CENTRAL LIMIT THEOREMS FOR COUPLED PARTICLE FILTERS

AJAY JASRA* AND FANGYUAN YU, * King Abdullah University of Science & Technology

Abstract

In this article we prove new central limit theorems (CLTs) for several coupled particle filters (CPFs). CPFs are used for the sequential estimation of the difference of expectations with respect to filters which are in some sense close. Examples include the estimation of the filtering distribution associated to different parameters (finite difference estimation) and filters associated to partially observed discretized diffusion processes (PODDP) and the implementation of the multilevel Monte Carlo (MLMC) identity. We develop new theory for CPFs, and based upon several results, we propose a new CPF which approximates the maximal coupling (MCPF) of a pair of predictor distributions. In the context of ML estimation associated to PODDP with time-discretization $\Delta t = 2^{-l}, l \in \{0, 1, \ldots\}$, we show that the MCPF and the approach of Jasra, Ballesio, et al. (2018) have, under certain assumptions, an asymptotic variance that is bounded above by an expression that is of (almost) the order of $O(\Delta t)$, uniformly in time. The $O(\Delta t)$ bound preserves the so-called forward rate of the diffusion in some scenarios, which is not the case for the CPF in Jasra et al. (2017).

Keywords: Coupled particle filter; central limit theorem; multilevel Monte Carlo

2010 Mathematics Subject Classification: Primary 65C05 Secondary 62F15

1. Introduction

The filtering problem is ubiquitous in statistics, applied probability, and applied mathematics, with far-reaching applications in weather prediction, finance, and engineering; see [4] and [7], for example. In most cases of practical interest, the filter must be numerically approximated, and a popular method for doing so is the particle filter (PF) (see e.g. [4], [8], and the references therein). The PF generates $N \geq 1$ samples in parallel and uses a combination of sampling, importance sampling, and resampling to approximate the filter. There is a substantial literature on the convergence of PFs (e.g. [8]) and in particular there are central limit theorems (CLTs) which allow one to understand the errors associated to estimation. Under certain assumptions, the associated asymptotic variance is bounded uniformly in time; see e.g. [5].

In this article, we are concerned with the filtering problem where one seeks to estimate the difference of expectations of two different but ‘close’ filters. As an example, suppose one observes data at discrete and regular time points associated to an unobserved diffusion process. In many cases, one must time-discretize the diffusion process, if the transition density is not available up to a nonnegative unbiased estimator. In such scenarios it is well known that the
cost of estimating the filter using a PF can be significantly reduced using a collapsing sum representation of the expectation associated to the filter with the most precise discretization and estimating each difference independently using a coupled particle filter (CPF). In other applications, one can approximate differences of expectations of filters with different parameter values as a type of finite difference approximations; see for instance [16] and [28].

The CPF developed in [6] (see also [16], [17], [22], [23], and [28]) is used in several applications as discussed above and various other contexts. It consists of a particle filter which runs on the product space of the two filters. The sampling operation often consists of simulating from a coupling of the Markov transitions of the hidden dynamics, which are often available in applications. Resampling proceeds by sampling the maximal coupling associated to the probability distributions of particle indices. The maximal coupling between two probability measures on the same space is the coupling which maximizes the probability that random variables drawn from the two probabilities are equal. The use of correlating the PFs is vital, for instance in multilevel (ML) applications (see [13], [15], and [22]), as it is this property which allows a variance reduction relative to a single PF. As has been noted by several authors, e.g. [28], unless the coupling of the Markov transitions is particularly strong, one expects the maximally coupled resampling operation to ultimately decorrelate the pairs of particles exponentially fast in time. As a result, the benefits of running such algorithms may have a minimal effect for long time intervals.

In this article we consider four CPFs. The first, the independent resampling CPF (IRCPF), is the case where the resampling is independent for each pair of particles. The second has maximally coupled index resampling; that is, particle pairs are resampled according to the coupling which maximizes the probability that the index (in the set \{1, \ldots, N\}) sampled for both particle pairs is equal. The CPF that uses this resampling operation is called the maximally coupled index resampling CPF (MCIRCPF). The third algorithm, which to our knowledge is new, is a CPF which approximates the sequence of maximal couplings of the predictors, which we call the maximally coupled particle filter (MCPF). The motivation for this method is associated to the fact that the sequence of limiting couplings approximated by the MCIRCPF does not seem to have any optimality properties in terms of coupling. The MCPF requires the Markov transition of the filter to have a density which is known pointwise. It is worth noting that the MCIRCPF is a CPF which maximizes the probability that resampled indices are equal, which, as we will discuss, may not be useful for approximating the filter/predictor for large time horizons (time of the stochastic process). The MCPF is approximating the maximal coupling of the predictor, which, as we will show, can be particularly useful for approximating the filter/predictor for large time horizons. The fourth algorithm in [20] is based on multinomial resampling which uses the same uniform random variable for each particle pair; we call this the Wasserstein CPF (WCPF). In general, all four algorithms can be used in each of the examples, with the constraint for the MCPF mentioned above. We remark that there are CPFs in [14] and [28], but they are not considered here, as their analysis involves even more mathematical complexity.

We prove a CLT for the first three algorithms (the CLT for the WCPF, when the state dynamics are one-dimensional, is in [20]) associated to the difference of estimates of expectations of the same function with respect to the predictors, which is extended to multiple dimensions. The asymptotic variance expression is directly related to the properties of the limiting coupling one approximates. Under certain assumptions (of the type in [30]), for partially observed discretized diffusion processes (PODDP) with (Euler) discretization \(\Delta_l\) (typically \(\Delta_l = 2^{-l}, l \in \{0, 1, \ldots \}\)), we show that the MCPF (resp. WCPF) has an asymptotic variance that is
bounded above by an expression that is of order $\Delta t \, O(\Delta t)$ (resp. $O(\Delta t^{1-\lambda})$, with $\lambda$ arbitrarily close to, but not equal to, zero), uniformly in time, which preserves the so-called forward rate of the diffusion in some scenarios. This is reassuring as it shows that filtering for difference estimation in the PODDP case can be effectively performed. This time and rate stability is associated to the fact that the limiting coupling on product spaces is associated to the optimal $L_0$ and $L_2$ Wasserstein couplings of the predictor/filter. For the IRCPF one does not recover the coupling rate of the diffusion process, and this poor performance is well-known in the literature. In the case of the MCIRCPF we show that even in a favourable situation, it can have an asymptotic variance, at time $n$, that is $O(e^{n} \Delta t)$; we also identify when one can expect the algorithm to work well. As was seen in the empirical results of [20] and [22], the time and rate behavior of the MCIRCPF in general does not seem to be as good as for the MCPF and WCPF. The assumptions used for our asymptotic variance results are also verified in a real example. Our CLTs are, to the best of our knowledge, the first results of these types for CPFs and require nonstandard proofs, especially in the case of the MCIRCPF. To summarize, the main results of the article are the following:

- Theorem 3.1 gives a WLLN for the MCIRCPF.
- Theorem 4.1 gives a CLT for the IRCPF, MCIRCPF, and MCPF.
- Theorem 4.3 gives a general bound on the asymptotic variance for each of the methods: the IRCPF, MCIRCPF, MCPF, and WCPF.
- Propositions 5.2 and 5.3 give the time-uniform bounds on the asymptotic variance for the MCPF and WCPF noted above (i.e. for PODDPs).

This paper is structured as follows. In Section 2 we give our notation, models, and the motivating example of a PODDP with MLMC. In Section 3 the algorithms are presented. In Section 4 our theoretical results are stated. Our CLTs are given and a general bound on the asymptotic variance is provided. In Section 5 our results are applied to a practical model in the context of using coupled particle filters for PODDP with MLMC. The article is summarized in Section 6. The appendix contains the proofs of our theoretical results.

2. Notation and models

2.1. Notation

Let $(X, \mathcal{A})$ be a measurable space. For a given function $v : X \to [1, +\infty)$ and for measurable $\varphi : X \to \mathbb{R}$,

$$
\|\varphi\|_v := \sup_{x \in X} \frac{|\varphi(x)|}{v(x)}.
$$

We define $\mathcal{L}_v(X) = \{\varphi : X \to \mathbb{R} : \|\varphi\|_v < +\infty\}$. When $v = 1$ we write $\|\varphi\| := \sup_{x \in X} |\varphi(x)|$. We write $B_b(X), C_b(X)$ for the bounded measurable and continuous, bounded measurable real-valued functions respectively. $C^2(X)$ is the collection of twice continuously differentiable real-valued functions on $X$. Let $d$ be a metric on $X$; then for $\varphi \in \mathcal{L}_v(X)$, we say that $\varphi \in \text{Lip}_{v, d}(X)$ if there exists a $C < +\infty$ such that for every $(x, y) \in X \times X$,

$$
|\varphi(x) - \varphi(y)| \leq C d(x, y)v(x)v(y).
$$

If $v = 1$, we write $\varphi \in \text{Lip}_d(X)$ and write $\|\varphi\|_{\text{Lip}}$ for the Lipschitz constant. $\mathcal{P}(X)$ denotes the collection of probability measures on $(X, \mathcal{A})$. We also write $\|\mu\|_v := \sup_{|\varphi| \leq v} |\mu(\varphi)|$ for
\( \mu \in \mathcal{P}(X) \). For a measure \( \mu \) on \((X, \mathcal{X})\) and a \( \mu \)-integrable function \( \varphi : X \to \mathbb{R} \), the notation \( \mu(\varphi) = \int_X \varphi(x) \mu(dx) \) is used. Let \( K : X \times \mathcal{X} \to (0, +\infty) \) be a nonnegative kernel and \( \mu \) a measure; then we use the notation \( \mu K(dy) = \int_X \mu(dx) K(x, dy) \), and for \( K(x, \cdot) \)-integrable \( \varphi : X \to \mathbb{R} \), we write \( K(\varphi)(x) = \int_X \varphi(y) K(x, dy) \). For \( \mu, \nu \in \mathcal{P}(X) \), the total variation distance is denoted \( \|\mu - \nu\|_{tv} = \sup_{B \in \mathcal{X}} |\mu(B) - \nu(B)| \). For \( B \in \mathcal{X} \) the indicator function is written \( \mathbb{1}_B(x) \) and the Dirac measure \( \delta_B(dx) \). For two measures \( \mu, \nu \) on \((X, \mathcal{X})\), the product measure is denoted \( \mu \otimes \nu \). For two measurable functions \( \varphi, \psi \) on \((X, \mathcal{X})\), the tensor product of functions is denoted \( \varphi \otimes \psi \). \( U_A \) denotes the uniform distribution on the set \( A \). \( \mathcal{N}(a, b) \) is the \( t \)-dimensional Gaussian distribution of mean \( a \) and covariance \( b \) (if \( t = 1 \) the subscript is dropped from \( \mathcal{N} \)). \( \mathbb{P} \) and \( \mathbb{E} \) are used to denote probability and expectation with respect to the law of the specified algorithm—the context will be clear in each instance. The symbols \( \Rightarrow \) and \( \to \mathbb{P} \) are used to denote convergence in distribution and probability respectively. In the context of the article, this is as \( N \to +\infty \).

### 2.2. Models

Let \((X, \mathcal{X})\) be a measurable space and \( \{G_n\}_{n \geq 0} \) a sequence of nonnegative, bounded, and measurable functions such that \( G_n : X \to \mathbb{R}_+ \). Let \( \eta_n^f, \eta_n^c \in \mathcal{P}(X) \) and let \( \{M_n^f\}_{n \geq 1}, \{M_n^c\}_{n \geq 1} \) be two sequences of Markov kernels, i.e. \( M_n^f : X \to \mathcal{P}(X), M_n^c : X \to \mathcal{P}(X) \). For \( s \in \{f, c\}, B \in \mathcal{X} \), define

\[
\gamma_n^s(B) = \int_{X^{n+1}} \mathbb{1}_B(x_n) \left( \prod_{p=0}^{n-1} G_p^s(x_p) \right) \eta_0^s(dx_0) \prod_{p=1}^{n} M_p^s(x_{p-1}, dx_p)
\]

and

\[
\eta_n^s(B) = \frac{\gamma_n^s(B)}{\gamma_n^s(1)}.
\]

The objective is to consider Monte-Carlo-type algorithms which will approximate, for \( \varphi \in \mathcal{B}_b(X) \) and any \( n \geq 0 \), quantities such as

\[
\eta_n^f(\varphi) - \eta_n^c(\varphi) \tag{1}
\]

or

\[
\frac{\eta_n^f(G_n \varphi)}{\eta_n^f(G_n)} - \frac{\eta_n^c(G_n \varphi)}{\eta_n^c(G_n)} \tag{2}
\]

(1) corresponds to a predictor of a state-space model and (2) to the filter. We focus explicitly on the predictor from here on.

**Remark 2.1.** We can make the \( G_n \) also depend on \( \{f, c\} \), which may be of importance in applications. In the subsequent development, this is not done, but could be, at a cost of slightly longer mathematical arguments and notational complications.

The major point is that one would like to approximate couplings of \( (\eta_n^f, \eta_n^c) \), say \( \tilde{\eta}_n \in \mathcal{P}(X \times X) \), i.e. that for any \( B \in \mathcal{X} \) and every \( n \geq 0 \),

\[
\tilde{\eta}_n(B \times X) = \eta_n^f(B), \quad \tilde{\eta}_n(X \times B) = \eta_n^c(B),
\]

and we consider approximating

\[
\tilde{\eta}_n(\varphi \otimes 1) - \tilde{\eta}_n(1 \otimes \varphi).
\]
An explanation of why coupling the pairs is of interest has been given in the introduction and will be further illuminated in Section 2.3.

Throughout the article, it is assumed that \( \tilde{\eta}_0 \in \mathcal{P}(X \times X) \) is such that for any \( B \in \mathcal{X} \),

\[
\tilde{\eta}_0(B \times X) = \eta^0_0(B), \quad \tilde{\eta}_0(X \times B) = \eta^c_0(B),
\]

and moreover for any \( n \geq 1 \) there exist Markov kernels \( \{\tilde{M}_n\} \), \( \tilde{M}_n : X \times X \rightarrow \mathcal{P}(X \times X) \), such that for any \( B \in \mathcal{X}, (x, x') \in X \times X \),

\[
\tilde{M}_n(B \times X)(x, x') = M^\ell_n(B)(x), \quad \tilde{M}_n(X \times B)(x, x') = M^c_n(B)(x'). \tag{3}
\]

2.3. Example

The following example is from [22] and there is some overlap with the presentation in that article. We start with a diffusion process

\[
dZ_t = a(Z_t)dt + b(Z_t)dW_t \tag{4}
\]

with \( Z_t \in \mathbb{R}^d = X \), \( a : \mathbb{R}^d \rightarrow \mathbb{R}^d \) (\( j \)th element denoted \( a^j \)), \( b : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d} \) (\( ij \), \( k \)th element denoted \( b^{ij,k} \)), \( t \geq 0 \), and \( \{W_t\}_{t \geq 0} \) a \( d \)-dimensional Brownian motion. The following assumption is made and is referred to as (D). We set \( Z_0 = x^* \in X \).

The coefficients \( a^j, b^{ij,k} \) belong to \( C^2(X) \), for \( j, k = 1, \ldots, d \). Also, \( a \) and \( b \) satisfy the following properties:

(i) the uniform ellipticity property: \( b(z)b(z)^T \) is uniformly positive definite;

(ii) the globally Lipschitz property: there is a \( C > 0 \) such that \(|a^j(z) - a^j(y)| + |b^{ij,k}(z) - b^{ij,k}(y)| \leq C|z - y| \) for all \((z, y) \in X \times X, (j, k) \in \{1, \ldots, d\}^2 \).

The data are observed at regular unit time-intervals (i.e. in discrete time) \( y_1, y_2, \ldots, y_k \in Y \). It is assumed that conditional on \( Z_k, Y_k \) is independent of all other random variables with density \( G(z_k, y_k) \). Let \( M(z, dy) \) be the transition of the diffusion process (over unit time) and consider a discrete-time Markov chain \( X_0, X_1, \ldots \) with initial distribution \( M(x^*, \cdot) \) and transition \( M(x, dy) \). Here we are creating a discrete-time Markov chain that corresponds to the discrete-time skeleton of the diffusion process at a time lag of 1. Now we write \( G_k(x_k) \) instead of \( G(x_k, y_{k+1}) \). Then we define, for \( B \in \mathcal{X} \),

\[
\gamma_n(B) := \int_{X^{n+1}} \mathbb{I}_B(x_n) \left( \prod_{p=0}^{n-1} G_p(x_p) \right) \cdot M(x^*, dx_0) \prod_{p=1}^{n} M(x_{p-1}, dx_p).
\]

The predictor is \( \eta_n(B) = \gamma_n(B)/\gamma_n(1) \), which corresponds to the distribution associated to \( Z_{n+1} | y_1, \ldots, y_n \).

In many applications, one must time-discretize the diffusion to use the model in practice. We suppose an Euler discretization with discretization \( \Delta_l = 2^{-l}, l \geq 0 \), and write the associated transition kernel over unit time as \( M^\ell(x, dy) \). Note that, in practice, one may not be able to compute the density of the kernel, as it is a composition of \( \Delta_l^{-1} - 1 \) Gaussians; however, one can certainly sample from \( M^\ell \) in most cases. Hence we are interested in the Feynman–Kac model for \( B \in \mathcal{X} \),

\[
\gamma_n^\ell(B) := \int_{X^{n+1}} \mathbb{I}_B(x_n) \left( \prod_{p=0}^{n-1} G_p(x_p) \right) \cdot M^\ell(x^*, dx_0) \prod_{p=1}^{n} M^\ell(x_{p-1}, dx_p),
\]
with associated predictor \( \eta_n^l(B) = \gamma_n^l(B)/\gamma_n^l(1) \).

Below, we will explain why one may wish to compute, for \( l \geq 1 \), \( \eta_n^l(\varphi) - \eta_n^{l-1}(\varphi), \varphi \in B_b(X) \). That is, \( f \) as used above relates to the predictor associated to the discretization \( \Delta_l \) and \( c \) to the predictor with discretization \( \Delta_{l-1} \). A natural coupling of \( M_n^l = M^l \) and \( M_n^{l-1} = M^{l-1} \) exists (e.g. [13]) so one also has a given \( \tilde{\eta}_n \) and \( \tilde{M}_n \).

Before continuing, we note some results which will help in our discussion below. As established in [22, Eq. (32)], for \( C < +\infty \) one has

\[
\sup_{\varphi \in \mathcal{A}} \sup_{x \in X} \left| M_n^l(\varphi)(x) - M_n^c(\varphi)(x) \right| \leq C \Delta_l, \tag{5}
\]

where \( \mathcal{A} = \{ \varphi \in B_b(X) \cap \text{Lip}_{d_1}(X) : \| \varphi \| \leq 1 \} \), with \( d_1 \) the \( L_1 \)-norm. In addition, [22, Proposition D.1] states that for \( p > 0, C < +\infty \), and any \( (x, y) \in X \times X \),

\[
\int_{X \times X} \| u - v \|^p \tilde{M}_n((x, y), d(u, v)) \leq C(\| x - y \| + \Delta_l^{1/2} \varphi). \tag{6}
\]

When \( p = 2 (\| u - v \| \) is the \( L_2 \)-norm), the term \( \Delta_l \) is the so-called *forward strong error rate*. In the proof of [22, Theorem 4.3], it is shown that for any \( n \geq 0 \) there exists a \( C < +\infty \) such that for \( \varphi \in \mathcal{A}, l \geq 0 \), one has

\[
| \eta_n(\varphi) - \eta_n^l(\varphi) | \leq C \Delta_l. \tag{7}
\]

Note that this bound can be deduced from (5), but that \( C \) may depend on \( n \); this latter point is ignored for now.

2.3.1. Multilevel Monte Carlo. Suppose one can exactly sample from \( \eta_n^l \) for any \( l, n \geq 0 \). The Monte Carlo estimate of \( \eta_n^l(\varphi) \), with \( \varphi \in \mathcal{A} \), is then of course \( \frac{1}{N} \sum_{i=1}^N \varphi(x_i^l) \), where the \( X_i^l \) are independent and identically distributed (i.i.d.) from \( \eta_n^l \). One has the mean square error (MSE)

\[
E \left[ \left( \frac{1}{N} \sum_{i=1}^N \varphi(x_i^l) - \eta_n(\varphi) \right)^2 \right] = \frac{\text{Var}_{\eta_n} [\varphi(X)]}{N} + | \eta_n(\varphi) - \eta_n^l(\varphi) |^2,
\]

where \( \text{Var}_{\eta_n} [\varphi(X)] \) is the variance of \( \varphi(X) \) with respect to \( \eta_n^l \). Then, for \( \varepsilon > 0 \) given, to target an MSE of \( \mathcal{O}(\varepsilon^2) \), by (7), one chooses \( l = \mathcal{O}(| \log(\varepsilon) |) \) (as one sets \( \Delta_l^2 = \varepsilon^2 \)). Then one must choose \( N = \mathcal{O}(\varepsilon^{-2}) \), and we suppose the cost of simulating one sample is \( \mathcal{O}(\Delta_l^{-1}) \) (see [22] for a justification), again ignoring \( n \). Then the cost of achieving an MSE of \( \mathcal{O}(\varepsilon^2) \) is \( \mathcal{O}(\varepsilon^{-3}) \).

Now, for any \( L \geq 1 \), one has the multilevel identity ([13], [15])

\[
\eta_n^L(\varphi) = \eta_n^0(\varphi) + \sum_{l=1}^L \left( \eta_n^l - \eta_n^{l-1} \right)(\varphi).
\]

Suppose that, for \( l \geq 1 \), it is possible to exactly sample a coupling of \( (\eta_n^l, \eta_n^{l-1}) \)—let us denote it \( \tilde{\eta}_n^l \)—so that one has

\[
\int_{X \times X} \| x - y \|^2 \tilde{\eta}_n^l(d(x, y)) \leq C \Delta_l,
\]

and the cost of such a simulation is \( \mathcal{O}(\Delta_l^{-1}) \). The rate \( \Delta_l \) has been taken from the strong error rate in (6). Now, to estimate \( [\eta_n^l - \eta_n^{l-1}](\varphi) \), one simulates

\[
(X_n^{1, l}, X_n^{1, l-1}), \ldots, (X_n^{N_l, l}, X_n^{N_l, l-1})
\]
For $1 \leq l \leq L$ this is repeated independently for estimating $[\eta^l_n - \eta^{l-1}_n](\varphi)$, and the case $l = 0$ is performed, independently, using the Monte Carlo method above. One can then easily show that the MSE of the estimate of $\eta^L_n$ is bounded above by

$$\text{Var}_{\tilde{\eta}_n}[\varphi(X)]/N_0 + C \sum_{l=1}^{L} \Delta_l/N_l + |\eta_n(\varphi) - \eta^l_n(\varphi)|^2$$

(recall $\varphi \in A$). Then, for $\epsilon > 0$ given, to target an MSE of $O(\epsilon^2)$, set $L = O(|\log(\epsilon)|)$ and $N_l = O(\epsilon^{-2}|\log(\epsilon)|\Delta_l)$ (see [13]); then the MSE is $O(\epsilon^2)$. The overall computational effort is $O(\sum_{l=0}^{L} N_l \Delta_l^{-1}) = O(\epsilon^{-2}|\log(\epsilon)|^2)$; if $\epsilon$ is suitably small this is a significant reduction in computational effort relative to the MC method above.

The main point here is that, at present, there are no computational methods to perform the exact simulation mentioned, and thus we focus on particle filters, which have been developed in the literature. The estimates (variance/cost) above typically depend on $n$, and our objective is to consider whether the variance of PF estimates can be $O(\Delta_l)$ uniformly in time, so that the ML gain is retained uniformly in time. We focus on the asymptotic variance in the CLT (versus the finite sample variance) as this is more straightforward to deal with.

3. Algorithms

3.1. Independent pair resampling

The first procedure we consider is as follows. Let $n \geq 1$, $B \in \mathcal{X} \times \mathcal{X}$, and $\mu \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$, and define the probability measure $\tilde{\eta}_n^l(\mu)(B) = \mu(G_{n-1} \otimes G_{n-1}) M_n(B)$. Set $u_n = (x^f_n, x^c_n) \in \mathcal{X} \times \mathcal{X}$, then consider the joint probability measure on $(\mathcal{X} \times \mathcal{X})^n$ given by

$$\mathbb{P}(d(u_0, \ldots, u_n)) = \tilde{\eta}_0(du_0) \prod_{p=1}^{n} \tilde{\Phi}_p^f(\eta_{p-1}^{f} \otimes \eta_{p-1}^{c})(du_p).$$

Noting (3), it is easily checked that the marginal of $x^f_n$ (resp. $x^c_n$) is $\eta^f_n$ (resp. $\eta^c_n$). Denote by $\tilde{\eta}_n^l$ the marginal of $u_n$ induced by this joint probability measure. Thus, if one could sample a trajectory of $(u_0, \ldots, u_n)$ from $\mathbb{P}(d(u_0, \ldots, u_n))$ one could easily approximate quantities such as (1) or (2) using Monte Carlo methods. In most practical applications of interest, this is not possible.

The particle approximation of (8) is taken as

$$\mathbb{P}(d(u_0^{1:N}, \ldots, u_n^{1:N})) = \left( \prod_{i=1}^{N} \tilde{\eta}_0(du_0^i) \right) \left( \prod_{p=1}^{n} \prod_{i=1}^{N} \tilde{\Phi}_p^f(\eta_{p-1}^{f} \otimes \eta_{p-1}^{c})(du_p) \right),$$

i.i.d. from $\tilde{\eta}_n^l$ for some $N_l \geq 1$, and the approximation is

$$\frac{1}{N_l} \sum_{i=1}^{N_l} \{ \varphi(x^{i,l}_n) - \varphi(x^{i,l-1}_n) \}.$$
where for \( p \geq 1 \), \( s \in \{ f, c \} \),

\[
\eta_{p-1}^{N,s}(dx) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\Delta_{u_i^{\ell}}} (dx).
\]

The IRCPF algorithm which simulates from \( \mathbb{P}(d(u_0^1, \ldots)) \) is presented in Algorithm 1. The key point is that in this algorithm, the resampled indices for a pair of particles are generated (conditionally) independently. Set

\[
\tilde{\eta}_{p}^{N,1}(du) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\nu_i^{\ell}} (du).
\]

As has been mentioned by many authors (e.g. [28]), one does not expect this procedure to be effective, in the sense that \( \tilde{\eta}_n^f \) would not provide an appropriate dependence between \( (\eta_n^f, \eta_n^c) \).

