Stability properties of some particle filters

Nick Whiteley*

October 4, 2011

Abstract

Under multiplicative drift and other regularity conditions, it is established that the asymptotic variance associated with a particle filter approximation of the prediction filter is bounded uniformly in time, and the non-asymptotic, relative variance associated with a particle approximation of the normalizing constant is bounded linearly in time. The conditions are demonstrated to hold for some hidden Markov models on non-compact state spaces. The particle stability results are obtained by proving v-norm multiplicative stability and exponential moment results for the underlying Feynman-Kac formulae.

1 Introduction

Particle filters have become very popular devices for approximate solution of non-linear filtering problems in hidden Markov models (HMM's) and various aspects of their theoretical properties are now well understood. However, there are still very few results which establish some form of stability over time of particle filtering methods on non-compact spaces, at least without resorting to algorithmic modifications which involve a random computational expense. The aim of the present work is to establish theoretical guarantees about some stability properties of a standard particle filter, under assumptions which are verifiable for some HMM's with non-compact state spaces.

It is now well known that, under mild conditions, the error associated with particle approximation of filtering distributions satisfies a central limit theorem. The first stability property we obtain is a time-uniform bound on the corresponding asymptotic variance. Making use of some recent results on functional expansions for particle approximation measures, the second stability property we obtain is a linear-in-time bound on the non-asymptotic, relative variance of the particle approximations of normalizing constants. These two properties are established by first proving some multiplicative stability and exponential moment results for the Feynman-Kac formulae underlying the particle filter. The adopted approach involves Lyapunov function, multiplicative stability ideas in a weighted ∞-norm setting, which allows treatment of a non-compact state space. We thus obtain stability results which hold under weaker assumptions than those existing in the literature. The main restriction is that our assumptions are typically satisfied under some constraints on the observation component of the HMM and/or the observation sequence driving the filter. On the other hand, subject to these constraints, our stability results

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

hold uniformly over observation records and without any stochasticity necessarily present in the observation process.

The rest of this paper is structured as follows. Section 2 briefly introduces filtering in HMM's, particle filters, and comments on some existing stability results. Section 3 gives some applications of the main particle stability results to classes of hidden Markov models. The hope is that sections 2 and 3 can be read without the reader necessarily delving into the main results of section 4 or the corresponding proofs and auxiliary results of section 5, which are obtained in the more abstract setting of interacting particle approximations of Feynman-Kac formulae.

2 Setting

2.1 Hidden Markov models and filtering

A hidden Markov model is a bi-variate, discrete time Markov chain $((X_n, Y_n); n \geq 0)$ where the signal process (X_n) is also a Markov chain and each observation Y_n is conditionally independent of the rest of the bi-variate process given X_n. Each X_n is valued in a state-space X and each Y_n is valued in the observation space Y. The present work focuses on the case where X is non-compact, and we are typically interested in the case that X is some subset of $\mathbb{R}^d$. In any case, throughout the following we assume that X and Y are Polish spaces endowed with their respective Borel σ-algebras, $\mathcal{B}(\mathsf{X})$ and $\mathcal{B}(\mathsf{Y})$. Our main stability results, presented in section 4, are in the setting of Feynman-Kac formulae which can be considered as underlying the filtering problem of interest. In that section, more precise definitions are given. In the present section, we consider the HMM directly.

Let μ be a probability distribution on X, let f be a Markov kernel acting from X to itself and let g be a Markov kernel acting from X to Y, with $g(x,\cdot)$ admitting density, similarly denoted by $g(x,y)$, with respect to some dominating σ-finite measure. We will assume that $g(x,y) > 0$ and, for now, that $\sup_{x,y} g(x,y) < \infty$. Loosely speaking, the task of filtering is to compute some conditional distributions of the (X_n) process given the observations (Y_n), under an assumed model:

$$\begin{aligned}
&(X_0, Y_0) \sim \mu(dx_0)\, g(x_0, dy_0), \\
&(X_n, Y_n) | \{X_{n-1} = x_{n-1}\} \sim f(x_{n-1}, dx_n)\, g(x_n, dy_n), \quad n \geq 1, \qquad (2.1)
\end{aligned}$$

For a realization of observations $(y_0, y_1, \ldots,)$, we may take as a recursive definition of the (one-step-ahead) *prediction filters*, the sequence of distributions $(\pi_n; n \geq 0)$ following

$$\begin{aligned}
\pi_0(dx_0) &:= \mu(dx_0), \\
\pi_n(dx_n) &:= \frac{\int_{\mathsf{X}} \pi_{n-1}(dx_{n-1})\, g(x_{n-1}, y_{n-1})\, f(x_{n-1}, dx_n)}{\int_{\mathsf{X}} \pi_{n-1}(dx_{n-1})\, g(x_{n-1}, y_{n-1})}, \quad n \geq 1. \qquad (2.2)
\end{aligned}$$

We also define the sequence $(Z_n; n \geq 0)$ by

$$Z_0 := 1, \qquad Z_n := Z_{n-1} \int_{\mathsf{X}} \pi_{n-1}(dx_{n-1})\, g(x_{n-1}, y_{n-1}), \quad n \geq 1. \qquad (2.3)$$

Note that the dependence of π_n and Z_n on $y_{0:n-1} = (y_0, \ldots y_{n-1})$ is suppressed from the notation. Unless stated otherwise, whenever (π_n) or (Z_n) appear below it should be understood that they depend on an arbitrary but fixed and deterministic Y-valued sequence $(y_0, y_1, \ldots)$. The same applies for the particle approximations introduced in section 2.2. The set of observation sequences for which our particle variance results hold is made precise and discussed in section 3.

Under the model (2.1), π_n is the conditional distribution of X_n given $\{Y_{0:n-1} = y_{0:n-1}\}$; and Z_n is the joint density of $Y_{0:n-1}$ evaluated at $y_{0:n-1}$. The convention of working with the one-step-ahead quantities is mostly for simplicity of presentation in the following.

In applications there typically will be some degree of model mis-specification; perhaps the data generating process (X_n, Y_n) is not distributed according to (2.1) with this particular μ, f and g, or perhaps (Y_n) are not the observations from an HMM at all (for ease of presentation we purposefully avoid giving a name to a "true" distribution for (Y_n)). Never-the-less, as $(y_0, y_1, \ldots)$ arrive our aim is to compute, or well-approximate (π_n) and (Z_n) as per (2.2)-(2.3) with some μ, f and g of our choosing.

HMM's are simple and yet flexible models which have found countless applications. However, under choices of μ, f and g which are desirable in many practical situations, (π_n) and (Z_n) are not available in closed form.

2.2 Particle filtering

Particle filters [Gordon et al., 1993] are a class of stochastic algorithms which yield approximations of (π_n) and (Z_n) using a population of N samples which interact over time. These approximations will be denoted by $\left(\pi_n^N\right)$ and $\left(Z_n^N\right)$. Algorithm 1 is perhaps the most simple generic particle filtering scheme (a more precise probabilistic definition is considered in section 4). At time $n \geq 1$, the sampling step performs a selection-mutation operation and is equivalent to choosing, with replacement, N individuals from the population on the basis of their fitness, proportional to $g(\cdot, y_{n-1})$, followed by them each mutating in a conditionally independent manner according to f .

Algorithm 1

For n=0,

Sample $\left(\xi_0^i\right)_{i=1}^N \overset{\text{iid}}{\sim} \mu$,

Report $\pi_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_0^i}, \qquad Z_0^N = 1.$

For $n \geq 1$,

Report $Z_n^N = Z_{n-1}^N \frac{1}{N}\sum_{j=1}^N g\left(\xi_{n-1}^j, y_{n-1}\right)$,

Sample $\left(\xi_n^i\right)_{i=1}^N \Big| \left(\xi_{n-1}^i\right)_{i=1}^N \quad \overset{\text{iid}}{\sim} \quad \frac{\sum_{j=1}^N g\left(\xi_{n-1}^j, y_{n-1}\right) f\left(\xi_{n-1}^j, \cdot\right)}{\sum_{j=1}^N g\left(\xi_{n-1}^j, y_{n-1}\right)}$,

Report $\pi_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i}$.

A large number of variations and extensions of this algorithm have been developed. A full survey is well beyond the scope of the present work, but a few comments are called for. Firstly, Algorithm 1 implicitly uses "multinomial resampling" at every time step. It would be interesting to investigate similar results to those presented here for other resampling schemes, for example via the analyses of Chopin [2004], Del Moral et al. [2011]. Secondly, Algorithm 1 involves mutation at every time step according to the Markov kernel f. Again, various alternative schemes have been devised. Mutation according to f is not an essential characteristic of the main results of section 4 and it is only for simplicity that the results of section 3 are presented in this context. Thirdly, the results presented here are likely to be relevant to related classes of sequential Monte Carlo methods, for example the smoothing algorithms treated by Del Moral et al. [2010] and Douc et al. [2011].

2.3 Existing stability results for particle filters

One of the first and most influential works on stability of particle filters is that of Del Moral and Guionnet [2001] who established time-uniform convergence properties of the particle approximations. They required uniform upper and lower bounds on g and stability of the corresponding exact filter, in turn derived using quite strong assumptions on f involving simultaneous, uniform minorization and majorization, which are rarely satisfied then X is non-compact. Similar mixing assumptions have been employed in LeGland and Oudjane [2004], Chopin [2004], Künsch [2005], Cérou et al. [2011] in order to establish (respectively) uniform convergence of particle filtering approximations; a time-uniform bound on the asymptotic variance; and linear-in-time bounds on the non-asymptotic variance of the normalizing constant. All also consider variants of the standard particle filter in Algorithm 1.

LeGland and Oudjane [2003] developed truncation ideas in order to achieve uniform particle approximations without mixing assumptions, but with random computational cost and/or proposals restricted to compact sets. A further development was made by Oudjane and Rubenthaler [2005], allowing treatment of some non-ergodic signals via a particle filter incorporating an accept/reject step. Truncation ideas have also been used in Heine and Crisan [2008] in order to obtain uniform convergence of particle filter approximations for HMM's on non-compact state-spaces with quite specific structure (including X and Y being of the same dimension). van Handel [2009] has established uniform convergence of time-averaged filters under tightness assumptions on non-compact spaces. Del Moral and Jacod [2001] proved tightness of the sequence of asymptotic variances (as a function of random observations) in the linear-Gaussian case. Favetto [2009] has proved tightness of the same for a class of HMM's, but subject to a mixing assumption on f.

It is stressed that: 1) a time-uniform bound on the asymptotic variance for π_n^N and 2) a linear-in-time bound on the relative variance for Z_n^N , as pursued here, are different properties from the time-uniform convergence results proved in most of the above. The existing works featuring the most similar type of results to those considered here are [Chopin, 2004], [Künsch, 2005], [Favetto, 2009] and [Cérou et al., 2011], all of which rely on strong mixing assumptions, at least on f, which we do not invoke.

The overall approach used in the present work to express Feynman-Kac formulae and associated functionals is the semigroup formulation of Del Moral [2004], but the stability ideas are different and are based around a weighted ∞-norm function space setting. In Theorem 1, the decomposition idea of Kleptsyna and Veretennikov [2008] and some technical approaches from Douc et al. [2009] are employed.

For completeness we also mention the following. Whiteley [2011] considered stability properties of a related class of sequential Monte Carlo methods which are not used for filtering and operate in a different structural regime, where the number of distributions involved may be considered a parameter of the algorithm. Whiteley et al. [2011] considered relative variance for Z_n^N in the context of time-homogeneous Feynman-Kac models (obtained in the present setting by setting all $y_0, y_1, \ldots$ to a constant), appealing to spectral properties of the integral kernel involved. There is nothing explicitly spectral about the present work, but there are some related structural ideas involved (see section 4). For example, Theorem 1 is expressed in such a way that it may be viewed as an non-homogeneous analogue of the v-norm multiplicative ergodicity results of Kontoyiannis and Meyn [2005], in the context of positive operators. The assumptions in the present work also allow the treatment of time-homogeneous Feynman-Kac models, and in that setting are actually stronger than the assumptions of Whiteley et al. [2011] (because in assumption (H3)-(H4) of section 4.2 here, we require a simultaneous local minorization/majorization condition), but on the other hand the approach of Whiteley et al. [2011] is specific to the time-homogeneous setting.

3 Summary and application of some results

In this section, the results of section 4 are summarized and applied to some specific hidden Markov models and the particle filter of Algorithm 1. To this end we consider the following assumptions on μ, f and g which serve as an intermediate layer of abstraction and which together imply that assumptions (H1)-(H5) of section 4 are satisfied. Discussion of the latter assumptions and their relation to the existing literature is given in section 4.1.1.

Consider the following:

- $\mathsf{Y}_\star \subseteq \mathsf{Y}$ is measurable, and the quantities in the below conditions may depend on $\mathsf{Y}_\star$.
- There exists $V : \mathsf{X} \to [1, \infty)$ unbounded, $\underline{d} \in [1, \infty)$ and $\delta > 0$ with the following properties. For each $d \in [\underline{d}, \infty)$,

$$g(x, y) \int_{C_d} f(x, dx') > 0, \quad \forall x \in \mathsf{X},\ y \in \mathsf{Y}_\star, \tag{3.1}$$

where $C_d := \{x : V(x) \leq d\}$, and there exists $b_d < \infty$ such that

$$\begin{aligned} &\sup_{y \in \mathsf{Y}_\star} g(x, y) \int_{\mathsf{X}} f(x, dx') \exp\left[V(x')\right] \\ &\leq \exp\left[V(x)(1 - \delta) + b_d \mathbb{I}_{C_d}(x)\right], \quad \forall x \in \mathsf{X}, \end{aligned} \tag{3.2}$$

and there exists a probability measure ν_d and $0 < \epsilon_d^- \leq \epsilon_d^+ < \infty$ such that

$$\begin{aligned} & \epsilon_d^- \nu_d \left(dx'\right) \mathbb{I}_{C_d}\left(x'\right) \\ & \leq g\left(x,y\right) f\left(x,dx'\right) \mathbb{I}_{C_d}\left(x'\right) \\ & \leq \epsilon_d^+ \nu_d \left(dx'\right) \mathbb{I}_{C_d}\left(x'\right), \quad \forall x \in C_d,\ y \in \mathsf{Y}_\star, \end{aligned} \tag{3.3}$$

with $\nu_d\left(C_r\right) > 0$ for all $r \in [\underline{d}, d]$

- $\int \exp\left[V(x)\right] \mu\left(dx\right) < \infty$
- Although not required for all results of section 4, in the present section it is also assumed that

$$\sup_{(x,y)\in\mathsf{X}\times\mathsf{Y}_\star} g(x,y) < \infty. \tag{3.4}$$

The condition of (3.2) is a multiplicative drift condition. Similar conditions have been used in the study of stability of exact filters [Douc et al., 2009] and can hold when $\mathsf{Y}_\star = \mathsf{Y}$ is non-compact. It may be the case that f alone satisfies such a multiplicative condition (see section 3.1 below), in which case (3.2) can be satisfied when $\sup_{y\in\mathsf{Y}} g(x,y)$ is not bounded above in x. When (3.4) holds, then (3.2) can hold even when f is not ergodic, but it is then typically required that $\mathsf{Y}_\star \subset \mathsf{Y}$ is compact (see section 3.2). The conditions of (3.3) and (3.4) together imply that for all $d \in [\underline{d}, \infty)$,

$$\sup_{y\in\mathsf{Y}_\star} \sup_{(x,x')\in C_d\times C_d} \frac{g\left(x,y\right)}{g\left(x',y\right)} < \infty,$$

which can, loosely, be interpreted as a constraint on the amount of information which any single observation in $\mathsf{Y}_\star$ can provide about the hidden state in each C_d. For the example of section 3.1.1 we are able to satisfy the assumptions when $\mathsf{Y}_\star = \mathsf{Y}$ is compact. For non-compact Y in the examples below, we resort to taking $\mathsf{Y}_\star$ compact.

