{0 \leq \ell < m} v(x) \check{P} \check{P}_{\ell+1,m}(v^*)(x) \leq c_1 v(x) \check{P}(v^*)(x) \leq c v^*(x), \quad x \in \mathsf{X}, \tag{B.6}$$

where $c := c_1 e^{\bar{b}_d}$.

In order to prove (3.5) first fix arbitrarily $n \geq 1$, $1 \leq s \leq n+1$ and $(i_1, \ldots, i_s) \in \mathcal{I}_{n,s}$. The proof is via a backward inductive argument through the coalescent time indices. Assume that at rank $1 < j < s$,

$$v(x) \check{\mathbb{E}}_x \left[\prod_{k \in \{i_{j+1} - i_j, \ldots, i_s - i_j\}} v\left(\check{X}_k\right) v^*\left(\check{X}_{n+1-i_j}\right) \right] \quad \leq \quad c^{s+1-j} v^*(x). \tag{B.7}$$

Assuming (B.7) is true, then at rank $j-1$,

$$\begin{aligned}
&v(x)\check{\mathbb{E}}_x\left[\prod_{k\in\{i_j-i_{j-1},\ldots,i_s-i_{j-1}\}} v\left(\check{X}_k\right)v^*\left(\check{X}_{n+1-i_{j-1}}\right)\right]\\
&= v(x)\int \check{P}_{i_{j-1},i_j}\left(x,dx'\right)v(x')\check{\mathbb{E}}_{x'}\left[\prod_{k\in\{i_{j+1}-i_j,\ldots,i_s-i_j\}} v\left(\check{X}_k\right)v^*\left(\check{X}_{n+1-i_j}\right)\right]\\
&\leq c^{s+1-j}v(x)\int \check{P}_{i_{j-1},i_j}\left(x,dx'\right)v^*(x')\\
&\leq c^{s+1-(j-1)}v^*(x),
\end{aligned}$$

where the final inequality is due to equation (B.6). Furthermore

$$v(x)\check{\mathbb{E}}_x\left[v^*\left(\check{X}_{n+1-i_s}\right)\right] \quad = \quad v(x)\check{P}_{i_s,n+1}\left(v^*\right)(x)\leq cv^*(x),$$

where the inequality is again due to (B.6) and therefore at rank $j=s-1$,

$$\begin{aligned}
v(x)\check{\mathbb{E}}_x\left[\prod_{k=(i_s-i_{s-1})} v\left(\check{X}_k\right)v^*\left(\check{X}_{n+1-i_{s-1}}\right)\right] &= v(x)\int \check{P}_{i_{s-1},i_s}\left(x,dx'\right)v(x')\check{\mathbb{E}}_{x'}\left[v^*\left(\check{X}_{n-i_s}\right)\right]\\
&\leq cv(x)\int \check{P}_{i_{s-1},i_s}\left(x,dx'\right)v^*(x')\\
&\leq c^2v^*(x').
\end{aligned}$$

The above arguments prove that (B.7) holds at rank $j=1$ and the proof of the Lemma is then also complete as $n+1$, $1\leq s\leq n+1$ and $(i_1,\ldots,i_s)\in\mathcal{I}_{n,s}$ were arbitrary. □

B.2 Technical Results

Lemma B.1. *Assume (H1)-(H5).Then there exists $\bar{\rho}<1$, $d_0\geq 1$ and for any $d\geq d_0$ there exists $\bar{b}_d<\infty$ and $\bar{b}'_d$ such that*

$$\check{P}\left(e^{V^*}\right) \leq e^{V^*-V+\bar{b}_d\mathbb{I}_{C_d}} \tag{B.8}$$

$$\check{P}\left(e^{V^*}\right) \leq \bar{\rho}e^{V^*}+\bar{b}'_d\mathbb{I}_{C_d}, \tag{B.9}$$

where V^* is as in equation (3.3).

Proof. Under the assumptions of the lemma, Theorem 2.2 holds, the eigenfunction $h_0\in\mathcal{L}_v$, and the twisted kernel is well defined. Then under (H5), we have for any $d\geq\underline{d}$,

$$\begin{aligned}
\check{P}\left(\frac{e^{V(1+\epsilon)}}{h_0}\right) &= \lambda^{-1}h_0^{-1}Q\left(e^{V(1+\epsilon)}\right)\\
&\leq \exp\left(V\left(1+\epsilon\right)-\log h_0-V-\epsilon_0V-\log\lambda+b_d^*\mathbb{I}_{C_d}\right).
\end{aligned}$$

As V is unbounded, there exists d_0 such that for all $d\geq d_0$, equation (B.8) holds with $\bar{b}_d:=b_d^*-\log\lambda$.

