A Sequential Particle Algorithm that Keeps the Particle System Alive

François LeGland[1] and Nadia Oudjane[2,3]

[1] IRISA / INRIA, Campus de Beaulieu, 35042 RENNES Cédex, France, legland@irisa.fr
[2] EDF, Division R&D, 92141 CLAMART Cédex, France, nadia.oudjane@edf.fr
[3] Département de Mathématiques, Université Paris XIII, 93430 VILLETANEUSE, France, n_oudjane@math.univ-paris13.fr

Summary. A sequential particle algorithm proposed by Oudjane (2000) is studied here, which uses an adaptive random number of particles at each generation and guarantees that the particle system never dies out. This algorithm is especially useful for approximating a nonlinear (normalized) Feynman–Kac flow, in the special case where the selection functions can take the zero value, e.g. in the simulation of a rare event using an *importance splitting* approach. Among other results, a central limit theorem is proved by induction, based on the result of Rényi (1957) for sums of a random number of independent random variables. An alternate proof is also given, based on an original central limit theorem for triangular arrays of martingale increments spread across generations with different random sizes.

1 Introduction

The problem considered here is the particle approximation of the linear (unnormalized) flow and of the associated nonlinear (normalized) flow, defined by

$$\langle \gamma_n, \phi \rangle = \mathbb{E}[\phi(X_n) \prod_{k=0}^{n} g_k(X_k)\,] \qquad \text{and} \qquad \langle \mu_n, \phi \rangle = \frac{\langle \gamma_n, \phi \rangle}{\langle \gamma_n, 1 \rangle}\ ,$$

for any bounded measurable function ϕ, where $\{X_k\,,\ k = 0, 1, \cdots, n\}$ is a Markov chain with initial probability distribution η_0 and transition kernels $\{Q_k\,,\ k = 1, \cdots, n\}$, and where $\{g_k\,,\ k = 0, 1, \cdots, n\}$ are given bounded measurable nonnegative functions, known as selection or fitness functions. This problem has been widely studied [3] and the focus here is on the special case where the selection functions can possibly take the zero value. Clearly

$$\gamma_k = g_k\ (\gamma_{k-1}\, Q_k) = g_k\ (\mu_{k-1}\, Q_k)\ \langle \gamma_{k-1}, 1 \rangle \qquad \text{and} \qquad \gamma_0 = g_0\ \eta_0\ , \tag{1}$$

or equivalently $\gamma_k = \gamma_{k-1}\, R_k$ where the nonnegative (unnormalized) kernel R_k is defined by $R_k(x, dx') = Q_k(x, dx')\, g_k(x')$, hence

H.A.P. Blom, J. Lygeros (Eds.): Stochastic Hybrid Systems, LNCIS 337, pp. 351–389, 2006.

$$\langle \gamma_k, 1 \rangle = \langle \mu_{k-1} Q_k, g_k \rangle \langle \gamma_{k-1}, 1 \rangle \qquad \text{and} \qquad \langle \gamma_0, 1 \rangle = \langle \eta_0, g_0 \rangle \ , \tag{2}$$

and the minimal assumption made throughout this paper is that $\langle \gamma_n, 1 \rangle > 0$, or equivalently that $\langle \eta_0, g_0 \rangle > 0$ and $\langle \mu_{k-1} Q_k, g_k \rangle > 0$ for any $k = 1, \cdots, n$, otherwise the problem is not well defined. There are many practical situations where the selection functions can possibly take the zero value

- simulation of a rare event using an *importance splitting* approach [3, Section 12.2], [1, 6],
- simulation of a Markov chain conditionned or constrained to visit a given sequence of subspaces of the state space (this includes tracking a mobile in the presence of obstacles : when the mobile is hidden behind an obstacle, occlusion occurs and no observation is available at all, however this information can still be used, with a selection function equal to the indicator function of the region hidden by the obstacle),
- simulation of a r.v. in the tail of a given probability distribution,
- nonlinear filtering with bounded observation noise,
- implementation of a *robustification* approach in nonlinear filtering, using truncation of the likelihood function [10, 16],
- algorithms of approximate nonlinear filtering, where hidden state and observation are simulated jointly, and where the simulated observation is validated against the actual observation [4, 5, 18, 19], e.g. when there is no explicit expression available for the likelihood function, or when a likelihood function does not even exist (nonadditive observation noise, noise–free observations, etc.).

This work has been announced in [12], and it is organized as follows. In Section 2, the (usual) nonsequential particle algorithm is presented, and the potential difficulty that arises if the selection functions can possibly take the zero value, i.e. the possible extinction of the particle system, is addressed. Refined $\mathbb{L}^1$ error estimates are stated in Theorem 1, for the purpose of comparison with the sequential particle algorithm, and the central limit theorem proved in [3, Section 9.4] is recalled. In Section 3, the sequential particle algorithm already proposed in [15, 11] is introduced, which uses an adaptive random number of particles at each generation and automatically keeps the particle system alive, i.e. which ensures its non–extinction. The main contributions of this work are $\mathbb{L}^1$ error estimates stated in Theorem 3 and a central limit theorem stated in Theorem 4. An interesting feature of the sequential particle algorithm is that a fixed performance can be guaranteed in advance, at the expense of a random computational effort : this could be seen as an adaptive rule to automatically choose the number of particles. To get a fair comparison of the nonsequential and sequential particle algorithms, the time–averaged random number of simulated particles, which is an indicator of how much computational effort has been used, is used as a normalizing factor. The different behaviour of the two particle algorithms is illustrated on the simple example of binary selection functions (taking only the value 0 or 1). The proof

of Theorem 4 relies on results stated in Section 4 for sums of a random number of i.i.d. random variables, especially when this random number is a stopping time. A conditional version of the central limit theorem, known in sequential analysis as the Anscombe theorem and proved in [17], is stated in Theorem 6, and a central limit theorem for triangular arrays of martingale increments spread across generations with different random sizes, is stated in Theorem 7.

The remaining part of this work is devoted to proofs of the main results. The central limit theorem for the sequential particle algorithm, stated in Theorem 4, is proved in Section 5 by induction, based on the central limit theorem stated in Theorem 6 for sums of a random number of i.i.d. random variables, and an alternate proof is given in Section 6, based on the central limit theorem stated in Theorem 7 for triangular arrays of martingale increments spread across generations with different random sizes. Finally, Theorems 6 and 7 are proved in Appendices A and B respectively. Further details, including the proofs of Theorems 1 and 3 can be found in [13, Sections 5 and 6].

2 Nonsequential Particle Algorithm

The evolution of the normalized (nonlinear) flow $\{\mu_k\,,\, k = 0, 1, \cdots, n\}$ is described by the following diagram

$$\mu_{k-1} \longrightarrow \eta_k = \mu_{k-1}\, Q_k \longrightarrow \mu_k = g_k \cdot \eta_k \ ,$$

with initial condition $\mu_0 = g_0 \cdot \eta_0$, where the notation $\cdot$ denotes the projective product. It follows from (2) and from the definition that

$$\mathbb{E}[\prod_{k=0}^{n} g_k(X_k)\,] = \langle \gamma_n, 1\rangle = \prod_{k=0}^{n} \langle \eta_k, g_k\rangle \ ,$$

i.e. the expectation of a product is replaced by the product of expectations. Notice that the ratio

$$\rho_k = \frac{\sup\limits_{x\in E} g_k(x)}{\langle \eta_k, g_k\rangle}$$

is an indicator of how difficult a given problem is : indeed, a large value of ρ_k means that regions where the selection function g_k is large have a small probability under η_k. The idea behind the particle approach is to look for an approximation

$$\mu_k \approx \mu_k^N = \sum_{i=1}^{N} w_k^i\, \delta_{\xi_k^i} \ ,$$

in the form of the weighted empirical probability distribution associated with the particle system $(\xi_k^i, w_k^i\,,\, i = 1, \cdots, N)$, where N denotes the number of particles. The weights and positions of the particles are chosen is such a way

that the evolution of the approximate sequence $\{\mu_k^N\ ,\ k = 0, 1, \cdots, n\}$ is described by the following diagram

$$\mu_{k-1}^N \longrightarrow \eta_k^N = S^N(\mu_{k-1}^N\, Q_k) \longrightarrow \mu_k^N = g_k \cdot \eta_k^N\ ,$$

with initial condition defined by $\mu_0^N = g_0 \cdot \eta_0^N$ and $\eta_0^N = S^N(\eta_0)$, where the notation $S^N(\mu)$ denotes the empirical probability distribution associated with an N–sample with common probability distribution μ. In practice, particles

- are selected according to their respective weights $(w_{k-1}^i\ ,\ i = 1, \cdots, N)$ (selection step),
- move according to the Markov kernel Q_k (mutation step),
- are weighted by evaluating the fitness function g_k (weighting step).

Starting from (1) and introducing the particle approximation

$$\gamma_k^N = g_k\ S^N(\mu_{k-1}^N\, Q_k)\ \langle \gamma_{k-1}^N, 1\rangle = g_k\, \eta_k^N\ \langle \gamma_{k-1}^N, 1\rangle\ ,$$

and

$$\gamma_0^N = g_0\ S^N(\eta_0) = g_0\, \eta_0^N\ ,$$

for the unnormalized (linear) flow, it is easily seen that

$$\langle \gamma_k^N, 1\rangle = \langle \eta_k^N, g_k\rangle\ \langle \gamma_{k-1}^N, 1\rangle \qquad \text{and} \qquad \langle \gamma_0^N, 1\rangle = \langle \eta_0^N, g_0\rangle\ , \tag{3}$$

hence

$$\frac{\gamma_k^N}{\langle \gamma_k^N, 1\rangle} = g_k \cdot \eta_k^N = \mu_k^N \qquad \text{and} \qquad \frac{\gamma_0^N}{\langle \gamma_0^N, 1\rangle} = g_0 \cdot \eta_0^N = \mu_0^N\ .$$

However, if the function g_k can possibly take the zero value, and even if $\langle \eta_k, g_k\rangle > 0$, it can happen that $\langle \eta_k^N, g_k\rangle = 0$, i.e. it can happen that the evaluation of the function g_k returns the zero value for all the particles generated at the end of the mutation step : in such a situation, the particle systems dies out and the algorithm cannot continue. A reinitialization procedure has been proposed and studied in [5], in which the particle system is generated afresh from an arbitrary restarting probability distribution ν. Alternatively, one could be interested by the behavior of the algorithm until the extinction time of the particle system, defined by

$$\tau^N = \inf\{k \geq 0\ :\ \langle \eta_k^N, g_k\rangle = 0\}\ .$$

Under the assumption that $\langle \gamma_n, 1\rangle > 0$, the probability $\mathbb{P}[\tau^N \leq n]$ that the algorithm cannot continue up to the time instant n goes to zero with exponential rate [3, Theorem 7.4.1].

Example 1 (Binary selection). In the special case of binary selection functions (taking only the value 0 or 1), i.e. indicator functions $g_k = 1_{A_k}$ of Borel subsets for any $k = 0, 1, \cdots, n$, it holds

$$p_0 = \langle \eta_0, g_0 \rangle = \mathbb{P}[X_0 \in A_0] \ ,$$

and

$$p_k = \langle \eta_k, g_k \rangle = \mathbb{P}[X_k \in A_k \mid X_0 \in A_0, \cdots, X_{k-1} \in A_{k-1}] \ ,$$

for any $k = 1, \cdots, n$, and it follows from (2) that

$$P_n = \mathbb{P}[X_0 \in A_0, \cdots, X_n \in A_n] = \langle \gamma_n, 1 \rangle = \prod_{k=0}^{n} \langle \eta_k, g_k \rangle \ .$$

On the *good set* $\{\tau^N > n\}$, the nonsequential particle algorithm results in the following approximations

$$p_k \approx p_k^N = \langle \eta_k^N, g_k \rangle = \frac{|I_k^N|}{N} \qquad \text{where} \qquad I_k^N = \{i = 1, \cdots, N \ : \ \xi_k^i \in A_k\} \ ,$$

denotes the set of successful particles within an N–sample with common probability distribution η_0 (for $k = 0$) and $\mu_{k-1}^N Q_k$ (for $k = 1, \cdots, n$), and it follows from (3) that

$$P_n \approx P_n^N = \langle \gamma_n^N, 1 \rangle = \prod_{k=0}^{n} \langle \eta_k^N, g_k \rangle = \prod_{k=0}^{n} \frac{|I_k^N|}{N} \ .$$

In other words, the probability P_n of a successful sequence is approximated as the product of the fraction of successful particles at each generation, and each transition probability p_k separately is approximated as the fraction of successful particles at the corresponding generation. Notice that the computational effort, i.e. the number N of simulated particles at each generation, is *fixed* in advance, whereas the number $|I_k^N|$ of successful particles at the k–th generation is *random*, and could even be zero.

The following results have been obtained for the nonsequential particle algorithm with a constant number N of particles : a nonasymptotic estimate [3, Theorem 7.4.3]

$$\sup_{\phi \,:\, \|\phi\| = 1} \mathbb{E} |\, 1_{\{\tau^N > n\}} \, \langle \mu_n^N, \phi \rangle - \langle \mu_n, \phi \rangle \,| \leq \frac{c_n^0}{\sqrt{N}} + \mathbb{P}[\tau^N \leq n] \ ,$$

and a central limit theorem (see [3, Section 9.4] for a slightly different algorithm)

$$\sqrt{N} \, [\, 1_{\{\tau^N > n\}} \, \langle \mu_n^N, \phi \rangle - \langle \mu_n, \phi \rangle \,] \Longrightarrow \mathcal{N}(0, v_n^0(\phi)) \ ,$$

in distribution as $N \uparrow \infty$, with an explicit expression for the asymptotic variance. In the simple case where the fitness functions are positive, i.e. cannot take the zero value, these results are well–known and can be found in [7, Proposition 2.9, Corollary 2.20], where the proof relies on a central limit theorem for triangular arrays of martingale increments, or in [9, Theorem 4], where the same central limit theorem is obtained by induction.

For the purpose of comparison with the sequential particle algorithm, the following nonasymptotic error estimates are proved in [13, Section 5].

Theorem 1. *With the extinction time τ^N defined by*

$$\tau^N = \inf\{k \geq 0 \,:\, \langle \eta_k^N, g_k \rangle = 0\} \ ,$$

it holds

$$\mathbb{E}\,|\,1_{\{\tau^N > n\}} \, \frac{\langle \gamma_n^N, 1\rangle}{\langle \gamma_n, 1\rangle} - 1\,| \leq z_n^N + \mathbb{P}[\tau^N \leq n] \ , \tag{4}$$

and

$$\sup_{\phi\,:\,\|\phi\|=1} \mathbb{E}|\,1_{\{\tau^N > n\}} \, \langle \mu_n^N, \phi\rangle - \langle \mu_n, \phi\rangle\,| \leq 2\, z_n^N + \mathbb{P}[\tau^N \leq n] \ , \tag{5}$$

where the sequence $\{z_k^N \,,\, k = 0, 1, \cdots, n\}$ satisfies the linear recursion

$$z_k^N \leq \rho_k \, (1 + \frac{\sqrt{\rho_k}}{\sqrt{N}}) \, z_{k-1}^N + \frac{\sqrt{\rho_k}}{\sqrt{N}} \qquad \textit{and} \qquad z_0^N \leq \frac{\sqrt{\rho_0}}{\sqrt{N}} \ . \tag{6}$$

Remark 1. The forcing term in (6) is $\sqrt{\rho_k}/\sqrt{N}$, and

$$\limsup_{N\uparrow\infty} \, [\sqrt{N}\, z_k^N] \leq \rho_k \, \limsup_{N\uparrow\infty} \, [\sqrt{N}\, z_{k-1}^N] + \sqrt{\rho_k} \ ,$$

and

$$\limsup_{N\uparrow\infty} \, [\sqrt{N}\, z_0^N] \leq \sqrt{\rho_0} \ .$$

Notice that with a fixed number N of simulated particles, the performance is $\sqrt{\rho_k}/\sqrt{N}$ and depends on ρ_k : as a result, it is not possible to guarantee in advance a fixed performance, since ρ_k is not known.

