The Annals of Statistics
2008, Vol. 36, No. 5, 2344–2376
DOI: 10.1214/07-AOS514

LIMIT THEOREMS FOR WEIGHTED SAMPLES WITH APPLICATIONS TO SEQUENTIAL MONTE CARLO METHODS

BY RANDAL DOUC AND ERIC MOULINES

École Polytechnique and ENST

In the last decade, sequential Monte Carlo methods (SMC) emerged as a key tool in computational statistics [see, e.g., *Sequential Monte Carlo Methods in Practice* (2001) Springer, New York, *Monte Carlo Strategies in Scientific Computing* (2001) Springer, New York, *Complex Stochastic Systems* (2001) 109–173]. These algorithms approximate a sequence of distributions by a sequence of weighted empirical measures associated to a weighted population of particles, which are generated recursively.

Despite many theoretical advances [see, e.g., *J. Roy. Statist. Soc. Ser. B* **63** (2001) 127–146, *Ann. Statist.* **33** (2005) 1983–2021, *Feynman–Kac Formulae. Genealogical and Interacting Particle Systems with Applications* (2004) Springer, *Ann. Statist.* **32** (2004) 2385–2411], the large-sample theory of these approximations remains a question of central interest. In this paper we establish a law of large numbers and a central limit theorem as the number of particles gets large. We introduce the concepts of *weighted sample* consistency and asymptotic normality, and derive conditions under which the transformations of the weighted sample used in the SMC algorithm preserve these properties. To illustrate our findings, we analyze SMC algorithms to approximate the filtering distribution in state-space models. We show how our techniques allow to relax restrictive technical conditions used in previously reported works and provide grounds to analyze more sophisticated sequential sampling strategies, including branching, resampling at randomly selected times, and so on.

1. Introduction.

Sequential Monte Carlo (SMC) refer to a class of methods designed to approximate a *sequence of probability distributions* over a *sequence of probability space* by a set of points, termed *particles* that each have an assigned non-negative weight and are updated recursively in time. SMC methods can be seen as a combination of the sequential importance sampling introduced method in [8] and the sampling importance resampling algorithm proposed in [9]. In the *importance sampling* step, the particles are propagated forward in time using proposal kernels and their importance weights are updated taking into account the targeted distribution. In the *resampling* or the *branching* step, particles multiply or die depending on their importance weights. Many algorithms have been proposed

Received June 2007; revised June 2007.

AMS 2000 subject classifications. Primary 60F05, 62L10, 65C05; secondary 65C35, 65C60.

Key words and phrases. Branching, conditional central limit theorems, particle filtering, sequential importance sampling, sequential Monte Carlo.

since, which differ in the way the particles and the importance weights evolve and adapt.

SMC methods have a long history in molecular simulations, where they have been found to be one of the most powerful means for the simulation and optimization of chain polymers (see, e.g., [10]). SMC methods have more recently emerged as a key tool to solve *on-line* prediction/filtering/smoothing problems in a dynamic system. Simple yet flexible SMC methods have been shown to overcome the numerical difficulties and pitfalls typically encountered with traditional methods based on approximate nonlinear filtering (such as the extended Kalman filter or Gaussian-sum filters); see, for instance, [1, 2, 11, 12] and the references therein. More recently, SMC methods have been shown to be a promising alternative to Markov chain Monte Carlo techniques for sampling complex distributions over large dimensional spaces; see, for instance, [4] and [13].

In this paper we study the large sample properties of *weighted particle* approximations as the number of particles tend to infinity. Because the particles interact during the resampling/branching steps, they are not independent, which make the analysis of particle approximation a challenging area of research. This topic has attracted in recent years a great deal of efforts, making it a daunting task to give credit to every contributor. The first rigorous convergence result was obtained in [14], who established the almost-sure convergence of an elementary SMC algorithm (the so-called *bootstrap filter*). A central limit theorem for this algorithm was derived in [15] and later refined in [16]. The proof of the CLT was later simplified and extended to more general SMC algorithms by [5] and [7]. Bounds on the fluctuations of the particle approximations for different norms were reported in [16], [17] and [18]. [6] provides an up-to-date and thorough coverage of recent theoretical developments in this area.

With few exceptions (see [7] and, to a lesser extent, [5] and [18]), these results apply under simplifying assumptions on the way importance sampling and resampling/branching is performed, which restrict the scope of applicability of the results only to the most elementary SMC implementations. In particular, all these results assume that resampling/branching is performed at each iteration, which implies that the weights are not propagated. This is clearly an annoying limitation since it has been noticed by many practitioners that resampling the particle system at each time step is most often not a clever choice.

The main purpose of this paper is to derive an asymptotic theory of weighted system of particles. To the best of our knowledge, limit theorems for such weighted approximations were only considered in [11], who mostly sketched consistency proofs. In this paper we establish both the law of large numbers and central limit theorems, under assumptions that are presumably closed from being minimal. These results apply not only to the many different implementations of the SMC algorithms, including rather sophisticated schemes such as the resample and move algorithm [19] or the auxiliary particle filter by [20], but they also cover resampling schedules (when to resample) that can be either deterministic or dynamic, that is, based on the distribution of the importance weights at the current iteration. They also cover sampling schemes that can be either simple random sampling (with weights) or also residual sampling [11] or auxiliary sampling [20]. We do not impose a specific structure on the sequence of the target probability measure; therefore, our results apply not only to sequential filtering or smoothing of state-space contexts, but also to recent algorithms developed for a population Monte-Carlo or for molecular simulation.

The paper is organized as follows. In Section 2 we introduce the definitions of weighted sample consistency and asymptotic normality; we then discuss the conditions upon which consistency or/and asymptotic normality of a weighted sample is preserved by the importance sampling, resampling and branching steps. In Section 3.2 we apply the result to the estimation of the joint smoothing distribution for a state-space model. In particular, we establish a central limit theorem for a SMC method involving a dynamic resampling scheme. These results are based on new results on conditional limit theorems for a triangular array of dependent data which are established in Appendix A.

2. Notation and main results.

2.1. *Notation.* All the random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathrm{P})$. A state space X is said to be *general* if it is equipped with a countably generated σ-field $\mathcal{B}(\mathsf{X})$. For a general state space X, we denote by $\mathcal{P}(\mathsf{X})$ the set of probability measures on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$ and $\mathbb{B}(\mathsf{X})$ [resp. $\mathbb{B}^+(\mathsf{X})$] the set of all $\mathcal{B}(\mathsf{X})/\mathcal{B}(\mathbb{R})$-measurable (resp. nonnegative) functions from X to $\mathbb{R}$ equipped with the Borel σ-field $\mathcal{B}(\mathbb{R})$. A subset $\mathsf{C} \subseteq \mathsf{X}$ is said to be *proper* if the following conditions are satisfied: (i) C is a linear space: for any f and g in C and reals α and β, $\alpha f + \beta g \in \mathsf{C}$; (ii) if $g \in \mathsf{C}$ and f is measurable with $|f| \le |g|$, then $|f| \in \mathsf{C}$; (iii) for all c, the constant function $f \equiv c$ belongs to C.

For any $\mu \in \mathcal{P}(\mathsf{X})$ and $f \in \mathbb{B}(\mathsf{X})$ satisfying $\int_{\mathsf{X}} \mu(dx)|f(x)| < \infty$, $\mu(f)$ denotes $\int_{\mathsf{X}} f(x)\mu(dx)$. Let X and Y be two general state spaces. A kernel V from $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$ to $(\mathsf{Y}, \mathcal{B}(\mathsf{Y}))$ is a map from $\mathsf{X} \times \mathcal{B}(\mathsf{Y})$ into $[0, 1]$ such that, for each $A \in \mathcal{B}(\mathsf{Y})$, $x \mapsto V(x, A)$ is a nonnegative bounded measurable function on X and, for each $x \in \mathsf{X}$, $A \mapsto V(x, A)$ is a measure on $\mathcal{B}(\mathsf{Y})$. We say that V is finite if $V(x, \mathsf{Y}) < \infty$ for any $x \in \mathsf{X}$; it is Markovian if $V(x, \mathsf{X}) \equiv 1$ for any $x \in \mathsf{X}$. For any function $f \in \mathbb{B}(\mathsf{X} \times \mathsf{Y})$ such that $\int_{\mathsf{Y}} V(x, dy)|f(x, y)| < \infty$, we denote by $V(\cdot, f)$ or $Vf(\cdot)$ the function $x \mapsto V(x, f) \stackrel{\text{def}}{=} \int_{\mathsf{Y}} V(x, dy) f(x, y)$. For ν a measure on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$, we denote by νV the measure on $(\mathsf{Y}, \mathcal{B}(\mathsf{Y}))$ defined for any $A \in \mathcal{B}(\mathsf{Y})$ by $\nu V(A) = \int_{\mathsf{X}} \nu(dx) V(x, A)$.

Throughout the paper, we denote by Ξ, μ a probability measure on $(\Xi, \mathcal{B}(\Xi))$, $\{M_N\}_{N \ge 0}$ a sequence integer-valued random variable, C a proper subset of Ξ. We approximate the probability measure μ by points $\xi_{N,i} \in \Xi$, $i = 1, \dots, M_N$ associated to nonnegative weights $\omega_{N,i} \ge 0$.

DEFINITION 1. A weighted sample $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ on Ξ is said to be *consistent* for the probability measure μ and the (proper) set C if, for any $f \in \mathsf{C}$, as $N \to \infty$, $\Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f(\xi_{N,i}) \stackrel{\mathrm{P}}{\longrightarrow} \mu(f)$ and $\Omega_N^{-1} \max_{i=1}^{M_N} \omega_{N,i} \stackrel{\mathrm{P}}{\longrightarrow} 0$, where $\Omega_N = \sum_{i=1}^{M_N} \omega_{N,i}$.

This definition of weighted sample consistency is similar to the notion of *properly weighted sample* introduced in [11]. The difference stems from the smallness condition which states that the contribution of each individual term in the sum vanishes in the limit as $N \to \infty$.

We denote by γ a finite measure on $(\Xi, \mathcal{B}(\Xi))$, with A and W proper sets of Ξ, and σ a real nonnegative function on A, and $\{a_N\}$ a nondecreasing real sequence diverging to infinity.

DEFINITION 2. A weighted sample $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ on Ξ is said to be *asymptotically normal* for $(\mu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$ if

$$a_N \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} \{f(\xi_{N,i}) - \mu(f)\} \stackrel{\mathcal{D}}{\longrightarrow} \mathrm{N}\{0, \sigma^2(f)\} \qquad \text{for any } f \in \mathsf{A}, \tag{1}$$

$$a_N^2 \Omega_N^{-2} \sum_{i=1}^{M_N} \omega_{N,i}^2 f(\xi_{N,i}) \stackrel{\mathrm{P}}{\longrightarrow} \gamma(f) \qquad \text{for any } f \in \mathsf{W}, \tag{2}$$

$$a_N \Omega_N^{-1} \max_{1 \le i \le M_N} \omega_{N,i} \stackrel{\mathrm{P}}{\longrightarrow} 0. \tag{3}$$

Note that these definitions implicitly imply that the sets C, A and W are proper.

To analyze the sequential Monte Carlo methods discussed in the Introduction, we now need to study how the importance sampling and the resampling steps affect the consistent or/and asymptotically normal weighted sample.

2.2. *Importance sampling*. We will show that the importance sampling step transforms a weighted sample consistent (or asymptotically normal) for a distribution ν on a general state space $(\Xi, \mathcal{B}(\Xi))$ into a weighted sample consistent (or asymptotically normal) for a distribution μ on $(\tilde{\Xi}, \mathcal{B}(\tilde{\Xi}))$. Let L be a Markov kernel from $(\Xi, \mathcal{B}(\Xi))$ to $(\tilde{\Xi}, \mathcal{B}(\tilde{\Xi}))$ such that, for any $f \in \mathbb{B}(\tilde{\Xi})$,

$$\mu = \frac{\nu L f}{\nu L(\tilde{\Xi})}. \tag{4}$$

We wish to transform a weighted sample $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ targeting the distribution ν on $(\Xi, \mathcal{B}(\Xi))$ into a weighted sample $\{(\tilde{\xi}_{N,i}, \tilde{\omega}_{N,i})\}_{i=1}^{\tilde{M}_N}$ targeting μ on $(\tilde{\Xi}, \mathcal{B}(\tilde{\Xi}))$, where $\tilde{M}_N = \alpha M_N$ (α denoting the number of offsprings of each particle). The use of multiple offsprings has been suggested by [9]: when the importance sampling step is followed by a resampling step, an increase in the number of distinct particles will increase the number of distinct particles **after** the resampling step. In the sequential context, this operation is a practical mean for contending particle impoverishment. These offsprings are proposed using a Markov kernel denoted R from $(\Xi, \mathcal{B}(\Xi))$ to $(\tilde{\Xi}, \mathcal{B}(\tilde{\Xi}))$. We assume that, for any $\xi \in \Xi$, the probability measure $L(\xi, \cdot)$ on $(\tilde{\Xi}, \mathcal{B}(\tilde{\Xi}))$ is absolutely continuous with respect to R, which we denote $L(\xi, \cdot) \ll R(\xi, \cdot)$ and define

$$W(\xi, \tilde{\xi}) \stackrel{\text{def}}{=} \frac{dL(\xi, \cdot)}{dR(\xi, \cdot)}(\tilde{\xi}). \tag{5}$$

The new weighted sample $\{(\tilde{\xi}_{N,i}, \tilde{\omega}_{N,i})\}_{i=1}^{\tilde{M}_N}$ is constructed as follows. We draw new particle positions $\{\tilde{\xi}_{N,j}\}_{j=1}^{\tilde{M}_N}$ conditionally independent, given

$$\mathcal{F}_{N,0} \stackrel{\text{def}}{=} \sigma(M_N, \{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}), \tag{6}$$

with distribution given for $i = 1, \dots, M_N$, $k = 1, \dots, \alpha$ and $A \in \mathcal{B}(\tilde{\Xi})$ by

$$\mathrm{P}\big(\tilde{\xi}_{N,\alpha(i-1)+k} \in A | \mathcal{F}_{N,0}\big) = R(\xi_{N,i}, A), \tag{7}$$

and associate to each new particle positions the importance weight

$$\tilde{\omega}_{N,\alpha(i-1)+k} = \omega_{N,i} W\big(\xi_{N,i}, \tilde{\xi}_{N,\alpha(i-1)+k}\big), \tag{8}$$

for $i = 1, \dots, M_N$ and $k = 1, \dots, \alpha$. The importance sampling step is *unbiased* in the sense that, for any $f \in \mathcal{B}(\tilde{\Xi})$ and $i = 1, \dots, M_N$,

$$\sum_{j=\alpha(i-1)+1}^{\alpha i} \mathrm{E}[\tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) | \mathcal{F}_{N,j-1}] = \alpha \omega_{N,i} L(\xi_{N,i}, f), \tag{9}$$

where for $j = 1, \dots, \tilde{M}_N$, $\mathcal{F}_{N,j} \stackrel{\text{def}}{=} \mathcal{F}_{N,0} \vee \sigma(\{\tilde{\xi}_{N,l}\}_{1 \le l \le j})$. The following theorems state conditions under which the importance sampling step described above preserves the weighted sample consistency. Denote by

$$\tilde{\mathsf{C}} \stackrel{\text{def}}{=} \{f \in \mathsf{L}^1(\tilde{\Xi}, \mu), L(\cdot, |f|) \in \mathsf{C}\}. \tag{10}$$

THEOREM 1. *Assume that the weighted sample $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is consistent for (ν, C) and that $L(\cdot, \tilde{\Xi})$ belongs to C. Then, the set $\tilde{\mathsf{C}}$ defined in (10) is a proper set and the weighted sample $\{(\tilde{\xi}_{N,i}, \tilde{\omega}_{N,i})\}_{i=1}^{\tilde{M}_N}$ defined by (7) and (8) is consistent for $(\mu, \tilde{\mathsf{C}})$.*

We now turn to prove the asymptotic normality. Define

$$\begin{aligned} \tilde{\mathsf{A}} &\stackrel{\text{def}}{=} \{f : L(\cdot, |f|) \in \mathsf{A}, R(\cdot, W^2 f^2) \in \mathsf{W}\}, \\ \tilde{\mathsf{W}} &\stackrel{\text{def}}{=} \{f : R(\cdot, W^2 |f|) \in \mathsf{W}\}. \end{aligned} \tag{11}$$

THEOREM 2. *Suppose that the assumptions of Theorem 1 hold. Assume in addition, that the weighted sample* $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ *is asymptotically normal for* $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$, *and that the function* $R(\cdot, W^2)$ *belongs to* W.

