Central limit theorem for the stratified resampling mechanism

Roberta Flenghi, Benjamin Jourdain

Central limit theorem for the stratified resampling mechanism

Roberta Flenghi[1] roberta.flenghi@enpc.fr

Benjamin Jourdain[1] benjamin.jourdain@enpc.fr

[1]Cermics, École des Ponts, INRIA, Marne-la-Vallée, France.

Abstract

The stratified resampling mechanism is one of the resampling schemes commonly used in the resampling steps of particle filters. In the present paper, we prove a central limit theorem for this mechanism under the assumption that the initial positions are independent and identically distributed and the weights proportional to a positive function of the positions such that the image of their common distribution by this function has a non zero component absolutely continuous with respect to the Lebesgue measure. This result relies on the convergence in distribution of the fractional part of partial sums of the normalized weights to some random variable uniformly distributed on $[0, 1]$, which is established in the companion paper [25] by overcoming the difficulty raised by the coupling through the normalization. Under the conjecture that a similar convergence in distribution remains valid at the next steps of a particle filter which alternates selections according to the stratified resampling mechanism and mutations according to Markov kernels, we provide an inductive formula for the asymptotic variance of the resampled population after n steps. We perform numerical experiments which support the validity of this formula.

1 Introduction

Particle filtering, also known as Sequential Monte Carlo methods (see Chapter 4 in [3] and [24] for a general introduction), is a powerful method to estimate the evolving state of a system over time, even when the state cannot be directly observed but can only be inferred through noisy measurements or observations. It has become a very popular class of numerical methods for the solution of optimal estimation problems in non-linear non-Gaussian scenarios. This kind of method is used in real-time applications appearing in fields such as chemical engineering, computer vision, financial econometrics, target tracking, robotics and statistics (see among others [11] and [10]). The use of Monte Carlo methods for nonlinear filtering problems can be traced back to Handschin [5] and Mayne [4]. They introduced a sequential version of the importance sampling method (see [1], [2] as references for the importance sampling) and the corresponding algorithm is known as sequential importance sampling, often abbreviated SIS. A drawback of the latter was identified by Gordon et al. in [9]: as the number of iterations increases, the importance weights tend to degenerate (a phenomenon usually known as *weight degeneracy*). This means that after a certain number of iterations, some weights tend to become very small so that the corresponding positions no longer contribute to the estimation. Thus, Gordon et al. introduced the resampling step where, in view of stabilizing the Monte Carlo error over time, the key idea is to eliminate the particles having low weights and to replicate the particles having high weights. Therefore, by propagating M particles through weighting, resampling and mutation steps (each particle evolves randomly according to a given transition probability kernel), particle

"This work is supported by the french National Research Agency under the grant ANR-21-CE40-0006 (SINEQ)."

filters can be used to numerically estimate the state of the system given the observations (see also [10], [24]).

Let us henceforth concentrate on the resampling step. Let $M \geq 1$. Given a sequence of $\mathbb{R}^d$-valued random vectors $(X_m)_{1\leq m\leq M}$ with associated random weights $(w_m^M)_{1\leq m\leq M}$ such that $w_m^M > 0$ and $\sum_{m=1}^{M} w_m^M = 1$, a resampling scheme defines the resampled sequence $(Y_m)_{1\leq m\leq M}$ such that

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m}\,\Big|\,\mathcal{G}\right) = \sum_{m=1}^{M} w_m^M \delta_{X_m}. \tag{1}$$

where $\mathcal{G} := \sigma\left((X_m, w_m^M)_{1\leq m\leq M}\right)$. Resampling schemes [8] permit to replace the probability measure $\sum_{m=1}^{M} w_m^M \delta_{X_m}$ with non equal weights by some empirical measure $\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m}$ with the same conditional expectation given $\mathcal{G}$. Depending on the definition used for the random variables Y_m, several resampling schemes can be considered. The most common resampling techniques are of the following types: multinomial, residual , stratified [6] and systematic [7]. See [8] for a brief description of such methods. The simplest approach is the multinomial resampling. It consists in drawing, conditionally upon $\mathcal{G}$, the new positions $(Y_m)_{1\leq m\leq M}$ independently from the common distribution $\sum_{m=1}^{M} w_m^M \delta_{X_m}$. While the residual resampling consists in replicating $\lfloor M w_m^M \rfloor$-times X_m for $m \in \{1, \cdots, M\}$ and the remaining $M - \sum_{m=1}^{M} \lfloor M w_m^M \rfloor$ variables Y_m are drawn, conditionally upon $\mathcal{G}$, independently from the common distribution $\frac{1}{M-\sum_{m=1}^{M}\lfloor M w_m^M\rfloor}\sum_{\ell=1}^{M}\left\{M w_\ell^M\right\}\delta_{X_\ell}$. Concerning the stratified resampling, it is not straightforward to understand its behaviour, notably due to the complicated structure in the definition of the Y_m which, however, continue to be conditionally independent given $\mathcal{G}$. The systematic resampling is even more complicated to understand since the Y_m are no longer conditionally independent given $\mathcal{G}$.

Resampling schemes have been largely studied in the literature, we now present a selection of such results. The asymptotic behaviour for the multinomial resampling scheme has been extensively studied in [22] (see Corollary 7.4.2 and Section 9.4.2). Douc et al. in [8] showed that residual and stratified resampling improve over multinomial resampling in the sense that they have a lower conditional variance (with respect to the σ-algebra generated by $(X_m)_{1\leq m\leq M}$). They also proved, by means of a counter-example, that the same property does not hold for systematic resampling. Furthermore, they established a central limit theorem for the residual resampling approach suggesting that a similar result should be obtained for the stratified resampling scheme. One of the last contributions concerning resampling schemes, is given by Gerber et al. [19]. Using the notion of negative association [16], they first provided a general consistency result for resampling. An application of this theorem gives the proof of almost sure weak convergence of $\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m}$ in the stratified resampling method. Moreover they provide a counter-example to almost sure weak convergence for the systematic resampling method. More recently, Chopin et al. [17] studied the resampling schemes for particle filters with weakly informative observations. Empirical evidence indicates that when the weights used in resampling exhibit high variability, the selection of the resampling strategy tends to have a weak impact. However, in cases where the weights are close to being uniform, the performance differences between the different resampling methods can be substantial. By keeping M fixed, they also considered the asymptotic behaviour of the resampling schemes as the weights become less and less informative. See also [12], [13], [14], [15] for additional references.

In this paper, we focus on the stratified resampling scheme for the weights $w_m^M = g(X_m)/\sum_{\ell=1}^{M} g(X_\ell)$ where g is a positive measurable function. Our purpose is to study the asymptotic behaviour of the method as the number of particles M goes to ∞. Let f be a real-valued measurable function. To compute the

asympotic variance of $\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m)$, it turns out to be essential to understand the behaviour as $M \to \infty$ of

$$\left\{ \frac{M \sum_{\ell=1}^{\lceil \alpha M \rceil - 1} g(X_\ell)}{\sum_{\ell=1}^{M} g(X_\ell)} \right\} \tag{2}$$

where $0 < \alpha < 1$ and where $\{x\}$ denotes the fractional part of $x \in \mathbb{R}$. Under the assumption that the X_i are i.i.d. such that the law of $g(X_i)$ has an absolutely continuous component, we prove in the companion paper [25] the convergence in distribution of (2) to a random variable uniformly distributed on $[0,1]$. Under this assumption, we explicit the limit of $\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m) \right)$ and prove that $\frac{1}{\sqrt{M}} \sum_{m=1}^{M} \left(f(Y_m^M) - \frac{\mathbb{E}(f(X_1)g(X_1))}{\mathbb{E}(g(X_1))} \right)$ converges in distribution to a centered Gaussian random variable with variance equal to the limit. The proof of this Central Limit Theorem relies on the asymptotic behaviour as $M \to \infty$ of a vector composed of $\sqrt{M} \left(\frac{\sum_{m=1}^{M} g(X_m) f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)} - \mathbb{E}\left(f(X_1) g(X_1) \right) \right)$ and the fractional parts (2) for $\alpha \in \{\alpha_1, \cdots, \alpha_s\}$ with $0 < \alpha_1 < \cdots < \alpha_s < 1$. In the companion paper [25], we check that this vector converges in distribution to a vector with centered Gaussian first component and independent s last components uniformly distributed on $[0,1]^s$. Under the conjecture that a similar convergence in distribution remains valid at the next steps of a particle filter which alternates selections according to the stratified resampling mechanism and mutations according to Markov kernels, we provide an inductive formula for the asymptotic variance of the resampled population after n steps.

The paper is organized as follows. In Section 2, we recall the definition of the stratified sampling scheme and the statement of the main result (Theorem 2.2) is given. In Section 3, the asymptotic variance is derived and in Section 4, the proof of the Central Limit Theorem is given. In Section 5, the proof of some ausiliary results is provided. In Section 6, we consider a particle filter which alternates selections according to the stratified resampling mechanism and mutations according to Markov kernels and we provide an inductive formula for the asymptotic variance of the resampled sequence. We perform numerical experiments which support the validity of this formula.

Notation We denote by $\lfloor x \rfloor$ the integer j such that $j \leq x < j+1$ and by $\{x\} = x - \lfloor x \rfloor$ the fractional part of $x \in \mathbb{R}$. We denote the set of real-valued bounded measurable functions on $\mathbb{R}^d$ by $\mathcal{B}_b\left(\mathbb{R}^d\right)$. Given μ a positive measure on $\mathbb{R}^d$ endowed with the Borel sigma algebra and $\phi : \mathbb{R}^d \to \mathbb{R}$ a measurable function that is either positive or such that $\int_{\mathbb{R}^d} |\phi(x)| \, \mu(dx) < \infty$, we denote $\mu(\phi) = \int_{\mathbb{R}^d} \phi(x) \, \mu(dx)$.

2 Statement of the Main Result

Given $M > 0$, let $(X_m)_{1 \leq m \leq M}$ be a sequence of i.i.d. $\mathbb{R}^d$-valued random vectors following the law η and let $g : \mathbb{R}^d \to (0, \infty)$ be a measurable function such that $0 < \inf_{x \in \mathbb{R}^d} g(x) \leq \sup_{x \in \mathbb{R}^d} g(x) < \infty$. In what follows we denote $\bar{g} := \sup_{x \in \mathbb{R}^d} g(x)$ and $\underline{g} := \inf_{x \in \mathbb{R}^d} g(x)$.
We now generate the sequence $(Y_m^M)_{1 \leq m \leq M}$ according to the selection step of the stratified sampling. We recall that starting from M random variables $(U_m)_{1 \leq m \leq M}$ i.i.d. distributed according to the uniform law on $(0,1)$ and independent of $(X_m)_{1 \leq m \leq M}$, the sequence $(Y_m^M)_{1 \leq m \leq M}$ is defined in the following way

$$Y_m^M = \sum_{\ell=1}^{M} 1_{\left\{ \sum_{j=1}^{\ell-1} w_j^M < m - U_m \leq \sum_{j=1}^{\ell} w_j^M \right\}} X_\ell \ \textit{for} \, m \in \{1, \cdots, M\} \tag{3}$$

where $w_m^M = \frac{Mg(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)}$ for $m = 1, \cdots, M$. Since the weights are preserved when multiplied by a positive constant, up to dividing g by $\mathbb{E}(g(X_1))$, without loss of generality we may suppose that

$$\mathbb{E}(g(X_1)) = 1.$$

Let us observe that the random vectors $(Y_m^M)_{1\leq m\leq M}$ are conditionally independent given $\mathcal{F}$, the σ-algebra generated by the sequence $(X_m)_{1\leq m\leq M}$. Moreover one has

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m^M}|\mathcal{F}\right) = \frac{1}{\sum_{m=1}^{M} g(X_m)}\sum_{m=1}^{M} g(X_m)\delta_{X_m}. \tag{4}$$

Given $f : \mathbb{R}^d \to \mathbb{R}$ a measurable function, our purpose is to provide a central limit theorem for $\frac{1}{M}\sum_{m=1}^{M} f(Y_m^M)$. If we start looking at the variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)$, using (4) one has

$$\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\right) = \mathrm{Var}\left(\mathbb{E}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\,\Big|\,\mathcal{F}\right)\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\,\Big|\,\mathcal{F}\right)\right) \tag{5}$$

$$= \mathrm{Var}\left(\sqrt{M}\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)}\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\,\Big|\,\mathcal{F}\right)\right). \tag{6}$$

Thus, in particular, we will prove respectively the convergence of the first and second term in (6). The first term is common to all the resampling schemes while the second one really depends on the considered resampling scheme.

Let us consider the following hypotheses:

I1 $\sup_{x\in\mathbb{R}^d} |f(x)| < \infty$

I2 the law of $g(X_1)$ has an absolutely continuous component with respect to the Lebesgue measure on $\mathbb{R}$

Before providing the statement of the central limit theorem, to express the asymptotic variance we introduce the real-valued functions β_0 and β_1 respectively given by

$$\beta_0(x, y_1) = \{x + y_1\}\,(1 - \{x + y_1\}) + x(1 - x) - 2x(1 - x - y_1)1_{\{y_1<1-x\}} \tag{7}$$

$$\begin{aligned}\beta_1(x, y_1, y_2, y_3) = 2\Bigg(&\{x + y_1\}\left(1 - \{x + y_1\} - y_2\right)1_{\{y_2<1-\{x+y_1\}\}} \\ &- \{x + y_1\}\left(1 - \{x + y_1\} - y_2 - y_3\right)1_{\{y_2+y_3<1-\{x+y_1\}\}} \\ &- x\left(1 - x - y_1 - y_2\right)1_{\{y_1+y_2<1-x\}} + x\,(1 - x - y_1 - y_2 - y_3)\,1_{\{y_1+y_2+y_3<1-x\}}\Bigg).\end{aligned} \tag{8}$$

Remark 2.1. *Let us observe that β_0 is continuous. Indeed, the fractional part is continuous apart from the integers where its left-hand limit is equal to 1 and its right-hand limit is equal to 0. The composition*

with the function $f(z) = z(1-z)$ which is such that $f(0) = f(1) = 0$, allows us to conclude that $(x, y_1) \mapsto \{x + y_1\}(1 - \{x + y_1\})$ is continuous. Moreover $(x, y_1) \mapsto x(1 - x - y_1)1_{\{y_1<1-x\}}$ is continuous since the function $\bar{f}(x, y_1) = 1 - x - y_1$ is equal to 0 on the set $\left\{(x, y_1) \in \mathbb{R}^2 : y_1 = 1 - x\right\}$ of discontinuity points of the indicator function. Similarly it is possible to prove that also β_1 is a continuous function on $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+$.

The following result holds.

Theorem 2.2. *Under the notation introduced above and under **I1-2** we have*

$$\lim_{M\to\infty} Var\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\right) = \sigma_1^2(f) + \sigma_2^2(f) \tag{9}$$

where

$$\sigma_1^2(f) = \eta\left(\left(g\left(f - \eta(fg)\right)\right)^2\right) \tag{10}$$

and

$$\sigma_2^2(f) := \sum_{k=0}^{\left\lceil \frac{\bar{g}}{g} \right\rceil} \mathbb{E}\left(F_k\right) \tag{11}$$

with $(F_k)_{k\in\mathbb{N}}$ given by

$$F_k = \begin{cases} f^2(X_1)\beta_0\left(U_1, g(X_1)\right) & k = 0 \\ -f(X_1)f(X_{k+1})\beta_1\left(U_1, g(X_1), \sum_{\ell=2}^{k} g(X_\ell), g(X_{k+1})\right) & k > 0 \end{cases} \tag{12}$$

where $U_1 \sim \mathcal{U}(0, 1)$ is independent of $X_1, \cdots, X_{k+1}$.

Moreover the following convergence in distribution holds

$$\sqrt{M}\left(\frac{1}{M}\sum_{m=1}^{M} f(Y_m^M) - \eta(fg)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \sigma_1^2(f) + \sigma_2^2(f)\right). \tag{13}$$

We will split the proof of the theorem into two parts: in Section 3, we are going to prove the result about the asymptotic variance and in Section 4, the proof of the Central Limit Theorem is provided.

3 Asymptotic Variance

The following proposition provides the asymptotic behaviour of the first term in (6). We provide its proof for the sake of completeness.

Proposition 3.1. *Under **I1**, the following convergence in distribution holds*

$$\sqrt{M}\left(\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)} - \eta(fg)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \sigma_1^2(f)\right) \tag{14}$$

where $\sigma_1^2(f)$ has been defined in (10). Moreover

$$\lim_{M\to\infty} Var\left(\sqrt{M}\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)}\right) = \sigma_1^2(f). \tag{15}$$

Proof. We recall that we suppose that $\eta(g) = 1$. Let us define $h := g\,(f - \eta(fg))$. By observing that $\eta(h) = 0$ and

$$\sqrt{M}\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)} - \sqrt{M}\eta(fg) = \frac{M}{\sum\limits_{m=1}^{M} g(X_m)} \times \frac{1}{\sqrt{M}} \sum_{m=1}^{M} h(X_m), \tag{16}$$

by Slutsky's theorem we have

$$\sqrt{M}\left(\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)} - \eta(fg)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \eta(h^2)\right).$$

Let us now prove (15). Let us preliminary study the following quantity

$$\mathbb{E}\left(\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} h(X_m)\right)^4\right) = \frac{1}{M^2} \sum_{m_1,m_2,m_3,m_4=1}^{M} \mathbb{E}\left(\prod_{i=1}^{4} h(X_{m_i})\right).$$

Among the M^4 expectations appearing in the sum on the right-hand side, the ones where one index is different from the other three are equal to 0 by the independence of the X_i and the fact that $\eta(h) = 0$. Therefore

$$\mathbb{E}\left(\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} h(X_m)\right)^4\right) = \frac{1}{M}\eta(h^4) + \frac{3(M-1)}{M}\eta(h^2)^2 \leq D, \tag{17}$$

where D is a finite constant depending on g and f but not on M.
Since (16) holds, to prove (15) we can study the convergence as M goes to infinity of the variance of $\sqrt{M}\dfrac{1}{\bar{g}_M}\bar{h}_M$ given by

$$\mathbb{E}\left(\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)^2\right) - \mathbb{E}\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)^2$$

where we denote by $\bar{g}_M$ the empirical mean $\dfrac{1}{M}\sum\limits_{m=1}^{M} g(X_m)$ and by $\bar{h}_M$ the empirical mean $\dfrac{1}{M}\sum\limits_{m=1}^{M} h(X_m)$.

