Variance Estimation of Adaptive Sequential Monte Carlo

Qiming Du and Arnaud Guyader

Abstract

This article is concerned with the variance estimation in Adaptive Sequential Monte Carlo (ASMC) methods. We deal with the case where the asymptotic variance coincides with the one of a "limiting" Sequential Monte Carlo (SMC) algorithm. We prove that, under mild assumptions, the estimator introduced by Lee and Whiteley in [9] in the non adaptive case, which can be obtained by a single simulation, is also a consistent estimator of the asymptotic variance for ASMC methods. To do this, we provide a different consistent estimator in terms of coalescent tree-typed measures. This estimator is constructed by tracing the genealogy of the associated Interactive Particle System. The tools we use connect the study of Particle Markov Chain Monte Carlo methods [1] and the variance estimation problem of SMC methods. As such, they may give some new insights when dealing with complex genealogy-involved problems of Interactive Particle System in more general scenarios.

Contents

Introduction

Sequential Monte Carlo (SMC) methods are among the most widely used Monte Carlo techniques in filtering, Bayesian inference, rare events simulations and many other disciplines (see for example [8] and references therein). The algorithm is designed to approximate a sequence of probability measures $(\eta_n)_{n\geq 0}$ in high dimension state spaces, by simulating an Interactive Particle System (IPS) via an importance sampling and resampling mechanism. The flow of measures is then approximated by the empirical version $(\eta_n^N)_{n\geq 0}$. A lot of convergence results when the sample size N goes to infinity can be found in the literature, see for example [6].

In practice, when applying these SMC methods, it is also very important to have a control on the constructed estimators, such as confidential intervals. As usual, if one has a CLT-typed theorem for the test function f such as

$$\sqrt{N}\left(\eta_n^N(f)-\eta_n(f)\right) \xrightarrow[N\to\infty]{\mathrm{d}} \mathcal{N}(0,\sigma_n^2(f)),$$

it suffices to provide a consistent estimator $\hat{\sigma}_n^2(f)$ of $\sigma_n^2(f)$. Indeed, Slutsky's lemma then ensures that

$$\frac{\sqrt{N}\left(\eta_n^N(f)-\eta_n(f)\right)}{\hat{\sigma}_n(f)} \xrightarrow[N\to\infty]{\mathrm{d}} \mathcal{N}(0,1). \tag{1}$$

The standard way to achieve this is by resimulating the IPS independently many times and by estimating $\sigma_n^2(f)$ with the crude variance estimator. However, since a single run of the algorithm may take a lot of time, this is usually intractable. In addition, as the estimator $\eta_n^N(f)$ of $\eta_n(f)$ provided by SMC is typically biased, it is also nontrivial to implement the parallel computing for a large number of IPS with N relatively small. As a consequence, a variance estimator available with a single run of the simulation is of crucial interest for applications.

The first consistent variance estimator of this type was proposed by Chan and Lai in [5], by using the ancestral information encoded in the genealogy of the associated IPS. Then, Lee and Whiteley proposed an unbiased variance estimator for the unnormalized measures γ_n^N and a term by term estimator [9], with lots of insights on the genealogy of the IPS. All of these estimators are studied in the classical SMC framework, meaning in a non adaptive setting: The weight functions and the Markov proposal kernels are fixed.

In this article, we deal with *adaptive* SMC methods: At each resampling step, the weight functions and/or Markov proposal kernels depend upon the history of the simulated process. The idea is to approximate an ideal "limiting" SMC algorithm, which is usually out of reach, by exploiting the induced information tracked by some summary statistics. Such approaches are expected to be more efficient and more automated than the non-adaptive ones, which contain more user-specified tuning parameters.

Specifically, we are interested in the case where the adaptive SMC algorithm is asymptotically "identical" to a "limiting" SMC algorithm. More precisely, we expect the asymptotic variance of the adaptive SMC algorithm to be identical to the "ideal" non-adaptive one. This kind of stability property is at the core of the pair of articles [2] and [4]. The framework discussed in the present article is just a slightly generalized version of the one presented in Section 2 in [2] but still ensures the stability property of their Theorem 2.3. However, we do not discuss here the variance estimation problem in the case where such stability does not exist.

Another remark concerns the Adaptive Multilevel Splitting (AMS) methods under multinomial selection scheme, such as the one discussed in [4]. Despite the fact that our assumptions are not verified by the AMS framework, we expect that the variance estimator would also work under the assumptions that a CLT-typed result is available. Nonetheless, we believe that this case requires a specific analysis and different assumptions. To account for this, one can notice that the proofs in [4] and [2] differ in many points, although the take-home message is the same.

From a theoretical viewpoint, to prove the consistency of the variance estimator proposed in [9], we were not able to adapt their technical tools. This is due to the additional randomness brought by the weight functions and Markov kernels at each resampling step. As a consequence, we propose to develop new techniques in order to estimate the terms Γ_n^b that appear in the expansion of the variance given in [3]. The mains ideas are: first, our term by term estimator is consistent and, second, the difference between our estimator and the one of Lee and Whiteley goes to 0 in probability when the sample size N goes to infinity. However, in practice, one uses the estimator proposed by Lee and Whiteley, while the one we introduce here may be seen as a handy tool to prove the consistency of the former.

Interestingly, the construction of our estimators $\Gamma^b_{n,N}$ for Γ^b_n uses the idea of many-body Feynman-Kac models, which were designed in [7] to study the propagation of chaos properties of Conditional Particle Markov Chain Monte Carlo methods. Above the specific context of the present article, these connections may give some insights on how to deal with complex genealogy-involved problems in more general settings.

Notations and conventions

Before going into details, let us provide a few notation which is useful in the following.

- The underlying probability space is denoted by $(\Omega, \mathscr{F}, \mathbf{P})$. For σ-fields $\mathcal{E}, \mathcal{G} \subset \mathscr{F}$, $\mathcal{E} \vee \mathcal{G}$ denotes the smallest σ-field on Ω containing $\mathcal{E}$ and $\mathcal{G}$. For any $x, y \in \mathbf{R}$, we denote $x \wedge y := \min\{x, y\}$ and $x \vee y := \max\{x, y\}$. We use standard conventions $\inf \emptyset = \infty$.
- For the notation $\mathbf{R}^d$, we consider the Euclidean space, where the notation $|\cdot|$ denote the Euclidean norm and "$\cdot$" denote the dot product on Euclidean space.
- Random variables take values in Polish spaces. For example, a topological space E which is metrizable, separable and complete for some distance d_E. It is endowed with the Borel σ-algebra generated by d_E, denoted by $\mathcal{B}(E)$. In particular, for multiple particles such as a particle layer $\mathbf{X_n} := (X^1_n, \cdots, X^N_n) \in E^N$ in an IPS, we consider the Polish space $(E^N, \mathcal{B}(E)^{\otimes N})$.
- We denote respectively by $\mathcal{M}(E)$, $\mathcal{M}_+(E)$ and $\mathcal{P}(E)$ the set of all signed finite measures, the subset of all nonnegative finite measures and the subset of all the probability measures on $(E, \mathcal{B}(E))$. The set $\mathcal{P}(E)$ is endowed with the Prohorov-Lévy metric, *i.e.* the weak convergence is tested with continuous bounded test functions.
- For any Polish space E, $\mathcal{B}_b(E)$ denotes the collection of of all the bounded measurable functions from $(E, \mathcal{B}(E))$ to $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ equipped with uniform norm $\|\cdot\|_\infty$, among which the constant function will be denoted by 1 with a slight abuse of notation. L^p denotes the standard random variable space $L^p(\Omega, \mathscr{F}, \mathbf{P})$.
- For all $\mu \in \mathcal{M}(E)$ and for all test functions $f \in \mathcal{B}_b(E)$, $\mu(f)$ denotes the integration

  $$\int_E f(x)\mu(dx).$$

  A finite transition kernel Q from Polish space $(E, \mathcal{B}(E))$ to $(F, \mathcal{B}(F))$ is a function

  $$Q : E \times \mathcal{B}(F) \mapsto \mathbf{R}_+.$$

  More precisely, for all $x \in E$, $Q(x, \cdot)$ is a finite nonnegative measure in $\mathcal{M}_+(F)$ and for all $A \in \mathcal{B}(F)$, $Q(x, A)$ is a $\mathcal{B}(E)$-measurable function. We say a Q is a Markov kernel if Q is a finite transition kernel and for all $x \in E$, $Q(x, \cdot)$ is a probability measure in $\mathcal{P}(F)$. For a signed measure $\mu \in \mathcal{M}(E)$ and a test function $f \in \mathcal{B}_b(F)$, we denote $\mu Q \in \mathcal{M}(E)$ and $Q(f) \in \mathcal{B}_b(E)$, respectively, by

  $$\mu Q(A) := \int_E \mu(dx)Q(x, A), \qquad \forall A \in \mathcal{B}(F)$$

  and

  $$Q(f)(x) := \int_F Q(x, dy)f(y), \qquad \forall x \in E.$$

  Given two finite transition kernels Q_1 and Q_2 respectively from E_0 to E_1 and E_1 to E_1, we denote Q_1Q_2 the transition kernel from E_0 to E_2 defined by

  $$Q_1Q_2(x, A) := \int_{E_1} Q_1(x, dy)Q_2(y, A), \qquad \forall (x, A) \in E_1 \times \mathcal{B}(E_2).$$

  In general, there is no reason that Q_1Q_2 is still a finite transition kernel.

- For two test functions $f, g \in \mathcal{B}_b(E)$, we denote

$$f \otimes g : E^2 \ni (x, y) \mapsto f(x)g(y) \in \mathbf{R}$$

and

$$f \cdot g : E \ni x \mapsto f(x)g(x) \in \mathbf{R}.$$

In particular, we denote

$$f^{\otimes 2} := f \otimes f.$$

For two finite transition kernels Q and H from $(E, \mathcal{B}(E))$ to $(F, \mathcal{B}(F))$, we denote

$$Q \otimes H\,((x, y), (A, B)) := Q(x, A) \times H(y, B)$$

for all $(x, y) \in E \times E$ and for all $(A, B) \in \mathcal{B}(F) \otimes \mathcal{B}(F)$. Similarly, we also denote

$$Q^{\otimes 2} := Q \otimes Q.$$

- In order to define the coalescent tree-typed measures of size 2, we introduce two special integral operators. We denote

$$C_1((x^1, x^2), d(y^1, y^2)) := \delta_{(x^1, x^1)} d(y^1, y^2)$$

and

$$C_0((x^1, x^2), d(y^1, y^2)) := \delta_{(x^1, x^2)} d(y^1, y^2).$$

More precisely, let E be a Polish spaces, for any measurable functions $H : E \mapsto \mathbf{R}$, C_0 and C_1 are Markov transition kernels such that

$$C_0(H)(x, y) = H(x, y) \qquad \text{and} \qquad C_1(H)(x, y) = H(x, x).$$

We call $b := (b_0, \dots, b_n) \in \{0, 1\}^{n+1}$ a coalescence indicator where $b_p = 1$ indicates that there is a coalescence at level p.

- To describe the simulation algorithm, we use the notation $X \sim \mu(\cdot)$ for a given probability measure $\mu \in \mathcal{P}(E)$. Rigorously, it means that given U, an uniformly distributed random variable on $[0, 1]$, there exists a measurable function $H : [0, 1] \mapsto E$ such that $\mathrm{Law}(H(U)) = \mu$. The X is then defined by $X := H(U)$. The existence of X and such representation is given by Kuratowski's Theorem as E is assumed to be a Polish space. Similarly, given a random variable X taking values on E, the notation $Y \sim Q(X, \cdot)$ or $Y \sim Q(X, dY)$ means that there exists some measurable function $H : E \times [0, 1] \mapsto E$ such that $\mathrm{Law}(H(x, U)) = Q(x, \cdot)$, for $\mathrm{Law}(X)$-almost every $x \in E$.

- For the notation concerning the indices of the particles in the IPS, we use

$$[N] := \{1, 2, \dots, N\} \qquad \text{and} \qquad [N]^q_p := \left\{(i_1, \dots, i_q) \in [N]^q : \mathrm{Card}\{i_1, \dots, i_q\} = p\right\}$$

to denote the indices sets useful in the following sections. In particular, we denote $(N)^q := [N]^q_q$. In order to differentiate $(N)^q$ from the Cartesian product when dealing with the indices sets of a particle block, such as

$$\underbrace{(N)^2 \times (N)^2 \times \cdots \times (N)^2}_{q \text{ times}},$$

we use the notation $\left((N)^2\right)^{\times q}$ as an abbreviation.

- For the particle block of size 2, we consider an IPS with $n + 1$ layers and N particles at each layer, we use the notation $\ell_p^{[2]} := (\ell_p^1, \ell_p^2) \in (N)^2$ to track the indices of two different particles at level p for $p \leq n$. Similarly, we denote

$$\ell_{0:n}^{[2]} := \left((\ell_0^1, \ell_0^2), \cdots, (\ell_n^1, \ell_n^2)\right) \in \left((N)^2\right)^{\times(n+1)}$$

for the particle block of size 2 from layer 0 to layer n. With a slight abuse of notation, we admit that

$$((i,j),(k,l)) = (i,j,k,l).$$

We denote $\boldsymbol{x_p} := (x_p^1, \ldots, x_p^N)$ all the particles at layer p, where x_p^i denotes the i-th particle of this layer. To deal with the genealogy of the particle system, we denote $\boldsymbol{a_p} := (a_p^1, \ldots, a_p^N)$ all the parent indices of layer p, where $a_p^i \in [N]$ denotes the index of parent of i-th particle at layer $p+1$. We also denote $a_p^{(i,j)} := (a_p^i, a_p^j) \in [N]^2$ the indices of parents of i-th and j-th particles at layer $p+1$. In combination with the former notation, we use frequently

$$a_p^{\ell_{p+1}^{[2]}} := (a_p^{\ell_{p+1}^1}, a_p^{\ell_{p+1}^2})$$

to track the genealogy within a certain particle block. The notation for the particle block of size $q \in [N]$ is similar. The capital letters such as $\mathbf{X_p}$, X_p^i, $\mathbf{A_p}$ and A_p^j are used to denote the associated random variables.

- For all $\boldsymbol{x} = (x^1, \ldots, x^N) \in E^N$, we define the empirical measure associated to $\boldsymbol{x}$ by

$$m : \boldsymbol{x} \mapsto m(\boldsymbol{x}) := \frac{1}{N} \sum_{i=1}^{N} \delta_{x^i} \in \mathcal{P}(E^2).$$

We also denote

$$m^{\otimes 2} : \boldsymbol{x} \mapsto m^{\otimes 2}(\boldsymbol{x}) := \frac{1}{N^2} \sum_{i,j} \delta_{(x^i, x^j)} \in \mathcal{P}(E^2)$$

and

$$m^{\odot 2} : \boldsymbol{x} \mapsto m^{\odot 2}(\boldsymbol{x}) := \frac{1}{N(N-1)} \sum_{i \neq j} \delta_{(x^i, x^j)} \in \mathcal{P}(E^2).$$

Elementary calculation gives that

$$m^{\otimes 2}(\boldsymbol{x}) = \frac{N-1}{N} \cdot m^{\odot 2}(\boldsymbol{x}) C_0 + \frac{1}{N} \cdot m^{\odot 2}(\boldsymbol{x}) C_1.$$

With a slight abuse of notation, we denote

$$m([N]) := \frac{1}{N} \sum_{i=1}^{N} \delta_i \quad \text{and} \quad m^{\otimes 2}([N]) := m([N]) \otimes m([N])$$

1 ASMC framework

In this section, we give the formal definition and the regularity assumptions of the ASMC framework studied in this article. The motivation is mainly from the ASMC via summary statistics introduced in Section 2 of [2], where a list of adaptive SMC models are referred there.