Then, the sets $\tilde{\mathsf{A}}$ *and* $\tilde{\mathsf{W}}$ *defined in* (11) *are proper and the weighted sample* $\{(\tilde{\xi}_{N,i}, \tilde{\omega}_{N,i})\}_{i=1}^{\tilde{M}_N}$ *is asymptotically normal for* $(\mu, \tilde{\mathsf{A}}, \tilde{\mathsf{W}}, \tilde{\sigma}, \tilde{\gamma}, \{a_N\})$ *with* $\tilde{\gamma}(f) \stackrel{\text{def}}{=} \alpha^{-1}\gamma R(W^2 f)/(\nu L(\tilde{\Xi}))^2$ *and*

$$\tilde{\sigma}^2(f) \stackrel{\text{def}}{=} \sigma^2\{L[f - \mu(f)]\}/(\nu L(\tilde{\Xi}))^2 + \alpha^{-1}\gamma R\{[W[f - \mu(f)] - R(\cdot, W[f - \mu(f)])]^2\}/(\nu L(\tilde{\Xi}))^2.$$

2.3. *Resampling*. Resampling converts a weighted sample $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ targeting a distribution $\nu(\Xi, \mathcal{B}(\Xi))$ into an equally weighted sample $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ targeting the same distribution ν. The resampling step is an essential ingredient in the sequential context because it removes particles with small weights and produces multiple copies of particles with large weights. Denote by $G_{N,i}$ the number of times the ith particle is replicated. The number of particles after resampling $\tilde{M}_N = \sum_{i=1}^{M_N} G_{N,i}$ is supposed to be an $\mathcal{F}_{N,0}$-measurable integer-valued random variable, where $\mathcal{F}_{N,0}$ is given in (6); it might differ from the initial number of particles M_N, but will generally be a (deterministic) function of it. There are many different resampling procedures described in the literature. The simplest is the *multinomial* resampling, in which the distribution of $(G_{N,1}, \dots, G_{N,M_N})$ conditionally to $\mathcal{F}_{N,0}$ is multinomial:

$$(G_{N,1}, \dots, G_{N,M_N})|\mathcal{F}_{N,0} \sim \mathsf{Mult}(\tilde{M}_N, \{\Omega_N^{-1}\omega_{N,i}\}_{i=1}^{M_N}). \tag{12}$$

Another possible solution is the *deterministic-plus-residual multinomial resampling*, introduced in [21]. Denote by $\lfloor x \rfloor$ the integer part of x and by $\langle x \rangle$ denote the fractional part of x, $\langle x \rangle \stackrel{\text{def}}{=} x - \lfloor x \rfloor$. This scheme consists in retaining at least $\lfloor \Omega_N^{-1} \tilde{M}_N \omega_{N,i} \rfloor$, $i = 1, \dots, M_N$, copies of the particles and then reallocating the remaining particles by applying the multinomial resampling procedure with the residual importance weights defined as $\langle \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rangle$, that is, $G_{N,i} = \lfloor \Omega_N^{-1} \tilde{M}_N \omega_{N,i} \rfloor + H_{N,i}$, where

$$\begin{aligned}(H_{N,1}, \dots, H_{N,M_N})&|\mathcal{F}_{N,0} \\ &\sim \mathsf{Mult}\left(\sum_{i=1}^{M_N} \langle \Omega_N^{-1} \tilde{M}_N \omega_{N,i} \rangle, \left\{ \frac{\langle \Omega_N^{-1} \tilde{M}_N \omega_{N,i} \rangle}{\sum_{i=1}^{M_N} \langle \Omega_N^{-1} \tilde{M}_N \omega_{N,i} \rangle} \right\}_{i=1}^{M_N}\right).\end{aligned} \tag{13}$$

If the weighted sample $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is consistent for (ν, C), where C is a proper subset of $\mathbb{B}(\mathsf{X})$, it is a natural question to ask whether the uniformly weighted sample $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ is consistent for ν and, if so, what an appropriately defined class of functions on Ξ might be. It happens that a fairly general result can be obtained in this case.

THEOREM 3. *Assume that the weighted sample* $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ *is consistent for* (ν, C). *Then, the uniformly weighted sample* $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ *obtained using either* (12) *or* (13) *is consistent for* (ν, C).

It is also sensible to strengthen the requirement of consistency into asymptotic normality, and prove that the resampling procedures (12) and (13) transform an asymptotically normal weighted sample for ν into an asymptotically normal sample for ν. We consider first the multinomial sampling algorithm. We define

$$\tilde{\mathsf{A}} \stackrel{\text{def}}{=} \{f \in \mathsf{A},\ f^2 \in \mathsf{C}\}, \tag{14}$$

THEOREM 4. *Assume the following*:

(i) $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ *is consistent for* (ν, C) *and asymptotically normal for* $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$; *in addition*, $a_N^{-2} M_N \stackrel{\mathrm{P}}{\longrightarrow} \beta^{-1}$ *for some* $\beta \in [0, \infty)$.

(ii) $\tilde{M}_N$ *is* $\mathcal{F}_{N,0}$-*measurable, where* $\mathcal{F}_{N,0}$ *is defined in* (6), *and* $\tilde{M}_N / M_N \stackrel{\mathrm{P}}{\longrightarrow} \ell$ *where* $\ell \in [0, \infty]$.

Then $\tilde{\mathsf{A}}$ *is a proper set and the equally weighted particle system* $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ *obtained using* (12) *is asymptotically normal for* $(\nu, \tilde{\mathsf{A}}, \mathsf{C}, \tilde{\sigma}, \tilde{\gamma}, \{a_N\})$ *with* $\tilde{\sigma}^2(f) = \beta \ell^{-1} \operatorname{Var}_\nu(f) + \sigma^2(f)$ *and* $\tilde{\gamma} = \beta \ell^{-1} \nu$.

The analysis of the deterministic-plus-multinomial residual sampling is more involved. To carry out the analysis, it is required to consider situations where the importance weights are a function of the particle position, that is, $\omega_{N,i} = \Phi(\xi_{N,i})$, where $\Phi \in \mathbb{B}^+(\Xi)$. This condition is fulfilled in most applications of sequential Monte Carlo methods and should therefore not be considered as a stringent limitation. For $\ell \in \mathbb{R}^+$, and ν a probability measure on Ξ, define $\nu_{\ell,\Phi}$ the measure $\nu_{\ell,\Phi}(f) = \nu(\frac{\langle \ell \nu(\Phi^{-1}) \Phi \rangle}{\ell \nu(\Phi^{-1}) \Phi} f)$ for $f \in \mathbb{B}^+(\Xi)$.

THEOREM 5. *Assume the following*:

(i) $\{(\xi_{N,i}, \Phi(\xi_{N,i}))\}_{i=1}^{M_N}$ *is consistent for* (ν, C) *and asymptotically normal for* $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$; *in addition*, $a_N^{-2} M_N \stackrel{\mathrm{P}}{\longrightarrow} \beta^{-1}$ *for some* $\beta \in [0, \infty)$.

(ii) $\tilde{M}_N$ *is* $\mathcal{F}_{N,0}$-*measurable, where* $\mathcal{F}_{N,0}$ *is defined in* (6), *and* $\tilde{M}_N / M_N \stackrel{\mathrm{P}}{\longrightarrow} \ell$ *where* $\ell \in [0, \infty]$.

(iii) $\Phi^{-1} \in \mathsf{C}$, *and* $\nu(\ell \nu(\Phi^{-1}) \Phi \in \mathbb{N} \cup \{\infty\}) = 0$.

Then, the uniformly weighted sample $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ *obtained using* (13) *is asymptotically normal for* $(\nu, \tilde{\mathsf{A}}, \mathsf{C}, \tilde{\sigma}, \tilde{\gamma}, \{a_N\})$, *where* $\tilde{\mathsf{A}}$ *is given by* (14), $\tilde{\gamma} \stackrel{\text{def}}{=} \beta \ell^{-1} \nu$, *and*

$$\tilde{\sigma}^2(f) \stackrel{\text{def}}{=} \beta \ell^{-1} \nu_{\ell,\Phi}\{(f - \nu_{\ell,\Phi}(f)/\nu_{\ell,\Phi}(\mathbf{1}))^2\} + \sigma^2(f) \qquad \text{for } f \in \tilde{\mathsf{A}}.$$

REMARK 1. Because $\langle \ell\nu(\Phi^{-1})\Phi\rangle/\ell\nu(\Phi^{-1})\Phi \le 1$, for any $f \in \tilde{\mathsf{A}}$,

$$\nu_{\ell,\Phi}\{(f - \nu_{\ell,\Phi}(f)/\nu_{\ell,\Phi}(\mathbf{1}))^2\} = \inf_{c\in\mathbb{R}} \nu_{\ell,\Phi}\{(f-c)^2\} \le \inf_{c\in\mathbb{R}} \nu\{(f-c)^2\} = \operatorname{Var}_\nu(f),$$

showing that the variance of the residual-plus-deterministic sampling is always lower than that of the multinomial sampling. These results extend [7], Theorem 2 to derive an expression of the variance of the residual sampling in a specific case. Note, however, the assumption Theorem 5(iii) is missing in the statement of [7], Theorem 2. This assumption cannot be relaxed, as shown in Appendix D.

2.4. *Branching*. Branching procedures have been considered as an alternative to resampling procedure (see [16, 22, 23] and [6], Chapter 11); these procedures are easier to implement than resampling and are popular among practitioners. In the branching procedures, the number of times each particle is replicated $(G_{N,1}, \dots, G_{N,M_N})$ are independent conditionally to $\mathcal{F}_{N,0}$ and are distributed in such a way that $\mathrm{E}[G_{N,i}|\mathcal{F}_{N,0}] = \tilde{m}_N \Omega_N^{-1} \omega_{N,i}$, $i = 1, \dots, M_N$, where $\tilde{m}_N$ is the targeted number of particles, assumed to be a $\mathcal{F}_{N,0}$ random variable. Most often, $\tilde{m}_N$ is chosen to be a deterministic function of the current number of particles M_N, for example, $\tilde{m}_N = M_N$ or $\tilde{m}_N = N$ (in which case we target a "deterministic" number of particles). Contrary to the resampling procedures, the number of particles $\tilde{M}_N$ after branching is no longer $\mathcal{F}_{N,0}$-measurable, that is, the actual number of particles $\tilde{M}_N$ is different from the targeted number $\tilde{m}_N$ and cannot be predicted before the branching numbers $\{G_{N,i}\}_{i=1}^{M_N}$ are drawn. There are of course many different ways to select the branching numbers. In the Poisson branching, the branching numbers $\{G_{N,i}\}_{i=0}^{M_N}$ are conditionally independent given $\mathcal{F}_{N,0}$ with Poisson distribution with parameters $\{\tilde{m}_N \Omega_N^{-1} \omega_{N,i}\}_{i=1}^{M_N}$,

$$\{G_{N,i}\}_{i=1}^{M_N} | \mathcal{F}_{N,0} \sim \bigotimes_{i=1}^{M_N} \operatorname{Pois}(\tilde{m}_N \Omega_N^{-1} \omega_{N,i}), \tag{15}$$

where $\otimes$ denotes the tensor product of measures. Similarly, in the binomial branching, the branching numbers $\{G_{N,i}\}_{i=0}^{M_N}$ are conditionally independent given $\mathcal{F}_{N,0}$ with binomial distribution of parameters $\{(\tilde{m}_N, \Omega_N^{-1}\omega_{N,i})\}_{i=1}^{M_N}$,

$$\{G_{N,i}\}_{i=1}^{M_N} | \mathcal{F}_{N,0} \sim \bigotimes_{i=1}^{M_N} \operatorname{Bin}(\tilde{m}_N, \Omega_N^{-1} \omega_{N,i}), \tag{16}$$

The third branching algorithm, referred to as the Bernoulli branching algorithm, shares similarities with the deterministic-plus-residual multinomial sampling. In this case, for each ith, $\lfloor \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rfloor$ are retained; to correct for the truncation, an additional particle is eventually added, that is, $G_{N,i} = \lfloor \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rfloor + H_{N,i}$, where $\{H_{N,i}\}_{i=1}^{M_N}$ are conditionally independent given $\mathcal{F}_{N,0}$ with Bernoulli distribution of parameter $\{\langle \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rangle\}_{i=1}^{M_N}$,

$$G_{N,i} = \lfloor \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rfloor + H_{N,i}, \qquad (17)$$
$$\{H_{N,i}\}_{i=1}^{M_N} | \mathcal{F}_{N,0} \sim \bigotimes_{i=1}^{M_N} \mathrm{Ber}(\langle \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rangle).$$

As above, it may be shown that these branching algorithms preserve consistency.

THEOREM 6. *Assume that the weighted sample* $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ *is consistent for* (ν, C). *Then*, $\tilde{M}_N / \tilde{m}_N \xrightarrow{\mathrm{P}} 1$ *and the uniformly weighted sample* $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ *obtained using either* (15), (16) *and* (17) *is consistent for* (ν, C).

We may also strengthen the conditions to establish the asymptotic normality. For the Poisson and the binomial branching, the asymptotic normality is satisfied under almost the same conditions as for the multinomial sampling (see Theorem 4); in addition, the asymptotic variance of these procedures are equal.

THEOREM 7. *Assume the following*:

(i) $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ *is consistent for* (ν, C) *and asymptotically normal for* $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$; *in addition*, $a_N^{-2} M_N \xrightarrow{\mathrm{P}} \beta^{-1}$ *for some* $\beta \in [0, \infty)$.

(ii) $\tilde{m}_N$ *is* $\mathcal{F}_{N,0}$-*measurable, where* $\mathcal{F}_{N,0}$ *is defined in* (6), *and* $M_N^{-1} \tilde{m}_N \xrightarrow{\mathrm{P}} \ell$ *where* $\ell \in [0, \infty]$.

Then the equally weighted particle system $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ *obtained using either* (15) *or* (16) *is asymptotically normal for* $(\nu, \tilde{\mathsf{A}}, \mathsf{C}, \tilde{\sigma}, \tilde{\gamma}, \{a_N\})$, *with* $\tilde{\mathsf{A}} \stackrel{\text{def}}{=} \{f, f^2 \in \mathsf{C} \cap \mathsf{W}\}$, $\tilde{\sigma}^2(f) = \beta \ell^{-1} \mathrm{Var}_\nu(f) + \sigma^2(f)$, *and* $\tilde{\gamma} = \beta \ell^{-1} \nu$.

We now consider the case of the Bernoulli branching. As for the deterministic-plus-residual sampling, it is here required to assume that the weights are a function of the particle positions, that is, $\omega_{N,i} = \Phi(\xi_{N,i})$.

THEOREM 8. *Assume the following*:

(i) $\{(\xi_{N,i}, \Phi(\xi_{N,i}))\}_{i=1}^{M_N}$ *is consistent for* (ν, C) *and asymptotically normal for* $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$; *in addition*, $a_N^{-2} M_N \xrightarrow{\mathrm{P}} \beta^{-1}$ *for some* $\beta \in [0, \infty)$.

(ii) $\tilde{m}_N$ *is* $\mathcal{F}_{N,0}$-*measurable, where* $\mathcal{F}_{N,0}$ *is defined in* (6), *and* $\tilde{m}_N / M_N \xrightarrow{\mathrm{P}} \ell$ *where* $\ell \in [0, \infty]$,

(iii) $\Phi^{-1} \in \mathsf{C}$, *and* $\nu(\ell \nu(\Phi^{-1}) \Phi \in \mathbb{N} \cup \{\infty\}) = 0$.

Then, the uniformly weighted sample $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ *defined by* (17) *is asymptotically normal for* $(\nu, \tilde{\mathsf{A}}, \mathsf{C}, \tilde{\sigma}, \tilde{\gamma}, \{a_N\})$, *where* $\tilde{\mathsf{A}} \stackrel{\text{def}}{=} \{f \in \mathsf{A}, (1+\Phi) f^2 \in \mathsf{C}\}$, $\tilde{\gamma} \stackrel{\text{def}}{=} \beta \ell^{-1} \nu$, *and*

$$\tilde{\sigma}^2(f) \stackrel{\text{def}}{=} \beta \ell^{-1} \nu \left(\frac{\langle \ell \nu(\Phi^{-1}) \Phi \rangle (1 - \langle \ell \nu(\Phi^{-1}) \Phi \rangle)}{\ell \nu(\Phi^{-1}) \Phi} (f - \nu f)^2 \right) + \sigma^2(f), \qquad f \in \tilde{\mathsf{A}}.$$

REMARK 2. Since $\frac{\langle \ell \nu(\Phi^{-1}) \Phi \rangle (1 - \langle \ell \nu(\Phi^{-1}) \Phi \rangle)}{\ell \nu(\Phi^{-1}) \Phi} \leq 1$, the asymptotic variance of the Bernoulli branching is always lower than the asymptotic variance of the multinomial resampling. Compared with the deterministic-plus-residual sampling, the two quantities are not ordered uniformly w.r.t. f.

3. Applications.

3.1. *Fractional reweighting.* It has been advocated (see, e.g., [24]) that it could be advantageous when resampling to keep a fraction of the weight. The success of this approach has been mainly motivated on heuristic grounds. The results developed above allow to easily obtain an expression of the asymptotic variance of this scheme. The procedure goes as follows. Let ν be a distribution on $(\Xi, \mathcal{B}(\Xi))$ and assume that the weighted sample $\{(\xi_{N,i}, \Phi(\xi_{N,i}))\}_{i=1}^{M_N}$ targets ν. Let $\kappa \in [0,1]$. In a first step, we modify the weight pattern, that is, we consider the weighted sample $\{(\xi_{N,i}, \Phi^{1-\kappa}(\xi_{N,i}))\}_{i=1}^{M_N}$. This weighted sample targets the distribution $\nu_\kappa(\cdot) \stackrel{\text{def}}{=} \nu(\Phi^{-\kappa} \cdot)/\nu(\Phi^{-\kappa})$. In a second step, we resample the weighted sample $\{(\xi_{N,i}, \Phi^{1-\kappa}(\xi_{N,i}))\}_{i=1}^{M_N}$. Resampling produces an equally weighted sample denoted $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{M_N}$, also targeting ν_κ. To state the results, we consider multinomial resampling, but similar results can be obtained for other forms of resampling and branching. In a third step, we affect to the resampled particles the weights $\Phi^\kappa(\tilde{\xi}_{N,i})$, that is, we consider the weighted sample $\{(\tilde{\xi}_{N,i}, \Phi^\kappa(\tilde{\xi}_{N,i}))\}_{i=1}^{M_N}$, which obviously targets ν.