For completeness, the central limit theorem obtained in [3, Section 9.4] for a slightly different algorithm is recalled below.

Theorem 2 (Del Moral). *With the extinction time τ^N defined by*

$$\tau^N = \inf\{k \geq 0 \,:\, \langle \eta_k^N, g_k \rangle = 0\} \ ,$$

it holds

$$\sqrt{N}\,[\,1_{\{\tau^N > n\}} \, \frac{\langle \gamma_n^N, 1\rangle}{\langle \gamma_n, 1\rangle} - 1\,] \Longrightarrow \mathcal{N}(0, V_n^0) \ ,$$

and

$$\sqrt{N}\,[\,1_{\{\tau^N > n\}} \, \langle \mu_n^N, \phi\rangle - \langle \mu_n, \phi\rangle\,] \Longrightarrow \mathcal{N}(0, v_n^0(\phi)) \ ,$$

in distribution as $N \uparrow \infty$, for any bounded measurable function ϕ, with the asymptotic variance

$$V_n^0 = \sum_{k=0}^{n} \frac{\mathrm{var}(g_k \, R_{k+1:n}\, 1, \eta_k)}{\langle \eta_k, g_k \, R_{k+1:n}\, 1\rangle^2} \ ,$$

and

$$v_n^0(\phi) = \sum_{k=0}^{n} \frac{\mathrm{var}(g_k\, R_{k+1:n}\,(\phi - \langle \mu_n, \phi\rangle), \eta_k)}{\langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle^2} \ ,$$

respectively, where

$$R_{k+1:n}\,\phi(x) = R_{k+1}\cdots R_n\,\phi(x) = \mathbb{E}[\phi(X_n) \prod_{p=k+1}^{n} g_p(X_p) \mid X_k = x] \ ,$$

for any $k = 0, 1, \cdots, n$, *with the convention* $R_{n+1:n}\,\phi(x) = \phi(x)$, *for any* $x \in E$.

Remark 2. Notice that

$$\langle \eta_0, g_0\, R_{1:n}\,(\phi - \langle \mu_n, \phi\rangle)\,\rangle = \langle \gamma_0\, R_{1:n}, \phi - \langle \mu_n, \phi\rangle\,\rangle = \langle \gamma_n, \phi - \langle \mu_n, \phi\rangle\,\rangle = 0 \ ,$$

and

$$\langle \eta_k, g_k\, R_{k+1:n}\,(\phi - \langle \mu_n, \phi\rangle)\,\rangle = \langle \mu_{k-1}\, R_{k:n}, \phi - \langle \mu_n, \phi\rangle\,\rangle = \frac{\langle \gamma_n, \phi - \langle \mu_n, \phi\rangle\,\rangle}{\langle \gamma_{k-1}, 1\rangle} = 0 \ ,$$

for any $k = 1, \cdots, n$, hence the following equivalent expression holds for the asymptotic variance

$$v_n^0(\phi) = \sum_{k=0}^{n} \frac{\langle \eta_k, |g_k\, R_{k+1:n}\,[\phi - \langle \mu_n, \phi\rangle]\,|^2\,\rangle}{\langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle^2} \ .$$

Example 2 (Binary selection). In the special case of binary selection functions, it holds

$$R_{k+1:n}\, 1(x) = \mathbb{P}[X_{k+1} \in A_{k+1}, \cdots, X_n \in A_n \mid X_k = x] \ ,$$

for any $k = 0, 1, \cdots, n$, with the convention $R_{n+1:n}\, 1(x) = 1$, for any $x \in E$, and it follows from Theorem 2 that

$$\sqrt{N}\,[\,1_{\{\tau^N > n\}}\, \frac{P_n^N}{P_n} - 1] \Longrightarrow \mathcal{N}(0, V_n^0) \ ,$$

in distribution as $N \uparrow \infty$, with the asymptotic variance

$$V_n^0 = \sum_{k=0}^{n} (\frac{1}{p_k} - 1) + \sum_{k=0}^{n} \frac{1}{p_k}\, \frac{\mathrm{var}(R_{k+1:n}\, 1, \mu_k)}{\langle \mu_k, R_{k+1:n}\, 1\rangle^2} \ .$$

Indeed, since $g_k^2 = g_k$, it holds

$$\begin{aligned}
\frac{\mathrm{var}(g_k\, R_{k+1:n}\, 1, \eta_k)}{\langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle^2} &= \frac{\langle \eta_k, g_k\, |R_{k+1:n}\, 1|^2\rangle}{\langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle^2} - 1 \\
&= \frac{1}{p_k}\, \frac{\langle \mu_k, |R_{k+1:n}\, 1|^2\rangle}{\langle \mu_k, R_{k+1:n}\, 1\rangle^2} - 1 \\
&= (\frac{1}{p_k} - 1) + \frac{1}{p_k}\, [\, \frac{\langle \mu_k, |R_{k+1:n}\, 1|^2\rangle}{\langle \mu_k, R_{k+1:n}\, 1\rangle^2} - 1]\ ,
\end{aligned}$$

for any $k = 0, 1, \cdots, n$.

3 Sequential Particle Algorithm

The purpose of this work is to study a sequential particle algorithm, already proposed in [15, 11], which automatically keeps the particle system alive, i.e. which ensures its non–extinction. For any level $H > 0$, and for any $k = 0, 1, \cdots, n$, define the random number of particles

$$N_k^H = \inf\{N \geq 1 \,:\, \sum_{i=1}^{N} g_k(\xi_k^i) \geq H \sup_{x \in E} g_k(x)\}\ ,$$

where the random variables $\xi_0^1, \cdots, \xi_0^i, \cdots$ are i.i.d. with common probability distribution η_0 (for $k = 0$), and where, conditionally w.r.t. the σ–algebra $\mathcal{H}_{k-1}^H$ generated by the particle system until the $(k-1)$–th generation, the random variables $\xi_k^1, \cdots, \xi_k^i, \cdots$ are i.i.d. with common probability distribution $\mu_{k-1}^H\, Q_k$ (for $k = 1, \cdots, n$). The particle approximation $\{\mu_k^H\ ,\ k = 0, 1, \cdots, n\}$ is now parameterized by the level $H > 0$, and its evolution is described by the following diagram

$$\mu_{k-1}^H \longrightarrow \eta_k^H = S^{N_k^H}(\mu_{k-1}^H\, Q_k) \longrightarrow \mu_k^H = g_k \cdot \eta_k^H\ ,$$

with initial condition defined by $\mu_0^H = g_0 \cdot \eta_0^H$ and $\eta_0^H = S^{N_0^H}(\eta_0)$. Starting from (1) and introducing the particle approximation

$$\gamma_k^H = g_k\, S^{N_k^H}(\mu_{k-1}^H\, Q_k)\, \langle \gamma_{k-1}^H, 1\rangle = g_k\, \eta_k^H\, \langle \gamma_{k-1}^H, 1\rangle\ ,$$

and

$$\gamma_0^H = g_0\, S^{N_0^H}(\eta_0) = g_0\, \eta_0^H\ ,$$

for the unnormalized (linear) flow, it is easily seen that

$$\langle \gamma_k^H, 1\rangle = \langle \eta_k^H, g_k\rangle\, \langle \gamma_{k-1}^H, 1\rangle \qquad \text{and} \qquad \langle \gamma_0^H, 1\rangle = \langle \eta_0^H, g_0\rangle\ , \tag{7}$$

hence

$$\frac{\gamma_k^H}{\langle \gamma_k^H, 1\rangle} = g_k \cdot \eta_k^H = \mu_k^H \qquad \text{and} \qquad \frac{\gamma_0^H}{\langle \gamma_0^H, 1\rangle} = g_0 \cdot \eta_0^H = \mu_0^H \ .$$

Clearly, $N_k^H \geq H$ and if $\langle \mu_{k-1}^H Q_k, g_k\rangle > 0$ — a sufficient condition for which is

$$\widehat{g}_k(x) = Q_k\, g_k(x) = \mathbb{E}[g_k(X_k) \mid X_{k-1} = x] > 0 \ ,$$

for any x in the support of μ_{k-1}^H — then the random number N_k^H of particles is a.s. finite, see Section 4 below. Moreover

$$\langle \eta_0^H, g_0\rangle = \langle S^{N_0^H}(\eta_0), g_0\rangle = \frac{1}{N_0^H}\sum_{i=1}^{N_0^H} g_0(\xi_0^i) \geq \frac{H}{N_0^H}\sup_{x\in E} g_0(x) > 0 \ ,$$

and

$$\langle \eta_k^H, g_k\rangle = \langle S^{N_k^H}(\mu_{k-1}^H Q_k), g_k\rangle = \frac{1}{N_k^H}\sum_{i=1}^{N_k^H} g_k(\xi_k^i) \geq \frac{H}{N_k^H}\sup_{x\in E} g_k(x) > 0 \ ,$$

for any $k = 1, \cdots, n$, i.e. the particle system never dies out and the algorithm can always continue, by construction.

Remark 3. It follows from Lemma 3 below that $\dfrac{N_0^H}{H} \to \rho_0$ in probability, and in view of Remark 8 (ii) below, if $\langle \mu_{k-1}^H Q_k, g_k\rangle > 0$ then $\dfrac{N_k^H}{H\,\rho_k^H} \to 1$ in probability as $H \uparrow \infty$, where

$$\rho_k^H = \frac{\sup_{x\in E} g_k(x)}{\langle \mu_{k-1}^H Q_k, g_k\rangle} \ ,$$

for any $k = 1, \cdots, n$.

Remark 4. For any $k = 0, 1, \cdots, n$ and any integer $i \geq 1$, let $\mathcal{F}_{k,i}^H = \mathcal{F}_{k,0}^H \vee \sigma(\xi_k^1, \cdots, \xi_k^i)$, where $\mathcal{F}_{0,0}^H = \{\emptyset, \Omega\}$ (for $k = 0$) and $\mathcal{F}_{k,0}^H = \mathcal{H}_{k-1}^H$ (for $k = 1, \cdots, n$) by convention. The random number N_k^H is a stopping time w.r.t. $\mathcal{F}_k^H = \{\mathcal{F}_{k,i}^H, \, i \geq 0\}$, which allows to define the σ–algebra $\mathcal{F}_{k,N_k^H}^H = \mathcal{H}_k^H$: clearly N_k^H is measurable w.r.t. $\mathcal{H}_k^H$, and therefore the random variable

$$\sigma_k^H = N_0^H + \cdots + N_k^H \ ,$$

is measurable w.r.t. $\mathcal{H}_k^H$.

Example 3 (Binary selection). In the special case of binary selection functions, the sequential particle algorithm results in the following approximations

$$p_k \approx p_k^H = \langle \eta_k^H, g_k\rangle = \frac{H}{N_k^H} \qquad \text{where} \qquad N_k^H = \inf\{N \geq 1 \, : \, |I_k^N| = H\} \ ,$$

for any integer $H \geq 1$, and where for any integer $N \geq 1$

$$I_k^N = \{i = 1, \cdots, N \,:\, \xi_k^i \in A_k\} \ ,$$

denotes the set of successful particles within an N–sample with common probability distribution η_0 (for $k = 0$) and $\mu_{k-1}^H Q_k$ (for $k = 1, \cdots, n$), and it follows from (7) that

$$P_n \approx P_n^H = \langle \gamma_n^H, 1 \rangle = \prod_{k=0}^{n} \langle \eta_k^H, g_k \rangle = \prod_{k=0}^{n} \frac{H}{N_k^H} \ .$$

Notice that the approximation $\mu_k^H = g_k \cdot \eta_k^H$ obtained here is exactly the empirical probability distribution associated with an H–sample that would be obtained using the rejection method, with common probability distribution $g_0 \cdot \eta_0$ (for $k = 0$) and $g_k \cdot (\mu_{k-1}^H Q_k)$ (for $k = 1, \cdots, n$). Here again, the probability P_n of a successful sequence is approximated as the product of the fraction of successful particles at each generation, and each transition probability p_k separately is approximated as the fraction of successful particles at the corresponding generation. In opposition to the nonsequential particle algorithm, notice that the number H of successful particles at each generation is *fixed* in advance, whereas the computational effort, i.e. the number N_k^H of simulated particles needed to get H successful particles exactly at the k–th generation, is *random*.

The main contributions of this paper are the following results for the sequential particle algorithm with a random number of particles, defined by the level $H > 0$: a nonasymptotic estimate (which was already obtained in [11, Theorem 5.4] in a different context), see Theorem 3 below

$$\sup_{\phi \,:\, \|\phi\| = 1} \mathbb{E} | \langle \mu_n^H - \mu_n, \phi \rangle | \leq \frac{c_n}{\sqrt{H}} \ ,$$

and a central limit theorem, see Theorem 4 below

$$\sqrt{H} \, \langle \mu_n^H - \mu_n, \phi \rangle \Longrightarrow \mathcal{N}(0, v_n(\phi)) \ ,$$

in distribution as $H \uparrow \infty$, with an explicit expression for the asymptotic variance.

Theorem 3. *If $\langle \mu_{k-1}^H Q_k, g_k \rangle > 0$ for any $k = 1, \cdots, n$ — a sufficient condition for which is*

$$\widehat{g}_k(x) = Q_k \, g_k(x) = \mathbb{E}[g_k(X_k) \mid X_{k-1} = x] > 0 \ ,$$

for any x in the support of μ_{k-1}^H — then

$$\mathbb{E} \left| \frac{\langle \gamma_n^H, 1 \rangle}{\langle \gamma_n, 1 \rangle} - 1 \right| \leq z_n^H \qquad \textit{and} \qquad \sup_{\phi \,:\, \|\phi\| = 1} \mathbb{E} | \langle \mu_n^H - \mu_n, \phi \rangle | \leq 2 \, z_n^H \ , \tag{8}$$

where the sequence $\{z_k^H\ ,\ k=0,1,\cdots,n\}$ *satisfies the linear recursion*

$$z_k^H \leq \rho_k\,(1+\omega_H+\omega_H^2)\,z_{k-1}^H+\omega_H\,(1+\omega_H\,\rho_k)\ , \tag{9}$$

and

$$z_0^H \leq \omega_H\,(1+\omega_H\,\rho_0)\ ,$$

where $\omega_H=\dfrac{1}{H}\,\sqrt{H+1}$ *is of order* $1/\sqrt{H}$.

The proof of Theorem 3 can be found in [13, Section 6].

Remark 5. Up to higher order terms, the forcing term in (9) is $1/\sqrt{H}$ exactly (which is equivalent to $\overline{\rho_k^H}/\ \overline{N_k^H}$, in view of Remark 3), and

$$\limsup_{H\uparrow\infty}\,[\sqrt{H}\,z_k^H]\leq\rho_k\ \limsup_{H\uparrow\infty}\,[\sqrt{H}\,z_{k-1}^H]+1 \qquad \text{and} \qquad \limsup_{H\uparrow\infty}\,[\sqrt{H}\,z_0^H]\leq 1\,.$$

In opposition to the nonsequential particle algorithm, notice that it is possible here to guarantee in advance a fixed performance of $1/\sqrt{H}$ exactly, without any knowledge of ρ_k, at the expense of using an adaptive random number N_k^H of simulated particles : this could be seen as an adaptive rule to automatically choose the number of particles.

