



# Asymptotic Error Analysis of Multilevel Stochastic Approximations for the Value-at-Risk and Expected Shortfall

Stéphane Crépey, Noufel Frikha, Azar Louzi, Gilles Pagès

## ► To cite this version:

Stéphane Crépey, Noufel Frikha, Azar Louzi, Gilles Pagès. Asymptotic Error Analysis of Multilevel Stochastic Approximations for the Value-at-Risk and Expected Shortfall. 2023. hal-04304985v1

**HAL Id: hal-04304985**

**<https://hal.science/hal-04304985v1>**

Preprint submitted on 24 Nov 2023 (v1), last revised 19 Jul 2024 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Asymptotic Error Analysis of Multilevel Stochastic Approximations for the Value-at-Risk and Expected Shortfall

Stéphane Crépey\*, Noufel Frikha<sup>†</sup>, Azar Louzi<sup>‡</sup>, Gilles Pagès<sup>§</sup>

November 24, 2023

## Abstract

This article is a follow up to Crépey, Frikha, and Louzi (2023), where we introduced a nested stochastic approximation algorithm and its multilevel acceleration for computing the value-at-risk and expected shortfall of a random financial loss. We establish central limit theorems for the renormalized errors associated with both algorithms and their averaged variations. Our findings are substantiated through numerical examples.

**Keywords:** value-at-risk, expected shortfall, nested stochastic approximation, multilevel Monte Carlo, Ruppert & Polyak averaging, convergence rate, central limit theorem.

**MSC:** 65C05, 62L20, 62G32, 91Gxx.

## 1 Introduction

Stochastic approximation (SA) methods aim at finding a root of a given function for which only noisy observations are available. When only approximate samples of the innovation can be generated, multilevel stochastic approximation (MLSA) accelerates the latter by first producing a rough estimate of the root and then adding a series of refinements referred to as levels (Frikha, 2016). They often score a significantly lower complexity than their naive counterparts when used to achieve a given prescribed error (Dereich & Müller-Gronbach, 2019; Frikha, 2016; Giles & Haji-Ali, 2019).

An MLSA scheme was introduced in Crépey, Frikha, and Louzi (2023) for the estimation of the value-at-risk (VaR) and expected shortfall (ES) of a random loss that writes as a conditional expectation, where the multilevel feature deals with the number of inner Monte Carlo samplings for the loss (assumed not available in closed form). VaR and ES currently stand as the most widely used risk metrics in the fields of finance (Basel Committee on Banking Supervision, 2013) and insurance (European Parliament and the Council, 2009; Swiss Federal Office of Private Insurance, 2004). A theoretical analysis of the  $L^2$  error and computational complexity,

---

\*Université Paris Cité, Laboratoire de Probabilités, Statistique et Modélisation (LPSM), CNRS UMR 8001 ([stephane.crepey@lpsm.paris](mailto:stephane.crepey@lpsm.paris)). The research of S. Crépey has benefited from the support of the Chair *Capital Markets Tomorrow: Modeling and Computational Issues* under the aegis of the Institut Europlace de Finance, a joint initiative of Laboratoire de Probabilités, Statistique et Modélisation (LPSM) / Université Paris Cité and Crédit Agricole CIB.

<sup>†</sup>Université Paris 1 Panthéon-Sorbonne, Centre d'Economie de la Sorbonne (CES), 106 Boulevard de l'Hôpital, 75642 Paris Cedex 13 ([noufel.frikha@univ-paris1.fr](mailto:noufel.frikha@univ-paris1.fr)). The research of N. Frikha has benefited from the support of the Institut Europlace de Finance.

<sup>‡</sup>Université Paris Cité, Laboratoire de Probabilités, Statistique et Modélisation (LPSM), CNRS UMR 8001 ([azar.louzi@lpsm.paris](mailto:azar.louzi@lpsm.paris)).

<sup>§</sup>Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation (LPSM), CNRS UMR 8001, case 158, 4 place Jussieu, F-75252 Paris Cedex 5 ([gilles.pages@sorbonne-universite.fr](mailto:gilles.pages@sorbonne-universite.fr)).

along with numerical case studies, have demonstrated the superior effectiveness of the MLSA algorithm when compared to the naively nested approach. However, Crépey, Frikha, and Louzi (2023) does not delve into the joint asymptotic distribution of the renormalized VaR and ES estimation errors. Such asymptotics allow deriving trust regions and confidence intervals for the proposed algorithms. A generic central limit theorem (CLT) for MLSA is already available in Frikha (2016, Theorem 2.11). However, unlike Frikha (2016), our objective function is not strongly convex, rendering the analysis of the convergence rate more delicate. Furthermore, unlike Barrera, Crépey, Diallo, Fort, Gobet, and Stazhynski (2019, Theorems 3.2 and 3.3) (relying on Fort (2015, Theorems 2.1 and 3.2)) and Crépey, Frikha, and Louzi (2023), our MLSA method employs a two-time-scale scheme for the VaR and ES computation. In line with Konda and Tsitsiklis (2004) and Mokkadem and Pelletier (2006), our two-time-scale SA scheme uses a slower learning rate sequence  $(\gamma_n)_{n \geq 1}$  for the VaR estimation, which requires more accuracy and is independent of the ES estimation, and a faster learning rate for the ES estimation to speed it up. But due to the nested nature of our framework, the results of the two aforementioned papers are not directly applicable to our scheme. A thorough analysis is thus necessary to derive the asymptotic joint distribution of our renormalized nested SA (NSA) and MLSA estimators for the (VaR, ES) pair. The optimal rate of convergence is attained by selecting the VaR learning rate as  $\gamma_n = \gamma_1 n^{-1}$  with a constraint on  $\gamma_1$  that involves the density of the financial loss evaluated at the VaR itself which is generally unknown in practice. The Ruppert & Polyak averaging principle as introduced in Polyak and Juditsky (1992) and Ruppert (1991) addresses this classical limitation. In the quest of optimal complexity, we also provide central limit theorems (CLT) associated with the averaged variants of our algorithms.

The paper is structured as follows. After a concise overview of the VaR and ES estimation problem as per Crépey, Frikha, and Louzi (2023), Section 2 establishes a CLT for the renormalized error associated with the nested SA algorithm. The CLT for the averaged version is introduced in Section 3. Section 4 addresses the asymptotic distribution of the renormalized error linked to the MLSA algorithm, which averaged counterpart is dealt with in Section 5. Section 6 discusses the optimal complexities of the various SA schemes. Section 7 presents a numerical case study supporting the theoretical properties established in the paper.

## 2 A Nested Stochastic Approximation for the Value-at-Risk and Expected Shortfall

### 2.1 Approximating the Value-at-Risk and Expected Shortfall

We consider a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  rich enough to support all the random variables defined below. The VaR and ES of a random loss represented by an  $L^1(\mathbb{P})$  real-valued random variable  $X_0$ , at a given confidence level  $\alpha \in (0, 1)$ , can be retrieved as solutions to the problem

$$\min_{\xi \in \mathbb{R}} V_0(\xi), \quad \text{where} \quad V_0(\xi) := \xi + \frac{1}{1-\alpha} \mathbb{E}[(X_0 - \xi)^+]. \quad (2.1)$$

More precisely, if the cdf  $F_{X_0}$  of  $X_0$  is continuous, then, according to Rockafellar and Uryasev (2000),  $V'_0(\xi) = (1-\alpha)^{-1}(F_{X_0}(\xi) - \alpha)$ ,  $\xi \in \mathbb{R}$ , and

$$\xi_\star := \text{VaR}_\alpha(X_0) = \min \text{Argmin } V_0, \quad \chi_\star := \text{ES}_\alpha(X_0) = \min V_0. \quad (2.2)$$

if  $X_0$  admits a continuous pdf  $f_{X_0}$ , then  $V_0$  is twice continuously differentiable on  $\mathbb{R}$  with  $V''_0(\xi) = (1-\alpha)^{-1}f_{X_0}(\xi)$ ,  $\xi \in \mathbb{R}$ .

Hereafter, we assume that

$$X_0 = \mathbb{E}[\varphi(Y, Z)|Y], \quad (2.3)$$

where  $Y$  and  $Z$  are independent  $\mathbb{R}^d$  valued random variables and  $\varphi: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is a measurable function such that  $\varphi(Y, Z) \in L^1(\mathbb{P})$ . In financial applications,  $\varphi(Y, Z)$  would model the future

cash flows of a portfolio,  $Y$  the underlying risk factors at a certain point in time, and  $X_0$  the value of the portfolio at that time.

Under the approach initiated by Bardou, Frikha, and Pagès (2009a) (see also Bardou, Frikha, and Pagès (2009b, 2016), as well as Frikha (2014) for similar analysis on shortfall risk measures), assuming that exact iid samples of  $X_0$  are available, one may compute the couple  $(\xi_\star, \chi_\star)$  using a two-time-scale stochastic approximation algorithm. However, we do not assume a known distribution of  $\varphi(Y, Z)$  conditionally on  $Y$ , hence the conditional expectation is not computable in closed form, and no iid exact simulation of  $X_0$  is available. The brute force, naively nested SA algorithm introduced in Crépey, Frikha, and Louzi (2023) for solving (2.1)–(2.3) is based on the following natural approximation to the problem:

$$\min_{\xi \in \mathbb{R}} V_h(\xi), \quad \text{where} \quad V_h(\xi) := \xi + \frac{1}{1-\alpha} \mathbb{E}[(X_h - \xi)^+]. \quad (2.4)$$

Here  $X_h$  is the empirical mean approximation of  $X_0$ , for some bias parameter  $h = \frac{1}{K} \in \mathcal{H} = \{\frac{1}{K'}, K' \in \mathbb{N}^*\}$ , defined by

$$X_h := \frac{1}{K} \sum_{k=1}^K \varphi(Y, Z^{(k)}), \quad (2.5)$$

where  $\{Z^{(k)}, 1 \leq k \leq K\} \stackrel{\text{iid}}{\sim} Z$  are independent from  $Y$ . If the cdf of  $X_h$  is continuous and increasing, solutions

$$\xi_\star^h = \operatorname{argmin} V_h \quad \text{and} \quad \chi_\star^h = \min V_h \quad (2.6)$$

provide biased estimates of  $(\xi_\star, \chi_\star)$ . For  $h \in \mathcal{H}$ , if the cdf  $F_{X_h}$  of  $X_h$  is continuous, then  $V_h' = (1-\alpha)^{-1}(F_{X_h} - \alpha)$  (and  $V_h'' = (1-\alpha)^{-1}f_{X_h}$  if  $X_h$  admits a continuous pdf  $f_{X_h}$ ).

Defining

$$H_1(\xi, x) = 1 - \frac{1}{1-\alpha} \mathbb{1}_{x \geq \xi} \quad \text{and} \quad H_2(\xi, \chi, x) = \chi - \left( \xi + \frac{1}{1-\alpha} (x - \xi)^+ \right) \quad (2.7)$$

for  $\xi, \chi, x \in \mathbb{R}$ , the couple  $(\xi_\star^h, \chi_\star^h)$  can be approximated by the following nested SA scheme driven by an innovation sequence  $\{X_h^{(n)}, n \geq 1\} \stackrel{\text{iid}}{\sim} X_h$ :

$$\xi_{n+1}^h = \xi_n^h - \gamma_{n+1} H_1(\xi_n^h, X_h^{(n+1)}), \quad (2.8)$$

$$\chi_{n+1}^h = \chi_n^h - \frac{1}{n+1} H_2(\xi_n^h, \chi_n^h, X_h^{(n+1)}), \quad (2.9)$$

$n \in \mathbb{N}$ , starting from real valued random variables  $\xi_0^h$  and  $\chi_0^h = 0$ ,  $\xi_0^h$  being independent from  $\{X_h^{(n)}, n \geq 1\}$  and such that  $\mathbb{E}[|\xi_0^h|^2] < \infty$ , for a positive deterministic sequence of step sizes  $\{\gamma_n, n \geq 1\}$  satisfying  $\sum_{n=1}^\infty \gamma_n = \infty$  and  $\sum_{n=1}^\infty \gamma_n^2 < \infty$ .

## 2.2 Convergence Rate Analysis

**Lemma 2.1** (Giorgi, Lemaire, and Pagès (2020, Lemma 3.2 and Proposition 5.2)).

(i) If  $\mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^p] < \infty$  holds for some  $p > 1$ , then

$$\mathbb{E}[|X_h - X_{h'}|^p] \leq C|h - h'|^\frac{p}{2}, \quad h, h' \in \mathcal{H}.$$

(ii) If the  $X_h$  admit densities  $f_{X_h}$ , bounded uniformly in  $h \in \mathcal{H}$ , and  $\mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^p] < \infty$  for some  $p > 1$ , then, for any  $\xi \in \mathbb{R}$ ,

$$\mathbb{E}[|\mathbb{1}_{X_h > \xi} - \mathbb{1}_{X_{h'} > \xi}|] \leq C(h \vee h')^\frac{p}{2(1+p)}, \quad h, h' \in \mathcal{H}.$$

We now put forward some important assumptions concerning the sequence  $\{X_h, h \in \mathcal{H}\}$ .

**Assumption 1.** (i) For any  $h \in \mathcal{H} \cup \{0\}$ , the cdf  $F_{X_h}$  of  $X_h$  admits the first order Taylor expansion

$$F_{X_h}(\xi) - F_{X_0}(\xi) = v(\xi)h + \epsilon(\xi, h)h, \quad \xi \in \mathbb{R},$$

for some functions  $v, \epsilon(\cdot, h): \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $\int_{\xi_*}^{\infty} v(\xi) d\xi < \infty$  and  $\lim_{\mathcal{H} \ni h \downarrow 0} \epsilon(\xi_*, h) = \lim_{\mathcal{H} \ni h \downarrow 0} \int_{\xi_*}^{\infty} |\epsilon(\xi, h)| d\xi = 0$  for any  $\xi_* \in \text{Argmin } V_0$ .

(ii) For any  $h \in \mathcal{H} \cup \{0\}$ , the law of  $X_h$  admits a continuous density  $f_{X_h}$  with respect to the Lebesgue measure. Moreover, the sequence of functions  $\{f_{X_h}, h \in \mathcal{H}\}$  converges locally uniformly towards  $f_{X_0}$ .

(iii) For any  $R > 0$ , the sequence of functions  $\{f_{X_h}, h \in \mathcal{H}\}$  satisfies

$$\inf_{\substack{h \in \mathcal{H} \\ \xi \in B(\xi_*, R)}} f_{X_h}(\xi) > 0.$$

(iv) The density functions  $\{f_{X_h}, h \in \mathcal{H}\}$  are uniformly bounded and uniformly Lipschitz:

$$\sup_{h \in \mathcal{H}} \{\|f_{X_h}\|_{\infty} + [f_{X_h}]_{\text{Lip}}\} < \infty,$$

where  $[f_{X_h}]_{\text{Lip}}$  denotes the Lipschitz constant of  $f_{X_h}$ .

**Remark 2.1.** Assumptions 1(ii) and (iii) imply that, for any  $\xi_* \in \text{Argmin } V_0$ ,  $f_{X_0}(\xi_*) > 0$ , hence  $V_0''(\xi_*) > 0$  and  $\text{Argmin } V_0$  is reduced to a singleton  $\{\xi_*\}$ .

We recall the following results concerning the weak error and statistical error of the nested SA scheme (2.8)-(2.9).

**Lemma 2.2** (Crépey, Frikha, and Louzi (2023, Proposition 3.1 and Theorem 3.1)).

(i) Suppose that Assumption 1 is satisfied and that the density function  $f_{X_0}$  is positive. Then, as  $\mathcal{H} \ni h \downarrow 0$ , for any  $\xi_*^h \in \text{Argmin } V_h$ ,

$$\xi_*^h - \xi_* = -\frac{v(\xi_*)}{f_{X_0}(\xi_*)}h + o(h), \quad \chi_*^h - \chi_* = -h \int_{\xi_*}^{\infty} \frac{v(\xi)}{1-\alpha} d\xi + o(h).$$

(ii) Suppose that Assumption 1 is satisfied, that the random variable  $\varphi(Y, Z)$  is in  $L^2(\mathbb{P})$ , that

$$\sup_{h \in \mathcal{H}} \mathbb{E} \left[ |\xi_0^h|^2 \exp \left( \frac{4}{1-\alpha} c_{\alpha} \sup_{h \in \mathcal{H}} \|f_{X_h}\|_{\infty} |\xi_0^h| \right) \right] < \infty$$

holds for  $c_{\alpha} = 1 \vee \frac{\alpha}{1-\alpha}$ , and that  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $\gamma_1 > 0$  and  $\beta \in (\frac{1}{2}, 1]$ . Setting

$$\lambda = \inf_{h \in \mathcal{H}} \lambda_h \quad \text{with} \quad \lambda_h = \frac{3}{8} V_h''(\xi_*^h) \wedge \|V_h''\|_{\infty} \frac{V_h''(\xi_*^h)^4}{4[V_h''']_{\text{Lip}}^2} \quad (2.10)$$

and assuming that  $\lambda \gamma_1 > 1$  if  $\beta = 1$ , there exists a constant  $C < \infty$  such that, for any positive integer  $n$ ,

$$\sup_{h \in \mathcal{H}} \mathbb{E}[(\xi_n^h - \xi_*^h)^2] \leq C \gamma_n.$$

We can now state the main result of this part.

**Theorem 2.1.** Suppose that Assumption 1 is satisfied, that

$$\sup_{h \in \mathcal{H}} \mathbb{E} \left[ |\xi_0^h|^2 \exp \left( \frac{4}{1-\alpha} c_\alpha \sup_{h \in \mathcal{H}} \|f_{X_h}\|_\infty |\xi_0^h| \right) \right] < \infty, \quad (2.11)$$

that  $\mathbb{E}[|X_0|^{2+\delta}] < \infty$  and  $\mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{2+\delta}] < \infty$  hold for some  $\delta > 0$ , and that  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $\beta \in (\frac{1}{2}, 1]$ , with  $\lambda\gamma_1 > 1$  if  $\beta = 1$ . Then

$$\begin{pmatrix} h^{-\beta}(\xi_{\lceil h^{-2} \rceil}^h - \xi_\star^h) \\ h^{-1}(\chi_{\lceil h^{-2} \rceil}^h - \chi_\star^h) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_\beta) \quad \text{as } \mathcal{H} \ni h \downarrow 0,$$

where

$$\Sigma_\beta = \begin{pmatrix} \frac{\alpha\gamma_1}{2f_{X_0}(\xi_\star) - \gamma_1^{-1}(1-\alpha)\mathbb{1}_{\beta=1}} & \alpha \frac{\chi_\star - \xi_\star}{f_{X_0}(\xi_\star)} \mathbb{1}_{\beta=1} \\ \alpha \frac{\chi_\star - \xi_\star}{f_{X_0}(\xi_\star)} \mathbb{1}_{\beta=1} & \frac{\mathbb{V}\text{ar}((X_0 - \xi_\star)^+)}{(1-\alpha)^2} \end{pmatrix}. \quad (2.12)$$

*Proof.* We denote by  $C$  a finite positive constant that may change from line to line and does not depend upon  $h$  or  $n$ . We follow a strategy similar to the one used for proving Frikha (2016, Theorem 2.7). For any  $h \in \mathcal{H}$ , let  $\{\mathcal{F}_n^h, n \geq 0\}$  be the filtration defined by  $\mathcal{F}_0^h = \sigma(\xi_0^h, \chi_0^h)$  and, for  $n \geq 1$ ,  $\mathcal{F}_n^h = \sigma(\xi_0^h, \chi_0^h, X_h^{(1)}, \dots, X_h^{(n)})$ . Let

$$g_n^h = (V_0''(\xi_\star) - V_h''(\xi_\star^h))(\xi_{n-1}^h - \xi_\star^h), \quad (2.13)$$

$$r_n^h = V_h''(\xi_\star^h)(\xi_{n-1}^h - \xi_\star^h) - V_h'(\xi_{n-1}^h), \quad (2.14)$$

$$\rho_n^h = V_h'(\xi_{n-1}^h) - V_h'(\xi_\star^h) - (H_1(\xi_{n-1}^h, X_h^{(n)}) - H_1(\xi_\star^h, X_h^{(n)})), \quad (2.15)$$

$$e_n^h = V_h'(\xi_\star^h) - H_1(\xi_\star^h, X_h^{(n)}) = -H_1(\xi_\star^h, X_h^{(n)}). \quad (2.16)$$

The sequence  $\{\xi_n^h, n \geq 0\}$  with dynamics (2.8) can be decomposed as

$$\xi_n^h - \xi_\star^h = (1 - \gamma_n V_0''(\xi_\star))(\xi_{n-1}^h - \xi_\star^h) + \gamma_n g_n^h + \gamma_n r_n^h + \gamma_n \rho_n^h + \gamma_n e_n^h, \quad (2.17)$$

hence

$$\begin{aligned} \xi_n^h - \xi_\star^h &= (\xi_0^h - \xi_\star^h) \Pi_{1:n} + \sum_{k=1}^n \gamma_k \Pi_{k+1:n} g_k^h \\ &\quad + \sum_{k=1}^n \gamma_k \Pi_{k+1:n} r_k^h + \sum_{k=1}^n \gamma_k \Pi_{k+1:n} \rho_k^h + \sum_{k=1}^n \gamma_k \Pi_{k+1:n} e_k^h, \end{aligned} \quad (2.18)$$

where

$$\Pi_{k:n} = \prod_{j=k}^n (1 - \gamma_j V_0''(\xi_\star)) \quad (2.19)$$

(with the convention  $\prod_{\emptyset} = 1$ ). Also, the sequence  $\{\chi_n^h, n \geq 0\}$  with dynamics (2.9) satisfies

$$\begin{aligned} \chi_n^h - \chi_\star^h &= \frac{1}{n} \sum_{k=1}^n \left( \xi_{k-1}^h + \frac{1}{1-\alpha} (X_h^{(k)} - \xi_{k-1}^h)^+ \right) - V_h(\xi_\star^h) \\ &= \frac{1}{n} \sum_{k=1}^n \theta_k^h + \frac{1}{n} \sum_{k=1}^n \zeta_k^h + \frac{1}{n} \sum_{k=1}^n \eta_k^h, \end{aligned} \quad (2.20)$$

where

$$\theta_k^h = \xi_{k-1}^h - \xi_\star^h + \frac{1}{1-\alpha} ((X_h^{(k)} - \xi_{k-1}^h)^+ - (X_h^{(k)} - \xi_\star^h)^+) - (V_h(\xi_{k-1}^h) - V_h(\xi_\star^h)), \quad (2.21)$$

$$\zeta_k^h = V_h(\xi_{k-1}^h) - V_h(\xi_\star^h), \quad (2.22)$$

$$\eta_k^h = \frac{1}{1-\alpha} ((X_h^{(k)} - \xi_\star^h)^+ - \mathbb{E}[(X_h^{(k)} - \xi_\star^h)^+]). \quad (2.23)$$

Before studying each term appearing in (2.18)-(2.20), we provide a useful upper bound on  $\Pi_{k:n}$ . Since  $\lim_{k \rightarrow \infty} \gamma_k = 0$ , there exists  $k_0 \geq 0$  such that for  $j \geq k_0$ ,  $(1 - \gamma_j V_0''(\xi_\star)) > 0$ . Thus, using the inequality  $1 + x \leq \exp(x)$ , for  $x \in \mathbb{R}$ , we obtain

$$\begin{aligned} |\Pi_{k:n}| &= |\Pi_{k:k_0-1}| \prod_{j=k_0 \vee k}^n (1 - \gamma_j V_0''(\xi_\star)) \\ &\leq |\Pi_{k:k_0-1}| \exp \left( -V_0''(\xi_\star) \sum_{j=k_0 \vee k}^n \gamma_j \right) \leq K \exp \left( -V_0''(\xi_\star) \sum_{j=k}^n \gamma_j \right) \end{aligned} \quad (2.24)$$

for  $n$  large enough, where  $K = 1 \vee \max_{1 \leq k \leq k_0} |\Pi_{k:k_0-1}| \exp(V_0''(\xi_\star) \sum_{j=k_0 \wedge k}^{k-1} \gamma_j)$  (with the convention  $\sum_{\emptyset} = 0$ ).

We now let  $n = \lceil h^{-2} \rceil$  and deal with each term appearing in (2.18)-(2.20).