Under the above assumptions, the main conclusions of Propositions 3 and 4, section 4.5, may be summarized as follows.

Uniformly bounded variance in the CLT for π_n^N

It is known (e.g. [Del Moral, 2004, Section 9.4.2]) that under (3.4), for any $\varphi : \mathsf{X} \to \mathbb{R}$ bounded, measurable, $n \geq 1$ and any $y_{0:n} \in \mathsf{Y}_\star^{n+1}$,

$$\sqrt{N} \int_{\mathsf{X}} \left[\pi_n^N(dx) - \pi_n(dx)\right] \varphi(x) \longrightarrow \mathcal{N}\left(0, \sigma_n^2\left(y_{0:n}\right)\right)$$

in distribution as $N \to \infty$. Under the conditions of (3.1)-(3.4), Proposition 3 may be applied to establish there exists $c_\mu < \infty$ depending on $\mathsf{Y}_\star$, such that for all such φ and $n \geq 0$

$$\sigma_n^2\left(y_{0:n}\right) \leq \mathrm{Var}_{\pi_n}\left(\varphi\right) + \left\|\varphi\right\|^2 c_\mu, \quad \forall y_{0:n} \in \mathsf{Y}_\star^{n+1}, \tag{3.5}$$

with $\|\cdot\|$ the sup norm. Discussion of a CLT for other classes of φ is given in section 4.5.1.

Linearly bounded relative variance for Z_n^N

Under the conditions of (3.1)-(3.4), Proposition 4 may be applied to establish that there exists $c'_\mu < \infty$ depending on $\mathsf{Y}_\star$ such that for all $n \geq 0$,

$$N > c'_\mu (n+1) \quad \Longrightarrow \quad \mathbb{E}_\mu \left[\left(\frac{Z_n^N}{Z_n} - 1 \right)^2 \right] \leq c'_\mu \frac{4}{N} (n+1), \quad \forall y_{0:n} \in \mathsf{Y}_\star^{n+1}. \tag{3.6}$$

where $\mathbb{E}_\mu$ is expectation with respect to the law of the N-particle filtering algorithm initialized using μ.

3.1 A class of ergodic signal models

The following class of signal model has been considered by Kleptsyna and Veretennikov [2008] and Douc et al. [2009] in the context of stability of exact filters (i.e. without particle approximation). We have $\mathsf{X} = \mathbb{R}^{d_x}$ for some $d_x \geq 1$. The transition kernel f corresponds to the signal model

$$X_{n+1} = X_n + B(X_n) + \sigma(X_n) W_n, \qquad (W_n; n \geq 1) \overset{\text{iid}}{\sim} \mathcal{N}(0, I_{d_x}), \tag{3.7}$$

with

- B is a d_x-dimensional vector function, locally bounded and

$$\lim_{r \to \infty} \sup_{|x| \geq r} |x + B(x)| - |x| = -\infty \tag{3.8}$$

- σ is a $d_x \times d_x$ matrix function, and has the so-called non-degenerate noise variance property

$$0 < \inf_{x \in \mathbb{R}^{d_x}} \inf_{\lambda \in \mathbb{R}^{d_x}, |\lambda|=1} \lambda^T \sigma(x) \sigma^T(x) \lambda \leq \sup_{x \in \mathbb{R}^{d_x}} \sup_{\lambda \in \mathbb{R}^{d_x}, |\lambda|=1} \lambda^T \sigma(x) \sigma^T(x) \lambda < \infty. \tag{3.9}$$

As per Lemma 4 in the appendix, f in this case itself satisfies a multiplicative drift condition with $v(x) := \exp(1 + c|x|)$ for c a positive constant. An example of a possible signal model with non-Gaussian transition probability and f itself satisfying a multiplicative drift condition is the discretely sampled Cox-Ingersoll-Ross process, see [Whiteley et al., 2011].

We now discuss some observation models which may be combined with the signal model above.

3.1.1 Discrete-valued observations

With $\mathsf{Y} = \{0, 1\}^{d_x}$, consider the multivariate binary observation model

$$\left(Y_n^1, \ldots, Y_n^{d_x}\right) \Big| \{X_n = x_n\} \sim \mathcal{B}e\left(p\left(x_n^1\right)\right) \otimes \cdots \otimes \mathcal{B}e\left(p\left(x_n^{d_x}\right)\right),$$

where $\mathcal{B}e$ denotes the Bernoulli distribution, $p(x) := 1/(1+e^{-x})$ and $Y_n = \left(Y_n^1, \ldots, Y_n^{d_x}\right)$, $x_n = \left(x_n^1, \ldots, x_n^{d_x}\right)$. This corresponds to

$$g(x,y) = \prod_{j=1}^{d_x} p\left(x^j\right)^{\mathbb{I}[y^j=1]} \left(1 - p\left(x^j\right)\right)^{\mathbb{I}[y^j=0]}.$$

Clearly $\sup_{x,y} g(x,y) = 1$ and for any compact $C \subset \mathbb{R}^{d_x}$, $\inf_{x \in C} \inf_{y \in \mathsf{Y}} g(x,y) > 0$. Combined with Lemma 4, this establishes that the assumptions of equations (3.1), (3.2) and (3.3) are satisfied when this observation model is combined with the signal model of equation (3.7)-(3.9).

3.1.2 Uninformative observations in $\mathbb{R}^d$

With $\mathsf{Y} = \mathbb{R}^{d_y}$, $d_y \geq 1$, Consider the observation model

$$Y_n = H\left(X_n\right) + \zeta_n, \qquad (\zeta_n; n \geq 1) \overset{\text{iid}}{\sim} \mathcal{N}\left(0, I_{d_y}\right),$$

with H a bounded, vector-function. That the disturbance terms are standard normal here is only for simplicity of presentation. Obviously we have

$$g(x,y) = \frac{1}{(2\pi)^{d_y/2}} \exp\left(-\frac{1}{2}\left[y - H(x)\right]^T \left[y - H(x)\right]\right)$$

so that $\sup_{(x,y) \in (\mathsf{X},\mathsf{Y})} g(x,y) = (2\pi)^{-d_y/2}$. In this case the observations may be considered uninformative as for each y, $\inf_{x \in \mathsf{X}} g(x,y) > 0$. In light of Lemma 4, standard calculations show that this observation model combined with f of (3.7)-(3.9) satisfies the drift condition of (3.2) with $\mathsf{Y}_\star = \mathsf{Y}$ and $\underline{d}$ chosen large enough. However, when we attempt to verify (3.3) (via (5.2) in Lemma 4) by incorporating $g(x,y)$, the minorization part of (3.3) is not satisfied with $\mathsf{Y}_\star = \mathsf{Y}$, due to the requirement of uniformity in y. We may satisfy (3.3) by taking $\mathsf{Y}_\star \subset \mathsf{Y}$ a compact set, and the constants involved will then depend on $\mathsf{Y}_\star$.

3.1.3 Stochastic volatility observations

With $\mathsf{Y} = \mathbb{R}$ and $d_x = 1$, consider the stochastic volatility observation model (considered in [Douc et al., 2009, Section 4.3]),

$$Y_n = \beta \exp\left(X_n/2\right) \epsilon_n, \qquad (\epsilon_n; n \geq 0) \overset{\text{iid}}{\sim} \mathcal{N}\left(0, I_{d_x}\right)$$

where $\beta > 0$ is a fixed parameter of the model. The corresponding likelihood is

$$g(x,y) = \frac{1}{(2\pi)^{1/2}\beta} \exp\left[-y^2 \exp\left(-x\right)/(2\beta^2) - x/2\right],$$

which is not uniformly upper-bounded on $\mathsf{X} \times \mathsf{Y}$. But, as noted in [Douc et al., 2009, Section 4.3], $\sup_{x \in \mathsf{X}} g(x,y) \leq (2\pi e)^{-1/2} |y|^{-1}$. For $0 < \underline{y} < \bar{y} < \infty$, take $\mathsf{Y}_\star := \left[-\bar{y}, -\underline{y}\right] \cup \left[\underline{y}, \bar{y}\right]$. Then (3.4) is satisfied and using Lemma 4, the drift condition of (3.2) and the upper bound of (3.3) are satisfied with $\underline{d}$ large enough. The lower bound of (3.3) is also satisfied because for $d < \infty$, $\inf_{(x,y) \in C_d \times \mathsf{Y}_\star} g(x,y) > 0$.

3.2 A class of possibly non-ergodic signal models

We now consider a class of signal model which includes some non-ergodic f and point out how characteristics of the observation model can be used to satisfy the drift condition (3.2).

Take $\mathsf{X} = \mathbb{R}^{d_x}$ for some $d_x \geq 1$ and consider the signal model

$$X_{n+1} = B(X_n) + W_n, \qquad (W_n; n \geq 0) \overset{\text{iid}}{\sim} \mathcal{N}(0, I_{d_x}), \tag{3.10}$$

with B is a d_x-dimensional vector function, locally bounded. That the disturbance terms (W_n) are standard Normal is only for simplicity of presentation; one can draw analogous conclusions under conditions such as (3.9), but we focus here on the interplay between V, $\mathsf{Y}_\star$, B and g. For some $\delta_0 > 1$, take $V(x) := \dfrac{x^T x}{2(1+\delta_0)} + 1$.

Assuming that $\mathsf{Y}_\star$, B and g are such that, for some $\delta_1 \in (0,1)$,

$$\lim_{r\to\infty} \sup_{|x|\geq r} \sup_{y \in \mathsf{Y}_\star} -\frac{(1-\delta_1)x^T x}{2(1+\delta_0)} + \frac{1}{2\delta_0} B^T(x) B(x) + \log g(x,y) < 0. \tag{3.11}$$

standard manipulations then establish that the drift condition of (3.2) is satisfied with $\delta < \delta_1$ and $\underline{d}$ large enough. For the condition of (3.2), again with $\underline{d}$ large enough we can take ν_d the normalized restriction of Lebesgue measure to C_d if it is the case that

$$\inf_{(x,y) \in C_d \times \mathsf{Y}_\star} g(x,y) > 0. \tag{3.12}$$

The conditions (3.11) and (3.12) are satisfied, for example, when

- The signal model is a random walk, $B(x) := x$
- $\mathsf{Y} = \mathbb{R}^{d_y}$,

$$Y_n = H(x) + \sigma_y \zeta_n, \qquad (\zeta_n; n \geq 0) \overset{\text{iid}}{\sim} \mathcal{N}\left(0, I_{d_y}\right)$$

  with $\sigma_y > 0$, so that

$$g(x,y) = \frac{1}{(2\pi)^{d_y/2} \sigma_y^{d_y}} \exp\left(-\frac{1}{2\sigma_y^2} [y - H(x)]^T [y - H(x)]\right)$$

- $\mathsf{Y}_\star$ is compact
- H is locally bounded and such that

$$\lim_{r\to\infty} \sup_{|x|\geq r} \quad \frac{x^T x}{2} \frac{(1+\delta_1)}{\delta_0(1+\delta_0)} + \frac{1}{\sigma_y^2} \left(\sup_{y\in\mathsf{Y}_\star} |y|\right) \left(\sup_{|\lambda|=1} \lambda^T H(x)\right) - \frac{H(x)^T H(x)}{2\sigma_y^2} < 0.$$

For the latter condition, we observe a trade-off in terms of δ_0 (which defines V), the constant δ_1 ($\delta < \delta_1$ appears in the drift condition), the observation noise variance σ_y^2 and the growth of $H(x)$.

4 $\mathcal{L}_v$-stability of Feynman-Kac formulae and particle approximations

4.1 Definitions and assumptions

As per the introduction, let the Polish state space X be non-compact and endowed with its Borel σ-algebra $\mathcal{B}(\mathsf{X})$ (the observation space Y will not feature explicitly in the following Feynman-Kac formulation, see Remark 2 below). 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 η denote $\eta(\varphi) := \int \varphi(x)\eta(dx)$, $K\varphi(x) = K(\varphi)(x) := \int K(x, dy)\varphi(y)$ and $\eta K(\cdot) := \int \eta(dx)K(x, \cdot)$. 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 η such that $\eta(v) < \infty$.

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 corresponding v-norm on signed measures is $\|\eta\|_v := \sup_{|\varphi|\leq v} \eta(v)$. For any $n \geq 1$ and $1 \leq s \leq (n + 1)$, define $\mathcal{I}_{n,s} := \{(i_1, ..., i_s) \in \mathbb{N}_0^s; 0 \leq i_1 < \ldots < i_s \leq n\}$.

Let $\mu \in \mathcal{P}$ be an initial distribution and for each $n \in \mathbb{N}$ let $(M_n; n \geq 1)$ be a collection of Markov kernels, each kernel acting $\mathsf{X} \times \mathcal{B}(\mathsf{X}) \to [0, 1]$. Let $(G_n; n \geq 0)$ be a collection of $\mathcal{B}(\mathsf{X})$-measurable, real-valued, strictly positive functions on X.

Next let $(Q_n; n \geq 1)$ be the collection of integral kernels defined by

$$Q_n(x, dx') := G_{n-1}(x)M_n(x, dx').$$

For $1 \leq p \leq n$, let $Q_{p,n}$ be the semigroup defined by

$$Q_{p,n} \quad := \quad Q_{p+1} \ldots Q_n, \quad p < n, \tag{4.1}$$

$Q_{n,n} = Id$ and by convention $Q_{n+1,n} = Id$.

We now introduce our first two assumptions, which will be called upon in the following.

(H1) There exists $V : \mathsf{X} \to [1, \infty)$ unbounded and constants $\delta > 0$ and $\underline{d} \geq 1$ with the following properties. For each each $d \in [\underline{d}, \infty)$ there exists $b_d < \infty$ such that the following multiplicative drift condition holds:

$$\sup_{n\geq 1} Q_n\left(e^V\right) \quad \leq \quad e^{V(1-\delta)+b_d \mathbb{I}_{C_d}}, \tag{4.2}$$

where $C_d := \{x \in \mathsf{X}; V(x) \leq d\}$.

Whenever (H1) holds we may also consider:

(H2) $\mu \in \mathcal{P}_v$, where $v = e^V$ is as in (H1).

We may now proceed with some further definitions. Define the collection of measures $(\gamma_n; n \geq 0)$ and probability measures $(\eta_n; n \geq 0)$

$$\gamma_n(A) := \mu Q_{0,n}(A), \qquad \eta_n(A) := \frac{\gamma_n(A)}{\gamma_n(1)}, \quad A \in \mathcal{B}(\mathsf{X}), \tag{4.3}$$

where the dependence of (γ_n) and (η_n) on the initial distribution μ is suppressed from the notation.

