



**HAL**  
open science

# Adaptive Multilevel Stochastic Approximation of the Value-at-Risk

Stéphane Crépey, Noufel Frikha, Azar Louzi, Jonathan Spence

► **To cite this version:**

Stéphane Crépey, Noufel Frikha, Azar Louzi, Jonathan Spence. Adaptive Multilevel Stochastic Approximation of the Value-at-Risk. 2026. <hal-04670735v2>

**HAL Id: hal-04670735**

**<https://hal.science/hal-04670735v2>**

Preprint submitted on 12 Apr 2026

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.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

# Adaptive Multilevel Stochastic Approximation of the Value-at-Risk

Stéphane Crépey\*    Noufel Frikha<sup>†</sup>    Azar Louzi<sup>‡</sup>    Jonathan Spence<sup>§</sup>

April 12, 2026

**Abstract.** Crépey, Frikha, and Louzi (2025) introduced a multilevel stochastic approximation scheme to compute the value-at-risk of a financial loss that is only simulatable by Monte Carlo. The best complexity of the scheme is in  $O(\varepsilon^{-\frac{5}{2}})$ ,  $\varepsilon > 0$  being a prescribed accuracy, which is suboptimal compared to the canonical multilevel Monte Carlo performance. This suboptimality stems from the discontinuity of the Heaviside function involved in the biased stochastic gradient that is recursively evaluated to derive the value-at-risk. To mitigate this issue, this paper proposes and analyzes a multilevel stochastic approximation algorithm that adaptively selects the number of inner samples at each level, and proves that its best complexity is in  $O(\varepsilon^{-2} |\ln \varepsilon|^{\frac{5}{2}})$ . Our theoretical analysis is exemplified through numerical experiments.

**Keywords.** stochastic approximation ; value-at-risk ; nested Monte Carlo ; multilevel Monte Carlo ; adaptive Monte Carlo

**MSC.** Primary 65C05, 62L20; secondary 62G32, 91G60

**arXiv.** 2408.06531    **HAL.** 04670735

Accepted for publication in Annals of Applied Probability

## 1 Introduction

The value-at-risk (VaR) is the predominantly used risk metric in finance [8]. It stands for the quantile, at some confidence level, of the loss of a portfolio. The probability that such loss exceeds the VaR expresses the chances of occurrence of large losses on the portfolio. Evaluating a portfolio's VaR is thus paramount to assessing its risk. Note that a large class of financial portfolios can only be valued by Monte Carlo (MC), which introduces a bias when estimating the loss, hence increasing the VaR computation's complexity. Determining a portfolio's VaR is thus considered a challenging task.

Following the celebrated works on VaR convexification (Rockafellar and Uryasev [37], Bertal and Teboulle [10] and Bardou et al. [4]), a nascent line of research has adopted a stochastic approximation (SA) viewpoint to estimate the VaR (Bardou et al. [4, 5, 6], Costa and Gadat [13] and Bercu et al. [11]). Recent developments (Barrera et al. [7] and Crépey et al. [14, 15]) have focused on the nested nature of the problem. Notably, Barrera et al. [7] propose a nested SA (NSA) method that couples an outer SA scheme with an inner nested MC scheme to compute the VaR. It is shown by Crépey et al. [14] that such a scheme runs in  $O(\varepsilon^{-3})$  time,  $\varepsilon > 0$  being

---

\*Université Paris Cité, CNRS, Laboratoire de Probabilités, Statistique et Modélisation (LPSM, UMR 8001), 75013 Paris. [stephane.crepey@lpsm.paris](mailto:stephane.crepey@lpsm.paris)

<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)

<sup>‡</sup>Université Paris Cité, CNRS, Laboratoire de Probabilités, Statistique et Modélisation (LPSM, UMR 8001), 75013 Paris. [azar.louzi@lpsm.paris](mailto:azar.louzi@lpsm.paris)

<sup>§</sup>Maxwell Institute for Mathematical Sciences, School of Mathematics, University of Edinburgh, Edinburgh, UK, EH9 3FD. [jspence5@ed.ac.uk](mailto:jspence5@ed.ac.uk)

a prescribed accuracy. Meanwhile, subsequently to the wide success of multilevel MC (MLMC) methods (Heinrich [32, 33], Giles [24] and Dereich [16]), multilevel SA (MLSA) algorithms (Frikha [21] and Dereich [17]) have emerged as a means to accelerate biased SA schemes such as NSA. While the latter injects biased nested simulations directly into an iterative SA scheme, MLSA telescopes a sequence of correlated estimate pairs of lower and lower biases, hence reducing the complexity by an order of magnitude. Crépey et al. [14] leverage the multilevel acceleration proposed by Frikha [21] to reduce the complexity of the VaR NSA scheme of Barrera et al. [7] to  $O(\varepsilon^{-2-\delta})$ ,  $\frac{1}{2} \leq \delta < 1$  being a small number governed by the integrability degree of the loss. They also identify a parametrization for their MLSA algorithm that achieves a complexity for the estimation of the expected shortfall (ES), a risk measure closely linked to the VaR, in  $O(\varepsilon^{-2} |\ln \varepsilon|^2)$ . Crépey et al. [15] further obtain central limit theorems (CLTs) for VaR and ES NSA and MLSA schemes, as well as their Polyak-Ruppert variants that enjoy stronger numerical stability properties.

Albeit significant, the performance gain achieved by MLSA over NSA in estimating the VaR [14] is to be nuanced. According to Giles [24] and Frikha [21], the canonical optimal complexity for multilevel techniques is of order  $\varepsilon^{-2}$ . The suboptimality of the VaR MLSA scheme arises from its inner recursions relying on the evaluation of a discontinuous gradient akin to a Heaviside function ( $x \mapsto \mathbb{1}_{x>0}$ ). When a biased simulation of the loss falls on the opposite side of the discontinuity relative to the exact target loss, the multilevel recursion incurs an  $O(1)$  update error. The accumulation of these errors propagates to the final estimator, leading to the suboptimality.

Motivated by valuing a digital option with Heaviside payoff in a local volatility model, Giles [25] surveys multiple approaches to the above discontinuity issue in a multilevel Monte Carlo (MLMC) context. Most techniques described therein emanate from the derivative sensitivity computation literature and their common goal is to lower the variance of the paired estimators at each level of the multilevel simulation. We retain three ideas that may apply to our nested SA setting. Firstly, a natural idea is to smoothen the Heaviside payoff, but this only attains a complexity in  $O(\varepsilon^{-2-\frac{1}{4}})$  due to a large smoothing Lipschitz constant that is involved in. Secondly, Malliavin calculus could be used in conjunction with a cubic spline interpolation, which however propagates regression noises to the final estimator. Finally, levelwise adaptive refinement of inner Monte Carlo samplings seems to be a versatile technique that achieves the desired performance gain. It advocates to dynamically refine the biased simulations in order to deal with the discontinuity issue (Giles and Haji-Ali [26], Haji-Ali et al. [31] and Giles et al. [27]). Let us provide a brief overview of this technique.

An early example of adaptive nested Monte Carlo simulation can be traced back to Broadie et al. [12], who endeavor to efficiently evaluate the probability of occurrence of large losses within the nested Monte Carlo setting of Gordy and Juneja [30]. The latter had proposed an algorithm for estimating the VaR, which computes an order statistic out of multiple loss scenarios, each one being obtained by averaging out multiple inner MC samples. The authors in [12] are interested in estimating the probability that such nested loss, simulated with the help of inner MC samples, exceeds the VaR. This estimation involves evaluating a Heaviside function that is centered on the VaR itself. The adaptive strategy developed in [12] consists in refining simulations of risk scenarios that fall too close to the VaR to reduce their chances of ending up on the wrong side thereof, by increasing the amount of inner simulations involved. In so doing, evaluating the Heaviside function at the nested loss becomes almost equivalent to evaluating it at the true loss. Giles and Haji-Ali [26] and Haji-Ali et al. [31] revisit the adaptive refinement concept of [12] in the context of nested MLMC, whereby the adaptive refinement is used to increase the number of inner simulations at each level according to how close their functional falls with respect to the Heaviside discontinuity.

We note that the work of Elfverson et al. [19] has led to similar adaptive ideas in a partial differential equation Monte Carlo approximation setting, where the Monte Carlo estimation er-

ror is assumed to be almost surely bounded. The terminology “*adaptive MLMC*” has also been employed in other meanings, such as adaptive path simulation (Hoel et al. [34]) and importance sampling (Ben Alaya et al. [9]).

In this article, we develop an adaptive refinement strategy to efficiently estimate the VaR using the MLSA algorithm derived in [14]. As in MLMC, our adaptive strategy prioritizes aligning the samples with their targets relative to the discontinuities of the update functions, which are centred at the SA iterates. Note that the strategy that was developed for MLMC in [26, 31] does not directly apply to our SA setting, since refining the samples that drive an SA scheme shifts their distributions relatively to the SA iterates, who themselves evolve from one iteration to the next. This is problematic since the shift in sample distribution affects the original SA problem formulation, rendering it possibly non-convex and possessing multiple stationary points. To address this issue, we carefully incorporate a saturation factor that varies along the iterations in order to mitigate the undesirable effects of the change in sample distribution.

We apply our strategy to both NSA and MLSA paradigms and provide sharp  $L^2(\mathbb{P})$ -controls. For a prescribed accuracy  $\varepsilon > 0$ , their optimal complexities are shown to be in  $O(\varepsilon^{-\frac{5}{2}} |\ln \varepsilon|^{\frac{1}{2}})$  and in  $O(\varepsilon^{-2} |\ln \varepsilon|^{\frac{5}{2}})$  respectively, resulting in an order of magnitude  $\varepsilon^{\frac{1}{2}}$  speed-up on their non-adaptive counterparts, up to a logarithmic factor. The adaptive MLSA algorithm largely attains a complexity order that is comparable to the standard unbiased Robbins-Monro algorithm, thus demonstrating the narrowing of the performance gap between nested MLSA schemes and Robbins-Monro schemes in the context of Heaviside-type update functions. Our numerical analyses, conducted on a toy model as well as a more concrete financial setup, strongly support our theoretical findings.

The paper is structured as follows. Section 2 recalls the problem and main results in [14] regarding the estimation of the VaR using NSA and MLSA. Section 3 develops an adaptive refinement strategy to enhance the efficiency of the Monte Carlo sampling that is nested within the MLSA approach. Sections 4 and 5 exploit this strategy to reduce the complexities of NSA and MLSA and provide subsequent  $L^2(\mathbb{P})$ -controls and complexity rates. Section 7 presents numerical studies to demonstrate the performance improvement resulting from the adaptive strategy.

## 2 Stochastic Approximation Approach

This section recalls the setting and main results in [14] on the stochastic approximation of the VaR. They will serve as a baseline for the adaptive strategy that we develop in Section 3.

### 2.1 Unbiased Stochastic Approximation Algorithm

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space accommodating all of the subsequent random variables. Let  $X^0 \in L^1(\mathbb{P})$  be an  $\mathbb{R}$ -valued random loss of a portfolio, defined at some time horizon  $\tau > 0$ . Following Acerbi and Tasche [2] and Föllmer and Schied [23], the VaR of  $X^0$  at some confidence level  $\alpha \in (0, 1)$ , denoted  $\xi_\star^0$ , is defined as

$$\xi_\star^0 := \inf \{ \xi \in \mathbb{R} : \mathbb{P}(X^0 \leq \xi) \geq \alpha \}. \quad (2.1)$$

As stated in [37, 10, 4], if the cdf  $F_{X^0}$  of  $X^0$  is continuous, then  $\xi_\star^0$  is the left-end solution to

$$\min_{\xi \in \mathbb{R}} V_0(\xi), \quad \text{where} \quad V_0(\xi) := \xi + \frac{1}{1 - \alpha} \mathbb{E}[(X^0 - \xi)^+], \quad \xi \in \mathbb{R}. \quad (2.2)$$

Moreover,  $V_0$  is convex and continuously differentiable on  $\mathbb{R}$ , with

$$V_0'(\xi) = \frac{F_{X^0}(\xi) - \alpha}{1 - \alpha} = \mathbb{E}[H(\xi, X^0)], \quad \xi \in \mathbb{R},$$

where

$$H(\xi, x) = 1 - \frac{1}{1 - \alpha} \mathbf{1}_{x \geq \xi}, \quad \xi, x \in \mathbb{R}. \quad (2.3)$$

If  $F_{X^0}$  is additionally increasing, then  $V_0$  is strictly convex and  $\xi_\star^0$  is the unique minimizer of  $V_0$ :

$$\xi_\star^0 = \operatorname{argmin} V_0. \quad (2.4)$$

Besides, if  $X^0$  admits a continuous pdf  $f_{X^0}$ , then  $V_0$  is twice continuously differentiable on  $\mathbb{R}$ , with

$$V_0''(\xi) = \frac{1}{1 - \alpha} f_{X^0}(\xi), \quad \xi \in \mathbb{R}.$$

When i.i.d. samples of  $X^0$  are available, [4] proposes to estimate  $\xi_\star^0$  using the unbiased SA scheme of dynamics

$$\xi_{n+1}^0 = \xi_n^0 - \gamma_{n+1} H(\xi_n^0, X_{n+1}^0), \quad n \in \mathbb{N}, \quad (2.5)$$

where  $(X_n^0)_{n \geq 1} \stackrel{\text{i.i.d.}}{\sim} X^0$ ,  $\xi_0^0$  is a real-valued random initialization independent of the samples  $(X_n^0)_{n \geq 1}$ , and  $(\gamma_n)_{n \geq 1}$  is a positive non-increasing sequence such that  $\sum_{n=1}^{\infty} \gamma_n = \infty$  and  $\lim_{n \rightarrow \infty} \gamma_n = 0$ .

We are interested in the setting where  $X^0$  can be written as a conditional expectation:

$$X^0 = \mathbb{E}[\varphi(Y, Z) | Y]. \quad (2.6)$$

Here,  $Y$  and  $Z$  are two independent random variables taking values in  $\mathbb{R}^d$  and  $\mathbb{R}^q$  respectively, and  $\varphi : \mathbb{R}^d \times \mathbb{R}^q \rightarrow \mathbb{R}$  is a measurable function such that  $\varphi(Y, Z) \in L^1(\mathbb{P})$ . From a financial standpoint,  $Y$  typically models the dynamics of the portfolio's risk factors up to the time horizon  $\tau$  of the loss simulation,  $Z$  their dynamics beyond  $\tau$  and  $\varphi(Y, Z)$  the subsequent future cash flows.