**Step 1. Study of  $\{h^{-\beta}(\xi_0^h - \xi_\star^h) \Pi_{1:\lceil h^{-2} \rceil}, h \in \mathcal{H}\}$ .**

Since, by Assumption 1(ii), the sequence of functions  $\{V_h'' = (1 - \alpha)^{-1} f_{X_h}, h \in \mathcal{H}\}$  converges locally uniformly to  $V_0''$  as  $\mathcal{H} \ni h \downarrow 0$ , and/, by Lemma 2.2(i),  $\lim_{\mathcal{H} \ni h \downarrow 0} \xi_\star^h = \xi_\star$ , we obtain  $\lim_{\mathcal{H} \ni h \downarrow 0} V_h''(\xi_\star^h) = V_0''(\xi_\star)$ . In particular, via Lemma 2.2(ii),

$$2\gamma_1 V_0''(\xi_\star) \geq 2\gamma_1 \inf_{h \in \mathcal{H}} V_h''(\xi_\star^h) \geq \frac{16}{3} \gamma_1 \inf_{h \in \mathcal{H}} \lambda_h \geq \gamma_1 \lambda > 1 \quad \text{if} \quad \beta = 1. \quad (2.25)$$

We then deduce, using (2.24), the inequality  $\gamma_1^{\frac{1}{2}} \gamma_{\lceil h^{-2} \rceil}^{-\frac{1}{2}} \geq h^{-\beta}$  and Lemma A.1(ii), that

$$\limsup_{\mathcal{H} \ni h \downarrow 0} h^{-\beta} |\Pi_{1:\lceil h^{-2} \rceil}| \leq C \gamma_1^{\frac{1}{2}} \limsup_{n \uparrow \infty} \gamma_n^{-\frac{1}{2}} e^{-V_0''(\xi_\star) \sum_{j=1}^n \gamma_j} = 0.$$

Owing to Lemma 2.2(i),  $\lim_{\mathcal{H} \ni h \downarrow 0} \xi_\star^h = \xi_\star$ , so that  $\{\xi_\star^h, h \in \mathcal{H}\}$  is bounded which, combined with (2.11), gives  $\sup_{h \in \mathcal{H}} \mathbb{E}[|\xi_0^h - \xi_\star^h|] \leq \sup_{h \in \mathcal{H}} \mathbb{E}[|\xi_0^h|] + \sup_{h \in \mathcal{H}} |\xi_\star^h| < \infty$ . We thus conclude

$$h^{-\beta} (\xi_0^h - \xi_\star^h) \Pi_{1:\lceil h^{-2} \rceil} \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as} \quad \mathcal{H} \ni h \downarrow 0.$$

**Step 2. Study of  $\{h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} g_k^h, h \in \mathcal{H}\}$ .**

It follows from Lemma 2.2(ii) and (2.24) that

$$\mathbb{E} \left[ \left| \sum_{k=1}^n \gamma_k \Pi_{k+1:n} g_k^h \right| \right] \leq C |V_0''(\xi_\star) - V_h''(\xi_\star^h)| \sum_{k=1}^n \gamma_k^{\frac{3}{2}} e^{-V_0''(\xi_\star) \sum_{j=k+1}^n \gamma_j}.$$

Since  $2\gamma_1 V_0''(\xi_\star) > 1$  if  $\beta = 1$ , recalling (2.25), one can thus apply Lemma A.1(i), so that

$$\limsup_{\mathcal{H} \ni h \downarrow 0} h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k^{\frac{3}{2}} e^{-V_0''(\xi_\star) \sum_{j=k+1}^{\lceil h^{-2} \rceil} \gamma_j} \leq C \gamma_1^{\frac{1}{2}} \limsup_{n \uparrow \infty} \gamma_n^{-\frac{1}{2}} \sum_{k=1}^n \gamma_k^{\frac{3}{2}} e^{-V_0''(\xi_\star) \sum_{j=k+1}^n \gamma_j} \leq C.$$

Now, using the fact that  $\lim_{\mathcal{H} \ni h \downarrow 0} |V_0''(\xi_\star) - V_h''(\xi_\star^h)| = 0$ , we conclude

$$h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} g_k^h \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as} \quad \mathcal{H} \ni h \downarrow 0.$$

**Step 3. Study of  $\{h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} r_k^h, h \in \mathcal{H}\}$ .**

Using the fact that  $V_h'(\xi_\star^h) = 0$ , a first order Taylor expansion gives

$$V_h'(\xi) = V_h''(\xi_\star^h)(\xi - \xi_\star^h) + (\xi - \xi_\star^h) \int_0^1 (V_h''(\xi_\star^h + t(\xi - \xi_\star^h)) - V_h''(\xi_\star^h)) dt, \quad (2.26)$$

so that the uniform Lipschitz regularity of  $\{V_h'' = (1 - \alpha)^{-1} f_{X_h}, h \in \mathcal{H}\}$  yields

$$\mathbb{E}[|r_k^h|] \leq \frac{\sup_{h' \in \mathcal{H}} [f_{X_{h'}}]_{\text{Lip}}}{2(1 - \alpha)} \mathbb{E}[(\xi_{k-1}^h - \xi_\star^h)^2]. \quad (2.27)$$

Hence, by Lemma 2.2(ii), for  $k \geq 2$ ,  $\mathbb{E}[|r_k^h|] \leq C\gamma_k$ . Recalling (2.25) if  $\beta = 1$ , using the inequality (2.24) and Lemma A.1(i) gives

$$\limsup_{\mathcal{H} \ni h \downarrow 0} \mathbb{E} \left[ \left| h^{-\beta} \sum_{k=2}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} r_k^h \right| \right] \leq C \limsup_{n \uparrow \infty} \gamma_n^{-\frac{1}{2}} \sum_{k=2}^n \gamma_k^2 e^{-V_0''(\xi_\star) \sum_{j=k+1}^n \gamma_j} = 0.$$

Notice that (2.11) gives  $\sup_{h \in \mathcal{H}} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2] \leq 2(\sup_{h \in \mathcal{H}} \mathbb{E}[|\xi_0^h|^2] + \sup_{h \in \mathcal{H}} |\xi_\star^h|^2) < \infty$ . Hence, the inequalities (2.27) and (2.24) and Lemma A.1(ii) guarantee that

$$\limsup_{\mathcal{H} \ni h \downarrow 0} \mathbb{E}[|h^{-\beta} \gamma_1 \Pi_{1:\lceil h^{-2} \rceil} r_1^h|] \leq C \sup_{h \in \mathcal{H}} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2] \limsup_{n \uparrow \infty} \gamma_n^{-\frac{1}{2}} e^{-V_0''(\xi_\star) \sum_{j=1}^n \gamma_j} = 0.$$

Putting together the two previous limits

$$h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} r_k^h \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

**Step 4. Study of  $\{h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} \rho_k^h, h \in \mathcal{H}\}$ .**

As  $V_h'(\xi) = \mathbb{E}[H_1(\xi, X_h)]$ , we note that (2.15) implies that  $\mathbb{E}[\rho_k^h | \mathcal{F}_{k-1}^h] = 0$ . Hence,  $\{\rho_k^h, k \geq 1\}$  is a sequence of  $\{\mathcal{F}_k^h, k \geq 1\}$ -martingale increments. We thus obtain

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_{k=1}^n \gamma_k \Pi_{k+1:n} \rho_k^h \right)^2 \right] &= \sum_{k=1}^n \gamma_k^2 |\Pi_{k+1:n}|^2 \mathbb{E}[|\rho_k^h|^2] \\ &\leq \sum_{k=1}^n \gamma_k^2 |\Pi_{k+1:n}|^2 \mathbb{E}[(H_1(\xi_{k-1}^h, X_h^{(k)}) - H_1(\xi_\star^h, X_h^{(k)}))^2]. \end{aligned}$$

Recalling the definition (2.7), by the uniform boundedness of  $\{f_{X_h}, h \in \mathcal{H}\}$ ,

$$\begin{aligned} &\mathbb{E}[(H_1(\xi_{k-1}^h, X_h^{(k)}) - H_1(\xi_\star^h, X_h^{(k)}))^2] \\ &= \frac{1}{(1 - \alpha)^2} \mathbb{E} \left[ \left| \mathbb{1}_{X_h^{(k)} \geq \xi_{k-1}^h} - \mathbb{1}_{X_h^{(k)} \geq \xi_\star^h} \right|^2 \right] \\ &= \frac{1}{(1 - \alpha)^2} \mathbb{E} \left[ \mathbb{E} \left[ \mathbb{1}_{\xi_{k-1}^h \leq X_h^{(k)} < \xi_\star^h} + \mathbb{1}_{\xi_\star^h \leq X_h^{(k)} < \xi_{k-1}^h} \middle| \mathcal{F}_{k-1}^h \right] \right] \\ &= \frac{1}{(1 - \alpha)^2} \mathbb{E}[|F_{X_h}(\xi_{k-1}^h) - F_{X_h}(\xi_\star^h)|] \\ &\leq \frac{\sup_{h' \in \mathcal{H}} \|f_{X_{h'}}\|_\infty}{(1 - \alpha)^2} \mathbb{E}[(\xi_{k-1}^h - \xi_\star^h)^2]^{\frac{1}{2}}. \end{aligned} \quad (2.28)$$



Thus, recalling that  $\sup_{h \in \mathcal{H}} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2] < \infty$  and using Lemma 2.2(ii), we obtain  $\mathbb{E}[(H_1(\xi_{k-1}^h, X_h^{(k)}) - H_1(\xi_\star^h, X_h^{(k)}))^2] \leq C\gamma_k^{\frac{1}{2}}$ , for  $k \geq 1$ . Since  $\gamma_1^{\frac{1}{2}}\gamma_{\lceil h^{-2} \rceil}^{-\frac{1}{2}} \geq h^{-\beta}$  and  $2\gamma_1 V_0''(\xi_\star) > 1$  if  $\beta = 1$ , recalling (2.25), by combining (2.24) with Lemma A.1(i), we obtain

$$\limsup_{\mathcal{H} \ni h \downarrow 0} \mathbb{E} \left[ \left( h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1: \lceil h^{-2} \rceil} \rho_k^h \right)^2 \right] \leq C \limsup_{n \uparrow \infty} \gamma_n^{-1} \sum_{k=1}^n \gamma_k^{\frac{5}{2}} e^{-2V_0''(\xi_\star) \sum_{j=1}^n \gamma_j} = 0,$$

which eventually yields

$$h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1: \lceil h^{-2} \rceil} \rho_k^h \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

**Step 5 Study of  $\left\{ \frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \theta_k^h, h \in \mathcal{H} \right\}$ .**

Observe that from the very definition (2.4) of  $V_h$ , (2.21) gives  $\mathbb{E}[\theta_k^h | \mathcal{F}_{k-1}^h] = 0$ , i.e.  $\{\theta_k^h, k \geq 1\}$  is a sequence of  $\{\mathcal{F}_k^h, k \geq 1\}$ -martingale increments. Hence, using Lemma 2.2(ii), we obtain

$$\begin{aligned} \mathbb{E} \left[ \left( \frac{1}{\sqrt{n}} \sum_{k=1}^n \theta_k^h \right)^2 \right] &= \frac{1}{n} \sum_{k=1}^n \mathbb{E}[|\theta_k^h|^2] \\ &\leq \frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ \left| \xi_{k-1}^h - \xi_\star^h + \frac{1}{1-\alpha} ((X_h^{(k)} - \xi_{k-1}^h)^+ - (X_h^{(k)} - \xi_\star^h)^+) \right|^2 \right] \\ &\leq \frac{C}{n} \sum_{k=0}^{n-1} \mathbb{E}[(\xi_k^h - \xi_\star^h)^2] \\ &\leq C \left( \frac{1}{n} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2] + \frac{1}{n} \sum_{k=1}^{n-1} \gamma_k \right). \end{aligned} \tag{2.29}$$

A comparison between series and integrals together with the fact that  $\sup_{h \in \mathcal{H}} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2] < \infty$  yield

$$\frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \theta_k^h \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

**Step 6. Study of  $\left\{ \frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \zeta_k^h, h \in \mathcal{H} \right\}$ .**

Using a second order Taylor expansion, the uniform boundedness of  $\{V_h'' = (1-\alpha)^{-1} f_{X_h}, h \in \mathcal{H}\}$  and Lemma 2.2(ii), we obtain

$$\begin{aligned} \mathbb{E} \left[ \left| \frac{1}{\sqrt{n}} \sum_{k=1}^n \zeta_k^h \right|^2 \right] &\leq \frac{\sup_{h' \in \mathcal{H}} \|f_{X_{h'}}\|_\infty}{2(1-\alpha)} \left( \frac{1}{\sqrt{n}} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2] + \frac{1}{\sqrt{n}} \sum_{k=2}^n \mathbb{E}[(\xi_{k-1}^h - \xi_\star^h)^2] \right) \\ &\leq C \left( \frac{1}{\sqrt{n}} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2] + \frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} \gamma_k \right). \end{aligned} \tag{2.30}$$

A comparison between series and integrals gives  $\lim_{n \uparrow \infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} \gamma_k = 0$ . Since, in addition,  $\sup_{h \in \mathcal{H}} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2] < \infty$ , we conclude that

$$\frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \zeta_k^h \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

**Step 7. Study of**  $\left\{ \left( h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} e_k^h, \frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \eta_k^h \right), h \in \mathcal{H} \right\}$ .

The purpose here is to apply the central limit theorem (Hall & Heyde, 1980, Corollary 3.1, page 58) to the martingale array  $\left\{ \left( h^{-\beta} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} e_k^h, \frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \eta_k^h \right), h \in \mathcal{H} \right\}$ . We first check the conditional Lindeberg condition. By the assumption on the random variables  $\{X_h, h \in \mathcal{H}\}$ , there exists  $\delta > 0$  such that

$$\sup_{h \in \mathcal{H}} \mathbb{E}[|X_h - \xi_\star^h|^{2+\delta}] \leq 2^{1+\delta} \left( \sup_{h \in \mathcal{H}} \mathbb{E}[|X_h|^{2+\delta}] + \sup_{h \in \mathcal{H}} |\xi_\star^h|^{2+\delta} \right) < \infty.$$

From the definition (2.7) of  $H_1$ , applying (2.24) and then Lemma A.1(i) (note in particular that  $(2+\delta)V_0''(\xi_\star)\gamma_1 > 1 + \delta/2$  if  $\beta = 1$ , recalling (2.25)) gives

$$\begin{aligned} \limsup_{\mathcal{H} \ni h \downarrow 0} \sum_{k=1}^{\lceil h^{-2} \rceil} \mathbb{E}[|h^{-\beta} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} H_1(\xi_\star^h, X_h^{(k)})|^{2+\delta}] \\ \leq \gamma_1^{1+\frac{\delta}{2}} c_\alpha^{2+\delta} \limsup_{n \uparrow \infty} \gamma_n^{-(1+\frac{\delta}{2})} \sum_{k=1}^n \gamma_k^{2+\delta} e^{-(2+\delta)V_0''(\xi_\star) \sum_{j=k+1}^n \gamma_j} = 0, \end{aligned}$$

where  $c_\alpha = 1 \vee \alpha/(1-\alpha)$ . Moreover, it holds

$$\sum_{k=1}^{\lceil h^{-2} \rceil} \mathbb{E}\left[\left|\frac{h^{-1}}{\lceil h^{-2} \rceil} \eta_k^h\right|^{2+\delta}\right] \leq C h^{2+\delta} \sum_{k=1}^{\lceil h^{-2} \rceil} \mathbb{E}[|X_h - \xi_\star^h|^{2+\delta}] \leq C \sup_{h \in \mathcal{H}} \mathbb{E}[|X_h - \xi_\star^h|^{2+\delta}] h^\delta,$$

so that  $\limsup_{\mathcal{H} \ni h \downarrow 0} \sum_{k=1}^{\lceil h^{-2} \rceil} \mathbb{E}\left[\left|\frac{h^{-1}}{\lceil h^{-2} \rceil} \eta_k^h\right|^{2+\delta}\right] = 0$ . Hence, the conditional Lindeberg condition is satisfied.

We now prove the convergence, as  $\mathcal{H} \ni h \downarrow 0$ , of the conditional covariance matrices sequence  $\{S_h = (S_h^{i,j})_{1 \leq i,j \leq 2}, h \in \mathcal{H}\}$  defined by

$$\begin{aligned} S_h^{1,1} &:= \sum_{k=1}^{\lceil h^{-2} \rceil} h^{-2\beta} \gamma_k^2 \Pi_{k+1:\lceil h^{-2} \rceil}^2 \mathbb{E}[H_1(\xi_\star^h, X_h)^2] = \frac{\alpha}{1-\alpha} h^{-2\beta} \gamma_{\lceil h^{-2} \rceil} \Sigma_{\lceil h^{-2} \rceil}, \\ S_h^{2,2} &:= \sum_{k=1}^{\lceil h^{-2} \rceil} \mathbb{E}\left[\left(\frac{h^{-1}}{\lceil h^{-2} \rceil} \eta_k^h\right)^2 \middle| \mathcal{F}_{k-1}^h\right] = \frac{h^{-2}}{\lceil h^{-2} \rceil} \frac{\text{Var}((X_h - \xi_\star^h)^+)}{(1-\alpha)^2}, \\ S_h^{1,2} = S_h^{2,1} &:= \frac{h^{-(1+\beta)}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} \mathbb{E}[e_k^h \eta_k^h | \mathcal{F}_{k-1}^h], \end{aligned}$$

with

$$\Sigma_n = \frac{1}{\gamma_n} \sum_{k=1}^n \gamma_k^2 \Pi_{k+1:n}^2. \quad (2.31)$$

We first study the convergence of  $\{S_h^{1,1}, h \in \mathcal{H}\}$ . We first remark that  $\lim_{\mathcal{H} \ni h \downarrow 0} h^{-2\beta} \gamma_{\lceil h^{-2} \rceil} = \gamma_1$ ,  $\lim_{\mathcal{H} \ni h \downarrow 0} S_h^{1,1} = \frac{\gamma_1 \alpha}{1-\alpha} \lim_{n \uparrow \infty} \Sigma_n$ , provided that the latter limit exists. Now observe that

$$\begin{aligned} \Sigma_{n+1} &= \gamma_{n+1} + \frac{\gamma_n}{\gamma_{n+1}} (1 - \gamma_{n+1} V_0''(\xi_\star))^2 \Sigma_n \\ &= \Sigma_n + \frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}} \Sigma_n + \gamma_n \gamma_{n+1} V_0''(\xi_\star)^2 \Sigma_n + (\gamma_{n+1} - \gamma_n) + \gamma_n (1 - 2V_0''(\xi_\star) \Sigma_n). \end{aligned}$$

Asymptotically,

$$\frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}} = \frac{\mathbf{1}_{\beta=1}}{\gamma_1} \gamma_n + o(\gamma_n), \quad \gamma_{n+1} - \gamma_n = o(\gamma_n) \quad \text{and} \quad \gamma_n \gamma_{n+1} = o(\gamma_n).$$

Hence

$$\Sigma_{n+1} = \Sigma_n + \left(1 - \left(2V_0''(\xi_\star) - \frac{\mathbb{1}_{\beta=1}}{\gamma_1}\right)\Sigma_n\right)\gamma_n + (\Sigma_n + 1)o(\gamma_n).$$

The best candidate limit for  $\{\Sigma_n, n \geq 1\}$  is

$$\Sigma_\star := \frac{1}{2V_0''(\xi_\star) - \frac{\mathbb{1}_{\beta=1}}{\gamma_1}}. \quad (2.32)$$

Let  $\Delta\Sigma_n := \Sigma_n - \Sigma_\star$ . One computes

$$\begin{aligned} \Delta\Sigma_{n+1} &= \Delta\Sigma_n + \left(\frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}} - \frac{\mathbb{1}_{\beta=1}}{\gamma_1}\gamma_n\right)\Delta\Sigma_n + \gamma_n\gamma_{n+1}\Delta\Sigma_n - \gamma_n\left(2V_0''(\xi_\star) - \frac{\mathbb{1}_{\beta=1}}{\gamma_1}\right)\Delta\Sigma_n \\ &\quad + \left(\frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}} - \frac{\mathbb{1}_{\beta=1}}{\gamma_1}\gamma_n\right)\Sigma_\star + \gamma_n\gamma_{n+1}\Sigma_\star + (\gamma_{n+1} - \gamma_n) \\ &= (1 - \mu\gamma_n + o(\gamma_n))\Delta\Sigma_n + o(\gamma_n), \end{aligned}$$

where  $\mu := 2V_0''(\xi_\star) - \frac{\mathbb{1}_{\beta=1}}{\gamma_1} > 0$ . Let  $\varepsilon > 0$ . There exists  $n_0 \geq 0$  such that, for  $n \geq n_0$ ,  $1 - (\mu + \varepsilon)\gamma_n > 0$  and

$$|\Delta\Sigma_{n+1}| \leq (1 - (\mu + \varepsilon)\gamma_n)|\Delta\Sigma_n| + \varepsilon\gamma_n. \quad (2.33)$$

Thus, for  $n \geq n_0$ ,

$$|\Delta\Sigma_n| \leq |\Delta\Sigma_{n_0}| \exp\left(-(\mu + \varepsilon) \sum_{k=n_0}^n \gamma_k\right) + \varepsilon \sum_{k=n_0}^n \gamma_k \exp\left(-(\mu + \varepsilon) \sum_{j=k}^n \gamma_j\right).$$

By Lemma 2.1(i),  $\limsup_{n \uparrow \infty} |\Delta\Sigma_n| \leq C\varepsilon$ , so that

$$\lim_{n \uparrow \infty} \Sigma_n = \Sigma_\star. \quad (2.34)$$

Hence

$$S_h^{1,1} \xrightarrow{\mathbb{P}\text{-as}} \frac{\alpha}{1 - \alpha} \frac{\gamma_1}{2V_0''(\xi_\star) - \frac{\mathbb{1}_{\beta=1}}{\gamma_1}} \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

We then study the convergence of  $\{S_h^{2,2}, h \in \mathcal{H}\}$ . Given that

$$\mathbb{E}[|(X_h - \xi_\star^h)^+ - (X_0 - \xi_\star)^+|^2] \leq 2(\mathbb{E}[(X_h - X_0)^2] + (\xi_\star^h - \xi_\star)^2),$$

with  $\lim_{\mathcal{H} \ni h \downarrow 0} \xi_\star^h = \xi_\star$  and

$$\mathbb{E}[(X_h - X_0)^2] = h \mathbb{E}[\text{Var}(\varphi(Y, Z)|Y)] \rightarrow 0 \quad \text{as } \mathcal{H} \ni h \downarrow 0,$$

it follows that  $(X_h - \xi_\star^h)^+$  converges to  $(X_0 - \xi_\star)^+$  in  $L^2(\mathbb{P})$  as  $\mathcal{H} \ni h \downarrow 0$ , so that

$$\text{Var}((X_h - \xi_\star^h)^+) \rightarrow \text{Var}((X_0 - \xi_\star)^+) \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

Finally

$$S_h^{2,2} \xrightarrow{\mathbb{P}\text{-as}} \frac{\text{Var}((X_0 - \xi_\star)^+)}{(1 - \alpha)^2} \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

We conclude by studying the convergence of  $\{S_h^{1,2}, h \in \mathcal{H}\}$ . Recalling that  $\{(X_h - \xi_\star^h)^+, h \in \mathcal{H}\}$  converges to  $(X_0 - \xi_\star)^+$  in  $L^2(\mathbb{P})$ , we deduce that

$$\mathbb{E}[e_k^h \eta_k^h | \mathcal{F}_{k-1}^h] = \frac{\alpha}{(1 - \alpha)^2} \mathbb{E}[(X_h - \xi_\star^h)^+] \rightarrow \frac{\alpha}{(1 - \alpha)^2} \mathbb{E}[(X_0 - \xi_\star)^+] \quad \text{as } \mathcal{H} \ni h \downarrow 0. \quad (2.35)$$

Besides, a direct computation gives  $\lim_{\mathcal{H} \ni h \downarrow 0} \frac{h^{-(1+\beta)}}{\lceil h^{-2} \rceil} = 1_{\beta=1}$ . Hence, by Lemma A.1(i),

$$\lim_{\mathcal{H} \ni h \downarrow 0} \frac{h^{-(1+\beta)}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} \left| \mathbb{E}[e_k^h \eta_k^h | \mathcal{F}_{k-1}^h] - \frac{\alpha}{(1-\alpha)^2} \mathbb{E}[(X_0 - \xi_\star)^+] \right| = 0,$$

so that

$$\lim_{\mathcal{H} \ni h \downarrow 0} \frac{h^{-(1+\beta)}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \gamma_k \Pi_{k+1:\lceil h^{-2} \rceil} \mathbb{E}[e_k^h \eta_k^h | \mathcal{F}_{k-1}^h] = \frac{\alpha \mathbb{E}[(X_0 - \xi_\star)^+] \mathbb{1}_{\beta=1}}{(1-\alpha)^2} \lim_{n \uparrow \infty} \tilde{\Sigma}_n,$$

where

$$\tilde{\Sigma}_n := \sum_{k=1}^n \gamma_k \Pi_{k+1:n}.$$

Observe now that

$$\tilde{\Sigma}_{n+1} = \gamma_{n+1} + (1 - \gamma_{n+1} V_0''(\xi_\star)) \tilde{\Sigma}_n = \tilde{\Sigma}_n + \gamma_{n+1} (1 - V_0''(\xi_\star) \tilde{\Sigma}_n).$$

We let  $\tilde{\Sigma}_\star := \frac{1}{V_0''(\xi_\star)}$  and  $\Delta \tilde{\Sigma}_n := \tilde{\Sigma}_n - \tilde{\Sigma}_\star$ . Hence,

$$\Delta \tilde{\Sigma}_{n+1} = (1 - \gamma_{n+1} V_0''(\xi_\star)) \Delta \tilde{\Sigma}_n = -\tilde{\Sigma}_\star \prod_{k=1}^{n+1} (1 - \gamma_k V_0''(\xi_\star)),$$

so that

$$|\Delta \tilde{\Sigma}_n| \leq |\tilde{\Sigma}_\star| \exp \left( -V_0''(\xi_\star) \sum_{k=1}^n \gamma_k \right).$$

Since  $\sum_{n \geq 1} \gamma_n = \infty$ ,

$$\lim_{n \uparrow \infty} \sum_{k=1}^n \gamma_k \Pi_{k+1:n} = \tilde{\Sigma}_\star,$$

which eventually yields

$$S_h^{1,2} = S_h^{2,1} \xrightarrow{\mathbb{P}\text{-as}} \frac{\alpha \mathbb{E}[(X_0 - \xi_\star)^+] \mathbb{1}_{\beta=1}}{V_0''(\xi_\star)(1-\alpha)^2} \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

The proof is now complete. □

**Corollary 2.1.** *Assume that the hypotheses of Theorem 2.1 are satisfied. Then*

$$\begin{pmatrix} h^{-\beta} (\xi_{\lceil h^{-2} \rceil}^h - \xi_\star) \\ h^{-1} (\chi_{\lceil h^{-2} \rceil}^h - \chi_\star) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N} \left( \begin{pmatrix} -\frac{v(\xi_\star)}{f_{X_0}(\xi_\star)} \mathbb{1}_{\beta=1} \\ -\int_{\xi_\star}^\infty \frac{v(\xi)}{1-\alpha} d\xi \end{pmatrix}, \Sigma_\beta \right) \quad \text{as } \mathcal{H} \ni h \downarrow 0,$$

where we recall that  $\Sigma_\beta$  is given by (2.12).