By the Cauchy–Schwarz inequality we have

$$\begin{aligned}\mathbb{E}\left(\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M} - \sqrt{M}\bar{h}_M\right)^2\right) &= \mathbb{E}\left(\left(\sqrt{M}\bar{h}_M\right)^2\left(\frac{1}{\bar{g}_M} - 1\right)^2\right) \\ &\leq \mathbb{E}^{\frac{1}{2}}\left(\left(\sqrt{M}\bar{h}_M\right)^4\right)\mathbb{E}^{\frac{1}{2}}\left(\left(\frac{1}{\bar{g}_M} - 1\right)^4\right) \leq D^{\frac{1}{2}} \cdot \mathbb{E}^{\frac{1}{2}}\left(\left(\frac{1}{\bar{g}_M} - 1\right)^4\right),\end{aligned}$$

where the right-hand side converges to 0 as M goes to infinity by the Strong Law of Large Numbers and by Lebesgue's theorem. With $\eta(h) = 0$, this in particular implies that

$$0 = \lim_{M\to\infty}\left|\mathbb{E}^{\frac{1}{2}}\left(\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)^2\right) - \mathbb{E}^{\frac{1}{2}}\left(\left(\sqrt{M}\bar{h}_M\right)^2\right)\right| = \lim_{M\to\infty}\left|\mathbb{E}^{\frac{1}{2}}\left(\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)^2\right) - \eta(h^2)^{\frac{1}{2}}\right|$$

and

$$0 = \lim_{M\to\infty}\left|\mathbb{E}\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right) - \mathbb{E}\left(\sqrt{M}\bar{h}_M\right)\right| = \lim_{M\to\infty}\left|\mathbb{E}\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)\right|$$

and so the proof is complete. □

Let us now study the second term in the expression (6). Our purpose is to prove the following result:

Theorem 3.2. *Under **I1-2**, the following convergence holds*

$$\lim_{M\to\infty} \mathbb{E}\left(Var\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) | \mathcal{F} \right) \right) = \sigma_2^2(f) \tag{18}$$

where $\sigma_2^2(f)$ has been defined in (11).

To prove Theorem 3.2 we need some preliminary results the proofs of which are given in Section 5. By the conditional independence of the random vectors $(Y_m^M)_{1\le m\le M}$ with respect to the σ-algebra $\mathcal{F}$, we have

$$\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) \,\Big|\, \mathcal{F} \right) = \frac{1}{M} \sum_{m=1}^{M} \mathrm{Var}\left(f(Y_m^M) \,\Big|\, \mathcal{F} \right) \tag{19}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \mathbb{E}\left(f^2(Y_m^M) \,\Big|\, \mathcal{F} \right) - \frac{1}{M} \sum_{m=1}^{M} \mathbb{E}\left(f(Y_m^M) \,\Big|\, \mathcal{F} \right)^2. \tag{20}$$

We introduce the following notation that will be useful in what follows: for $i = 1, \cdots, M$, let us denote $u_i^M := \left\{ w_1^M + \cdots + w_i^M \right\}$, $\mu_i^M := \lfloor w_1^M + \cdots + w_i^M \rfloor + 1$ where $w_i^M = \frac{Mg(X_i)}{\sum_{\ell=1}^{M} g(X_\ell)}$ has been introduced in (3) and by convention $u_0^M = 0$, $\mu_0^M = 1$.

We now rewrite in a more explicit way the conditional variance $\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) | \mathcal{F} \right)$. The proof is given in Section 5.

Proposition 3.3. *Under **I1**, we have*

$$Var\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) \,\Big|\, \mathcal{F} \right) \tag{21}$$

$$= \frac{1}{M} \sum_{i=1}^{M} f^2(X_i) \beta_0(u_{i-1}^M, w_i^M) - \frac{1}{M} \sum_{k=1}^{(M-1)\wedge \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \sum_{i=1}^{M-k} f(X_i) f(X_{i+k}) \beta_1(u_{i-1}^M, w_i^M, \sum_{\ell=2}^{k} w_{i+\ell-1}^M, w_{i+k}^M) \tag{22}$$

where β_0 and β_1 are respectively defined in (7) and (8).

In what follows, without specifying it every time, we will suppose that $M \ge 1 + \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil$ so that $(M-1) \wedge \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil = \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil$.

Let us observe that it is possible to rewrite (22) in the following way

$$\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) \,\Big|\, \mathcal{F} \right) = \sum_{k=0}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \int_0^1 F_{k,\alpha_1}^M \, d\alpha_1 \tag{23}$$

where

$$F_{k,\alpha_1}^M = \begin{cases} f^2(X_{\lceil \alpha_1 M \rceil}) \beta_0(u_{\lceil \alpha_1 M \rceil - 1}^M, w_{\lceil \alpha_1 M \rceil}^M) & k = 0 \\ -1_{\{\lceil \alpha_1 M \rceil \le M-k\}} f(X_{\lceil \alpha_1 M \rceil}) f(X_{\lceil \alpha_1 M \rceil + k}) \beta_1(u_{\lceil \alpha_1 M \rceil - 1}^M, w_{\lceil \alpha_1 M \rceil}^M, \sum_{\ell=2}^{k} w_{\lceil \alpha_1 M \rceil + \ell - 1}^M, w_{\lceil \alpha_1 M \rceil + k}^M) & k > 0. \end{cases}$$

Given $0 < \alpha_1 < 1$ and $k \in \mathbb{N}$, F^M_{k,α_1} is a function of the vector

$$T^M_{\alpha_1,k} := \left(f\left(X_{\lceil \alpha_1 M \rceil}\right), f\left(X_{\lceil \alpha_1 M \rceil + k}\right), u^M_{\lceil \alpha_1 M \rceil - 1}, w^M_{\lceil \alpha_1 M \rceil}, \sum_{\ell=2}^{k} w^M_{\lceil \alpha_1 M \rceil + \ell - 1}, w^M_{\lceil \alpha_1 M \rceil + k} \right)$$

which is well defined for M big enough. Therefore to study the limit behavior as $M \to \infty$ of the expectation of F^M_{k,α_1}, we are going to study the limit behavior of the vector $T^M_{\alpha_1,k}$. To obtain a much more general formulation useful in the next section, we introduce also the following family of vectors well defined for M big enough. Let $s \geq 1$ and $0 < \alpha_1 < \cdots < \alpha_s < 1$. For $k := (k_1, \cdots, k_s) \in \mathbb{N}^s$, let us define $\forall i = 1, \cdots, s$

$$T^M_{\alpha_i,k_i} := \left(f\left(X_{\lceil \alpha_i M \rceil}\right), f\left(X_{\lceil \alpha_i M \rceil + k_i}\right), u^M_{\lceil \alpha_i M \rceil - 1}, w^M_{\lceil \alpha_i M \rceil}, \sum_{\ell=2}^{k_i} w^M_{\lceil \alpha_i M \rceil + \ell - 1}, w^M_{\lceil \alpha_i M \rceil + k_i} \right). \tag{24}$$

Moreover let

$$H^M := \sqrt{M} \left(\frac{\sum_{m=1}^{M} (fg)(X_m)}{\sum_{m=1}^{M} g(X_m)} - \eta(fg) \right) \tag{25}$$

that thanks to (14) converges in distribution to a centered Gaussian random variable with variance $\sigma_1^2(f)$. The purpose of the next proposition is therefore to study the convergence in distribution of $\left(H^M, T^M_{\alpha_1,k_1}, \cdots, T^M_{\alpha_s,k_s}\right)$.

Proposition 3.4. *Assume **I1-2**. Let $s \in \mathbb{N}^*$ and $k := (k_1, \cdots, k_s) \in \mathbb{N}^s$. We set $K_i := k_1 + \cdots + k_{i-1} + i - 1$ for $i \in \{2, \cdots, s\}$ and $K_1 := 0$. Let $(U_j)_{1 \leq j \leq s}$ be a sequence of i.i.d random variables distributed according to the uniform law on $(0,1)$ and independent of $H \sim \mathcal{N}\left(0, \sigma_1^2(f)\right)$ with $(U_1, \cdots, U_s, H)$ independent from $(X_j)_{1 \leq j \leq K_{s+1}}$. For $0 < \alpha_1 < \cdots < \alpha_s < 1$, as M goes to infinity, the following convergence in distribution holds*

$$\left(H^M, T^M_{\alpha_1,k_1}, \cdots, T^M_{\alpha_s,k_s}\right) \overset{d}{\Longrightarrow} (H, T_{1,k}, \cdots, T_{s,k}) \tag{26}$$

where

$$T_{i,k} := \left(f\left(X_{K_i+1}\right), f\left(X_{K_{i+1}}\right), U_i, g(X_{K_i+1}), \sum_{\ell=K_i+2}^{K_{i+1}-1} g(X_\ell), g(X_{K_{i+1}}) \right) \quad \forall i = 1, \cdots, s. \tag{27}$$

Proof of Proposition 3.4. Let $M \geq \max\left(\frac{2}{\alpha_1}, \frac{k_s}{1-\alpha_s}, \max_{1 \leq i \leq s-1} \frac{k_i+1}{\alpha_{i+1}-\alpha_i} \right)$ so that $2 \leq \lceil \alpha_i M \rceil \leq M - k_i \ \forall i = 1, \cdots, s$ and $k_i < \lceil \alpha_{i+1} M \rceil - \lceil \alpha_i M \rceil \ \forall i = 1, \cdots, s-1$. The last condition allows us to separate the variables: $\left(X_{\lceil \alpha_i M \rceil}, \cdots, X_{\lceil \alpha_i M \rceil + k_i}\right)$ is independent of $\left(X_{\lceil \alpha_j M \rceil}, \cdots, X_{\lceil \alpha_j M \rceil + k_j}\right)$ for i, j distinct elements of $\{1, \cdots, s\}$.
By defining $J^M_{\alpha_i,k_i} := \{\lceil \alpha_i M \rceil, \cdots, \lceil \alpha_i M \rceil + k_i\}$ for $i = 1, \cdots, s$, let $J^M := \bigcup_{i=1}^s J^M_{\alpha_i,k_i}$.
Moreover let $(\tilde{X}_m)_{m \geq 1}$ be a sequence of i.i.d. $\mathbb{R}^d$-valued random vectors following the law η independent of $(X_m)_{1 \leq m \leq M}$. Given $x := (x_j)_{1 \leq j \leq K_{s+1}} \in \mathbb{R}^{d(K_{s+1})}$, we define

$$H^M_x := \sqrt{M} \left(\frac{\sum_{j=1}^{K_{s+1}} (fg)(x_j) + \sum_{j \notin J^M_s} (fg)(X_j)}{\sum_{j=1}^{K_{s+1}} g(x_j) + \sum_{j \notin J^M_s} g(X_j)} - \eta(fg) \right)$$

and

$$\tilde{H}_x^M := \sqrt{M}\left(\frac{\sum_{j=1}^{K_{s+1}}(fg)(x_j)+\sum_{j=1}^{M-K_{s+1}}(fg)(\tilde{X}_j)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}-\eta(fg)\right).$$

Similarly for $i=1,\cdots,s$ we define the vector $T^M_{\alpha_i,k_i,x}$ by

$$\left(f\big(x_{K_i+1}\big),f\big(x_{K_{i+1}}\big),\left\{M\times\frac{\sum_{j=1}^{\lceil\alpha_iM\rceil-1}g(X_j)+\sum_{j=1}^{K_i}g(x_j)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j\notin J_s^M}g(X_j)}\right\},\frac{Mg(x_{K_i+1})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j\notin J_s^M}g(X_j)},\frac{M\sum_{\ell=K_i+2}^{K_{i+1}-1}g(x_\ell)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j\notin J_s^M}g(X_j)},\frac{Mg(x_{K_{i+1}})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j\notin J_s^M}g(X_j)}\right)$$

and the vector $\tilde{T}^M_{\alpha_i,k_i,x}$ by

$$\left(f\big(x_{K_i+1}\big),f\big(x_{K_{i+1}}\big),\left\{M\times\frac{\sum_{j=1}^{\lceil\alpha_iM\rceil-K_i-1}g(\tilde{X}_j)+\sum_{j=1}^{K_i}g(x_j)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}\right\},\frac{Mg(x_{K_i+1})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)},\frac{M\sum_{\ell=K_i+2}^{K_{i+1}-1}g(x_\ell)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)},\frac{Mg(x_{K_{i+1}})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}\right).$$

We notice that

$$\left(H_x^M,T^M_{\alpha_1,k_1,x},\cdots,T^M_{\alpha_s,k_s,x}\right)\overset{\mathcal{L}}{=}\left(\tilde{H}_x^M,\tilde{T}^M_{\alpha_1,k_1,x},\cdots,\tilde{T}^M_{\alpha_s,k_s,x}\right). \tag{28}$$

Let us observe that by hypothesis **I2** the law of $g\left(\tilde{X}_i\right)$ has a component absolutely continuous. Moreover since $\int_{\mathbb{R}}|1+y|^{s+1}e^{-y^2}dy<\infty\ \forall s\geq 1$, we can apply Theorem 1 of [25] with $q=s$ and

- (Y_i,Z_i) equal to $\left(g\left(\tilde{X}_i\right),(fg)(\tilde{X}_i)\right)$
- $\left(\beta_M^1,\cdots,\beta_M^s,\beta_M^{s+1}\right)$ equal to $(\lceil\alpha_1M\rceil-K_1-1,\cdots,\lceil\alpha_sM\rceil-K_s-1,M-K_{s+1})$
- $\phi(x)=\frac{1}{x}$
- x equal to $\sum_{j=1}^{K_{s+1}}g(x_j)$, z equal to $\sum_{j=1}^{K_{s+1}}(fg)(x_j)$, $(y_1,\cdots,y_s)$ equal to $\left(\sum_{j=1}^{K_1}g(x_j),\cdots,\sum_{j=1}^{K_s}g(x_j)\right)$ and θ equal to $\eta(fg)$
- $T:=H\sim\mathcal{N}\left(0,\sigma_1^2(f)\right)$ by (14)

and deduce the following convergence in distribution as $M\to\infty$

$$\left(\tilde{H}_x^M,\left(\left\{M\times\frac{\sum_{j=1}^{\lceil\alpha_iM\rceil-K_i-1}g(\tilde{X}_j)+\sum_{j=1}^{K_i}g(x_j)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}\right\}\right)_{1\leq i\leq s}\right)\overset{d}{\Longrightarrow}(H,U_1,\cdots,U_s). \tag{29}$$

Moreover by the Strong Law of Large Numbers

$$\lim_{M\to\infty}\frac{Mg(x_{K_i+1})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}=g(x_{K_i+1}),\quad\lim_{M\to\infty}\frac{M\sum_{\ell=K_i+2}^{K_{i+1}-1}g(x_\ell)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}=\sum_{\ell=K_i+2}^{K_{i+1}-1}g(x_\ell)$$

and

$$\lim_{M\to\infty} \frac{Mg(x_{K_{i+1}})}{\sum_{j=1}^{K_{s+1}} g(x_j) + \sum_{j=1}^{M-K_{s+1}} g(\tilde{X}_j)} = g(x_{K_{i+1}}).$$

Therefore by Slutsky's theorem, for each $x := (x_j)_{1\le j\le K_{s+1}} \in \mathbb{R}^{d(K_{s+1})}$ we can conclude that the following random vector

$$\left(\tilde{H}_x^M, \tilde{T}^M_{\alpha_1,k_1,x}, \cdots, \tilde{T}^M_{\alpha_s,k_s,x}\right)$$

converges in distribution as M goes to infinity to

$$(H, T_{1,k,x}, \cdots, T_{s,k,x}) \tag{30}$$

where the vector $T_{i,k,x}$ is given by

$$\left(f(x_{K_i+1}), f(x_{K_{i+1}}), U_i, g(x_{K_i+1}), \sum_{\ell=K_i+2}^{K_{i+1}-1} g(x_\ell), g(x_{K_{i+1}})\right).$$

In particular, given $\phi : \mathbb{R}^{1+6s} \to \mathbb{R}$ a continuous bounded function one has

$$\lim_{M\to\infty} \mathbb{E}\left(\phi\left(\tilde{H}_x^M, \tilde{T}^M_{\alpha_1,k_1,x}, \cdots, \tilde{T}^M_{\alpha_s,k_s,x}\right)\right) = \mathbb{E}\left(\phi\left(H, T_{1,k,x}, \cdots, T_{s,k,x}\right)\right). \tag{31}$$

We are now ready to prove (26). By conditioning with respect to $\sigma\left((X_i)_{i\in J_s^M}\right)$ and by applying the Freezing Lemma, (28) and (31), one has

$$\begin{aligned}
&\mathbb{E}\left(\phi\left(H^M, T^M_{\alpha_1,k_1}, \cdots, T^M_{\alpha_s,k_s}\right)\right) \\
&\quad = \int_{\mathbb{R}^{d(k_1+\cdots+k_s+s)}} \mathbb{E}\left(\phi\left(H_x^M, T^M_{\alpha_1,k_1,x}, \cdots, T^M_{\alpha_s,k_s,x}\right)\right) \prod_{i=1}^{K_{s+1}} \eta(dx_i) \\
&\quad = \int_{\mathbb{R}^{d(k_1+\cdots+k_s+s)}} \mathbb{E}\left(\phi\left(\tilde{H}_x^M, \tilde{T}^M_{\alpha_1,k_1,x}, \cdots, \tilde{T}^M_{\alpha_s,k_s,x}\right)\right) \prod_{i=1}^{K_{s+1}} \eta(dx_i) \underset{M\to\infty}{\longrightarrow} \mathbb{E}\left(\phi\left(H, T_{1,k}, \cdots, T_{s,k}\right)\right)
\end{aligned}$$

where $T_{i,k}$, $i = 1, \cdots, s$ is defined in (27) and so the proof is complete. □

We are now ready to prove Theorem 3.2 .

Proof of Theorem 3.2. By (23), it is enough to study the convergence as M goes to infinity of

$$\sum_{k=0}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \int_0^1 \mathbb{E}\left(F^M_{k,\alpha_1}\right) d\alpha_1. \tag{32}$$

Let us observe that we can replace β_0 and β_1 (introduced in (7) and (8)) in the definition of F^M_{k,α_1} respectively by the bounded functions $\beta_0\left(0 \vee x \wedge 1, 0 \vee y_1\right)$ and $\beta_1\left(0 \vee x \wedge 1, 0 \vee y_1, 0 \vee y_2, 0 \vee y_3\right)$ that to simplify the notation we will keep calling β_0 and β_1.