*Remark 2.1.* Generally, the risk factors  $Y$  and  $Z$  can be sampled from a model at a fixed cost  $c_0 > 0$ , and the cash flow structure  $\varphi$  is known. If  $Y$  and  $Z$  are simulated by discretizing a stochastic differential equation (SDE), one must take into account additional terms in our error controls, as well as extra factors in our complexity formulas. We refer for instance to Rhee and Glynn [35] for a multilevel method to simulate the functional  $\varphi(Y, Z)$  based on an SDE.

Hereafter, we are interested in the case where  $X^0$  cannot be sampled exactly. One must then leverage the nested formulation of  $X^0$  in (2.6) to produce suitable samples for the SA approach.

## 2.2 Nested Stochastic Approximation Algorithm

Let  $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$ . A natural idea is to approximate  $X^0$  by the nested Monte Carlo estimator

$$X^h := \frac{1}{K} \sum_{k=1}^K \varphi(Y, Z^k), \quad \text{where } h := \frac{1}{K} \in \mathcal{H} = \left\{ \frac{1}{K'} : K' \in \mathbb{N}^* \right\} \quad (2.7)$$

and  $Y$  and  $(Z^k)_{1 \leq k \leq K} \stackrel{\text{i.i.d.}}{\sim} Z$  are independent.  $h$  is termed *bias parameter*, since it helps control the bias associated with the VaR estimation. For some  $h \in \mathcal{H}$ , we swap  $X^0$  by  $X^h$  in the original problem (2.2) and obtain the approximate problem

$$\min_{\xi \in \mathbb{R}} V_h(\xi), \quad \text{where } V_h(\xi) := \xi + \frac{1}{1 - \alpha} \mathbb{E}[(X^h - \xi)^+], \quad \xi \in \mathbb{R}. \quad (2.8)$$

Again, assuming that  $\varphi(Y, Z) \in L^1(\mathbb{P})$ , if the cdf  $F_{X^h}$  of  $X^h$  is continuous, then  $V_h$  is convex and continuously differentiable on  $\mathbb{R}$ , with

$$V_h'(\xi) = \frac{F_{X^h}(\xi) - \alpha}{1 - \alpha} = \mathbb{E}[H(\xi, X^h)], \quad \xi \in \mathbb{R},$$

recalling the definition (2.3). If  $F_{X^h}$  is additionally increasing, then  $V_h$  is strictly convex and

$$\xi_\star^h = \operatorname{argmin} V_h \quad (2.9)$$

is well defined and constitutes a biased estimator of  $\xi_\star^0$ . Finally, if  $X^h$  admits a continuous pdf  $f_{X^h}$ , then  $V_h$  is twice continuously differentiable on  $\mathbb{R}$ , with

$$V_h''(\xi) = \frac{1}{1-\alpha} f_{X^h}(\xi), \quad \xi \in \mathbb{R}.$$

Assuming  $F_{X^h}$  continuous and increasing,  $\xi_\star^0$  can be estimated with bias  $h \in \mathcal{H}$  using the NSA scheme of dynamics

$$\xi_{n+1}^h = \xi_n^h - \gamma_{n+1} H(\xi_n^h, X_{n+1}^h), \quad n \in \mathbb{N}, \quad (2.10)$$

where  $(X_n^h)_{n \geq 1} \stackrel{\text{i.i.d.}}{\sim} X^h$ ,  $\xi_0^h$  is a real-valued random initialization independent of the samples  $(X_n^h)_{n \geq 1}$ , and  $(\gamma_n)_{n \geq 1}$  is a positive non-increasing sequence such that  $\sum_{n=1}^{\infty} \gamma_n = \infty$  and  $\lim_{n \rightarrow \infty} \gamma_n = 0$ .

### 2.2.1 Convergence Analysis

Let  $\overline{\mathcal{H}} := \mathcal{H} \cup \{0\}$ . The global error of NSA can be written as a sum of statistical and bias errors as such:

$$\xi_n^h - \xi_\star^0 = (\xi_n^h - \xi_\star^h) + (\xi_\star^h - \xi_\star^0). \quad (2.11)$$

*Remark 2.2.* In the MC realm, a statistical error has zero expectation, which is not the case for the quantity  $\xi_n^h - \xi_\star^h$ . Our terminology here is borrowed from the SA literature (e.g. Dufflo [18] or Frikha [21]). Nevertheless, as recalled later in Lemma 2.5(ii), we do have  $\xi_n^h - \xi_\star^h \rightarrow 0$  in  $L^2(\mathbb{P})$  as  $n \uparrow \infty$ , exactly as in the MC framework.

**Assumption 2.3** ([14, Assumptions 3.2 & 3.5]).

(i) For all  $h \in \mathcal{H}$ ,  $F_{X^h}$  admits the first order Taylor expansion

$$F_{X^h}(\xi) - F_{X^0}(\xi) = v(\xi)h + w(\xi, h)h, \quad \xi \in \mathbb{R},$$

for some functions  $v, w(\cdot, h) : \mathbb{R} \rightarrow \mathbb{R}$  satisfying, for all  $\xi_\star^0 \in \operatorname{Argmin} V_0$ ,

$$\int_{\xi_\star^0}^{\infty} v(\xi) d\xi < \infty, \quad \lim_{\mathcal{H} \ni h \downarrow 0} w(\xi_\star^0, h) = \lim_{\mathcal{H} \ni h \downarrow 0} \int_{\xi_\star^0}^{\infty} w(\xi, h) d\xi = 0. \quad (2.12)$$

(ii) For all  $h \in \overline{\mathcal{H}}$ , the law of  $X^h$  admits a continuous pdf  $f_{X^h}$  with respect to the Lebesgue measure. Moreover, the pdf  $(f_{X^h})_{h \in \mathcal{H}}$  converge locally uniformly to  $f_{X^0}$ .

(iii) For all  $R > 0$ ,

$$\inf_{h \in \overline{\mathcal{H}}, \xi \in B(\xi_\star^0, R)} f_{X^h}(\xi) > 0.$$

(iv) The pdf  $(f_{X^h})_{h \in \overline{\mathcal{H}}}$  are uniformly bounded and uniformly Lipschitz, namely,

$$\sup_{h \in \overline{\mathcal{H}}} (\|f_{X^h}\|_\infty + [f_{X^h}]_{\text{Lip}}) < \infty,$$

where  $[f_{X^h}]_{\text{Lip}}$  denotes the Lipschitz constant of  $f_{X^h}$ ,  $h \in \overline{\mathcal{H}}$ .

*Remark 2.4* ([15, Remark 1.2]). Giorgi et al.'s [29, Propositions 5.1(a,b)] provide guarantees on the postulates of Assumptions 2.3(i)-(ii). The second part of (2.12) also reads  $w(\xi_\star^0, h) = \int_{\xi_\star^0}^\infty w(\xi, h)d\xi = o(1)$  as  $\mathcal{H} \ni h \downarrow 0$ ,  $\xi_\star^0 \in \text{Argmin } V_0$ . Assumptions 2.3(iii)-(iv) are natural in view of Assumption 2.3(ii) and the increasing nature of  $F_{X^0}$ . By Assumptions 2.3(ii)-(iii), for all  $h \in \overline{\mathcal{H}}$  and all  $\xi_\star^h \in \text{Argmin } V_h$ ,  $V_h''(\xi_\star^h) = (1 - \alpha)^{-1} f_{X^h}(\xi_\star^h) > 0$ , subsequently reducing  $\text{Argmin } V_h$  to a singleton  $\{\xi_\star^h\}$ . Assumptions 2.3(iii) may be easily relaxed to supposing  $f_{X^0}(\xi_\star^0) > 0$ , but the proposed formulation facilitates the derivation of the Lyapunov properties reprised in Lemma A.1, which are fundamental to the statistical error controls of Lemma 2.5(ii) below.

We refer to Lemma A.1 for important consequences of Assumption 2.3. For  $q \in \mathbb{N}^*$ , we let

$$\bar{\lambda}_q := \inf_{h \in \overline{\mathcal{H}}} \bar{\lambda}_{h,q} > 0, \quad (2.13)$$

where the  $(\bar{\lambda}_{h,q})_{h \in \overline{\mathcal{H}}, q \geq 1}$  are defined in Lemma A.1(iii). Note that  $(\bar{\lambda}_q)_{q \geq 1}$  are non-decreasing with respect to  $q \in \mathbb{N}^*$ .

**Lemma 2.5** ([14, Proposition 3.4 & Theorem 3.7(i)]). *(i) Suppose that  $\varphi(Y, Z) \in L^1(\mathbb{P})$ , that Assumptions 2.3(i)-(iii) hold, and that  $f_{X^0} > 0$ . Then, as  $\mathcal{H} \ni h \downarrow 0$ ,*

$$\xi_\star^h - \xi_\star^0 = -\frac{v(\xi_\star^0)}{f_{X^0}(\xi_\star^0)} h + o(h).$$

Thus, for all  $h \in \mathcal{H}$ ,

$$|\xi_\star^h - \xi_\star^0| \leq Ch, \quad \text{where } C = \left| \frac{v(\xi_\star^0)}{f_{X^0}(\xi_\star^0)} \right| + \sup_{h \in \mathcal{H}} \left\{ \left| \frac{\xi_\star^h - \xi_\star^0}{h} + \frac{v(\xi_\star^0)}{f_{X^0}(\xi_\star^0)} \right| \right\}. \quad (2.14)$$

*(ii) Suppose that  $\varphi(Y, Z) \in L^2(\mathbb{P})$ , that Assumption 2.3 holds, and that*

$$\sup_{h \in \overline{\mathcal{H}}} \mathbb{E} \left[ |\xi_0^h|^2 \exp \left( \frac{2k_\alpha}{1 - \alpha} \sup_{h' \in \mathcal{H}} \|f_{X^{h'}}\|_\infty |\xi_0^h| \right) \right] < \infty, \quad (2.15)$$

where  $k_\alpha = 1 \vee \frac{\alpha}{1 - \alpha}$ . If  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\gamma_1 > 0$  and  $\beta \in (0, 1]$ , and if  $\bar{\lambda}_1 \gamma_1 > 1$  when  $\beta = 1$ , then, for all  $h \in \mathcal{H}$  and all  $n \in \mathbb{N}^*$ ,

$$\mathbb{E}[(\xi_n^h - \xi_\star^h)^2] \leq C \gamma_n,$$

where

$$C = \sup_{h \in \mathcal{H}} \kappa_{h,1} (C_{h,1,1} + C_{h,2}), \quad (2.16)$$

and, for  $h \in \mathcal{H}$  and  $q \in \mathbb{N}^*$ ,  $C_{h,1,q} =$

$$\begin{cases} \gamma_1^{-1} (\mathbb{E}[\mathcal{L}_{h,1}^{\mu_{h,q}}(\xi_0^h)] + \frac{e+1}{e\mu_{h,q}+1}) \\ \quad \times \sup_{m \geq 1} \{ m^\beta \exp(2^{2\beta+1}(1 + e\mu_{h,q})\zeta_{h,1}^{\mu_{h,q}} \gamma_1^2 \Phi_{1-2\beta}(m+1) - \frac{\lambda_{h,1}^{\mu_{h,q}}}{2} \gamma_1 \Phi_{1-\beta}(m+1)) \} \\ \quad + (1 + e)\zeta_{h,1}^{\mu_{h,q}} (2^{2\beta} \gamma_1 \sup_{m \geq 1} \{ m^\beta \exp(-2^{-(\beta+2)} \lambda_{h,1}^{\mu_{h,q}} \gamma_1 m^{1-\beta}) \Phi_{1-2\beta}(m+1) \} + \frac{2^{\beta+1}}{\lambda_{h,1}^{\mu_{h,q}}}) \\ \hspace{15em} \text{if } \beta \in (0, 1), \\ \gamma_1^{-1} \exp((1 + e\mu_{h,q})\zeta_{h,1}^{\mu_{h,q}} \frac{\pi^2}{6} \gamma_1^2 + \frac{\gamma_1 \lambda_{h,1}^{\mu_{h,q}}}{2}) (\mathbb{E}[\mathcal{L}_{h,1}^{\mu_{h,q}}(\xi_0^h)] + \gamma_1^2 (1 + e)\zeta_{h,1}^{\mu_{h,q}} 2^{(\gamma_1 \lambda_{h,1}^{\mu_{h,q}})^2}) \\ \hspace{15em} \text{if } \beta = 1, \end{cases}$$

and  $C_{h,2} =$

$$\left\{ \begin{array}{l} (\gamma_1^{-1} \mathbb{E}[\bar{\mathcal{L}}_{h,2}(\xi_0^h)] + \frac{1+2^\beta C_{h,1,2}}{1+e\mu_{h,2}^2}) \\ \quad \times \sup_{m \geq 1} \{m^\beta \exp(2^{2\beta+1} \bar{\zeta}_{h,2} (1 + e\mu_{h,2}^2) \gamma_1^2 \Phi_{1-2\beta}(m+1) - \frac{\bar{\lambda}_{h,2}}{2} \gamma_1 \Phi_{1-\beta}(m+1))\} \\ + (2^\beta C_{h,1,2} + 1) \\ \quad \times \bar{\zeta}_{h,2} (\gamma_1^2 2^{2\beta} \sup_{m \geq 1} \{m^\beta \exp(-2^{-(\beta+2)} \bar{\lambda}_{h,2} \gamma_1 m^{1-\beta}) \Phi_{1-2\beta}(m+1)\} + \gamma_1 \frac{2^{2\beta+1}}{\bar{\lambda}_{h,2}}), \\ \hspace{20em} \text{if } \beta \in (0, 1), \\ \gamma_1^{-1} \exp((1 + e\mu_{h,2}^2) \bar{\zeta}_{h,2} \frac{\pi^2}{6} \gamma_1^2 + \frac{\gamma_1 \bar{\lambda}_{h,2}}{2}) \\ \quad \times (\mathbb{E}[\bar{\mathcal{L}}_{h,2}(\xi_0^h)] + \gamma_1^3 (2C_{h,1,2} + 1) \bar{\zeta}_{h,2} 2^{(\gamma_1 \bar{\lambda}_{h,2}) \vee 3} \sup_{m \geq 1} \{ \frac{\Phi_{\bar{\lambda}_{h,2} \gamma_1^{-2}(m+1)}}{(m+1)^{\bar{\lambda}_{h,2} \gamma_1^{-1}}} \}) \\ \hspace{20em} \text{if } \beta = 1, \end{array} \right.$$

and also, for  $\mu \geq 0$ ,

$$\begin{aligned} \bar{\zeta}_{h,q} &= \zeta_{h,q}^{\mu_{h,q}}, \quad \zeta_{h,q}^\mu = \frac{1}{2} \chi_{h,q}^\mu k_\alpha^2 ((\gamma_1 \sigma_q^\mu + \sigma_{q-1}^\mu) \vee \sigma_q^\mu), \\ \sigma_{h,q}^\mu &= 2^{q-1} e^{\mu k_\alpha^2 \gamma_1} ((1 + e\mu^q k_\alpha^2 \gamma_1^q) \vee e k_\alpha^2), \end{aligned}$$

with the constants  $\mu_{h,q}$ ,  $\kappa_{h,q}$ ,  $\lambda_{h,q}^\mu$ ,  $\bar{\lambda}_{h,q}$  and  $\chi_{h,q}^\mu$ , and the functions  $\mathcal{L}$  and  $\Phi$ , being defined in Lemmas A.1 and A.2.

