

# Central Limit Theorems for Coupled Particle Filters

BY AJAY JASRA & FANGYUAN YU

Department of Statistics & Applied Probability, National University of Singapore, Singapore, 117546, SG.

E-Mail: *staja@nus.edu.sg*, *e0204835@u.nus.edu*

## Abstract

In this article we prove new central limit theorems (CLT) for several coupled particle filters (CPFs). CPFs are used for the sequential estimation of the difference of expectations w.r.t. 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 discretization  $\Delta_t$  we show that the MCPF and the approach in [19] have, under assumptions, an asymptotic variance that is upper-bounded by an expression that is (almost)  $\mathcal{O}(\Delta_t)$ , uniformly in time. The  $\mathcal{O}(\Delta_t)$  rate preserves the so-called forward rate of the diffusion in some scenarios which is not the case for the CPF in [21].

**Key words:** Coupled Particle Filter, Central Limit Theorem, Multilevel Monte Carlo.

## 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, 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 (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 CLTs which allow one to understand the errors associated to estimation. Under 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, if one observes data in 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 non-negative 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. 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, 27].

The CPF developed in [6] (see also [16, 17, 21, 22, 27]) 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 use of correlating the PFs is vital, for instance in ML applications (see [13, 15, 21]), as it is this property which allows a variance reduction relative to a single PF. As has been noted by several authors, e.g. [27], unless the coupling of the Markov transitions is particularly strong, one expects that 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 is the case where the resampling is independent for each pair (independent resampling CPF (IRCPF)), the second with the maximally coupled resampling (the maximally coupled resampling PF (MCRPF)). The third algorithm, which to our knowledge is new, is based upon a weak law of large numbers for the MCRPF on the product space. This result shows that the limiting coupling on product space does not correspond to the maximal coupling of the filter (or predictor). This coupling does not seem to have any optimality properties, so we suggest a new CPF, the MCPF which approximates the maximal coupling of the predictors. This algorithm requires that the Markov transition of the filter to have a density which is known pointwise and essentially samples from the maximal coupling of particle filters. The fourth algorithm in [19] 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, 27], but they are not considered here as they require even more mathematical complexity for their analysis.

We prove a CLT for the first three algorithms (the CLT for the WCPF, when the state-dynamics are one dimensional is in [19]) associated to the difference of estimates of expectations of the same function w.r.t. 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 assumptions (of the type in [29]), for the PODDP with (Euler) discretization  $\Delta_t$ , we show that the MCPF (resp. WCPF) has an asymptotic variance that is upper-bounded by an expression that is  $\mathcal{O}(\Delta_t)$  (resp.  $\mathcal{O}(\Delta_t^{1-\lambda})$ ,  $\lambda$  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 are 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 MCRPF we show even in a favourable case, that it can have an asymptotic variance, at time  $n$ , that is  $\mathcal{O}(e^n \Delta_t)$  and identify when one can expect the algorithm to work well. As was seen in the empirical results of [19, 21] the time and rate behaviour of the MCRPF 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 non-standard proofs, especially in the case of the MCRPF. To summarize, the main results of the article are:

- Theorem 3.1 gives a WLLN for the MCRPF.
- Theorem 4.1 gives a CLT for the IRCPF, MCRPF & MCPF.
- Theorem 4.3 gives a general bound on the asymptotic variance for each of the methods, IRCPF, MCRPF, MCPF & 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 notations, models and the motivating example of a PODDP with MLMC is given. 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 Notations

Let  $(\mathsf{X}, \mathcal{X})$  be a measurable space. For a given function  $v : \mathsf{X} \rightarrow [1, \infty)$  we denote by  $\mathcal{L}_v(\mathsf{X})$  the class of functions  $\varphi : \mathsf{X} \rightarrow \mathbb{R}$  for which

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

When  $v \equiv 1$  we write  $\|\varphi\| := \sup_{x \in \mathsf{X}} |\varphi(x)|$ . If  $v = 1$  we write  $\mathcal{B}_b(\mathsf{X})$ ,  $\mathcal{C}_b(\mathsf{X})$  for the bounded measurable and continuous, bounded measurable functions respectively.  $\mathcal{C}^2(\mathsf{X})$  are the collection of twice continuously differentiable real-valued functions on  $\mathsf{X}$ . Let  $d$  be a metric on  $\mathsf{X}$  then for  $\varphi \in \mathcal{L}_v(\mathsf{X})$ , we say that  $\varphi \in \text{Lip}_{v,d}(\mathsf{X})$  if there exist a  $C < +\infty$  such that for every  $(x, y) \in \mathsf{X} \times \mathsf{X}$ :

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

If  $v = 1$ , we write  $\varphi \in \text{Lip}_d(\mathsf{X})$  and write  $\|\varphi\|_{\text{Lip}}$  for the Lipschitz constant.  $\mathcal{P}(\mathsf{X})$  denotes the collection of probability measures on  $(\mathsf{X}, \mathcal{X})$ . We also denote, for  $\mu \in \mathcal{P}(\mathsf{X})$ ,  $\|\mu\|_v := \sup_{|\varphi| \leq v} |\mu(\varphi)|$ . For a measure  $\mu$  on  $(\mathsf{X}, \mathcal{X})$  and a  $\mu$ -integrable,  $\varphi : \mathsf{X} \rightarrow \mathbb{R}$ , the notation  $\mu(\varphi) = \int_{\mathsf{X}} \varphi(x) \mu(dx)$  is used. Let  $K : \mathsf{X} \times \mathcal{X} \rightarrow (0, \infty)$  be a non-negative kernel and  $\mu$  be a measure then we use the notations  $\mu K(dy) = \int_{\mathsf{X}} \mu(dx) K(x, dy)$  and for  $K(x, \cdot)$ -integrable,  $\varphi : \mathsf{X} \rightarrow \mathbb{R}$   $K(\varphi)(x) = \int_{\mathsf{X}} \varphi(y) K(x, dy)$ . For  $\mu, \nu \in \mathcal{P}(\mathsf{X})$ , the total variation distance is denoted  $\|\mu - \nu\|_{\text{tv}} = \sup_{B \in \mathcal{X}} |\mu(B) - \nu(B)|$ . For  $B \in \mathcal{X}$  the indicator is written  $\mathbb{1}_B(x)$  and the dirac measure  $\delta_B(dx)$ . For two measures  $\mu, \nu$  on  $(\mathsf{X}, \mathcal{X})$ , the product measure is  $\mu \otimes \nu$ . For two measurable functions  $\varphi, \psi$  on  $(\mathsf{X}, \mathcal{X})$ , the tensor product of functions is  $\varphi \otimes \psi$ .  $\mathcal{U}_A$  denotes the uniform distribution on the set  $A$ .  $\mathcal{N}_t(a, b)$  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 w.r.t. the law of the specified algorithm - the context will be clear in each instant.  $\Rightarrow$  and  $\rightarrow_{\mathbb{P}}$  are used to denote convergence in distribution and probability respectively. In the context of the article, this is as  $N \rightarrow \infty$ .

## 2.2 Models

Let  $(\mathbf{X}, \mathcal{X})$  be a measurable space and  $\{G_n\}_{n \geq 0}$  a sequence of non-negative, bounded and measurable functions such that  $G_n : \mathbf{X} \rightarrow \mathbb{R}_+$ . Let  $\eta_0^f, \eta_0^c \in \mathcal{P}(\mathbf{X})$  and  $\{M_n^f\}_{n \geq 1}, \{M_n^c\}_{n \geq 1}$  be two sequences of Markov kernels, i.e.  $M_n^f : \mathbf{X} \rightarrow \mathcal{P}(E), M_n^c : \mathbf{X} \rightarrow \mathcal{P}(E)$ . Define, for  $s \in \{f, c\}, B \in \mathcal{X}$

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

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 quantities, for  $\varphi \in \mathcal{B}_b(\mathbf{X})$  any  $n \geq 0$ , such as

$$\eta_n^f(\varphi) - \eta_n^c(\varphi) \quad (1)$$

or

$$\frac{\eta_n^f(G_n \varphi)}{\eta_n^f(G_n)} - \frac{\eta_n^c(G_n \varphi)}{\eta_n^c(G_n)}. \quad (2)$$

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

**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  $\check{\eta}_n \in \mathcal{P}(\mathbf{X} \times \mathbf{X})$  i.e. that for any  $B \in \mathcal{X}$  and every  $n \geq 0$

$$\check{\eta}_n(B \times \mathbf{X}) = \eta_n^f(B) \quad \check{\eta}_n(\mathbf{X} \times B) = \eta_n^c(B)$$

and consider approximating

$$\check{\eta}_n(\varphi \otimes 1) - \check{\eta}_n(1 \otimes \varphi)$$

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  $\check{\eta}_0 \in \mathcal{P}(\mathbf{X} \times \mathbf{X})$  is such that for any  $B \in \mathcal{X}$

$$\check{\eta}_0(B \times \mathbf{X}) = \eta_0^f(B) \quad \check{\eta}_0(\mathbf{X} \times B) = \eta_0^c(B)$$

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

$$\check{M}_n(B \times \mathbf{X})(x, x') = M_n^f(B)(x) \quad \check{M}_n(\mathbf{X} \times B)(x, x') = M_n^c(B)(x'). \quad (3)$$

## 2.3 Example

The following example is from [21] and there is some overlap with the presentation in that article. We start with a diffusion process:

$$dZ_t = a(Z_t)dt + b(Z_t)dW_t \quad (4)$$

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

The coefficients  $a^j, b^{j,k} \in \mathcal{C}^2(\mathbf{X})$ , for  $j, k = 1, \dots, d$ . Also,  $a$  and  $b$  satisfy

- (i) **uniform ellipticity:**  $b(z)b(z)^T$  is uniformly positive definite;
- (ii) **globally Lipschitz:** there is a  $C > 0$  such that  $|a^j(z) - a^j(y)| + |b^{j,k}(z) - b^{j,k}(y)| \leq C|z - y|$  for all  $(z, y) \in \mathbb{X} \times \mathbb{X}$ ,  $(j, k) \in \{1, \dots, d\}^2$ .

The data are observed at regular unit time-intervals (i.e. in discrete time)  $y_1, y_2, \dots, y_k \in \mathbb{Y}$ . 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  $X_0, X_1, \dots$  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  $B \in \mathcal{X}$

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

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

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  $M^l(x, dy)$ . 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  $M^l$  in most cases. Hence we are interested in the Feynman-Kac model for  $B \in \mathcal{X}$

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

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  $l \geq 1$ ,  $\eta_n^l(\varphi) - \eta_n^{l-1}(\varphi)$ ,  $\varphi \in \mathcal{B}_b(\mathbb{X})$ . That is,  $f$  as used above, relates to the predictor associated to the discretization  $\Delta_l$  and  $c$ , the predictor with discretization  $\Delta_{l-1}$ . A natural coupling of  $M_n^f = M^l$  and  $M_n^c = M^{l-1}$  exists (e.g. [13]) so one also has a given  $\check{\eta}_0$  and  $\check{M}_n$ .

Before continuing, we note some results which will help in our discussion below. As established in [21, eq. (32)] one has for  $C < +\infty$

$$\sup_{\mathcal{A}} \sup_{x \in \mathbb{X}} |M_n^f(\varphi)(x) - M_n^c(\varphi)(x)| \leq C\Delta_l \quad (5)$$

where  $\mathcal{A} = \{\varphi \in \mathcal{B}_b(\mathbb{X}) \cap \text{Lip}_{d_1}(\mathbb{X}) : \|\varphi\| \leq 1\}$ ,  $d_1$  is the  $L_1$ -norm. In addition, [21, Proposition D.1.] states that for  $p > 0$ ,  $C < +\infty$  and any  $(x, y) \in \mathbb{X} \times \mathbb{X}$

$$\int_{\mathbb{X} \times \mathbb{X}} \|u - v\|^p \check{M}_n((x, y), d(u, v)) \leq C(|x - y| + \Delta_l^{1/2})^p. \quad (6)$$

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 [21, 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

$$|\eta_n(\varphi) - \eta_n^l(\varphi)| \leq C\Delta_l. \quad (7)$$

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  $\eta_n^l(\varphi)$ , set  $\varphi \in \mathcal{A}$ , is then of course  $\frac{1}{N} \sum_{i=1}^N \varphi(x_n^i)$ ,  $X_n^i$  are i.i.d. from  $\eta_n^l$ . One has the mean square error (MSE)

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

where  $\text{Var}_{\eta_n^l}[\varphi(X)]$  is the variance of  $\varphi(X)$  w.r.t.  $\eta_n^l$ . Then, for  $\varepsilon > 0$  given, to target a mean square error 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 [21] for a justification), again ignoring  $n$ . Then the cost of achieving a MSE of  $\mathcal{O}(\varepsilon^2)$  is hence  $\mathcal{O}(\varepsilon^{-3})$ .

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

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

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

$$\int_{\mathbb{X} \times \mathbb{X}} \|x - y\|^2 \check{\eta}_n^l(dx, dy) \leq C \Delta_l$$

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  $((X_n^{1,l}, X_n^{1,l-1}), \dots, (X_n^{N_l,l}, X_n^{N_l,l-1}))$  i.i.d. from  $\check{\eta}_n^l$  for some  $N_l \geq 1$  and the approximation is  $\frac{1}{N_l} \sum_{i=1}^{N_l} \{\varphi(x_n^{i,l}) - \varphi(x_n^{i,l-1})\}$ . For  $1 \leq l \leq L$  this is repeated independently for estimating  $[\eta_n^l - \eta_n^{l-1}](\varphi)$  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  $\eta_n^L$  is upper-bounded by (recall  $\varphi \in \mathcal{A}$ ) by

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

Then, for  $\varepsilon > 0$  given, to target a mean square error of  $\mathcal{O}(\varepsilon^2)$ , set  $L = \mathcal{O}(|\log(\varepsilon)|)$ , and  $N_l = \mathcal{O}(\varepsilon^{-2} |\log(\varepsilon)| \Delta_l)$  (see [13]), then the MSE is  $\mathcal{O}(\varepsilon^2)$ . The overall computational effort is  $\mathcal{O}(\sum_{l=0}^L N_l \Delta_l^{-1}) = \mathcal{O}(\varepsilon^{-2} |\log(\varepsilon)|^2)$ ; if  $\varepsilon$  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, developed in the literature. The estimates (variance/cost) above typically depend on  $n$  and our objective is to consider if the variance of PF estimates can be  $\mathcal{O}(\Delta_l)$  uniformly in time and hence 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 can be more straightforward to deal with.

## 3 Algorithms

### 3.1 Independent Pair Resampling

The first procedure we consider is as follows. Let  $n \geq 1$ ,  $B \in \mathcal{X} \vee \mathcal{X}$  and  $\mu \in \mathcal{P}(\mathbb{X} \times \mathbb{X})$  and define the probability measure:

$$\check{\Phi}_n^I(\mu)(B) = \frac{\mu([G_{n-1} \otimes G_{n-1}] \check{M}_n(B))}{\mu(G_{n-1} \otimes G_{n-1})}.$$

Set  $u_n = (x_n^f, x_n^c) \in \mathbb{X} \times \mathbb{X}$  then consider the joint probability measure on  $(\mathbb{X} \times \mathbb{X})^{n+1}$

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

Noting (3), 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  $\check{\eta}_n^I$  the marginal of  $u_n$  induced by this joint probability measure. Thus, if one could sample a trajectory of  $(u_0, \dots, u_n)$  from  $\mathbb{P}(d(u_0, \dots, u_n))$  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 is taken as:

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

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

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

The key point is that in this algorithm, the resampled indices for a pair of particles are generated (conditionally) independently. Set

$$\check{\eta}_p^{N,I}(du) = \frac{1}{N} \sum_{i=1}^N \delta_{u_p^i}(du).$$

As has been mentioned by many authors (e.g. [27]) one does not expect this procedure to be effective, in the sense that  $\check{\eta}_n^I$  would not provide an appropriate dependence between  $(\eta_n^f, \eta_n^c)$ .

### 3.2 Maximally Coupled Resampling

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

$$\begin{aligned} \check{\Phi}_n^M(\mu)(B) &= \mu\left(\{F_{n-1,\mu,f} \wedge F_{n-1,\mu,c}\} \check{M}_n(B)\right) + \left(1 - \mu\left(\{F_{n-1,\mu,f} \wedge F_{n-1,\mu,c}\}\right)\right) \times \\ &\quad (\mu \otimes \mu)\left(\{\bar{F}_{n-1,\mu,f} \otimes \bar{F}_{n-1,\mu,c}\} \bar{M}_n(B)\right) \end{aligned}$$

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

$$\begin{aligned} \bar{F}_{n-1,\mu,f}(x, y) &= \frac{F_{n-1,\mu,f}(x, y) - \{F_{n-1,\mu,f}(x, y) \wedge F_{n-1,\mu,c}(x, y)\}}{\mu(F_{n-1,\mu,f} - \{F_{n-1,\mu,f} \wedge F_{n-1,\mu,c}\})} \\ \bar{F}_{n-1,\mu,c}(x, y) &= \frac{F_{n-1,\mu,c}(x, y) - \{F_{n-1,\mu,f}(x, y) \wedge F_{n-1,\mu,c}(x, y)\}}{\mu(F_{n-1,\mu,c} - \{F_{n-1,\mu,f} \wedge F_{n-1,\mu,c}\})} \\ F_{n-1,\mu,f}(x, y) &= \check{G}_{n-1,\mu,f}(x) \otimes 1 \\ F_{n-1,\mu,c}(x, y) &= 1 \otimes \check{G}_{n-1,\mu,c}(y) \\ \check{G}_{n-1,\mu,f}(x) &= \frac{G_{n-1}(x)}{\mu(G_{n-1} \otimes 1)} \\ \check{G}_{n-1,\mu,c}(y) &= \frac{G_{n-1}(y)}{\mu(1 \otimes G_{n-1})} \end{aligned}$$

and for  $((x, y), (u, v)) \in \mathbb{X}^2 \times \mathbb{X}^2$  and  $B \in \mathcal{X} \vee \mathcal{X}$

$$\bar{M}_n(B)((x, y), (u, v)) = \check{M}_n(B)(x, v).$$

Now set  $\check{\eta}_0^M = \check{\eta}_0$  and we define, recursively, for any  $n \geq 1$ ,  $B \in \mathcal{X} \vee \mathcal{X}$

$$\check{\eta}_n^M(B) = \check{\Phi}_n^M(\check{\eta}_{n-1}^M)(B).$$

Note that by [21, Proposition A.1.] we have for  $B \in \mathcal{X}$

$$\check{\eta}_n^M(B \times \mathbb{X}) = \eta_n^f(B) \quad \text{and} \quad \check{\eta}_n^M(\mathbb{X} \times B) = \eta_n^c(B).$$

The algorithm used here is:

$$\mathbb{P}(d(u_0^{1:N}, \dots, u_n^{1:N})) = \left(\prod_{i=1}^N \check{\eta}_0(du_0^i)\right) \left(\prod_{p=1}^n \prod_{i=1}^N \check{\Phi}_p^M(\check{\eta}_{p-1}^{N,M})(du_p^i)\right)$$

where for  $p \geq 1$ ,

$$\check{\eta}_{p-1}^{N,M}(du) = \frac{1}{N} \sum_{i=1}^N \delta_{u_p^i}(du).$$

and we set for  $s \in \{f, c\}$

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

This procedure is as in [6] and adopted in, for instance, [17, 21]. The idea is to provide a local optimality procedure w.r.t. 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 MCRPF, we present a preliminary result, which will prove to be of interest. The following assumptions are used and note that in (A3) there is a metric  $d$  on  $\mathbb{X} \times \mathbb{X}$  implicit in the assumption.

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

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

(A3)  $\mathsf{X}$  is a locally compact and separable metric space.

The assumptions are adopted due to the complexity of the operator  $\check{\Phi}_n^M$ . As can be seen in Appendix C, where the proofs for the following result are given, it is non-trivial to work with  $\check{\Phi}_n^M$ . To weaken these assumptions would lead to further calculations, which would essentially confirm the same result.

**Theorem 3.1.** *Assume (A1-3). Then for any  $\varphi \in \mathcal{C}_b(\mathsf{X} \times \mathsf{X})$ ,  $n \geq 0$  we have*

$$\check{\eta}_n^{N,M}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^M(\varphi).$$

What is interesting here, is that on the *product space*  $\mathsf{X} \times \mathsf{X}$  the target  $\check{\eta}_n^M$  is a coupling of  $(\eta_n^f, \eta_n^c)$ , but the actual maximal coupling of  $(\eta_n^f, \eta_n^c)$  is not sampled from. 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 w.r.t. 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 [21] 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  $(\eta_n^f, \eta_n^c)$ , 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  $(\eta_n^f, \eta_n^c)$ . We will assume that for  $s \in \{f, c\}$  the Markov kernels  $M_n^s$ , as well as  $\eta_0^s$ , admit a density w.r.t. a  $\sigma$ -finite measure  $dx$ . The densities are denoted  $M_n^s$  and  $\eta_0^s$ ,  $s \in \{f, c\}$  and we assume that the densities can be evaluated numerically. To remove this latter requirement is left to future work.

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

$$\check{\Phi}_n^C(\mu, \nu)(B) = \frac{\int_{\mathsf{X} \times \mathsf{X}} \mathbb{I}_B(x, y) \left[ \int_{\mathsf{X}} F_{n-1, \mu, f}(u) \wedge F_{n-1, \nu, c}(u) \delta_{\{u, u\}}(d(x, y)) du + \right]}{1 - \int_{\mathsf{X}} F_{n-1, \mu, f}(u) \wedge F_{n-1, \nu, c}(u) du} \bar{F}_{n-1, \mu, \nu, f}(x) \bar{F}_{n-1, \nu, \mu, c}(y) dx dy$$

where, for  $x \in \mathsf{X}$

$$\begin{aligned} \bar{F}_{n-1, \mu, \nu, f}(x) &= F_{n-1, \mu, f}(x) - F_{n-1, \mu, f}(x) \wedge F_{n-1, \nu, c}(x) \\ \bar{F}_{n-1, \nu, \mu, c}(x) &= F_{n-1, \nu, c}(x) - F_{n-1, \mu, f}(x) \wedge F_{n-1, \nu, c}(x) \\ F_{n-1, \mu, f}(x) &= \frac{\mu(G_{n-1} M_n^f(\cdot, x))}{\mu(G_{n-1})} \\ F_{n-1, \nu, c}(x) &= \frac{\nu(G_{n-1} M_n^c(\cdot, x))}{\nu(G_{n-1})}. \end{aligned}$$

Now set  $\check{\eta}_0^C$  as the maximal coupling of  $(\eta_0^f, \eta_0^c)$  and set for  $B \in \mathcal{X} \vee \mathcal{X}$

$$\check{\eta}_n^C(B) = \check{\Phi}_n^C(\eta_{n-1}^f, \eta_{n-1}^c)(B).$$

We have for  $B \in \mathcal{X}$

$$\check{\eta}_n^C(B \times \mathsf{X}) = \eta_n^f(B) \quad \text{and} \quad \check{\eta}_n^C(\mathsf{X} \times B) = \eta_n^c(B).$$

Moreover  $\check{\eta}_n^C$  is the maximal coupling of  $(\eta_n^f, \eta_n^c)$ .