Before going further we note the following elementary implications of the assumptions (H1) and (H2) introduced so far. Assumption (H1) implies that for all $n \geq 1$ and $x \in \mathsf{X}$, $Q_n\left(e^V\right)(x)/e^{V(x)} \leq e^{b_{\underline{d}}} < \infty$ and thus for all $0 \leq p \leq n$ and $x \in \mathsf{X}$,

$$Q_{p,n}\left(e^V\right)(x) \quad < \quad \infty, \tag{4.4}$$

Combined with Assumption (H2), we also observe that that for all $n \geq 0$, $\eta_n \in \mathcal{P}_v$.

It is straightforward to verify that the unnormalized measures (γ_n) have the following product representation:

$$\gamma_n(A) = \prod_{p=0}^{n-1} \eta_p(G_p)\, \eta_n(A), \quad n \geq 1. \tag{4.5}$$

We denote by $\mathbb{E}_\mu$ the expectation w.r.t. to the canonical law of the non-homogeneous Markov chain $(X_n; n \geq 0)$ where $X_0 \sim \mu$ and $X_n | \{X_{n-1} = x_{n-1}\} \sim M_n(x_{n-1}, \cdot)$. For $p \leq n$ and a suitable test function φ we abuse notation by writing

$$\mathbb{E}_{p,x}\left[\varphi(X_p, \ldots, X_n)\right] \quad := \quad \mathbb{E}_\mu\left[\varphi(X_p, \ldots, X_n) | X_p = x\right],$$

and for a probability measure η we write

$$\mathbb{E}_{p,\eta}\left[\varphi(X_p, \ldots, X_n)\right] \quad := \quad \int_{\mathsf{X}} \eta(dx)\, \mathbb{E}_{p,x}\left[\varphi(X_p, \ldots, X_n)\right].$$

Under these notational conventions we have, for $0 \leq p < n$ and $\eta \in \mathcal{P}$, the identity

$$\eta Q_{p,n}(A) = \mathbb{E}_{p,\eta}\left[\prod_{q=p}^{n-1} G_q(X_q)\, \mathbb{I}[X_n \in A]\right].$$

In particular,

$$\eta_p Q_{p,n}(A) = \mathbb{E}_{p,\eta_p}\left[\prod_{q=p}^{n-1} G_q(X_q)\, \mathbb{I}[X_n \in A]\right] = \prod_{q=p}^{n-1} \eta_q(G_q)\, \eta_n(A),$$

due to (4.3) and (4.5), which will be used repeatedly.

Definition 1. *(λ-values and h-functions).* For $n \geq 0$ let

$$\lambda_n := \eta_n(G_n),$$

and for $0 \leq p \leq n$ let $h_{n,p} : \mathsf{X} \to (0, \infty)$ be the function defined by

$$h_{n,n}(x) := 1, \qquad h_{p,n}(x) := \frac{Q_{p,n}(1)(x)}{\prod_{q=p}^{n-1} \lambda_q}, \quad p < n. \tag{4.6}$$

Remark 1. It is stressed that each λ_p, and therefore each $h_{p,n}$, depends implicitly on the initial distribution μ. This plays a key structural role in the proofs which follow. With the exception of Corollary 1, throughout the following μ should be understood as arbitrary but fixed.

The two other main assumptions are the following.

(**H3**) With $\underline{d}$ as in (H1), for each $d \in [\underline{d}, \infty)$,

$$Q_n(x, C_d) > 0 \quad \forall x \in \mathsf{X},\ n \geq 1,$$

and there exists $\epsilon_d^- > 0$ and $\nu_d \in \mathcal{P}_v$, such that

$$\inf_{n \geq 1} Q_n(x, C_d \cap A) \quad \geq \quad \epsilon_d^- \nu_d(C_d \cap A), \quad \forall x \in C_d,\ A \in \mathcal{B}(\mathsf{X}).$$

with $\nu_d(C_r) > 0$, for all $r \in [\underline{d}, d]$.

When (H1) and (H3) hold, we may also consider:

(**H4**) With $\underline{d}$ as in (H1) and (ν_d), (ϵ_d^-) as in (H3), for each $d \in [\underline{d}, \infty)$ there exists $\epsilon_d^+ \in [\epsilon_d^-, \infty)$ such that

$$\sup_{n \geq 1} Q_n(x, C_d \cap A) \quad \leq \quad \epsilon_d^+ \nu_d(C_d \cap A), \quad \forall x \in C_d,\ A \in \mathcal{B}(\mathsf{X}).$$

4.1.1 Comments on the assumptions

Assumptions (H3)-(H4) taken together are more specific than the local-Doeblin condition of Douc et al. [2009] (when the latter is considered as holding for non-negative kernels) because they are phrased in terms of the level sets for V and hold time-simultaneously. It is possible to obtain results which are the analogue of those presented herein under multi-step versions of (H3)-(H4), but this involves substantial notational complications which would obscure presentation.

Assumption (H1) is a type of multiplicative drift condition involving the Markov kernels (M_n) and the potential functions (G_n). A notable characteristic of this assumption is that it implies that for all $\epsilon > 0$ there exists $d \geq \underline{d}$ such that $|||Q_n - \mathbb{I}_{C_d} Q_n|||_v < \epsilon$ for all $n \geq 1$, which is itself time-simultaneous variation of [Douc et al., 2009, condition H2]. In the present work, the sublevel sets of V play a central role in the proof of Theorem 2 and thus Proposition 4.

In the above definitions the functions (G_n) have been taken as strictly positive. It would be interesting to also consider vanishing potential functions, but that situation is more complicated as the particle system may become extinct.

4.1.2 Particle system

The particle system may be considered a canonical non-homogeneous Markov chain and therefore its definition is only sketched. For $N \geq 1$, and each $n \geq 0$ let $\xi_n = \left(\xi_n^1, \ldots, \xi_n^N\right)$, be a X^N-valued and then define

$$\eta_n^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\xi_n^i},\ n \geq 0,$$
$$\gamma_0^N := \eta_0^N,$$
$$\gamma_n^N := \left[\prod_{p=0}^{n-1} \eta_p^N \left(G_p\right)\right] \eta_n^N,\ n \geq 1.$$

The particle system of population size N is the X^N- valued Markov chain with transitions given symbolically by

$$\left(\xi_0^1, \ldots \xi_0^N\right) \overset{\text{iid}}{\sim} \mu, \qquad \left(\xi_n^1, \ldots \xi_n^N\right) \middle| \, \xi_{n-1} \overset{\text{iid}}{\sim} \frac{\eta_{n-1}^N Q_n\left(\cdot\right)}{\eta_{n-1}^N\left(G_{n-1}\right)},\ n \geq 1.$$

Remark 2. In order to obtain algorithm 1 take $G_n(x) := g(x, y_n)$, $M_n(x, dx') := f(x, dx')$. In that case $\eta_n^N \equiv \pi_n^N$ and $\gamma_n^N(1) \equiv Z_n$, and similarly, $\eta_n \equiv \pi_n$, $\gamma_n(1) \equiv Z_n$. Other particle filters (such as the "fully-adapted" auxiliary particle filter of Pitt and Shephard [1999]) arise from other choices of G_n and M_n. More generally, the state-space X may be augmented, e.g. to X^2, in order to accommodate M_n corresponding to other choices of proposal kernel and corresponding importance weight, see for example Doucet et al. [2000]. In such cases one would need multi-step versions of (H3)-(H4).

4.2 Uniform v-controls

The main results of this section are Propositions 1 and 2, which establish uniform controls on the measures (η_n), the λ-values and the h-functions. The uniform bounds of these propositions play a key role in the stability results which then follow.

The first key ingredient is the following Lemma, which establishes some relationships between the measures (η_n), the λ-values and the h-functions.

Lemma 1. *Assume (H1)-(H2). The measures (η_n), h-functions and λ-values satisfy, for any $n \geq 1$ and $0 \leq p < n$, the recursive formulae*

$$\eta_p Q_{p+1} = \lambda_p \eta_{p+1}, \quad Q_{p+1}\left(h_{p+1,n}\right) = \lambda_p h_{p,n}, \tag{4.7}$$

and

$$\eta_p\left(h_{p,n}\right) = 1.$$

Furthermore $h_{p,n} \in \mathcal{L}_v$ where $v = e^V$ is as in (H1).

Proof. For the measure equation,

$$\eta_n\left(A\right) = \frac{\gamma_n\left(A\right)}{\gamma_n\left(1\right)} = \frac{\gamma_{n-1} Q_n\left(A\right)}{\gamma_{n-1} Q_n\left(1\right)} = \frac{\eta_{n-1} Q_n\left(A\right)}{\eta_{n-1} Q_n\left(1\right)} = \frac{\eta_{n-1} Q_n\left(A\right)}{\eta_{n-1}\left(G_{n-1}\right)},$$

where the third equality is due to the product formula (4.5). For the h-function equation, using Definition 1,

$$h_{p-1,n} = \frac{Q_{p-1,n}(1)}{\prod_{q=p-1}^{n-1}\lambda_q} = \frac{1}{\lambda_{p-1}}\frac{Q_p Q_{p,n}(1)}{\prod_{q=p}^{n-1}\lambda_q} = \frac{1}{\lambda_{p-1}}Q_p(h_{p,n}).$$

The equality $\eta_p(h_{p,n}) = 1$ is direct from (4.3) and the definition of $h_{p,n}$. The assertion $h_{p,n} \in \mathcal{L}_v$ follows immediately from Definition 1 and (4.4). □

The second key ingredient is the collection of kernels and drift functions identified in the following definition (that these kernels are Markov is a consequence of Lemma 1).

Definition 2. *(S-kernels and drift functions).* For $n \geq 1$, $1 \leq p \leq n$ let $S_{p,n} : \mathsf{X} \times \mathcal{B}(\mathsf{X}) \to \mathbb{R}_+$ be the Markov kernel defined by

$$S_{p,n}(x, A) := \frac{Q_p(\mathbb{I}_A h_{p,n})(x)}{\lambda_{p-1}h_{p-1,n}(x)}, \tag{4.8}$$

and let $v_{p,n} : \mathsf{X} \to [1.\infty)$ be defined by

$$v_{p,n}(x) := \frac{v(x)}{h_{p,n}(x)}\|h_{p,n}\|_v,$$

where v is as in (H1).

For each $n \geq 1$ and $\eta \in \mathcal{P}$, We denote by $\check{\mathbb{E}}_\eta^{(n)}$ expectation w.r.t. the canonical law of the $(n+1)$-step non-homogeneous Markov chain $\{\check{X}_{p,n}; 0 \leq p \leq n\}$ with $\check{X}_{0,n} \sim \eta$ and for $1 \leq p \leq n, \check{X}_{p,n} \big| \{\check{X}_{p-1,n} = \check{x}_{p-1,n}\} \sim S_{p,n}(\check{x}_{p-1,n}, \cdot)$. By analogy to the definitions of section 4.1, for each $n \geq 1$ we write

$$\check{\mathbb{E}}_{p,x}^{(n)}\left[\varphi\left(\check{X}_{p,n}, \ldots, \check{X}_{n,n}\right)\right] := \check{\mathbb{E}}_\eta^{(n)}\left[\varphi\left(\check{X}_{p,n}, \ldots, \check{X}_{n,n}\right)\big| \check{X}_{p,n} = x\right].$$

The S-kernels and the corresponding expectations are of interest due to the following change-of-measure identity.

Lemma 2. *Assume (H1)-(H2). For any $n \geq 1$, $0 \leq p < n$, a suitable test function φ and $x \in \mathsf{X}$,*

$$\frac{\mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1} G_q(X_q)\varphi(X_p, \ldots, X_n)\right]}{\mathbb{E}_{p,\eta_p}\left[\prod_{q=p}^{n-1} G_q(X_q)\right]} = h_{p,n}(x)\check{\mathbb{E}}_{p,x}^{(n)}\left[\varphi\left(\check{X}_{p,n}, \ldots, \check{X}_{n,n}\right)\right].$$

Proof. From Definitions 1 and 2,

$$\begin{aligned}
&\frac{\mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1} G_q(X_q)\varphi(X_p, \ldots, X_n)\right]}{\eta_p Q_{p,n}(1)} \\
&= h_{p,n}(x)\mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1}\frac{G_q(X_q)}{\lambda_q}\frac{h_{q+1,n}(X_{q+1})}{h_{q,n}(X_q)}\varphi(X_0, \ldots, X_n)\frac{1}{h_{n,n}(X_n)}\right] \\
&= h_{p,n}(x)\check{\mathbb{E}}_{p,x}^{(n)}\left[\varphi\left(\check{X}_{p,n}, \ldots, \check{X}_{n,n}\right)\right].
\end{aligned}$$

□

Remark 3. The S-kernels have previously been identified as playing a key role when analyzing stability properties of Feynman-Kac formulae and particle systems, see [Del Moral and Guionnet, 2001], albeit written in a slightly different form. From Definition 1 we have immediately that

$$S_{p,n}(x, A) = \frac{Q_p(\mathbb{I}_A h_{p,n})(x)}{\lambda_{p-1} h_{p-1,n}(x)} = \frac{Q_p(\mathbb{I}_A Q_{p,n}(1))(x)}{Q_{p-1,n}(1)(x)},$$

and it is in the latter form that these kernel are usually considered. However, in the context of the Lyapunov drift techniques employed here, (4.8) expressed in terms of the λ-values and h-functions plays a central role in proofs of the two following propositions. The main theme of the proof of Proposition 1 is to obtain uniform bounds on $\|\eta_n\|_v$ via the representation of Lemma 2, the identity $\check{\mathbb{E}}_{p,x}^{(n)}\left[v\left(\check{X}_{n,n}\right)\right] = S_{p+1,n} \cdots S_{n,n}(v)(x)$ and the drift functions $(v_{p,n})$.

Note that Proposition 1 does not require the majorization-type assumption (H4).

Proposition 1. *Assume (H1)-(H3) and let v be as therein. Then there exists a finite constant c_μ depending on μ and the quantities in (H1) and (H3), such that*

$$\sup_{n\geq 0} \|\eta_n\|_v \leq c_\mu \mu(v).$$

Proof. See section 5. □

The interest in the uniform bound of Proposition 1 is that, via the following proposition, we obtain some uniform bounds on the λ-values and h-functions.

Proposition 2. *Assume (H1)-(H3) and let v be as therein. Then 1)-2) below are equivalent.*

1) $\sup_{n\geq 0} \|\eta_n\|_v < \infty$

2) $\inf_{n\geq 0} \lambda_n > 0$

If additionally (H4) holds then *1) and 2) are equivalent to 3):*

3) $\sup_{n\geq 1} \sup_{0\leq p\leq n} \|h_{p,n}\|_v < \infty$

Proof. Lemmata 7, 8 and 9. See section 5. □

Before proceeding further, note that in the results from this point on, the statements often feature a constant c_μ. The value of this constant may change from one result to the next.

4.3 A multiplicative stability theorem

The form of the following result can be interpreted as a non-homogeneous analogue of the multiplicative ergodic theorem of Kontoyiannis and Meyn [2005] in the context of positive operators, for direct comparison the reader is referred to [Whiteley et al., 2011, Theorem 2.2, equation (2.10)]. This proposition will be applied in section 4.5 to bound the asymptotic variance associated with $\left(\eta_n^N\right)$. The proof is postponed.