*Remark 2.6.* (i) The value of the constant  $C$  in Lemma 2.5(ii) is retrieved from the proof of [14, Theorem 3.7(i)]. Its value is based on the results of Lemmas A.1 and A.2.

(ii) The proof of Lemma 2.5(ii) in [14] can be replicated line by line by taking  $h = 0$ , which corresponds to the unbiased case of Section 2.1. We can thereby obtain the following  $L^2(\mathbb{P})$ -control over the statistical error of the SA algorithm (2.5):

$$\mathbb{E}[(\xi_n^0 - \xi_\star^0)^2] \leq C \gamma_n, \quad n \in \mathbb{N}, \quad (2.17)$$

where

$$C = \kappa_{0,1} (C_{0,1} + C_{0,2}),$$

with  $C_{0,1}$  and  $C_{0,2}$  being defined respectively like  $C_{h,1}$  and  $C_{h,2}$  in Lemma 2.5(ii), once  $h$  is substituted with 0, and  $\kappa_{0,1}$  being defined in Lemma A.1(v).

## 2.2.2 Complexity Analysis

The controls of Lemma 2.5 on the bias and statistical errors in (2.11) result in the following complexities.

**Proposition 2.7** ([14, Theorem 3.8]). *Let  $\varepsilon > 0$  be a prescribed accuracy. Within the framework of Lemma 2.5(ii), setting*

$$h = \frac{1}{\lceil C \varepsilon^{-1} \rceil} \quad \text{and} \quad n = \lceil \bar{C}^{\frac{1}{\beta}} \gamma_1^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} \rceil,$$

where  $C, \bar{C} > 0$  are the constants described in (2.14) and (2.16) respectively, yields a global  $L^2(\mathbb{P})$  error for NSA of order  $\varepsilon$  as  $\varepsilon \downarrow 0$ . The corresponding computational cost verifies

$$\text{Cost}_{\text{NSA}}^\beta \leq C' n h^{-1} \sim C'' \varepsilon^{-\frac{2}{\beta}-1} \quad \text{as } \varepsilon \downarrow 0,$$

for some constants  $C', C'' > 0$  independent of  $\varepsilon$ . The minimal corresponding computational cost satisfies

$$\text{Cost}_{\text{NSA}}^1 \leq C'' \varepsilon^{-3} \quad \text{as } \varepsilon \downarrow 0,$$

and is attained for  $\gamma_n = \gamma_1 n^{-1}$ , i.e. for  $\beta = 1$  under the constraint  $\bar{\lambda}_1 \gamma_1 > 1$ .

### 2.3 Multilevel Stochastic Approximation Algorithm

With a complexity ceiling of  $O(\varepsilon^{-3})$  for nested VaR estimation, as achieved by NSA, a multilevel approach is proposed in [14] to accelerate the latter scheme. We recall below the MLSA scheme and the associated  $L^2(\mathbb{P})$ -control and complexity.

Take  $h_0 := \frac{1}{K} \in \mathcal{H}$ ,  $M \geq 2$  an integer and  $L \in \mathbb{N}^*$  a number of levels, and let

$$h_\ell := \frac{h_0}{M^\ell} = \frac{1}{KM^\ell} \in \mathcal{H}, \quad \ell \in \llbracket 0, L \rrbracket. \quad (2.18)$$

To each bias parameter  $h_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$ , corresponds an approximate problem  $\min_{\xi \in \mathbb{R}} V_{h_\ell}(\xi)$  to (2.2) with unique solution  $\xi_\star^{h_\ell}$ . These solutions can be telescoped into

$$\xi_\star^{h_L} = \xi_\star^{h_0} + \sum_{\ell=1}^L \xi_\star^{h_\ell} - \xi_\star^{h_{\ell-1}}. \quad (2.19)$$

By denoting  $\mathbf{N} := (N_\ell)_{0 \leq \ell \leq L}$  a sequence of iteration amounts, the MLSA scheme [14, 21] consists in approximating  $\xi_\star^0 \approx \xi_\star^{h_L}$  with

$$\xi_{\mathbf{N}}^{\text{ML}} = \xi_{N_0}^{h_0} + \sum_{\ell=1}^L \xi_{N_\ell}^{h_\ell} - \xi_{N_\ell}^{h_{\ell-1}}, \quad (2.20)$$

where each level  $\ell \in \llbracket 0, L \rrbracket$  is simulated independently. More precisely, once  $\xi_{N_0}^{h_0}$  is simulated using  $N_0$  iterations of (2.10) with bias  $h_0$ , at each level  $\ell \in \llbracket 1, L \rrbracket$ , for  $j \in \{(\ell-1), \ell\}$ ,  $\xi_{N_\ell}^{h_j}$  is obtained by iterating

$$\xi_{n+1}^{h_j} = \xi_n^{h_j} - \gamma_{n+1} H(\xi_n^{h_j}, X_{n+1}^{h_j}), \quad 0 \leq n \leq N_\ell - 1, \quad (2.21)$$

where  $(X_n^{h_{\ell-1}}, X_n^{h_\ell})_{1 \leq n \leq N_\ell} \stackrel{\text{i.i.d.}}{\sim} (X^{h_{\ell-1}}, X^{h_\ell})$  and  $(\xi_0^{h_{\ell-1}}, \xi_0^{h_\ell})$  is an  $\mathbb{R}^2$ -valued random initialization that is independent of the samples  $(X_n^{h_{\ell-1}}, X_n^{h_\ell})_{1 \leq n \leq N_\ell}$ . Crucially, at each level  $\ell \in \llbracket 1, L \rrbracket$ ,  $X^{h_\ell}$  and  $X^{h_{\ell-1}}$  are strongly correlated, in the sense that

$$X^{h_{\ell-1}} = \frac{1}{KM^{\ell-1}} \sum_{k=1}^{KM^{\ell-1}} \varphi(Y, Z^k), \quad X^{h_\ell} = \frac{1}{M} X^{h_{\ell-1}} + \frac{1}{KM^\ell} \sum_{k=KM^{\ell-1}+1}^{KM^\ell} \varphi(Y, Z^k), \quad (2.22)$$

where  $(Z^k)_{1 \leq k \leq KM^\ell} \stackrel{\text{i.i.d.}}{\sim} Z$  are independent of  $Y$ .

At each level  $\ell \in \llbracket 1, L \rrbracket$ ,  $\xi_{N_\ell}^{h_\ell}$  and  $\xi_{N_\ell}^{h_{\ell-1}}$  are referred to respectively as the fine and coarse approximations.  $\xi_{N_0}^{h_0}$  is referred to as the initial (or level 0) approximation. We can infer from (2.21) that MLSA correlates multiple NSA pairs, each pair being set with consecutive bias parameters on the geometric scale  $\mathcal{H}_0 := \{h_\ell, \ell \in \mathbb{N}\}$ . The produced effect is an incremental bias correction of the level 0 approximation with bias parameter  $h_0$ . Conversely, NSA can be viewed as an MLSA instance with 0 higher levels.

The correlation between  $X^{h_\ell}$  and  $X^{h_{\ell-1}}$  is primarily induced by their shared  $Y$  simulation. A quick look at the proof of [14, Theorem 4.6] (recalled below in Theorem 2.12) reveals that their overlap in simulations of  $Z$  has little impact on the error controls on the MLSA algorithm. Indeed, in MLMC, antithetic sampling [28] which seeks to exploit the correlation between  $X^{h_\ell}$  and  $X^{h_{\ell-1}}$  in terms of  $Z$  simulations, only achieves an improvement on standard MLMC in the order of  $|\ln \varepsilon|$ , for a prescribed accuracy  $\varepsilon > 0$ .

### 2.3.1 Convergence Analysis

The following lemma delimits three frameworks under which we derive  $L^2(\mathbb{P})$ -controls and complexities for the MLSA scheme.

**Lemma 2.8** ([14, Proposition 4.2]).

(i) Assume that the real-valued random variables  $X^h$  admit pdf  $f_{X^h}$  that are bounded uniformly in  $h \in \overline{\mathcal{H}}$ .

a. If there exists  $p_\star > 1$  such that

$$\mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{p_\star}] < \infty, \quad (2.23)$$

then, for all  $h, h' \in \overline{\mathcal{H}}$  such that  $0 \leq h \leq h'$ , and all  $\xi \in \mathbb{R}$ ,

$$\mathbb{E}[|\mathbb{1}_{\{X^h > \xi\}} - \mathbb{1}_{\{X^{h'} > \xi\}}|] \leq C(h' - h)^{\frac{p_\star}{2(p_\star+1)}},$$

where

$$C = B_{p_\star} \mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{p_\star}]^{\frac{1}{p_\star+1}} \left( \sup_{h \in \overline{\mathcal{H}}} \|f_{X^h}\|_\infty \right)^{\frac{p_\star}{p_\star+1}},$$

with  $B_{p_\star}$  being a positive constant that depends only on  $p_\star$ .

b. Assume that there exists a constant  $\mathfrak{g} \geq 0$  such that, for all  $u \in \mathbb{R}$ ,

$$\mathbb{E}\left[\exp\left(u(\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y])\right) \middle| Y\right] \leq e^{\mathfrak{g}u^2} \quad \mathbb{P}\text{-as.} \quad (2.24)$$

Then, for all  $h, h' \in \overline{\mathcal{H}}$  such that  $0 \leq h < h'$  and all  $\xi \in \mathbb{R}$ ,

$$\begin{aligned} & \mathbb{E}[|\mathbb{1}_{\{X^h > \xi\}} - \mathbb{1}_{\{X^{h'} > \xi\}}|] \\ & \leq 2\sqrt{\mathfrak{g}(h' - h)} \left( 1 + \sqrt{2} \left( \sup_{h'' \in \overline{\mathcal{H}}} \|f_{X^{h''}}\|_\infty \right) |\ln(\mathfrak{g}(h' - h))|^{\frac{1}{2}} \right). \end{aligned}$$

(ii) For  $\ell \in \mathbb{N}^*$ , let

$$G^\ell := \frac{X^{h_\ell} - X^{h_{\ell-1}}}{\sqrt{h_\ell}},$$

and

$$F_{X^{h_{\ell-1}}|G^\ell=g} : x \mapsto \mathbb{P}(X^{h_{\ell-1}} \leq x | G^\ell = g), \quad g \in \text{supp}(\mathbb{P}_{G^\ell}).$$

Consider the sequence of random variables  $(K^\ell)_{\ell \geq 1}$  given by  $K^\ell := K^\ell(G^\ell)$ , where

$$K^\ell(g) := \sup_{x \neq y} \frac{|F_{X^{h_{\ell-1}}|G^\ell=g}(x) - F_{X^{h_{\ell-1}}|G^\ell=g}(y)|}{|x - y|}, \quad g \in \text{supp}(\mathbb{P}_{G^\ell}), \quad \ell \in \mathbb{N}^*.$$

Assume that the  $(K^\ell)_{\ell \geq 1}$  satisfy

$$\sup_{\ell \geq 1} \mathbb{E}[|G^\ell|K^\ell] < \infty. \quad (2.25)$$

Then, for all  $\xi \in \mathbb{R}$ ,

$$\mathbb{E}[|\mathbb{1}_{\{X^{h_\ell} > \xi\}} - \mathbb{1}_{\{X^{h_{\ell-1}} > \xi\}}|] \leq Ch_\ell^{\frac{1}{2}},$$

where

$$C = \sup_{\ell \geq 1} \mathbb{E}[|G^\ell|K^\ell].$$

*Remark 2.9.* Lemma 2.8's frameworks are ordered from weakest to strongest.

- (i) The constant  $B_{p_\star}$  in the postulate of Lemma 2.8(i)a emanates from a Burkholder-Davis-Gundy inequality. According to Revuz and Yor [36, Chapter IV, Section 4],

$$B_{p_\star} = \begin{cases} \frac{4-p_\star}{2-p_\star} 2^{p_\star} & \text{if } p_\star < 2, \\ \left(\frac{p_\star}{p_\star-1}\right)^{\frac{p_\star}{2}} & \text{if } p_\star \geq 2. \end{cases}$$

- (ii) According to Fathi and Frikha [20] and Frikha and Menozzi [22], if

$$\mathbb{E} \left[ \exp \left( u_0 (\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y])^2 \right) \right] < \infty$$

for some  $u_0 > 0$ , then there exists  $\mathfrak{g} > 0$  such that, for all  $u \geq 0$ ,

$$\mathbb{E} \left[ \exp \left( u (\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]) \right) \right] \leq e^{\mathfrak{g}u^2}.$$

This is analogous to our Gaussian concentration framework (2.24).

Note that the framework (2.24) entails the framework (2.23) for all  $p_\star > 1$ , which can be shown via an exponential series expansion.

- (iii) ([14, Remark 4.3]) The framework (2.25) relaxes Gordy and Juneja's [30, Assumption 1], and is the most advantageous computationally (c.f. Proposition 2.13(iii)). It holds for instance if  $(g, \ell) \mapsto F_{X^{h\ell-1}|G^\ell=g}$  is uniformly Lipschitz in  $g$  and  $\ell$ .

**Assumption 2.10** ([14, Assumption 4.4]). There exist some constants  $c, \delta_0 > 0$  such that, for all  $h \in \mathcal{H}$  and for any compact set  $\mathcal{K} \subset \mathbb{R}$ ,

$$\sup_{\xi \in \mathcal{K}} |f_{X^h}(\xi) - f_{X^0}(\xi)| \leq ch^{\frac{1}{4} + \delta_0}.$$

*Remark 2.11* ([14, Remark 4.5]). Assumption 2.10 holds within the framework of [29, Proposition 5.1(a)].