**Remark 2.2.** (Barrera et al., 2019, Algorithm 1) differs from our NSA scheme in two ways. First, they use a single-time-scale scheme for both (VaR, ES) components, while we use a two-time-scale scheme with a slower, more precise VaR component independent of the ES scheme and then a faster ES component with optimal convergence rate. Second, the  $n$ th iteration of their NSA algorithm uses  $h_n$  instead of the positive constant  $h$  in (2.8)-(2.9), for some sequence  $\{h_n, n \geq 1\}$  that tends to 0 as  $n \uparrow \infty$ . Their approach builds iterates such that  $(\xi_n^{h_n}, \chi_n^{h_n}) \rightarrow (\xi_\star, \chi_\star)$  as  $n \uparrow \infty$ , allowing them to use (Fort, 2015, Theorem 2.1) for obtaining their unbiased

CLT (Barrera et al., 2019, Theorem 3.2). However, their approach results in an algorithm with a substantially increased complexity. Indeed, in view of Barrera et al. (2019, Theorem 3.2), the error of their algorithm after  $n$  iterations is of order  $\sqrt{\gamma_n} = \gamma_1^{1/2} n^{-\beta/2}$ ,  $\beta \in (\frac{1}{2}, 1]$ . For a prescribed error order  $\varepsilon > 0$ , one has to choose  $n = \lceil C\varepsilon^{-2/\beta} \rceil$ . The discussion of (Barrera et al., 2019, Section 3.2) advises to take  $h_n = \lceil Cn^{-\beta'} \rceil$ ,  $\beta' > \beta$ . This results in a complexity of order  $C \sum_{k=1}^n h_k^{-1} = C\varepsilon^{-2(1+\beta')/\beta}$ , which is optimal for  $\beta' \rightarrow \beta = 1$ , yielding  $C\varepsilon^{-4}$ . As will be discussed in Section 6, the NSA algorithm proposed in Crépey, Frikha, and Louzi (2023), instead, has an optimal complexity of  $C\varepsilon^{-3}$ .

*Proof.* For  $h \in \mathcal{H}$ , we write

$$\begin{pmatrix} h^{-\beta}(\xi_{\lceil h^{-2} \rceil}^h - \xi_\star) \\ h^{-1}(\chi_{\lceil h^{-2} \rceil}^h - \chi_\star) \end{pmatrix} = \begin{pmatrix} h^{-\beta}(\xi_{\lceil h^{-2} \rceil}^h - \xi_\star^h) \\ h^{-1}(\chi_{\lceil h^{-2} \rceil}^h - \chi_\star^h) \end{pmatrix} + \begin{pmatrix} h^{-\beta}(\xi_\star^h - \xi_\star) \\ h^{-1}(\chi_\star^h - \chi_\star) \end{pmatrix}.$$

From Lemma 2.2(i),

$$\begin{pmatrix} h^{-\beta}(\xi_\star^h - \xi_\star) \\ h^{-1}(\chi_\star^h - \chi_\star) \end{pmatrix} \rightarrow - \begin{pmatrix} \frac{v(\xi_\star)}{f_{X_0}(\xi_\star)} \mathbf{1}_{\beta=1} \\ \int_{\xi_\star}^\infty \frac{v(\xi)}{1-\alpha} d\xi \end{pmatrix} \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

Combining the previous result and Theorem 2.1 yields the desired result.  $\square$

### 3 Averaged Nested Stochastic Approximation Algorithm

According to Theorem 2.1 and Corollary 2.1, the best rate of convergence in the CLT is achieved by setting  $\beta = 1$ , that is, by taking  $\gamma_n = \gamma_1 n^{-1}$  for the nested VaR stochastic algorithm. Note carefully nonetheless that the choice  $\beta = 1$  and the ensuing rate are only available under the constraint  $\lambda\gamma_1 > 1$ . However, one does not have access to the value of  $\lambda$ . So typically,  $\gamma_1$  must be fine-tuned empirically, adding a computational burden to the implementation of the algorithm.

To avoid this fine tuning issue of the VaR SA scheme, we propose to explore the Ruppert-Polyak averaging principle (Frikha, 2016; Polyak & Juditsky, 1992; Ruppert, 1991). We follow the footsteps of (Barrera et al., 2019; Frikha, 2016) and define, for a given bias parameter  $h \in \mathcal{H}$ , the new sequence of averaged VaR estimators

$$\bar{\xi}_n^h = \frac{1}{n} \sum_{k=1}^n \xi_k^h = \left(1 - \frac{1}{n}\right) \bar{\xi}_{n-1}^h + \frac{1}{n} \xi_n^h, \quad n \geq 1, \quad (3.1)$$

where  $\bar{\xi}_0^h = 0$ ,  $\xi_0^h$  is a real valued random variable satisfying  $\mathbb{E}[|\xi_0^h|^2] < \infty$ , and the estimates  $\{\xi_n^h, n \geq 0\}$  are obtained similarly to the scheme (2.8) with a step sequence  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $\beta \in (\frac{1}{2}, 1)$ .

**Remark 3.1.** Unlike (Barrera et al., 2019, Theorem 3.3), we do not average out the ES nested SA estimators  $\{\chi_n^h, 0 \leq n \leq \lceil h^{-2} \rceil\}$  inasmuch as, in view of Theorem 2.1, their convergence rate of order  $h$  is already optimal.

**Theorem 3.1.** Assume that the hypotheses of Theorem 2.1 are satisfied and that  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $\gamma_1 > 0$  and  $\beta \in (\frac{1}{2}, 1)$ . Then

$$h^{-1} \begin{pmatrix} \bar{\xi}_{\lceil h^{-2} \rceil}^h - \xi_\star^h \\ \chi_{\lceil h^{-2} \rceil}^h - \chi_\star^h \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \bar{\Sigma}) \quad \text{as } \mathcal{H} \ni h \downarrow 0,$$

where

$$\bar{\Sigma} = \begin{pmatrix} \frac{\alpha(1-\alpha)}{f_{X_0}(\xi_\star)^2} & \frac{\alpha}{1-\alpha} \frac{\mathbb{E}[(X_0 - \xi_\star)^+]}{f_{X_0}(\xi_\star)} \\ \frac{\alpha}{1-\alpha} \frac{\mathbb{E}[(X_0 - \xi_\star)^+]}{f_{X_0}(\xi_\star)} & \frac{\mathbb{V}\text{ar}((X_0 - \xi_\star)^+)}{(1-\alpha)^2} \end{pmatrix}. \quad (3.2)$$

*Proof.* We employ the same notation as in the proof of Theorem 2.1. From the decomposition (2.17), we have

$$\xi_n^h - \xi_\star^h = \frac{1}{V_h''(\xi_\star^h)} a_n^h + \frac{1}{V_h''(\xi_\star^h)} r_n^h + \frac{1}{V_h''(\xi_\star^h)} \rho_n^h + \frac{1}{V_h''(\xi_\star^h)} e_n^h, \quad (3.3)$$

where the definitions for  $\{r_n^h, n \geq 1\}$ ,  $\{\rho_n^h, n \geq 1\}$  and  $\{e_n^h, n \geq 1\}$  are provided by (2.14)–(2.16) and  $\{a_n^h, n \geq 1\}$  is given by

$$a_n^h = -\frac{1}{\gamma_n} (\xi_n^h - \xi_\star^h - (\xi_{n-1}^h - \xi_\star^h)). \quad (3.4)$$

From (3.1) and (3.3), we readily get that, for  $n \geq 1$ ,

$$\bar{\xi}_n^h - \xi_\star^h = \frac{1}{V_h''(\xi_\star^h)n} \sum_{k=1}^n a_k^h + \frac{1}{V_h''(\xi_\star^h)n} \sum_{k=1}^n r_k^h + \frac{1}{V_h''(\xi_\star^h)n} \sum_{k=1}^n \rho_k^h + \frac{1}{V_h''(\xi_\star^h)n} \sum_{k=1}^n e_k^h. \quad (3.5)$$

We now let  $n = \lceil h^{-2} \rceil$  and study each term of the above decomposition.

**Step 1. Study of  $\left\{ \frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} a_k^h, h \in \mathcal{H} \right\}$ .**

By summing by parts we obtain

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n \frac{1}{\gamma_k} (\xi_k^h - \xi_\star^h - (\xi_{k-1}^h - \xi_\star^h)) &= \frac{1}{n} \left( \frac{1}{\gamma_n} (\xi_n^h - \xi_\star^h) - \frac{1}{\gamma_1} (\xi_0^h - \xi_\star^h) \right) \\ &\quad + \frac{1}{n} \sum_{k=2}^n \left( \frac{1}{\gamma_{k-1}} - \frac{1}{\gamma_k} \right) (\xi_{k-1}^h - \xi_\star^h). \end{aligned} \quad (3.6)$$

We deal with each term on the right hand side separately. On the one hand, recalling that  $\sup_{h \in \mathcal{H}} \mathbb{E}[|\xi_0^h - \xi_\star^h|] < \infty$ , it follows from Lemma 2.2(ii) that

$$\begin{aligned} \mathbb{E} \left[ \left| \frac{1}{n} \left( \frac{1}{\gamma_n} (\xi_n^h - \xi_\star^h) - \frac{1}{\gamma_1} (\xi_0^h - \xi_\star^h) \right) \right| \right] &\leq \frac{C}{n} \left( \frac{1}{\gamma_n} \mathbb{E}[(\xi_n^h - \xi_\star^h)^2]^{\frac{1}{2}} + \mathbb{E}[|\xi_0^h - \xi_\star^h|] \right) \\ &\leq \frac{C}{n} \left( \frac{1}{\sqrt{\gamma_n}} + 1 \right). \end{aligned}$$

On the other hand, using again Lemma 2.2(ii) and a comparison between series and integrals, we obtain

$$\begin{aligned} \mathbb{E} \left[ \left| \frac{1}{n} \sum_{k=2}^n \left( \frac{1}{\gamma_{k-1}} - \frac{1}{\gamma_k} \right) (\xi_{k-1}^h - \xi_\star^h) \right| \right] &\leq \frac{1}{n} \sum_{k=2}^n \left( \frac{1}{\gamma_{k-1}} - \frac{1}{\gamma_k} \right) \mathbb{E}[(\xi_{k-1}^h - \xi_\star^h)^2]^{\frac{1}{2}} \\ &\leq \frac{C}{n} \sum_{k=2}^n \left( \frac{1}{\gamma_{k-1}} - \frac{1}{\gamma_k} \right) \gamma_{k-1}^{\frac{1}{2}} \\ &\leq \frac{C}{n\sqrt{\gamma_n}}. \end{aligned}$$

Gathering the previous upper bounds,

$$\mathbb{E} \left[ \left| \frac{1}{n} \sum_{k=1}^n \frac{1}{\gamma_k} a_k^h \right| \right] \leq \frac{C}{n\sqrt{\gamma_n}}. \quad (3.7)$$

Now, recalling that  $\lim_{\mathcal{H} \ni h \downarrow 0} V_h''(\xi_\star^h) = V_0''(\xi_\star)$  and that  $\beta \in (\frac{1}{2}, 1)$ , we obtain

$$\frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} a_k^h \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

**Step 2. Study of**  $\left\{ \frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} r_k^h, h \in \mathcal{H} \right\}$ .

It follows from the inequality (2.27) and Lemma 2.2(ii) that for  $k \geq 1$

$$\mathbb{E}[|r_k^h|] \leq C \mathbb{E}[(\xi_{k-1}^h - \xi_\star^h)^2] \leq C\gamma_k, \quad (3.8)$$

so that

$$\mathbb{E} \left[ \left| \frac{1}{\sqrt{n}} \sum_{k=1}^n r_k^h \right| \right] \leq \frac{C}{\sqrt{n}} \sum_{k=1}^n \gamma_k.$$

A comparison between series and integrals eventually yields

$$\frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} r_k^h \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as} \quad \mathcal{H} \ni h \downarrow 0.$$

**Step 3. Study of**  $\left\{ \frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \rho_k^h, h \in \mathcal{H} \right\}$ .

Using (2.28) and Lemma 2.2(ii) yields

$$\mathbb{E}[|\rho_k^h|^2] \leq C \mathbb{E}[(\xi_{k-1}^h - \xi_\star^h)^2]^{\frac{1}{2}} \leq C\gamma_k^{\frac{1}{2}}. \quad (3.9)$$

Recalling that  $\{\rho_k^h, k \geq 1\}$  is a sequence of martingale increments, this implies that

$$\mathbb{E} \left[ \left| \frac{1}{\sqrt{n}} \sum_{k=1}^n \rho_k^h \right|^2 \right] \leq \frac{C}{n} \sum_{k=1}^n \gamma_k^{\frac{1}{2}}.$$

Using comparison between series and integrals, we conclude that

$$\frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \rho_k^h \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as} \quad \mathcal{H} \ni h \downarrow 0.$$

In view of the study of the different terms of the decomposition (2.20) in the proof of Theorem 2.1, it suffices to establish a CLT for the corresponding martingale arrays. This is the goal of the next step.

**Step 4. Study of**  $\left\{ \left( \frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} e_k^h, \frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \eta_k^h \right), h \in \mathcal{H} \right\}$ .

We apply again the CLT for martingale arrays (Hall & Heyde, 1980, Corollary 3.1) to the sequence

$$\left\{ \left( \frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} e_k^h, \frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \eta_k^h \right), h \in \mathcal{H} \right\},$$

recalling that the sequence  $\{\eta_n^h, n \geq 1\}$  is defined by (2.23). We start by checking the conditional Lindeberg condition. Note that Step 7 of the proof of Theorem 2.1 already guarantees that the sequence  $\left\{ \frac{h^{-1}}{\lceil h^{-2} \rceil} \sum_{k=1}^{\lceil h^{-2} \rceil} \eta_k^h, h \in \mathcal{H} \right\}$  satisfies this condition. From the definition (2.7) of  $H_1$  and (2.16), for any  $\delta > 0$ , we have

$$\begin{aligned} \sum_{k=1}^{\lceil h^{-2} \rceil} \mathbb{E} \left[ \left| \frac{h^{-1}}{V_h''(\xi_\star^h) \lceil h^{-2} \rceil} e_k^h \right|^{2+\delta} \right] &\leq \sum_{k=1}^{\lceil h^{-2} \rceil} \frac{\mathbb{E}[|e_k^h|^{2+\delta}]}{|V_h''(\xi_\star^h)|^{2+\delta} \lceil h^{-2} \rceil^{1+\frac{\delta}{2}}} \\ &\leq \frac{c_\alpha^{2+\delta}}{|V_h''(\xi_\star^h)|^{2+\delta} \lceil h^{-2} \rceil^{\frac{\delta}{2}}} \rightarrow 0 \quad \text{as} \quad \mathcal{H} \ni h \downarrow 0, \end{aligned}$$

recalling that  $c_\alpha = 1 \vee \alpha/(1 - \alpha)$ . We now prove the convergence of the conditional covariance matrices sequence  $\{S_h = (S_h^{i,j})_{1 \leq i,j \leq 2}, h \in \mathcal{H}\}$ , defined by

$$\begin{aligned} S_h^{1,1} &:= \sum_{k=1}^{\lceil h^{-2} \rceil} \frac{h^{-2}}{V_h''(\xi_\star^h)^2 \lceil h^{-2} \rceil^2} \mathbb{E}[|e_k^h|^2 | \mathcal{F}_{k-1}^h] = \frac{\alpha}{1 - \alpha} \sum_{k=1}^{\lceil h^{-2} \rceil} \frac{h^{-2}}{V_h''(\xi_\star^h)^2 \lceil h^{-2} \rceil^2}, \\ S_h^{2,2} &:= \sum_{k=1}^{\lceil h^{-2} \rceil} \mathbb{E} \left[ \left( \frac{h^{-1}}{\lceil h^{-2} \rceil} \eta_k^h \right)^2 \middle| \mathcal{F}_{k-1}^h \right] = \frac{h^{-2}}{\lceil h^{-2} \rceil} \frac{\mathbb{V}\text{ar}((X_h - \xi_\star^h)^+)}{(1 - \alpha)^2}, \\ S_h^{1,2} = S_h^{2,1} &:= \frac{h^{-2}}{V_h''(\xi_\star^h)^2 \lceil h^{-2} \rceil^2} \sum_{k=1}^{\lceil h^{-2} \rceil} \mathbb{E}[e_k^h \eta_k^h | \mathcal{F}_{k-1}^h]. \end{aligned}$$

Recalling again that  $\lim_{\mathcal{H} \ni h \downarrow 0} V_h''(\xi_\star^h) = V_0''(\xi_\star) = \frac{f_{X_0}(\xi_\star)}{1 - \alpha}$ , we readily get

$$S_h^{1,1} \xrightarrow{\mathbb{P}\text{-as}} \frac{\alpha(1 - \alpha)}{f_{X_0}(\xi_\star)^2} \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

Note that the asymptotic behavior of the sequence  $\{S_h^{2,2}, h \in \mathcal{H}\}$  has been studied in Step 7 of the proof of Theorem 2.1.

From (2.35) and the fact that  $\lim_{\mathcal{H} \ni h \downarrow 0} V_h''(\xi_\star^h) = V_0''(\xi_\star) = \frac{f_{X_0}(\xi_\star)}{1 - \alpha}$ , Cesàro's lemma yields

$$S_h^{1,2} = S_h^{2,1} \xrightarrow{\mathbb{P}\text{-as}} \frac{\alpha}{1 - \alpha} \frac{\mathbb{E}[(X_0 - \xi_\star)^+]}{f_{X_0}(\xi_\star)} \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

The proof is now complete. □

**Corollary 3.1.** *Assume that the hypotheses of Theorem 3.1 are satisfied. Then*

$$h^{-1} \begin{pmatrix} \bar{\xi}_{\lceil h^{-2} \rceil}^h - \xi_\star \\ \chi_{\lceil h^{-2} \rceil}^h - \chi_\star \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N} \left( \begin{pmatrix} -\frac{v(\xi_\star)}{f_{X_0}(\xi_\star)} \\ -\int_{\xi_\star}^\infty \frac{v(\xi)}{1 - \alpha} d\xi \end{pmatrix}, \bar{\Sigma} \right) \quad \text{as } \mathcal{H} \ni h \downarrow 0,$$

where  $\bar{\Sigma}$  is given by (3.2).

See Remark 2.2 for a comment on the absence of bias in (Barrera et al., 2019, Theorem 3.3) and their reliance on the unbiased averaged CLT (Fort, 2015, Theorem 3.2).

*Proof.* For  $h \in \mathcal{H}$ , we decompose

$$\begin{pmatrix} \bar{\xi}_{\lceil h^{-2} \rceil}^h - \xi_\star \\ \chi_{\lceil h^{-2} \rceil}^h - \chi_\star \end{pmatrix} = \begin{pmatrix} \bar{\xi}_{\lceil h^{-2} \rceil}^h - \xi_\star^h \\ \chi_{\lceil h^{-2} \rceil}^h - \chi_\star^h \end{pmatrix} + \begin{pmatrix} \xi_\star^h - \xi_\star \\ \chi_\star^h - \chi_\star \end{pmatrix}.$$

By Lemma 2.2(i),

$$h^{-1} \begin{pmatrix} \xi_\star^h - \xi_\star \\ \chi_\star^h - \chi_\star \end{pmatrix} \rightarrow - \begin{pmatrix} \frac{v(\xi_\star)}{f_{X_0}(\xi_\star)} \\ \int_{\xi_\star}^\infty \frac{v(\xi)}{1 - \alpha} d\xi \end{pmatrix} \quad \text{as } \mathcal{H} \ni h \downarrow 0.$$

Using Theorem 3.1 and the result above allows concluding the proof. □



## 4 Multilevel Stochastic Approximation Algorithm

Let  $h_0 = \frac{1}{K} \in \mathcal{H}$ . Define the sequence of bias parameters  $\{h_\ell, 0 \leq \ell \leq L\}$  by

$$h_\ell = \frac{h_0}{M^\ell} = \frac{1}{KM^\ell} \in \mathcal{H}, \quad 0 \leq \ell \leq L, \quad (4.1)$$

for some fixed integer  $M > 1$ . The MLSA algorithm is based on the telescopic summations

$$\xi_\star^{h_L} = \xi_\star^{h_0} + \sum_{\ell=1}^L \xi_\star^{h_\ell} - \xi_\star^{h_{\ell-1}}, \quad (4.2)$$

$$\chi_\star^{h_L} = \chi_\star^{h_0} + \sum_{\ell=1}^L \chi_\star^{h_\ell} - \chi_\star^{h_{\ell-1}}. \quad (4.3)$$

Following Crépey, Frikha, and Louzi (2023) and Frikha (2016), for a sequence of positive integers  $\mathbf{N} = \{N_\ell, 0 \leq \ell \leq L\}$ , we define the MLSA estimators

$$\xi_{\mathbf{N}}^{\text{ML}} = \xi_{N_0}^{h_0} + \sum_{\ell=1}^L \xi_{N_\ell}^{h_\ell} - \xi_{N_\ell}^{h_{\ell-1}}, \quad (4.4)$$

$$\chi_{\mathbf{N}}^{\text{ML}} = \chi_{N_0}^{h_0} + \sum_{\ell=1}^L \chi_{N_\ell}^{h_\ell} - \chi_{N_\ell}^{h_{\ell-1}}. \quad (4.5)$$

Each level  $0 \leq \ell \leq L$  is simulated independently: for  $j \in \{(\ell-1)^+, \ell\}$ , given  $\{(X_{h_{\ell-1}}^{(n)}, X_{h_\ell}^{(n)}), 1 \leq n \leq N_\ell\} \stackrel{\text{iid}}{\sim} (X_{h_{\ell-1}}, X_{h_\ell})$ , we iterate for  $0 \leq n \leq N_\ell - 1$

$$\xi_{n+1}^{h_j} = \xi_n^{h_j} - \gamma_{n+1} H_1(\xi_n^{h_j}, X_{h_j}^{(n+1)}), \quad (4.6)$$

$$\chi_{n+1}^{h_j} = \chi_n^{h_j} - \frac{1}{n+1} H_2(\chi_n^{h_j}, \xi_n^{h_j}, X_{h_j}^{(n+1)}), \quad (4.7)$$

starting from real valued random variables  $\xi_0^{h_j}$  and  $\chi_0^{h_j} = 0$ ,  $\xi_0^{h_j}$  being independent from  $\{X_{h_j}^{(n)}, 1 \leq n \leq N_\ell\}$  and such that  $\mathbb{E}[|\xi_0^{h_j}|^2] < \infty$ . To obtain perfectly correlated  $X_{h_\ell}$  and  $X_{h_{\ell-1}}$ , we first sample  $X_{h_{\ell-1}}$  and then use the decomposition

$$X_{h_\ell} = \frac{1}{M} X_{h_{\ell-1}} + \frac{1}{KM^\ell} \sum_{k=KM^{\ell-1}+1}^{KM^\ell} \varphi(Y, Z^{(k)}) \quad (4.8)$$

to obtain  $X_{h_\ell}$  by sampling iid random variables  $\{Z^{(k)}, KM^{\ell-1} < k \leq KM^\ell\}$ .