Theorem 1. *Assume (H1)-(H4). Then there exists $\rho < 1$ depending only on μ and the constants in (H1), (H3) and (H4) and $c_\mu < \infty$ depending on the quantities in (H1)-(H4) such that for any $\varphi \in \mathcal{L}_v$, $n \geq 1$ and $0 \leq p < n$,*

$$\left| \frac{Q_{p,n}(\varphi)(x)}{\prod_{q=p}^{n-1} \lambda_q} - h_{p,n}(x)\eta_n(\varphi) \right| \leq \rho^{n-p} \|\varphi\|_v c_\mu v(x) \mu(v), \quad \forall x \in \mathsf{X}.$$

Proof. See section 5. ☐

As a consequence of this theorem we obtain v-norm exponential stability with respect to initial condition for measures (η_n).

Corollary 1. *Assume (H1)-(H4), then with ρ and μ as in Theorem 1, for any $\mu' \in \mathcal{P}_v$, there exists $c_{\mu,\mu'} < \infty$ such that*

$$\left\| \eta_n^{(\mu)} - \eta_n^{(\mu')} \right\|_v \leq \rho^n c_{\mu,\mu'} \mu(v) \mu'(v),$$

where $\eta_n^{(\mu)} := \dfrac{\mu Q_{0,n}}{\mu Q_{0,n}(1)}$ *and* $\eta_n^{(\mu')} := \dfrac{\mu' Q_{0,n}}{\mu' Q_{0,n}(1)}$.

Proof. Taking the bound of Theorem 1 and integrating w.r.t. μ' gives

$$\left| \frac{\mu' Q_{0,n}(\varphi)}{\prod_{p=0}^{n-1} \lambda_p} - \mu'(h_{0,n}) \eta_n^{(\mu)}(\varphi) \right| \leq \rho^n \|\varphi\|_v c_\mu \mu(v) \mu'(v).$$

It is stressed that in the above display λ_p and $h_{0,n}$ are as in definition 1, i.e. dependent on μ, but not on μ'. Now as $\mu' \in \mathcal{P}_v$, for any $d \in [\underline{d}, \infty)$, $\mu'(C_d^c) \leq \mu'\left(\mathbb{I}_{C_d^c} e^V\right)/e^d \leq \mu\left(e^V\right)/e^d$ so there exists $d \in [\underline{d}, \infty)$ such that $\mu'(C_d) > 0$. Then dividing through by $\mu'(h_{0,n}) = \mu' Q_{0,n}(1)/\prod_{p=0}^{n-1} \lambda_p$,

$$\begin{aligned}
\left| \frac{\mu' Q_{0,n}(\varphi)}{\mu' Q_{0,n}(1)} - \eta_n^{(\mu)}(\varphi) \right| &\leq \rho^n \|\varphi\|_v \frac{c_\mu}{\mu'(h_{0,n})} \mu(v) \mu'(v) \\
&\leq \rho^n \|\varphi\|_v \frac{c_\mu}{\mu'(C_d) \inf_{x \in C_d} h_{0,n}(x)} \mu(v) \mu'(v) \\
&\leq \rho^n \|\varphi\|_v c_{\mu,\mu'} \mu(v) \mu'(v),
\end{aligned}$$

where the final inequality holds due to Lemma 10. ☐

4.4 Exponential moments for additive functionals

We now present a result on finite exponential moments for a class of additive, possibly unbounded path space functionals. It will be applied in section 4.5 to bounds on the relative variance associated with $\gamma_n^N(1)$. The proof is mostly technical and is given in the appendix.

Theorem 2. *Assume (H1)-(H4) and let δ and v be as therein. Then there exists a finite constant c_μ depending on μ and the quantities in (H1)-(H4) such that for any collection of measurable functions $\{F_n; n \geq 1\}$ with each $F_n : \mathsf{X} \to \mathbb{R}$ and $\sup_x (|F_n(x)| - \delta V(x)) < \infty$; any $n \geq 1$,*

$0 \le s \le n+1$, *and* $(i_1, \ldots, i_s) \in \mathcal{I}_{n,s}$,

$$\frac{\mathbb{E}_\mu \left[\prod_{p=0}^{n-1} G_p(X_p) \exp\left(\sum_{k \in \{i_1,\ldots,i_s\}} |F_k(X_k)| \right) \right]}{\mathbb{E}_\mu \left[\prod_{p=0}^{n-1} G_p(X_p) \right]} \le c_\mu^s \mu(v) \prod_{k \in \{i_1,\ldots,i_s\}} \left\| e^{|F_k|} \right\|_{v^\delta},$$

with the conventions that, when $s = 0$, the summation on the left hand side is zero and the product on the right hand side is unity.

Proof. See section 5. □

4.5 Variance bounds

Remark 4. At this point we introduce a further assumption, (H5) below. This assumption is not necessary for all of the results of this section but is employed for the following three reasons: 1) it is not restrictive in filtering applications; 2) it allows Lemma 3 below to be invoked (an equivalent result can also be obtained without (H5), but subject to constraints on the growth rates of (G_n) and the assumption that the Markov kernels (M_n) themselves obey a suitable simultaneous multiplicative drift condition); and 3) it allows an existing CLT for particle systems to be simply stated below without proof (see also remark 6).

(H5) $\sup_{n\ge0} \sup_{x\in\mathsf{X}} G_n(x) < \infty$

The following lemma plays an important technical role in the variance results which follow.

Lemma 3. *Assume (H1)-(H5) with v the drift function in (H1)–(H4). Then for any $\alpha \in (0,1)$, the statements of (H1)-(H4) also hold for the drift function $v_1 := v^\alpha$ and with α-dependent constants.*

Proof. Let $\bar{G} := \sup_{n\ge0} \sup_{x\in\mathsf{X}} G_n(x)$. Then for all $x \in \mathsf{X}$, and any $d \in [\underline{d}, \infty)$ as in (H1),

$$\begin{aligned}
\sup_{n\ge1} Q_n\left(e^{\alpha V}\right)(x) &\le \bar{G} \sup_{n\ge1} \left[\frac{G_{n-1}(x)}{\bar{G}} M_n\left(e^V\right)(x) \right]^\alpha \\
&= \bar{G}^{1-\alpha} \sup_{n\ge1} \left[Q_n\left(e^V\right)(x) \right]^\alpha \\
&\le \exp\left[\alpha V(x)(1-\delta) + \alpha b_d \mathbb{I}_{C_d}(x) + (1-\alpha)\log \bar{G} \right],
\end{aligned}$$

where Jensen's inequality and (H1) have been applied, and δ, b_d and $C_d = \{x \in \mathsf{X}; V(x) \le d\}$ are as in (H1). Then for any $\delta_0 \in (0, \delta)$ and $\bar{G} < \infty$ there exists $\underline{d}_\alpha \in [\underline{d}, \infty)$ such that for any $d \in [\underline{d}_\alpha, \infty)$ and $x \notin \{x \in \mathsf{X}; \alpha V(x) \le d\}$,

$$\begin{aligned}
\sup_{n\ge1} Q_n\left(e^{\alpha V}\right)(x) &\le \exp\left[\alpha V(x)(1-\delta_0) - \alpha d(\delta-\delta_0) + (1-\alpha)\log \bar{G} \right] \\
&\le \exp\left[\alpha V(x)(1-\delta_0) \right], \qquad (4.9)
\end{aligned}$$

and for $x \in \{x \in \mathsf{X}; \alpha V(x) \leq d\}$,

$$\begin{aligned}\sup_{n\geq 1} Q_n\left(e^{\alpha V}\right)(x) &\leq \exp\left[\alpha d(1-\delta)+\alpha b_d+(1-\alpha)\log \bar{G}\right] \\ &=: \exp\left(b_{d,\alpha}\right). \end{aligned} \tag{4.10}$$

The statement of (H1) holds with the drift function $v_1 := v^{\alpha}$ because equations (4.9)-(4.10) show that we may replace $\underline{d}$, δ, b_d, C_d in the corresponding statements with $\underline{d}_{\alpha}$, δ_0, $b_{d,\alpha}$,$\{x \in \mathsf{X}; \alpha V(x) \leq d\}$, respectively.

It is immediate that (H2) holds for v^{α} , because $v \geq 1$. (H3)-(H4) also hold for v^{α}, by replacing $\underline{d}$, C_d,ϵ_d^-, ϵ_d^+, ν_d with $\underline{d}_{\alpha}$, $\{x \in \mathsf{X}; \alpha V(x) \leq d\}$, $\epsilon_{d/\alpha}^-$, $\epsilon_{d/\alpha}^+$, $\nu_{d/\alpha}$, respectively. □

4.5.1 Asymptotic variance for η_n^N

Remark 5. There are several existing CLT results for the particle systems in question, see for example [Chopin, 2004, Douc and Moulines, 2008]. We choose to present that of Del Moral [2004, Proposition 9.4.2], as it holds immediately under (H5), and we may state also the corresponding asymptotic variance expression with essentially no further work. The restriction is that the stated result holds only for bounded functions. It is of interest to investigate whether the same result holds for a suitable class of possibly unbounded functions in terms of v, for example via the techniques of Chopin [2004] or Douc and Moulines [2008], but this is beyond the scope of the present article.

The following CLT holds for errors associated with the particle approximation measures $\left(\eta_n^N\right)$. Straightforward manipulations of the asymptotic variance expression of [Del Moral, 2004, Proposition 9.4.2] show that it can be written as in (4.11) below, in terms of the h-functions and λ-values.

Theorem 3. *[Del Moral, 2004, Proposition 9.4.2]. Assume (H5). Then for $\varphi : \mathsf{X} \to \mathbb{R}$ bounded and measurable and any $n \geq 1$,*

$$\sqrt{N}\left(\eta_n^N - \eta_n\right)(\varphi) \quad\to\quad \mathcal{N}\left(0, \sigma_n^2\right)$$

in distribution as $N \to \infty$, where

$$\sigma_n^2 := \eta_n\left[\left(\varphi - \eta_n(\varphi)\right)^2\right] + \sum_{p=0}^{n-1} \eta_p\left[\left(\frac{Q_{p,n}(\varphi)}{\prod_{q=p}^{n-1}\lambda_q} - h_{p,n}\eta_n(\varphi)\right)^2\right]. \tag{4.11}$$

We can readily apply the result of Theorem 1 to obtain a time-uniform bound on the asymptotic variance.

Proposition 3. *Assume (H1)-(H5). Then there exists $c_\mu < \infty$ depending only on μ and the quantities in (H1)-(H5) such that for any $n \geq 1$,*

$$\sigma_n^2 \leq \eta_n\left[\left(\varphi - \eta_n(\varphi)\right)^2\right] + c_\mu \left\|\varphi\right\|_1^2 \mu(v)^2,$$

where v is as in (H1) and φ and σ_n^2 are as in Theorem 3.

Proof. As (H1)-(H4) are assumed to hold with some drift function v , then by Lemma 3, the same assumptions hold with the drift function $v^{1/2}$ and suitable constants. Then applying Theorem 1 (using the drift $v^{1/2}$ and the corresponding instances (H1)-(H4)), and then Proposition 1 (using the drift v), we find that there is $c_\mu < \infty$ such that

$$\begin{aligned}\eta_p\left[\left(\frac{Q_{p,n}(\varphi)}{\prod_{q=p}^{n-1}\lambda_q} - h_{p,n}\eta_n(\varphi)\right)^2\right] &\leq \rho^{2(n-p)}c_\mu \|\varphi\|_1^2 \mu\left(v^{1/2}\right)^2 \eta_p(v) \\ &\leq \rho^{2(n-p)}\|\varphi\|_1^2 c_\mu \mu(v)^2,\end{aligned}$$

and the statement of the Theorem follows by summing. □

4.5.2 Non-asymptotic variance for $\gamma_n^N(1)$

For $n \geq 1$ and $1 \leq s \leq n+1$, define

$$\Upsilon_n^{(i_1,\dots i_s)} := \frac{\mu Q_{0,i_1}(1)\mathbb{E}_\mu\left[\prod_{p=0}^{n-1} G_p(X_p)\prod_{j=1}^{s} Q_{i_j,i_{j+1}}(1)\left(X_{i_j}\right)\right]}{\left[\gamma_n(1)\right]^2},$$

with the convention that $i_{s+1} = n$.

Building from Del Moral et al. [2009], Cérou et al. [2011] obtained a non-asymptotic functional expansion of the relative variance associated with $\gamma_n^N(1)$. Elementary manipulations of this relative variance show that it may be written in terms of the quantities $\left(\Upsilon_n^{(i_1,\dots i_s)}\right)$ as follows, and as we assume (H5), the quantities involved are well defined (although this is not a necessary condition, one may alternatively assume (H1)-(H2)).

Theorem 4. [Cérou et al., 2011, Proposition 3.4]Assume *(H5).* Then for any $n \geq 1$,

$$\begin{aligned}&\mathbb{E}_\mu\left[\left(\frac{\gamma_n^N(1)}{\gamma_n(1)} - 1\right)^2\right] \\ &= \sum_{s=1}^{n+1}\left(1 - \frac{1}{N}\right)^{(n+1)-s}\frac{1}{N^s}\sum_{(i_1,\dots,i_s)\in\mathcal{I}_{n,s}}\left[\Upsilon_n^{(i_1,\dots i_s)} - 1\right],\end{aligned}$$

where the expectation is with respect to the law of the N-particle system initialized from μ.

We may now apply Theorem 2 in order to obtain the following linear-in-n bound on the relative variance.

Proposition 4. *Assume (H1)-(H5) and let v be as therein. Then there exists a finite constant c_μ depending on μ and the quantities in (H1)-(H5) such that for any $n \geq 1$,*

$$N > c_\mu(n+1) \quad \Longrightarrow \quad \mathbb{E}_\mu\left[\left(\frac{\gamma_n^N(1)}{\gamma_n(1)} - 1\right)^2\right] \leq c_\mu\frac{4}{N}(n+1)\mu(v)^2.$$

Proof. Throughout the proof c is a finite constant depending on μ and the quantities in (H1)-(H5) whose value may change on each appearance.

First notice that by definition 1 and the product formula (4.5) we may write

$$\begin{aligned}\Upsilon_n^{(i_1,\ldots i_s)} &= \frac{\mu Q_{0,i_1}(1)}{\mu Q_{0,i_1}(1)}\frac{1}{\gamma_n(1)}\mathbb{E}_\mu\left[\prod_{p=0}^{n-1}G_p(X_p)\prod_{j=1}^{s}\left(\frac{Q_{i_j,i_{j+1}}(1)\left(X_{i_j}\right)}{\prod_{k=i_j}^{i_{j+1}-1}\lambda_k}\right)\right] \\ &= \frac{1}{\gamma_n(1)}\mathbb{E}_\mu\left[\prod_{p=0}^{n-1}G_p(X_p)\prod_{j=1}^{s}h_{i_j,i_{j+1}}\left(X_{i_j}\right)\right], \end{aligned} \tag{4.12}$$

with the convention that $\prod_n^{n-1}=1$ in the first equality to deal with the case $i_s=n$.