**Theorem 2.12** ([14, Theorem 4.6]). Suppose that  $\varphi(Y, Z) \in L^2(\mathbb{P})$ , that Assumptions 2.3 and 2.10 hold, and that

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

where  $k_\alpha = 1 \vee \frac{\alpha}{1-\alpha}$ . If  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\gamma_1 > 0$  and  $\beta \in (0, 1]$ , and if  $\bar{\lambda}_2 \gamma_1 > 2$  when  $\beta = 1$ , then, there exists some constant  $C > 0$  such that, for all  $L \in \mathbb{N}^*$  and all  $\mathbf{N} = (N_0, \dots, N_L) \in (\mathbb{N}^*)^{L+1}$ ,

$$\mathbb{E}[(\xi_{\mathbf{N}}^{\text{ML}} - \xi_{\star}^{h_L})^2] \leq C \left( \gamma_{N_0} + \left( \sum_{\ell=1}^L \gamma_{N_\ell} \right)^2 + \sum_{\ell=1}^L \gamma_{N_\ell}^{\frac{3}{2}} + \sum_{\ell=1}^L \gamma_{N_\ell} \epsilon(h_\ell) \right), \quad (2.26)$$

where

$$\epsilon(h) := \begin{cases} h^{\frac{p_\star}{2(1+p_\star)}} & \text{if (2.23) holds,} \\ h^{\frac{1}{2}} |\ln h|^{\frac{1}{2}} & \text{if (2.24) holds,} \\ h^{\frac{1}{2}} & \text{if (2.25) holds,} \end{cases} \quad h \in \mathcal{H}. \quad (2.27)$$

### 2.3.2 Complexity Analysis

The global error of MLSA can be decomposed into statistical and bias errors:

$$\xi_{\mathbf{N}}^{\text{ML}} - \xi_{\star}^0 = (\xi_{\mathbf{N}}^{\text{ML}} - \xi_{\star}^{h_L}) + (\xi_{\star}^{h_L} - \xi_{\star}^0). \quad (2.28)$$

**Proposition 2.13** ([14, Proposition 4.7, Lemma 4.8 & Theorem 4.9]).

(i) Suppose that Assumption 2.3(i) is satisfied. Let  $\varepsilon > 0$  be a prescribed accuracy and  $C > 0$  the constant defined in (2.14). If  $h_0 > C^{-1}\varepsilon$ , then, setting

$$L = \left\lceil \frac{\ln(Ch_0\varepsilon^{-1})}{\ln M} \right\rceil \geq 1 \quad (2.29)$$

achieves a bias error for MLSA of order  $\varepsilon$ .

(ii) The computational cost of MLSA satisfies

$$\text{Cost}_{\text{MLSA}}^{\beta} \leq \bar{C} \sum_{\ell=0}^L \frac{N_{\ell}}{h_{\ell}},$$

for some constant  $\bar{C} > 0$ .

(iii) Let  $\varepsilon > 0$  be a prescribed accuracy. Within the framework of Theorem 2.12, there exists a constant  $C' > 0$  such that, setting

$$N_{\ell} = \left\lceil (C'\gamma_1)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{\beta}{1+\beta}} \epsilon(h_{\ell'})^{\frac{1}{1+\beta}} \right)^{\frac{1}{\beta}} h_{\ell}^{\frac{1}{1+\beta}} \epsilon(h_{\ell})^{\frac{1}{1+\beta}} \right\rceil, \quad \ell \in \llbracket 0, L \rrbracket,$$

i.e.

$$N_{\ell} = \begin{cases} \left\lceil [(C'\gamma_1)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} h_{\ell}^{\frac{1}{1+\beta}(1+\frac{p^*}{2(1+p^*)})} (\sum_{\ell'=0}^L h_{\ell'}^{-\frac{1}{1+\beta}(\beta-\frac{p^*}{2(1+p^*)})})^{\frac{1}{\beta}}] \right\rceil & \text{if (2.23) holds,} \\ \left\lceil [(C'\gamma_1)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} h_{\ell}^{\frac{3}{2(1+\beta)}} |\ln h_{\ell}|^{\frac{1}{2(1+\beta)}} (\sum_{\ell'=0}^L h_{\ell'}^{-\frac{2\beta-1}{2(1+\beta)}} |\ln h_{\ell'}|^{\frac{1}{2(1+\beta)}})^{\frac{1}{\beta}}] \right\rceil & \text{if (2.24) holds,} \\ \left\lceil [(C'\gamma_1)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} h_{\ell}^{\frac{3}{2(1+\beta)}} (\sum_{\ell'=0}^L h_{\ell'}^{-\frac{2\beta-1}{2(1+\beta)}})^{\frac{1}{\beta}}] \right\rceil & \text{if (2.25) holds,} \end{cases}$$

achieves a statistical error on the estimation of  $\xi_{\star}^0$  of order  $\varepsilon$ . The optimal computational cost is attained when  $\beta = 1$  under the constraint  $\lambda_2\gamma_1 > 2$ , in which case

$$\text{Cost}_{\text{MLSA}}^1 \leq C'' \begin{cases} \varepsilon^{-3+\frac{p^*}{2(1+p^*)}} & \text{if (2.23) holds,} \\ \varepsilon^{-\frac{5}{2}} |\ln \varepsilon|^{\frac{1}{2}} & \text{if (2.24) holds,} \\ \varepsilon^{-\frac{5}{2}} & \text{if (2.25) holds,} \end{cases}$$

for some constant  $C'' > 0$ .

*Remark 2.14.* The advantageous framework described by (2.25) results in its best computational cost in  $O(\varepsilon^{-\frac{5}{2}})$ , which remains suboptimal compared to the canonical optimum of  $O(\varepsilon^{-2})$  that is achievable by multilevel techniques [24, 21].

The next section devises a strategy to adaptively refine the inner Monte Carlo samples of NSA and MLSA.

### 3 Adaptive Refinement Strategy

The suboptimality of MLSA is linked to the discontinuity of the gradient  $H$  (2.3) intervening in the VaR recursion (2.21). Indeed, for  $n \in \mathbb{N}^*$ , if the simulated loss

$$X_n^h = \frac{1}{K} \sum_{k=1}^K \varphi(Y_n, Z_n^k)$$

is too close to the estimate  $\xi_{n-1}^h$  but falls on its opposite side with respect to its sampling target

$$X_n^0 := \mathbb{E}[\varphi(y, Z)]|_{y=Y_n}, \quad (3.1)$$

an  $O(1)$  error is registered due to the discontinuity of  $H(\xi_{n-1}^h, \cdot)$  at  $\xi_{n-1}^h$ , thus lowering the overall performance of the multilevel algorithm. To address this issue, we investigate the incorporation of an adaptive refinement layer into MLSA.

The following steps elucidate the intuition behind our adaptive refinement strategy. For simplicity, we consider a nested simulation  $X^{h_\ell}$ , targeting an unbiased simulation  $X^0$ , that we want to refine adaptively given a current iterate  $\xi$  at the iteration number  $n$ . Assuming that  $Y$  and  $Z^1, \dots, Z^{KM^\ell}$  *i.i.d.*  $Z$  were used to simulate  $X^{h_\ell}$  according to (2.7), refining the latter to  $X^{h_{\ell+1}}$  consists in simulating additional samples  $Z^{KM^\ell+1}, \dots, Z^{KM^{\ell+1}}$  *i.i.d.*  $Z$  independently from  $Y$  and  $Z^1, \dots, Z^{KM^\ell}$  and setting

$$X^{h_{\ell+1}} = \frac{1}{M} X^{h_\ell} + \frac{1}{KM^{\ell+1}} \sum_{k=KM^\ell+1}^{KM^{\ell+1}} \varphi(Y, Z^k). \quad (3.2)$$

*Step 1. A confidence based heuristic.*

We loosely adapt the reasonings employed in [26, Section 2.3.1] and [31, Section 3] to derive a preliminary refinement strategy. Roughly speaking, considering a refinement amount  $\eta \in \mathbb{N}$ , we want to ensure that  $H(\xi, X^{h_{\ell+\eta}}) = H(\xi, X^0)$  with high probability by aligning  $X^{h_{\ell+\eta}}$  with  $X^0$  on the same side of the discontinuity of  $H(\xi, \cdot)$  at  $\xi$ . Via a conditional CLT,

$$X^{h_{\ell+\eta}} \approx \mathcal{N}\left(X^0, h_{\ell+\eta} \text{Var}(\varphi(Y, Z)|Y)\right). \quad (3.3)$$

To achieve the desired alignment, we consider a critical value  $C_{\text{ad}}$  corresponding to a confidence level  $p \in (0, 1)$ , and choose  $\eta$  minimal such that  $|X^{h_{\ell+\eta}} - X^0| \leq |X^0 - \xi|$  with confidence  $p$ , which can be expressed as

$$C_{\text{ad}} \sqrt{h_{\ell+\eta} \text{Var}(\varphi(Y, Z)|Y)} \leq |X^0 - \xi|, \quad \text{or} \quad |X^0 - \xi| \geq C_{\text{ad}} h_{\ell+\eta}^{\frac{1}{2}}, \quad (3.4)$$

up to a modification of  $C_{\text{ad}}$  conditionally on  $Y$ . However, this approach is impractical as  $X^0$  is inaccessible.

We thus change our perspective and require instead that  $|X^{h_{\ell+\eta}} - X^0| \leq |X^{h_{\ell+\eta}} - \xi|$  with confidence  $p$ , to ensure that  $X^0$  and  $X^{h_{\ell+\eta}}$  are on the same side relative to  $\xi$ . We then select  $\eta$  minimal such that

$$C_{\text{ad}} \sqrt{h_{\ell+\eta} \text{Var}(\varphi(Y, Z)|Y)} \leq |X^{h_{\ell+\eta}} - \xi|, \quad \text{i.e.} \quad |X^{h_{\ell+\eta}} - \xi| \geq C_{\text{ad}} h_{\ell+\eta}^{\frac{1}{2}}, \quad (3.5)$$

allowing for a modification of  $C_{\text{ad}}$  conditional on  $Y$ . In effect, we augment the number of inner simulations of  $X^{h_\ell}$  by  $\eta$  refinements until the threshold  $C_{\text{ad}} h_{\ell+\eta}^{\frac{1}{2}}$  around  $\xi$  is crossed. [26] views this process as estimating  $X^0$  by  $X^{h_{\ell+\eta}}$  in the original strategy (3.4), while [27] interprets the

evaluation of the criterion (3.5) as performing a Student t-test on the null hypothesis “ $X^{h_{\ell+\eta}} = \xi$ ”.

For more flexibility, we introduce two parameters  $r > 1$  and  $0 < \theta \leq 1$  that respectively control the strictness and budgeting of the refinement strategy.

*Step 2. Refinement strictness.*

We redefine the refinement strategy as choosing  $\eta$  minimal such that

$$|X^{h_{\ell+\eta}} - \xi| \geq C_{\text{ad}} h_{\ell+\eta}^{\frac{1}{r}}. \quad (3.6)$$

The parameter  $r$  allows us to adjust the strategy’s propensity to refine. Larger  $r$  values impose more strictness. Setting  $r = 2$  retrieves the previous baseline strategy.

*Step 3. Refinement budgeting.*

The strategy outlined above may be risky, as the refinement amount  $\eta$  required to satisfy (3.6) could be excessively large, hence increasing the subsequent complexity. To address this, we impose an upper limit on  $\eta$  at  $\lceil \theta \ell \rceil$ :

$$\eta(\xi) = \lceil \theta \ell \rceil \wedge \min\{k : |X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}} h_{\ell+k}^{\frac{1}{r}}\}. \quad (3.7)$$

Note importantly the dependency of  $\eta$  on  $\xi$ . For this strategy to be computationally efficient, the entailed number of inner simulations should, on average, be comparable to the number of inner simulations absent the strategy. To ensure this, we relax the constant  $C_{\text{ad}}$  to a normalization factor  $c(h_\ell)$  that is calibrated such that  $\mathbb{E}[h_{\ell+\eta(\xi)}^{-1}] = O(h_\ell^{-1})$ . We have

$$\mathbb{E}[h_{\ell+\eta(\xi)}^{-1}] = \sum_{k=0}^{\lceil \theta \ell \rceil} h_{\ell+k}^{-1} \mathbb{P}(\eta(\xi) = k) \leq h_\ell^{-1} + \sum_{k=1}^{\lceil \theta \ell \rceil} h_{\ell+k}^{-1} \mathbb{P}(\eta(\xi) = k).$$

From (3.7) and Assumption 2.3(iv),

$$\mathbb{P}(\eta(\xi) = k) \leq \mathbb{P}(|X^{h_{\ell+k-1}} - \xi| < c(h_\ell) h_{\ell+k-1}^{\frac{1}{r}}) \leq 2M^{\frac{1}{r}} \sup_{h \in \bar{H}} \|f_{X^h}\|_\infty c(h_\ell) h_{\ell+k}^{\frac{1}{r}}.$$

Extending the definition of  $s \mapsto h_s$  to  $\mathbb{R}$  with  $h_s := \frac{h_0}{M^s}$ , we obtain, for some constant  $C > 0$ ,

$$\mathbb{E}[h_{\ell+\eta(\xi)}^{-1}] \leq C \left( h_\ell^{-1} + c(h_\ell) \sum_{k=1}^{\lceil \theta \ell \rceil} h_{\ell+k}^{\frac{1}{r}} \right) \leq C (h_\ell^{-1} + c(h_\ell) h_{(1+\theta)\ell(\frac{1}{r}-1)}).$$

To balance the terms on the right-hand side, we normalize with

$$c(h_\ell) = h_{\theta\ell(r-1)-\ell}^{\frac{1}{r}},$$

yielding the refinement strategy

$$\eta(\xi) = \lceil \theta \ell \rceil \wedge \min\{k : |X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}}\}.$$

This approach is essentially the same as the one delineated in [31]. The amount of inner simulations is increased until the varying threshold  $C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}}$  around  $\xi$  is crossed by the sample  $X^{h_{\ell+k}}$ .