Provided that $\{(\xi_{N,i}, \Phi(\xi_{N,i}))\}_{i=1}^{M_N}$ is asymptotically normal for ν, the following result shows that $\{(\tilde{\xi}_{N,i}, \Phi^\kappa(\tilde{\xi}_{N,i}))\}_{i=1}^{M_N}$ is also asymptotically normal and provides an explicit expression for the asymptotic variance.

THEOREM 9. *Assume that* $\{(\xi_{N,i}, \Phi(\xi_{N,i}))\}_{i=1}^{M_N}$ *is consistent for* (ν, C) *and asymptotically normal for* $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$. *Assume, in addition, that* $a_N^{-2} \times M_N \stackrel{\mathrm{P}}{\longrightarrow} \beta$ *and that* $\Phi^\kappa, \Phi^{-\kappa} \in \mathsf{C}$, $\Phi^{-2\kappa} \in \mathsf{W}$. *Then*, $\{(\tilde{\xi}_{N,i}, \Phi^\kappa(\tilde{\xi}_{N,i}))\}_{i=1}^{M_N}$ *is consistent for* (ν, C) *and asymptotically normal for* $(\nu, \mathsf{A}_\kappa, \mathsf{W}_\kappa, \sigma_\kappa, \gamma_\kappa, \{a_N\})$, *where* $\mathsf{A}_\kappa \stackrel{\text{def}}{=} \{f \in \mathsf{A}, f^2 \in \mathsf{W}, f^2 \Phi^\kappa \in \mathsf{C}\}$, $\mathsf{W}_\kappa \stackrel{\text{def}}{=} \{f : \Phi^\kappa f \in \mathsf{C}\}$, $\gamma_\kappa(\cdot) \stackrel{\text{def}}{=} \beta \nu(\Phi^{-\kappa}) \nu(\Phi^\kappa \cdot)$

and

$$\sigma_\kappa^2(f) \stackrel{\text{def}}{=} \beta\nu(\Phi^{-\kappa})\nu\big[\Phi^\kappa\big(f-\nu(f)\big)^2\big] + \sigma^2[f-\nu(f)].$$

PROOF. The proof follows by applying Theorems 1 and 2 [with $R(\xi, d\tilde{\xi}) = \delta_\xi(d\tilde{\xi})$, $L(\xi, d\tilde{\xi}) = \Phi^{-\kappa}(\xi)\delta_\xi(d\tilde{\xi})$ and $\alpha = 1$] to show the consistency and asymptotic normality of the weighted sample $\{(\xi_{N,i}, \Phi^{1-\kappa}(\xi_{N,i}))\}_{i=1}^{M_N}$, then Theorems 3 and 4 to prove the consistency and asymptotic normality of the equally weighted sample $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{M_N}$ and again Theorems 1 and 2 [this time with $R(\xi, d\tilde{\xi}) = \delta_\xi(d\tilde{\xi})$, $L(\xi, d\tilde{\xi}) = \Phi^\kappa(\xi)\delta_\xi(d\tilde{\xi})$ and $\alpha = 1$] to prove the asymptotic normality of $\{(\tilde{\xi}_{N,i}, \Phi^\kappa(\tilde{\xi}_{N,i}))\}_{i=1}^{M_N}$. □

The asymptotic variance of the weighted sample $\{(\tilde{\xi}_{N,i}, \Phi^\kappa(\tilde{\xi}_{N,i}))\}_{i=1}^{M_N}$ after "weighted" resampling is higher than the variance σ^2 of the "original" weighted sample $\{(\xi_{N,i}, \Phi(\xi_{N,i}))\}_{i=1}^{M_N}$. There is no obvious "optimal" choice for the exponent κ, and it is easy to find functions f for which "unweighted" resampling performs better. For example, assume that $\{\xi_{N,i}\}_{1\le i\le M_N}$ is an i.i.d. sample from a distribution with density $f(x) \propto \exp(-x^2/2 - |x|)$ and take $\Phi(x) = \exp|x|$; ν is the Gaussian distribution $\mathcal{N}(0,1)$. Take $f_{a,b} = \mathbf{1}\{a \le |X| \le a+b\}$ for $a, b > 0$. Then, $\nu(f_{a,b}) = 0$ and by straightforward calculation,

$$\left.\frac{d\sigma_\kappa^2(f_{a,b})}{d\kappa}\right|_{\kappa=0} = \beta\nu(\mathbf{1}\{a \le |X| \le a+b\})\left(-\nu(|X|) + \frac{\nu(|X|\mathbf{1}\{a \le |X| \le a+b\})}{\nu(\mathbf{1}\{a \le |X| \le a+b\})}\right),$$

which can thus be negative or positive depending on the values of a and b.

3.2. *State-space models.* In this section we apply the results to state-space models (see, e.g., [2], Chapters 3, 4 and [3] for an introduction to that field). The state process $\{X_k\}_{k\ge1}$ is a Markov chain on a general state space X with initial distribution χ and kernel Q. The observations $\{Y_k\}_{k\ge1}$ take values in Y that are independent conditionally on the state sequence $\{X_k\}_{k\ge1}$; in addition, there exists a measure λ on $(\mathsf{Y}, \mathcal{B}(\mathsf{Y}))$, and a transition density function $x \mapsto g(x,y)$, referred to as the likelihood, such that $\mathrm{P}(Y_k \in A | X_k = x) = \int_A g(x,y)\lambda(dy)$, for all $A \in \mathsf{Y}$. The kernel Q and the likelihood functions $x \mapsto g(x,y)$ are assumed to be known. These quantities could be time-dependent. The (joint) smoothing distribution $\phi_{\chi,k}(y_{1:k}, \cdot)$ is defined as

$$\phi_{\chi,k}(y_{1:k}, f) \stackrel{\text{def}}{=} \mathrm{E}[f(X_{1:k})|Y_{1:k} = y_{1:k}], \qquad f \in \mathbb{B}(\mathsf{X}^k), \tag{18}$$

where for any sequence $\{a_i\}_{1\le i\le k}$, $a_{1:k} \stackrel{\text{def}}{=} (a_1, \ldots, a_k)$. We shall consider the case in which the observations have an arbitrary but fixed value $y_{1:k}$ and denote $g_k(x) =$

$g(x, y_k)$. In the Monte Carlo framework, we approximate the posterior distribution $\phi_{\chi,k}$ using a weighted sample $\{(\xi_{N,i}^{(k)}, \omega_{N,i}^{(k)})\}_{1\le k\le M_N}$, where $\xi_{N,i}^{(k)} \in \mathsf{X}^k$. $\xi_{N,i}^{(k)}$ are referred to as *path particles* in the literature; see [6].

We will apply the results presented in Section 2. It is first required to define a transition kernel L_{k-1} satisfying (4) with $\nu = \phi_{\chi,k-1}$, $(\Xi, \mathcal{B}(\Xi)) = (\mathsf{X}^{k-1}, \mathcal{B}(\mathsf{X}^{k-1}))$, $\mu = \phi_{\chi,k}$ and $(\tilde{\Xi}, \mathcal{B}(\tilde{\Xi})) = (\mathsf{X}^k, \mathcal{B}(\mathsf{X}^k))$, that is, for any $f \in \mathbb{B}^+(\mathsf{X}^k)$,

$$\phi_{\chi,k}(f) = \frac{\int\cdots\int \phi_{\chi,k-1}(dx_{1:k-1}) L_{k-1}(x_{1:k-1}, d\tilde{x}_{1:k}) f(\tilde{x}_{1:k})}{\int\cdots\int \phi_{\chi,k-1}(dx_{1:k-1}) L_{k-1}(dx_{1:k-1}, \mathsf{X}^k)}. \tag{19}$$

In the second step, we must choose a proposal kernel R_{k-1} satisfying

$$L_{k-1}(x_{1:k-1}, \cdot) \ll R_{k-1}(x_{1:k-1}, \cdot) \qquad \text{for any } x_{1:k-1} \in \mathsf{X}^{k-1}. \tag{20}$$

We proceed from the weighted sample $\{(\xi_{N,i}^{(k-1)}, \omega_{N,i}^{(k-1)})\}_{i=1}^{M_N}$ targeting $\phi_{\chi,k-1}$ to $\{(\xi_{N,i}^{(k)}, \omega_{N,i}^{(k)})\}_{i=1}^{M_N}$ targeting $\phi_{\chi,k}$ as follows. To keep the discussion simple, it is assumed that each particle gives birth to a single offspring. In the proposal step, we draw $\{\tilde{\xi}_{N,i}^{(k)}\}_{i=1}^{M_N}$ conditionally independent given $\mathcal{F}_N^{(k-1)}$ with distribution given, for any $f \in \mathbb{B}^+(\mathsf{X}^k)$, by

$$\begin{aligned} \mathrm{E}\big[f(\tilde{\xi}_{N,i}^{(k)}) | \mathcal{F}_N^{(k-1)}\big] &= R_{(k-1)}(\xi_{N,i}^{k-1}, f) \\ &= \int\cdots\int R_{k-1}(\xi_{N,i}^{(k-1)}, dx_{1:k}) f(x_{1:k}), \end{aligned} \tag{21}$$

where $i = 1, \ldots, M_N$. Next we assign to the particle $\tilde{\xi}_{N,i}^{(k)}$, $i = 1, \ldots, M_N$, the importance weight $\tilde{\omega}_{N,i}^{(k)} = \omega_{N,i}^{(k-1)} W_{k-1}(\xi_{N,i}^{(k-1)}, \tilde{\xi}_{N,i}^{(k)})$ with

$$W_{k-1}(x_{1:k-1}, \tilde{x}_{1:k}) = \frac{dL_{k-1}(x_{1:k-1}, \cdot)}{dR_{k-1}(x_{1:k-1}, \cdot)}(\tilde{x}_{1:k}). \tag{22}$$

Instead of resampling at each iteration (which is the assumption upon which most of the asymptotic analysis have been carried out so far), we rejuvenate the particle system only when the importance weights are too skewed. As discussed in [25], Section 4 a sensible approach is to control the coefficient of variations of weights, defined by

$$\big[\mathrm{CV}_N^{(k)}\big]^2 \stackrel{\text{def}}{=} \frac{1}{M_N}\sum_{i=1}^{M_N}\big(M_N \tilde{\omega}_{N,i}^{(k)}/\tilde{\Omega}_N^{(k)} - 1\big)^2.$$

The coefficient of variation is minimal when the normalized importance weights $\tilde{\omega}_{N,i}^{(k)}/\tilde{\Omega}_N^{(k)}$, $i = 1, \ldots, M_N$, are all equal to $1/M_N$, in which case $\mathrm{CV}_N^{(k)} = 0$. The maximal value of $\mathrm{CV}_N^{(k)}$ is $\sqrt{M_N - 1}$, which corresponds to one of the normalized weights being one and all others being null. Therefore, the coefficient of variation is often interpreted as a measure of the number of ineffective particles.

When the coefficient of variation $\mathrm{CV}_N^{(k)} \geq \kappa$ falls below a threshold κ, that is, we draw $I_k^{N,1}, \ldots, I_k^{N,M_N}$ conditionally independent given $\tilde{\mathcal{F}}_N^{(k)} = \mathcal{F}_N^{(k-1)} \vee \sigma(\{\tilde{\xi}_{N,i}^{(k)}, \tilde{\omega}_{N,i}^{(k)}\}_{i=1}^{M_N})$, with distribution

$$\mathrm{P}\big(I_k^{N,i} = j \,|\, \tilde{\mathcal{F}}_N^{(k)}\big) = \tilde{\omega}_{N,i}^{(k)} / \widetilde{\Omega}_N^{(k)}, \qquad i = 1, \ldots, M_N,\ j = 1, \ldots, M_N \tag{23}$$

and we set $\bar{\xi}_{N,i}^{(k)} = \tilde{\xi}_{N,I_k^{N,i}}^{(k)}$ and $\tilde{\omega}_{N,i}^{(k)} = 1$ for $i = 1, \ldots, M_N$. If $\mathrm{CV}_N^{(k)} < \kappa$, we copy the path particles: for $i = 1, \ldots, M_N$,

$$\big(\xi_{N,i}^{(k)}, \omega_{N,i}^{(k)}\big) = \big(\tilde{\xi}_{N,i}^{(k)}, \tilde{\omega}_{N,i}^{(k)}\big)\mathbf{1}\{\mathrm{CV}_N^{(k)} \leq \kappa\} + \big(\bar{\xi}_{N,i}^{(k)}, 1\big)\mathbf{1}\{\mathrm{CV}_N^{(k)} > \kappa\}. \tag{24}$$

In both cases, we set $\tilde{\mathcal{F}}_N^{(k)} = \mathcal{F}_N^{(k)} \vee \sigma(\{(\xi_{N,i}^{(k)}, \omega_{N,i}^{(k)})\}_{i=1}^{M_N}$. We consider here only multinomial resampling, but the deterministic-plus-residual sampling or branching alternatives can be applied as well.

THEOREM 10. *For any $k > 0$, let L_k and R_k be transition kernels from $(\mathsf{X}^k, \mathcal{B}(\mathsf{X}^k))$ to $(\mathsf{X}^{k+1}, \mathcal{B}(\mathsf{X}^{k+1}))$ satisfying (19) and (20), respectively. Assume that the equally weighted sample $\{(\xi_{N,i}^{(1)}, 1)\}_{i=1}^{M_N}$ is consistent for $\{\phi_{\chi,1}, \mathsf{L}^1(\mathsf{X}, \phi_{\chi,1})\}$ and asymptotically normal for $(\phi_{\chi,1}, \mathsf{A}_1, \mathsf{W}_1, \sigma_1, \phi_{\chi,1}, \{M_N^{1/2}\})$, where A_1 and W_1 are proper sets, and define recursively (A_k) and (W_k) by*

$$\mathsf{A}_k \stackrel{\text{def}}{=} \{f \in \mathsf{L}^2(\mathsf{X}^k, \phi_{\chi,k}),\ L_{k-1}(\cdot, f) \in \mathsf{A}_{k-1}, R_{k-1}(\cdot, W_{k-1}^2 f^2) \in \mathsf{W}_{k-1}\},$$

$$\mathsf{W}_k \stackrel{\text{def}}{=} \{f \in \mathsf{L}^1(\mathsf{X}^k, \phi_{\chi,k}),\ R_{k-1}(\cdot, W_{k-1}^2 |f|) \in \mathsf{W}_{k-1}\}.$$

Assume, in addition, that, for any $k \geq 1$, $R_k(\cdot, W_k^2) \in \mathsf{W}_k$. Then for any $k \geq 1$, (A_k) and (W_k) are proper sets and $\{(\xi_{N,i}^k, \omega_{N,i}^k)\}_{i=1}^{M_N}$ is consistent for $\{\phi_{\chi,k}, \mathsf{L}^1(\mathsf{X}, \phi_{\chi,k})\}$ and asymptotically normal for $(\phi_{\chi,k}, \mathsf{A}_k, \mathsf{W}_k, \sigma_k, \gamma_k, \{M_N^{1/2}\})$, where the functions σ_k and the measure γ_k are given by

$$\sigma_k^2(f) = \varepsilon_k \operatorname{Var}_{\phi_{\chi,k}}(f) + \frac{\sigma_{k-1}^2(L_{k-1} f_{\chi,k}) + \gamma_{k-1} R_{k-1}[\{W_{k-1} f_{\chi,k} - R_{k-1}(\cdot, W_{k-1} f_{\chi,k})\}^2]}{\{\phi_{\chi,k-1} L_{k-1}(\mathsf{X}^k)\}^2}$$

$$\gamma_k(f) = \varepsilon_k \phi_{\chi,k} + (1 - \varepsilon_k) \frac{\gamma_{k-1} R_{k-1}(W_{k-1}^2 f)}{[\phi_{\chi,k-1} L_{k-1}(\mathsf{X}^k)]^2},$$

where $f_{\chi,k} \stackrel{\text{def}}{=} f - \phi_{\chi,k}(f)$, W_k is defined in (22), and

$$\varepsilon_k \stackrel{\text{def}}{=} \mathbf{1}\{[\phi_{\chi,k-1} L_{k-1}(\mathsf{X}^k)]^{-2} \gamma_{k-1} R_{k-1}(W_{k-1}^2) \geq 1 + \kappa^2\}.$$