For what has been said there exists a constant $\tilde{C} < \infty$ such that

$$\sup_{(x,y_1)\in\mathbb{R}^2} |\beta_0(x, y_1)| \vee \sup_{(x,y_1,y_2,y_3)\in\mathbb{R}^4} |\beta_1(x, y_1, y_2, y_3)| \le \tilde{C}. \tag{33}$$

By Proposition 3.4 applied with $s = 1$ and by Remark 2.1, we have that for each $\alpha_1 \in (0,1)$

$$F^M_{0,\alpha_1} \overset{d}{\Longrightarrow} F_0$$

with F_0 defined in (12). Moreover by the hypothesis of boundedness of f and (33) we have

$$\lim_{M\to\infty} \mathbb{E}\left(F^M_{0,\alpha_1}\right) = \mathbb{E}\left(F_0\right).$$

By Lebesgue's theorem, $\int_0^1 \mathbb{E}\left(F^M_{0,\alpha_1}\right) d\alpha_1$ will converge to the same limit.

Let us now study the convergence of $\sum_{k=1}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \int_0^1 \mathbb{E}\left(F^M_{k,\alpha_1}\right) d\alpha_1$. By Proposition 3.4 applied with $s = 1$ and by Remark 2.1, we have that for each $\alpha_1 \in (0,1)$ and $k \geq 1$

$$F^M_{k,\alpha_1} \overset{d}{\Longrightarrow} F_k$$

with F_k defined in (12). Moreover by the hypothesis of boundedness of f and (33),

$$\lim_{M\to\infty} \mathbb{E}\left(F^M_{k,\alpha_1}\right) = \mathbb{E}\left(F_k\right).$$

By Lebesgue's theorem, $\lim_{M\to\infty} \int_0^1 \mathbb{E}\left(F^M_{k,\alpha_1}\right) d\alpha_1 = \mathbb{E}\left(F_k\right)$ for each $k = 1, \cdots, \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil$ and so the proof is complete. □

We are now ready to provide a central limit theorem for $\frac{1}{M}\sum_{m=1}^{M} f(Y^M_m)$ knowing its asymptotic variance.

4 Central Limit Theorem

Theorem 4.1. *Let assume **I1-2**. The following convergence in distribution holds*

$$\sqrt{M}\left(\frac{1}{M}\sum_{m=1}^{M} f(Y^M_m) - \eta(fg)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \sigma_1^2(f) + \sigma_2^2(f)\right) \tag{34}$$

where $\sigma_1^2(f)$ and $\sigma_2^2(f)$ are respectively defined in (10) and (11).

Proof of Theorem 4.1. Let $u \in \mathbb{R}$. By introducing the notation

$$a^M_m := e^{\frac{iu}{\sqrt{M}}\left(f(Y^M_m) - \mathbb{E}\left(f(Y^M_m)|\mathcal{F}\right)\right)} \; for \; m = 1, \cdots, M \tag{35}$$

and by applying (4), we have

$$\begin{aligned}
\mathbb{E}\left(e^{\frac{iu}{\sqrt{M}}\sum_{m=1}^{M}\left(f(Y^M_m)-\eta(fg)\right)}\right) &= \mathbb{E}\left(e^{iu\sqrt{M}\left(\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} f(Y^M_m)\middle|\mathcal{F}\right)-\eta(fg)\right)}\mathbb{E}\left(\prod_{m=1}^{M} a^M_m \;\middle|\; \mathcal{F}\right)\right) \\
&= \mathbb{E}\left(e^{iuH^M}\mathbb{E}\left(\prod_{m=1}^{M} a^M_m \;\middle|\; \mathcal{F}\right)\right)
\end{aligned} \tag{36}$$

where H^M is defined in (25). The purpose of what follows is therefore to study the convergence as M goes to infinity of (36).

First step

We are first going to study the conditional expectation $\mathbb{E}\left(\prod_{m=1}^{M} a_m^M \;\middle|\; \mathcal{F}\right)$.

By using that for each $u, y \in \mathbb{R}$ $\left|e^{iuy} - 1 - iuy + \frac{(uy)^2}{2}\right| \leq \frac{|uy|^3}{6}$, one has

$$\left|a_m^M - 1 - \frac{iu}{\sqrt{M}}\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right) + \frac{u^2\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right)^2}{2M}\right|$$
$$\leq \frac{|u|^3\left|f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right|^3}{6M^{\frac{3}{2}}} \leq \frac{4\left\|f\right\|_\infty^3 |u|^3}{3M^{\frac{3}{2}}}.$$

Therefore for $m = 1, \cdots, M$, a_m^M can be rewritten as

$$a_m^M = 1 + \frac{iu}{\sqrt{M}}\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right) - \frac{u^2\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right)^2}{2M} + r_m^M$$

with $\left|r_m^M\right| \leq \frac{4\|f\|_\infty^3 |u|^3}{3M^{\frac{3}{2}}}$.

By defining for $m = 1, \cdots, M$

$$b_m^M := 1 - \frac{u^2\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right)^2}{2M}, \tag{37}$$

let us now study the difference

$$\left|\mathbb{E}\left(\prod_{m=1}^{M} a_m^M \;\middle|\; \mathcal{F}\right) - \mathbb{E}\left(\prod_{m=1}^{M} b_m^M \;\middle|\; \mathcal{F}\right)\right|. \tag{38}$$

By rewriting $\prod_{m=1}^{M} a_m^M - \prod_{m=1}^{M} b_m^M = \sum_{m=1}^{M} \prod_{\ell=1}^{m-1} a_\ell^M \left(a_m^M - b_m^M\right) \prod_{\ell=m+1}^{M} b_\ell^M$, where by convention the empty product is equal to one, by using that the random vectors $(Y_m^M)_{1\leq m\leq M}$ are conditionally independent given $\mathcal{F}$ and that $\left|r_m^M\right| \leq \frac{4\|f\|_\infty^3 |u|^3}{3M^{\frac{3}{2}}}$, $\left|a_\ell^M\right| \leq 1$ and $\left|b_\ell^M\right| \leq 1 + \frac{2\|f\|_\infty^2 u^2}{M}$, (38) becomes

$$\begin{aligned}
&\left|\mathbb{E}\left(\prod_{m=1}^{M} a_m^M \;\middle|\; \mathcal{F}\right) - \mathbb{E}\left(\prod_{m=1}^{M} b_m^M \;\middle|\; \mathcal{F}\right)\right| \\
&= \left|\sum_{m=1}^{M} \mathbb{E}\left(\prod_{\ell=1}^{m-1} a_\ell^M \left(\frac{iu}{\sqrt{M}}\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right) + r_m^M\right) \prod_{\ell=m+1}^{M} b_\ell^M \;\middle|\; \mathcal{F}\right)\right| \\
&= \left|\sum_{m=1}^{M} \mathbb{E}\left(\prod_{\ell=1}^{m-1} a_\ell^M \;\middle|\; \mathcal{F}\right) \mathbb{E}\left(\frac{iu}{\sqrt{M}}\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right) + r_m^M \;\middle|\; \mathcal{F}\right) \mathbb{E}\left(\prod_{\ell=m+1}^{M} b_\ell^M \;\middle|\; \mathcal{F}\right)\right| \\
&= \left|\sum_{m=1}^{M} \mathbb{E}\left(\prod_{\ell=1}^{m-1} a_\ell^M \;\middle|\; \mathcal{F}\right) \mathbb{E}\left(r_m^M \;\middle|\; \mathcal{F}\right) \mathbb{E}\left(\prod_{\ell=m+1}^{M} b_\ell^M \;\middle|\; \mathcal{F}\right)\right| \leq M \times \frac{4\left\|f\right\|_\infty^3 |u|^3}{3M^{\frac{3}{2}}} \times \left(1 + \frac{2\left\|f\right\|_\infty^2 u^2}{M}\right)^M \underset{M\to\infty}{\longrightarrow} 0.
\end{aligned}$$

This implies that

$$\limsup_{M\to\infty} \left|\mathbb{E}\left(e^{iuH^M}\mathbb{E}\left(\prod_{m=1}^{M} a_m^M \;\middle|\; \mathcal{F}\right)\right) - \mathbb{E}\left(e^{iuH^M}\mathbb{E}\left(\prod_{m=1}^{M} b_m^M \;\middle|\; \mathcal{F}\right)\right)\right| = 0. \tag{39}$$

Second step

By the previous step, it is enough to study the limit behaviour as $M \to \infty$ of $\mathbb{E}\left(e^{iuH^M}\mathbb{E}\left(\prod_{m=1}^{M} b_m^M \;\middle|\; \mathcal{F}\right)\right)$ where b_m^M is defined in (37) .

By using the general fact that $\prod_{m=1}^{M}(1-c_m) = \sum_{s=0}^{M}(-1)^s \sum_{1\le m_1<\cdots<m_s\le M}\prod_{j=1}^{s} c_{m_j}$ for any real sequence $(c_m)_{1\le m\le M}$, we have

$$\begin{aligned}\mathbb{E}\left(\prod_{m=1}^{M} b_m^M \;\middle|\; \mathcal{F}\right) &= \sum_{s=0}^{M}\left(\frac{-u^2}{2M}\right)^s \sum_{1\le m_1<\cdots<m_s\le M} \mathbb{E}\left(\prod_{j=1}^{s}\left(f(Y_{m_j}^M) - \mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2 \;\middle|\; \mathcal{F}\right)\\ &= \sum_{s=0}^{M}\left(\frac{-u^2}{2M}\right)^s \sum_{1\le m_1<\cdots<m_s\le M}\prod_{j=1}^{s}\mathbb{E}\left(\left(f(Y_{m_j}^M) - \mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2 \;\middle|\; \mathcal{F}\right) \qquad (40)\end{aligned}$$

where for the last equality we use that the random vectors $(Y_m^M)_{1\le m\le M}$ are conditionally independent given $\mathcal{F}$. Let us introduce $B_{m_1,\cdots,m_s} := \prod_{j=1}^{s}\mathbb{E}\left(\left(f(Y_{m_j}^M) - \mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2 \;\middle|\; \mathcal{F}\right)$. Since $B_{m_1,\cdots,m_s}$ is symmetric in its indexes and satisfies $|B_{m_1,\cdots,m_s}| \le \left(4\|f\|_\infty^2\right)^s$, for each $s \ge 1$ we have

$$\left|\frac{1}{M^s}\left(\frac{1}{s!}\sum_{m_1,\cdots,m_s=1}^{M} B_{m_1,\cdots,m_s} - \sum_{1\le m_1<\cdots<m_s\le M} B_{m_1,\cdots,m_s}\right)\right| \le \frac{\left(4\|f\|_\infty^2\right)^s}{M^s}\left(\frac{M^s}{s!} - \binom{M}{s}\right).$$

It is therefore possible to rewrite (40) as

$$\begin{aligned}\mathbb{E}\left(\prod_{m=1}^{M} b_m^M \;\middle|\; \mathcal{F}\right) &= \sum_{s=0}^{M}\left(-\frac{u^2}{2M}\right)^s \sum_{1\le m_1<\cdots<m_s\le M}\prod_{j=1}^{s}\mathbb{E}\left(\left(f(Y_{m_j}^M) - \mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2 \;\middle|\; \mathcal{F}\right)\\ &= \sum_{s=0}^{M}\left(-\frac{u^2}{2M}\right)^s \frac{1}{s!}\sum_{m_1,\cdots,m_s=1}^{M}\prod_{j=1}^{s}\mathbb{E}\left(\left(f(Y_{m_j}^M) - \mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2 \;\middle|\; \mathcal{F}\right) + \bar{r}^M\\ &= \sum_{s=0}^{M}\left(-\frac{u^2}{2}\right)^s \frac{1}{s!}\left(\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left(\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M)\mid\mathcal{F}\right)\right)^2 \;\middle|\; \mathcal{F}\right)\right)^s + \bar{r}^M\\ &= \sum_{s=0}^{M}\left(-\frac{u^2}{2}\right)^s \frac{1}{s!}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M) \;\middle|\; \mathcal{F}\right)\right)^s + \bar{r}^M\end{aligned}$$

with $\bar{r}^M$ such that

$$\left|\bar{r}^M\right| \le \sum_{s\ge 1}\frac{\left(2u^2\|f\|_\infty^2\right)^s}{s!}\left(1 - \frac{M(M-1)\cdots(M-s+1)}{M^s}\right) \underset{M\to\infty}{\longrightarrow} 0.$$

With $\left|e^{iuH^M}\right| \le 1$, (36) and (39), this implies that

$$\left|\mathbb{E}\left(e^{\frac{iu}{\sqrt{M}}\sum_{m=1}^{M}\left(f(Y_m^M)-\eta(fg)\right)}\right) - \mathbb{E}\left(e^{iuH^M}\sum_{s=0}^{M}\left(-\frac{u^2}{2}\right)^s\frac{1}{s!}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M) \;\middle|\; \mathcal{F}\right)\right)^s\right)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

Third step

The purpose of what follows is therefore to study the convergence as M goes to infinity of

$$\mathbb{E}\left(e^{iuH^M}\sum_{s=0}^{M}\left(-\frac{u^2}{2}\right)^s\frac{1}{s!}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M}f(Y_m^M)\,\Big|\,\mathcal{F}\right)\right)^s\right). \tag{41}$$

By (23), (41) can be rewritten as

$$\begin{aligned}&\mathbb{E}\left(e^{iuH^M}\right)+\sum_{s=1}^{M}\left(-\frac{u^2}{2}\right)^s\frac{1}{s!}\sum_{k_1,\cdots,k_s=0}^{\left\lceil\frac{\bar{g}}{\underline{g}}\right\rceil}\int_{[0,1]^s}\mathbb{E}\left(e^{iuH^M}F^M_{k_1,\alpha_1}\cdots F^M_{k_s,\alpha_s}\right)d\alpha_1\cdots d\alpha_s\\&=\mathbb{E}\left(e^{iuH^M}\right)+\sum_{s\geq 1}1_{\{s\leq M\}}z(s,M)\end{aligned} \tag{42}$$

where for $s\geq 1$ and $M\geq 1$

$$z(s,M)=\left(-\frac{u^2}{2}\right)^s\sum_{k_1,\cdots,k_s=0}^{\left\lceil\frac{\bar{g}}{\underline{g}}\right\rceil}\int_{[0,1]^s}1_{\{0<\alpha_1<\cdots<\alpha_s<1\}}\mathbb{E}\left(e^{iuH^M}F^M_{k_1,\alpha_1}\cdots F^M_{k_s,\alpha_s}\right)d\alpha_1\cdots d\alpha_s$$

and where we recall that for $i=1,\cdots,s$

$$F^M_{k_i,\alpha_i}=\begin{cases}f^2(X_{\lceil\alpha_iM\rceil})\beta_0(u^M_{\lceil\alpha_iM\rceil-1},w^M_{\lceil\alpha_iM\rceil}) & k_i=0\\ -1_{\{\lceil\alpha_iM\rceil\leq M-k_i\}}f(X_{\lceil\alpha_iM\rceil})f(X_{\lceil\alpha_iM\rceil+k_i})\beta_1(u^M_{\lceil\alpha_iM\rceil-1},w^M_{\lceil\alpha_iM\rceil},\sum_{\ell=2}^{k_i}w^M_{\lceil\alpha_iM\rceil+\ell-1},w^M_{\lceil\alpha_iM\rceil+k_i}) & k_i>0.\end{cases}$$

Let us now study the convergence as $M\to\infty$ of (42): $\lim_{M\to\infty}\mathbb{E}\left(e^{iuH^M}\right)=e^{-\frac{u^2\sigma_1^2(f)}{2}}$ by (14) and to study the convergence of the sum we will apply Lebesgue's theorem.
Given $s\geq 1$, we first compute the pointwise convergence as $M\to\infty$ of $z(s,M)$ and then we show the existence of a function $\tilde{z}$ independent of M that dominates $|z|$ and such that $\sum_{s\geq 1}\tilde{z}(s)<\infty$.