### 3.2. Maximally coupled resampling

Let \( n \geq 1 \), \( B \in \mathcal{X} \times \mathcal{X} \), and \( \mu \in \mathcal{P}(X \times X) \), and define the probability measure

\[
\Phi_n^B(\mu)(B) = \mu \left( \left( F_{n-1,\mu,f} \wedge F_{n-1,\mu,c} \right) \tilde{M}_n(B) \right) + \left( 1 - \mu \left( F_{n-1,\mu,f} \wedge F_{n-1,\mu,c} \right) \right) \\
\times (\mu \otimes \mu) \left( \left( \mathcal{F}_{n-1,\mu,f} \otimes \mathcal{F}_{n-1,\mu,c} \right) \tilde{M}_n(B) \right),
\]

where for \( (x, y) \in \mathcal{X} \times \mathcal{X} \),

\[
\mathcal{F}_{n-1,\mu,f}(x, y) = \frac{F_{n-1,\mu,f}(x, y) - \{ F_{n-1,\mu,f}(x, y) \wedge F_{n-1,\mu,c}(x, y) \}}{\mu(F_{n-1,\mu,f} - \{ F_{n-1,\mu,f} \wedge F_{n-1,\mu,c} \} )},
\]

\[
\mathcal{F}_{n-1,\mu,c}(x, y) = \frac{F_{n-1,\mu,c}(x, y) - \{ F_{n-1,\mu,f}(x, y) \wedge F_{n-1,\mu,c}(x, y) \}}{\mu(F_{n-1,\mu,c} - \{ F_{n-1,\mu,f} \wedge F_{n-1,\mu,c} \} )},
\]

\[
F_{n-1,\mu,f}(x, y) = \mathcal{G}_{n-1,\mu,f}(x) \otimes 1,
\]

\[
F_{n-1,\mu,c}(x, y) = 1 \otimes \mathcal{G}_{n-1,\mu,c}(y),
\]

\[
\mathcal{G}_{n-1,\mu,f}(x) = \frac{G_{n-1}(x)}{\mu(G_{n-1} \otimes 1)},
\]

\[
\mathcal{G}_{n-1,\mu,c}(y) = \frac{G_{n-1}(y)}{\mu(1 \otimes G_{n-1})}.
\]
and for \((x, y), (u, v)\) \(\in X \times X\) and \(B \in \mathcal{X} \times \mathcal{X}\),
\[
\tilde{M}_n(B) ((x, y), (u, v)) = \tilde{M}_n(B)(x, v).
\]

Now set \(\tilde{\eta}_{n}^M = \tilde{\eta}_0\); we define, recursively, for any \(n \geq 1\), \(B \in \mathcal{X} \times \mathcal{X}\),
\[
\tilde{\eta}_{n}^M (B) = \Phi_1^M (\tilde{\eta}_{n-1}^M)(B).
\]

Note that by [22, Proposition A.1] we have, for \(B \in \mathcal{X}\),
\[
\tilde{\eta}_{n}^M (B \times \mathcal{X}) = \eta_n^f (B) \quad \text{and} \quad \tilde{\eta}_{n}^M (\mathcal{X} \times B) = \eta_n^c (B).
\] (9)

To generate an MCIRCPF, we consider the following probability measure:
\[
\mathbb{P}(d(u_{1:0}^1, \ldots, u_{n}^1:N)) = \left( \prod_{i=1}^{N} \tilde{\eta}_0(du_i^0) \right) \left( \prod_{p=1}^{n} \prod_{i=1}^{N} \Phi_p^M (\tilde{\eta}_{n-1}^{N,M})(du_i^p) \right),
\]
where for \(p \geq 1\),
\[
\tilde{\eta}_{p-1}^{N,M} (du) = \frac{1}{N} \sum_{i=1}^{N} \delta_{u^p_i}(du),
\]
and for \(s \in \{f, c\}\) we set
\[
\tilde{\eta}_{p-1}^{N,s} (dx) = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^p}(dx).
\]

The way in which one can generate from \(\mathbb{P}(d(u_{1:0}^1, \ldots, ))\) is given in Algorithm 2. To explain further, the simulation from \(\Phi_p^M (\tilde{\eta}_{p-1}^{N,M})\) consists of whether the resampled indices for the particle pair are equal (second bullet in Algorithm 2) or different (third bullet in Algorithm 2). This procedure is as in [6] and adopted in, for instance, [17] and [22]. The idea is to provide a local optimality procedure with respect to the resampling operation. This is in the sense that for any fixed \(N\), and conditional on the information generated so far, one will maximize the probability that the resampled indices are equal.

For the MCIRCPF, we present a preliminary result which will prove to be of interest. The following assumptions are used; note that in (A3) there is a metric \(d\) on \(X \times X\) implicit in the assumption.

(A1) For every \(n \geq 0\), \(G_n \in \mathcal{C}_b(X)\).

(A2) For every \(n \geq 1\), \(\tilde{M}_n\) is Feller.

(A3) \(X\) is a locally compact and separable metric space.

These assumptions are adopted because of the complexity of the operator \(\Phi_n^M\). As can be seen in Appendix C, where the proofs for the following result are given, it is nontrivial to work with \(\Phi_n^M\). To weaken these assumptions would lead to further calculations, which would essentially confirm the same result.

**Theorem 3.1.** Assume (A1)–(A3). Then for any \(\varphi \in \mathcal{C}_b(X \times X), n \geq 0, we have
\[
\tilde{\eta}_n^{N,M}(\varphi) \to \mathbb{P} \tilde{\eta}_n^M(\varphi).
\]
Algorithm 2: The MCIRCPF Algorithm.

1. For $i \in \{1, \ldots, N\}$, independently sample $u^i_0$ from $\tilde{\eta}_0(\cdot)$. Set $p = 1$.

2. For $i \in \{1, \ldots, N\}$ sample $u^i_p$ from $\Phi_p^M(\tilde{\eta}^{N,M}_{p-1})(\cdot)$. This can be achieved as follows:
   - Generate $W \sim U[0,1]$.
   - If
     \[
     W < \sum_{i=1}^{N} \left( \left( \frac{G_{p-1}(x^f_{p-1})}{\sum_{j=1}^{N} G_{p-1}(x^f_{p-1})} \right) \wedge \left( \frac{G_{p-1}(x^c_{p-1})}{\sum_{j=1}^{N} G_{p-1}(x^c_{p-1})} \right) \right)
     \]
     then generate $u^i_p$ according to
     \[
     \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left( \frac{G_{p-1}(x^f_{p-1})}{\sum_{k=1}^{N} G_{p-1}(x^f_{p-1})} \wedge \left( \frac{G_{p-1}(x^c_{p-1})}{\sum_{k=1}^{N} G_{p-1}(x^c_{p-1})} \right) \right) \tilde{M}_p\left( (x^f_{p-1}, x^c_{p-1}), \cdot \right).
     \]
   - Otherwise generate $u^i_p$ according to
     \[
     \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left( \frac{G_{p-1}(x^f_{p-1})}{\sum_{k=1}^{N} G_{p-1}(x^{f,j}_{p-1})} \wedge \left( \frac{G_{p-1}(x^c_{p-1})}{\sum_{k=1}^{N} G_{p-1}(x^{c,j}_{p-1})} \right) \right) \tilde{M}_p\left( (x^f_{p-1}, x^c_{p-1}), \cdot \right).
     \]

Set $p = p + 1$ and go to the start of 2.

What is interesting here is that Theorem 3.1 verifies what is expected, given the construction above: on the product space $X \times X$ the MCIRCPF approximates the target $\tilde{\eta}^M_\eta$. As noted in (9), $\tilde{\eta}^M_\eta$ is a coupling of $(\eta^f_\eta, \eta^c_\eta)$, but it is not the actual maximal coupling of $(\eta^f_\eta, \eta^c_\eta)$; this can be easily checked to be the case in most problems of practical interest. As is well known, the maximal coupling will maximize the probability that two random variables are equal, with specified marginals, and is the optimal coupling of two probability measures with respect to the $L_0$–Wasserstein distance. If this former property is desirable from a practical perspective, then the above algorithm should not be used. The maximally coupled resampling operation is, for a finite number of samples (particles), the optimal (in the above sense) way to couple the resampling operation, but may not lead to large sample ‘good’ couplings. This is manifested in [22], where the forward error rate (6) is lost for the diffusion problem in Section 2.3. As mentioned above, the limit is a coupling of $(\eta^f_\eta, \eta^c_\eta)$, but in general there is no reason to suspect that it is optimal in any sense.

3.3. Maximal coupling

We now present an algorithm which can sample (in the limit) from the maximal coupling of $(\eta^f_\eta, \eta^c_\eta)$. We will assume that for $s \in \{f, c\}$ the Markov kernels $M^s_\eta$, as well as $\eta^s_0$, admit a density with respect to a $\sigma$–finite measure $dx$. The densities are denoted $M^s_{\eta_0}$ and $\eta^s_0$, $s \in \{f, c\}$, and we assume that the densities can be evaluated numerically. To remove this latter requirement is left to future work.
Let \( n \geq 1 \), \( B \in \mathcal{X} \times \mathcal{X} \), and \((\mu, \nu) \in \mathcal{P}(X) \times \mathcal{P}(X)\), and define the probability measure
\[
\hat{\Phi}^C_n(\mu, \nu)(B) = \int_{\mathcal{X} \times \mathcal{X}} \mathbb{I}_B(x, y) \left[ \int_{\mathcal{X}} F_{n-1, \mu, f}(u) \wedge F_{n-1, \nu, c}(u) \delta_{(u, u)}(d(x, y)) du + \frac{1}{1 - \int_{\mathcal{X}} F_{n-1, \mu, f}(u) \wedge F_{n-1, \nu, c}(u) du} \mathcal{F}_{n-1, \nu, f}(x) \mathcal{F}_{n-1, \nu, c}(y) dxdy \right],
\]
where, for \( x \in \mathcal{X} \),
\[
\mathcal{F}_{n-1, \mu, f}(x) = F_{n-1, \mu, f}(x) - F_{n-1, \mu, f}(x) \wedge F_{n-1, \nu, c}(x),
\]
\[
\mathcal{F}_{n-1, \nu, c}(x) = F_{n-1, \nu, c}(x) - F_{n-1, \mu, f}(x) \wedge F_{n-1, \nu, c}(x),
\]
\[
F_{n-1, \mu, f}(x) = \frac{\mu(G_{n-1}M^{f, c}_n(\cdot, x))}{\mu(G_{n-1})},
\]
\[
F_{n-1, \nu, c}(x) = \frac{\nu(G_{n-1}M^{f, c}_n(\cdot, x))}{\nu(G_{n-1})}.
\]

Now set \( \tilde{\eta}^C_0 \) as the maximal coupling of \((\eta^f_0, \eta^c_0)\), and for \( B \in \mathcal{X} \times \mathcal{X} \) set
\[
\tilde{\eta}^C_n(B) = \hat{\Phi}^C_n(\eta^f_{n-1}, \eta^c_{n-1})(B).
\]
We have, for \( B \in \mathcal{X} \),
\[
\tilde{\eta}^C_n(B \times \mathcal{X}) = \eta^f_n(B) \quad \text{and} \quad \tilde{\eta}^C_n(\mathcal{X} \times B) = \eta^c_n(B).
\]
Moreover, \( \tilde{\eta}^C_n \) is the maximal coupling of \((\eta^f_n, \eta^c_n)\). To see why this is the case, we note that our assumptions imply that \((\eta^f_n, \eta^c_n)\) have densities with respect to a \(\sigma\)-finite measure and that these densities are simply \((F_{n-1, \eta^f_{n-1}, f}, F_{n-1, \eta^c_{n-1}, c})\); hence the expression \( \hat{\Phi}^C_n(\eta^f_{n-1}, \eta^c_{n-1})(\cdot) \) corresponds exactly to the definition of the maximal coupling of \((\eta^f_n, \eta^c_n)\).

To generate an MCPF, we consider the following probability measure:
\[
\mathbb{P}(d(u_0^{1:N}, \ldots, u_n^{1:N})) = \left( \prod_{i=1}^{N} \tilde{\eta}^C_0(du^i_0) \right) \left( \prod_{p=1}^{N} \prod_{i=1}^{N} \hat{\Phi}^C_p(\eta^f_{p-1, \eta^f_{n-1}}, \eta^c_{p-1, \eta^c_{n-1}})(du^i_p) \right),
\]
where for \( p \geq 1, s \in \{f, c\}, \)
\[
\eta^s_{p-1}(dx) = \frac{1}{N} \sum_{i=p}^{N} \delta_{x^i}(dx).
\]
For \( \mu \in \mathcal{P}(X), s \in \{f, c\}, B \in \mathcal{X}, n \geq 1 \), we set
\[
\Phi^s_n(\mu)(B) = \frac{\mu(G_{n-1}M^{s}_n(B))}{\mu(G_{n-1})}.
\]

In Algorithm 3 we present how one can simulate from \( \mathbb{P}(d(u_0^{1:N}, \ldots, \eta^c_n)) \). We remark that as \( M^{s}_n, n \geq 1, \) and \( \eta^O_s, s \in \{f, c\}, \) can be evaluated numerically, we can sample from \( \tilde{\eta}^C_0 \) and \( \hat{\Phi}^C_n(\eta^f_{n-1}, \eta^c_{n-1}) \), the latter of which is the maximal coupling of \( \Phi^f_n(\eta^f_{n-1}) \) and \( \Phi^c_n(\eta^c_{n-1}) \), using the algorithm in [29], which is presented in the bullet points in Algorithm 3. The algorithm
\textbf{Algorithm 3:} The MCPF Algorithm.

1. For $i \in \{1, \ldots, N\}$, independently sample $u_i^f$ from $\tilde{\eta}_0(\cdot)$. Set $p = 1$.

2. For $i \in \{1, \ldots, N\}$ sample $u_p^f$ from $\tilde{\Phi}_p^{N,f}(\eta_{p-1}^N, \eta_{p-1}^N, \eta_{p-1}^N)(\cdot)$. This can be achieved as follows:
   - Sample $X_p^f \sim \Phi_p^{N,f}(\eta_{p-1}^N)$ and $W|X_p^f \sim \mathcal{U}_{[0,F_{p-1}^{N,c}]}(x_p^f)$. If $W < F_{p-1}^{N,c}(x_p^f)$, output $(x_p^f, x_p^c)$, otherwise go to the next step.
   - Sample $X_p^c \sim \Phi_p^{N,c}(\eta_{p-1}^N)$ and $\tilde{W}|X_p^c \sim \mathcal{U}_{[0,F_{p-1}^{N,c}]}(x_p^c)$, if $\tilde{W} > F_{p-1}^{N,c}(x_p^c)$, output $(x_p^f, x_p^c)$, otherwise repeat this step again.

Set $p = p + 1$ and go to the start of 2.

in [29] is a rejection sampler, which would add a random running time element per time-step, which may not be desirable for some applications. As noted in [18], it can be shown that the expected running time is at most 2, but that the variance of the running time depends upon the closeness, in terms of the total variation distance, of the two probability measures that one is trying to couple. In general, in the limiting case, we expect that the total variation distance between $(\eta^f_n, \eta^c_n)$ will be small (in the diffusion case, under certain assumptions, it can be $O(\Delta t)$) and hence that the rejection sampler will perform well.

### 3.4. Wasserstein coupled resampling

We describe the WCPF used in [20]. For this case we restrict our attention to the case that $X = \mathbb{R}$. The reason for this simplification is that the approach is based upon an optimal coupling result which appears to be explicit only in the one-dimensional case. In the approach to be considered, one might consider the multi-dimensional case using, for instance Hilbert space-filling curves, but we have not tested such a method empirically and it is unclear how useful it is in practice. It is explicitly assumed that the cumulative distribution function (CDF) and its inverse associated to the probability

$$\eta_n^s(B) = \eta_n^s(G_n\mathbb{1}_B) / \eta_n^s(G_n),$$

where $s \in \{f, c\}$, $n \geq 0$, and $B \in \mathcal{X}$, exist and are continuous functions. We denote the CDF (resp. inverse) of $\eta_n^f$ as $F_{\eta_n^f}$ (resp. $F_{\eta_n^f}^{-1}$). In general we write probability measures on $X$ for which the CDF and inverse are well-defined as $\mathcal{P}_F(X)$ with the associated CDF $F_\mu$. Let $n \geq 1$, $B \in \mathcal{X} \times \mathcal{X}$, and $\mu, v \in \mathcal{P}_F(X)$, and define the probability measure

$$\tilde{\Phi}_n^W(\mu, v)(B) = \int_{X \times X} \left( \int_0^1 \delta_{[F_\mu^{-1}(w), F_v^{-1}(w)]} (du) dw \right) \tilde{M}_n(B)(u).$$

Consider the joint probability measure on $(X \times X)^{n+1}$ given by

$$\mathbb{P}(du_0, \ldots, u_n) = \tilde{\eta}_0(du_0) \prod_{p=1}^n \tilde{\Phi}_p^{N,f}(\eta_{p-1}^N, \eta_{p-1}^N)(du_p).$$
Algorithm 4: The WCPF Algorithm.

1. For $i \in \{1, \ldots, N\}$, independently sample $u^i_0$ from $\tilde{\eta}_0(\cdot)$. Set $p = 1$.

2. For $i \in \{1, \ldots, N\}$ sample $u^i_p$ from $\tilde{\Phi}_p^{W}(\tilde{\eta}_{p-1}^{N,f}, \tilde{\eta}_{p-1}^{N,c})(\cdot)$. This can be achieved as follows:
   - Sort the samples $(x^1_{p-1}, \ldots, x^N_{p-1})$ to construct an approximation of the CDFs $\tilde{\eta}_{p-1}^{f,s}$, $s \in \{f, c\}$.
   - Generate, by inversion with the same uniform, from the empirical CDFs $\tilde{\eta}_{p-1}^{N,f,s}$. Given the two samples, generate from $\tilde{M}_p$.

Set $p = p + 1$ and go to the start of 2.

It is easily checked that the marginal of $x_n^f$ (resp. $x_n^c$) is $\eta_n^f$ (resp. $\eta_n^c$). Denote by $\tilde{\eta}_n^W$ the marginal of $u_n$ induced by this joint probability measure.

To generate a WCPF, we consider the following:

$$
\mathbb{P}(d(u^{1:N}_0, \ldots, u^{1:N}_n)) = \left( \prod_{i=1}^{N} \tilde{\eta}_0(du^i_0) \right) \left( \prod_{p=1}^{n} \prod_{i=1}^{N} \tilde{\Phi}_p^{W}(\tilde{\eta}_{p-1}^{N,f}, \tilde{\eta}_{p-1}^{N,c})(du^i_p) \right),
$$

where for $p \geq 1$, $s \in \{f, c\}$,

$$
\tilde{\eta}_{p-1}^{N,s}(dx) = \sum_{i=1}^{N} \frac{G_{p-1}(x^i_{p-1})}{\sum_{j=1}^{N} G_{p-1}(x^j_{p-1})} \delta_{x^i_{p-1}}(dx).
$$

As before, for $p \geq 1$, $s \in \{f, c\}$,

$$
\eta_{p-1}^{N,s}(dx) = \frac{1}{N} \sum_{i=1}^{N} \delta_{x^i_{p-1}}(dx),
$$

and for $p \geq 0$,

$$
\tilde{\eta}_p^{N,W}(du) = \frac{1}{N} \sum_{i=1}^{N} \delta_{u^i_p}(du).
$$

In Algorithm 4 we present how one can simulate from $\mathbb{P}(d(u^{1:N}_0, \ldots ))$. It should be noted that whilst the sorting operation in Algorithm 4 is $O(N \log(N))$ in the worst case, it is seen in [20] that this step is often $O(N)$ in practice and is less expensive than the sampling of $\tilde{M}_p$ in the diffusion case.

4. Theoretical results

This section is split into two. We give our CLTs in Section 4.1 and bounds on the asymptotic variance in Section 4.2.
4.1. Central limit theorems

Define the sequence of nonnegative kernels \( \{Q^s_n\}_{n \geq 1} \), \( s \in \{f, c\} \), by

\[
Q^s_n(x, dy) = G_{n-1}(x)M^s_n(dy),
\]

for \( B \in \mathcal{X} \), \( x, y \in X \), define

\[
Q^s_{p,n}(B)(x_p) = \int_{\mathcal{X}^{n-p}} \prod_{q=p}^{n-1} Q^s_{q+1}(x_q, dx_{q+1})
\]

for \( 0 \leq p < n \); in the case \( p = n \), let \( Q^s_{p,n} \) be the identity operator. Now, for \( 0 \leq p < n \), \( s \in \{f, c\} \),

\[
D^s_{p,n}(B)(x_p) = \frac{Q^s_{p,n}(\mathbb{I}_B - \eta^s_n(B))}{\eta^s_n(Q^s_{p,n}(1))};
\]

in the case \( p = n \), \( D^s_{p,n}(B)(x) = \mathbb{I}_B - \eta^s_n(B) \).

We then have our CLTs, where the proofs are given in Appendix A. Recall that for the MCPF \( M^s_n \) is used as a notation for the kernel and density of \( M^s_n \).

**Theorem 4.1.** The following statements hold.

1. For the IRCPF: For any \( \varphi \in \mathcal{B}_b(X) \), \( n \geq 0 \), we have

\[
\sqrt{N}[(\bar{\eta}^N_n - \tilde{\eta}^I_n)(\varphi \otimes 1 - 1 \otimes \varphi)] \Rightarrow \mathcal{N}(0, \sigma^2_{n,I}(\varphi)),
\]

where

\[
\sigma^2_{n,I}(\varphi) = \sum_{p=0}^{n} \tilde{\eta}^I_p((D^f_{p,n}(\varphi) \otimes 1 - 1 \otimes D^c_{p,n}(\varphi)))^2).
\]

2. For the MCIRCPF: Assume (A1)–(A3); then for any \( \varphi \in \mathcal{C}_b(X) \), \( n \geq 0 \), we have

\[
\sqrt{N}[(\bar{\eta}^N_n - \bar{\eta}^M_n)(\varphi \otimes 1 - 1 \otimes \varphi)] \Rightarrow \mathcal{N}(0, \sigma^2_{n,M}(\varphi)),
\]

where

\[
\sigma^2_{n,M}(\varphi) = \sum_{p=0}^{n} \bar{\eta}^M_p((D^f_{p,n}(\varphi) \otimes 1 - 1 \otimes D^c_{p,n}(\varphi)))^2).
\]

3. For the MCPF: Suppose that for \( s \in \{f, c\} \), \( n \geq 1 \), \( M^s_n \in \mathcal{B}_b(X \times X) \). Then for any \( \varphi \in \mathcal{B}_b(X) \), \( n \geq 0 \), we have

\[
\sqrt{N}[(\bar{\eta}^N_n - \bar{\eta}^C_n)(\varphi \otimes 1 - 1 \otimes \varphi)] \Rightarrow \mathcal{N}(0, \sigma^2_{n,C}(\varphi)),
\]

where

\[
\sigma^2_{n,C}(\varphi) = \sum_{p=0}^{n} \bar{\eta}^C_p((D^f_{p,n}(\varphi) \otimes 1 - 1 \otimes D^c_{p,n}(\varphi)))^2).
\]

For Wasserstein resampling, the following result is established in [20].

**Theorem 4.2.** Suppose that \( \bar{M}_n, M^s_n, s \in \{f, c\} \), are Feller for every \( n \geq 1 \) and \( G_n \in \mathcal{C}_b(X) \) for every \( n \geq 0 \). Then for any \( \varphi \in \mathcal{C}_b(X) \), \( n \geq 0 \), we have

\[
\sqrt{N}[(\bar{\eta}^N_n - \bar{\eta}^W_n)(\varphi \otimes 1 - 1 \otimes \varphi)] \Rightarrow \mathcal{N}(0, \sigma^2_{n,W}(\varphi)),
\]

where

\[
\sigma^2_{n,W}(\varphi) = \sum_{p=0}^{n} \bar{\eta}^W_p((D^f_{p,n}(\varphi) \otimes 1 - 1 \otimes D^c_{p,n}(\varphi)))^2).
\]
where

\[ \sigma^2_{n,W}(\varphi) = \sum_{p=0}^{n} \tilde{\eta}_p^W((D_{p,n}^f(\varphi) \otimes 1 - 1 \otimes D_{p,n}^c(\varphi))^2). \]

**Remark 4.1.** One can also prove a multivariate CLT using the Cramér–Wold device. Consider \( 1 \leq t < +\infty, (\varphi_1, \ldots, \varphi_t, \psi_1, \ldots, \psi_t) \in B_0(X)^{2t} \) (or if required \( C_b(X)^{2t} \)). Consider the \( t \times t \) positive definite and symmetric matrix \( \Sigma^s_n(\varphi_{1:t}, \psi_{1:t}) \), \( s \in \{ I, M, C, W \} \), with \((i,j)\)th entry denoted \( \Sigma^s_{n,(i,j)}(\varphi_{1:t}, \psi_{1:t}) \). Using the Cramér–Wold device under the various assumptions of each the algorithms one can easily deduce that for each \( s \in \{ I, M, C, W \} \),

\[ \sqrt{n}((\tilde{\eta}_{n}^{N,s} - \tilde{\eta}_{n}^s)(\varphi_1 \otimes 1 - 1 \otimes \psi_1), \ldots, (\tilde{\eta}_{n}^{N,s} - \tilde{\eta}_{n}^s)(\varphi_t \otimes 1 - 1 \otimes \psi_t)) \Rightarrow \mathcal{N}_t(0, \Sigma^s_n(\varphi_{1:t}, \psi_{1:t})), \]

where

\[ \Sigma^s_{n,(i,j)}(\varphi_{1:t}, \psi_{1:t}) = \sum_{p=0}^{n} \tilde{\eta}_p^s(D_{p,n}^f(\varphi_i) \otimes 1 - 1 \otimes D_{p,n}^c(\psi_j))[D_{p,n}^f(\varphi_j) \otimes 1 - 1 \otimes D_{p,n}^c(\psi_j)]. \]

**4.2. Asymptotic variance**

We now consider bounding the asymptotic variance in the general case. The result below (Theorem 4.3), will allow one to understand when one can expect time-uniformly ‘close’ errors in approximations of \( (\tilde{\eta}_{n}^{N,s} - \tilde{\eta}_{n}^s)(\varphi) \). The following assumptions are used, which are essentially those in [30] (see also [19]) with some additional assumptions ((H6)–(H7) below) as we are treating a more delicate case than that of [30]. The assumptions can hold on unbounded state spaces, as will be the case for our applications.

**(H1)** There exist an unbounded \( \tilde{V}: X \to [1, +\infty) \) and constants \( \delta \in (0, 1) \) and \( d \geq 1 \) with the following properties. For each \( d \in (d, +\infty) \) there exists a \( b_d < +\infty \) such that \( \forall x \in X \) and any \( s \in \{ f, c \}, \)

\[ \sup_{n \geq 1} Q_n^d(e^{\tilde{V}})(x) \leq e^{(1-\delta)\tilde{V}(x) + b_d\tilde{C}_r(x)}, \]

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

**(H2)**

1. There exists a \( C < +\infty \) such that for any \( s \in \{ f, c \} \), \( \eta_0^s(v) \leq C \), with \( v = e^{\tilde{V}} \), with \( \tilde{V} \) as in (H1).

2. For any \( r > 1 \) there exists a \( C < +\infty \) such that for any \( s \in \{ f, c \} \), \( \eta_0^s(C_r)^{-1} \leq C \).

**(H3)** With \( d \) as in (H1), for each \( d \in [d, +\infty) \), \( s \in \{ f, c \}, \)

\[ \int_{C_d} G_{n-1}(x)M_n^s(x, dy) > 0 \ \forall x \in X, n \geq 1, \]

and there exist \( \tilde{\varepsilon} > 0, v_d \in \mathcal{P}_v \) such that for each \( A \in X, s \in \{ f, c \}, \)

\[ \inf_{n \geq 1} \int_{C_d \cap A} Q_n^s(x, dy) \geq \tilde{\varepsilon} \ v_d(C_d \cap A), \ \forall x \in C_d, \]

with \( C_d \) as in (H1).
(H4) With $d$ as in (H1), and $\tilde{\epsilon}_d^-$, $\nu_d$ as in (H2), for each $d \in [d, +\infty)$ there exists $\tilde{\epsilon}_d^+ \in [\tilde{\epsilon}_d^-, +\infty)$ such that for each $A \in \mathcal{X}$, $s \in \{f, c\}$,

$$\sup_{n \geq 1} \int_{C_d \cap A} Q_n^c(x, dy) \leq \tilde{\epsilon}_d^+ \nu(C_d \cap A), \quad \forall x \in C_d,$$

with $C_d$ as in (H1).

(H5)

1. \[ \sup_{n \geq 0} \sup_{x \in X} G_n(x) < +\infty. \]

2. For any $r > 1$ there exists a $C < +\infty$ such that for any $\inf_{x \in C_r} G_0(x) \geq C$, with $C_r$ as in (H1).

(H6) With $d$ as in (H1), for each $d \in [d, +\infty)$, we have for each $n \geq 0$, $s \in \{f, c\}$, that

$$\frac{1}{G_n M_{n+1}(\|C_d\)} \in \mathcal{L}_v(X)$$

and $\sup_{n \geq 0} \max_{s \in \{f, c\}} \|1/G_n M_{n+1}(\|C_d\)} \|v < +\infty$.

(H7) Let $d$ be a given metric on $X$. For any $\xi \in (0, 1)$, there exists a $C < +\infty$ such that for $\varphi \in \mathcal{L}_{v\xi}(X) \cap \text{Lip}_{v\xi,d}(X)$, $(x, y) \in X \times X$, $s \in \{f, c\}$,

$$\sup_{n \geq 1} |Q_n^\xi(\varphi)(x) - Q_n^\xi(\varphi)(y)| \leq C \|\varphi\|_{v\xi} d(x, y) [v(x)v(y)]^\xi.$$

Let $A = \{(x, y) \in X \times X : x \neq y\}$. For $n \geq 1$, $0 \leq p \leq n$, $s \in \{f, c\}$, $x \in X$, define

$$h_{p,n}^s(x) := \frac{Q_{p,n}(1)(x)}{\eta_p^s(Q_{p,n}(1))}$$

and, with $B \in \mathcal{X}$,

$$S_{p,n}^{f,c}(B)(x) := \frac{Q_{p,n}^f(B)(x)}{Q_{p,n}^f(1)(x)} - \frac{Q_{p,n}^c(B)(x)}{Q_{p,n}^c(1)(x)}.$$

Set

$$B(n, f, c, \varphi, \xi) = \|\varphi\| \sum_{p=0}^{n-1} \rho^{n-p} \left\{ ||h_{p,n}\xi||_{v\xi} + ||\eta_n^f - \eta_n^c||(\varphi) + ||\varphi||||h_{p,n}^f - h_{p,n}^c||_{v\xi} \right\}.$$

Below is our main result, whose proof can be found in Appendix E.