Let v and δ be as in (H1). Then by Lemma 3, the statements of (H1)-(H4) also hold for the drift function v^δ and with constants which depend on δ. Then Proposition 1 and Proposition 2 both applied with the drift function v^δ and the corresponding instances of (H1)-(H4) of show that

$$\sup_{n\geq 1}\sup_{0\leq p\leq n}\|h_{p,n}\|_{v^\delta}<\infty,$$

so that, using the representation (4.12), and applying Theorem 2 with the drift function v and the corresponding instances of (H1)-(H4), there exists a finite constant c such that

$$\begin{aligned}\Upsilon_n^{(i_1,\ldots i_s)} &\leq c^s\frac{1}{\gamma_n(1)}\mathbb{E}_\mu\left[\prod_{p=0}^{n-1}G_p(X_p)\prod_{j=1}^{s}v^\delta\left(X_{i_j}\right)\right] \\ &\leq c^s\mu(v).\end{aligned}$$

Therefore by Theorem 4,

$$\mathbb{E}_\mu\left[\left(\frac{\gamma_n^N(1)}{\gamma_n(1)}-1\right)^2\right]\leq\mu(v)^2\sum_{s=1}^{n+1}\left(1-\frac{1}{N}\right)^{(n+1)-s}\frac{1}{N^s}\sum_{(i_1,\ldots,i_s)\in\mathcal{I}_{n,s}}c^s$$

The remainder of the proof then follows by the same arguments as [Cérou et al., 2011, Proofs of Theorem 5.1 and Corollary 5.2], so the details are omitted. □

5 Proofs and auxiliary results

Auxiliary result for section 3.1

Lemma 4. *When f is the transition kernel corresponding to the model of equations (3.7)-(3.9), there exists $\underline{d}<\infty$ and $\delta>0$ such that, for any $d\in[\underline{d},\infty)$ there exists $b_d<\infty$ and*

$$\int_{\mathsf{X}}f\left(x,dx'\right)v(x')\leq v(x)^{1-\delta}\exp\left[b_d\mathbb{I}_{C_d}(x)\right],\quad x\in\mathsf{X}, \tag{5.1}$$

where $v(x):=\exp(1+c|x|)$ for c a positive constant, and furthermore for each such d there exists $0<\epsilon_d^-<\epsilon_d^+<\infty$ such that

$$\epsilon_d^-\nu_d(A\cap C_d)\leq f(x,A\cap C_d)\leq\epsilon_d^+\nu_d(A\cap C_d),\quad x\in C_d, A\in\mathcal{B}(\mathsf{X}), \tag{5.2}$$

with ν_d the normalized restriction of Lebesgue measure to C_d. Furthermore $\int_{C_d} f(x, dx') > 0$, $\forall x \in \mathsf{X}$.

Proof. As per Douc et al. [2009], under the assumptions on the model, there exists $\beta < \infty$ such that

$$\begin{aligned} \frac{\int_{\mathsf{X}} f\left(x, dx'\right) v(x')}{v(x)} &\leq \beta \exp\left[c\left(|x + B(x)| - |x|\right)\right] \\ &= \beta \exp\left[-c\,|x|\left(1 - \frac{|x + B(x)|}{|x|}\right)\right], \end{aligned}$$

and then using (3.8), there exists $\delta_1 > 0$ such that for $|x|$ sufficiently large,

$$\left(1 - \frac{|x + B(x)|}{|x|}\right) \geq \delta_1,$$

so for such $|x|$ and $\delta \in (0, \delta_1)$,

$$\frac{\int_{\mathsf{X}} f\left(x, dx'\right) v(x')}{v(x)} \leq \exp\left[-V(x)\delta - c\,|x|\left(\delta_1 - \delta\right) + \log\beta + 1\right],$$

and by increasing $|x|$ further if necessary, we conclude that the result holds with $b_d := d + \log\beta$. (5.2) and $\int_{C_d} f(x, dx') > 0$ hold immediately. ☐

Proofs and results for section 4.2

The proof of Proposition 1 is given after Lemma 5 and Lemma 6.

Lemma 5. *Assume (H1)-(H3). Then for any $d \in [\underline{d}, \infty)$, any $n \geq 1$ and $1 \leq p \leq n$, the following inequalities hold,*

$$S_{p,n}\left(v_{p,n}\right) \leq \rho_{p,n} v_{p-1,n} + B_{p,n}\mathbb{I}_{C_d}, \tag{5.3}$$

where

$$\rho_{p,n} := \frac{e^{-\delta d}}{\lambda_{p-1}} \frac{\left\|h_{p,n}\right\|_v}{\left\|h_{p-1,n}\right\|_v} < \infty \tag{5.4}$$

$$B_{p,n} := \frac{e^{d(1-\delta)+b_d}}{\epsilon_d^-}\left\|h_{p,n}\right\|_v \frac{1}{\nu_d\left(\mathbb{I}_{C_d} h_{p,n}\right)} < \infty, \tag{5.5}$$

and with the dependence of $\rho_{p,n}$ and $B_{p,n}$ on d suppressed from the notation.

Proof. For $x \notin C_d$,

$$\begin{aligned} S_{p,n}\left(v_{p,n}\right)(x) &= \frac{Q_p\left(v\right)(x)}{\lambda_{p-1} h_{p-1,n}(x)}\left\|h_{p,n}\right\|_v \\ &\leq \frac{v(x)}{\lambda_{p-1} h_{p-1,n}(x)} e^{-\delta d}\left\|h_{p,n}\right\|_v \\ &= v_{p,n-1}(x)\frac{e^{-\delta d}}{\lambda_{p-1}}\frac{\left\|h_{p,n}\right\|_v}{\left\|h_{p-1,n}\right\|_v}, \end{aligned}$$

where (H1) has been applied.

For $x \in C_d$, from Lemma 1 and (H3),

$$\lambda_{p-1} h_{p-1,n}(x) = Q_p\left(h_{p,n}\right)(x) \geq \epsilon_d^- \nu_d\left(\mathbb{I}_{C_d} h_{p,n}\right),$$

and thus using (H1),

$$\begin{aligned} S_{p,n}\left(v_{p,n}\right)(x) &\leq e^{d(1-\delta)+b_d} \frac{\left\|h_{p,n}\right\|_v}{\lambda_{p-1}} \frac{1}{h_{p-1,n}(x)} \\ &\leq \frac{e^{d(1-\delta)+b_d}}{\epsilon_d^-} \left\|h_{p,n}\right\|_v \frac{1}{\nu_d\left(\mathbb{I}_{C_d} h_{p,n}\right)}. \end{aligned}$$

We have $\rho_{p,n} < \infty$ and $B_{p,n} < \infty$ because for any $p \leq n$, $\lambda_{p-1} > 0$, $h_{p,n} \in \mathcal{L}_v$, $h_{p,n}(x) > 0$ for all $x \in \mathsf{X}$, and for any $d \geq \underline{d}$, $\nu_d\left(C_d\right) > 0$. ☐

Lemma 6. *Assume (H1)-(H3). Then for any $d \in [\underline{d}, \infty)$, $0 \leq p < q \leq n$, and $x \in \mathsf{X}$,*

$$\begin{aligned} &\check{\mathbb{E}}_{p,x}^{(n)}\left[v_{q,n}\left(\check{X}_{q,n}\right)\right] \\ &\leq \frac{e^{-\delta d(q-p)}}{\prod_{k=p}^{q-1} \lambda_k} \frac{\left\|h_{q,n}\right\|_v}{\left\|h_{p,n}\right\|_v} v_{p,n}(x) \\ &+\frac{e^{d(1-\delta)+b_d}}{\epsilon_d^-} \left\|h_{q,n}\right\|_v \left[\frac{1}{\nu_d\left(\mathbb{I}_{C_d} h_{q,n}\right)} + \sum_{k=p+1}^{q-1} \frac{e^{-\delta d(q-k)}}{\prod_{j=k}^{q-1} \lambda_j} \frac{1}{\nu_d\left(\mathbb{I}_{C_d} h_{k,n}\right)}\right], \end{aligned} \tag{5.6}$$

with the convention that the sum is zero when $p = q-1$.

Proof. For each n, p and q in the specified ranges, the proof begins by recursive application of the drift inequalities of Lemma 5. A simple induction yields

$$\begin{aligned} &\check{\mathbb{E}}_{p,x}^{(n)}\left[v_{q,n}\left(\check{X}_{q,n}\right)\right] \\ &\leq \left(\prod_{k=p+1}^{q} \rho_{k,n}\right) v_{p,n}(x) + \sum_{k=p+1}^{q} \left(\prod_{j=k+1}^{q} \rho_{j,n}\right) B_{k,n}, \end{aligned} \tag{5.7}$$

with the convention that the right-most product is equal to 1 when $p = q-1$.

By the definitions of $(h_{p,n})$, $(\rho_{p,n})$ and $(B_{p,n})$,

$$\begin{aligned} \prod_{k=p+1}^{q} \rho_{k,n} &= \prod_{k=p+1}^{q} \frac{e^{-\delta d}}{\lambda_{k-1}} \frac{\left\|h_{k,n}\right\|_v}{\left\|h_{k-1,n}\right\|_v} \\ &= \frac{e^{-\delta d(q-p)}}{\prod_{k=p+1}^{q} \lambda_{k-1}} \frac{\left\|h_{q,n}\right\|_v}{\left\|h_{p,n}\right\|_v}, \end{aligned} \tag{5.8}$$

and for $k < q$,

$$\begin{aligned}
&\left(\prod_{j=k+1}^{q} \rho_{j,n}\right) B_{k,n} \\
&= \left(\frac{e^{-\delta d(q-k)}}{\prod_{j=k+1}^{q} \lambda_{j-1}} \frac{\|h_{q,n}\|_v}{\|h_{k,n}\|_v}\right) \frac{e^{d(1-\delta)+b_d}}{\epsilon_d^-} \|h_{k,n}\|_v \frac{1}{\nu_d\left(\mathbb{I}_{C_d} h_{k,n}\right)} \\
&= \frac{e^{d(1-\delta)+b_d}}{\epsilon_d^-} \|h_{q,n}\|_v \frac{e^{-\delta d(q-k)}}{\prod_{j=k}^{q-1} \lambda_j} \frac{1}{\nu_d\left(\mathbb{I}_{C_d} h_{k,n}\right)}. \qquad (5.9)
\end{aligned}$$

The proof is complete upon combining (5.7), (5.8), (5.9) and applying the definition of $B_{q,n}$ for the case $q = k$. ☐

Proof. (Proposition 1). For $n = 0$ we have trivially $\eta_0(v) = \mu(v)$.

For $n \geq 1$, by Lemma 2,

$$\begin{aligned}
\eta_n(v) &= \frac{\mathbb{E}_\mu\left[\prod_{q=0}^{n-1} G_q(X_q)\, v(X_n)\right]}{\mathbb{E}_\mu\left[\prod_{q=0}^{n-1} G_q(X_q)\right]}. \\
&= \int \mu(dx)\, h_{0,n}(x) \check{\mathbb{E}}_x^{(n)}\left[v\left(\check{X}_{n,n}\right)\right] \\
&\leq \int \mu(dx)\, h_{0,n}(x) \check{\mathbb{E}}_x^{(n)}\left[v_{n,n}\left(\check{X}_{n,n}\right)\right], \qquad (5.10)
\end{aligned}$$

where the inequality is due to $h_{n,n} = 1$ and $\|h_{n,n}\|_v \leq 1$. The proof proceeds by bounding the expectation.

Fix $d \in [\underline{d}, \infty)$ arbitrarily. Applying Lemma 6 with $q = n$ and $p = 0$, and again noting $h_{n,n} = 1$, $\|h_{n,n}\|_v \leq 1$, we obtain

$$\begin{aligned}
\check{\mathbb{E}}_x^{(n)}\left[v_{n,n}\left(\check{X}_{n,n}\right)\right] &\leq \frac{e^{-\delta d n}}{\prod_{k=0}^{n-1} \lambda_k} \frac{1}{\|h_{0,n}\|_v} v_{0,n}(x) \\
&\quad + \frac{e^{d(1-\delta)+b_d}}{\epsilon_d^-}\left[\frac{1}{\nu_d(C_d)} + \sum_{k=1}^{n-1} \frac{e^{-\delta d(n-k)}}{\prod_{j=k}^{n-1} \lambda_j} \frac{1}{\nu_d\left(\mathbb{I}_{C_d} h_{k,n}\right)}\right], \\
&= \frac{e^{-\delta d n}}{\mu Q_{0,n}(1)} \frac{1}{\|h_{0,n}\|_v} v_{0,n}(x) \\
&\quad + \frac{e^{d(1-\delta)+b_d}}{\epsilon_d^-}\left[\frac{1}{\nu_d(C_d)} + \sum_{k=1}^{n-1} \frac{e^{-\delta d(n-k)}}{\nu_d\left[\mathbb{I}_{C_d} Q_{k,n}(1)\right]}\right], \qquad (5.11)
\end{aligned}$$

with the convention (as per Lemma 6), that the summation is equal to zero when $n = 1$. The equality is due to the definitions of the λ-values and h-functions.

We now obtain lower bounds in order to treat the $\mu Q_{0,n}(1)$ and $\nu_d\left[\mathbb{I}_{C_d} Q_{k,n}(1)\right]$ terms. Recall that $d \in [\underline{d}, \infty)$ was arbitrary. Now choose arbitrarily $r \in [\underline{d}, d]$. Then under (H3), for any $\eta \in \mathcal{P}_v$

and any $0 \leq k < n$,

$$\begin{aligned}
\eta\left[\mathbb{I}_{C_d} Q_{k,n}(1)\right] &= \mathbb{E}_{k,\eta}\left[\mathbb{I}_{C_d}(X_k)\prod_{q=k}^{n-1} G_q(X_q)\right] \\
&\geq \mathbb{E}_{k,\eta}\left[\mathbb{I}_{C_r}(X_k)\prod_{q=k}^{n-1} G_q(X_q)\,\mathbb{I}_{C_r}(X_q)\,\mathbb{I}_{C_r}(X_n)\right] \\
&\geq \eta(C_r)\left[\epsilon_r^- \nu_r(C_r)\right]^{n-k}. \qquad (5.12)
\end{aligned}$$

Under (H2), for r and d increased if necessary, but still subject to $r \leq d$, we have $\mu(C_r) = 1 - \mu(C_r^c) \geq 1 - \mu\left(\mathbb{I}_{C_r^c} e^V\right) e^{-r} \geq 1 - \mu\left(e^V\right) e^{-r} > 0$. Now hold r constant and if necessary, increase d so that $e^{-\delta d} < \left[\epsilon_r^- \nu_r(C_r)\right]^{-1}$. Equation (5.12) then gives

$$\sup_{n\geq 1} \frac{e^{-\delta dn}}{\mu Q_{0,n}(1)} \leq \sup_{n\geq 1} \frac{e^{-\delta dn}}{\mu\left[\mathbb{I}_{C_d} Q_{0,n}(1)\right]} \leq \frac{1}{\mu(C_r)} < \infty. \qquad (5.13)$$

Then under (H1), noting $\nu_d(C_r) > 0$ and applying (5.12),

$$\begin{aligned}
&\sup_{n\geq 1}\left[\frac{1}{\nu_d(C_d)} + \sum_{k=1}^{n-1} \frac{e^{-\delta d(n-k)}}{\nu_d\left[\mathbb{I}_{C_d} Q_{k,n}(1)\right]}\right] \\
&\leq \frac{1}{\nu_d(C_d)} + \frac{1}{\nu_d(C_r)} \sup_{n\geq 1}\left[\sum_{k=1}^{n-1} \frac{e^{-\delta d(n-k)}}{\left[\epsilon_r^- \nu_r(C_r)\right]^{(n-k)}}\right] < \infty. \qquad (5.14)
\end{aligned}$$