The particle approximation used is:

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

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

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

We set for  $\mu \in \mathcal{P}(\mathbf{X})$ ,  $s \in \{f, c\}$ ,  $B \in \mathcal{X}$ ,  $n \geq 1$

$$\Phi_n^s(\mu)(B) = \frac{\mu(G_{n-1} M_n^s(B))}{\mu(G_{n-1})}.$$

We remark that as  $M_n^s$ ,  $n \geq 1$  and  $\eta_0^s$ ,  $s \in \{f, c\}$ , can be evaluated numerically we can sample from  $\check{\eta}_0^C$  and  $\check{\Phi}_n^C(\eta_{n-1}^{N,f}, \eta_{n-1}^{N,c})$ , the latter of which is the maximal coupling of  $\Phi_n^f(\eta_{n-1}^{N,f})$  and  $\Phi_n^c(\eta_{n-1}^{N,c})$ , using the algorithm in [28]. This is as follows:

1. Sample  $X_n^f \sim \Phi_n^f(\eta_{n-1}^{N,f})$  and  $W|x_n^f \sim \mathcal{U}_{[0, F_{n-1, \eta_{n-1}^{N,f}, f}(x_n^f)]}$ . If  $W < F_{n-1, \eta_{n-1}^{N,c}, c}(x_n^f)$ , output  $(x_n^f, x_n^f)$ , otherwise go-to 2..
2. Sample  $X_n^c \sim \Phi_n^c(\eta_{n-1}^{N,c})$  and  $\tilde{W}|x_n^c \sim \mathcal{U}_{[0, F_{n-1, \eta_{n-1}^{N,c}, c}(x_n^c)]}$ , if  $\tilde{W} > F_{n-1, \eta_{n-1}^{N,f}, f}(x_n^c)$ , output  $(x_n^f, x_n^c)$ , otherwise start 2. again.

This algorithm is a rejection sampler, which would add a random running time element per time-step, which may not be desirable for some applications. In the limiting case (i.e. sampling the maximal coupling of  $(\eta_n^f, \eta_n^c)$ ) the expected number of steps to acceptance is

$$1 + \frac{\int_{\mathbf{X}} |\eta_n^c(x) - \eta_n^f(x)| dx}{2 \int_{\mathbf{X}} [\eta_n^c(x) - \eta_n^f(x)] \mathbb{I}_{\{x \in \mathbf{X}: \eta_n^c(x) \geq \eta_n^f(x)\}}(x) dx} = 1 + R_n.$$

One might expect that  $R_n$  above is  $\mathcal{O}(1)$  in applications. We expect  $(\eta_n^f, \eta_n^c)$  to be close and possibly that  $\eta_n^c$  is more diffuse than  $\eta_n^f$ , for instance that  $\eta_n^c(x)/\eta_n^f(x) \geq C$ , for some constant  $C \leq 1$ , hence one may find  $R_n = \mathcal{O}(1)$ . In general, one might want to confirm this conjecture before implementing this approach.

### 3.4 Wasserstein Coupled Resampling

We describe the WCPF used in [19]. For this case we restrict our attention to the case that  $\mathbf{X} = \mathbb{R}$ . It is explicitly assumed that the cumulative distribution function (CDF) and its inverse associated to the probability,  $s \in \{f, c\}$ ,  $n \geq 0$ ,  $B \in \mathcal{X}$

$$\bar{\eta}_n^s(B) = \frac{\eta_n^s(G_n \mathbb{I}_B)}{\eta_n^s(G_n)}$$

exists and are continuous functions. We denote the CDF (resp. inverse) of  $\bar{\eta}_n^s$  as  $F_{\bar{\eta}_n^s}$  (resp.  $F_{\bar{\eta}_n^s}^{-1}$ ). In general we write probability measures on  $\mathbf{X}$  for which the CDF and inverse are well-defined as  $\mathcal{P}_F(\mathbf{X})$  with the associated CDF  $F_\mu$ . Let  $n \geq 1$ ,  $B \in \mathcal{X} \vee \mathcal{X}$  and  $\mu, \nu \in \mathcal{P}_F(\mathbf{X})$  and define the probability measure:

$$\check{\Phi}_n^W(\mu, \nu)(B) = \int_{\mathbf{X} \times \mathbf{X}} \left( \int_0^1 \delta_{\{F_\mu^{-1}(w), F_\nu^{-1}(w)\}}(du) dw \right) \check{M}_n(B)(u).$$

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

$$\mathbb{P}(d(u_0, \dots, u_n)) = \check{\eta}_0(du_0) \prod_{p=1}^n \check{\Phi}_p^W(\bar{\eta}_{p-1}^f, \bar{\eta}_{p-1}^c)(du_p).$$

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  $\check{\eta}_n^W$  the marginal of  $u_n$  induced by this joint probability measure.

The algorithm used here is:

$$\mathbb{P}(d(u_0^{1:N}, \dots, u_n^{1:N})) = \left( \prod_{i=1}^N \check{\eta}_0(du_0^i) \right) \left( \prod_{p=1}^n \prod_{i=1}^N \check{\Phi}_p^W(\bar{\eta}_{p-1}^{N,f}, \bar{\eta}_{p-1}^{N,c})(du_p^i) \right)$$



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

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

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

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

and for  $p \geq 0$

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

## 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

Denote the sequence of non-negative kernels  $\{Q_n^s\}_{n \geq 1}$ ,  $s \in \{f, c\}$ ,  $Q_n^s(x, dy) = G_{n-1}(x)M_n^s(x, dy)$  and for  $B \in \mathcal{X}$ ,  $x_p \in \mathbf{X}$

$$Q_{p,n}^s(B)(x_p) = \int_{\mathbf{X}^{n-p}} \mathbb{I}_B(x_n) \prod_{q=p}^{n-1} Q_{q+1}^s(x_q, dx_{q+1})$$

$0 \leq p < n$  and in the case  $p = n$ ,  $Q_{p,n}^s$  is the identity operator. Now denote for  $0 \leq p < n$ ,  $s \in \{f, c\}$ ,  $B \in \mathcal{X}$ ,  $x_p \in \mathbf{X}$

$$D_{p,n}^s(B)(x_p) = \frac{Q_{p,n}^s(\mathbb{I}_B - \eta_n^s(B))}{\eta_p^s(Q_{p,n}^s(1))}$$

in the case  $p = n$ ,  $D_{p,n}^s(B)(x) = \mathbb{I}_B - \eta_n^s(B)$ .

We then have our CLTs, where the proofs are given in Appendix A. Recall that for the MCPF  $M_n^s$  is used as a notation for the kernel and density of  $M_n^s$ .

**Theorem 4.1.** *We have*

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

$$\sqrt{n}[\check{\eta}_n^{N,I} - \check{\eta}_n^I](\varphi \otimes 1 - 1 \otimes \varphi) \Rightarrow \mathcal{N}(0, \sigma_n^{2,I}(\varphi))$$

where

$$\sigma_n^{2,I}(\varphi) = \sum_{p=0}^n \check{\eta}_p^I(\{(D_{p,n}^f(\varphi) \otimes 1 - 1 \otimes D_{p,n}^c(\varphi))\}^2).$$

2. *For the MCRPF: Assume (A1-3), then for any  $\varphi \in \mathcal{C}_b(\mathbf{X})$ ,  $n \geq 0$  we have*

$$\sqrt{n}[\check{\eta}_n^{N,M} - \check{\eta}_n^M](\varphi \otimes 1 - 1 \otimes \varphi) \Rightarrow \mathcal{N}(0, \sigma_n^{2,M}(\varphi))$$

where

$$\sigma_n^{2,M}(\varphi) = \sum_{p=0}^n \check{\eta}_p^M(\{(D_{p,n}^f(\varphi) \otimes 1 - 1 \otimes D_{p,n}^c(\varphi))\}^2).$$

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

$$\sqrt{n}[\check{\eta}_n^{N,C} - \check{\eta}_n^C](\varphi \otimes 1 - 1 \otimes \varphi) \Rightarrow \mathcal{N}(0, \sigma_n^{2,C}(\varphi))$$

where

$$\sigma_n^{2,C}(\varphi) = \sum_{p=0}^n \check{\eta}_p^C(\{(D_{p,n}^f(\varphi) \otimes 1 - 1 \otimes D_{p,n}^c(\varphi))\}^2).$$

For Wasserstein resampling, the following result is established in [19].

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

$$\sqrt{N}[\check{\eta}_n^{N,W} - \check{\eta}_n^W](\varphi \otimes 1 - 1 \otimes \varphi) \Rightarrow \mathcal{N}(0, \sigma_n^{2,W}(\varphi))$$

where

$$\sigma_n^{2,W}(\varphi) = \sum_{p=0}^n \check{\eta}_p^W (\{(D_{p,n}^f(\varphi) \otimes 1 - 1 \otimes D_{p,n}^c(\varphi))\}^2).$$

**Remark 4.1.** *One can also prove a multivariate CLT using the Cramer-Wold device. Consider  $1 \leq t < +\infty, (\varphi_1, \dots, \varphi_t, \psi_1, \dots, \psi_t) \in \mathcal{B}_b(\mathbf{X})^{2t}$  (or if required  $\mathcal{C}_b(\mathbf{X})^{2t}$ ). Consider the  $t \times t$  positive definite and symmetric matrix  $\Sigma_n^s(\varphi_{1:t}, \psi_{1:t}), s \in \{I, M, C, W\}$ , with  $(i, j)$ th entry denoted  $\Sigma_{n,(ij)}^s(\varphi_{1:t}, \psi_{1:t})$ . Using the Cramer-Wold device under the various assumptions of each the algorithms one can easily deduce that for each  $s \in \{I, M, C, W\}$*

$$\sqrt{N}([\check{\eta}_n^{N,s} - \check{\eta}_n^s](\varphi_1 \otimes 1 - 1 \otimes \psi_1), \dots, [\check{\eta}_n^{N,s} - \check{\eta}_n^s](\varphi_t \otimes 1 - 1 \otimes \psi_t)) \Rightarrow \mathcal{N}_t(0, \Sigma_n^s(\varphi_{1:t}, \psi_{1:t}))$$

where

$$\Sigma_{n,(ij)}^s(\varphi_{1:t}, \psi_{1:t}) = \sum_{p=0}^n \check{\eta}_p^s ([D_{p,n}^f(\varphi_i) \otimes 1 - 1 \otimes D_{p,n}^c(\psi_i)][D_{p,n}^f(\varphi_j) \otimes 1 - 1 \otimes D_{p,n}^c(\psi_j)]).$$

## 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  $[\eta_n^f - \eta_n^c](\varphi)$ . The following assumptions are used, which are essentially those in [29] (see also [18]) with some additional assumptions ((H6-7) below) as we are treating a more delicate case than in [29]. The assumptions can hold on unbounded state-spaces as will be the case for our applications.

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

$$\sup_{n \geq 1} Q_n^s(e^{\tilde{V}})(x) \leq e^{(1-\delta)\tilde{V}(x) + b_d \mathbb{1}_{C_d}(x)}$$

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

**(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  $\underline{d}$  as in (H1), for each  $d \in [\underline{d}, \infty), s \in \{f, c\}$ ,

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

and there exist  $\tilde{\epsilon}_d^- > 0, \nu_d \in \mathcal{P}_v$  such that for each  $A \in \mathcal{X}, s \in \{f, c\}$ ,

$$\inf_{n \geq 1} \int_{C_d \cap A} Q_n^s(x, dy) \geq \tilde{\epsilon}_d^- \nu_d(C_d \cap A), \quad \forall x \in C_d.$$

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

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

**(H5)** 1.

$$\sup_{n \geq 0} \sup_{x \in \mathbf{X}} G_n(x) < +\infty.$$

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

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

$$\frac{1}{G_n M_{n+1}^s(\mathbb{I}_{C_d})} \in \mathcal{L}_v(\mathbf{X})$$

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

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

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

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

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

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

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

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

**Theorem 4.3.** *Assume (H1-6). Then for any  $\xi \in (0, 1/4)$  there exists a  $\rho < 1$  and  $C < +\infty$  depending on the constants in (H1-6) such that for any  $\varphi \in \mathcal{B}_b(\mathbf{X})$ ,  $n \geq 1$ ,  $s \in \{I, M, C, W\}$  we have*

$$\begin{aligned} \sigma_n^{2,s}(\varphi) &\leq \check{\eta}_n^s \left( \{\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))\}^2 \right) + \\ &C \left\{ B(n, f, c, \varphi, \xi) + \|\varphi\|^2 \sum_{p=0}^{n-1} \rho^{n-p} \left( \check{\eta}_p^s(\mathbb{I}_A(v \otimes v)^{2\xi})(\rho^{n-p} + 1) \right) \right\}. \end{aligned}$$

Additionally assume (H7). Then if  $\mathbf{d}^2 \in \mathcal{L}_{(v \otimes v)^\xi}$ ,  $\tilde{\xi} \in (0, 1/2)$ , for any  $\xi \in (0, (1 - 2\tilde{\xi})/12)$  there exists a  $\rho < 1$  and  $C < +\infty$  depending on the constants in (H1-7) such that for any  $\varphi \in \mathcal{B}_b(\mathbf{X}) \cap \text{Lip}_{\mathbf{d}}(\mathbf{X})$ ,  $n \geq 1$ ,  $s \in \{I, M, C, W\}$  we have

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

**Remark 4.2.** *The range of  $\xi$  are such that the upper-bounds are finite. This is because  $\check{\eta}_p^s(v)$  is upper-bounded by a constant that does not depend on  $p$  (see [29, 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. That

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

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. That

$$\tilde{\eta}_p^s(\mathbb{I}_A(v \otimes v)^{2\xi}) \quad \text{or} \quad \tilde{\eta}_p^s(\mathbf{d}^2(v^{4\xi} \otimes v^{8\xi}))$$

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

We note that in 1. strictly 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  $\mathbf{d}$  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 (H1-8) can be verified. We begin with a control of the term  $B(n, l, l-1, \varphi, \xi)$ .

### 5.1 A General Result

Define the set, for  $x \in \mathbf{X}$

$$\mathbf{B}_l(x) := \{y \in \mathbf{X} : \|y - x\| > 2\Delta_l\}$$

where  $\|\cdot\|$  is the  $L_2$ -norm. We write the transition densities w.r.t. Lebesgue measure  $dy$  of the diffusion as  $M(x, y)$  and the Euler approximation as  $M^l(x, y)$ ,  $(x, y) \in \mathbf{X} \times \mathbf{X}$ . We note the following two equations quoted in [11]: there exists a  $0 < C, C' < +\infty$  (which depend on the drift and diffusion coefficients of (4)) such that for any  $l \geq 0$

$$M(x, y) + M^l(x, y) \leq C \exp\{-C'\|y - x\|^2\} \quad \forall (x, y) \in \mathbf{X} \times \mathbf{X} \quad (8)$$

$$|M(x, y) - M^l(x, y)| \leq C\Delta_l \exp\{-C'\|y - x\|^2\} \quad \forall x \in \mathbf{X}, y \in \mathbf{B}_l(x). \quad (9)$$

We add an additional assumption:

(H8) We have, for  $C'$  as in (8)-(9):

1. For any  $\xi \in (0, 1/2)$ , there exists a  $C < +\infty$  such that for any  $x \in \mathbf{X}$ ,  $l \geq 0$

$$\int_{\mathbf{B}_l(x)^c} v(y)^\xi dy \leq C v(x)^{2\xi} \int_{\mathbf{B}_l(x)^c} dy.$$

2. For any  $\xi \in (0, 1/2)$ , there exists a  $C < +\infty$  such that for any  $x \in \mathbf{X}$ ,  $l \geq 0$

$$\int_{\mathbf{B}_l(x)} v(y)^\xi \exp\{-C'\|y - x\|^2\} dy \leq C v(x)^{2\xi}.$$

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

**Proposition 5.1.** *Assume (H1-6), (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 (H1-6) (H8) such that for any  $\varphi \in \mathcal{B}_b(\mathbf{X})$   $n \geq 1$ ,  $1 \leq l \leq L$ , we have*

$$B(n, l, l-1, \varphi, \xi) \leq C \|\varphi\|^2 \Delta_l v(x^*)^{2\xi + \hat{\xi}}.$$

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  $\tilde{\eta}_n^s$ .

### 5.2 IRCPF

We consider the term

$$\tilde{\eta}_n^I(\{\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))\}^2)$$

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

$$\tilde{\eta}_n^I(\{\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))\}^2) =$$

$$\frac{(\eta_{n-1}^f \otimes \eta_{n-1}^c)([G_{n-1} \otimes G_{n-1}]\check{M}_n([\varphi \otimes 1 - 1 \otimes \varphi]^2))}{(\eta_{n-1}^f \otimes \eta_{n-1}^c)(G_{n-1} \otimes G_{n-1})} - [\eta_n^f - \eta_n^c](\varphi)^2.$$

Then using  $\varphi \in \mathcal{A}$  (along with the  $C_2$ -inequality) and [21, Lemma D.2.], there is a finite constant  $C < +\infty$  that depends on  $n$  such that

$$\check{\eta}_n^I \left( \{\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))\}^2 \right) \leq C \left( \frac{(\eta_{n-1}^f \otimes \eta_{n-1}^c)([G_{n-1} \otimes G_{n-1}]\check{M}_n(\tilde{\varphi}))}{(\eta_{n-1}^f \otimes \eta_{n-1}^c)(G_{n-1} \otimes G_{n-1})} + \Delta_l^2 \right)$$

where  $\tilde{\varphi}(x, y) = \|x - y\|^2$ . Now applying (6) one has the upper-bound

$$C \left( \frac{(\eta_{n-1}^f \otimes \eta_{n-1}^c)([G_{n-1} \otimes G_{n-1}]\tilde{\varphi})}{(\eta_{n-1}^f \otimes \eta_{n-1}^c)(G_{n-1} \otimes G_{n-1})} + \Delta_l + \Delta_l^2 \right).$$

In general, there is no reason to expect that  $(\eta_{n-1}^f \otimes \eta_{n-1}^c)([G_{n-1} \otimes G_{n-1}]\tilde{\varphi})$  is small as a function of  $\Delta_l$ . 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 MCRPF

To start our discussion, suppose first that  $G_n$  is constant for each  $n \geq 0$ . This represents the most favourable scenario for the MCRPF, 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  $\mathbf{d}_1$  the  $L_1$ -norm and set  $\varphi \in \mathcal{B}_b(\mathbb{X}) \cap \text{Lip}_{\mathbf{d}_1}(\mathbb{X})$  ( $\mathbb{X} = \mathbb{R}^d$ ). Then setting  $\tilde{\varphi}(x, y) = \|x - y\|^2$

$$\check{\eta}_n^M \left( \{\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))\}^2 \right) \leq C \|\varphi\|_{\text{Lip}}^2 \check{\eta}_n^M(\tilde{\varphi})$$

where  $C < +\infty$  does not depend on  $n$ . Now as  $G_n$  constant, one can deduce (e.g. by [24, Theorem 2.7.3], recall (D) is assumed) that

$$\check{\eta}_n^M \left( \{\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi))\}^2 \right) \leq Cn(n+4) \exp\{C'n^2\} \Delta_l \quad (10)$$

where  $C, C' < +\infty$  does not depend on  $n$ . Now (10) is not necessarily tight in  $n$  for every diffusion that satisfies (D), for instance if one has  $dZ_t = a dt + b dW_t$ , for  $b > 0$ ,  $a \in (-1, 0)$ , then there is no dependence on  $n$  on the R.H.S. of (10), however, it is worrying that the coupling can be exponentially bad in time, in this favourable case. The main point here, is that the principal source of coupling in this algorithm is  $\check{M}$  and if this coupling deteriorates when iterating  $\check{M}$ , then for the MCRPF one *cannot hope to obtain time uniform couplings*.

More generally, if  $G_n$  is non-constant, we first remark that the upper-bound (10) is likely to be  $\mathcal{O}(\Delta_l^{1/2})$  as it is the resampling operation where the forward rate is lost (see [21]). Secondly, one expects to find *examples* where  $\sigma_n^{2,s}(\varphi)$  is time-uniform or grows slowly as a function of  $n$  (due to the empirical results in [21]). However, due to the highly non-linear and complex expression for  $\check{\Phi}_n^M$ , 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 (H1-6), (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 (H1-6) (H8) such that for any  $\varphi \in \mathcal{B}_b(\mathbb{X})$ ,  $n \geq 0$ ,  $1 \leq l \leq L$  we have*

$$\sigma_n^{2,C}(\varphi) \leq C \|\varphi\|^2 \Delta_l v(x^*)^{32\xi + 2\hat{\xi}}.$$

### 5.5 WCPF

Note that  $\mathbb{X} = \mathbb{R}$  and, for the metric in (H7), we take  $\mathbf{d}_1$  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 (H1-8). Then, let  $\lambda \in (0, 1)$  be given, if  $\tilde{\varphi} \in \mathcal{L}_{(v \otimes v)^\xi}(\mathbf{X})$ , for any  $\tilde{\xi} \in (0, 1/(16(1 + \lambda)))$  and set  $(\xi, \hat{\xi}) \in (0, \min\{1/32, \lambda/(16(1 + \lambda)), (1 - 2\tilde{\xi})/12\}) \times (0, 1/2)$  then there exists a  $C < +\infty$  depending on the constants in (H1-8) such that for any  $\varphi \in \mathcal{B}_b(\mathbf{X}) \cap \text{Lip}_{d_1}(\mathbf{X})$ ,  $n \geq 0$ ,  $1 \leq l \leq L$  we have*

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

As  $\lambda > 0$  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 case and discuss strategies to establish this in Remark F.2 in Appendix F, Section F.3.

## 5.6 Example

We consider  $\mathbf{X} = \mathbb{R}$ , the diffusion  $dZ_t = -\frac{3}{2}Z_t dt + dW_t$  and  $\mathbf{Y} = \{0, 1\}$  and  $G(x, y) = \left(\frac{be^x + a}{1 + e^x}\right)^y \left(1 - \left(\frac{be^x + a}{1 + e^x}\right)\right)^{1-y}$ ,  $y \in \mathbf{Y}$  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 a  $L^* > L$  for one 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  $\alpha_l = (1 - (3/2)\Delta_l)^{\Delta_l^{-1}}$ ,  $\beta_l = \Delta_l(1 - (1 - (3/2)\Delta_l)^{2\Delta_l^{-1}})/(1 - (1 - (3/2)\Delta_l)^2)$ . Due to our assumption concerning  $L^*$ , one can easily find constants  $0 < \underline{\alpha} < \bar{\alpha} < 1$ ,  $0 < \underline{\beta} < \bar{\beta} \leq 1$ , that do not depend on  $l$ , with  $\underline{\alpha} \leq |\alpha_l| \leq \bar{\alpha}$ ,  $\underline{\beta} \leq \beta_l \leq \bar{\beta}$  for each  $l \in \{0, \dots, L\}$ .