The proof of (B.9) then follows exactly as in the proof of Lemma A.1. □

Lemma B.2. *Assume (H1)-(H5). Then there exists $c < \infty$ depending only on the quantities in (H1)-(H5) such that for any $\ell < m$ and $\varphi : \mathsf{X} \to \mathbb{R}_0^+$,*

$$\lambda^{-2(m-\ell)} DQ_{\ell,m}^{\otimes 2} (\varphi \otimes v)(x,x') \le c v(x) h_0(x) \check{P}_{\ell,m} \left(\frac{\varphi}{h_0} \right)(x), \quad (x,x') \in \mathsf{X}, \tag{B.10}$$

where v is as in (H2), and λ and $h_0 \in \mathcal{L}_v$ are respectively the eigenvalue and eigenfunction as in Theorem 2.2.

Proof. By standard iteration of the geometric drift condition in equation (A.5) of Lemma A.1, there is a finite constant c such that

$$\sup_{0 \le \ell < m} \check{P}_{\ell,m} (\check{v})(x) \le c \check{v}(x), \quad x \in \mathsf{X}. \tag{B.11}$$

Then due to the definition of the twisted kernel and $\check{v}$ (see Lemma A.1), there exists a constant c such that for any $0 \le \ell < m$, and $\varphi : \mathsf{X} \to \mathbb{R}_0^+$,

$$\begin{aligned} \lambda^{-2(m-\ell)} Q_{\ell,m}^{\otimes 2} (\varphi \otimes v)(x,x') &= h_0(x) h_0(x') \check{P}_{\ell,m}^{\otimes 2} \left(\frac{\varphi}{h_0} \otimes \frac{v}{h_0} \right)(x,x') \\ &\le c h_0(x) h_0(x') \check{P}_{\ell,m}^{\otimes 2} \left(\frac{\varphi}{h_0} \otimes \check{v} \right)(x,x') \\ &\le c h_0(x) \check{P}_{\ell,m} \left(\frac{\varphi}{h_0} \right)(x)\, v(x'), \quad (x,x') \in \mathsf{X}^2, \end{aligned} \tag{B.12}$$

where the final inequality is due to (B.11). □

Lemma B.3. *Assume (H1)-(H5). Then there exists $c < \infty$ depending only on the quantities in (H1)-(H5) such that for any $0 \le k < \ell \le m$ and $(x,x') \in \mathsf{X}^2$,*

$$\lambda^{-2(m-k)} DQ_{k,\ell}^{\otimes 2} DQ_{\ell,m}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2} \right)(x,x') \le c v(x) h_0(x) \check{\mathbb{E}}_x \left[v^* \left(\check{X}_{\ell-k} \right) \right].$$

Proof. Throughout the proof c is a finite constant whose value may change on each appearance. First consider the case $\ell < m$. Noting that $v \ge 1$ and applying Lemma B.2 with $\varphi = v$ gives the following bound,

$$\begin{aligned} \lambda^{-2(m-\ell)} DQ_{\ell,m}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2} \right)(x,x') &\le c v(x) h_0(x) \check{P}_{\ell,m} \left(\frac{v^{(1+\epsilon_0)}}{h_0} \right)(x) \\ &\le c v(x) h_0(x) \check{P}_{\ell,m} (v^*)(x) \\ &\le c v(x) h_0(x) \check{\mathbb{E}}_x \left[v^* \left(\check{X}_{m-\ell} \right) \right] v(x'). \end{aligned}$$

A further application of Lemma B.2 with $\varphi(x) = v(x) h_0(x) \check{\mathbb{E}}_x \left[v^* \left(\check{X}_{m-\ell} \right) \right]$ and an application of

Lemma 3.1 yields:

$$\begin{aligned}\lambda^{-2(m-k)} DQ_{k,\ell}^{\otimes 2} DQ_{\ell,m}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2}\right)(x,x) &\leq c^2 v(x) h_0(x) \check{\mathbb{E}}_x \left[\check{\mathbb{E}}_{\check{X}_{\ell-k}} \left[v\left(\check{X}_0\right) v^*\left(\check{X}_{m-\ell}\right)\right]\right] \\ &\leq c^2 v(x) h_0(x) \check{\mathbb{E}}_x \left[v^*\left(\check{X}_{\ell-k}\right)\right].\end{aligned}$$