### 4.1 Convergence Rate Analysis

**Lemma 4.1.** (i) Assume that  $\text{Var}(\varphi(Y, Z)) < \infty$ . Then

$$h_\ell^{-\frac{1}{2}} (X_{h_\ell} - X_{h_{\ell-1}}) =: G_\ell \xrightarrow{\mathcal{L}} G := ((M-1)\text{Var}(\varphi(Y, Z)|Y))^{\frac{1}{2}} \mathfrak{N} \quad \text{as } \ell \uparrow \infty, \quad (4.9)$$

where  $\mathfrak{N}$  is a standard normal random variable independent of  $Y$  defined on a copy  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  of the original probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ .

(ii) Assume that, for all  $\ell \geq 1$ , the function  $F_{X_{h_{\ell-1}}|G_\ell}$  is continuously differentiable  $\mathbb{P}$ -as with derivative  $f_{X_{h_{\ell-1}}|G_\ell}$ , such that the sequence of functions  $\{(x, g) \mapsto f_{X_{h_{\ell-1}}|G_\ell=g}(x), \ell \geq 1\}$  is bounded uniformly in  $\ell \geq 1$  and converges locally uniformly to some bounded and continuous function  $(x, g) \mapsto f_g(x)$ . Then, for any  $\xi \in \mathbb{R}$ ,

$$h_\ell^{-\frac{1}{2}} \mathbb{E}[|\mathbf{1}_{X_{h_\ell} > \xi} - \mathbf{1}_{X_{h_{\ell-1}} > \xi}|] \rightarrow \mathbb{E}[|G|f_G(\xi)] \quad \text{as } \ell \uparrow \infty.$$

(iii) For any  $\xi \in \mathbb{R}$ , we have

$$h_\ell^{-\frac{1}{2}}((X_{h_\ell} - \xi)^+ - (X_{h_{\ell-1}} - \xi)^+) \xrightarrow{\mathcal{L}} \mathbf{1}_{X_0 > \xi} G \quad \text{as } \ell \uparrow \infty.$$

*Proof.* (i) We let  $G_\ell = h_\ell^{-\frac{1}{2}}(X_{h_\ell} - X_{h_{\ell-1}})$ , and introduce

$$\Phi^{(k)}(y) := \varphi(y, Z^{(k)}) - \mathbb{E}[\varphi(y, Z)], \quad y \in \mathbb{R}^d,$$

where  $\{Z^{(k)}, k \geq 1\}$  is an iid sequence of random variables independent of  $Y$  with the same law as  $Z$ . We can then write

$$\begin{aligned} X_{h_\ell} - X_{h_{\ell-1}} &= \frac{1}{KM^\ell} \sum_{k=1}^{KM^\ell} \Phi^{(k)}(Y) - \frac{1}{KM^{\ell-1}} \sum_{k=1}^{KM^{\ell-1}} \Phi^{(k)}(Y) \\ &= \left(1 - \frac{1}{M}\right) \left( \frac{1}{KM^{\ell-1}(M-1)} \sum_{k=KM^{\ell-1}+1}^{KM^\ell} \Phi^{(k)}(Y) - \frac{1}{KM^{\ell-1}} \sum_{k=1}^{KM^{\ell-1}} \Phi^{(k)}(Y) \right), \end{aligned}$$

so that

$$G_\ell = \left(1 - \frac{1}{M}\right)^{\frac{1}{2}} \left( \frac{1}{(KM^{\ell-1}(M-1))^{\frac{1}{2}}} \sum_{k=KM^{\ell-1}+1}^{KM^\ell} \Phi^{(k)}(Y) - \frac{(M-1)^{\frac{1}{2}}}{(KM^{\ell-1})^{\frac{1}{2}}} \sum_{k=1}^{KM^{\ell-1}} \Phi^{(k)}(Y) \right).$$

Now, conditionally on  $Y$ , the following central limit theorem holds:

$$U_\ell := \left( \frac{\frac{1}{(KM^{\ell-1}(M-1))^{\frac{1}{2}}} \sum_{k=KM^{\ell-1}+1}^{KM^\ell} \Phi^{(k)}(Y)}{\frac{1}{(KM^{\ell-1})^{\frac{1}{2}}} \sum_{k=1}^{KM^{\ell-1}} \Phi^{(k)}(Y)} \right) \xrightarrow[\ell \uparrow \infty]{\mathcal{L}} \mathcal{N}(0, \mathbb{E}[\Phi^{(1)}(Y)^2 | Y] I_2),$$

where  $I_2$  stands for the  $2 \times 2$  identity matrix. Observing that  $G_\ell = (1 - 1/M)^{\frac{1}{2}} \langle u_M, U_\ell \rangle$ , with  $u_M := (1, -(M-1)^{\frac{1}{2}})^\top$ , we conclude that, conditionally on  $Y$ ,

$$h_\ell^{-\frac{1}{2}}(X_{h_\ell} - X_{h_{\ell-1}}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (M-1) \mathbb{V}\text{ar}(\varphi(Y, Z) | Y)) \quad \text{as } \ell \uparrow \infty.$$

The proof of (4.9) is thus complete.

(ii) We have

$$\mathbb{E}[|\mathbf{1}_{X_{h_\ell} > \xi} - \mathbf{1}_{X_{h_{\ell-1}} > \xi}|] = \mathbb{P}(X_{h_{\ell-1}} \leq \xi < X_{h_\ell}) + \mathbb{P}(X_{h_\ell} \leq \xi < X_{h_{\ell-1}}).$$

Introducing the random variable  $G_\ell$ , we compute

$$\begin{aligned} \mathbb{P}(X_{h_{\ell-1}} \leq \xi < X_{h_\ell}) &= \mathbb{P}(X_{h_{\ell-1}} \leq \xi < X_{h_{\ell-1}} + h_\ell^{\frac{1}{2}} G_\ell) \\ &= \mathbb{P}(X_{h_{\ell-1}} \leq \xi < X_{h_{\ell-1}} + h_\ell^{\frac{1}{2}} G_\ell, G_\ell > 0) \\ &= \mathbb{E}[\mathbb{P}(\xi - h_\ell^{\frac{1}{2}} G_\ell < X_{h_{\ell-1}} \leq \xi, G_\ell > 0 | G_\ell)] \\ &= \mathbb{E}[\mathbf{1}_{G_\ell > 0} (F_{X_{h_{\ell-1}} | G_\ell}(\xi) - F_{X_{h_{\ell-1}} | G_\ell}(\xi - h_\ell^{\frac{1}{2}} G_\ell))] \\ &= h_\ell^{\frac{1}{2}} \mathbb{E}[G_\ell^+ f_{X_{h_{\ell-1}} | G_\ell}(\xi)] + h_\ell^{\frac{1}{2}} r_\ell^+, \end{aligned}$$

where

$$r_\ell^+ := \int_0^1 \mathbb{E}[G_\ell^+ (f_{X_{h_{\ell-1}} | G_\ell}(\xi - th_\ell^{\frac{1}{2}} G_\ell) - f_{X_{h_{\ell-1}} | G_\ell}(\xi))] dt.$$

Similarly,

$$\mathbb{P}(X_{h_\ell} \leq \xi < X_{h_{\ell-1}}) = h_\ell^{\frac{1}{2}} \mathbb{E}[G_\ell^- f_{X_{h_{\ell-1}}|G_\ell}(\xi)] + h_\ell^{\frac{1}{2}} r_\ell^-,$$

where

$$r_\ell^- := \int_0^1 \mathbb{E}[G_\ell^- (f_{X_{h_{\ell-1}}|G_\ell}(\xi - th_\ell^{\frac{1}{2}} G_\ell) - f_{X_{h_{\ell-1}}|G_\ell}(\xi))] dt.$$

Hence

$$h_\ell^{-\frac{1}{2}} \mathbb{E}[\mathbb{1}_{X_{h_\ell} > \xi} - \mathbb{1}_{X_{h_{\ell-1}} > \xi}] = \mathbb{E}[|G_\ell| f_{X_{h_{\ell-1}}|G_\ell}(\xi)] + r_\ell^+ + r_\ell^-.$$

We now prove that the first term in the right hand side of the above decomposition converges towards  $\mathbb{E}[|G|]f_G(\xi)$  and that the two other terms vanish as  $\ell \uparrow \infty$ . For a fixed  $K > 0$ , we decompose the first term as

$$\begin{aligned} \mathbb{E}[|G_\ell| f_{X_{h_{\ell-1}}|G_\ell}(\xi)] &= \mathbb{E}[|G_\ell| (f_{X_{h_{\ell-1}}|G_\ell}(\xi) - f_{G_\ell}(\xi)) \mathbb{1}_{|G_\ell| \leq K}] \\ &\quad + \mathbb{E}[|G_\ell| (f_{X_{h_{\ell-1}}|G_\ell}(\xi) - f_{G_\ell}(\xi)) \mathbb{1}_{|G_\ell| > K}] + \mathbb{E}[|G_\ell| f_{G_\ell}(\xi)]. \end{aligned} \quad (4.10)$$

For the first term, we use the fact that the sequence of functions  $\{g \mapsto f_{X_{h_{\ell-1}}|G_\ell=g}(\xi), \ell \geq 1\}$  converges locally uniformly towards  $g \mapsto f_g(\xi)$ . We thus deduce

$$\lim_{\ell \uparrow \infty} \mathbb{E}[|G_\ell| (f_{X_{h_{\ell-1}}|G_\ell}(\xi) - f_{G_\ell}(\xi)) \mathbb{1}_{|G_\ell| \leq K}] = 0.$$

For the second term, since  $g \mapsto f_g(\xi)$  is bounded and the functions  $\{g \mapsto f_{X_{h_{\ell-1}}|G_\ell=g}(\xi), \ell \geq 1\}$  are bounded uniformly in  $\ell \geq 1$ , and given that

$$\mathbb{E}[G_\ell^2] \leq 2h_\ell^{-1} (\mathbb{E}[(X_{h_\ell} - X_0)^2] + \mathbb{E}[(X_{h_{\ell-1}} - X_0)^2]) \leq 2(1 + M^{-1}) \mathbb{E}[\text{Var}(\varphi(Y, Z)|Y)] < \infty,$$

so that  $\sup_{\ell \geq 1} \mathbb{E}[G_\ell^2] < \infty$ , we have the upper bound

$$\mathbb{E}[|G_\ell| (f_{X_{h_{\ell-1}}|G_\ell}(\xi) - f_{G_\ell}(\xi)) \mathbb{1}_{|G_\ell| > K}] \leq C \mathbb{E}[G_\ell^2]^{\frac{1}{2}} \mathbb{P}(|G_\ell| > K)^{\frac{1}{2}} \leq CK^{-1} \sup_{\ell \geq 1} \mathbb{E}[G_\ell^2].$$

For the last term, since  $\sup_{\ell \geq 1} \mathbb{E}[G_\ell^2] < \infty$ , the function  $g \mapsto f_g(\xi)$  is continuous and bounded and  $G_\ell \xrightarrow[\ell \uparrow \infty]{\mathcal{L}} G$ , we have

$$\lim_{\ell \uparrow \infty} \mathbb{E}[|G_\ell| f_{G_\ell}(\xi)] = \mathbb{E}[|G| f_G(\xi)].$$

Coming back to the decomposition (4.10), letting  $\ell \uparrow \infty$  and then  $M \uparrow \infty$  eventually yields

$$\lim_{\ell \uparrow \infty} \mathbb{E}[|G_\ell| f_{X_{h_{\ell-1}}|G_\ell}(\xi)] = \mathbb{E}[|G| f_G(\xi)].$$

Similar lines of reasonings using the tightness of  $\{G_\ell, \ell \geq 1\}$ , the uniform boundedness and the local uniform convergence of the functions  $\{(x, g) \mapsto f_{X_{h_{\ell-1}}|G_\ell=g}(x), \ell \geq 1\}$ , and the continuity of  $(x, g) \mapsto f_g(x)$  give

$$\lim_{\ell \uparrow \infty} r_\ell^+ = \lim_{\ell \uparrow \infty} r_\ell^- = 0.$$

(iii) Since  $\mathbb{P}(X_0 = \xi) = 0$ , a first order Taylor's expansion gives

$$(X_h - \xi)^+ = (X_0 - \xi)^+ + \mathbb{1}_{X_0 > \xi} (X_h - X_0) + a(X_0, X_h - X_0)(X_h - X_0),$$

where

$$a(X_0, X_h - X_0) = \int_0^1 (\mathbb{1}_{X_0 + t(X_h - X_0) > \xi} - \mathbb{1}_{X_0 > \xi}) dt.$$

Since  $X_h \xrightarrow{\mathbb{P}} X_0$  as  $\mathcal{H} \ni h \downarrow 0$  and  $\mathbb{P}(X_0 = \xi) = 0$ ,

$$a(X_0, X_h - X_0) \xrightarrow[\mathcal{H} \ni h \downarrow 0]{\mathbb{P}} 0.$$

The above expansion gives, for  $\ell \geq 1$ ,

$$\begin{aligned} (X_{h_\ell} - \xi)^+ - (X_{h_{\ell-1}} - \xi)^+ &= \mathbf{1}_{X_0 > \xi}(X_{h_\ell} - X_{h_{\ell-1}}) \\ &\quad + a(X_0, X_{h_\ell} - X_0)(X_{h_\ell} - X_0) - a(X_0, X_{h_{\ell-1}} - X_0)(X_{h_{\ell-1}} - X_0). \end{aligned}$$

The standard central limit theorem guarantees that both sequences  $\{h_\ell^{-\frac{1}{2}}(X_{h_\ell} - X_0), \ell \geq 1\}$  and  $\{h_\ell^{-\frac{1}{2}}(X_{h_{\ell-1}} - X_0), \ell \geq 1\}$  are tight. Hence,

$$h_\ell^{-\frac{1}{2}}(a(X_0, X_{h_\ell} - X_0)(X_{h_\ell} - X_0) - a(X_0, X_{h_{\ell-1}} - X_0)(X_{h_{\ell-1}} - X_0)) \xrightarrow[\ell \uparrow \infty]{\mathbb{P}} 0.$$

It follows from the proof of Lemma 4.1(i) that  $h_\ell^{-\frac{1}{2}}(X_{h_\ell} - X_{h_{\ell-1}}) \xrightarrow{\mathcal{L}} G$  conditionally on  $Y$  as  $\ell \uparrow \infty$ . Since  $X_0$  is  $\sigma(Y)$ -measurable, then, conditionally on  $Y$ ,

$$h_\ell^{-\frac{1}{2}} \mathbf{1}_{X_0 > \xi}(X_{h_\ell} - X_{h_{\ell-1}}) \xrightarrow[\ell \uparrow \infty]{\mathcal{L}} \mathbf{1}_{X_0 > \xi} G.$$

Gathering the previous asymptotic results yields, by Slutsky's theorem,

$$h_\ell^{-\frac{1}{2}}((X_{h_\ell} - \xi)^+ - (X_{h_{\ell-1}} - \xi)^+) \xrightarrow{\mathcal{L}} \mathbf{1}_{X_0 > \xi} G \quad \text{as } \ell \uparrow \infty.$$

□

Finally, we obtain the CLT for the MLSA estimators of the VaR and ES. Depending on the context, we use the notation  $N_\ell$  to designate both  $N_\ell$  and  $\lceil N_\ell \rceil$  interchangeably.

**Assumption 2.** *There exist  $C < \infty$  and  $\delta_0 > 0$  such that, for any  $h \in \mathcal{H}$  and any compact set  $\mathcal{K} \subset \mathbb{R}$ ,*

$$\sup_{\xi \in \mathcal{K}} |f_{X_h}(\xi) - f_{X_0}(\xi)| \leq Ch^{\frac{1}{4} + \delta_0}.$$

**Theorem 4.1.** *Suppose that the assumptions of Theorem 2.1 are satisfied and that Assumption 2 holds. Assume also that  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $\beta \in (\frac{1}{2}, 1]$ , with  $\lambda\gamma_1 > 1$  if  $\beta = 1$ . Let*

$$N_\ell = h_L^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{\frac{1}{\beta}} h_\ell^{\frac{3}{2(1+\beta)}}, \quad 0 \leq \ell \leq L. \quad (4.11)$$

Then

$$\begin{pmatrix} h_L^{-1}(\xi_{\mathbf{N}}^{\text{ML}} - \xi_\star^{h_L}) \\ h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}}(\chi_{\mathbf{N}}^{\text{ML}} - \chi_\star^{h_L}) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_\beta^{\text{ML}}) \quad \text{as } L \uparrow \infty,$$

where  $\Sigma_\beta^{\text{ML}} =$

$$\begin{pmatrix} \frac{\gamma_1 \mathbb{E}[G|f_G(\xi_\star)]}{(1-\alpha)(2f_{X_0}(\xi_\star) - (1-\alpha)\gamma_1^{-1}\mathbf{1}_{\beta=1})} & 0 \\ 0 & \frac{h_0^{\frac{2\beta-1}{2(1+\beta)}} \left( M^{\frac{2\beta-1}{2(1+\beta)}} - 1 \right)^{\frac{1}{\beta}} \left( \frac{h_0^{-1} \text{Var}((X_{h_0} - \xi_\star^{h_0})^+)}{M^{\frac{2\beta-1}{2\beta(1+\beta)}}} + \frac{\text{Var}(\mathbf{1}_{X_0 > \xi_\star} G)}{M^{\frac{2\beta-1}{2(1+\beta)}} - 1} \right)}{(1-\alpha)^2} \end{pmatrix}. \quad (4.12)$$

**Remark 4.1.** We use the optimal iteration parameters (4.11) obtained for the VaR MLSA in Crépey, Frikha, and Louzi (2023, Section 4.2). These were obtained by optimizing the MLSA complexity while constraining the global  $L^2$  error of the MLSA VaR estimator to a prescribed error of order  $\varepsilon^2 > 0$ .

*Proof.* In the following developments, we denote by  $C$  a positive constant whose value may change from line to line and does not depend upon  $L$ .

According to (4.4), (4.2) and (2.18),

$$\begin{aligned}
\xi_{\mathbf{N}}^{\text{ML}} - \xi_{\star}^{h_L} &= \xi_{N_0}^{h_0} - \xi_{\star}^{h_0} + \sum_{\ell=1}^L (\xi_{N_{\ell}}^{h_{\ell}} - \xi_{\star}^{h_{\ell}} - (\xi_{N_{\ell}}^{h_{\ell-1}} - \xi_{\star}^{h_{\ell-1}})) \\
&= \xi_{N_0}^{h_0} - \xi_{\star}^{h_0} + \sum_{\ell=1}^L (\xi_0^{h_{\ell}} - \xi_{\star}^{h_{\ell}} - (\xi_0^{h_{\ell-1}} - \xi_{\star}^{h_{\ell-1}})) \Pi_{1:N_{\ell}} \\
&\quad + \sum_{\ell=1}^L \sum_{k=1}^{N_{\ell}} \gamma_k \Pi_{k+1:N_{\ell}} (g_k^{h_{\ell}} - g_k^{h_{\ell-1}}) + \sum_{\ell=1}^L \sum_{k=1}^{N_{\ell}} \gamma_k \Pi_{k+1:N_{\ell}} (r_k^{h_{\ell}} - r_k^{h_{\ell-1}}) \\
&\quad + \sum_{\ell=1}^L \sum_{k=1}^{N_{\ell}} \gamma_k \Pi_{k+1:N_{\ell}} (\rho_k^{h_{\ell}} - \rho_k^{h_{\ell-1}}) + \sum_{\ell=1}^L \sum_{k=1}^{N_{\ell}} \gamma_k \Pi_{k+1:N_{\ell}} (e_k^{h_{\ell}} - e_k^{h_{\ell-1}}),
\end{aligned} \tag{4.13}$$

recalling the definitions (2.13)–(2.19). Similarly, from (2.20), (4.3) and (4.5),

$$\begin{aligned}
\chi_{\mathbf{N}}^{\text{ML}} - \chi_{\star}^{h_L} &= \chi_{N_0}^{h_0} - \chi_{\star}^{h_0} + \sum_{\ell=1}^L (\chi_{N_{\ell}}^{h_{\ell}} - \chi_{\star}^{h_{\ell}} - (\chi_{N_{\ell}}^{h_{\ell-1}} - \chi_{\star}^{h_{\ell-1}})) \\
&= \chi_{N_0}^{h_0} - \chi_{\star}^{h_0} + \sum_{\ell=1}^L \frac{1}{N_{\ell}} \sum_{k=1}^{N_{\ell}} \theta_k^{h_{\ell}} - \theta_k^{h_{\ell-1}} \\
&\quad + \sum_{\ell=1}^L \frac{1}{N_{\ell}} \sum_{k=1}^{N_{\ell}} \zeta_k^{h_{\ell}} - \zeta_k^{h_{\ell-1}} + \sum_{\ell=1}^L \frac{1}{N_{\ell}} \sum_{k=1}^{N_{\ell}} \eta_k^{h_{\ell}} - \eta_k^{h_{\ell-1}},
\end{aligned} \tag{4.14}$$

recalling the definitions (2.21)–(2.23).

We will study each part of the decompositions (4.13)–(4.14) independently.

**Step 1. Study of  $\{h_L^{-1}(\xi_{N_0}^{h_0} - \xi_{\star}^{h_0}), L \geq 1\}$ .**

Using Lemma 2.2(ii) and (4.11), we obtain

$$\mathbb{E}[|h_L^{-1}(\xi_{N_0}^{h_0} - \xi_{\star}^{h_0})|^2] \leq Ch_L^{-2} \gamma_{N_0} \leq Ch_L^{\frac{2\beta-1}{2(1+\beta)}},$$

so that

$$h_L^{-1}(\xi_{N_0}^{h_0} - \xi_{\star}^{h_0}) \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 2. Study of  $\{h_L^{-1} \sum_{\ell=1}^L (\xi_0^{h_{\ell}} - \xi_{\star}^{h_{\ell}} - (\xi_0^{h_{\ell-1}} - \xi_{\star}^{h_{\ell-1}})) \Pi_{1:N_{\ell}}, L \geq 1\}$ .**

Using the facts that  $\sup_{h \in \mathcal{H}} \mathbb{E}[|\xi_0^h - \xi_{\star}^h|] < \infty$ , recalling (2.11), and that  $\limsup_{n \uparrow \infty} \gamma_n^{-1} |\Pi_{1:n}| =$

0, which stems from (2.24) and Lemma A.1(ii), we obtain

$$\begin{aligned}
\mathbb{E} \left[ \left| h_L^{-1} \sum_{\ell=1}^L (\xi_0^{h_\ell} - \xi_\star^{h_\ell} - (\xi_0^{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})) \Pi_{1:N_\ell} \right| \right] \\
\leq 2 \sup_{h \in \mathcal{H}} \mathbb{E}[|\xi_0^h - \xi_\star^h|] h_L^{-1} \sum_{\ell=1}^L |\Pi_{1:N_\ell}| \\
\leq C h_L^{-1} \sum_{\ell=1}^L \gamma_{N_\ell} \\
\leq C h_L^{\frac{1}{2}},
\end{aligned}$$

where we used (4.11) for the last inequality. Hence

$$h_L^{-1} \sum_{\ell=1}^L (\xi_0^{h_\ell} - \xi_\star^{h_\ell} - (\xi_0^{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})) \Pi_{1:N_\ell} \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 3. Study of**  $\left\{ h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (g_k^{h_\ell} - g_k^{h_{\ell-1}}), L \geq 1 \right\}$ .