**Theorem 4.3.** Assume (H1)–(H6). Then for any $\xi \in (0, 1/4)$ there exist $\rho < 1$ and $C < +\infty$ depending on the constants in (H1)–(H6) such that for any $\varphi \in \mathcal{B}_b(X)$, $n \geq 1$, $s \in \{I, M, C, W\}$, we have

$$\sigma_n^{2,s}(\varphi) \leq \tilde{\eta}_{n}^s \left( \{\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))^2 \right) + C \left\{ B(n, f, c, \varphi, \xi) + ||\varphi||^2 \sum_{p=0}^{n-1} \rho^{n-p} \left( \tilde{\eta}_{p}^s I_A (\nu \otimes \nu)^2 + \rho^{n-p} \right). \right\}.$$
Additionally assume (H7). Then if \( d^2 \in \mathcal{L}_{(v \otimes v)\xi} \), \( \xi \in (0, 1/2) \), for any \( \xi \in (0, (1 - 2\xi)/12) \) there exist \( r < 1 \) and \( C < +\infty \) depending on the constants in (H1)–(H7) such that for any \( \varphi \in \mathcal{B}_{b}(X) \cap \text{Lip}_d(X) \), \( n \geq 1 \), \( s \in \{I, M, C, W\} \), we have

\[
\sigma_n^2(\varphi) \leq \tilde{\eta}_n^s \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))^2 \right) \\
+ C \left\{ B(n, f, c, \varphi, \xi) + \|\varphi\|^2 \sum_{p=0}^{n-1} r^p \left( \tilde{\eta}_p^s(d^2(v^{4\xi} \otimes v^{8\xi})) (\rho^n - p + 1) \right) \right\}.
\]

**Remark 4.2.** The range of \( \xi \) is such that the upper bounds are finite. This is because \( \tilde{\eta}_p^s(v) \) is bounded above by a constant that does not depend on \( p \) (see [30, Proposition 1]).

The main conclusion of this theorem is that one expects that the (variances of) approximations of \( [\eta_n^f - \eta_n^c](\varphi) \) are uniformly ‘close’ in time if the following are satisfied:

1. The quantities

\[
\tilde{\eta}_n^s \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))^2 \right) \quad \text{and} \quad |[\eta_n^f - \eta_n^c](\varphi)|
\]

are small uniformly in time.

2. The differences

\[
\|h_{p,n}^f, S_{p,n}^f, c(\varphi)\|_{v^\xi} \quad \text{and} \quad \|h_{p,n}^f - h_{p,n}^c\|_{v^\xi}
\]

are small at a rate which decays more slowly than exponentially in time.

3. The quantities

\[
\tilde{\eta}_p^s(I_A(v \otimes v)^{2\xi}) \quad \text{or} \quad \tilde{\eta}_p^s(d^2(v^{4\xi} \otimes v^{8\xi}))
\]

are small at a rate which decays more slowly than exponentially in time.

We note that in 1, strictly speaking, one could have \( |[\eta_n^f - \eta_n^c](\varphi)| \) close at a rate which decays more slowly than exponentially in time, but as one requires the time-uniform closeness of \( \tilde{\eta}_n^s((\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi)))^2) \), one expects this property for the former. The use of \( A \) and \( d \) are linked to the coupling properties associated to \( \tilde{\eta}_p^s \), as we consider in the next section.

### 5. Application to partially observed diffusions

We now return to the PODDP model considered in Section 2.3. Throughout, \( 1 \leq l \leq L \) is fixed, with \( L > 1 \) fixed. We will consider the significance of the results in Theorems 4.1, 4.2, and 4.3 for each of the four CPFs. We also present an example in Section 5.6 where (H1)–(H8) can be verified. We begin with a control of the term \( B(n, l, l - 1, \varphi, \xi) \).

#### 5.1. A general result

For \( x \in X \), define the set

\[
B_l(x) := \{ y \in X : \| y - x \| > 2\Delta_l \},
\]

where \( \| \cdot \| \) is the \( L_2 \) norm. We write the transition densities with respect to Lebesgue measure \( dy \) of the diffusion as \( M(x, y) \) and the Euler approximation as \( M'(x, y) \), \( (x, y) \in X \times X \). We note
the following two equations quoted in [11]: there exist $0 < C, C' < +\infty$ (which depend on the drift and diffusion coefficients of (4)) such that for any $l \geq 0$,
\begin{align}
M(x, y) + M^I(x, y) & \leq C \exp\{-C'\|y - x\|^2\} \quad \forall (x, y) \in \mathbb{X} \times \mathbb{X}, \\
|M(x, y) - M^I(x, y)| & \leq C \Delta_l \exp\{-C'\|y - x\|^2\} \quad \forall x \in \mathbb{X}, y \in B_l(x).
\end{align}
(10)

We add an additional assumption:

(H8) For $C'$ as in (10)–(11), we have the following:

1. For any $\xi \in (0, 1/2)$, there exists a $C < +\infty$ such that for any $x \in \mathbb{X}, l \geq 0$,
   \[
   \int_{B_l(x)} v(y)^{\xi} dy \leq C v(x)^{2\xi} \int_{B_l(x)} dy.
   \]
2. For any $\xi \in (0, 1/2)$, there exists a $C < +\infty$ such that for any $x \in \mathbb{X}, l \geq 0$,
   \[
   \int_{B_l(x)} v(y)^{\xi} \exp\{-C'\|y - x\|^2\} dy \leq C v(x)^{2\xi}.
   \]

We have the following result, whose proof is in Appendix F, Section F.1.

**Proposition 5.1.** Assume (H1)–(H6), (H8). Then for any $\xi, \tilde{\xi} \in (0, 1/8) \times (0, 1/2)$ there exists a $C < +\infty$ depending on the constants in (H1)–(H6) and (H8) such that for any $\varphi \in B_l(\mathbb{X}), n \geq 1, 1 \leq l \leq L$, we have
\[
B(n, l, l - 1, \varphi, \xi) \leq C\|\varphi\|^2 \Delta_l v(x^*)^{\xi} + \tilde{\xi}.
\]

The main point of this result is that now the time-uniform coupling ability of each algorithm will now rely on the properties of the limiting coupling $\tilde{\eta}_n^\xi$.

### 5.2. IRCPF

We consider the term
\[
\tilde{\eta}_n^f\left(\left\{\varphi \otimes 1 - 1 \otimes \varphi - \left(\left(\eta_n^f - \eta_n^c\right)(\varphi)\right)\right\}^2\right)
\]
in the upper bound of Theorem 4.3. Let us suppose that $\varphi \in \mathcal{A}$, where $\mathcal{A}$ is defined below (5). One has that
\[
\tilde{\eta}_n^f\left(\left\{\varphi \otimes 1 - 1 \otimes \varphi - \left(\left(\eta_n^f - \eta_n^c\right)(\varphi)\right)\right\}^2\right) = \frac{(\eta_{n-1}^f \otimes \eta_{n-1}^c)([G_{n-1} \otimes G_{n-1}][\tilde{M}_n([\varphi \otimes 1 - 1 \otimes \varphi])])}{(\eta_{n-1}^f \otimes \eta_{n-1}^c)(G_{n-1} \otimes G_{n-1})} - [\eta_n^f - \eta_n^c](\varphi)^2.
\]

Then using $\varphi \in \mathcal{A}$ (along with the $C_2$–inequality) and [22, Lemma D.2], there is a finite constant $C < +\infty$ that depends on $n$ such that
\[
\tilde{\eta}_n^f\left(\left\{\varphi \otimes 1 - 1 \otimes \varphi - \left(\left(\eta_n^f - \eta_n^c\right)(\varphi)\right)\right\}^2\right) \leq C \left(\frac{(\eta_{n-1}^f \otimes \eta_{n-1}^c)([G_{n-1} \otimes G_{n-1}][\tilde{M}_n(\varphi)])}{(\eta_{n-1}^f \otimes \eta_{n-1}^c)(G_{n-1} \otimes G_{n-1})} + \Delta_l^2\right),
\]
where \( \bar{\varphi}(x, y) = \|x - y\|^2 \). Now, applying (6), one has the upper bound

\[
C \left( \frac{(\eta_{n-1}^f \otimes \eta_{n-1}^c) \{(G_{n-1} \otimes G_{n-1})\bar{\varphi}\}}{(\eta_{n-1}^f \otimes \eta_{n-1}^c)(G_{n-1} \otimes G_{n-1})} + \Delta_I + \Delta_{\tilde{\varphi}}^I \right).
\]

In general, there is no reason to expect that \((\eta_{n-1}^f \otimes \eta_{n-1}^c)(G_{n-1} \otimes G_{n-1})\bar{\varphi}\) is small as a function of \(\Delta_I\). This is unsurprising, because one uses an independent coupling in the resampling operation. As a result, in the sense of Section 2.3.1 for ML estimation, the IRCPF is not useful.

### 5.3. MCIRCPF

To start our discussion, suppose first that \(G_n\) is constant for each \(n \geq 0\). This represents the most favorable scenario for the MCIRCPF, because in the resampling operation, the resampled indices for each pair are equal (of course one would not use resampling in this case). For the metric in (H7), we take \(d_1\) to be the \(L_1\) - norm and set \(\varphi \in B_b(X) \cap \text{Lip}_{d_1}(X)\) (\(X = \mathbb{R}^d\)). Then setting \(\bar{\varphi}(x, y) = \|x - y\|^2\), we have

\[
\eta_n^M \left( \{\varphi \otimes 1 - 1 \otimes \varphi - (\eta_n^f - \eta_n^c)(\varphi)\}^2 \right) \leq C\|\varphi\|^2 \text{Lip} \eta_n^M(\bar{\varphi}),
\]

where \(C < +\infty\) does not depend on \(n\). Now, as \(G_n\) is constant, one can deduce (e.g. by [25, Theorem 2.7.3]; recall (D) is assumed) that

\[
\eta_n^M \left( \{\varphi \otimes 1 - 1 \otimes \varphi - (\eta_n^f - \eta_n^c)(\varphi)\}^2 \right) \leq Cn(n + 4) \exp(C'n^2)\Delta_I,
\]

where \(C, C' < +\infty\) do not depend on \(n\). Now (12) is not necessarily tight in \(n\) for every diffusion that satisfies (D); for instance if one has \(dZ_t = adt + bdW_t\), for \(b > 0, a \in (-1, 0)\), then there is no dependence on \(n\) in the right-hand side of (12). However, it is worrying that the coupling can be exponentially bad in time, in this favorable case. The main point here is that the principal source of coupling in this algorithm is \(\tilde{M}\), and if this coupling deteriorates when iterating \(M\), then for the MCIRCPF one cannot hope to obtain time-uniform couplings.

More generally, if \(G_n\) is nonconstant, we first remark that the upper bound (12) is likely to be \(O(\Delta_I^{1/2})\), as it is the resampling operation where the forward rate is lost (see [22]). Secondly, one expects to find examples where \(\sigma_n^{2,\varphi}(\varphi)\) is time-uniform or grows slowly as a function of \(n\) (due to the empirical results in [22]). However, because of the highly nonlinear and complex expression for \(\Phi_n^M\), we expect a general result to be particularly arduous to obtain; this is left to future work.

### 5.4. MCPF

We have the following result which establishes the time-uniform coupling of the MCPF. The proof can be found in Appendix F, Section F.2.

**Proposition 5.2.** Assume (H1)–(H6), (H8). Then for any \((\xi, \hat{\xi}) \in (0, 1/32) \times (0, 1/2)\) there exists a \(C < +\infty\) depending on the constants in (H1)–(H6) and (H8) such that for any \(\varphi \in B_b(X), n \geq 0, 1 \leq l \leq L\), we have

\[
\sigma_n^{2,\varphi}(\varphi) \leq C\|\varphi\|^2 \Delta_I \nu(x^*)^{32\xi + 2\hat{\xi}}.
\]
5.5. WCPF

Note that $X = \mathbb{R}$ and, for the metric in (H7), we take $d_1$ to be the $L_1$-norm. Let $\tilde{\varphi}(x, y) = (x - y)^2$. The proof of the following result, which establishes the time-uniform coupling of the WCPF, can be found in Appendix F, Section F.3.

**Proposition 5.3.** Assume (H1)–(H8). Then let $L \geq 1$ that do not depend on $l$, with $\alpha \leq |\alpha_l| < \alpha_l, \beta_l \leq \beta$. Then there exists a $C$ depending on the constants in (H1)–(H8) such that for any $\varphi \in B_\delta(X) \cap Lip_{d_1}(X)$, $n \geq 0$, $1 \leq l \leq L$, we have

$$\sigma_n^{2,W}(\varphi) \leq C\|\varphi\|_{Lip}\|\tilde{\varphi}\|_{\tilde{\varphi}}(\Delta_t)^{1/(1+\lambda)}\lambda(x^*)^{20\xi + (16\lambda + 1)\xi + 2\epsilon}/(1+\lambda).$$

As $\lambda$ is greater than 0 in Proposition 5.3 (and can be made close to zero), one almost has the (time-uniform) forward error rate for the WCPF. We believe that $\lambda = 0$ is the desired case and discuss strategies to establish this in Remark F.2 in Appendix F, Section F.3.

5.6. Example

We consider $X = \mathbb{R}$, the diffusion $dZ_t = -\frac{3}{2}Z_t dt + dW_t$ and $Y = \{0, 1\}$ and

$$G(x, y) = \left(\frac{be^x + a}{1 + e^x}\right)^y\left(1 - \left(\frac{be^x + a}{1 + e^x}\right)^{1-y}\right)$$

for $y \in Y$ and any $0 < a < b < 1$. The metric in (H7) is the $L_1$-norm. In practice the CPFs can only be run (with currently available computational power) with $L \leq 20$. So we will assume that there is an $L^* > L$ for which one cannot run the algorithm (say $L^* = 50$). This will reduce the complexity of the forthcoming discussion. For the Euler discretization (although of course it is not required here) one can determine that the transition kernel $M_l(x, y)$ is the density of a $\mathcal{N}(\alpha_l; x, \beta_l)$ distribution with $\alpha_l = (1 - (3/2)\Delta_t)\Delta_t^{-1}$ and

$$\beta_l = \Delta_t(1 - (1 - (3/2)\Delta_t)2\Delta_t^{-1})/(1 - (1 - (3/2)\Delta_t)^2).$$

Thanks to our assumption concerning $L^*$, one can easily find constants $0 < \alpha < \alpha_l < 1, 0 < \beta_l < \beta_l < 1$ that do not depend on $l$, with $\alpha \leq |\alpha_l| < \alpha_l, \beta_l \leq \beta_l$ for each $l \in \{0, \ldots, L\}$.

To verify (H1) we use, as in [31, Example 4.2.1], the Lyapunov function

$$V(x) = \frac{1}{2(1 + \delta_0)}x^2 + 1$$

for some $\delta_0 > 0$. One can show that for any $1 \leq l \leq L, n \geq 1, y \in X$,

$$Q_n'(e^{V}(y)) \leq \exp\left\{\frac{\alpha_l^2(1 + \delta_0)}{1 + \delta_0 - \beta_l}(V(y) - 1) + C\right\},$$

where $[\alpha_l^2(1 + \delta_0)]/[1 + \delta_0 - \beta_l] < 1$ for every $0 \leq l \leq L, \delta_0 > 0$, and $0 < C < +\infty$ is a constant that does not depend on $l, n, \delta_0$ and can be made arbitrarily large. One can easily verify
(H1) for any \( \delta_0 > 0 \) with \( d > 1 \) large enough. (H2)–(H5) follow by elementary calculations associated to the choice of \( \hat{V}, M_l, \) and \( G_n \) and are omitted.

To verify (H6) as \( G_n \) is bounded below, we consider \( M_l(\hat{V}) \) and suppose \( x > 0, l \geq 1, d > 1 \) (the case \( x \leq 0, l = 0 \) can be verified in a similar manner). We have for any \( x > 0 \) that

\[
\nu(x)M_l(\hat{V})(x) \geq C \exp \left\{ \frac{x^2}{2} \left( \frac{1}{\delta_0 + 1} - \frac{\alpha_l^2}{\beta_l} \right) - \frac{\alpha_l x \tilde{d}}{\beta_l} \right\},
\]

where \( \tilde{d} = \sqrt{(d - 1)(1 + \delta_0)} \). As \( \delta_0 \) is arbitrary one can choose \( \delta_0 > 0 \) so that

\[
\frac{1}{\delta_0 + 1} - \frac{\alpha_l^2}{\beta_l} \geq 0
\]

for any \( 1 \leq l \leq L \). Checking calculations on a computer for \( L^* = 50 \) yields that \( 0 < \delta_0 \leq 5 \) suffices.

For (H7) one has the following result, whose proof is in Appendix F, Section F.4.

**Lemma 5.1.** (H7) is satisfied in this example.

To verify (H8), we shall suppose that \( C' \) in (14) is at least \( 1 + 1/(2(1 + \delta_0)) \) (i.e. \( C' > 1/12 \)). It is noted that it is nontrivial to check the value of \( C' \) in [1], but such a choice seems reasonable. In the given situation, it is then straightforward to verify (H8).

### 6. Summary

We have considered CLTs for coupled particle filters and the associated asymptotic variance for applications. The main message is that it can be nontrivial to construct CPFs for which one inherits the appropriate ‘closeness’ uniformly in time. The MCPF and WCPF seem to be the best options, but suffer from the fact that (at least as considered in this paper) one either requires the density of the transition and has a random running time per time step (MCPF), or is constrained to the case that \( X \) is one-dimensional (WCPF). Nonetheless these are still useful algorithms when they can be implemented.

There are a number of possible extensions of this work. First, one could extend these results to the context of normalization constant estimation (e.g. [21]) and the associated asymptotic variance. Second, one could perform a more in-depth analysis and implementation of the MCPF, which to our knowledge has not been done in the literature.

### Appendix A. Common proofs for the CLT

In each of the appendices, the proof of the main result appears at the beginning. The associated technical results are given in such a way that, for the overall proof to be understood, they should be read in order.

For \( p \geq 0, \varphi \in \mathcal{B}_b(X), s \in \{f, c\}, \)

\[
V_p^{N,s}(\varphi) = \sqrt{N} [\eta_p^{N,s} - \Phi_p^{s}(\eta_{p-1}^{N,s})](\varphi),
\]

with the convention that \( V_0^{N,s}(\varphi) = \sqrt{N} [\eta_0^{N,s} - \eta_0](\varphi) \). For \( p \geq 0, p \leq n, n > 0, \varphi \in \mathcal{B}_b(X), s \in \{f, c\}, \)

\[
R_p^{N,s}(\varphi) = \frac{\eta_p^{N,s}(D_p^{s,n}(\varphi))}{\eta_p^{N,s}(G_p)} [\eta_p^{s}(G_p) - \eta_p^{N,s}(G_p)].
\]
We adopt the convention $R_{n+1}^{N,s}(\varphi) = 0$.

Now we note that, using the calculations in [2] and [10], for $t \in \{I, M, C, W\}$ we have

$$\sqrt{N}[\tilde{\eta}^N_{n,t} - \tilde{\eta}^f_n](\varphi \otimes 1 - 1 \otimes \varphi) = \sum_{p=0}^n [V_p^{N,f}(D^f_{p,n}(\varphi)) + \sqrt{N}R_{p+1}^{N,f}(D^f_{p,n}(\varphi))]$$

$$- \sum_{p=0}^n (V_p^{N,c}(D^c_{p,n}(\varphi)) + \sqrt{N}R_{p+1}^{N,c}(D^c_{p,n}(\varphi))). \tag{13}$$

Proof of Theorem 4.1. The proof follows immediately from (13), Lemma A.1, and Proposition A.1.

A.1. Technical results

The following results can be established for all four algorithms, but as the result for the Wasserstein method is given in [20], only the other three cases are considered. When required we specify particular conditions required for a given algorithm. By default proofs are specified for the IRCPF case, and the MCIRCPF and MCPF are mentioned where required.

Lemma A.1. For $\varphi \in \mathcal{B}_b(X)$, $n > 0$, $0 \leq p \leq n$, $s \in \{f, c\}$, we have

$$\sqrt{N}R_{p+1}^{N,s}(\varphi) \to 0.$$

Proof. By Proposition B.1 (for MCIRCPF this is [22, Proposition C.6], and for MCPF it is Proposition D.1), $\eta_{p}^{N,s}(G_{p})$ converges in probability to a well-defined limit. Hence we need only to show that

$$\sqrt{N}\eta_{p}^{N,s}(D^{s}_{p,n}(\varphi))[\eta_{p}^{s}(G_{p}) - \eta_{p}^{N,s}(G_{p})]$$

will converge in probability to zero. By Cauchy–Schwarz,

$$\sqrt{N}\mathbb{E}[|\eta_{p}^{N,s}(D^{s}_{p,n}(\varphi))[\eta_{p}^{s}(G_{p}) - \eta_{p}^{N,s}(G_{p})]|]$$

$$\leq \sqrt{N}\mathbb{E}[|\eta_{p}^{N,s}(D^{s}_{p,n}(\varphi))|^{2}]^{1/2}\mathbb{E}[|\eta_{p}^{s}(G_{p}) - \eta_{p}^{N,s}(G_{p})|^{2}]^{1/2}.$$

Applying Proposition B.1 (for MCIRCPF [22, Proposition C.6], for MCPF Proposition D.1), it easily follows that there is a finite constant $C < +\infty$ that does not depend upon $N$, such that

$$\sqrt{N}\mathbb{E}[|\eta_{p}^{N,s}(D^{s}_{p,n}(\varphi))[\eta_{p}^{s}(G_{p}) - \eta_{p}^{N,s}(G_{p})]|] \leq \frac{C}{\sqrt{N}}.$$

This bound allows one to easily conclude.

Lemma A.2. For $\varphi \in \mathcal{B}_b(X)$, $p \geq 0$, $s \in \{f, c\}$, we have

$$\mathbb{E}[V_{p}^{N,s}(\varphi)] = 0,$$

$$\lim_{N \to +\infty} \mathbb{E}[V_{p}^{N,s}(\varphi)^{2}] = \eta_{p}^{s}((\varphi - \eta_{p}^{s}(\varphi))^{2}).$$

Proof. $\mathbb{E}[V_{p}^{N,s}(\varphi)] = 0$ follows immediately from the expression, so we focus on the second property. We have

$$\mathbb{E}[V_{p}^{N,s}(\varphi)^{2}] = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}[(\varphi(\chi_{p}^{1,s}) - \Phi_{p}^{s}(\eta_{p-1})(\varphi))^{2}]$$

$$= \mathbb{E}[(\varphi(\chi_{p}^{1,s}))^{2}] - \mathbb{E}[\Phi_{p}^{s}(\eta_{p-1})(\varphi)^{2}].$$
\( \Phi^s_p(\eta^{N,s}_{p-1})(\psi) \) is a bounded random quantity, and moreover, by Proposition B.1 (for MCIRCPF [22, Proposition C.6], for MCPF Proposition D.1), it converges in probability to \( \eta^s_p(\psi) \). Hence by [3, Theorem 25.12], \( \lim_{N \to +\infty} \mathbb{E}[\Phi^s_p(\eta^{N,s}_{p-1})(\psi)^2] = \eta^s_p(\psi)^2 \). Hence we consider

\[
\mathbb{E}[\varphi(X^{1,s}_p)^2] = \mathbb{E}\left[ \frac{1}{N} \sum_{i=1}^{N} \varphi(X^{i,s}_p)^2 - \eta^s_p(\psi)^2 \right] + \eta^s_p(\psi)^2.
\]

By Jensen,

\[
\mathbb{E}\left[ \frac{1}{N} \sum_{i=1}^{N} \varphi(X^{i,s}_p)^2 - \eta^s_p(\psi)^2 \right] \leq \left( \mathbb{E}\left[ \frac{1}{N} \sum_{i=1}^{N} \varphi(X^{i,s}_p)^2 - \eta^s_p(\psi)^2 \right]^2 \right)^{1/2},
\]

and hence we conclude via Proposition B.1 (for MCIRCPF [22, Proposition C.6], for MCPF Proposition D.1) that

\[
\lim_{N \to +\infty} \mathbb{E}[\varphi(X^{1,s}_p)^2] = \eta^s_p(\psi)^2.
\]

The result follows. \( \square \)

**Lemma A.3.** The following statements hold.

1. For the IRCPF: For \((\varphi, \psi) \in \mathcal{B}_b(\mathbb{X})^2, p \geq 0, \) we have

\[
\lim_{N \to +\infty} \mathbb{E}[V^N(f)(\psi)V^{N,c}_p(\psi)] = \gamma^f_p(\psi \otimes \psi) - \eta^c_p(\psi).
\]

2. For the MCIRCPF: Assume (A1)–(A3), \((\varphi, \psi) \in \mathcal{C}_b(\mathbb{X})^2, p \geq 0; \) then we have

\[
\lim_{N \to +\infty} \mathbb{E}[V^N(f)(\psi)V^{N,c}_p(\psi)] = \gamma^M_p(\psi \otimes \psi) - \eta^c_p(\psi).
\]

3. For the MCPF: Suppose that for \(s \in \{f, c\}, n \geq 1, M^s_n \in \mathcal{B}_b(\mathbb{X} \times \mathbb{X}). \) For \((\varphi, \psi) \in \mathcal{B}_b(\mathbb{X})^2, p \geq 0, \) we have

\[
\lim_{N \to +\infty} \mathbb{E}[V^N(f)(\psi)V^{N,c}_p(\psi)] = \eta^c_p(\psi \otimes \psi) - \eta^c_p(\psi).
\]

**Proof.** For all methods, we have

\[
\mathbb{E}[V^N(f)(\psi)V^{N,c}_p(\psi)] = N\left( \mathbb{E}[\eta^N_p(\varphi)\eta^{N,c}_p(\psi)] - \mathbb{E}[\eta^N_p(\varphi)\Phi_p(\eta^{N,c}_p)(\psi)] \right)
\]

\[
- \mathbb{E}[\Phi_p(\eta^{N,f}_{p-1})(\varphi)\Phi_p(\eta^{N,c}_p)(\psi)] + \mathbb{E}[\Phi_p(\eta^{N,f}_{p-1})(\varphi)\Phi_p(\eta^{N,c}_p)(\psi)]\right).
\]

\[
= N\left( \mathbb{E}[\eta^N_p(\varphi)\eta^{N,c}_p(\psi)] - \mathbb{E}[\Phi_p(\eta^{N,f}_{p-1})(\varphi)\Phi_p(\eta^{N,c}_p)(\psi)] \right).
\]  

(14)

We consider calculations to verify Parts 1–3 in the statement of the Lemma by using (14).

**Proof of Part 1:** Now

\[
N\mathbb{E}[\eta^N_p(\varphi)\eta^{N,c}_p(\psi)] = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E}[\varphi(X^{i,f}_p)\psi(X^{i,c}_p)]
\]

\[
= \mathbb{E}[\Phi_p(\eta^{N,f}_{p-1})(\varphi)\Phi_p(\eta^{N,c}_p)(\psi)] + (N-1)\mathbb{E}[\Phi_p(\eta^{N,f}_{p-1})(\varphi)\Phi_p(\eta^{N,c}_p)(\psi)].
\]
Thus,
\[ \mathbb{E}[V_p^{N,f}(\psi) V_p^{N,c}(\psi)] = \mathbb{E}[\tilde{\Phi}_p^{N,f}(\eta_{p-1}^{N,f} \otimes \eta_{p-1}^{N,c})(\psi) - \mathbb{E}[\Phi_p^{f}(\eta_{p-1}^{N,f})(\psi) \Phi_p^{c}(\eta_{p-1}^{N,c})(\psi)]. \]

Similarly, via Proposition B.1 and [3, Theorem 25.12],
\[ \tilde{\Phi}_p^{f}(\eta_{p-1}^{N,f} \otimes \eta_{p-1}^{N,c})(\psi) \] is a bounded random quantity, and by Propositions B.1 and B.2 it converges in probability to \( \eta_p^f(\psi) \); hence by [3, Theorem 25.12],
\[ \lim_{N \to +\infty} \mathbb{E}[\tilde{\Phi}_p^{f}(\eta_{p-1}^{N,f} \otimes \eta_{p-1}^{N,c})(\psi)] = \eta_p^f(\psi). \]

Similarly, via Proposition B.1 and [3, Theorem 25.12],
\[ \lim_{N \to +\infty} \mathbb{E}[\tilde{\Phi}_p^{c}(\eta_{p-1}^{N,c})(\psi)] = \eta_p^c(\psi), \]
and hence we can conclude the proof of Part 1.