To verify (H1) we use, as in [30, Example 4.2.1], the Lyapunov function  $\tilde{V}(x) = \frac{1}{2(1+\delta_0)}x^2 + 1$  for some  $\delta_0 > 0$ . One can show that for any  $1 \leq l \leq L$ ,  $n \geq 1$ ,  $y \in \mathbf{X}$

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

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  $\underline{d} > 1$  large enough. (H2-5) follow by elementary calculations associated to the choice of  $\tilde{V}$ ,  $M^l$  and  $G_n$  and are omitted.

To verify (H6) as  $G_n$  is lower-bounded, we consider  $M^l(C_d)(x)$  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  $x > 0$ :

$$v(x)M^l(C_d)(x) \geq C \exp \left\{ \frac{x^2}{2} \left( \frac{1}{\delta_0 + 1} - \frac{\alpha_l^2}{\beta_l} \right) - \frac{\alpha_l}{\beta_l} x \tilde{d} \right\}$$

where  $\tilde{d} = \sqrt{(d-1)2(1+\delta_0)}$ . 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, where the 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 (9) is at least  $1 + 1/(2(1 + \delta_0))$  (i.e.  $C' > 1/12$ ). It is noted that it is non-trivial 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 non-trivial 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 requires either the density of the transition and have a random running time per time step (MCPF) or is constrained to the case that  $\mathbf{X}$  is one-dimensional (WCPF). None-the-less these are still useful algorithms when they can be implemented.

There are a number of extensions to this work. First one may extend these results to the context of normalization constant estimation (e.g. [20]) and the associated asymptotic variance. Secondly, is a more in depth analysis and implementation of the MCPF, which to our knowledge has not been performed in the literature.

## Acknowledgements

AJ & FY were supported by an AcRF tier 2 grant: R-155-000-161-112. AJ 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.

## A Common Proofs for the CLT

The structure of all the appendices are to give the proof of the main result at the beginning. The associated technical results are given in such a way that they should be read in order to understand the overall proof.

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

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

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

$$R_{p+1}^{N,s}(\varphi) = \frac{\eta_p^{N,s}(D_{p,n}^s(\varphi))}{\eta_p^{N,s}(G_p)}[\eta_p^s(G_p) - \eta_p^{N,s}(G_p)].$$

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

$$\begin{aligned} \sqrt{N}[\check{\eta}_m^{N,t} - \check{\eta}_m^t](\varphi \otimes 1 - 1 \otimes \varphi) &= \sum_{p=0}^n \{V_p^{N,f}(D_{p,n}^f(\varphi)) + \sqrt{N}R_{p+1}^{N,f}(D_{p,n}^f(\varphi))\} - \\ &\quad \sum_{p=0}^n \{V_p^{N,c}(D_{p,n}^c(\varphi)) + \sqrt{N}R_{p+1}^{N,c}(D_{p,n}^c(\varphi))\}. \end{aligned} \quad (11)$$

*Proof of Theorem 4.1.* The proof follows immediately from (11), Lemma A.1, and Proposition A.1.  $\square$

### 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 [19], 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 MCRPF & MCPF is mentioned where required.

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

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

*Proof.* By Proposition B.1 (for MCRPF it is [21, Proposition C.6], 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 show that

$$\sqrt{N}\eta_p^{N,s}(D_{p,n}^s(\varphi))[\eta_p^s(G_p) - \eta_p^{N,s}(G_p)]$$

will converge in probability to zero. By Cauchy-Schwarz:

$$\begin{aligned} \sqrt{N}\mathbb{E}[|\eta_p^{N,s}(D_{p,n}^s(\varphi))[\eta_p^s(G_p) - \eta_p^{N,s}(G_p)]|] &\leq \\ \sqrt{N}\mathbb{E}[|\eta_p^{N,s}(D_{p,n}^s(\varphi))|^2]^{1/2}\mathbb{E}[|\eta_p^s(G_p) - \eta_p^{N,s}(G_p)|^2]^{1/2}. \end{aligned}$$

Applying Proposition B.1 (for MCRPF [21, Proposition C.6], for MCPF it is Proposition D.1) it easily follows that there is a finite constant  $C < +\infty$  that does not depend upon  $N$  such that

$$\sqrt{N}\mathbb{E}[|\eta_p^{N,s}(D_{p,n}^s(\varphi))[\eta_p^s(G_p) - \eta_p^{N,s}(G_p)]|] \leq \frac{C}{\sqrt{N}}.$$

This bound allows one to easily conclude.  $\square$

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

$$\begin{aligned}\mathbb{E}[V_p^{N,s}(\varphi)] &= 0 \\ \lim_{N \rightarrow +\infty} \mathbb{E}[V_p^{N,s}(\varphi)^2] &= \eta_p^s((\varphi - \eta_p^s(\varphi))^2).\end{aligned}$$

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

$$\begin{aligned}\mathbb{E}[V_p^{N,s}(\varphi)^2] &= \frac{1}{N} \sum_{i=1}^N \mathbb{E}[(\varphi(X_p^{i,s}) - \Phi_p^s(\eta_{p-1}^{N,s})(\varphi))^2] \\ &= \mathbb{E}[\varphi(X_p^{1,s})^2] - \mathbb{E}[\Phi_p^s(\eta_{p-1}^{N,s})(\varphi)^2]\end{aligned}$$

$\Phi_p^s(\eta_{p-1}^{N,s})(\varphi)$  is a bounded random quantity and moreover by Proposition B.1 (MCRPF [21, Proposition C.6], MCPF Proposition D.1) it converges in probability to  $\eta_p^s(\varphi)$ . Hence by [3, Theorem 25.12],  $\lim_{N \rightarrow +\infty} \mathbb{E}[\Phi_p^s(\eta_{p-1}^{N,s})(\varphi)^2] = \eta_p^s(\varphi)^2$ . Hence we consider

$$\mathbb{E}[\varphi(X_p^{1,s})^2] = \mathbb{E}\left[\frac{1}{N} \sum_{i=1}^N \varphi(X_p^{i,s})^2 - \eta_p^s(\varphi^2)\right] + \eta_p^s(\varphi^2).$$

By Jensen

$$\mathbb{E}\left[\frac{1}{N} \sum_{i=1}^N \varphi(X_p^{i,s})^2 - \eta_p^s(\varphi^2)\right] \leq \mathbb{E}\left[\left|\frac{1}{N} \sum_{i=1}^N \varphi(X_p^{i,s})^2 - \eta_p^s(\varphi^2)\right|^{1/2}\right]$$

and hence we conclude via Proposition B.1 (MCRPF [21, Proposition C.6], MCPF Proposition D.1) that

$$\lim_{N \rightarrow +\infty} \mathbb{E}[\varphi(X_p^{1,s})^2] = \eta_p^s(\varphi^2)$$

and the result thus follows.  $\square$

**Lemma A.3.** We have:

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

$$\lim_{N \rightarrow +\infty} \mathbb{E}[V_p^{N,f}(\varphi)V_p^{N,c}(\psi)] = \check{\eta}_p^f(\varphi \otimes \psi) - \eta_p^f(\varphi)\eta_p^c(\psi).$$

2. For the MCRPF: Assume (A1-3),  $(\varphi, \psi) \in \mathcal{C}_b(\mathbf{X})^2$ ,  $p \geq 0$  then we have

$$\lim_{N \rightarrow +\infty} \mathbb{E}[V_p^{N,f}(\varphi)V_p^{N,c}(\psi)] = \check{\eta}_p^M(\varphi \otimes \psi) - \eta_p^f(\varphi)\eta_p^c(\psi).$$

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

$$\lim_{N \rightarrow +\infty} \mathbb{E}[V_p^{N,f}(\varphi)V_p^{N,c}(\psi)] = \check{\eta}_p^C(\varphi \otimes \psi) - \eta_p^f(\varphi)\eta_p^c(\psi).$$

*Proof.* We have for all methods:

$$\begin{aligned}\mathbb{E}[V_p^{N,f}(\varphi)V_p^{N,c}(\psi)] &= N \left( \mathbb{E}[\eta_p^{N,f}(\varphi)\eta_p^{N,c}(\psi)] - \mathbb{E}[\eta_p^{N,f}(\varphi)\Phi_p^c(\eta_{p-1}^{N,c})(\psi)] - \right. \\ &\quad \left. \mathbb{E}[\Phi_p^f(\eta_{p-1}^{N,f})(\varphi)\eta_p^{N,c}(\psi)] + \mathbb{E}[\Phi_p^f(\eta_{p-1}^{N,f})(\varphi)\Phi_p^c(\eta_{p-1}^{N,c})(\psi)] \right) \\ &= N \left( \mathbb{E}[\eta_p^{N,f}(\varphi)\eta_p^{N,c}(\psi)] - \mathbb{E}[\Phi_p^f(\eta_{p-1}^{N,f})(\varphi)\Phi_p^c(\eta_{p-1}^{N,c})(\psi)] \right).\end{aligned}\tag{12}$$

We consider calculations for (i)-(iii) associated to (12).

**Proof for (i):** Now

$$\begin{aligned}N\mathbb{E}[\eta_p^{N,f}(\varphi)\eta_p^{N,c}(\psi)] &= \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \mathbb{E}[\varphi(X_p^{i,f})\psi(X_p^{j,c})] \\ &= \mathbb{E}[\check{\Phi}_p^f(\eta_{p-1}^{N,f} \otimes \eta_{p-1}^{N,c})(\varphi \otimes \psi)] + (N-1)\mathbb{E}[\Phi_p^f(\eta_{p-1}^{N,f})(\varphi)\Phi_p^c(\eta_{p-1}^{N,c})(\psi)].\end{aligned}$$



Thus,

$$\mathbb{E}[V_p^{N,f}(\varphi)V_p^{N,c}(\psi)] = \mathbb{E}[\check{\Phi}_p^I(\eta_{p-1}^{N,f} \otimes \eta_{p-1}^{N,c})(\varphi \otimes \psi)] - \mathbb{E}[\Phi_p^f(\eta_{p-1}^{N,f})(\varphi)\Phi_p^c(\eta_{p-1}^{N,c})(\psi)].$$

$\check{\Phi}_p^I(\eta_{p-1}^{N,f} \otimes \eta_{p-1}^{N,c})(\varphi \otimes \psi)$  is a bounded random quantity and by Propositions B.1 and B.2 it converges in probability to  $\check{\eta}_p^I(\varphi \otimes \psi)$  hence by [3, Theorem 25.12],

$$\lim_{N \rightarrow +\infty} \mathbb{E}[\check{\Phi}_p^I(\eta_{p-1}^{N,f} \otimes \eta_{p-1}^{N,c})(\varphi \otimes \psi)] = \check{\eta}_p^I(\varphi \otimes \psi).$$

Similarly via Proposition B.1 and [3, Theorem 25.12]

$$\lim_{N \rightarrow +\infty} \mathbb{E}[\Phi_p^f(\eta_{p-1}^{N,f})(\varphi)\Phi_p^c(\eta_{p-1}^{N,c})(\psi)] = \eta_p^f(\varphi)\eta_p^c(\psi)$$

and hence we can conclude the proof of (i).

**Proof for (ii):** Using similar calculations as for (i) we have

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

The proof of Theorem 3.1 establishes that

$$\check{\Phi}_p^M(\check{\eta}_{p-1}^{N,M})(\varphi \otimes \psi) \rightarrow_{\mathbb{P}} \check{\eta}_p^M(\varphi \otimes \psi).$$

The proof can then be completed in much the same way as for (i).

**Proof for (iii):** Using similar calculations as for (i) we have

$$\mathbb{E}[V_p^{N,f}(\varphi)V_p^{N,c}(\psi)] = \mathbb{E}[\check{\Phi}_p^C(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\varphi \otimes \psi)] - \mathbb{E}[\Phi_p^f(\eta_{p-1}^{N,f})(\varphi)\Phi_p^c(\eta_{p-1}^{N,c})(\psi)].$$

The proof of Theorem D.1 establishes that

$$\check{\Phi}_p^C(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\varphi \otimes \psi) \rightarrow_{\mathbb{P}} \check{\eta}_p^C(\varphi \otimes \psi).$$

The proof can then be completed in much the same way as for (i). □

Define for  $p \geq 0$ ,  $\varphi, \psi \in \mathcal{B}_b(\mathbf{X})$ .

$$V_p^N(\varphi, \psi) = V_p^{N,f}(\varphi) - V_p^{N,c}(\psi).$$

**Proposition A.1.** *We have:*

1. *For the IRCPF: Let  $n \geq 0$ , then for any  $(\varphi_0, \dots, \varphi_n) \in \mathcal{B}_b(\mathbf{X})^{n+1}$ ,  $(\psi_0, \dots, \psi_n) \in \mathcal{B}_b(\mathbf{X})^{n+1}$ ,  $(V_0^N(\varphi_0, \psi_0), \dots, V_n^N(\varphi_n, \psi_n))$  converges in distribution to a  $(n+1)$ -dimensional Gaussian random variable with zero mean and diagonal covariance matrix, with the  $p \in \{0, \dots, n\}$  diagonal entry*

$$\check{\eta}_p^I(\{(\varphi_p \otimes 1 - 1 \otimes \psi_p) - \check{\eta}_p^I(\varphi_p \otimes 1 - 1 \otimes \psi_p)\}^2).$$

2. *For the MCRPF: Assume (A1-3), then for  $n \geq 0$  and any  $(\varphi_0, \dots, \varphi_n) \in \mathcal{C}_b(\mathbf{X})^{n+1}$ ,  $(\psi_0, \dots, \psi_n) \in \mathcal{C}_b(\mathbf{X})^{n+1}$ ,  $(V_0^N(\varphi_0, \psi_0), \dots, V_n^N(\varphi_n, \psi_n))$  converges in distribution to a  $(n+1)$ -dimensional Gaussian random variable with zero mean and diagonal covariance matrix, with the  $p \in \{0, \dots, n\}$  diagonal entry*

$$\check{\eta}_p^M(\{(\varphi_p \otimes 1 - 1 \otimes \psi_p) - \check{\eta}_p^M(\varphi_p \otimes 1 - 1 \otimes \psi_p)\}^2).$$

3. *For the MCPF: Suppose that for  $s \in \{f, c\}$ ,  $n \geq 1$ ,  $M_n^s \in \mathcal{B}_b(\mathbf{X} \times \mathbf{X})$ , then for  $n \geq 0$  and any  $(\varphi_0, \dots, \varphi_n) \in \mathcal{B}_b(\mathbf{X})^{n+1}$ ,  $(\psi_0, \dots, \psi_n) \in \mathcal{B}_b(\mathbf{X})^{n+1}$ ,  $(V_0^N(\varphi_0, \psi_0), \dots, V_n^N(\varphi_n, \psi_n))$  converges in distribution to a  $(n+1)$ -dimensional Gaussian random variable with zero mean and diagonal covariance matrix, with the  $p \in \{0, \dots, n\}$  diagonal entry*

$$\check{\eta}_p^C(\{(\varphi_p \otimes 1 - 1 \otimes \psi_p) - \check{\eta}_p^C(\varphi_p \otimes 1 - 1 \otimes \psi_p)\}^2).$$

*Proof.* This follows by using almost the same exposition and proofs as [8], pp. 293-294, Theorem 9.3.1 and Corollary 9.3.1 and the results (of this paper) Lemmata A.2-A.3. The proof is thus omitted. □

## 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  $\varphi \in \mathcal{B}_b(X)$ ,  $N \geq 1$  we have*

$$\mathbb{E}[|[\eta_n^{N,s} - \eta_n^s](\varphi)|^p]^{1/p} \leq \frac{C\|\varphi\|}{\sqrt{N}}.$$

*Proof.* This can be proved easily by induction. For instance using the strategy in [21, Proposition C.6] and is hence omitted.  $\square$

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

$$(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi) \rightarrow_{\mathbb{P}} (\eta_n^f \otimes \eta_n^c)(\varphi).$$

*Proof.* Our proof is by induction, consider the case  $n = 0$  and let  $\epsilon > 0$  be arbitrary, then

$$\begin{aligned} & \mathbb{P}(|(\eta_n^{N,f} \otimes \eta_n^{N,c}) - (\eta_n^f \otimes \eta_n^c)|(\varphi)| > \epsilon) \leq \\ & \mathbb{P}(|[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi) - \mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)]]| > \epsilon/2) + \mathbb{P}(|\mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)] - (\eta_n^f \otimes \eta_n^c)(\varphi)| > \epsilon/2) \leq \\ & \mathbb{P}(|[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi) - \mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)]]| > \epsilon/2) + \mathcal{O}(N^{-1}) \end{aligned}$$

where the last line follows as

$$\mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)] = \frac{N-1}{N}(\eta_n^f \otimes \eta_n^c)(\varphi) + \frac{1}{N}\tilde{\eta}_n^f(\varphi).$$

To deal with the term  $\mathbb{P}(|[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi) - \mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)]]| > \epsilon/2)$  we will apply the bounded difference inequality (see [25]). To this end, first note that  $(u_0^1, \dots, u_0^N)$  are i.i.d. and setting  $f(u_0^1, \dots, u_0^N) = (\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)$  we note that for any  $1 \leq k \leq N$  (the notational convention is clear if  $k = 1$  or  $k = N$ ) and any  $(u_0^1, \dots, u_0^N) \in (X \times X)^{n+1}$ ,  $\tilde{u}_0^k \in X \times X$

$$\begin{aligned} & |f(u_0^1, \dots, u_0^N) - f(u_0^1, \dots, u_0^{k-1}, \tilde{u}_0^k, u_0^{k+1}, \dots, u_0^N)| = \frac{1}{N^2} |[\varphi(u_0^k) - \varphi(\tilde{u}_0^k)] + \\ & \sum_{i \in \{1, \dots, N\} \setminus \{k\}} [\varphi(x_0^{i,f}, x_0^{i,c}) - \varphi(x_0^{i,f}, \tilde{x}_0^{i,c})] + \sum_{i \in \{1, \dots, N\} \setminus \{k\}} [\varphi(x_0^{k,f}, x_0^{i,c}) - \varphi(\tilde{x}_0^{k,f}, x_0^{i,c})] \leq \\ & \frac{4\|\varphi\|}{N}. \end{aligned} \tag{13}$$

Thus, by the bounded difference inequality:

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

Hence we can easily conclude the result when  $n = 0$  as  $\epsilon > 0$  was arbitrary.

Now we assume the result for  $n - 1$  and consider  $n$ , we have

$$\begin{aligned} & \mathbb{P}(|[(\eta_n^{N,f} \otimes \eta_n^{N,c}) - (\eta_n^f \otimes \eta_n^c)](\varphi)| > \epsilon) \leq \mathbb{P}(|[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi) - \mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)|\mathcal{F}_{n-1}^N]| > \epsilon/2) + \\ & \mathbb{P}(|\mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)|\mathcal{F}_{n-1}^N] - (\eta_n^f \otimes \eta_n^c)(\varphi)| > \epsilon/2) \end{aligned}$$

where  $\mathcal{F}_{n-1}^N$  is the  $\sigma$ -algebra generated by the particle system up-to time  $n - 1$ . For the term  $\mathbb{P}(|[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi) - \mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi)|\mathcal{F}_{n-1}^N]| > \epsilon/2)$ , on conditioning upon  $\mathcal{F}_{n-1}^N$  one can apply the above bounded difference inequality, as  $(u_n^1, \dots, u_n^N)$  are (conditionally) i.i.d. and that the bound in (13) can also be obtained for any  $n$ . Hence one has

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

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

$$\begin{aligned}\mathbb{E}[(\eta_n^{N,f} \otimes \eta_n^{N,c})(\varphi) | \mathcal{F}_{n-1}^N] &= \frac{1}{N^2} \left( \sum_{i=1}^N \mathbb{E}[\varphi(U_n^i) | \mathcal{F}_{n-1}^N] + \sum_{i \neq j} \mathbb{E}[\varphi(X_n^{i,f}, X_n^{i,c}) | \mathcal{F}_{n-1}^N] \right) \\ &= \frac{1}{N} \check{\Phi}_n^I(\eta_{n-1}^{N,f} \otimes \eta_{n-1}^{N,c})(\varphi) + \frac{(N-1)}{N} (\Phi_n^f(\eta_{n-1}^{N,f}) \otimes \Phi_n^c(\eta_{n-1}^{N,c}))(\varphi).\end{aligned}$$

Now by the induction hypothesis and Proposition B.1  $\check{\Phi}_n^I(\eta_{n-1}^{N,f} \otimes \eta_{n-1}^{N,c})(\varphi)$  converges in probability to

$$\frac{(\eta_{n-1}^f \otimes \eta_{n-1}^c)((G_{n-1} \otimes G_{n-1})\check{M}_n(\varphi))}{\eta_{n-1}^f(G_{n-1})\eta_{n-1}^c(G_{n-1})}$$

hence the term  $\frac{1}{N} \check{\Phi}_n^I(\eta_{n-1}^{N,f} \otimes \eta_{n-1}^{N,c})(\varphi)$  goes to zero. Then, again by the induction hypothesis and Proposition B.1  $(\Phi_n^f(\eta_{n-1}^{N,f}) \otimes \Phi_n^c(\eta_{n-1}^{N,c}))(\varphi)$  converges in probability to

$$\frac{(\eta_{n-1}^f \otimes \eta_{n-1}^c)((G_{n-1} \otimes G_{n-1})(M_n^f \otimes M_n^c)(\varphi))}{\eta_{n-1}^f(G_{n-1})\eta_{n-1}^c(G_{n-1})} = (\eta_n^f \otimes \eta_n^c)(\varphi)$$

and hence we conclude the result.  $\square$

## C Technical Results for the MCRPF

*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

$$\check{\eta}_n^{N,M}(\varphi) - \check{\eta}_n^M(\varphi) = \check{\eta}_n^{N,M}(\varphi) - \check{\Phi}_n^M(\check{\eta}_{n-1}^{N,M})(\varphi) + \check{\Phi}_n^M(\check{\eta}_{n-1}^{N,M})(\varphi) - \check{\Phi}_n^M(\check{\eta}_{n-1}^M)(\varphi).$$

One can easily prove that

$$|\check{\eta}_n^{N,M}(\varphi) - \check{\Phi}_n^M(\check{\eta}_{n-1}^{N,M})(\varphi)| \rightarrow_{\mathbb{P}} 0$$

by using the (conditional) Marcinkiewicz-Zygmund inequality, so we focus on the latter term. Define

$$\begin{aligned}T_1^N &:= \check{\eta}_{n-1}^{N,M} \left( \{F_{n-1, \check{\eta}_{n-1}^{N,M}, f} \wedge F_{n-1, \check{\eta}_{n-1}^{N,M}, c}\} \check{M}_n(\varphi) \right) - \check{\eta}_{n-1}^M \left( \{F_{n-1, \check{\eta}_{n-1}^M, f} \wedge F_{n-1, \check{\eta}_{n-1}^M, c}\} \check{M}_n(\varphi) \right) \\ T_2^N &:= \left( 1 - \check{\eta}_{n-1}^{N,M} \left( \{F_{n-1, \check{\eta}_{n-1}^{N,M}, f} \wedge F_{n-1, \check{\eta}_{n-1}^{N,M}, c}\} \right) \right) (\check{\eta}_{n-1}^{N,M} \otimes \check{\eta}_{n-1}^{N,M}) \left( \left\{ \bar{F}_{n-1, \check{\eta}_{n-1}^{N,M}, f} \otimes \bar{F}_{n-1, \check{\eta}_{n-1}^{N,M}, c} \right\} \bar{M}_n(\varphi) \right) \\ T_3 &:= \left( 1 - \check{\eta}_{n-1}^M \left( \{F_{n-1, \check{\eta}_{n-1}^M, f} \wedge F_{n-1, \check{\eta}_{n-1}^M, c}\} \right) \right) (\check{\eta}_{n-1}^M \otimes \check{\eta}_{n-1}^M) \left( \left\{ \bar{F}_{n-1, \check{\eta}_{n-1}^M, f} \otimes \bar{F}_{n-1, \check{\eta}_{n-1}^M, c} \right\} \bar{M}_n(\varphi) \right)\end{aligned}$$

Then

$$T_1^N + T_2^N - T_3 = \check{\Phi}_n^M(\check{\eta}_{n-1}^{N,M})(\varphi) - \check{\Phi}_n^M(\check{\eta}_{n-1}^M)(\varphi).$$

By Lemma C.1  $T_1^N \rightarrow_{\mathbb{P}} 0$  and by Lemma C.6,  $T_2^N - T_3 \rightarrow_{\mathbb{P}} 0$ . This concludes the proof.  $\square$

**Lemma C.1.** *Assume (A1). Then if for any  $\varphi \in \mathcal{C}_b(\mathbb{X} \times \mathbb{X})$ ,  $n \geq 0$*

$$\check{\eta}_n^{N,M}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^M(\varphi)$$

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

$$\check{\eta}_n^{N,M} \left( \{F_{n, \check{\eta}_n^{N,M}, f} \wedge F_{n, \check{\eta}_n^{N,M}, c}\} \psi \right) \rightarrow_{\mathbb{P}} \check{\eta}_n^M \left( \{F_{n, \check{\eta}_n^M, f} \wedge F_{n, \check{\eta}_n^M, c}\} \psi \right).$$

*Proof.* Set

$$T_1^N := \check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}\} \psi - \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \psi \right) \quad (14)$$

$$T_2^N := \left( \check{\eta}_n^{N,M} - \check{\eta}_n^N \right) \left( \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \psi \right). \quad (15)$$

Then we have the decomposition

$$T_1^N + T_2^N = \check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}\} \psi \right) - \check{\eta}_n^M \left( \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \psi \right) =$$

We will show that (14)-(15) converge in probability to zero.