Combining (5.13), (5.14) and (5.11), establishes that there exists a finite constant c_μ, independent of n such

$$\check{\mathbb{E}}_x^{(n)}\left[v_{n,n}\left(\check{X}_{n,n}\right)\right] \leq \frac{1}{\mu(C_r)} \frac{1}{\|h_{0,n}\|_v} v_{0,n}(x) + c_\mu.$$

and then returning to (5.10), we have shown that

$$\begin{aligned}
\eta_n(v) &\leq \frac{1}{\mu(C_r)} \frac{1}{\|h_{0,n}\|_v} \int h_{0,n}(x) v_{0,n}(x)\,\mu(\mathrm{d}x) \\
&\quad + c_\mu \int h_{0,n}(x)\mu(\mathrm{d}x) \\
&= \frac{\mu(v)}{\mu(C_r)} + c_\mu,
\end{aligned}$$

where the final equality uses the definition of $v_{0,n}$ and the property $\mu(h_{0,n}) = \eta(h_{0,n}) = 1$ as in Lemma 1. Thus there exists a finite constant c'_μ such that

$$\sup_{n\geq 1} \eta_n(v) \leq c'_\mu \mu(v),$$

which completes the proof. □

Lemma 7. *Assume (H1)-(H3) and let v be as therein. Then*

$$\sup_{n\geq 0}\|\eta_n\|_v<\infty \quad\Longleftrightarrow\quad \inf_{n\geq 0}\lambda_n>0. \tag{5.15}$$

Proof. ($\Rightarrow$). Suppose $\sup_{n\geq 0}\|\eta_n\|_v<\infty$. Then there exists a finite constant $\bar{\eta}$ such that for any $d\geq \underline{d}$,

$$\sup_{n\geq 0}\eta_n\left(C_d^c\right)\leq \sup_{n\geq 0}\frac{\eta_n\left(\mathbb{I}_{C_d^c}e^V\right)}{e^d}\leq \sup_{n\geq 0}\frac{\eta_n\left(e^V\right)}{e^d}\leq \bar{\eta}e^{-d}.$$

Thus for all $\beta<1$, there exists $d\geq \underline{d}$ such that $\sup_{n\geq 0}\eta_n\left(C_d^c\right)<\beta$. Thus for $\beta\in(0,1)$ there exists $r\geq \underline{d}$ such that

$$\inf_{n\geq 0}\lambda_n\geq \inf_{n\geq 0}\eta_n\left(\mathbb{I}_{C_r}Q_{n+1}\left(\mathbb{I}_{C_r}\right)\right)\geq \epsilon_r^-\nu_r\left(C_r\right)\inf_{n\geq 0}\eta_n\left(C_r\right)\geq \epsilon_r^-\nu_r\left(C_r\right)(1-\beta),$$

where the second inequality is due to (H3).

($\Leftarrow$). Suppose $\inf_{n\geq 0}\lambda_n>0$. Then there exists $\underline{\lambda}>0$ such that for any $n\geq 1$,

$$\eta_n\left(e^V\right)=\frac{\eta_{n-1}Q_n\left(e^V\right)}{\eta_{n-1}\left(G_{n-1}\right)}\leq \frac{\eta_{n-1}Q_n\left(e^V\right)}{\underline{\lambda}},$$

where (4.7) has been used. Now set $d>\underline{d}\vee\left(-\frac{1}{\delta}\log\underline{\lambda}\right)$. Then under (H1),

$$\begin{aligned}
\eta_n\left(e^V\right) &\leq \frac{\eta_{n-1}\left[\mathbb{I}_{C_d^c}Q_n\left(e^V\right)\right]}{\underline{\lambda}}+\frac{\eta_{n-1}\left[\mathbb{I}_{C_d}Q_n\left(e^V\right)\right]}{\underline{\lambda}}\\
&\leq \frac{e^{-\delta d}}{\underline{\lambda}}\eta_{n-1}\left(e^V\right)+\frac{e^{d(1-\delta)+b_d}}{\underline{\lambda}}\\
&=: \rho\eta_{n-1}\left(e^V\right)+B,
\end{aligned} \tag{5.16}$$

for some $\rho<1$ and $B<\infty$. Iteration of (5.16) establishes 1). □

Lemma 8. *Assume (H1)-(H4) and let v be as therein. Then*

$$\inf_{n\geq 0}\lambda_n>0 \qquad\Longrightarrow\qquad \sup_{n\geq 1}\sup_{0\leq p\leq n}\|h_{p,n}\|_v<\infty.$$

Proof. Recall the definition

$$h_{p,n}(x)=\frac{Q_{p.n}\left(1\right)(x)}{\eta_pQ_{p,n}\left(1\right)}. \tag{5.17}$$

For the case $p=n$, $h_{p,n}=1$. For other cases we proceed by decomposing and then bounding the numerator.

Set $d\in[\underline{d},\infty)$ arbitrarily, let $n\geq 1$, $0\leq p<n$ and define $\tau_p^{(d)}:=\inf\left\{q\geq p; X_q\in C_d, X_{q+1}\in C_d\right\}$.

Now consider the decomposition:

$$\begin{aligned} Q_{p,n}(1)(x) &= \sum_{k=p}^{n-1} \mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1} G_q(X_q)\,\mathbb{I}\left\{\tau_p^{(d)} = k\right\}\right] \\ &\quad + \mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1} G_q(X_q)\,\mathbb{I}\left\{\tau_p^{(d)} \geq n\right\}\right] \end{aligned} \tag{5.18}$$

and define

$$A_p := |||Q_p \mathbb{I}_{C_d^c}|||_v, \qquad B_p := |||Q_p \mathbb{I}_{C_d}|||_v,$$

$$\Xi_0 := v(X_p), \qquad \Xi_j := \left[\prod_{q=p}^{p+j-1} \frac{G_q(X_q)}{A_{q+1}^{\mathbb{I}_{C_d^c}(X_q)} B_{q+1}^{\mathbb{I}_{C_d}(X_q)}}\right] v(X_{p+j}), \quad 1 \leq j \leq n-p.$$

Assumption (H1) implies that, for $1 \leq j \leq n-p$, $\mathbb{E}_{p+j-1, X_{p+j-1}}[\Xi_j] \leq \Xi_{j-1}$, so that

$$\mathbb{E}_{p,x}[\Xi_{n-p}] \;\leq\; \mathbb{E}_{p,x}[\Xi_0] = v(x). \tag{5.19}$$

For $k > p$, define $M_{p,k}^{(d)} := \sum_{q=p}^{k-1} \mathbb{I}_{C_d^c}(X_q)$. Then the following bound holds under (H1):

$$\begin{aligned} &\left[\prod_{q=p}^{k-1} A_{q+1}^{\mathbb{I}_{C_d^c}(X_q)} B_{q+1}^{\mathbb{I}_{C_d}(X_q)}\right] \mathbb{I}\left\{M_{p,k}^{(d)} \geq (k-p)/2\right\} \\ &\leq \left(\sup_{q\geq 1} |||Q_q \mathbb{I}_{C_d^c}|||_v\right)^{M_d^{(p,k)}} \mathbb{I}\left\{M_{p,k}^{(d)} \geq (k-p)/2\right\} \left(1 \vee \sup_{q\geq 1} |||Q_q|||_v\right)^{(k-p)/2} \\ &\leq \exp\left[-\delta d (k-p)/2\right] \exp\left[b_{\underline{d}} (k-p)/2\right]. \end{aligned} \tag{5.20}$$

where $|||Q_q \mathbb{I}_{C_d}|||_v \leq |||Q_q|||_v$ has been used.

Consider one term from the summation in (5.18) with $p < k < n$. By [Douc et al., 2009, Lemma 17], $\mathbb{I}\left\{\tau_p^{(d)} \geq k\right\} = \mathbb{I}\left\{\sum_{q=p}^{k-1} \mathbb{I}_{C_d}(X_q)\mathbb{I}_{C_d}(X_{q+1}) = 0\right\} \leq \mathbb{I}\left\{M_{p,k}^{(d)} \geq (k-p)/2\right\}$. Then combining (5.19) and (5.20) and using (H4),

$$\begin{aligned} &\mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1} G_q(X_q)\,\mathbb{I}\left\{\tau_p^{(d)} = k\right\}\right] \\ &\leq \epsilon_d^+ \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right] \mathbb{E}_{p,x}\left[\prod_{q=p}^{k-1} G_q(X_q)\,\mathbb{I}\left\{M_{p,k}^{(d)} \geq (k-p)/2\right\} v(X_k)\right] \\ &\leq \epsilon_d^+ \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right] v(x) \\ &\quad \cdot \exp\left[-\delta d (k-p)/2\right] \exp\left[b_{\underline{d}} (k-p)/2\right], \quad k > p, \end{aligned} \tag{5.21}$$

similarly

$$
\begin{aligned}
&\mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1} G_q\left(X_q\right) \mathbb{I}\left[\tau_d^{(p)} \geq n\right]\right] \\
&\leq \mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1} G_q\left(X_q\right) \mathbb{I}\left\{M_{p,k}^{(d)} \geq (k-p)/2\right\} v\left(X_n\right)\right] \\
&\leq v(x) \exp\left[-\delta d\left(n-p\right)/2\right] \exp\left[b_{\underline{d}}\left(n-p\right)/2\right], \qquad (5.22)
\end{aligned}
$$

and also by (H4),

$$
\mathbb{E}_{p,x}\left[\prod_{q=p}^{n-1} G_q\left(X_q\right) \mathbb{I}\left[\tau_d^{(p)} = p\right]\right] \leq \epsilon_d^{+} \nu_d\left[\mathbb{I}_{C_d} Q_{p+1,n}(1)\right] v(x), \qquad (5.23)
$$

recalling from section 4.1 the convention $Q_{n+1,n} = Id$. Returning to (5.18), the bounds of (5.21)-(5.23) show that for $p < n$,

$$
\begin{aligned}
&Q_{p,n}(1)(x) \\
&\leq \epsilon_d^{+} v(x) \sum_{k=p}^{n-1} \exp\left[-\delta d\left(k-p\right)/2\right] \exp\left[b_{\underline{d}}\left(k-p\right)/2\right] \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right] \\
&\quad + v(x) \exp\left[-\delta d\left(n-p\right)/2\right] \exp\left[b_{\underline{d}}\left(n-p\right)/2\right]. \qquad (5.24)
\end{aligned}
$$

We now turn to the denominator of (5.17) and stress that we are continuing to use the same arbitrary value of d as above.

As per the statement of the Lemma, suppose $\underline{\lambda} := \inf_{n\geq 0} \lambda_n > 0$. Then by Lemma 7, $\bar{\eta} := \sup_{n\geq 0} \eta_n\left(e^V\right) < \infty$ and choosing independently $\epsilon \in (0,1)$, by (H1) d may then be chosen large enough that

$$
\inf_{n\geq 0} \eta_n\left(C_d\right) = \inf_{n\geq 0} 1 - \eta_n\left(C_d^c\right) \geq \inf_{n\geq 0} 1 - \eta_n\left(\mathbb{I}_{C_d^c} e^V\right) e^{-d} \geq 1 - \bar{\eta} e^{-d} \geq 1 - \epsilon =: \underline{\eta}.
$$

Then for $p < k < n$,

$$
\begin{aligned}
\eta_p Q_{p,n}(1) &= \left(\prod_{q=p}^{k-1} \lambda_q\right) \eta_k Q_{k,n}(1) \\
&\geq \underline{\lambda}^{(k-p)} \eta_k\left[\mathbb{I}_{C_d} Q_{k,n}(1)\right] \\
&\geq \epsilon_d^{-} \underline{\lambda}^{(k-p)} \eta_k\left(C_d\right) \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right] \\
&\geq \epsilon_d^{-} \underline{\lambda}^{(k-p)} \underline{\eta} \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right], \qquad (5.25)
\end{aligned}
$$

and for $p < n$,

$$
\eta_p Q_{p,n}(1) \geq \epsilon_d^{-} \eta_p\left(C_d\right) \nu_d\left[\mathbb{I}_{C_d} Q_{p+1,n}(1)\right] \geq \epsilon_d^{-} \underline{\eta} \nu_d\left[\mathbb{I}_{C_d} Q_{p+1,n}(1)\right], \qquad (5.26)
$$

and also

$$
\eta_p Q_{p,n}(1) \geq \underline{\lambda}^{(n-p)}. \qquad (5.27)
$$

Combining (5.25)-(5.27) with (5.24) and (5.17) we finally obtain, for $p < n$,

$$\begin{aligned} h_{p,n}(x) &\leq \frac{\epsilon_d^+}{\epsilon_d^- \underline{\eta}} v(x) \sum_{k=p}^{n-1} \exp\left[-\delta d\left(k-p\right)/2\right] \exp\left[(k-p)\left(b_{\underline{d}}/2 - \log \underline{\lambda}\right)\right] \\ &+ v(x) \exp\left[-\delta d\left(n-p\right)/2\right] \exp\left[(n-p)\left(b_{\underline{d}}/2 - \log \underline{\lambda}\right)\right]. \end{aligned}$$

Then increasing d further if necessary, we conclude that there exists $c < \infty$ such that for any $x \in \mathsf{X}$, $\sup_{n\geq 1} \sup_{0\leq p\leq n} h_{p,n}(x) \leq cv(x)$ and this completes the proof. □

Lemma 9. *Assume (H1)-(H3) and let v be as therein. Then*

$$\sup_{n\geq 1} \sup_{0\leq p\leq n} \|h_{p,n}\|_v < \infty \qquad \Longrightarrow \qquad \inf_{n\geq 0} \lambda_n > 0.$$

Proof. Suppose $\sup_{n\geq 1} \sup_{0\leq p\leq n} \|h_{p,n}\|_v < \infty$. Then by Lemma 1 and (H3), for any $x \in C_d$,

$$\begin{aligned} \inf_{n\geq 0} \lambda_n &= \inf_{n\geq 0} \frac{Q_{n+1}\left(h_{n+1,n+1}\right)(x)}{h_{n,n+1}(x)} \\ &\geq \inf_{n\geq 0} \frac{Q_{n+1}\left(C_d\right)(x)}{\|h_{n,n+1}\|_v v(x)} \\ &\geq \frac{\epsilon_d^-}{e^d} \frac{\nu_d\left(C_d\right)}{\sup_{n\geq 0} \|h_{n,n+1}\|_v} \\ &> 0. \end{aligned}$$

□

Proofs for section 4.3

The following Lemma will be used in the proofs of Theorems 1 and 2.

Lemma 10. *Assume (H1)-(H4) and let $\underline{d}$ be as therein. Then for any $d \in [\underline{d}, \infty)$,*

$$\inf_{n\geq 1} \inf_{0\leq p\leq n} \inf_{x\in C_d} h_{p,n}(x) > 0.$$

Proof. We will prove a finite, uniform upper bound on

$$\begin{aligned} \sup_{x\in C_d} \frac{1}{h_{p,n}(x)} &= \sup_{x\in C_d} \frac{\eta_p Q_{p,n}(1)}{Q_{p,n}(1)(x)} \\ &= \frac{\eta_p Q_{p,n}(1)}{\inf_{x\in C_d} Q_{p,n}(1)(x)}. \end{aligned} \tag{5.28}$$

The proof uses the same approach as in the proof of Lemma 9 and therefore some steps are omitted for brevity. For the case $p = n$, $\eta_n Q_{n,n}(1) = 1$ and $Q_{n,n}(1)(x) = 1$ for all x. For the remaining cases we proceed by considering the numerator of 5.28.

Set $d \in [\underline{d}, \infty)$ arbitrarily, let $n \geq 1$ and $p < n$ and define $\tau_p^{(d)} := \inf\{q \geq p; X_q \in C_d, X_{q+1} \in C_d\}$.