**Proof of Part 2:** Using similar calculations as for Part 1 we have
\[ \mathbb{E}[V_p^{N,f}(\psi) V_p^{N,c}(\psi)] = \mathbb{E}[\tilde{\Phi}_p^{M,f}(\eta_{p-1}^{N,f} \otimes \eta_{p-1}^{N,c})(\psi) - \mathbb{E}[\Phi_p^{f}(\eta_{p-1}^{N,f})(\psi) \Phi_p^{c}(\eta_{p-1}^{N,c})(\psi)]. \]

The proof of Theorem 3.1 establishes that
\[ \tilde{\Phi}_p^{M}(\eta_{p-1}^{N,f})(\psi) \to \eta_p^{M}(\psi). \]

The proof can then be completed in much the same way as for Part 1.

**Proof of Part 3:** Using similar calculations as for Part 1 we have
\[ \mathbb{E}[V_p^{N,f}(\psi) V_p^{N,c}(\psi)] = \mathbb{E}[\tilde{\Phi}_p^{C,f}(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\psi) - \mathbb{E}[\Phi_p^{f}(\eta_{p-1}^{N,f})(\psi) \Phi_p^{c}(\eta_{p-1}^{N,c})(\psi)]. \]

The proof of Theorem D.1 establishes that
\[ \tilde{\Phi}_p^{C}(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\psi) \to \eta_p^{C}(\psi). \]
The proof can then be completed in much the same way as for (i).

For \( p \geq 0, \psi \in B_b(X) \), define
\[ V_p^N(\psi) = V_p^{N,f}(\psi) - V_p^{N,c}(\psi). \]

**Proposition A.1.** The following statements hold.

1. For the IRCPF: Let \( n \geq 0 \); then for any \( (\psi_0, \ldots, \psi_n) \in B_b(X)^{n+1}, (\psi_0, \ldots, \psi_n) \in B_b(X)^{n+1}, \)
\[ (V_0^N(\psi_0), \ldots, V_n^N(\psi_n)) \]
converges in distribution to an \((n+1)-\)dimensional Gaussian random variable with zero mean and diagonal covariance matrix, with the \( p \)th diagonal entry for \( p \in \{0, \ldots, n\} \) given by
\[ \eta_p^f(\{(\psi_p \otimes 1 - 1 \otimes \psi_p) - \eta_p^f(\psi_p \otimes 1 - 1 \otimes \psi_p)^2\}). \]
2. For the MCIRCPF: Assume (A1)–(A3); then for \( n \geq 0 \) and any \( (\varphi_0, \ldots, \varphi_n) \in \mathcal{C}_b(\mathbf{X})^{n+1}, \) \( (\psi_0, \ldots, \psi_n) \in \mathcal{C}_b(\mathbf{X})^{n+1}, \)

\[
(V^N_0(\varphi_0, \psi_0), \ldots, V^N_n(\varphi_n, \psi_n))
\]

converges in distribution to an \( (n+1) \)-dimensional Gaussian random variable with zero mean and diagonal covariance matrix, with the \( p \)th diagonal entry for \( p \in \{0, \ldots, n\} \) given by

\[
\tilde{\eta}^M_p ((\varphi_p \otimes 1 - 1 \otimes \psi_p) - \tilde{\eta}^M_p (\varphi_p \otimes 1 - 1 \otimes \psi_p))^2).
\]

3. For the MCPF: Suppose that for \( s \in \{f, c\}, n \geq 1, M^s_n \in \mathcal{B}_b(\mathbf{X} \times \mathbf{X}); \) then for \( n \geq 0 \) and any \( (\varphi_0, \ldots, \varphi_n) \in \mathcal{B}_b(\mathbf{X})^{n+1}, (\psi_0, \ldots, \psi_n) \in \mathcal{B}_b(\mathbf{X})^{n+1}, \)

\[
(V^N_0(\varphi_0, \psi_0), \ldots, V^N_n(\varphi_n, \psi_n))
\]

converges in distribution to an \( (n+1) \)-dimensional Gaussian random variable with zero mean and diagonal covariance matrix, with the \( p \)th diagonal entry for \( p \in \{0, \ldots, n\} \) given by

\[
\tilde{\eta}^C_p ((\varphi_p \otimes 1 - 1 \otimes \psi_p) - \tilde{\eta}^C_p (\varphi_p \otimes 1 - 1 \otimes \psi_p))^2).
\]

Proof. This follows from almost the same exposition and proofs as [8, pp. 293–294, Theorem 9.3.1, Corollary 9.3.1] and Lemmata A.2–A.3 of the present paper. The proof is thus omitted.

Appendix B. Technical results for the IRCPF

**Proposition B.1.** For any \( n \geq 0, s \in \{f, c\}, p \geq 1 \) there exists a \( C < +\infty \) such that for any \( \varphi \in \mathcal{B}_b(\mathbf{X}), N \geq 1 \) we have

\[
\mathbb{E}[|\tilde{\eta}^N_{n,s} - \tilde{\eta}^C_n(\varphi)|^p]^{1/p} \leq \frac{C\|\varphi\|}{\sqrt{N}}.
\]

Proof. This can be proved easily by induction, for instance using the strategy in [22, Proposition C.6]; the proof is hence omitted.

**Proposition B.2.** For any \( \varphi \in \mathcal{B}_b(\mathbf{X} \times \mathbf{X}), n \geq 0 \) we have

\[
(\eta^N_{n,f} \otimes \eta^N_{n,c})(\varphi) \rightarrow \mathbb{P} (\eta^C_n \otimes \eta^C_n)(\varphi).
\]

Proof. Our proof is by induction. Consider the case \( n = 0 \) and let \( \epsilon > 0 \) be arbitrary; then

\[
\mathbb{P}(|(\eta^N_{n,f} \otimes \eta^N_{n,c}) - (\eta^C_n \otimes \eta^C_n)|(\varphi)| > \epsilon)
\]

\[
\leq \mathbb{P}(|(|(\eta^N_{n,f} \otimes \eta^N_{n,c})(\varphi) - \mathbb{E}[(\eta^N_{n,f} \otimes \eta^N_{n,c})(\varphi)])| > \epsilon/2) + \mathbb{P}(|\mathbb{E}[(\eta^N_{n,f} \otimes \eta^N_{n,c})(\varphi)] - (\eta^C_n \otimes \eta^C_n)(\varphi)| > \epsilon/2) + O(N^{-1}),
\]

where the last line follows from

\[
\mathbb{E}[(\eta^N_{n,f} \otimes \eta^N_{n,c})(\varphi)] = \frac{N-1}{N} (\eta^C_n \otimes \eta^C_n)(\varphi) + \frac{1}{N} \tilde{\eta}^C_n(\varphi).
\]
To deal with the term
\[
\mathbb{P}(\|(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi) - \mathbb{E}[(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi)]\| > \varepsilon/2)
\]
we will apply the bounded difference inequality (see [26]). To this end, first note that \((u_{0}^{1}, \ldots, u_{0}^{N})\) are i.i.d., and setting \(f(u_{0}^{1}, \ldots, u_{0}^{N}) = (\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi)\), we note that for any \(1 \leq k \leq N\) (the notational convention is clear if \(k = 1\) or \(k = N\)) and any \((u_{0}^{1}, \ldots, u_{0}^{N}) \in (X \times X)^{n+1}\), \(\tilde{u}_{0}^{k} \in X \times X\),

\[
[f(u_{0}^{1}, \ldots, u_{0}^{N}) - f(u_{0}^{1}, \ldots, u_{0}^{k-1}, u_{0}^{k}, u_{0}^{k+1}, \ldots, u_{0}^{N})] = \frac{1}{N} (\varphi(u_{0}^{k}) - \varphi(\tilde{u}_{0}^{k}))
\]

\[+ \sum_{i \in \{1, \ldots, N\} \setminus \{k\}} \left[ \varphi(X_{0}^{i,f}, X_{0}^{i,c}) - \varphi(X_{0}^{k,f}, X_{0}^{k,c}) \right] + \sum_{i \in \{1, \ldots, N\} \setminus \{k\}} \left[ \varphi(X_{0}^{i,f}, X_{0}^{i,c}) - \varphi(X_{0}^{k,f}, X_{0}^{k,c}) \right]
\]

\[
\leq \frac{4\|\varphi\|}{N}.
\]

Thus, by the bounded difference inequality,

\[
\mathbb{P}(\|(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi) - \mathbb{E}[(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi)]\| > \varepsilon/2) \leq 2 \exp\left\{-\frac{\varepsilon^{2}N}{16\|\varphi\|^{2}}\right\}.
\]

Hence we can easily conclude the result when \(n = 0\) as \(\varepsilon > 0\) was arbitrary.

Now we assume the result for \(n = n - 1\) and consider \(n\). We have

\[
\mathbb{P}(\|(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c}) - (\eta_{n}^{f} \otimes \eta_{n}^{c})\|(\varphi) > \varepsilon) \leq \mathbb{P}(\|(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi) - \mathbb{E}[(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi)]\| > \varepsilon/2) + \mathbb{P}(\|\mathbb{E}[(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi)]\|_{\mathcal{F}_{n-1}^{N}} - (\eta_{n}^{f} \otimes \eta_{n}^{c})(\varphi)) > \varepsilon/2),
\]

where \(\mathcal{F}_{n-1}^{N}\) is the \(\sigma\)-algebra generated by the particle system up to time \(n - 1\). For the term

\[
\mathbb{P}(\|(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi) - \mathbb{E}[(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi)]\|_{\mathcal{F}_{n}^{N}} > \varepsilon/2,\mathcal{F}_{n-1}) \leq 2 \exp\left\{-\frac{\varepsilon^{2}N}{16\|\varphi\|^{2}}\right\},
\]

and thus one can conclude the result if

\[
\mathbb{E}[(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi)|_{\mathcal{F}_{n-1}^{N}}] \rightarrow \mathbb{P}(\eta_{n}^{f} \otimes \eta_{n}^{c})(\varphi).
\]

Now

\[
\mathbb{E}[(\eta_{n}^{N,f} \otimes \eta_{n}^{N,c})(\varphi)|_{\mathcal{F}_{n-1}^{N}}] = \frac{1}{N^{2}} \left( \sum_{i=1}^{N} \mathbb{E}[\varphi(U_{i}^{f})|_{\mathcal{F}_{n-1}^{N}}] + \sum_{i \neq j} \mathbb{E}[(\varphi(X_{n}^{i,f}, X_{n}^{i,c})|_{\mathcal{F}_{n-1}^{N}})\right)
\]

\[= \frac{1}{N_{n-1}} \Phi_{n}(\eta_{n-1}^{N,f} \otimes \eta_{n-1}^{N,c})(\varphi) + \frac{(N - 1)}{N} (\Phi_{n}(\eta_{n-1}^{N,f}) \otimes \Phi_{n}(\eta_{n-1}^{N,c}))(\varphi).
\]
Now by the induction hypothesis and Proposition B.1, \( \hat{\Phi}^f_n(\eta_{n-1} \otimes \eta_{n-1}^c)(\varphi) \) converges in probability to
\[
\left( \eta^f_{n-1} \otimes \eta^c_{n-1} \right) \left( (G_{n-1} \otimes G_{n-1}) \hat{M}_n(\varphi) \right) / \eta^f_{n-1}(G_{n-1}) \eta^c_{n-1}(G_{n-1})
\]
hence the term \( \frac{1}{N} \hat{\Phi}^f_n(\eta_{n-1} \otimes \eta_{n-1}^c)(\varphi) \) goes to zero. Then, again by the induction hypothesis and Proposition B.1, \( (\Phi^f_n(\eta_{n-1}^f) \otimes \Phi^f_n(\eta_{n-1}^c))(\varphi) \) converges in probability to
\[
\left( \eta^f_{n-1} \otimes \eta^c_{n-1} \right) \left( (G_{n-1} \otimes G_{n-1}) \left( M^f_{n-1} \otimes M^c_{n-1} \right) (\varphi) \right) / \eta^f_{n-1}(G_{n-1}) \eta^c_{n-1}(G_{n-1})
\]
and hence we conclude the result. \( \square \)

**Appendix C. Technical results for the MCIRCPF**

**Proof of Theorem 3.1.** The proof is by induction. The case \( n = 0 \) follows by the weak law of large numbers for i.i.d. random variables, so we assume the result at time \( n - 1 \). We have
\[
\eta_{n,M}^N(\varphi) - \eta_{n,M}^N(\varphi) = \eta_{n,M}^N(\varphi) - \Phi^N_n(\eta_{n-1,M}^N)(\varphi) + \Phi^N_n(\eta_{n-1}^c)(\varphi) - \Phi^N_n(\eta_{n-1}^c)(\varphi).
\]
One can easily prove that
\[
|\eta_{n,M}^N(\varphi) - \Phi^N_n(\eta_{n-1,M})(\varphi)| \to \mathbb{P} 0
\]
by using the (conditional) Marcinkiewicz–Zygmund inequality, so we focus on the latter term. Define
\[
T_1^N := \left( F_{n-1,\eta_{n-1,M}^N} \land F_{n-1,\eta_{n-1}^c} \right) \hat{M}_n(\varphi) - \eta_{n-1}^M \left( F_{n-1,\eta_{n-1,M}^N} \land F_{n-1,\eta_{n-1}^c} \right) \hat{M}_n(\varphi)
\]
\[
T_2^N := \left( 1 - \eta_{n-1}^M \left( F_{n-1,\eta_{n-1,M}^N} \land F_{n-1,\eta_{n-1}^c} \right) \right) \left( \eta_{n-1}^M \otimes \eta_{n-1}^c \right) \left( F_{n-1,\eta_{n-1,M}^N} \land F_{n-1,\eta_{n-1}^c} \right) \hat{M}_n(\varphi)
\]
\[
T_3^N := \left( 1 - \eta_{n-1}^M \left( F_{n-1,\eta_{n-1,M}^N} \land F_{n-1,\eta_{n-1}^c} \right) \right) \left( \eta_{n-1}^M \otimes \eta_{n-1}^c \right) \left( F_{n-1,\eta_{n-1,M}^N} \land F_{n-1,\eta_{n-1}^c} \right) \hat{M}_n(\varphi)
\]
Then
\[
T_1^N + T_2^N - T_3^N = \Phi^N_n(\eta_{n-1,M})(\varphi) - \Phi^N_n(\eta_{n-1}^c)(\varphi).
\]
By Lemma C.1, \( T_1^N \to \mathbb{P} 0 \) and by Lemma C.6, \( T_2^N - T_3 \to \mathbb{P} 0 \). This concludes the proof. \( \square \)

**Lemma C.1.** Assume (A1). Then if for any \( \varphi \in C_b(X \times X), n \geq 0 \)
\[
\eta_{n,M}^N(\varphi) \to \mathbb{P} \eta_{n,M}^N(\varphi),
\]
we have for any \( \psi \in C_b(X \times X), n \geq 0 \) that
\[
\eta_{n,M}^N \left( \left( F_{n,\eta_{n,M}^N} \land F_{n,\eta_{n}^c} \right) \psi \right) \to \mathbb{P} \eta_{n,M}^N \left( \left( F_{n,\eta_{n,M}^N} \land F_{n,\eta_{n}^c} \right) \psi \right).
\]

**Proof.** Set
\[
T_1^N := \eta_{n,M}^N \left( \left( F_{n,\eta_{n,M}^N} \land F_{n,\eta_{n}^c} \right) \psi \right) - \left( F_{n,\eta_{n,M}^N} \land F_{n,\eta_{n}^c} \right) \psi.
\]
\[ T^N_2 := \left( \tilde{\eta}^N_n - \check{\eta}^N_n \right) \left( \{ F_{n, \tilde{\eta}^N_n} f \wedge F_{n, \check{\eta}^N_n} \} \psi \right). \]  

(17)

Then we have the decomposition

\[ T^N_1 = T^N_2 + T^N_3 + T^N_4 + T^N_5. \]

To show that \( T^N_1 \) converges in probability to zero we will show that (18)–(20) each converge in probability to zero. For (18) we have

\[ \tilde{\eta}^N_n \left( \{ F_{n, \tilde{\eta}^N_n} f - F_{n, \check{\eta}^N_n} f \} \psi \right) = \frac{\eta_{n}^f(G_n) - \eta^N_n(G_n)}{\eta^f_n(G_n)} \tilde{\eta}^N_n \psi \left( (G_n \otimes 1) \psi \right). \]

By [22, Proposition C.6],

\[ \eta^f_n(G_n) \rightarrow 0, \]

and by hypothesis \( \tilde{\eta}^N_n(G_n \otimes 1) \psi \) converges in probability; hence (18) converges in probability to zero. For (19) this term converges in probability to zero by an almost identical argument to (18) and is hence omitted. For (20),

\[ \tilde{\eta}^N_n \left( \{ |F_{n, \tilde{\eta}^N_n} f - F_{n, \check{\eta}^N_n} f| \} \psi \right) \]

\[ = \frac{\eta^f_n(G_n)\eta^N_n(G_n) - \eta_n^f(G_n)\eta_n^N_n(G_n)}{\eta^f_n(G_n)\eta_n^f(G_n)\eta^N_n(G_n) - \eta_n^f(G_n)\eta_n^N_n(G_n)} \tilde{\eta}^N_n \left( \{ |\eta^N_n(G_n)(G_n \otimes 1) - \eta^f_n(G_n)(1 \otimes G_n)\} \psi \right) \]

\[ - \frac{1}{\eta^f_n(G_n)\eta^N_n(G_n)} \tilde{\eta}^N_n \left( \{ |\eta^N_n(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)\} \psi \right). \]

(21)

The first term on the right-hand side converges in probability to zero by [22, Proposition C.6] and the hypothesis. Hence, as

\[ \frac{1}{\eta^f_n(G_n)\eta^N_n(G_n)} \]

converges in probability, we need only show that

\[ \tilde{\eta}^N_n \left( \{ |\eta^N_n(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)\} \psi \right) \]
converges in probability to zero to conclude. Using standard algebra, we have almost surely that
\[
\left| \eta_n^{c,G_n}(G_n \otimes 1) - \eta_n^{c,f}(G_n)(1 \otimes G_n) - |\eta_n^{c,G_n}(G_n \otimes 1) - \eta_n^{f}(G_n)(1 \otimes G_n)| \right|
\leq |[\eta_n^{c,G_n} - \eta_n^{c,f}(G_n)(1 \otimes G_n)]| + |[\eta_n^{c,f} - \eta_n^{f}(G_n)(1 \otimes G_n)]|.
\]
(22)

Hence
\[
\tilde{\eta}_n^{N,M}(\psi) \leq \eta_n^{N,c}(G_n)\tilde{\eta}_n^{N,M}((G_n \otimes 1)(\psi)) + |[\eta_n^{c,f} - \eta_n^{f}(G_n)]\tilde{\eta}_n^{N,M}((1 \otimes G_n)(\psi)),
\]
and one can easily conclude by the above arguments.

**Term (17):**

\[
\{F_{n,\tilde{\eta}_n^{M,c}} \land F_{n,\tilde{\eta}_n^{M,c}}\} \psi = \frac{1}{2} \left( F_{n,\tilde{\eta}_n^{M,f}} + F_{n,\tilde{\eta}_n^{M,c}} - |F_{n,\tilde{\eta}_n^{M,f}} - F_{n,\tilde{\eta}_n^{M,c}}| \right) \psi,
\]
we have \(\{F_{n,\tilde{\eta}_n^{M,f}} \land F_{n,\tilde{\eta}_n^{M,c}}\} \psi \in C_b(\mathbb{X} \times \mathbb{X})\), so (17) converges in probability to zero. \(\square\)

**Lemma C.2.** Assume (A1). Then if for any \(\varphi \in C_b(\mathbb{X} \times \mathbb{X})\), \(n \geq 0\)
\[
\tilde{\eta}_n^{N,M}(\varphi) \to \tilde{\eta}_n^{M}(\varphi),
\]
we have for any \(\varphi \in C_b(\mathbb{X} \times \mathbb{X})\), \(n \geq 0\) that
\[
(\tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M}) \left( \left\{ \frac{F_{n,\tilde{\eta}_n^{M,f}} - F_{n,\tilde{\eta}_n^{M,c}}}{\tilde{\eta}_n^{N,M}(F_{n,\tilde{\eta}_n^{M,f}} - F_{n,\tilde{\eta}_n^{M,c}})} \otimes \tilde{F}_{n,\tilde{\eta}_n^{M,c}} \right\} \tilde{M}_{n+1}(\psi) \right).
\]
(23)

Then we have
\[
T_1^N + T_2^N = (\tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M}) \left( \left\{ \frac{F_{n,\tilde{\eta}_n^{M,f}} - F_{n,\tilde{\eta}_n^{M,c}}}{\tilde{\eta}_n^{N,M}(F_{n,\tilde{\eta}_n^{M,f}} - F_{n,\tilde{\eta}_n^{M,c}})} \otimes \tilde{F}_{n,\tilde{\eta}_n^{M,c}} \right\} \tilde{M}_{n+1}(\psi) \right).
\]
(24)
We will show that (23) and (24) converge in probability to zero.

**Term (23):** For (23) we note that

\[
\tilde{\eta}_n^{N,M}(F_{n,\tilde{\eta}_n^{N,M},c} - F_{n,\tilde{\eta}_n^{N,M},c} \land F_{n,\tilde{\eta}_n^{N,M},c}) = 1 - \tilde{\eta}_n^{N,M}(F_{n,\tilde{\eta}_n^{N,M},c} \land F_{n,\tilde{\eta}_n^{N,M},c}).
\]

By Lemma C.1 this converges in probability to a constant, so we only consider the convergence to zero of

\[
(\tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M})(\{(F_{n,\tilde{\eta}_n^{N,M},c} - F_{n,\tilde{\eta}_n^{N,M},c}) \otimes \tilde{F}_{n,\tilde{\eta}_n^{N,M},c}\}) = (\tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M})(\{(G_n \otimes 1) \otimes \tilde{F}_{n,\tilde{\eta}_n^{N,M},c}\}) = \tilde{M}_{n+1}(\psi).
\]

By [22, Proposition C.6],

\[
\left(\frac{1}{\eta_n^{N,f}(G_n)} - \frac{1}{f_n(G_n)}\right) \to 0,
\]

so if we can show that the remaining term on the right-hand side is, in absolute value, almost surely bounded above by a convergent (in probability) random variable, we have shown that (23) converges in probability to zero. We have

\[
|\tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M}| \leq \|\psi\| \|G_n\| \tilde{\eta}_n^{N,M}(\{F_{n,\tilde{\eta}_n^{N,M},c}\}).
\]

Then almost surely,

\[
\tilde{\eta}_n^{N,M}(\{F_{n,\tilde{\eta}_n^{N,M},c}\}) \leq \frac{\tilde{\eta}_n^{N,M}(F_{n,\tilde{\eta}_n^{N,M},c}) + \tilde{\eta}_n^{N,M}(\{F_{n,\tilde{\eta}_n^{N,M},c} \land F_{n,\tilde{\eta}_n^{N,M},c}\})}{\tilde{\eta}_n^{N,M}(F_{n,\tilde{\eta}_n^{N,M},c})}.
\]

By the above arguments, both the denominator and the numerator will converge in probability to a finite constant, and hence we have shown that (23) converges in probability to zero.

**Term (24):** We note that

\[
\frac{\tilde{\eta}_n^{M}(F_{n,\tilde{\eta}_n^{M},c} - F_{n,\tilde{\eta}_n^{M},c} \land F_{n,\tilde{\eta}_n^{M},c}) - \tilde{\eta}_n^{M}(F_{n,\tilde{\eta}_n^{M},c} - F_{n,\tilde{\eta}_n^{M},c} \land F_{n,\tilde{\eta}_n^{M},c})}{\tilde{\eta}_n^{M}(F_{n,\tilde{\eta}_n^{M},c} - F_{n,\tilde{\eta}_n^{M},c} \land F_{n,\tilde{\eta}_n^{M},c})} = \frac{\tilde{\eta}_n^{M}(F_{n,\tilde{\eta}_n^{M},c} \land F_{n,\tilde{\eta}_n^{M},c}) - \tilde{\eta}_n^{M}(F_{n,\tilde{\eta}_n^{M},c} \land F_{n,\tilde{\eta}_n^{M},c})}{(1 - \tilde{\eta}_n^{M}(F_{n,\tilde{\eta}_n^{M},c} \land F_{n,\tilde{\eta}_n^{M},c}))}. \tag{26}
\]

By Lemma C.1 this converges in probability to zero. Thus, we need only show that

\[
|\tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M}| \leq \|\psi\| \tilde{\eta}_n^{N,M}(\{F_{n,\tilde{\eta}_n^{N,M},c}\})
\]

is (almost surely) bounded above by a convergent (in probability) random variable. Almost surely,

\[
|\tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M}| \leq \|\psi\| \tilde{\eta}_n^{N,M}(\{F_{n,\tilde{\eta}_n^{N,M},c}\})
\]
As $F_{n,\eta_n^f} \in C_b(X \times X)$, $\hat{\eta}_n^{N,M}(F_{n,\eta_n^f})$ converges in probability to a finite constant and our proof is concluded by the argument associated to (26).

Lemma C.3. Assume (A1). Then if for any $\varphi \in C_b(X \times X)$, $n \geq 0$

$$\hat{\eta}_n^{N,M}(\varphi) \rightarrow P \hat{\eta}_n^{M}(\varphi),$$

we have for any $\psi \in C_b(X \times X)$, $n \geq 0$ that

$$\left(\hat{\eta}_n^{N,M} \otimes \hat{\eta}_n^{N,M}\right) \left(\left[\frac{F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c} - F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c}}{\hat{\eta}_n^{M}(F_{n,\eta_n^f} - F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c})}\right] \otimes \overline{F}_{n,\eta_n^{N,M,c}} M_{n+1}(\psi)\right) \rightarrow P 0.

Proof. Define

$$C := \frac{1}{\hat{\eta}_n^{M}(F_{n,\eta_n^f} - F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c})},$$

$$T_1^N := -\frac{C}{2} \left(\hat{\eta}_n^{N,M} \otimes \hat{\eta}_n^{N,M}\right) \left(\left[\frac{F_{n,\eta_n^f} - F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c}}{\hat{\eta}_n^{M}(F_{n,\eta_n^f} - F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c})}\right] \otimes \overline{F}_{n,\eta_n^{N,M,c}} M_{n+1}(\psi)\right),$$

$$T_2^N := -\frac{C}{2} \left(\hat{\eta}_n^{N,M} \otimes \hat{\eta}_n^{N,M}\right) \left(\left[|F_{n,\eta_n^f} - F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c}| - |F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c}|\right] \otimes \overline{F}_{n,\eta_n^{N,M,c}} M_{n+1}(\psi)\right),$$

$$T_3^N := -\frac{C}{2} \left(\hat{\eta}_n^{N,M} \otimes \hat{\eta}_n^{N,M}\right) \left(\left[\frac{F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c} - F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c}}{\hat{\eta}_n^{M}(F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c})}\right] \otimes \overline{F}_{n,\eta_n^{N,M,c}} M_{n+1}(\psi)\right).$$

Then we have

$$T_1^N + T_2^N + T_3^N = \left(\hat{\eta}_n^{N,M} \otimes \hat{\eta}_n^{N,M}\right) \left(\left[\frac{F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c} - F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c}}{\hat{\eta}_n^{M}(F_{n,\eta_n^f} \wedge F_{n,\eta_n^M,c})}\right] \otimes \overline{F}_{n,\eta_n^{N,M,c}} M_{n+1}(\psi)\right).$$

We will show that (27), (28), and (29) will converge in probability to zero to conclude the proof.

Term (27): We have that (27) is equal to

$$-\frac{C}{2} \frac{\eta_n^{N,f}(G_n) - \eta_n(G_n)}{\eta_n^{N,f}(G_n)\eta_n(G_n)} \left(\hat{\eta}_n^{N,M} \otimes \hat{\eta}_n^{N,M}\right) \left(\left[\frac{G_n \otimes 1}{1} \otimes \overline{F}_{n,\eta_n^{N,M,c}} M_{n+1}(\psi)\right]\right).$$

By [22, Proposition C.6],

$$\frac{\eta_n^{N,f}(G_n) - \eta_n(G_n)}{\eta_n^{N,f}(G_n)\eta_n(G_n)} \rightarrow P 0,$$

so if we can show that the remaining term on the right-hand side is, in absolute value, almost surely bounded above by a convergent (in probability) random variable, we have shown that (27) converges in probability to zero. This can be achieved using the argument for (25) in the proof of Lemma C.2, and hence we have shown that (27) converges in probability to zero.