Remark 6. It follows from Theorem 3 that $\langle\mu_{k-1}^H\,Q_k,g_k\rangle\rightarrow\langle\eta_k,g_k\rangle$ in probability, hence $\rho_k^H\rightarrow\rho_k$ in probability, for any $k=1,\cdots,n$, and it follows from Remark 3 that $\dfrac{N_k^H}{H}\rightarrow\rho_k$ in probability as $H\uparrow\infty$, for any $k=0,1,\cdots,n$.

Theorem 4. *If* $\langle\mu_{k-1}^H\,Q_k,g_k\rangle>0$ *for any* $k=1,\cdots,n$ *— a sufficient condition for which is*

$$\widehat{g}_k(x)=Q_k\,g_k(x)=\mathbb{E}[g_k(X_k)\mid X_{k-1}=x]>0\ ,$$

for any x *in the support of* μ_{k-1}^H *— then*

$$\sqrt{H}\,[\,\frac{\langle\gamma_n^H,1\rangle}{\langle\gamma_n,1\rangle}-1\,]\Longrightarrow\mathcal{N}(0,V_n)\ , \tag{10}$$

and

$$\sqrt{H}\,\langle\mu_n^H-\mu_n,\phi\rangle\Longrightarrow\mathcal{N}(0,v_n(\phi))\ , \tag{11}$$

in distribution as $H\uparrow\infty$, *for any bounded measurable function* ϕ, *with the asymptotic variance*

$$V_n=\sum_{k=0}^{n}\frac{\mathrm{var}(g_k\,R_{k+1:n}\,1,\eta_k)}{\langle\eta_k,g_k\,R_{k+1:n}\,1\rangle^2}\ \frac{1}{\rho_k}\ ,$$

and

$$v_n(\phi) = \sum_{k=0}^{n} \frac{\mathrm{var}(g_k\, R_{k+1:n}\,(\phi - \langle \mu_n, \phi\rangle), \eta_k)}{\langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle^2}\, \frac{1}{\rho_k}\ ,$$

respectively, where

$$R_{k+1:n}\,\phi(x) = R_{k+1}\cdots R_n\,\phi(x) = \mathbb{E}[\phi(X_n)\ \prod_{p=k+1}^{n} g_p(X_p)\mid X_k = x]\ ,$$

for any $k = 0, 1, \cdots, n$, *with the convention* $R_{n+1:n}\,\phi(x) = \phi(x)$, *for any* $x \in E$.

In view of Remark 2, the following equivalent expression holds for the asymptotic variance

$$v_n(\phi) = \sum_{k=0}^{n} \frac{\langle \eta_k, |g_k\, R_{k+1:n}\,[\phi - \langle \mu_n, \phi\rangle]\,|^2\,\rangle}{\langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle^2}\, \frac{1}{\rho_k}\ .$$

Remark 7. To prove Theorem 4, it is enough to prove that

$$\sqrt{H}\ \frac{\langle \gamma_n^H - \gamma_n, \phi\rangle}{\langle \gamma_n, 1\rangle} \Longrightarrow \mathcal{N}(0, V_n(\phi))\ , \tag{12}$$

for any bounded measurable function ϕ, where the asymptotic variance $V_n(\phi)$ is defined by

$$\begin{aligned} &V_n(\phi)\ \langle \gamma_n, 1\rangle^2 \\ &\quad = \mathrm{var}(g_0\, R_{1:n}\,\phi, \eta_0)\ \frac{1}{\rho_0} + \sum_{k=1}^{n} \mathrm{var}(g_k\, R_{k+1:n}\,\phi, \eta_k)\ \frac{\langle \gamma_{k-1}, 1\rangle^2}{\rho_k}\ , \end{aligned} \tag{13}$$

or equivalently by

$$V_n(\phi) = \sum_{k=0}^{n} \frac{\mathrm{var}(g_k\, R_{k+1:n}\,\phi, \eta_k)}{\langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle^2}\, \frac{1}{\rho_k}\ ,$$

since

$$\langle \gamma_n, 1\rangle = \langle \gamma_0\, R_{1:n}, 1\rangle = \langle \eta_0, g_0\, R_{1:n}\, 1\rangle\ ,$$

and since

$$\langle \gamma_n, 1\rangle = \langle \gamma_{k-1}\, R_{k:n}, 1\rangle = \langle \gamma_{k-1}, 1\rangle\ \langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle\ ,$$

for any $k = 1, \cdots, n$. Indeed, notice that

$$\langle \mu_n^H - \mu_n, \phi\rangle = \langle \frac{\gamma_n^H}{\langle \gamma_n^H, 1\rangle}, \phi - \langle \mu_n, \phi\rangle\,\rangle = \frac{\langle \gamma_n, 1\rangle}{\langle \gamma_n^H, 1\rangle}\ \langle \frac{\gamma_n^H - \gamma_n}{\langle \gamma_n, 1\rangle}, \phi - \langle \mu_n, \phi\rangle\,\rangle\ ,$$

for any bounded measurable function ϕ, and it follows from Theorem 3 that $\langle \gamma_n^H, 1\rangle \to \langle \gamma_n, 1\rangle$ in probability as $H \uparrow \infty$, hence (10) and (11) follow from (12) and from the Slutsky lemma, with $V_n = V_n(1)$ and $v_n(\phi) = V_n(\phi - \langle \mu_n, \phi\rangle)$, respectively.

Moreover, if (12) holds, then

$$(\sqrt{H}\,[\,\frac{\langle \gamma_n^H, 1\rangle}{\langle \gamma_n, 1\rangle} - 1\,], \sqrt{H}\,\langle \mu_n^H - \mu_n, \phi_1\rangle, \cdots, \sqrt{H}\,\langle \mu_n^H - \mu_n, \phi_d\rangle)\ ,$$

converge jointly in distribution as $H \uparrow \infty$ to a Gaussian limit, for any bounded measurable functions $\phi_1, \cdots, \phi_d$, using the Cramér–Wold device.

Two different proofs of (12) are given in Sections 5 and 6, respectively. A first proof follows the approach of [9, Theorem 4] by induction, and relies on an extension of a central limit theorem for sums of a random number of i.i.d. random variables, see Theorem 6 below. An alternate proof follows the approach of [3, Chapter 9], see also [7, Proposition 2.9, Corollary 2.20], and relies on an original central limit theorem for triangular arrays of martingale increments spread across generations with different random sizes, see Theorem 7 below.

To get a fair comparison of the nonsequential and sequential particle algorithms, the time–averaged random number of simulated particles, which is an indicator of how much computational effort has been used, can be used as a normalizing factor instead of the level $H > 0$. It follows from Remark 6 that

$$\frac{1}{H}\,[\,\frac{1}{n+1}\,\sum_{k=0}^{n} N_k^H\,] \longrightarrow \frac{1}{n+1}\,\sum_{k=0}^{n} \rho_k\ ,$$

in probability as $H \uparrow \infty$, hence under the assumptions of Theorem 4, and using the Slutsky lemma

$$[\,\frac{1}{n+1}\,\sum_{k=0}^{n} N_k^H\,]^{1/2}\,\frac{\langle \gamma_n^H - \gamma_n, 1\rangle}{\langle \gamma_n, 1\rangle} \Longrightarrow \mathcal{N}(0, V_n^*)\ ,$$

and

$$[\,\frac{1}{n+1}\,\sum_{k=0}^{n} N_k^H\,]^{1/2}\,\langle \mu_n^H - \mu_n, \phi\rangle \Longrightarrow \mathcal{N}(0, v_n^*(\phi))\ ,$$

in distribution as $H \uparrow \infty$, with the asymptotic variance

$$V_n^* = [\,\frac{1}{n+1}\,\sum_{k=0}^{n} \rho_k\,]\,V_n \qquad \text{and} \qquad v_n^*(\phi) = [\,\frac{1}{n+1}\,\sum_{k=0}^{n} \rho_k\,]\,v_n(\phi)\ ,$$

respectively, where V_n and $v_n(\phi)$ are defined in Theorem 4. Notice that the asymptotic variances V_n^0 and $v_n^0(\phi)$ defined in Theorem 2 for the nonsequential particle algorithm coincide with the asymptotic variances V_n^* and $v_n^*(\phi)$ for the renormalized sequential particle algorithm respectively, in the special case where $\rho_0 = \rho_1 = \cdots = \rho_n$.

Example 4 (Binary selection). In the special case of binary selection functions, the support of μ_{k-1}^H is contained in A_{k-1}, and if

$$Q_k(x, A_k) = \mathbb{P}[X_k \in A_k \mid X_{k-1} = x] > 0 \ ,$$

for any $x \in A_{k-1}$, then it follows from Theorem 4 that

$$\sqrt{H}\,(\frac{P_n^H}{P_n} - 1) \Longrightarrow \mathcal{N}(0, V_n) \ ,$$

in distribution as $H \uparrow \infty$, with the asymptotic variance

$$V_n = \sum_{k=0}^{n}(1 - p_k) + \sum_{k=0}^{n} \frac{\mathrm{var}(R_{k+1:n}\,1, \mu_k)}{\langle \mu_k, R_{k+1:n}\,1\rangle^2} \ ,$$

since $1/\rho_k = \langle \eta_k, g_k\rangle = p_k$ for any $k = 0, 1, \cdots, n$.

4 Limit Theorems in Sequential Analysis

In this section, some basic properties are proved for sums of a random number of i.i.d. random variables, especially when this random number is a stopping time. Let $\xi_1, \cdots, \xi_i, \cdots$ be i.i.d. random variables with common probability distribution μ, and let Λ be a nonnegative bounded measurable function, possibly taking the zero value. For any $H > 0$, consider the stopping time

$$N_H = \inf\{N \geq 1 \ : \ \sum_{i=1}^{N} \Lambda(\xi_i) \geq H\,\lambda\} \qquad \text{where} \qquad \lambda = \sup_{x \in E} \Lambda(x) \ .$$

Lemma 1. *If $\langle \mu, \Lambda\rangle > 0$, then the stopping time N_H is a.s. finite and integrable.*

Proof. By the strong law of large numbers, it follows that

$$\frac{1}{N} \sum_{i=1}^{N} \Lambda(\xi_i) \longrightarrow \langle \mu, \Lambda\rangle \ ,$$

a.s. as $N \uparrow \infty$, and if $\langle \mu, \Lambda\rangle > 0$, then

$$\sum_{i=1}^{N} \Lambda(\xi_i) \longrightarrow \infty \ ,$$

and the finite level $H\,\lambda$ is reached after a finite number of steps, i.e. the stopping time N_H is a.s. finite. In addition, for any $a > 0$

$$\begin{aligned}\mathbb{P}[N_H > N] = \mathbb{P}[\sum_{i=1}^{N} \Lambda(\xi_i) < H\,\lambda] &= \mathbb{P}[\exp\{-a \sum_{i=1}^{N} \Lambda(\xi_i)\} > e^{-a\,H\,\lambda}] \\ &\leq e^{a\,H\,\lambda}\, r^N \ ,\end{aligned}$$

by independence, where

$$r = \mathbb{E}[\exp\{-a\,\Lambda(\xi)\}] = \int_E e^{-a\,\Lambda(x)}\,\mu(dx) = \langle \mu, e^{-a\,\Lambda}\rangle \ ,$$

and $r < 1$ if and only if $\langle \mu, \Lambda\rangle > 0$. This proves that the stopping time N_H is integrable, and the estimate

$$\mathbb{E}[N_H] = \sum_{N=0}^{\infty} \mathbb{P}[N_H > N] \leq e^{a\,H\,\lambda} \sum_{N=0}^{\infty} r^N \leq \frac{e^{a\,H\,\lambda}}{1-r} < \infty \ ,$$

holds. □

Lemma 2. *If $\langle \mu, \Lambda\rangle > 0$, then the rough estimate*

$$\sup_{\phi\,:\,\|\phi\|=1} \{\, \mathbb{E}\,|\,\langle S^{N_H}(\mu) - \mu, \Lambda\,\phi\rangle\,|^2 \,\}^{1/2} \leq \omega_H\,\lambda \ ,$$

and the refined estimate

$$\sup_{\phi\,:\,\|\phi\|=1} \mathbb{E}\,|\,\langle S^{N_H}(\mu) - \mu, \Lambda\,\phi\rangle\,| \leq \omega_H\,[\,\langle \mu, \Lambda\rangle + \omega_H\,\lambda\,] \ ,$$

hold, where $\omega_H = \dfrac{1}{H}\sqrt{H+1}$ is of order $1/\sqrt{H}$.

Proof. Let

$$\delta_H = \Lambda\,(S^{N_H}(\mu) - \mu) \qquad \text{and} \qquad \delta'_H = \frac{\Lambda\,(S^{N_H}(\mu) - \mu)}{\langle S^{N_H}(\mu), \Lambda\rangle} \ .$$

Notice that

$$\begin{aligned}\delta_H = \delta'_H\,\langle S^{N_H}(\mu), \Lambda\rangle &= \delta'_H\,[\,\langle \mu, \Lambda\rangle + \langle \delta_H, 1\rangle\,] \\ &= \delta'_H\,[\,\langle \mu, \Lambda\rangle + \langle \delta'_H, 1\rangle\,\langle S^{N_H}(\mu), \Lambda\rangle\,] \ ,\end{aligned}$$

hence

$$|\,\langle \delta_H, \phi\rangle\,| \leq |\,\langle \delta'_H, \phi\rangle\,|\,\lambda \ ,$$

and

$$|\,\langle \delta_H, \phi\rangle\,| \leq |\,\langle \delta'_H, \phi\rangle\,|\,[\,\langle \mu, \Lambda\rangle + |\,\langle \delta'_H, 1\rangle\,|\,\lambda\,] \ ,$$

for any bounded measurable function ϕ. It follows from (the proof of) Lemma 5.4 in [11] that

$$\sup_{\phi\,:\,\|\phi\|=1} \{\, \mathbb{E}\,|\,\langle \delta'_H, \phi\rangle\,|^2 \,\}^{1/2} \leq \frac{1}{H}\,\sqrt{H+1} = \omega_H \ ,$$

which immediately proves the rough estimate, and using the Cauchy–Schwartz inequality and the Minkowski triangle inequality yields

$$\begin{aligned}
\mathbb{E}\,|\,\langle \delta_H, \phi\rangle\,| \quad &\leq \quad \mathbb{E}[\,|\,\langle \delta'_H, \phi\rangle\,|\,[\,\langle \mu, \Lambda\rangle + |\,\langle \delta'_H, 1\rangle\,|\,\lambda\,]\,] \\
&\leq \quad \{\, \mathbb{E}\,|\,\langle \delta'_H, \phi\rangle\,|^2 \,\}^{1/2}\,[\,\langle \mu, \Lambda\rangle + \{\, \mathbb{E}\,|\,\langle \delta'_H, 1\rangle\,|^2 \,\}^{1/2}\,\lambda\,] \\
&\leq \quad \omega_H\,[\,\langle \mu, \Lambda\rangle + \omega_H\,\lambda\,]\,\|\phi\| \ ,
\end{aligned}$$

which proves the refined estimate. □

Lemma 3. *If* $\langle \mu, \Lambda\rangle > 0$, *then* $\dfrac{N_H}{H\,\rho} \to 1$ *and* $\dfrac{H}{N_H} \to \dfrac{1}{\rho}$ *in* $\mathbb{L}^2$ *as* $H \uparrow \infty$, *with rate* $1/\sqrt{H}$, *where* $\rho = \dfrac{\lambda}{\langle \mu, \Lambda\rangle}$.