Let us start by computing the pointwise convergence. Let $s\geq 1$ and $0\leq k_1,k_2,\cdots,k_s\leq\left\lceil\frac{\bar{g}}{\underline{g}}\right\rceil$.
By applying the reasoning based on Proposition 3.4 done in the proof of Theorem 3.2, $\forall\ 0<\alpha_1<\cdots<\alpha_s<1$ we check that

$$e^{iuH^M}F^M_{k_1,\alpha_1}\cdots F^M_{k_s,\alpha_s}\overset{d}{\Longrightarrow}e^{iuH}F_{1,k}\cdots F_{s,k}$$

where, recalling that k denotes the multiindex $(k_1,\cdots,k_s)$ and $K_i:=k_1+\cdots+k_{i-1}+i-1\ \forall i=1,\cdots,s$ with the convention $K_1:=0$, $F_{i,k}$ for $i=1,\cdots,s$ is given by

$$F_{i,k}=\begin{cases}f^2(X_{K_i+1})\beta_0\left(U_i,g(X_{K_i+1})\right) & k_i=0\\ -f(X_{K_i+1})f(X_{K_{i+1}})\beta_1\left(U_i,g(X_{K_i+1}),\sum_{\ell=K_i+2}^{K_{i+1}-1}g(X_\ell),g(X_{K_{i+1}})\right) & k_i>0\end{cases}$$

and $H\sim\mathcal{N}\left(0,\sigma_1^2(f)\right)$ independent of $(U_j)_{1\leq j\leq s}$ and $(X_j)_{1\leq j\leq K_{s+1}}$.
Moreover by the hypothesis of boundedness of f, (33) and the fact that $\left|e^{iuH^M}\right|\leq 1$,

$$\lim_{M\to\infty}\mathbb{E}\left(e^{iuH^M}F^M_{k_1,\alpha_1}\cdots F^M_{k_s,\alpha_s}\right)=\mathbb{E}\left(e^{iuH}F_{1,k}\cdots F_{s,k}\right).$$

By Lebesgue's theorem and by observing that $F_{1,k}, \cdots, F_{s,k}$ are independent, respectively distributed as $F_{k_1}, \cdots, F_{k_s}$ and independent of H, $z(s,M)$ converges to

$$\left(-\frac{u^2}{2}\right)^s \sum_{k_1,\cdots,k_s=0}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \int_{[0,1]^s} 1_{\{0<\alpha_1<\cdots<\alpha_s<1\}} \mathbb{E}\left(e^{iuH} F_{1,k}\cdots F_{s,k}\right) d\alpha_1 \cdots d\alpha_s$$

$$= e^{-\frac{u^2\sigma_1^2(f)}{2}} \frac{1}{s!} \left(-\frac{u^2}{2}\right)^s \sum_{k_1,\cdots,k_s=0}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \mathbb{E}\left(F_{k_1}\right)\cdots\mathbb{E}\left(F_{k_s}\right)$$

$$= e^{-\frac{u^2\sigma_1^2(f)}{2}} \frac{1}{s!} \left(-\frac{u^2}{2}\right)^s \left(\sum_{k_1=0}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \mathbb{E}\left(F_{k_1}\right)\right)^s = e^{-\frac{u^2\sigma_1^2(f)}{2}} \frac{1}{s!} \left(-\frac{u^2}{2}\right)^s \left(\sigma_2^2(f)\right)^s.$$

Let us now show that z is dominated by a summable function $\tilde{z}$. By using that f, β_0 and β_1 are bounded functions, for $s \geq 1$ one has

$$|z(s,M)| = \left| \left(-\frac{u^2}{2}\right)^s \sum_{k_1,\cdots,k_s=0}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \int_{[0,1]^s} 1_{\{0<\alpha_1<\cdots<\alpha_s<1\}} \mathbb{E}\left(e^{iuH} F_{1,k}\cdots F_{s,k}\right) d\alpha_1 \cdots d\alpha_s \right|$$

$$\leq C \frac{1}{s!} \left(\frac{u^2\left(\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil + 1\right)}{2}\right)^s$$

for a given finite constant C. It is now sufficient to observe that $\sum_{s\geq 1} \frac{1}{s!}\left(\left(\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil + 1\right) \times \frac{u^2}{2}\right)^s = e^{\left(\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil + 1\right)\frac{u^2}{2}} < \infty$. In conclusion we have proved that (42) converges as $M \to \infty$ to

$$e^{-\frac{u^2\sigma_1^2(f)}{2}} \sum_{s\geq 0} \frac{1}{s!} \left(-\frac{u^2}{2}\right)^s \left(\sigma_2^2(f)\right)^s = e^{-\frac{u^2\left(\sigma_1^2(f)+\sigma_2^2(f)\right)}{2}}.$$

□

5 Proof of Proposition 3.3

We recall that for $i = 1, \cdots, M$ we denote $u_i^M := \left\{w_1^M + \cdots + w_i^M\right\}$ and $\mu_i^M := \lfloor w_1^M + \cdots + w_i^M \rfloor + 1$, where by convention $u_0^M = 0$, $\mu_0^M = 1$. The following technical result holds.

Lemma 5.1. *If $1 \leq p < \ell \leq M$ then*

1. *$\mu_p^M < \mu_\ell^M$ if and only if $w_{p+1}^M + \cdots + w_\ell^M \geq 1 - u_p^M$*
2. *$\mu_p^M = \mu_\ell^M$ if and only if $w_{p+1}^M + \cdots + w_\ell^M < 1 - u_p^M$*
3. *$\mu_p^M = \mu_\ell^M$ if and only if $u_\ell^M = u_p^M + w_{p+1}^M + \cdots + w_\ell^M$.*

Proof of Lemma 5.1. Let us observe that 2. follows directly from 1. and 3. is a direct consequence of the definition of the integer and fractional part. Therefore it is enough to prove 1.

If $\mu_p^M < \mu_\ell^M$, then

$$w_1^M + \cdots + w_p^M - u_p^M + 1 = \mu_p^M \leq \mu_\ell^M - 1 \leq w_1^M + \cdots + w_\ell^M$$

that implies

$$1 - u_p^M \leq w_{p+1}^M + \cdots + w_\ell^M.$$

Let us now prove the other implication. If $1 - u_p^M \leq w_{p+1}^M + \cdots + w_\ell^M$, then

$$\mu_p^M = w_1^M + \cdots + w_p^M + 1 - u_p^M \leq w_1^M + \cdots + w_\ell^M < \mu_\ell^M.$$

□

The following lemma provides an explicit expression for the conditional expectation $\mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)$ for $m = 1, \cdots, M$ which appears in (20) and so it allows to prove Proposition 3.3.

Lemma 5.2. *Given $\varrho \in \mathcal{B}_b\left(\mathbb{R}^d\right)$, for $m = 1, \cdots, M$ we have*

$$\mathbb{E}\left(\varrho(Y_m^M) \mid \mathcal{F}\right) = \sum_{i=1}^{M} \varrho(X_i) q_{m,i}^M 1_{\left\{\mu_{i-1}^M \leq m \leq \mu_i^M\right\}}$$

where for $i = 1, \cdots, M$

$$q_{m,i}^M = 1_{\left\{\mu_{i-1}^M < \mu_i^M\right\}}(1_{\left\{\mu_{i-1}^M < m < \mu_i^M\right\}} + 1_{\left\{m = \mu_{i-1}^M\right\}}(1 - u_{i-1}^M) + 1_{\left\{m = \mu_i^M\right\}} u_i^M) + 1_{\left\{\mu_{i-1}^M = m = \mu_i^M\right\}} w_i^M. \quad (43)$$

Proof of Lemma 5.2. For $m = 1, \cdots, M$, by (3), by observing that $m - U_m$ is uniformly distributed on $[m-1, m]$ and by the Freezing Lemma, we have

$$\mathbb{E}\left(\varrho(Y_m^M) \mid \mathcal{F}\right) = \sum_{i=1}^{M} \varrho(X_i) \int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du. \quad (44)$$

By observing that for $i = 1, \cdots, M$

$$1 = 1_{\left\{m < \mu_{i-1}^M\right\}} + 1_{\left\{m > \mu_i^M\right\}} + 1_{\left\{\mu_{i-1}^M < \mu_i^M\right\}} \left(1_{\left\{m = \mu_{i-1}^M\right\}} + 1_{\left\{m = \mu_i^M\right\}} + 1_{\left\{\mu_{i-1}^M < m < \mu_i^M\right\}}\right) + 1_{\left\{m = \mu_{i-1}^M = \mu_i^M\right\}},$$

let us study the value of the integral in (44) according to this partition. If $m < \mu_{i-1}^M$ then $m \leq \mu_{i-1}^M - 1 \leq w_1^M + \cdots + w_{i-1}^M$ and if $m > \mu_i^M$ then $m \geq \mu_i^M + 1 > w_1^M + \cdots + w_i^M + 1$ and so in both cases the integral is zero. Let us now suppose that $\mu_{i-1}^M < \mu_i^M$. If $\mu_{i-1}^M < m < \mu_i^M$, then $m \geq \mu_{i-1}^M + 1 > w_1^M + \cdots + w_{i-1}^M + 1$ and $m \leq w_1^M + \cdots + w_i^M$. Therefore

$$\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du = 1.$$

If $m = \mu_{i-1}^M$:

$$\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du = \int_{w_1^M + \cdots + w_{i-1}^M}^{\mu_{i-1}^M} du = 1 - u_{i-1}^M.$$

If $m = \mu_i^M$:

$$\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du = \int_{\mu_i^M - 1}^{w_1^M + \cdots + w_i^M} du = u_i^M.$$

Finally if $m = \mu_i^M = \mu_{i-1}^M$:

$$\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du = \int_{w_1^M+\cdots+w_{i-1}^M}^{w_1^M+\cdots+w_i^M} du = w_i^M.$$

To sum up we have obtained that for $i = 1, \cdots, M$

$$\begin{aligned}
&\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du \\
&\quad = 1_{\{\mu_{i-1}^M < \mu_i^M\}} \big(1_{\{\mu_{i-1}^M < m < \mu_i^M\}} + 1_{\{m=\mu_{i-1}^M\}}(1-u_{i-1}^M) + 1_{\{m=\mu_i^M\}} u_i^M\big) + 1_{\{\mu_{i-1}^M = m = \mu_i^M\}} w_i^M.
\end{aligned}$$

□

Proof of Proposition 3.3. Let us rewrite (20). By Lemma 5.2 we have

$$\begin{aligned}
&\frac{1}{M}\sum_{m=1}^{M} \mathbb{E}\left(f^2(Y_m^M)|\mathcal{F}\right) - \frac{1}{M}\sum_{m=1}^{M} \mathbb{E}\left(f(Y_m^M)|\mathcal{F}\right)^2 \\
&\quad = \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{M} f^2(X_i) q_{m,i}^M 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} - \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{M} f^2(X_i)(q_{m,i}^M)^2 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} \\
&\qquad - \frac{2}{M}\sum_{m=1}^{M} \sum_{1\leq i<j\leq M} f(X_i) f(X_j) q_{m,i}^M q_{m,j}^M 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} 1_{\{\mu_{j-1}^M \leq m \leq \mu_j^M\}} \\
&\quad = \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{M} f^2(X_i)\left(q_{m,i}^M - (q_{m,i}^M)^2\right) 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} - \frac{2}{M}\sum_{m=1}^{M} \sum_{1\leq i<j\leq M} f(X_i) f(X_j) q_{m,i}^M q_{m,j}^M 1_{\{m=\mu_i^M=\mu_{j-1}^M\}}.
\end{aligned} \tag{45}$$

We are now going first to rewrite the first component of the right-hand side and then the second one.

First term:

Let us observe that by Lemma 5.1 applied to the couple (p, l) equal to $(i-1, i)$, we have $1_{\{\mu_i^M > \mu_{i-1}^M\}} = 1_{\{w_i^M \geq 1-u_{i-1}^M\}}$ and $1_{\{\mu_i^M = \mu_{i-1}^M\}} = 1_{\{w_i^M < 1-u_{i-1}^M\}}$ for $i = 1 \cdots, M$. Therefore by Lemma 5.2 and by using that $\mu_i^M - \mu_{i-1}^M = u_{i-1}^M - u_i^M + w_i^M$, we have

$$\begin{aligned}
&\sum_{m=1}^{M} q_{m,i}^M 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} \\
&= \sum_{m=1}^{M} \Big(1_{\{w_i^M \geq 1-u_{i-1}^M\}} \big(1_{\{\mu_{i-1}^M < m < \mu_i^M\}} + 1_{\{m=\mu_{i-1}^M\}}(1-u_{i-1}^M) + 1_{\{m=\mu_i^M\}} u_i^M\big) + 1_{\{w_i^M < 1-u_{i-1}^M\}} 1_{\{\mu_{i-1}^M = m = \mu_i^M\}} w_i^M\Big) \\
&= 1_{\{w_i^M \geq 1-u_{i-1}^M\}}\left(\mu_i^M - \mu_{i-1}^M - u_{i-1}^M + u_i^M\right) + 1_{\{w_i^M < 1-u_{i-1}^M\}} w_i^M = w_i^M 1_{\{w_i^M \geq 1-u_{i-1}^M\}} + w_i^M 1_{\{w_i^M < 1-u_{i-1}^M\}} = w_i^M.
\end{aligned}$$

Let us observe that since, by (4), $\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left(f^2(Y_m^M)|\mathcal{F}\right)=\frac{1}{M}\sum_{i=1}^{M}f^2(X_i)w_i^M$, we got the expected result. Similarly, by using that $\mu_i^M-\mu_{i-1}^M=u_{i-1}^M-u_i^M+w_i^M$, we have

$$\begin{aligned}
&\sum_{m=1}^{M}(q_{m,i}^M)^2 1_{\left\{\mu_{i-1}^M\le m\le\mu_i^M\right\}}\\
&\quad=1_{\left\{w_i^M\ge 1-u_{i-1}^M\right\}}(\mu_i^M-\mu_{i-1}^M-1+(1-u_{i-1}^M)^2+(u_i^M)^2)+1_{\left\{w_i^M<1-u_{i-1}^M\right\}}(w_i^M)^2\\
&\quad=(w_i^M-u_i^M(1-u_i^M)-u_{i-1}^M(1-u_{i-1}^M))1_{\left\{w_i^M\ge 1-u_{i-1}^M\right\}}+(w_i^M)^2 1_{\left\{w_i^M<1-u_{i-1}^M\right\}}\\
&\quad=w_i^M-u_i^M(1-u_i^M)-u_{i-1}^M(1-u_{i-1}^M)-(w_i^M(1-w_i^M)-u_i^M(1-u_i^M)-u_{i-1}^M(1-u_{i-1}^M))1_{\left\{w_i^M<1-u_{i-1}^M\right\}}.
\end{aligned}$$

Substrating the two quantities and using Lemma 5.1 to rewrite u_i^M as $w_i^M+u_{i-1}^M$ if $w_i^M<1-u_{i-1}^M$, we have

$$\begin{aligned}
&\sum_{m=1}^{M}q_{m,i}^M 1_{\left\{\mu_{i-1}^M\le m\le\mu_i^M\right\}}-\sum_{m=1}^{M}(q_{m,i}^M)^2 1_{\left\{\mu_{i-1}^M\le m\le\mu_i^M\right\}}\\
&=u_i^M(1-u_i^M)+u_{i-1}^M(1-u_{i-1}^M)+(w_i^M(1-w_i^M)-u_i^M(1-u_i^M)-u_{i-1}^M(1-u_{i-1}^M))1_{\left\{w_i^M<1-u_{i-1}^M\right\}}\\
&=u_i^M(1-u_i^M)+u_{i-1}^M(1-u_{i-1}^M)+(w_i^M(1-w_i^M)-(w_i^M+u_{i-1}^M)(1-w_i^M-u_{i-1}^M)-u_{i-1}^M(1-u_{i-1}^M))1_{\left\{w_i^M<1-u_{i-1}^M\right\}}\\
&=u_i^M(1-u_i^M)+u_{i-1}^M(1-u_{i-1}^M)-(2u_{i-1}^M(1-u_{i-1}^M)-2w_i^M u_{i-1}^M)1_{\left\{w_i^M<1-u_{i-1}^M\right\}}.
\end{aligned}$$

Therefore by observing that for $i=1\cdots,M$ $u_i^M=\left\{u_{i-1}^M+w_i^M\right\}$, the first term can be rewritten as

$$\frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{M}f^2(X_i)\left(q_{m,i}^M-(q_{m,i}^M)^2\right)1_{\left\{\mu_{i-1}^M\le m\le\mu_i^M\right\}}=\frac{1}{M}\sum_{i=1}^{M}f^2(X_i)\beta_0(u_{i-1}^M,w_i^M) \tag{46}$$

with β_0 defined in (7).

Second term:

By Lemma 5.2 and by rewriting $1_{\left\{\mu_{i-1}^M<\mu_i^M\right\}}=1-1_{\left\{\omega_i^M<1-u_{i-1}^M\right\}}$, we can rewrite the right-hand side of (45) in the following way

$$\begin{aligned}
&\frac{2}{M}\sum_{m=1}^{M}\sum_{1\le i<j\le M}f(X_i)f(X_j)q_{m,i}^M q_{m,j}^M 1_{\left\{m=\mu_i^M=\mu_{j-1}^M\right\}}\\
&=\frac{2}{M}\sum_{1\le i<j\le M}f(X_i)f(X_j)\left(1_{\left\{\mu_{i-1}^M<\mu_i^M\right\}}u_i^M+1_{\left\{\mu_{i-1}^M=\mu_i^M\right\}}w_i^M\right)\\
&\qquad\qquad\times\left(1_{\left\{\mu_{j-1}^M<\mu_j^M\right\}}(1-u_{j-1}^M)+1_{\left\{\mu_{j-1}^M=\mu_j^M\right\}}w_j^M\right)1_{\left\{\mu_i^M=\mu_{j-1}^M\right\}}\\
&=\frac{2}{M}\sum_{k=1}^{M-1}\sum_{i=1}^{M-k}f(X_i)f(X_{i+k})\left(\rho_{i,k}^M+\tilde{\rho}_{i,k}^M\right)
\end{aligned}$$

where for $k=1,\cdots,M-1$ and $i=1,\cdots,M-k$

$$\begin{aligned}
\rho_{i,k}^M&=u_i^M\left(1_{\left\{\mu_{i+k-1}^M<\mu_{i+k}^M\right\}}(1-u_{i+k-1}^M)+1_{\left\{\mu_{i+k-1}^M=\mu_{i+k}^M\right\}}w_{i+k}^M\right)1_{\left\{\mu_i^M=\mu_{i+k-1}^M\right\}}\\
\tilde{\rho}_{i,k}^M&=\left(w_i^M-u_i^M\right)\left(1_{\left\{\mu_{i+k-1}^M<\mu_{i+k}^M\right\}}(1-u_{i+k-1}^M)+1_{\left\{\mu_{i+k-1}^M=\mu_{i+k}^M\right\}}w_{i+k}^M\right)1_{\left\{\mu_{i-1}^M=\mu_i^M\right\}}1_{\left\{\mu_i^M=\mu_{i+k-1}^M\right\}}.
\end{aligned}$$

Let us now rewrite $\rho_{i,k}^M$ and $\tilde{\rho}_{i,k}^M$. To simplify the notation, we will denote $\sum_{\ell=j_1}^{j_2} w_\ell^M$ for $1 \le j_1 \le j_2 \le M$ by $s_{j_1}^{j_2}$. We can apply Lemma 5.1 to the couples