There exists a compact set  $\mathcal{K} \subset \mathbb{R}$  such that  $\xi_\star^{h_\ell} \in \mathcal{K}$ , for all  $\ell \geq 0$ . Thus, Assumptions 1(iv) and 2 and Lemma 2.2(i) imply

$$\begin{aligned}
|V_0''(\xi_\star) - V_{h_\ell}''(\xi_\star^{h_\ell})| &\leq \frac{1}{1-\alpha} \left( [f_{X_0}]_{\text{Lip}} |\xi_\star^{h_\ell} - \xi_\star| + \sup_{\xi \in \mathcal{K}} |f_{X_0}(\xi) - f_{X_{h_\ell}}(\xi)| \right) \\
&\leq C(h_\ell + h_\ell^{\frac{1}{4} + \delta_0}) \leq C h_\ell^{(\frac{1}{4} + \delta_0) \wedge 1}.
\end{aligned}$$

From the definition (2.13), by using Lemma 2.2(ii), we deduce that, for  $\ell \geq 1$ ,

$$\mathbb{E}[|g_k^{h_\ell}|] \leq |V_0''(\xi_\star) - V_{h_\ell}''(\xi_\star^{h_\ell})| \mathbb{E}[(\xi_{k-1}^{h_\ell} - \xi_\star^{h_\ell})^2]^{\frac{1}{2}} \leq C h_\ell^{(\frac{1}{4} + \delta_0) \wedge 1} \gamma_k^{\frac{1}{2}}. \quad (4.15)$$

Via (2.24), (4.11) and Lemma A.1(i), by distinguishing the two cases  $\delta_0 < \frac{3}{4} \wedge \frac{2\beta-1}{4(1+\beta)} = \frac{2\beta-1}{4(1+\beta)}$  and  $\delta_0 \geq \frac{2\beta-1}{4(1+\beta)}$ , we obtain

$$\begin{aligned}
\mathbb{E} \left[ \left| h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (g_k^{h_\ell} - g_k^{h_{\ell-1}}) \right| \right] \\
\leq h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k |\Pi_{k+1:N_\ell}| (\mathbb{E}[|g_k^{h_\ell}|] + \mathbb{E}[|g_k^{h_{\ell-1}}|]) \\
\leq C h_L^{-1} \sum_{\ell=1}^L h_\ell^{(\frac{1}{4} + \delta_0) \wedge 1} \sum_{k=1}^{N_\ell} \gamma_k^{\frac{3}{2}} |\Pi_{k+1:N_\ell}| \\
\leq C h_L^{-1} \sum_{\ell=1}^L h_\ell^{(\frac{1}{4} + \delta_0) \wedge 1} \gamma_{N_\ell}^{\frac{1}{2}} \\
\leq C h_L^{\delta_0 \wedge \frac{2\beta-1}{4(1+\beta)}},
\end{aligned}$$

so that

$$h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (g_k^{h_\ell} - g_k^{h_{\ell-1}}) \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 4. Study of**  $\left\{ h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (r_k^{h_\ell} - r_k^{h_{\ell-1}}), L \geq 1 \right\}$ .

It follows from (3.8), (2.24), Lemma A.1(i) and (4.11) that

$$\begin{aligned}
\mathbb{E} \left[ \left| h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (r_k^{h_\ell} - r_k^{h_{\ell-1}}) \right| \right] \\
\leq h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k |\Pi_{k+1:N_\ell}| (\mathbb{E}[|r_k^{h_\ell}|] + \mathbb{E}[|r_k^{h_{\ell-1}}|]) \\
\leq C h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k^2 |\Pi_{k+1:N_\ell}| \\
\leq C h_L^{-1} \sum_{\ell=1}^L \gamma_{N_\ell} \\
\leq C h_L^{\frac{1}{2}},
\end{aligned}$$

so that

$$h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (r_k^{h_\ell} - r_k^{h_{\ell-1}}) \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 5. Study of**  $\left\{ h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}), L \geq 1 \right\}$ .

Given that the innovations of the MLSA scheme are independent levelwise, the random variables  $\left\{ \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}), \ell \geq 1 \right\}$  are independent and with zero mean. Besides, at each level  $\ell \geq 1$ , the sequence  $\{\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}, k \geq 1\}$  is an  $\{\mathcal{F}_k^{h_\ell}, k \geq 1\}$ -martingale increment sequence. Thus, from (3.9), (2.24), Lemma A.1(i) (recalling that  $\lambda \gamma_1 > 1$ ) and (4.11), we have

$$\begin{aligned}
\mathbb{E} \left[ \left( h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}) \right)^2 \right] \\
= h_L^{-2} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k^2 \Pi_{k+1:N_\ell}^2 \mathbb{E}[(\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}})^2] \\
\leq 2 h_L^{-2} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k^2 \Pi_{k+1:N_\ell}^2 (\mathbb{E}[|\rho_k^{h_\ell}|^2] + \mathbb{E}[|\rho_k^{h_{\ell-1}}|^2]) \\
\leq C h_L^{-2} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k^{\frac{5}{2}} \Pi_{k+1:N_\ell}^2 \\
\leq C h_L^{-2} \sum_{\ell=1}^L \gamma_{N_\ell}^{\frac{3}{2}} \\
\leq C h_L^{\frac{1}{4}},
\end{aligned}$$

so that

$$h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}) \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 6. Study of**  $\left\{ h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}}, L \geq 1 \right\}$ .

Define

$$\bar{\gamma}_n = \frac{1}{n} \sum_{k=1}^n \gamma_k, \quad n \geq 1.$$

The random variables  $\left\{ \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}}, \ell \geq 1 \right\}$  are independent and centered. Moreover, for any fixed level  $\ell \geq 1$ , the sequence  $\left\{ \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}}, k \geq 1 \right\}$  is an  $\{\mathcal{F}_k^{h_\ell}, k \geq 1\}$ -martingale increment sequence. Hence, using (2.29), a comparison between series and integrals and then (4.11), we obtain

$$\begin{aligned} & \mathbb{E} \left[ \left| h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}} \right|^2 \right] \\ & \leq 2h_L^{-\frac{2}{\beta} - \frac{2\beta-1}{2\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \left( \mathbb{E} \left[ \left( \frac{1}{\sqrt{N_\ell}} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} \right)^2 \right] + \mathbb{E} \left[ \left( \frac{1}{\sqrt{N_\ell}} \sum_{k=1}^{N_\ell} \theta_k^{h_{\ell-1}} \right)^2 \right] \right) \\ & \leq Ch_L^{-\frac{2}{\beta} - \frac{2\beta-1}{2\beta(1+\beta)}} \sum_{\ell=1}^L \frac{\bar{\gamma}_{N_\ell}}{N_\ell} \\ & \leq C \begin{cases} h_L^{\frac{3\beta}{2(1+\beta)}}, & \beta \in (\frac{1}{2}, 1), \\ h_L^{\frac{3}{4}} |\ln h_L|, & \beta = 1. \end{cases} \end{aligned}$$

Hence

$$h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}} \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 7. Study of  $\left\{ h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \zeta_k^{h_\ell} - \zeta_k^{h_{\ell-1}}, L \geq 1 \right\}$ .**

Using (2.30), a comparison between series and integrals and then (4.11), we obtain

$$\begin{aligned} & \mathbb{E} \left[ \left| h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \zeta_k^{h_\ell} - \zeta_k^{h_{\ell-1}} \right| \right] \\ & \leq h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{\sqrt{N_\ell}} \left( \mathbb{E} \left[ \left| \frac{1}{\sqrt{N_\ell}} \sum_{k=1}^{N_\ell} \zeta_k^{h_\ell} \right| \right] + \mathbb{E} \left[ \left| \frac{1}{\sqrt{N_\ell}} \sum_{k=1}^{N_\ell} \zeta_k^{h_{\ell-1}} \right| \right] \right) \\ & \leq Ch_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \bar{\gamma}_{N_\ell} \\ & \leq C \begin{cases} h_L^{\frac{9}{4\beta(1+\beta)}}, & \beta \in (\frac{1}{2}, 1), \\ h_L^{\frac{3}{4}} |\ln h_L|, & \beta = 1. \end{cases} \end{aligned}$$

Hence

$$h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \zeta_k^{h_\ell} - \zeta_k^{h_{\ell-1}} \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 8. Study of  $\left\{ h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} (\chi_{N_0}^{h_0} - \chi_\star^{h_0}), L \geq 1 \right\}$ .**

It follows from the definition of  $N_0$  in (4.11) that

$$h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} (\chi_{N_0}^{h_0} - \chi_\star^{h_0}) = h_0^{-\frac{3}{4(1+\beta)}} \left( 1 - M^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{\frac{1}{2\beta}} \frac{\sqrt{N_0} (\chi_{N_0}^{h_0} - \chi_\star^{h_0})}{\left( 1 - M^{-\frac{2\beta-1}{2(1+\beta)}} (L+1) \right)^{\frac{1}{2\beta}}}.$$



According to the decomposition (2.20),

$$\sqrt{N_0}(\chi_{N_0}^{h_0} - \chi_\star^{h_0}) = \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \theta_k^{h_0} + \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \zeta_k^{h_0} + \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \eta_k^{h_0}.$$

From (2.29) and (2.30), we obtain

$$\frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \theta_k^{h_0} \xrightarrow{L^2(\mathbb{P})} 0, \quad \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \zeta_k^{h_0} \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } N_0 \uparrow \infty.$$

Since  $\{\eta_k^{h_0}, k \geq 1\}$  is an iid sequence satisfying  $\mathbb{E}[|\eta_1^{h_0}|^2] < \infty$ , the classical CLT yields

$$\frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \eta_k^{h_0} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\text{Var}((X_{h_0} - \xi_\star^{h_0})^+)}{(1 - \alpha)^2}\right) \quad \text{as } N_0 \uparrow \infty.$$

Hence

$$h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} (\chi_{N_0}^{h_0} - \chi_\star^{h_0}) \xrightarrow{L \uparrow \infty} \mathcal{N}\left(0, h_0^{-\frac{3}{2(1+\beta)}} (1 - M^{-\frac{2\beta-1}{2(1+\beta)}})^{\frac{1}{\beta}} \frac{\text{Var}((X_{h_0} - \xi_\star^{h_0})^+)}{(1 - \alpha)^2}\right).$$

### Step 9. Study of

$$\left\{ \left( h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}), h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \eta_k^{h_\ell} - \eta_k^{h_{\ell-1}} \right), L \geq 1 \right\}.$$

Since  $N_1 \geq \dots \geq N_L$ , we have

$$\sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) = \sum_{k=1}^{N_1} \sum_{\ell=1}^L \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell},$$

and

$$\sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \eta_k^{h_\ell} - \eta_k^{h_{\ell-1}} = \sum_{k=1}^{N_1} \sum_{\ell=1}^L \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell}.$$

We now apply again the CLT for martingale arrays (Hall & Heyde, 1980, Corollary 3.1) to the martingale array

$$\left\{ \left( h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}), h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \right), L \geq 1 \right\}.$$

We first check the conditional Lindeberg condition. Let  $\delta > 0$ . By the levelwise independence of the innovations of the MLSA scheme and given that, for all  $k \geq 1$ , the random variables  $\{(e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell}, 1 \leq \ell \leq L\}$  are independent and centered, applying successively the Marcinkiewicz-Zygmund and the Jensen inequalities, we obtain

$$\begin{aligned} & \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| h_L^{-1} \sum_{\ell=1}^L \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] \\ & \leq C h_L^{-(2+\delta)} \sum_{k=1}^{N_1} \gamma_k^{2+\delta} \mathbb{E} \left[ \left| \sum_{\ell=1}^L \Pi_{k+1:N_\ell}^2 (e_k^{h_\ell} - e_k^{h_{\ell-1}})^2 \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{1+\frac{\delta}{2}} \right] \\ & \leq C h_L^{-(2+\delta)} \sum_{k=1}^{N_1} \gamma_k^{2+\delta} L^{\frac{\delta}{2}} \sum_{\ell=1}^L |\Pi_{k+1:N_\ell}|^{2+\delta} \mathbb{E} [|e_k^{h_\ell} - e_k^{h_{\ell-1}}|^{2+\delta}] \mathbb{1}_{1 \leq k \leq N_\ell} \\ & = C h_L^{-(2+\delta)} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k^{2+\delta} |\Pi_{k+1:N_\ell}|^{2+\delta} \mathbb{E} [|e_k^{h_\ell} - e_k^{h_{\ell-1}}|^{2+\delta}]. \end{aligned} \quad (4.16)$$

It follows from Lemma 4.1(ii) and the uniform boundedness of  $\{f_{X_{h_{\ell-1}}}, \ell \geq 1\}$  that

$$\begin{aligned}
& \mathbb{E}[|e_k^{h_\ell} - e_k^{h_{\ell-1}}|^{2+\delta}] \\
&= \mathbb{E}[|H_1(\xi_\star^{h_\ell}, X_{h_\ell}^{(k)}) - H_1(\xi_\star^{h_{\ell-1}}, X_{h_{\ell-1}}^{(k)})|^{2+\delta}] \\
&\leq \frac{1}{(1-\alpha)^{2+\delta}} \left( \mathbb{E}[|\mathbb{1}_{X_{h_\ell} > \xi_\star^{h_\ell}} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star^{h_\ell}}|] + \mathbb{E}[|\mathbb{1}_{X_{h_{\ell-1}} > \xi_\star^{h_\ell}} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star^{h_{\ell-1}}}|] \right) \\
&\leq C \left( \mathbb{E}[|\mathbb{1}_{X_{h_\ell} > \xi_\star} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star}|] + \mathbb{E}[|\mathbb{1}_{X_{h_\ell} > \xi_\star^{h_\ell}} - \mathbb{1}_{X_{h_\ell} > \xi_\star}|] \right. \\
&\quad \left. + \mathbb{E}[|\mathbb{1}_{X_{h_{\ell-1}} > \xi_\star^{h_{\ell-1}}} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star}|] + \mathbb{E}[|\mathbb{1}_{X_{h_{\ell-1}} > \xi_\star^{h_\ell}} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star^{h_{\ell-1}}}|] \right) \\
&\leq C(h_\ell^{\frac{1}{2}} + \mathbb{E}[|F_{X_{h_\ell}}(\xi_\star^{h_\ell}) - F_{X_{h_\ell}}(\xi_\star)|] + \mathbb{E}[|F_{X_{h_{\ell-1}}}(\xi_\star^{h_{\ell-1}}) - F_{X_{h_{\ell-1}}}(\xi_\star)|] \\
&\quad + \mathbb{E}[|F_{X_{h_{\ell-1}}}(\xi_\star^{h_\ell}) - F_{X_{h_{\ell-1}}}(\xi_\star^{h_{\ell-1}})|]) \\
&\leq C(h_\ell^{\frac{1}{2}} + |\xi_\star^{h_\ell} - \xi_\star| + |\xi_\star^{h_{\ell-1}} - \xi_\star| + |\xi_\star^{h_\ell} - \xi_\star^{h_{\ell-1}}|) \\
&\leq Ch_\ell^{\frac{1}{2}},
\end{aligned} \tag{4.17}$$

where we used Lemma 2.2(i) for the last inequality. Plugging the previous estimate into (4.16) and then using Lemma A.1(i) and then (4.11), we obtain

$$\begin{aligned}
& \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-1} \gamma_k \Pi_{k+1:N_\ell}(e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] \\
&\leq Ch_L^{-(2+\delta)} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \gamma_{N_\ell}^{1+\delta} h_\ell^{\frac{1}{2}} \\
&\leq Ch_L^{\frac{\delta}{2}} L^{\frac{\delta}{2}}.
\end{aligned}$$

Hence

$$\limsup_{L \uparrow \infty} \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-1} \gamma_k \Pi_{k+1:N_\ell}(e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] = 0.$$

Similarly,

$$\begin{aligned}
& \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] \\
&\leq Ch_L^{-(2+\delta)(\frac{1}{\beta} + \frac{2\beta-1}{4\beta(1+\beta)})} \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L \frac{(\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}})^2}{N_\ell^2} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{1+\frac{\delta}{2}} \right] \\
&\leq Ch_L^{-(2+\delta)(\frac{1}{\beta} + \frac{2\beta-1}{4\beta(1+\beta)})} \sum_{k=1}^{N_1} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \frac{\mathbb{E}[|\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}|^{2+\delta}]}{N_\ell^{2+\delta}} \mathbb{1}_{1 \leq k \leq N_\ell} \\
&= Ch_L^{-(2+\delta)(\frac{1}{\beta} + \frac{2\beta-1}{4\beta(1+\beta)})} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \frac{1}{N_\ell^{2+\delta}} \sum_{k=1}^{N_\ell} \mathbb{E}[|\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}|^{2+\delta}].
\end{aligned} \tag{4.18}$$

Now note that

$$\begin{aligned}
& \mathbb{E}[|\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}|^{2+\delta}] \\
&\leq \frac{2^{1+\delta}}{(1-\alpha)^{2+\delta}} \mathbb{E}[|(X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+|^{2+\delta}] \\
&\leq \frac{2^{2+2\delta}}{(1-\alpha)^{2+\delta}} (\mathbb{E}[|X_{h_\ell} - X_{h_{\ell-1}}|^{2+\delta}] + |\xi_\star^{h_\ell} - \xi_\star^{h_{\ell-1}}|^{2+\delta}) \\
&\leq Ch_\ell^{1+\frac{\delta}{2}},
\end{aligned} \tag{4.19}$$

where we used the facts that  $x \mapsto x^+$  is 1-Lipschitz, as well as Lemmas 2.1(i) and 2.2(i). Therefore, plugging the previous estimate into (4.18) and using (4.11), we obtain

$$\begin{aligned} \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] \\ \leq C h_L^{-(2+\delta)(\frac{1}{\beta} + \frac{2\beta-1}{4\beta(1+\beta)})} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \frac{h_\ell^{1+\frac{\delta}{2}}}{N_\ell^{1+\delta}} \\ \leq C h_L^{\delta(\frac{1}{\beta} + \frac{2\beta-1}{4\beta(1+\beta)})} |\ln h_L|^{\frac{\delta}{2}} \sum_{\ell=1}^L h_\ell^{\frac{-(2-\beta)\delta+2\beta-1}{2(1+\beta)}}. \end{aligned}$$

By taking  $\delta \in (0, \frac{2\beta-1}{2-\beta})$ , we obtain

$$\sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] \leq C h_L^{\delta(\frac{1}{\beta} + \frac{2\beta-1}{4\beta(1+\beta)})} |\ln h_L|^{\frac{\delta}{2}},$$

so that

$$\limsup_{L \uparrow \infty} \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] = 0.$$

The conditional Lindeberg condition is thus satisfied.

We now prove the convergence of the conditional covariance matrices sequence  $\{S_L = (S_L^{i,j})_{1 \leq i,j \leq 2}, L \geq 1\}$  defined by

$$\begin{aligned} S_L^{1,1} &:= \sum_{k=1}^{N_1} \mathbb{E} \left[ \left( \sum_{\ell=1}^L h_L^{-1} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell} \right)^2 \middle| \mathcal{F}_{k-1}^{h_L} \right], \\ S_L^{2,2} &:= \sum_{k=1}^{N_1} \mathbb{E} \left[ \left( \sum_{\ell=1}^L h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right)^2 \middle| \mathcal{F}_{k-1}^{h_L} \right], \\ S_L^{1,2} = S_L^{2,1} &:= \sum_{k=1}^{N_1} \mathbb{E} \left[ \sum_{\ell=1}^L \frac{h_L^{-1 - \frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}}}{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) (\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell} \middle| \mathcal{F}_{k-1}^{h_L} \right]. \end{aligned}$$

The sequence of random variables  $\{\sum_{\ell=1}^L \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell}, k \geq 1\}$  is an  $\{\mathcal{F}_k^{h_L}, k \geq 1\}$ -martingale increment sequence. Also, the random variables  $\{(e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell}, 1 \leq \ell \leq L\}$  are independent and centered. Hence

$$S_L^{1,1} = \sum_{k=1}^{N_1} \mathbb{E} \left[ \left( \sum_{\ell=1}^L h_L^{-1} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell} \right)^2 \right] = \sum_{\ell=1}^L U_\ell,$$

with

$$\begin{aligned} U_\ell &:= \sum_{k=1}^{N_\ell} h_L^{-2} \gamma_k^2 \Pi_{k+1:N_\ell}^2 \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] \\ &= \gamma_1 \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{-1} h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \left( \gamma_{N_\ell}^{-1} \sum_{k=1}^{N_\ell} \gamma_k^2 \Pi_{k+1:N_\ell}^2 \right) h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] \\ &= \gamma_1 \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{-1} h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \Sigma_{N_\ell} h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2], \end{aligned}$$

where we used (4.11) and the definition (2.31) of  $\Sigma_n$ . The uniform boundedness of  $\{f_{X_h}, h \in \mathcal{H}\}$  and Lemma 2.2(i) give

$$\begin{aligned}
& h_\ell^{-\frac{1}{2}} \left| \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] - \frac{1}{(1-\alpha)^2} \mathbb{E}[\mathbb{1}_{X_{h_\ell} > \xi_\star} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star}] \right| \\
& \leq \frac{h_\ell^{-\frac{1}{2}}}{(1-\alpha)^2} (\mathbb{E}[\mathbb{1}_{X_{h_\ell} > \xi_\star^{h_\ell}} - \mathbb{1}_{X_{h_\ell} > \xi_\star}] + \mathbb{E}[\mathbb{1}_{X_{h_{\ell-1}} > \xi_\star^{h_{\ell-1}}} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star}]) \\
& \leq \frac{h_\ell^{-\frac{1}{2}}}{(1-\alpha)^2} (|F_{X_{h_\ell}}(\xi_\star^{h_\ell}) - F_{X_{h_\ell}}(\xi_\star)| + |F_{X_{h_{\ell-1}}}(\xi_\star^{h_{\ell-1}}) - F_{X_{h_{\ell-1}}}(\xi_\star)|) \\
& \leq \frac{\sup_{h \in \mathcal{H}} \|f_{X_h}\|_\infty}{(1-\alpha)^2} h_\ell^{-\frac{1}{2}} (|\xi_\star^{h_\ell} - \xi_\star| + |\xi_\star^{h_{\ell-1}} - \xi_\star|) \\
& \leq C h_\ell^{\frac{1}{2}}.
\end{aligned}$$

Hence, by Lemma 4.1(ii),

$$\lim_{\ell \uparrow \infty} h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] = \frac{1}{(1-\alpha)^2} \lim_{\ell \uparrow \infty} h_\ell^{-\frac{1}{2}} \mathbb{E}[\mathbb{1}_{X_{h_\ell} > \xi_\star} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star}] = \frac{\mathbb{E}[|G|f_G(\xi_\star)]}{(1-\alpha)^2}. \quad (4.20)$$

Now, using the definitions (2.31) and (2.32),

$$\begin{aligned}
S_L^{1,1} &= \gamma_1 \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{-1} \sum_{\ell=1}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \Sigma_{N_\ell} h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] \\
&= \gamma_1 \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{-1} \sum_{\ell=1}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \Sigma_\star h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] \\
&\quad + \gamma_1 \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{-1} \sum_{\ell=1}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \Delta \Sigma_{N_\ell} h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2],
\end{aligned}$$

where  $\Delta \Sigma_{N_\ell} = \Sigma_{N_\ell} - \Sigma_\star$ . Reusing the notation from the result (2.33), there exists  $\mu > 0$  such that for  $\varepsilon > 0$ , there exists  $n_0 \geq 0$  such that, for  $n \geq n_0$ ,  $1 - (\mu + \varepsilon)\gamma_n > 0$  and  $|\Delta \Sigma_n| \leq \varepsilon$ . One further has the existence of  $L_0 \geq 1$  such that, for  $L \geq L_0$ ,  $N_1 \geq \dots \geq N_L \geq n_0$  so that  $|\Delta \Sigma_{N_L}| \leq \varepsilon$ . By the inequality (2.33) and Lemma A.1(i), for  $n \geq N_L$ ,

$$\begin{aligned}
|\Delta \Sigma_n| &\leq |\Delta \Sigma_{N_L}| \exp \left( -(\mu + \varepsilon) \sum_{k=N_L}^n \gamma_k \right) + \varepsilon \sum_{k=N_L}^n \gamma_k \exp \left( -(\mu + \varepsilon) \sum_{j=k}^n \gamma_j \right) \\
&\leq |\Delta \Sigma_{N_L}| + C\varepsilon \leq C\varepsilon.
\end{aligned}$$

In particular,  $\sup_{1 \leq \ell \leq L} |\Delta \Sigma_{N_\ell}| \leq C\varepsilon$ . Thus

$$\limsup_{L \uparrow \infty} \sup_{1 \leq \ell \leq L} |\Delta \Sigma_{N_\ell}| = 0.$$

Using the previous limit, (4.20), and the ‘./ $\infty$ ’ case of Cesàro’s lemma, we obtain

$$\begin{aligned}
\limsup_{L \uparrow \infty} \left| \gamma_1 \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{-1} \sum_{\ell=1}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \Delta \Sigma_{N_\ell} h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] \right| \\
\leq C \limsup_{L \uparrow \infty} \sup_{1 \leq \ell \leq L} |\Delta \Sigma_{N_\ell}| = 0.
\end{aligned}$$

Finally, again by (4.20) and Cesàro’s lemma (‘./ $\infty$ ’ version),

$$\begin{aligned}
\lim_{L \uparrow \infty} S_L^{1,1} &= \lim_{L \uparrow \infty} \gamma_1 \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \right)^{-1} \left( \sum_{\ell=1}^L h_\ell^{-\frac{2\beta-1}{2(1+\beta)}} \Sigma_\star h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_k^{h_\ell} - e_k^{h_{\ell-1}})^2] \right) \\
&= \frac{\gamma_1 \mathbb{E}[|G|f_G(\xi_\star)]}{(1-\alpha)^2 (2V_0''(\xi_\star) - \frac{1}{\gamma_1})}.
\end{aligned}$$