1.1 Adaptive SMC Settings

Let $(E_n, \mathcal{B}(E_n))_{n \geq 0}$ be a sequence of Polish spaces endowed with Borel σ-algebra. For each level $n \geq 1$, we consider a family of potential functions $G_{n-1,z} : E_{n-1} \mapsto \mathbf{R}_+$ and Markov kernels $M_{n,z} : (E_{n-1}, \mathcal{B}(E_n)) \mapsto [0,1]$ parameterized by a parameter $z \in \mathbf{R}^d$. Accordingly, we define a family of transition kernels $Q_{n,z}$ by

$$Q_{n,z}(x, A) := G_{n-1,z}(x) \times M_{n,z}(x, A),$$

with the convention $Q_{0,z}(x, A) := \delta_x(A)$. We suppose that there exists a sequence of reference parameters $(z_n^*)_{n \geq 0}$ and for each $n \geq 1$, we denote

$$G_{n-1} := G_{n-1, z_{n-1}^*}, \qquad M_n := M_{n, z_{n-1}^*} \qquad \text{and} \qquad Q_n := Q_{n, z_{n-1}^*}.$$

Starting with a known probability measure $\gamma_0 := \eta_0 \in \mathcal{P}(E_0)$, the Feynman-Kac measures are defined by

$$\gamma_n := \gamma_0 Q_1 \cdots Q_n, \qquad n \geq 1$$

along with the normalized measures defined by

$$\eta_n := \frac{1}{\gamma_n(1)} \cdot \gamma_n,$$

In order to have a meaningful simulation problem in practice, we suppose that

$$\gamma_n(1) = \prod_{p=0}^{n-1} \eta_p(G_p) > 0$$

for all $n \geq 0$. We define the Feynman-Kac semigroup

$$Q_{p,n} := Q_{p+1} \cdots Q_n$$

for $p < n$ and $Q_{n,n}(x, A) := \delta_x(A)$. ASMC algorithms aim at approximating the sequences of measures $(\gamma_n)_{n\geq 0}$ and $(\eta_n)_{n\geq 0}$ by exploiting summary statistics

$$\zeta_n : E_n \mapsto \mathbf{R}^d$$

such that for $n \geq 0$, we have

$$\eta_n(\zeta_n) = z_n^*.$$

1.2 ASMC algorithm

Roughly speaking, ASMC and SMC algorithm share the same selection mutation-typed mechanism. However, as the parameters $(z_n^*)_{n\geq 0}$ are not analytically tractable in practice, the potential functions $(G_n)_{n\geq 0}$ and transition kernels $(M_n)_{n\geq 1}$ are estimated "on the fly" by the design of an adaptive algorithm.

IPS with genealogy Let $N \in \mathbf{N}^*$ be the number of particles at each layer. The associated IPS of ASMC algorithm is a Markov chain $(\mathbf{X_n})_{n\geq 0}$ taking values in $(E_n^N, \mathcal{B}(E_n)^{\otimes N})_{n\geq 0}$ with genealogy $(\mathbf{A_n})_{n\geq 0}$ tracking the indices of parents of each particle at each level n. We recall that $A_{p-1}^i = j$ means that the parent of the particle X_p^i at layer p is X_{p-1}^j at layer $p-1$. For the estimation of η_n, we define

$$\eta_n^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_n^i}.$$

At each level $n \geq 0$, the estimated parameters are defined by $Z_n^N := \eta_n^N(\zeta_n)$. In order to simplify the notation, we denote

$$G_{n,N} := G_{n,Z_n^N}, \qquad M_{n,N} := M_{n,Z_{n-1}^N} \qquad \text{and} \qquad Q_{n,N} := Q_{n,Z_{n-1}^N}$$

Thus, the estimations of the unnormalized Feynman-Kac measures are defined by

$$\gamma_n^N := \prod_{p=0}^{n-1} \eta_p^N(G_{p,N}) \cdot \frac{1}{N} \sum_{i=1}^N \delta_{X_n^i}$$

In the following sections, we use the convention

$$\eta_{-1}^N = \gamma_{-1}^N := \eta_0.$$

Now, let us give a formal definition of the IPS associated with the ASMC algorithm.

(i) Initial distribution: At level 0, we let $\mathbf{X_0} \sim \eta_0^{\otimes N}$.

(ii) Transition kernels: For all $p \geq 0$, we let $Z_p^N = \eta_p^N(\zeta_p)$. The elementary transition $X_p^i \rightsquigarrow X_{p+1}^i$ is decomposed into two steps:

- *Selection:* Given $\mathbf{X_p} = \boldsymbol{x_p}$, we make an independent multinomial selection of parent for each particle by

$$S_{p,N}(\boldsymbol{x_p}, da_p^i) = \sum_{k=1}^{N} \frac{G_{p,N}(x_p^k)}{\sum_{j=1}^{N} G_{p,N}(x_p^j)} \cdot \delta_k(da_p^i). \tag{2}$$

Thus, the genealogy of level p to level $p+1$ is tracked by

$$\mathbf{A_p} \sim \bigotimes_{i=1}^{N} S_{p,N}(\mathbf{X_p}, \cdot)$$

- *Mutation:* Given the parent indices $\mathbf{A_p} = \boldsymbol{a_p}$, each particle at level p evolves independently according to the transition kernel $M_{p+1,N}$, i.e.,

$$X_{p+1}^i \sim M_{p+1,N}(X_p^{a_p^i}, \cdot).$$

Therefore, the transition of the IPS with its genealogy from level p to level $p+1$ is given by

$$(\mathbf{A_p}, \mathbf{X_{p+1}}) \sim \bigotimes_{i=1}^{N} \Phi_{p+1,N}(\mathbf{X_p}, d(A_p^i, X_{p+1}^i))$$

with $\Phi_{p,N}$ defined by

$$\Phi_{p+1,N}(\boldsymbol{x_p}, d(a_p^i, x_{p+1}^i)) = S_{p,N}(\boldsymbol{x_p}, da_p^i) \times M_{p+1,N}(x_p^{a_p^i}, dx_{p+1}^i).$$

Filtrations Finally, we define the filtrations $(\mathcal{G}_n^N)_{n\geq 0}$ (resp. $(\mathcal{F}_n^N)_{n\geq 0}$) with (resp. without) the genealogy of the IPS. For $n \geq 1$, we define

$$\mathcal{G}_n^N := \sigma(\mathbf{X_0}, \ldots, \mathbf{X_n}, \mathbf{A_0}, \ldots, \mathbf{A_{n-1}})$$

and

$$\mathcal{F}_n^N := \sigma(\mathbf{X_0}, \ldots, \mathbf{X_n}).$$

In particular, we let

$$\mathcal{G}_0^N := \sigma(\mathbf{X_0}) \qquad \text{and} \qquad \mathcal{G}_{-1}^N := \{\emptyset, \Omega\}.$$

1.3 Assumptions

Before listing our assumptions, we want to mention that they are introduced in a similar taste, but slightly weaker, as in [2]. The reason of making these slight generalizations is to guide the possible future applications. Despite the fact they are not direct to verify in practice, though, we try to throw light on possible modifications we could make in order to design more effective models. In the following sections, we use $\mathcal{A}$ as shorthand of *Assumption*.

Assumption 1. *For each $n \geq 0$, we assume that $G_{n,z}$ is strictly positive and bounded uniformly over $z \in \mathbf{R}^d$, i.e.*

$$\sup_{(x,z)\in E_n\times\mathbf{R}^d} G_{n,z}(x) = \left\|G_{n,\cdot}\right\|_\infty < +\infty.$$

We also assume that

$$\zeta_n = (\zeta_n^1, \ldots, \zeta_n^d) \quad \textit{with} \quad \forall k \in [d], \quad \zeta_n^k \in \mathcal{B}_b(E_n).$$

Assumption 2. *For any test function $f \in \mathcal{B}_b(E_n)$, we assume that there exists a function $h_n : E_{n-1} \times \mathbf{R}^d \mapsto \mathbf{R}^d$ such that*

$$Q_{n,z_{n-1}}(f)(x) - Q_n(f)(x) = h_n(x, z_{n-1}) \cdot (z_{n-1} - z_{n-1}^*).$$

The function h is assumed to have the following regularity properties:

- *h_n is uniformly bounded over $E_{n-1} \times \mathbf{R}^d$ by $\|h_n\|_\infty$.*
- *For any $z \in \mathbf{R}^d$, $h_n(\cdot, z)$ is $\mathcal{B}(E_{n-1})$-measurable.*
- *The application $z \mapsto h_n(x, z)$ is continuous at z^*_{n-1} uniformly over $x \in E_{n-1}$. More precisely, for any $\epsilon > 0$, there exists $\delta > 0$, such that $\left|z_{n-1} - z^*_{n-1}\right| < \delta$ implies that*

$$\left|h_n(x, z_{n-1}) - h_n(x, z^*_{n-1})\right| < \epsilon.$$

- *h satisfies the equality $\eta_{n-1}\left(h_n(\cdot, z^*_{n-1})\right) = 0$.*

Remark. $\mathcal{A}2$ guarantees some *continuity* properties of the transition kernels $Q_{n,z}$ over the parameter z and is just a intermediate step of the framework studied in Section 2 of [2]. In fact, we can take the ω defined in (2.17) as the function h_n above. All of the first three regularity properties are verified by Assumption 1-2 and the forth point is given by Theorem 2.13. We also want to mention that the third point is equivalent to

$$\begin{aligned} &\forall \epsilon > 0,\ \exists g \in \mathcal{B}_b(E_{n-1}),\ \exists \delta > 0, \\ &\qquad\qquad s.t.\ \left|z_{n-1} - z^*_{n-1}\right| < \delta \implies \left|h_n(x, z_{n-1}) - h_n(x, z^*_{n-1})\right| < g_n(x) \cdot \epsilon. \end{aligned}$$

We expect that the function g_n and h_n can be relaxed to some unbounded function such as $L^2(\eta_{n-1})$ functions, along with some more strict limitations to the test function f. We also believe that this is one of the essential differences of the ASMC framework studied in [2] and AMS framework studied in [4]. In general, it is not easy to verify the existence of such h_n. However, we have, at least, a direction to explore in the case where the $Q_{n,z}(f)$ is not differentiable w.r.t. z. We also remark that we did not study the consistency of $\gamma_n^N(f)$ and $\eta_n^N(f)$ separately with weaker assumptions, as our main interest is the CLT-typed result as (1).

1.4 Central limit theorems

In this article, we only deal with the case where the asymptotic variance equals to the "limiting" one, which is only a special case of the central limit theorem given in [2] under slightly weaker assumptions. This is why we used a totally different strategy for the proof, which can be found in Section 3.2.

Theorem 1.1. *Assume ($\mathcal{A}1$-$\mathcal{A}2$). For any test function $f \in \mathcal{B}_n(E_n)$, we have*

- $\sqrt{N}\left(\gamma_n^N(f) - \gamma_n(f)\right) \xrightarrow[N\to\infty]{\mathrm{d}} \mathcal{N}\left(0, \sigma^2_{\gamma_n}(f)\right)$
- $\sqrt{N}\left(\eta_n^N(f) - \eta_n(f)\right) \xrightarrow[N\to\infty]{\mathrm{d}} \mathcal{N}\left(0, \sigma^2_{\eta_n}(f - \eta_n(f))\right)$

with

$$\sigma^2_{\gamma_n}(f) := \sum_{p=0}^{n}\left(\gamma_p(1)\gamma_p(Q_{p,n}(f)^2) - \gamma_n(f)^2\right) \qquad \text{and} \qquad \sigma^2_{\eta_n}(f) := \sigma^2_{\gamma_n}(f)/\gamma_n(1)^2.$$

2 Variance estimations

In this section, we recall the coalescent tree-based expansion of the variance introduced in [3], as well as the variance estimator proposed by Lee and Whiteley in [9] for the Particle Filters and some of its properties.

2.1 Coalescent tree-based variance expansion

The main goal of this article is to estimate the form $\sigma^2_{\gamma_n}(f)$ and $\sigma^2_{\eta_n}(f - \eta_n(f))$.

Definition 2.1. *We associate with any coalescence indicator $b \in \{0,1\}^{n+1}$ the nonnegative measure $\Gamma_n^b \in \mathcal{M}_+(E_n)$ defined by*

$$\Gamma_n^b(F) := \eta_0^{\otimes 2} C_{b_0} Q_1^{\otimes 2} C_{b_1} \cdots Q_n^{\otimes 2} C_{b_n}(F)$$

and

$$\bar{\Gamma}_n^b(F) := \frac{1}{\gamma_n(1)^2} \cdot \Gamma_n^b(F)$$

for all $F \in \mathcal{B}_b(E_n^2)$. We also denote $\Gamma_n^{(p)}(F)$ where there is only one coalescence at level p. In particular, we denote

$$\Gamma_n^{(\emptyset)}(F) = \gamma_n^{\otimes 2}(F) \qquad \textit{and} \qquad \bar{\Gamma}_n^{(\emptyset)}(F) = \eta_n^{\otimes 2}(F)$$

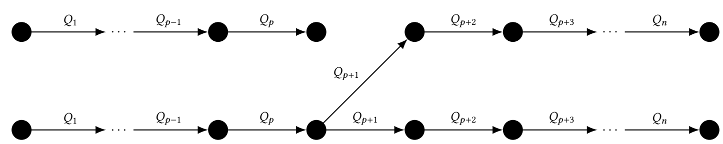

Figure 1: A representation of the coalescent tree-typed measure $\Gamma_n^{(p)}$.

By definition, it is easy to verify that we have

$$\Gamma_n^{(p)}(f^{\otimes 2}) = \gamma_p(1)\gamma_p\left(Q_{p,n}(f)^2\right) \tag{3}$$

which gives us an alternative representation of the asymptotic variance based on coalescent tree-typed measures:

$$\sigma^2_{\gamma_n}(f) = \sum_{p=0}^{n} \left(\Gamma_n^{(p)}(f^{\otimes 2}) - \Gamma_n^{(\emptyset)}(f^{\otimes 2})\right) \tag{4}$$

and

$$\sigma^2_{\eta_n}(f) = \sum_{p=0}^{n} \left(\bar{\Gamma}_n^{(p)}(f^{\otimes 2}) - \bar{\Gamma}_n^{(\emptyset)}(f^{\otimes 2})\right) \tag{5}$$

Definition 2.2. *For any test function F in $\mathcal{B}_b(E_n^2)$ and indicator of coalescence b, we define an estimator $\bar{\Gamma}_{n,N}^b$ of the measure $\bar{\Gamma}_n^b$ by*

$$\bar{\Gamma}_{n,N}^b(F) := \frac{N^{n-1}}{(N-1)^{n+1}} \sum_{\ell_{0:n}^{[2]} \in ((N)^2)^{\times(n+1)}} \prod_{p=0}^{n-1} \left\{ \lambda_p^b(A_p^{\ell_{p+1}^{[2]}}, \ell_p^{[2]}) \right\} \cdot C_{b_n}(F)(X_n^{\ell_n^{[2]}})$$

where $\lambda_p^b(\tilde{a}_p^{[2]}, \ell_p^{[2]})$ is an indicator function defined by

$$\lambda_p^b(\tilde{a}_p^{[2]}, \ell_p^{[2]}) := \mathbf{1}_{\{b_p=1, \tilde{a}_p^1=\tilde{a}_p^2=\ell_p^1\neq\ell_p^2\}} + \mathbf{1}_{\{b_p=0, \tilde{a}_p^1=\ell_p^1\neq\tilde{a}_p^2=\ell_p^2\}}$$

with the convention

$$\bar{\Gamma}_{0,N}^b(F) := \frac{1}{N(N-1)} \sum_{\ell_0^{[2]} \in (N)^2} C_{b_0}(F)(X_0^{\ell_0^{[2]}}) = \frac{1}{N(N-1)} \sum_{i\neq j} C_{b_0}(F)(X_0^i, X_0^j)$$

and

$$\bar{\Gamma}_{-1,N}^b(F) = \Gamma_{-1,N}^b(F) := \eta_0^{\otimes 2} C_{b_0}.$$

Also, we denote

$$\Gamma_{n,N}^b(F) = \gamma_n^N(1)^2 \cdot \bar{\Gamma}_{n,N}^b(F).$$

A concrete example As the definition of the estimator $\Gamma^{b}_{n,N}$ is relatively complex, we give an example in order to understand what it is and how to do the calculation in practice. Let us introduce some additional notation. Now, we consider an IPS as follow.