PROOF. The proof follows by induction. Assume that for some $k > 1$, the weighted sample $\{(\xi_{N,i}^{(k-1)}, \omega_{N,i}^{(k-1)})\}_{i=1}^{M_N}$ is consistent for $\{\phi_{\chi,k-1}, \mathsf{L}^1(\mathsf{X}, \phi_{\chi,k-1})\}$

and asymptotically normal for $(\phi_{\chi,k-1}, \mathsf{A}_{k-1}, \mathsf{W}_{k-1}, \sigma_{k-1}, \gamma_{k-1}, \{M_N^{1/2}\})$. By Theorems 1 and 2, $(\tilde{\xi}_{N,i}^{(k)}, \tilde{\omega}_{N,i}^{(k)})_{i=1}^{M_N}$ is consistent for $\{\phi_{\chi,k}, \mathsf{L}^1(\mathsf{X}, \phi_{\chi,k})\}$ and asymptotically normal for $(\phi_{\chi,k}, \tilde{\mathsf{A}}_k, \tilde{\mathsf{W}}_k, \tilde{\sigma}_k, \tilde{\gamma}_k, \{M_N^{1/2}\})$, where $\tilde{\mathsf{A}}_k, \tilde{\mathsf{W}}_k, \tilde{\sigma}_k, \tilde{\gamma}_k$, are defined from $\mathsf{A}_k, \mathsf{W}_k, \sigma_k, \gamma_k$, using Theorem 2. And by Theorems 3 and 4, $(\bar{\xi}_{N,i}^{(k)}, 1)_{i=1}^{M_N}$ is consistent for $\{\phi_{\chi,k}, \mathsf{L}^1(\mathsf{X}, \phi_{\chi,k})\}$ and asymptotically normal for $(\phi_{\chi,k}, \bar{\mathsf{A}}_k, \bar{\mathsf{W}}_k, \bar{\sigma}_k, \bar{\gamma}_k, \{M_N^{1/2}\})$, where $\bar{\mathsf{A}}_k, \bar{\mathsf{W}}_k, \bar{\sigma}_k, \bar{\gamma}_k$, are defined from $\tilde{\mathsf{A}}_k, \tilde{\mathsf{W}}_k, \tilde{\sigma}_k$, $\tilde{\gamma}_k$, using Theorem 4. The asymptotic normality of $(\tilde{\xi}_{N,i}^{(k)}, \tilde{\omega}_{N,i}^{(k)})_{i=1}^{M_N}$ and $(\bar{\xi}_{N,i}^{(k)}, 1)_{i=1}^{M_N}$, combined with

$$[\mathrm{CV}_N^{(k)}]^2 = M_N \sum_{i=1}^{M_N} \left(\frac{\tilde{\omega}_{N,i}^k}{\tilde{\Omega}_N^k} \right)^2 - 1 \xrightarrow{\mathrm{P}} \tilde{\gamma}_k(1) - 1 \quad \text{and} \quad \epsilon_k = \mathbf{1}\{\tilde{\gamma}_k(1) - 1 > \kappa^2\},$$

complete the proof. $\square$

APPENDIX A: CONDITIONAL LIMITS THEOREMS FOR TRIANGULAR ARRAY OF DEPENDENT RANDOM VARIABLES

In this section we derive limit theorems for triangular arrays of dependent random variables with a random number of terms. Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space, let X be a random variable, and let $\mathcal{G}$ be a sub-σ field of $\mathcal{F}$. Define $X^+ \stackrel{\text{def}}{=} \max(X, 0)$ and $X^- \stackrel{\text{def}}{=} -\min(X, 0)$. Following [26], Section II.7, if $\min(\mathrm{E}[X^+|\mathcal{G}], \mathrm{E}[X^-|\mathcal{G}]) < \infty$, P-a.s., the generalized conditional expectation of X given $\mathcal{G}$ is defined by $\mathrm{E}[X|\mathcal{G}] = \mathrm{E}[X^+|\mathcal{G}] - \mathrm{E}[X^-|\mathcal{G}]$, where, on the P-null-set of sample points for which $\mathrm{E}[X^+|\mathcal{G}] = \mathrm{E}[X^-|\mathcal{G}] = \infty$, the difference $\mathrm{E}[X^+|\mathcal{G}] - \mathrm{E}[X^-|\mathcal{G}]$ is given an arbitrary value, for instance, zero. Let $\{M_N\}_{N\geq 0}$ be a sequence of random positive integers, $\{U_{N,i}\}_{i=1}^{M_N}$ be a triangular array of random variables, and $\{\mathcal{F}_{N,i}\}_{0\leq i\leq M_N}$ be a triangular array of sub-sigma-fields of $\mathcal{F}$. Throughout this section, it is assumed that (i) M_N is $\mathcal{F}_{N,0}$-measurable, and (ii) $\mathcal{F}_{N,i-1} \subseteq \mathcal{F}_{N,i}$ and for each N and $i = 1, \ldots, M_N$, $U_{N,i}$ is $\mathcal{F}_{N,i}$-measurable.

THEOREM A.1. *Assume that* $\mathrm{E}[|U_{N,j}||\mathcal{F}_{N,j-1}] < \infty$ P-*a.s. for any* N *and any* $j = 1, \ldots, M_N$, *and*

$$\sup_N \mathrm{P}\left(\sum_{j=1}^{M_N} \mathrm{E}[|U_{N,j}||\mathcal{F}_{N,j-1}] \geq \lambda \right) \to 0 \qquad \text{as } \lambda \to \infty \tag{25}$$

$$\sum_{j=1}^{M_N} \mathrm{E}[|U_{N,j}|\mathbf{1}\{|U_{N,j}| \geq \epsilon\}|\mathcal{F}_{N,j-1}] \xrightarrow{\mathrm{P}} 0 \qquad \text{for any positive } \epsilon. \tag{26}$$

Then, $\max_{1\leq i\leq M_N} |\sum_{j=1}^i U_{N,j} - \sum_{j=1}^i \mathrm{E}[U_{N,j}|\mathcal{F}_{N,j-1}]| \xrightarrow{\mathrm{P}} 0$.

PROOF. Assume first that for each N and each $i = 1, \dots, M_N$, $U_{N,i} \geq 0$, P-a.s. By [27], Lemma 3.5, we have that, for any constants ϵ and $\eta > 0$,

$$\mathrm{P}\left[\max_{1 \leq i \leq M_N} U_{N,i} \geq \epsilon\right] \leq \eta + \mathrm{P}\left[\sum_{i=1}^{M_N} \mathrm{P}(U_{N,i} \geq \epsilon | \mathcal{F}_{N,i-1}) \geq \eta\right].$$

From the conditional version of the Chebyshev identity,

$$\mathrm{P}\left[\max_{1 \leq i \leq M_N} U_{N,i} \geq \epsilon\right] \leq \eta + \mathrm{P}\left[\sum_{i=1}^{M_N} \mathrm{E}[U_{N,i} \mathbf{1}\{U_{N,i} \geq \epsilon\} | \mathcal{F}_{N,i-1}] \geq \eta\epsilon\right]. \tag{27}$$

Let ϵ and $\lambda > 0$ and define $\bar{U}_{N,i} \stackrel{\text{def}}{=} U_{N,i} \mathbf{1}\{U_{N,i} < \epsilon\} \mathbf{1}\{\sum_{j=1}^{i} \mathrm{E}[U_{N,j} | \mathcal{F}_{N,j-1}] < \lambda\}$. For any $\delta > 0$,

$$\begin{aligned}
&\mathrm{P}\left(\max_{1 \leq i \leq M_N} \left|\sum_{j=1}^{i} U_{N,j} - \sum_{j=1}^{i} \mathrm{E}[U_{N,j} | \mathcal{F}_{N,j-1}]\right| \geq 2\delta\right) \\
&\quad \leq \mathrm{P}\left(\max_{1 \leq i \leq M_N} \left|\sum_{j=1}^{i} \bar{U}_{N,j} - \sum_{j=1}^{i} \mathrm{E}[\bar{U}_{N,j} | \mathcal{F}_{N,j-1}]\right| \geq \delta\right) \\
&\qquad + \mathrm{P}\left(\max_{1 \leq i \leq M_N} \left|\sum_{j=1}^{i} U_{N,j} - \bar{U}_{N,j} - \sum_{j=1}^{i} \mathrm{E}[U_{N,j} - \bar{U}_{N,j} | \mathcal{F}_{N,j-1}]\right| \geq \delta\right).
\end{aligned}$$

The second term in the RHS is bounded by

$$\begin{aligned}
&\mathrm{P}\left(\max_{1 \leq i \leq M_N} U_{N,i} \geq \epsilon\right) + \mathrm{P}\left(\sum_{j=1}^{M_N} \mathrm{E}[U_{N,j} | \mathcal{F}_{N,j-1}] \geq \lambda\right) \\
&\qquad + \mathrm{P}\left(\sum_{j=1}^{M_N} \mathrm{E}[U_{N,j} \mathbf{1}\{U_{N,j} \geq \epsilon\} | \mathcal{F}_{N,j-1}] \geq \delta\right).
\end{aligned}$$

Equations (26) and (27) imply that the first and last terms in the last expression converge to zero for any $\epsilon > 0$ and (25) implies that the second term may be arbitrarily small by choosing for λ sufficiently large. Now, by the Doob maximal inequality,

$$\begin{aligned}
&\mathrm{P}\left(\max_{1 \leq i \leq M_N} \left|\sum_{j=1}^{i} \bar{U}_{N,j} - \mathrm{E}[\bar{U}_{N,j} | \mathcal{F}_{N,j-1}]\right| \geq \delta\right) \\
&\qquad \leq \delta^{-2} \mathrm{E}\left[\sum_{j=1}^{M_N} \mathrm{E}\big[(\bar{U}_{N,j} - \mathrm{E}[\bar{U}_{N,j} | \mathcal{F}_{N,j-1}])^2 | \mathcal{F}_{N,0}\big]\right].
\end{aligned}$$

This last term does not exceed

$$\delta^{-2}\,\mathrm{E}\left[\sum_{i=1}^{M_N}\mathrm{E}[\bar{U}_{N,j}^2|\mathcal{F}_{N,0}]\right] \le \delta^{-2}\epsilon\,\mathrm{E}\left[\sum_{j=1}^{M_N}\mathrm{E}[\bar{U}_{N,j}|\mathcal{F}_{N,0}]\right]$$
$$\le \delta^{-2}\epsilon\,\mathrm{E}\left[\sum_{j=1}^{M_N}\mathrm{E}[\bar{U}_{N,j}|\mathcal{F}_{N,j-1}]\right] \le \delta^{-2}\epsilon\lambda.$$

Since ϵ is arbitrary, the proof follows for $U_{N,j} \ge 0$, P-a.s., for each N and $j = 1, \dots, M_N$. The proof extends to an arbitrary triangular array $\{U_{N,j}\}_{i=1}^{M_N}$ by applying the preceding result to $\{U_{N,j}^+\}_{i=1}^{M_N}$ and $\{U_{N,j}^-\}_{1\le j\le M_N}$. $\square$

LEMMA A.2. *Assume that for all* N, $\sum_{i=1}^{M_N}\mathrm{E}[U_{N,i}^2|\mathcal{F}_{N,i-1}] = 1$, *for* $i = 1, \dots, M_N$, $\mathrm{E}[U_{N,i}|\mathcal{F}_{N,i-1}] = 0$ *and for all* $\epsilon > 0$,

$$\sum_{i=1}^{M_N}\mathrm{E}[U_{N,i}^2\mathbf{1}\{|U_{N,i}| \ge \epsilon\}|\mathcal{F}_{N,0}] \xrightarrow{\mathrm{P}} 0. \tag{28}$$

Then, for any real u, $\mathrm{E}[\exp(iu\sum_{j=1}^{M_N}U_{N,j})|\mathcal{F}_{N,0}] - \exp(-u^2/2) \xrightarrow{\mathrm{P}} 0$.

PROOF. Denote $\sigma_{N,i}^2 \stackrel{\mathrm{def}}{=} \mathrm{E}[U_{N,i}^2|\mathcal{F}_{N,i-1}]$. Write the following decomposition (with the convention $\sum_{j=a}^{b} = 0$ if $a > b$):

$$e^{iu\sum_{j=1}^{M_N}U_{N,j}} - e^{-(u^2/2)\sum_{j=1}^{M_N}\sigma_{N,j}^2}$$
$$= \sum_{l=1}^{M_N} e^{iu\sum_{j=1}^{l-1}U_{N,j}}\big(e^{iuU_{N,l}} - e^{-(u^2/2)\sigma_{N,l}^2}\big)e^{-(u^2/2)\sum_{j=l+1}^{M_N}\sigma_{N,j}^2}.$$

Since $\sum_{j=1}^{l-1}U_{N,j}$ and $\sum_{j=l+1}^{M_N}\sigma_{N,j}^2 = 1 - \sum_{j=1}^{l}\sigma_{N,j}^2$ are $\mathcal{F}_{N,l-1}$-measurable,

$$\begin{aligned}&\left|\mathrm{E}\left[\exp\left(iu\sum_{j=1}^{M_N}U_{N,j}\right) - \exp\left(-(u^2/2)\sum_{j=1}^{M_N}\sigma_{N,j}^2\right)\middle|\mathcal{F}_{N,0}\right]\right|\\ &\quad\le \sum_{l=1}^{M_N}\mathrm{E}\big[\big|\mathrm{E}[\exp(\mathrm{i}uU_{N,l})|\mathcal{F}_{N,l-1}] - \exp(-u^2\sigma_{N,l}^2/2)\big||\mathcal{F}_{N,0}\big].\end{aligned} \tag{29}$$

For any $\epsilon > 0$, it is easily shown that

$$\begin{aligned}&\mathrm{E}\left[\sum_{l=1}^{M_N}\big|\mathrm{E}\big[\exp(iuU_{N,l}) - 1 + \tfrac{1}{2}u^2\sigma_{N,l}^2|\mathcal{F}_{N,l-1}\big]\big||\mathcal{F}_{N,0}\right]\\ &\quad\le \tfrac{1}{6}\epsilon|u|^3 + u^2\sum_{l=1}^{M_N}\mathrm{E}[U_{N,l}^2\mathbf{1}\{|U_{N,l}| \ge \epsilon\}|\mathcal{F}_{N,0}].\end{aligned} \tag{30}$$

Since $\epsilon > 0$ is arbitrary, it follows from (28) that the RHS tends in probability to 0 as $N \to \infty$. Finally, for all $\epsilon > 0$,

$$\mathrm{E}\left[\sum_{l=1}^{M_N}\left|\mathrm{E}\left[\exp(-u^2\sigma_{N,l}^2/2) - 1 + \frac{1}{2}u^2\sigma_{N,l}^2|\mathcal{F}_{N,l-1}\right]\right|\Big|\mathcal{F}_{N,0}\right]$$
$$\leq \frac{u^4}{8}\sum_{l=1}^{M_N}\mathrm{E}[\sigma_{N,l}^4|\mathcal{F}_{N,0}] \leq \frac{u^4}{8}\left(\epsilon^2 + \sum_{j=1}^{M_N}\mathrm{E}[U_{N,j}^2\mathbf{1}\{|U_{N,j}| \geq \epsilon\}|\mathcal{F}_{N,0}]\right).$$

(28) shows that the RHS of the previous equation tends in probability to 0 as $N \to \infty$. The proof follows. $\square$

THEOREM A.3. *Assume that for each N and $i = 1, \dots, M_N$, $\mathrm{E}[U_{N,i}^2|\mathcal{F}_{N,i-1}] < \infty$ and*

$$\sum_{i=1}^{M_N}\{\mathrm{E}[U_{N,i}^2|\mathcal{F}_{N,i-1}] - (\mathrm{E}[U_{N,i}|\mathcal{F}_{N,i-1}])^2\} \xrightarrow{\mathrm{P}} \sigma^2 \qquad \textit{for some } \sigma^2 > 0, \tag{31}$$

$$\sum_{i=1}^{M_N}\mathrm{E}\big[U_{N,i}^2\mathbf{1}_{\{|U_{N,i}|\geq\epsilon\}}|\mathcal{F}_{N,i-1}\big] \xrightarrow{\mathrm{P}} 0 \qquad \textit{for any } \epsilon > 0. \tag{32}$$

Then, for any real u,

$$\mathrm{E}\left[\exp\left(iu\sum_{i=1}^{M_N}\{U_{N,i} - \mathrm{E}[U_{N,i}|\mathcal{F}_{N,i-1}]\}\right)\Big|\mathcal{F}_{N,0}\right] \xrightarrow{\mathrm{P}} \exp(-(u^2/2)\sigma^2). \tag{33}$$

PROOF. We first assume that $\mathrm{E}[U_{N,i}|\mathcal{F}_{N,i-1}] = 0$ for all $i = 1, \dots, M_N$, and $\sigma^2 = 1$. Define the stopping time $\tau_N \stackrel{\mathrm{def}}{=} \max\{1 \leq k \leq M_N : \sum_{j=1}^k \sigma_{N,j}^2 \leq 1\}$, with the convention $\max\varnothing = 0$. Put $\bar{U}_{N,k} = U_{N,k}$ for $k \leq \tau_N$, $\bar{U}_{N,k} = 0$ for $\tau_N < k \leq M_N$ and $\bar{U}_{N,M_N+1} = (1 - \sum_{j=1}^{\tau_N}\sigma_{N,j}^2)^{1/2}Y_N$, where $\{Y_N\}$ are $\mathcal{N}(0,1)$ independent and independent of $\mathcal{F}_{N,M_N}$. Put

$$\sum_{j=1}^{M_N} U_{N,j} = \sum_{j=1}^{M_N+1}\bar{U}_{N,j} - \bar{U}_{N,M_N+1} + \sum_{j=\tau_N+1}^{M_N} U_{N,j}. \tag{34}$$