In order to deal with  $\{S_L^{2,2}, L \geq 1\}$ , we proceed similarly. In particular, one has

$$\begin{aligned} S_L^{2,2} &= \sum_{k=1}^{N_1} \mathbb{E} \left[ \left( \sum_{\ell=1}^L h_L^{-\frac{1}{\beta} - \frac{2\beta-1}{4\beta(1+\beta)}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbf{1}_{1 \leq k \leq N_\ell} \right)^2 \right] \\ &= \sum_{\ell=1}^L \frac{h_L^{-\frac{2}{\beta} - \frac{2\beta-1}{2\beta(1+\beta)}}}{N_\ell^2} \sum_{k=1}^{N_\ell} \mathbb{E}[(\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}})^2] \\ &= \sum_{\ell=1}^L W_\ell, \end{aligned}$$

where we introduced, for  $\ell \geq 1$

$$\begin{aligned} W_\ell &:= \frac{h_L^{-\frac{2}{\beta} - \frac{2\beta-1}{2\beta(1+\beta)}}}{N_\ell} \mathbb{E}[(\eta_1^{h_\ell} - \eta_1^{h_{\ell-1}})^2] \\ &= \frac{h_L^{-\frac{2}{\beta} - \frac{2\beta-1}{2\beta(1+\beta)}}}{(1-\alpha)^2} \frac{h_\ell}{N_\ell} h_\ell^{-1} \mathbb{V}\text{ar}((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+). \end{aligned}$$

By Lemma 2.2(i),

$$\begin{aligned} h_\ell^{-\frac{1}{2}} \mathbb{E}[|(X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+ - ((X_{h_\ell} - \xi_\star)^+ - (X_{h_{\ell-1}} - \xi_\star)^+)|] \\ \leq h_\ell^{-\frac{1}{2}} (|\xi_\star^{h_\ell} - \xi_\star| + |\xi_\star^{h_{\ell-1}} - \xi_\star|) \leq Ch_\ell^{\frac{1}{2}}, \end{aligned}$$

so that, by Lemma 4.1(iii) and Slutsky's theorem,

$$h_\ell^{-\frac{1}{2}} ((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+) \xrightarrow[\ell \uparrow \infty]{\mathcal{L}} \mathbf{1}_{X_0 > \xi_\star} G.$$

By Lemmas 2.1(i) and 2.2(i),

$$\sup_{\ell \geq 1} h_\ell^{-\frac{p}{2}} \mathbb{E}[|(X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+|^p] < \infty,$$

for some  $p > 2$ . The previous weak convergence together with the above uniform integrability yield

$$\lim_{\ell \uparrow \infty} h_\ell^{-1} \mathbb{V}\text{ar}((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+) = \mathbb{V}\text{ar}(\mathbf{1}_{X_0 > \xi_\star} G). \quad (4.21)$$

Eventually, (4.11) and Cesàro's lemma ('0/0' version) give

$$\begin{aligned} \lim_{L \uparrow \infty} S_L^{2,2} &= \lim_{L \uparrow \infty} \frac{h_L^{-\frac{2}{\beta} - \frac{2\beta-1}{2\beta(1+\beta)}}}{(1-\alpha)^2} \left( \sum_{\ell=1}^L \frac{h_\ell}{N_\ell} \right) \\ &\quad \times \lim_{L \uparrow \infty} \left( \sum_{\ell=1}^L \frac{h_\ell}{N_\ell} \right)^{-1} \left( \sum_{\ell=1}^L \frac{h_\ell}{N_\ell} h_\ell^{-1} \mathbb{V}\text{ar}((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+) \right) \\ &= \frac{h_0^{\frac{2\beta-1}{2(1+\beta)}} (M^{\frac{2\beta-1}{2(1+\beta)}} - 1)^{\frac{1}{\beta}-1}}{(1-\alpha)^2} \\ &\quad \times \lim_{L \uparrow \infty} \left( \sum_{\ell=1}^L h_\ell^{\frac{2\beta-1}{2(1+\beta)}} \right)^{-1} \left( \sum_{\ell=1}^L h_\ell^{\frac{2\beta-1}{2(1+\beta)}} h_\ell^{-1} \mathbb{V}\text{ar}((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+) \right) \\ &= \frac{h_0^{\frac{2\beta-1}{2(1+\beta)}} (M^{\frac{2\beta-1}{2(1+\beta)}} - 1)^{\frac{1}{\beta}-1}}{(1-\alpha)^2} \mathbb{V}\text{ar}(\mathbf{1}_{X_0 > \xi_\star} G). \end{aligned}$$

We now prove that  $\{S_L^{1,2}, L \geq 1\}$  converges towards zero in  $L^1(\mathbb{P})$ . It follows from (4.20) and (4.21) that

$$\begin{aligned}\mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] &= \frac{1}{(1-\alpha)^2} \mathbb{E}[|\mathbb{1}_{X_{h_\ell} > \xi_\star} - \mathbb{1}_{X_{h_{\ell-1}} > \xi_\star}|] \leq Ch_\ell^{\frac{1}{2}}, \\ \mathbb{E}[(\eta_1^{h_\ell} - \eta_1^{h_{\ell-1}})^2] &= \frac{1}{(1-\alpha)^2} \mathbb{V}\text{ar}((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+) \leq Ch_\ell,\end{aligned}\tag{4.22}$$

which, by the Cauchy-Schwarz inequality, yields

$$\mathbb{E}[|(e_k^{h_\ell} - e_k^{h_{\ell-1}})(\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}})|] \leq Ch_\ell^{\frac{3}{4}}.\tag{4.23}$$

Hence, using (4.23) then (2.24), Lemma A.1(i), and (4.11), we obtain

$$\begin{aligned}|S_L^{1,2}| &\leq h_L^{-1-\frac{1}{\beta}-\frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \gamma_k |\Pi_{k+1:N_\ell}| \mathbb{E}[|(e_k^{h_\ell} - e_k^{h_{\ell-1}})(\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}})|] \\ &\leq Ch_L^{-1-\frac{1}{\beta}-\frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{h_\ell^{\frac{3}{4}}}{N_\ell} \sum_{k=1}^{N_\ell} \gamma_k |\Pi_{k+1:N_\ell}| \\ &\leq Ch_L^{-1+\frac{1}{\beta}+\frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L h_\ell^{-\frac{3(1-\beta)}{4(1+\beta)}} \\ &\leq C \begin{cases} h_L^{\frac{-\beta^2-\beta+3}{4\beta(1+\beta)}}, & \beta \in (\frac{1}{2}, 1), \\ h_L^{\frac{1}{8}} |\ln h_L|, & \beta = 1. \end{cases}\end{aligned}$$

Hence

$$S_L^{1,2} = S_L^{2,1} \xrightarrow{\mathbb{P}\text{-as}} 0 \quad \text{as } L \uparrow \infty.$$

It follows that

$$\left( h_L^{-1} \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}), h_L^{-\frac{1}{\beta}-\frac{2\beta-1}{4\beta(1+\beta)}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \eta_k^{h_\ell} - \eta_k^{h_{\ell-1}} \right) \xrightarrow[L \uparrow \infty]{\mathcal{L}} \mathcal{N}(0, \tilde{\Sigma}_\beta),$$

where

$$\tilde{\Sigma}_\beta := \begin{pmatrix} \frac{\gamma_1 \mathbb{E}[|G|f_G(\xi_\star)]}{(1-\alpha)(2f_{X_0}(\xi_\star) - (1-\alpha)\gamma_1^{-1}\mathbb{1}_{\beta=1})} & 0 \\ 0 & \frac{h_0^{\frac{2\beta-1}{2(1+\beta)}} (M^{\frac{2\beta-1}{2(1+\beta)}} - 1)^{\frac{1}{\beta}} \mathbb{V}\text{ar}(\mathbb{1}_{X_0 > \xi_\star} G)}{(1-\alpha)^2 \frac{2\beta-1}{M^{\frac{2\beta-1}{2(1+\beta)}} - 1}} \end{pmatrix}.$$

**Step 10. Conclusion.**

Combining Steps 1 to 7, the CLTs of Steps 8 and 9, and noting that the two sequences introduced in the two last steps are independent (by independency of the levels), concludes the demonstration.  $\square$

## 5 Averaged Multilevel Stochastic Approximation Algorithm

For each level  $0 \leq \ell \leq L$  and for  $j \in \{(\ell-1)^+, \ell\}$ , we consider the estimate  $\bar{\xi}_{N_\ell}^{h_j}$  calculated by averaging out the simulations  $\xi_0^{h_j}, \dots, \xi_{N_\ell}^{h_j}$  from (4.6). The averaged MLSA (AMLSA) approximation for the VaR is

$$\bar{\xi}_{\mathbf{N}}^{\text{ML}} = \bar{\xi}_{N_0}^{h_0} + \sum_{\ell=1}^L \bar{\xi}_{N_\ell}^{h_\ell} - \bar{\xi}_{N_\ell}^{h_{\ell-1}}.\tag{5.1}$$

## 5.1 Convergence Rate Analysis

**Theorem 5.1.** *Suppose that the assumptions of Theorem 2.1 are satisfied, that Assumption 2 holds, that  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $\beta \in (\frac{8}{9}, 1)$ , and that  $\delta_0 \geq \frac{1}{8}$ . Let*

$$N_\ell = h_L^{-2} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{1}{4}} \right) h_\ell^{\frac{3}{4}}, \quad 0 \leq \ell \leq L. \quad (5.2)$$

Then

$$\begin{pmatrix} h_L^{-1} (\bar{\xi}_N^{\text{ML}} - \xi_\star^{h_L}) \\ h_L^{-\frac{9}{8}} (\chi_N^{\text{ML}} - \chi_\star^{h_L}) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \bar{\Sigma}^{\text{ML}}) \quad \text{as } L \uparrow \infty,$$

where

$$\bar{\Sigma}^{\text{ML}} = \begin{pmatrix} \frac{\mathbb{E}[|G|f_G(\xi_\star)]}{(1-\alpha)^2(1-M^{-\frac{1}{4}})} & 0 \\ 0 & \frac{h_0^{-\frac{3}{8}}(1-M^{-\frac{1}{4}})^{\frac{1}{2}} \text{Var}((X_{h_0} - \xi_\star^{h_0})^+)}{(1-\alpha)^2} + \frac{h_0^{\frac{1}{4}} \text{Var}(\mathbf{1}_{X_0 > \xi_\star} G)}{(1-\alpha)^2 M_0^{\frac{1}{4}}} \end{pmatrix}. \quad (5.3)$$

**Remark 5.1.** *Contrary to the classical Ruppert-Polyak averaging principle, the above CLT is not valid for any  $\beta \in (\frac{1}{2}, 1)$ , but requires  $\beta \in (\frac{8}{9}, 1)$ . This lower threshold on the admissible values of  $\beta$  appears when one studies the convergence of the different related terms in the proof below.*

The values of  $N_\ell$  in (5.2) are determined by minimizing the algorithm complexity while maintaining the  $L^2$  error at order  $\varepsilon^2$ , as detailed in Section B.2; the ensuing complexity is discussed in Section 6.

*Proof.* In the following developments, we denote by  $C$  a positive constant whose value may change from line to line and does not depend upon  $L$ .

Recalling the sequences  $\{a_n^h, n \geq 1\}$ ,  $\{g_n^h, n \geq 1\}$ ,  $\{r_n^h, n \geq 1\}$ ,  $\{\rho_n^h, n \geq 1\}$  and  $\{e_n^h, n \geq 1\}$  from (3.4) and (2.13)–(2.16), the decomposition (2.17) implies

$$\xi_n^h - \xi_\star^h = \frac{1}{V_0''(\xi_\star)} a_n^h + \frac{1}{V_0''(\xi_\star)} g_n^h + \frac{1}{V_0''(\xi_\star)} r_n^h + \frac{1}{V_0''(\xi_\star)} \rho_n^h + \frac{1}{V_0''(\xi_\star)} e_n^h. \quad (5.4)$$

Averaging the previous identity yields

$$\begin{aligned} \bar{\xi}_n^h - \xi_\star^h &= \frac{1}{V_0''(\xi_\star)n} \sum_{k=1}^n a_k^h + \frac{1}{V_0''(\xi_\star)n} \sum_{k=1}^n g_k^h \\ &\quad + \frac{1}{V_0''(\xi_\star)n} \sum_{k=1}^n r_k^h + \frac{1}{V_0''(\xi_\star)n} \sum_{k=1}^n \rho_k^h + \frac{1}{V_0''(\xi_\star)n} \sum_{k=1}^n e_k^h. \end{aligned} \quad (5.5)$$

From (5.1), (4.2), and the above equality, we obtain

$$\begin{aligned} \bar{\xi}_N^{\text{ML}} - \xi_\star^{h_L} &= \bar{\xi}_{N_0}^{h_0} - \xi_\star^{h_0} + \sum_{\ell=1}^L (\bar{\xi}_{N_\ell}^{h_\ell} - \xi_\star^{h_\ell} - (\bar{\xi}_{N_{\ell-1}}^{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})) \\ &= \bar{\xi}_{N_0}^{h_0} - \xi_\star^{h_0} + \frac{1}{V_0''(\xi_\star)} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_\ell} - a_k^{h_{\ell-1}} \end{aligned} \quad (5.6)$$

$$+ \frac{1}{V_0''(\xi_\star)} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}} + \frac{1}{V_0''(\xi_\star)} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} r_k^{h_\ell} - r_k^{h_{\ell-1}} \quad (5.7)$$

$$+ \frac{1}{V_0''(\xi_\star)} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \rho_k^{h_\ell} - \rho_k^{h_{\ell-1}} + \frac{1}{V_0''(\xi_\star)} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} e_k^{h_\ell} - e_k^{h_{\ell-1}}. \quad (5.8)$$

We now study each term of the above decomposition.

**Step 1. Study of  $\{h_L^{-1}(\bar{\xi}_{N_0}^{h_0} - \xi_\star^{h_0}), L \geq 1\}$ .**

Thanks to (3.1), Lemma 2.2(ii) and a comparison between series and integrals,

$$\mathbb{E}[(\bar{\xi}_{N_0}^{h_0} - \xi_\star^{h_0})^2] \leq \frac{1}{N_0} \sum_{k=1}^{N_0} \mathbb{E}[(\xi_k^{h_0} - \xi_\star^{h_0})^2] \leq \frac{C}{N_0} \sum_{k=1}^{N_0} \gamma_k \leq C\gamma_{N_0}.$$

Hence, using (5.2),

$$\mathbb{E}[|h_L^{-1}(\bar{\xi}_{N_0}^{h_0} - \xi_\star^{h_0})|^2] \leq Ch_L^{-2} \gamma_{N_0} \leq Ch_L^{-2+\frac{9\beta}{4}}.$$

Recalling that  $\beta \in (\frac{8}{9}, 1)$ , this implies that

$$h_L^{-1}(\bar{\xi}_{N_0}^{h_0} - \xi_\star^{h_0}) \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 2. Study of  $\{h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_\ell} - a_k^{h_{\ell-1}}, L \geq 1\}$ .**

From the inequality (3.7) and (5.2), we obtain

$$\mathbb{E}\left[\left|h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_\ell} - a_k^{h_{\ell-1}}\right|\right] \leq Ch_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell \sqrt{\gamma_{N_\ell}}} \leq Ch_L^{\frac{3\beta+2}{4}},$$

hence

$$h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_\ell} - a_k^{h_{\ell-1}} \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 3. Study of  $\{h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}}, L \geq 1\}$ .**

Using (4.15) and a comparison between series and integrals, we obtain

$$\begin{aligned} \mathbb{E}\left[\left|h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}}\right|\right] &\leq 2h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} (\mathbb{E}[|g_k^{h_\ell}|] + \mathbb{E}[|g_k^{h_{\ell-1}}|]) \\ &\leq Ch_L^{-1} \sum_{\ell=1}^L h_\ell^{(\frac{1}{4}+\delta_0)\wedge 1} \gamma_{N_\ell}^{\frac{1}{2}}. \end{aligned}$$

If  $\delta_0 \geq \frac{3}{4}$ , by (5.2),

$$\begin{aligned} \mathbb{E}\left[\left|h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}}\right|\right] &\leq Ch_L^{-1} \sum_{\ell=1}^L h_\ell \gamma_{N_\ell}^{\frac{1}{2}} \\ &\leq Ch_L^{-1+\frac{9}{8}\beta} \sum_{\ell=1}^L h_\ell^{1-\frac{3}{8}\beta} \\ &\leq Ch_L^{-1+\frac{9}{8}\beta}, \end{aligned}$$



where we used the fact that  $\beta < 1$ . Else, if  $\frac{1}{8} \leq \delta_0 < \frac{3}{4}$ , by (5.2) again,

$$\begin{aligned}
& \mathbb{E} \left[ \left| h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}} \right| \right] \\
& \leq C h_L^{-1} \sum_{\ell=1}^L h_\ell^{\frac{1}{4} + \delta_0} \gamma_{N_\ell}^{\frac{1}{2}} \\
& \leq C h_L^{-1 + \frac{9}{8}\beta} \sum_{\ell=1}^L h_\ell^{\frac{1}{4} + \delta_0 - \frac{3}{8}\beta} \\
& \leq C \begin{cases} h_L^{-1 + \frac{9}{8}\beta}, & \delta_0 > \frac{1}{8}, \\ h_L^{-1 + \frac{9}{8}\beta} |\ln h_L|, & \delta_0 = \frac{1}{8}. \end{cases}
\end{aligned}$$

Taking into account the fact that  $\beta \in (\frac{8}{9}, 1)$ , we thus obtain

$$h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}} \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 4. Study of**  $\left\{ h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} r_k^{h_\ell} - r_k^{h_{\ell-1}}, L \geq 1 \right\}$ .

Using (3.8), a comparison between series and integrals and (5.2), we obtain

$$\begin{aligned}
& \mathbb{E} \left[ \left| h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} r_k^{h_\ell} - r_k^{h_{\ell-1}} \right| \right] \\
& \leq 2 h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} (\mathbb{E}[|r_k^{h_\ell}|] + \mathbb{E}[|r_k^{h_{\ell-1}}|]) \\
& \leq C h_L^{-1} \sum_{\ell=1}^L \gamma_{N_\ell} \\
& \leq C h_L^{-1 + \frac{3}{2}\beta}.
\end{aligned}$$

In view of the fact that  $\beta \in (\frac{8}{9}, 1)$ , we obtain

$$h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} r_k^{h_\ell} - r_k^{h_{\ell-1}} \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 5. Study of**  $\left\{ h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}, L \geq 1 \right\}$ .

Using similar arguments to the ones used in Step 5 of the proof of Theorem 4.1, we obtain

from (3.9), (5.2), and a comparison between series and integrals, that

$$\begin{aligned}
& \mathbb{E} \left[ \left( h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \rho_k^{h_\ell} - \rho_k^{h_{\ell-1}} \right)^2 \right] \\
&= h_L^{-2} \sum_{\ell=1}^L \frac{1}{N_\ell^2} \sum_{k=1}^{N_\ell} \mathbb{E} [(\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}})^2] \\
&\leq 2h_L^{-2} \sum_{\ell=1}^L \frac{1}{N_\ell^2} \sum_{k=1}^{N_\ell} (\mathbb{E} [|\rho_k^{h_\ell}|^2] + \mathbb{E} [|\rho_k^{h_{\ell-1}}|^2]) \\
&\leq Ch_L^{-2} \sum_{\ell=1}^L \frac{\gamma_{N_\ell}^{\frac{1}{2}}}{N_\ell} \\
&\leq Ch_L^{\frac{3\beta-2}{4}}.
\end{aligned}$$

Hence, given that  $\beta \in (\frac{8}{9}, 1)$ ,

$$h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \rho_k^{h_\ell} - \rho_k^{h_{\ell-1}} \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 6. Study of**  $\left\{ h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}}, L \geq 1 \right\}$ .

The random variables  $\left\{ \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}}, \ell \geq 1 \right\}$  are independent and centered. Hence, using (2.29), a comparison between series and integrals, and then (5.2), we obtain

$$\begin{aligned}
& \mathbb{E} \left[ \left| h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}} \right|^2 \right] \\
&\leq 2h_L^{-\frac{9}{4}} \sum_{\ell=1}^L \frac{1}{N_\ell} \left( \mathbb{E} \left[ \left( \frac{1}{\sqrt{N_\ell}} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} \right)^2 \right] + \mathbb{E} \left[ \left( \frac{1}{\sqrt{N_\ell}} \sum_{k=1}^{N_\ell} \theta_k^{h_{\ell-1}} \right)^2 \right] \right) \\
&\leq Ch_L^{-\frac{9}{4}} \sum_{\ell=1}^L \frac{\gamma_{N_\ell}}{N_\ell} \\
&\leq Ch_L^{-\frac{3}{4} + \frac{3}{2}\beta}.
\end{aligned}$$

Therefore

$$h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \theta_k^{h_\ell} - \theta_k^{h_{\ell-1}} \xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 7. Study of**  $\left\{ h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \zeta_k^{h_\ell} - \zeta_k^{h_{\ell-1}}, L \geq 1 \right\}$ .

Using (2.30), we obtain

$$\begin{aligned}
& \mathbb{E} \left[ \left| h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \zeta_k^{h_\ell} - \zeta_k^{h_{\ell-1}} \right|^2 \right] \\
&\leq h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{\sqrt{N_\ell}} \left( \mathbb{E} \left[ \left| \frac{1}{\sqrt{N_\ell}} \sum_{k=1}^{N_\ell} \zeta_k^{h_\ell} \right|^2 \right] + \mathbb{E} \left[ \left| \frac{1}{\sqrt{N_\ell}} \sum_{k=1}^{N_\ell} \zeta_k^{h_{\ell-1}} \right|^2 \right] \right) \\
&\leq Ch_L^{-\frac{9}{8}} \sum_{\ell=1}^L \gamma_{N_\ell} \\
&\leq Ch_L^{-\frac{9}{8} + \frac{3}{2}\beta}.
\end{aligned}$$

Hence, since  $\beta \in (\frac{8}{9}, 1)$ ,

$$h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \zeta_k^{h_\ell} - \zeta_k^{h_{\ell-1}} \xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } L \uparrow \infty.$$

**Step 8. Study of  $\left\{ h_L^{-\frac{9}{8}} (\chi_{N_0}^{h_0} - \chi_\star^{h_0}), L \geq 1 \right\}$ .**

It follows from (5.2) that

$$h_L^{-\frac{9}{8}} (\chi_{N_0}^{h_0} - \chi_\star^{h_0}) = h_0^{-\frac{3}{8}} (1 - M^{-\frac{1}{4}})^{\frac{1}{2}} \frac{\sqrt{N_0} (\chi_{N_0}^{h_0} - \chi_\star^{h_0})}{(1 - M^{-\frac{1}{4}(L+1)})^{\frac{1}{2}}}.$$

According to the decomposition (2.20),

$$\sqrt{N_0} (\chi_{N_0}^{h_0} - \chi_\star^{h_0}) = \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \theta_k^{h_0} + \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \zeta_k^{h_0} + \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \eta_k^{h_0}.$$

From (2.29) and (2.30), we obtain

$$\begin{aligned} \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \theta_k^{h_0} &\xrightarrow{L^2(\mathbb{P})} 0 \quad \text{as } N_0 \uparrow \infty, \\ \frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \zeta_k^{h_0} &\xrightarrow{L^1(\mathbb{P})} 0 \quad \text{as } N_0 \uparrow \infty. \end{aligned}$$

Since  $\{\eta_k^{h_0}, k \geq 1\}$  is an iid sequence satisfying  $\mathbb{E}[|\eta_1^{h_0}|^2] < \infty$ , the classical CLT yields

$$\frac{1}{\sqrt{N_0}} \sum_{k=1}^{N_0} \eta_k^{h_0} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\mathbb{V}\text{ar}((X_{h_0} - \xi_\star^{h_0})^+)}{(1 - \alpha)^2}\right) \quad \text{as } N_0 \uparrow \infty,$$

so that

$$h_L^{-\frac{9}{8}} (\chi_{N_0}^{h_0} - \chi_\star^{h_0}) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, h_0^{-\frac{3}{4}} (1 - M^{-\frac{1}{4}}) \frac{\mathbb{V}\text{ar}((X_{h_0} - \xi_\star^{h_0})^+)}{(1 - \alpha)^2}\right) \quad \text{as } L \uparrow \infty.$$

**Step 9. Study of**

$$\left\{ \left( h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} e_k^{h_\ell} - e_k^{h_{\ell-1}}, h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \eta_k^{h_\ell} - \eta_k^{h_{\ell-1}} \right), L \geq 1 \right\}.$$

Since  $N_1 \geq \dots \geq N_L$ ,

$$\sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} e_k^{h_\ell} - e_k^{h_{\ell-1}} = \sum_{k=1}^{N_1} \sum_{\ell=1}^L \frac{e_k^{h_\ell} - e_k^{h_{\ell-1}}}{N_\ell} \mathbf{1}_{1 \leq k \leq N_\ell}.$$