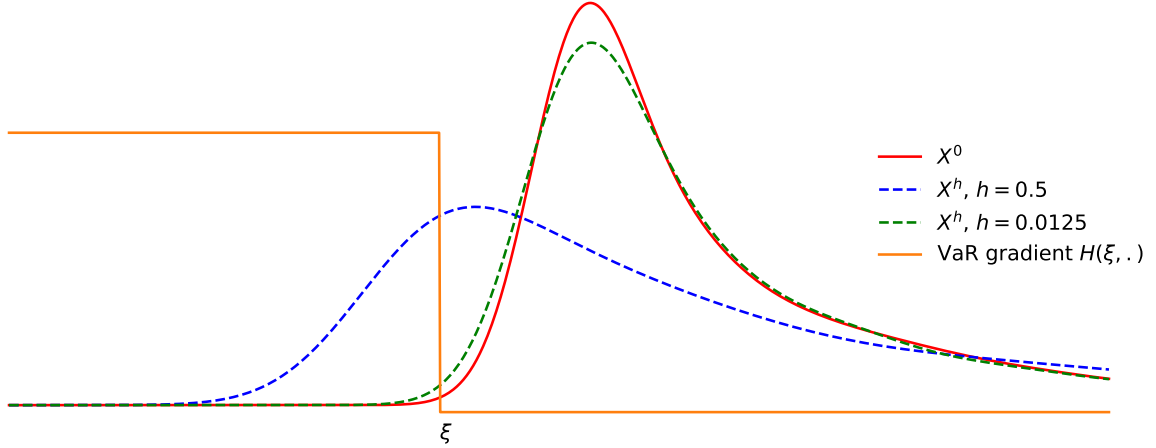
*Step 4. Refinement saturation.*

To illustrate the strategy thus far, let's test it along a single dynamic of (2.10) with bias parameter  $h_\ell$ :

$$\xi_{n+1}^{h_\ell} = \xi_n^{h_\ell} - \gamma_{n+1} H(\xi_n^{h_\ell}, X_{n+1}^{h_\ell + \eta(\xi_n^{h_\ell})}).$$

Unlike the MC setting [31] where the iterate  $\xi$  remains constant across the recursions, the iterates  $(\xi_n^{h_\ell})_{n \geq 1}$  above change from one step to the next. This causes the samples  $(X_n^{h_\ell})_{n \geq 1}$  to refine to  $(X_n^{h_\ell + \eta(\xi_{n-1}^{h_\ell})})_{n \geq 1}$ , de facto solving the root finding program  $\xi : \mathbb{E}[H(\xi, X^{h_\ell + \eta(\xi)})] = 0$ , or equivalently,  $\xi : \mathbb{E}[\mathbb{1}_{\{X^{h_\ell + \eta(\xi)} < \xi\}}] = \alpha$ . Since  $\eta(\xi)$  evolves with  $\xi$ ,  $\xi \mapsto \mathbb{E}[\mathbb{1}_{\{X^{h_\ell + \eta(\xi)} < \xi\}}]$  is no longer given by a cdf, as it may not be monotone and could have several roots.

To address this issue, we saturate the refinement amount  $\eta$  for large iteration numbers  $n$  (corresponding to terminal SA phases) to the cap  $[\theta \ell]$ , turning it into the convex program  $\min_\xi V_{h_\ell + [\theta \ell]}(\xi)$  for  $n$  large. This is achieved by incorporating an increasing dependency on  $n$  into the strategy's threshold. Such dependency was calibrated from the ensuing error controls provided in Proposition 4.4 and Theorem 5.4. We refer to Figure 1 for a visual representation of the adaptive refinement procedure, to (3.10) for our definitive strategy and to Remark 4.1 for additional related comments.



**Figure 1.** An illustration of the adaptive refinement procedure. The sample  $X^h$  is refined by lowering its bias parameter  $h$  from  $\frac{1}{2}$  to  $\frac{1}{8}$ , to ensure that  $H(\xi, X^h)$  is equal to  $H(\xi, X^0)$  with great probability. The plotted pdf of  $X^0$  and  $X^h$ ,  $h \in \{\frac{1}{2}, \frac{1}{8}\}$ , are based on the case study of Section 7.1. The gradient  $x \mapsto H(\xi, \cdot)$  is not represented at scale.

The following lemma revisits the framework (2.25) to grant us a flexible basis for our adaptive refinement strategy. Its proof is postponed to Appendix B.

**Lemma 3.1.** For  $0 \leq h < h' \in \overline{\mathcal{H}}$ , define

$$G_h^{h'} := \frac{X^{h'} - X^h}{\sqrt{h'}}, \quad (3.8)$$

and

$$F_{X^h | G_h^{h'} = g} : x \mapsto \mathbb{P}(X^h \leq x | G_h^{h'} = g), \quad g \in \text{supp}(\mathbb{P}_{G_h^{h'}}).$$

Assume that the random variables  $(K_h^{h'})_{0 \leq h < h' \in \overline{\mathcal{H}}}$ , defined by  $K_h^{h'} := K_h^{h'}(G_h^{h'})$ , where

$$K_h^{h'}(g) := \sup_{x \neq y} \frac{|F_{X^h | G_h^{h'} = g}(x) - F_{X^h | G_h^{h'} = g}(y)|}{|x - y|}, \quad g \in \text{supp}(\mathbb{P}_{G_h^{h'}}), \quad 0 \leq h < h' \in \overline{\mathcal{H}},$$

satisfy

$$\sup_{0 \leq h < h' \in \overline{\mathcal{H}}} \mathbb{E}[K_h^{h'} | G_h^{h'}] < \infty. \quad (3.9)$$

Then, for all  $0 \leq h < h' \in \overline{\mathcal{H}}$ ,

$$\mathbb{E}[|\mathbb{1}_{\{X^{h'} > \xi\}} - \mathbb{1}_{\{X^h > \xi\}}|] \leq C(h')^{\frac{1}{2}},$$

where

$$C = \sup_{0 \leq h_1 < h_2 \in \overline{\mathcal{H}}} \mathbb{E}[K_{h_1}^{h_2} | G_{h_1}^{h_2}].$$

*Remark 3.2.* The framework (2.25) can be viewed as a special case of (3.9) for consecutive bias pairs on the geometric scale  $\mathcal{H}_0 = \{h_\ell, \ell \in \mathbb{N}\} \subset \mathcal{H}$ . In standard MLSA, the fine and coarse approximations at each level are controlled by consecutive bias parameters in  $\mathcal{H}_0$ . If one is to refine either approximation separately, the refined fine and coarse approximations are no longer controlled by consecutive bias parameters in  $\mathcal{H}_0$ , hence the need for the generalized framework above.

Following the roadmap of [26, 31, 27], we define

$$\eta_n^\ell(\xi) := \lceil \theta \ell \rceil \wedge \min \{k \in \llbracket 0, \lceil \theta \ell \rceil \rrbracket : |X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}} \psi_n^{\ell, k}\}, \quad \ell, n \in \mathbb{N}^*, \quad (3.10)$$

with the convention  $\min \emptyset = +\infty$ , where  $C_{\text{ad}} > 0$ ,  $r > 1$ ,  $0 < \theta \leq 1$ ,

$$\psi_n^{\ell, k} := \begin{cases} u_n^{\frac{-1}{p^*}} h_{\theta \ell(r-1)+k}^{\frac{1}{r}} & \text{if (2.23) holds,} \\ \ln^{\frac{1}{2}}((\gamma_n/\gamma_1)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}) h_{\theta \ell(r-1)+k}^{\frac{1}{r}} & \text{if (2.24) or (3.9) holds,} \end{cases} \quad (3.11)$$

and

$$u_n = \gamma_1 n^{-\delta}, \quad \text{with } \delta \in (0, 1]. \quad (3.12)$$

Note the property

$$\begin{aligned} \sum_{k=0}^{\lceil \theta \ell \rceil} \psi_n^{\ell, k} &\leq \begin{cases} (1 - M^{-\frac{1}{r}})^{-1} \psi_n^{\ell, 0} & \text{if (2.23) holds,} \\ (1 - M^{-\frac{1}{r}})^{-1} \sqrt{2 + \theta} \psi_n^{\ell, 0} & \text{if (2.24) or (3.9) holds,} \end{cases} \\ &\leq \frac{\sqrt{3}}{1 - M^{-\frac{1}{2}}} \psi_n^{\ell, 0}, \end{aligned} \quad (3.13)$$

where we used in the last inequality the constraints  $\theta \in (0, 1]$  and  $r \in (1, 2]$ , which stem from the frameworks of Lemmas 4.2 and 5.2.

We refer to  $C_{\text{ad}}$  as the confidence constant [26, 31],  $r$  as the refinement strictness parameter [31, 27],  $\theta$  as the budgeting parameter, the  $n$ -dependent factor in (3.11) as the saturation factor, and the quantity  $\frac{|X^{h_{\ell+k}} - \xi|}{\psi_n^{\ell, k}}$ , that our adaptive strategy seeks to greedily make as large as possible, as the error margin [12].

---

**Algorithm 1** Adaptive refinement strategy

---

**Require:** A sample  $X^{h_\ell}$ , an iterate  $\xi$ , refinement parameters  $C_{\text{ad}} > 0$ ,  $r > 1$  and  $0 < \theta \leq 1$ , a level  $\ell \in \mathbb{N}^*$ , an iteration number  $n \in \mathbb{N}^*$

- 1:  $\eta \leftarrow 0$
  - 2: **while**  $\eta < \lceil \theta \ell \rceil$  **and**  $|X^{h_{\ell+\eta}} - \xi| < C_{\text{ad}} \psi_n^{\ell, \eta}$  **do**
  - 3:     Refine  $X^{h_{\ell+\eta}}$  to  $X^{h_{\ell+\eta+1}}$
  - 4:      $\eta \leftarrow \eta + 1$
  - 5: **end while**
  - 6: **return**  $X^{h_{\ell+\eta}}$
-

We refer to [31, Algorithm 3.1] for an analogous refinement strategy in an MLMC setting, that is independent of the iteration number  $n$ .

*Remark 3.3.* (i) At each level  $\ell$  and iteration number  $n$ , the adaptive strategy refines the simulation  $X_n^{h_\ell}$  used in (2.21) to  $X_n^{h_{\ell+\eta}}$ , with the aim of escaping the region  $[\xi \pm C_{\text{ad}}\psi_n^{\ell,\eta}]$  around the discontinuity of  $x \mapsto H(\xi, x)$  at  $x = \xi$ .

(ii) No refinement is applied at the level  $\ell = 0$ , as  $\eta_n^0 = 0$  for all  $n \in \mathbb{N}^*$ .

(iii) The thresholds  $C_{\text{ad}}\psi_n^{\ell,k}$  depend on the hyperparameters  $r > 1$  and  $0 < \theta \leq 1$ . It decreases for  $r$  or  $\theta$  large. It also depends on the level  $\ell$  and the iteration number  $n$ . It decreases for  $\ell$  large and increases for  $n$  large.

(iv) Recalling the definition (3.1), by virtue of the almost sure convergence of  $X_n^{h_\ell}$  to  $X_n^0$  as  $\ell \uparrow \infty$  (by the conditional law of large numbers, under suitable assumptions), the refined simulation  $X_n^{h_{\ell+\eta_n^\ell(\varepsilon)}}$  is in theory closer to the actual simulation target  $X_n^0$  than  $X_n^{h_\ell}$ . It is therefore expected that  $H(\xi, X_n^{h_{\ell+\eta_n^\ell(\varepsilon)}}) = H(\xi, X_n^0)$  with high probability. See Remark 4.1 for a complementary comment.

(v) In our derivation of the adaptive refinement strategy, the confidence constant  $C_{\text{ad}}$  took in fact the form of  $C_p\sqrt{\text{Var}(\varphi(Y, Z)|Y)}$ , where  $C_p$  is a critical value of the law  $\mathcal{N}(0, 1)$  at some confidence level  $p$  and  $\sqrt{\text{Var}(\varphi(Y, Z)|Y)}$  is the sampling standard deviation. For a 99%-confidence level, a natural choice for  $C_{\text{ad}}$  is three standard deviations, i.e.  $C_p = 3$ . The standard deviation can be estimated empirically. However, the ensuing convergence and complexity analyses show that our choice of  $C_{\text{ad}}$  constant for the adaptive strategy (3.10) accomplishes the desired performance gain. In practice,  $C_{\text{ad}}$  should be fine-tuned to avoid over- or under-adapting.

## 4 Adaptive Nested Stochastic Approximation Algorithm

Before delving into the adaptive refinement of MLSA, let us first investigate the influence of our refinement strategy on NSA, which we recall can be considered an instance of 0 level MLSA.

Consider a bias parameter  $h_\ell$ ,  $\ell \in \mathbb{N}^*$ , on the geometric scale  $\mathcal{H}_0 = \{h_{\ell'}, \ell' \in \mathbb{N}\}$ , which allows resorting to single-level refinements (3.2). The adaptively refined nested SA algorithm for estimating the VaR can be written as

$$\tilde{\xi}_{n+1}^{h_\ell} = \tilde{\xi}_n^{h_\ell} - \gamma_{n+1}H(\tilde{\xi}_n^{h_\ell}, \tilde{X}_{n+1}^{h_\ell}), \quad n \in \mathbb{N}, \quad (4.1)$$

where

$$\tilde{X}_{n+1}^{h_\ell} := X_{n+1}^{h_{\ell+\tilde{\eta}_{n+1}^\ell}}, \quad \tilde{\eta}_{n+1}^\ell := \eta_{n+1}^\ell(\tilde{\xi}_n^{h_\ell}), \quad n \in \mathbb{N}, \quad (4.2)$$

and  $\tilde{\xi}_0^{h_\ell}$  is a random real-valued initialization that is independent of the samples  $(\tilde{X}_n^{h_\ell})_{n \geq 1}$ .