We will prove that (a) $\{\bar{U}_{N,j}\}_{1\leq j\leq M_N+1}$ satisfies the assumptions of Lemma A.2, (b) $\bar{U}_{N,M_N+1} \xrightarrow{\mathrm{P}} 0$, and (c) $\sum_{j=\tau_N+1}^{M_N} U_{N,j} \xrightarrow{\mathrm{P}} 0$. If $\tau_N < M_N$, then for any $\epsilon > 0$,

$$0 \leq 1 - \sum_{j=1}^{\tau_N}\sigma_{N,j}^2 \leq \sigma_{N,\tau_N+1}^2$$
$$\leq \max_{1\leq j\leq M_N}\sigma_{N,j}^2 \leq \epsilon^2 + \sum_{j=1}^{M_N}\mathrm{E}[U_{N,j}^2\mathbf{1}\{|U_{N,j}| \geq \epsilon\}|\mathcal{F}_{N,j-1}].$$

Since $\epsilon > 0$ is arbitrary, it follows from (32) that $1 - \sum_{j=1}^{\tau_N} \sigma_{N,j}^2 \xrightarrow{\mathrm{P}} 0$, which implies that $\mathrm{E}[\bar{U}_{N,M_N+1}^2 | \mathcal{F}_{N,0}] \xrightarrow{\mathrm{P}} 0$, showing (a) and (b). It remains to prove (c). We have

$$\sum_{j=\tau_N+1}^{M_N} \sigma_{N,j}^2 = \sum_{j=1}^{M_N} \sigma_{N,j}^2 - 1 + \left(1 - \sum_{j=1}^{\tau_N} \sigma_{N,j}^2\right) \xrightarrow{\mathrm{P}} 0. \tag{35}$$

For any $\lambda > 0$,

$$\mathrm{E}\left(\sum_{j=\tau_N+1}^{M_N} U_{N,j} \mathbf{1}\left\{\sum_{i=\tau_N+1}^{j} \sigma_{N,i}^2 \le \lambda\right\}\right)^2 = \mathrm{E}\left(\sum_{j=\tau_N+1}^{M_N} \sigma_{N,j}^2 \mathbf{1}\left\{\sum_{i=\tau_N+1}^{j} \sigma_{N,i}^2 \le \lambda\right\}\right).$$

The term between braces converges to 0 in probability by (35) and its value is bounded by λ, which shows that $\sum_{j=\tau_N+1}^{M_N} U_{N,j} \mathbf{1}\{\sum_{i=\tau_N+1}^{j} \sigma_{N,i}^2 \le \lambda\} \xrightarrow{\mathrm{P}} 0$. Moreover,

$$\mathrm{P}\left(\sum_{j=\tau_N+1}^{M_N} U_{N,j} \mathbf{1}\left\{\sum_{i=\tau_N+1}^{j} \sigma_{N,i}^2 > \lambda\right\} \neq 0\right) \le \mathrm{P}\left(\sum_{i=\tau_N+1}^{M_N} \sigma_{N,i}^2 > \lambda\right),$$

which converges to 0 by (35). The proof is completed when $\mathrm{E}[U_{N,i} | \mathcal{F}_{N,i-1}] = 0$. To deal with the general case, it suffices to set $\bar{U}_{N,i} = U_{N,i} - \mathrm{E}[U_{N,i} | \mathcal{F}_{N,i-1}]$ and to use the following technical lemma. □

LEMMA A.4. *Let $\mathcal{G}$ be a σ-field and X a random variable such that* $\mathrm{E}[X^2 | \mathcal{G}] < \infty$. *Then, for any $\epsilon > 0$,*

$$4\,\mathrm{E}[|X|^2 \mathbf{1}\{|X| \ge \epsilon\} | \mathcal{G}] \ge \mathrm{E}\big[|X - \mathrm{E}[X|\mathcal{G}]|^2 \mathbf{1}\{|X - \mathrm{E}[X|\mathcal{G}]| \ge 2\epsilon\} | \mathcal{G}\big].$$

PROOF. Let $Y = X - \mathrm{E}[X|\mathcal{G}]$. We have $\mathrm{E}[Y|\mathcal{G}] = 0$. It is equivalent to show that, for any $\mathcal{G}$-measurable random variable Z,

$$\mathrm{E}[Y^2 \mathbf{1}\{|Y| \ge 2\epsilon\} | \mathcal{G}] \le 4\,\mathrm{E}[|Y + Z|^2 \mathbf{1}\{|Y + Z| \ge \epsilon\} | \mathcal{G}].$$

On the set $\{|Z| < \epsilon\}$,

$$\begin{aligned}
\mathrm{E}[Y^2 \mathbf{1}\{|Y| \ge 2\epsilon\} | \mathcal{G}] &\le 2\,\mathrm{E}\big[((Y+Z)^2 + Z^2) \mathbf{1}\{|Y+Z| \ge \epsilon\} | \mathcal{G}\big] \\
&\le 2(1 + Z^2/\epsilon^2)\,\mathrm{E}[(Y+Z)^2 \mathbf{1}\{|Y+Z| \ge \epsilon\} | \mathcal{G}] \\
&\le 4\,\mathrm{E}[(Y+Z)^2 \mathbf{1}\{|Y+Z| \ge \epsilon\} | \mathcal{G}].
\end{aligned}$$

Moreover, on the set $\{|Z| \ge \epsilon\}$, using that $\mathrm{E}[ZY|\mathcal{G}] = Z\,\mathrm{E}[Y|\mathcal{G}] = 0$,

$$\begin{aligned}
\mathrm{E}[Y^2 \mathbf{1}\{|Y| \ge 2\epsilon\} | \mathcal{G}] &\le \mathrm{E}[Y^2 + Z^2 - \epsilon^2 | \mathcal{G}] \\
&\le \mathrm{E}[(Y+Z)^2 - \epsilon^2 | \mathcal{G}] \le \mathrm{E}[(Y+Z)^2 \mathbf{1}\{|Y+Z| \ge \epsilon\} | \mathcal{G}].
\end{aligned}$$

The proof is completed. □

APPENDIX B: PROOF OF THEOREMS 1 AND 2

PROOF OF THEOREM 1. For $j = 1, \dots, \tilde{M}_N$, we set $\mathcal{F}_{N,j} = \mathcal{F}_{N,0} \vee \sigma(\{\tilde{\xi}_{N,k}\}_{1 \le k \le j})$, where $\mathcal{F}_{N,0}$ is given in (6). Checking that $\tilde{\mathsf{C}}$ is proper is straightforward, so we turn to the consistency. We show first that, for any $f \in \tilde{\mathsf{C}}$,

$$\frac{1}{\alpha \Omega_N} \sum_{j=1}^{\tilde{M}_N} \tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) \xrightarrow{\mathrm{P}} \nu L(f), \tag{36}$$

where $\tilde{\xi}_{N,j}$ and $\tilde{\omega}_{N,j}$ are defined in (7) and (8), respectively. The unbiasedness condition (9) implies

$$(\alpha \Omega_N)^{-1} \sum_{j=1}^{\tilde{M}_N} \mathrm{E}[\tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) | \mathcal{F}_{N,j-1}] = \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} L(\xi_{N,i}, f).$$

Because the weighted sample $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is consistent for (ν, C) and for $f \in \tilde{\mathsf{C}}$, the function $L(\cdot, f) \in \mathsf{C}$, $\Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} L(\xi_{N,i}, f) \xrightarrow{\mathrm{P}} \nu L(f)$, it suffices to show that

$$(\alpha \Omega_N)^{-1} \sum_{j=1}^{\tilde{M}_N} \{\tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) - \mathrm{E}[\tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) | \mathcal{F}_{N,j-1}]\} \xrightarrow{\mathrm{P}} 0. \tag{37}$$

Put $U_{N,j} = (\alpha \Omega_N)^{-1} \tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j})$ for $j = 1, \dots, \tilde{M}_N$ and appeal to Theorem A.1. Just as above,

$$\sum_{j=1}^{\tilde{M}_N} \mathrm{E}[|U_{N,j}| | \mathcal{F}_{N,j-1}] = \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} L(\xi_{N,i}, |f|) \xrightarrow{\mathrm{P}} \nu L(|f|),$$

showing that the sequence $\{\sum_{j=1}^{\tilde{M}_N} \mathrm{E}[|U_{N,j}| | \mathcal{F}_{N,j-1}]\}_{N \ge 0}$ is tight [Theorem A.1, equation (25)]. For any $\epsilon > 0$, put $A_N \stackrel{\text{def}}{=} \sum_{j=1}^{\tilde{M}_N} \mathrm{E}[|U_{N,j}| \mathbf{1}\{|U_{N,j}| \ge \epsilon\} | \mathcal{F}_{N,j-1}]$. We need to show that $A_N \xrightarrow{\mathrm{P}} 0$ [Theorem A.1, equation (26)]. For any positive $C, \xi \in \Xi$, $R(\xi, W|f| \mathbf{1}\{W|f| \ge C\}) \le R(\xi, W|f|) = L(\xi, |f|)$. Because the function $L(\cdot, |f|)$ belongs to the proper set C, the function $R(\cdot, W|f| \mathbf{1}\{W|f| \ge C\})$ belongs to C. Hence, for all $C, \epsilon > 0$,

$$\begin{aligned} &A_N \mathbf{1}\left\{(\alpha \Omega_N)^{-1} \max_{1 \le i \le M_N} \omega_{N,i} \le \epsilon / C\right\} \\ &\quad \le \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} \left[\alpha^{-1} \sum_{k=1}^{\alpha} R(\xi_{N,i}, W|f| \mathbf{1}\{W|f| \ge C\})\right] \\ &\quad \xrightarrow{\mathrm{P}} \nu R(W|f| \mathbf{1}\{W|f| \ge C\}). \end{aligned}$$

By dominated convergence, the RHS can be made arbitrarily small by letting $C \to \infty$. Combining with $\Omega_N^{-1} \max_{1 \le i \le M_N} \omega_{N,i} \xrightarrow{\mathrm{P}} 0$, this shows that A_N tends to zero in probability, showing (26). Thus, Theorem A.1 applies and (36) holds. Under the stated assumptions, the function $L(\cdot, \tilde{\Xi})$ belongs to C, implying that the constant function $g \equiv 1$ satisfies (36); therefore, $(\alpha\Omega_N)^{-1} \sum_{j=1}^{\tilde{M}_N} \tilde{\omega}_{N,j} \xrightarrow{\mathrm{P}} \nu L(\tilde{\Xi})$. Combined with (36), this shows that, for any $f \in \tilde{\mathsf{C}}$,

$$\tilde{\Omega}_N^{-1} \sum_{j=1}^{\tilde{M}_N} \tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) \xrightarrow{\mathrm{P}} \mu(f).$$

To complete the proof, it remains to prove that $\tilde{\Omega}_N^{-1} \max_{1 \le j \le \tilde{M}_N} \tilde{\omega}_{N,j} \xrightarrow{\mathrm{P}} 0$. Since $(\alpha\Omega_N)^{-1}\tilde{\Omega}_N \xrightarrow{\mathrm{P}} \nu L(\tilde{\Xi})$, it suffices to show that $(\alpha\Omega_N)^{-1} \times \max_j \tilde{\omega}_{N,j} \xrightarrow{\mathrm{P}} 0$. For any $C > 0$,

$$(\alpha\Omega_N)^{-1} \max_{1 \le j \le \tilde{M}_N} \tilde{\omega}_{N,j} \mathbf{1}_{\{W(\tilde{\xi}_{N,j}) \le C\}} \le C(\alpha\Omega_N)^{-1} \max_{1 \le i \le M_N} \omega_{N,i} \xrightarrow{\mathrm{P}} 0,$$

$$\begin{aligned}(\alpha\Omega_N)^{-1} \max_{1 \le j \le \tilde{M}_N} \tilde{\omega}_{N,j} \mathbf{1}_{\{W(\tilde{\xi}_{N,j}) > C\}} &\le (\alpha\Omega_N)^{-1} \sum_{j=1}^{\tilde{M}_N} \tilde{\omega}_{N,j} \mathbf{1}_{\{W(\tilde{\xi}_{N,j}) > C\}} \\ &\xrightarrow{\mathrm{P}} \nu L(\{W > C\}).\end{aligned}$$

The term in the RHS of the last equation goes to zero as $C \to \infty$, which concludes the proof. □

PROOF OF THEOREM 2. Pick $f \in \tilde{\mathsf{A}}$ and assume, without loss of generality, that $\mu(f) = 0$. Write $\tilde{\Omega}_N^{-1} \sum_{i=1}^{\tilde{M}_N} \tilde{\omega}_{N,i} f(\tilde{\xi}_{N,i}) = (\alpha\Omega_N/\tilde{\Omega}_N)(A_N + B_N)$, with

$$A_N = (\alpha\Omega_N)^{-1} \sum_{j=1}^{\tilde{M}_N} \mathrm{E}[\tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) | \mathcal{F}_{N,j-1}] = \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} L(\xi_{N,i}, f),$$

$$B_N = (\alpha\Omega_N)^{-1} \sum_{j=1}^{\tilde{M}_N} \{\tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) - \mathrm{E}[\tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}) | \mathcal{F}_{N,j-1}]\}.$$

Because $\alpha\Omega_N/\tilde{\Omega}_N \xrightarrow{\mathrm{P}} 1$ (see Theorem 1), the conclusion of the theorem follows from Slutsky's theorem if we prove that $a_N(A_N + B_N) \xrightarrow{\mathcal{D}} \mathrm{N}(0, \sigma^2(Lf) + \eta^2(f))$, where

$$\eta^2(f) \stackrel{\text{def}}{=} \alpha^{-1}\gamma R\{[Wf - R(\cdot, Wf)]^2\}, \tag{38}$$

with W given in (5). The function $L(\cdot, f)$ belongs to A and $\nu L(f) = \mu(f) \times \nu L(\tilde{\boldsymbol{\Xi}}) = 0$. Because $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is asymptotically normal for $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$, $a_N A_N \xrightarrow{\mathcal{D}} \mathrm{N}(0, \sigma^2(Lf))$. Next we prove that, for any real u,

$$\mathrm{E}[\exp(iua_N B_N)|\mathcal{F}_{N,0}] \xrightarrow{\mathrm{P}} \exp(-(u^2/2)\eta^2(f)),$$

where $\eta^2(f)$ is defined in (38). For that purpose, we use Theorem A.3, and we thus need to check (31)–(32) with

$$U_{N,j} \stackrel{\text{def}}{=} (\alpha\Omega_N)^{-1} a_N \tilde{\omega}_{N,j} f(\tilde{\xi}_{N,j}), \qquad j = 1, \ldots, \tilde{M}_N.$$

Under the stated assumptions, for $f \in \mathsf{A}$, the function $R(\cdot, W^2 f^2)$ belongs to W. Because the W is proper and the function $R(\cdot, W^2 f^2) \in \mathsf{W}$, the relation $\{L(\cdot, f)\}^2 = \{R(\xi, Wf)\}^2 \le R(\cdot, W^2 f^2)$ implies that the function $\{L(\cdot, f)\}^2$ also belongs to W. Because $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is asymptotically normal for $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$, (2) implies

$$\sum_{j=1}^{\tilde{M}_N} \mathrm{E}[U_{N,j}^2|\mathcal{F}_{N,j-1}] = \alpha^{-1} \frac{a_N^2}{\Omega_N^2} \sum_{i=1}^{M_N} \omega_{N,i}^2 R(\xi_{N,i}, W^2 f^2) \xrightarrow{\mathrm{P}} \alpha^{-1}\gamma\, R(W^2 f^2),$$

$$\sum_{j=1}^{\tilde{M}_N} (\mathrm{E}[U_{N,j}|\mathcal{F}_{N,j-1}])^2 = \alpha^{-1} \frac{a_N^2}{\Omega_N^2} \sum_{i=1}^{M_N} \omega_{N,i}^2 \{L(\xi_{N,i}, f)\}^2 \xrightarrow{\mathrm{P}} \alpha^{-1}\gamma\{R(Wf)\}^2.$$

These displays imply that (31) holds. It remains to check (32). For $\epsilon > 0$, denote $C_N \stackrel{\text{def}}{=} (\sum_{j=1}^{\tilde{M}_N} \mathrm{E}[U_{N,j}^2 \mathbf{1}_{\{|U_{N,j}| \ge \epsilon\}}|\mathcal{F}_{N,j-1}])$. For all $C > 0$, it is easily shown that

$$C_N \le \frac{a_N^2}{\alpha\Omega_N^2} \sum_{i=1}^{M_N} \omega_{N,i}^2 R(\xi_{N,i}, W^2 f^2 \mathbf{1}\{|Wf| \ge C\})$$
$$+ \mathbf{1}\left\{\frac{a_N \max_{1\le i\le M_N} \omega_{N,i}}{\alpha\Omega_N} \ge \frac{\epsilon}{C}\right\} \sum_{j=1}^{\tilde{M}_N} \mathrm{E}[U_{N,j}^2|\mathcal{F}_{N,j-1}].$$

Since $a_N \Omega_N^{-1} \max_{1\le i\le M_N} \omega_{N,i} \xrightarrow{\mathrm{P}} 0$ and the function $R(\cdot, W^2 f^2) \in \mathsf{W}$, the RHS of the previous display converges in probability to $\gamma R(W^2 f^2 \mathbf{1}\{|Wf| \ge C\})$, which can be made arbitrarily small by taking C sufficiently large. Therefore, condition (32) is satisfied and Theorem A.3 applies, showing that $a_N(A_N + B_N) \xrightarrow{\mathcal{D}} \mathrm{N}(0, \sigma^2(Lf) + \eta^2(f))$.