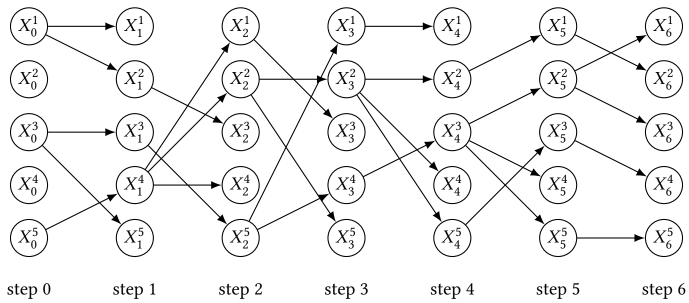

Figure 2: An IPS with 7 levels and 5 particles at each level.

We are going to estimate $\Gamma^{(3)}_{6}(f \otimes g)$. We denote $b^* = (0, 0, 0, 1, 0, 0, 0)$. In the associated IPS, we find the choices of $\ell^{[2]}_{0:6}$ such that

$$\prod_{p=0}^{5} \lambda_p^{b^*}(A_p^{\ell^{[2]}_{p+1}}, \ell^{[2]}_p) = 1 \tag{6}$$

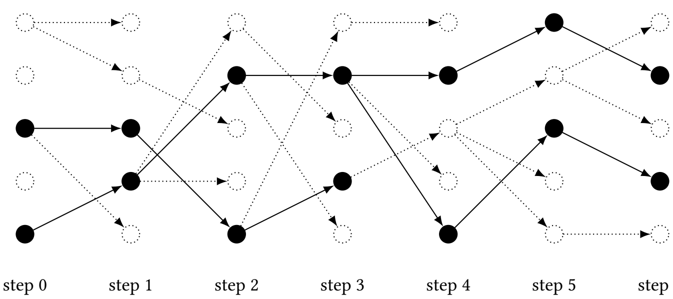

Figure 3: The first family of $\ell^{[2]}_{0:6}$ such that (6) is verified.

There are two choices possible:

- $\ell^{[2]}_{0:6} = ((5, 3), (4, 3), (2, 5), (2, 4), (2, 5), (1, 3), (2, 4))$
- $\ell^{[2]}_{0:6} = ((5, 3), (4, 3), (2, 5), (2, 4), (5, 2), (3, 1), (4, 2))$

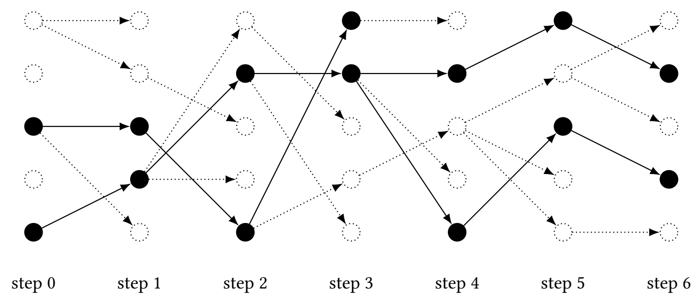

Figure 4: The second family of $\ell_{0:6}^{[2]}$ such that (6) is verified.

There are another two choices possible:

- $\ell_{0:6}^{[2]} = ((5,3),(4,3),(2,5),(2,1),(2,5),(1,3),(2,4))$
- $\ell_{0:6}^{[2]} = ((5,3),(4,3),(2,5),(2,1),(5,2),(3,1),(4,2))$

Hence, the number of choices of $\ell_{0:6}^{[2]}$ where $\ell_6^{[2]} = (2,4)$ is 2, and the number of choices of $\ell_{0:6}^{[2]}$ where $\ell_6^{[2]} = (4,2)$ is also 2. As a consequence, we have

$$\Gamma_{6,5}^{(3)}(f^{\otimes 2}) = \left(\frac{6}{6-1}\right)^5 \cdot \gamma_6^5(1)^2 \cdot 2 \cdot \left(f(X_6^2)g(X_6^4) + f(X_6^4)g(X_6^2)\right).$$

Theorem 2.1 (Approximation of coalescent tree-typed measures)**.** *Assume ($\mathcal{A}1$-$\mathcal{A}2$). For any test functions $\phi, \psi \in \mathcal{B}_b(E_n)$ and for any coalescence indicator $b \in \{0,1\}^{n+1}$, we have*

$$\Gamma_{n,N}^b(\phi \otimes \psi) \xrightarrow[N\to\infty]{\mathbf{P}} \Gamma_n^b(\phi \otimes \psi)$$

2.2 Term by term estimator

We provide a term by term variance estimator in this section. This is in fact a by-product of the analysis on the disjoint ancestral lines-based estimator. Honestly speaking, as we see in Lemma 2.1, the difference of these two estimator is $O_{\mathbf{P}}(1/N)$, which is negligible. We do not recommend to use this estimator to construct the confidential interval. However, this estimator does use the different genealogical information of the particle system, and it is different from the original term by term estimator firstly introduced in [9], we decide to give a brief discussion here. It has also some potential values to construct the adaptive particle system as is briefly discussed in [9]. Inspired by the coalescent tree-typed variance expansion in (4) and (5), we construct the following term by term variance estimator. Given a test function $f \in \mathcal{B}_b(E_n)$, we let

$$\sigma_{\gamma_{n,N}}^2(f) := \sum_{p=0}^{n} \left(\Gamma_{n,N}^{(p)}(f^{\otimes 2}) - \Gamma_{n,N}^{(\emptyset)}(f^{\otimes 2})\right)$$

and

$$\sigma_{\eta_{n,N}}^2(f) := \sum_{p=0}^{n} \left(\bar{\Gamma}_{n,N}^{(p)}(f^{\otimes 2}) - \bar{\Gamma}_{n,N}^{(\emptyset)}(f^{\otimes 2})\right).$$

Theorem 2.2. *Assume ($\mathcal{A}1$-$\mathcal{A}2$). For $f \in \mathcal{B}_b(E_n)$, we have*

$$\sigma_{\gamma_{n,N}}^2(f) \xrightarrow[N\to\infty]{\mathbf{P}} \sigma_{\gamma_n}^2(f)$$

and

$$\sigma_{\eta_{n,N}}^2(f - \eta_n^N(f)) \xrightarrow[N\to\infty]{\mathbf{P}} \sigma_{\eta_n}^2(f - \eta_n(f)).$$

Proof. Theorem 2.1 ensures the consistency of both $\Gamma^{(p)}_{n,N}(f^{\otimes 2})$ and $\Gamma^{(\emptyset)}_{n,N}(f^{\otimes 2})$, which amounts to say that

$$\sigma^2_{\gamma_n,N}(f) = \sum_{p=0}^{n} \left(\Gamma^{(p)}_{n,N}(f^{\otimes 2}) - \Gamma^{(\emptyset)}_{n,N}(f^{\otimes 2})\right)$$

is a consistent estimator of $\sigma^2_{\gamma_n}(f)$. Similarly, for the consistency of $\sigma^2_{\eta_n,N}(f - \eta_n^N(f))$, it is sufficient to show that

$$\bar{\Gamma}^b_{n,N}\left(\left[f - \eta_n^N(f)\right]^{\otimes 2}\right) \xrightarrow[N\to\infty]{\mathbf{P}} \bar{\Gamma}^b_n\left(\left[f - \eta_n^N(f)\right]^{\otimes 2}\right). \tag{7}$$

In fact, by elementary calculations, we have

$$\begin{aligned}&\bar{\Gamma}^b_{n,N}\left(\left[f - \eta_n^N(f)\right]^{\otimes 2}\right)\\ &\qquad = \bar{\Gamma}^b_{n,N}(f^{\otimes 2}) + \eta_n^N(f)^2 \cdot \bar{\Gamma}^b_{n,N}(1) - \eta_n^N(f) \cdot \left(\bar{\Gamma}^b_{n,N}(1 \otimes f) + \bar{\Gamma}^b_{n,N}(f \otimes 1)\right)\end{aligned}$$

The verification of (7) ends with Theorem 2.1 and Theorem 3.1.

□

2.3 Disjoint ancestral lines-based estimator

Now, let us introduce the variance estimator proposed in [9]. Given a test function $f \in \mathcal{B}_b(E_n)$, we let

$$V_n^N(f) := \eta_n^N(f)^2 - \frac{N^{n-1}}{(N-1)^{n+1}} \sum_{E_n^i \neq E_n^j} f(X_n^i) f(X_n^j)$$

where E_n^i is the ancestor index of X_n^i at level 0. This is the variance estimator introduced by Lee and Whiteley in [9] in the case that the number of the particles N are the same at each step. By combining Lemma 2.1 and Theorem 2.2, we have the following convergence result.

Theorem 2.3. *Assume ($\mathcal{A}$1-$\mathcal{A}$2). For any test function $f \in \mathcal{B}_b(E_n)$, we have*

$$N\gamma_n^N(1)^2 V_n^N(f) \xrightarrow[N\to\infty]{\mathbf{P}} \sigma^2_{\gamma_n}(f)$$

and

$$N V_n^N(f - \eta_n^N(f)) \xrightarrow[N\to\infty]{\mathbf{P}} \sigma^2_{\eta_n}(f - \eta_n(f)).$$

2.4 Connection between the two estimators

In this section, we give some combinatorial results of the coalescent tree-typed measures $\Gamma^b_{n,N}$ and $\bar{\Gamma}^b_{n,N}$. These results connect the variance estimator of Lee & Whiteley, i.e. the disjoint ancestral lines-based estimator and the term by term estimator. Another remark is that these results do not depend on $\mathcal{A}$2: They are provided by the structure of the IPS and the underlying multinomial selection scheme.

Proposition 2.1. *Assume ($\mathcal{A}$1). For any test function $F \in \mathcal{B}_b(E_n^2)$, we have the following decompositions.*

$$(\gamma_n^N)^{\otimes 2}(F) = \sum_{b\in\{0,1\}^{n+1}} \left\{\prod_{p=0}^{n} \frac{(N-1)^{1-b_p}}{N}\right\} \cdot \Gamma^b_{n,N}(F) \qquad a.s.$$

and

$$(\eta_n^N)^{\otimes 2}(F) = \sum_{b\in\{0,1\}^{n+1}} \left\{\prod_{p=0}^{n} \frac{(N-1)^{1-b_p}}{N}\right\} \cdot \bar{\Gamma}^b_{n,N}(F) \qquad a.s.$$

Proof. By definition, we have almost surely

$$\begin{aligned}
&\sum_{b\in\{0,1\}^{n+1}}\prod_{p=0}^{n}\left\{\frac{(N-1)^{1-b_p}}{N}\right\}\cdot\bar{\Gamma}^{b}_{n,N}(F)\\
&=\frac{N^{n-1}}{(N-1)^{n+1}}\sum_{b\in\{0,1\}^{n+1}}\sum_{\ell^{[2]}_{0:n}\in((N)^2)^{\times(n+1)}}\prod_{p=0}^{n}\left\{\frac{(N-1)^{1-b_p}}{N}\right\}\prod_{p=0}^{n-1}\lambda^{b}_{p}(A_p^{\ell^{[2]}_{p+1}},\ell^{[2]}_{p})C_{b_n}(F)(X_n^{\ell^{[2]}_n})\\
&=\sum_{\ell^{[2]}_0\in(N)^2}\cdots\sum_{\ell^{[2]}_{n-1}\in(N)^2}\left\{\prod_{p=0}^{n-1}\left(\mathbf{1}_{\{A_p^{\ell^1_{p+1}}=A_p^{\ell^2_{p+1}}=\ell^1_p\neq\ell^2_p\}}\cdot\frac{1}{N}+\mathbf{1}_{\{A_p^{\ell^1_{p+1}}=\ell^1_p\neq A_p^{\ell^2_{p+1}}=\ell^2_p\}}\cdot\frac{N-1}{N}\right)\right\}\\
&\cdot\left(\frac{N}{N-1}\right)^{n}\cdot\left\{\frac{1}{N}\cdot m^{\odot 2}(\mathbf{X_n})C_1(F)+\frac{N-1}{N}\cdot m^{\odot 2}(\mathbf{X_n})C_0(F)\right\}\\
&=\left(\frac{N}{N-1}\right)^{n}\cdot\left(\frac{N-1}{N}\right)^{n}\cdot m^{\otimes 2}(\mathbf{X_n})(F)\\
&=(\eta^N_n)^{\otimes 2}(F)
\end{aligned}$$

By multiplying $\gamma^N_n(1)^2$ for both sides, we have

$$(\gamma^N_n)^{\otimes 2}(F)=\sum_{b\in\{0,1\}^{n+1}}\left\{\prod_{p=0}^{n}\frac{(N-1)^{1-b_p}}{N}\right\}\cdot\Gamma^{b}_{n,N}(F)\qquad a.s.$$

□

Lemma 2.1. *Assume ($\mathcal{A}$1). For any test function $f\in\mathcal{B}_b(E_n)$, we have*

$$NV^N_n(f)-\sum_{p=0}^{n}\left\{\bar{\Gamma}^{(p)}_{n,N}(f^{\otimes 2})-\bar{\Gamma}^{(\emptyset)}_{n,N}(f^{\otimes 2})\right\}=O_{\mathbf{P}}(\frac{1}{N})$$

and

$$NV^N_n(f-\eta^N_n(f))-\sum_{p=0}^{n}\left\{\bar{\Gamma}^{(p)}_{n,N}\left(\left[f-\eta^N_n(f)\right]^{\otimes 2}\right)-\bar{\Gamma}^{(\emptyset)}_{n,N}\left(\left[f-\eta^N_n(f)\right]^{\otimes 2}\right)\right\}=O_{\mathbf{P}}(\frac{1}{N}).$$

Proof. By construction, we have

$$V^N_n(f)=\eta^N_n(f)^2-\Gamma^{(\emptyset)}_{n,N}(f^{\otimes 2}).$$

Then, by the decomposition showed in Proposition 2.1 and the stochastic boundness given in Proposition 3.1, we have

$$\begin{aligned}
&NV^N_n(f)\\
&=N\cdot\left[\bar{\Gamma}^{(\emptyset)}_{n,N}(f^{\otimes 2})\left((\frac{N-1}{N})^{n+1}-1\right)+\left(\frac{N-1}{N}\right)^{n}\cdot\frac{1}{N}\sum_{p=0}^{n}\bar{\Gamma}^{(p)}_{n,N}(f^{\otimes 2})+O_{\mathbf{P}}(\frac{1}{N^2})\right]\\
&=-\bar{\Gamma}^{(\emptyset)}_{n,N}(f^{\otimes 2})\sum_{p=0}^{n}\left(\frac{N-1}{N}\right)^{p}+\left(\frac{N-1}{N}\right)^{n}\sum_{p=0}^{n}\bar{\Gamma}^{(p)}_{n,N}(f^{\otimes 2})+O_{\mathbf{P}}(\frac{1}{N})\\
&=\left(\frac{N-1}{N}\right)^{n}\sum_{p=0}^{n}\left\{\bar{\Gamma}^{(p)}_{n,N}(f^{\otimes 2})-\left(\frac{N}{N-1}\right)^{p}\bar{\Gamma}^{(\emptyset)}_{n,N}(f^{\otimes 2})\right\}+O_{\mathbf{P}}(\frac{1}{N})\\
&=\sum_{p=0}^{n}\left\{\bar{\Gamma}^{(p)}_{n,N}(f^{\otimes 2})-\bar{\Gamma}^{(\emptyset)}_{n,N}(f^{\otimes 2})\right\}+O_{\mathbf{P}}(\frac{1}{N})
\end{aligned}$$

Hence, we have

$$NV_n^N(f) - \sum_{k=0}^{n} \left\{ \bar{\Gamma}_{n,N}^{(k)}(f^{\otimes 2}) - \Gamma_{n,N}^{(\emptyset)}(f^{\otimes 2}) \right\} = O_{\mathbf{P}}(\frac{1}{N}).$$

Similarly, since

$$\bar{\Gamma}_{n,N}^{b} \left(\left[f - \eta_n^N(f) \right]^{\otimes 2} \right) = O_{\mathbf{P}}(1),$$

the same algebraic manipulation yields that

$$NV_n^N(f - \eta_n^N(f)) - \sum_{p=0}^{n} \left\{ \bar{\Gamma}_{n,N}^{(p)} \left(\left[f - \eta_n^N(f) \right]^{\otimes 2} \right) - \bar{\Gamma}_{n,N}^{(\emptyset)} \left(\left[f - \eta_n^N(f) \right]^{\otimes 2} \right) \right\} = O_{\mathbf{P}}(\frac{1}{N}).$$

□

3 Proofs

3.1 Basic convergence results

In this section, we provide some well-known convergence results on SMC framework under our specific parametrizations, i.e., with adaptive potential functions and transition kernels. We concentrate on the properties that do not use the additional information tracked in the genealogy of the associated IPS. Therefore, we give a rougher (classical) definition of the associated IPS without considering the genealogy in order to simplify the story. It is easy to check that the distributions of the particles are identical to the ones defined in Section 1.2.