We apply again the CLT for martingale arrays (Hall & Heyde, 1980, Corollary 3.1) to the sequence

$$\left\{ \left( h_L^{-1} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} e_k^{h_\ell} - e_k^{h_{\ell-1}}, h_L^{-\frac{9}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \eta_k^{h_\ell} - \eta_k^{h_{\ell-1}} \right), L \geq 1 \right\}.$$

We first check the conditional Lindeberg condition. Let  $\delta > 0$ . Since, for all  $k \geq 1$ , the sequence  $\left\{ \frac{1}{N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) \mathbb{1}_{1 \leq k \leq N_\ell}, 1 \leq \ell \leq L \right\}$  is a sequence of independent centered random variables, using first the Marcinkiewicz-Zygmund inequality and then Jensen's inequality,

$$\begin{aligned}
& \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-1} \frac{e_k^{h_\ell} - e_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] \\
& \leq C h_L^{-(2+\delta)} \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L \frac{(e_k^{h_\ell} - e_k^{h_{\ell-1}})^2}{N_\ell^2} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{1+\frac{\delta}{2}} \right] \\
& \leq C h_L^{-(2+\delta)} \sum_{k=1}^{N_1} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \frac{\mathbb{E} [|e_k^{h_\ell} - e_k^{h_{\ell-1}}|^{2+\delta}]}{N_\ell^{2+\delta}} \mathbb{1}_{1 \leq k \leq N_\ell} \\
& \leq C h_L^{-(2+\delta)} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \frac{h_\ell^{\frac{1}{2}}}{N_\ell^{1+\delta}},
\end{aligned}$$

where we used (4.17) for the last inequality. From the previous upper bound and (5.2), we obtain

$$\begin{aligned}
& \sum_{k=0}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-1} \frac{e_k^{h_\ell} - e_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] \\
& \leq C h_L^{-(2+\delta)} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \frac{h_\ell^{\frac{1}{2}}}{N_\ell^{1+\delta}} \\
& \leq C h_L^{\frac{1+5\delta}{4}} L^{\frac{\delta}{2}} \sum_{\ell=1}^L h_\ell^{\frac{1-3\delta}{4}} \\
& \leq C \begin{cases} h_L^{\frac{1+5\delta}{4}} L^{\frac{\delta}{2}}, & \delta < \frac{1}{3}, \\ h_L^{\frac{1+5\delta}{4}} L^{1+\frac{\delta}{2}}, & \delta = \frac{1}{3}, \\ h_L^{\frac{1}{2}(1+\delta)} L^{\frac{\delta}{2}}, & \delta > \frac{1}{3}. \end{cases}
\end{aligned}$$

We thus conclude

$$\limsup_{L \uparrow \infty} \sum_{k=0}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-1} \frac{e_k^{h_\ell} - e_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] = 0,$$

for any  $\delta > 0$ . Proceeding as in (4.18), using in particular (4.19) together with the fact that  $\sup_{h \in \mathcal{H}} \mathbb{E} [|X_h|^{2+\delta}] < \infty$  for some  $\delta > 0$  and (5.2),

$$\begin{aligned}
& \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-\frac{9}{8}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbb{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] \\
& \leq C h_L^{-\frac{9}{8}(2+\delta)} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \frac{1}{N_\ell^{2+\delta}} \sum_{k=1}^{N_\ell} \mathbb{E} [| \eta_k^{h_\ell} - \eta_k^{h_{\ell-1}} |^{2+\delta}] \\
& \leq C h_L^{-\frac{9}{8}(2+\delta)} L^{\frac{\delta}{2}} \sum_{\ell=1}^L \frac{h_\ell^{1+\frac{\delta}{2}}}{N_\ell^{1+\delta}} \\
& \leq C \begin{cases} h_L^{\frac{9}{8}\delta} L^{\frac{\delta}{2}}, & \delta < 1, \\ h_L^{\frac{9}{8}\delta} L^{1+\frac{\delta}{2}}, & \delta = 1, \\ h_L^{\frac{1}{4}+\frac{7}{8}\delta} L^{\frac{\delta}{2}}, & \delta > 1. \end{cases}
\end{aligned}$$

Thus

$$\limsup_{L \uparrow \infty} \sum_{k=1}^{N_1} \mathbb{E} \left[ \left| \sum_{\ell=1}^L h_L^{-\frac{9}{8}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbf{1}_{1 \leq k \leq N_\ell} \right|^{2+\delta} \right] = 0.$$

The conditional Lindeberg condition is thus satisfied.

It remains to investigate the asymptotic behavior of the conditional covariance matrices sequence  $\{S_L = (S_L^{i,j})_{1 \leq i,j \leq 2}, L \geq 1\}$  defined by

$$\begin{aligned} S_L^{1,1} &:= \sum_{k=1}^{N_1} \mathbb{E} \left[ \left( \sum_{\ell=1}^L h_L^{-1} \frac{e_k^{h_\ell} - e_k^{h_{\ell-1}}}{N_\ell} \mathbf{1}_{1 \leq k \leq N_\ell} \right)^2 \middle| \mathcal{F}_{k-1}^{h_L} \right], \\ S_L^{2,2} &:= \sum_{k=1}^{N_1} \mathbb{E} \left[ \left( \sum_{\ell=1}^L h_L^{-\frac{9}{8}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbf{1}_{1 \leq k \leq N_\ell} \right)^2 \middle| \mathcal{F}_{k-1}^{h_L} \right], \\ S_L^{1,2} = S_L^{2,1} &:= \sum_{k=1}^{N_1} \mathbb{E} \left[ \sum_{\ell=1}^L \frac{h_L^{-\frac{17}{8}}}{N_\ell^2} (e_k^{h_\ell} - e_k^{h_{\ell-1}})(\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}) \mathbf{1}_{1 \leq k \leq N_\ell} \middle| \mathcal{F}_{k-1}^{h_L} \right]. \end{aligned}$$

In order to deal with  $\{S_L^{1,1}, L \geq 1\}$ , we write

$$S_L^{1,1} = \sum_{k=1}^{N_1} \mathbb{E} \left[ \left( \sum_{\ell=1}^L h_L^{-1} \frac{e_k^{h_\ell} - e_k^{h_{\ell-1}}}{N_\ell} \mathbf{1}_{1 \leq k \leq N_\ell} \right)^2 \right] = \sum_{\ell=1}^L \bar{U}_\ell,$$

where, for  $\ell \geq 1$ ,

$$\bar{U}_\ell := \frac{h_L^{-2}}{N_\ell^2} \sum_{k=1}^{N_\ell} \mathbb{E}[(e_k^{h_\ell} - e_k^{h_{\ell-1}})^2] = \frac{h_L^{-2}}{N_\ell} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2].$$

By the ‘./ $\infty$ ’ case of Cesàro’s lemma, (4.20) and (5.2),

$$\begin{aligned} \lim_{L \uparrow \infty} S_L^{1,1} &= \lim_{L \uparrow \infty} h_L^{-2} \left( \sum_{\ell=1}^L \frac{h_\ell^{\frac{1}{2}}}{N_\ell} \right) \times \lim_{L \uparrow \infty} \left( \sum_{\ell=1}^L \frac{h_\ell^{\frac{1}{2}}}{N_\ell} \right)^{-1} \left( \sum_{\ell=1}^L \frac{h_\ell^{\frac{1}{2}}}{N_\ell} h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] \right) \\ &= \frac{1}{(1 - M^{-\frac{1}{4}})} \times \lim_{L \uparrow \infty} \left( \sum_{\ell=1}^L h_\ell^{-\frac{1}{4}} \right)^{-1} \left( \sum_{\ell=1}^L h_\ell^{-\frac{1}{4}} h_\ell^{-\frac{1}{2}} \mathbb{E}[(e_1^{h_\ell} - e_1^{h_{\ell-1}})^2] \right) \\ &= \frac{\mathbb{E}[|G|f_G(\xi_*)]}{(1 - \alpha)^2 (1 - M^{-\frac{1}{4}})}. \end{aligned}$$

As for  $\{S_L^{2,2}, L \geq 1\}$ , we have

$$S_L^{2,2} = \sum_{k=1}^{N_1} \mathbb{E} \left[ \left( \sum_{\ell=1}^L h_L^{-\frac{9}{8}} \frac{\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}}}{N_\ell} \mathbf{1}_{1 \leq k \leq N_\ell} \right)^2 \right] = \sum_{\ell=1}^L \bar{W}_\ell,$$

with, for  $\ell \geq 1$ ,

$$\begin{aligned} \bar{W}_\ell &:= \frac{h_L^{-\frac{9}{4}}}{N_\ell} \mathbb{E}[(\eta_1^{h_\ell} - \eta_1^{h_{\ell-1}})^2] \\ &= \frac{h_L^{-\frac{9}{4}}}{(1 - \alpha)^2} \frac{h_\ell}{N_\ell} h_\ell^{-1} \mathbb{V}\text{ar}((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+). \end{aligned}$$

In accordance with the convergence result (4.21), the ‘0/0’ case of Cesàro’s lemma and (5.2) give

$$\begin{aligned}
\lim_{L \uparrow \infty} S_L^{2,2} &= \lim_{L \uparrow \infty} \frac{h_L^{-\frac{9}{4}}}{(1-\alpha)^2} \left( \sum_{\ell=1}^L \frac{h_\ell}{N_\ell} \right) \\
&\quad \times \lim_{L \uparrow \infty} \left( \sum_{\ell=1}^L \frac{h_\ell}{N_\ell} \right)^{-1} \left( \sum_{\ell=1}^L \frac{h_\ell}{N_\ell} h_\ell^{-1} \mathbb{V}\text{ar}((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+) \right) \\
&= \frac{h_0^{\frac{1}{4}}}{(1-\alpha)^2 M^{\frac{1}{4}}} \\
&\quad \times \lim_{L \uparrow \infty} \left( \sum_{\ell=1}^L h_\ell^{\frac{1}{4}} \right)^{-1} \left( \sum_{\ell=1}^L h_\ell^{\frac{1}{4}} h_\ell^{-1} \mathbb{V}\text{ar}((X_{h_\ell} - \xi_\star^{h_\ell})^+ - (X_{h_{\ell-1}} - \xi_\star^{h_{\ell-1}})^+) \right) \\
&= \frac{h_0^{\frac{1}{4}}}{(1-\alpha)^2 M^{\frac{1}{4}}} \mathbb{V}\text{ar}(\mathbb{1}_{X_0 > \xi_\star} G).
\end{aligned}$$

From (4.23), (5.2), we obtain

$$\begin{aligned}
|S_L^{1,2}| &\leq h_L^{-\frac{17}{8}} \sum_{\ell=1}^L \frac{1}{N_\ell^2} \sum_{k=1}^{N_\ell} \mathbb{E}[|(e_k^{h_\ell} - e_k^{h_{\ell-1}})(\eta_k^{h_\ell} - \eta_k^{h_{\ell-1}})|] \\
&\leq C h_L^{-\frac{17}{8}} \sum_{\ell=1}^L \frac{h_\ell^{\frac{3}{4}}}{N_\ell} \\
&\leq C h_L^{\frac{1}{8}} L,
\end{aligned}$$

yielding

$$S_L^{1,2} = S_L^{2,1} \xrightarrow{\mathbb{P}\text{-as}} 0 \quad \text{as } L \uparrow \infty.$$

The proof is now complete.  $\square$

## 6 Complexity Analysis

We briefly discuss here the optimal complexities of the SA schemes that stem from our results. Let  $\varepsilon > 0$  denote a target total error order.

By Corollary 2.1, the convergence rate of the nested SA scheme is of order  $h^\beta$  as  $\mathcal{H} \ni h \downarrow 0$ . To achieve a prescribed error of order  $\varepsilon$ , we have to choose a bias parameter  $h = \varepsilon^{1/\beta}$ ,  $\beta \in (\frac{1}{2}, 1]$ , and a number of iterations  $n = \lceil h^{-2} \rceil = \lceil \varepsilon^{-2/\beta} \rceil$ . The ensuing complexity is

$$\text{Cost}_{\text{NSA}}^\beta = C \frac{n}{h} = C \varepsilon^{-\frac{3}{\beta}}. \quad (6.1)$$

The optimal complexity is reached for  $\beta = 1$  under the constraint  $\lambda\gamma_1 > 1$ , for which

$$\text{Cost}_{\text{NSA}}^1 = C \varepsilon^{-3}. \quad (6.2)$$

The complexity of the averaged nested SA (ANSA) scheme (cf. Corollary 3.1) with a bias parameter  $h = \varepsilon$  and  $n = \lceil h^{-2} \rceil = \lceil \varepsilon^{-2} \rceil$  SA iterations is

$$\text{Cost}_{\text{ANSA}} = C \frac{n}{h} = C \varepsilon^{-3}, \quad (6.3)$$

regardless of the value of  $\beta \in (\frac{1}{2}, 1)$ , which allows to circumvent the constraint on  $\gamma_1 > 0$  that appears in the case of the nested SA scheme.

Concerning the MLSA scheme, Lemma 2.2(i) and Theorem 4.1 show that the convergence rate of this algorithm is of order  $h_L$  as  $L \uparrow \infty$ , so that to achieve a prescribed error  $\varepsilon$ , we must choose  $h_L$  of the order of  $\varepsilon$ , i.e.

$$L = \left\lceil \frac{\ln h_0 \varepsilon^{-1}}{\ln M} \right\rceil. \quad (6.4)$$

Via (4.11), this results in the complexity

$$\text{Cost}_{\text{MLSA}}^\beta = C \sum_{\ell=0}^L \frac{N_\ell}{h_\ell} = C \varepsilon^{-1 - \frac{3}{2\beta}}. \quad (6.5)$$

This complexity is minimal for  $\beta = 1$ , provided that the condition  $\lambda\gamma_1 > 1$  is satisfied, and is worth

$$\text{Cost}_{\text{MLSA}}^1 = C \varepsilon^{-\frac{5}{2}}. \quad (6.6)$$

Finally, regarding the averaged MLSA scheme, Theorem 5.1 and Lemma 2.2(i) show that, we must choose  $L$  as in (6.4) to obtain  $h_L$  of the order of  $\varepsilon$ . Using (5.2), the ensuing complexity is

$$\text{Cost}_{\text{AMLSA}} = C \sum_{\ell=0}^L \frac{N_\ell}{h_\ell} = C \varepsilon^{-\frac{5}{2}}. \quad (6.7)$$

This complexity is independent of  $\beta \in (\frac{8}{9}, 1)$  and bears no additional condition on  $\gamma_1 > 0$ .

## 7 Financial Case Study

The code for this numerical case study is available at [github.com/azarlouzi/avg\\_mlsa](https://github.com/azarlouzi/avg_mlsa).

We consider a long position in a swap of strike  $\bar{K}$ , maturity  $T$  on some underlying (e.g. FX or interest) rate. The swap is issued at par. The rate's risk neutral model  $\{S_t, 0 \leq t \leq T\}$  is a Bachelier process of initial value  $S_0$ , drift  $\kappa$  and volatility  $\sigma$ . The swap pays at reset dates  $0 < T_1 < \dots < T_d = T$  the cash flows  $\Delta T_i (S_{T_{i-1}} - \bar{K})$ , where  $\Delta T_i = T_i - T_{i-1}$ . The swap's nominal  $\text{Nom}$  is set so that each leg is worth 100 at inception. The risk-free rate is  $r$ . For  $t \in [0, T]$ , we let  $i_t$  denote the integer such that  $t \in [T_{i_t-1}, T_{i_t})$  if  $t \in [0, T)$ , and  $i_t = +\infty$  otherwise.

Hence

$$dS_t = \kappa S_t dt + \sigma dW_t, \quad \text{i.e.} \quad S_t = S_0 e^{\kappa t} + \sigma \int_0^t e^{-\kappa(s-t)} dW_s, \quad (7.1)$$

where  $\{W_t, 0 \leq t \leq T\}$  is a standard Brownian motion. The fair value of the swap at time  $t \in [0, T]$  is

$$P_t = \text{Nom} \times \mathbb{E} \left[ \sum_{i=i_t}^d e^{-r(T_i-t)} \Delta T_i (S_{T_{i-1}} - \bar{K}) \middle| \mathcal{F}_t \right]$$

The loss, at some short time horizon  $\delta \in (0, T_1)$ , on a short position on the swap, is

$$X_0 = e^{-r\delta} P_\delta.$$

We are interested in retrieving the VaR  $\xi_\star$  and ES  $\chi_\star$ , at some confidence level  $\alpha \in (0, 1)$ , of the position.

## 7.1 Analytical and Simulation Formulas

The swap is issued at par, i.e.  $P_0 = 0$ , hence

$$\bar{K} = S_0 \frac{\sum_{i=1}^d e^{-rT_i} \Delta T_i e^{\kappa T_{i-1}}}{\sum_{i=1}^d e^{-rT_i} \Delta T_i}. \quad (7.2)$$

Note that  $i_\delta = 1$  so that, by (7.1)-(7.2) and the fact that  $\pm \int_0^\delta e^{-\kappa s} dW_s \sim \mathcal{N}(0, \int_0^\delta e^{-2\kappa s} ds)$ ,

$$X_0 \stackrel{\mathcal{L}}{=} \eta Y, \quad \text{where} \quad \eta = \text{Nom} \times \sigma \sqrt{\frac{1 - e^{-2\kappa\delta}}{2\kappa}} \sum_{i=2}^d e^{-rT_i} \Delta T_i e^{\kappa T_{i-1}} \quad (7.3)$$

and  $Y$  is a standard Gaussian random variable independent of  $W$ . This allows one to simulate  $X_0$  exactly.

On this basis, the values of  $\xi_\star$  and  $\chi_\star$  can be obtained analytically via

$$\alpha = \mathbb{P}(X_0 \leq \xi_\star) = \mathbb{P}(\eta Y \leq \xi_\star), \quad \text{i.e.} \quad \xi_\star = \eta F^{-1}(\alpha), \quad (7.4)$$

where  $F$  denotes the standard Gaussian cdf, and

$$\chi_\star = \mathbb{E}[X_0 | X_0 \geq \xi_\star] = \frac{\eta}{1 - \alpha} \mathbb{E}[Y \mathbb{1}_{\eta Y \geq \xi_\star}], \quad \text{i.e.} \quad \chi_\star = \frac{\eta}{1 - \alpha} f\left(\frac{\xi_\star}{\eta}\right), \quad (7.5)$$

where  $f$  is the standard Gaussian pdf.

We also have

$$X_0 = \text{Nom} \times \sigma \mathbb{E} \left[ \sum_{i=2}^d e^{-rT_i} \Delta T_i e^{\kappa T_{i-1}} \int_0^{T_{i-1}} e^{-\kappa s} dW_s \middle| \mathcal{F}_\delta \right] \stackrel{\mathcal{L}}{=} \mathbb{E}[\varphi(Y, Z) | Y], \quad (7.6)$$

where

$$\begin{aligned} Y &= \sqrt{\frac{1 - e^{-2\kappa\delta}}{2\kappa}} U_0 \sim \pm \int_0^\delta e^{-\kappa s} dW_s, \\ Z_1 &= \sqrt{\frac{1 - e^{-2\kappa(T_1 - \delta)}}{2\kappa}} U_1 \sim \pm \int_\delta^{T_1} e^{-\kappa s} dW_s, \\ Z_i &= \sqrt{\frac{1 - e^{-2\kappa\Delta T_i}}{2\kappa}} U_i \sim \pm \int_{T_{i-1}}^{T_i} e^{-\kappa s} dW_s, \quad 2 \leq i \leq d-1, \\ \varphi(y, z) &= \text{Nom} \times \sigma \sum_{i=2}^d e^{-rT_i} \Delta T_i e^{\kappa T_{i-1}} \left( y + \sum_{j=1}^{i-1} z_j \right), \quad y \in \mathbb{R}, z = (z_1, \dots, z_{d-1}) \in \mathbb{R}^{d-1}, \end{aligned}$$

with  $\{U_i, 0 \leq i \leq d-1\} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$ . On this basis,  $X_h$  can be simulated as

$$X_h = \frac{1}{K} \sum_{k=1}^K \varphi(Y, Z^{(k)}), \quad (7.7)$$

where  $h = \frac{1}{K} \in \mathcal{H}$  and  $\{Z^{(k)}, 1 \leq k \leq K\} \stackrel{\text{iid}}{\sim} Z$  are independent from  $Y$ .

## 7.2 Numerical Results

We aim at comparing numerically the four schemes of Sections 2–5 in terms of asymptotic error distribution. To this end, we fix a small prescribed accuracy  $\varepsilon > 0$ , set  $h = \varepsilon$  as in (6.1)-(6.3) for the nested and the averaged nested SA schemes and set  $h_L = \varepsilon$  as in (6.5)-(6.7) for the multilevel



and the averaged multilevel SA schemes. We run each scheme 5000 times and then we plot, for each algorithm separately, the empirical joint distribution of the normalized errors of the corresponding (VaR, ES) estimates. The error normalizations are calculated as they appear in Corollaries 2.1 and 3.1 and Theorems 4.1 and 5.1. The unbiased SA scheme of (Bercu et al., 2021, Theorem 3.4) and the unbiased averaged SA scheme of (Bardou et al., 2009a, Theorem 2.4), based on simulating the loss  $X_0$  exactly via (7.3), are also run for benchmarking purposes. A small difference between our unbiased averaged SA scheme and Bardou, Frikha, and Pagès (2009a, Theorem 2.4) is that we do not average out the ES component. Qualitatively however, the resulting outputs should be similar.

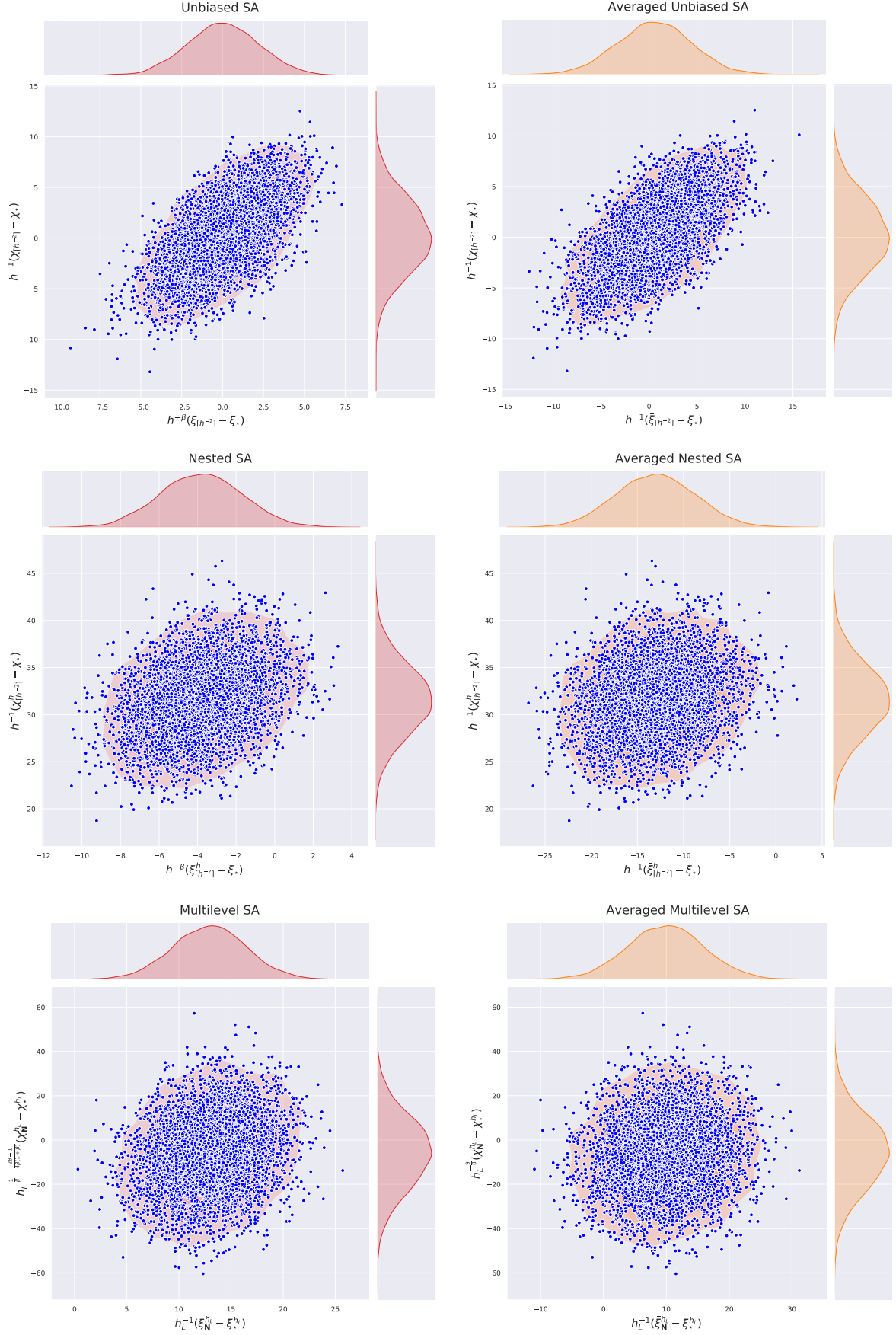
For the case study, we set  $S_0 = 1$ ,  $r = 2\%$ ,  $\kappa = 12\%$ ,  $\sigma = 20\%$ ,  $T = 1$  year,  $\Delta T_i = 3$  months,  $\delta = 7$  days and  $\alpha = 85\%$ . We use a 30/360 day count fraction convention. (7.4) and (7.5) yield  $\xi_\star \approx 2.19$  and  $\chi_\star \approx 3.29$ . The biased risk measures  $\xi_\star^{h_L}$  and  $\chi_\star^{h_L}$  needed for the CLTs of Theorems 4.1 and 5.1 are computed by running the nested SA scheme with the bias  $h_L$  and  $10^5$  iterations. We set  $\varepsilon = \frac{1}{256}$  and  $\beta = 0.9$  for all SA schemes. We adopt the hyperparameters  $\gamma_n = n^{-\beta}$  for the unbiased SA and unbiased averaged SA algorithms,  $\gamma_n = 0.1(250 + n)^{-\beta}$  for the NSA and ANSA algorithms, and  $h_0 = \frac{1}{32}$ ,  $M = 2$  and  $\gamma_n = 0.1(1500 + n)^{-\beta}$  for the MLSA and AMLSA algorithms.