Consider now (2). Recalling that $\tilde{\Omega}_N/(\alpha\Omega_N) \xrightarrow{\mathrm{P}} \nu L(\tilde{\boldsymbol{\Xi}})$, it is sufficient to show that, for $f \in \tilde{\mathsf{W}}$,

$$\frac{a_N^2}{(\alpha\Omega_N)^2} \sum_{j=1}^{\tilde{M}_N} \tilde{\omega}_{N,j}^2 f(\tilde{\xi}_{N,j}) \xrightarrow{\mathrm{P}} \alpha^{-1}\gamma R(h), \tag{39}$$

where $h \stackrel{\text{def}}{=} W^2 f$. Define $U_{N,j} = (\alpha\Omega_N)^{-2} a_N^2 \tilde{\omega}_{N,j}^2 f(\tilde{\xi}_{N,j})$. Under the stated assumptions, for any $f \in \tilde{\mathsf{W}}$, the functions $R(\cdot, |h|)$ and $R(\cdot, h)$ belong to W. Because $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is asymptotically normal for $(\nu, \mathsf{A}, \mathsf{W}, \sigma, \gamma, \{a_N\})$,

$$\sum_{j=1}^{\tilde{M}_N} \mathrm{E}[|U_{N,j}||\mathcal{F}_{N,j-1}] = \frac{a_N^2}{\alpha\Omega_N^2} \sum_{i=1}^{M_N} \omega_{N,i}^2 R(\xi_{N,i}, |h|) \xrightarrow{\mathrm{P}} \alpha^{-1}\gamma R(|h|), \tag{40}$$

$$\sum_{j=1}^{\tilde{M}_N} \mathrm{E}[U_{N,j}|\mathcal{F}_{N,j-1}] = \frac{a_N^2}{\alpha\Omega_N^2} \sum_{i=1}^{M_N} \omega_{N,i}^2 R(\xi_{N,i}, h) \xrightarrow{\mathrm{P}} \alpha^{-1}\gamma R(h). \tag{41}$$

We appeal to Theorem A.1. Equation (40) shows the tightness condition (25). For $\epsilon > 0$, set $C_N = \sum_{j=1}^{\tilde{M}_N} \mathrm{E}[|U_{N,j}|\mathbf{1}_{\{|U_{N,j}| \geq \epsilon\}}|\mathcal{F}_{N,j-1}]$. Since $R(\cdot, |h|)$ belongs to W, the function $R(\cdot, |h|\mathbf{1}\{|h| \geq C\})$ belongs to W. For all $C > 0$,

$$\begin{aligned} C_N \leq{}& \frac{a_N^2}{\Omega_N^2} \sum_{i=1}^{M_N} \omega_{N,i}^2 R(\xi_{N,i}, |h|\mathbf{1}\{|h| \geq C\}) \\ &+ \mathbf{1}\left\{\max_i \frac{M_N \omega_{N,i}^2}{(\alpha\Omega_N)^2} \geq \frac{\epsilon}{C}\right\} \frac{a_N^2}{\Omega_N^2} \sum_{i=1}^{M_N} \omega_{N,i}^2 R(\xi_{N,i}, |h|). \end{aligned}$$

Proceeding as above, $C_N \xrightarrow{\mathrm{P}} 0$. Thus, Theorem A.1 applies and condition (2) is proved. Consider finally (3). Combining with $\tilde{\Omega}_N/(\alpha\Omega_N) \xrightarrow{\mathrm{P}} \nu L(\tilde{\boldsymbol{\Xi}})$ (see proof of Theorem 1) and $\tilde{M}_N = \alpha M_N$, it is actually sufficient to show that $C_N \stackrel{\text{def}}{=} (\alpha\Omega_N)^{-2} a_N^2 \max_{1 \leq j \leq \tilde{M}_N} \tilde{\omega}_{N,j}^2 \xrightarrow{\mathrm{P}} 0$. For any $C > 0$,

$$C_N \leq C^2 a_N^2 \frac{\max_{1 \leq i \leq M_N} \omega_{N,i}^2}{\alpha^2 \Omega_N^2} + \frac{a_N^2}{\alpha^2\Omega_N^2} \sum_{j=1}^{\tilde{M}_N} \tilde{\omega}_{N,j}^2 \mathbf{1}_{\{W(\xi_{N,\lfloor j/\alpha \rfloor + 1}, \tilde{\xi}_{N,j}) \geq C\}}.$$

Applying (39) with $f \equiv 1$, the RHS of the previous display converges in probability to $\gamma R(W^2 \mathbf{1}\{W \geq C\})$. The proof follows since C is arbitrary. □

APPENDIX C: PROOF OF THEOREM 3 AND 4

PROOF OF THEOREM 3. As above, we set for $j = 1, \ldots, \tilde{M}_N$, $\mathcal{F}_{N,j} = \mathcal{F}_{N,0} \vee \sigma(\{\tilde{\xi}_{N,k}\}_{1 \leq k \leq j})$, where $\mathcal{F}_{N,0}$ is given in (6). It is easily shown that the resampling procedures (12) and (13) satisfy, for any $f \in \mathbb{B}(\boldsymbol{\Xi})$,

$$\tilde{M}_N^{-1} \sum_{i=1}^{\tilde{M}_N} \mathrm{E}[f(\tilde{\xi}_{N,i})|\mathcal{F}_{N,i-1}] = \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f(\xi_{N,i}),$$

whatever the choice of the labels of the particles are (and that in both cases, these quantities are independent conditionally to $\mathcal{F}_{N,0}$).

Pick f in C. Since C is proper, $|f|\mathbf{1}\{|f| \ge C\} \in \mathsf{C}$ for any $C \ge 0$. Because $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is consistent for (ν, C), and C is a proper set of functions,

$$
\begin{aligned}
&\tilde{M}_N^{-1} \sum_{i=1}^{\tilde{M}_N} \mathrm{E}\big[|f(\tilde{\xi}_{N,i})| \mathbf{1}_{\{|f(\tilde{\xi}_{N,i})| \ge C\}} | \mathcal{F}_{N,i-1}\big] \\
&\qquad = \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} |f(\xi_{N,i})| \mathbf{1}_{\{|f(\xi_{N,i})| \ge C\}} \xrightarrow{\mathrm{P}} \nu\big(|f| \mathbf{1}_{\{|f| \ge C\}}\big).
\end{aligned} \tag{42}
$$

We now check (25)–(26) of Theorem A.1. For any $i = 1, \dots, M_N$, put $U_{N,i} \stackrel{\text{def}}{=} \tilde{M}_N^{-1} f(\tilde{\xi}_{N,i})$. Taking $C = 0$ in (42),

$$
\sum_{i=1}^{\tilde{M}_N} \mathrm{E}[|U_{N,i}| | \mathcal{F}_{N,i-1}] = \tilde{M}_N^{-1} \sum_{i=1}^{\tilde{M}_N} \mathrm{E}[|f(\tilde{\xi}_{N,i})| | \mathcal{F}_{N,i-1}] \xrightarrow{\mathrm{P}} \nu(|f|) < \infty,
$$

whence the sequence $\{\sum_{i=1}^{\tilde{M}_N} \mathrm{E}[|U_{N,i}| | \mathcal{F}_{N,i-1}]\}_{N \ge 0}$ is tight. Next, for any positive ϵ and C, we have for sufficiently large N

$$
\begin{aligned}
&\sum_{i=1}^{\tilde{M}_N} \mathrm{E}\big[|U_{N,i}| \mathbf{1}_{\{|U_{N,i}| \ge \epsilon\}} | \mathcal{F}_{N,i-1}\big] \\
&\qquad = \frac{1}{\tilde{M}_N} \sum_{i=1}^{\tilde{M}_N} \mathrm{E}\big[|f(\tilde{\xi}_{N,i})| \mathbf{1}_{\{|f|(\tilde{\xi}_{N,i}) \ge \epsilon \tilde{M}_N\}} | \mathcal{F}_{N,i-1}\big] \\
&\qquad \le \tilde{M}_N^{-1} \sum_{i=1}^{\tilde{M}_N} \mathrm{E}\big[|f(\tilde{\xi}_{N,i})| \mathbf{1}_{\{|f|(\tilde{\xi}_{N,i}) \ge C\}} | \mathcal{F}_{N,i-1}\big] \xrightarrow{\mathrm{P}} \mu\big(|f| \mathbf{1}_{\{|f| \ge C\}}\big).
\end{aligned}
$$

By dominated convergence, the RHS of this display tends to zero as $C \to \infty$. Thus, the LHS of the display converges to zero in probability, showing (26). □

PROOF OF THEOREM 4. Let $f \in \tilde{\mathsf{A}}$ and rewrite the sum $\tilde{M}_N^{-1} \sum_{i=1}^{\tilde{M}_N} f(\tilde{\xi}_{N,i}) - \nu(f) = A_N + B_N$, where $A_N = \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} \{f(\xi_{N,i}) - \nu(f)\}$ and $B_N = \tilde{M}_N^{-1} \sum_{i=1}^{\tilde{M}_N} \{f(\tilde{\xi}_{N,i}) - \mathrm{E}[f(\tilde{\xi}_{N,i}) | \mathcal{F}_{N,i-1}]\}$. We first prove that

$$
\mathrm{E}[\exp(iua_N B_N) | \mathcal{F}_{N,0}] \xrightarrow{\mathrm{P}} \exp(-(u^2/2)\beta \operatorname{Var}_\nu(f)). \tag{43}
$$

We will appeal to Theorem A.3 and, hence, need to check (31)–(32) with $U_{N,i} \stackrel{\text{def}}{=} a_N \tilde{M}_N^{-1} f(\tilde{\xi}_{N,i})$. First, because $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is consistent for (ν, C) and since for $f \in \tilde{\mathsf{A}}$, $f^2 \in \mathsf{C}$,

$$\sum_{j=1}^{\tilde{M}_N} \{\mathrm{E}[U_{N,j}^2 | \mathcal{F}_{N,j-1}] - (\mathrm{E}[U_{N,j} | \mathcal{F}_{N,j-1}])^2\}$$
$$= a_N \tilde{M}_N^{-1} \left(\Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f^2(\xi_{N,i}) - \left\{ \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f(\xi_{N,i}) \right\}^2 \right)$$
$$\stackrel{\mathrm{P}}{\longrightarrow} \beta\big(\nu(f^2) - \{\nu(f)\}^2\big),$$

showing (31). Pick $\epsilon > 0$. For any $C > 0$, there exists N_C sufficiently large such that, for all $N \geq N_C$,

$$\sum_{j=1}^{\tilde{M}_N} \mathrm{E}\big[U_{N,j}^2 \mathbf{1}_{\{|U_{N,j}| \geq \epsilon\}} | \mathcal{F}_{N,j-1}\big]$$
$$\leq a_N^2 \tilde{M}_N^{-2} \sum_{i=1}^{\tilde{M}_N} \mathrm{E}[f^2(\tilde{\xi}_{N,j}) \mathbf{1}\{|f(\tilde{\xi}_{N,j})| \geq C\} | \mathcal{F}_{N,j-1}]$$
$$= a_N^2 \tilde{M}_N^{-1} \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f^2(\xi_{N,i}) \mathbf{1}\{|f(\xi_{N,i})| \geq C\}.$$

Since f^2 belongs to the proper set C, the function $f^2 \mathbf{1}\{|f| \geq C\}$ also belongs to C and the RHS of the above display converges in probability to $\nu(f^2 \mathbf{1}\{|f| \geq C\})$. (32) follows because C is arbitrary. Condition (1) follows by combining (43) with $a_N A_N \stackrel{\mathcal{D}}{\longrightarrow} \mathrm{N}(0, \sigma^2(f))$. Conditions (2) and (3) are trivially satisfied. □

APPENDIX D: PROOF OF THEOREM 5

PROOF OF THEOREM 5. Pick $f \in \tilde{\mathsf{A}}$. To apply Theorem A.3, we just have to check (31) and (32) where $U_{N,i} = a_N \tilde{M}_N^{-1} f(\tilde{\xi}_{N,i})$. Set $A_N \stackrel{\mathrm{def}}{=} \sum_{i=1}^{\tilde{M}_N} \{\mathrm{E}[U_{N,i}^2 | \mathcal{F}_{N,i-1}] - (\mathrm{E}[U_{N,i} | \mathcal{F}_{N,i-1}])^2\}$. Note that

$$A_N = \frac{a_N^2}{\tilde{M}_N^2} \left(\sum_{i=1}^{M_N} \langle \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rangle \right)$$
$$\times \left\{ \sum_{i=1}^{M_N} \tilde{\omega}_{N,i} f^2(\xi_{N,i}) - \left(\sum_{i=1}^{M_N} \tilde{\omega}_{N,i} f(\xi_{N,i}) \right)^2 \right\},$$

where the weights $\tilde{\omega}_{N,i}$ are given by $\tilde{\omega}_{N,i} = \frac{\langle \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rangle}{\sum_{i=1}^{M_N} \langle \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rangle}$. With this notation, A_N may be rewritten as

$$A_N = \frac{a_N^2}{\tilde{M}_N^2} \sum_{i=1}^{M_N} \langle \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rangle f^2(\xi_{N,i}) - \frac{a_N^2}{\tilde{M}_N} \frac{(\tilde{M}_N^{-1} \sum_{i=1}^{M_N} \langle \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rangle f(\xi_{N,i}))^2}{\tilde{M}_N^{-1} \sum_{i=1}^{M_N} \langle \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rangle},$$

and Lemma A.5 shows that $A_N \xrightarrow{\mathrm{P}} \tilde{\beta} \ell^{-1} \nu_{\ell,\Phi}\{(f - \nu_{\ell,\Phi}(f)/\nu_{\ell,\Phi}(\mathbf{1}))^2\}$, since $f^2 \in \mathsf{C}$. Conditions (2) and (3) are trivially satisfied and the theorem follows. $\square$

LEMMA A.5. *Under the assumptions of Theorem 5, for any* $f \in \mathsf{C}$,

$$\frac{1}{\tilde{M}_N} \sum_{i=1}^{M_N} \langle \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rangle f(\xi_{N,i}) \xrightarrow{\mathrm{P}} \nu\biggl(f \frac{\langle \ell \nu(\Phi^{-1}) \Phi \rangle}{\ell \nu(\Phi^{-1}) \Phi}\biggr). \tag{44}$$

PROOF. For any $K \geq 1$, denote

$$\mathsf{B}_K \stackrel{\text{def}}{=} [K, \infty] \cup \bigcup_{j=0}^{K} [j - 1/K, j + 1/K]. \tag{45}$$

Because the weighted sample $\{(\xi_{N,i}, \Phi(\xi_{N,i})\}_{i=1}^{M_N}$ is consistent for (ν, C) and $\lfloor \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rfloor \leq \tilde{M}_N \Omega_N^{-1} \omega_{N,i}$, we have, for any $f \in \mathsf{C}$,

$$\begin{aligned}
&\tilde{M}_N^{-1} \sum_{i=1}^{M_N} \lfloor \tilde{M}_N \Omega_N^{-1} \omega_{N,i} \rfloor |f(\xi_{N,i})| \mathbf{1}\{\ell \nu(\Phi^{-1}) \omega_{N,i} \in \mathsf{B}_K\} \\
&\qquad \leq \frac{1}{\Omega_N} \sum_{i=1}^{M_N} \omega_{N,i} |f(\xi_{N,i})| \mathbf{1}\{\ell \nu(\Phi^{-1}) \Phi(\xi_{N,i}) \in \mathsf{B}_K\} \\
&\qquad \xrightarrow{\mathrm{P}} \nu\bigl(|f| \mathbf{1}\{\ell \nu(\Phi^{-1}) \Phi \in \mathsf{B}_K\}\bigr).
\end{aligned}$$

The RHS of the previous display can be made arbitrarily small by taking K sufficiently large because $\int f(\xi) \mathbf{1}\{\ell \nu(\Phi^{-1}) \Phi(\xi) \in \{\infty\} \cup \mathbb{N}\} \nu(d\xi) = 0$. For any given $K > 0$, since Φ^{-1} belongs to C and $\{\xi_{N,i}, \omega_{N,i}\}_{i=1}^{M_N}$ is consistent for (ν, C),

$$\tilde{M}_N \Omega_N^{-1} = \frac{\tilde{M}_N}{M_N} \Biggl(\Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} \Phi^{-1}(\xi_{N,i})\Biggr) \xrightarrow{\mathrm{P}} \ell \nu(\Phi^{-1}).$$