- $\mathbf{X_0} \sim \eta_0^{\otimes N}$
- For $p \geq 1$, we let

  $$\mathbf{X_p} \sim \bigotimes_{i=1}^{N} K_{p,\eta_{p-1}^N}(X_{p-1}^i, dX_p^i)$$

  with K_{p,η_{p-1}^N} a Markov kernel defined by

  $$\forall (x, A) \in E_{p-1} \times \mathcal{B}(E_p), \qquad K_{p,\eta_{p-1}^N}(x, A) := \frac{\eta_{p-1}^N Q_{p,N}(x, A)}{\eta_{p-1}^N \left(G_{p-1,N} \right)}.$$

Theorem 3.1. *Assume ($\mathcal{A}$1-$\mathcal{A}$2). For any $f \in \mathcal{B}_b(E_n)$, we have*

$$\gamma_n^N(f) \xrightarrow[N\to\infty]{a.s.} \gamma_n(f)$$

and

$$\eta_n^N(f) \xrightarrow[N\to\infty]{a.s.} \eta_n(f).$$

In particular, we also have

$$Z_n^N \xrightarrow[N\to\infty]{a.s.} z_n^*.$$

Proof. By definition, it is clear that the convergence of γ_n^N indicates the convergence of η_n^N. Therefore, it is sufficient to prove the former one. We prove that

$$Z_n^N \xrightarrow[N\to\infty]{a.s.} z_n^* \qquad \text{and} \qquad \forall f \in \mathcal{B}_b(E_n), \quad \gamma_n^N(f) \xrightarrow[N\to\infty]{a.s.} \gamma_n(f)$$

by induction.

Step 0:

The almost surely convergence of γ_0^N to γ_0 w.r.t. a test function in $\mathcal{B}_b(E_0)$ is given by strong law of large numbers. The a.s. convergence of Z_0^N to z_0^* is verified as $\zeta_0^k \in \mathcal{B}_b(E_0)$ for any $k \in [d]$.

Step $n \geq 1$:

We assume that

$$Z_{n-1}^N \xrightarrow[N\to\infty]{a.s.} z_{n-1}^* \tag{8}$$

and for any $\phi \in \mathcal{B}_b(E_{n-1})$, we have

$$\gamma_{n-1}^N(\phi) \xrightarrow[N\to\infty]{a.s.} \gamma_{n-1}(\phi).$$

For any $f \in \mathcal{B}_b(E_n)$, we have

$$\begin{aligned}&\left|\gamma_n^N(f) - \gamma_n(f)\right|\\ &\leq \underbrace{\left|\gamma_n^N(f) - \gamma_{n-1}^N Q_{n,N}(f)\right|}_{P_1(N)} + \underbrace{\left|\gamma_{n-1}^N Q_{n,N}(f) - \gamma_{n-1}^N Q_n(f)\right|}_{P_2(N)} + \underbrace{\left|\gamma_{n-1}^N Q_n(f) - \gamma_{n-1} Q_n(f)\right|}_{P_3(N)}.\end{aligned} \tag{9}$$

- For $P_1(N)$, we denote

  $$U_{n,N}^i := \eta_{n-1}^N(G_{n-1,N}) f(X_n^i) - \eta_{n-1}^N Q_{n,N}(f).$$

  with $U_{n,N}^i := 0$. It is readily checked that

  $$P_1(N) = \gamma_{n-1}^N(1) \cdot \frac{1}{N}\sum_{i=1}^N U_{n,N}^i$$

  We consider a filtration $(\mathcal{E}_n^i)_{0\leq i\leq N}$ defined by

  $$\mathcal{E}_n^i := \mathcal{F}_{n-1}^N \vee \sigma(X_n^1, \ldots, X_n^i)$$

  with $\mathcal{E}_n^0 := \mathcal{F}_{n-1}^N$. By definition, we have

  $$\mathbf{E}\left[U_{n,N}^i \,\middle|\, \mathcal{E}_n^{i-1}\right] = \eta_{n-1}^N(G_{n-1,N}) \cdot \frac{\eta_{n-1}^N Q_{n,N}(f)(x)}{\eta_{n-1}^N\left(G_{n-1,N}\right)} - \eta_{n-1}^N Q_{n,N}(f) = 0. \qquad a.s.$$

  Thus, $(U_{n,N}^i)_{0\leq i\leq N}$ is a $(\mathcal{E}_n^i)_{0\leq i\leq N}$-martingale difference array. Under ($\mathcal{A}1$), we have

  $$\left|U_{n,N}^i\right| \leq C_n := 2\left\|G_{n-1,\cdot}\right\|_\infty \cdot \|f\|_\infty . \qquad a.s.$$

  Therefore, for any $\epsilon > 0$, Heoffding-Azuma inequality gives

  $$\mathbf{P}\left(\left|\sum_{i=1}^N U_{n,N}^i\right| \geq N\cdot\epsilon\right) \leq 2\exp\left(\frac{-N\cdot\epsilon^2}{2\cdot C_n^2}\right)$$

  Consequently, the choice $\epsilon_N := N^{-1/4}$ and Borel-Cantelli Lemma give the following convergence.

  $$\frac{1}{N}\sum_{i=1}^N U_{n,N}^i \xrightarrow[N\to\infty]{a.s.} 0$$

  Combined with induction hypothesis, we have

  $$P_1(N) \xrightarrow[N\to\infty]{a.s.} 0.$$

- Then, we consider the following subset of Ω:

$$\Omega_{n-1} := \left\{ \omega \in \Omega : Z_{n-1}^N \xrightarrow[N\to\infty]{} z_{n-1}^* \right\}.$$

By induction hypothesis, we have $\mathbf{P}(\Omega_{n-1}) = 1$. By applying $\mathcal{A}2$, there exists a function h such that for any $\epsilon > 0$, and for any $\omega \in \Omega_{n-1}$, there exists $N(\epsilon, \omega)$, such that for all $N > N(\epsilon, \omega)$, we have

$$Q_{n,N}(f)(x)(\omega) - Q_n(f)(x) = h_n(x, Z_{n-1}^N(\omega)) \cdot (Z_{n-1}^N(\omega) - z_{n-1}^*).$$

This is equivalent to saying that

$$Q_{n,N}(f)(x) - Q_n(f)(x) = h_n(x, Z_{n-1}^N) \cdot (Z_{n-1}^N - z_{n-1}^*). \qquad a.s.$$

Hence, we deduce that

$$P_2(N) \leq d \cdot \left| \gamma_{n-1}^N(1) \right| \cdot \| h_n \|_\infty \cdot \left| Z_{n-1}^N - z_{n-1}^* \right| \leq d \cdot \| h_n \|_\infty \cdot \left\{ \prod_{q=0}^{p-2} \| G_{q,\cdot} \|_\infty \right\} \cdot \left| Z_{n-1}^N - z_{n-1}^* \right|. \qquad a.s.$$

By induction hypothesis, we conclude that

$$P_2(N) \xrightarrow[N\to\infty]{a.s.} 0.$$

- Under ($\mathcal{A}1$), we have $Q_n(f) \in \mathcal{B}_b(E_{n-1})$. Thus, induction hypothesis gives directly

$$P_3(N) \xrightarrow[N\to\infty]{a.s.} 0.$$

The verification of the convergence

$$\forall f \in \mathcal{B}_b(E_n), \quad \gamma_n^N(f) \xrightarrow[N\to\infty]{a.s.} \gamma_n(f)$$

is then finished. This also give the a.s. convergence of Z_n^N as $\zeta_n^k \in \mathcal{B}_b(E_n)$ for any $k \in [d]$.

$\square$

3.2 Proof of Theorem 1.1

Proof. We prove

$$\forall 0 \leq p \leq n, \quad \left| Z_p^N - z_p^* \right| = O_{\mathbf{P}}(1/\sqrt{N})$$

and

$$\sqrt{N} \left(\gamma_n^N(f) - \gamma_n(f) \right) \xrightarrow[N\to\infty]{\mathrm{d}} \mathcal{N}\left(0, \sigma_{\gamma_n}^2(f)\right)$$

by induction. The verification of step 0 is direct by central limite theorem for i.i.d. random variables. For step $n \geq 1$, we suppose that

$$\forall 0 \leq p \leq n-1, \quad \left| Z_p^N - z_p^* \right| = O_{\mathbf{P}}(1/\sqrt{N})$$

and

$$\sqrt{N} \left(\gamma_{n-1}^N(f) - \gamma_{n-1}(f) \right) \xrightarrow[N\to\infty]{\mathrm{d}} \mathcal{N}\left(0, \sigma_{\gamma_{n-1}}^2(f)\right).$$

For any test function $f \in \mathcal{B}_b(E_n)$, we denote

$$f_p := Q_{p,n}(f).$$

We start with the following decomposition

$$
\begin{aligned}
&\gamma_n^N(f) - \gamma_n(f) \\
&= \sum_{p=0}^{n} \left(\gamma_p^N(f_p) - \gamma_{p-1}^N Q_p(f_p)\right) \\
&= \frac{1}{N} \sum_{p=0}^{n} \sum_{i=1}^{N} \left\{\left(\gamma_p^N(1) f_p(X_p^i) - \gamma_{p-1}^N Q_{p,N}(f_p)\right) + \left(\gamma_{p-1}^N Q_{p,N}(f_p) - \gamma_{p-1}^N Q_p(f_p)\right)\right\}
\end{aligned}
\tag{10}
$$

with the convention $\gamma_{-1}^N = \gamma_0$. For $k \in [(n+1)N]$, we denote

$$
p_k := \left\lfloor \frac{k}{N} \right\rfloor \qquad \text{and} \qquad i_k := k - p_k \times N.
$$

We define the filtration

$$
\forall k \in [(n+1)N], \quad \mathcal{E}_k^N = \mathcal{F}_{p_k}^N \vee \sigma(X_{p_k}^1, \cdots, X_{p_k}^{i_k}).
$$

Then, we denote

$$
U_k^N := \frac{1}{\sqrt{N}} \left(\gamma_{p_k}^N(1) f_{p_k}(X_{p_k}^{i_k}) - \gamma_{p_k-1}^N Q_{p_k,N}(f_{p_k})\right)
$$

and

$$
D_p^N := \sqrt{N} \left(\gamma_{p-1}^N Q_{p,N}(f_p) - \gamma_{p-1}^N Q_p(f_p)\right)
$$

Therefore, we have

$$
\sqrt{N}\left(\gamma_n^N(f) - \gamma_n(f)\right) = \sum_{k=1}^{(n+1)N} \left(U_k^N + \frac{1}{N} \cdot D_{p_k}^N\right) = \sum_{k=1}^{(n+1)N} U_k^N + \sum_{p=0}^{n} D_p^N.
$$

Then, let us consider the subsets of Ω

$$
\Omega_{p-1} = \left\{\omega \in \Omega : Z_{p-1}^N \xrightarrow[N\to\infty]{} z_{p-1}^*\right\}
$$

By $\mathcal{A}2$, we deduce that there exist a function h_p such that for any $\epsilon > 0$ and for any $\omega \in \Omega_{p-1}$, we have

$$
\begin{aligned}
D_p^N(\omega) =& \sqrt{N} \cdot \gamma_{p-1}^N(\omega) \left(h_p(\cdot, Z_{p-1}^N(\omega))\right) \cdot (Z_{p-1}^N(\omega) - z_{p-1}^*). \\
=& \sqrt{N} \cdot \gamma_{p-1}^N(\omega) \left(h_p(\cdot, Z_{p-1}^N(\omega)) - h_p(\cdot, z_{p-1}^*)\right) \cdot (Z_{p-1}^N(\omega) - z_{p-1}^*). \\
&+ \sqrt{N} \cdot \gamma_{p-1}^N(\omega) \left(h_p(\cdot, z_{p-1}^*)\right) \cdot (Z_{p-1}^N(\omega) - z_{p-1}^*).
\end{aligned}
$$

- For the first part, we have

$$
\begin{aligned}
&\sqrt{N} \cdot \gamma_{p-1}^N(\omega) \left(h_p(\cdot, Z_{p-1}^N(\omega)) - h_p(\cdot, z_{p-1}^*)\right) \cdot (Z_{p-1}^N(\omega) - z_{p-1}^*) \\
\leq& \sqrt{N} d \cdot \gamma_{p-1}^N(\omega)(1) \cdot \sup_{x \in E_{n-1}} \left|h_p(x, Z_{p-1}^N(\omega)) - h_p(x, z_{p-1}^*)\right| \cdot \left|Z_{p-1}^N(\omega) - z_{p-1}^*\right|,
\end{aligned}
$$

  which amounts to say that

$$
\begin{aligned}
&\sqrt{N} \gamma_{p-1}^N \left(h_p(\cdot, Z_{p-1}^N) - h_p(\cdot, z_{p-1}^*)\right) \cdot (Z_{p-1}^N - z_{p-1}^*) \\
\leq& d \cdot \sqrt{N} \left|Z_{p-1}^N - z_{p-1}^*\right| \cdot \gamma_{p-1}^N(1) \cdot \sup_{x \in E_{n-1}} \left|h_p(\cdot, Z_{p-1}^N) - h_p(\cdot, z_{p-1}^*)\right| \qquad a.s.
\end{aligned}
$$

Since $z \mapsto h_p(x,z)$ is continuous uniformly over $x \in E_{p-1}$, we have, for any $\epsilon > 0$, there exists $N(\omega, \epsilon) > 0$ such that for all $N > N(\omega, \epsilon)$, we have

$$\left| h_p(\cdot, Z_{p-1}^N(\omega)) - h_p(\cdot, z_{p-1}^*) \right| \le \epsilon.$$

This verifies that

$$\sup_{x \in E_{n-1}} \left| h_p(\cdot, Z_{p-1}^N) - h_p(\cdot, z_{p-1}^*) \right| \xrightarrow[N\to\infty]{a.s.} 0.$$

Then, as is guaranteed by the induction hypothesis, we have

$$\sqrt{N} \left| Z_{p-1}^N - z_{p-1}^* \right| = \mathcal{O}_{\mathbf{P}}(1), \tag{11}$$

from which we deduce that

$$\sqrt{N} \cdot \gamma_{p-1}^N \left(h_p(\cdot, Z_{p-1}^N) - h_p(\cdot, z_{p-1}^*) \right) \cdot (Z_{p-1}^N - z_{p-1}^*) \xrightarrow[N\to\infty]{\mathbf{P}} 0.$$

- For the second part, since $\mathcal{A}2$ and Theorem 3.1 imply that

$$\gamma_{p-1}^N(h_p(\cdot, z_{p-1}^*)) \xrightarrow[N\to\infty]{a.s.} \gamma_{p-1}(h_p(\cdot, z_{p-1}^*)) = 0,$$

we conclude, by induction hypothesis (11), that

$$\sqrt{N} \cdot \gamma_{p-1}^N \left(h_p(\cdot, z_{p-1}^*) \right) \cdot (Z_{p-1}^N - z_{p-1}^*) \xrightarrow[N\to\infty]{\mathbf{P}} 0.$$

Finally, we conclude that

$$D_p^N \xrightarrow[N\to\infty]{\mathbf{P}} 0,$$

which yields that

$$\sum_{p=0}^{n} D_p^N \xrightarrow[N\to\infty]{\mathbf{P}} 0.$$

Next, it is easy to check that $(U_k^N)_{1\le k\le (n+1)N}$ is a $(\mathcal{E}_k^N)_{1\le k\le (n+1)N}$-martingale difference array. By Theorem 2.3 in [10], we verify that

- Since $\mathcal{A}1$ gives that

$$\max_{1\le k\le (n+1)N} (U_k^N) \le \frac{2}{\sqrt{N}} \cdot \|f\|_\infty \cdot \sum_{p=1}^{n} \prod_{q=0}^{p-1} \left\| G_{q,\cdot} \right\|_\infty \qquad a.s., \tag{12}$$

it is also uniformly bounded in L^2-norm.