Figure 1, page 41, plots kernel fittings of the joint distributions of the normalized errors for each algorithm. For a sufficiently small prescribed accuracy  $\varepsilon > 0$ , the joint probability distribution of the normalized (VaR, ES) error appears to be centrosymmetric unimodal for all algorithms. This is in line with the CLTs proven in the previous sections as well as with (Bercu et al., 2021, Theorem 3.4). Consistent with the expressions (3.2) and (5.3) for the covariance matrices of the corresponding limiting normalized errors, the directions of the axes of the ellipsoidal fitted densities suggest some asymptotic correlation between the VaR and ES errors for the ANSA algorithm and an asymptotic decorrelation for the AMLSA algorithm. The MLSA panel shows no asymptotic correlation, in accordance with the corresponding CLT. The NSA panel does show some correlation, but this is due to the choice  $\beta = 0.9$  that is too close to 1, for which correlation does exist. Similar comments are applicable to the unbiased SA and the averaged unbiased SA panels.

## 8 Conclusion

In Crépey, Frikha, and Louzi (2023), a nested stochastic approximation algorithm for VaR and ES, as well as a multilevel acceleration of the latter, were presented and compared in terms of non asymptotic  $L^2$  errors. The present article complements the latter by analyzing the corresponding asymptotic error distributions, as required for delimiting VaR and ES trust regions and confidence intervals. Further averaged extensions of these algorithms are also presented and shown to achieve better convergence rates than their original counterparts. A financial case study where exact VaR and ES values, as well as unbiased SA schemes, are available for benchmarking purposes, validates our theoretical findings.

The optimal complexity attained by the presented algorithms, for some prescribed accuracy  $\varepsilon > 0$ , is  $O(\varepsilon^{-\frac{5}{2}})$ . Given a step size  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $\beta \in (\frac{1}{2}, 1]$ , this complexity is achieved by the MLSA algorithm for  $\beta = 1$  under the constraint  $\lambda\gamma_1 > 1$ , where the constant  $\lambda > 0$  in (2.10) is explicit but tedious to compute. This complexity is also achieved by the AMLSA algorithm for  $\beta \in (\frac{8}{9}, 1)$  without any constraint on  $\gamma_1$ . However,  $O(\varepsilon^{-\frac{5}{2}})$  remains higher than the theoretical optimum of  $O(\varepsilon^{-2})$  for such multilevel algorithms (Frikha, 2016; Giles, 2008). The gap in performance stems from the discontinuity of the gradient used in the updating formula (2.8), causing an  $O(1)$  error when the generated loss  $X_h$  is too close to the estimate  $\xi_n^h$  but falls on the opposite side of the discontinuity with respect to the simulation target  $X_0$  (Giles, Haji-Ali, & Spence, 2023; Haji-Ali, Spence, & Teckentrup, 2022). This limitation should be addressed in future research.



**Figure 1:** Joint distributions of the normalized VaR and ES errors.

## A Auxiliary Result

**Lemma A.1.** Let  $\gamma_n = \gamma_1 n^{-\beta}$ , with  $\gamma_1 > 0$  and  $\beta \in (\frac{1}{2}, 1]$ .

- (i) Let  $\{e_n, n \geq 1\}$  be a nonnegative sequence. Then, for  $b \geq 0$  and  $\lambda > 0$ , with  $\gamma_1 > b/\lambda$  if  $\beta = 1$ ,

$$\limsup_{n \rightarrow \infty} \gamma_n^{-b} \sum_{k=1}^n \gamma_k^{1+b} e_k \exp \left( -\lambda \sum_{j=k+1}^n \gamma_j \right) \leq \frac{1}{C} \limsup_{n \rightarrow \infty} e_n,$$

where

$$C := \begin{cases} \lambda - b/\gamma_1 & \text{if } \beta = 1 \text{ and } \gamma_1 > b/\lambda, \\ \lambda & \text{if } \beta \in (\frac{1}{2}, 1). \end{cases}$$

- (ii) For  $b \geq 0$  and  $\lambda > 0$ , with  $\gamma_1 > b/\lambda$  if  $\beta = 1$ ,

$$\limsup_{n \rightarrow \infty} \gamma_n^{-b} \exp \left( -\lambda \sum_{j=1}^n \gamma_j \right) = 0.$$

Lemma A.1(i) is a special case of Fort (2015, Lemma 5.9), and Lemma A.1(ii) is a special case of Lemma A.1(i).

## B Square Convergence of the Averaged MLSA Scheme

We study the  $L^2$  error of the AMLSA algorithm (5.1) to fully apprehend its convergence and complexity.

### B.1 Convergence Rate Analysis

**Lemma B.1** (Crépey, Frikha, and Louzi (2023, Theorem 3.1)). Suppose that Assumptions 1 and 2 are satisfied, that  $\varphi(Y, Z)$  is in  $L^2(\mathbb{P})$ , that

$$\sup_{h \in \mathcal{H}} \mathbb{E} \left[ |\xi_0^h|^4 \exp \left( \frac{16}{1-\alpha} c_\alpha \sup_{h \in \mathcal{H}} \|f_{X_h}\|_\infty |\xi_0^h| \right) \right] < \infty,$$

where  $c_\alpha = 1 \vee \frac{\alpha}{1-\alpha}$ , and that  $\gamma_n = \gamma_1 n^{-\beta}$ , with  $\beta \in (0, 1]$  and  $\lambda_2 \gamma_1 > 2$  if  $\beta = 1$  where

$$\lambda_2 = \inf_{h \in \mathcal{H}} \frac{3}{4} V_h''(\xi_\star^h) \wedge \|V_h''\|_\infty \frac{V_h''(\xi_\star^h)^4}{[V_h''']_{\text{Lip}}^2}.$$

Then, for any positive integer  $n$ ,

$$\sup_{h \in \mathcal{H}} \mathbb{E}[(\xi_n^h - \xi_\star^h)^4] \leq C \gamma_n^2. \quad (\text{B.1})$$

**Theorem B.1.** Within the framework of Lemma B.1 and assuming that  $\gamma_n = \gamma_1 n^{-\beta}$ , with  $\beta \in (\frac{1}{2}, 1)$ , for any positive integer  $L$  and any  $\mathbf{N} = \{N_\ell, 0 \leq \ell \leq L\} \in \mathbb{N}_0^{L+1}$ , it holds

$$\begin{aligned} \mathbb{E}[(\bar{\xi}_{\mathbf{N}}^{\text{ML}} - \xi_\star^{h_L})^2] &\leq C \left( \gamma_{N_0} + \left( \sum_{\ell=1}^L \frac{1}{N_\ell \sqrt{\gamma_{N_\ell}}} \right)^2 \right. \\ &\quad \left. + \left( \sum_{\ell=1}^L h_\ell^{\frac{1+4\delta_0 \wedge 3}{4}} \sqrt{\gamma_{N_\ell}} \right)^2 + \left( \sum_{\ell=1}^L \gamma_{N_\ell} \right)^2 + \sum_{\ell=1}^L \frac{\sqrt{\gamma_{N_\ell}}}{N_\ell} + \sum_{\ell=1}^L \frac{h_\ell^{\frac{1}{2}}}{N_\ell} \right). \end{aligned} \quad (\text{B.2})$$

*Proof.* In the following developments, we denote by  $C$  a positive constant whose value may vary from line to line but does not depend upon  $L$ . We come back to the decomposition (5.8) and analyze each term separately.

**Step 1. Study of  $\{\bar{\xi}_{N_0}^{h_0} - \xi_\star^{h_0}, L \geq 1\}$ .**

By Lemma 2.2(ii) and a comparison between series and integrals, we have that

$$\mathbb{E}[(\bar{\xi}_{N_0}^{h_0} - \xi_\star^{h_0})^2] \leq \frac{1}{N_0} \sum_{k=1}^{N_0} \mathbb{E}[(\xi_k^{h_0} - \xi_\star^{h_0})^2] \leq \frac{C}{N_0} \sum_{k=1}^{N_0} \gamma_k \leq C\gamma_{N_0}.$$

**Step 2. Study of  $\left\{ \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_\ell} - a_k^{h_{\ell-1}}, L \geq 1 \right\}$ .**

We have that

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_\ell} - a_k^{h_{\ell-1}} \right)^2 \right]^{\frac{1}{2}} \leq \sum_{\ell=1}^L \mathbb{E} \left[ \left( \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_\ell} \right)^2 \right]^{\frac{1}{2}} + \mathbb{E} \left[ \left( \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_{\ell-1}} \right)^2 \right]^{\frac{1}{2}}.$$

It follows from (3.4), (3.6) and Lemma 2.2(ii) that

$$\begin{aligned} \mathbb{E} \left[ \left( \frac{1}{n} \sum_{k=1}^n a_k^h \right)^2 \right]^{\frac{1}{2}} &\leq \frac{1}{n} \left( \frac{1}{\gamma_n} \mathbb{E}[(\xi_n^h - \xi_\star^h)^2]^{\frac{1}{2}} + \frac{1}{\gamma_1} \mathbb{E}[(\xi_0^h - \xi_\star^h)^2]^{\frac{1}{2}} \right) \\ &\quad + \frac{1}{n} \sum_{k=1}^{n-1} \left( \frac{1}{\gamma_{k+1}} - \frac{1}{\gamma_k} \right) \mathbb{E}[(\xi_k^h - \xi_\star^h)^2]^{\frac{1}{2}} \\ &\leq C \left( \frac{1}{n} \left( \frac{1}{\sqrt{\gamma_n}} + 1 \right) + \frac{1}{n} \sum_{k=1}^{n-1} \left( \frac{1}{\gamma_{k+1}} - \frac{1}{\gamma_k} \right) \sqrt{\gamma_k} \right) \\ &\leq \frac{C}{n\sqrt{\gamma_n}}. \end{aligned}$$

Therefore

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} a_k^{h_\ell} - a_k^{h_{\ell-1}} \right)^2 \right] \leq C \left( \sum_{\ell=1}^L \frac{1}{N_\ell \sqrt{\gamma_{N_\ell}}} \right)^2.$$

**Step 3. Study of  $\left\{ \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}}, L \geq 1 \right\}$ .**

We have

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}} \right)^2 \right]^{\frac{1}{2}} \leq \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \mathbb{E}[(g_k^{h_\ell})^2]^{\frac{1}{2}} + \mathbb{E}[(g_k^{h_{\ell-1}})^2]^{\frac{1}{2}}.$$

It follows from Assumption 2, the uniform Lipschitz regularity of  $V_h$  combined with Lemma 2.2

(i) that  $|V_0''(\xi_\star) - V_{h_\ell}''(\xi_\star^{h_\ell})| \leq C(h_\ell^{(\frac{1}{4} + \delta_0) \wedge 1} + h_\ell)$ , which in turn using also (2.13) and Lemma 2.2(ii) give for  $\ell \geq 1$ ,

$$\mathbb{E}[(g_k^{h_\ell})^2]^{\frac{1}{2}} \leq |V_0''(\xi_\star) - V_{h_\ell}''(\xi_\star^{h_\ell})| \mathbb{E}[(\xi_{k-1}^{h_\ell} - \xi_\star^{h_\ell})^2]^{\frac{1}{2}} \leq C h_\ell^{(\frac{1}{4} + \delta_0) \wedge 1} \gamma_k^{\frac{1}{2}}.$$

By a comparison between series and integrals, we thus deduce

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} g_k^{h_\ell} - g_k^{h_{\ell-1}} \right)^2 \right] \leq C \left( \sum_{\ell=1}^L h_\ell^{\frac{1}{4} + (\delta_0 \wedge \frac{3}{4})} \gamma_{N_\ell}^{\frac{1}{2}} \right)^2.$$

**Step 4. Study of**  $\left\{ \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} r_k^{h_\ell} - r_k^{h_{\ell-1}}, L \geq 1 \right\}$ .

Note that

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} r_k^{h_\ell} - r_k^{h_{\ell-1}} \right)^2 \right]^{\frac{1}{2}} \leq \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \mathbb{E}[(r_k^{h_\ell})^2]^{\frac{1}{2}} + \mathbb{E}[(r_k^{h_{\ell-1}})^2]^{\frac{1}{2}}.$$

By (2.14), (2.26), the uniform Lipschitz regularity of  $V_h''$  (stemming from Assumption 1(iv)) and eventually Lemma B.1, we get

$$\mathbb{E}[|r_k^h|^2] \leq C \mathbb{E}[(\xi_{k-1}^h - \xi_\star^h)^4] \leq C \gamma_k^2.$$

Combining the two previous inequalities and using a comparison between series and integrals, we conclude

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} r_k^{h_\ell} - r_k^{h_{\ell-1}} \right)^2 \right] \leq C \left( \sum_{\ell=1}^L \gamma_{N_\ell} \right)^2.$$

**Step 5. Study of**  $\left\{ \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}, L \geq 1 \right\}$ .

We recall that the random variables  $\left\{ \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} (\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}), \ell \geq 1 \right\}$  are independent and with zero mean and that at each level  $\ell \geq 1$ , the sequence  $\{\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}, k \geq 1\}$  is an  $\{\mathcal{F}_k^{h_\ell}, k \geq 1\}$ -martingale increment sequence. Using (3.9) and a comparison between series and integrals, we obtain

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \rho_k^{h_\ell} - \rho_k^{h_{\ell-1}} \right)^2 \right] \leq 2 \sum_{\ell=1}^L \frac{1}{N_\ell^2} \sum_{k=1}^{N_\ell} \mathbb{E}[|\rho_k^{h_\ell}|^2] + \mathbb{E}[|\rho_k^{h_{\ell-1}}|^2] \leq C \sum_{\ell=1}^L \frac{\sqrt{\gamma_{N_\ell}}}{N_\ell}.$$

**Step 6. Study of**  $\left\{ \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} e_k^{h_\ell} - e_k^{h_{\ell-1}}, L \geq 1 \right\}$ .

Again, the random variables  $\left\{ \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}), \ell \geq 1 \right\}$  are independent and centered and, for any fixed  $\ell \geq 1$ ,  $\{e_k^{h_\ell} - e_k^{h_{\ell-1}}, k \geq 1\}$  is a sequence of  $\{\mathcal{F}_k^{h_\ell}, k \geq 1\}$ -martingale increments. From (4.22), we thus get

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} e_k^{h_\ell} - e_k^{h_{\ell-1}} \right)^2 \right] = \sum_{\ell=1}^L \frac{1}{N_\ell^2} \sum_{k=1}^{N_\ell} \mathbb{E}[(e_k^{h_\ell} - e_k^{h_{\ell-1}})^2] \leq C \sum_{\ell=1}^L \frac{h_\ell^{\frac{1}{2}}}{N_\ell}.$$

□

## B.2 Complexity Analysis

The global error of estimating  $\xi_\star$  by  $\bar{\xi}_\mathbf{N}^{\text{ML}}$  can be decomposed as the sum of a statistical and a bias error

$$\bar{\xi}_\mathbf{N}^{\text{ML}} - \xi_\star = (\bar{\xi}_\mathbf{N}^{\text{ML}} - \xi_\star^{h_L}) + (\xi_\star^{h_L} - \xi_\star).$$

Let  $\varepsilon^2$  be a fixed prescribed  $L^2$  error,  $\varepsilon \in (0, 1)$ . Lemma 2.2(i) guarantees that the bias of estimating  $\xi_\star$  by  $\bar{\xi}_\mathbf{N}^{\text{ML}}$  to be of order  $h_L$ , so that we have to choose the number of levels

$$L = \left\lceil \frac{\ln h_0 \varepsilon^{-1}}{\ln M} \right\rceil \quad \text{so that} \quad h_L = \frac{h}{M^L} \leq \varepsilon. \quad (\text{B.3})$$

The cost of the AMLSA scheme is

$$\text{Cost}_{\text{AMLSA}} = C \sum_{\ell=0}^L \frac{N_\ell}{h_\ell},$$

for some fixed constant  $C$  independent of  $L$ . In order to optimize the number of iterations  $\mathbf{N} = \{N_\ell, 0 \leq \ell \leq L\}$ , we try out several candidates as a leading term in the upper bound (B.2). We retain three particularly compelling cases. As usual, we minimize the above computational cost under the constraint that the aforementioned candidate is of order  $\varepsilon^2$ .

We first consider the optimization problem

$$\begin{aligned} & \text{minimize}_{N_0, \dots, N_L > 0} && \sum_{\ell=0}^L N_\ell h_\ell^{-1} \\ & \text{subject to} && \sum_{\ell=0}^L h_\ell^{(1+4\delta_0 \wedge 3)/4} \sqrt{\gamma_{N_\ell}} = C^{-1} \varepsilon, \end{aligned}$$

yielding

$$N_\ell = \left[ C^{\frac{2}{\beta}} \gamma_1^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{2\beta-1-4\delta_0 \wedge 3}{2(2+\beta)}} \right)^{\frac{2}{\beta}} h_\ell^{\frac{5+4\delta_0 \wedge 3}{2(2+\beta)}} \right], \quad 0 \leq \ell \leq L.$$

Such a choice adds the constraints

$$\frac{1}{2} < \beta \leq \frac{2}{3} \quad \text{and} \quad \delta_0 \geq \frac{3}{4}$$

to guarantee that the upper estimate (B.2) is of order  $\varepsilon^2$ . The corresponding complexity is

$$\text{Cost}_{\text{AMLSA}}^\beta = C \varepsilon^{-\frac{2}{\beta}} \xrightarrow{\beta \uparrow \frac{2}{3}} \varepsilon^{-3}.$$

We now consider the optimization problem

$$\begin{aligned} & \text{minimize}_{N_0, \dots, N_L > 0} && \sum_{\ell=0}^L N_\ell h_\ell^{-1} \\ & \text{subject to} && \sum_{\ell=0}^L N_\ell^{-1} \gamma_{N_\ell}^{-\frac{1}{2}} = C^{-1} \varepsilon. \end{aligned}$$

The minimizers are given by

$$N_\ell = \left[ C^{\frac{2}{2-\beta}} \gamma_1^{-\frac{1}{2-\beta}} \varepsilon^{-\frac{2}{2-\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{2-\beta}{4-\beta}} \right)^{\frac{2}{2-\beta}} h_\ell^{\frac{2}{4-\beta}} \right], \quad 0 \leq \ell \leq L.$$

If the additional conditions

$$\frac{2}{3} \leq \beta < 1 \quad \text{and} \quad \delta_0 > \frac{1}{12},$$

are met, one attains the order  $\varepsilon^2$  for the global  $L^2$  error (B.2) with an optimal cost satisfying

$$\text{Cost}_{\text{AMLSA}}^\beta = C \varepsilon^{-\frac{2}{2-\beta}-1} \xrightarrow{\beta \downarrow \frac{2}{3}} \varepsilon^{-\frac{5}{2}}.$$

The final case worth exploring consists in solving the optimization problem

$$\begin{aligned} & \text{minimize}_{N_0, \dots, N_L > 0} && \sum_{\ell=0}^L N_\ell h_\ell^{-1} \\ & \text{subject to} && \sum_{\ell=0}^L h_\ell^{\frac{1}{2}} N_\ell^{-1} = C^{-1} \varepsilon^2. \end{aligned}$$

This yields, under constraints stated in Theorem 5.1,

$$N_\ell = \left[ C^2 \varepsilon^{-2} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{1}{4}} \right) h_\ell^{\frac{3}{4}} \right], \quad 0 \leq \ell \leq L.$$

The convergence analysis for this last heuristic is provided by Theorem 5.1, and the ensuing complexity is discussed in Section 6.

## Bibliography

- Bardou, O., Frikha, N., & Pagès, G. (2009a). Computing VaR and CVaR using stochastic approximation and adaptive unconstrained importance sampling. *Monte Carlo Methods and Applications*, 15(3), 173–210.
- Bardou, O., Frikha, N., & Pagès, G. (2009b). Recursive computation of value-at-risk and conditional value-at-risk using MC and QMC. In P. L’Ecuyer & A. B. Owen (Eds.), *Monte Carlo and Quasi-Monte Carlo Methods 2008* (pp. 193–208). Springer.
- Bardou, O., Frikha, N., & Pagès, G. (2016). CVaR hedging using quantization-based stochastic approximation algorithm. *Mathematical Finance*, 26(1), 184–229.
- Barrera, D., Crépey, S., Diallo, B., Fort, G., Gobet, E., & Staszynski, U. (2019). Stochastic approximation schemes for economic capital and risk margin computations. *ESAIM: Proceedings and Surveys*, 65, 182–218.
- Basel Committee on Banking Supervision. (2013). *Consultative document: Fundamental Review of the Trading Book: A revised market risk framework*.
- Bercu, B., Costa, M., & Gadat, S. (2021). Stochastic approximation algorithms for superquantiles estimation. *Electronic Journal of Probability*, 26(none), 1–29.
- Crépey, S., Frikha, N., & Louzi, A. (2023). A multilevel stochastic approximation algorithm for value-at-risk and expected shortfall estimation. *arXiv:2304.01207*.
- Dereich, S., & Müller-Gronbach, T. (2019). General multilevel adaptations for stochastic approximation algorithms of Robbins–Monro and Polyak–Ruppert type. *Numerische Mathematik*, 142(2), 279–328.
- European Parliament and the Council. (2009). *Directive of the European Parliament and of the Council on the taking-up and pursuit of the business of insurance and reinsurance (Solvency II)*.
- Fort, G. (2015). Central limit theorems for stochastic approximation with controlled Markov chain dynamics. *ESAIM: PS*, 19, 60–80.
- Frikha, N. (2014). Shortfall risk minimization in discrete time financial market models. *SIAM Journal on Financial Mathematics*, 5(1), 384–414.
- Frikha, N. (2016). Multi-level stochastic approximation algorithms. *The Annals of Applied Probability*, 26(2), 933–985.
- Giles, M. B. (2008). Multilevel Monte Carlo path simulation. *Operations Research*, 56(3), 607–617.
- Giles, M. B., & Haji-Ali, A.-L. (2019). Multilevel nested simulation for efficient risk estimation. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2), 497–525.
- Giles, M. B., Haji-Ali, A.-L., & Spence, J. (2023). Efficient risk estimation for the credit valuation adjustment.
- Giorgi, D., Lemaire, V., & Pagès, G. (2020). Weak error for nested multilevel Monte Carlo. *Methodology and Computing in Applied Probability*, 22, 1325–1348.
- Haji-Ali, A.-L., Spence, J., & Teckentrup, A. L. (2022). Adaptive multilevel Monte Carlo for probabilities. *SIAM Journal on Numerical Analysis*, 60(4), 2125–2149.
- Hall, P., & Heyde, C. C. (1980). *Martingale limit theory and its application*. Academic Press New York.
- Konda, V. R., & Tsitsiklis, J. N. (2004). Convergence rate of linear two-time-scale stochastic approximation. *The Annals of Applied Probability*, 14(2), 796–819.
- Mokkadem, A., & Pelletier, M. (2006). Convergence rate and averaging of nonlinear two-time-scale stochastic approximation algorithms. *The Annals of Applied Probability*, 16(3), 1671–1702.
- Polyak, B. T., & Juditsky, A. B. (1992). Acceleration of stochastic approximation by averaging. *SIAM Journal on Control and Optimization*, 30(4), 838–855.
- Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21–41.

- Ruppert, D. (1991). Handbook of sequential analysis. *Stochastic Approximation*, 503–529.
- Swiss Federal Office of Private Insurance. (2004, September 17). *The risk margin for the Swiss solvency test*.