We have the decomposition

$$\eta_p Q_{p,n}(1) = \sum_{k=p}^{n-1} \mathbb{E}_{p,\eta_p}\left[\prod_{q=p}^{n-1} G_q(X_q)\,\mathbb{I}\left[\tau_p^{(d)} = k\right]\right] + \mathbb{E}_{p,\eta_p}\left[\prod_{q=p}^{n-1} G_q(X_q)\,\mathbb{I}\left[\tau_p^{(d)} \geq n\right]\right], \tag{5.29}$$

This is of exactly the same form as in equation (5.18) in the proof of Lemma 9, except for the initial measure η_p. Thus by exactly the same arguments (integrate equation (5.24) w.r.t. η_p) we obtain the bound

$$\begin{aligned}
&\eta_p Q_{p,n}(1) \\
&\leq \epsilon_d^{+} \eta_p(v) \sum_{k=p}^{n-1} \exp\left[-\delta d (k-p)/2\right] \exp\left[b_{\underline{d}} (k-p)/2\right] \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right] \\
&\quad + \eta_p(v) \exp\left[-\delta d(n-p)/2\right] \exp\left[b_{\underline{d}}(n-p)/2\right].
\end{aligned} \tag{5.30}$$

Now set $r \in [\underline{d}, d]$. For the denominator of (5.28) we have by (H3),

$$\begin{aligned}
\inf_{x \in C_d} Q_{p,n}(1)(x) &\geq \inf_{x \in C_d} Q_p\left[\mathbb{I}_{C_d} Q_{p+1,n}(1)\right](x) \\
&\geq \epsilon_d^{-} \nu_d\left[\mathbb{I}_{C_r} Q_{p+1,n}(1)\right],
\end{aligned} \tag{5.31}$$

also

$$\epsilon_d^{-} \nu_d\left[\mathbb{I}_{C_r} Q_{p+1,n}(1)\right] \geq \epsilon_d^{-} \nu_d(C_r)\left[\epsilon_r^{-} \nu_r(C_r)\right]^{n-p-1}$$

and for $p < k < n$,

$$\begin{aligned}
&\epsilon_d^{-} \nu_d\left[\mathbb{I}_{C_r} Q_{p+1,n}(1)\right] \\
&= \epsilon_d^{-} \mathbb{E}_{p+1,\nu_d}\left[\mathbb{I}_{C_r}(X_{p+1}) \prod_{q=p+1}^{n-1} G_q(X_q)\right] \\
&\geq \epsilon_d^{-} \mathbb{E}_{p+1,\nu_d}\left[\mathbb{I}_{C_r}(X_{p+1}) \prod_{q=p+1}^{n-1} G_q(X_q)\,\mathbb{I}_{C_d}(X_k)\,\mathbb{I}_{C_d}(X_{k+1})\right] \\
&\geq \epsilon_d^{-} \mathbb{E}_{p+1,\nu_d}\left[\mathbb{I}_{C_r}(X_{p+1}) \prod_{q=p+1}^{k-1} G_q(X_q)\right] \epsilon_d^{-} \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right] \\
&\geq \epsilon_d^{-} \mathbb{E}_{p+1,\nu_d}\left[\mathbb{I}_{C_r}(X_{p+1}) \prod_{q=p+1}^{k-1} G_q(X_q)\,\mathbb{I}_{C_r}(X_{q+1})\right] \epsilon_d^{-} \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right] \\
&\geq \epsilon_d^{-} \nu_d(C_r)\left[\epsilon_r^{-} \nu_r(C_r)\right]^{k-p-1} \epsilon_d^{-} \nu_d\left[\mathbb{I}_{C_d} Q_{k+1,n}(1)\right].
\end{aligned} \tag{5.32}$$

Combining (5.28), (5.30), (5.31) and (5.32) gives for $p < n$

$$\begin{aligned}
&\sup_{x\in C_d}\frac{1}{h_{p,n}(x)}\\
&\leq \frac{\epsilon_d^+}{\epsilon_d^-}\eta_p(v)\\
&+\frac{\epsilon_d^+}{\epsilon_d^-}\frac{\eta_p(v)}{\epsilon_d^-\nu_d(C_r)}\\
&\cdot\left(\sum_{k=p+1}^{n-1}\exp\left[-\delta d(k-p)/2\right]\exp\left[(k-p)b_{\underline{d}}/2+(k-p-1)\log\left[\epsilon_r^-\nu_r(C_r)\right]\right]\right)\\
&+\frac{1}{\epsilon_d^-}\frac{\eta_p(v)}{\nu_d(C_r)}\exp\left[-\delta d(n-p)/2\right]\exp\left[(n-p)\left(b_{\underline{d}}/2+\log\left[\epsilon_r^-\nu_r(C_r)\right]\right)\right],
\end{aligned}$$

with the convention that the summation is zero when $p = n-1$. With r kept fixed, increasing d and noting that under the assumptions of the lemma, Proposition 1 holds, we conclude that there exists a finite constant $c_\mu(d)$ such that

$$\sup_{n\geq 1}\sup_{0\leq p\leq n}\sup_{x\in C_d}\frac{1}{h_{p,n}(x)}\leq c_\mu(d).$$

The proof is complete because $d_1 \leq d_2 \Rightarrow C_{d_1} \subseteq C_{d_2}$. □

Proof. (Theorem 1). The proof is based directly on those of [Douc et al., 2009, Proposition 12 and Lemma 15], which are in turn developments from the decomposition ideas of Kleptsyna and Veretennikov [2008]. However, there are some crucial differences here: the focus of the present work is on the v-norm on measures, as opposed to total variation, and different techniques will be used to deal with and control denominator terms in equation (5.33) below, by way of Propositions 1 and 2.

Throughout the proof, c is a finite constant whose value depends on μ and the quantities in (H1)-(H4) and whose value may change on each appearance.

Let $\left(\bar{X}_n; n\geq 0\right)$ be the bi-variate Markov chain on X^2 with

$$\bar{X}_n\big|\left\{\bar{X}_{n-1}=\left(x_{n-1},x'_{n-1}\right)\right\}\quad\sim\quad M_n\left(x_{n-1},\cdot\right)\otimes M_n\left(x'_{n-1},\cdot\right)$$

and for some distribution H on X^2 we denote by $\bar{\mathbb{E}}_H$ the expectation with respect to the law of this bi-variate chain initialized by $\bar{X}_0 \sim H$. In line with previous definitions, for η a distribution on X we write $\bar{\mathbb{E}}_{p,\delta_x\otimes\eta} := \int\delta_x(\mathrm{d}x)\,\eta(\mathrm{d}x')\,\bar{\mathbb{E}}_H\left[\varphi\left(\bar{X}_p,\ldots,\bar{X}_n\right)\big|\left\{\bar{X}_p=(x,x')\right\}\right]$. Also define $\bar{C}_d := C_d\times C_d$ and throughout the following writing $\bar{x} = (x,x')$ for a point in X^2, define $\bar{G}_n(\bar{x}) := G_n(x)\,G_n(x')$ and $\bar{v}(\bar{x}) := v(x)\,v(x')$.

For each $n\geq 1$ define the tensor-product kernel $\bar{Q}_n(\bar{x},d\bar{y}) := Q_n(x,dy)\otimes Q_n(x',dy')$ and let $\left(\bar{Q}_{p,n}\right)$ be the semigroup defined in the same fashion as (4.1).

Now fix arbitrarily $d\in[\underline{d},\infty)$ and define, for $n\geq 1$,

$$\bar{R}_n(\bar{x},d\bar{y})\quad:=\quad\bar{Q}_n(\bar{x},d\bar{y})-\mathbb{I}_{\bar{C}_d}(\bar{x})\left(\epsilon_d^-\right)^2\nu_d\otimes\nu_d(d\bar{y})$$

and $\left(\bar{R}_{p,n}\right)$ in the same way. The dependence of $\bar{R}_n$ on d is suppressed from the notation.

First set $n \geq 1$ and $0 \leq p \leq n$ arbitrarily. We have from the above definitions,

$$
\begin{aligned}
&\left|\frac{Q_{p,n}(\varphi)(x)}{h_{p,n}(x)\prod_{q=p}^{n-1}\lambda_q} - \eta_n(\varphi)\right| \\
&= \left|\frac{Q_{p,n}(\varphi)(x)}{Q_{p,n}(1)(x)} - \frac{\eta_p Q_{p,n}(\varphi)}{\eta_p Q_{p,n}(1)}\right| \\
&= \left|\frac{(\delta_x \otimes \eta_p)\bar{Q}_{p,n}(\varphi\otimes 1) - (\eta_p \otimes \delta_x)\bar{Q}_{p,n}(\varphi \otimes 1)}{Q_{p,n}(1)(x)\eta_p Q_{p,n}(1)}\right| \\
&= \left|\frac{(\delta_x \otimes \eta_p)\bar{R}_{p,n}(\varphi\otimes 1 - 1\otimes\varphi)}{Q_{p,n}(1)(x)\eta_p Q_{p,n}(1)}\right| \\
&\leq 2\|\varphi\|_v \frac{(\delta_x\otimes\eta_p)\bar{R}_{p,n}(\bar{v})}{Q_{p,n}(1)(x)\eta_p Q_{p,n}(1)} =: 2\|\varphi\|_v \frac{\Delta_{p,n}(x,\eta_p)}{Q_{p,n}(1)(x)\eta_p Q_{p,n}(1)}, \qquad (5.33)
\end{aligned}
$$

where the third equality is due to the decomposition of Kleptsyna and Veretennikov [2008].

Now define $\rho_d := 1 - \left(\frac{\epsilon_d^-}{\epsilon_d^+}\right)^2 < 1$ and $\bar{M}_{p,n}^{(d)} := \sum_{k=p}^{n-1}\mathbb{I}_{\bar{C}_d}\left(\bar{X}_k\right)\mathbb{I}_{\bar{C}_d}\left(\bar{X}_{k+1}\right)$. Following essentially the same argument as [Douc et al., 2009, Proof of Proposition 12] then gives, for any $\beta \in (0,1)$,

$$
\begin{aligned}
\Delta_{p,n}(x,\eta_p) &\leq \bar{\mathbb{E}}_{\delta_x\otimes\eta_p}\left[\prod_{q=p}^{n-1}\bar{G}_q\left(\bar{X}_q\right)\rho_d^{\bar{M}_{p,n}^{(d)}}\bar{v}\left(\bar{X}_n\right)\right] \\
&= \bar{\mathbb{E}}_{\delta_x\otimes\eta_p}\left[\prod_{q=p}^{n-1}\bar{G}_q\left(\bar{X}_q\right)\rho_d^{\bar{M}_{p,n}^{(d)}}\mathbb{I}\left\{\bar{M}_{p,n}^{(d)} \geq \beta(n-p)\right\}\bar{v}\left(\bar{X}_n\right)\right] \\
&\quad + \bar{\mathbb{E}}_{\delta_x\otimes\eta_p}\left[\prod_{q=p}^{n-1}\bar{G}_q\left(\bar{X}_q\right)\rho_d^{\bar{M}_{p,n}^{(d)}}\mathbb{I}\left\{\bar{M}_{p,n}^{(d)} < \beta(n-p)\right\}\bar{v}\left(\bar{X}_n\right)\right] \\
&=: \Delta_{p,n}^{(1)}(x,\eta_p) + \Delta_{p,n}^{(2)}(x,\eta_p).
\end{aligned}
$$

We first consider $\Delta_{p,n}^{(1)}(x,\eta_p)$. As $\rho_d < 1$, we have the bound

$$
\frac{\Delta_{p,n}^{(1)}(x,\eta_p)}{Q_{p,n}(1)(x)\eta_p Q_{p,n}(1)} \leq \rho_d^{\beta(n-p)}\left[\frac{Q_{p,n}(v)(x)}{Q_{p,n}(1)(x)}\right]\eta_n(v),
$$

but using Lemma 6, Proposition 1, Proposition 2 and Lemma 10 show that for r large enough, but then fixed,

$$
\begin{aligned}
\frac{Q_{p,n}(v)(x)}{Q_{p,n}(1)(x)} &\leq \check{\mathbb{E}}_{p,x}^{(n)}\left[v_{n,n}\left(\check{X}_{n,n}\right)\right] \\
&\leq \frac{e^{-\delta r(n-p)}}{\underline{\lambda}^{(n-p)}}\frac{1}{\|h_{p,n}\|_v}v_{p,n}(x) \\
&\quad + \frac{e^{r(1-\delta)+b_r}}{\epsilon_r^-}\left[\frac{1}{\nu_r\left(\mathbb{I}_{C_r}h_{p,n}\right)} + \sum_{k=p+1}^{n-1}\frac{e^{-\delta r(n-k)}}{\underline{\lambda}^{(n-k)}}\frac{1}{\nu_r\left(\mathbb{I}_{C_r}h_{k,n}\right)}\right] \\
&\leq c\frac{v_{p,n}(x)}{\|h_{p,n}\|_v}, \qquad (5.34)
\end{aligned}
$$

so

$$\begin{aligned}\frac{\Delta_{p,n}^{(1)}(x,\eta_p)}{Q_{p,n}(1)(x)\eta_p Q_{p,n}(1)} &\leq c\rho_d^{\beta(n-p)}\frac{v_{p,n}(x)}{\left\|h_{p,n}\right\|_v}\eta_n(v)\\ &\leq c\rho_d^{\beta(n-p)}\frac{v_{p,n}(x)}{\left\|h_{p,n}\right\|_v}\mu(v). \qquad (5.35)\end{aligned}$$

where the second inequality is due to Proposition 1.

Now consider $\Delta_{p,n}^{(2)}(x,\eta_p)$. The main idea for treating this term is that of [Douc et al., 2009, Proof of Lemma 15]. There are some cosmetic differences of indexing, but some intermediate steps are omitted for brevity. Define

$$\widetilde{M}_{p,n}^{(d)} := \sum_{k=p}^{n-1}\mathbb{I}_{\bar{C}_d^c}\left(\bar{X}_k\right), \qquad a_{p,n} := \lfloor (n-p)(1-\beta)/2 - 1/2\rfloor,$$

$$A_p := |||\bar{Q}_p\mathbb{I}_{\bar{C}_d^c}|||_{v\otimes v}, \qquad B_p := |||\bar{Q}_p\mathbb{I}_{\bar{C}_d}|||_{v\otimes v},$$

$$\Xi_0 := \bar{v}\left(\bar{X}_p\right), \qquad \Xi_k := \left[\prod_{q=p}^{p+k-1}\frac{\bar{G}_q\left(\bar{X}_q\right)}{A_{q+1}^{\mathbb{I}_{\bar{C}_d^c}(\bar{X}_q)}B_{q+1}^{\mathbb{I}_{\bar{C}_d}(\bar{X}_q)}}\right]\bar{v}\left(\bar{X}_{p+k}\right),\ 1\leq k\leq n-p.$$

Then for $1\leq k\leq n-p$, $\bar{\mathbb{E}}_{p+k-1,\bar{X}_{p+k-1}}\left[\Xi_k\right]\leq \Xi_{k-1}$, so that

$$\bar{\mathbb{E}}_{p,\delta_x\otimes\eta_p}\left[\Xi_{n-p}\right] \leq \bar{\mathbb{E}}_{p,\delta_x\otimes\eta_p}\left[\Xi_0\right] = v(x)\eta_p(v)\leq cv(x)\mu(v), \qquad (5.36)$$

where the last inequality is due to Proposition 1.