Proof. For any $N \geq 1$, define

$$D_N = \sum_{i=1}^{N} \Lambda(\xi_i) \qquad \text{and} \qquad M_N = \sum_{i=1}^{N} [\Lambda(\xi_i) - \langle \mu, \Lambda\rangle\,] = D_N - N\,\langle \mu, \Lambda\rangle \ .$$

By definition of the stopping time N_H, it holds

$$H\,\lambda \leq D_{N_H} = D_{N_H - 1} + \Lambda(\xi_{N_H}) \leq (H+1)\,\lambda \ ,$$

hence, upon subtracting $H\,\lambda$ throughout

$$0 \leq D_{N_H} - H\,\lambda \leq \lambda \ .$$

Using the decomposition

$$N_H\,\langle \mu, \Lambda\rangle - H\,\lambda = D_{N_H} - H\,\lambda - M_{N_H} \ ,$$

and the triangle inequality yields

$$|N_H\,\langle \mu, \Lambda\rangle - H\,\lambda| \leq |D_{N_H} - H\,\lambda| + |M_{N_H}| \leq \lambda + |M_{N_H}| \ .$$

Since $\langle \mu, \Lambda\rangle > 0$, it follows from Lemma 1 that the stopping time N_H is integrable, and it follows from the Wald identity, see e.g. [14, Proposition IV–4–21], that

$$\mathbb{E}[D_{N_H}] = \mathbb{E}[N_H]\,\langle \mu, \Lambda\rangle \qquad \text{and} \qquad \mathbb{E}|M_{N_H}|^2 = \mathbb{E}[N_H]\,\mathrm{var}(\Lambda, \mu) \ ,$$

hence

$$\mathbb{E}|M_{N_H}|^2 = \frac{\mathrm{var}(\Lambda,\mu)}{\langle \mu,\Lambda\rangle}\,\mathbb{E}[D_{N_H}] \le (H+1)\,\lambda^2 \ ,$$

since $\mathrm{var}(\Lambda,\mu) = \langle \mu,\Lambda^2\rangle - \langle \mu,\Lambda\rangle^2 \le \langle \mu,\Lambda^2\rangle \le \lambda\,\langle \mu,\Lambda\rangle$, and since $D_{N_H} \le (H+1)\,\lambda$. Using the Minkowski triangle inequality yields

$$\{\mathbb{E}|N_H\,\langle \mu,\Lambda\rangle - H\,\lambda|^2\}^{1/2} \le \lambda + \{\mathbb{E}|M_{N_H}|^2\}^{1/2} \le (\sqrt{H+1}+1)\,\lambda \ ,$$

and, upon dividing by $H\,\lambda$ throughout

$$\{\mathbb{E}|\frac{N_H}{H\,\rho} - 1|^2\}^{1/2} \le \frac{1}{H}\,(\sqrt{H+1}+1) = \omega_H + \frac{1}{H} \ ,$$

where ω_H is of order $1/\sqrt{H}$. Since $N_H \ge H$, it holds

$$|\frac{H}{N_H} - \frac{1}{\rho}| \le \frac{N_H}{H}\,|\frac{H}{N_H} - \frac{1}{\rho}| = \frac{1}{H\,\lambda}\,|N_H\,\langle \mu,\Lambda\rangle - H\,\lambda| \ ,$$

hence

$$\{\mathbb{E}|\frac{H}{N_H} - \frac{1}{\rho}|^2\}^{1/2} \le \frac{1}{H}\,(\sqrt{H+1}+1) = \omega_H + \frac{1}{H} \ . \qquad \square$$

Remark 8. A direct look into the proofs of Lemma 2 and Lemma 3 shows that a conditional version of the same results holds under the following assumptions. For any $H > 0$, let $\xi_1^H,\cdots,\xi_i^H,\cdots$ be i.i.d. random variables conditionally w.r.t. the σ–algebra $\mathcal{F}^H$, with common conditional probability distribution μ_H, let Λ be a nonnegative bounded measurable function, possibly taking the zero value, and consider the stopping time

$$N_H = \inf\{N \ge 1 \,:\, \sum_{i=1}^{N} \Lambda(\xi_i) \ge H\,\lambda\} \qquad \text{where} \qquad \lambda = \sup_{x\in E} \Lambda(x) \ .$$

If $\langle \mu_H,\Lambda\rangle > 0$, then (i) the rough estimate

$$\sup_{\phi\,:\,\|\phi\|=1} \{\,\mathbb{E}[\,|\,\langle S^{N_H}(\mu_H) - \mu_H, \Lambda\,\phi\rangle\,|^2 \mid \mathcal{F}^H\,]\,\}^{1/2} \le \omega_H\,\lambda \ ,$$

and the refined estimate

$$\sup_{\phi\,:\,\|\phi\|=1} \mathbb{E}[\,|\,\langle S^{N_H}(\mu_H) - \mu_H, \Lambda\,\phi\rangle\,|\mid \mathcal{F}^H] \le \omega_H\,[\,\langle \mu_H,\Lambda\rangle + \omega_H\,\lambda\,] \ ,$$

hold, where ω_H of order $1/\sqrt{H}$, and (ii)

$$\{\mathbb{E}[\,|\frac{N_H}{H\rho_H} - 1|^2 \mid \mathcal{F}^H]\}^{1/2} \le \omega_H + \frac{1}{H} \ ,$$

and

$$\{\mathbb{E}[\,|\frac{H}{N_H} - \frac{1}{\rho_H}|^2 \mid \mathcal{F}^H]\}^{1/2} \le \omega_H + \frac{1}{H} \ ,$$

with $\rho_H = \dfrac{\lambda}{\langle \mu_H,\Lambda\rangle}$.

The following central limit theorem, known in sequential analysis as the Anscombe theorem, has been proved in [17] for sums of a random number of i.i.d. random variables, see also [8, Theorem I.3.1] or [20, Theorem 2.40].

Theorem 5 (Anscombe). *For any $H > 0$, let $\rho_H > 0$ be a deterministic constant, and let $X_1^H, \cdots, X_i^H, \cdots$ be i.i.d. random variables with zero mean and variance σ_H^2. If $r_H = \lfloor H\rho_H \rfloor \to \infty$, if*

$$\frac{N_H}{H\rho_H} \longrightarrow 1 \ ,$$

in probability, and if the Lindeberg condition

$$\mathbb{E}[1_{\{|\frac{X_i^H}{\sigma_H}| \geq c\sqrt{r_H}\}} |\frac{X_i^H}{\sigma_H}|^2] \longrightarrow 0 \ ,$$

holds for any $c > 0$, then

$$\frac{1}{\sqrt{N_H}} \sum_{i=1}^{N_H} \frac{X_i^H}{\sigma_H} \Longrightarrow \mathcal{N}(0,1) \qquad \textit{and} \qquad \frac{1}{\sqrt{H\rho_H}} \sum_{i=1}^{N_H} \frac{X_i^H}{\sigma_H} \Longrightarrow \mathcal{N}(0,1) \ ,$$

in distribution as $H \uparrow \infty$.

Remark 9. Using the Slutsky lemma, and since $\frac{\sqrt{N_H}}{\sqrt{H\rho_H}} \to 1$ in probability as $H \uparrow \infty$, the two convergence results are indeed equivalent.

The next theorem provides a stronger result, with a precise statement on the convergence of conditional characteristic functions, in a special case where both σ_H^2 and ρ_H are random variables. It is used in an essential way in Section 5, in the proof of Theorem 4 by induction.

Theorem 6. *For any $H > 0$, let $X_1^H, \cdots, X_i^H, \cdots$ be i.i.d. random variables conditionally w.r.t. the σ–algebra $\mathcal{F}^H$, with zero conditional mean and conditional variance σ_H^2, and let $\rho_H > 0$ be a $\mathcal{F}^H$–measurable r.v. If $r_H = \lfloor H\rho_H \rfloor \to \infty$ in probability, if*

$$F_H(d) = \mathbb{P}[\,|\frac{N_H}{H\rho_H} - 1| > d \mid \mathcal{F}^H\,] \longrightarrow 0 \ ,$$

in probability for any $d > 0$, and if the conditional Lindeberg condition

$$R_H(c) = \mathbb{E}[1_{\{|\frac{X_i^H}{\sigma_H}| \geq c\sqrt{r_H}\}} |\frac{X_i^H}{\sigma_H}|^2 \mid \mathcal{F}^H] \longrightarrow 0 \ ,$$

holds in probability for any $c > 0$, then for any fixed real number u

$$\mathbb{E}[\exp\{i\,\frac{u}{\sqrt{H\rho_H}}\sum_{j=1}^{N_H}\frac{X_j^H}{\sigma_H}\}\mid\mathcal{F}^H]\longrightarrow\exp\{-\tfrac{1}{2}\,u^2\}\ , \tag{14}$$

in $\mathbb{L}^1$ *as* $H\uparrow\infty$. *If in addition* $\dfrac{\sigma_H}{\sqrt{\rho_H}}\to\dfrac{\sigma}{\sqrt{\rho}}$ *in probability, then for any fixed real number* u

$$\mathbb{E}[\exp\{i\,u\sqrt{H}\,\frac{1}{N_H}\sum_{j=1}^{N_H}X_j^H\}\mid\mathcal{F}^H]\longrightarrow\exp\{-\tfrac{1}{2}\,\frac{u^2\,\sigma^2}{\rho}\}\ , \tag{15}$$

in $\mathbb{L}^1$ *as* $H\uparrow\infty$.

Using the Lebesgue dominated convergence theorem, it is sufficient to prove that (14) and (15) hold in probability. The proof of Theorem 6 is postponed to Appendix A.

Remark 10. If (14) holds, then in particular

$$Z_H=\frac{1}{\sqrt{H\rho_H}}\sum_{j=1}^{N_H}\frac{X_j^H}{\sigma_H}\Longrightarrow\mathcal{N}(0,1)\ ,$$

in distribution as $H\uparrow\infty$.

Remark 11. If $F_H(d)\to 0$ in probability for any $d>0$, (or equivalently in $\mathbb{L}^1$ using the Lebesgue dominated convergence theorem), then equivalently $\mathbb{E}[F_H(d)]\to 0$ for any $d>0$, since these r.v.'s are nonnegative, which means that $\dfrac{N_H}{H\rho_H}\to 1$ in probability as $H\uparrow\infty$.

The last result of this section is a central limit theorem for triangular arrays of martingale increments spread across generations with random sizes. It is used in an essential way in Section 6, in an alternate proof of Theorem 4.

Theorem 7. *For any* $k=0,1,\cdots,n$, *let* $\mathcal{F}_k^H=\{\mathcal{F}_{k,i}^H\,,\ i\geq 0\}$ *be an increasing sequence of* σ*–algebras, let* N_k^H *be a stopping time w.r.t.* $\mathcal{F}_k^H$, *which allows to define the* σ*–algebra* $\mathcal{H}_k^H=\mathcal{F}_{k,N_k^H}^H$, *assume that* $\mathcal{F}_{0,0}^H=\{\emptyset,\Omega\}$ *(for* $k=0$*) and* $\mathcal{F}_{k,0}^H=\mathcal{H}_{k-1}^H$ *(for* $k=1,\cdots,n$*), and let* $\{X_{k,i}^H\,,\ i\geq 1\}$ *be a sequence of square integrable random variables adapted to* $\mathcal{F}_k^H$, *such that*

$$\mathbb{E}[X_{k,i}^H\mid\mathcal{F}_{k,i-1}^H]=0\ , \tag{16}$$

$$\mathbb{E}[\,|X_{k,i}^H|^2\mid\mathcal{F}_{k,i-1}^H]=V_{k,0}^H\ , \tag{17}$$

and

$$\mathbb{E}[\,|X_{k,i}^H|^2\,1_{\{|X_{k,i}^H|>\varepsilon\}}\mid\mathcal{F}_{k,i-1}^H]\leq Y_{k,0}^{H,\varepsilon}\ , \tag{18}$$

for any $i\geq 1$, *where* $V_{k,0}^H$ *and* $Y_{k,0}^{H,\varepsilon}$ *are measurable w.r.t.* $\mathcal{F}_{k,0}^H$. *If for any* $\varepsilon>0$

$$\sum_{k=0}^{n} N_k^H \, V_{k,0}^H \longrightarrow W_n \qquad \textit{and} \qquad \sum_{k=0}^{n} N_k^H \, Y_{k,0}^{H,\varepsilon} \longrightarrow 0 \ , \tag{19}$$

in probability, then

$$S_n^H = \sum_{k=0}^{n} \sum_{i=1}^{N_k^H} X_{k,i}^H \Longrightarrow \mathcal{N}(0, W_n) \ ,$$

in distribution as $H \uparrow \infty$.

The proof of Theorem 7 is postponed to Appendix B. The idea is to rewrite S_n^H as a single sum across all generations, and to use a central limit theorem for triangular arrays of martingale increments [2, Theorem 2.8.42].

5 Proof of Theorem 4 by induction

In view of Remark 7 above, the problem reduces to prove (12), i.e. to prove asymptotic normality for the unnormalized linear flow. The proof given below follows the approach of [9, Theorem 4] by induction.

Proof of Theorem 4. Notice first that

$$\langle \gamma_0^H - \gamma_0, \phi \rangle = \langle S^{N_0^H}(\eta_0) - \eta_0, g_0 \, \phi \rangle = \frac{1}{N_0^H} \sum_{j=1}^{N_0^H} X_{0,j}^H(\phi) \ ,$$

where

$$X_{0,j}^H(\phi) = g_0(\xi_0^j) \, \phi(\xi_0^j) - \langle \eta_0, g_0 \, \phi \rangle \ ,$$

for any $j = 1, \cdots, N_0^H$, and where $\xi_0^1, \cdots, \xi_0^j, \cdots$ are i.i.d. random variables with common probability distribution η_0, hence the random variables $X_{0,1}^H(\phi), \cdots, X_{0,j}^H(\phi), \cdots$ are i.i.d. with zero mean and with variance $\mathrm{var}(g_0 \, \phi, \eta_0)$ independent of $H > 0$. It follows from Lemma 3 that $\dfrac{N_0^H}{H} \to \rho_0$ in probability as $H \uparrow \infty$, hence the assumptions of Theorem 5 are satisfied, and the induction assumption (12) holds at step 0, with

$$V_0(\phi) \, \langle \gamma_0, 1 \rangle^2 = \mathrm{var}(g_0 \, \phi, \eta_0) \, \frac{1}{\rho_0} \ .$$

Assume now that the induction assumption (12) holds at step $(k-1)$. Notice that

$$\gamma_k^H - \gamma_k = \gamma_k^H - \gamma_{k-1}^H \, R_k + (\gamma_{k-1}^H - \gamma_{k-1}) \, R_k \ ,$$

hence

$$\langle \gamma_k^H - \gamma_k, \phi \rangle = \langle \gamma_k^H - \gamma_{k-1}^H \, R_k, \phi \rangle + \langle \gamma_{k-1}^H - \gamma_{k-1}, R_k \, \phi \rangle \ ,$$

for any bounded measurable function ϕ, and the last term goes to zero in probability as $H \uparrow \infty$. Notice also that

$$\begin{aligned}\langle \gamma_k^H - \gamma_{k-1}^H R_k, \phi\rangle &= \langle S^{N_k^H}(\mu_{k-1}^H Q_k) - \mu_{k-1}^H Q_k, g_k\,\phi\rangle\,\langle \gamma_{k-1}^H, 1\rangle \\ &= \frac{1}{N_k^H}\sum_{j=1}^{N_k^H} X_{k,j}^H(\phi)\ ,\end{aligned}$$

where

$$X_{k,j}^H(\phi) = [\,g_k(\xi_k^j)\,\phi(\xi_k^j) - \langle \mu_{k-1}^H Q_k, g_k\,\phi\rangle\,]\,\langle \gamma_{k-1}^H, 1\rangle\ ,$$

for any $j = 1, \cdots, N_k^H$, and where, conditionally w.r.t. the σ–algebra $\mathcal{H}_{k-1}^H$ generated by the particle system up to the $(k-1)$–th generation, the random variables $\xi_k^1, \cdots, \xi_k^j, \cdots$ are i.i.d. with common probability distribution $\mu_{k-1}^H Q_k$, hence the random variables $X_{k,1}^H(\phi), \cdots, X_{k,j}^H(\phi), \cdots$ are i.i.d. with zero conditional mean and with conditional variance

$$(\sigma_k^H(\phi))^2 = \mathrm{var}(g_k\,\phi, \mu_{k-1}^H Q_k)\,\langle \gamma_{k-1}^H, 1\rangle^2\ .$$