When $\ell = m$, we have again by Lemma B.2,

$$\begin{aligned}\lambda^{-2(m-k)} DQ_{k,\ell}^{\otimes 2} DQ_{\ell,m}^{\otimes 2} \left(v^{1/2} \otimes v^{1/2}\right)(x,x) &= \lambda^{-2(m-k)} Q_{k,m}^{\otimes 2} \left(D\left(v^{1/2} \otimes v^{1/2}\right)\right)(x,x) \\ &= \lambda^{-2(m-k)} DQ_{k,m}^{\otimes 2} (v \otimes 1)(x,x) \\ &\leq cv(x) h_0(x) \check{P}_{k,m}\left(\frac{v}{h_0}\right)(x) \\ &\leq cv(x) h_0(x) \check{P}_{k,m}(v^*)(x) \\ &= cv(x) h_0(x) \check{\mathbb{E}}_x \left[v^*\left(\check{X}_{\ell-k}\right)\right].\end{aligned}$$

This completes the proof. □

References

A. Beskos, D. Crisan, and A. Jasra. On the stability of sequential Monte Carlo methods in high-dimensions. 2011.

F. Cérou, P. Del Moral, and A Guyader. A nonasymptotic variance theorem for unnormalized Feynman Kac particle models. *Annales de l'Institut Henri Poincaré*, 47(3), 2011.

H. P. Chan and T. Lai. A sequential Monte Carlo approach to computing tail probabilities in stochastic models. *The Annals of Applied Probability*, 14(1):(to appear), 2011.

N. Chopin, P. Del Moral, and S. Rubenthaler. Stability of Feynman Kac formulae with path-dependent potentials. *Stochastic Processes and their Applications*, 121(1):38–60, 2011.

J.C. Cox, J.E. Ingersoll Jr, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 7(2):385–407, 1985.

D. Crisan and A. Bain. *Fundamentals of Stochastic Filtering*. Stochastic Modelling and Applied Probability. Springer, 2008.

P. Del Moral. *Feynman-Kac Formulae. Genealogical and interacting particle systems with applications*. Probability and its Applications. Springer Verlag, New York, 2004.

P. Del Moral and A. Doucet. Particle motions in absorbing medium with hard and soft obstacles. *Stochastic Analysis and Applications*, 22:1175–1207, 2004.

P. Del Moral and A. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 37(2), 2001.

P. Del Moral and L. Miclo. Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman Kac semigroups. *ESAIM: Probability and Statistics*, 7:171– 208, March 2003.

P. Del Moral, Patras, and S. Rubenthaler. Tree based functional expansions for Feynman Kac particle models. *Annals of Applied Probability*, 19(2):778–825, 2009.

P. Del Moral, A. Doucet, and A. Jasra. On adaptive resampling strategies for sequential Monte Carlo methods. *Bernoulli*, 2011. To appear.

G. B. Di Masi and L. Stettner. Risk-sensitive control of discrete-time Markov processes with infinite horizon. *SIAM J. Control Optim.*, 38:61–78, November 1999.

A. Doucet, N. De Freitas, and N. Gordon, editors. *Sequential Monte Carlo methods in practice*. Springer, New York, 2001.

A. Jasra and P. Del Moral. Sequential Monte Carlo methods for option pricing. *Stochastic Analysis and Applications*, 29(2):292–317, 2011.

A. Jasra and A. Doucet. Sequential Monte Carlo methods for diffusion processes. *Proceedings of the Royal Society A*, 465:3709–3727, 2009.

I. Kontoyiannis and S.P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *The Annals of Applied Probability*, 13(1):304–362, 2003.

I. Kontoyiannis and S.P. Meyn. Large deviation asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electronic Journal of Probability*, 10(3):61–123, 2005.

F. Le Gland and N. Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filter. *The Annals of Applied Probability*, 14(1):144–187, 2004.

S. Meyn and R L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, 2nd edition, 2009.

S.P. Meyn. Large deviation asymptotics and control variates for simulating large functions. *Annals of Applied Probability*, 16(1):310–339, 2006.

E. Numellin. *General irreducible Markov chains and non-negative operators*. Cambridge Tracts in Mathematics. Cambridge University Press, 2004.

R. van Handel. Uniform time average consistency of Monte Carlo particle filters. *Stochastic Processes and their Applications*, 119(11), 2009.

P. Whittle. *Risk-Sensitive Optimal Control*. John Wiley and Sons, 1990.