**Term (14):** Define

$$T_3^N := \frac{1}{2} \check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^M,f}\} \psi \right) \quad (16)$$

$$T_4^N := \frac{1}{2} \check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},c} - F_{n,\check{\eta}_n^M,c}\} \psi \right) \quad (17)$$

$$T_5^N := \frac{1}{2} \check{\eta}_n^{N,M} \left( \{|F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,c}| - |F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},c}|\} \psi \right) \quad (18)$$

Then we have

$$T_1^N = T_3^N + T_4^N + T_5^N.$$

To show that  $T_1^N$  converges in probability to zero we will show that (16)-(18) each converge in probability to zero. For (16) we have

$$\check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^M,f}\} \psi \right) = \frac{\eta_n^f(G_n) - \eta_n^{N,f}(G_n)}{\eta_n^{N,f}(G_n)\eta_n^f(G_n)} \check{\eta}_n^{N,M} ((G_n \otimes 1)\psi).$$

By [21, Proposition C.6]

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

and by hypothesis  $\check{\eta}_n^{N,M}((G_n \otimes 1)\psi)$  converges in probability, hence (16) converges in probability to zero. For (17) this term converges in probability to zero by an almost identical argument to (16) and is hence omitted. For (18)

$$\begin{aligned} & \check{\eta}_n^{N,M} \left( \{|F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,c}| - |F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},c}|\} \psi \right) = \\ & \frac{\eta_n^{N,f}(G_n)\eta_n^{N,c}(G_n) - \eta_n^f(G_n)\eta_n^c(G_n)}{\eta_n^f(G_n)\eta_n^c(G_n)\eta_n^{N,f}(G_n)\eta_n^{N,c}(G_n)} \check{\eta}_n^{N,M} (\{|\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)|\} \psi) - \\ & \frac{1}{\eta_n^{N,f}(G_n)\eta_n^{N,c}(G_n)} \check{\eta}_n^{N,M} \left( \{|\eta_n^{N,c}(G_n)(G_n \otimes 1) - \eta_n^{N,f}(G_n)(1 \otimes G_n)| - \right. \\ & \left. |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)|\} \psi \right). \end{aligned} \quad (19)$$

The first term on the R.H.S. converges in probability to zero by [21, Proposition C.6] and hypothesis. Hence as

$$\frac{1}{\eta_n^{N,f}(G_n)\eta_n^{N,c}(G_n)}$$

converges in probability, we need only show that

$$\check{\eta}_n^{N,M} \left( \{|\eta_n^{N,c}(G_n)(G_n \otimes 1) - \eta_n^{N,f}(G_n)(1 \otimes G_n)| - |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)|\} \psi \right)$$

converges in probability to zero to conclude. Using standard algebra, we have almost surely, that

$$\begin{aligned} & \left| \eta_n^{N,c}(G_n)(G_n \otimes 1) - \eta_n^{N,f}(G_n)(1 \otimes G_n) - |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)| \right| \leq \\ & |[\eta_n^{N,c} - \eta_n^c](G_n)(G_n \otimes 1)| + |[\eta_n^{N,f} - \eta_n^f](G_n)(1 \otimes G_n)|. \end{aligned} \quad (20)$$

Hence

$$\begin{aligned} & \check{\eta}_n^{N,M} \left( \{ |\eta_n^{N,c}(G_n)(G_n \otimes 1) - \eta_n^{N,f}(G_n)(1 \otimes G_n)| - |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)| \} \psi \right) \leq \\ & \quad [|\eta_n^{N,c} - \eta_n^c|(G_n)] \check{\eta}_n^{N,M}((G_n \otimes 1)|\psi|) + [|\eta_n^{N,f} - \eta_n^f|(G_n)] \check{\eta}_n^{N,M}((1 \otimes G_n)|\psi|) \end{aligned}$$

and one can easily conclude by the above arguments.

**Term (15):** As

$$\{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \psi = \frac{1}{2} \left( F_{n,\check{\eta}_n^M,f} + F_{n,\check{\eta}_n^M,c} - |F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,c}| \right) \psi$$

we have  $\{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \psi \in \mathcal{C}_b(\mathbb{X} \times \mathbb{X})$ , so (15) converges in probability to zero.  $\square$

**Lemma C.2.** *Assume (A1). Then if for any  $\varphi \in \mathcal{C}_b(\mathbb{X} \times \mathbb{X})$ ,  $n \geq 0$*

$$\check{\eta}_n^{N,M}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^M(\varphi)$$

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

$$\begin{aligned} & (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \frac{F_{n,\check{\eta}_n^{N,M},f}}{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \\ & - (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \frac{F_{n,\check{\eta}_n^M,f}}{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \rightarrow_{\mathbb{P}} 0. \end{aligned}$$

*Proof.* Define

$$T_1^N := (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \frac{F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^M,f}}{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \quad (21)$$

$$T_2^N := \frac{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}) - \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})}{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}) \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})} \times (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ F_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \quad (22)$$

then we have

$$\begin{aligned} T_1^N + T_2^N &= (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \frac{F_{n,\check{\eta}_n^{N,M},f}}{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) - \\ & \quad - (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \frac{F_{n,\check{\eta}_n^M,f}}{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right). \end{aligned}$$

We will show that (21) and (22) will converge in probability to zero.

**Term (21):** For (21) we note that

$$\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}) = 1 - \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}).$$

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

$$\begin{aligned} & (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ (F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^M,f}) \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) = \\ & \left( \frac{1}{\eta_n^{N,f}(G_n)} - \frac{1}{\eta_n^f(G_n)} \right) (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ (G_n \otimes 1) \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right). \end{aligned}$$

By [21, Proposition C.6]

$$\left( \frac{1}{\eta_n^{N,f}(G_n)} - \frac{1}{\eta_n^f(G_n)} \right) \rightarrow_{\mathbb{P}} 0$$

so if we can show that the remaining term on the R.H.S. is, in absolute value, (almost surely) upper-bounded by a convergent (in probability) random variable, we have shown that (21) converges in probability to zero. We have

$$|(\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M})\left(\{(G_n \otimes 1) \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c}\} \bar{M}_{n+1}(\psi)\right)| \leq \|\psi\| \|G_n\| \check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|). \quad (23)$$

Then almost surely

$$\check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|) \leq \frac{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},c}) + \check{\eta}_n^{N,M}(\{F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}\})}{|\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},c} - \{F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}\})|} \quad (24)$$

By the above arguments, both the denominator and numerator will converge in probability to a finite constant and hence we have shown that (21) converges in probability to zero.

**Term (22):** We note

$$\begin{aligned} & \frac{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}) - \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})}{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}) \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})} = \\ & \frac{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}) - \check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})}{(1 - \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}))(1 - \check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}))}. \end{aligned}$$

By Lemma C.1 this converges in probability to zero. Thus, we need only show that  $|(\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M})\left(\{F_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c}\} \bar{M}_{n+1}(\psi)\right)|$  is (almost surely) upper-bounded by a convergent (in probability) random variable. Almost surely,

$$|(\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M})\left(\{F_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c}\} \bar{M}_{n+1}(\psi)\right)| \leq \|\psi\| \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^M,f}) \check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|).$$

As  $F_{n,\check{\eta}_n^M,f} \in \mathcal{C}_b(\mathbf{X} \times \mathbf{X})$ ,  $\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^M,f})$  converges in probability to a finite constant and our proof is concluded by the argument associated to (24).  $\square$

**Lemma C.3.** *Assume (A1). Then if for any  $\varphi \in \mathcal{C}_b(\mathbf{X} \times \mathbf{X})$ ,  $n \geq 0$*

$$\check{\eta}_n^{N,M}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^M(\varphi)$$

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

$$(\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M})\left(\left(\left\{\frac{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}}{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})}\right\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c}\right) \bar{M}_{n+1}(\psi)\right) \rightarrow_{\mathbb{P}} 0.$$

*Proof.* Define

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

$$T_1^N := -\frac{C}{2} (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M})\left(\left(\{F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^M,f}\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c}\right) \bar{M}_{n+1}(\psi)\right) \quad (25)$$

$$T_2^N := -\frac{C}{2} (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M})\left(\left(\{F_{n,\check{\eta}_n^{N,M},c} - F_{n,\check{\eta}_n^M,c}\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c}\right) \bar{M}_{n+1}(\psi)\right) \quad (26)$$

$$T_3^N := -\frac{C}{2} (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M})\left(\left(\{|F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,c}| - |F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},c}|\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c}\right) \bar{M}_{n+1}(\psi)\right). \quad (27)$$

Then we have

$$T_1^N + T_2^N + T_3^N = (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M})\left(\left(\left\{\frac{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}}{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})}\right\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c}\right) \bar{M}_{n+1}(\psi)\right)$$

We will show that (25), (26) and (27) will converge in probability to zero to conclude the proof.

**Term (25):** We have that (25) is equal to

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

By [21, Proposition C.6]

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

so if we can show that the remaining term on the R.H.S. is, in absolute value, (almost surely) upper-bounded by a convergent (in probability) random variable, we have shown that (25) converges in probability to zero. This can be achieved using the argument for (23) in the proof of Lemma C.2 and hence we have shown that (25) converges in probability to zero.

**Term (26):** The argument is almost identical to (25) and is omitted.

**Term (27):** Using a decomposition similar to (19), we define:

$$T_4^N := -\frac{C}{2} \frac{\eta_n^f(G_n)\eta_n^c(G_n) - \eta_n^{N,f}(G_n)\eta_n^{N,c}(G_n)}{\eta_n^f(G_n)\eta_n^c(G_n)\eta_n^{N,f}(G_n)\eta_n^{N,c}(G_n)} (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left( \{|\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)|\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right) \bar{M}_{n+1}(\psi) \right) \quad (28)$$

$$T_5^N := -\frac{C}{2} \frac{1}{\eta_n^{N,f}(G_n)\eta_n^{N,c}(G_n)} (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left( \{|\eta_n^{N,c}(G_n)(G_n \otimes 1) - \eta_n^{N,f}(G_n)(1 \otimes G_n)| - |\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)|\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right) \bar{M}_{n+1}(\psi) \right) \quad (29)$$

then we have

$$T_3^N = -T_4^N - T_5^N.$$

We will show that (28) and (29) will converge to zero in probability. For (28), by [21, Proposition C.6]

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

and

$$\left| (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left( \{|\eta_n^c(G_n)(G_n \otimes 1) - \eta_n^f(G_n)(1 \otimes G_n)|\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right) \bar{M}_{n+1}(\psi) \right) \right| \leq \|G_n\| \|\psi\| (\eta_n^c(G_n) + \eta_n^f(G_n)) \check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|)$$

thus, using the argument for (23) in the proof of Lemma C.2 we have shown that (28) converges in probability to zero. For (29) as  $(\eta_n^{N,f}(G_n)\eta_n^{N,c}(G_n))^{-1}$  converges in probability to finite constant, we need only show that the remaining term converges in probability to zero. Using (20) we have that

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

Then (29) converges in probability to zero by using the argument for (23) in the proof of Lemma C.2 for  $\check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|)$ . For the remaining term one can apply the (last) argument for (18) in Lemma C.1. Hence (27) converges in probability to zero and the proof is concluded.  $\square$

**Lemma C.4.** Assume (A1). Then if for any  $\varphi \in \mathcal{C}_b(\mathbb{X} \times \mathbb{X})$ ,  $n \geq 0$

$$\check{\eta}_n^{N,M}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^M(\varphi)$$

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

$$\frac{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}) - \check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})}{(1 - \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}))(1 - \check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}))} \times (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left( \left\{ F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c} \right\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right) \bar{M}_{n+1}(\psi) \right) \rightarrow_{\mathbb{P}} 0.$$

*Proof.* The proof concerning (22) of Lemma C.2 establishes that

$$\frac{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}) - \check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})}{(1 - \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}))(1 - \check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}))} \rightarrow_{\mathbb{P}} 0.$$

Then, almost surely

$$\begin{aligned} & \left| (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left( \left\{ F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c} \right\} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right) \bar{M}_{n+1}(\psi) \right) \right| \leq \\ & \quad \|\psi\| \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}) \check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|). \end{aligned}$$

By Lemma C.1  $\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})$  converges in probability to a constant and by using the argument for (23) in the proof of Lemma C.2 for  $\check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|)$  we conclude the proof.  $\square$

**Lemma C.5.** *Assume (A3). Then if for any  $n \geq 0$ ,  $\varphi \in \mathcal{C}_b(\mathbf{X} \times \mathbf{X})$*

$$\check{\eta}_n^{N,M}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^M(\varphi)$$

*we have for any  $\psi \in \mathcal{C}_b(\mathbf{X}^2 \times \mathbf{X}^2)$*

$$\left( \check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M} \right) (\psi) \rightarrow_{\mathbb{P}} \left( \check{\eta}_n^M \otimes \check{\eta}_n^M \right) (\psi).$$

*Proof.* We will use a density argument. Define

$$\mathcal{F} = \{f : \mathbf{X}^2 \times \mathbf{X}^2 \rightarrow \mathbb{R} : f(x, y) = g(x) \otimes h(y), (g, h) \in \mathcal{C}_b(\mathbf{X} \times \mathbf{X}) \times \mathcal{C}_b(\mathbf{X} \times \mathbf{X})\}.$$

Denote  $\mathcal{G}$  as the set of functions which are finite, linear combinations of functions in  $\mathcal{F}$ . As we have that for some  $n \geq 0$ ,  $\varphi \in \mathcal{C}_b(\mathbf{X} \times \mathbf{X})$ ,  $\check{\eta}_n^{N,M}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^M(\varphi)$ , it follows that for  $\tilde{\psi} \in \mathcal{G}$

$$\left( \check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M} \right) (\tilde{\psi}) \rightarrow_{\mathbb{P}} \left( \check{\eta}_n^M \otimes \check{\eta}_n^M \right) (\tilde{\psi})$$

and thus, by [3, Theorem 25.12]

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \left\{ \left( \check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M} \right) - \left( \check{\eta}_n^M \otimes \check{\eta}_n^M \right) \right\} (\tilde{\psi}) \right| \right] = 0. \quad (30)$$

Let  $\epsilon > 0$  be arbitrary and  $\psi \in \mathcal{C}_b(\mathbf{X}^2 \times \mathbf{X}^2)$ . By the Stone-Weierstrass theorem,  $\mathcal{G}$  is dense in  $\mathcal{C}_b(\mathbf{X}^2 \times \mathbf{X}^2)$ , hence there exist a  $\tilde{\psi} \in \mathcal{G}$  such that

$$\|\psi - \tilde{\psi}\| < \epsilon/3. \quad (31)$$

Then we have for any  $N \geq 1$

$$\begin{aligned} & \mathbb{E} \left[ \left| \left\{ \left( \check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M} \right) - \left( \check{\eta}_n^M \otimes \check{\eta}_n^M \right) \right\} (\psi) \right| \right] \leq \\ & \mathbb{E} \left[ \left( \check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M} \right) (|\psi - \tilde{\psi}|) \right] + \left( \check{\eta}_n^M \otimes \check{\eta}_n^M \right) (|\psi - \tilde{\psi}|) + \mathbb{E} \left[ \left| \left\{ \left( \check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M} \right) - \left( \check{\eta}_n^M \otimes \check{\eta}_n^M \right) \right\} (\tilde{\psi}) \right| \right]. \end{aligned}$$

By (30) there exist a  $N^* \geq 1$  such that for every  $N \geq N^*$

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

Hence applying (31) to the terms

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

we have for every  $N \geq N^*$

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

and as  $\epsilon > 0$  is arbitrary we have

$$\left( \check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M} \right) (\psi) \rightarrow_{\mathbb{P}} \left( \check{\eta}_n^M \otimes \check{\eta}_n^M \right) (\psi)$$

as was to be proved.  $\square$



**Lemma C.6.** Assume (A1-3). Then if for any  $\varphi \in \mathcal{C}_b(\mathbb{X} \times \mathbb{X})$ ,  $n \geq 0$

$$\check{\eta}_n^{N,M}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^M(\varphi)$$

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

$$\begin{aligned} & \left(1 - \check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}\} \right)\right) (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \bar{F}_{n,\check{\eta}_n^{N,M},f} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \rightarrow_{\mathbb{P}} \\ & \left(1 - \check{\eta}_n^M \left( \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \right)\right) (\check{\eta}_n^M \otimes \check{\eta}_n^M) \left( \left\{ \bar{F}_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^M,c} \right\} \bar{M}_{n+1}(\psi) \right). \end{aligned}$$

*Proof.* We make the definitions

$$\begin{aligned} T_1^N & := \left( \check{\eta}_n^M \left( \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \right) - \check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}\} \right) \right) \times \\ & \quad (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \bar{F}_{n,\check{\eta}_n^{N,M},f} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \end{aligned} \quad (32)$$

$$\begin{aligned} T_2^N & := \left( 1 - \check{\eta}_n^M \left( \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \right) \right) (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \bar{F}_{n,\check{\eta}_n^{N,M},f} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} - \right. \right. \\ & \quad \left. \left. \bar{F}_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^M,c} \right\} \bar{M}_{n+1}(\psi) \right) \end{aligned} \quad (33)$$

$$T_3^N = \left[ (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) - (\check{\eta}_n^M \otimes \check{\eta}_n^M) \right] \left( \left\{ \bar{F}_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^M,c} \right\} \bar{M}_{n+1}(\psi) \right) \quad (34)$$

Then we have

$$\begin{aligned} T_1^N + T_2^N + T_3^N & = \left( 1 - \check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}\} \right) \right) (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \bar{F}_{n,\check{\eta}_n^{N,M},f} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \times \right. \\ & \quad \left. \bar{M}_{n+1}(\psi) \right) - \left( 1 - \check{\eta}_n^M \left( \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \right) \right) (\check{\eta}_n^M \otimes \check{\eta}_n^M) \left( \left\{ \bar{F}_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^M,c} \right\} \bar{M}_{n+1}(\psi) \right) \end{aligned}$$

We will show that (32)-(34) will converge in probability to zero.

**Term (32):** By Lemma C.1

$$\left| \check{\eta}_n^M \left( \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \right) - \check{\eta}_n^{N,M} \left( \{F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}\} \right) \right| \rightarrow_{\mathbb{P}} 0$$

so if we can show that

$$T_1^N := \left| (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \bar{F}_{n,\check{\eta}_n^{N,M},f} \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \right|$$

is (almost-surely) upper-bounded by a term which converges in probability to a positive constant, we have shown that that (32) converges in probability to zero. Clearly, almost-surely

$$T_1^N \leq \|\psi\| \check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},f}|) \check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|)$$

we focus on showing that  $\check{\eta}_n^{N,M}(|\bar{F}_{n,\check{\eta}_n^{N,M},c}|)$  is upper-bounded by a term which converges in probability to a finite constant. This can be verified in an almost identical manner to the approach for (24) in the proof of Lemma C.2; hence we have verified that (32) converges in probability to zero.