---

**Algorithm 2** Adaptive Nested SA for estimating the VaR
 

---

**Require:**  $K, M, \ell, N \in \mathbb{N}^*$  with  $M \geq 2$ , a positive non-increasing sequence  $(\gamma_n)_{n \geq 1}$  such that

$$\sum_{n=1}^{\infty} \gamma_n = \infty \text{ and } \lim_{n \rightarrow \infty} \gamma_n = 0, \text{ refinement parameters } C_{\text{ad}} > 0, r > 1 \text{ and } 0 < \theta \leq 1$$

- 1: Sample  $\xi_0^{h_\ell}$  randomly
  - 2: **for**  $n = 0 \dots N - 1$  **do**
  - 3:   Simulate  $Y_{n+1} \sim Y$  and  $Z_{n+1}^1, \dots, Z_{n+1}^{KM^\ell} \stackrel{\text{i.i.d.}}{\sim} Z$  independently of  $Y_{n+1}$
  - 4:    $X_{n+1}^{h_\ell} \leftarrow \frac{1}{KM^\ell} \sum_{k=1}^{KM^\ell} \varphi(Y_{n+1}, Z_{n+1}^k)$
  - 5:    $\tilde{X}_{n+1}^{h_\ell} \leftarrow \text{Refine}_{C_{\text{ad}}, r, \theta, \ell, n}(X_{n+1}^{h_\ell}, \tilde{\xi}_n^{h_\ell})$
  - 6:    $\tilde{\xi}_{n+1}^{h_\ell} \leftarrow \tilde{\xi}_n^{h_\ell} - \gamma_{n+1} H(\tilde{\xi}_n^{h_\ell}, \tilde{X}_{n+1}^{h_\ell})$
  - 7: **end for**
  - 8: **return**  $\tilde{\xi}_N^{h_\ell}$
- 

*Remark 4.1.* Unlike in [26, 31, 27], the adaptive refinement (3.10) depends on the iteration number  $n$ . As previously discussed, using an adaptive refinement independent of  $n$  can be viewed as seeking a root of  $\xi \mapsto \mathbb{E}[H(\xi, X^{h_{\ell+\eta^\ell(\xi)}})]$ , which is not guaranteed to retrieve a problem like (2.8), as it could have multiple roots. The introduced dependency on  $n$  saturates the refinement amount to  $\lceil \theta \ell \rceil$  for large  $n$ , thus aligning  $X_n^{h_{\ell+\eta_n^\ell(\xi)}}$  with  $X_n^{h_{\ell+\lceil \theta \ell \rceil}}$  with high probability and practically solving the strictly convex nested SA problem  $\min_{\xi} V_{h_{\ell+\lceil \theta \ell \rceil}}(\xi)$  with bias parameter  $h_{\ell+\lceil \theta \ell \rceil}$ .

This behavior can be verified by calculation. Assuming that  $\varphi(Y, Z) \in L^1(\mathbb{P})$ , by Markov's inequality,

$$\begin{aligned} \mathbb{P}(\eta_n^\ell(\xi) < \lceil \theta \ell \rceil) &= \mathbb{P}(\exists k \in \llbracket 0, \lceil \theta \ell \rceil - 1 \rrbracket, |X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}} \psi_n^{\ell, k}) \\ &\leq \sum_{k=0}^{\lceil \theta \ell \rceil - 1} \mathbb{P}(|X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}} \psi_n^{\ell, k}) \\ &\leq \frac{1}{C_{\text{ad}} \psi_n^{\ell, k}} \sum_{k=0}^{\lceil \theta \ell \rceil - 1} \mathbb{E}[|X^{h_{\ell+k}} - \xi|] \rightarrow 0 \text{ as } n \uparrow \infty. \end{aligned} \quad (4.3)$$

Observe that

$$\mathbb{P}(\exists n_0 \in \mathbb{N}, \forall n \geq n_0, \eta_n^\ell(\xi) = \lceil \theta \ell \rceil) = 1 - \mathbb{P}(\limsup_{n \uparrow \infty} \{\eta_n^\ell(\xi) < \lceil \theta \ell \rceil\}).$$

Given the increasing monotonicity of  $n \mapsto \psi_n^{\ell, k}$  following Remark 3.3(iii), the events  $(\{\eta_n^\ell(\xi) < \lceil \theta \ell \rceil\})_{n \geq 0}$  are decreasing, so that by Fatou's lemma and (4.3),

$$\mathbb{P}(\limsup_{n \uparrow \infty} \{\eta_n^\ell(\xi) < \lceil \theta \ell \rceil\}) \leq \liminf_{n \uparrow \infty} \mathbb{P}(\eta_n^\ell(\xi) < \lceil \theta \ell \rceil) = 0.$$

Finally,

$$\mathbb{P}(\exists n_0 \in \mathbb{N}, \forall n \geq n_0, \eta_n^\ell(\xi) = \lceil \theta \ell \rceil) = 1.$$

In words, since  $(\eta_n^\ell(\xi))_{n \geq 1}$  are discrete random variables taking values in  $\llbracket 0, \lceil \theta \ell \rceil \rrbracket$ , considering (4.3), there exists a  $\mathbb{P}$ -as-finite random time  $n_0$  that depends on  $\ell$  and  $\xi$ , such that for  $n \in \mathbb{N}$ , the equality  $\eta_n^\ell(\xi) = \lceil \theta \ell \rceil$  holds on the event  $\{n \geq n_0\}$ . Hence the desired saturation effect for  $n$  large.

Of course, the iterate  $\xi$  depends on the iteration number  $n$ . The boundedness of the iterates in some  $L^p(\mathbb{P})$  space,  $0 < p \leq 2$ , is hence essential to retrieve the asymptotic behavior described in (4.3). Such boundedness is established in Proposition 4.4.

## 4.1 Convergence Analysis

The proofs of the ensuing results are postponed to Appendix B.

**Lemma 4.2.** *Let  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\gamma_1 > 0$  and  $\beta \in (0, 1]$ .*

(i) *Assume that the real-valued random variables  $X^h$  admit pdf  $f_{X^h}$  that are bounded uniformly in  $h \in \overline{\mathcal{H}}$ .*

a. *If (2.23) is satisfied, with*

$$p_\star > 2, \quad 1 < r < 2, \quad \text{and} \quad 0 < \theta \leq \frac{p_\star - 2}{p_\star + 2}, \quad (4.4)$$

*then, for all  $\ell, n \in \mathbb{N}^*$ ,*

$$|\mathbb{E}[\mathbf{1}_{\{X^{h_{\ell+\eta_n^\ell(\xi)}} > \xi\}} - \mathbf{1}_{\{X^{h_{\ell+\lceil \theta \ell \rceil}} > \xi\}}]| \leq C_1 h_\ell^{1+\theta} u_n,$$

*where*

$$C_1 = \frac{h_0^{-\frac{2p_\star}{p_\star+2}} h_1^{-p_\star(\frac{1}{r}-\frac{1}{2})}}{M^{p_\star(\frac{1}{r}-\frac{1}{2})} - 1} \sqrt{2} C_{\text{ad}}^{-p_\star} B_{p_\star} \mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{p_\star}]^{\frac{1}{p_\star}},$$

*with  $B_{p_\star}$  being a positive constant that depends only on  $p_\star$ .*

b. *If (2.24) is satisfied for some  $\mathfrak{g} > 0$ , with*

$$h_0 \geq \left( \frac{8\mathfrak{g}}{C_{\text{ad}}^2} \right)^{\frac{r}{2-r}}, \quad 1 < r \leq 2, \quad \text{and} \quad 0 < \theta \leq 1, \quad (4.5)$$

*then, for all  $\ell, n \in \mathbb{N}^*$ ,*

$$|\mathbb{E}[\mathbf{1}_{\{X^{h_{\ell+\eta_n^\ell(\xi)}} > \xi\}} - \mathbf{1}_{\{X^{h_{\ell+\lceil \theta \ell \rceil}} > \xi\}}]| \leq C_2 h_\ell^{1+\theta} \gamma_n,$$

*where*

$$C_2 = \frac{2}{\gamma_1(1 - M^{-1})}.$$

(ii) *If (3.9) is satisfied with, for some  $v_0 > 0$ ,*

$$\sup_{0 \leq h < h' \in \overline{\mathcal{H}}} \mathbb{E}[\exp(v_0 |G_h^{h'}|^2)] < \infty, \quad h_0 \geq \left( \frac{2}{v_0 C_{\text{ad}}^2} \right)^{\frac{r}{2-r}}, \quad (4.6)$$

$$1 < r \leq 2, \quad \text{and} \quad 0 < \theta \leq 1,$$

*then, for all  $\ell, n \in \mathbb{N}^*$ ,*

$$|\mathbb{E}[\mathbf{1}_{\{X^{h_{\ell+\eta_n^\ell(\xi)}} > \xi\}} - \mathbf{1}_{\{X^{h_{\ell+\lceil \theta \ell \rceil}} > \xi\}}]| \leq C_3 h_\ell^{1+\theta} \gamma_n,$$

*where*

$$C_3 = \frac{\sup_{0 \leq h < h' \in \overline{\mathcal{H}}} \mathbb{E}[\exp(v_0 |G_h^{h'}|^2)]}{\gamma_1(1 - M^{-1})}.$$

**Remark 4.3.** (i) We refer to Remark 2.9(i) for an elaboration on the value of  $B_{p_\star}$  in Lemma 4.2(i)a.

(ii) Unlike [31, Lemma 3.9], the bias controls above display a dependency on  $n$  that decays roughly in the order of the step size  $\gamma_n$ . This property will prove useful to control the numerical error induced by the refinement strategy (3.10) in the adaptive nested scheme (4.1).

- (iii) Using the special values of  $h = 0$  and  $h' = 1$  within the first postulate of the framework (4.6) yields that the quantity

$$\mathbb{E} \left[ \exp \left( v_0 (\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y])^2 \right) \right]$$

is bounded. As discussed in Remark 2.9(ii), this entails a Gaussian concentration of the random variable  $\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]$ .

Define

$$\tilde{\gamma}_n^\ell = \begin{cases} (u_n h_\ell^{1+\theta}) \vee \gamma_n & \text{if (4.4) holds,} \\ \gamma_n & \text{if (4.5) or (4.6) holds,} \end{cases} \quad \ell, n \in \mathbb{N}^*. \quad (4.7)$$

Recalling the definition of  $(\lambda_q)_{q \geq 1}$  in (2.13), the next lemma provides an adaptive counterpart to Lemma 2.5(ii).

**Proposition 4.4.** *Let  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\gamma_1 > 0$  and  $\beta \in (0, 1]$ . Suppose that Assumption 2.3 holds.*

- (i) *If  $\bar{\lambda}_1 \gamma_1 > 1$  when  $\beta = 1$ , and*

$$\sup_{\ell \geq 0} \mathbb{E} \left[ |\tilde{\xi}_0^{h_\ell}|^2 \exp \left( \frac{2k_\alpha}{1-\alpha} \sup_{\ell' \geq 1} \|f_{X^{h_{\ell'}}}\|_\infty |\tilde{\xi}_0^{h_\ell}| \right) \right] < \infty,$$

where  $k_\alpha = 1 \vee \frac{\alpha}{1-\alpha}$ , then, for all  $\ell, n \in \mathbb{N}^*$ ,

$$\mathbb{E}[(\tilde{\xi}_n^{h_\ell} - \xi_\star^{h_{\ell+\lceil \theta \ell \rceil}})^2] \leq C^\beta \tilde{\gamma}_n^\ell, \quad \text{where } C^\beta = \sup_{\ell \geq 1} C^{\ell, \beta}, \quad (4.8)$$

with the  $(C^{\ell, \beta})_{\ell \geq 1}$  being the positive constants given in (B.47).

- (ii) *If  $\bar{\lambda}_2 \gamma_1 > 2$  when  $\beta = 1$ , and*

$$\sup_{\ell \geq 0} \mathbb{E} \left[ |\tilde{\xi}_0^{h_\ell}|^4 \exp \left( \frac{4k_\alpha}{1-\alpha} \sup_{\ell' \geq 1} \|f_{X^{h_{\ell'}}}\|_\infty |\tilde{\xi}_0^{h_\ell}| \right) \right] < \infty,$$

where  $k_\alpha = 1 \vee \frac{\alpha}{1-\alpha}$ , then, for all  $\ell, n \in \mathbb{N}^*$ ,

$$\mathbb{E}[(\tilde{\xi}_n^{h_\ell} - \xi_\star^{h_{\ell+\lceil \theta \ell \rceil}})^4] \leq \bar{C}^\beta (\tilde{\gamma}_n^\ell)^2, \quad \text{where } \bar{C}^\beta = \sup_{\ell \geq 1} \bar{C}^{\ell, \beta}, \quad (4.9)$$

with the  $(\bar{C}^{\ell, \beta})_{\ell \geq 1}$  being the positive constants given in (B.57).

*Remark 4.5.* The above results are comparable to the non-adaptive case that is dealt with in Lemma 2.5(ii). The  $L^4(\mathbb{P})$ -control is necessary to our study of the adaptive MLSA scheme in Section 5.

## 4.2 Complexity Analysis

Proposition 4.4(i) provides an insight into the behavior of adNSA. Fixing  $\ell \in \mathbb{N}^*$  and running (4.1)  $n \in \mathbb{N}^*$  times results in a global error

$$\tilde{\xi}_n^{h_\ell} - \xi_\star^0 = (\tilde{\xi}_n^{h_\ell} - \xi_\star^{h_{\ell+\lceil \theta \ell \rceil}}) + (\xi_\star^{h_{\ell+\lceil \theta \ell \rceil}} - \xi_\star^0),$$

where the first term represents the statistical error and the second the bias error. The proofs of the following results are available in Appendix D.

**Proposition 4.6.** *Suppose that Assumption 2.3(i) holds. Let  $\varepsilon > 0$  be a fixed prescribed accuracy, and let  $C > 0$  denote the constant described in (2.14). If  $h_0 > C^{-1}\varepsilon$ , then, setting*

$$\ell = \left\lceil \frac{\ln(Ch_0\varepsilon^{-1})}{(1+\theta)\ln M} \right\rceil \geq 1 \quad (4.10)$$

*achieves a bias error for adNSA of order  $\varepsilon$ .*

*Remark 4.7.* According to the decomposition (2.11), when using NSA with a bias  $h_\ell$  on the geometric scale  $\mathcal{H}_0 = \{h_{\ell'}, \ell' \in \mathbb{N}\}$ , one must set

$$\ell = \left\lceil \frac{\ln(Ch_0\varepsilon^{-1})}{\ln M} \right\rceil \geq 1$$

to achieve a bias error of order  $\varepsilon$ . adNSA requires a smaller  $\ell$ , which in turn translates into a less expensive, albeit biased, simulation  $X^{h_\ell}$  that is then refined. This adjustment contributes to a reduction in the complexity of adNSA.

The next lemma quantifies the average simulation amounts performed per iteration under the adaptive strategy.

**Lemma 4.8.** *Under Assumptions 2.3(ii)-(iv), by setting  $r > 1$ , for all  $\ell, n \in \mathbb{N}^*$ ,*

$$\mathbb{E}[h_{\ell+\tilde{\eta}_n}^{-1}] \leq C \begin{cases} h_\ell^{-1} n^{\frac{\delta}{p^*}} & \text{if (2.23) holds,} \\ h_\ell^{-1} (\ell + \ln n)^{\frac{1}{2}} & \text{if (2.24) or (3.9) holds,} \end{cases}$$

where  $C > 0$  is described in (D.3).

*Remark 4.9.* Note the analogous [31, Proposition 3.1] for adaptive Monte Carlo, independent nonetheless of the iteration number  $n$ . The average amount of inner simulations therein is in the order of  $h_\ell^{-1}$ . The saturation of our adaptive strategy for  $n$  large is responsible for expanding these amounts by a small order depending on  $n$ . We refer to Remark 4.1 for further comments on this dependence on  $n$ .