- By (12), we also have

$$\max_{1\le k\le (n+1)N} (U_k^N) \xrightarrow[N\to\infty]{\mathbf{P}} 0.$$

- Standard calculation gives that

$$\begin{aligned} &\sum_{k=1}^{(n+1)N} \left(U_k^N \right)^2 \\ &= \sum_{p=0}^{n} \left(\gamma_p^N(1)^2 \eta_p^N(f_p)^2 + \gamma_{p-1}^N Q_{p,N}(f_p)^2 - 2 \cdot \gamma_p^N(1) \eta_p^N(f_p) \cdot \gamma_{p-1}^N Q_{p,N}(f_p) \right). \end{aligned} \tag{13}$$

As is shown above, the convergence of D_p^N indicates that

$$\gamma_{p-1}^N Q_{p,N}(f_p) - \gamma_{p-1}^N Q_p(f_p) \xrightarrow[N\to\infty]{\mathbf{P}} 0.$$

Then, by applying Theorem 3.1, we obtain that

$$\sum_{k=1}^{(n+1)N} \left(U_k^N\right)^2 \xrightarrow[N\to\infty]{\mathbf{P}} \sigma_{\gamma_n}^2(f).$$

Therefore, we have the following central limit theorem

$$\sum_{k=1}^{(n+1)N} U_k^N \xrightarrow[N\to\infty]{\mathrm{d}} \mathcal{N}\left(0, \sigma_{\gamma_n}^2(f)\right).$$

The conclusion follows by Slutsky's Lemma. Finally, the case for the normalized measure η_n^N is easily deduced by the following decomposition and Slutsky's Lemma.

$$\eta_n^N(f) - \eta_n(f) = \frac{\gamma_n(1)}{\gamma_n^N(1)} \cdot \frac{1}{\gamma_n(1)} \cdot \sqrt{N}\, \gamma_n^N\left(f - \eta_n(f)\right)$$

□

3.3 Proof of Theorem 2.1

Before starting, we introduce some additional notation. Fixing two test functions $\phi, \psi \in \mathcal{B}_b(E_n)$ and coalescence indicator b. With a slight abuse of notation, for a coalescence indicator $b \in \{0,1\}^{n+1} = (b_1, \ldots, b_n)$, we denote, for all $0 \le p \le n$, that

$$\Gamma_p^b := \Gamma_p^{(b_1,\ldots,b_p)} \qquad \text{and} \qquad \Gamma_{p,N}^b := \Gamma_{p,N}^{(b_1,\ldots,b_p)}$$

with the convention

$$\Gamma_{-1,N}^b := \eta_0^{\otimes 2} C_{b_0}.$$

We also remark for $b_n \in \{0,1\}$, and for any $\phi, \psi \in \mathcal{B}_b(E_n)$, there exists $f, g \in \mathcal{B}_b(E_{n-1})$, such that

$$Q_{n,z}^{\otimes 2} C_{b_n}(\phi \otimes \psi) = f \otimes g.$$

In fact, for $b_n = 0$, we take

$$f = Q_{n,z}(\phi) \quad \text{and} \quad g = Q_{n,z}(\psi).$$

Similarly, for $b_n = 1$, we take

$$f = Q_{n,z}(\phi \cdot \psi) \quad \text{and} \quad g = Q_{n,z}(1).$$

Proof. The proof will be done by induction.

step 0:

- For $b_0 = 1$, the weak law of large numbers gives

$$\frac{1}{N} \sum_{i=1}^{N} \phi(X_0^i)\psi(X_0^i) \xrightarrow[N\to\infty]{\mathbf{P}} \eta_0(\phi \cdot \psi).$$

- For $b_0 = 0$, we have

$$\frac{1}{N^2} \sum_{i \neq j} \phi(X_0^i)\psi(X_0^j) = \frac{1}{N} \sum_{i=i}^{N} \phi(X_0^i) \cdot \frac{1}{N} \sum_{j=i}^{N} \psi(X_0^j) - \frac{1}{N} \cdot \frac{1}{N} \sum_{i=1}^{N} \phi(X_0^i)\psi(X_0^i).$$

  Hence, by Weak Law of Large Numbers, we have

$$\frac{1}{N(N-1)} \sum_{i \neq j} \phi(X_0^i)\psi(X_0^j) \xrightarrow[N\to\infty]{\mathbf{P}} \eta_0(\phi)\eta_0(\psi).$$

step $n \geq 1$:

We suppose that for any test functions $f, g \in \mathcal{B}_b(E_{n-1})$ and coalescence indicator b, we have

$$\Gamma_{n-1,N}^{b}(f \otimes g) \xrightarrow[N\to\infty]{\mathbf{P}} \Gamma_{n-1}^{b}(f \otimes g).$$

Then, we consider the following decomposition

$$\begin{aligned}
&\Gamma_{n,N}^{b}(\phi \otimes \psi) - \Gamma_{n}^{b}(\phi \otimes \psi) \\
&= \underbrace{\Gamma_{n,N}^{b}(\phi \otimes \psi) - \Gamma_{n-1,N}^{b} Q_{n,N}^{\otimes 2} C_{b_n}(\phi \otimes \psi)}_{R_1(N)} + \underbrace{\Gamma_{n-1,N}^{b} Q_{n,N}^{\otimes 2} C_{b_n}(\phi \otimes \psi) - \Gamma_{n-1,N}^{b} Q_{n}^{\otimes 2} C_{b_n}(\phi \otimes \psi)}_{R_2(N)} \\
&\qquad + \underbrace{\Gamma_{n-1,N}^{b} Q_{n}^{\otimes 2} C_{b_n}(\phi \otimes \psi) - \Gamma_{n-1}^{b} Q_{n}^{\otimes 2} C_{b_n}(\phi \otimes \psi)}_{R_3(N)}
\end{aligned}$$

- The convergence
$$R_1(N) \xrightarrow[N\to\infty]{\mathbf{P}} 0$$
is given by Lemma 3.1.
- The convergence
$$R_2(N) \xrightarrow[N\to\infty]{\mathbf{P}} 0$$
is given by Lemma 3.4.
- The convergence
$$R_3(N) \xrightarrow[N\to\infty]{\mathbf{P}} 0$$
is a direct consequence of induction hypothesis.

The proof is then finished. □

3.4 Technical results

In this section, we list some supporting technical results. Before going further, we recall the definition of $\Gamma_{n,N}^{b}(1)$:

$$\Gamma_{n,N}^{b}(1) := \gamma_n^N(1)^2 \cdot \frac{N^{n-1}}{(N-1)^{n+1}} \sum_{\ell_{0:n}^{[2]} \in ((N)^2)^{\times(n+1)}} \prod_{p=0}^{n-1} \left\{ \lambda_p^b(A_p^{\ell_{p+1}^{[2]}}, \ell_p^{[2]}) \right\}$$

We denote

$$\Lambda_n^{\ell_n^{[2]}} := \sum_{\ell_{0:n-1}^{[2]} \in ((N)^2)^{\times n}} \prod_{p=0}^{n-1} \left\{ \lambda_p^b(A_p^{\ell_{p+1}^{[2]}}, \ell_p^{[2]}) \right\} \tag{14}$$

with the convention $\Lambda_0^{\ell_0^{[2]}} := 1$. Hence, we have

$$\Gamma_{n,N}^{b}(1) := \gamma_n^N(1)^2 \cdot \frac{N^{n-1}}{(N-1)^{n+1}} \sum_{\ell_n^{[2]} \in (N)^2} \Lambda_n^{\ell_n^{[2]}}. \tag{15}$$

We remark that (15) covers the case when $n = 0$. Then, by definition, we have, for $n \geq 1$, that

$$\Lambda_n^{\ell_n^{[2]}} = \sum_{\ell_{n-1}^{[2]} \in (N)^2} \Lambda_{n-1}^{\ell_{n-1}^{[2]}} \cdot \lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}). \tag{16}$$

This decomposition will appear several times in our proofs. This is one of the core techniques we have implemented in order to study the behavior of the coalescent tree-typed measures.

Proposition 3.1. *Assume ($\mathcal{A}$1). For any coalescence indicator b, we have*

$$\sup_{N>1} \mathbf{E}\left[\Gamma^b_{n,N}(1)^2\right] < +\infty$$

Remark. By considering, for e.g., the Markov inequality, the L^2 bound also verifies the stochastic boundness of the coalescent tree-typed measures. One have

$$\forall f, g \in \mathcal{B}_b(E_n), \quad \Gamma^b_{n,N}(f \otimes g) = \mathcal{O}_{\mathbf{P}}(1).$$

In addition, we also want to mention that the analysis for the case 1, and case 2 in the proof is relatively rough, which may gives some wrong intuition for understanding of our strategy. Although, it is enough to have the boundness for our reasoning. A similar but more careful analysis can be found in the proof of Lemma 3.3, where the corresponding conditional expectations parts found in (18) are proved to be almost surely negligible with respect to N.

Proof. We give a proof by induction. The verification for step 0 is trivial as $\Gamma^b_{0,N}(1) = 1$. For $n \geq 1$, we suppose that

$$\sup_{N>1} \mathbf{E}\left[\Gamma^b_{n-1,N}(1)^2\right] < +\infty$$

By combining (15) and (16), we have

$$\begin{aligned}
&\mathbf{E}\left[\Gamma^b_{n,N}(1)^2 \,\middle|\, \mathcal{G}^N_{n-1}\right] \\
&\quad = \gamma^N_n(1)^4 \cdot \left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2 \sum_{(\ell^{[2]}_n, \ell'^{[2]}_n) \in ((N)^2)^{\times 2}} \ \sum_{(\ell^{[2]}_{n-1}, \ell'^{[2]}_{n-1}) \in ((N)^2)^{\times 2}} \Lambda^{\ell^{[2]}_{n-1}}_{n-1} \Lambda^{\ell'^{[2]}_{n-1}}_{n-1} \\
&\qquad \cdot \mathbf{E}\left[\lambda^b_{n-1}(A^{\ell^{[2]}_n}_{n-1}, \ell^{[2]}_{n-1}) \lambda^b_{n-1}(A^{\ell'^{[2]}_n}_{n-1}, \ell'^{[2]}_{n-1}) \,\middle|\, \mathcal{G}^N_{n-1}\right]
\end{aligned} \tag{17}$$

Now, for $N > 4$, we consider the decomposition

$$\begin{aligned}
&((N)^2)^{\times 2} \\
&= \left(((N)^2)^{\times 2} \cap [N]^4_1\right) \cup \left(((N)^2)^{\times 2} \cap [N]^4_2\right) \cup \left(((N)^2)^{\times 2} \cap [N]^4_3\right) \cup \left(((N)^2)^{\times 2} \cap [N]^4_4\right) \\
&= \left(((N)^2)^{\times 2} \cap [N]^4_2\right) \cup \left(((N)^2)^{\times 2} \cap [N]^4_3\right) \cup (N)^4.
\end{aligned}$$

- Case 1: $(\ell^{[2]}_n, \ell'^{[2]}_n) \in ((N)^2)^{\times 2} \cap [N]^4_2$.
  By the definition of $S_{n-1,N}$ in (2), we have

$$\mathbf{E}\left[\lambda^b_{n-1}(A^{\ell^{[2]}_n}_{n-1}, \ell^{[2]}_{n-1}) \lambda^b_{n-1}(A^{\ell'^{[2]}_n}_{n-1}, \ell'^{[2]}_{n-1}) \,\middle|\, \mathcal{G}^N_{n-1}\right] \leq \left(\frac{\|G_{n-1,\cdot}\|_\infty}{N \cdot m(X_{n-1})(G_{n-1,N})}\right)^2 \qquad a.s.$$

  Meanwhile, we noticed that

$$\#\left(((N)^2)^{\times 2} \cap [N]^4_2\right) = 2N(N-1).$$

  Hence, we have

$$\begin{aligned}
&\sum_{(\ell^{[2]}_n, \ell'^{[2]}_n) \in ((N)^2)^{\times 2} \cap [N]^4_2} m(X_{n-1})(G_{n-1,N})^4 \cdot \mathbf{E}\left[\lambda^b_{n-1}(A^{\ell^{[2]}_n}_{n-1}, \ell^{[2]}_{n-1}) \lambda^b_{n-1}(A^{\ell'^{[2]}_n}_{n-1}, \ell'^{[2]}_{n-1}) \,\middle|\, \mathcal{G}^N_{n-1}\right] \\
&\qquad \leq \frac{2N(N-1)}{N^2} \cdot \|G_{n-1,\cdot}\|^4_\infty \qquad a.s.
\end{aligned}$$

- Case 2: $(\ell_n^{[2]}, \ell_n^{'[2]}) \in ((N)^2)^{\times 2} \cap [N]_3^4$.
  Similarly, we have

$$\mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]})\lambda_{n-1}^b(A_{n-1}^{\ell_n^{'[2]}}, \ell_{n-1}^{'[2]}) \,\middle|\, \mathcal{G}_{n-1}^N\right] \le \left(\frac{\|G_{n-1,\cdot}\|_\infty}{N \cdot m(X_{n-1})(G_{n-1,N})}\right)^3 \qquad a.s.$$

  and

$$\#\left(((N)^2)^{\times 2} \cap [N]_3^4\right) = 4N(N-1)(N-2).$$

  These imply that

$$\sum_{(\ell_n^{[2]}, \ell_n^{'[2]}) \in ((N)^2)^{\times 2} \cap [N]_3^4} m(X_{n-1})(G_{n-1,N})^4 \cdot \mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]})\lambda_{n-1}^b(A_{n-1}^{\ell_n^{'[2]}}, \ell_{n-1}^{'[2]}) \,\middle|\, \mathcal{G}_{n-1}^N\right]$$
$$\le \frac{4N(N-1)(N-2)}{N^3} \cdot \|G_{n-1,\cdot}\|_\infty^4 \qquad a.s.$$