**Term (33):** Set

$$\begin{aligned} C & := \check{\eta}_n^M \left( \{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}\} \right) \\ T_4^N & := C (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ [\bar{F}_{n,\check{\eta}_n^{N,M},f} - \bar{F}_{n,\check{\eta}_n^M,f}] \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \end{aligned} \quad (35)$$

$$T_5^N := C (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \bar{F}_{n,\check{\eta}_n^M,f} \otimes [\bar{F}_{n,\check{\eta}_n^{N,M},c} - \bar{F}_{n,\check{\eta}_n^M,c}] \right\} \bar{M}_{n+1}(\psi) \right) \quad (36)$$

then

$$T_2^N = T_4^N + T_5^N.$$

We need to show that (35) and (36) converge in probability to zero. As the proof for (36) is similar and easier than that for (35), we focus on the latter. Define

$$\begin{aligned}
T_6^N &:= C(\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \left( \frac{F_{n,\check{\eta}_n^{N,M},f}}{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c})} \right) \otimes \right. \right. \\
&\quad \left. \left. \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \\
T_7^N &:= C(\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \left( \frac{F_{n,\check{\eta}_n^M,f}}{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})} \right) \otimes \right. \right. \\
&\quad \left. \left. \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \\
T_8^N &:= C(\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \left( \frac{F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c} - F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}}{\check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} - F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})} \right) \otimes \right. \right. \\
&\quad \left. \left. \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right) \\
T_9^N &:= \frac{\check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}) - \check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c})}{(1 - \check{\eta}_n^{N,M}(F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c}))(1 - \check{\eta}_n^M(F_{n,\check{\eta}_n^M,f} \wedge F_{n,\check{\eta}_n^M,c}))} \times \\
&\quad C(\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) \left( \left\{ \left( F_{n,\check{\eta}_n^{N,M},f} \wedge F_{n,\check{\eta}_n^{N,M},c} \right) \otimes \bar{F}_{n,\check{\eta}_n^{N,M},c} \right\} \bar{M}_{n+1}(\psi) \right)
\end{aligned}$$

Then we have that

$$T_4^N = T_6^N - T_7^N + T_8^N - T_9^N.$$

By Lemma C.2  $T_6^N - T_7^N \rightarrow_{\mathbb{P}} 0$ , by Lemma C.3  $T_8^N \rightarrow_{\mathbb{P}} 0$  and by Lemma C.4  $T_9^N \rightarrow_{\mathbb{P}} 0$ , hence we conclude that  $T_4^N \rightarrow_{\mathbb{P}} 0$ .

**Term (34):** As

$$\left\{ \bar{F}_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^M,c} \right\} \bar{M}_{n+1}(\psi) \in \mathcal{C}_b(\mathbb{X}^2 \times \mathbb{X}^2)$$

it easily follows that by Lemma C.5

$$\left[ (\check{\eta}_n^{N,M} \otimes \check{\eta}_n^{N,M}) - (\check{\eta}_n^M \otimes \check{\eta}_n^M) \right] \left( \left\{ \bar{F}_{n,\check{\eta}_n^M,f} \otimes \bar{F}_{n,\check{\eta}_n^M,c} \right\} \bar{M}_{n+1}(\psi) \right) \rightarrow_{\mathbb{P}} 0.$$

□

## 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*

$$\mathbb{E}[|\eta_n^{N,s} - \eta_n^s(\varphi)|^p]^{1/p} \leq \frac{C\|\varphi\|}{\sqrt{N}}.$$

*Proof.* As for Proposition B.1. □

Recall that we are assuming that  $M_n^s$  has a density and we are denoting the density also as  $M_n^s$ .

**Theorem D.1.** *Suppose that for  $n \geq 1$ ,  $s \in \{f, c\}$ ,  $M_n^s \in \mathcal{B}_b(\mathbb{X} \times \mathbb{X})$ . Then for any  $\varphi \in \mathcal{B}_b(\mathbb{X} \times \mathbb{X})$ ,  $n \geq 0$  we have*

$$\check{\eta}_n^{N,C}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^C(\varphi).$$

*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

$$\check{\eta}_n^{N,C}(\varphi) - \check{\eta}_n^C(\varphi) = \check{\eta}_n^{N,C}(\varphi) - \check{\Phi}_p^C(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\varphi) + \check{\Phi}_p^C(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\varphi) - \check{\eta}_n^C(\varphi).$$

As in the proof of Theorem 3.1, one can easily prove that

$$|\check{\eta}_n^{N,C}(\varphi) - \check{\Phi}_p^C(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\varphi)| \rightarrow_{\mathbb{P}} 0$$

by using the (conditional) Marcinkiewicz-Zygmund inequality, so we focus on

$$\check{\Phi}_p^C(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\varphi) - \check{\eta}_n^C(\varphi).$$

Define:

$$\begin{aligned} T_1^N &:= \int_{\mathbf{X}} \varphi(x, x) F_{n-1, \eta_{n-1}^{N,f}, f}(x) \wedge F_{n, \eta_{n-1}^{N,c}, c}(x) dx - \int_{\mathbf{X}} \varphi(x, x) F_{n-1, \eta_{n-1}^f, f}(x) \wedge F_{n-1, \eta_{n-1}^c, c}(x) dx \\ T_2^N &:= \left(1 - \int_{\mathbf{X}} F_{n-1, \eta_{n-1}^{N,f}, f}(x) \wedge F_{n-1, \eta_{n-1}^{N,c}, c}(x) dx\right)^{-1} \int_{\mathbf{X} \times \mathbf{X}} \varphi(x, y) \overline{F}_{n-1, \eta_{n-1}^{N,f}, \eta_{n-1}^{N,c}, f}(x) \times \\ &\quad \overline{F}_{n-1, \eta_{n-1}^{N,c}, \eta_{n-1}^{N,f}, c}(x) dx dy - \left(1 - \int_{\mathbf{X}} F_{n-1, \eta_{n-1}^f, f}(x) \wedge F_{n-1, \eta_{n-1}^c, c}(x) dx\right)^{-1} \times \\ &\quad \int_{\mathbf{X} \times \mathbf{X}} \varphi(x, y) \overline{F}_{n-1, \eta_{n-1}^f, \eta_{n-1}^c, f}(x) \overline{F}_{n-1, \eta_{n-1}^c, \eta_{n-1}^f, c}(x) dx dy \end{aligned}$$

Recalling that  $\check{\eta}_n^C(\varphi) = \check{\Phi}_n^C(\eta_{n-1}^f, \eta_{n-1}^c)(\varphi)$  we have that

$$T_1^N + T_2^N = \check{\Phi}_p^C(\eta_{p-1}^{N,f}, \eta_{p-1}^{N,c})(\varphi) - \check{\eta}_n^C(\varphi).$$

By Lemma D.1 1.  $T_1^N \rightarrow_{\mathbb{P}} 0$  and by Lemma D.1 2.  $T_2^N \rightarrow_{\mathbb{P}} 0$  which concludes the proof.  $\square$

**Lemma D.1.** *Suppose that for  $n \geq 1$ ,  $s \in \{f, c\}$ ,  $M_n^s \in \mathcal{B}_b(\mathbf{X} \times \mathbf{X})$  and that for  $\varphi \in \mathcal{B}_b(\mathbf{X} \times \mathbf{X})$ ,  $n \geq 0$*

$$\check{\eta}_n^{N,C}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^C(\varphi).$$

Then for any  $\psi \in \mathcal{B}_b(\mathbf{X} \times \mathbf{X})$ ,  $n \geq 0$  we have

1. 
$$\int_{\mathbf{X}} \psi(x, x) F_{n, \eta_n^{N,f}, f}(x) \wedge F_{n, \eta_n^{N,c}, c}(x) dx \rightarrow_{\mathbb{P}} \int_{\mathbf{X}} \psi(x, x) F_{n, \eta_n^f, f}(x) \wedge F_{n, \eta_n^c, c}(x) dx.$$
2. 
$$\begin{aligned} &\left(1 - \int_{\mathbf{X}} F_{n, \eta_n^{N,f}, f}(x) \wedge F_{n, \eta_n^{N,c}, c}(x) dx\right)^{-1} \int_{\mathbf{X} \times \mathbf{X}} \psi(x, y) \overline{F}_{n, \eta_n^{N,f}, \eta_n^{N,c}, f}(x) \overline{F}_{n, \eta_n^{N,c}, \eta_n^{N,f}, c}(x) dx dy \rightarrow_{\mathbb{P}} \\ &\left(1 - \int_{\mathbf{X}} F_{n, \eta_n^f, f}(x) \wedge F_{n, \eta_n^c, c}(x) dx\right)^{-1} \int_{\mathbf{X} \times \mathbf{X}} \psi(x, y) \overline{F}_{n, \eta_n^f, \eta_n^c, f}(x) \overline{F}_{n, \eta_n^c, \eta_n^f, c}(x) dx dy. \end{aligned}$$

*Proof.* We begin by noting that for any  $x \in \mathbf{X}$ , by the assumption that  $\check{\eta}_n^{N,C}(\varphi) \rightarrow_{\mathbb{P}} \check{\eta}_n^C(\varphi)$ , we have

$$\begin{aligned} F_{n, \eta_n^{N,f}, f}(x) &\rightarrow_{\mathbb{P}} F_{n, \eta_n^f, f}(x) \\ F_{n, \eta_n^{N,c}, c}(x) &\rightarrow_{\mathbb{P}} F_{n, \eta_n^c, c}(x) \end{aligned}$$

and hence that

$$F_{n, \eta_n^{N,f}, f}(x) \wedge F_{n, \eta_n^{N,c}, c}(x) \rightarrow_{\mathbb{P}} F_{n, \eta_n^f, f}(x) \wedge F_{n, \eta_n^c, c}(x) \quad (37)$$

$$\overline{F}_{n, \eta_n^{N,f}, \eta_n^{N,c}, f}(x) \rightarrow_{\mathbb{P}} \overline{F}_{n, \eta_n^f, \eta_n^c, f}(x) \quad (38)$$

$$\overline{F}_{n, \eta_n^{N,c}, \eta_n^{N,f}, c}(x) \rightarrow_{\mathbb{P}} \overline{F}_{n, \eta_n^c, \eta_n^f, c}(x). \quad (39)$$

For 1. we have

$$\begin{aligned} &\mathbb{E} \left[ \left| \int_{\mathbf{X}} \psi(x, x) \left\{ F_{n, \eta_n^{N,f}, f}(x) \wedge F_{n, \eta_n^{N,c}, c}(x) - F_{n, \eta_n^f, f}(x) \wedge F_{n, \eta_n^c, c}(x) \right\} dx \right| \right] \leq \\ &\|\psi\| \int_{\mathbf{X}} \left\{ \mathbb{E} \left[ \left| F_{n, \eta_n^{N,f}, f}(x) \wedge F_{n, \eta_n^{N,c}, c}(x) - F_{n, \eta_n^f, f}(x) \wedge F_{n, \eta_n^c, c}(x) \right| \right] \right\} dx. \end{aligned}$$

Then as  $F_{n, \eta_n^{N,f}, f}(x) \wedge F_{n, \eta_n^{N,c}, c}(x) \leq \|M_{n+1}^f\| \wedge \|M_{n+1}^c\|$ , by (37) and [3, Theorem 25.12] for any  $x \in \mathbf{X}$

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| F_{n, \eta_{n-1}^{N,f}, f}(x) \wedge F_{n, \eta_{n-1}^{N,c}, c}(x) - F_{n, \eta_{n-1}^f, f}(x) \wedge F_{n-1, \eta_{n-1}^c, c}(x) \right| \right] = 0.$$

Hence by the bounded convergence theorem

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[ \left| \int_{\mathcal{X}} \psi(x, x) \left\{ F_{n, \eta_n^{N, f}}(x) \wedge F_{n, \eta_n^{N, c}}(x) - F_{n, \eta_n^f}(x) \wedge F_{n-1, \eta_n^c}(x) \right\} dx \right| \right] = 0$$

and thus the result follows.

For 2. using almost the same proof as in 1. except using (38)-(39) in place of using (37), we have

$$\begin{aligned} & \int_{\mathcal{X} \times \mathcal{X}} \psi(x, y) \bar{F}_{n, \eta_n^{N, f}, \eta_n^{N, c}}(x) \bar{F}_{n, \eta_n^{N, c}, \eta_n^{N, f}}(x) dx dy \rightarrow_{\mathbb{P}} \\ & \int_{\mathcal{X} \times \mathcal{X}} \psi(x, y) \bar{F}_{n, \eta_n^f, \eta_n^c}(x) \bar{F}_{n, \eta_n^c, \eta_n^f}(x) dx dy. \end{aligned} \quad (40)$$

Then define

$$\begin{aligned} T_1^N & := \left( 1 - \int_{\mathcal{X}} F_{n, \eta_n^{N, f}}(x) \wedge F_{n, \eta_n^{N, c}}(x) dx \right)^{-1} - \left( 1 - \int_{\mathcal{X}} F_{n, \eta_n^f}(x) \wedge F_{n, \eta_n^c}(x) dx \right)^{-1} \times \\ & \quad \int_{\mathcal{X} \times \mathcal{X}} \psi(x, y) \bar{F}_{n, \eta_n^{N, f}, \eta_n^{N, c}}(x) \bar{F}_{n, \eta_n^{N, c}, \eta_n^{N, f}}(x) dx dy \\ T_2^N & := \left( 1 - \int_{\mathcal{X}} F_{n, \eta_n^f}(x) \wedge F_{n, \eta_n^c}(x) dx \right)^{-1} \left( \int_{\mathcal{X} \times \mathcal{X}} \psi(x, y) \bar{F}_{n, \eta_n^{N, f}, \eta_n^{N, c}}(x) \times \right. \\ & \quad \left. \bar{F}_{n, \eta_n^{N, c}, \eta_n^{N, f}}(x) dx dy - \int_{\mathcal{X} \times \mathcal{X}} \psi(x, y) \bar{F}_{n, \eta_n^f, \eta_n^c}(x) \bar{F}_{n, \eta_n^c, \eta_n^f}(x) dx dy \right). \end{aligned}$$

Then we have

$$\begin{aligned} & T_1^N + T_2^N = \\ & \left( 1 - \int_{\mathcal{X}} F_{n, \eta_n^{N, f}}(x) \wedge F_{n, \eta_n^{N, c}}(x) dx \right)^{-1} \int_{\mathcal{X} \times \mathcal{X}} \psi(x, y) \bar{F}_{n, \eta_n^{N, f}, \eta_n^{N, c}}(x) F_{n, \eta_n^{N, c}, \eta_n^{N, f}}(x) dx dy - \\ & \left( 1 - \int_{\mathcal{X}} F_{n, \eta_n^f}(x) \wedge F_{n, \eta_n^c}(x) dx \right)^{-1} \int_{\mathcal{X} \times \mathcal{X}} \psi(x, y) \bar{F}_{n, \eta_n^f, \eta_n^c}(x) F_{n, \eta_n^c, \eta_n^f}(x) dx dy. \end{aligned}$$

By 1. and (40),  $T_1^N \rightarrow_{\mathbb{P}} 0$  and by (40)  $T_2^N \rightarrow_{\mathbb{P}} 0$  and hence the proof is complete.  $\square$

## E Proofs for the Asymptotic Variance

Throughout  $v$  is as in (H1). We will frequently quote results from [29] which are proved when one considers  $v$  not  $v^\xi$ ,  $\xi \in (0, 1]$ . The case  $\xi \in (0, 1)$  for these quoted results apply using [29, Lemma 3]. Note that [29, Lemma 3] implies that (H1-4) 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 does not depend on  $n, p, f, c$ . If there are any important dependencies of  $C$  they will be stated.

*Proof of Theorem 4.3.* We have

$$\sigma_n^{2, s}(\varphi) = \check{\eta}_n^s \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^f - \eta_n^c](\varphi)))^2 \right) + \sum_{p=0}^{n-1} \check{\eta}_p^s \left( \{(D_{p, n}^f(\varphi) \otimes 1 - 1 \otimes D_{p, n}^c(\varphi))\}^2 \right).$$

Define

$$T_1 := 2\eta_p^f \left( (D_{p, n}^f(\varphi) - D_{p, n}^c(\varphi))^2 \right) \quad (41)$$

$$T_2 := 2\check{\eta}_p^s \left( (D_{p, n}^c(\varphi) \otimes 1 - 1 \otimes D_{p, n}^c(\varphi))^2 \right) \quad (42)$$

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

$$T_1 + T_2 \geq \check{\eta}_p^s \left( \{(D_{p, n}^f(\varphi) \otimes 1 - 1 \otimes D_{p, n}^c(\varphi))\}^2 \right) \quad (43)$$

Our proof for the case when only (H1-6) holds, versus when (H1-7) holds is the same when considering  $T_1$  in (41), versus  $T_2$  in (42). Thus our proof is split into three parts of controlling  $T_1$  and then  $T_2$  first under (H1-6) and then

under (H1-7). The proof is concluded by adding and then summing the bounds (and then summing over  $p$ ) for the relevant case.

**Term (41):** We have for any  $x \in X$

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

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

$$(D_{p,n}^f(\varphi)(x) - D_{p,n}^c(\varphi)(x))^2 \leq \\ C\|\varphi\|\rho^{n-p} \left( \|h_{p,n}^f S_{p,n}^{f,c}(\varphi)\|_{v^\xi} + |[\eta_n^f - \eta_n^c](\varphi)| + \|\varphi\| \|h_{p,n}^f - h_{p,n}^c\|_{v^\xi} \rho^{n-p} \right) v(x)^{4\xi}.$$

Hence, as  $\eta_p^f(v^{4\xi}) \leq C$  [29, Proposition 1] (the bound in that Proposition does not depend on  $\eta_0^f$  by (H2) and (H5) 2.), it follows that

$$|T_1| \leq C\|\varphi\|\rho^{n-p} \left\{ \|h_{p,n}^f S_{p,n}^{f,c}(\varphi)\|_{v^\xi} + |[\eta_n^f - \eta_n^c](\varphi)| + \|\varphi\| \|h_{p,n}^f - h_{p,n}^c\|_{v^\xi} \rho^{n-p} \right\}. \quad (44)$$

**Term (42), case only (H1-6) hold:** We have for any  $(x, y) \in A$  (if  $(x, y) \in A^c$  the term is zero)

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

Applying Lemmata E.1 and E.6 yields the upper-bound for any  $(x, y) \in A$ :

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

Hence it follows that

$$|T_2| \leq C\|\varphi\|^2 \rho^{n-p} \left( \tilde{\eta}_p^s(\mathbb{I}_A(v \otimes v)^{2\xi})(\rho^{n-p} + 1) \right). \quad (45)$$

Noting (43), (44) and (45) the proof can be concluded in the case that only (H1-6) hold.

**Term (42), case (H1-7) hold:** We have by Lemma E.8

$$(D_{p,n}^c(\varphi)(x) - D_{p,n}^c(\varphi)(y))^2 \leq C\|\varphi\|^2 (\rho)^{n-p} \mathbf{d}(x, y)^2 v(x)^{4\xi} v(y)^{8\xi}$$

where  $\rho$  is the square of the  $\rho$  in Lemma E.8. Then

$$|T_2| \leq C\|\varphi\|^2 \rho^{n-p} \left( \tilde{\eta}_p^s(\mathbf{d}^2(v^{4\xi} \otimes v^{8\xi}))(\rho^{n-p} + 1) \right). \quad (46)$$

Noting (43), (44) and (46) the proof can be concluded in the case that (H1-7) hold.  $\square$

The following result is essentially [29, Theorem 1].

**Lemma E.1.** *Assume (H1-5). Then for any  $\xi \in (0, 1]$  there exists  $\rho < 1$  and  $C < +\infty$  also depending on the constants in (H1-5) such that for any  $\varphi \in \mathcal{L}_{v^\xi}(X)$ ,  $n \geq 1$  and  $0 \leq p < n$ ,  $x \in X$ ,  $s \in \{f, c\}$  we have*

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

*Proof.* The proof for  $\xi = 1$  is exactly as in [29, Theorem 1], except noting that (H2) and (H5) 2. establish that the constants in [29, Theorem 1] do not depend on  $\eta_0^s$  ( $\mu$  in that paper).  $\square$

**Lemma E.2.** *Assume (H1-6). Then for any  $\xi \in (0, 1]$ ,  $s \in \{f, c\}$  we have*

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

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

**Lemma E.3.** Assume (H1-6). Then for any  $\xi \in (0, 1]$  there exists a  $\rho < 1$  and  $C < +\infty$  also depending on the constants in (H1-6) such that for any  $\varphi \in \mathcal{B}_b(\mathsf{X})$ ,  $n \geq 1$  and  $0 \leq p < n$   $x \in \mathsf{X}$  we have

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

*Proof.* Define

$$T_1 := \left( S_{p,n}^{f,c}(\varphi)(x) + [\eta_n^f - \eta_n^c](\varphi) \right) h_{p,n}^f(x) \quad (47)$$

$$T_2 := D_{p,n}^c(\varphi)(x) \frac{1}{h_{p,n}^c(x)} [h_{p,n}^f(x) - h_{p,n}^c(x)] \quad (48)$$

Then we have

$$T_1 + T_2 = D_{p,n}^f(\varphi)(x) - D_{p,n}^c(\varphi)(x). \quad (49)$$

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

**Term (47):** We have

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

Then by [29, Proposition 2]  $\|h_{p,n}^f\|_{v^\xi}$  is upper-bounded by a finite constant that does not depend on  $p, n, f$ , so we easily have that

$$|T_1| \leq C \left( \|h_{p,n}^f S_{p,n}^{f,c}(\varphi)\|_{v^\xi} + |[\eta_n^f - \eta_n^c](\varphi)| \right) v(x)^{2\xi}. \quad (50)$$

**Term (48):** We have

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

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

$$|T_2| \leq C \|\varphi\| \|h_{p,n}^f - h_{p,n}^c\|_{v^\xi} \rho^{n-p} v(x)^{3\xi}. \quad (51)$$

Noting (49), (50) and (51) the proof can be concluded.  $\square$

Define for  $n \geq 1$ ,  $0 \leq p < n$ ,  $x \in \mathsf{X}$ ,  $s \in \{f, c\}$

$$R_{p+1}^{(n),s}(x, dy) := \frac{1}{Q_{p,n}^s(1)(x)} Q_{p+1,n}^s(1)(y) Q_{p+1}^s(x, dy). \quad (52)$$

**Lemma E.4.** Assume (H1-5), (H7). Then for any  $\xi \in (0, 1]$  there exists a  $C < +\infty$  depending on the constants in (H1-5), (H7), such that for any  $\varphi \in \mathcal{B}_b(\mathsf{X}) \cap \text{Lip}_d(\mathsf{X})$ ,  $(x, y) \in \mathsf{X} \times \mathsf{X}$ ,  $s \in \{f, c\}$  we have

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

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

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

By [29, Proposition 2 (2)],

$$\inf_{p \geq 0} \min_{s \in \{f, c\}} \eta_p^s(G_p) \geq C > 0 \quad (53)$$

for some constant  $C$ . So applying (53) along with (H7) we have

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

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

$$\frac{Q_{p,n}^s(\varphi)(x)}{\eta_p^s(Q_{p,n}^s(1))} - \frac{Q_{p,n}^s(\varphi)(y)}{\eta_p^s(Q_{p,n}^s(1))} = \frac{1}{\eta_p^s(G_p)} \left( \frac{Q_{p+1}^s(Q_{p+1,n}^s(\varphi))(x)}{\eta_{p+1}^s(Q_{p+1,n}^s(1))} - \frac{Q_{p+1}^s(Q_{p+1,n}^s(\varphi))(y)}{\eta_{p+1}^s(Q_{p+1,n}^s(1))} \right)$$