1. $(p,\ell) = (i, i+k-1)$ so that $1_{\{\mu_i^M = \mu_{i+k-1}^M\}} = 1_{\{s_{i+1}^{i+k-1} < 1-u_i^M\}}$
2. $(p,\ell) = (i-1, i)$ so that $1_{\{\mu_{i-1}^M = \mu_i^M\}} = 1_{\{\omega_i^M < 1-u_{i-1}^M\}}$
3. $(p,\ell) = (i+k-1, i+k)$ so that $1_{\{\mu_{i+k-1}^M = \mu_{i+k}^M\}} = 1_{\{\omega_{i+k}^M < 1-u_{i+k-1}^M\}}$ and $1_{\{\mu_{i+k-1}^M < \mu_{i+k}^M\}} = 1_{\{\omega_{i+k}^M \ge 1-u_{i+k-1}^M\}}$

so that, by observing that when $s_{i+1}^{i+k-1} < 1-u_i^M$ we can rewrite u_{i+k-1}^M as $u_i^M + s_{i+1}^{i+k-1}$, one has

$$\begin{aligned}
\rho_{i,k}^M &= u_i^M \left(1_{\{\omega_{i+k}^M \ge 1-u_{i+k-1}^M\}}(1-u_{i+k-1}^M) + w_{i+k}^M 1_{\{\omega_{i+k}^M < 1-u_{i+k-1}^M\}}\right) 1_{\{s_{i+1}^{i+k-1} < 1-u_i^M\}} \\
&= u_i^M \left(1_{\{\omega_{i+k}^M \ge 1-u_i^M - s_{i+1}^{i+k-1}\}}(1-u_i^M - s_{i+1}^{i+k-1}) + w_{i+k}^M 1_{\{\omega_{i+k}^M < 1-u_i^M - s_{i+1}^{i+k-1}\}}\right) 1_{\left\{s_{i+1}^{i+k_1-1} < 1-u_i^M\right\}} \\
&= u_i^M (1-u_i^M - s_{i+1}^{i+k-1}) 1_{\{s_{i+1}^{i+k-1} < 1-u_i^M \le s_{i+1}^{i+k}\}} + u_i^M w_{i+k}^M 1_{\{s_{i+1}^{i+k} < 1-u_i^M\}} \\
&= u_i^M (1-u_i^M - s_{i+1}^{i+k-1}) 1_{\{s_{i+1}^{i+k-1} < 1-u_i^M\}} - u_i^M (1-u_i^M - s_{i+1}^{i+k}) 1_{\{s_{i+1}^{i+k} < 1-u_i^M\}}
\end{aligned}$$

and, by observing that when $w_i^M < 1-u_{i-1}^M$ and $s_{i+1}^{i+k-1} < 1-u_i^M$ we can rewrite $u_i^M = u_{i-1}^M + w_i^M$ and $u_{i+k-1}^M = u_{i-1}^M + s_i^{i+k-1}$, one has

$$\begin{aligned}
\tilde{\rho}_{i,k}^M &= \left(w_i^M - u_i^M\right) \left(1_{\{\omega_{i+k}^M \ge 1-u_{i+k-1}^M\}}(1-u_{i+k-1}^M) + w_{i+k}^M 1_{\{\omega_{i+k}^M < 1-u_{i+k-1}^M\}}\right) 1_{\{\omega_i^M < 1-u_{i-1}^M\}} 1_{\{s_{i+1}^{i+k-1} < 1-u_i^M\}} \\
&= -\left(u_{i-1}^M + w_i^M\right) \left(1_{\{s_i^{i+k} \ge 1-u_{i-1}^M\}}(1-u_{i-1}^M - s_i^{i+k-1}) + w_{i+k}^M 1_{\{s_i^{i+k} < 1-u_{i-1}^M\}}\right) 1_{\{s_i^{i+k-1} < 1-u_{i-1}^M\}} \\
&\quad + w_i^M \left(1_{\{s_i^{i+k} \ge 1-u_{i-1}^M\}}(1-u_{i-1}^M - s_i^{i+k-1}) + w_{i+k}^M 1_{\{s_i^{i+k} < 1-u_{i-1}^M\}}\right) 1_{\{s_i^{i+k-1} < 1-u_{i-1}^M\}} \\
&= -u_{i-1}^M (1-u_{i-1}^M - s_i^{i+k-1}) 1_{\{s_i^{i+k-1} < 1-u_{i-1}^M \le s_i^{i+k}\}} - u_{i-1}^M w_{i+k}^M 1_{\{s_i^{i+k} < 1-u_{i-1}^M\}} \\
&= -u_{i-1}^M (1-u_{i-1}^M - s_i^{i+k-1}) 1_{\{s_i^{i+k-1} < 1-u_{i-1}^M\}} + u_{i-1}^M (1-u_{i-1}^M - s_i^{i+k}) 1_{\{s_i^{i+k} < 1-u_{i-1}^M\}}.
\end{aligned}$$

Therefore we have obtained that

$$\begin{aligned}
&\frac{1}{M} \sum_{m=1}^M \sum_{1 \le i < j \le M} f(X_i) f(X_j) q_{m,i}^M q_{m,j}^M 1_{\{m = \mu_i^M = \mu_{j-1}^M\}} \\
&\qquad = \frac{1}{M} \sum_{k=1}^{M-1} \sum_{i=1}^{M-k} f(X_i) f(X_{i+k}) \beta_1 \left(u_{i-1}^M, w_i^M, \sum_{\ell=2}^k w_{i+\ell-1}^M, w_{i+k}^M\right)
\end{aligned}$$

with β_1 defined as in (8). Let us now observe that there exists a finite constant C_1 such that for $k = 1, \cdots, M-1$, $i = 1, \cdots, M-k$

$$\left|\beta_1 \left(u_{i-1}^M, w_i^M, \sum_{\ell=2}^k w_{i+\ell-1}^M, w_{i+k}^M\right)\right| \le C_1 1_{\left\{\sum_{\ell=2}^k w_{i+\ell-1}^M < 1\right\}} \le C_1 1_{\left\{k < 1 + \frac{\bar{g}}{\underline{g}}\right\}}. \tag{47}$$

Therefore

$$\frac{1}{M}\sum_{k=1}^{M-1}\sum_{i=1}^{M-k} f(X_i)f(X_{i+k})\beta_1\left(u_{i-1}^M, w_i^M, \sum_{\ell=2}^{k} w_{i+\ell-1}^M, w_{i+k}^M\right)$$
$$= \frac{1}{M}\sum_{k=1}^{(M-1)\wedge\left\lceil\frac{\bar{g}}{\underline{g}}\right\rceil}\sum_{i=1}^{M-k} f(X_i)f(X_{i+k})\beta_1\left(u_{i-1}^M, w_i^M, \sum_{\ell=2}^{k} w_{i+\ell-1}^M, w_{i+k}^M\right).$$

□

6 Asymptotic Variance for the Next Steps

What we have seen so far is the study of the stratified sampling selection step: it is actually part of a more general algorithm where each step consists of a selection part and a mutation part that we are now going to describe more in details (see for instance [22] for a more general version of the algorithm).

For each step $n \geq 0$, we are going to recursively define the selection sequence $\left(Y_n^{M,m}\right)_{1\leq m\leq M}$ and the mutation sequence $\left(X_n^{M,m}\right)_{1\leq m\leq M}$. Let $g_n : \mathbb{R}^d \to (0,\infty), n \geq 0$ be a family of measurable functions such that $0 < \inf_{x\in\mathbb{R}^d} g_n(x) \leq \sup_{x\in\mathbb{R}^d} g_n(x) < \infty$. In what follows we denote $\bar{g}_n := \sup_{x\in\mathbb{R}^d} g_n(x)$ and $\underline{g}_n := \inf_{x\in\mathbb{R}^d} g_n(x)$. Let for $n \geq 0$ and $k \geq 0$

$$\phi_n(k) := \left\lceil \frac{\bar{g}_n}{\underline{g}_n}(1+k)\right\rceil. \tag{48}$$

Moreover let $(Z_k)_{k\geq 0}$ be a $\mathbb{R}^d$-valued time-inhomogeneous Markov chain that is for each bounded measurable function $h : \mathbb{R}^d \to \mathbb{R}$, $\mathbb{E}\left(h\left(Z_k\right)|\sigma\left(Z_0,\cdots,Z_{k-1}\right)\right) = \int_{\mathbb{R}^d} h(x)P_k\left(Z_{k-1},dx\right)$ for a transition kernel P_k possibly depending on k. We denote the law of Z_0 by η.

Initialization: we generate a sequence $(X_0^{M,m})_{1\leq m\leq M}$ of i.i.d. $\mathbb{R}^d$-valued random vectors following the law η. By convention $Y_0^{M,m} := X_0^{M,m}$ for $1 \leq m \leq M$.

From n to n+1: the transition from $\left(Y_n^{M,m}, X_n^{M,m}\right)_{1\leq m\leq M}$ to $\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)_{1\leq m\leq M}$ for $n \geq 0$ consists of two steps.

Selection: we generate the random vectors $\left(Y_{n+1}^{M,m}\right)_{1\leq m\leq M}$ conditionally independent given $\mathcal{F}^n := \sigma\left(\left(X_i^{M,m}, Y_i^{M,m}\right)_{1\leq m\leq M, 0\leq i\leq n}\right)$ by

$$Y_{n+1}^{M,m} = \sum_{\ell=1}^{M} 1_{\left\{\frac{\sum_{j=1}^{\ell-1} g_n\left(X_n^{M,j}\right)}{\sum_{j=1}^{M} g_n\left(X_n^{M,j}\right)} < m - U_n^m \leq \frac{\sum_{j=1}^{\ell} g_n\left(X_n^{M,j}\right)}{\sum_{j=1}^{M} g_n\left(X_n^{M,j}\right)}\right\}} X_n^{M,\ell} \quad m = 1,\cdots,M \tag{49}$$

where $(U_n^m)_{1\leq m\leq M}$ is a sequence of i.i.d. random variables independent of $\mathcal{F}^n$ and distributed according to the uniform law on $(0,1)$.

Mutation: given $\mathcal{G}^{n+1} := \sigma\left(\left(X_i^{M,m}, Y_i^{M,m}\right)_{1\leq m\leq M, 0\leq i\leq n}, \left(Y_{n+1}^{M,m}\right)_{1\leq m\leq M}\right)$ we generate the random vectors $X_{n+1}^{M,m}, 1\leq m\leq M$ conditionally independent and respectively distributed according to the probability measure $P_{n+1}\left(Y_{n+1}^{M,m}, \cdot\right)$

For each $n\geq 0$ we denote the weights by

$$w_n^{M,m} = \frac{Mg_n(X_n^{M,m})}{\sum_{\ell=1}^{M} g_n(X_n^{M,\ell})} \quad m=1,\cdots,M.$$

Moreover let us observe that the selection property for $n=0$ (4) remains valid for each n: $\forall n\geq 0$

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_{n+1}^{M,m}}|\mathcal{F}^n\right) = \sum_{m=1}^{M}\frac{w_n^{M,m}}{M}\delta_{X_n^{M,m}}. \tag{50}$$

The study of the asymptotic variance of the selection part of the first step that is $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f\left(Y_1^{M,m}\right)\right)$ for a given bounded measurable function $f:\mathbb{R}^d\to\mathbb{R}$ was the object of Section 3. Our purpose now is to generalize this result by studying the asymptotic variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f\left(Y_{n+1}^{M,m}\right)$ for $n\geq 1$.
Before doing it let us begin with some notation: for $n\geq 0$, $M\geq 1$ and $k=0,\cdots,M-1$ let

$$\widetilde{\eta}_n^{k,M} := \frac{1}{M}\sum_{i=1}^{M-k}\delta_{\left(Y_n^{M,i},\cdots,Y_n^{M,i+k}\right)}$$

$$\bar{\eta}_n^{k,M} := \frac{1}{M}\sum_{i=1}^{M-k}\delta_{\left(X_n^{M,i},\cdots,X_n^{M,i+k}\right)}.$$

Remark 6.1. *Let $f\in\mathcal{B}_b\left(\mathbb{R}^d\right)$. If $k=0$ and $n=0$, by the classical Strong Law of Large Numbers we have*

$$\lim_{M\to\infty}\widetilde{\eta}_0^{0,M}(f) = \lim_{M\to\infty}\bar{\eta}_0^{0,M}(f) = \bar{\eta}_0^0(f) := \eta(f) = \mathbb{E}\left(f\left(Z_0\right)\right) \quad a.s.$$

It is possible to prove that a Strong Law of Large Numbers holds (see [22, Corollary 7.4.2]) also in the case $k=0$ and $n\geq 1$:

$$\lim_{M\to\infty}\widetilde{\eta}_n^{0,M}(f) = \lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M} f\left(Y_n^{M,i}\right) = \widetilde{\eta}_n^0(f) := \frac{\mathbb{E}\left(f\left(Z_{n-1}\right)\prod_{p=0}^{n-1} g_p\left(Z_p\right)\right)}{\mathbb{E}\left(\prod_{p=0}^{n-1} g_p\left(Z_p\right)\right)} \quad a.s.$$

$$\lim_{M\to\infty}\bar{\eta}_n^{0,M}(f) = \lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M} f\left(X_n^{M,i}\right) = \bar{\eta}_n^0(f) := \frac{\mathbb{E}\left(f\left(Z_n\right)\prod_{p=0}^{n-1} g_p\left(Z_p\right)\right)}{\mathbb{E}\left(\prod_{p=0}^{n-1} g_p\left(Z_p\right)\right)} \quad a.s..$$

We are now going to prove that given $h:\mathbb{R}^{2d}\to\mathbb{R}$ a bounded measurable function, the study of the asymptotic variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)$ depends on the study of the asymptotic variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f\left(Y_{n+1}^{M,m}\right)$ for $f:\mathbb{R}^d\to\mathbb{R}$ a bounded measurable function.

Proposition 6.2. *Let $h \in \mathcal{B}_b\left(\mathbb{R}^{2d}\right)$. The following convergence holds*

$$\left| Var\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)\right) - Var\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} P_{n+1}h\left(Y_{n+1}^{M,m}\right)\right) - \frac{\bar{\eta}_n^0\left(g_n\left(P_{n+1}h^2 - \left(P_{n+1}h\right)^2\right)\right)}{\bar{\eta}_n^0(g_n)}\right| \underset{M\to\infty}{\longrightarrow} 0.$$

Proof of Proposition 6.2. By using that the $X_{n+1}^{M,m}$ are conditionally independent given $\mathcal{G}^{n+1}$ and that the $Y_{n+1}^{M,m}$ are $\mathcal{G}^{n+1}$ measurable, one has

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)\right)\\
&= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} \mathbb{E}\left(h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)\right) + \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} \mathrm{Var}\left(h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)\right)\\
&= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} \mathbb{E}\left(h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)\right) + \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} \mathbb{E}\left(h^2\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)\right)\\
&\quad - \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} \mathbb{E}\left(h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)^2\right).
\end{aligned}$$

Since $\mathcal{L}\left(X_{n+1}^{M,m} \mid \mathcal{G}^{n+1}\right) \sim P_{n+1}\left(Y_{n+1}^{M,m}, \cdot\right)$, for each bounded measurable function $\tilde{h} : \mathbb{R}^{2d} \to \mathbb{R}$ one has

$$\mathbb{E}\left(\tilde{h}(Y_{n+1}^{M,m}, X_{n+1}^{M,m}) \mid \mathcal{G}^{n+1}\right) = P_{n+1}\tilde{h}\left(Y_{n+1}^{M,m}\right) := \int_{\mathbb{R}^d} \tilde{h}(Y_{n+1}^{M,m}, x) P_{n+1}\left(Y_{n+1}^{M,m}, dx\right). \tag{51}$$

Therefore we obtain that

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)\right)\\
&= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} P_{n+1}h\left(Y_{n+1}^{M,m}\right)\right) + \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} P_{n+1}h^2\left(Y_{n+1}^{M,m}\right)\right) - \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\left(P_{n+1}h\left(Y_{n+1}^{M,m}\right)\right)^2\right)\\
&= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} P_{n+1}h\left(Y_{n+1}^{M,m}\right)\right) + \mathbb{E}\left(\sum_{m=1}^{M} \frac{w_n^{M,m}}{M} P_{n+1}h^2\left(X_n^{M,m}\right)\right) - \mathbb{E}\left(\sum_{m=1}^{M} \frac{w_n^{M,m}}{M}\left(Ph\left(X_n^{M,m}\right)\right)^2\right)
\end{aligned}$$

where to obtain the last equality we use the selection property (50). By Remark 6.1,

$$\begin{aligned}
&\lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} g_n\left(X_n^{M,m}\right) P_{n+1}h^2\left(X_n^{M,m}\right) = \bar{\eta}_n^0(g_n P_{n+1}h^2) \quad a.s.,\\
&\lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} g_n\left(X_n^{M,m}\right)\left(P_{n+1}h\left(X_n^{M,m}\right)\right)^2 = \bar{\eta}_n^0(g_n\left(P_{n+1}h\right)^2) \quad a.s.,\\
&\lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} g_n\left(X_n^{M,m}\right) = \bar{\eta}_n^0(g_n) \quad a.s.
\end{aligned}$$

Thus by Lebesgue theorem,

$$\lim_{M\to\infty}\mathbb{E}\left(\sum_{m=1}^{M}\frac{w_n^{M,m}}{M}P_{n+1}h^2\left(X_n^{M,m}\right)\right)=\frac{\bar{\eta}_n^0(g_nP_{n+1}h^2)}{\bar{\eta}_n^0(g_n)},$$
$$\lim_{M\to\infty}\mathbb{E}\left(\sum_{m=1}^{M}\frac{w_n^{M,m}}{M}\left(P_{n+1}h\left(X_n^{M,m}\right)\right)^2\right)=\frac{\bar{\eta}_n^0(g_n\left(P_{n+1}h\right)^2)}{\bar{\eta}_n^0(g_n)}$$

and this concludes the proof. □

We are now ready to study for $n\geq 1$ and $f\in\mathcal{B}_b\left(\mathbb{R}^d\right)$ the convergence as $M\to\infty$ of

$$V_{n+1}^M(f):=\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M}f\left(Y_{n+1}^{M,m}\right)\right).$$

With the aim of studying $\lim_{M\to\infty}V_{n+1}^M(f)$, we will assume the following conjectures:

Conjecture 1. *For each $f\in\mathcal{B}_b\left(\mathbb{R}^d\right)$ one has*

$$\left|Var\left(\sqrt{M}\frac{\bar{\eta}_n^{0,M}(g_nf)}{\bar{\eta}_n^{0,M}(g_n)}\right)-\frac{V_n^M(P_nf_n)}{\left(\bar{\eta}_n^0(g_n)\right)^4}-\frac{\bar{\eta}_{n-1}^0\left(g_{n-1}\left(P_nf_n^2-\left(P_nf_n\right)^2\right)\right)}{\left(\bar{\eta}_n^0(g_n)\right)^4\bar{\eta}_{n-1}^0(g_{n-1})}\right|\underset{M\to\infty}{\longrightarrow}0$$

where

$$f_n:=g_n\left(\bar{\eta}_n^0(g_n)f-\bar{\eta}_n^0(g_nf)\right).$$

Remark 6.3. *In Section 3 we proved that in the case $n=0$, by observing that $Var\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}(g_0f)}{\bar{\eta}_0^{0,M}(g_0)}\right)=Var\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}\left(\frac{g_0}{\eta(g_0)}f\right)}{\bar{\eta}_0^{0,M}\left(\frac{g_0}{\eta(g_0)}\right)}\right)$, one has $\left|Var\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}(g_0f)}{\bar{\eta}_0^{0,M}(g_0)}\right)-\frac{\eta(f_0^2)}{\eta(g_0)^4}\right|\underset{M\to\infty}{\longrightarrow}0.$*

For $n\geq 1$ the idea underlying the Conjecture 1 is the following: by considering

$$Var\left(\sqrt{M}\left(\frac{\bar{\eta}_n^{0,M}(g_nf)}{\bar{\eta}_n^{0,M}(g_n)}-\frac{\bar{\eta}_n^0(g_nf)}{\bar{\eta}_n^0(g_n)}\right)\right)=Var\left(\sqrt{M}\frac{\bar{\eta}_n^{0,M}(f_n)}{\bar{\eta}_n^{0,M}(g_n)\bar{\eta}_n^0(g_n)}\right),$$

we asymptotically replace in the denominator of the right-hand side $\bar{\eta}_n^{0,M}(g_n)$ by $\bar{\eta}_n^0(g_n)$. Thus we have

$$\left|Var\left(\sqrt{M}\frac{\bar{\eta}_n^{0,M}(f_n)}{\bar{\eta}_n^{0,M}(g_n)\bar{\eta}_n^0(g_n)}\right)-\frac{1}{\left(\bar{\eta}_n^0(g_n)\right)^4}Var\left(\sqrt{M}\bar{\eta}_n^{0,M}(f_n)\right)\right|\underset{M\to\infty}{\longrightarrow}0.$$

We then rewrite $Var\left(\sqrt{M}\bar{\eta}_n^{0,M}(f_n)\right)$ following the same reasoning used in the proof of Proposition 6.2.