By [Douc et al., 2009, Lemma 19], $\bar{M}_{p,n}^{(d)} < \beta(n-p)$ implies $\widetilde{M}_{p,n}^{(d)} \geq a_{p,n}$, and then

$$\begin{aligned}&\left[\prod_{q=p}^{p+k-1}A_{q+1}^{\mathbb{I}_{\bar{C}_d^c}(\bar{X}_q)}B_{q+1}^{\mathbb{I}_{\bar{C}_d}(\bar{X}_q)}\right]\mathbb{I}\left\{\bar{M}_{p,n}^{(d)} < \beta(n-p)\right\}\\ &\leq \left(\sup_{q\geq 1}A_q\right)^{a_{p,n}}\left(1\vee\sup_{q\geq 1}|||Q_q|||_v\right)^{2(n-p-a_{p,n})}\\ &\leq \left(\sup_{q\geq 1}|||Q_q\mathbb{I}_{C_d^c}|||_v\right)^{a_{p,n}}\left(1\vee\sup_{q\geq 1}|||Q_q|||_v\right)^{2(n-p)}\\ &\leq e^{-\delta d a_{p,n}}\left(1\vee\sup_{q\geq 1}|||Q_q|||_v\right)^{2(n-p)}\\ &\leq \exp\left(-\delta d a_{p,n}\right)\exp\left[0\vee 2b_{\underline{d}}(n-p)\right]. \qquad (5.37)\end{aligned}$$

where (H1) has been used. For the remainder of the proof we may assume without loss of generality that $b_{\underline{d}} > 0$.

Combining (5.36) and (5.37) then gives

$$\begin{aligned}\Delta^{(2)}_{p,n}(x,\eta_p) &\leq \bar{\mathbb{E}}_{\delta_x\otimes\eta_p}\left[\prod_{q=p}^{n-1}\bar{G}_q\left(\bar{X}_q\right)\mathbb{I}\left\{\bar{M}^{(d)}_{p,n}<\beta(n-p)\right\}\bar{v}\left(\bar{X}_n\right)\right] \\ &\leq \left(\sup_{q\geq 1}A_q\right)^{a_{p,n}}\left(\sup_{q\geq 1}B_q\right)^{n-p-a_{p,n}}\bar{\mathbb{E}}_{p,\delta_x\otimes\eta_p}\left[\Xi_{n-p}\right] \\ &\leq c\exp\left[-\delta d a_{p,n}+2b_{\underline{d}}(n-p)\right]v(x)\mu(v),\end{aligned}$$

and therefore

$$\begin{aligned}&\frac{\Delta^{(2)}_{p,n}(x,\eta_p)}{Q_{p,n}(1)(x)\eta_p Q_{p,n}(1)} \\ &= \frac{\Delta^{(2)}_{p,n}(x,\eta_p)}{h_{p,n}(x)\left(\prod_{q=p}^{n-1}\lambda_q\right)^2} \\ &\leq c\exp\left[-\delta d a_{p,n}+2(n-p)\left(b_{\underline{d}}-\log\underline{\lambda}\right)\right]\frac{v_{p,n}(x)}{\|h_{p,n}\|_v}\mu(v),\end{aligned} \tag{5.38}$$

where Proposition 1 and Proposition 2 have been applied and $\underline{\lambda}=\inf_{n\geq 0}\lambda_n>0$.

Collecting the bounds of (5.35), (5.38) and returning to (5.33), we establish that

$$\begin{aligned}&\left|\frac{Q_{p,n}(\varphi)(x)}{h_{p,n}(x)\prod_{q=p}^{n-1}\lambda_q}-\eta_n(\varphi)\right| \\ &\leq 2c\|\varphi\|_v\frac{v_{p,n}(x)}{\|h_{p,n}\|_v}\mu(v)\left[\rho_d^{\beta(n-p)}+\exp\left[-\delta d a_{p,n}+2(n-p)\left(b_{\underline{d}}-\log\underline{\lambda}\right)\right]\right] \\ &\leq 2c\|\varphi\|_v\frac{v_{p,n}(x)}{h_{p,n}(x)}\mu(v) \\ &\quad\cdot\left[\rho_d^{\beta(n-p)}+\exp\left[-(n-p)\left(\delta d(1-\beta)/2-2b_{\underline{d}}+2\log\underline{\lambda}\right)+3\delta d/2\right]\right],\end{aligned}$$

where for the second inequality, $\lfloor a\rfloor\geq a-1$ has been used. The proof is complete upon recalling that $d\in[\underline{d},\infty)$ was arbitrary, $\rho_d<1$, $\beta\in(0,1)$ and multiplying through by $h_{p,n}(x)$. $\square$

Proofs for section 4.4

Proof. (Theorem 2). Throughout the proof c denotes a finite constant whose value may change on each appearance but which depends only on μ and the quantities in (H1)-(H4). Also, throughout the proof we take by convention that for any $j<k$, $\sum_k^j\equiv 0$.

First consider the case $s>0$. By Lemma 2,

$$\frac{\mathbb{E}_{\mu}\left[\prod_{p=0}^{n-1}G_{p}\left(X_{p}\right)\exp\left(\sum_{k\in\{i_{1},\ldots,i_{s}\}}\left|F_{k}\left(X_{k}\right)\right|\right)\right]}{\mathbb{E}_{\mu}\left[\prod_{p=0}^{n-1}G_{p}\left(X_{p}\right)\right]}$$
$$=\int\mu\left(dx\right)h_{0,n}\left(x\right)\check{\mathbb{E}}_{x}^{(n)}\left[\exp\left(\sum_{k\in\{i_{1},\ldots,i_{s}\}}\left|F_{k}\left(\check{X}_{k,n}\right)\right|\right)\right]$$
$$\leq\left(\prod_{k\in\{i_{1},\ldots,i_{s}\}}\left\|e^{|F_{k}|}\right\|_{v^{\delta}}\right)\int\mu(dx)h_{0,n}(x)\check{\mathbb{E}}_{x}^{(n)}\left[\prod_{k\in\{i_{1},\ldots,i_{s}\}}v^{\delta}\left(\check{X}_{k,n}\right)\right]. \tag{5.39}$$

We now obtain some bounds which will be used to control the expectation in (5.39). Proposition 1 holds under the assumptions of the theorem so we may apply the upper and lower bounds of Proposition 2 and Lemma 10 to the bound of Lemma 6 and choose d therein large enough, in order to establish that there exists a finite constant c independent of $1\leq p<q\leq n$ and $x\in\mathsf{X}$ such that

$$\begin{aligned}
&\check{\mathbb{E}}_{p,x}^{(n)}\left[v_{q,n}\left(\check{X}_{q,n}\right)\right]\\
&\leq\frac{e^{-\delta d(q-p)}}{\underline{\lambda}^{(q-p)}}\frac{\left\|h_{q,n}\right\|_{v}}{\left\|h_{p,n}\right\|_{v}}v_{p,n}\left(x\right)\\
&\quad+\frac{e^{d(1-\delta)+b_{d}}}{\epsilon_{d}^{-}}\left\|h_{q,n}\right\|_{v}\left[\frac{1}{\nu_{d}\left(\mathbb{I}_{C_{d}}h_{q,n}\right)}+\sum_{k=p+1}^{q-1}\frac{e^{-\delta d(q-k)}}{\underline{\lambda}^{(q-k)}}\frac{1}{\nu_{d}\left(\mathbb{I}_{C_{d}}h_{k,n}\right)}\right]\\
\leq\;&c\frac{\left\|h_{q,n}\right\|_{v}}{\left\|h_{p,n}\right\|_{v}}v_{p,n}\left(x\right),
\end{aligned}\tag{5.40}$$

where $\underline{\lambda}=\inf_{n\geq0}\lambda_{n}>0$. Therefore by (H1), for $p\leq q$,

$$\begin{aligned}
v^{\delta}(x)\check{\mathbb{E}}_{p-1,x}^{(n)}\left[v_{q,n}\left(\check{X}_{q,n}\right)\right] &\leq cv^{\delta}(x)\frac{\left\|h_{q,n}\right\|_{v}}{\left\|h_{p,n}\right\|_{v}}S_{p,n}\left(v_{p,n}\right)(x)\\
&= cv^{\delta}(x)\left\|h_{q,n}\right\|_{v}\frac{Q_{p}\left(v\right)(x)}{\lambda_{p-1}h_{p-1,n}(x)}\\
&\leq c\left\|h_{q,n}\right\|_{v}\frac{v\left(x\right)}{h_{p-1,n}(x)}\frac{e^{b_{\underline{d}}}}{\underline{\lambda}}\\
&\leq c\frac{\left\|h_{q,n}\right\|_{v}}{\left\|h_{p-1,n}\right\|_{v}}v_{p-1,n}(x).
\end{aligned}\tag{5.41}$$

Now fix $n\geq1$, $1\leq s\leq n+1$, $(i_{1},\ldots,i_{s})\in\mathcal{I}_{n,s}$ arbitrarily and define $(\Xi_{k,n};0\leq k\leq s)$ by

$$\begin{aligned}
\Xi_{0,n}&:=\frac{v_{0,n}\left(\check{X}_{0,n}\right)}{\left\|h_{0,n}\right\|_{v}},\\
\Xi_{k,n}&:=\frac{v_{i_{k},n}\left(\check{X}_{i_{k},n}\right)}{\left\|h_{i_{k},n}\right\|_{v}}\exp\left[\sum_{j=1}^{k-1}\left(\delta V\left(\check{X}_{i_{j},n}\right)-\log c\right)\right],\ 1\leq k\leq s.
\end{aligned}$$

where c is as in (5.41). We then have

$$
\begin{aligned}
&\check{\mathbb{E}}^{(n)}_{i_{k-1},\check{X}_{i_{k-1},n}}\left[\Xi_{k,n}\right] \\
&= \frac{1}{\left\|h_{i_k,n}\right\|_v}\check{\mathbb{E}}^{(n)}_{i_{k-1},\check{X}_{i_{k-1},n}}\left[v_{i_k,n}\left(\check{X}_{i_k,n}\right)\right]\exp\left[\sum_{j=1}^{k-1}\left(\delta V\left(\check{X}_{i_j,n}\right)-\log c\right)\right] \\
&\leq c\frac{v_{i_{k-1},n}\left(\check{X}_{i_{k-1},n}\right)}{\left\|h_{i_{k-1},n}\right\|_v}v^{\delta}\left(\check{X}_{i_{k-1},n}\right)\exp\left[\sum_{j=1}^{k-1}\left(\delta V\left(\check{X}_{i_j,n}\right)-\log c\right)\right] \\
&= \Xi_{k-1,n},
\end{aligned}
$$

where the inequality is due to (5.41). Thus $\left(\Xi_{k,n},\check{\mathcal{F}}_{k,n};0\leq k\leq s\right)$ is a super-Martingale, with $\check{\mathcal{F}}_{k,n}:=\sigma\left(\check{X}_{0,n},\ldots,\check{X}_{i_k-1,n},\check{X}_{i_k,n}\right)$. Therefore

$$
\begin{aligned}
\check{\mathbb{E}}^{(n)}_x\left[\prod_{k\in\{i_1,\ldots,i_s\}}v^{\delta}\left(\check{X}_{k,n}\right)\right] &\leq c^s\check{\mathbb{E}}^{(n)}_x\left[\Xi_{s,n}\right]\left\|h_{i_s,n}\right\|_v \\
&\leq c^s\frac{v_{0,n}(x)}{\left\|h_{0,n}\right\|_v}\left\|h_{i_k,n}\right\|_v \\
&\leq c^s\frac{v(x)}{h_{0,n}(x)} \qquad (5.42)
\end{aligned}
$$

where Propositions 1-2 have been used for the last inequality.

The proof is completed upon combining (5.42) with (5.39) and noting that the result holds trivially when $s=0$. □

References

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

N. Chopin. Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Ann. Statist.*, 32(6):2385–2411, 2004.

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. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. *Ann. Inst. Henri Poincaré Probab. Stat.*, 37(2):155–194, 2001.

P. Del Moral and J. Jacod. Interacting particle filtering with discrete time observations: Asymptotic behaviour in the Gaussian case. In T. Hida, R.L. Karandikar, H. Kunita, B.S. Rajput, S. Watanabe, and J. Xiong, editors, *Stochastics in Finite and Infinite Dimensions: In Honor of Gopinath Kallianpur.* 2001.

P. Del Moral, F. Patras, and S. Rubenthaler. Tree based functional expansions for Feynman Kac particle models. *Ann. Appl. Probab.*, 19(2):778–825, 2009.

P. Del Moral, A. Doucet, and S.S. Singh. A backward particle interpretation of Feynman-Kac formulae. *ESAIM: Mathematical Modelling and Numerical Analysis*, 44(05):947–975, 2010.

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

R. Douc and E. Moulines. Limit theorems for weighted samples with applications to sequential Monte Carlo methods. *Ann. Statist.*, 36(5):2344–2376, 2008.

R. Douc, G. Fort, E. Moulines, and P. Priouret. Forgetting the initial distribution for hidden Markov models. *Stochastic Process. Appl.*, 119:1235–1256, 2009.

R Douc, A. Garivier, E. Moulines, and J. Olsson. Sequential Monte Carlo smoothing for general state space hidden Markov models. *Ann. Appl. Probab.*, 2011. To appear.

A. Doucet, S. Godsill, and C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.*, 10(3):197–208, 2000.

B. Favetto. On the asymptotic variance in the central limit theorem for particle filters. *ESAIM: Probab. Stat.*, 1(1), 2009.

N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *Radar and Signal Processing, IEE Proceedings F*, volume 140, pages 107–113. IET, 1993.

K. Heine and D. Crisan. Uniform approximations of discrete-time filters. *Adv. in Appl. Probab.*, 40(4):979–1001, 2008.

M. L. Kleptsyna and A. Y. Veretennikov. On discrete time ergodic filters with wrong initial data. *Probab. Theory Related Fields*, 141(3-4):411–444, 2008.

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

H.R. Künsch. Recursive Monte Carlo filters: algorithms and theoretical analysis. *Ann. Statist.*, 33(5):1983–2021, 2005.

F. LeGland and N. Oudjane. A robustification approach to stability and to uniform particle approximation of nonlinear filters: the example of pseudo-mixing signals. *Stochastic Process. Appl.*, 106(2):279–316, 2003.

F. LeGland and N. Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. *Ann. Appl. Probab.*, 14(1):144–187, 2004.

N. Oudjane and S. Rubenthaler. Stability and uniform particle approximation of nonlinear filters in case of non ergodic signals. *Stoch. Anal. Appl.*, 23(3):421–448, 2005.

M.K. Pitt and N. Shephard. Filtering via simulation: Auxiliary particle filters. *J. Amer. Statist. Assoc.*, pages 590–599, 1999.

R. van Handel. Uniform time average consistency of Monte Carlo particle filters. *Stochastic Process. Appl.*, 119(11):3835–3861, 2009.

N. Whiteley. Sequential Monte Carlo samplers: error bounds and insensitivity to initial conditions. *ArXiv preprint*, 1103.3970, March 2011.

N. Whiteley, N. Kantas, and A. Jasra. Linear variance bounds for particle approximations of time-homogeneous Feynman-Kac formulae. *ArXiv preprint*, 1108.3988, August 2011.