Now we have, for any  $x \in \mathbf{X}$

$$\left| \frac{Q_{p+1,n}^s(\varphi)(x)}{\eta_{p+1}^s(Q_{p+1,n}^s(1))} \right| \frac{1}{v(x)^\xi} \leq \|\varphi\| \|h_{p+1,n}^s\|_{v^\xi} \leq C \|\varphi\|$$

where we have used [29, Proposition 2 (3)] to get to the second inequality. Then applying (53), the induction hypothesis and (H7) we have

$$\begin{aligned} \left| \frac{Q_{p,n}^s(\varphi)(x)}{\eta_p^s(Q_{p,n}^s(1))} - \frac{Q_{p,n}^s(\varphi)(y)}{\eta_p^s(Q_{p,n}^s(1))} \right| &\leq C \left\| \frac{Q_{p+1,n}^s(\varphi)}{\eta_{p+1}^s(Q_{p+1,n}^s(1))} \right\|_{v^\xi} \mathbf{d}(x,y) v(x)^\xi v(y)^\xi \\ &\leq C \|\varphi\| \mathbf{d}(x,y) v(x)^\xi v(y)^\xi. \end{aligned}$$

□

**Lemma E.5.** *Assume (H1-7). Then for any  $\xi \in (0, 1/2)$  there exists a  $C < +\infty$  depending on the constants in (H1-7) such that for any  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X}) \cap \text{Lip}_{v^\xi, \mathbf{d}}(\mathbf{X})$ ,  $n \geq 1$  and  $0 \leq p < n$  ( $x, y \in \mathbf{X} \times \mathbf{X}$ ,  $s \in \{f, c\}$ ) we have*

$$\left| h_{p,n}^s(x) \left( R_{p+1}^{(n),s}(\varphi)(x) - R_{p+1}^{(n),s}(\varphi)(y) \right) \right| \leq C \|\varphi\|_{v^\xi} \mathbf{d}(x,y) v(x)^{2\xi} v(y)^{4\xi}.$$

*Proof.* Define

$$T_1 := \frac{h_{p,n}^s(x)}{Q_{p,n}^s(1)(x)} \left( Q_{p+1}^s(\varphi Q_{p+1,n}^s(1))(x) - Q_{p+1}^s(\varphi Q_{p+1,n}^s(1))(y) \right) \quad (54)$$

$$T_2 := \frac{h_{p,n}^s(x) Q_{p+1}^s(\varphi Q_{p+1,n}^s(1))(y)}{Q_{p,n}^s(1)(x) Q_{p,n}^s(1)(y)} \left( Q_{p,n}^s(1)(y) - Q_{p,n}^s(1)(x) \right) \quad (55)$$

Then we have

$$T_1 + T_2 = h_{p,n}^s(x) \left( R_{p+1}^{(n),s}(\varphi)(x) - R_{p+1}^{(n),s}(\varphi)(y) \right). \quad (56)$$

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

**Term (54):** We have

$$\begin{aligned} &\frac{h_{p,n}^s(x)}{Q_{p,n}^s(1)(x)} \left( Q_{p+1}^s(\varphi Q_{p+1,n}^s(1))(x) - Q_{p+1}^s(\varphi Q_{p+1,n}^s(1))(y) \right) = \\ &\frac{1}{\eta_p^s(G_p)} \left( Q_{p+1}^s(\varphi h_{p+1,n}^s)(x) - Q_{p+1}^s(\varphi h_{p+1,n}^s)(y) \right). \end{aligned}$$

Then by Lemma E.4 and [29, Proposition 2 (3)],

$$h_{p+1,n}^s \in \mathcal{L}_{v^{\tilde{\xi}}}(\mathbf{X}) \cap \text{Lip}_{v^{\tilde{\xi}}, \mathbf{d}}(\mathbf{X}) \quad \forall \tilde{\xi} \in (0, 1] \quad (57)$$

so, applying (53) and (H7) we have

$$|T_1| \leq C \|\varphi h_{p+1,n}^s\|_{v^{2\xi}} \mathbf{d}(x,y) [v(x)v(y)]^{2\xi}$$

By [29, Proposition 2 (3)]

$$\|\varphi h_{p+1,n}^s\|_{v^{2\xi}} \leq C \|\varphi\|_{v^\xi}$$

so

$$|T_1| \leq C \|\varphi\|_{v^\xi} \mathbf{d}(x, y) [v(x)v(y)]^{2\xi}. \quad (58)$$

**Term (55):** We have

$$\begin{aligned} & \frac{h_{p,n}^s(x) Q_{p+1}^s(\varphi Q_{p+1,n}^s(1))(y)}{Q_{p,n}^s(1)(x) Q_{p,n}^s(1)(y)} \left( Q_{p,n}^s(1)(y) - Q_{p,n}^s(1)(x) \right) = \\ & \frac{Q_{p+1}^s(\varphi Q_{p+1,n}^s(1))(y)}{Q_{p,n}^s(1)(y)} \left( Q_{p+1}^s(h_{p+1,n}^s)(y) - Q_{p+1}^s(h_{p+1,n}^s)(x) \right) = \\ & \frac{Q_{p+1}^s(\varphi h_{p+1,n}^s)(y) v(y)^\xi}{\eta_p^s(G_p) h_{p,n}^s(y) v(y)^\xi} \left( Q_{p+1}^s(h_{p+1,n}^s)(y) - Q_{p+1}^s(h_{p+1,n}^s)(x) \right). \end{aligned}$$

Applying (53) and Lemma E.2 gives the upper-bound

$$|T_2| \leq C Q_{p+1}^s(|\varphi| h_{p+1,n}^s)(y) v(y)^\xi \left| Q_{p+1}^s(h_{p+1,n}^s)(y) - Q_{p+1}^s(h_{p+1,n}^s)(x) \right|.$$

Now  $\varphi \in \mathcal{L}_{v^\xi}(\mathsf{X})$  and by [29, Proposition 2 (3)]  $\sup_{n \geq 1} \sup_{0 \leq p \leq n} \max_{s \in \{f, c\}} \|h_{p+1,n}^s\|_{v^\xi} < +\infty$ , then, noting (57) and thus applying (H7)

$$|T_2| \leq C \|\varphi\|_{v^\xi} Q_{p+1}^s(v^{2\xi})(y) v(y)^\xi \|h_{p+1,n}^s\|_{v^\xi} |x - y| [v(x)v(y)]^\xi.$$

Again applying [29, Proposition 2 (3)] and (H1) we have finally that

$$|T_2| \leq C \|\varphi\|_{v^\xi} \mathbf{d}(x, y) v(x)^\xi v(y)^{4\xi}. \quad (59)$$

Noting (56), (58) and (59) the proof can be concluded.  $\square$

Recall  $A = \{(x, y) \in \mathsf{X} \times \mathsf{X} : x \neq y\}$ .

**Lemma E.6.** *Assume (H1-6). Then for any  $\xi \in (0, 1]$  there exists a  $\rho < 1$  and  $C < +\infty$  depending on the constants in (H1-6) such that for any  $\varphi \in \mathcal{B}_b(\mathsf{X})$ ,  $n \geq 1$  and  $0 \leq p < n$ ,  $(x, y) \in \mathsf{X} \times \mathsf{X}$  we have*

$$|D_{p,n}^c(\varphi)(x) - D_{p,n}^c(\varphi)(y)| \leq C \mathbb{1}_A(x, y) \|\varphi\| \left( \rho^{n-p} + 1 \right) v(x)^\xi v(y)^{2\xi}.$$

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

$$T_1 := \left( \frac{Q_{p,n}^c(\varphi)(x)}{Q_{p,n}^c(1)(x)} - \frac{Q_{p,n}^c(\varphi)(y)}{Q_{p,n}^c(1)(y)} \right) h_{p,n}^c(x) \quad (60)$$

$$T_2 := D_{p,n}^c(\varphi)(y) \left( \frac{h_{p,n}^c(x)}{h_{p,n}^c(y)} - 1 \right) \quad (61)$$

then we have

$$T_1 + T_2 = D_{p,n}^c(\varphi)(x) - D_{p,n}^c(\varphi)(y). \quad (62)$$

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

**Term (60):** We have

$$\begin{aligned} & \left( \frac{Q_{p,n}^c(\varphi)(x)}{Q_{p,n}^c(1)(x)} - \frac{Q_{p,n}^c(\varphi)(y)}{Q_{p,n}^c(1)(y)} \right) h_{p,n}^c(x) = \\ & \frac{1}{\eta_p^c(G_p)} \left( Q_{p+1}^c \left( \frac{Q_{p+1,n}^c(\varphi)}{\eta_{p+1}^c(Q_{p+1,n}^c(1))} \right) (x) - Q_{p+1}^c \left( \frac{Q_{p+1,n}^c(\varphi)}{\eta_{p+1}^c(Q_{p+1,n}^c(1))} \right) (y) \right) + \\ & \frac{Q_{p,n}^c(\varphi)(y)}{Q_{p,n}^c(1)(y)} \frac{1}{\eta_p^c(G_p)} \left( Q_{p+1}^c(h_{p+1,n}^c)(y) - Q_{p+1}^c(h_{p+1,n}^c)(x) \right). \end{aligned}$$

Applying (53) and  $\varphi \in \mathcal{B}_b(\mathsf{X})$  we have

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



By [29, Proposition 2 (3)]  $\sup_{n \geq 1} \sup_{0 \leq p \leq n} \max_{s \in \{f, c\}} \|h_{p+1, n}^s\|_{v^\xi} < +\infty$ , then in addition applying (H1) we have

$$|T_1| \leq C \|\varphi\| (v(x)^\xi + v(y)^\xi). \quad (63)$$

**Term (61):** We have

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

Then applying Lemma E.1 for  $D_{p, n}^c(\varphi)(y)$ , Lemma E.2 for  $1/(h_{p, n}^c(y)v(y)^\xi)$  and [29, Proposition 2 (3)] as above, we have

$$|T_2| \leq C \|\varphi\| \rho^{n-p} v(x)^\xi v(y)^{3\xi}. \quad (64)$$

Noting (62), (63) and (64) the proof can be concluded.  $\square$

Define for  $B \in \mathcal{X}$ ,  $x \in \mathbf{X}$ ,  $s \in \{f, c\}$

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

**Lemma E.7.** *Assume (H1-7). Then for any  $\xi \in (0, 1)$  there exists a  $C < +\infty$  depending on the constants in (H1-7) such that for any  $\varphi \in \mathcal{B}_b(\mathbf{X}) \cap \text{Lip}_d(\mathbf{X})$ ,  $(x, y) \in \mathbf{X} \times \mathbf{X}$ ,  $s \in \{f, c\}$  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

$$T_1 = \frac{1}{h_{p, n}^s(x)} \left( \frac{Q_{p, n}^s(\varphi)(x)}{\eta_p^s(Q_{p, n}^s(1))} - \frac{Q_{p, n}^s(\varphi)(y)}{\eta_p^s(Q_{p, n}^s(1))} \right) \quad (65)$$

$$T_2 = \frac{Q_{p, n}^s(\varphi)(y)}{Q_{p, n}^s(1)(y)} \frac{1}{h_{p, n}^s(x)} \left( \frac{Q_{p, n}^s(1)(y)}{\eta_p^s(Q_{p, n}^s(1))} - \frac{Q_{p, n}^s(1)(x)}{\eta_p^s(Q_{p, n}^s(1))} \right) \quad (66)$$

then we have

$$T_1 + T_2 = P_{p, n}^s(\varphi)(x) - P_{p, n}^s(\varphi)(y) \quad (67)$$

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

**Term (65):** We have, for any  $\tilde{\xi} \in (0, 1]$

$$\frac{1}{h_{p, n}^s(x)} \left( \frac{Q_{p, n}^s(\varphi)(x)}{\eta_p^s(Q_{p, n}^s(1))} - \frac{Q_{p, n}^s(\varphi)(y)}{\eta_p^s(Q_{p, n}^s(1))} \right) = \frac{v(x)^{\tilde{\xi}}}{h_{p, n}^s(x)v(x)^{\tilde{\xi}}} \left( \frac{Q_{p, n}^s(\varphi)(x)}{\eta_p^s(Q_{p, n}^s(1))} - \frac{Q_{p, n}^s(\varphi)(y)}{\eta_p^s(Q_{p, n}^s(1))} \right).$$

Then applying Lemmata E.2 and E.7, we have the upper-bound for any  $\tilde{\xi} \in (0, 1]$

$$|T_1| \leq C \|\varphi\| d(x, y) v(x)^{2\tilde{\xi}} v(y)^{\tilde{\xi}}. \quad (68)$$

**Term (66):** We have, for any  $\tilde{\xi} \in (0, 1]$

$$\begin{aligned} & \frac{Q_{p, n}^s(\varphi)(y)}{Q_{p, n}^s(1)(y)} \frac{1}{h_{p, n}^s(x)} \left( \frac{Q_{p, n}^s(1)(y)}{\eta_p^s(Q_{p, n}^s(1))} - \frac{Q_{p, n}^s(1)(x)}{\eta_p^s(Q_{p, n}^s(1))} \right) = \\ & \frac{Q_{p, n}^s(\varphi)(y)}{Q_{p, n}^s(1)(y)} \frac{v(x)^{\tilde{\xi}}}{h_{p, n}^s(x)v(x)^{\tilde{\xi}}} \left( \frac{Q_{p, n}^s(1)(y)}{\eta_p^s(Q_{p, n}^s(1))} - \frac{Q_{p, n}^s(1)(x)}{\eta_p^s(Q_{p, n}^s(1))} \right) \end{aligned}$$

Then applying,  $\varphi \in \mathcal{B}_b(\mathbf{X})$ , Lemmata E.2 and E.7, we have the upper-bound for any  $\tilde{\xi} \in (0, 1]$

$$|T_2| \leq C \|\varphi\| d(x, y) v(x)^{2\tilde{\xi}} v(y)^{\tilde{\xi}}. \quad (69)$$

Noting (67), (68) and (69) (the latter two with  $\tilde{\xi} = \xi/2$ ) the proof can be concluded.  $\square$

**Lemma E.8.** *Assume (H1-7). Then for any  $\xi \in (0, 1/2)$  there exists a  $\rho < 1$  and  $C < +\infty$  depending on the constants in (H1-7) such that for any  $\varphi \in \mathcal{B}_b(\mathsf{X}) \cap \text{Lip}_d(\mathsf{X})$ ,  $n \geq 1$  and  $0 \leq p < n$ ,  $(x, y) \in \mathsf{X} \times \mathsf{X}$  we have*

$$|D_{p,n}^c(\varphi)(x) - D_{p,n}^c(\varphi)(y)| \leq C\|\varphi\|\rho^{n-p}d(x,y)v(x)^{2\xi}v(y)^{4\xi}.$$

*Proof.* The proof uses the same decomposition (62) as in the proof of Lemma E.6. We use the same bound (64) as in the proof of Lemma E.6 and focus on the term  $T_1$  as defined in (60). We note by [8, pp. 133]

$$P_{p,n}^c(\varphi)(x) = R_{p+1}^{(n),c}(P_{p+1,n}^c(\varphi))(x).$$

where  $R_{p+1}^{(n),c}$  is defined in (52).

Now  $T_1$  as in the proof of Lemma E.6, can be written as

$$\begin{aligned} T_1 &= h_{p,n}^c(x) \left( P_{p,n}^c(\varphi)(x) - P_{p,n}^c(\varphi)(y) \right) \\ &= h_{p,n}^c(x) \left( P_{p,n}^c(\varphi - \eta_n^c(\varphi))(x) - P_{p,n}^c(\varphi - \eta_n^c(\varphi))(y) \right) \\ &= h_{p,n}^c(x) \left( R_{p+1}^{(n),s}(P_{p+1,n}^c(\varphi - \eta_n^c(\varphi)))(x) - R_{p+1}^{(n),s}(P_{p+1,n}^c(\varphi - \eta_n^c(\varphi)))(y) \right). \end{aligned}$$

Now, noting Lemma E.7, we apply Lemma E.5 to yield the upper-bound

$$|T_1| \leq C \left\| (P_{p+1,n}^c(\varphi - \eta_n^c(\varphi))) \right\|_{v^\xi} d(x,y)v(x)^{2\xi}v(y)^{4\xi}. \quad (70)$$

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

$$\left\| (P_{p+1,n}^c(\varphi - \eta_n^c(\varphi))) \right\|_{v^\xi} \leq C\|\varphi\|\rho^{n-p}. \quad (71)$$

Then using the above inequality in (70) and noting (63) and (64) the proof can be concluded.  $\square$

## 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.* Follows easily from Propositions F.1, F.2 and F.3.  $\square$

**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  $\varphi \in \mathcal{L}_{v^\xi}(\mathsf{X})$ ,  $1 \leq l \leq L$ ,  $x \in \mathsf{X}$  we have*

$$|M^l(\varphi)(x) - M^{l-1}(\varphi)(x)| \leq C\|\varphi\|_{v^\xi} \Delta_l v(x)^{2\xi}.$$

*Proof.* Define

$$T_1 := \int_{\mathbf{B}_l(x)} \varphi(y) [M^l(x,y) - M(x,y) - \{M^{l-1}(x,y) - M(x,y)\}] dy \quad (72)$$

$$T_2 := \int_{\mathbf{B}_l(x)^c} \varphi(y) [M^l(x,y) + M(x,y) - \{M^{l-1}(x,y) + M(x,y)\}] dy. \quad (73)$$

Then we have

$$T_1 + T_2 = M^l(\varphi)(x) - M^{l-1}(\varphi)(x) \quad (74)$$

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

**Term (72):** We have, by applying (9),

$$\begin{aligned} & \left| \int_{\mathbf{B}_l(x)} \varphi(y) [M^l(x, y) - M(x, y) - \{M^{l-1}(x, y) - M(x, y)\}] dy \right| \leq \\ & C \Delta_l \int_{\mathbf{B}_l(x)} |\varphi(y)| \exp\{-C'\|y-x\|^2\} dy. \end{aligned}$$

Using  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$  and (H8) 2. we have

$$|T_1| \leq C \|\varphi\|_{v^\xi} \Delta_l v(x)^{2\xi}. \quad (75)$$

**Term (73):** We have, by applying (8)

$$\begin{aligned} & \left| \int_{\mathbf{B}_l(x)^c} \varphi(y) [M^l(x, y) + M(x, y) - \{M^{l-1}(x, y) + M(x, y)\}] dy \right| \leq \\ & C \int_{\mathbf{B}_l(x)^c} |\varphi(y)| \exp\{-C'\|y-x\|^2\} dy \end{aligned}$$

Using  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$ , the fact

$$\int_{\mathbf{B}_l(x)^c} v(y)^\xi \exp\{-C'\|y-x\|^2\} dy \leq C \int_{\mathbf{B}_l(x)^c} v(y)^\xi dy$$

along with (H8) and that  $\int_{\mathbf{B}_l(x)^c} dy \leq C \Delta_l$ , we have

$$|T_2| \leq C \|\varphi\|_{v^\xi} \Delta_l v(x)^{2\xi}. \quad (76)$$

Noting (74), (75) and (76) the proof can be concluded.  $\square$

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

**Lemma F.2.** *Assume (H5) 1. (H8). Then for any  $\xi \in (0, 1/2)$  there exists a  $C < +\infty$  depending on the constants in (H5) 1. (H8) such that for any  $\mu \in \mathcal{P}_{v^\xi}(\mathbf{X})$ ,  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$ ,  $p \geq 1$ ,  $1 \leq l \leq L$  we have*

$$|[\Phi_p^l(\mu) - \Phi_p^{l-1}(\mu)](\varphi)| \leq \frac{C \|\varphi\|_{v^\xi} \Delta_l \mu(v^{2\xi})}{\mu(G_{p-1})}.$$

*Proof.* We have for  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$

$$[\Phi_p^l(\mu) - \Phi_p^{l-1}(\mu)](\varphi) = \frac{1}{\mu(G_{p-1})} \mu(G_{p-1} \{M^l(\varphi) - M^{l-1}(\varphi)\}).$$

Applying (H5) 1. along with Lemma F.1 allows one to conclude.  $\square$

**Lemma F.3.** *Assume (H1-5). Then for any  $\xi \in (0, 1]$  there exists a  $C < +\infty$  depending on the constants in (H1-5) such that for any  $1 \leq l \leq L$  we have*

$$\sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q_{p,n}^{l-1}(1)}{\eta_p^l(Q_{p,n}^{l-1}(1))} \right\|_{v^\xi} \leq C \quad (77)$$

$$\sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q_{p,n}^{l-1}(1)}{\Phi_p^{l-1}(\eta_{p-1}^l(Q_{p,n}^{l-1}(1)))} \right\|_{v^\xi} \leq C \quad (78)$$

$$\sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q_{p,n}^{l-1}(1)}{\Phi_p^l(\eta_{p-1}^l(Q_{p,n}^{l-1}(1)))} \right\|_{v^\xi} \leq C \quad (79)$$

and if additionally  $\xi \in (0, 1/2)$  there exists a  $C < +\infty$  depending on the constants in (H1-5) such that for any  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$ ,  $l \geq 1$  we have

$$\sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q_{p,n}^{l-1}(\varphi)}{\eta_p^l(Q_{p,n}^{l-1}(1))} \right\|_{v^{2\xi}} \leq C \|\varphi\|_{v^\xi}. \quad (80)$$

*Proof.* The proof of (77) (resp. (78)) and (80) (resp. (79)) are very similar, so we only give the proofs of (78) and (80).

**Proof of (78):** We have for any  $x \in \mathsf{X}$

$$\frac{Q_{p,n}^{l-1}(1)(x)}{\Phi_p^{l-1}(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))} = \frac{h_{p,n}^{l-1}(x)}{\Phi_p^{l-1}(\eta_{p-1}^l)(h_{p,n}^{l-1})}.$$

Now

$$\begin{aligned} \Phi_p^{l-1}(\eta_{p-1}^l)(h_{p,n}^{l-1}) &= \frac{\eta_{p-1}^l(Q_p^{l-1}(h_{p,n}^{l-1}))}{\eta_{p-1}^l(G_{p-1})} \\ &\geq \frac{\eta_{p-1}^l(\mathbb{I}_{C_d} Q_p^{l-1}(\mathbb{I}_{C_d} h_{p,n}^{l-1}))}{\sup_{n \geq 0} \|G_n\|} \\ &\geq C \eta_{p-1}^l(C_d) \varepsilon_d^- \nu(\mathbb{I}_{C_d} h_{p,n}^{l-1}) \\ &\geq C \nu(C_d) \inf_{n \geq 1} \inf_{0 \leq p \leq n} \inf_{x \in C_d} h_{p,n}^{l-1} \\ &\geq C \end{aligned} \tag{81}$$

where we have applied (H5) 1. and (H3) to go to line 3, [29, pp. 2527] (last displayed equation) to go line 4 and then [29, Lemma 10] to go to the final line. Then using [29, Proposition 2], it thus follows that for any  $x \in \mathsf{X}$

$$\left| \frac{Q_{p,n}^{l-1}(1)(x)}{\Phi_p^{l-1}(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))} \right| \leq C \nu(x)^\xi$$

which completes the proof of (78).