Conjecture 2. *Let $t\in\mathbb{N}$. Given $\psi\in\mathcal{B}_b\left(\mathbb{R}^{t+2}\right)$ continuous and $h\in\mathcal{B}_b\left(\mathbb{R}^{d(t+1)}\right)$, the following convergence holds*

$$\left|\frac{1}{M}\sum_{m=1}^{M-t}\mathbb{E}\left(h\left(X_n^{M,m},\cdots,X_n^{M,m+t}\right)\psi\left(u_n^{M,m-1},w_n^{M,m},\cdots,w_n^{M,m+t}\right)\right)\right.$$
$$\left.-\frac{1}{M}\sum_{m=1}^{M-t}\mathbb{E}\left(h\left(X_n^{M,m},\cdots,X_n^{M,m+t}\right)\int_0^1\psi\left(u,\tilde{g}_n(X_n^{M,m}),\cdots,\tilde{g}_n(X_n^{M,m+t})\right)du\right)\right|\underset{M\to\infty}{\longrightarrow}0$$

where $u_n^{M,m-1} = \left\{ \frac{M \sum_{i=1}^{m-1} g_n\left(X_n^{M,i}\right)}{\sum_{i=1}^{M} g_n\left(X_n^{M,i}\right)} \right\}$ *and* $\tilde{g}_n(x) = \frac{g_n(x)}{\bar{\eta}_n^0(g_n)}$.

Remark 6.4. *In Section 3 we have provided a formal proof of Conjecture 2 in the case* $n = 0$. *In that case we strongly used the fact that the random variables we are working with* $\left(\left(X_0^{M,m}\right)_{m\geq 1}\right)$ *are i.i.d. and the fact that the law of* $g_0\left(X_0^{M,m}\right)$ *has an absolutely continuous component to prove that is possible to asymptotically replace* $u_0^{M,i-1}$ *with a uniformly distributed random variable independent of* $\mathcal{F}^0$ *and* $w_0^{M,i}$ *with* $\frac{g_0\left(X_0^{M,i}\right)}{\bar{\eta}_0^0(g_0)}$.

Before providing the main result of this section, let us introduce some notation.
For $k \geq 0$ let $P_n^{\otimes(k+1)} : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right)$ be defined by

$$P_n^{\otimes(k+1)} h\left(x_0, \cdots, x_k\right) := \int_{\mathbb{R}^{d(k+1)}} h\left(y_0, \cdots, y_k\right) P_n\left(x_0, dy_0\right) \cdots P_n\left(x_k, dy_k\right)$$

and let for $s_k = 0, \cdots, \phi_{n-1}(k)$, $T_{n-1}^{k\to s_k} : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(s_k+1)}\right)$ given by

$$T_{n-1}^{k\to s_k} f\left(x_0, x_1, \cdots, x_{s_k}\right) = \sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k} f\left(x_0, x_{s_1}, \cdots, x_{s_k}\right) \int_0^1 \psi_{s_1:s_k}\left(u, \tilde{g}_{n-1}(x_0), \tilde{g}_{n-1}(x_1), \cdots, \tilde{g}_{n-1}(x_{s_k})\right) du$$

and where given $0 \leq s_k \leq \phi_{n-1}(k)$ and $s_0 := 0 \leq s_1 \leq s_2 \leq \cdots \leq s_k$:

$$\psi_{s_1:s_k}\left(u, y_0, y_1, \cdots, y_{s_k}\right) = \sum_{i=1}^{\lceil 1+\bar{g}_n/\underline{g}_n \rceil} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{u+\sum_{j=0}^{s_q-1} y_j < u' \leq u+\sum_{j=0}^{s_q} y_j\right\}} du'.$$

Finally let $\mathcal{T}_{n-1}^k : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(\phi_{n-1}(k)+1)}\right)$ defined by

$$\mathcal{T}_{n-1}^k h\left(x_0, \cdots, x_{\phi_{n-1}(k)}\right) = \sum_{s_k=0}^{\phi_{n-1}(k)} T_{n-1}^{k\to s_k}\left(P_n^{\otimes(k+1)} h\right)\left(x_0, x_1, \cdots, x_{s_k}\right).$$

Theorem 6.5. *Let us assume Conjecture 1 and Conjecture 2. Then* $\forall n \geq 1$ *and* $\forall f \in \mathcal{B}\left(\mathbb{R}^d\right)$, $V_n^M(f)$ *converges as* M *goes to infinity and we denote its limit by* $V_n(f)$. *Moreover, by defining* $\phi_{m:n}(0) := \phi_m\left(\phi_{m+1}\left(\cdots\left(\phi_n(0)\right)\right)\right)$ *for* $m \leq n$, $m, n \in \mathbb{N}$, $V_n(f)$ *is defined by the following recursive formula:*

$$\begin{aligned} V_{n+1}(f) = \frac{V_n\left(P_n f_n\right)}{\left(\bar{\eta}_n^0\left(g_n\right)\right)^4} + \frac{\bar{\eta}_{n-1}^0\left(g_{n-1}\left(P_n f_n^2 - \left(P_n f_n\right)^2\right)\right)}{\left(\bar{\eta}_n^0\left(g_n\right)\right)^4 \bar{\eta}_{n-1}^0\left(g_{n-1}\right)} \\ + \int \mathcal{T}_0^{\phi_{1:n}(0)} \cdots \mathcal{T}_{n-2}^{\phi_{n-1:n}(0)} \mathcal{T}_{n-1}^{\phi_{n:n}(0)} \bar{f}_n(x)\, \eta^{\otimes\phi_{0:n}(0)}(dx) \end{aligned}$$

where $\bar{f}_n \in \mathcal{B}_b\left(\mathbb{R}^{d(1+\phi_n(0))}\right)$ *is given by*

$$\bar{f}_n\left(x_0, \cdots, x_{\phi_n(0)}\right) := \sum_{k=0}^{\phi_n(0)} f\left(x_0\right) f\left(x_k\right) \int_0^1 \bar{\beta}_k\left(u, \tilde{g}_n\left(x_0\right), \cdots, \tilde{g}_n\left(x_k\right)\right) du$$

with $\bar{\beta}_k\left(u, y_0, \cdots, y_k\right) = \beta_0(u, y_0) 1_{\{k=0\}} - \beta_1\left(u, y_0, \sum_{\ell=1}^{k-1} y_\ell, y_k\right) 1_{\{k\neq 0\}}$ *where* β_0 *and* β_1 *are respectively defined in (7) and (8).*

Remark 6.6. *Let us observe that by Theorem 2.2,* $\lim_{M\to\infty} V_1^M(h) = V_1(h) = \sigma_1^2(h) + \sigma_2^2(h)\ \forall h \in \mathcal{B}_b\left(\mathbb{R}^d\right)$.

The proof of Theorem 6.5 relies on the following proposition the proof of which is provided after the proof of the theorem.

Proposition 6.7. *Let us assume Conjecture 2. Given* $n \geq 1$ *and* $h \in \mathcal{B}_b\left(\mathbb{R}^{d(\phi_n(0)+1)}\right)$ *one has*

$$\left|\mathbb{E}\left(\bar{\eta}_n^{\phi_n(0),M}(h)\right) - \int \mathcal{T}_0^{\phi_{1:n}(0)} \cdots \mathcal{T}_{n-2}^{\phi_{n-1:n}(0)} \mathcal{T}_{n-1}^{\phi_{n:n}(0)} h(x)\, \eta^{\otimes\phi_{0:n}(0)}(dx)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

Proof of Theorem 6.5. Similarly to what we have done in (5) and (6), we can rewrite the variance in the following way

$$\begin{aligned} V_{n+1}^M(f) &= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m})\right) \\ &= \mathrm{Var}\left(\mathbb{E}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right) \\ &= \mathrm{Var}\left(\sqrt{M}\frac{\sum_{m=1}^M g_n(X_n^{M,m}) f(X_n^{M,m})}{\sum_{m=1}^M g_n(X_n^{M,m})}\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right) \end{aligned} \tag{52}$$

where to obtain the last equality we use the selection property (50). Conjecture 1 gives the asymptotic behavior of the first term of (52).
Let us then study the second term of (52). Using the definition of $Y_{n+1}^{M,m}$ and following the same reasoning done to prove Proposition 3.3, we can rewrite the expression inside the expectation: for $M \geq 1 + \phi_n(0)$

$$\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right) \tag{53}$$

$$= \frac{1}{M}\sum_{k=0}^{\phi_n(0)}\sum_{i=1}^{M-k} f(X_n^{M,i}) f(X_n^{M,i+k}) \bar{\beta}_k\left(u_n^{M,i-1}, w_n^{M,i}, \cdots, w_n^{M,i+k}\right) \tag{54}$$

with $\bar{\beta}_k(u, y_0, \cdots, y_k) := \beta_0(u, y_0) 1_{\{k=0\}} - \beta_1\left(u, y_0, \sum_{\ell=1}^{k-1} y_\ell, y_k\right) 1_{\{k\neq 0\}}$ where we recall that β_0 and β_1 are respectively defined in (7) and (8). Given $k = 0, \cdots, \phi_n(0)$, we can apply Conjecture 2 with $t = k$, $h(x_0, \cdots, x_k) = f(x_0) f(x_k)$ and $\psi(u, y_0, \cdots, y_k) = \bar{\beta}_k(u, y_0, \cdots, y_k)$ and obtain that

$$\left|\mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right) - \sum_{k=0}^{\phi_n(0)} \mathbb{E}\left(\bar{\eta}_n^{k,M}\left(\tilde{f}_k\right)\right)\right| \underset{M\to\infty}{\longrightarrow} 0 \tag{55}$$

where for $k = 0, \cdots, \phi_n(0)$, $\tilde{f}_k(x_0, \cdots, x_k) = f(x_0) f(x_k) \int_0^1 \bar{\beta}_k(u, \tilde{g}_n(x_0), \cdots, \tilde{g}_n(x_k))\, du$.
We now observe that by defining $\bar{f}_n\left(x_0, \cdots, x_{\phi_n(0)}\right) = \sum_{k=0}^{\phi_n(0)} \tilde{f}_k(x_0, \cdots, x_k)$, one has

$$\left|\sum_{k=0}^{\phi_n(0)} \bar{\eta}_n^{k,M}\left(\tilde{f}_k\right) - \bar{\eta}_n^{\phi_n(0),M}\left(\bar{f}_n\right)\right|$$
$$= \frac{1}{M}\left|\sum_{k=0}^{\phi_n(0)}\left(\sum_{i=1}^{M-k}\tilde{f}_k\left(X_n^{M,i},\cdots,X_n^{M,i+k}\right) - \sum_{i=1}^{M-\phi_n(0)}\tilde{f}_k\left(X_n^{M,i},\cdots,X_n^{M,i+k}\right)\right)\right|$$
$$\leq \frac{C}{M}\sum_{k=0}^{\phi_n(0)}\left(\phi_n(0)-k\right) \underset{M\to\infty}{\longrightarrow} 0$$

for a finite constant C. Thus we can combine it with (55) and obtain that

$$\left|\mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_{n+1}^{M,m}) \;\middle|\; \mathcal{F}^n\right)\right) - \mathbb{E}\left(\bar{\eta}_n^{\phi_n(0),M}\left(\bar{f}_n\right)\right)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

We can now apply Proposition 6.7 with $h = \bar{f}_n$ and obtain that

$$\left|\mathbb{E}\left(\bar{\eta}_n^{\phi_n(0),M}\left(\bar{f}_n\right)\right) - \int \mathcal{T}_0^{\phi_{1:n}(0)}\cdots\mathcal{T}_{n-2}^{\phi_{n-1:n}(0)}\mathcal{T}_{n-1}^{\phi_{n:n}(0)}\bar{f}_n(x)\,\eta^{\otimes\phi_{0:n}(0)}(dx)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

Thus we have obtained that $\forall n \geq 1$, $\forall f \in \mathcal{B}_b\left(\mathbb{R}^d\right)$

$$\left|V_{n+1}^M(f) - \frac{V_n^M(P_n f_n)}{\left(\bar{\eta}_n^0(g_n)\right)^4} - \frac{\bar{\eta}_{n-1}^0\left(g_{n-1}\left(P_n f_n^2 - (P_n f_n)^2\right)\right)}{\left(\bar{\eta}_n^0(g_n)\right)^4\bar{\eta}_{n-1}^0(g_{n-1})}\right.$$
$$\left. - \int \mathcal{T}_0^{\phi_{1:n}(0)}\cdots\mathcal{T}_{n-2}^{\phi_{n-1:n}(0)}\mathcal{T}_{n-1}^{\phi_{n:n}(0)}\bar{f}_n(x)\,\eta^{\otimes\phi_{0:n}(0)}(dx)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

To conclude the proof, it is now sufficient to observe that by Theorem 2.2, $\forall h \in \mathcal{B}_b\left(\mathbb{R}^d\right)$ $V_1^M(h)$ converges as M goes to infinity and its limit is given by $V_1(h) = \sigma_1^2(h) + \sigma_2^2(h)$. □

The proof of Proposition 6.7 is a direct consequence of the following Lemma.

Lemma 6.8. *Let $n \geq 0$ and $k \geq 0$. Given $h \in \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right)$ one has*

$$\left|\mathbb{E}\left(\bar{\eta}_{n+1}^{k,M}(h)\right) - \mathbb{E}\left(\bar{\eta}_n^{\phi_n(k),M}(\mathcal{T}_n^k h)\right)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

Proof of Lemma 6.8. Since the $X_{n+1}^{M,m}$ are conditionally independent given $\mathcal{G}^{n+1}$ and $\mathcal{L}\left(X_{n+1}^{M,m} \mid \mathcal{G}^{n+1}\right) \sim P_{n+1}\left(Y_{n+1}^{M,m},\cdot\right)$, for each bounded measurable function $h : \mathbb{R}^{d(k+1)} \to \mathbb{R}$ one has

$$\mathbb{E}\left(h(X_{n+1}^{M,i},\cdots,X_{n+1}^{M,i+k}) \mid \mathcal{G}^{n+1}\right) = P_{n+1}^{\otimes(k+1)}h\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right) \tag{56}$$
$$:= \int_{\mathbb{R}^{d(k+1)}} h(x_0,\cdots,x_k)P_{n+1}\left(Y_{n+1}^{M,i},dx_0\right)\cdots P_{n+1}\left(Y_{n+1}^{M,i+k},dx_k\right). \tag{57}$$

Therefore for $M \geq k+1$

$$\mathbb{E}\left(\bar{\eta}_{n+1}^{k,M}(h)\right) = \frac{1}{M}\sum_{i=1}^{M-k}\mathbb{E}\left(h\left(X_{n+1}^{M,i},\cdots,X_{n+1}^{M,i+k}\right)\right) \tag{58}$$
$$= \frac{1}{M}\sum_{i=1}^{M-k}\mathbb{E}\left(P_{n+1}^{\otimes(k+1)}h\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right)\right) = \mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(P_{n+1}^{\otimes(k+1)}h)\right). \tag{59}$$

Thus our purpose now becomes to study the asymptotic behaviour of $\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right)$ for a given $f \in \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right)$. Recalling the definition (49) of $Y_{n+1}^{M,m}$, let us observe that if we denote by ℓ_m the random index in $\{1,\cdots,M\}$ such that $1_{\left\{\sum_{j=1}^{\ell_m-1} w_n^{M,j} < m - U_n^m \leq \sum_{j=1}^{\ell_m} w_n^{M,j}\right\}} = 1$ so that $Y_{n+1}^{M,m} = X_n^{M,\ell_m}$, one has that $m \mapsto \ell_m$ is non decreasing.
Therefore given $k \geq 0$, $M \geq k+1$ and $1 \leq i \leq M-k$ one has:

$$\begin{aligned}
&\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right)\\
&=\sum_{\ell_0=1}^{M}\cdots\sum_{\ell_k=\ell_{k-1}}^{M}\left(X_n^{M,\ell_0},\cdots,X_n^{M,\ell_k}\right)\prod_{q=0}^{k}1_{\left\{\sum_{j=1}^{\ell_{q-1}} w_n^{M,j} < i+q-U_n^{i+q} \leq \sum_{j=1}^{\ell_q} w_n^{M,j}\right\}}.
\end{aligned}$$

We can now apply the following change of variables $m = \ell_0, s_1 = \ell_1 - \ell_0, s_2 = \ell_2 - \ell_0, \cdots, s_k = \ell_k - \ell_0$ and set $s_0 := 0$ so that the above expression becomes

$$\sum_{0\leq s_k\leq M-1}\ \sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k}\ \sum_{m=1}^{M-s_k}\left(X_n^{M,m},X_n^{M,m+s_1},\cdots,X_n^{M,m+s_k}\right)\prod_{q=0}^{k}1_{\left\{\sum_{j=1}^{m+s_q-1} w_n^{M,j} < i+q-U_n^{i+q} \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}}.$$