Term (28): The argument is almost identical to (27) and is omitted.
**Term (29):** Using a decomposition similar to (21), we define

\[
T^N_4 := -C \frac{\eta^n(G_n) \eta^n_c(G_n) - \eta^{N, f}_n(G_n) \eta^{N, c}_n(G_n)}{2 \eta^n(G_n) \eta^n_c(G_n)} (\tilde{\eta}^N_{n, M} \otimes \tilde{\eta}^N_{n, M}) \\
\quad \left( \left( \left( |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)| \right) \otimes \tilde{F}_{n, \tilde{\eta}^N_{n, M, c}} \right) \tilde{M}_{n+1}(\psi) \right),
\]

(30)

\[
T^N_5 := -C \frac{1}{2 \eta^n(G_n) \eta^n_c(G_n)} \left( \tilde{\eta}^N_{n, M} \otimes \tilde{\eta}^N_{n, M} \right) \left( \left( |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^{N, f}(G_n)(1 \otimes G_n)| \right) \otimes \tilde{F}_{n, \tilde{\eta}^N_{n, M, c}} \right) \tilde{M}_{n+1}(\psi),
\]

(31)

Then we have

\[
T^N_3 = -T^N_4 - T^N_5.
\]

We will show that (30) and (31) will converge to zero in probability. For (30), by [22, Proposition C.6],

\[
\eta_n^f(G_n) \eta_n^c(G_n) - \eta^{N, f}_n(G_n) \eta^{N, c}_n(G_n) \to_p 0
\]

and

\[
\left| \left( \tilde{\eta}^N_{n, M} \otimes \tilde{\eta}^N_{n, M} \right) \left( \left( |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)| \right) \otimes \tilde{F}_{n, \tilde{\eta}^N_{n, M, c}} \right) \tilde{M}_{n+1}(\psi) \right|
\]

\[
\leq \|G_n\|_p \|\psi\| (\eta_n^f(G_n) \eta_n^c(G_n)) \tilde{\eta}^N_{n, M} (\tilde{F}_{n, \tilde{\eta}^N_{n, M, c}});
\]

thus, using the argument for (25) in the proof of Lemma C.2, we have shown that (30) converges in probability to zero. For (31), as \((\eta^{N, f}_n(G_n) \eta^{N, c}_n(G_n))^{-1}\) converges in probability to a finite constant, we need only show that the remaining term converges in probability to zero. Using (22) we have that

\[
\left| \left( \tilde{\eta}^N_{n, M} \otimes \tilde{\eta}^N_{n, M} \right) \left( \left( |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^{N, f}(G_n)(1 \otimes G_n)| \right) \otimes \tilde{F}_{n, \tilde{\eta}^N_{n, M, c}} \right) \tilde{M}_{n+1}(\psi) \right|
\]

\[
\leq \|\psi\| \tilde{\eta}^N_{n, M} \left( \left( |\eta_n^c(G_n)(G_n \otimes 1)| + |\eta_n^{N, f}(G_n)(1 \otimes G_n)| \right) \tilde{\eta}^N_{n, M} (\tilde{F}_{n, \tilde{\eta}^N_{n, M, c}}) \right).
\]

Then (31) converges in probability to zero by the argument for (25) in the proof of Lemma C.2 for \(\tilde{\eta}^N_{n, M} (\tilde{F}_{n, \tilde{\eta}^N_{n, M, c}})\). For the remaining term one can apply the (last) argument for (20) in Lemma C.1. Hence (29) converges in probability to zero and the proof is concluded.

**Lemma C.4.** Assume (A1). Then if for any \(\psi \in C_0(X \times X)\), \(n \geq 0\)

\[
\tilde{\eta}^N_{n, M} (\psi) \to_p \tilde{\eta}^M_{n, M} (\psi),
\]

we have for any \(\psi \in C_0(X \times X)\), \(n \geq 0\) that

\[
\frac{\tilde{\eta}^N_{n, M} (F_{n, \tilde{\eta}^N_{n, M, f}} \wedge F_{n, \tilde{\eta}^N_{n, M, c}}) - \tilde{\eta}^M_{n, M} (F_{n, \tilde{\eta}^M_{n, M, f}} \wedge F_{n, \tilde{\eta}^M_{n, M, c}})}{(1 - \tilde{\eta}^N_{n, M} (F_{n, \tilde{\eta}^N_{n, M, f}} \wedge F_{n, \tilde{\eta}^N_{n, M, c}}))(1 - \tilde{\eta}^M_{n, M} (F_{n, \tilde{\eta}^M_{n, M, f}} \wedge F_{n, \tilde{\eta}^M_{n, M, c}}))}
\]

\[
\times (\tilde{\eta}^N_{n, M} \otimes \tilde{\eta}^N_{n, M} \left( \left( \left| F_{n, \tilde{\eta}^N_{n, M, f}} \wedge F_{n, \tilde{\eta}^N_{n, M, c}} \right| \otimes \tilde{F}_{n, \tilde{\eta}^N_{n, M, c}} \right) \tilde{M}_{n+1}(\psi) \right) \to_p 0.
\]
Proof. The proof concerning (24) of Lemma C.2 establishes that
\[
\frac{\eta_n^{N,M}(F_{n,\tilde{\eta}_n^{N,M},f} \land F_{n,\tilde{\eta}_n^{N,M},c}) - \eta_n^{N,M}(F_{n,\tilde{\eta}_n^{N,M},f} \land F_{n,\tilde{\eta}_n^{N,M},c})}{(1 - \eta_n^{N,M}(F_{n,\tilde{\eta}_n^{N,M},f} \land F_{n,\tilde{\eta}_n^{N,M},c}))} \rightarrow \mathbb{P} \ 0.
\]
Then, almost surely
\[
\left| (\eta_n^{N,M} \otimes \tilde{\eta}_n^{N,M}) \left( \left\{ F_{n,\tilde{\eta}_n^{N,M},f} \land F_{n,\tilde{\eta}_n^{N,M},c} \right\} \otimes F_{n,\tilde{\eta}_n^{N,M},c} \right) \right| \rightarrow 0.
\]
By Lemma C.1, \( \eta_n^{N,M}(F_{n,\tilde{\eta}_n^{N,M},f} \land F_{n,\tilde{\eta}_n^{N,M},c}) \) converges in probability to a constant, and using the argument for (25) in the proof of Lemma C.2 for \( \eta_n^{N,M}(|F_{n,\tilde{\eta}_n^{N,M},c}|) \), we conclude the proof.

Lemma C.5. Assume (A3). Then if for any \( n \geq 0, \varphi \in C_b(X \times X) \)
\[\eta_n^{N,M}(\varphi) \rightarrow \mathbb{P} \tilde{\eta}_n^{M}(\varphi),\]
we have for any \( \psi \in C_b(X^2 \times X^2) \) that
\[\left( \tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M} \right)(\psi) \rightarrow \mathbb{P} \left( \tilde{\eta}_n^{M} \otimes \tilde{\eta}_n^{M} \right)(\psi).
\]

Proof. We will use a density argument. Define
\[\mathcal{F} = \{ f : X^2 \times X^2 \rightarrow \mathbb{R} : f(x, y) = g(x) \otimes h(y), (g, h) \in C_b(X \times X) \times C_b(X \times X) \}.
\]
Denote by \( \mathcal{G} \) the set of functions which are finite linear combinations of functions in \( \mathcal{F} \). As we have that for some \( n \geq 0, \varphi \in C_b(X \times X), \eta_n^{N,M}(\varphi) \rightarrow \mathbb{P} \tilde{\eta}_n^{M}(\varphi) \), it follows that for \( \tilde{\psi} \in \mathcal{G} \),
\[\left( \tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M} \right)(\tilde{\psi}) \rightarrow \mathbb{P} \left( \tilde{\eta}_n^{M} \otimes \tilde{\eta}_n^{M} \right)(\tilde{\psi}),
\]
and thus, by [3, Theorem 25.12],
\[
\lim_{N \rightarrow +\infty} \mathbb{E} \left[ \left| \left( \tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M} \right) - \left( \tilde{\eta}_n^{M} \otimes \tilde{\eta}_n^{M} \right) \right| (\tilde{\psi}) \right] = 0. \tag{32}
\]
Let \( \epsilon > 0 \) be arbitrary and \( \psi \in C_b(X^2 \times X^2) \). By the Stone–Weierstrass theorem, \( \mathcal{G} \) is dense in \( C_b(X^2 \times X^2) \); hence there exists a \( \tilde{\psi} \in \mathcal{G} \) such that
\[\| \psi - \tilde{\psi} \| < \epsilon / 3. \tag{33}\]
Then we have that for any \( N \geq 1, \)
\[
\mathbb{E} \left[ \left| \left( \tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M} \right) - \left( \tilde{\eta}_n^{M} \otimes \tilde{\eta}_n^{M} \right) \right| (\psi) \right] \leq \mathbb{E} \left[ \left( \tilde{\eta}_n^{N,M} \otimes \tilde{\eta}_n^{N,M} \right)(|\psi - \tilde{\psi}|) \right] + \mathbb{E} \left[ \left| \left( \tilde{\eta}_n^{M} \otimes \tilde{\eta}_n^{M} \right) - \left( \tilde{\eta}_n^{M} \otimes \tilde{\eta}_n^{M} \right) \right| (\tilde{\psi}) \right].
\]
By (32) there exists an \( N^* \geq 1 \) such that for every \( N \geq N^* \),

\[
\mathbb{E}\left[ \left\| \left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n - \left( \tilde{\eta}^M_n \otimes \tilde{\eta}^M_n \right) \left( \tilde{\psi} \right) \right) \right\| \right] < \epsilon/3.
\]

Hence, applying (33) to the terms

\[
\mathbb{E}\left[ \left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n \right) (|\psi - \tilde{\psi}|) \right] \quad \text{and} \quad \mathbb{E}\left[ \left( \tilde{\eta}^M_n \otimes \tilde{\eta}^M_n \right) (|\psi - \tilde{\psi}|) \right],
\]

we have for every \( N \geq N^* \) that

\[
\mathbb{E}\left[ \left\| \left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n - \left( \tilde{\eta}^M_n \otimes \tilde{\eta}^M_n \right) \right) (\psi) \right\| \right] < \epsilon,
\]

and as \( \epsilon > 0 \) is arbitrary we have

\[
\left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n \right)(\psi) \to \mathbb{P} \left( \tilde{\eta}^M_n \otimes \tilde{\eta}^M_n \right)(\psi),
\]

as was to be proved. \( \square \)

**Lemma C.6.** Assume \((A1)\)--\((A3)\). Then if for any \( \varphi \in \mathcal{C}_b(X \times X) \), \( n \geq 0 \)

\[
\tilde{\eta}^{N,M}_n(\varphi) \to \mathbb{P} \tilde{\eta}^M_n(\varphi),
\]

we have for any \( \psi \in \mathcal{C}_b(X \times X) \), \( n \geq 0 \) that

\[
\left( 1 - \tilde{\eta}^{N,M}_n \left( \{F_{n,\tilde{\eta}^{N,M}_n,f} \wedge F_{n,\tilde{\eta}^{N,M}_n,c} \} \right) \right) \left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n \right) \left( \{F_{n,\tilde{\eta}^{N,M}_n,f} \otimes F_{n,\tilde{\eta}^{N,M}_n,c} \} \tilde{M}_{n+1}(\psi) \right) \to \mathbb{P}
\]

\[
\left( 1 - \tilde{\eta}^M_n \left( \{F_{n,\tilde{\eta}^M_n,f} \wedge F_{n,\tilde{\eta}^M_n,c} \} \right) \right) \left( \tilde{\eta}^M_n \otimes \tilde{\eta}^M_n \right) \left( \{F_{n,\tilde{\eta}^M_n,f} \otimes F_{n,\tilde{\eta}^M_n,c} \} \tilde{M}_{n+1}(\psi) \right).
\]

**Proof.** We make the definitions

\[
T^N_1 := \left( \tilde{\eta}^M_n \left( \{F_{n,\tilde{\eta}^M_n,f} \wedge F_{n,\tilde{\eta}^M_n,c} \} \right) \right) \left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n \right) \left( \{F_{n,\tilde{\eta}^{N,M}_n,f} \otimes F_{n,\tilde{\eta}^{N,M}_n,c} \} \tilde{M}_{n+1}(\psi) \right),
\]

\[
T^N_2 := \left( 1 - \tilde{\eta}^M_n \left( \{F_{n,\tilde{\eta}^M_n,f} \wedge F_{n,\tilde{\eta}^M_n,c} \} \right) \right) \left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n \right) \left( \{F_{n,\tilde{\eta}^{N,M}_n,f} \otimes F_{n,\tilde{\eta}^{N,M}_n,c} \} - \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n \right) \tilde{M}_{n+1}(\psi),
\]

\[
T^N_3 := \left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n \right) \left( \{F_{n,\tilde{\eta}^{N,M}_n,f} \otimes F_{n,\tilde{\eta}^{N,M}_n,c} \} \tilde{M}_{n+1}(\psi) \right).
\]

Then we have

\[
T^N_1 + T^N_2 + T^N_3 = \left( 1 - \tilde{\eta}^M_n \left( \{F_{n,\tilde{\eta}^M_n,f} \wedge F_{n,\tilde{\eta}^M_n,c} \} \right) \right) \left( \tilde{\eta}^{N,M}_n \otimes \tilde{\eta}^{N,M}_n \right) \left( \{F_{n,\tilde{\eta}^{N,M}_n,f} \otimes F_{n,\tilde{\eta}^{N,M}_n,c} \} \right) \tilde{M}_{n+1}(\psi) - \left( 1 - \tilde{\eta}^M_n \left( \{F_{n,\tilde{\eta}^M_n,f} \wedge F_{n,\tilde{\eta}^M_n,c} \} \right) \right) \tilde{\eta}^M_n \otimes \tilde{\eta}^M_n \left( \{F_{n,\tilde{\eta}^M_n,f} \otimes F_{n,\tilde{\eta}^M_n,c} \} \tilde{M}_{n+1}(\psi) \right).
\]

We will show that (34)–(36) will converge in probability to zero.
We focus on showing that

\[ TN \]

so if we can show that

\[ \eta_n \]

we have shown that (34) converges in probability to zero. Clearly, almost surely we have

\[ T_n^N \]

is (almost surely) bounded above by a term which converges in probability to a positive constant, we have shown that (34) converges in probability to zero. Clearly, almost surely we have

\[ T_n^N \]

We focus on showing that \( \eta_n^N, M (|F_{n, \eta_n^N, c}^M|) \) is bounded above by a term which converges in probability to a finite constant. This can be verified in an almost identical manner to the approach for (26) in the proof of Lemma C.2; hence we have verified that (34) converges in probability to zero.

**Term (35):** Set

\[ C := \eta_n^M (\{ F_{n, \eta_n^N, f} \land F_{n, \eta_n^M, c} \}) , \]

\[ T_4^N := C (\eta_n^N \land \eta_n^M) \left\{ \left( \frac{F_{n, \eta_n^M, f} - F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c}}{\eta_n^M (F_{n, \eta_n^M, f} - F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c})} \right) \right\} \tilde{M}_{n+1} (\psi) , \]

\[ T_5^N := C (\eta_n^N \land \eta_n^M) \left\{ \left( \frac{F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c}}{\eta_n^M (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c})} \right) \right\} \tilde{M}_{n+1} (\psi) ; \]

then

\[ T_2^N = T_4^N + T_5^N . \]

We need to show that (37) and (38) converge in probability to zero. As the proof for (38) is similar to and easier than that for (37), we focus on the latter. Define

\[ T_6^N := C (\eta_n^N \land \eta_n^M) \left\{ \left( \frac{F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c}}{\eta_n^M (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c})} \right) \right\} \tilde{M}_{n+1} (\psi) , \]

\[ T_7^N := C (\eta_n^N \land \eta_n^M) \left\{ \left( \frac{F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c}}{\eta_n^M (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c})} \right) \right\} \tilde{M}_{n+1} (\psi) , \]

\[ T_8^N := C (\eta_n^N \land \eta_n^M) \left\{ \left( \frac{F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c}}{\eta_n^M (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c})} \right) \right\} \tilde{M}_{n+1} (\psi) , \]

\[ T_9^N := \frac{\eta_n^N (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c}) - \eta_n^M (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c})}{(1 - \eta_n^M (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c}))(1 - \eta_n^M (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c})))} \times C (\eta_n^N \land \eta_n^M) \left\{ \left( \frac{F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c}}{\eta_n^M (F_{n, \eta_n^M, f} \land F_{n, \eta_n^M, c})} \right) \right\} \tilde{M}_{n+1} (\psi) . \]
Then we have that
\[ T_4^N = T_6^N - T_7^N + T_8^N - T_9^N. \]
By Lemma C.2 we have \( T_6^N - T_7^N \rightarrow \mathbb{P} 0 \), by Lemma C.3 we have \( T_8^N \rightarrow \mathbb{P} 0 \), and by Lemma C.4 we have \( T_9^N \rightarrow \mathbb{P} 0 \); hence we conclude that \( T_4^N \rightarrow \mathbb{P} 0 \).

**Term (36):** As
\[
\left\{ F_{n, \hat{\eta}^M_{n, \cdot}} \otimes \tilde{F}_{n, \hat{\eta}^M_{n, \cdot}, \cdot} \right\} \tilde{M}_{n+1} (\psi) \in C_b (\mathbb{X}^2 \times \mathbb{X}^2),
\]
it easily follows by Lemma C.5 that
\[
\left[ (\hat{\eta}^M_{n, \cdot} \otimes \hat{\eta}^M_{n, \cdot}) - (\hat{\eta}^M_{n, \cdot} \otimes \hat{\eta}^M_{n, \cdot}) \right] \left( \left\{ F_{n, \hat{\eta}^M_{n, \cdot}} \otimes \tilde{F}_{n, \hat{\eta}^M_{n, \cdot}, \cdot} \right\} \tilde{M}_{n+1} (\psi) \right) \rightarrow \mathbb{P} 0. \]

**Appendix D. Technical results for the MCPF**

**Proposition D.1.** For any \( n \geq 0 \), \( s \in \{ f, c \}, p \geq 1 \) there exists a \( C < +\infty \) such that for any \( \varphi \in \mathcal{B}_b (\mathbb{X}) \), \( N \geq 1 \) we have
\[
\mathbb{E} [ | (\hat{\eta}^{N,s}_{n} - \eta^{s}_{n}) (\varphi) |^p ]^{1/p} \leq \frac{C \| \varphi \|}{\sqrt{N}}.
\]

**Proof:** As for Proposition B.1. \( \square \)

Recall that we are assuming that \( M^s_n \) has a density and we are denoting the density by \( M^s_n \) as well.

**Theorem D.1.** Suppose that for \( n \geq 1 \), \( s \in \{ f, c \}, M^s_n \in \mathcal{B}_b (\mathbb{X} \times \mathbb{X}) \). Then for any \( \varphi \in \mathcal{B}_b (\mathbb{X} \times \mathbb{X}) \), \( n \geq 0 \) we have
\[
\hat{\eta}^{N,s}_{n} (\varphi) \rightarrow \mathbb{P} \hat{\eta}^s (\varphi).
\]

**Proof:** The proof is by induction, the initialization following by the WLLN for i.i.d. random variables. The result is assumed at time \( n - 1 \), and we have the decomposition
\[
\hat{\eta}^{N,s}_{n} (\varphi) - \hat{\eta}^s (\varphi) = \hat{\Phi}^C (\eta^f_{n-1}, \eta^c_{n-1}) (\varphi) + \hat{\Phi}^C (\eta^f_{n-1}, \eta^c_{n-1}) (\varphi) - \hat{\eta}^s (\varphi).
\]
As in the proof of Theorem 3.1, one can easily prove that
\[
| \hat{\eta}^{N,s}_{n} (\varphi) - \hat{\Phi}^C (\eta^f_{n-1}, \eta^c_{n-1}) (\varphi) - \hat{\eta}^s (\varphi) |ightarrow \mathbb{P} 0
\]
by using the (conditional) Marcinkiewicz–Zygmund inequality, so we focus on
\[
\hat{\Phi}^C (\eta^f_{n-1}, \eta^c_{n-1}) (\varphi) - \hat{\eta}^s (\varphi).
\]
Define
\[
T_1^N := \int_{\mathbb{X}} \varphi (x, x) F_{n-1, \hat{\eta}^{N,f}_{n-1} \cdot} (x) \wedge F_{n, \hat{\eta}^{N,c}_{n, \cdot}} (x) dx - \int_{\mathbb{X}} \varphi (x, x) F_{n-1, \eta^{N,f}_{n-1} \cdot} (x) \wedge F_{n-1, \eta^{N,c}_{n-1, \cdot}} (x) dx,
\]
\[
T_2^N := \left( 1 - \int_{\mathbb{X}} F_{n-1, \hat{\eta}^{N,f}_{n-1} \cdot} (x) \wedge F_{n-1, \hat{\eta}^{N,c}_{n-1, \cdot}} (x) dx \right)^{-1} \int_{\mathbb{X} \times \mathbb{X}} \varphi (x, y) \tilde{F}_{n-1, \hat{\eta}^{N,f}_{n-1} \cdot, \cdot} (x) dy - \left( 1 - \int_{\mathbb{X}} F_{n-1, \eta^{N,f}_{n-1} \cdot} (x) \wedge F_{n-1, \eta^{N,c}_{n-1, \cdot}} (x) dx \right)^{-1} \int_{\mathbb{X} \times \mathbb{X}} \varphi (x, y) \tilde{F}_{n-1, \eta^{N,f}_{n-1} \cdot, \cdot} (x) dy.
\]
Recalling that \( \tilde{\eta}^C_n(\psi) = \Phi^C_n(\eta^f_{n-1}, \eta^c_{n-1})(\psi) \), we have that
\[
T^N_1 + T^N_2 = \Phi^C_p(\eta^f_{n-1}, \eta^c_{n-1})(\psi) - \tilde{\eta}^C_n(\psi).
\]
By Lemma D.1 Part 1 we have \( T^N_1 \to_P 0 \), and by Lemma D.1 Part 2 we have \( T^N_2 \to_P 0 \), which concludes the proof.

\[\square\]

**Lemma D.1.** Suppose that for \( n \geq 1, s \in \{f, c\} \), \( M^s_n \in \mathcal{B}_b(\mathbb{X} \times \mathbb{X}) \) and that for \( \varphi \in \mathcal{B}_b(\mathbb{X} \times \mathbb{X}) \), \( n \geq 0 \)
\[
\tilde{\eta}^N_{n_s}(\varphi) \to_P \tilde{\eta}^C_n(\varphi).
\]
Then for any \( \psi \in \mathcal{B}_b(\mathbb{X} \times \mathbb{X}) \), \( n \geq 0 \) we have the following:

1. \[
\int_{\mathbb{X}} \psi(x, x) F_{n, \eta_{n_s}} x \wedge F_{n, \eta_{n_s}} c(x) dx \to P \int_{\mathbb{X}} \psi(x, x) F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) dx.
\]
2. \[
\left( 1 - \int_{\mathbb{X}} F_{n, \eta_{n_s}} x \wedge F_{n, \eta_{n_s}} c(x) dx \right)^{-1} \int_{\mathbb{X} \times \mathbb{X}} \psi(x, y) F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} f(y) dx dy \to_P \int_{\mathbb{X} \times \mathbb{X}} \psi(x, y) F_{n, \eta_{n_s}} c(x) \wedge F_{n, \eta_{n_s}} c(y) dx dy.
\]

**Proof.** We begin by noting that for any \( x \in \mathbb{X} \), by the assumption that \( \tilde{\eta}^N_{n_s}(\varphi) \to_P \tilde{\eta}^C_n(\varphi) \), we have
\[
F_{n, \eta_{n_s}} f(x) \to_P F_{n, \eta_{n_s}} f(x),
\]
\[
F_{n, \eta_{n_s}} c(x) \to_P F_{n, \eta_{n_s}} c(x),
\]
and hence that
\[
F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) \to_P F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x),
\]
\[
F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) \to_P F_{n, \eta_{n_s}} f(x),
\]
\[
F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) \to_P F_{n, \eta_{n_s}} f(x).
\]

For Part 1 we have \[
\mathbb{E} \left[ \left| \int_{\mathbb{X}} \psi(x, x) \left( F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) - F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) \right) dx \right] \right] \leq \|\psi\| \int_{\mathbb{X}} \mathbb{E} \left[ \left| F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) - F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) \right| \right] dx.
\]
Then, as \[
F_{n, \eta_{n_s}} f(x) \wedge F_{n, \eta_{n_s}} c(x) \leq \|M^f_{n+1}\| \wedge \|M^c_{n+1}\|,
\]
by (39) and [3, Theorem 25.12] we have that for any \( x \in X \),
\[
\lim_{N \to +\infty} \mathbb{E} \left[ F_{n, \eta_{n-1}^f}(x) \wedge F_{n, \eta_{n-1}^c}(x) - F_{n, \eta_{n}^f}(x) \wedge F_{n-1, \eta_{n-1}^c}(x) \right] = 0.
\]
Hence, by the bounded convergence theorem,
\[
\lim_{N \to +\infty} \mathbb{E} \left[ \left| \int_X \psi(x, x) \left( F_{n, \eta_{n-1}^f}(x) \wedge F_{n, \eta_{n-1}^c}(x) - F_{n, \eta_{n}^f}(x) \wedge F_{n-1, \eta_{n-1}^c}(x) \right) dx \right| \right] = 0,
\]
and the result follows.

For Part 2, using almost the same proof as for Part 1 except with (40)–(41) in place of (39), we have
\[
\int_{X \times X} \psi(x, y) \overline{F}_{n, \eta_{n-1}^f, \eta_{n-1}^c}(x) \overline{F}_{n, \eta_{n-1}^f, \eta_{n-1}^c}(x) dx dy \to \mathbb{P}
\]
\[
\int_{X \times X} \psi(x, y) \overline{F}_{n, \eta_{n-1}^f, \eta_{n-1}^c}(x) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) dx dy.
\]
Define
\[
T_1^N := \left( 1 - \int_X F_{n, \eta_{n}^f}(x) \wedge F_{n, \eta_{n}^c}(x) dx \right)^{-1} \left( 1 - \int_X F_{n, \eta_{n}^f}(x) \wedge F_{n, \eta_{n}^c}(x) dx \right)^{-1}
\]
\[
\times \int_{X \times X} \psi(x, y) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) dx dy,
\]
\[
T_2^N := \left( 1 - \int_X F_{n, \eta_{n}^f}(x) \wedge F_{n, \eta_{n}^c}(x) dx \right)^{-1} \left( \int_{X \times X} \psi(x, y) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) dx dy \right)
\]
\[
\times \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) dx dy - \int_{X \times X} \psi(x, y) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) dx dy.
\]
Then we have
\[
T_1^N + T_2^N
\]
\[
= \left( 1 - \int_X F_{n, \eta_{n}^f}(x) \wedge F_{n, \eta_{n}^c}(x) dx \right)^{-1} \int_{X \times X} \psi(x, y) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) dx dy
\]
\[
- \left( 1 - \int_X F_{n, \eta_{n}^f}(x) \wedge F_{n, \eta_{n}^c}(x) dx \right)^{-1} \int_{X \times X} \psi(x, y) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) \overline{F}_{n, \eta_{n}^f, \eta_{n}^c}(x) dx dy.
\]
By Part 1 and (42), \( T_1^N \to \mathbb{P} 0 \), and by (42) \( T_2^N \to \mathbb{P} 0 \); hence the proof is complete. \( \square \)

**Appendix E. Proofs for the asymptotic variance**

Throughout, \( v \) is as in (H1). We will frequently quote results from [30] whose proofs are given for \( v \) rather than \( v^\xi, \xi \in (0, 1) \). The proofs of these quoted results can be extended to the case \( \xi \in (0, 1) \) using [30, Lemma 3]. Note that [30, Lemma 3] implies that (H1)–(H4) apply in the case for \( \xi \in (0, 1) \), with Lyapunov function \( v^\xi \), with constants depending upon \( \xi \). In our proofs \( C \) is a finite and positive constant whose value may change from line to line, and it does not depend on \( n, p, f, c \). Any important dependencies of \( C \) will be stated.
\textbf{Proof of Theorem 4.3.} We have

\[ \sigma_{n\to}(\varphi) = \eta_n^{\varphi}\left( (\varphi \otimes 1 - 1 \otimes \varphi - (\eta_n^\varphi - \eta_n^\varphi)(\varphi)) \right) + \sum_{p=0}^{n-1} \eta_p^\varphi((D_{p,n}^f(\varphi) \otimes 1 - 1 \otimes D_{p,n}^c(\varphi))^2). \]

Define

\[ T_1 := 2\eta_p^\varphi((D_{p,n}^f(\varphi) - D_{p,n}^c(\varphi))^2), \]

\[ T_2 := 2\eta_p^\varphi((D_{p,n}^c(\varphi) \otimes 1 - 1 \otimes D_{p,n}^c(\varphi))^2). \]

We focus on the summand and note that by the $C_2$—inequality, one has

\[ T_1 + T_2 \geq \eta_p^\varphi((D_{p,n}^f(\varphi) \otimes 1 - 1 \otimes D_{p,n}^c(\varphi))^2). \]

Our proof for the case when only (H1)–(H6) hold, versus when (H1)–(H7) hold, is the same when considering $T_1$ in (43), versus $T_2$ in (44). Thus our proof is split into three parts, controlling $T_1$ and then $T_2$ first under (H1)–(H6) and then under (H1)–(H7). The proof is concluded by adding and then summing the bounds (and then summing over $p$) for the relevant case.