**Proposition 4.10.** *Suppose that Assumptions 2.3(ii)-(iv) hold. By setting  $r > 1$ , the average complexity of adNSA satisfies*

$$\text{Cost}_{\text{adNSA}} \leq C \begin{cases} n^{1+\frac{\delta}{p^*}} h_\ell^{-1} & \text{if (2.23) holds,} \\ n(\ell^{\frac{1}{2}} + \ln^{\frac{1}{2}} n) h_\ell^{-1} & \text{if (2.24) or (3.9) holds,} \end{cases}$$

where  $C > 0$  is the constant described in (D.4).

*Remark 4.11.* Absent the saturation factor, the complexity would simply be in  $O(nh_\ell^{-1})$ , similar to the complexity order of NSA, as per Proposition 2.7.

**Theorem 4.12.** *Let  $\varepsilon > 0$  be a prescribed accuracy. Assume that  $h_0 > (C')^{-1}\varepsilon$ , where  $C' > 0$  is the constant described in (2.14), and that  $\bar{\lambda}_1\gamma_1 > 1$  if  $\beta = 1$ , and set  $\ell$  as in (4.10). Then, denoting  $C > 0$  the constant defined in (D.5), by taking*

$$n = \begin{cases} \lceil C^{\frac{2}{\delta}} \varepsilon^{-\frac{1}{\delta}} \rceil & \text{if (4.4) holds with } \delta \leq \frac{\beta}{2}, \\ \lceil C^{\frac{2}{\beta}} \varepsilon^{-\frac{2}{\beta}} \rceil & \text{if (4.4) holds with } \delta \geq \frac{\beta}{2}, \text{ or if (4.5) or (4.6) holds,} \end{cases}$$

*the statistical  $L^2(\mathbb{P})$ -error for adNSA is of order  $\varepsilon$  as  $\varepsilon \downarrow 0$ .*

*Under this setting, by taking  $r > 1$ , the computational cost of adNSA satisfies*

$$\text{Cost}_{\text{adNSA}} \leq \bar{C} \begin{cases} \varepsilon^{-\frac{p^*+\delta}{p^*\delta} - \frac{1}{1+\theta}} & \text{if (4.4) holds with } \delta \leq \frac{\beta}{2}, \\ \varepsilon^{-\frac{2(\delta+p^*)}{\beta p^*} - \frac{1}{1+\theta}} & \text{if (4.4) holds with } \delta \geq \frac{\beta}{2}, \\ \varepsilon^{-\frac{2}{\beta} - \frac{1}{1+\theta}} |\ln \varepsilon|^{\frac{1}{2}} & \text{if (4.5) or (4.6) holds,} \end{cases} \quad \text{as } \varepsilon \downarrow 0,$$

where  $\bar{C} > 0$  is the constant defined in (D.9). The minimal computational cost, which is attained under the constraint  $\bar{\lambda}_1 \gamma_1 > 1$ , satisfies

$$\text{Cost}_{\text{adNSA}} \leq \bar{C} \begin{cases} \varepsilon^{-\frac{5p_*+4}{2p_*}} & \text{if (4.4) holds, } \beta = 1, \delta = \frac{\beta}{2} = \frac{1}{2}, \text{ and } \theta = \frac{p_*-2}{p_*+2}, \\ \varepsilon^{-\frac{5}{2}} |\ln \varepsilon|^{\frac{1}{2}} & \text{if (4.5) or (4.6) holds, and } \beta = \theta = 1, \end{cases} \quad \text{as } \varepsilon \downarrow 0.$$

*Remark 4.13.* (i) Within the framework of Assumption (4.4), adNSA achieves a speed-up over NSA when  $p_* > 4$ . By contrast, heavier-tailed distributions, corresponding to an integrability exponent  $p_* \leq 4$ , do not appear to benefit from adaptive refinement. This may be explained by the fact that adaptivity is primarily sensitive to higher-order moments.

(ii) By comparison with Proposition 2.7, the complexities outlined in Theorem 4.12 are an order of magnitude of  $\varepsilon^{\frac{1}{2}}$  lower than those of NSA for large  $p_*$  under the framework (4.4) and up to a logarithmic factor under the frameworks (4.5) and (4.6). As we will see in the next section, adNSA serves as a fundamental component of the adaptive MLSA scheme. Consequently, the performance improvements noted on the former scheme should result in an acceleration of the latter.

The table below recaps the differences between NSA and adNSA.

Algorithm	NSA	adNSA
Bias parameter	$h \sim \varepsilon$	$h_\ell \sim \varepsilon^{\frac{1}{1+\theta}} \Leftrightarrow \ell \sim \frac{ \ln \varepsilon }{(1+\theta) \ln M}$
Iteration amount	$n \sim \varepsilon^{-\frac{2}{\beta}}$	$n \sim \begin{cases} \varepsilon^{-\frac{1}{\delta}} & \text{if (4.4) holds with } \delta \leq \frac{\beta}{2}, \\ \varepsilon^{-\frac{2}{\beta}} & \text{if (4.4) holds with } \delta \geq \frac{\beta}{2}, \\ & \text{or if (4.5) or (4.6) holds,} \end{cases}$
Best complexity	$\text{Cost}_{\text{NSA}} = O(\varepsilon^{-3})$	$\text{Cost}_{\text{adNSA}} = \begin{cases} O(\varepsilon^{-\frac{5p_*+4}{2p_*}}) & \text{if (4.4) holds,} \\ O(\varepsilon^{-\frac{5}{2}}  \ln \varepsilon ^{\frac{1}{2}}) & \text{if (4.5) or (4.6) holds.} \end{cases}$

**Table 1.** Comparison of NSA and adNSA.  $\varepsilon_0$  designates the prescribed accuracy. The statements refer to Propositions 2.7 and 4.12.

## 5 Adaptive Multilevel Stochastic Approximation Algorithm

Recalling that MLSA telescopes multiple paired NSA schemes, the previous development on adNSA provides a framework for the extension of the adaptive refinement strategy to the multilevel paradigm. We define the adaptive multilevel SA (adMLSAs) estimator for the VaR as

$$\tilde{\xi}_{\mathbf{N}}^{\text{ML}} = \xi_{N_0}^{h_0} + \sum_{\ell=1}^L \tilde{\xi}_{N_\ell}^{h_\ell} - \tilde{\xi}_{N_\ell}^{h_{\ell-1}}, \quad (5.1)$$

where  $\mathbf{N} := (N_0, \dots, N_L) \in (\mathbb{N}^*)^{L+1}$  represents the number of iterations at each level. Each level  $\ell \in \llbracket 0, L \rrbracket$  is simulated independently. As detailed in Remark 3.3(ii), the level 0 estimator is not refined, resulting in  $N_0$  iterations of NSA. Each of the remaining levels  $\ell \in \llbracket 1, L \rrbracket$  is obtained as follows: after initializing  $(\tilde{\xi}_0^{h_{\ell-1}}, \tilde{\xi}_0^{h_\ell})$ , for each  $n = 0, \dots, N_\ell - 1$ , once the components of  $(X_{n+1}^{h_{\ell-1}}, X_{n+1}^{h_\ell})$  have been simulated according to (2.22),  $X_{n+1}^{h_{\ell-1}}$  and  $X_{n+1}^{h_\ell}$  are separately refined as in (4.2), relative to the fine and coarse iterates  $\tilde{\xi}_n^{h_{\ell-1}}$  and  $\tilde{\xi}_n^{h_\ell}$ , into  $\tilde{X}_{n+1}^{h_{\ell-1}}$  and  $\tilde{X}_{n+1}^{h_\ell}$  respectively. These refined samples are then injected into separate single updates of (4.1) to obtain the following iterates  $(\tilde{\xi}_{n+1}^{h_{\ell-1}}, \tilde{\xi}_{n+1}^{h_\ell})$ .

---

**Algorithm 3** Adaptive Multilevel SA for estimating the VaR
 

---

**Require:** A number of levels  $L \geq 1$ , a bias parameter  $h_0 = \frac{1}{K} \in \mathcal{H}$ , an integer  $M \geq 2$ ,  $N_0, \dots, N_L \in \mathbb{N}^*$ , a positive non-increasing sequence  $(\gamma_n)_{n \geq 1}$  such that  $\sum_{n=1}^{\infty} \gamma_n = \infty$  and  $\lim_{n \rightarrow \infty} \gamma_n = 0$ , refinement parameters  $C_{\text{ad}} > 0$ ,  $r > 1$  and  $0 < \theta \leq 1$

- 1: Sample  $\xi_0^{h_0}$  randomly
- 2:  $\xi_{N_0}^{h_0} \leftarrow \text{NSA}(h_0, N_0)$
- 3: **for**  $\ell = 1 \dots L$  **do**
- 4:   Set  $h_\ell \leftarrow \frac{h_0}{M^\ell}$
- 5:   Sample  $(\tilde{\xi}_0^{h_{\ell-1}}, \tilde{\xi}_0^{h_\ell})$  randomly
- 6:   **for**  $n = 0 \dots N_\ell - 1$  **do**
- 7:     Simulate  $Y_{n+1} \sim Y$  and  $Z_{n+1}^1, \dots, Z_{n+1}^{KM^\ell} \stackrel{\text{i.i.d.}}{\sim} Z$  independently of  $Y_{n+1}$
- 8:      $X_{n+1}^{h_{\ell-1}} \leftarrow \frac{1}{KM^{\ell-1}} \sum_{k=1}^{KM^{\ell-1}} \varphi(Y_{n+1}, Z_{n+1}^k)$
- 9:      $X_{n+1}^{h_\ell} \leftarrow \frac{1}{M} X_{n+1}^{h_{\ell-1}} + \frac{1}{KM^\ell} \sum_{k=KM^{\ell-1}+1}^{KM^\ell} \varphi(Y_{n+1}, Z_{n+1}^k)$
- 10:     **for**  $j = \ell - 1 \dots \ell$  **do**
- 11:       $\tilde{X}_{n+1}^{h_j} \leftarrow \text{Refine}_{C_{\text{ad}}, r, \theta, j, n}(X_{n+1}^{h_j}, \tilde{\xi}_n^{h_j})$
- 12:       $\tilde{\xi}_{n+1}^{h_j} \leftarrow \tilde{\xi}_n^{h_j} - \gamma_{n+1} H(\tilde{\xi}_n^{h_j}, \tilde{X}_{n+1}^{h_j})$
- 13:     **end for**
- 14:   **end for**
- 15: **end for**
- 16:  $\tilde{\xi}_{\mathbf{N}}^{\text{ML}} \leftarrow \xi_{N_0}^{h_0} + \sum_{\ell=1}^L \tilde{\xi}_{N_\ell}^{h_\ell} - \tilde{\xi}_{N_\ell}^{h_{\ell-1}}$
- 17: **return**  $\tilde{\xi}_{\mathbf{N}}^{\text{ML}}$

---

*Remark 5.1.* In line 5 of Algorithm 3, the initializations  $\tilde{\xi}_0^{h_{\ell-1}}$  and  $\tilde{\xi}_0^{h_\ell}$  may be set equal, or even be set deterministically.

## 5.1 Convergence Analysis

The proofs of the following results are postponed to Appendix B.

The next result guarantees stronger error controls for adMLSA comparatively with its non-adaptive counterpart, Lemma 2.8, for MLSA.

**Lemma 5.2.** *Let  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\gamma_1 > 0$  and  $\beta \in (0, 1]$ .*

(i) *Assume that the  $X^h$  admit pdf  $f_{X^h}$  that are bounded uniformly in  $h \in \overline{\mathcal{H}}$ .*

a. *If (2.23) is satisfied, with*

$$1 < r < 2, \quad \text{and} \quad 0 < \theta \leq \frac{p_\star}{p_\star + 2}, \quad (5.2)$$

*then, for all  $\ell, n \in \mathbb{N}^*$ ,*

$$\mathbb{E}[|\mathbf{1}_{\{X^{h_{\ell+n_n^\ell(\xi)}} > \xi\}} - \mathbf{1}_{\{X^0 > \xi\}}|] \leq C_1 h_\ell^{(1+\theta)\frac{p_\star}{2(p_\star+1)}},$$

*where*

$$\begin{aligned} C_1 = & h_0^{-\frac{p_\star}{2(p_\star+1)}} B_{p_\star} \mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{p_\star}]^{\frac{1}{p_\star+1}} \left( \sup_{h \in \overline{\mathcal{H}}} \|f_{X^h}\|_\infty \right)^{\frac{p_\star}{p_\star+1}} \\ & + \frac{\gamma_1 h_0^{-\frac{2(p_\star+1)}{p_\star+2}} h_1^{-p_\star(\frac{1}{r}-\frac{1}{2})}}{M^{p_\star(\frac{1}{r}-\frac{1}{2})} - 1} \sqrt{2} C_{\text{ad}}^{-p_\star} B_{p_\star} \mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{p_\star}]^{\frac{1}{p_\star}}, \end{aligned}$$

*with  $B_{p_\star}$  being a positive constant that depends only on  $p_\star$ .*

b. If (2.24) is satisfied for  $\mathfrak{g} > 0$ , with

$$h_0 \geq \left( \frac{4\mathfrak{g}}{C_{\text{ad}}^2} \right)^{\frac{r}{2-r}}, \quad 1 < r \leq 2, \quad \text{and} \quad 0 < \theta \leq 1, \quad (5.3)$$

then, for all  $\ell, n \in \mathbb{N}^*$ ,

$$\mathbb{E}[|\mathbb{1}_{\{X^{h_{\ell+n}}(\xi) > \xi\}} - \mathbb{1}_{\{X^0 > \xi\}}|] \leq C_2 h_\ell^{\frac{1+\theta}{2}} |\ln h_\ell|^{\frac{1}{2}},$$

where

$$C_2 = 4\mathfrak{g}^{\frac{1}{2}} h_0^{-\frac{1}{2}} \left( 1 + \left( \sup_{h \in \overline{\mathcal{H}}} \|f_{X^h}\|_\infty \right) \left( 2(2 \vee |\ln(\mathfrak{g} h_0^{-1})|) \right)^{\frac{1}{2}} \right) + \frac{2}{1 - M^{-\frac{1}{2}}}.$$

(ii) If (3.9) is satisfied, with, for some  $v_0 > 0$ ,

$$\sup_{h \in \overline{\mathcal{H}}} \mathbb{E}[\exp(v_0 |G_0^h|^2)] < \infty, \quad h_0 \geq \left( \frac{1}{v_0 C_{\text{ad}}^2} \right)^{\frac{r}{2-r}}, \quad (5.4)$$

$$1 < r \leq 2, \quad \text{and} \quad 0 < \theta \leq 1,$$

then, for all  $\ell, n \in \mathbb{N}^*$ ,

$$\mathbb{E}[|\mathbb{1}_{\{X^{h_{\ell+n}}(\xi) > \xi\}} - \mathbb{1}_{\{X^0 > \xi\}}|] \leq C_3 h_\ell^{\frac{1+\theta}{2}},$$

where

$$C_3 = h_0^{-\frac{1}{2}} \sup_{0 \leq h_1 < h_2 \in \overline{\mathcal{H}}} \mathbb{E}[K_{h_1}^{h_2} |G_{h_1}^{h_2}|] + \frac{\sup_{h \in \overline{\mathcal{H}}} \mathbb{E}[\exp(v_0 |G_0^h|^2)]}{1 - M^{-\frac{1}{2}}}.$$

*Remark 5.3.* (i) Although the frameworks of Lemmas 4.2 and 5.2 are similar, those of Lemma 4.2 are stronger.