- Case 3: $(\ell_n^{[2]}, \ell_n^{'[2]}) \in (N)^4$.
  We have

$$\mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]})\lambda_{n-1}^b(A_{n-1}^{\ell_n^{'[2]}}, \ell_{n-1}^{'[2]}) \,\middle|\, \mathcal{G}_{n-1}^N\right] \le \left(\frac{\|G_{n-1,\cdot}\|_\infty}{N \cdot m(X_{n-1})(G_{n-1,N})}\right)^4 \qquad a.s.$$

  and

$$\#\left((N)^4\right) = N(N-1)(N-2)(N-3).$$

  Again, we have

$$\sum_{(\ell_n^{[2]}, \ell_n^{'[2]}) \in (N)^4} m(X_{n-1})(G_{n-1,N})^4 \cdot \mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]})\lambda_{n-1}^b(A_{n-1}^{\ell_n^{'[2]}}, \ell_{n-1}^{'[2]}) \,\middle|\, \mathcal{G}_{n-1}^N\right]$$
$$\le \frac{N(N-1)(N-2)(N-3)}{N^4} \cdot \|G_{n-1,\cdot}\|_\infty^4 \qquad a.s.$$

As a consequence, we have almost surely

$$\begin{aligned}
&\mathbf{E}\left[\Gamma_{n,N}^b(1)^2 \mid \mathcal{G}_{n-1}^N\right] \\
\le&\gamma_{n-1}^N(1)^4 \cdot \left(\frac{N^{n-2}}{(N-1)^n}\right)^2 \sum_{(\ell_{n-1}^{[2]}, \ell_{n-1}^{'[2]}) \in ((N)^2)^{\times 2}} \Lambda_{n-1}^{\ell_{n-1}^{[2]}} \Lambda_{n-1}^{\ell_{n-1}^{'[2]}} \\
&\cdot \frac{N}{N-1}\left(2 \cdot \|G_{n-1,\cdot}\|_\infty^4 + \frac{4(N-2)}{N} \cdot \|G_{n-1,\cdot}\|_\infty^4 + \frac{(N-2)(N-3)}{N^2} \cdot \|G_{n-1,\cdot}\|_\infty^4\right) \\
=&\Gamma_{n-1,N}^b(1)^2 \cdot \frac{N}{N-1} \cdot \left(2 \cdot \|G_{n-1,\cdot}\|_\infty^4 + \frac{4(N-2)}{N} \cdot \|G_{n-1,\cdot}\|_\infty^4 + \frac{(N-2)(N-3)}{N^2} \cdot \|G_{n-1,\cdot}\|_\infty^4\right)
\end{aligned} \tag{18}$$

Finally, by applying induction hypothesis, we have

$$\sup_{N>4} \mathbf{E}\left[\Gamma_{n,N}^b(1)^2\right] = \sup_{N>4} \mathbf{E}\left[\mathbf{E}\left[\Gamma_{n,N}^b(1)^2 \mid \mathcal{G}_{n-1}^N\right]\right] \le 7 \cdot \frac{5}{4} \cdot \|G_{n-1,\cdot}\|_\infty^4 \cdot \sup_{N>4} \mathbf{E}\left[\Gamma_{n-1,N}^b(1)^2\right] < +\infty,$$

which gives

$$\sup_{N>1} \mathbf{E}\left[\Gamma_{n,N}^b(1)^2\right] < +\infty.$$

$\square$

Lemma 3.1. *Assume ($\mathcal{A}1$). For any test function $f, g \in \mathcal{B}_b(E_n)$, we have, for all $n \geq 1$, that*

$$\mathbf{E}\left[\Gamma^b_{n,N}(f\otimes g)\,\middle|\,\mathcal{G}^N_{n-1}\right] = \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g) \qquad a.s. \tag{19}$$

and

$$\Gamma^b_{n,N}(f\otimes g) - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g) = o_{\mathbf{P}}(1).$$

Proof. By exploiting the notation defined in (14), we have

$$\Gamma^b_{n,N}(f\otimes g) := \gamma^N_n(1)^2\cdot\frac{N^{n-1}}{(N-1)^{n+1}}\sum_{\ell^{[2]}_n\in(N)^2}\Lambda^{\ell^{[2]}_n}_n\cdot f(X^{\ell^1_n}_n)g(X^{\ell^2_n}_n).$$

In fact, (19) is a direct consequence of Lemma A.2. Thanks to Chebyshev's inequality, it is sufficient to verify that

$$\mathrm{Var}\left[\Gamma^b_{n,N}(f\otimes g) - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g)\right] \xrightarrow[N\to\infty]{} 0 \tag{20}$$

for the convergence in probability. By the decomposition of conditional variance, we have

$$\begin{aligned}
&\mathrm{Var}\left[\Gamma^b_{n,N}(f\otimes g) - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g)\right]\\
&\qquad= \mathbf{E}\left[\mathrm{Var}\left[\Gamma^b_{n,N}(f\otimes g) - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g)\mid\mathcal{G}^N_{n-1}\right]\right]\\
&\qquad\qquad+ \mathrm{Var}\left[\mathbf{E}\left[\Gamma^b_{n,N}(f\otimes g) - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g)\mid\mathcal{G}^N_{n-1}\right]\right]
\end{aligned}$$

Hence, (19) yields that

$$\begin{aligned}
&\mathrm{Var}\left[\Gamma^b_{n,N}(f\otimes g) - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g)\right]\\
&\qquad= \mathbf{E}\left[\mathrm{Var}\left[\Gamma^b_{n,N}(f\otimes g) - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g)\mid\mathcal{G}^N_{n-1}\right]\right]\\
&\qquad\qquad= \mathbf{E}\left[\mathbf{E}\left[\Gamma^b_{n,N}(f\otimes g)^2 - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g)^2\mid\mathcal{G}^N_{n-1}\right]\right]
\end{aligned}$$

Finally, (20) is verified by Lemma 3.2. ◻

Lemma 3.2. *Assume ($\mathcal{A}1$). For any test function $f, g \in \mathcal{B}_b(E_n)$, we have, for all $n \geq 1$, that*

$$\mathbf{E}\left[\Gamma^b_{n,N}(f\otimes g)^2 - \Gamma^b_{n-1,N}Q^{\otimes 2}_{n,N}C_{b_n}(f\otimes g)^2\right] \xrightarrow[N\to\infty]{} 0.$$

Proof. We recall that we have, by definition, that

$$\begin{aligned}
&\Gamma^b_{n,N}(f\otimes g)^2\\
&=\gamma^N_n(1)^4\cdot\left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2\sum_{(\ell^{[2]}_n,\ell'^{[2]}_n)\in((N)^2)^{\times 2}}\Lambda^{\ell^{[2]}_n}_n\Lambda^{\ell'^{[2]}_n}_n\cdot f(X^{\ell^1_n}_n)g(X^{\ell^2_n}_n)\cdot f(X^{\ell'^1_n}_n)g(X^{\ell'^2_n}_n).\\
&=\underbrace{\gamma^N_n(1)^4\cdot\left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2\sum_{(\ell^{[2]}_n,\ell'^{[2]}_n)\in(N)^4}\Lambda^{\ell^{[2]}_n}_n\Lambda^{\ell'^{[2]}_n}_n\cdot f(X^{\ell^1_n}_n)g(X^{\ell^2_n}_n)\cdot f(X^{\ell'^1_n}_n)g(X^{\ell'^2_n}_n)}_{R_1(N)}\\
&\quad+\underbrace{\gamma^N_n(1)^4\cdot\left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2\sum_{(\ell^{[2]}_n,\ell'^{[2]}_n)\in((N)^2)^{\times 2}\setminus(N)^4}\Lambda^{\ell^{[2]}_n}_n\Lambda^{\ell'^{[2]}_n}_n\cdot f(X^{\ell^1_n}_n)g(X^{\ell^2_n}_n)\cdot f(X^{\ell'^1_n}_n)g(X^{\ell'^2_n}_n)}_{R_2(N)}.
\end{aligned}$$

- For $R_1(N)$, we prove that

$$\mathbf{E}\left[R_1(N) - \Gamma^b_{n-1,N} Q^{\otimes 2}_{n,N} C_{b_n}(f\otimes g)^2\right] \xrightarrow[N\to\infty]{} 0. \tag{21}$$

In fact, for any $(\ell^{[2]}_n, \ell'^{[2]}_n) \in (N)^4$, we have

$$(A^{\ell^1_n}_n, X^{\ell^1_n}_n, A^{\ell^1_n}_n, X^{\ell^2_n}_n) \qquad \text{and} \qquad (A^{\ell'^1_n}_n, X^{\ell'^1_n}_n, A^{\ell'^1_n}_n, X^{\ell'^2_n}_n)$$

are conditionally independent given $\mathcal{G}^N_{n-1}$ by the construction of IPS. Hence, by applying Lemma A.2 for $\ell^{[2]}_n$ and respectively for $\ell'^{[2]}_n$, we have

$$\mathbf{E}\left[\gamma^N_n(1)^4 \cdot \left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2 \Lambda^{\ell^{[2]}_n}_n \Lambda^{\ell'^{[2]}_n}_n \cdot f(X^{\ell^1_n}_n) g(X^{\ell^2_n}_n) \cdot f(X^{\ell'^1_n}_n) g(X^{\ell'^2_n}_n) \,\middle|\, \mathcal{G}^N_{n-1}\right]$$
$$= \frac{1}{N^2(N-1)^2} \cdot \Gamma^N_{n-1} Q^{\otimes 2}_{n,N} C_{b_n}(f\otimes g)^2.$$

This equality can also be obtained by doing the calculation with Lemma A.1 for $q = 4$ and for $(\ell^1_n, \ell^2_n, \ell'^1_n, \ell'^2_n)$, in a similar way as in the proof of Lemma A.2, which is more consistent with our techniques but more lengthy. Hence, we deduce that

$$\mathbf{E}\left[R_1(N) - \Gamma^b_{n-1,N} Q^{\otimes 2}_{n,N} C_{b_n}(f\otimes g)^2 \,\middle|\, \mathcal{G}^N_{n-1}\right]$$
$$= \left(\frac{N(N-1)(N-2)(N-3)}{N^2(N-1)^2} - 1\right) \cdot \Gamma^N_{n-1} Q^{\otimes 2}_{n,N} C_{b_n}(f\otimes g)^2 \xrightarrow[N\to\infty]{a.s.} 0.$$

Thanks to Proposition 3.1 and dominated convergence theorem, we finally have the convergence given in (21).

- For $R_2(N)$, since

$$\Lambda^{\ell^{[2]}_n}_n \geq 0 \qquad a.s.,$$

we deduce that

$$\mathbf{E}[R_2(N)] \leq \mathbf{E}\left[\gamma^N_n(1)^4 \cdot \left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2 \sum_{(\ell^{[2]}_n, \ell'^{[2]}_n) \in ((N)^2)^{\times 2} \setminus (N)^4} \Lambda^{\ell^{[2]}_n}_n \Lambda^{\ell'^{[2]}_n}_n\right] \cdot \|f\|^2_\infty \cdot \|g\|^2_\infty .$$

Hence, thanks to Lemma 3.3, one has

$$\mathbf{E}[R_2(N)] \xrightarrow[N\to\infty]{} 0.$$

The proof is then complete. □

Lemma 3.3. *Assume ($\mathcal{A}$1). For any test functions $f, g \in \mathcal{B}_b(E_n)$, we have*

$$\mathbf{E}\left[\gamma^N_n(1)^4 \cdot \left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2 \sum_{(\ell^{[2]}_n, \ell'^{[2]}_n) \in ((N)^2)^{\times 2} \setminus (N)^4} \Lambda^{\ell^{[2]}_n}_n \Lambda^{\ell'^{[2]}_n}_n\right] \xrightarrow[N\to\infty]{} 0.$$

Proof. The proof will be done by induction. At step 0, we have

$$\frac{1}{N^2(N-1)^2} \sum_{(\ell^{[2]}_0, \ell'^{[2]}_0) \in ((N)^2)^{\times 2} \setminus (N)^4} 1 = 1 - \frac{N(N-1)(N-2)(N-3)}{N^2(N-1)^2} \xrightarrow[N\to\infty]{} 0.$$

The convergence of its expectation is also verified. For step $n \geq 1$, we suppose that

$$\mathbf{E}\left[\gamma_{n-1}^N(1)^4 \cdot \left(\frac{N^{n-2}}{(N-1)^n}\right)^2 \sum_{(\ell_{n-1}^{[2]}, \ell_{n-1}^{\prime[2]}) \in ((N)^2)^{\times 2} \setminus (N)^4} \Lambda_{n-1}^{\ell_{n-1}^{[2]}} \Lambda_{n-1}^{\ell_{n-1}^{\prime[2]}}\right] \xrightarrow[N \to \infty]{} 0.$$

Let us consider the following decomposition:

$$\begin{aligned}
&\mathbf{E}\left[\gamma_n^N(1)^4 \cdot \left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2 \sum_{(\ell_n^{[2]}, \ell_n^{\prime[2]}) \in ((N)^2)^{\times 2} \setminus (N)^4} \Lambda_n^{\ell_n^{[2]}} \Lambda_n^{\ell_n^{\prime[2]}} \;\middle|\; \mathcal{G}_{n-1}^N\right] \\
&\qquad = \gamma_n^N(1)^4 \cdot \left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2 \sum_{(\ell_{n-1}^{[2]}, \ell_{n-1}^{\prime[2]}) \in ((N)^2)^{\times 2}} \Lambda_{n-1}^{\ell_{n-1}^{[2]}} \Lambda_{n-1}^{\ell_{n-1}^{\prime[2]}} \\
&\qquad\quad \cdot \sum_{(\ell_n^{[2]}, \ell_n^{\prime[2]}) \in ((N)^2)^{\times 2} \setminus (N)^4} \mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}) \lambda_{n-1}^b(A_{n-1}^{\ell_n^{\prime[2]}}, \ell_{n-1}^{\prime[2]}) \;\middle|\; \mathcal{G}_{n-1}^N\right]
\end{aligned}$$

Now, for $N > 4$, we consider the decomposition

$$((N)^2)^{\times 2} \setminus (N)^4 = \left(((N)^2)^{\times 2} \cap [N]_2^4\right) \cup \left(((N)^2)^{\times 2} \cap [N]_3^4\right)$$

- Case 1: $(\ell_n^{[2]}, \ell_n^{\prime[2]}) \in ((N)^2)^{\times 2} \cap [N]_2^4$.

  By the definition of $S_{n-1,N}$ in (2), we have

$$\begin{aligned}
&\mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}) \lambda_{n-1}^b(A_{n-1}^{\ell_n^{\prime[2]}}, \ell_{n-1}^{\prime[2]}) \;\middle|\; \mathcal{G}_{n-1}^N\right] \\
&\leq \left(\frac{\left\|G_{n-1,\cdot}\right\|_\infty}{N \cdot m(X_{n-1})(G_{n-1,N})}\right)^2 \cdot \left(\mathbf{1}_{\{b_{n-1}=1, \ell_{n-1}^1 = \ell_{n-1}^{\prime 1}\}} + \mathbf{1}_{\{b_{n-1}=0, \ell_{n-1}^1 = \ell_{n-1}^{\prime 1}, \ell_{n-1}^2 = \ell_{n-1}^{\prime 2}\}}\right) \\
&\leq \left(\frac{\left\|G_{n-1,\cdot}\right\|_\infty}{N \cdot m(X_{n-1})(G_{n-1,N})}\right)^2 \cdot \mathbf{1}_{\{\#\{\ell_{n-1}^1, \ell_{n-1}^{\prime 1}, \ell_{n-1}^2, \ell_{n-1}^{\prime 2}\} < 4\}}. \qquad a.s.
\end{aligned}$$

  Meanwhile, we noticed that

$$\#\left(((N)^2)^{\times 2} \cap [N]_2^4\right) = 2N(N-1).$$

  Hence, we have

$$\begin{aligned}
&\sum_{(\ell_n^{[2]}, \ell_n^{\prime[2]}) \in ((N)^2)^{\times 2} \cap [N]_2^4} m(X_{n-1})(G_{n-1,N})^4 \cdot \mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}) \lambda_{n-1}^b(A_{n-1}^{\ell_n^{\prime[2]}}, \ell_{n-1}^{\prime[2]}) \;\middle|\; \mathcal{G}_{n-1}^N\right] \\
&\qquad \leq \frac{2N(N-1)}{N^2} \cdot \left\|G_{n-1,\cdot}\right\|_\infty^4 \cdot \mathbf{1}_{\{\#\{\ell_{n-1}^1, \ell_{n-1}^{\prime 1}, \ell_{n-1}^2, \ell_{n-1}^{\prime 2}\} < 4\}}. \qquad a.s.
\end{aligned}$$

- Case 2: $(\ell_n^{[2]}, \ell_n^{\prime[2]}) \in ((N)^2)^{\times 2} \cap [N]_3^4$.