Therefore one has

$$\begin{aligned}
\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right) &= \frac{1}{M}\sum_{i=1}^{M-k}\mathbb{E}\left(f\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right)\right) = \frac{1}{M}\sum_{i=1}^{M-k}\mathbb{E}\left(\mathbb{E}\left(f\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right)\mid\mathcal{F}^n\right)\right)\\
&= \frac{1}{M}\sum_{0\leq s_k\leq M-1}\ \sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k}\ \sum_{m=1}^{M-s_k}\mathbb{E}\Bigg(f\left(X_n^{M,m},X_n^{M,m+s_1},\cdots,X_n^{M,m+s_k}\right)\\
&\qquad\times\sum_{i=1}^{M-k}\prod_{q=0}^{k}\int_{i+q-1}^{i+q}1_{\left\{\sum_{j=1}^{m+s_q-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}}du\Bigg)
\end{aligned}$$

where to obtain the last equality we apply the Freezing Lemma and the fact that the sequence $(U_n^m)_{1\leq m\leq M}$ is independent of $\mathcal{F}^n$ and the sequence $\left(X_n^{M,m}\right)_{1\leq m\leq M}$ is $\mathcal{F}^n$ - measurable.
Given $k \geq 0$, $M \geq k+1$, $0 \leq s_k \leq M-1$, $0 \leq s_1 \leq s_2 \leq \cdots \leq s_k$ and $1 \leq m \leq M - s_k$, let us now focus on the sum over i appearing in the above expression:

$$\sum_{i=1}^{M-k}\prod_{q=0}^{k}\int_{i+q-1}^{i+q}1_{\left\{\sum_{j=1}^{m+s_q-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}}du.$$

Let us first observe that if the quantity inside the sum is different from zero, then in particular $\int_{i-1}^{i}1_{\left\{\sum_{j=1}^{m-1} w_n^{M,j} < u \leq \sum_{j=1}^{m} w_n^{M,j}\right\}}du \neq 0$ and $\int_{i+k-1}^{i+k}1_{\left\{\sum_{j=1}^{m+s_k-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_k} w_n^{M,j}\right\}}du \neq 0$. This implies that $\sum_{j=1}^{m} w_n^{M,j} > i-1$ and $\sum_{j=1}^{m+s_k-1} w_n^{M,j} < i+k$. Thus

$$\frac{g_n}{\bar{g}_n}(s_k-1) \leq \sum_{j=m+1}^{m+s_k-1} w_n^{M,j} = \sum_{j=1}^{m+s_k-1} w_n^{M,j} - \sum_{j=1}^{m} w_n^{M,j} < i+k-i+1 = k+1$$

and so we have obtained an upper bound for s_k: $s_k \leq \left\lceil \frac{\bar{g}_n}{\underline{g}_n}(k+1) \right\rceil = \phi_n(k)$ where we use the notation introduced in (48).
Moreover we observe that we can replace the finite sum over i with an infinite sum: if $i \geq M-k+1$

$$\int_{i+k-1}^{i+k} 1_{\left\{ \sum_{j=1}^{m+s_k-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_k} w_n^{M,j} \right\}} du = 0$$

since $\sum_{j=1}^{m+s_k} w_n^{M,j} \leq \sum_{j=1}^{M} w_n^{M,j} = M$.
Therefore after all one has

$$\begin{aligned}
&\sum_{i=1}^{M-k} \int_{i-1}^{i} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{ \sum_{j=1}^{m+s_q-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j} \right\}} du \\
&\qquad = 1_{\{s_k \leq \phi_n(k)\}} \sum_{i \geq 1} \prod_{q=1}^{k} \int_{i+q-1}^{i+q} 1_{\left\{ \sum_{j=1}^{m+s_q-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j} \right\}} du.
\end{aligned}$$

We can now apply the change of variable $u := u' + \left\lfloor \sum_{j=1}^{m-1} w_n^{M,j} \right\rfloor$ in each of the above integrals so to obtain

$$\begin{aligned}
&\sum_{i \geq 1} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{ \sum_{j=1}^{m+s_q-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j} \right\}} du \\
&= \sum_{i \geq 1} \prod_{q=0}^{k} \int_{i+q-1-\left\lfloor \sum_{j=1}^{m-1} w_n^{M,j} \right\rfloor}^{i+q-\left\lfloor \sum_{j=1}^{m-1} w_n^{M,j} \right\rfloor} 1_{\left\{ u_n^{M,m-1} + \sum_{j=m}^{m+s_q-1} w_n^{M,j} < u' \leq u_n^{M,m-1} + \sum_{j=m}^{m+s_q} w_n^{M,j} \right\}} du' \\
&= \sum_{i \geq 1} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{ u_n^{M,m-1} + \sum_{j=m}^{m+s_q-1} w_n^{M,j} < u' \leq u_n^{M,m-1} + \sum_{j=m}^{m+s_q} w_n^{M,j} \right\}} du' \qquad (60)
\end{aligned}$$

where to obtain the last equality we use the fact that if $i < 1$, $\int_{i-1}^{i} 1_{\left\{ u_n^{M,m-1} + \sum_{j=m}^{m-1} w_n^{M,j} < u' \leq u_n^{M,m-1} + \sum_{j=m}^{m} w_n^{M,j} \right\}} du' = 0$. We observe that if the quantity inside the sum is different from zero, then in particular

$$\int_{i-1}^{i} 1_{\left\{ u_n^{M,m-1} < u' \leq u_n^{M,m-1} + w_n^{M,m} \right\}} du' \neq 0.$$

This implies that $i - 1 < u_n^{M,m-1} + w_n^{M,m} \leq 1 + \frac{\bar{g}_n}{\underline{g}_n}$ and so $i \leq \left\lceil 1 + \frac{\bar{g}_n}{\underline{g}_n} \right\rceil$. In conclusion, (60) can be rewritten as

$$\sum_{i=1}^{\lceil 1+\bar{g}_n/\underline{g}_n \rceil} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{ u_n^{M,m-1} + \sum_{j=m}^{m+s_q-1} w_n^{M,j} < u' \leq u_n^{M,m-1} + \sum_{j=m}^{m+s_q} w_n^{M,j} \right\}} du'. \qquad (61)$$

We denote (61) by $\psi_{s_1:s_k}\left(u_n^{M,m-1}, w_n^{M,m}, w_n^{M,m+1}, \cdots, w_n^{M,m+s_k}\right)$ where $\psi_{s_1:s_k} : \mathbb{R}^{s_k+2} \to \mathbb{R}$. Therefore if $M \geq \max\left(1+k, 1+\phi_n(k)\right)$ we have obtained that

$$\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right) \tag{62}$$

$$= \sum_{s_k=0}^{\phi_n(k)} \sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k} \frac{1}{M} \sum_{m=1}^{M-s_k} \mathbb{E}\Bigg(f\left(X_n^{M,m}, X_n^{M,m+s_1}, \cdots, X_n^{M,m+s_k}\right) \tag{63}$$

$$\times \psi_{s_1:s_k}\left(u_n^{M,m-1}, w_n^{M,m}, w_n^{M,m+1}, \cdots, w_n^{M,m+s_k}\right)\Bigg). \tag{64}$$

We can therefore apply Conjecture 2 with $t = s_k$, $h\left(x_0, x_1, \cdots, x_{s_k}\right) = f\left(x_0, x_{s_1}, x_{s_2}, \cdots, x_{s_k}\right)$, $\psi\left(u, y_0, \cdots, y_{s_k}\right) = \psi_{s_1:s_k}\left(u, y_0, \cdots, y_{s_k}\right)$.

$$\left|\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right) - \sum_{s_k=0}^{\phi_n(k)} \mathbb{E}\left(\bar{\eta}_n^{s_k,M}(T_n^{k\to s_k} f)\right)\right| \underset{M\to\infty}{\longrightarrow} 0 \tag{65}$$

where $T_n^{k\to s_k} : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(s_k+1)}\right)$ is defined by

$$T_n^{k\to s_k} f\left(x_0, x_1, \cdots, x_{s_k}\right) = \sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k} f\left(x_0, x_{s_1}, \cdots, x_{s_k}\right) \int_0^1 \psi_{s_1:s_k}\left(u, \tilde{g}_n(x_0), \tilde{g}_n(x_1), \cdots, \tilde{g}_n(x_{s_k})\right) du$$

and where given $0 \leq s_k \leq \phi_n(k)$ and $0 \leq s_1 \leq s_2 \leq \cdots \leq s_k$:

$$\psi_{s_1:s_k}\left(u, y_0, y_1, \cdots, y_{s_k}\right) = \sum_{i=1}^{\lceil 1+\bar{g}_n/\underline{g}_n\rceil} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{u+\sum_{j=0}^{s_{q-1}} y_j < u' \leq u+\sum_{j=0}^{s_q} y_j\right\}} du'.$$

We now observe that by defining $T_n^k : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(\phi_n(k)+1)}\right)$ by

$$T_n^k f\left(x_0, \cdots, x_{\phi_n(k)}\right) = \sum_{s_k=0}^{\phi_n(k)} T_n^{k\to s_k} f\left(x_0, \cdots, x_{s_k}\right),$$

one has

$$\left|\sum_{s_k=0}^{\phi_n(k)} \bar{\eta}_n^{s_k,M}(T_n^{k\to s_k} f) - \bar{\eta}_n^{\phi_n(k),M}(T_n^k f)\right|$$

$$= \frac{1}{M}\left|\sum_{s_k=0}^{\phi_n(k)} \left(\sum_{i=1}^{M-s_k} \left(T_n^{k\to s_k} f\right)\left(X_n^{M,i}, \cdots, X_n^{M,i+s_k}\right) - \sum_{i=1}^{M-\phi_n(k)} \left(T_n^{k\to s_k} f\right)\left(X_n^{M,i}, \cdots, X_n^{M,i+s_k}\right)\right)\right|$$

$$\leq \frac{1}{M} \sum_{s_k=0}^{\phi_n(k)} \left\|T_n^{k\to s_k} f\right\|_\infty \left(\phi_n(k) - s_k\right) \underset{M\to\infty}{\longrightarrow} 0.$$

Combining the above estimate with (65) we obtain that

$$\left|\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right) - \mathbb{E}\left(\bar{\eta}_n^{\phi_n(k),M}(T_n^k f)\right)\right| \underset{M\to\infty}{\longrightarrow} 0. \tag{66}$$

In conclusion, combining (59) with (66) by taking $f = P_{n+1}^{\otimes(k+1)} h$, we obtain that

$$\left| \mathbb{E}\left(\bar{\eta}_{n+1}^{k,M}(h) \right) - \mathbb{E}\left(\bar{\eta}_{n}^{\phi_n(k),M}(\mathcal{T}_n^k h) \right) \right| \underset{M\to\infty}{\longrightarrow} 0$$

where $\mathcal{T}_n^k = T_n^k P_{n+1}^{\otimes(k+1)}$ and this concludes the proof. □

7 Numerical Results

In what follows we fix $d = 1$ and given a sequence $(W_n)_{n\geq 1}$ of independent real-valued random variables distributed according to the uniform law on $(0,1)$, let $Z_{n+1} = Z_n + W_{n+1}$ for $n \geq 0$ with Z_0 distributed according to the uniform law on $(0,1)$ $\left(\eta(dx) = 1_{[0,1]}(x)\, dx\right)$. Thus in this case the transition kernel is given by $P(x, dy) = 1_{[x,x+1]}(y)dy$. Moreover we will fix $g_n(x) = f(x) = e^x \ \forall n \geq 0, \forall x \in \mathbb{R}$.

In this section, we are first going to numerically verify the Conjecture 1 in the case $n = 1$ and the Conjecture 2 in the case $n = 1$ and $n = 2$.
In the second place, we will test the two conjectures together by directly studying the asymptotic variance. As done in the theory, we study numerically the asymptotic behaviour of following expression:

$$V_{n+1}^M(f) = \mathrm{Var}\left(\sqrt{M} \frac{\bar{\eta}_n^{0,M}(g_n f)}{\bar{\eta}_n^{0,M}(g_n)} \right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n \right) \right). \tag{67}$$

We will consider separately the case $n = 0$ and the case $n = 1$. We recall that the first case has been fully studied in Section 3 without the need to introduce any conjecture.

7.1 Verification of the Conjectures

7.1.1 Conjecture 1

We recall the notation

$$f_1 := g_1 \left(\bar{\eta}_1^0(g_1) f - \bar{\eta}_1^0(g_1 f) \right).$$

Since we have fixed $g_n(x) = f(x) = e^x \ \forall n \geq 0, \forall x \in \mathbb{R}$, one has $\bar{\eta}_0^0(g_0) = e - 1$, $\bar{\eta}_1^0(g_1) = \frac{e^2-1}{2}$, $\bar{\eta}_1^0(g_1 f) = \frac{1}{6}(e^3 - 1)(e + 1)$. Moreover $Pf_1(x) = \bar{\eta}_1^0(g_1) \frac{e^{2(x+1)} - e^{2x}}{2} - \bar{\eta}_1^0(fg_1)\left(e^{x+1} - e^x\right)$ and $Pf_1^2(x) = \left(\bar{\eta}_1^0(g_1)\right)^2 \frac{e^{4(x+1)} - e^{4x}}{4} + \left(\bar{\eta}_1^0(fg_1)\right)^2 \frac{e^{2(x+1)} - e^{2x}}{2} - 2\bar{\eta}_1^0(g_1)\bar{\eta}_1^0(fg_1) \frac{e^{3(x+1)} - e^{3x}}{3}$ and

$$\begin{aligned}\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right) = \left(\bar{\eta}_1^0(g_1)\right)^2 \left(\frac{e^5 - 1}{5}\right)\left(\frac{e^2 - 1}{2}\right) + \left(\bar{\eta}_1^0(g_1 f)\right)^2 \left(\frac{e^3 - 1}{3}\right)\left(\frac{e^2 - 1}{2} - (e-1)^2\right) \\ - \bar{\eta}_1^0(g_1)\bar{\eta}_1^0(g_1 f)\left(\frac{e^4 - 1}{4}\right)\left(\frac{2}{3}\left(e^3 - 1\right) - \left(e^2 - 1\right)(e - 1)\right).\end{aligned}$$

We first observe that given a sequence $(T_i)_{i\geq 1}$ of square integrable i.i.d. random variables, by using the delta method it is possible to prove that as M goes to infinity the following convergence in distribution holds:

$$\sqrt{M}\left(\left(\frac{1}{M}\sum_{i=1}^{M} T_i^2 - \left(\frac{1}{M}\sum_{i=1}^{M} T_i \right)^2 \right) - \mathrm{Var}(T_1) \right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \mathrm{Var}\left((T_1 - \mathbb{E}(T_1))^2 \right)\right). \tag{68}$$

The general strategy will be the following: we fix $M = 10000$ and

1. we simulate $n_1 = 10^7$ independent samples T_j of $\sqrt{M}\dfrac{\bar{\eta}_1^{0,M}(g_1 f)}{\bar{\eta}_1^{0,M}(g_1)} = \sqrt{M}\dfrac{\sum_{i=1}^{M}(g_1 f)\left(X_1^{M,i}\right)}{\sum_{i=1}^{M} g_1\left(X_1^{M,i}\right)}$ and we compute, by using (68), the estimator

$$\hat{v}_1 := \frac{1}{n_1}\sum_{i=1}^{n_1} T_i^2 - \left(\frac{1}{n_1}\sum_{i=1}^{n_1} T_i\right)^2 \tag{69}$$

of $\mathrm{Var}\left(\sqrt{M}\dfrac{\bar{\eta}_1^{0,M}(g_1 f)}{\bar{\eta}_1^{0,M}(g_1)}\right)$ with relative 95% confidence interval $[a_1, b_1]$ given by

$$a_1 := \hat{v}_1 - \frac{1.96}{\sqrt{n_1}}\sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}\left(T_i - \frac{1}{n_1}\sum_{j=1}^{n_1} T_j\right)^4 - \left(\frac{1}{n_1}\sum_{i=1}^{n_1}\left(T_i - \frac{1}{n_1}\sum_{j=1}^{n_1} T_j\right)^2\right)^2} \tag{70}$$

$$b_1 := \hat{v}_1 + \frac{1.96}{\sqrt{n_1}}\sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}\left(T_i - \frac{1}{n_1}\sum_{j=1}^{n_1} T_j\right)^4 - \left(\frac{1}{n_1}\sum_{i=1}^{n_1}\left(T_i - \frac{1}{n_1}\sum_{j=1}^{n_1} T_j\right)^2\right)^2}. \tag{71}$$

2. we simulate $n_1 = 10^7$ independent samples H_j of $\dfrac{1}{\sqrt{M}}\sum_{m=1}^{M} Pf_1(Y_1^{M,m})$ and we compute the estimator $\hat{v}_2$ of $V_1^M(Pf_1) = \mathrm{Var}\left(\dfrac{1}{\sqrt{M}}\sum_{m=1}^{M} Pf_1(Y_1^{M,m})\right)$ with relative 95% confidence interval $[a_2, b_2]$ with $\hat{v}_2$, a_2, and b_2 respectively defined as in (69), (70) and (71).
Therefore

$$\frac{\hat{v}_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$$

is an estimator of $\dfrac{V_1^M(Pf_1)}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \dfrac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ with relative 95% confidence interval

$$\left[\frac{a_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}, \frac{b_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}\right].$$

3. we check that $\hat{v}_1$ is close to $\dfrac{\hat{v}_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \dfrac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ as expected

RESULTS

n=1	value	CI (95%)
$\hat{v}_1$	2.8021446	[2.7775619 , 2.8267273]
$\dfrac{\hat{v}_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \dfrac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$	2.7932862	[2.7831461 , 2.8034263]

7.1.2 Conjecture 2

Let $n \in \{1,2\}$ and $t \in \{1,2\}$. Moreover let $h(x_0, \cdots, x_t) = x_0 + \cdots + x_t$ and $\psi(u_0, w_0, \cdots, w_{t+1}) = u_0 + w_0 + \cdots + w_{t+1}$.
The general strategy will be the following: we fix $M = 10000$ and

1. we simulate $n_1 = 10^5$ independent samples T_j of $\frac{1}{M}\sum_{m=1}^{M-t} h\left(X_n^{M,m}, \cdots, X_n^{M,m+t}\right)\psi\left(u_n^{M,m-1}, w_n^{M,m}, \cdots, w_n^{M,m+t}\right)$ and we compute the standard estimator

$$\hat{v}_1 := \frac{1}{n_1}\sum_{j=1}^{n_1} T_j \tag{72}$$

of $\frac{1}{M}\sum_{m=1}^{M-t}\mathbb{E}\left(h\left(X_n^{M,m}, \cdots, X_n^{M,m+t}\right)\psi\left(u_n^{M,m-1}, w_n^{M,m}, \cdots, w_n^{M,m+t}\right)\right)$ with relative 95% confidence interval $[a_1, b_1]$ given by

$$a_1 := \hat{v}_1 - \frac{1.96}{\sqrt{n_1}} * \sqrt{\frac{1}{n_1}\sum_{j=1}^{n_1}\left(T_j - \hat{v}_1\right)^2} \tag{73}$$

$$b_1 := \hat{v}_1 + \frac{1.96}{\sqrt{n_1}} * \sqrt{\frac{1}{n_1}\sum_{j=1}^{n_1}\left(T_j - \hat{v}_1\right)^2} \tag{74}$$

2. we simulate $n_1 = 10^5$ independent samples H_j of

$$\frac{1}{M}\sum_{m=1}^{M-t} h\left(X_n^{M,m}, \cdots, X_n^{M,m+t}\right)\psi\left(U, \tilde{g}_n(X_n^{M,m}), \cdots, \tilde{g}_n(X_n^{M,m+t})\right)$$

where U is a random variable uniformly distributed on $(0,1)$ independent of $\mathcal{F}^n$ and $\tilde{g}_n = \frac{g_n}{\eta_n^0(g_n)}$ $\left(\eta_1^0(g_1) = \frac{e^2-1}{2}, \eta_2^0(g_2) = \frac{e^3-1}{3}\right)$. We then compute the standard estimator $\hat{v}_2$ of

$$\frac{1}{M}\sum_{m=1}^{M-t}\mathbb{E}\left(h\left(X_n^{M,m}, \cdots, X_n^{M,m+t}\right)\psi\left(U, \tilde{g}_n(X_n^{M,m}), \cdots, \tilde{g}_n(X_n^{M,m+t})\right)\right)$$

with relative 95% confidence interval $[a_2, b_2]$ with $\hat{v}_2, a_2, b_2$ respectively defined as in (72), (73) and (74).