**Proof of (80):** We have for any  $x \in \mathsf{X}$

$$\frac{Q_{p,n}^{l-1}(\varphi)(x)}{\Phi_p^l(Q_{p,n}^{l-1}(1))} = \frac{Q_{p,n}^{l-1}(\varphi)(x)}{Q_{p,n}^{l-1}(1)(x)} \frac{Q_{p,n}^{l-1}(1)(x)}{\eta_p^{l-1}(Q_{p,n}^{l-1}(1))} \frac{1}{\eta_p^l(h_{p,n}^{l-1})}$$

By using a very similar proof to (81), one can deduce that  $\eta_p^l(h_{p,n}^{l-1}) \geq C$ , as frequently noted  $\|h_{p,n}^{l-1}\|_{v^\xi} \leq C$  and by a similar proof to [29, Proposition 1]  $|Q_{p,n}^{l-1}(\varphi)(x)|/Q_{p,n}^{l-1}(1)(x) \leq C \|\varphi\|_{v^\xi} \nu(x)^\xi$ . Hence one can complete the proof of (80).  $\square$

Let  $\mu \in \mathcal{P}(\mathsf{X})$ ,  $s \in \{l, l-1\}$ ,  $l \geq 1$ ,  $n \geq 1$ ,  $0 \leq p \leq n$ , and denote  $\Phi_{(p,n)}^s(\mu) = \Phi_n^s \circ \dots \circ \Phi_{p+1}^s(\mu)$ , with the convention that if  $n = p$ ,  $\Phi_{(p,n)}^s(\mu) = \mu$  and  $\Phi_0^s(\mu) = \mu$ .

**Lemma F.4.** *Assume (H1-5), (H8). Then for any  $\xi \in (0, 1/4)$  there exists a  $C < +\infty$  depending on the constants in (H1-5) (H8) such that for any  $\varphi \in \mathcal{L}_{v^\xi}(\mathsf{X})$ ,  $n \geq 1$ ,  $1 \leq l \leq L$  we have*

$$|[\Phi_{(0,n)}^{l-1}(\eta_0^l) - \Phi_{(0,n)}^{l-1}(\eta_0^{l-1})](\varphi)| \leq C \|\varphi\|_{v^\xi} \Delta_l \nu(x^*)^{4\xi}.$$

*Proof.* Define

$$T_1 := [\eta_0^l - \eta_0^{l-1}] \left( \frac{Q_{0,n}^{l-1}(\varphi)}{\eta_0^l(Q_{0,n}^{l-1}(1))} \right) \tag{82}$$

$$T_2 := \left( \frac{\eta_0^{l-1}(Q_{0,n}^{l-1}(\varphi))}{\eta_0^{l-1}(Q_{0,n}^{l-1}(1))} \right) [\eta_0^{l-1} - \eta_0^l] \left( \frac{Q_{0,n}^{l-1}(1)}{\eta_0^l(Q_{0,n}^{l-1}(1))} \right). \tag{83}$$

Then we have

$$T_1 + T_2 = \Phi_{(0,n)}^{l-1}(\eta_0^l) - \Phi_{(0,n)}^{l-1}(\eta_0^{l-1})(\varphi). \tag{84}$$

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

**Term (82):** We have by Lemma F.3 (80) and Lemma F.1 that

$$|T_1| \leq C \|\varphi\|_{v^\xi} \Delta_l \nu(x^*)^{4\xi}. \tag{85}$$

**Term (83):** We have

$$\left(\frac{\eta_0^{l-1}(Q_{0,n}^{l-1}(\varphi))}{\eta_0^{l-1}(Q_{0,n}^{l-1}(1))}\right)[\eta_0^{l-1} - \eta_0^l] \left(\frac{Q_{0,n}^l(1)}{\eta_0^l(Q_{0,n}^l(1))}\right) = \eta_n^{l-1}(\varphi)[\eta_0^{l-1} - \eta_0^l] \left(\frac{Q_{0,n}^l(1)}{\eta_0^l(Q_{0,n}^l(1))}\right).$$

Then using  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$  along with [29, Proposition 1] for  $\eta_n^{l-1}(\varphi)$  and Lemma F.3 (77) and Lemma F.1 for the other term on the R.H.S. we have

$$|T_2| \leq C\|\varphi\|_{v^\xi} \Delta_l v(x^*)^{2\xi}. \quad (86)$$

Noting (84), (85) and (86) the proof can be concluded.  $\square$

**Lemma F.5.** *Assume (H1-6), (H8). Then for any  $(\xi, \hat{\xi}) \in (0, 1/8) \times (0, 1/2)$  there exists a  $\rho < 1$  and  $C < +\infty$  depending on the constants in (H1-6) (H8) such that for any  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$ ,  $n \geq 1$ ,  $0 \leq p \leq n$ ,  $1 \leq l \leq L$  we have*

$$|\Phi_{(p,n)}^{l-1}(\Phi_p^l(\eta_{p-1}^l)) - \Phi_{(p,n)}^{l-1}(\Phi_p^{l-1}(\eta_{p-1}^l))(\varphi)| \leq C\|\varphi\|_{v^\xi} \Delta_l \rho^{n-p} v(x^*)^{8\xi+2\hat{\xi}}.$$

*Proof.* Define

$$T_1 := [\Phi_p^l(\eta_{p-1}^l) - \Phi_p^{l-1}(\eta_{p-1}^l)] \left(\frac{Q_{p,n}^{l-1}(1)P_{p,n}^{l-1}(\bar{\varphi}_n)}{\Phi_p^l(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))}\right) \quad (87)$$

$$T_2 := \left(\frac{\Phi_p^{l-1}(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1)P_{p,n}^{l-1}(\bar{\varphi}_n))}{\Phi_p^{l-1}(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))}\right) \times \\ [\Phi_p^{l-1}(\eta_{p-1}^l) - \Phi_p^l(\eta_{p-1}^l)] \left(\frac{Q_{p,n}^{l-1}(1)}{\Phi_p^l(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))}\right) \quad (88)$$

where  $\bar{\varphi}_n = \varphi - \eta_n^{l-1}(\varphi)$ . Then we have

$$T_1 + T_2 = [\Phi_{(p,n)}^{l-1}(\Phi_p^l(\eta_{p-1}^l)) - \Phi_{(p,n)}^{l-1}(\Phi_p^{l-1}(\eta_{p-1}^l))(\varphi)]. \quad (89)$$

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

**Term (87):** We have by (a similar proof to) (71) that

$$\|P_{p,n}^{l-1}(\bar{\varphi}_n)\|_{v^{2\xi}} \leq C\|\varphi\|_{v^\xi} \rho^{n-p} \quad (90)$$

and Lemma F.3 (79)  $\|Q_{p,n}^{l-1}(1)/\Phi_p^l(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))\|_{2\xi} \leq C$ , so by Lemma F.2

$$\left|[\Phi_p^l(\eta_{p-1}^l) - \Phi_p^{l-1}(\eta_{p-1}^l)] \left(\frac{Q_{p,n}^{l-1}(1)P_{p,n}^{l-1}(\bar{\varphi}_n)}{\Phi_p^l(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))}\right)\right| \leq C\|\varphi\|_{v^\xi} \Delta_l \rho^{n-p} \frac{\eta_{p-1}^l(v^{8\xi})}{\eta_{p-1}^l(G_{p-1})}.$$

Applying [29, Proposition 1, Proposition 2 (2)] yields

$$|T_1| \leq C\|\varphi\|_{v^\xi} \Delta_l \rho^{n-p} v(x^*)^{8\xi}. \quad (91)$$

**Term (88):** We have by Lemma F.3 (78)  $\|Q_{p,n}^{l-1}(1)/\Phi_p^{l-1}(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))\|_{2\xi} \leq C$ , so by (90)

$$|T_2| \leq C\|\varphi\|_{v^\xi} \rho^{n-p} \left(\frac{\eta_{p-1}^l(Q_p^{l-1}(v^{4\xi}))}{\eta_{p-1}^l(G_{p-1})}\right) \left|[\Phi_p^{l-1}(\eta_{p-1}^l) - \Phi_p^l(\eta_{p-1}^l)] \left(\frac{Q_{p,n}^{l-1}(1)}{\Phi_p^l(\eta_{p-1}^l)(Q_{p,n}^{l-1}(1))}\right)\right|$$

Applying (H1), [29, Proposition 1, Proposition 2 (2)] to  $\eta_{p-1}^l(Q_p^{l-1}(v^{4\xi}))/\eta_{p-1}^l(G_{p-1})$  and Lemma F.2 along with Lemma F.3 (79) and [29, Proposition 1, Proposition 2 (2)] to the remaining term gives

$$|T_2| \leq C\|\varphi\|_{v^\xi} \Delta_l \rho^{n-p} v(x^*)^{4\xi+2\hat{\xi}}. \quad (92)$$

Noting (89), (91) and (92) the proof can be concluded.  $\square$

**Proposition F.1.** *Assume (H1-6), (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 (H1-6) (H8) such that for any  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$ ,  $n \geq 0$ ,  $1 \leq l \leq L$  we have*

$$|[\eta_n^l - \eta_n^{l-1}](\varphi)| \leq C\|\varphi\|_{v^\xi} \Delta_l v(x^*)^{8\xi+2\hat{\xi}}.$$

*Proof.* The case  $n = 0$  follows by Lemma F.1, so we suppose  $n \geq 1$ . Define

$$T_1 := [\Phi_{(0,n)}^l(\eta_0^l) - \Phi_{(0,n)}^{l-1}(\eta_0^l)](\varphi) \quad (93)$$

$$T_2 := [\Phi_{(0,n)}^{l-1}(\eta_0^l) - \Phi_{(0,n)}^{l-1}(\eta_0^{l-1})](\varphi). \quad (94)$$

Then we have

$$T_1 + T_2 = [\eta_n^l - \eta_n^{l-1}](\varphi). \quad (95)$$

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

**Term (93):** We have the standard telescoping sum identity

$$[\Phi_{(0,n)}^l(\eta_0^l) - \Phi_{(0,n)}^{l-1}(\eta_0^l)](\varphi) = \sum_{p=1}^n \left\{ [\Phi_{(p,n)}^{l-1}(\Phi_p^l(\eta_{p-1}^l)) - \Phi_{(p,n)}^{l-1}(\Phi_p^{l-1}(\eta_{p-1}^l))](\varphi) \right\}. \quad (96)$$

Applying Lemma F.5 gives

$$|T_1| \leq C \|\varphi\|_{v^\xi} \Delta_l v(x^*)^{8\xi+2\hat{\xi}}. \quad (97)$$

**Term (94):** We have by Lemma F.4

$$|T_2| \leq C \|\varphi\|_{v^\xi} \Delta_l v(x^*)^{4\xi}. \quad (98)$$

Noting (95), (97) and (98) the proof can be concluded.  $\square$

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

**Proposition F.2.** Assume (H1-6), (H8). Then for any  $\xi \in (0, 1/4)$  there exists a  $C < +\infty$  depending on the constants in (H1-6) (H8) such that for any  $\varphi \in \mathcal{B}_b(\mathbf{X})$ ,  $n \geq 1$ ,  $0 \leq p \leq n$ ,  $1 \leq l \leq L$  we have

$$\|h_{p,n}^l S_{p,n}^{l-1}(\varphi)\|_{v^\xi} \leq C \|\varphi\| \Delta_l.$$

*Proof.* For any  $x \in \mathbf{X}$ ,  $0 < \kappa < \xi$  one has

$$\frac{h_{p,n}^l(x) S_{p,n}^{l-1}(\varphi)(x)}{v(x)^\xi} = \frac{h_{p,n}^l(x)}{v(x)^\kappa} \frac{1}{v(x)^{\xi-\kappa}} \left( P_{p,n}^l(\varphi)(x) - P_{p,n}^{l-1}(\varphi)(x) \right)$$

and

$$P_{p,n}^l(\varphi)(x) - P_{p,n}^{l-1}(\varphi)(x) = \sum_{q=p+1}^n \left\{ [\Phi_{(q,n)}^{l-1}(\Phi_{(p,q)}^l(\delta_x)) - \Phi_{(q,n)}^{l-1}(\Phi_q^{l-1}\{\Phi_{(p,q-1)}^l(\delta_x)\})](\varphi) \right\}.$$

Noting (99), 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  $(\tilde{\kappa}, \hat{\kappa}) \in (0, 1/2) \times (0, 1]$

$$\left| \frac{1}{v(x)^{\xi-\kappa}} \left( P_{p,n}^l(\varphi)(x) - P_{p,n}^{l-1}(\varphi)(x) \right) \right| \leq C \|\varphi\| \Delta_l \frac{v(x)^{2\tilde{\kappa}+\hat{\kappa}}}{v(x)^{\xi-\kappa}}$$

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

$$\begin{aligned} \left| \frac{1}{v(x)^{\xi-\kappa}} \left( P_{p,n}^l(\varphi)(x) - P_{p,n}^{l-1}(\varphi)(x) \right) \right| &\leq C \|\varphi\| \Delta_l \\ \frac{h_{p,n}^l(x)}{v(x)^\kappa} &\leq C \end{aligned}$$

and the proof is concluded.  $\square$

**Lemma F.6.** Assume (H1-5). Then for any  $\xi \in (0, 1]$  there exists a  $C < +\infty$  depending on the constants in (H1-5) such that for any  $1 \leq l \leq L$  we have

$$\sup_{n \geq 1} \sup_{0 \leq p \leq n} \left\| \frac{Q_{p,n}^{l-1}(1)}{\eta_p^l(Q_{p,n}^l(1))} \right\|_{v^\xi} \leq C.$$

*Proof.* The proof is the same as [29, Lemma 8] as the constants in (H1-5) do not depend upon  $l$ .  $\square$

**Proposition F.3.** Assume (H1-6), (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 (H1-6) (H8) such that for any  $n \geq 1, 0 \leq p \leq n, 1 \leq l \leq L$ , we have

$$\|h_{p,n}^l - h_{p,n}^{l-1}\|_{v^\xi} \leq C(n-p)\Delta_l v(x^*)^{\hat{\xi}}.$$

*Proof.* Define for any  $x \in X$

$$T_1 := \frac{1}{\eta_p^l(Q_{p,n}^l(1))} (Q_{p,n}^l(1)(x) - Q_{p,n}^{l-1}(1)(x)) \quad (99)$$

$$T_2 := \frac{Q_{p,n}^{l-1}(1)(x)}{\eta_p^l(Q_{p,n}^l(1))} [\eta_p^{l-1} - \eta_p^l](h_{p,n}^{l-1}) \quad (100)$$

$$T_3 := h_{p,n}^{l-1}(x) \eta_p^l \left( \frac{1}{\eta_p^l(Q_{p,n}^l(1))} (Q_{p,n}^l(1) - Q_{p,n}^{l-1}(1)) \right). \quad (101)$$

Then we have

$$T_1 + T_2 - T_3 = h_{p,n}^l(x) - h_{p,n}^{l-1}(x). \quad (102)$$

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

**Term (99):** We have

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

Considering the summand we have

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

Now applying Lemma F.1, [29, Proposition 2 (3)] (with  $v^{\xi/4}$ ) and (H5) 1. we have

$$|Q_{p,q}^l(Q_{q,n}^{l-1}(1))(x) - Q_{p,q-1}^l(Q_{q-1,n}^{l-1}(1))(x)| \leq C\Delta_l \eta_{q+1}^{l-1}(Q_{q+1,n}^{l-1}(1)) Q_{p,q-1}^l(v^{\xi/2})(x).$$

Thus

$$|T_1| \leq \frac{C\Delta_l}{\eta_p^l(Q_{p,n}^l(1))} \sum_{q=p+1}^n \left\{ \eta_{q+1}^{l-1}(Q_{q+1,n}^{l-1}(1)) Q_{p,q-1}^l(v^{\xi/2})(x) \right\}. \quad (103)$$

Now,

$$\frac{\eta_{q+1}^{l-1}(Q_{q+1,n}^{l-1}(1)) Q_{p,q-1}^l(v^{\xi/2})(x)}{\eta_p^l(Q_{p,n}^l(1))} = \eta_{q+1}^{l-1} \left( \frac{Q_{q+1,n}^{l-1}(1)}{\eta_q^l(G_q) \eta_{q+1}^l(Q_{q+1,n}^l(1))} \right) \frac{Q_{p,q-1}^l(v^{\xi/2})(x)}{\eta_p^l(Q_{p,q-1}^l(1))} \frac{1}{\eta_{q-1}^l(G_{q-1})}. \quad (104)$$

Applying Lemma F.6, [29, Proposition 2 (3)] and then [29, Proposition 1] gives for any  $\hat{\xi} \in (0, 1]$

$$\eta_{q+1}^{l-1} \left( \frac{Q_{q+1,n}^{l-1}(1)}{\eta_q^l(G_q) \eta_{q+1}^l(Q_{q+1,n}^l(1))} \right) \leq C v(x^*)^{\hat{\xi}}. \quad (105)$$

In addition

$$\frac{Q_{p,q-1}^l(v^{\xi/2})(x)}{\eta_p^l(Q_{p,q-1}^l(1))} = \frac{Q_{p,q-1}^l(v^{\xi/2})(x)}{Q_{p,q-1}^l(1)(x)} h_{p,q-1}^l(x)$$

By a similar proof to [29, Proposition 1]  $Q_{p,q-1}^l(v^{\xi/2})(x)/Q_{p,q-1}^l(1)(x) \leq C v(x)^{\xi/2}$  and thus by [29, Proposition 2 (3)]  $h_{p,q-1}^l(x) \leq C v(x)^{\xi/2}$ , so

$$\frac{Q_{p,q-1}^l(v^{\xi/2})(x)}{\eta_p^l(Q_{p,q-1}^l(1))} \leq C v(x)^\xi. \quad (106)$$

Noting (103)-(106), we have shown that for any  $(\xi, \hat{\xi}) \in (0, 1/8) \times (0, 1/2)$

$$|T_1| \leq C(n-p)\Delta_l v(x)^\xi v(x^*)^{\hat{\xi}}. \quad (107)$$

**Term (100):** We have, by Lemma F.6 and Proposition F.1, for any  $(\xi, \kappa, \hat{\kappa}) \in (0, 1/8)^2 \times (0, 1/2)$

$$|T_2| \leq C v(x)^\xi \|h_{p,n}^{l-1}\|_{v^\kappa} \Delta_l v(x^*)^{8\kappa+2\hat{\kappa}}.$$

One can choose  $0 < \kappa = \hat{\kappa} < 1/20$ ,  $\hat{\xi} = 10\kappa$  and applying [29, Proposition 2 (3)], we have

$$|T_2| \leq C \Delta_l v(x)^\xi v(x^*)^{\hat{\xi}}. \quad (108)$$

**Term (101):** We have by [29, Proposition 2 (3)] and (107), for any  $(\kappa, \hat{\kappa}) \in (0, 1/8) \times (0, 1/2)$

$$|T_3| \leq C v(x)^\xi \times (n-p) \Delta_l \eta_p^l(v^\kappa) v(x^*)^{\hat{\kappa}}.$$

Applying [29, Proposition 1] and choosing  $(\kappa, \hat{\kappa})$ , so that  $\hat{\xi} = \kappa + \hat{\kappa} < 1/2$ , we have

$$|T_3| \leq C(n-p) \Delta_l v(x)^\xi v(x^*)^{\hat{\xi}}. \quad (109)$$

Noting (102), (107), (108) and (109) the proof can be concluded.  $\square$

## F.2 Proofs for Proposition 5.2

*Proof of Proposition 5.2.* We focus on the first bound in Theorem 4.3. Clearly

$$\check{\eta}_n^C \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}](\varphi)))^2 \right) \leq 2 \|\varphi\|^2 \check{\eta}_n^C(\mathbb{I}_A).$$

As  $\check{\eta}_n^C$  is the maximal coupling of  $(\eta_n^l, \eta_n^{l-1})$  we have  $\check{\eta}_n^C(\mathbb{I}_A) = \|\eta_n^l - \eta_n^{l-1}\|_{\text{tv}}$ . Hence on applying Proposition F.1 (noting Remark F.1), we have that

$$\check{\eta}_n^C \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}](\varphi)))^2 \right) \leq C \|\varphi\|^2 \Delta_l v(x^*)^{2\xi+\hat{\xi}}.$$

The proof is completed by using Proposition 5.1 and Proposition F.4.  $\square$

**Proposition F.4.** *Assume (H1-6), (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 (H1-6) (H8) such that for any  $n \geq 0$ ,  $1 \leq l \leq L$  we have*

$$\check{\eta}_n^C(\mathbb{I}_A(v \otimes v)^{2\xi}) \leq C \Delta_l v(x^*)^{32\xi+2\hat{\xi}}.$$

*Proof.* We will use  $(\eta_n^l, \eta_n^{l-1})$  to denote both the probability measure and density associated to  $(\eta_n^l, \eta_n^{l-1})$ . We have for any  $n \geq 0$

$$\begin{aligned} \check{\eta}_n^C(\mathbb{I}_A(v \otimes v)^{2\xi}) &= \left( 1 - \int_{\mathcal{X}} \eta_n^l(x) \wedge \eta_n^{l-1}(x) dx \right) \int_{\mathcal{X} \times \mathcal{X}} \mathbb{I}_A(x, y) v(x)^{2\xi} v(y)^{2\xi} \times \\ &\quad \frac{(\eta_n^l(x) - \eta_n^l(x) \wedge \eta_n^{l-1}(x)) (\eta_n^{l-1}(y) - \eta_n^l(y) \wedge \eta_n^{l-1}(y))}{1 - \int_{\mathcal{X}} \eta_n^l(x) \wedge \eta_n^{l-1}(x) dx} dx dy. \end{aligned}$$

Then by upper-bounding the indicator by 1 and applying Cauchy-Schwarz:

$$\check{\eta}_n^C(\mathbb{I}_A(v \otimes v)^{2\xi}) \leq T_1 T_2 \quad (110)$$

where

$$\begin{aligned} T_1 &:= \left( \int_{\mathcal{X}} v(x)^{4\xi} (\eta_n^l(x) - \eta_n^l(x) \wedge \eta_n^{l-1}(x)) dx \right)^{1/2} \\ T_2 &:= \left( \int_{\mathcal{X}} v(x)^{4\xi} (\eta_n^{l-1}(x) - \eta_n^l(x) \wedge \eta_n^{l-1}(x)) dx \right)^{1/2}. \end{aligned}$$

We now just deal with  $T_1$  as the proof for  $T_2$  is almost identical.