  First, we suppose that $\ell_n^1 = \ell_n^{\prime 1}$. Similar to the former case, we have

$$\begin{aligned}
&\mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}) \lambda_{n-1}^b(A_{n-1}^{\ell_n^{\prime[2]}}, \ell_{n-1}^{\prime[2]}) \;\middle|\; \mathcal{G}_{n-1}^N\right] \\
&\leq \left(\frac{\left\|G_{n-1,\cdot}\right\|_\infty}{N \cdot m(X_{n-1})(G_{n-1,N})}\right)^3 \cdot \left(\mathbf{1}_{\{b_{n-1}=1, \ell_{n-1}^1 = \ell_{n-1}^{\prime 1}\}} + \mathbf{1}_{\{b_{n-1}=0, \ell_{n-1}^1 = \ell_{n-1}^{\prime 1}\}}\right) \\
&\leq \left(\frac{\left\|G_{n-1,\cdot}\right\|_\infty}{N \cdot m(X_{n-1})(G_{n-1,N})}\right)^3 \cdot \mathbf{1}_{\{\#\{\ell_{n-1}^1, \ell_{n-1}^{\prime 1}, \ell_{n-1}^2, \ell_{n-1}^{\prime 2}\} < 4\}}. \qquad a.s.
\end{aligned}$$

By the same reasoning, we have, for $\ell_n^1 = \ell_n'^2$, $\ell_n^2 = \ell_n'^1$ and $\ell_n^2 = \ell_n'^2$, we also have

$$\begin{aligned}
&\mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]})\lambda_{n-1}^b(A_{n-1}^{\ell_n'^{[2]}}, \ell_{n-1}'^{[2]}) \,\middle|\, \mathcal{G}_{n-1}^N\right] \\
&\leq \left(\frac{\|G_{n-1,\cdot}\|_\infty}{N \cdot m(X_{n-1})(G_{n-1,N})}\right)^3 \cdot \mathbf{1}_{\{\#\{\ell_{n-1}^1, \ell_{n-1}'^1, \ell_{n-1}^2, \ell_{n-1}'^2\}<4\}}. \qquad a.s.
\end{aligned}$$

In addition, since

$$\#\left(((N)^2)^{\times 2} \cap [N]_3^4\right) = 4N(N-1)(N-2).$$

We deduce that

$$\begin{aligned}
&\sum_{(\ell_n^{[2]}, \ell_n'^{[2]}) \in ((N)^2)^{\times 2} \cap [N]_3^4} m(X_{n-1})(G_{n-1,N})^4 \cdot \mathbf{E}\left[\lambda_{n-1}^b(A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]})\lambda_{n-1}^b(A_{n-1}^{\ell_n'^{[2]}}, \ell_{n-1}'^{[2]}) \,\middle|\, \mathcal{G}_{n-1}^N\right] \\
&\qquad \leq \frac{4N(N-1)(N-2)}{N^3} \cdot \|G_{n-1,\cdot}\|_\infty^4 \cdot \mathbf{1}_{\{\#\{\ell_{n-1}^1, \ell_{n-1}'^1, \ell_{n-1}^2, \ell_{n-1}'^2\}<4\}}. \qquad a.s.
\end{aligned}$$

By combining the two cases, we have

$$\begin{aligned}
&\mathbf{E}\left[\gamma_n^N(1)^4 \cdot \left(\frac{N^{n-1}}{(N-1)^{n+1}}\right)^2 \sum_{(\ell_n^{[2]}, \ell_n'^{[2]}) \in ((N)^2)^{\times 2} \setminus (N)^4} \Lambda_n^{\ell_n^{[2]}} \Lambda_n^{\ell_n'^{[2]}} \,\middle|\, \mathcal{G}_{n-1}^N\right] \\
\leq& \gamma_{n-1}^N(1)^4 \cdot \left(\frac{N^{n-2}}{(N-1)^n}\right)^2 \sum_{(\ell_{n-1}^{[2]}, \ell_{n-1}'^{[2]}) \in ((N)^2)^{\times 2}} \Lambda_{n-1}^{\ell_{n-1}^{[2]}} \Lambda_{n-1}^{\ell_{n-1}'^{[2]}} \cdot \mathbf{1}_{\{\#\{\ell_{n-1}^1, \ell_{n-1}'^1, \ell_{n-1}^2, \ell_{n-1}'^2\}<4\}} \\
&\cdot \frac{N^2}{(N-1)^2} \cdot \left(\frac{2N(N-1)}{N^2} + \frac{4N(N-1)(N-2)}{N^3}\right) \cdot \|G_{n-1,\cdot}\|_\infty^4 \\
\leq& 6 \cdot \frac{N^2}{(N-1)^2} \cdot \|G_{n-1,\cdot}\|_\infty^4 \cdot \gamma_{n-1}^N(1)^4 \cdot \left(\frac{N^{n-2}}{(N-1)^n}\right)^2 \sum_{(\ell_{n-1}^{[2]}, \ell_{n-1}'^{[2]}) \in ((N)^2)^{\times 2} \setminus (N)^4} \Lambda_{n-1}^{\ell_{n-1}^{[2]}} \Lambda_{n-1}^{\ell_{n-1}'^{[2]}}
\end{aligned}$$

The conclusion follows from the induction hypothesis by taking the expectation for both sides. ◻

Lemma 3.4. *Assume ($\mathcal{A}$1-$\mathcal{A}$2). For any test functions $f, g \in \mathcal{B}_b(E_n)$, we have, for all $n \geq 1$, that*

$$\Gamma_{n-1,N}^b Q_{n,N}^{\otimes 2}(f \otimes g) - \Gamma_{n-1,N}^b Q_n^{\otimes 2}(f \otimes g) \xrightarrow[N \to \infty]{\mathbf{P}} 0$$

Proof. The verification is similar to the convergence of $P_2(N)$ showed in the proof of Theorem 3.1 by exploiting the following decomposition:

$$\begin{aligned}
&\Gamma_{n-1,N}^b Q_{n,N}^{\otimes 2}(f \otimes g) - \Gamma_{n-1,N}^b Q_n^{\otimes 2}(f \otimes g) \\
&\qquad = \underbrace{\Gamma_{n-1,N}^b Q_{n,N}^{\otimes 2}(f \otimes g) - \Gamma_{n-1,N}^b Q_{n,N} \otimes Q_n(f \otimes g)}_{D_1(N)} \\
&\qquad\qquad + \underbrace{\Gamma_{n-1,N}^b Q_{n,N} \otimes Q_n(f \otimes g) - \Gamma_{n-1,N}^b Q_n^{\otimes 2}(f \otimes g)}_{D_2(N)}.
\end{aligned}$$

We consider

$$\Omega_{n-1} := \left\{\omega \in \Omega : Z_{n-1}^N(\omega) \xrightarrow[N \to \infty]{} z_{n-1}^*\right\}. \tag{22}$$

By definition, it is clear that $\mathbf{P}(\Omega_{n-1}) = 1$. Now, let us fix $\omega \in \Omega_{n-1}$. By ($\mathcal{A}2$), we have, for any $g \in \mathcal{B}_b(E_n)$, there exists a function h, such that

$$Q_{n,Z_{n-1}^N(\omega)}(g)(x) - Q_n(g)(x) = h_n(x, Z_{n-1}^N(\omega)) \cdot (Z_{n-1}^N(\omega) - z_{n-1}^*).$$

This is equivalent to say that we have

$$Q_{n,N}(g)(x) - Q_n(g)(x) = h_n(x, Z_{n-1}^N) \cdot (Z_{n-1}^N - z_{n-1}^*). \leq d \cdot \|h_n\|_\infty \cdot \left|Z_{n-1}^N - z_{n-1}^*\right|. \qquad a.s.$$

Therefore, we deduce that

$$D_1(N) \leq 2 \cdot d \cdot \Gamma_{n-1,N}^b(1) \cdot \|f\|_\infty \cdot \left\|G_{n-1,\cdot}\right\|_\infty \cdot \|h_n\|_\infty \cdot \left|Z_{n-1}^N - z_{n-1}^*\right|. \qquad a.s.$$

Since by Proposition 3.1, one has

$$\Gamma_{n-1,N}^b(1) = O_{\mathbf{P}}(1),$$

we finally conclude by applying Theorem 3.1 that

$$D_1(N) \xrightarrow[N\to\infty]{\mathbf{P}} 0.$$

The verification for the convergence of $D_2(N)$ is similar. □

Appendix

A Many-body Feynman-Kac models

The many-body Feynman-Kac model was firstly introduced in [7] to study de propagation of chaos property of the Conditional Particle Markov Chain Monte Carlo. The idea is to trace the information of all particles in the particles system along with its genealogy. We also use the same kind of freezing or coupling techniques in order to construct a duality formula between the particle system and the coupled particle block.

We define

$$\mathcal{M}_p(\boldsymbol{x_{p-1}}, d(\boldsymbol{a_{p-1}}, \boldsymbol{x_p})) := \bigotimes_{i=1}^{N} \Phi_{p,N}(\boldsymbol{x_{p-1}}, d(a_{p-1}^i, x_p^i))$$

the transition kernel of the IPS with genealogy and

$$\mathcal{G}_{p-1}^{(q)}(\boldsymbol{x_{p-1}}) := m(\boldsymbol{x_{p-1}})(G_{p-1,N})^q$$

the potential function of the particle block of size 2. We denote the associated Feynman-Kac transition kernel

$$\mathbf{Q}_p^{(q)}(\boldsymbol{x_{p-1}}, d(\boldsymbol{a_{p-1}}, \boldsymbol{x_p})) := \mathcal{G}_{p-1}^{(q)}(\boldsymbol{x_{p-1}}) \times \mathcal{M}_p(\boldsymbol{x_{p-1}}, d(\boldsymbol{a_{p-1}}, \boldsymbol{x_p})).$$

Given $\ell_p^{[q]} \in (N)^q$, $\tilde{a}_{p-1}^{[q]} \in [N]^q$ and $\tilde{x}_p^{[q]} \in E_p^q$, we define

$$\begin{aligned}\mathbb{M}_p^{\tilde{a}_{p-1}^{[q]}, \ell_p^{[q]}, \tilde{x}_p^{[q]}}&(\boldsymbol{x_{p-1}}, d(\boldsymbol{a_{p-1}}, \boldsymbol{x_p})) \\ &:= \bigotimes_{i\in[N]\setminus\{\ell_p^1,\ldots,\ell_p^q\}} \left\{\Phi_{p,N}(\boldsymbol{x_{p-1}}, d(a_{p-1}^i, x_p^i))\right\} \otimes \delta_{\tilde{x}_p^{[q]}}(dx_p^{\ell_p^{[q]}}) \otimes \delta_{\tilde{a}_{p-1}^{[q]}}(da_{p-1}^{\ell_p^{[q]}})\end{aligned}$$

the conditional evolution transition kernel for the particle system given the coupled particle block $\tilde{X}_p^{[q]} = \tilde{x}_p^{[q]}$ at position $\ell_p^{[q]}$ with frozen genealogy $\tilde{A}_{p-1}^{[q]} = \tilde{a}_{p-1}^{[q]}$. In particular, we denote

$$\mathbb{M}_0^{\ell_0^{[q]},\tilde{x}_0^{[q]}}(d\boldsymbol{x_0}) := \left\{ \bigotimes_{i\in[N]\setminus\{\ell_0^1,\ldots,\ell_0^q\}} \eta_0(dx_0^i) \right\} \otimes \delta_{\tilde{x}_0^{[q]}}(dx_0^{\ell_0^{[q]}})$$

We also define

$$\mathbb{Q}_p^{(q)}(\boldsymbol{x_{p-1}}, d(\tilde{a}_{p-1}^{[q]}, \tilde{x}_p^{[q]})) := m([N])^{\otimes q}(d\tilde{a}_{p-1}^{[q]}) Q_{p,N}^{\otimes q}(x_{p-1}^{\tilde{a}_{p-1}^{[q]}}, d\tilde{x}_p^{[q]})$$

and

$$\mathbf{M}_p^{\ell_p^{[q]}}\left((\boldsymbol{a_{p-1}}, \boldsymbol{x_p}), d(\tilde{a}_{p-1}^{[q]}, \tilde{x}_p^{[q]})\right) := \delta_{x_p^{\ell_p^{[q]}}}(d\tilde{x}_p^{[q]}) \cdot \delta_{a_{p-1}^{\ell_p^{[q]}}}(d\tilde{a}_{p-1}^{[q]})$$

For the coupled particle block with size q, we denote respectively

$$(\tilde{X}_0^{[q]}, \ldots, \tilde{X}_n^{[q]}) \qquad \text{and} \qquad (\tilde{A}_0^{[q]}, \ldots, \tilde{A}_{n-1}^{[q]})$$

the coupled particle block and its genealogy. Then, we introduce the pivotal duality formula.