\textbf{Term (43):} We have for any $x \in X$ that

\[ (D_{p,n}^f(\varphi)(x) - D_{p,n}^c(\varphi)(x))^2 = D_{p,n}^f(\varphi)(x)[D_{p,n}^f(\varphi)(x) - D_{p,n}^c(\varphi)(x)] + D_{p,n}^c(\varphi)(x)[D_{p,n}^c(\varphi)(x) - D_{p,n}^f(\varphi)(x)]. \]

Applying Lemmata E.1 and E.3 yields the upper bound

\[ (D_{p,n}^f(\varphi)(x) - D_{p,n}^c(\varphi)(x))^2 \leq C(\varphi)\rho^{n-p}\left( \|h_{p,n}^{f,c}(\varphi, c_{p,n})(v)\| + \|\eta_n^\varphi - \eta_n^\varphi\| + \|\varphi\|\|h_{p,n}^f - h_{p,n}^c\|\rho^{n-p} \right)v(x)^{2\xi}. \]

Hence, as $\eta_p^\varphi(v^{\lambda_{\xi}}) \leq C [30, Proposition 1]$ (note that by (H2) and (H5) Part 2, the bound in the latter proposition does not depend on $\eta_p^\varphi$), it follows that

\[ |T_1| \leq C(\varphi)\rho^{n-p}\left( \|h_{p,n}^{f,c}(\varphi, c_{p,n})(v)\| + \|\eta_n^\varphi - \eta_n^\varphi\| + \|\varphi\|\|h_{p,n}^f - h_{p,n}^c\|\rho^{n-p} \right). \]

\textbf{Term (44), assuming only (H1)–(H6) hold:} We have for any $(x, y) \in A$ that

\[ (D_{p,n}^c(\varphi)(x) - D_{p,n}^c(\varphi)(y))^2 = D_{p,n}^c(\varphi)(x)[D_{p,n}^c(\varphi)(x) - D_{p,n}^c(\varphi)(y)] + D_{p,n}^c(\varphi)(y)[D_{p,n}^c(\varphi)(y) - D_{p,n}^c(\varphi)(x)] \]

(if $(x, y) \in A^c$ the term is zero). Applying Lemmata E.1 and E.6 yields the upper bound

\[ (D_{p,n}^c(\varphi)(x) - D_{p,n}^c(\varphi)(y))^2 \leq C(\varphi)\rho^{n-p}(v(x)v(y))^{2\xi}(\rho^{n-p} + 1) \]

for any $(x, y) \in A$. Hence it follows that

\[ |T_2| \leq C(\varphi)\rho^{n-p}\left( \eta_p^\varphi((1_A(v \otimes v))^{2\xi})(\rho^{n-p} + 1) \right). \]
Noting (45), (46), and (47), the proof can be concluded in the case that only \((H1)-(H6)\) hold.

**Lemma E.1.** Assume \((H1)-(H5)\). Then for any \(\xi \in (0, 1)\) there exist \(\rho < 1\) and \(C < +\infty\) also depending on the constants in \((H1)-(H5)\) such that for any \(\varphi \in \mathcal{L}_{\psi}(X)\), \(n \geq 1\) and \(0 \leq p < n\), \(x \in X\), \(s \in \{f, c\}\) we have

\[
|D_{p,n}^{s}(\varphi)(x)| \leq C\|\varphi\|_{\psi} \rho^{n-p} v(x)^{\xi}.
\]

**Proof.** The proof for \(\xi = 1\) is exactly as in [30, Theorem 1], except noting that \((H2)\) and \((H5)\) Part 2 establish that the constants in [30, Theorem 1] do not depend on \(\eta_0\) (the quantity denoted by \(\mu\) in [30]). \(\square\)

**Lemma E.2.** Assume \((H1)-(H6)\). Then for any \(\xi \in (0, 1)\), \(s \in \{f, c\}\) we have

\[
\sup_{n \geq 1} \sup_{0 \leq p \leq n} \max_{s \in \{f, c\}} \|1/h_{p,n}^{s}\|_{\psi} < +\infty.
\]

**Proof.** This is [12, Lemma B.2], except with the fix of using \((H5)\) Part 1 to go to the second line of that proof. \(\square\)

**Lemma E.3.** Assume \((H1)-(H6)\). Then for any \(\xi \in (0, 1)\) there exist \(\rho < 1\) and \(C < +\infty\) also depending on the constants in \((H1)-(H6)\) such that for any \(\varphi \in \mathcal{B}_{\rho}(X)\), \(n \geq 1\) and \(0 \leq p < n\), \(x \in X\) we have

\[
|D_{p,n}^{f}(\varphi)(x) - D_{p,n}^{c}(\varphi)(x)| \leq C\left(||h_{p,n}^{f} - h_{p,n}^{c}\||_{\psi} + ||\varphi\||h_{p,n}^{f} - h_{p,n}^{c}\|_{\psi} \rho^{n-p}\right)v(x)^{3\xi}.
\]

**Proof.** Define

\[
T_1 := \left(S_{p,n}^{f,c}(\varphi)(x) + \eta_n^{f} - \eta_n^{c}\right)h_{p,n}^{f}(x), \quad \text{(49)}
\]

\[
T_2 := D_{p,n}^{c}(\varphi)(x)\frac{1}{h_{p,n}^{c}(x)}\left[h_{p,n}^{f}(x) - h_{p,n}^{c}(x)\right]. \quad \text{(50)}
\]

Then we have

\[
T_1 + T_2 = D_{p,n}^{f}(\varphi)(x) - D_{p,n}^{c}(\varphi)(x). \quad \text{(51)}
\]

We will bound \(|T_1|\) and \(|T_2|\) and then sum the bounds to conclude.

**Term (49):** We have

\[
\left(S_{p,n}^{f,c}(\varphi)(x) + \eta_n^{f} - \eta_n^{c}\right)h_{p,n}^{f}(x) = \left(S_{p,n}^{f,c}(\varphi)(x)\right)(x)^{\xi} + \left(\eta_n^{f} - \eta_n^{c}\right)(x)h_{p,n}^{f}(x)(x)^{\xi}.
\]
Then by [30, Proposition 2], $\|h_{p,n}^f\|_{p,\xi}$ is bounded above by a finite constant that does not depend on $p,n,f$, so we easily have that

$$|T_1| \leq C\left(\|h_{p,n}^f\|_{p,\xi} + \|\eta_{n}^f - \eta_{n}^c\|_{p,\xi}\right)v(x)^{2\xi}. \quad (52)$$

**Term (50):** We have

$$D_{p,n}^c(\varphi)(x) = \frac{1}{h_{p,n}^c(x)}[h_{p,n}^f(x) - h_{p,n}^c(x)]$$

$$= \left(D_{p,n}^c(\varphi)(x)\frac{1}{h_{p,n}^c(x)} \frac{[h_{p,n}^f(x) - h_{p,n}^c(x)]}{v(x)^\xi}\right)v(x)^{2\xi}.$$

Applying Lemma E.1 to $D_{p,n}^c(\varphi)(x)$ and Lemma E.2 to $1/(h_{p,n}^c(x)v(x)^\xi)$ yields that

$$|T_2| \leq C\|\varphi\|\|h_{p,n}^f - h_{p,n}^c\|_{p,\xi}\rho^{n-p}v(x)^{3\xi}. \quad (53)$$

Noting (51), (52), and (53), the proof can be concluded. \hfill \Box

For $n \geq 1$, $0 \leq p < n$, $x \in X$, $s \in \{f,c\}$, define

$$R_{p+1}^{(n,s)}(x, dy) := \frac{1}{Q_{p,n}^s(1)(x)}Q_{p+1,n}^s(1)(y)Q_{p+1}^s(x, dy). \quad (54)$$

**Lemma E.4.** Assume (H1)–(H5), (H7). Then for any $\xi \in (0, 1]$ there exists a $C < +\infty$, depending on the constants in (H1)–(H5) and (H7), such that for any $\varphi \in B_0(X) \cap Lip_d(X)$, $(x, y) \in X \times X$, $s \in \{f,c\}$ we have

$$\sup_{n \geq 1} \sup_{0 \leq p \leq n} \frac{Q_{p,n}^s(\varphi)(x)}{\eta_{p}^s(Q_{p,n}^s(1))} - \frac{Q_{p,n}^s(\varphi)(y)}{\eta_{p}^s(Q_{p,n}^s(1))} \leq C\|\varphi\|d(x, y)v(x)^{\xi}v(y)^{\xi}. \quad (55)$$

**Proof:** We proceed by backward induction, starting with the case $p = n - 1$ (the case $p = n$ is trivial). We have

$$\frac{Q_{n-1,n}^s(\varphi)(x)}{\eta_{n-1}^s(Q_{n-1,n}^s(1))} - \frac{Q_{n-1,n}^s(\varphi)(y)}{\eta_{n-1}^s(Q_{n-1,n}^s(1))} = \frac{Q_{p}(\varphi)(x)}{\eta_{n-1}^s(G_{n-1})} - \frac{Q_{p}(\varphi)(y)}{\eta_{n-1}^s(G_{n-1})}. \quad (56)$$

By [30, Proposition 2(2)],

$$\inf_{p \geq 0} \min_{s \in \{f,c\}} \eta_{p}^s(G_{p}) \geq C > 0 \quad (55)$$

for some constant $C$. So applying (55) along with (H7) we have

$$\left|\frac{Q_{n-1,n}^s(\varphi)(x)}{\eta_{n-1}^s(Q_{n-1,n}^s(1))} - \frac{Q_{n-1,n}^s(\varphi)(y)}{\eta_{n-1}^s(Q_{n-1,n}^s(1))}\right| \leq C\|\varphi\|d(x, y)v(x)^{\xi}v(y)^{\xi}.$$

Now assuming the result for some $p + 1 < n$, we have

$$\frac{Q_{p,n}^s(\varphi)(x)}{\eta_{p}^s(Q_{p,n}^s(1))} - \frac{Q_{p,n}^s(\varphi)(y)}{\eta_{p}^s(Q_{p,n}^s(1))} = \frac{1}{\eta_{p}^s(G_{p})} \left(\frac{Q_{p+1,n}^s(Q_{p+1,n}^s(\varphi))(x)}{\eta_{p+1}^s(Q_{p+1,n}^s(1))} - \frac{Q_{p+1,n}^s(Q_{p+1,n}^s(\varphi))(y)}{\eta_{p+1}^s(Q_{p+1,n}^s(1))}\right).$$
Now, for any \( x \in X \), we have
\[
\frac{Q_{p+1,n}(\varphi)(x)}{\eta_{p+1}^s(Q_{p+1,n}(1))} \leq \|\varphi\| \|h_{p+1,n}^c\|_{v^\xi} \leq C\|\varphi\|,
\]
where we have used [30, Proposition 2(3)] to get to the second inequality. Then applying (55), the induction hypothesis, and (H7), we have
\[
\left| \frac{Q_{s,n}(\varphi)(x)}{\eta_{s}^s(Q_{s,n}(1))} - \frac{Q_{p,n}(\varphi)(y)}{\eta_{p}^s(Q_{p,n}(1))} \right| \leq C \left\| \frac{Q_{p+1,n}(\varphi)}{\eta_{p+1}^s(Q_{p+1,n}(1))} \right\| \|d(x,y)v^\xi v(y)^\xi \leq C\|\varphi\|d(x,y)v^\xi v(y)^\xi.
\]
\[\square\]

**Lemma E.5.** Assume (H1)–(H7). Then for any \( \xi \in (0,1/2) \) there exists a \( C < +\infty \) depending on the constants in (H1)–(H7) such that for any \( \varphi \in \mathcal{L}_{v^\xi}(X) \cap \operatorname{Lip}_{v^\xi,d}(X) \), \( n \geq 1 \) and \( 0 \leq p < n \), \( (x,y) \in X \times X \), \( s \in \{f,c\} \) we have
\[
\left| h_{p,n}^s(\varphi(x) - R_{p+1}^{(n),s}(\varphi)(y)) \right| \leq C\|\varphi\|_{v^\xi} d(x,y)v(x)^{2\xi}v(y)^{4\xi}.
\]

**Proof.** Define
\[
T_1 := \frac{h_{p,n}^s(x)}{Q_{p,n}(1)(x)} \left( Q_{p+1,n}^s(\varphi Q_{p+1,n}(1))(x) - Q_{p+1,n}^s(\varphi Q_{p+1,n}(1))(y) \right),
\]
\[
T_2 := \frac{h_{p,n}^s(x)Q_{p+1,n}^s(\varphi Q_{p+1,n}(1))(y)}{Q_{p,n}(1)(x)Q_{p,n}^s(1)(y)} \left( Q_{p,n}(1)(y) - Q_{p,n}^s(1)(x) \right).
\]
Then we have
\[
T_1 + T_2 = h_{p,n}^s(x) \left( R_{p+1}^{(n),s}(\varphi)(x) - R_{p+1}^{(n),s}(\varphi)(y) \right).
\]
We will bound \(|T_1|\) and \(|T_2|\) and then sum the bounds to conclude.

**Term (56):** We have
\[
\frac{h_{p,n}^s(x)}{Q_{p,n}(1)(x)} \left( Q_{p+1,n}^s(\varphi Q_{p+1,n}(1))(x) - Q_{p+1,n}^s(\varphi Q_{p+1,n}(1))(y) \right) = \frac{1}{\eta_{p}^s(G_p)} \left( Q_{p+1,n}^s(\varphi h_{p+1,n}^c)(x) - Q_{p+1,n}^s(\varphi h_{p+1,n}^c)(y) \right).
\]
Then by Lemma E.4 and [30, Proposition 2(3)],
\[
h_{p+1,n}^c \in \mathcal{L}_{v^\xi}(X) \cap \operatorname{Lip}_{v^\xi,d}(X) \quad \forall \xi \in (0,1],
\]
so, applying (55) and (H7), we have
\[
|T_1| \leq C\|\varphi h_{p+1,n}^c\|_{v^\xi} d(x,y)[v(x)v(y)]^{2\xi}.
\]
By [30, Proposition 2(3)],
\[
\|\varphi h_{p+1,n}^c\|_{v^\xi} \leq C\|\varphi\|_{v^\xi},
\]
so

\[ |T_1| \leq C \|\varphi\|_\infty d(x, y) [v(x)v(y)]^{2\xi}. \]  

(60)

**Term (57):** We have

\[
\frac{h^s_{p,n}(x)Q^s_{p+1} (\varphi Q^s_{p+1,n}(1))(y)}{Q^s_{p,n}(1)(x)Q^s_{p,n}(1)(y)} \left( Q^s_{p,n}(1)(y) - Q^s_{p,n}(1)(x) \right) 
\]

\[
= \frac{Q^s_{p+1} (\varphi Q^s_{p+1,n}(1))(y)}{Q^s_{p,n}(1)(y)} \left( Q^s_{p+1} (h^s_{p+1,n}(y) - Q^s_{p+1,n}(1)(x)) \right) 
\]

\[
= \frac{Q^s_{p+1} (\varphi h^s_{p+1,n}(y)v(y)^{\xi})}{\eta^p_{p,n}(G_p) h^s_{p,n}(y)v(y)^{\xi}} \left( Q^s_{p+1} (h^s_{p+1,n}(y) - Q^s_{p+1,n}(1)(x)) \right). 
\]

Applying (55) and Lemma E.2 gives the upper bound

\[ |T_2| \leq C Q^s_{p+1} (\varphi h^s_{p+1,n}(y)v(y)^{\xi}) \left| Q^s_{p+1} (h^s_{p+1,n}(y) - Q^s_{p+1,n}(1)(x)) \right|. \]

Now \( \varphi \in L_{\infty}(X) \), and by [30, Proposition 2(3)],

\[
\sup_{n \geq 1} \sup_{0 \leq p \leq \xi} \sup_{s \in \{f, c\}} \|h^s_{p+1,n}\|_{\infty} < +\infty; 
\]

thus, noting (59) and applying (H7), we have

\[ |T_2| \leq C \|\varphi\|_\infty Q^s_{p+1} (v^{2\xi})(y)v(y)^{\xi} \|h^s_{p+1,n}\|_{\infty} |x - y| [v(x)v(y)]^{\xi}. \]

Again applying [30, Proposition 2(3)] and (H1), we have finally that

\[ |T_2| \leq C \|\varphi\|_\infty d(x, y)v(x)v(y)^{\xi} |x - y|^{\xi} [v(x)v(y)]^{4\xi}. \]

(61)

Noting (58), (60), and (61), the proof can be concluded.

\[ \square \]

**Lemma E.6.** Assume (H1)–(H6). Then for any \( \xi \in (0, 1] \) there exist \( \rho < 1 \) and \( C < +\infty \) depending on the constants in (H1)–(H6) such that for any \( \varphi \in B_b(X) \), \( n \geq 1 \) and \( 0 \leq p < n \), \( (x, y) \in X \times X \) we have

\[ |D^c_{p,n}(\varphi)(x) - D^c_{p,n}(\varphi)(y)| \leq C \|\varphi\|_0 d(x, y)^{\rho^{n-p} + 1} v(x)^{\xi} v(y)^{2\xi}. \]

**Proof.** Throughout, we assume that \( x \neq y \). Define

\[
T_1 := \left( \frac{Q^c_{p,n}(\varphi)(x)}{Q^c_{p,n}(1)(x)} - \frac{Q^c_{p,n}(\varphi)(y)}{Q^c_{p,n}(1)(y)} \right) h^c_{p,n}(x),
\]

(62)

\[
T_2 := D^c_{p,n}(\varphi)(y) \left( h^c_{p,n}(x) \frac{h^c_{p,n}(x)}{h^c_{p,n}(y)} - 1 \right);
\]

(63)

then we have

\[ T_1 + T_2 = D^c_{p,n}(\varphi)(x) - D^c_{p,n}(\varphi)(y). \]

(64)

We will bound \( |T_1| \) and \( |T_2| \) and then sum the bounds to conclude.
Lemma E.7. The constants in (H1)–(H7) such that for any $\varphi \in B_b(X)$, we have

$$|T_1| \leq C\|\varphi\|(|Q_{p+1}^c(h_{p+1,n}^c)(x)| + |Q_{p+1}^c(h_{p+1,n}^c)(y)|).$$

By [30, Proposition 2(3)],

$$\sup_{n \geq 1} \sup_{0 \leq p \leq n, s \in \{f, c\}} \sup_{x \in [t, c]} \|h_{p+1,n}^c\|_{v^\xi} < +\infty;$$

then in addition applying (H1) we have

$$|T_1| \leq C\|\varphi\|(v(x)^\xi + v(y)^\xi).$$  \hspace{1cm} (65)

**Term (63):** We have

$$D_{p,n}^c(\varphi)(y) \left( \frac{h_{p,n}^c(x)}{h_{p,n}^c(y)} - 1 \right) = D_{p,n}^c(\varphi)(y) \left( \frac{1}{h_{p,n}^c(y)v(x)^\xi} \right) \frac{h_{p,n}^c(x) - h_{p,n}^c(y)}{v(x)^\xi v(y)^\xi} v(x)^\xi v(y)^\xi.$$

Then, applying Lemma E.1 for $D_{p,n}^c(\varphi)(y)$, Lemma E.2 for $1/(h_{p,n}^c(y)v(x)^\xi)$, and [30, Proposition 2(3)] as above, we have

$$|T_2| \leq C\|\varphi\|\rho^{n-p} v(x)^\xi v(y)^3\xi.$$  \hspace{1cm} (66)

Noting (64), (65), and (66), the proof can be concluded. \hfill \square

For $B \in \mathcal{X}$, $x \in X$, $s \in \{f, c\}$, define

$$P_{p,n}^s(B)(x) = \frac{Q_{p,n}^s(B)(x)}{Q_{p,n}^s(1)(x)}.$$

**Lemma E.7.** Assume (H1)–(H7). Then for any $\xi \in (0, 1)$ there exists a $C < +\infty$ depending on the constants in (H1)–(H7) such that for any $\varphi \in B_b(X) \cap Lip_d(X)$, $(x, y) \in X \times X$, $s \in \{f, c\}$ we have

$$\sup_{n \geq 1} \sup_{0 \leq p \leq n} |P_{p,n}^s(\varphi)(x) - P_{p,n}^s(\varphi)(y)| \leq C\|\varphi\|d(x, y)v(x)^\xi v(y)^\xi.$$

**Proof.** We assume $p < n$ as the case $p = n$ is trivial. Define

$$T_1 = \frac{1}{h_{p,n}^c(x)} \left( \frac{Q_{p,n}^s(\varphi)(x)}{\eta_{p}^s(Q_{p,n}^s(1))} - \frac{Q_{p,n}^s(\varphi)(y)}{\eta_{p}^s(Q_{p,n}^s(1))} \right),$$  \hspace{1cm} (67)

$$T_2 = \frac{Q_{p,n}^s(\varphi)(y)}{Q_{p,n}^s(1)(y) h_{p,n}^c(x)} \left( \frac{1}{\eta_{p}^s(Q_{p,n}^s(1))} - \frac{Q_{p,n}^s(1)(x)}{\eta_{p}^s(Q_{p,n}^s(1))} \right);$$  \hspace{1cm} (68)
then we have
\[ T_1 + T_2 = P^s_{p, n}(\varphi)(x) - P^s_{p, n}(\varphi)(y) \] (69)
We will bound \(|T_1|\) and \(|T_2|\) and then sum the bounds to conclude.

**Term (67):** We have, for any \(\tilde{\xi} \in (0, 1]\),
\[
\frac{1}{h^s_{p, n}(x)} \left( \frac{Q^s_{p, n}(\varphi)(x)}{\eta^s_p(Q^s_{p, n}(1))} - \frac{Q^s_{p, n}(\varphi)(y)}{\eta^s_p(Q^s_{p, n}(1))} \right) = \frac{v(x)^{\tilde{\xi}}}{h^s_{p, n}(x)} \left( \frac{Q^s_{p, n}(\varphi)(x)}{\eta^s_p(Q^s_{p, n}(1))} - \frac{Q^s_{p, n}(\varphi)(y)}{\eta^s_p(Q^s_{p, n}(1))} \right).
\]
Then applying Lemmata E.2 and E.7, for any \(\tilde{\xi} \in (0, 1]\) we have the upper bound
\[
|T_1| \leq C\|\varphi\|d(x, y)v(x)^{2\tilde{\xi}}v(y)^{\tilde{\xi}}. \tag{70}
\]

**Term (68):** We have, for any \(\tilde{\xi} \in (0, 1]\),
\[
\frac{Q^s_{p, n}(\varphi)(y)}{Q^s_{p, n}(1)(y)} \frac{1}{h^s_{p, n}(x)} \left( \frac{Q^s_{p, n}(1)(y)}{\eta^s_p(Q^s_{p, n}(1))} - \frac{Q^s_{p, n}(1)(x)}{\eta^s_p(Q^s_{p, n}(1))} \right) = \frac{Q^s_{p, n}(\varphi)(y)}{Q^s_{p, n}(1)(y)} \frac{v(x)^{\tilde{\xi}}}{h^s_{p, n}(x)} \left( \frac{Q^s_{p, n}(1)(y)}{\eta^s_p(Q^s_{p, n}(1))} - \frac{Q^s_{p, n}(1)(x)}{\eta^s_p(Q^s_{p, n}(1))} \right).
\]
Then applying \(\varphi \in B_b(X)\) and Lemmata E.2 and E.7, for any \(\tilde{\xi} \in (0, 1]\) we have the upper bound
\[
|T_2| \leq C\|\varphi\|d(x, y)v(x)^{2\tilde{\xi}}v(y)^{\tilde{\xi}}. \tag{71}
\]
Noting (69), (70), and (71) (the latter two with \(\tilde{\xi} = \xi/2\)), the proof can be concluded. \(\square\)

**Lemma E.8.** Assume (H1)–(H7). Then for any \(\xi \in (0, 1/2]\) there exist \(\rho < 1\) and \(C < +\infty\) depending on the constants in (H1)–(H7) such that for any \(\varphi \in B_b(X) \cap \text{Lip}_d(X), n \geq 1\) and \(0 \leq p < n, (x, y) \in X \times X\) we have
\[
|D^c_{p, n}(\varphi)(x) - D^c_{p, n}(\varphi)(y)| \leq C\|\varphi\|\rho^{n-p}d(x, y)v(x)^{2\tilde{\xi}}v(y)^{4\xi}.
\]

**Proof:** The proof uses the same decomposition (64) as the proof of Lemma E.6. We use the same bound (66) as in the proof of Lemma E.6 and focus on the term \(T_1\) as defined in (62). We note by [8, p.133] that
\[
P^c_{p, n}(\varphi)(x) = R^{(n), c}_{p+1} (P^c_{p+1, n}(\varphi))(x),
\]
where \(R^{(n), c}_{p+1}\) is defined in (54).

Now, \(T_1\) as in the proof of Lemma E.6 can be written as
\[
T_1 = h^c_{p, n}(x) \left( P^c_{p, n}(\varphi)(x) - P^c_{p, n}(\varphi)(y) \right)
\]
\[
= h^c_{p, n}(x) \left( P^c_{p, n}(\varphi - \eta^c_n(\varphi))(x) - P^c_{p, n}(\varphi - \eta^c_n(\varphi))(y) \right)
\]
\[
= h^c_{p, n}(x) \left( R^{(n), c}_{p+1, n}(\varphi - \eta^c_n(\varphi))(x) - R^{(n), c}_{p+1, n}(\varphi - \eta^c_n(\varphi))(y) \right).
\]
Noting Lemma E.7, we apply Lemma E.5 to yield the upper bound
\[
|T_1| \leq C \left\| (P^c_{p+1, n}(\varphi - \eta^c_n(\varphi))) \right\| \nu^{\tilde{\xi}}d(x, y)v(x)^{2\tilde{\xi}}v(y)^{4\xi}. \tag{72}
\]
Now
\[
\frac{(P_{p+1,n}^c(\varphi - \eta_n^c(\varphi))(x))}{v(x)^\xi} = \frac{D_{p+1,n}^c(\varphi)(x)}{v(x)^{\xi/2}} \frac{1}{h_{p+1,n}^c(x)v(x)^{\xi/2}}.
\]
Applying Lemma E.1 and Lemma E.2 yields that
\[
\| (P_{p+1,n}^c(\varphi - \eta_n^c(\varphi)) \|_\xi \leq C \| \varphi \| v^\rho n^{-p}. \tag{73}
\]
Then, using the above inequality in (72) and noting (65) and (66), the proof can be concluded. □

**Appendix F. Proofs for the diffusion case**

The statements at the beginning of Appendix E also apply here. This appendix should be read after Appendix E.

**F.1. Proofs for Proposition 5.1**

**Proof of Proposition 5.1.** This follows easily from Propositions F.1, F.2, and F.3. □

**Lemma F.1.** Assume (H8). Then for any \( \xi \in (0, 1/2) \) there exists a \( C < +\infty \) depending on the constants in (H8) such that for any \( \varphi \in L_\psi^v(X) \), \( 1 \leq l \leq L \), \( x \in X \) we have
\[
|M^l(\varphi)(x) - M^{l-1}(\varphi)(x)| \leq C \| \varphi \|_\psi \Delta l v(x)^{2\xi}. \tag{77}
\]

**Proof.** Define
\[
T_1 := \int_{B_l(x)^c} \varphi(y)[M^l(x, y) - M(x, y) - \{M^{l-1}(x, y) - M(x, y)\}]dy, \tag{74}
\]
\[
T_2 := \int_{B_l(x)^c} \varphi(y)[M^l(x, y) + M(x, y) - \{M^{l-1}(x, y) + M(x, y)\}]dy. \tag{75}
\]
Then we have
\[
T_1 + T_2 = M^l(\varphi)(x) - M^{l-1}(\varphi)(x). \tag{76}
\]
We will bound \( |T_1| \) and \( |T_2| \) and then sum the bounds to conclude.