In view of Remark 8 (ii)

$$F_H(d) = \mathbb{P}[\,|\frac{N_k^H}{H\rho_k^H} - 1| > d \mid \mathcal{H}_{k-1}^H\,] \longrightarrow 0 \qquad \text{with} \qquad \rho_k^H = \frac{\sup_{x\in E} g_k(x)}{\langle \mu_{k-1}^H Q_k, g_k\rangle}\ ,$$

in probability for any $d > 0$, as $H \uparrow \infty$. It follows from Theorem 3 that $\langle \gamma_{k-1}^H, 1\rangle \to \langle \gamma_{k-1}, 1\rangle$, $\langle \mu_{k-1}^H Q_k, g_k\rangle \to \langle \eta_k, g_k\rangle$ and $\mathrm{var}(g_k\,\phi, \mu_{k-1}^H Q_k) \to \mathrm{var}(g_k\,\phi, \eta_k)$ in probability, hence $\rho_k^H \to \rho_k$ and $\sigma_k^H(\phi) \to \sigma_k(\phi)$ in probability as $H \uparrow \infty$, with

$$\sigma_k^2(\phi) = \mathrm{var}(g_k\,\phi, \mu_{k-1} Q_k)\,\langle \gamma_{k-1}, 1\rangle^2\ .$$

Therefore, the assumptions of Theorem 6 are satisfied, and for any fixed real number u, it holds

$$\begin{aligned}&\mathbb{E}[\,\exp\{i\,u\sqrt{H}\,\langle \gamma_k^H - \gamma_{k-1}^H R_k, \phi\rangle\}\mid \mathcal{H}_{k-1}^H] \\ &= \mathbb{E}[\,\exp\{i\,u\sqrt{H}\,\frac{1}{N_k^H}\sum_{j=1}^{N_k^H} X_{k,j}^H(\phi)\}\mid \mathcal{H}_{k-1}^H] \\ &\longrightarrow \exp\{-\tfrac{1}{2}\,u^2\,\frac{\sigma_k^2(\phi)}{\rho_k}\}\ ,\end{aligned} \tag{20}$$

in $\mathbb{L}^1$ as $H \uparrow \infty$. Notice that

$$\mathbb{E}[\exp\{i\,u\sqrt{H}\,\langle\gamma_k^H-\gamma_k,\phi\rangle\}]$$

$$-\exp\{-\tfrac{1}{2}\,u^2\,\frac{\sigma_k^2(\phi)}{\rho_k}-\tfrac{1}{2}\,u^2\,V_{k-1}(R_k\,\phi)\,\langle\gamma_{k-1},1\rangle^2\}$$

$$=\mathbb{E}[\,[\,\mathbb{E}[\exp\{i\,u\sqrt{H}\,\langle\gamma_k^H-\gamma_{k-1}^H\,R_k,\phi\rangle\}\mid\mathcal{H}_{k-1}^H]-\exp\{-\tfrac{1}{2}\,u^2\,\frac{\sigma_k^2(\phi)}{\rho_k}\}\,]$$

$$\exp\{i\,u\sqrt{H}\,\langle\gamma_{k-1}^H-\gamma_{k-1},R_k\,\phi\rangle\}\,]$$

$$+\exp\{-\tfrac{1}{2}\,u^2\,\frac{\sigma_k^2(\phi)}{\rho_k}\}\,\mathbb{E}[\exp\{i\,u\sqrt{H}\,\langle\gamma_{k-1}^H-\gamma_{k-1},R_k\,\phi\rangle\}\,]$$

$$-\exp\{-\tfrac{1}{2}\,u^2\,\frac{\sigma_k^2(\phi)}{\rho_k}\}\;\exp\{-\tfrac{1}{2}\,u^2\,V_{k-1}(R_k\,\phi)\,\langle\gamma_{k-1},1\rangle^2\}\;,$$

and the triangle inequality yields

$$|\,\mathbb{E}[\exp\{i\,u\sqrt{H}\,\langle\gamma_k^H-\gamma_k,\phi\rangle\}\,]$$

$$-\exp\{-\tfrac{1}{2}\,u^2\,\frac{\sigma_k^2(\phi)}{\rho_k}-\tfrac{1}{2}\,u^2\,V_{k-1}(R_k\,\phi)\,\langle\gamma_{k-1},1\rangle^2\}\,|$$

$$\leq\mathbb{E}|\,\mathbb{E}[\exp\{i\,u\sqrt{H}\,\langle\gamma_k^H-\gamma_{k-1}^H\,R_k,\phi\rangle\}\mid\mathcal{H}_{k-1}^H]-\exp\{-\tfrac{1}{2}\,u^2\,\frac{\sigma_k^2(\phi)}{\rho_k}\}\,|$$

$$+\,|\,\mathbb{E}[\exp\{i\,u\sqrt{H}\,\langle\gamma_{k-1}^H-\gamma_{k-1},R_k\,\phi\rangle\}\,]$$

$$-\exp\{-\tfrac{1}{2}\,u^2\,V_{k-1}(R_k\,\phi)\,\langle\gamma_{k-1},1\rangle^2\}\,|\;,$$

where the first term goes to zero using (20), and the second term goes to zero since the induction assumption (12) holds at step $(k-1)$, as $H\uparrow\infty$. Therefore, the induction assumption (12) holds at step k, with

$$V_k(\phi)\,\langle\gamma_k,1\rangle^2=\frac{\sigma_k^2(\phi)}{\rho_k}+V_{k-1}(R_k\,\phi)\,\langle\gamma_{k-1},1\rangle^2\;,$$

and iterating the above relation yields

$$\begin{aligned}V_n(\phi)\,\langle\gamma_n,1\rangle^2 &= V_0(R_{1:n}\,\phi)\,\langle\gamma_0,1\rangle^2+\sum_{k=1}^n\sigma_k^2(R_{k+1:n}\,\phi)\,\frac{1}{\rho_k}\\ &= \mathrm{var}(g_0\,R_{1:n}\,\phi,\eta_0)\,\frac{1}{\rho_0}+\sum_{k=1}^n\mathrm{var}(g_k\,R_{k+1:n}\,\phi,\eta_k)\,\frac{\langle\gamma_{k-1},1\rangle^2}{\rho_k}\;,\end{aligned}$$

which is the expression given in (13) for the asymptotic variance. In view of Remark 7, this finishes the proof of Theorem 4. □

6 Alternate Proof of Theorem 4

The alternate proof given below follows the approach of [3, Chapter 9], see also [7, Proposition 2.9, Corollary 2.20], and relies on an approximate decomposition of $\sqrt{H}\,\langle \gamma_n^H - \gamma_n, \phi\rangle$ in terms of a triangular array of martingale increments, with a different random number $\sigma_n^H = N_0^H + \cdots + N_n^H$ of such increments on each different row of the array. This requires a specific central limit theorem, see Theorem 8 below, which is of independent interest.

Let $f = (f_0, f_1, \cdots, f_n)$ be an arbitrary collection of bounded measurable functions. For any $i = 1, \cdots, N_0^H$, the random variable

$$X_{0,i}^H(f) = \frac{1}{\rho_0\sqrt{H}}\,[f_0(\xi_0^i) - \langle \eta_0, f_0\rangle] \qquad \text{where} \qquad \rho_0 = \frac{\sup_{x\in E} g_0(x)}{\langle \eta_0, g_0\rangle}\ ,$$

is measurable w.r.t. $\mathcal{F}_{0,i}^H$, where $\xi_0^1, \cdots, \xi_0^i, \cdots$ are i.i.d. random variables with common probability distribution η_0. Moreover

$$\mathbb{E}[X_{0,i}^H(f) \mid \mathcal{F}_{0,i-1}^H] = 0\ , \tag{21}$$

and

$$\mathbb{E}[\,|X_{0,i}^H(f)|^2 \mid \mathcal{F}_{0,i-1}^H] = \frac{1}{(\rho_0\sqrt{H})^2}\ \mathrm{var}(f_0, \eta_0) = V_{0,0}^H(f)\ , \tag{22}$$

with the convention $\mathcal{F}_{0,0}^H = \{\emptyset, \Omega\}$. Notice that

$$|X_{0,i}^H(f)| \leq \frac{1}{\rho_0\sqrt{H}}\,2\,\|f_0\| \qquad \text{and} \qquad \rho_0 \geq 1\ ,$$

hence for any $\varepsilon > 0$

$$\begin{aligned}
&\mathbb{E}[\,|X_{0,i}^H(f)|^2\ 1_{\{|X_{0,i}^H(f)|\,>\,\varepsilon\}} \mid \mathcal{F}_{0,i-1}^H] \\
&\qquad \leq \frac{1}{\rho_0\,H}\,(2\,\|f_0\|)^2\ 1_{\{\frac{1}{\sqrt{H}}\,2\,\|f_0\|\,>\,\varepsilon\}} = Y_{0,0}^{H,\varepsilon}(f)\ .
\end{aligned} \tag{23}$$

For any $k = 1, \cdots, n$, the random variable

$$\rho_k^H = \frac{\sup_{x\in E} g_k(x)}{\langle \mu_{k-1}^H\, Q_k, g_k\rangle}\ ,$$

is measurable w.r.t. $\mathcal{H}_{k-1}^H = \mathcal{F}_{k,0}^H$, and for any $i = 1, \cdots, N_k^H$, the random variable

$$X_{k,i}^H(f) = \frac{\langle \gamma_{k-1}^H, 1\rangle}{\rho_k^H\sqrt{H}}\ [f_k(\xi_k^i) - \langle \mu_{k-1}^H\, Q_k, f_k\rangle]\ ,$$

is measurable w.r.t. $\mathcal{F}^H_{k,i}$, where, conditionally w.r.t. the σ–algebra $\mathcal{H}^H_{k-1}$ generated by the particle system up to the $(k-1)$–th generation, the random variables $\xi^1_k, \cdots, \xi^i_k, \cdots$ are i.i.d. with common probability distribution $\mu^H_{k-1}\, Q_k$. Moreover

$$\mathbb{E}[X^H_{k,i}(f) \mid \mathcal{F}^H_{k,i-1}] = 0 \ , \tag{24}$$

and

$$\mathbb{E}[\,|X^H_{k,i}(f)|^2 \mid \mathcal{F}^H_{k,i-1}] = \frac{\langle \gamma^H_{k-1}, 1\rangle^2}{(\rho^H_k \sqrt{H})^2}\ \mathrm{var}(f_k, \mu^H_{k-1}\, Q_k) = V^H_{k,0}(f) \ , \tag{25}$$

where the random variable $V^H_{k,0}(f)$ is measurable w.r.t. $\mathcal{H}^H_{k-1} = \mathcal{F}^H_{k,0}$. Notice that

$$|X^H_{k,i}(f)| \leq \frac{\langle \gamma^H_{k-1}, 1\rangle}{\rho^H_k \sqrt{H}}\, 2\, \|f_k\| \qquad \text{and} \qquad \rho^H_k \geq 1 \ ,$$

hence for any $\varepsilon > 0$

$$\begin{aligned} &\mathbb{E}[\,|X^H_{k,i}(f)|^2\, 1_{\{|X^H_{k,i}(f)|\, >\, \varepsilon\}} \mid \mathcal{F}^H_{k,i-1}] \\ &\qquad \leq \frac{\langle \gamma^H_{k-1}, 1\rangle^2}{\rho^H_k\, H}\, (2\, \|f_k\|)^2\, 1_{\{\frac{\langle \gamma^H_{k-1}, 1\rangle}{\sqrt{H}}\, 2\, \|f_k\|\, >\, \varepsilon\}} = Y^{H,\varepsilon}_{k,0}(f) \ , \end{aligned} \tag{26}$$

where the random variable $Y^{H,\varepsilon}_{k,0}(f)$ is measurable w.r.t. $\mathcal{H}^H_{k-1} = \mathcal{F}^H_{k,0}$.

Theorem 8. *For any collection $f = (f_0, f_1, \cdots, f_n)$ of bounded measurable functions*

$$S^H_n(f) = \sum_{k=0}^{n} \sum_{i=1}^{N^H_k} X^H_{k,i}(f) \Longrightarrow \mathcal{N}(0, W_n(f)) \ ,$$

in distribution as $H \uparrow \infty$, with asymptotic variance

$$W_n(f) = \frac{1}{\rho_0}\ \mathrm{var}(f_0, \eta_0) + \sum_{k=1}^{n} \frac{\langle \gamma_{k-1}, 1\rangle^2}{\rho_k}\ \mathrm{var}(f_k, \eta_k) \ .$$

Remark 12. Since the mapping $f \longmapsto S^H_n(f)$ is linear (which incidentally implies that the mapping $f \longmapsto W_n(f)$ is quadratic), the result of Theorem 8 is easily extended to any collection $f = (f_0, f_1, \cdots, f_n)$ of d–dimensional bounded measurable functions, using the Cramér–Wold device, and it follows from the structure of the asymptotic variance that the random variables

$$(\sum_{i=1}^{N^H_0} X^H_{0,i}(f), \cdots, \sum_{i=1}^{N^H_k} X^H_{k,i}(f), \cdots, \sum_{i=1}^{N^H_n} X^H_{n,i}(f)) \ ,$$

are mutually independent, asymptotically as $H \uparrow \infty$.