Because Φ is a proper set, for any $f \in \mathsf{C}$ and $K \geq 0$, the function

$$g_K \stackrel{\text{def}}{=} \frac{\lfloor \ell \nu(\Phi^{-1}) \Phi(\cdot) \rfloor}{\ell \nu(\Phi^{-1}) \Phi(\cdot)} f(\cdot) \mathbf{1}\{\ell \nu(\Phi^{-1}) \Phi(\cdot) \in \mathsf{B}_K^c\}$$

also belongs to C and the consistency of the weighted sample $\{\xi_{N,i}, \omega_{N,i}\}_{i=1}^{M_N}$ therefore implies

$$\tilde{M}_N^{-1} \sum_{i=1}^{M_N} \lfloor \ell\nu(\Phi^{-1})\Phi(\xi_{N,i})\rfloor f(\xi_{N,i}) \mathbf{1}\{\ell\nu(\Phi^{-1})\Phi(\xi_{N,i}) \in \mathsf{B}_K^c\}$$

$$= \left(\frac{\ell\nu(\Phi^{-1})}{\tilde{M}_N \Omega_N^{-1}}\right) \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} g_K(\xi_{N,i}) \xrightarrow{\mathrm{P}} \nu(g_K).$$

For any $K \geq 1$, it can be easily checked that if $b \in \mathsf{B}_K^c$ and $c \in [1 - 1/K^2, 1 + 1/K^2]$, then $\lfloor b \rfloor = \lfloor bc \rfloor$. By applying this relation for each $i \in \{1, \ldots, M_N\}$, with $b \stackrel{\text{def}}{=} \ell\nu(\Phi^{-1})\Phi(\xi_{N,i})$ and $c \stackrel{\text{def}}{=} \tilde{M}_N \Omega_N^{-1} / \ell\nu(\Phi^{-1})$ and using that $\tilde{M}_N \Omega_N^{-1} / \ell\nu(\Phi^{-1}) \xrightarrow{\mathrm{P}} 1$, we therefore obtain

$$\tilde{M}_N^{-1} \sum_{i=1}^{M_N} \lfloor \tilde{M}_N \Omega_N^{-1} \Phi(\xi_{N,i})\rfloor f(\xi_{N,i}) \mathbf{1}\{\ell\nu(\Phi^{-1})\Phi(\xi_{N,i}) \in \mathsf{B}_K^c\} \xrightarrow{\mathrm{P}} \nu(g_K).$$

The proof of (44) follows by letting $K \to \infty$. □

The condition $\nu\{\ell\nu(\Phi^{-1})\Phi \in \mathbb{N} \cup \{\infty\}\} = 0$ in Proposition 5 and Lemma A.5 is crucial. Assume that $\{\xi_{N,i}\}_{1 \leq i \leq N}$ is an i.i.d. μ-distributed sample where μ is the distribution on the set $\{1/2, 2\}$ given by $\mu(\{1/2\}) = 2/3$ and $\mu(\{2\}) = 1/3$. Let ν be the distribution on $\{1/2, 2\}$ given by $\nu(\{1/2\}) = 1/3$ and $\nu(\{2\}) = 2/3$. The weighted sample $\{(\xi_{N,i}, \xi_{N,i})\}_{1 \leq i \leq N}$ [i.e., where we have set $\Phi(\xi) = \xi$] is a consistent sample for ν: for any function $f \in \mathbb{B}(\{1/2, 2\}) \stackrel{\text{def}}{=} \{f : \{1/2, 2\} \to \mathbb{R}, |f(1/2)| < \infty \text{ and } |f(2)| < \infty\}$,

$$\begin{aligned} &\frac{\sum_{i=1}^{M_N} \xi_{N,i} f(\xi_{N,i})}{\sum_{i=1}^{M_N} \xi_{N,i}} \\ &\qquad \xrightarrow{\mathrm{P}} \frac{(1/2) f(1/2)\mu(1/2) + 2f(2)\mu(2)}{(1/2)\mu(1/2) + 2\mu(2)} \\ &\qquad = (1/2) f(1/2) + 1/3 f(2) = \nu(f). \end{aligned}$$

In this example, $\ell = 1$ and obviously $\nu(\Phi^{-1}) = 1$. Moreover,

$$\nu\{\Phi \in \{\infty\} \cup \mathbb{N}\} = \nu\{\{1/2, 2\} \cap \mathbb{N}\} = \nu(\{2\}) = 2/3 \neq 0.$$

We will show that the convergence in Lemma A.5 fails. More precisely, setting $f(\xi) = \xi$, we will show that

$$\begin{aligned} \frac{1}{M_N} \sum_{i=1}^{M_N} \left\lfloor M_N \frac{\omega_{N,i}}{\Omega_N} \right\rfloor f(\xi_{N,i}) &= \frac{1}{M_N} \sum_{i=1}^{M_N} \left\lfloor M_N \frac{\xi_{N,i}}{\sum_{j=1}^{M_N} \xi_{N,j}} \right\rfloor f(\xi_{N,i}) \\ &\xrightarrow{\mathcal{D}} 4/3 - (2Z)/3, \end{aligned} \tag{46}$$

where Z is a Bernoulli variable with parameter $1/2$. This would imply that $M_N^{-1}\sum_{i=1}^{M_N}\lfloor M_N \frac{\omega_{N,i}}{\Omega_N}\rfloor f(\xi_{N,i})$ does not converge in probability to a constant. The LLN and CLT for i.i.d. random variables imply that

$$M_N^{-1}\Omega_N = M_N^{-1}\sum_{i=1}^{M_N}\xi_{N,i} \xrightarrow{\mathrm{P}} 1,$$

$$\begin{bmatrix}\mathbf{1}\{M_N\Omega_N^{-1} < 1\}\\ \mathbf{1}\{M_N\Omega_N^{-1} \geq 1\}\end{bmatrix} = \begin{bmatrix}\mathbf{1}\{(M_N)^{1/2}(M_N^{-1}\Omega_N - 1) > 0\}\\ \mathbf{1}\{(M_N)^{1/2}(M_N^{-1}\Omega_N - 1) \leq 0\}\end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} Z \\ 1-Z\end{bmatrix},$$

where Z is a Bernoulli random variable with parameter $1/2$. Since $\omega_{N,i} = \Phi(\xi_{N,i}) = \xi_{N,i} \in \{1/2, 2\}$ and $f(\xi) = \xi$,

$$\begin{aligned}
&\mathbf{1}\left\{\frac{1}{2} < \Omega_N^{-1}M_N < \frac{3}{2}\right\}\frac{1}{M_N}\sum_{i=1}^{M_N}\lfloor M_N\Omega_N^{-1}\omega_{N,i}\rfloor f(\xi_{N,i})\\
&\quad = \mathbf{1}\left\{\frac{1}{2} < \Omega_N^{-1}M_N < \frac{3}{2}\right\}\frac{2}{M_N}\sum_{i=1}^{M_N}\lfloor 2\Omega_N^{-1}M_N\rfloor \mathbf{1}\{\xi_{N,i} = 2\}\\
&\quad = \mathbf{1}\left\{\frac{1}{2} < \frac{M_N}{\Omega_N} < 1\right\}\frac{2}{M_N}\sum_{i=1}^{M_N}\mathbf{1}\left\{\xi_{N,i} = 2\right\}\\
&\qquad + \mathbf{1}\left\{1 \leq \frac{M_N}{\Omega_N} < \frac{3}{2}\right\}\frac{4}{M_N}\sum_{i=1}^{M_N}\mathbf{1}\{\xi_{N,i} = 2\}\\
&\quad \xrightarrow{\mathcal{D}} (2Z)/3 + 4(1-Z)/3 = 4/3 - (2Z)/3.
\end{aligned}$$

The proof of (46) is concluded by noting that $M_N\Omega_N^{-1} \xrightarrow{\mathrm{P}} 1$.

APPENDIX E: PROOF OF THEOREMS 6, 7 AND 8

PROOF OF THEOREM 6. Let $f \in \mathsf{C}$. We set for $j = 1, \ldots, M_N$, $\mathcal{F}_{N,j} \stackrel{\text{def}}{=} \mathcal{F}_{N,0} \vee \sigma(\{G_{N,i}\}_{i=1}^{j})$, where $\mathcal{F}_{N,0}$ is defined in (6), and $U_{n,j} \stackrel{\text{def}}{=} \tilde{m}_N^{-1}G_{N,j}f(\xi_{N,j})$. Note that

$$\begin{aligned}
\sum_{i=1}^{M_N}\mathrm{E}[|U_{N,i}||\mathcal{F}_{N,i-1}] &= \tilde{m}_N^{-1}\sum_{i=1}^{M_N}\mathrm{E}[G_{N,i}|\mathcal{F}_{N,i-1}]|f(\xi_{N,i})|\\
&= \Omega_N^{-1}\sum_{i=1}^{M_N}\omega_{N,i}|f(\xi_{N,i})| \xrightarrow{\mathrm{P}} \nu(|f|) < \infty,
\end{aligned}$$

showing that the sequence $\{\sum_{i=1}^{M_N}\mathrm{E}[|U_{N,i}||\mathcal{F}_{N,i-1}]\}$ is tight. For any $\epsilon > 0$ and $C > 0$, using that $\{\tilde{m}_N^{-1}G_{N,i}|f(\xi_{N,i})| \geq \epsilon\} \subseteq \{\tilde{m}_N^{-1}G_{N,i} \geq \epsilon/C\} \cup \{|f(\xi_{N,i})| \geq$

$C\}$, we obtain $\sum_{i=1}^{M_N} \mathrm{E}[|U_{N,i}|\mathbf{1}\{|U_{N,i}| \geq \epsilon\}|\mathcal{F}_{N,i-1}] \leq A_N(\epsilon, C) + B_N(C)$, where

$$A_N(\epsilon, C) \stackrel{\text{def}}{=} \tilde{m}_N^{-1} \sum_{i=1}^{M_N} |f(\xi_{N,i})| \, \mathrm{E}[G_{N,i}\mathbf{1}\{\tilde{m}_N^{-1} G_{N,i} \geq \epsilon/C\}|\mathcal{F}_{N,i-1}]$$

$$B_N(C) \stackrel{\text{def}}{=} \tilde{m}_N^{-1} \sum_{i=1}^{M_N} \mathrm{E}[G_{N,i}|\mathcal{F}_{N,i-1}]|f(\xi_{N,i})|\mathbf{1}\{|f(\xi_{N,i})| \geq C\}.$$

Note that $\mathrm{E}[G_{N,i}\mathbf{1}\{\tilde{m}_N^{-1} G_{N,i} \geq \epsilon/C\}|\mathcal{F}_{N,i-1}] \leq (C/\epsilon)\tilde{m}_N^{-1}\,\mathrm{E}[G_{N,i}^2|\mathcal{F}_{N,i-1}]$. In addition, it can easily be checked that for the Poisson, binomial and Bernoulli branching,

$$\begin{aligned}\tilde{m}_N^{-1}\,\mathrm{E}[G_{N,i}^2|\mathcal{F}_{N,i-1}] &\leq \Omega_N^{-1}\omega_{N,i} + \tilde{m}_N(\Omega_N^{-1}\omega_{N,i})^2 \\ &\leq \Omega_N^{-1}\omega_{N,i}\Bigl(1 + \tilde{m}_N\Omega_N^{-1} \sup_{1\leq i\leq M_N} \omega_{N,i}\Bigr),\end{aligned}$$

which implies, using that $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ is consistent for (ν, C),

$$A_N(\epsilon, C) \leq (C/\epsilon)\Bigl(\tilde{m}_N^{-1} + \Omega_N^{-1} \sup_{1\leq i\leq M_N} \omega_{N,i}\Bigr) \sum_{i=1}^{M_N} \Omega_N^{-1}\omega_{N,i}|f(\xi_{N,i})| \stackrel{\mathrm{P}}{\longrightarrow} 0.$$

On the other hand, since $B_N(C) = \Omega_N^{-1}\sum_{i=1}^{M_N} \omega_{N,i}|f(\xi_{N,i})|\mathbf{1}\{|f(\xi_{N,i})| \geq C\}$, the consistency of the weighted sample $\{(\xi_{N,i}, \omega_{N,i})\}_{i=1}^{M_N}$ implies $B_N(C) \stackrel{\mathrm{P}}{\longrightarrow} \nu(|f|\mathbf{1}\{|f| \geq C\}$. Since C can be chosen arbitrarily large, the two previous relations show the negligibility, $\sum_{i=1}^{M_N} \mathrm{E}[|U_{N,i}|\mathbf{1}\{|U_{N,i}| \geq \epsilon\}|\mathcal{F}_{N,i-1}] \stackrel{\mathrm{P}}{\longrightarrow} 0$. Theorem A.1 therefore shows that, for any $f \in \mathsf{C}$, $\tilde{m}_N^{-1}\sum_{i=1}^{M_N} G_{N,i} f(\xi_{N,i}) \stackrel{\mathrm{P}}{\longrightarrow} \nu(f)$. Applying this relation to $f \equiv 1$ shows that $\tilde{M}_N/\tilde{m}_N \stackrel{\mathrm{P}}{\longrightarrow} 1$. □

PROOF OF THEOREM 7. We appeal to Theorem A.3 and, hence, need to check (31) and (32). Set $U_{N,i} = a_N\tilde{m}_N^{-1} G_{N,i} f(\xi_{N,i})$. Let $f \in \tilde{\mathsf{A}}$ such that $\nu(f) = 0$. Therefore,

$$\begin{aligned}&\sum_{i=1}^{M_N}\{\mathrm{E}[U_{N,i}^2|\mathcal{F}_{N,i-1}] - (\mathrm{E}[U_{N,i}|\mathcal{F}_{N,i-1}])^2\} \\ &\qquad = a_N^2\tilde{m}_N^{-2}\sum_{i=1}^{M_N} f^2(\xi_{N,i})\,\mathrm{Var}[G_{N,i}|\mathcal{F}_{N,i-1}].\end{aligned}$$

The conditional variance is given by $\mathrm{Var}[G_{N,i}|\mathcal{F}_{N,i-1}] = \tilde{m}_N\Omega_N^{-1}\omega_{N,i}$ for the Poisson branching and $\mathrm{Var}[G_{N,i}|\mathcal{F}_{N,i-1}] = \tilde{m}_N\Omega_N^{-1}\omega_{N,i}(1 - \Omega_N^{-1}\omega_{N,i})$ for the binomial branching. Note that $a_N^2\tilde{m}_N^{-1}\sum_{i=1}^{M_N} f^2(\xi_{N,i})\Omega_N^{-1}\omega_{N,i} \stackrel{\mathrm{P}}{\longrightarrow} \beta\ell^{-1}\nu(f^2)$ and, for $p \geq 2$, using that $\Omega_N^{-1} \sup_{1 \leq i \leq M_N} \omega_{N,i} \xrightarrow{\mathrm{P}} 0$. Therefore, in both cases,

$$a_N^2 \tilde{m}_N^{-2} \sum_{i=1}^{M_N} f^2(\xi_{N,i}) \operatorname{Var}[G_{N,i} | \mathcal{F}_{N,i-1}] \xrightarrow{\mathrm{P}} \beta \ell^{-1} \operatorname{Var}_\nu(f).$$

We now check the tightness condition (32). For any $\epsilon > 0$ and $C > 0$, using that $\{a_N \tilde{m}_N^{-1} G_{N,i} | f(\xi_{N,i})| \geq \epsilon\} \subseteq \{a_N \tilde{m}_N^{-1} G_{N,i} \geq \epsilon / C\} \cup \{|f(\xi_{N,i})| \geq C\}$, we obtain $\sum_{i=1}^{M_N} \mathrm{E}[|U_{N,i}|^2 \mathbf{1}\{|U_{N,i}| \geq \epsilon\} | \mathcal{F}_{N,i-1}] \leq A_N(\epsilon, C) + B_N(C)$, where

$$A_N(\epsilon, C) \stackrel{\text{def}}{=} a_N^2 \tilde{m}_N^{-2} \sum_{i=1}^{M_N} |f(\xi_{N,i})|^2 \mathrm{E}[G_{N,i}^2 \mathbf{1}\{a_N \tilde{m}_N^{-1} G_{N,i} \geq \epsilon / C\} | \mathcal{F}_{N,i-1}],$$

$$B_N(C) \stackrel{\text{def}}{=} a_N^2 \tilde{m}_N^{-2} \sum_{i=1}^{M_N} \mathrm{E}[G_{N,i}^2 | \mathcal{F}_{N,i-1}] |f(\xi_{N,i})|^2 \mathbf{1}\{|f(\xi_{N,i})| \geq C\}.$$