Lemma A.1. *For $p \geq 1$, $q \in [N]$ and $\ell_p^{[q]} \in (N)^q$, we have the following duality formula between integral operators*

$$\begin{aligned}\mathbf{Q}_p^{(q)}(\boldsymbol{x_{p-1}}, d(\boldsymbol{a_{p-1}}, \boldsymbol{x_p})) \cdot \mathbf{M}_p^{\ell_p^{[q]}}\left((\boldsymbol{a_{p-1}}, \boldsymbol{x_p}), d(\tilde{a}_{p-1}^{[q]}, \tilde{x}_p^{[q]})\right)& \\ = \mathbb{Q}_p^{(q)}(\boldsymbol{x_{p-1}}, d(\tilde{x}_p^{[q]}, \tilde{a}_{p-1}^{[q]})) \cdot \mathbb{M}_p^{\tilde{a}_{p-1}^{[q]},\ell_p^{[q]},\tilde{x}_p^{[q]}}(\boldsymbol{x_{p-1}}, d(\boldsymbol{a_{p-1}}, \boldsymbol{x_p}))&\end{aligned}$$

and

$$\eta_0^{\otimes N}(d\boldsymbol{x_0}) \cdot \delta_{x_0^{\ell_0^{[q]}}}(d\tilde{x}_0^{[q]}) = \eta_0^{\otimes q}(d\tilde{x}_0^{[q]}) \mathcal{M}_0^{\ell_0^{[q]},\tilde{x}_0^{[q]}}(d\boldsymbol{x_0})$$

Proof. The step 0 is trivial. For $p \geq 1$, fixing $\ell_p^{[q]} \in (N)^q$, for any functions $f_i \in \mathcal{B}_b([N])$ and $g_i \in \mathcal{B}_b(E_p)$ for $i \in [q]$, we denote $\mathbf{F} := f_1 \otimes \cdots \otimes f_q \otimes g_1 \otimes \cdots \otimes g_q$. Then, we have

$$\begin{aligned}&\int m(\boldsymbol{x_{p-1}})(G_{p-1,N})^q \mathbf{F}(\tilde{a}_{p-1}^{[q]}, a_{p-1}^{\ell_p^{[q]}}, x_p^{\ell_p^{[2]}}, \tilde{x}_p^{[2]}) \\ &\left[\sum_{k=1}^{n} \frac{G_{p-1,N}(x_{p-1}^k)}{N \cdot m(\boldsymbol{x_{p-1}})(G_{p-1,N})} \cdot \delta_k\right]^{\otimes q} (da_{p-1}^{\ell_p^{[q]}}) \cdot M_{p,N}^{\otimes q}(x_{p-1}^{a_{p-1}^{\ell_p^{[q]}}}, dx_p^{\ell_p^{[q]}}) \delta_{x_p^{\ell_p^{[q]}}}(d\tilde{x}_p^{[q]}) \delta_{a_{p-1}^{\ell_p^{[q]}}}(d\tilde{a}_{p-1}^{[q]}) \\ &= \int \mathbf{F}(k^{[q]}, k^{[q]}, x_p^{\ell_p^{[q]}}, \tilde{x}_p^{[q]}) \cdot m([N])^{\otimes q}(dk^{[q]}) G_{p-1,N}^{\otimes q}(x_{p-1}^{k^{[q]}}) \cdot M_{p,N}^{\otimes q}(x_{p-1}^{k^{[q]}}, dx_p^{\ell_p^{[q]}}) \delta_{x_p^{\ell_p^{[q]}}}(d\tilde{x}_p^{[q]}) \\ &= \int \mathbf{F}(k^{[q]}, k^{[q]}, x_p^{\ell_p^{[q]}}, \tilde{x}_p^{[q]}) \cdot m([N])^{\otimes q}(dk^{[q]}) \cdot Q_{p,N}^{\otimes q}(x_{p-1}^{k^{[q]}}, dx_p^{\ell_p^{[q]}}) \delta_{x_p^{\ell_p^{[q]}}}(d\tilde{x}_p^{[q]}) \\ &= \int \mathbf{F}(k^{[q]}, k^{[q]}, x_p^{\ell_p^{[q]}}, x_p^{\ell_p^{[q]}}) \cdot m([N])^{\otimes q}(dk^{[q]}) \cdot Q_{p,N}^{\otimes q}(x_{p-1}^{k^{[q]}}, dx_p^{\ell_p^{[q]}}) \\ &= \int \mathbf{F}(\tilde{a}_{p-1}^{[q]}, a_{p-1}^{\ell_p^{[q]}}, x_p^{\ell_p^{[q]}}, \tilde{x}_p^{[q]}) m([N])^{\otimes q}(d\tilde{a}_{p-1}^{[q]}) Q_{p,N}^{\otimes q}(x_{p-1}^{\tilde{a}_{p-1}^{[q]}}, d\tilde{x}_p^{[q]}) \delta_{\tilde{x}_p^{[q]}}(dx_p^{\ell_p^{[q]}}) \delta_{\tilde{a}_{p-1}^{[q]}}(da_{p-1}^{\ell_p^{[q]}}).\end{aligned}$$

This gives the equality of the nonidentical part of the duality formula. ◻

We recall the notation we defined in (14):

$$\Lambda_n^{\ell_n^{[2]}} := \sum_{\ell_{0:n-1}^{[2]} \in ((N)^2)^{\times n}} \prod_{p=0}^{n-1} \left\{ \lambda_p^b (A_p^{\ell_{p+1}^{[2]}}, \ell_p^{[2]}) \right\}$$

with the convention $\Lambda_0^{\ell_0^{[2]}} := 1$. In fact, this gives another representation of the approximation of the coalescent tree-typed measures:

$$\Gamma_{n,N}^b (f \otimes g) = \gamma_n^N (1)^2 \cdot \frac{N^{n-1}}{(N-1)^{n+1}} \sum_{\ell_n^{[2]} \in (N)^2} \Lambda_n^{\ell_n^{[2]}} \cdot f(X_n^{\ell_n^1}) g(X_n^{\ell_n^2}).$$

We introduce the following Lemma which is useful in the proof of the consistency of the variance estimator.

Lemma A.2. *Assume ($\mathcal{A}$1). Fixing $\ell_n^{[2]} \in (N)^2$, for any coalescence indicator b, and for any test functions f and g in $\mathcal{B}_b(E_n)$, we have*

$$\mathbf{E}\left[\gamma_n^N (1)^2 \cdot \frac{N^{n-1}}{(N-1)^{n+1}} \cdot \Lambda_n^{\ell_n^{[2]}} \cdot f(X_n^{\ell_n^1}) g(X_n^{\ell_n^2}) \,\middle|\, \mathcal{G}_{n-1}^N \right] = \frac{1}{N(N-1)} \cdot \Gamma_{n-1}^N Q_{n,N}^{\otimes 2} C_{b_n} (f \otimes g)$$

Proof. By applying the decomposition in (16), we have

$$\begin{aligned} &\gamma_n^N (1)^2 \cdot \frac{N^{n-1}}{(N-1)^{n+1}} \cdot \Lambda_n^{\ell_n^{[2]}} \cdot f(X_n^{\ell_n^1}) g(X_n^{\ell_n^2}) \\ &\qquad = \gamma_n^N (1)^2 \cdot \frac{N^{n-1}}{(N-1)^{n+1}} \sum_{\ell_{n-1}^{[2]} \in (N)^2} \Lambda_{n-1}^{\ell_{n-1}^{[2]}} \cdot \lambda_{n-1}^b (A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}). \cdot f(X_n^{\ell_n^1}) g(X_n^{\ell_n^2}) \end{aligned}$$

Hence, it is sufficient to show that for each $\ell_{n-1}^{[2]} \in (N)^2$, we have

$$\begin{aligned} &\mathbf{E}\left[\frac{N}{(N-1)} \cdot m(\mathbf{X_{n-1}})(G_{n-1,N})^2 \cdot \lambda_{n-1}^b (A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}). \cdot f(X_n^{\ell_n^1}) g(X_n^{\ell_n^2}) \,\middle|\, \mathcal{G}_{n-1}^N \right] \\ =&\mathbf{E}\left[\frac{N}{(N-1)} \cdot m(\mathbf{X_{n-1}})(G_{n-1,N})^2 \cdot \lambda_{n-1}^b (A_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}). \cdot f(X_n^{\ell_n^1}) g(X_n^{\ell_n^2}) \,\middle|\, \sigma(\mathbf{X_{n-1}}) \right] \\ =&\frac{1}{N(N-1)} \cdot Q_{n,N}^{\otimes 2} C_{b_n} (f \otimes g)(X_{n-1}^{\ell_{n-1}^{[2]}}) \end{aligned} \tag{23}$$

For any test function $F := \bigotimes_{i=1}^N f_i$, where $f_i \in \mathcal{B}_b(E_{n-1})$, by applying Lemma A.1, we have

$$\begin{aligned} &\int F(\boldsymbol{x_{n-1}}) \mathbf{Q}_n (x_{n-1}, d(\boldsymbol{a_{n-1}}, \boldsymbol{x_n})) \mathbf{M}_n^{\ell_n^{[2]}} \left((\boldsymbol{a_{n-1}}, \boldsymbol{x_n}), d(\tilde{a}_{n-1}^{[2]}, \tilde{x}_n^{[2]}) \right) \lambda_{n-1}(a_{n-1}^{\ell_n^{[2]}}, \ell_{n-1}^{[2]}) f \otimes g(x_n^{\ell_n^{[2]}}) \mu_{n-1}(d\boldsymbol{x_{n-1}}) \\ =& \int F(\boldsymbol{x_{n-1}}) \mathbf{Q}_n (x_{n-1}, d(\boldsymbol{a_{n-1}}, \boldsymbol{x_n})) \mathbf{M}_n^{\ell_n^{[2]}} \left((\boldsymbol{a_{n-1}}, \boldsymbol{x_n}), d(\tilde{a}_{n-1}^{[2]}, \tilde{x}_n^{[2]}) \right) \lambda_{n-1}(\tilde{a}_{n-1}^{[2]}, \ell_{n-1}^{[2]}) f \otimes g(\tilde{x}_n^{\ell_n^{[2]}}) \mu_{n-1}(d\boldsymbol{x_{n-1}}) \\ =& \int F(\boldsymbol{x_{n-1}}) \mathbb{Q}_n (x_{n-1}, d(\tilde{a}_{n-1}^{[2]}, \tilde{x}_n^{[2]})) \mathbb{M}_n^{\tilde{a}_{n-1}^{[2]}, \ell_n^{[2]}, \tilde{x}_n^{[2]}} (\boldsymbol{x_{n-1}}, d(\boldsymbol{a_{n-1}}, \boldsymbol{x_n})) \lambda_{n-1}(\tilde{a}_{n-1}^{[2]}, \ell_{n-1}^{[2]}) f \otimes g(\tilde{x}_n^{[2]}) \mu_{n-1}(d\boldsymbol{x_{n-1}}) \\ =& \frac{1}{N^2} \cdot \int F(\boldsymbol{x_{n-1}}) Q_{n,N}^{\otimes 2} C_{b_n} (f \otimes g)(x_{n-1}^{\ell_{n-1}^{[2]}}) \mu_{n-1}(d\boldsymbol{x_{n-1}}), \end{aligned}$$

where μ_{n-1} denotes the underlying probability measure of $\mathbf{X_{n-1}}$. Therefore, (23) is verified.

$\square$

Some intuitions for the construction of $\Gamma^b_{n,N}$ In general, the coupled particle block does not necessarily have the parents-children relations. Let us see a representation of the duality formula recursively applied in a mini IPS from level 0 to level 4. Every black arrow pointing to the particles in the original IPS represents a Markov transition $M_{p,N}$ and the ones pointing to the particles in the coupled particle block represent the Feynman-Kac transition kernels $Q_{p,N}$. The red dotted arrows are identities.

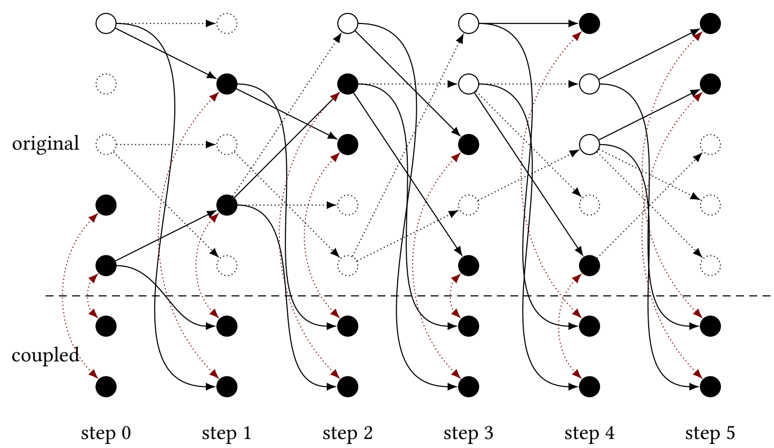

Figure 5: An illustration of the duality formula recursively applied to a mini IPS of 6 levels with 5 paricles at each level. The indices of the coupled particle block are $\ell_0^{[2]} = (4, 5)$, $\ell_1^{[2]} = (2, 4)$, $\ell_2^{[2]} = (2, 3)$, $\ell_3^{[2]} = (3, 5)$, $\ell_4^{[2]} = (1, 5)$ and $\ell_5^{[2]} = (1, 2)$.

However, we can get any ancestral relations or coalesent tree-typed form by manipulating the genealogical information encoded in the coupled genealogy. This is the essential idea we used by introducing many-body Feynman-Kac models. To make it clearer, we define an event

$$\left\{\ell_{p-2}^{[2]} = \tilde{A}_{p-2}^{[2]}, \ell_{p-1}^{1} = \tilde{A}_{p-1}^{1} = \tilde{A}_{p-1}^{2} \neq \ell_{p-1}^{2}, \ell_{p}^{[2]} = \tilde{A}_{p}^{[2]}\right\} \tag{24}$$

On this event, we are able to track the coalescent tree-typed form as follows

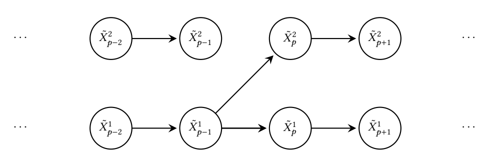

Figure 6: The coupled particle block tracked by the event defined by (24).

As the coupled particle block and its genealogy are defined as the copies of certain particles and parent indices in the associated IPS, connected by the duality formula, we could then use the information in the IPS to estimate those certain coalescent tree-typed particle blocks.

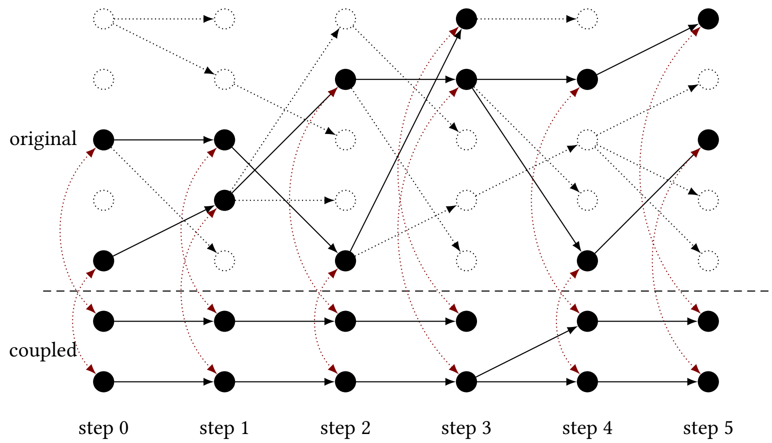

Figure 7: An illustration of the duality formula recursively applied to a mini IPS of 6 levels with 5 paricles at each level. The indices of the coupled particle block are $\ell_0^{[2]} = (3, 5)$, $\ell_1^{[2]} = (3, 4)$, $\ell_2^{[2]} = (2, 5)$, $\ell_3^{[2]} = (1, 2)$, $\ell_4^{[2]} = (2, 5)$ and $\ell_5^{[2]} = (1, 3)$.

The duality formula gives a way to deal with the adaptive versions of the coalescent tree-typed measures Γ_n^b (i.e. all the Feynman-Kac transition kernels Q_p in the definition are replaced by the adaptive version $Q_{p,N}$) by using the information in the associated IPS. This is the original idea of the relatively complex construction of the estimators $\Gamma_{n,N}^b$.

References

[1] C. Andrieu, A. Doucet, and R. Holenstein. Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(3):269–342, 2010.

[2] A. Beskos, A. Jasra, N. Kantas, and A. Thiery. On the convergence of adaptive sequential Monte Carlo methods. *Ann. Appl. Probab.*, 26(2):1111–1146, 2016.

[3] F. Cérou, P. Del Moral, and A. Guyader. A nonasymptotic theorem for unnormalized Feynman-Kac particle models. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(3):629–649, 2011.

[4] F. Cérou and A. Guyader. Fluctuation analysis of adaptive multilevel splitting. *Ann. Appl. Probab.*, 26(6):3319–3380, 2016.

[5] H. P. Chan and T. L. Lai. A general theory of particle filters in hidden Markov models and some applications. *Ann. Statist.*, 41(6):2877–2904, 2013.

[6] P. Del Moral. *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York, 2004. Genealogical and interacting particle systems with applications.

[7] P. Del Moral, R. Kohn, and F. Patras. On particle Gibbs samplers. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1687–1733, 2016.

[8] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo methods in practice*. Statistics for Engineering and Information Science. Springer-Verlag, New York, 2001.

[9] A. Lee and N. Whiteley. Variance estimation in the particle filter. *Biometrika*, 105(3):609–625, 2018.

[10] D. L. McLeish. Dependent central limit theorems and invariance principles. *Ann. Probability*, 2:620–628, 1974.