3. we check that $\hat{v}_1$ is close to $\hat{v}_2$ as expected

RESULTS

n=1

t=1	value	CI (95%)
$\hat{v}_1$	5.7510732	[5.7509322 , 5.7512143]
$\hat{v}_2$	5.7509738	[5.7470892 , 5.7548583]

t=2	value	CI (95%)
$\hat{v}_1$	11.8853516	[11.8850608 , 11.8856425]
$\hat{v}_2$	11.8835799	[11.8777417 , 11.8894182]

n=2

t=1	value	CI (95%)
$\hat{v}_1$	9.2154319	[9.2152408 , 9.215623]
$\hat{v}_2$	9.2150201	[9.2087834 , 9.2212568]

t=2	value	CI (95%)
$\hat{v}_1$	19.0901421	[19.0897513 , 19.0905329]
$\hat{v}_2$	19.0895089	[19.0801501 , 19.0988676]

7.2 Focus on the Variance

7.2.1 case $n = 0$

We recall the notation $\bar{\eta}_0^0(h) := \eta(h)\ \forall h \in \mathcal{B}_b(\mathbb{R})$ and $f_0 := g_0\left(\bar{\eta}_0^0(g_0)\, f - \bar{\eta}_0^0(g_0 f)\right)$.

By (15) and by observing that $\mathrm{Var}\left(\sqrt{M}\dfrac{\bar{\eta}_0^{0,M}(g_0 f)}{\bar{\eta}_0^{0,M}(g_0)}\right) = \mathrm{Var}\left(\sqrt{M}\dfrac{\bar{\eta}_0^{0,M}\left(\frac{g_0}{\eta(g_0)} f\right)}{\bar{\eta}_0^{0,M}\left(\frac{g_0}{\eta(g_0)}\right)}\right)$, one has

$$\left|\mathrm{Var}\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}(g_0 f)}{\bar{\eta}_0^{0,M}(g_0)}\right) - \frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}\right| \underset{M\to\infty}{\longrightarrow} 0$$

where with the choices made $\bar{\eta}_0^0(g_0) = e-1$, $\bar{\eta}_0^0(g_0 f) = \frac{e^2-1}{2}$, $\eta\left(f_0^2\right) = \left(\bar{\eta}_0^0(g_0)\right)^2 \frac{e^4-1}{4} + \left(\bar{\eta}_0^0(g_0 f)\right)^2 \frac{e^2-1}{2} - 2\bar{\eta}_0^0(g_0)\,\bar{\eta}_0^0(g_0 f)\,\frac{e^3-1}{3}$.

The application of Theorem 3.2 with $g = \tilde{g}_0 = \frac{g_0}{\bar{\eta}_0^0(g_0)}$ gives

$$\left|\mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_1^{M,m}) \;\middle|\; \mathcal{F}^0\right)\right) - \sum_{k=0}^{\phi_0(0)} \mathbb{E}(F_k)\right| \underset{M\to\infty}{\longrightarrow} 0$$

with $(F_k)_{k\in\mathbb{N}}$ given by

$$F_k = \begin{cases} f^2(X_0^{M,1})\beta_0\left(U_1, \tilde{g}_0(X_0^{M,1})\right) & k = 0 \\ -f(X_0^{M,1}) f(X_0^{M,k+1})\beta_1\left(U_1, \tilde{g}_0(X_0^{M,1}), \sum_{\ell=2}^{k} \tilde{g}_0(X_0^{M,\ell}), \tilde{g}_0(X_0^{M,k+1})\right) & k > 0 \end{cases} \tag{75}$$

where $U_1 \sim \mathcal{U}(0,1)$ is independent of $X_0^{M,1}, \cdots, X_0^{M,k+1}$ and β_0 and β_1 are respectively defined in (7) and (8). By using that when U_1 is uniformly distributed on $(0,1)$, $\{U_1 + r\}$ is uniformly distributed on $(0,1)$ for each $r \geq 0$, we can apply the Freezing Lemma to rewrite $\mathbb{E}(F_k)\ \forall k = 0, \cdots, \phi_0(0)$:

$$\sum_{k=0}^{\phi_0(0)} \mathbb{E}(F_k) = \sum_{k=0}^{\phi_0(0)} \mathbb{E}\left(\phi_{k,\tilde{g}_0}\left(X_0^{M,1}, \cdots, X_0^{M,k+1}\right)\right) \tag{76}$$

for measurable functions $\phi_{k,\tilde{g}_0} : \mathbb{R}^{(k+1)} \to \mathbb{R}$ given by

$$\phi_{0,\tilde{g}_0}(x) = \frac{1}{3}\left(f^2(x)\left(1 - 1_{\{\tilde{g}_0(x)<1\}}\left(1 - \tilde{g}_0(x)\right)^3\right)\right) \tag{77}$$

and for each $k \geq 1$

$$\phi_{k,\tilde{g}_0}(x) = \frac{1}{2}\left(f(x_1)f(x_{k+1})1_{\left\{\sum_{i=2}^{k}\tilde{g}_0(x_i)<1\right\}}1_{\left\{\sum_{i=1}^{k+1}\tilde{g}_0(x_i)<1\right\}}\tilde{g}_0(x_1)\tilde{g}_0(x_{k+1})\left(2-2\sum_{i=2}^{k}\tilde{g}_0(x_i)-(\tilde{g}_0(x_1)+\tilde{g}_0(x_{k+1}))\right)\right)$$
$$+\frac{1}{6}\left(f(x_1)f(x_{k+1})\cdot 1_{\left\{\sum_{i=2}^{k}\tilde{g}_0(x_i)<1\right\}}1_{\left\{\sum_{i=1}^{k+1}\tilde{g}_0(x_i)\geq 1\right\}}\left(1-\sum_{i=2}^{k}\tilde{g}_0(x_i)\right)^3\right)$$
$$-\frac{1}{6}\left(f(x_1)f(x_{k_1+1})\cdot 1_{\left\{\sum_{i=2}^{k}\tilde{g}_0(x_i)<1\right\}}1_{\left\{\sum_{i=1}^{k+1}\tilde{g}_0(x_i)\geq 1\right\}}\left(\left(1-\sum_{i=1}^{k}\tilde{g}_0(x_i)\right)^3 1_{\left\{\sum_{i=1}^{k}\tilde{g}_0(x_i)<1\right\}}+\left(1-\sum_{i=2}^{k+1}\tilde{g}_0(x_i)\right)^3 1_{\left\{\sum_{i=2}^{k+1}\tilde{g}_0(x_i)<1\right\}}\right)\right). \tag{78}$$

Thus using numerical methods we are going to check what we already know theoretically that is

$$\left| V_1^M(f) - \frac{\eta\left(f_0^2\right)}{\eta\left(g_0\right)^4} - \sum_{k=0}^{\phi_0(0)}\mathbb{E}\left(\phi_{k,\tilde{g}_0}\left(X_0^{M,1},\cdots,X_0^{M,k+1}\right)\right)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

The general strategy will be the following: we fix $M = 10000$ and

1. we simulate $n_1 = 10^7$ independent samples T_j of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_1^{M,m})$ and we compute the estimator $\hat{v}_1$ of $V_1^M(f) = \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_1^{M,m})\right)$ with relative 95% confidence interval $[a_1, b_1]$ with $\hat{v}_1, a_1, b_1$ respectively defined as in (69), (70) and (71).

   Therefore $\hat{v}_1 - \frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}$ is an estimator of $V_1^M(f) - \frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}$ with relative 95% confidence interval $\left[a_1 - \frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}, b_1 - \frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}\right]$.

2. We simulate $n_2 = 10^5$ independent samples Z_j of $\sum_{k=0}^{\phi_0(0)}\phi_{k,\tilde{g}_0}\left(X_0^{M,1},\cdots,X_0^{M,k+1}\right)$ and we compute the standard estimator $\hat{v}_2$ of $\sum_{k=0}^{\phi_0(0)}\mathbb{E}\left(\phi_{k,\tilde{g}_0}\left(X_0^{M,1},\cdots,X_0^{M,k+1}\right)\right)$ with relative 95% confidence interval $[a_2, b_2]$ with $\hat{v}_2, a_2, b_2$ respectively defined as in (72), (73) and (74).

3. We check that $\hat{v}_1 - \frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}$ is close to $\hat{v}_2$ as expected.

RESULTS

n=0	value	CI (95%)
$\hat{v}_1 - \frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}$	0.07943	[0.079127 , 0.079733]
$\hat{v}_2$	0.0793412	[0.0790773 , 0.0796051]

7.2.2 case $n = 1$

In this case the asymptotic behaviour of the first term of the right-hand side of (67) is given by Conjecture 1

$$\left|\mathrm{Var}\left(\sqrt{M}\frac{\bar{\eta}_1^{0,M}(g_1 f)}{\bar{\eta}_1^{0,M}(g_1)}\right) - \frac{V_1^M(Pf_1)}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2-(Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4\bar{\eta}_0^0(g_0)}\right| \underset{M\to\infty}{\longrightarrow} 0$$

and by (55) the asymptotic behaviour of the second term of the right-hand side of (67) is given by

$$\left|\mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_2^{M,m}) \;\middle|\; \mathcal{F}^1\right)\right) - \sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}\left(\tilde{f}_k\right)\right)\right| \underset{M\to\infty}{\longrightarrow} 0 \tag{79}$$

where for $k = 0, \cdots, \phi_1(0)$,

$$\tilde{f}_k(x_0, \cdots, x_k) = f(x_0) f(x_k) \int_0^1 \bar{\beta}_k(u, \tilde{g}_1(x_0), \cdots, \tilde{g}_1(x_k))\, du$$

with $\bar{\beta}_k(u, y_0, \cdots, y_k) := \beta_0(u, y_0) 1_{\{k=0\}} - \beta_1\left(u, y_0, \sum_{\ell=1}^{k-1} y_\ell, y_k\right) 1_{\{k\neq 0\}}$. We recall that β_0 and β_1 are respectively defined in (7) and (8) and $\tilde{g}_1 = \frac{g_1}{\bar{\eta}_1^0(g_1)}$.
We now observe that

$$\sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}\left(\tilde{f}_k\right)\right) = \sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}\left(\phi_{k,\tilde{g}_1}\right)\right) \tag{80}$$

where the $\phi_{k,\tilde{g}_1} : \mathbb{R}^{(k+1)} \to \mathbb{R}$ are measurable functions defined in (77) and (78).

Thus using numerical methods we are going to check that

$$\left|V_2^M(f) - \frac{V_1^M(Pf_1)}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)} - \sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}\left(\phi_{k,\tilde{g}_1}\right)\right)\right| \underset{M\to\infty}{\longrightarrow} 0$$

so that Conjecture 1 and Conjecture 2 are numerically verified at the same time.

The general strategy will be the following:

1. (a) we simulate $n_1 = 10^7$ independent samples T_j of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_2^{M,m})$ and we compute the estimator $\hat{v}_{1,1}$ of $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_2^{M,m})\right)$ with relative 95% confidence interval $[a_{1,1}, b_{1,1}]$ with $\hat{v}_{1,1}, a_{1,1}$, and $b_{1,1}$ respectively defined as in (69), (70) and (71).

   (b) we simulate $n_1 = 10^7$ independent samples H_j of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} Pf_1(Y_1^{M,m})$ independent of $(T_j)_{1\le j\le n_1}$. We compute the estimator $\hat{v}_{1,2}$ of $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} Pf_1(Y_1^{M,m})\right)$ with relative 95% confidence interval $[a_{1,2}, b_{1,2}]$ with $\hat{v}_{1,2}$, $a_{1,2}$, and $b_{1,2}$ respectively defined as in (69), (70) and (71).

   (c) we compute the estimator $\hat{v}_{1,1} - \frac{\hat{v}_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2-(Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ of

$$V_2^M(f) - \frac{V_1^M(Pf_1)}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$$

   with 90% confidence interval

$$\left[a_{1,1} - \frac{b_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}, b_{1,1} - \frac{a_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}\right].$$

2. We simulate $n_2 = 10^5$ independent samples Z_j of $\sum_{k=0}^{\phi_1(0)} \bar{\eta}_1^{k,M}(\phi_{k,\tilde{g}_1})$ and we compute the standard estimator $\hat{v}_2$ of $\sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}(\phi_{k,\tilde{g}_1})\right)$ with relative 95% confidence interval $[a_2, b_2]$ with $\hat{v}_2, a_2, b_2$ respectively defined as in (72), (73) and (74).

3. We check that $\hat{v}_{1,1} - \frac{\hat{v}_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ is close to $\hat{v}_2$ as expected.

RESULTS

n=1	value	CI
$\hat{v}_{1,1} - \frac{\hat{v}_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$	0.4729737	[0.4690806 , 0.4768669]
$\hat{v}_2$	0.4725217	[0.4724719 , 0.4725714]

References

[1] M. Evans and T. Swartz. Methods for approximating integrals in Statistics with special emphasis on Bayesian integration problems. Statist. Sci., 10, 254–272, 1995.

[2] C. P. Robert and G.Casella. Monte Carlo Statistical Methods. Springer, 2nd ed., 2004.

[3] O. Cappé, E. Moulines and T. Rydén. Inference in Hidden Markov Models. Springer, 2005.

[4] J. Handschin and D. Mayne. Monte Carlo techniques to estimate the conditionnal expectation in multi-stage non-linear filtering. Int. J. Control, 9, 547–559, 1969.

[5] J. Handschin. Monte Carlo techniques for prediction and filtering of non- linear stochastic processes. Automatica, 6, 555–563, 1970.

[6] G. Kitagawa. Monte Carlo filter and smoother for non-gaussian state space models. Journal of Computational and Graphical Statistics, 5 (1), 1-25, 1996.

[7] J. Carpenter, P. Clifford and P. Fearnhead. Improved particle filter for nonlinear problems. IEE Proceedings Radar, Sonar Navigation, 146 (1), 2-7, 1999.

[8] R. Douc, O. Cappé and E. Moulines. Comparison of resampling schemes for particle filtering. Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 64-69, 2005.

[9] N. Gordon, D. Salmond and A. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proceedings F (Radar and Signal Processing), 140 (2), 107-113, 1993.

[10] A. Doucet, N. de Freitas and N.J. Gordon. Sequential Monte Carlo Methods in Practice. Springer-Verlag, 2001.

[11] B. Ristic, M. Arulampalam and A. Gordon. Beyond Kalman Filters: Particle Filters for Target Tracking. Artech House, 2004.

[12] J. Liu and R. Chen. Sequential Monte-Carlo methods for dynamic systems. Journal of the Royal Statistical Society B, 93, 1032–1044, 1998.

[13] P. Fearnhead. Sequential Monte Carlo methods in filter theory. PhD thesis, University of Oxford, 1998.

[14] H. R. Künsch. Recursive Monte-Carlo filters: algorithms and theoretical analysis. The Annals of Statistics, 33 (5), 1983–2021, 2005.

[15] N. Chopin. Central limit theorem for sequential monte carlo methods and its application to bayesian inference. The Annals of Statistics, 32 (6), 2385–2411, 2004.

[16] K. Joag-Dev and F. Proschan. Negative association of random variables with applications. The Annals of Statistics, 11 (1), 286-295, 1983.

[17] N. Chopin, S.S. Singh, T. Soto and M. Vihola. On resampling schemes for particle filters with weakly informative observations. ArXiv, 2022.

[18] P. L'Ecuyer and C. Lemieux. Variance reduction via lattice rules. Management Science, 46 (9), 1214-1235, 2000.

[19] M. Gerber, N. Chopin and N. Whiteley. Negative association, ordering and convergence of resampling methods. The Annals of Statistics, 47 (4), 2236–2260, 2019.

[20] P. Billingsley. Probability and Measure. Wiley series in probability and mathematical statistics: probability and mathematical statistics. A Wiley-Interscience Publication, 3rd edn., 1995.

[21] H. E. Akyuz and H. Gamgam. Robust confidence intervals for the difference of two independent population variances. Hacet. J. Math. Stat., 49 (1), 478-493, 2020.

[22] P. Del Moral. Feynman-Kac formulae. Genealogical and interacting particle systems with applications. Probability and its Applications. Springer, 2004.

[23] D. G. Bonett. Approximate confidence interval for standard deviation of non normal distributions. Computational Statistics and Data Analysis, 50, 775 – 782, 2006.

[24] N. Chopin and O. Papaspiliopoulos. An introduction to sequential Monte Carlo. Springer Series in Statistics. Springer, 2020.

[25] R. Flenghi and B. Jourdain. Convergence to the uniform distribution of vectors of partial sums modulo one with a common factor. ArXiv:2308.01874, 2023.