Proof of Theorem 8. It follows from (21) and (24), from (22) and (25), and from (23) and (26), that the assumptions (16), (17) and (18) of Theorem 7 are satisfied, respectively. It follows from Theorem 3 that $\langle \gamma^H_{k-1}, 1\rangle \to \langle \gamma_{k-1}, 1\rangle$ and $\mathrm{var}(f_k, \mu^H_{k-1}\, Q_k) \to \mathrm{var}(f_k, \eta_k)$ in probability as $H \uparrow \infty$, for any $k = 1, \cdots, n$. Therefore, it follows from Remark 3 that

$$\begin{aligned}
&\sum_{k=0}^{n} N^H_k\, V^H_{k,0}(f) \\
&\quad = \frac{N^H_0}{H\,\rho_0}\,\frac{1}{\rho_0}\,\mathrm{var}(f_0, \eta_0) + \sum_{k=1}^{n} \frac{N^H_k}{H\,\rho^H_k}\,\frac{\langle \gamma^H_{k-1}, 1\rangle^2}{\rho^H_k}\,\mathrm{var}(f_k, \mu^H_{k-1}\, Q_k) \\
&\quad \longrightarrow W_n(f) = \frac{1}{\rho_0}\,\mathrm{var}(f_0, \eta_0) + \sum_{k=1}^{n} \frac{\langle \gamma_{k-1}, 1\rangle^2}{\rho_k}\,\mathrm{var}(f_k, \eta_k)\ ,
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{k=0}^{n} N^H_k\, Y^{H,\varepsilon}_{k,0}(f) \\
&\quad \leq \frac{N^H_0}{\rho_0\, H}\,(2\,\|f_0\|)^2\, 1_{\{\frac{1}{\sqrt{H}}\,2\,\|f_0\| > \varepsilon\}} \\
&\qquad + \sum_{k=1}^{n} \frac{N^H_k}{\rho^H_k\, H}\,\langle \gamma^H_{k-1}, 1\rangle^2\,(2\,\|f_k\|)^2\, 1_{\{\frac{\langle \gamma^H_{k-1}, 1\rangle}{\sqrt{H}}\,2\,\|f_k\| > \varepsilon\}} \\
&\quad \longrightarrow 0\ ,
\end{aligned}$$

in probability as $H \uparrow \infty$, and the proof follows from Theorem 7. □

Proof of Theorem 4. For any bounded measurable function ϕ, the following decomposition holds

$$\begin{aligned}
\langle \gamma^H_n - \gamma_n, \phi\rangle &= \sum_{k=1}^{n} \langle \gamma^H_k - \gamma^H_{k-1}\, R_k, R_{k+1:n}\,\phi\rangle + \langle \gamma^H_0 - \gamma_0, R_{1:n}\,\phi\rangle \\
&= \sum_{k=1}^{n} \langle \gamma^H_{k-1}, 1\rangle\,\langle g_k\,(\eta^H_k - \mu^H_{k-1}\, Q_k), R_{k+1:n}\,\phi\rangle \\
&\qquad + \langle g_0\,(\eta^H_0 - \eta_0), R_{1:n}\,\phi\rangle \\
&= \sum_{k=1}^{n} \langle \gamma^H_{k-1}, 1\rangle\,\langle \eta^H_k - \mu^H_{k-1}\, Q_k, f_k\rangle + \langle \eta^H_0 - \eta_0, f_0\rangle\ ,
\end{aligned}$$

where the collection $f = (f_0, f_1, \cdots, f_n)$ of bounded measurable functions is defined by

$$f_k(x) = g_k(x)\, R_{k+1:n}\, \phi(x) \ ,$$

for any $k = 0, 1, \cdots, n$, with the convention $R_{n+1:n}\, \phi(x) = \phi(x)$, for any $x \in E$. Notice that

$$\begin{aligned} \langle \eta_0^H - \eta_0, f_0 \rangle &= \frac{1}{N_0^H} \sum_{i=1}^{N_0^H} [f_0(\xi_0^i) - \langle \eta_0, f_0 \rangle] \\ &= \frac{1}{\sqrt{H}} \sum_{i=1}^{N_0^H} X_{0,i}^H(f) + (1 - \frac{N_0^H}{H\,\rho_0})\, \langle \eta_0^H - \eta_0, f_0 \rangle \ , \end{aligned}$$

and

$$\begin{aligned} &\langle \gamma_{k-1}^H, 1 \rangle\, \langle \eta_k^H - \mu_{k-1}^H\, Q_k, f_k \rangle \\ &\quad = \frac{\langle \gamma_{k-1}^H, 1 \rangle}{N_k^H} \sum_{i=1}^{N_k^H} [f_k(\xi_k^i) - \langle \mu_{k-1}^H\, Q_k, f_k \rangle] \\ &\quad = \frac{1}{\sqrt{H}} \sum_{i=1}^{N_k^H} X_{k,i}^H(f) + \langle \gamma_{k-1}^H, 1 \rangle\, (1 - \frac{N_k^H}{H\,\rho_k^H})\, \langle \eta_k^H - \mu_{k-1}^H\, Q_k, f_k \rangle \ , \end{aligned}$$

for any $k = 1, \cdots, n$. Taking the sum of both sides for $k = 0, 1, \cdots, n$ yields

$$\sqrt{H}\, \langle \gamma_n^H - \gamma_n, \phi \rangle = S_n^H(f) + \varepsilon_0^H(f) + \sum_{k=1}^{n} \langle \gamma_{k-1}^H, 1 \rangle\, \varepsilon_k^H(f) \ ,$$

where

$$\varepsilon_0^H(f) = \sqrt{H}\, (1 - \frac{N_0^H}{H\,\rho_0})\, \langle \eta_0^H - \eta_0, f_0 \rangle \ ,$$

and where

$$\varepsilon_k^H(f) = \sqrt{H}\, (1 - \frac{N_k^H}{H\,\rho_k^H})\, \langle \eta_k^H - \mu_{k-1}^H\, Q_k, f_k \rangle \ ,$$

for any $k = 1, \cdots, n$. Using the Cauchy–Schwartz inequality, it follows from Lemma 3 and from the rough estimate of Lemma 2 that

$$\begin{aligned} \mathbb{E}|\varepsilon_0^H(f)| &\leq \sqrt{H}\, \{\mathbb{E}|1 - \frac{N_0^H}{H\,\rho_0}|^2\}^{1/2}\, \{\mathbb{E}\,|\langle \eta_0^H - \eta_0, f_0 \rangle|^2\}^{1/2} \\ &\leq \sqrt{H}\, (\omega_H + \frac{1}{H})\, \omega_H\, \|g_0\|\, \|R_{1:n}\, \phi\| \ , \end{aligned}$$

and in view of Remark 8 (i) and (ii)

$$\begin{aligned}\mathbb{E}[\,|\varepsilon_k^H(f)|\mid\mathcal{H}_{k-1}^H] &\leq \sqrt{H}\,\{\mathbb{E}[\,|1-\frac{N_k^H}{H\,\rho_k^H}|^2\mid\mathcal{H}_{k-1}^H]\,\}^{1/2}\\ &\qquad\{\mathbb{E}[\,|\langle\eta_k^H-\mu_{k-1}^H\,Q_k,f_k\rangle|^2\mid\mathcal{H}_{k-1}^H]\,\}^{1/2}\\ &\leq \sqrt{H}\,(\omega_H+\frac{1}{H})\,\omega_H\,\|g_k\|\,\|R_{k+1:n}\,\phi\|\;,\end{aligned}$$

for any $k=1,\cdots,n$, hence

$$\begin{aligned}&\mathbb{E}\,|\sqrt{H}\,\langle\gamma_n^H-\gamma_n,\phi\rangle-S_n^H(f)\,|\\ &\qquad\leq\omega'_H\,[\,\|g_0\|\,\|R_{1:n}\,\phi\|+\sum_{k=1}^n\mathbb{E}[\,\langle\gamma_{k-1}^H,1\rangle\,]\,\|g_k\|\,\|R_{k+1:n}\,\phi\|\,]\;,\end{aligned}$$

where $\omega'_H=\sqrt{H}\,(\omega_H+\frac{1}{H})\,\omega_H$ is of order $1/\sqrt{H}$. It follows from the above discussion that

$$\sqrt{H}\,\langle\gamma_n^H-\gamma_n,\phi\rangle-S_n^H(f)\longrightarrow 0\;,$$

in probability, and it follows from Theorem 7 that $\sqrt{H}\,\langle\gamma_n^H-\gamma_n,\phi\rangle$ converges in distribution as $H\uparrow\infty$ to a Gaussian random variable with zero mean and with variance

$$W_n(f)=\frac{1}{\rho_0}\,\mathrm{var}(g_0\,R_{1,n}\,\phi,\eta_0)+\sum_{k=1}^n\frac{\langle\gamma_{k-1},1\rangle^2}{\rho_k}\,\mathrm{var}(g_k\,R_{k+1,n}\,\phi,\eta_k)\;,$$

which proves (12) with the expression given in (13) for the asymptotic variance. In view of Remark 7, this finishes the proof of Theorem 4. □

A Proof of Theorem 6

Proof of Theorem 6. For any $H>0$, notice that

$$\begin{aligned}Z_H &= \frac{1}{\sqrt{H\rho_H}}\sum_{i=1}^{N_H}\frac{X_i^H}{\sigma_H}\\ &= \frac{\sqrt{r_H}}{\sqrt{H\rho_H}}\,[\,\frac{1}{\sqrt{r_H}}\sum_{i=1}^{r_H}\frac{X_i^H}{\sigma_H}+\frac{1}{\sqrt{r_H}}\,(\sum_{i=1}^{N_H}\frac{X_i^H}{\sigma_H}-\sum_{i=1}^{r_H}\frac{X_i^H}{\sigma_H})\,]\\ &= a_H\,(S_H+O_H)\;,\end{aligned}\tag{27}$$

with

$$a_H=\frac{\sqrt{r_H}}{\sqrt{H\rho_H}}\leq 1\;,\qquad O_H=\frac{1}{\sqrt{r_H}}\,(\sum_{i=1}^{N_H}\frac{X_i^H}{\sigma_H}-\sum_{i=1}^{r_H}\frac{X_i^H}{\sigma_H})\;,$$

and

$$S_H = \frac{1}{\sqrt{r_H}} \sum_{i=1}^{r_H} \frac{X_i^H}{\sigma_H} \ .$$

Notice that

$$0 \le 1 - a_H = 1 - \frac{\sqrt{r_H}}{\sqrt{H\rho_H}} \le 1 - \frac{\sqrt{r_H}}{\sqrt{r_H+1}} \le \frac{1}{2\,r_H+1} = \varepsilon(r_H) \ ,$$

where the straightforward estimate $1 - \dfrac{\sqrt{x}}{\sqrt{x+1}} \le \dfrac{1}{2\,x+1}$ holds for any non-negative real number $x \ge 0$, and

$$\mathbb{E}[\,|S_H| \mid \mathcal{F}^H] \le 1 \ ,$$

using the Cauchy–Schwartz inequality. Therefore, for any fixed real number u

$$\begin{aligned}
&\mathbb{E}[\exp\{i\,\frac{u}{\sqrt{H\rho_H}} \sum_{j=1}^{N_H} \frac{X_j^H}{\sigma_H}\} \mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,u^2\} \\
&\quad = \mathbb{E}[\exp\{i\,u\,Z_H\} \mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,u^2\} \\
&\quad = \mathbb{E}[\exp\{i\,u\,Z_H\} - \exp\{i\,u\,S_H\} \mid \mathcal{F}^H] \\
&\qquad + \mathbb{E}[\exp\{i\,\frac{u}{\sqrt{r_H}} \sum_{j=1}^{r_H} \frac{X_j^H}{\sigma_H}\} \mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,u^2\} \ .
\end{aligned}$$

The last term can be controlled easily, using classical estimates in the central limit theorem for sums of i.i.d. random variables. Moreover, for any $B > 0$ and any $0 < d < 1$

$$\begin{aligned}
&\mathbb{E}[\,|\exp\{i\,u\,Z_H\} - \exp\{i\,u\,S_H\}| \mid \mathcal{F}^H] \\
&\quad \le \mathbb{E}[\,1_{\{|O_H| \le B\}}\,|\exp\{i\,u\,Z_H\} - \exp\{i\,u\,S_H\}| \mid \mathcal{F}^H] \\
&\qquad + \mathbb{E}[\,1_{\{|O_H| > B\}}\,|\exp\{i\,u\,Z_H\} - \exp\{i\,u\,S_H\}| \mid \mathcal{F}^H] \\
&\quad \le |u|\,((1-a_H)\,\mathbb{E}[\,|S_H| \mid \mathcal{F}^H] + a_H\,B) + 2\,\mathbb{P}[\,|O_H| > B \mid \mathcal{F}^H] \\
&\quad \le |u|\,(\varepsilon(r_H) + B) + 2\,\mathbb{P}[\,|O_H| > B\,,\,|\frac{N_H}{H\rho_H} - 1| \le d \mid \mathcal{F}^H] \\
&\qquad + 2\,\mathbb{P}[\,|\frac{N_H}{H\rho_H} - 1| > d \mid \mathcal{F}^H] \ ,
\end{aligned}$$

and the second term can be controlled using the Kolmogorov maximal inequality, along the lines of the proof of [8, Theorem I.3.1].

► For any fixed real number u

$$\mathbb{E}[\exp\{i\,u\,S_H\} \mid \mathcal{F}^H] = \mathbb{E}[\exp\{i\,\frac{u}{\sqrt{r_H}}\sum_{j=1}^{r_H}\frac{X_j^H}{\sigma_H}\} \mid \mathcal{F}^H] = (\Phi_N(u))^{r_H}\ ,$$

by independence, where

$$\Phi_H(u) = \mathbb{E}[\exp\{i\,\frac{u}{\sqrt{r_H}}\,\frac{X_j^H}{\sigma_H}\} \mid \mathcal{F}^H]\ ,$$

does not depend on $j = 1, \cdots, r_H$, and it follows from [13, Lemma C.3] that

$$|\Phi_H(u) - (1 - \tfrac{1}{2}\,\frac{u^2}{r_H})| \le \tfrac{1}{6}\,c\,\frac{|u|^3}{r_H} + R_H(c)\,\frac{u^2}{r_H}\ ,$$

for any $c > 0$. Using the straightforward estimate $|x^r - y^r| \le r\,|x - y|$, which holds for any integer r and for any complex numbers x, y such that $|x| \le 1$ and $|y| \le 1$, yields

$$|(\Phi_H(u))^{r_H} - (1 - \tfrac{1}{2}\,\frac{u^2}{r_H})^{r_H}\,| \le r_H\,|\Phi_H(u) - (1 - \tfrac{1}{2}\,\frac{u^2}{r_H})\,| \le \tfrac{1}{6}\,c\,|u|^3 + R_H(c)\,u^2\ .$$

Using the same estimate again and the straightforward estimate $|e^{-x} - (1 - x)| \le \frac{1}{2}\,x^2$, which holds for any nonnegative real number $x \ge 0$, yields

$$|\exp\{-\tfrac{1}{2}\,u^2\} - (1 - \tfrac{1}{2}\,\frac{u^2}{r_H})^{r_H}\,| \le r_H\,|\exp\{-\tfrac{1}{2}\,\frac{u^2}{r_H}\} - (1 - \tfrac{1}{2}\,\frac{u^2}{r_H})\,| \le \tfrac{1}{8}\,\frac{u^4}{r_H}\ .$$

Combining the above estimates together and using the triangle inequality, yields

$$|\mathbb{E}[\exp\{i\,u\,S_H\} \mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,u^2\}| \le \tfrac{1}{6}\,c\,|u|^3 + R_H(c)\,u^2 + \tfrac{1}{8}\,\frac{u^4}{r_H}\ , \qquad (28)$$

and on the *good set* $\{r_H > r\}$

$$|\mathbb{E}[\exp\{i\,u\,S_H\} \mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,u^2\}| \le \tfrac{1}{6}\,c\,|u|^3 + R_H(c)\,u^2 + \tfrac{1}{8}\,\frac{u^4}{r}\ .$$