**Term (74):** We have, by applying (11),
\[
\left| \int_{B_l(x)} \varphi(y)[M^l(x, y) - M(x, y) - \{M^{l-1}(x, y) - M(x, y)\}]dy \right| \leq C \Delta_l \int_{B_l(x)} |\varphi(y)| \exp\{-C'\|y - x\|^2\} dy.
\]
Using \( \varphi \in L_\psi^v(X) \) and (H8) Part 2, we have
\[
|T_1| \leq C \| \varphi \|_\psi \Delta l v(x)^{2\xi}. \tag{77}
\]

**Term (75):** We have, by applying (10),
\[
\left| \int_{B_l(x)^c} \varphi(y)[M^l(x, y) + M(x, y) - \{M^{l-1}(x, y) + M(x, y)\}]dy \right| \leq C \int_{B_l(x)^c} |\varphi(y)| \exp\{-C'\|y - x\|^2\} dy.
\]
Using $\varphi \in \mathcal{L}_{\psi}(X)$ and the fact that
\[
\int_{B(x)^c} \nu(y)^\xi \exp\{-C\|y-x\|^2\} dy \leq C \int_{B(x)^c} \nu(y)^\xi dy,
\]
along with (H8) and the fact that $\int_{B(x)^c} dy \leq C \Delta_l$, we have
\[
|T_2| \leq C \|\varphi\|_{\psi} \Delta_l \nu(x)^{2\xi}.
\]  
Noting (76), (77), and (78) the proof can be concluded. \qed

Let $\xi \in (0, 1]$ and denote by $\mathcal{P}_{\psi}(X)$ the collection of probability measures for which $\mu(\nu^\xi) < +\infty$.

**Lemma F.2.** Assume (H5) Part 1 and (H8). Then for any $\xi \in (0, 1/2)$ there exists a $C < +\infty$ depending on the constants in (H5) Part 1 and (H8) such that for any $\mu \in \mathcal{P}_{\psi}(X)$, $\varphi \in \mathcal{L}_{\psi}(X)$, $p \geq 1$, $1 \leq l \leq L$ we have
\[
||\Phi_p^l(\mu) - \Phi_p^{l-1}(\mu)||_{\psi} \leq \frac{C\|\varphi\|_{\psi} \Delta_l \nu(x)^{2\xi}}{\mu(G_{p-1})}.
\]

**Proof.** We have for $\varphi \in \mathcal{L}_{\psi}(X)$ that
\[
[\Phi_p^l(\mu) - \Phi_p^{l-1}(\mu)](\varphi) = \frac{1}{\mu(G_{p-1})} \mu(G_{p-1}[M_p(\varphi) - M_p^{l-1}(\varphi)]).
\]
Applying (H5) Part 1 along with Lemma F.1 allows one to conclude. \qed

**Lemma F.3.** Assume (H1)–(H5). Then for any $\xi \in (0, 1]$ there exists a $C < +\infty$ depending on the constants in (H1)–(H5) such that for any $1 \leq l \leq L$ we have
\[
\sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q_{p,n}^{l-1}(1)}{\eta_p^l(Q_{p,n}^{l-1}(1))} \right\|_{\psi} \leq C, 
\]  
\[
\sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q_{p,n}^{l-1}(1)}{\Phi_p^l(\eta_{p-1}^l(Q_{p,n}^{l-1}(1)))} \right\|_{\psi} \leq C, 
\]  
If additionally $\xi \in (0, 1/2)$, then there exists a $C < +\infty$ depending on the constants in (H1)–(H5) such that for any $\varphi \in \mathcal{L}_{\psi}(X)$, $l \geq 1$ we have
\[
\sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q_{p,n}^{l-1}(\varphi)}{\eta_p^l(Q_{p,n}^{l-1}(1))} \right\|_{\psi} \leq C\|\varphi\|_{\psi}.
\]

**Proof.** The proofs of (79) (resp. (80)) and (82) (resp. (81)) are very similar, so we only give the proofs of (80) and (82).

**Proof of (80):** We have for any $x \in X$ that
\[
\frac{Q_{p,n}^{l-1}(1)(x)}{\Phi_p^{l-1}(\eta_{p-1}^l(Q_{p,n}^{l-1}(1)))} = \frac{h_{p,n}^{l-1}(x)}{\Phi_p^{l-1}(\eta_{p-1}^l(h_{p,n}^{l-1}(x)))}.
\]
Now
\[
\Phi_p^{l-1}(\eta_{p-1}^l)(h_{p,n}^{l-1}) = \frac{\eta_{p-1}^l(Q_p^{l-1}(h_{p,n}^{l-1}))}{\eta_{p-1}^l(G_{p-1})} \\
\geq \frac{\eta_{p-1}^l(\mathbb{I}_d Q_p^{l-1}(\mathbb{I}_d h_{p,n}^{l-1}))}{\sup_{n \geq 0} \| G_n \|} \\
\geq C\eta_{p-1}^l(C_d)\epsilon_\delta v(\mathbb{I}_d h_{p,n}^{l-1}) \\
\geq C\epsilon_\delta(C_d)\inf_{n \geq 1} \inf_{0 \leq p \leq n, x \in C_d} h_{p,n}^{l-1} \\
\geq C,
\]
where we have applied (H5) Part 1 and (H3) to obtain line 3, [30, p. 2527] (last displayed equation) to obtain line 4, and then [30, Lemma 10] to obtain the final line. Then, using [30, Proposition 2], it follows that for any \(x \in X\),
\[
\left| \frac{Q_{p,n}^{l-1}(1)(x)}{\Phi_p^{l-1}(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))} \right| \leq C\epsilon_\delta(x).
\]
which completes the proof of (80).

**Proof of (82):** We have for any \(x \in X\) that
\[
\frac{Q_{p,n}^{l-1}(\varphi)(x)}{\eta_p^{l-1}(Q_{p,n}^{l-1}(1))} = \frac{Q_{p,n}^{l-1}(\varphi)(x)}{Q_{p,n}^{l-1}(1)(x)} \cdot \frac{Q_{p,n}^{l-1}(1)(x)}{\eta_p^{l-1}(Q_{p,n}^{l-1}(1))} \cdot \frac{1}{\eta_p^{l-1}(h_{p,n}^{l-1})}.
\]
By using an argument very similar to (83), one can deduce that \(\eta_p^{l}(h_{p,n}^{l-1}) \geq C\), as frequently noted \(\| h_{p,n}^{l-1} \|_{\epsilon,\delta} \leq C\), and by a proof similar to that of [30, Proposition 1],
\[
|Q_{p,n}^{l-1}(\varphi)(x)|/Q_{p,n}^{l-1}(1)(x) \leq C\| \varphi \|_{\epsilon,\delta} v(x)^{\epsilon_\delta}.
\]
Hence one can complete the proof of (82). \(\square\)

Let \(\mu \in \mathcal{P}(X)\), \(s \in \{l, l-1\}\), \(l \geq 1\), \(n \geq 1\), \(0 \leq p \leq n\), and define \(\Phi_{s,(p,n)}^l(\mu) = \Phi_s^l \circ \cdots \circ \Phi_{s+1}^l(\mu)\), with the convention that if \(n = p\), \(\Phi_{s,(p,n)}^l(\mu) = \mu\) and \(\Phi_{s}^l(\mu) = \mu\).

**Lemma F.4.** Assume (H1)–(H5), (H8). Then for any \(\xi \in (0, 1/4)\) there exists a \(C < +\infty\) depending on the constants in (H1)–(H5) and (H8) such that for any \(\varphi \in \mathcal{L}_{\epsilon,\delta}(X)\), \(n \geq 1\), \(1 \leq l \leq L\) we have
\[
||\Phi_{s,(0,n)}^l(\eta_{0}^l) - \Phi_{s,(0,n)}^{l-1}(\eta_{0}^{l-1})(\varphi)|| \leq C\| \varphi \|_{\epsilon,\delta} \Delta_n v(x)^{4\xi}.
\]

**Proof:** Define
\[
T_1 := [\eta_{0}^l - \eta_{0}^{l-1}] \left( \frac{Q_{0,n}^{l-1}(\varphi)}{\eta_{0}^{l-1}(Q_{0,n}^{l-1}(1))} \right),
\]
\[
T_2 := \left( \frac{\eta_{0}^{l-1}(Q_{0,n}^{l-1}(\varphi))}{\eta_{0}^{l-1}(Q_{0,n}^{l-1}(1))} \right) [\eta_{0}^{l-1} - \eta_{0}^l] \left( \frac{Q_{0,n}^{l-1}(1)}{\eta_{0}^{l-1}(Q_{0,n}^{l-1}(1))} \right).
\]
Then we have
\[ T_1 + T_2 = \Phi_{(0,n)}^l(\eta_0^l) - \Phi_{(0,n)}^{l-1}(\eta_0^{l-1}))(\varphi). \quad (86) \]

We will bound \(|T_1|\) and \(|T_2|\) and then sum the bounds to conclude.

**Term (84):** We have by Lemma F.3 (82) and Lemma F.1 that
\[ |T_1| \leq C\|\varphi\|_{\tilde{v}} \Delta_l \nu(x)^{\xi}. \quad (87) \]

**Term (85):** We have
\[
\left(\eta_0^{l-1}(Q_0^{l-1}(\varphi))\right) \left(\eta_0^{l-1} - \eta_0^{l}\right) \left(\frac{Q_{0,n}^l(1)}{\eta_0(Q_{0,n}^l(1))}\right) = \eta_n^{l-1}(\varphi)[\eta_n^{l-1} - \eta_0^{l}](\frac{Q_{0,n}^l(1)}{\eta_0(Q_{0,n}^l(1))}).
\]

Then, using \(\varphi \in \mathcal{L}_{\tilde{v}}(X)\) along with [30, Proposition 1] for \(\eta_n^{l-1}(\varphi)\) and Lemma F.3 (79) and Lemma F.1 for the other term on the right-hand side, we have
\[ |T_2| \leq C\|\varphi\|_{\tilde{v}} \Delta_l \nu(x)^{2\xi}. \quad (88) \]

Noting (86), (87), and (88), the proof can be concluded. \(\square\)

**Lemma F.5.** Assume (H1)–(H6), (H8). Then for any \((\xi, \tilde{\xi}) \in (0, 1/8) \times (0, 1/2)\) there exist \(\rho < 1\) and \(C < +\infty\) depending on the constants in (H1)–(H6) and (H8) such that for any \(\varphi \in \mathcal{L}_{\tilde{v}}(X)\), \(n \geq 1\), \(0 \leq p \leq n\), \(1 \leq l \leq L\) we have
\[
|\Phi_{(p,n)}^{l-1}(\Phi_p^{l}(\eta_{p-1}^{l})) - \Phi_{(p,n)}^{l-1}(\Phi_p^{l-1}(\eta_{p-1}^{l}))|(\varphi)| \leq C\|\varphi\|_{\tilde{v}} \Delta_l \rho^{n-\rho} \nu(x)^{\xi + 2\xi}.
\]

**Proof.** Define
\[
T_1 := [\Phi_p^{l}(\eta_{p-1}^{l}) - \Phi_p^{l-1}(\eta_{p-1}^{l})](\frac{Q_{p,n}^{l-1}(1)\rho_{p,n}^{l-1}(\tilde{\varphi}_n)}{\Phi_p^{l}(\eta_{p-1}^{l})(Q_{p,n}^{l}(1))}), \quad (89)
\]
\[
T_2 := \left(\Phi_p^{l-1}(\eta_{p-1}^{l})(Q_{p,n}^{l-1}(1)\rho_{p,n}^{l-1}(\tilde{\varphi}_n))\right) \cdot \left(\Phi_p^{l-1}(\eta_{p-1}^{l})(Q_{p,n}^{l-1}(1))\right)
\times [\Phi_p^{l-1}(\eta_{p-1}^{l}) - \Phi_p^{l}(\eta_{p-1}^{l})](\frac{Q_{p,n}^{l-1}(1)}{\Phi_p^{l}(\eta_{p-1}^{l})(Q_{p,n}^{l}(1))}), \quad (90)
\]
where \(\tilde{\varphi}_n = \varphi - \eta_n^{l-1}(\varphi)\). Then we have
\[ T_1 + T_2 = [\Phi_{(p,n)}^{l-1}(\Phi_p^{l}(\eta_{p-1}^{l})) - \Phi_{(p,n)}^{l-1}(\Phi_p^{l-1}(\eta_{p-1}^{l}))|(\varphi)|. \quad (91) \]

We will bound \(|T_1|\) and \(|T_2|\) and then sum the bounds to conclude.

**Term (89):** We have by (a similar proof to) (73) that
\[ \|P_{p,n}^{l-1}(\tilde{\varphi}_n)\|_{2\xi} \leq C\|\varphi\|_{\tilde{v}} \rho^{n-\rho}, \quad (92) \]
and by Lemma F.3 (81),
\[ \|Q_{p,n}^{l-1}(1)\Phi_p^{l}(\eta_{p-1}^{l})(Q_{p,n}^{l-1}(1))\|_{2\xi} \leq C, \]

so by Lemma F.2,
\[ |\Phi^l_p (\eta^l_{p-1}) - \Phi^{l-1}_p (\eta^{l-1}_{p-1})| \left( \frac{Q^{l-1}_{p,n}(1) \eta^{l-1}_{p-1}(\tilde{\varphi}_n)}{\Phi^l_p (\eta^l_{p-1})(Q^{l-1}_{p,n}(1))} \right) \leq C \|\varphi\|_{\psi} \Delta_l \rho^{n-p} \eta^{l-1}_{p-1}(v^{8\xi}). \]
Applying [30, Proposition 1, Proposition 2(2)] yields
\[ |T_1| \leq C \|\varphi\|_{\psi} \Delta_l \rho^{n-p} v(x^*)^{8\xi}. \] (93)

**Term (90):** We have by Lemma F.3 (80) that
\[ \|Q^{l-1}_{p,n}(1) / \Phi^{l-1}_p (\eta^{l-1}_{p-1})(Q^{l-1}_{p,n}(1))\|_{2\xi} \leq C, \]
so by (92),
\[ |T_2| \leq C \|\varphi\|_{\psi} \rho^{n-p} \left( \frac{\eta^{l-1}_{p-1}(Q^{l-1}_p(\eta^{4\xi})))}{\eta^{l-1}_{p-1}(G_{p-1})} \right) \left[ \Phi^{l-1}_p (\eta^{l-1}_{p-1}) - \Phi^{l-1}_p (\eta^{l-1}_{p-1}) \right] \left( \frac{Q^{l-1}_{p,n}(1)}{\Phi^l_p (\eta^l_{p-1})(Q^{l-1}_{p,n}(1))} \right). \]
Applying (H1) and [30, Proposition 1, Proposition 2(2)] to \( \eta^{l-1}_{p-1}(Q^{l-1}_p(\eta^{4\xi}))) / \eta^{l-1}_{p-1}(G_{p-1}) \) and Lemma F.2 along with Lemma F.3 (81) and [30, Proposition 1, Proposition 2(2)] to the remaining term gives
\[ |T_2| \leq C \|\varphi\|_{\psi} \Delta_l \rho^{n-p} v(x^*)^{4\xi + 2\tilde{\xi}}. \] (94)
Noting (91), (93), and (94), the proof can be concluded. \( \square \)

**Proposition F.1.** Assume (H1)–(H6), (H8). Then for any \( \xi, \tilde{\xi} \in (0, 1/8) \times (0, 1/2) \) there exists a \( C < +\infty \) depending on the constants in (H1)–(H6) and (H8) such that for any \( \varphi \in L_{v^\xi}(X), n \geq 0, 1 \leq l \leq L \) we have
\[ ||\eta^l_{n} - \eta^{l-1}_{n}(\varphi)|| \leq C \|\varphi\|_{\psi} \Delta_l v(x^*)^{8\xi + 2\tilde{\xi}}. \]

**Proof.** The case \( n = 0 \) follows by Lemma F.1, so we suppose \( n \geq 1 \). Define
\[ T_1 := [\Phi^{l}_l(\eta^l_0) - \Phi^{l-1}_l(\eta^{l-1}_0)](\varphi), \] (95)
\[ T_2 := [\Phi^{l-1}_l(\eta^l_0) - \Phi^{l-1}_l(\eta^{l-1}_0)](\varphi). \] (96)
Then we have
\[ T_1 + T_2 = [\eta^l_n - \eta^{l-1}_n](\varphi). \] (97)
We will bound \( |T_1| \) and \( |T_2| \) and then sum the bounds to conclude.

**Term (95):** We have the standard telescoping sum identity
\[ [\Phi^{l}_l(\eta^l_0) - \Phi^{l-1}_l(\eta^{l-1}_0)](\varphi) = \sum_{p=1}^{n} \left\{ [\Phi^{l-1}_p(\Phi^{l}_p(\eta^l_{p-1})) - \Phi^{l-1}_p(\Phi^{l-1}_p(\eta^{l-1}_{p-1}))](\varphi) \right\}. \] (98)
Applying Lemma F.5 gives
\[ |T_1| \leq C \|\varphi\|_{\psi} \Delta_l v(x^*)^{8\xi + 2\tilde{\xi}}. \] (99)
Term (96): We have by Lemma F.4 that
\[ |T_2| \leq C \| \varphi \|_{\mathcal{I}} \Delta_I v(x^*)^{4\xi}. \]  
(100)
Noting (97), (99), and (100), the proof can be concluded. \( \square \)

**Remark F.1.** If \( \varphi \in B_b(X) \) in Lemma F.5 and Proposition F.1, one can take \((\xi, \hat{\xi}) \in (0, 1/2) \times (0, 1] \); then the upper bound in Lemma F.5 is \( C \| \varphi \| \Delta_I \rho^{n-p} v(x^*)^{2\xi+\hat{\xi}} \) and the one in Proposition F.1 is \( C \| \varphi \| \Delta_I v(x^*)^{2\xi+\hat{\xi}} \).

**Proposition F.2.** Assume \((H1)–(H6), (H8)\). Then for any \( \xi \in (0, 1/4) \) there exists a \( C < +\infty \) depending on the constants in \((H1)–(H6)\) and \((H8)\) such that for any \( \varphi \in B_b(X) \), \( n \geq 1, 0 \leq p \leq n, 1 \leq l \leq L \) we have
\[ \| h^l_{p,n} S^{l-1}_{p,n}(\varphi) \|_{\mathcal{I}} \leq C \| \varphi \| \Delta_I. \]

**Proof.** For any \( x \in X \), \( 0 < \kappa < \xi \) one has
\[ \frac{h^l_{p,n}(x) S^{l-1}_{p,n}(\varphi)(x)}{v(x)^\xi} = \frac{h^l_{p,n}(x)}{v(x)^\xi} \frac{1}{\xi - \kappa} \left( p^l_{p,n}(\varphi)(x) - p^{l-1}_{p,n}(\varphi)(x) \right) \]
and
\[ p^l_{p,n}(\varphi)(x) - p^{l-1}_{p,n}(\varphi)(x) = \sum_{q=p+1}^{n} \left\{ \left[ \Phi^l_{(q,n)}(\Phi^l_{(p,q)}(\delta_x)) - \Phi^{l-1}_{(q,n)}(\Phi^{l-1}_{(p,q-1)}(\delta_x))) \right](\varphi) \right\}. \]
Noting (98), one can repeat the proofs of Lemmata F.3 and F.5 in almost the same way for this case to deduce that for any \((\tilde{\kappa}, \hat{\kappa}) \in (0, 1/2) \times (0, 1], \)
\[ \left| \frac{1}{\xi - \kappa} \left( p^l_{p,n}(\varphi)(x) - p^{l-1}_{p,n}(\varphi)(x) \right) \right| \leq C \| \varphi \| \Delta_I \rho^{n-p} v(x)^{2\xi+\hat{\kappa}}, \]
where we have noted Remark F.1. Choose \( 0 < \tilde{\kappa} = \hat{\kappa} < 1/(24), \kappa = \xi/2; \) then one can set \( \xi = 6\tilde{\kappa} \) and \( 0 < \xi < 1/4 \). Hence
\[ \left| \frac{1}{\xi - \kappa} \left( p^l_{p,n}(\varphi)(x) - p^{l-1}_{p,n}(\varphi)(x) \right) \right| \leq C \| \varphi \| \Delta_I \]
and the proof is concluded. \( \square \)

**Lemma F.6.** Assume \((H1)–(H5)\). Then for any \( \xi \in (0, 1] \) there exists a \( C < +\infty \) depending on the constants in \((H1)–(H5)\) such that for any \( 1 \leq l \leq L \) we have
\[ \sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q^{l-1}_{p,n}(1)}{Q^l_{p,n}(1)} \right\|_{\mathcal{I}} \leq C. \]

**Proof.** The proof is the same as for [30, Lemma 8], as the constants in \((H1)–(H5)\) do not depend upon \( l \). \( \square \)

**Proposition F.3.** Assume \((H1)–(H6), (H8)\). Then for any \((\xi, \hat{\xi}) \in (0, 1/8) \times (0, 1/2) \) there exists a \( C < +\infty \) depending on the constants in \((H1)–(H6)\) and \((H8)\) such that for any \( n \geq 1, 0 \leq p \leq n, 1 \leq l \leq L \), we have
\[ \| h^l_{p,n} - h^{l-1}_{p,n} \|_{\mathcal{I}} \leq C(n - p) \Delta_I v(x^*)^{\xi}. \]
Proof: For any \( x \in \mathcal{X} \), define

\[
T_1 := \frac{1}{\eta_p^l(Q_{p,n}(1))} (Q_{p,q}^l(1)(x) - Q_{p,q-1}^{l-1}(1)(x)),
\]

\[
T_2 := \frac{Q_{p,n}^{l-1}(1)(x)}{\eta_p^l(Q_{p,n}(1))} [\eta_p^{l-1} - \eta_p^l](h_{p,n}^{l-1}),
\]

\[
T_3 := h_{p,n}^{l-1}(x) \eta_p^l \left( \frac{1}{\eta_p^l(Q_{p,n}(1))} (Q_{p,n}^l(1) - Q_{p,n}^{l-1}(1)) \right).
\]

Then we have

\[
T_1 + T_2 - T_3 = h_{p,n}^l(x) - h_{p,n}^{l-1}(x).
\]

We will bound \(|T_1|, |T_2|, \) and \(|T_3|\) and then sum the bounds to conclude.

**Term** (101): We have

\[
T_1 = \frac{1}{\eta_p^l(Q_{p,n}(1))} \sum_{q=p+1}^n \left\{ Q_{p,q}^l(Q_{q,n}^{l-1}(1))(x) - Q_{p,q-1}^l(Q_{q-1,n}^{l-1}(1))(x) \right\}.
\]

Considering the summand, we have

\[
Q_{p,q}^l(Q_{q,n}^{l-1}(1))(x) - Q_{p,q-1}^l(Q_{q-1,n}^{l-1}(1))(x)
= \eta_{q+1}^{l-1} Q_{q+1,n}^{l-1}(1) \left( Q_{p,q-1}^l(G_{q-1}[M_q^l - M_{q-1}^l]h_{q+1,n}^{l-1}) \right).
\]

Now, applying Lemma F.1, [30, Proposition 2(3)] (with \( \nu^{\xi/4} \)), and (H5) Part 1, we have

\[
|Q_{p,q}^l(Q_{q,n}^{l-1}(1))(x) - Q_{p,q-1}^l(Q_{q-1,n}^{l-1}(1))(x)| \leq C \Delta \eta_{q+1}^{l-1} (Q_{q+1,n}^{l-1}(1)) Q_{p,q-1}^l(\nu^{\xi/2})(x).
\]

Thus

\[
|T_1| \leq \frac{C \Delta l}{\eta_p^l(Q_{p,n}(1))} \sum_{q=p+1}^n \left\{ \eta_{q+1}^{l-1} (Q_{q+1,n}^{l-1}(1)) Q_{p,q-1}^l(\nu^{\xi/2})(x) \right\}.
\]

Now,

\[
\frac{\eta_{q+1}^{l-1} (Q_{q+1,n}^{l-1}(1)) Q_{p,q-1}^l(\nu^{\xi/2})(x)}{\eta_p^l(Q_{p,n}(1))}
= \eta_{q+1}^{l-1} \left( \frac{Q_{q+1,n}^{l-1}(1)}{\eta_q^l(G_q) \eta_{q+1}^l(Q_{q+1,n}^{l-1}(1))} \right) \frac{Q_{p,q-1}^l(\nu^{\xi/2})(x)}{\eta_p^l(Q_{p,q-1}^l(1))} \frac{1}{\eta_{q-1}^l(G_{q-1})}.
\]

Applying Lemma F.6, [30, Proposition 2(3)], and then [30, Proposition 1] gives for any \( \nu^{\xi} \in (0, 1) \) that

\[
\eta_{q+1}^{l-1} \left( \frac{Q_{q+1,n}^{l-1}(1)}{\eta_q^l(G_q) \eta_{q+1}^l(Q_{q+1,n}^{l-1}(1))} \right) \leq C \nu(\nu^{\xi})\nu^{\xi}.
\]

In addition,

\[
\frac{Q_{p,q-1}^l(\nu^{\xi/2})(x)}{\eta_p^l(Q_{p,q-1}^l(1))} = \frac{Q_{p,q-1}^l(\nu^{\xi/2})(x)}{Q_{p,q-1}^l(1)} h_{p,q-1}(x).
\]
By a proof similar to that of [30, Proposition 1],
\[ Q_{p,q-1}(\nu^{\xi/2})/Q_{p,q-1}(1) \leq C\nu(\xi)^{\xi/2}, \]
and thus by [30, Proposition 2(3)], \( h_{p,q-1}(x) \leq C\nu(x)^{\xi/2} \), so
\[ \frac{Q_{p,q-1}(\nu^{\xi/2})}{\eta_p'(Q_{p,q-1}(1))} \leq C\nu(x)^{\xi}. \] (108)

Noting (105)–(108), we have shown that for any \((\xi, \hat{\xi}) \in (0, 1/8) \times (0, 1/2)\),
\[ |T_1| \leq C(n - p)\Delta I \nu(x)^{\xi} \nu(x^*)^{\hat{\xi}}. \] (109)

**Term (102):** By Lemma F.6 and Proposition F.1, we have that for any \((\xi, \kappa, \hat{\kappa}) \in (0, 1/8)^2 \times (0, 1/2)\),
\[ |T_2| \leq C\nu(x)^{\xi} \| h_{p,n}^{l-1} \|_{\nu^\kappa} \Delta I \nu(x^*)^{8\kappa + 2\hat{\xi}}. \]
One can choose \(0 < \kappa = \hat{\kappa} < 1/20, \hat{\xi} = 10\kappa\), and applying [30, Proposition 2(3)], we have
\[ |T_2| \leq C \Delta I \nu(x)^{\xi} \nu(x^*)^{\hat{\xi}}. \] (110)

**Term (103):** We have by [30, Proposition 2(3)] and (109) that for any \((\kappa, \hat{\kappa}) \in (0, 1/8) \times (0, 1/2)\),
\[ |T_3| \leq C\nu(x)^{\xi} (n - p)\Delta I \nu(x^*)^{\xi}. \]
Applying [30, Proposition 1] and choosing \((\kappa, \hat{\kappa})\), so that \(\hat{\xi} = \kappa + \hat{\kappa} < 1/2\), we have
\[ |T_3| \leq C(n - p)\Delta I \nu(x)^{\xi} \nu(x^*)^{\hat{\xi}}. \] (111)

Noting (104), (109), (110), and (111), the proof can be concluded. \( \square \)

**F.2. Proofs for Proposition F.2**

**Proof of Proposition F.2.** We focus on the first bound in Theorem 4.3. Clearly
\[ \tilde{\eta}_n^C((\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}](\varphi)))^2) \leq 2\|\varphi\|^2 \tilde{\eta}_n^C(\mathbb{I}_A). \]

As \( \tilde{\eta}_n^C \) is the maximal coupling of \((\eta_n^l, \eta_n^{l-1})\), we have \( \tilde{\eta}_n^C(\mathbb{I}_A) = \|\eta_n^l - \eta_n^{l-1}\|_{\text{tv}} \). Hence, on applying Proposition F.1 (noting Remark F.1), we have that
\[ \tilde{\eta}_n^C((\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}](\varphi)))^2) \leq C\|\varphi\|^2 \Delta I \nu(x^*)^{2\xi + \hat{\xi}}. \]

The proof is completed by using Proposition F.1 and Proposition F.4. \( \square \)

**Proposition F.4.** Assume (H1)–(H6), (H8). Then for any \((\xi, \hat{\xi}) \in (0, 1/32) \times (0, 1/2)\) there exists a \(C < +\infty\) depending on the constants in (H1)–(H6) and (H8) such that for any \(n \geq 0\), \(1 \leq l \leq L\) we have
\[ \tilde{\eta}_n^C(\mathbb{I}_A(\nu \otimes \nu)^{2\xi}) \leq C\Delta I \nu(x^*)^{32\xi + 2\hat{\xi}}. \]
We now just deal with \( T \) where Proposition 5.1 and Lemmata F.7 and F.8.