We have, letting  $D_n^l = \{x : \eta_n^l(x) \geq \eta_n^{l-1}(x)\}$ ,  $\tilde{\varphi}(x) = v(x)^{4\xi} \mathbb{I}_{D_n^l}(x)$  and  $\hat{\varphi}(x) = v(x)^{4\xi} \mathbb{I}_{(D_n^l)^c}(x)$

$$\begin{aligned} T_1^2 &= \frac{1}{2} \left( \int_{\mathbf{X}} v(x)^{4\xi} [\eta_n^l(x) - \eta_n^{l-1}(x)] dx + \int_{\mathbf{X}} v(x)^{4\xi} |\eta_n^l(x) - \eta_n^{l-1}(x)| dx \right) \\ &= \frac{1}{2} \left( \int_{\mathbf{X}} v(x)^{4\xi} [\eta_n^l(x) - \eta_n^{l-1}(x)] dx + \int_{\mathbf{X}} \tilde{\varphi}(x) (\eta_n^l(x) - \eta_n^{l-1}(x)) dx + \right. \\ &\quad \left. \int_{\mathbf{X}} \hat{\varphi}(x) (\eta_n^{l-1}(x) - \eta_n^l(x)) dx \right) \\ &\leq \frac{1}{2} (|[\eta_n^l - \eta_n^{l-1}](v^{4\xi})| + |[\eta_n^l - \eta_n^{l-1}](\tilde{\varphi})| + |[\eta_n^l - \eta_n^{l-1}](\hat{\varphi})|) \end{aligned}$$

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

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

Similar calculations give  $T_2 \leq C(\Delta_l)^{1/2} v(x^*)^{16\xi + \hat{\xi}}$  and hence noting (111) one can conclude.  $\square$

### F.3 Proofs for Proposition 5.3

Recall here that  $\mathbf{X} = \mathbb{R}$  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 (H1-6), (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 (H1-6) (H8) such that for any  $\varphi \in \mathcal{L}_{v^\xi}(\mathbf{X})$ ,  $n \geq 0$ ,  $1 \leq l \leq L$  we have*

$$|[\bar{\eta}_n^l - \bar{\eta}_n^{l-1}](\varphi)| \leq C \|\varphi\|_{v^\xi} \Delta_l v(x^*)^{8\xi + 2\hat{\xi}}.$$

*Proof.* Define

$$T_1 := \frac{[\eta_n^l - \eta_n^{l-1}](G_n \varphi)}{\eta_n^l(G_n)} \quad (111)$$

$$T_2 := \frac{\eta_n^{l-1}(G_n \varphi)}{\eta_n^{l-1}(G_n) \eta_n^l(G_n)} [\eta_n^l - \eta_n^{l-1}](G_n). \quad (112)$$

Then we have

$$T_1 - T_2 = [\bar{\eta}_n^l - \bar{\eta}_n^{l-1}](\varphi). \quad (113)$$

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

**Term (111):** By Proposition F.1 and  $\eta_n^l(G_n) \geq C$  we have

$$|T_1| \leq C \|\varphi\|_{v^\xi} \Delta_l v(x^*)^{8\xi + 2\hat{\xi}}. \quad (114)$$

**Term (112):** By Proposition F.1, (H5) and  $\eta_n^s(G_n) \geq C$ ,  $s \in \{l, l-1\}$ , we have

$$|T_2| \leq C \|\varphi\|_{v^\xi} \Delta_l v(x^*)^{8\xi + 2\hat{\xi}}. \quad (115)$$

Noting (113), (114) and (115) the proof can be concluded.  $\square$

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

**Corollary F.1.** *Assume (H1-6), (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 (H1-6) (H8) such that for any  $n \geq 0$ ,  $1 \leq l \leq L$  we have*

$$\check{\eta}_n^C(\mathbb{I}_A(v \otimes v)^\xi) \leq C \Delta_l v(x^*)^{16\xi + 2\hat{\xi}}.$$

*Proof.* 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  $\check{\eta}_n^W$  the optimal Wasserstein coupling of  $(\check{\eta}_n^l, \check{\eta}_n^{l-1})$  that is, for  $B \in \mathcal{X} \vee \mathcal{X}$

$$\check{\eta}_n^W(B) = \int_0^1 \mathbb{I}_B(F_{\check{\eta}_n^l}^{-1}(u), F_{\check{\eta}_n^{l-1}}^{-1}(u)) du.$$

**Lemma F.7.** *Assume (H1-6), (H8). Then if  $\check{\varphi} \in \mathcal{L}_{(v \otimes v)^\xi}(\mathbf{X})$ , for any  $\check{\xi} \in (0, 1/16)$  and set  $\hat{\xi} \in (0, 1/2)$  then there exists a  $C < +\infty$  depending on the constants in (H1-6) (H8) such that for any  $\varphi \in \mathcal{B}_b(\mathbf{X}) \cap \text{Lip}_{d_1}(\mathbf{X})$   $n \geq 0$ ,  $1 \leq l \leq L$  we have*

$$\check{\eta}_n^W \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}] (\varphi)))^2 \right) \leq C \|\varphi\|_{\text{Lip}}^2 \|\check{\varphi}\|_{v^\xi} \Delta_l v(x^*)^{16\check{\xi} + 2\hat{\xi}}.$$

*Proof.* We assume  $n \geq 1$ , the case  $n = 0$  is noted below. As  $\varphi \in \mathcal{B}_b(\mathbf{X}) \cap \text{Lip}_{d_1}(\mathbf{X})$ , it easily follows that

$$\check{\eta}_n^W \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}] (\varphi)))^2 \right) \leq \|\varphi\|_{\text{Lip}}^2 \check{\eta}_{n-1}^W(\check{M}(\check{\varphi})).$$

Applying (6) (when  $p = 2$ ) and using the optimality of the Wasserstein coupling with  $\check{\varphi} \in \mathcal{L}_{(v \otimes v)^\xi}(\mathbf{X})$ , gives

$$\check{\eta}_n^W \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}] (\varphi)))^2 \right) \leq C \|\varphi\|_{\text{Lip}}^2 \|\check{\varphi}\|_{v^\xi} (\Delta_l + \check{\eta}_{n-1}^C(\mathbb{I}_A(v \otimes v)^{\hat{\xi}})).$$

Application of Corollary F.1 yields the desired result. The case  $n = 0$  follows as

$$\check{\eta}_n^W \left( (\varphi \otimes 1 - 1 \otimes \varphi - ([\eta_n^l - \eta_n^{l-1}] (\varphi)))^2 \right) \leq \|\varphi\|_{\text{Lip}}^2 \check{\eta}_n^W(\check{\varphi})$$

and the optimality of the Wasserstein coupling,  $\check{\varphi} \in \mathcal{L}_{(v \otimes v)^\xi}(\mathbf{X})$  with Corollary F.1.  $\square$

**Lemma F.8.** *Assume (H1-6), (H8). Then, let  $\lambda \in (0, 1)$  be given, if  $\check{\varphi} \in \mathcal{L}_{(v \otimes v)^\xi}(\mathbf{X})$ , for any  $\check{\xi} \in (0, 1/(16(1+\lambda)))$  and set  $(\xi, \hat{\xi}) \in (0, \min\{1/32, \lambda/(16(1+\lambda)), (1-2\check{\xi})/12\}) \times (0, 1/2)$  then there exists a  $C < +\infty$  depending on the constants in (H1-6) (H8) such that for any  $n \geq 0$ ,  $1 \leq l \leq L$  we have*

$$\check{\eta}_n^W(\check{\varphi}(v^{4\xi} \otimes v^{8\xi})) \leq C \|\check{\varphi}\|_{v^\xi} (\Delta_l)^{1/(1+\lambda)} v(x^*)^{20\xi + (16(\lambda+1)\check{\xi} + 2\hat{\xi})/(1+\lambda)}.$$

*Proof.* We assume  $n \geq 1$ , the case  $n = 0$  is similar and omitted for brevity. By Cauchy-Schwarz and the structure of  $\check{\eta}_n^W$ :

$$\check{\eta}_n^W(\check{\varphi}(v^{4\xi} \otimes v^{8\xi})) \leq \check{\eta}_{n-1}^W(\check{M}(\check{\varphi}^2)^{1/2} \check{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}).$$

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

$$\check{\eta}_n^W(\check{\varphi}(v^{4\xi} \otimes v^{8\xi})) \leq C \left( \check{\eta}_{n-1}^W(\check{\varphi} \check{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) + \Delta_l \check{\eta}_{n-1}^W(\check{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) \right).$$

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

Now, we have, by Hölder's inequality

$$\check{\eta}_{n-1}^W(\check{\varphi} \check{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) \leq \check{\eta}_{n-1}^W(\check{\varphi}^{1+\lambda})^{1/(1+\lambda)} \check{\eta}_{n-1}^W(\check{M}(v^{8\xi} \otimes v^{16\xi})^{(1+\lambda)/(2\lambda)})^{\lambda/(1+\lambda)}.$$

By the optimality of the Wasserstein coupling with  $\check{\varphi} \in \mathcal{L}_{(v \otimes v)^\xi}(\mathbf{X})$ , we have

$$\begin{aligned} & \check{\eta}_{n-1}^W(\check{\varphi} \check{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) \leq \\ & \left( \|\check{\varphi}\|_{v^\xi} \check{\eta}_{n-1}^C(\mathbb{I}_A(v \otimes v)^{\hat{\xi}(1+\lambda)}) \right)^{1/(1+\lambda)} \check{\eta}_{n-1}^W(\check{M}(v^{8\xi} \otimes v^{16\xi})^{(1+\lambda)/(2\lambda)})^{\lambda/(1+\lambda)} \end{aligned}$$

Then applying Corollary F.1 gives

$$\begin{aligned} & \check{\eta}_{n-1}^W(\check{\varphi} \check{M}(v^{8\xi} \otimes v^{16\xi})^{1/2}) \leq \\ & C \|\check{\varphi}\|_{v^\xi} (\Delta_l)^{1/(1+\lambda)} v(x^*)^{(16(\lambda+1)\check{\xi} + 2\hat{\xi})/(1+\lambda)} \check{\eta}_{n-1}^W(\check{M}(v^{8\xi} \otimes v^{16\xi})^{(1+\lambda)/(2\lambda)})^{\lambda/(1+\lambda)}. \end{aligned}$$

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

$$\begin{aligned} & \check{\eta}_{n-1}^W (\check{M}(v^{8\xi} \otimes v^{16\xi})^{(1+\lambda)/(2\lambda)})^{\lambda/(1+\lambda)} \leq \\ & \check{\eta}_{n-1}^l (M^l(v^{16\xi})^{(1+\lambda)/(2\lambda)})^{\lambda/(2(1+\lambda))} \check{\eta}_{n-1}^{l-1} (M^{l-1}(v^{32\xi})^{(1+\lambda)/(2\lambda)})^{\lambda/(2(1+\lambda))}. \end{aligned}$$

Applying Jensen twice gives

$$\check{\eta}_{n-1}^W (\check{M}(v^{8\xi} \otimes v^{16\xi})^{(1+\lambda)/(2\lambda)})^{\lambda/(1+\lambda)} \leq \eta_n^l(v)^{4\xi} \eta_n^{l-1}(v)^{16\xi} \leq Cv(x^*)^{20\xi}.$$

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 case, however, due to 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  $\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi})$  w.r.t. any coupling of  $(\check{\eta}_n^l, \check{\eta}_n^{l-1})$ . To see this, one can construct a functional covariance equality for  $\mu(\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi}))$  ( $\mu$  is any coupling of  $(\check{\eta}_n^l, \check{\eta}_n^{l-1})$  such that  $\mu(\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi})) < +\infty$ ) as in [23, Theorem 3.1] (that is, in terms of the CDFs of  $\mu$ ,  $\check{\eta}_n^l$  and  $\check{\eta}_n^{l-1}$  as in Hoeffding's Lemma) and then when centering with  $(\check{\eta}_n^l \otimes \check{\eta}_n^{l-1})(\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi}))$  and applying Hoeffding-Fréchet bounds, one observes that  $\check{\eta}_n^W$  is not optimal. As a result, one cannot transfer between  $\check{\eta}_n^W$  and  $\check{\eta}_n^C$  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 exist a  $\check{\eta}_n^W$ -integrable  $\hat{V} : \mathbf{X} \times \mathbf{X} \rightarrow (0, \infty)$  such that there exist a  $C < +\infty$  such that for every  $(x, y) \in \mathbf{X} \times \mathbf{X}$ ,  $\tilde{\varphi}(x, y)v^{4\xi}(x)v^{8\xi}(y) \leq C\hat{V}(x, y)$  and  $\hat{V}$  satisfies the Monge condition (e.g. [26, eq. (3.1.7)]).
2.  $v$  is bounded.
3. There exist a  $C < +\infty$  such that for every  $n \geq 0$ ,  $\check{\eta}_n^W(\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi})) \leq C\check{\eta}_n^C(\tilde{\varphi}(v^{4\xi} \otimes v^{8\xi}))$ .
4. There exist a  $C < +\infty$  such that for every  $n \geq 0$ ,  $1 \leq l \leq L$ ,  $\sup_{u \in [0, 1]} |F_{\check{\eta}_n^l}^{-1}(u) - F_{\check{\eta}_n^{l-1}}^{-1}(u)| \leq C\Delta_l$ .

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

## F.4 Proof of Lemma 5.1

*Proof of Lemma 5.1.* We begin by noting that for any  $B \in \mathcal{X}$ ,  $0 \leq l \leq L$ ,  $y \in \mathbf{X}$  we have

$$\begin{aligned} M^l(\mathbb{I}_B v)(y) &= \sqrt{\left(\frac{1 + \delta_0}{1 + \delta_0 - \beta_l}\right)} \exp\left\{\frac{\alpha_l^2 y^2}{2(1 + \delta_0 - \beta_l)} + 1\right\} \times \\ &\int_B \frac{1}{\sqrt{2\pi\beta_l}} \exp\left\{-\frac{1 + \delta_0 - \beta_l}{2\beta_l(1 + \delta_0)} \left(x - \frac{\alpha_l y(1 + \delta_0)}{1 + \delta_0 - \beta_l}\right)^2\right\} dx. \end{aligned} \quad (116)$$

This will be used below.

We give the proof for  $\xi = 1$  as it is similar in other cases. We have for any  $0 \leq l \leq L$ ,  $n \geq 1$ ,  $\varphi \in \mathcal{L}_v(\mathbf{X}) \cap \text{Lip}_{v, d_1}$

$$\begin{aligned} & |Q_n^l(\varphi)(x) - Q_n^l(\varphi)(y)| \leq \\ & \|\varphi\|_v M^l(v)(x) |G_{n-1}(x) - G_{n-1}(y)| + \|G_{n-1}\| \|\varphi\|_v \int_{\mathbf{X}} v(u) |M^l(x, u) - M^l(y, u)| du. \end{aligned}$$

As noted in Section 5.6,  $[\alpha_l^2(1 + \delta_0)]/[1 + \delta_0 - \beta_l] < 1$  for any  $0 \leq l \leq L$ ,  $\delta_0 > 0$ , so using (116) one can determine that  $M^l(v)(x) \in \mathcal{L}_v(\mathbf{X})$  for any  $\delta_0 > 0$ , with  $\|M^l(v)\|_v$  independent of  $l$ . In addition, for every  $y \in \mathbf{Y}$ ,  $G_{n-1} \in \text{Lip}_{d_1}(\mathbf{X})$  with Lipschitz constant independent of  $n$  so we need only consider the right-hand term in the above displayed equation.

Set  $D_l(x, y) := \{u \in \mathbf{X} : M^l(x, u) - M^l(y, u)\}$ , we consider only  $\int_{D_l(x, y)} v(u) |M^l(x, u) - M^l(y, u)| du$  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 (116) and  $x > y$

$$\int_{D_l(x, y)} v(u) |M^l(x, u) - M^l(y, u)| du = \sqrt{\left(\frac{1 + \delta_0}{1 + \delta_0 - \beta_l}\right)} \times$$

$$\left\{ \exp \left\{ \frac{\alpha_l^2 x^2}{2(1 + \delta_0 - \beta_l)} + 1 \right\} \int_{-\infty}^{c_l(x,y)} \frac{1}{\sqrt{2\pi\beta_l}} \exp \left\{ -\frac{1 + \delta_0 - \beta_l}{2\beta_l(1 + \delta_0)} \left( u - \frac{\alpha_l x(1 + \delta_0)}{1 + \delta_0 - \beta_l} \right)^2 \right\} du - \right. \\ \left. \exp \left\{ \frac{\alpha_l^2 y^2}{2(1 + \delta_0 - \beta_l)} + 1 \right\} \int_{-\infty}^{c_l(x,y)} \frac{1}{\sqrt{2\pi\beta_l}} \exp \left\{ -\frac{1 + \delta_0 - \beta_l}{2\beta_l(1 + \delta_0)} \left( u - \frac{\alpha_l y(1 + \delta_0)}{1 + \delta_0 - \beta_l} \right)^2 \right\} du \right\}$$

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

$$\begin{aligned} \bar{c}_{l,1}(x, y) &:= \sqrt{\left(\frac{1 + \delta_0 - \beta_l}{\beta_l(1 + \delta_0)}\right)} \left( c_l(x, y) - \frac{\alpha_l x(1 + \delta_0)}{1 + \delta_0 - \beta_l} \right) \\ \bar{c}_{l,2}(x, y) &:= \sqrt{\left(\frac{1 + \delta_0 - \beta_l}{\beta_l(1 + \delta_0)}\right)} \left( c_l(x, y) - \frac{\alpha_l y(1 + \delta_0)}{1 + \delta_0 - \beta_l} \right) \\ \tilde{v}(x) &:= \exp \left\{ \frac{\alpha_l^2 y^2}{2(1 + \delta_0 - \beta_l)} + 1 \right\} \end{aligned}$$

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

$$\begin{aligned} &\frac{\int_{D_l(x,y)} v(u) |M^l(x, u) - M^l(y, u)| du}{v(x)v(y)} = \\ &\sqrt{\left(\frac{1 + \delta_0}{1 + \delta_0 - \beta_l}\right)} \left( \left\{ \frac{\tilde{v}(x)}{v(x)v(y)} \{ \Theta(\bar{c}_{l,1}(x, y)) - \Theta(\bar{c}_{l,2}(x, y)) \} + \right. \right. \\ &\left. \left. \Theta(\bar{c}_{l,2}(x, y)) \left[ \frac{\tilde{v}(x)}{v(x)} \left\{ \frac{1}{v(y)} - \frac{1}{v(x)} \right\} + \frac{1}{v(x)} \left\{ \frac{\tilde{v}(x)}{v(x)} - \frac{\tilde{v}(y)}{v(y)} \right\} \right] \right) \right). \end{aligned}$$

From here, it is straightforward to establish that all the functions in the differences are in the set  $\text{Lip}_{d_1}(\mathbb{X})$  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.  $\square$

## References

- [1] BALLY, V. & 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., JASRA A., LAW, K. J. H., TEMPONE, R. & ZHOU, Y. (2017). Multilevel sequential Monte Carlo samplers. *Stoch. Proc. Appl.*, **127**, 1417–1440.
- [3] BILLINGSLEY, P. (1995). *Probability and Measure*. Wiley: New York.
- [4] CAPPE, O., MOULINES, E. & 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. & SINGH, S. S. (2015). On particle Gibbs sampling. *Bernoulli*, **21**, 1855–1883.
- [7] CRISAN, D. & 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. & MICLO, L. (2000). Branching and interacting particle systems approximation of Feynman-Kac formulae with applications to non-linear filtering. In Azema, J., Emery, M., Ledoux, M., Yor, M. (Eds) *Seminaire de Probabilites XXIV Lecture Notes in Mathematics*, **1729**, 1–145, Springer: Berlin.
- [10] DEL MORAL, P., DOUCET, A. & JASRA, A. (2012). On adaptive resampling procedures for sequential Monte Carlo methods. *Bernoulli*, **18**, 252–272.
- [11] DEL MORAL, P., JACOD, J. & PROTTER, P. (2001). The Monte Carlo method for filtering with discrete-time observations. *Probab. Theory Rel. Fields*, **120**, 346–368.

- [12] DEL MORAL, P., JASRA, A. & 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. *Op. Res.*, **56**, 607–617.
- [14] GREGORY, A., COTTER, C., & REICH, S. (2016). Multilevel ensemble transform particle filtering. *SIAM J. Sci. Comp.*, **38**, A1317–A1338.
- [15] HEINRICH, S. (2001). Multilevel Monte Carlo methods. In *Large-Scale Scientific Computing*, (eds. S. Margenov, J. Wasniewski & P. Yalamov), Springer: Berlin.
- [16] JACOB, P., LINDSTEN, F. & SCHÖN, T. (2016). Coupling of particle filters. arXiv preprint.
- [17] JACOB, P., LINDSTEN, F. & SCHÖN, T. (2018). Smoothing with couplings of conditional particle filters. *J. Amer. Statist. Assoc.* (to appear).
- [18] JASRA, A., (2015). On the behaviour of the backward interpretation of Feynman-Kac formulae under verifiable conditions. *J. Appl. Probab.*, **52**, 339–359.
- [19] JASRA, A., BALLELIO, M., VON SCHWERIN, E., & TEMPONE, R. (2018). A coupled particle filter for multilevel estimation. Technical Report.
- [20] JASRA, A., KAMATANI, K., OSEI, P. P., & ZHOU, Y. (2018). Multilevel particle filters: Normalizing Constant Estimation. *Statist. Comp.*, **28**, 47–60.
- [21] JASRA, A., KAMATANI, K., LAW K. J. H. & ZHOU, Y. (2017). Multilevel particle filters. *SIAM J. Numer. Anal.*, **55**, 3068–3096.
- [22] LEE, A., SINGH, S. S. & VIHOLA, M. (2018). Coupled conditional backward sampling particle filter. arXiv preprint.
- [23] LO, A. (2017). Functional generalizations of Hoeffding’s covariance lemma and a formula for Kendall’s tau. *Stat. Probab. Lett.*, **122**, 218–226.
- [24] MAO, X. (2007). *Stochastic Differential Equations and Applications*. 2nd Edition, Woodhead: Cambridge.
- [25] MCDIARMID, C. (1989). On the method of bounded differences. In J. Siemons (Ed.), *Surveys in Combinatorics*, Volume 141 of London Mathematical Society Lecture Notes, pp. 148–188. CUP: Cambridge.
- [26] RACHEV, T. & RÜSCHENDORF, S. (1998). *Mass Transportation Problems, Volume 1: Theory*. Springer: New York.
- [27] SEN, D., THIERY, A. & JASRA, A. (2018). On coupling particle filters. *Stat. Comp.*, **28**, 461–475.
- [28] THORISSON, H. (2000). *Coupling, stationarity, and regeneration*. Springer: New York.
- [29] WHITELEY, N. P. (2013). Stability properties of some particle filters. *Ann. Appl. Probab.*, **23**, 2500–2537.
- [30] WHITELEY, N. P., KANTAS, N., JASRA, A. (2012). Linear variance bounds for particle approximations of time homogenous Feynman-Kac formulae. *Stoch. Proc. Appl.*, **122**, 1840–1865.