► Notice that

$$|N_H - r_H| \le |N_H - H\rho_H| + 1 \le |\frac{N_H}{H\rho_H} - 1|\ H\rho_H + 1\ ,$$

hence if $|\dfrac{N_H}{H\rho_H} - 1| \le d$, then $|N_H - r_H| \le d\,(r_H + 1) + 1 = d_H$, and either $r_H - \lceil d_H \rceil \le N_H \le r_H$ or $r_H \le N_H \le r_H + \lceil d_H \rceil$. Therefore, for any $B > 0$ and any $0 < d < 1$

$$\mathbb{P}[\,|O_H| > B\,,\,|\frac{N_H}{H\rho_H} - 1| \le d \mid \mathcal{F}^H]$$

$$\le \mathbb{P}[\,|\sum_{i=1}^{N_H} \frac{X_i^H}{\sigma_H} - \sum_{i=1}^{r_H} \frac{X_i^H}{\sigma_H}| > B\sqrt{r_H}\,,\,|N_H - r_H| \le \lceil d_H \rceil \mid \mathcal{F}^H]$$

$$\le \mathbb{P}[\max_{r_H - \lceil d_H \rceil \le N \le r_H} |\sum_{i=1}^{r_H} \frac{X_i^H}{\sigma_H} - \sum_{i=1}^{N} \frac{X_i^H}{\sigma_H}| > B\sqrt{r_H} \mid \mathcal{F}^H]$$

$$+ \mathbb{P}[\max_{r_H \le N \le r_H + \lceil d_H \rceil} |\sum_{i=1}^{N} \frac{X_i^H}{\sigma_H} - \sum_{i=1}^{r_H} \frac{X_i^H}{\sigma_H}| > B\sqrt{r_H} \mid \mathcal{F}^H]$$

$$\le 2\,\frac{1}{B^2\,r_H}\,\lceil d_H \rceil$$

$$\le 2\,\frac{d\,(r_H + 1) + 2}{B^2\,r_H}\ ,$$

using the Kolmogorov maximal inequality, and on the *good set* $\{r_H > r\} \in \mathcal{F}^H$

$$\mathbb{P}[\,|O_H| > B\,,\,|\frac{N_H}{H\rho_H} - 1| \le d \mid \mathcal{F}^H] \le 2\,\frac{d\,(r+1) + 2}{B^2\,r}\ .$$

► Combining the above estimates, using the triangle inequality and taking $B = d^{1/3}$, yields

$$|\,\mathbb{E}[\exp\{i\,\frac{u}{\sqrt{H\rho_H}}\,\sum_{j=1}^{N_H} \frac{X_j^H}{\sigma_H}\} \mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,u^2\}\,|$$

$$\le 2 \cdot 1_{\{r_H\,\le\,r\}} + R_H(c)\,u^2 + 2\,F_H(d)$$

$$+ \tfrac{1}{6}\,c\,|u|^3 + \tfrac{1}{8}\,\frac{u^4}{r} + |u|\,(\varepsilon(r) + d^{1/3}) + 4\,\frac{d\,(r+1) + 2}{d^{2/3}\,r}\ .$$

Taking d so that $d \to 0$ and $d^{2/3}\,r \to \infty$ when $r \uparrow \infty$, it is possible for any $a > 0$, to find $r > 0$ large enough, $c > 0$ small enough, such that

$$\tfrac{1}{6}\,c\,|u|^3 + \tfrac{1}{8}\,\frac{u^4}{r} + |u|\,(\varepsilon(r) + d^{1/3}) + 4\,\frac{d\,(r+1) + 2}{d^{2/3}\,r} < \tfrac{1}{2}\,a\ ,$$

in which case

$$\mathbb{P}[\,|\,\mathbb{E}[\exp\{i\,\frac{u}{\sqrt{H\rho_H}}\,\sum_{j=1}^{N_H} \frac{X_j^H}{\sigma_H}\} \mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,u^2\}\,| > a]$$

$$\le \mathbb{P}[\,2 \cdot 1_{\{r_H\,\le\,r\}} + R_H(c)\,u^2 + 2\,F_H(d) > \tfrac{1}{2}\,a]\ ,$$

which goes to zero as $H \uparrow \infty$: this terminates the proof of (14).

To prove (15), notice that

$$\sqrt{H}\,\frac{1}{N_H}\sum_{j=1}^{N_H} X_j^H = \frac{H\rho_H}{N_H}\,\frac{\sigma_H}{\sqrt{\rho_H}}\,\frac{1}{\sqrt{H\rho_H}}\sum_{j=1}^{N_H}\frac{X_j^H}{\sigma_H} = c_H\,Z_H\ ,$$

where

$$c_H = \frac{H\rho_H}{N_H}\,\frac{\sigma_H}{\sqrt{\rho_H}} \qquad \text{and} \qquad c = \frac{\sigma}{\sqrt{\rho}}\ ,$$

and where Z_H is defined in (27), hence

$$\begin{aligned}
&\mathbb{E}[\exp\{i\,u\sqrt{H}\,\frac{1}{N_H}\sum_{j=1}^{N_H} X_j^H\}\mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,\frac{u^2\,\sigma^2}{\rho}\}\\
&\quad = \mathbb{E}[\exp\{i\,u\,c_H\,Z_H\}\,] - \exp\{-\tfrac{1}{2}\,u^2\,c^2\}\\
&\quad = \mathbb{E}[\exp\{i\,u\,c_H\,Z_H\} - \exp\{i\,u\,c\,Z_H\}\mid \mathcal{F}^H]\\
&\qquad + \mathbb{E}[\exp\{i\,\frac{u\,c}{\sqrt{H\,\rho_H}}\sum_{j=1}^{N_H}\frac{X_j^H}{\sigma_H}\}\mid \mathcal{F}^H] - \exp\{-\tfrac{1}{2}\,u^2\,c^2\}\ .
\end{aligned}$$

The last term goes to zero in $\mathbb{L}^1$ as $H \uparrow \infty$, using (14). Moreover, for any $b > 0$ and any $0 < d < 1$

$$\begin{aligned}
&\mathbb{E}|\,\mathbb{E}[\exp\{i\,u\,c_H\,Z_H\} - \exp\{i\,u\,c\,Z_H\}\mid \mathcal{F}^H]\,|\\
&\quad \le \mathbb{E}|\,\exp\{i\,u\,c_H\,Z_H\} - \exp\{i\,u\,c\,Z_H\}\,|\\
&\quad \le \mathbb{E}[\,1_{\{|\frac{\sigma_H}{\sqrt{\rho_H}} - \frac{\sigma}{\sqrt{\rho}}| \le b\,,\ |\frac{N_H}{H\rho_H} - 1| \le d\}}\\
&\qquad\qquad |\,\exp\{i\,u\,c_H\,Z_H\} - \exp\{i\,u\,c\,Z_H\}\,|\,]\\
&\qquad + 2\,\mathbb{P}[\,|\frac{N_H}{H\rho_H} - 1| > d] + 2\,\mathbb{P}[\,|\frac{\sigma_H}{\sqrt{\rho_H}} - \frac{\sigma}{\sqrt{\rho}}| > b]\ .
\end{aligned}$$

The last two terms go to zero as $H \uparrow \infty$, by assumption and in view of Remark 11. Next, if $|\frac{N_H}{H\rho_H} - 1| \le d$, then clearly $|\frac{H\rho_H}{N_H} - 1| \le \frac{d}{1-d}$, and since

$$c_H - c = \frac{H\rho_H}{N_H}\,\frac{\sigma_H}{\sqrt{\rho_H}} - \frac{\sigma}{\sqrt{\rho}} = (\frac{H\rho_H}{N_H} - 1)\,\frac{\sigma_H}{\sqrt{\rho_H}} + (\frac{\sigma_H}{\sqrt{\rho_H}} - \frac{\sigma}{\sqrt{\rho}})\ ,$$

then it holds

$$1_{\{|\frac{\sigma_H}{\sqrt{\rho_H}} - \frac{\sigma}{\sqrt{\rho}}| \le b\,,\, |\frac{N_H}{H\rho_H} - 1| \le d\}}\, |c_H - c|$$

$$\le \frac{d}{1-d}\,(\frac{\sigma}{\sqrt{\rho}} + b) + b = \frac{d\,\frac{\sigma}{\sqrt{\rho}} + b}{1-d}\ .$$

Therefore, using the straightforward estimate $|e^{i\,x} - e^{i\,x'}| \le \min(|x - x'|, 2)$ which holds for any real numbers x, x', yields

$$\mathbb{E}[\,1_{\{|\frac{\sigma_H}{\sqrt{\rho_H}} - \frac{\sigma}{\sqrt{\rho}}| \le b\,,\, |\frac{N_H}{H\rho_H} - 1| \le d\}}$$

$$|\,\exp\{i\,u\,c_H\,Z_H\} - \exp\{i\,u\,c\,Z_H\}\,|\,]$$

$$\le \mathbb{E}[\,1_{\{|\frac{\sigma_H}{\sqrt{\rho_H}} - \frac{\sigma}{\sqrt{\rho}}| \le b\,,\, |\frac{N_H}{H\rho_H} - 1| \le d\}}\, \min(|u|\,|c_H - c|\,|Z_H|, 2)\,]$$

$$\le \mathbb{E}\,\min(|u|\,\frac{d\,\frac{\sigma}{\sqrt{\rho}} + b}{1-d}\,|Z_H|, 2)\ ,$$

hence using (14) yields

$$\limsup_{H\uparrow\infty}\, \mathbb{E}|\,\mathbb{E}[\,\exp\{i\,u\,c_H\,Z_H\} - \exp\{i\,u\,c\,Z_H\}\mid \mathcal{F}^H]\,|$$

$$\le \limsup_{H\uparrow\infty}\, \mathbb{E}[\,1_{\{|\frac{\sigma_H}{\sqrt{\rho_H}} - \frac{\sigma}{\sqrt{\rho}}| \le b\,,\, |\frac{N_H}{H\rho_H} - 1| \le d\}}$$

$$|\,\exp\{i\,u\,c_H\,Z_H\} - \exp\{i\,u\,c\,Z_H\}\,|\,]$$

$$\le \mathbb{E}\,\min(|u|\,\frac{d\,\frac{\sigma}{\sqrt{\rho}} + b}{1-d}\,|Z|, 2)\ ,$$

in view of Remark 10, where Z is a standard Gaussian r.v. (with zero mean and unit variance). Finally, using the Lebesgue dominated convergence theorem, it follows that

$$\mathbb{E}[\,\exp\{i\,u\,c_H\,Z_H\} - \exp\{i\,u\,c\,Z_H\}\mid \mathcal{F}^H] \longrightarrow 0\ ,$$

in $\mathbb{L}^1$ as $H \uparrow \infty$, since $b > 0$ and $0 < d < 1$ are arbitrary : this terminates the proof of (15). □

B Proof of Theorem 7

By definition, S_n^H is written as a double sum over generations with different random sizes : the idea is to rewrite this as a single sum across all generations, and to use a central limit theorem for triangular arrays of martingale increments [2, Theorem 2.8.42]. Notice that the i–th particle within the k–th generation can be associated in a unique way with an integer p between 1 and $\sigma_n^H = N_0^H + \cdots + N_n^H$: clearly $p_{k,i} = \sigma_{k-1}^H + i$, and conversely the random integers k_p and i_p are defined by

$$k_p = \inf\{k \geq 0 \,:\, \sigma_k^H \geq p\} \qquad \text{and} \qquad i_p = p - \sigma_{k_p-1}^H \ ,$$

with the convention $\sigma_{-1}^H = 0$, or in other words $k_p = k$ and $i_p = i$ if and only if

$$\sigma_{k-1}^H + 1 \leq p = \sigma_{k-1}^H + i \leq \sigma_k^H \ ,$$

with $1 \leq i \leq N_k^H$, see Figure 1.

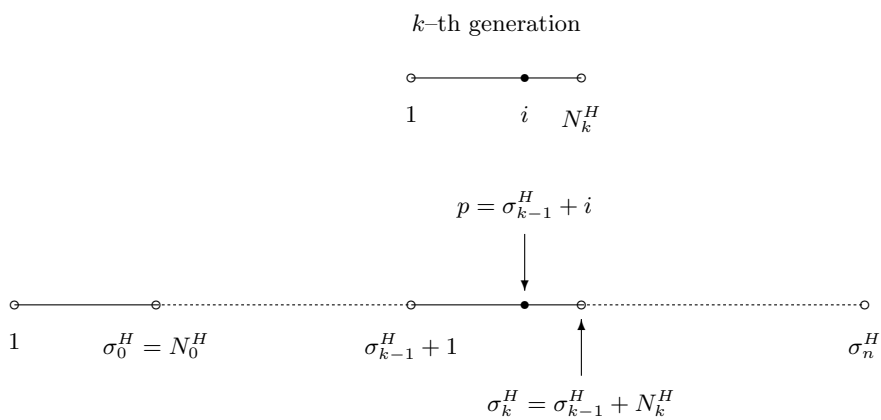

Fig. 1. The i–th particle within the k–th generation (above), seen as the p–th particle across all generations (below)

For any $k = 0, 1, \cdots, n$ and any integer $i \geq 1$

$$\{k_p = k, i_p = i\} = \{p = \sigma_{k-1}^H + i, i \leq N_k^H\} \in \mathcal{F}_{k,i-1}^H \subset \mathcal{F}_{k,i}^H \ ,$$

since $\{p = \sigma_{k-1}^H + i\} \in \mathcal{H}_{k-1}^H$ and $\{i \leq N_k^H\} \in \mathcal{F}_{k,i-1}^H$, which allows to define the σ–algebra $\mathcal{G}_p^H = \mathcal{F}_{k_p,i_p}^H$ in the usual way : by definition, $A \in \mathcal{G}_p^H$ if and only if $A \cap \{k_p = k, i_p = i\} \in \mathcal{F}_{k,i}^H$ for any $k = 0, 1, \cdots, n$ and any integer $i \geq 1$. Using this new labeling of the particle system yields

$$S_n^H = \sum_{k=0}^{n} \sum_{i=1}^{N_k^H} X_{k,i}^H = \sum_{p=1}^{\sigma_n^H} U_p^H \ ,$$

where the time changed random variable $U_p^H = X_{k_p,i_p}^H$ is measurable w.r.t. $\mathcal{G}_p^H$, for any $p = 1, \cdots, \sigma_n^H$: indeed for any Borel subset B, any $k = 0, 1, \cdots, n$ and any integer $i \geq 1$

$$\{U_p^H \in B\} \cap \{k_p = k, i_p = i\} = \{X_{k,i}^H \in B\} \cap \{k_p = k, i_p = i\} \ ,$$

hence $\{U_p^H \in B\} \in \mathcal{G}_p^H$, since $\{X_{k,i}^H \in B\} \in \mathcal{F}_{k,i}^H$ and $\{k_p = k, i_p = i\} \in \mathcal{F}_{k,i-1}^H$. Moreover, the random variable σ_n^H is a stopping time w.r.t. $\mathcal{G}^H = \{\mathcal{G}_p^H \ , p \geq 1\}$: indeed, for any integer $p \geq 1$, any $k = 0, 1, \cdots, n$ and any integer $i \geq 1$

$$\begin{aligned} \{\sigma_n^H = p\} \cap \{k_p = k, i_p = i\} &= \{\sigma_n^H = p\} \cap \{p = \sigma_{k-1}^H + i, 1 \leq i \leq N_k^H\} \\ &= \begin{cases} \emptyset \ , & \text{if } k \neq n, \\ \{p = \sigma_{k-1}^H + i\} \cap \{N_k^H = i\} \ , & \text{if } k = n, \end{cases} \end{aligned}$$

hence $\{\sigma_n^H = p\} \in \mathcal{G}_p^H$ since $\{p = \sigma_{k-1}^H + i\} \in \mathcal{H}_{k-1}^H$ and $\{N_k^H = i\} \in \mathcal{F}_{k,i}^H$.