Then, using the upper bound of 1 for the indicator and applying Cauchy–Schwarz, we have

\[
\hat{\eta}_n^C(\mathbb{I}_A(v \otimes v)^{2\xi}) = \left(1 - \int_X \eta_n^l(x) \otimes \eta_n^{-l}(x) dx \right) \int_{X \times X} \mathbb{I}_A(x, y)v(x)^{2\xi}v(y)^{2\xi} \times \left(\frac{\eta_n^l(x) - \eta_n^l(y)}{1 - \int_X \eta_n^l(x) \otimes \eta_n^{-l}(x) dx} - \frac{\eta_n^{-l}(y) - \eta_n^{-l}(y)}{1 - \int_X \eta_n^{-l}(x) \otimes \eta_n^{-l}(x) dx}\right) \, dx \, dy.
\]

Then, using the upper bound of 1 for the indicator and applying Cauchy–Schwarz, we have

\[
\hat{\eta}_n^C(\mathbb{I}_A(v \otimes v)^{2\xi}) \leq T_1T_2,
\]

where

\[
T_1 := \left(\int_X v(x)^{4\xi}(\eta_n^l(x) - \eta_n^{-l}(x)) dx\right)^{1/2},
\]

\[
T_2 := \left(\int_X v(x)^{4\xi}(\eta_n^{-l}(x) - \eta_n^l(x)) dx\right)^{1/2}.
\]

We now just deal with \( T_1 \), as the proof for \( T_2 \) is almost identical.

Letting \( D_n^l = \{ x : \eta_n^l(x) \geq \eta_n^{-l}(x) \} \), \( \tilde{\varphi}(x) = v(x)^{4\xi}\|D_n^l\|_v \), and \( \tilde{\varphi}(x) = v(x)^{4\xi}\|D_n^l\|_v \), we have

\[
T_1^2 = \frac{1}{2} \left( \int_X v(x)^{4\xi} |\eta_n^l(x) - \eta_n^{-l}(x)| dx + \int_X v(x)^{4\xi} |\eta_n^l(x) - \eta_n^{-l}(x)| dx \right) \leq \frac{1}{2}(|\eta_n^l - \eta_n^{-l}|(v^{4\xi}) + |\tilde{\varphi}(\eta_n^l) - \tilde{\varphi}(\eta_n^{-l})| + |\tilde{\varphi}(\eta_n^l) - \tilde{\varphi}(\eta_n^{-l})|)
\]

As \( \xi \in (0, 1/32) \), \( (v^{4\xi}, \tilde{\varphi}, \tilde{\varphi}) \in \mathcal{L}_v^\kappa(X)^3 \), with \( 0 < \kappa < 1/8 \), so applying Proposition F.1 we have

\[
T_1 \leq C(\Delta)^{1/2}v(x^*)^{16\xi + \hat{\xi}}.
\]

Similar calculations give \( T_2 \leq C(\Delta)^{1/2}v(x^*)^{16\xi + \hat{\xi}} \); hence, noting (112), one can conclude. \( \square \)

**F.3. Proofs for Proposition 5.3**

Recall here that \( X = \mathbb{R} \) and the metric in (H7) is \( d_1 \), the \( L_1 \)–norm. Let \( \tilde{\varphi}(x, y) = (x - y)^2 \).

**Proof of Proposition 5.3.** Considering the second bound in Theorem 4.3, the result follows by Proposition 5.1 and Lemmata F.7 and F.8. \( \square \)

**Proposition F.5.** Assume (H1)–(H6), (H8). Then for any \((\xi, \hat{\xi}) \in (0, 1/8) \times (0, 1/2)\) there exists a \( C < +\infty \) depending on the constants in (H1)–(H6) and (H8) such that for any \( \varphi \in \mathcal{L}_v^\kappa(X) \), \( n \geq 0 \), \( 1 \leq l \leq L \) we have

\[
||\eta_n^l - \eta_n^{-l}(\varphi)|| \leq C||\varphi||_{v, \xi} \Delta_1v(x^*)^{8\xi + 2\hat{\xi}}.
\]
\textit{Proof.} Define
\begin{align}
T_1 & := \frac{[\eta_n^l - \eta_n^{l-1}](G_n \varphi)}{\eta_n^l(G_n)}, \\
T_2 & := \frac{\eta_n^{l-1}(G_n \varphi)}{\eta_n^{l-1}(G_n)} - \frac{[\eta_n^l - \eta_n^{l-1}](G_n)}{\eta_n^{l-1}(G_n)}.
\end{align}
(113)

Then we have
\begin{align}
T_1 - T_2 &= [\bar{\eta}_n^l - \bar{\eta}_n^{l-1}](\varphi).
\end{align}
(115)

We will bound $|T_1|$ and $|T_2|$ and then sum the bounds to conclude.

**Term (113):** By Proposition F.1 and $\eta_n^l(G_n) \geq C$ we have
\begin{align}
|T_1| & \leq C\|\varphi\|_{v, \xi} \Delta_l \nu(x^*)^{8\xi + 2\hat{\xi}}.
\end{align}
(116)

**Term (114):** By Proposition F.1, (H5), and $\eta_n^s(G_n) \geq C$, $s \in \{l, l - 1\}$, we have
\begin{align}
|T_2| & \leq C\|\varphi\|_{v, \xi} \Delta_l \nu(x^*)^{8\xi + 2\hat{\xi}}.
\end{align}
(117)

Noting (115), (116), and (117), the proof can be concluded. $\square$

Denote by $\bar{\eta}_n^C$ the maximal coupling of $(\bar{\eta}_n^l, \bar{\eta}_n^{l-1})$.

**Corollary F.1.** Assume (H1)–(H6), (H8). Then for any $(\xi, \hat{\xi}) \in (0, 1/16) \times (0, 1/2)$ there exists a $C < +\infty$ depending on the constants in (H1)–(H6) and (H8) such that for any $n \geq 0$, $1 \leq l \leq L$ we have
\begin{align}
\bar{\eta}_n^C(\mathbb{I}_A(v \otimes v)^{\hat{\xi}}) & \leq C\Delta_l \nu(x^*)^{16\xi + 2\hat{\xi}}.
\end{align}

**Proof.** This follows by the same proof as for Proposition F.4, except using Proposition F.5 instead of Proposition F.1 in the proof. $\square$

Denote by $\bar{\eta}_n^W$ the optimal Wasserstein coupling of $(\bar{\eta}_n^l, \bar{\eta}_n^{l-1})$; that is, for $B \in \mathcal{X} \times \mathcal{X}$,
\begin{align}
\bar{\eta}_n^W(B) &= \int_0^1 \mathbb{I}_B(F_{\bar{\eta}_n^l}^{-1}(u), F_{\bar{\eta}_n^{l-1}}^{-1}(u))du.
\end{align}

**Lemma A.7.** Assume (H1)–(H6), (H8). If $\tilde{\varphi} \in \mathcal{L}_{(v \otimes v)^{\hat{\xi}}}(X)$ for any $\tilde{\xi} \in (0, 1/16)$, and $\hat{\xi} \in (0, 1/2)$, then there exists a $C < +\infty$ depending on the constants in (H1)–(H6) and (H8) such that for any $\varphi \in B_b(X) \cap \text{Lip}_d(X)$, $n \geq 0$, $1 \leq l \leq L$ we have
\begin{align}
\bar{\eta}_n^W((\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}](\varphi)))^2) & \leq C\|\varphi\|_{L_\text{lip}}^2\tilde{\varphi}_{v, \xi}(\Delta_l \nu(x^*)^{16\xi + 2\hat{\xi}}).
\end{align}

**Proof.** We assume $n \geq 1$; the case $n = 0$ is noted below. As $\varphi \in B_b(X) \cap \text{Lip}_d(X)$, it easily follows that
\begin{align}
\bar{\eta}_n^W((\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}](\varphi)))^2) \leq \|\varphi\|_{\text{Lip}}^2\bar{\eta}_n^W(\tilde{\varphi}(\Delta_l + \bar{\eta}_n^{C}(\mathbb{I}_A(v \otimes v)^{\hat{\xi}}))).
\end{align}

Applying (6) (when $p = 2$) and using the optimality of the Wasserstein coupling with $\tilde{\varphi} \in \mathcal{L}_{(v \otimes v)^{\hat{\xi}}}(X)$ gives
\begin{align}
\bar{\eta}_n^W((\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}](\varphi)))^2) \leq C\|\varphi\|_{\text{Lip}}^2\tilde{\varphi}_{v, \xi}(\Delta_l + \bar{\eta}_n^{C}(\mathbb{I}_A(v \otimes v)^{\hat{\xi}}))).
\end{align}
CLTs for coupled particle filters

Application of Corollary F.1 yields the desired result. The case \( n = 0 \) follows from

\[
\tilde{\eta}^W_n \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}](\varphi)))^2 \right) \leq \|\varphi\|_{\text{Lip}}^2 \tilde{\eta}^W_n (\tilde{\varphi})
\]

and the optimality of the Wasserstein coupling, \( \tilde{\varphi} \in \mathcal{L}_{(v \otimes v)}(X) \) with Corollary F.1.

**Lemma F.8.** Assume (H1)–(H6), (H8). Then let \( \lambda \in (0, 1) \) be given; assume \( \tilde{\varphi} \in \mathcal{L}_{(v \otimes v)}(X) \), for any \( \tilde{\xi} \in (0, 1/(16(1 + \lambda))) \), and set

\[
(\xi, \tilde{\xi}) \in (0, \min(1/32, \lambda/(16(1 + \lambda)), (1 - 2\tilde{\xi})/12)) \times (0, 1/2).
\]

Then there exists a \( C < +\infty \) depending on the constants in (H1)–(H6) and (H8) such that for any \( n \geq 0 \), \( 1 \leq l \leq L \) we have

\[
\tilde{\eta}^W_n (\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi})) \leq C \|\tilde{\varphi}\|_{v^{\xi}} (\Delta_l)^{1/(1 + \lambda)} v(x^*)^{20\xi + (16(\lambda + 1)\xi + 2\tilde{\xi})/(1 + \lambda)}.
\]

**Proof.** We assume \( n \geq 1 \); the case \( n = 0 \) is similar and omitted for brevity. By Cauchy–Schwarz and the structure of \( \tilde{\eta}^W_n \),

\[
\tilde{\eta}^W_n (\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi})) \leq \tilde{\eta}^W_{n-1}(\tilde{\varphi}\tilde{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}).
\]

Applying (6) (when \( p = 4 \)) gives the upper bound

\[
\tilde{\eta}^W_n (\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi})) \leq C \left( \tilde{\eta}^W_{n-1}(\tilde{\varphi}\tilde{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) + \Delta_l \tilde{\eta}^W_{n-1}(\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi}))^{1/2} \right).
\]

To complete the proof, we focus on \( \tilde{\eta}^W_{n-1}(\tilde{\varphi}\tilde{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) \), as the other term can be controlled using the arguments below.

By Hölder’s inequality, we have

\[
\tilde{\eta}^W_{n-1}(\tilde{\varphi}\tilde{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) \leq \tilde{\eta}^W_{n-1}(\tilde{\varphi}(1 + \lambda)^{1/(1 + \lambda)}\tilde{\eta}^W_{n-1}(\tilde{\varphi}(v^{8\xi} \otimes v^{16\xi})^{1/(1 + \lambda)})(1 + \lambda)/(2\lambda)^{\lambda}/(1 + \lambda).
\]

By the optimality of the Wasserstein coupling with \( \tilde{\varphi} \in \mathcal{L}_{(v \otimes v)}(X) \), we have

\[
\tilde{\eta}^W_{n-1}(\tilde{\varphi}\tilde{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) \leq \left( \|\tilde{\varphi}\|_{v^{\xi}} \tilde{\eta}^C_{n-1}(v \otimes v)^{\xi}(1 + \lambda) \right)^{1/(1 + \lambda)}\tilde{\eta}^W_{n-1}(\tilde{\varphi}(v^{8\xi} \otimes v^{16\xi})^{1/(1 + \lambda)})(1 + \lambda)/(2\lambda)^{\lambda}/(1 + \lambda).
\]

Then applying Corollary F.1 gives

\[
\tilde{\eta}^W_{n-1}(\tilde{\varphi}\tilde{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) \leq C \|\tilde{\varphi}\|_{v^{\xi}} (\Delta_l)^{1/(1 + \lambda)} v(x^*)^{(16(\lambda + 1)\xi + 2\tilde{\xi})/(1 + \lambda)}\tilde{\eta}^W_{n-1}(\tilde{\varphi}(v^{8\xi} \otimes v^{16\xi})^{1/(1 + \lambda)})(1 + \lambda)/(2\lambda)^{\lambda}/(1 + \lambda).
\]

Now for the remaining term, applying Cauchy–Schwarz twice gives

\[
\tilde{\eta}^W_{n-1}(\tilde{\varphi}\tilde{M}(v^{8\xi} \otimes v^{16\xi})^{1/(1 + \lambda)})(1 + \lambda)/(2\lambda)^{\lambda}/(1 + \lambda) \leq \tilde{\eta}^l_{n-1}(\tilde{M}^{l}(v^{16\xi})^{1/(1 + \lambda)})(1 + \lambda)/(2\lambda)^{\lambda}/(1 + \lambda).
\]
Applying Jensen twice gives
\[ \tilde{\eta}_{n-1}^{W}(\tilde{M}(v^{8\xi} \otimes v^{8\xi}))^{1/(1+\lambda)}/(2\lambda) \leq \eta_{n-1}^{l}(v)^{4\xi} \eta_{n-1}^{l}(v)^{16\xi} \leq C\alpha^{20\xi}. \]

The proof is thus complete. \( \square \)

**Remark F.2.** As \( \lambda > 0 \) in Lemma F.8, one almost has the (time-uniform) forward error rate for the WCPF. We believe that \( \lambda = 0 \) is the desired case; however, because of technical difficulties we have not obtained this. One of the issues of the proof is that the Wasserstein coupling is not the coupling which minimizes the expectation of \( \tilde{\phi}(v^{4\xi} \otimes v^{8\xi}) \) with respect to any coupling of \((\eta_{n}^{l}, \eta_{n}^{l-1})\). To see this, one can construct a functional covariance equality for \( \mu(\tilde{\phi}(v^{4\xi} \otimes v^{8\xi})) \) (\( \mu \) is any coupling of \((\eta_{n}^{l}, \eta_{n}^{l-1}) \)) such that \( \mu(\tilde{\phi}(v^{4\xi} \otimes v^{8\xi})) < +\infty \) as in [24, Theorem 3.1] (that is, in terms of the CDFs of \( \mu_{n}^{l} \) and \( \eta_{n}^{l-1} \), as in Hoeffding’s lemma), and then when centering with \((\tilde{\eta}_{n}^{l} \otimes \tilde{\eta}_{n}^{l-1})(\tilde{\phi}(v^{4\xi} \otimes v^{8\xi})) \) and applying Hoeffding–Fréchet bounds, one observes that \( \tilde{\eta}_{n}^{W} \) is not optimal. As a result, one cannot transfer between \( \tilde{\eta}_{n}^{W} \) and \( \tilde{\eta}_{n}^{C} \) as is done in the proofs of Lemmata F.7–F.8, nor can one use Kantorovich duality. However, one can obtain \( \lambda = 0 \) if any of the following hold true:

1. There exists an \( \tilde{\eta}_{n}^{W} \)–integrable \( \hat{V} : X \times X \rightarrow (0, +\infty) \) such that there exists a \( C < +\infty \) such that for every \((x, y) \in X \times X\), \( \tilde{\phi}(x, y)v^{4\xi}(x)v^{8\xi}(y) \leq C\hat{V}(x, y) \), and \( \hat{V} \) satisfies the Monge condition (e.g. [27, Eq. (3.1.7)]).
2. \( v \) is bounded.
3. There exists a \( C < +\infty \) such that for every \( n \geq 0 \),
\[ \tilde{\eta}_{n}^{W}(\tilde{\phi}(v^{4\xi} \otimes v^{8\xi})) \leq C\tilde{\eta}_{n}^{C}(\tilde{\phi}(v^{4\xi} \otimes v^{8\xi})). \]
4. There exists a \( C < +\infty \) such that for every \( n \geq 0 \), \( 1 \leq l \leq L \),
\[ \sup_{u \in [0, 1]} |F_{n}^{-1}(u) - F_{n-1}^{-1}(u)| \leq C\Delta_{l}. \]

In general we do not believe 1 can hold in practice, and 2 is not useful in the context of the article. We believe 3 and 4 to hold up to some conditions, but have not obtained the proofs.

**F.4. Proof of Lemma 5.1**

**Proof of Lemma 5.1.** We begin by noting that for any \( B \in \mathcal{X} \), \( 0 \leq l \leq L \), \( y \in X \) we have
\[ M^{l}(\mathbb{I}_{B}v)(y) = \sqrt{\frac{(1 + \delta_{0})}{(1 + \delta_{0} - \beta_{l})}} \exp \left\{ \frac{\alpha_{l}y^2}{2(1 + \delta_{0} - \beta_{l})} + 1 \right\} \times \int_{B} \frac{1}{\sqrt{2\pi\beta_{l}}} \exp \left\{ - \frac{1 + \delta_{0} - \beta_{l}}{2\beta_{l}(1 + \delta_{0})} \left( x - \frac{\alpha_{l}(1 + \delta_{0})}{1 + \delta_{0} - \beta_{l}} \right)^{2} \right\} dx. \] (118)

This will be used below.

We give the proof for \( \xi = 1 \), as it is similar in other cases. We have for any \( 0 \leq l \leq L \), \( n \geq 1 \), \( \varphi \in \mathcal{L}_{v}(X) \cap \text{Lip}_{v, \alpha_{l}} \) that
\[ |Q_{n}^{l}(\varphi)(x) - Q_{n}^{l}(\varphi)(y)| \leq \|\varphi\|_{v}M^{l}(v)(x)\left| G_{n-1}(x) - G_{n-1}(y) \right| + \|G_{n-1}\|\|\varphi\|_{v} \int_{X} v(u)|M^{l}(x, u) - M^{l}(y, u)| du. \]
As noted in Section 5.6, \( \alpha_l^2(1 + \delta_l)/(1 + \delta_l - \beta_l) < 1 \) for any \( 0 \leq l \leq L, \delta_l > 0 \), so using (118) one can determine that \( M^t(v)(x) \in \mathcal{L}_1(X) \) for any \( \delta_l > 0 \), with \( \|M^t(v)\|_v \) independent of \( l \). In addition, for every \( y \in Y \), we have \( G_{n-1} \in \text{Lip}_{d_l}(X) \) with Lipschitz constant independent of \( n \), so we need only consider the right-hand term in the above displayed equation.

Set \( D_l(x, y) \coloneqq \{ u \in X : M^t(x, u) - M^t(y, u) \} \). We consider only

\[
\int_{D_l(x, y)} v(u) |M^t(x, u) - M^t(y, u)| du,
\]

as the calculations on \( D_l(x, y)^C \) are very similar. We suppose \( x > y \), as the proof with \( x < y \) is analogous, so we have by (118) and \( x > y \) that

\[
\int_{D_l(x, y)} v(u) |M^t(x, u) - M^t(y, u)| du = \sqrt{\left( \frac{1 + \delta_l}{1 + \delta_l - \beta_l} \right)}
\times \left\{ \exp \left\{ \frac{\alpha_l^2 x^2}{2(1 + \delta_l - \beta_l) + 1} \right\} \int_{-\infty}^{c_l(x, y)} \frac{1}{\sqrt{2\pi\beta_l}} \exp \left\{ - \frac{1 + \delta_l - \beta_l}{2\beta_l(1 + \delta_l)} \left( u - \frac{\alpha_l x(1 + \delta_l)}{1 + \delta_l - \beta_l} \right)^2 \right\} du \right. \\
- \exp \left\{ \frac{\alpha_l^2 y^2}{2(1 + \delta_l - \beta_l) + 1} \right\} \int_{-\infty}^{c_l(x, y)} \frac{1}{\sqrt{2\pi\beta_l}} \exp \left\{ - \frac{1 + \delta_l - \beta_l}{2\beta_l(1 + \delta_l)} \left( u - \frac{\alpha_l y(1 + \delta_l)}{1 + \delta_l - \beta_l} \right)^2 \right\} du \right\},
\]

where \( c_l(x, y) = (\alpha_l^2 x^2 - \alpha_l^2 y^2)/(x - y) \). Define

\[
\bar{c}_{l,1}(x, y) := \sqrt{\left( \frac{1 + \delta_l - \beta_l}{\beta_l(1 + \delta_l)} \right)} \left( c_l(x, y) - \frac{\alpha_l x(1 + \delta_l)}{1 + \delta_l - \beta_l} \right),
\]

\[
\bar{c}_{l,2}(x, y) := \sqrt{\left( \frac{1 + \delta_l - \beta_l}{\beta_l(1 + \delta_l)} \right)} \left( c_l(x, y) - \frac{\alpha_l y(1 + \delta_l)}{1 + \delta_l - \beta_l} \right)
\]

\[
\tilde{v}(x) := \exp \left\{ \frac{\alpha_l^2 x^2}{2(1 + \delta_l - \beta_l) + 1} \right\},
\]

and define \( \Theta(x) \) as the CDF of a standard normal distribution; then we have

\[
\int_{D_l(x, y)} v(u) |M^t(x, u) - M^t(y, u)| du \\
\frac{v(x)v(y)}{v(x)v(y)}
\]

\[
= \sqrt{\left( \frac{1 + \delta_l}{1 + \delta_l - \beta_l} \right)} \left\{ \frac{\tilde{v}(x)}{v(x)v(y)} \left( \Theta(\bar{c}_{l,1}(x, y)) - \Theta(\bar{c}_{l,2}(x, y)) \right) \right. \\
+ \Theta(\bar{c}_{l,2}(x, y)) \left[ \frac{\tilde{v}(x)}{v(x)} \left\{ \frac{1}{v(y)} - \frac{1}{v(x)} \right\} + \frac{1}{v(x)} \left\{ \tilde{v}(x) - \tilde{v}(y) \right\} \right] \right\}.
\]

From here, it is straightforward to establish that all the functions in the differences are in the set \( \text{Lip}_{d_l}(X) \) with Lipschitz constants that do not depend on \( l \), and, moreover, that the other terms are uniformly bounded in \( x, y, l \) (where relevant)—the proofs are omitted as they are standard. \( \square \)
Acknowledgement

Both authors were supported by KAUST baseline funding. A. Jasra was supported under the KAUST Competitive Research Grants Program Round 4 (CRG4) project Advanced Multi-Level Sampling Techniques for Bayesian Inverse Problems with Applications to Subsurface, ref. 2584. We would like to thank Alexandros Beskos, Sumeetpal Singh, and Xin Tong for useful conversations associated to this work. We thank two referees, the associate editor, and the editor-in-chief for substantial comments which have led to improvements of the article.

References

[1] Bally, V. and Talay, D. (1996). The law of the Euler scheme for stochastic differential equations II: Approximation of the density. Monte Carlo Meth. Appl. 2, 93–128.
[2] Beskos, A. et al. (2017). Multilevel sequential Monte Carlo samplers. Stoch. Process. Appl. 127, 1417–1440.
[3] Billingsley, P. (1995). Probability and Measure. Wiley, New York.
[4] Cappe, O., Moulines, E. and Ryden, T. (2005). Inference in Hidden Markov Models. Springer, New York.
[5] Chopin, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. Ann. Statist. 32, 2385–2411.
[6] Chopin, N. and Singh, S. S. (2015). On particle Gibbs sampling. Bernoulli 21, 1855–1883.
[7] Crisan, D. and Bain, A. (2008). Fundamentals of Stochastic Filtering. Springer, New York.
[8] Del Moral, P. (2004). Feynman–Kac Formulae: Genealogical and Interacting Particle Systems with Applications. Springer, New York.
[9] Del Moral, P. and Miclo, L. (2000). Branching and interacting particle systems approximation of Feynman–Kac formulae with applications to non-linear filtering. In Séminaire de Probabilités XXXIV (Lecture Notes in Mathematics 1729), eds. J. Azéma, M. Émery, M. Ledoux, and M. Yor, Springer, Berlin, pp. 1–145.
[10] Del Moral, P., Doucet, A. and Jasra, A. (2012). On adaptive resampling procedures for sequential Monte Carlo methods. Bernoulli 18, 252–272.
[11] Del Moral, P., Jacod, J. and Protter, P. (2001). The Monte Carlo method for filtering with discrete-time observations. Prob. Theory Relat. Fields 120, 346–368.
[12] Del Moral, P., Jasra, A. and Law, K. J. H. (2017). Multilevel sequential Monte Carlo: Mean square error bounds under verifiable conditions. Stoch. Anal. 35, 478–498.
[13] Giles, M. B. (2008). Multilevel Monte Carlo path simulation. Operat. Res. 56, 607–617.
[14] Gregory, A., Cotter, C. and Reich, S. (2016). Multilevel ensemble transform particle filtering. SIAM J. Sci. Comput. 38, A1317–A1338.
[15] Heinrich, S. (2001). Multilevel Monte Carlo methods. In Large-Scale Scientific Computing, eds. S. Margenov, J. Wasniewski, and P. Yalamov, Springer, Berlin, pp. 58–67.
[16] Jacob, P., Lindsten, F. and Schon, T. (2016). Coupling of particle filters. Preprint. Available at https://arxiv.org/abs/1606.01156.
[17] Jacob, P., Lindsten, F. and Schon, T. (2019). Smoothing with couplings of conditional particle filters. To appear in J. Amer. Statist. Assoc.
[18] Jacob, P., O’Leary, J. and Atachde, Y. (2020). Unbiased Markov chain Monte Carlo with couplings. To appear in J. R. Statist. Soc. B.
[19] Jasra, A. (2015). On the behaviour of the backward interpretation of Feynman–Kac formulae under verifiable conditions. J. Appl. Prob. 52, 339–359.
[20] Jasra, A., Ballesio, M., von Schwerin, E. and Tempone, R. (2020). A Wasserstein coupled particle filter for multilevel estimation. Preprint. Available at https://arxiv.org/abs/2004.0398.
[21] Jasra, A., Kamatanzi, K., Osei, P. P. and Zhou, Y. (2018). Multilevel particle filters: Normalizing constant estimation. Statist. Comput. 28, 47–60.
[22] Jasra, A., Kamatanzi, K., Law, K. H. and Zhou, Y. (2017). Multilevel particle filters. SIAM J. Numer. Anal. 55, 3068–3096.
[23] Lee, A., Singh, S. S. and Vihola, M. (2019). Coupled conditional backward sampling particle filter. To appear in Ann. Statist.
[24] Lo, A. (2017). Functional generalizations of Hoeffding’s covariance lemma and a formula for Kendall’s tau. Statist. Prob. Lett. 122, 218–226.
[25] Mao, X. (2007). Stochastic Differential Equations and Applications, 2nd edn. Woodhead, Cambridge.
[26] McDiarmid, C. (1989). On the method of bounded differences. In Surveys in Combinatorics (London Math. Soc. Lecture Notes 141), ed. J. Siemons, Cambridge University Press, pp. 148–188.
[27] Rachev, T. and Rüschendorf, S. (1998). Mass Transportation Problems, Volume 1: Theory. Springer, New York.
[28] Sen, D., Thiery, A. and Jasra, A. (2018). On coupling particle filters. Statist. Comput. 28, 461–475.
[29] Thorisson, H. (2000). Coupling, Stationarity, and Regeneration. Springer, New York.
[30] Whiteley, N. P. (2013). Stability properties of some particle filters. Ann. Appl. Prob. 23, 2500–2537.
[31] Whiteley, N. P., Kantas, N. and Jasra, A. (2012). Linear variance bounds for particle approximations of time homogenous Feynman–Kac formulae. Stoch. Process. Appl. 122, 1840–1865.