We first prove that for any $\epsilon > 0$ and $C > 0$, $A_N(\epsilon, C) \xrightarrow{\mathrm{P}} 0$. If G is either a Poisson or a binomial variable, then $\mathrm{E}[G^3] \leq \mathrm{E}[G] + 3(\mathrm{E}[G])^2 + (\mathrm{E}[G])^3$. Thus,

$$\begin{aligned}
A_N(\epsilon, C) &\leq a_N^3 \tilde{m}_N^{-3} (C/\epsilon) \sum_{i=1}^{M_N} \mathrm{E}[G_{N,i}^3 | \mathcal{F}_{N,i-1}] f^2(\xi_{N,i}) \\
&\leq a_N^3 \tilde{m}_N^{-3} (C/\epsilon) \sum_{i=1}^{M_N} [\tilde{m}_N \Omega_N^{-1} \omega_{N,i} \\
&\qquad\qquad + 3(\tilde{m}_N \Omega_N^{-1} \omega_{N,i})^2 + (\tilde{m}_N \Omega_N^{-1} \omega_{N,i})^3] f^2(\xi_{N,i}) \\
&\leq a_N^3 \tilde{m}_N^{-2} (C/\epsilon) \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f^2(\xi_{N,i}) \\
&\quad + \Bigl(3 a_N \tilde{m}_N^{-1} + a_N \Omega_N^{-1} \max_{1 \leq i \leq M_N} \omega_{N,i}\Bigr) (C/\epsilon) a_N^2 \Omega_N^{-2} \\
&\quad \times \sum_{i=1}^{M_N} \omega_{N,i}^2 f^2(\xi_{N,i}),
\end{aligned}$$

and the proof of $A_N(\epsilon, C) \xrightarrow{\mathrm{P}} 0$ follows upon noting that $a_N^3 \tilde{m}_N^{-2} = O_{\mathrm{P}}(a_N^{-1})$, $a_N \tilde{m}_N^{-1} = O_{\mathrm{P}}(a_N^{-1})$, $a_N \Omega_N^{-1} \max_{1 \leq i \leq M_N} \omega_{N,i} = o_{\mathrm{P}}(1)$ and

$$\Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f^2(\xi_{N,i}) \xrightarrow{\mathrm{P}} \nu(f^2) \quad \text{and} \quad a_N^2 \Omega_N^{-2} \sum_{i=1}^{M_N} \omega_{N,i}^2 f^2(\xi_{N,i}) \xrightarrow{\mathrm{P}} \gamma(f^2).$$

We finally consider $B_N(C)$. As above, if G is either a Poisson or a binomial variable, then $\mathrm{E}[G^2] \le \mathrm{E}[G] + (\mathrm{E}[G])^2$. Therefore,

$$B_N(C) \le a_N^2 \tilde{m}_N^{-1} \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f^2(\xi_{N,i}) \mathbf{1}\{|f(\xi_{N,i})| \ge C\} + a_N^2 \Omega_N^{-2} \sum_{i=1}^{M_N} \omega_{N,i}^2 f^2(\xi_{N,i}) \mathbf{1}\{|f(\xi_{N,i})| \ge C\}.$$

Since $a_N^2 \tilde{m}_N^{-1} = O_{\mathrm{P}}(1)$, $\Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} f^2(\xi_{N,i}) \mathbf{1}\{|f(\xi_{N,i})| \ge C\} \xrightarrow{\mathrm{P}} \nu(f^2 \mathbf{1}\{|f| \ge C\})$ and $a_N^2 \Omega_N^{-2} \sum_{i=1}^{M_N} \omega_{N,i}^2 f^2(\xi_{N,i}) \mathbf{1}\{|f(\xi_{N,i})| \ge C\} \xrightarrow{\mathrm{P}} \gamma(f^2 \mathbf{1}\{|f| \ge C\})$, the RHS can be made arbitrarily small by taking C sufficiently large. □

PROOF OF THEOREM 8. Let f in $\tilde{\mathsf{A}}$ such that $\nu(f) = 0$. Note that

$$a_n \tilde{M}_N^{-1} \sum_{i=1}^{\tilde{M}_N} f(\tilde{\xi}_{N,i}) = (\tilde{m}_N \tilde{M}_N^{-1}) a_n \tilde{m}_N^{-1} \sum_{i=1}^{M_N} G_{N,i} f(\xi_{N,i}),$$

where $G_{N,i}$ is defined by (17) with $\omega_{N,i} = \Phi(\xi_{N,i})$. We have that $\tilde{M}_N / \tilde{m}_N \xrightarrow{\mathrm{P}} 1$ by Theorem 6. To apply Theorem A.3, we just have to check (31) and (32), where $U_{N,i} = a_n \tilde{m}_N^{-1} G_{N,i} f(\xi_{N,i})$ and $\{\mathcal{F}_{N,k}\}$ defined by $\mathcal{F}_{N,0} = \sigma\{(\xi_{N,i})_{i=1}^{M_N}\}$ and for all $1 \le k \le M_N$, $\mathcal{F}_{N,k} = \mathcal{F}_{N,0} \vee \sigma\{(U_{N,i})_{1 \le i \le k}\}$. Lemma A.6 shows that

$$\begin{aligned} A_N &= \sum_{i=1}^{M_N} \{\mathrm{E}[U_{N,i}^2 | \mathcal{F}_{N,i-1}] - (\mathrm{E}[U_{N,i} | \mathcal{F}_{N,i-1}])^2\} \\ &= a_N^2 \tilde{m}_N^{-2} \left\{ \sum_{i=1}^{M_N} (\langle \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rangle - \langle \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rangle^2) f^2(\xi_{N,i}) \right\} \qquad (47) \\ &\xrightarrow{\mathrm{P}} \beta \ell^{-1} \nu \left(\frac{\langle \ell \nu(\Phi^{-1}) \Phi \rangle (1 - \langle \ell \nu(\Phi^{-1}) \Phi \rangle)}{\ell \nu(\Phi^{-1}) \Phi} (f - \nu f)^2 \right). \end{aligned}$$

It remains to check (2) and (3). By Theorem 3, the weighted sample $\{(\tilde{\xi}_{N,i}, 1)\}_{i=1}^{\tilde{M}_N}$ is consistent for (ν, C), which implies

$$a_N^2 \tilde{M}_N^{-2} \sum_{i=1}^{\tilde{M}_N} f(\tilde{\xi}_{N,i}) \xrightarrow{\mathrm{P}} \beta \nu(f),$$

and, thus, (2) is satisfied. (3) is trivially satisfied. □

LEMMA A.6. *Under the assumptions of Theorem 8, for any function f such that $\Phi f \in \mathsf{C}$,*

$$\Omega_N^{-2} a_N^2 \sum_{i=1}^{M_N} \omega_{N,i}^2 f(\xi_{N,i}) \xrightarrow{\mathrm{P}} \beta \nu(\Phi^{-1}) \nu(\Phi f), \tag{48}$$

$$\tilde{m}_N^{-2} a_N^2 \sum_{i=1}^{M_N} \langle \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rangle^q f(\xi_{N,i}) \xrightarrow{\mathrm{P}} \beta \ell^{-1} \nu\left(f \frac{\langle \ell \nu(\Phi^{-1}) \Phi \rangle^q}{\ell \nu(\Phi^{-1}) \Phi} \right), \quad q = 1, 2. \tag{49}$$

PROOF. We first consider (48). Note that

$$\Omega_N^{-2} a_N^2 \sum_{i=1}^{M_N} \omega_{N,i}^2 f(\xi_{N,i}) = (M_N \Omega_N^{-1})(a_N^2 M_N^{-1}) \Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} \Phi(\xi_{N,i}) f(\xi_{N,i}),$$

and the proof follows since $M_N \Omega_N^{-1} \xrightarrow{\mathrm{P}} \nu(\Phi^{-1})$, $a_N^2 M_N^{-1} \xrightarrow{\mathrm{P}} \beta$, and

$$\Omega_N^{-1} \sum_{i=1}^{M_N} \omega_{N,i} \Phi(\xi_{N,i}) f(\xi_{N,i}) \xrightarrow{\mathrm{P}} \nu(\Phi f).$$

The proof of (49) with $q = 1$ can be done along the same lines as in Lemma A.5. To prove (49) with $q = 2$, some adaptations are required. We define by W the set of functions f such that $\Phi f \in \mathsf{C}$ and, for $f \in \mathsf{W}$, we set $\gamma(f) = \beta \nu(\Phi^{-1}) \nu(\Phi f)$. Since $\langle \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rangle \le \tilde{m}_N \Omega_N^{-1} \omega_{N,i}$, (48) shows that, for any $f \in \mathsf{W}$,

$$\begin{aligned} &a_N^2 \tilde{m}_N^{-2} \sum_{i=1}^{M_N} \langle \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rangle^2 |f(\xi_{N,i})| \mathbf{1}\{\ell \nu(\Phi^{-1}) \omega_{N,i} \in \mathsf{B}_K\} \\ &\quad \le a_N^2 \Omega_N^{-2} \sum_{i=1}^{M_N} \omega_{N,i}^2 |f(\xi_{N,i})| \mathbf{1}\{\ell \nu(\Phi^{-1}) \Phi(\xi_{N,i}) \in \mathsf{B}_K\} \\ &\quad \xrightarrow{\mathrm{P}} \gamma\big(|f| \mathbf{1}\{\ell \nu(\Phi^{-1}) \Phi \in \mathsf{B}_K\}\big), \end{aligned}$$

where B_K is defined in (45). We will now prove that, for any $f \in \mathsf{W}$,

$$\begin{aligned} &a_N^2 \tilde{m}_N^{-2} \sum_{i=1}^{M_N} \langle \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rangle^2 g_K(\xi_{N,i}) \\ &\quad \xrightarrow{\mathrm{P}} \gamma\left(\frac{\langle \ell \nu(\Phi^{-1}) \Phi \rangle^2}{(\ell \nu(\Phi^{-1}) \Phi)^2} g_K \right), \end{aligned} \tag{50}$$

where $g_K \stackrel{\text{def}}{=} f\mathbf{1}\{\ell\nu(\Phi^{-1})\Phi \in \mathsf{B}_K^c\}$. For that purpose, we establish that, for $p = 0, 1, 2$ and $f \in \mathsf{W}$,

$$\begin{aligned} A_N &\stackrel{\text{def}}{=} a_N^2 \tilde{m}_N^{-2} \sum_{i=1}^{M_N} (\tilde{m}_N \Omega_N^{-1} \omega_{N,i})^p \lfloor \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rfloor^{2-p} g_K(\xi_{N,i}) \\ &\xrightarrow{\text{P}} \gamma\biggl(\frac{\lfloor \ell\nu(\Phi^{-1})\Phi \rfloor^{2-p}}{(\ell\nu(\Phi^{-1})\Phi)^{2-p}} g_K \biggr). \end{aligned}$$

We may write

$$A_N = (\tilde{m}_N^{-1} \Omega_N \ell\nu(\Phi^{-1}))^{2-p} a_N^2 \Omega_N^{-2} \sum_{i=1}^{M_N} \biggl(\frac{\lfloor \tilde{m}_N \Omega_N^{-1} \Phi(\xi_{N,i}) \rfloor}{\ell\nu(\Phi^{-1})\Phi(\xi_{N,i})} \biggr)^{2-p} \omega_{N,i}^2 g_K(\xi_{N,i}).$$

Since W is a proper set, for any $f \in \mathsf{W}$, $(\frac{\lfloor \ell\nu(\Phi^{-1})\Phi \rfloor}{\ell\nu(\Phi^{-1})\Phi})^{2-p} g_K \in \mathsf{W}$,

$$\begin{aligned} a_N^2 \Omega_N^{-2} \sum_{i=1}^{M_N} \biggl(\frac{\lfloor \ell\nu(\Phi^{-1})\Phi(\xi_{N,i}) \rfloor}{\ell\nu(\Phi^{-1})\Phi(\xi_{N,i})} \biggr)^{2-p} \omega_{N,i}^2 g_K(\xi_{N,i}) \\ \xrightarrow{\text{P}} \gamma\biggl(\frac{\lfloor \ell\nu(\Phi^{-1})\Phi \rfloor^{2-p}}{(\ell\nu(\Phi^{-1})\Phi)^{2-p}} g_K \biggr). \end{aligned}$$

The proof of (49) follows since $\tilde{m}_N^{-1} \Omega_N \ell\nu(\Phi^{-1}) \xrightarrow{\text{P}} 1$ and $\lfloor \tilde{m}_N \Omega_N^{-1} \omega_{N,i} \rfloor = \lfloor \ell\nu(\Phi^{-1}) \omega_{N,i} \rfloor$ on the event $\{|\tilde{m}_N \Omega_N^{-1} / \ell\nu(\Phi^{-1}) - 1| \leq 1/K^2, \ell\nu(\Phi^{-1})\omega_{N,i} \in \mathsf{B}_K^c\}$. $\square$

REFERENCES

[1] DOUCET, A., DE FREITAS, N. and GORDON, N., eds. (2001). *Sequential Monte Carlo Methods in Practice*. Springer, New York. MR1847784

[2] LIU, J. (2001). *Monte Carlo Strategies in Scientific Computing*. Springer, New York. MR1842342

[3] KÜNSCH, H. R. (2001). State space and hidden Markov models. In *Complex Stochastic Systems* (O. E. Barndorff-Nielsen, D. R. Cox and C. Klueppelberg, eds.) 109–173. CRC Publisher, Boca Raton, FL. MR1893412

[4] GILKS, W. R. and BERZUINI, C. (2001). Following a moving target—Monte Carlo inference for dynamic Bayesian models. *J. Roy. Statist. Soc. Ser. B* **63** 127–146. MR1811995

[5] KÜNSCH, H. R. (2005). Recursive Monte Carlo filters: Algorithms and theoretical analysis. *Ann. Statist.* **33** 1983–2021. MR2211077

[6] DEL MORAL, P. (2004). *Feynman–Kac Formulae. Genealogical and Interacting Particle Systems with Applications*. Springer, Berlin. MR2044973

[7] CHOPIN, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to bayesian inference. *Ann. Statist.* **32** 2385–2411. MR2153989

[8] HANDSCHIN, J. and MAYNE, D. (1969). Monte Carlo techniques to estimate the conditional expectation in multi-stage non-linear filtering. *Internat. J. Control* **9** 547–559. MR0246490

[9] RUBIN, D. B. (1987). A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when the fraction of missing information is modest: The SIR algorithm (discussion of Tanner and Wong). *J. Amer. Statist. Assoc.* **82** 543–546.

[10] LANDAU, D. P. and BINDER, K. (2000). *A guide to Monte Carlo Simulations in Statistical Physics*. Cambridge Univ. Press. MR1781083

[11] LIU, J. and CHEN, R. (1998). Sequential Monte Carlo methods for dynamic systems. *J. Amer. Statist. Assoc.* **93** 1032–1044. MR1649198

[12] RISTIC, B., ARULAMPALAM, M. and GORDON, A. (2004). *Beyond Kalman Filters: Particle Filters for Target Tracking*. Artech House.

[13] CAPPÉ, O., GUILLIN, A., MARIN, J. M. and ROBERT, C. P. (2004). Population Monte Carlo. *J. Comput. Graph. Statist.* **13** 907–929. MR2109057

[14] DEL MORAL, P. (1996). Nonlinear filtering: Interacting particle solution. *Markov Process. Related Fields* **2** 555–579. MR1431187

[15] DEL MORAL, P. and GUIONNET, A. (1999). Central limit theorem for nonlinear filtering and interacting particle systems. *Ann. Appl. Probab.* **9** 275–297. MR1687359

[16] DEL MORAL, P. and MICLO, L. (2000). Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to nonlinear filtering. *Séminaire de Probabilités XXXIV*. *Lecture Notes in Math.* **1729** 1–145. Springer, Berlin. MR1768060

[17] CRISAN, D. and LYONS, T. (1997). Nonlinear filtering and measure-valued processes. *Probab. Theory Related Fields* **109** 217–244. MR1477650

[18] CRISAN, D. and DOUCET, A. (2002). A survey of convergence results on particle filtering methods for practitioners. *IEEE Trans. Signal Process.* **50** 736–746. MR1895071

[19] BERZUINI, C. and GILKS, W. R. (2001). Resample-move filtering with cross-model jumps. In *Sequential Monte Carlo Methods in Practice* (A. Doucet, N. De Freitas and N. Gordon, eds.) 117–138. Springer, Berlin. MR1847789

[20] PITT, M. K. and SHEPHARD, N. (1999). Filtering via simulation: Auxiliary particle filters. *J. Amer. Statist. Assoc.* **94** 590–599. MR1702328

[21] LIU, J. and CHEN, R. (1995). Blind deconvolution via sequential imputations. *J. Amer. Statist. Assoc.* **90** 567–576.

[22] CRISAN, D., GAINES, J. and LYONS, T. (1998). Convergence of a branching particle method to the solution of the Zakai equation. *SIAM J. Appl. Math.* **58** 1568–1590 (electronic). MR1637870

[23] CRISAN, D. (2003). Exact rates of convergence for a branching particle approximation to the solution of the Zakai equation. *Ann. Probab.* **31** 693–718. MR1964946

[24] LIU, J., CHEN, R. and LOGVINENKO, T. (2001). A theoretical framework for sequential importance sampling and resampling. In *Sequential Monte Carlo Methods in Practice* (A. Doucet, N. De Freitas and N. Gordon, eds.) 225–246. Springer, Berlin. MR1847794

[25] KONG, A., LIU, J. S. and WONG, W. (1994). Sequential imputation and Bayesian missing data problems. *J. Amer. Statist. Assoc.* **89** 278–288, 590–599.

[26] SHIRYAEV, A. N. (1996). *Probability*, 2nd ed. Springer, New York. MR1368405

[27] DVORETZKY, A. (1972). Asymptotic normality for sums of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Probab. II* (*Univ. California, Berkeley, Calif., 1970/1971*). *Probability Theory* 513–535. Univ. California Press, Berkeley. MR0415728

CMAP
ECOLE POLYTECHNIQUE
91128 PALAISEAU CEDEX
FRANCE
E-MAIL: douc@cmapx.polytechnique.fr

GET/TÉLÉCOM PARIS
PARIS
FRANCE
E-MAIL: moulines@tsi.enst.fr