Proof of Theorem 7. To apply Theorem 2.8.42 in [2] to

$$S_n^H = \sum_{p=1}^{\sigma_n^H} U_p^H \ ,$$

where σ_n^H is a stopping time w.r.t. $\mathcal{G}^H = \{\mathcal{G}_p^H \ , p \geq 1\}$, and where the random variable U_p^H is measurable w.r.t. $\mathcal{G}_p^H$ for any $p = 1, \cdots, \sigma_n^H$, the following three conditions have to be checked : a martingale increment property, the convergence of conditional variances, and a conditional Lindeberg condition. These three conditions follow immediately from (19) and from

$$\sum_{p=1}^{\sigma_n^H} \mathbb{E}[U_p^H \mid \mathcal{G}_{p-1}^H] = 0 \ ,$$

$$\sum_{p=1}^{\sigma_n^H} \mathbb{E}[\,|U_p^H|^2 \mid \mathcal{G}_{p-1}^H] = \sum_{k=0}^{n} N_k^H \, V_{k,0}^H \ ,$$

and

$$\sum_{p=1}^{\sigma_n^H} \mathbb{E}[\,|U_p^H|^2 \, 1_{\{|U_p^H| > \varepsilon\}} \mid \mathcal{G}_{p-1}^H] = \sum_{k=0}^{n} N_k^H \, Y_{k,0}^{H,\varepsilon} \ ,$$

which follow from (16), (17) and (18) respectively, using properties of the past σ–algebra $\mathcal{G}_{p-1}^H$, see [13, Lemmas B.1 and B.2], and using the preservation of

the martingale property under time change, see Lemma 4 below. Indeed, it follows from Lemma 4 and from (16), that

$$\mathbb{E}[U_p^H \mid \mathcal{G}_{p-1}^H] = 0 \ ,$$

for any integer $p \geq 1$, hence

$$\sum_{p=1}^{\sigma_n^H} \mathbb{E}[U_p^H \mid \mathcal{G}_{p-1}^H] = 0 \ .$$

Similarly, it follows from Corollary 1 and from (17), that

$$\mathbb{E}[\,|U_p^H|^2 \mid \mathcal{G}_{p-1}^H] = V_{k_p,0}^H \ ,$$

for any integer $p \geq 1$, hence

$$\sum_{p=1}^{\sigma_n^H} \mathbb{E}[\,|U_p^H|^2 \mid \mathcal{G}_{p-1}^H]$$

$$= \sum_{p=1}^{\sigma_n^H} V_{k_p,0}^H = \sum_{k=0}^{n} \sum_{p=\sigma_{k-1}^H+1}^{\sigma_k^H} V_{k_p,0}^H = \sum_{k=0}^{n} N_k^H \, V_{k,0}^H \ ,$$

since $k_p = k$ if $\sigma_{k-1}^H + 1 \leq p \leq \sigma_k^H$. Finally, it follows from Corollary 2 and from (18), that

$$\mathbb{E}[\,|U_p^H|^2 \, 1_{\{|U_p^H| > \varepsilon\}} \mid \mathcal{G}_{p-1}^H] \leq Y_{k_p,0}^{H,\varepsilon} \ ,$$

for any integer $p \geq 1$ and any $\varepsilon > 0$, hence

$$\sum_{p=1}^{\sigma_n^H} \mathbb{E}[\,|U_p^H|^2 \, 1_{\{|U_p^H| > \varepsilon\}} \mid \mathcal{G}_{p-1}^H]$$

$$\leq \sum_{p=1}^{\sigma_n^H} Y_{k_p,0}^{H,\varepsilon} = \sum_{k=0}^{n} \sum_{p=\sigma_{k-1}^H+1}^{\sigma_k^H} Y_{k_p,0}^{H,\varepsilon} = \sum_{k=0}^{n} N_k^H \, Y_{k,0}^{H,\varepsilon} \ ,$$

since $k_p = k$ if $\sigma_{k-1}^H + 1 \leq p \leq \sigma_k^H$. □

Lemma 4. *If for any $k = 0, 1, \cdots, n$ and any integer $i \geq 1$*

$$\mathbb{E}[F_{k,i} \mid \mathcal{F}_{k,i-1}^H] = 0 \ ,$$

then for any integer $p \geq 1$, the time changed random variable $G_p = F_{k_p,i_p}$ satisfies

$$\mathbb{E}[G_p \mid \mathcal{G}_{p-1}^H] = 0 \ .$$

Corollary 1. *If for any $k = 0, 1, \cdots, n$ and any integer $i \geq 1$*

$$\mathbb{E}[F_{k,i} \mid \mathcal{F}^H_{k,i-1}] = \widehat{F}^H_k \ ,$$

where the random variable $\widehat{F}^H_k$ is measurable w.r.t. $\mathcal{F}^H_{k,0}$, then for any integer $p \geq 1$, the time changed random variables $G_p = F_{k_p,i_p}$ and $\widehat{G}^H_p = \widehat{F}^H_{k_p}$ satisfy

$$\mathbb{E}[G_p \mid \mathcal{G}^H_{p-1}] = \widehat{G}^H_p \ .$$

Proof. Notice that

$$\mathbb{E}[F_{k,i} - \widehat{F}^H_k \mid \mathcal{F}^H_{k,i-1}] = 0 \ ,$$

under the assumption, and it follows from Lemma 4 above that

$$\mathbb{E}[G_p - \widehat{G}^H_p \mid \mathcal{G}^H_{p-1}] = 0 \ ,$$

or equivalently

$$\mathbb{E}[G_p \mid \mathcal{G}^H_{p-1}] = \widehat{G}^H_p \ ,$$

since $\widehat{G}^H_p$ is measurable w.r.t. $\mathcal{G}^H_{p-1}$, in view of [13, Lemma B.2]. □

Corollary 2. *If for any $k = 0, 1, \cdots, n$ and any integer $i \geq 1$*

$$F_{k,i} \leq F^*_k \ ,$$

where the random variable F^*_k is measurable w.r.t. $\mathcal{F}^H_{k,0}$, then for any integer $p \geq 1$, the time changed random variables $G_p = F_{k_p,i_p}$ and $G^*_p = F^*_{k_p}$ satisfy

$$\mathbb{E}[G_p \mid \mathcal{G}^H_{p-1}] \leq G^*_p \ .$$

Proof. Notice that

$$\mathbb{E}[\max(F_{k,i} - F^*_k, 0) \mid \mathcal{F}^H_{k,i-1}] = 0 \ ,$$

since $\max(F_{k,i} - F^*_k, 0) = 0$ under the assumption, and it follows from Lemma 4 above that

$$\mathbb{E}[\max(G_p - G^*_p, 0) \mid \mathcal{G}^H_{p-1}] = 0 \ ,$$

hence using the Jensen inequality yields

$$\max(\mathbb{E}[G_p - G^*_p \mid \mathcal{G}^H_{p-1}], 0) = 0 \qquad \text{i.e.} \qquad \mathbb{E}[G_p - G^*_p \mid \mathcal{G}^H_{p-1}] \leq 0 \ ,$$

or equivalently

$$\mathbb{E}[G_p \mid \mathcal{G}^H_{p-1}] \leq G^*_p \ ,$$

since G^*_p is measurable w.r.t. $\mathcal{G}^H_{p-1}$, in view of [13, Lemma B.2]. □

Proof of Lemma 4. First, recall the following identity

$$\sum_{i=1}^{N\wedge M} 1\{I=i\} = \sum_{i=1}^{M} 1\{N\geq i\}\, 1\{I=i\}\ ,$$

which is easily obtained using summation by parts. For any $A \in \mathcal{G}^H_{p-1}$, and any integer $M \geq 1$

$$\begin{aligned}
&\mathbb{E}[G_p\, 1_A \sum_{k=0}^{n} 1\{k_p=k\} \sum_{i=1}^{N_k^H\wedge M} 1\{i_p=i\}] \\
&\quad = \mathbb{E}[G_p\, 1_A \sum_{k=0}^{n} 1\{k_p=k\} \sum_{i=1}^{M} 1\{N_k^H\geq i\}\, 1\{i_p=i\}] \\
&\quad = \sum_{k=0}^{n}\sum_{i=1}^{M} \mathbb{E}[G_p\, 1_A\cap\{k_p=k, i_p=i\}\, 1\{N_k^H\geq i\}] \\
&\quad = \sum_{k=0}^{n}\sum_{i=1}^{M} \mathbb{E}[F_{k,i}\, 1_A\cap\{k_p=k, i_p=i\}\, 1\{N_k^H\geq i\}]\ .
\end{aligned}$$

Notice that $A\cap\{k_p=k, i_p=i\} \in \mathcal{F}^H_{k,i-1}$ in view of [13, Lemma B.1] and $\{N_k^H\geq i\} \in \mathcal{F}^H_{k,i-1}$, hence

$$\begin{aligned}
&\mathbb{E}[F_{k,i}\, 1_A\cap\{k_p=k, i_p=i\}\, 1\{N_k^H\geq i\}] \\
&\quad = \mathbb{E}[\,\mathbb{E}[F_{k,i}\mid \mathcal{F}^H_{k,i-1}]\, 1_A\cap\{k_p=k, i_p=i\}\, 1\{N_k^H\geq i\}] = 0\ ,
\end{aligned}$$

under the assumption, and

$$\mathbb{E}[G_p\, 1_A \sum_{k=0}^{n} 1\{k_p=k\} \sum_{i=1}^{N_k^H\wedge M} 1\{i_p=i\}] = 0\ ,$$

or equivalently

$$\begin{aligned}
&\mathbb{E}[\,G_p^+\, 1_A \sum_{k=0}^{n} 1\{k_p=k\} \sum_{i=1}^{N_k^H\wedge M} 1\{i_p=i\}] \\
&\quad = \mathbb{E}[\,G_p^-\, 1_A \sum_{k=0}^{n} 1\{k_p=k\} \sum_{i=1}^{N_k^H\wedge M} 1\{i_p=i\}]\ .
\end{aligned}$$

where $G_p^+ = \max(G_p, 0)$ and $G_p^- = \max(-G_p, 0)$. Finally, using the monotone convergence theorem yields

$$\begin{aligned}
\mathbb{E}[G_p^+ \, 1_A] &= \mathbb{E}[G_p^+ \, 1_A \sum_{k=0}^{n} 1_{\{k_p = k\}} \sum_{i=1}^{N_k^H} 1_{\{i_p = i\}}] \\
&= \lim_{M \uparrow \infty} \mathbb{E}[G_p^+ \, 1_A \sum_{k=0}^{n} 1_{\{k_p = k\}} \sum_{i=1}^{N_k^H \wedge M} 1_{\{i_p = i\}}] \\
&= \lim_{M \uparrow \infty} \mathbb{E}[G_p^- \, 1_A \sum_{k=0}^{n} 1_{\{k_p = k\}} \sum_{i=1}^{N_k^H \wedge M} 1_{\{i_p = i\}}] \\
&= \mathbb{E}[G_p^- \, 1_A \sum_{k=0}^{n} 1_{\{k_p = k\}} \sum_{i=1}^{N_k^H} 1_{\{i_p = i\}}] = \mathbb{E}[G_p^- \, 1_A] \, ,
\end{aligned}$$

or equivalently $\mathbb{E}[G_p \, 1_A] = 0$, hence $\mathbb{E}[G_p \mid \mathcal{G}_{p-1}^H] = 0$, since $A \in \mathcal{G}_{p-1}^H$ is arbitrary. □

Acknowledgments

The first author gratefully thanks Natacha Caylus for her careful reading of an earlier version of this work, and for suggesting the proof given in Appendix A for the second part of Theorem 6. We both thank Pierre Del Moral for his warm support, and for suggesting the alternate approach, based on a central limit theorem for triangular arrays of martingale increments, which is followed in Section 6 for the proof of Theorem 4.

References

1. Frédéric Cérou, Pierre Del Moral, François Le Gland, and Pascal Lezaud. Limit theorems for the multilevel splitting algorithm in the simulation of rare events. In Michael E. Kuhl, Natalie M. Steiger, Frank B. Armstrong, and Jeffrey A. Joines, editors, *Proceedings of the 2005 Winter Simulation Conference, Orlando 2005*, December 2005.
2. Didier Dacunha-Castelle and Marie Duflo. *Probability and Statistics II*. Springer–Verlag, Berlin, 1986.
3. Pierre Del Moral. *Feynman–Kac Formulae. Genealogical and Interacting Particle Systems with Applications*. Probability and its Applications. Springer–Verlag, New York, 2004.
4. Pierre Del Moral and Jean Jacod. Interacting particle filtering with discrete observations. In Arnaud Doucet, Nando de Freitas, and Neil Gordon, editors, *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science, chapter 3, pages 43–75. Springer–Verlag, New York, 2001.
5. Pierre Del Moral, Jean Jacod, and Philip Protter. The Monte Carlo method for filtering with discrete–time observations. *Probability Theory and Related Fields*, 120(3):346–368, 2001.

6. Pierre Del Moral and Pascal Lezaud. Branching and interacting particle interpretation of rare event probabilities. In Henk Blom and John Lygeros, editors, *Stochastic Hybrid Systems : Theory and Safety Critical Applications*, Lecture Notes in Control and Information Sciences, chapter 14. Springer–Verlag, Berlin. To appear.
7. Pierre Del Moral and Laurent Miclo. Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to nonlinear filtering. In Jacques Azéma, Michel Émery, Michel Ledoux, and Marc Yor, editors, *Séminaire de Probabilités XXXIV*, volume 1729 of *Lecture Notes in Mathematics*, pages 1–145. Springer–Verlag, Berlin, 2000.
8. Allan Gut. *Stopped Random Walks. Limit Theorems and Applications.* Applied Probability (A Series of the Applied Probability Trust). Springer–Verlag, New York, 1988.
9. Hans R. Künsch. Recursive Monte Carlo filters : Algorithms and theoretical analysis. *The Annals of Statistics*, 33(5):1983–2021, October 2005.
10. François Le Gland and Nadia Oudjane. A robustification approach to stability and to uniform particle approximation of nonlinear filters : the example of pseudo-mixing signals. *Stochastic Processes and their Applications*, 106(2):279–316, August 2003.
11. François Le Gland and Nadia Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric, and application to particle filters. *The Annals of Applied Probability*, 14(1):144–187, February 2004.
12. François Le Gland and Nadia Oudjane. A sequential particle algorithm that keeps the particle system alive. In *Proceedings of the 13th European Signal Processing Conference, Antalya 2005.* EURASIP, September 2005.
13. François Le Gland and Nadia Oudjane. A sequential algorithm that keeps the particle system alive. Rapport de Recherche 5826, INRIA, February 2006. `ftp://ftp.inria.fr/INRIA/publication/publi-pdf/RR/RR-5826.pdf`.
14. Jacques Neveu. *Discrete–Parameter Martingales*, volume 10 of *North–Holland Mathematical Library.* North–Holland, Amsterdam, 1975.
15. Nadia Oudjane. *Stabilité et Approximations Particulaires en Filtrage Non–Linéaire — Application au Pistage.* Thèse de Doctorat, Université de Rennes 1, Rennes, December 2000. `ftp://ftp.irisa.fr/techreports/theses/2000/oudjane.pdf`.
16. Nadia Oudjane and Sylvain Rubenthaler. Stability and uniform particle approximation of nonlinear filters in case of non ergodic signal. *Stochastic Analysis and Applications*, 23(3):421–448, May 2005.
17. Alfréd Rényi. On the asymptotic distribution of the sum of a random number of independent random variables. *Acta Mathematica Academiae Scientiarum Hungaricae*, 8:193–199, 1957.
18. Vivien Rossi. *Filtrage Non–Linéaire par Noyaux de Convolution — Application à un Procédé de Dépollution Biologique.* Thèse de Doctorat, École Nationale Supérieure Agronomique de Montpellier, Montpellier, December 2004.
19. Vivien Rossi and Jean-Pierre Vila. Nonlinear filtering in discrete time : A particle convolution approach. *Statistical Inference for Stochastic Processes.* To appear.
20. David Siegmund. *Sequential Analysis. Tests and Confidence Intervals.* Springer Series in Statistics. Springer–Verlag, New York, 1985.