(ii) The value of  $B_{p_\star}$  in Lemma 5.2(i)a is discussed in Remark 2.9(i).

(iii) In comparison with the non-adaptive frameworks of Lemma 2.8, the variance controls for adMLSA display an extra exponent factor of  $1+\theta$ , which should bring down the algorithm's statistical error faster to 0.

The main convergence result follows. Its proof is deferred to Appendix C.

**Theorem 5.4.** Suppose that  $\varphi(Y, Z) \in L^2(\mathbb{P})$ , that Assumptions 2.3 and 2.10 hold, and that

$$\sup_{\ell \geq 0} \mathbb{E} \left[ |\tilde{\xi}_0^{h_\ell}|^4 \exp \left( \frac{4k_\alpha}{1-\alpha} \sup_{\ell' \geq 1} \|f_{X^{h_{\ell'}}}\|_\infty |\tilde{\xi}_0^{h_{\ell'}}| \right) \right] < \infty,$$

where  $k_\alpha = 1 \vee \frac{\alpha}{1-\alpha}$ . If  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\gamma_1 > 0$  and  $\beta \in (0, 1]$ , and if  $\bar{\lambda}_2 \gamma_1 > 2$  when  $\beta = 1$ , then, for all  $L \in \mathbb{N}^*$  and all  $\mathbf{N} = (N_0, \dots, N_L) \in (\mathbb{N}^*)^{L+1}$ ,

$$\mathbb{E}[(\tilde{\xi}_{\mathbf{N}}^{\text{ML}} - \xi_\star^{h_{L+[\theta L]}})^2] \leq C \left( \gamma_{N_0} + \left( \sum_{\ell=1}^L \tilde{\gamma}_{N_\ell}^\ell \right)^2 + \sum_{\ell=1}^L \gamma_{N_\ell} (\tilde{\gamma}_{N_\ell}^\ell)^{\frac{1}{2}} + \sum_{\ell=1}^L \gamma_{N_\ell} \epsilon(h_\ell)^{1+\theta} \right), \quad (5.5)$$

where  $C > 0$  is the constant described in (C.13),  $(\tilde{\gamma}_n^\ell)_{\ell, n \geq 1}$  are defined in (4.7), and  $\epsilon$  is defined as in (2.27):

$$\epsilon(h) := \begin{cases} h^{\frac{p_\star}{2(1+p_\star)}} & \text{if (4.4) holds,} \\ h^{\frac{1}{2}} |\ln h|^{\frac{1}{2}} & \text{if (4.5) holds,} \\ h^{\frac{1}{2}} & \text{if (4.6) holds,} \end{cases} \quad h \in \mathcal{H}.$$

*Remark 5.5.* Similarly to Theorem 2.12, the upper  $L^2(\mathbb{P})$  control for the statistical error of adMLSA is comprised four terms: the first term governs the level 0 simulation and the remaining three govern the drifts and martingales arising from the multilevel linearization (C.7). The first difference with MLSA lies in the inclusion of  $\tilde{\gamma}_n^\ell$ , which emanates from the controls of Proposition 4.4. Additionally, the final term exhibits an extra exponent factor of  $1 + \theta$ , resulting in an accelerated convergence rate compared to MLSA.

The convergence rate speed-up, shown in the previous theorem, translates a significant performance gain that should reflect in the complexity of adMLSA, as we clarify next.

## 5.2 Complexity Analysis

Approximating  $\xi_\star^0$  by  $\tilde{\xi}_\mathbf{N}^{\text{ML}}$  results in a global error that decomposes into statistical and bias errors as such:

$$\tilde{\xi}_\mathbf{N}^{\text{ML}} - \xi_\star^0 = (\tilde{\xi}_\mathbf{N}^{\text{ML}} - \xi_\star^{h_{L+\lceil\theta L\rceil}}) + (\xi_\star^{h_{L+\lceil\theta L\rceil}} - \xi_\star^0). \quad (5.6)$$

The proofs related to this section are deferred to Appendix D.

**Proposition 5.6.** *Suppose that Assumption 2.3(i) holds. Let  $\varepsilon > 0$  be a prescribed accuracy and  $C > 0$  the constant described in (2.14). If  $h_0 > C^{-1}\varepsilon$ , then, setting the number of levels to*

$$L = \left\lceil \frac{\ln(C h_0 \varepsilon^{-1})}{(1 + \theta) \ln M} \right\rceil \geq 1 \quad (5.7)$$

*achieves a bias error for adMLSA of order  $\varepsilon$ .*

*Remark 5.7.* In view of Propositions 5.6 and 2.13(i), to achieve a comparable bias error order, adMLSA requires significantly less levels than MLSA. In fact, for  $\theta = 1$ , adMLSA requires half as many levels as MLSA to achieve a bias error of order  $\varepsilon$ . This property alone makes adMLSA much faster than MLSA. It is linked to the fact that, for a given number of levels  $L$ , adMLSA achieves a bias error of order  $h_L^{1+\theta}$ , an order of magnitude  $h_L^\theta$  lower than MLSA, which achieves a bias error of order  $h_L$ .

**Proposition 5.8.** *Suppose that Assumptions 2.3(ii)-(iv) hold. By setting  $r > 1$ , the average complexity of adMLSA satisfies*

$$\text{Cost}_{\text{adMLSA}} \leq C \begin{cases} \sum_{\ell=0}^L N_\ell^{1+\frac{\delta}{p_\star}} h_\ell^{-1} & \text{if (2.23) holds,} \\ \sum_{\ell=0}^L N_\ell (L^{\frac{1}{2}} + \ln^{\frac{1}{2}} N_\ell) h_\ell^{-1} & \text{if (2.24) or (3.9) holds,} \end{cases} \quad (5.8)$$

where  $C > 0$  is the constant given in (D.10).

*Remark 5.9.* Without a saturation factor in the refinement strategy, the computational cost would be lowered to  $O(\sum_{\ell=0}^L N_\ell h_\ell^{-1})$ , similar to both MLSA and the adaptive MLMC algorithm [31]. However, omitting this factor would adversely affect the convergence of the statistical error (5.5). See Remark 4.1 for further comments on the necessity to saturate the refinements.

**Theorem 5.10.** *Let  $C_{\text{ad}} > 0$ ,  $r > 1$  and  $0 < \theta \leq 1$ , and suppose that Assumptions 2.3 and 2.10 hold. Let  $\varepsilon > 0$  be a prescribed accuracy. Denote  $C, \bar{C} > 0$  the constants on the right-hand sides of (5.5) and (5.8) respectively.*

(i) *If (4.4) holds, then, by setting*

$$N_\ell = \left\lceil C_1^{\beta, \delta} \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{(2\beta-(1+\theta)p_\star+(2\beta-(1+\theta)\delta)}{2(p_\star+1)((1+\beta)p_\star+\delta)} p_\star} \right)^{\frac{1}{\beta}} h_\ell^{\frac{(3+\theta)p_\star+2}{2(p_\star+1)((1+\beta)p_\star+\delta)} p_\star} \right\rceil, \quad \ell \in \llbracket 0, L \rrbracket, \quad (5.9)$$

where

$$C_1^{\beta, \delta} = C^{\frac{p_\star}{(1+\beta)p_\star + \delta}} (\bar{C}\gamma_1)^{\frac{(p_\star + \delta)p_\star}{(1+\beta)p_\star + \delta}} \\ \times \left( \left( \frac{\beta p_\star}{p_\star + \delta} \right)^{\frac{p_\star + \delta}{(1+\beta)p_\star + \delta}} + \left( \frac{\beta p_\star}{p_\star + \delta} \right)^{-\frac{\beta p_\star}{(1+\beta)p_\star + \delta}} \right)^{\frac{1}{\beta}} \left( \frac{p_\star + \delta}{(1+\beta)p_\star + \delta} \right)^{\frac{1}{\beta}},$$

adMLSA scores a global  $L^2(\mathbb{P})$ -error of order  $\varepsilon$ . By taking  $L$  as in (5.7), the corresponding average complexity satisfies

$$\text{Cost}_{adMLSA} \leq \widehat{C}_1 \begin{cases} \varepsilon^{-\frac{2(p_\star + \delta)}{\beta p_\star}} & \text{if } \theta > \frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta}, \\ \varepsilon^{-\frac{2(p_\star + \delta)}{\beta p_\star}} |\ln \varepsilon|^{\frac{(1+\beta)p_\star + \delta}{\beta p_\star}} & \text{if } \theta = \frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta}, \text{ as } \varepsilon \downarrow 0, \\ \varepsilon^{-\frac{2(p_\star + \delta)}{\beta p_\star} - \frac{(2\beta - (1+\theta))p_\star + 2\beta - (1+\theta)\delta}{(1+\theta)\beta(p_\star + 1)}} & \text{if } \theta < \frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta}, \end{cases}$$

where  $\widehat{C}_1 > 0$  is the constant on the right hand side of (D.19). The best complexity is attained if  $\beta = 1$ ,  $\delta \downarrow 0$ , and  $\theta = \frac{p_\star - 2}{p_\star + 2}$ :

$$\inf_{\beta, \delta, \theta} \text{Cost}_{adMLSA} \leq \widehat{C}_1 \varepsilon^{-2 - \frac{3p_\star + 2}{2p_\star(p_\star + 1)} \iota}, \quad \text{for any } \iota > 0, \quad \text{as } \varepsilon \downarrow 0.$$

(ii) If (4.5) holds, then, taking  $L$  as in (5.7) and

$$N_\ell = \left[ (\bar{C}\gamma_1)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{2\beta - (1+\theta)}{2(1+\beta)}} |\ln h_{\ell'}|^{\frac{1+\theta}{2(1+\beta)}} \right)^{\frac{1}{\beta}} h_\ell^{\frac{3+\theta}{2(1+\beta)}} |\ln h_\ell|^{\frac{1+\theta}{2(1+\beta)}} \right], \quad \ell \in \llbracket 0, L \rrbracket,$$

leads to a global  $L^2(\mathbb{P})$  error for adMLSA of order  $\varepsilon$  as  $\varepsilon \downarrow 0$ . The corresponding average complexity satisfies

$$\text{Cost}_{adMLSA} \leq \widehat{C}_2 \begin{cases} \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} & \text{if } \theta > 2\beta - 1, \\ \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{3(1+\beta)+\theta}{2\beta}} & \text{if } \theta = 2\beta - 1, \text{ as } \varepsilon \downarrow 0, \\ \varepsilon^{-\frac{3(1+\theta)+2\beta}{2(1+\theta)\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} & \text{if } \theta < 2\beta - 1, \end{cases}$$

where  $\widehat{C}_2 > 0$  is the constant on the right hand side of (D.38). This complexity is minimized for  $\beta = \theta = 1$ , in which case

$$\inf_{\beta, \theta} \text{Cost}_{adMLSA} \leq \widehat{C}_2 \varepsilon^{-2} |\ln \varepsilon|^{\frac{7}{2}}, \quad \text{as } \varepsilon \downarrow 0.$$

(iii) If (4.6) holds, then, setting  $L$  as in (5.7) and

$$N_\ell = \left[ (\bar{C}\gamma_1)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{2\beta - (1+\theta)}{2(1+\beta)}} \right)^{\frac{1}{\beta}} h_\ell^{\frac{3+\theta}{2(1+\beta)}} \right], \quad \ell \in \llbracket 0, L \rrbracket,$$

gives a global  $L^2(\mathbb{P})$  error for adMLSA of order  $\varepsilon$  as  $\varepsilon \downarrow 0$ . The corresponding average complexity of

$$\text{Cost}_{adMLSA} \leq \widehat{C}_3 \begin{cases} \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} & \text{if } \theta > 2\beta - 1, \\ \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{3(1+\beta)+\theta}{2\beta}} & \text{if } \theta = 2\beta - 1, \text{ as } \varepsilon \downarrow 0, \\ \varepsilon^{-\frac{3(1+\theta)+2\beta}{2(1+\theta)\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} & \text{if } \theta < 2\beta - 1, \end{cases}$$

where  $\widehat{C}_3 > 0$  is the constant on the right hand side of (D.40). This complexity is minimized for  $\beta = \theta = 1$ , whereby

$$\inf_{\beta, \theta} \text{Cost}_{adMLSA} \leq \widehat{C}_3 \varepsilon^{-2} |\ln \varepsilon|^{\frac{5}{2}}, \quad \text{as } \varepsilon \downarrow 0.$$

*Remark 5.11.* (i) Within the frameworks of Theorems 5.10(ii)-(iii), adMLSA scores a significant performance gain over MLSA and retrieves the canonical multilevel performance [24, 21] of order  $\varepsilon^{-2}$ , up to a logarithmic factor.

(ii) As for the framework of Theorem 5.10(i), it is faster than its MLSA counterpart for any  $p_\star > 2$ . Better performance is expected by taking  $p_\star$  sufficiently large, recalling that a large class of portfolio payoffs are  $L^{p_\star}(\mathbb{P})$ -integrable for all  $p_\star > 2$ . Note that, asymptotically as  $p_\star \uparrow \infty$ ,  $\text{Cost}_{\text{adMLSA}} = O(\varepsilon^{-2-\iota})$ , for all  $\iota > 0$ , while  $\text{Cost}_{\text{MLSA}} = O(\varepsilon^{-\frac{5}{2}})$ , retrieving the same overperformance reported above.

The table below compares MLSA and adMLSA.

Algorithm	MLSA	adMLSA
$L^2(\mathbb{P})$ -control	$O(h_L^2 + \sum_{\ell=0}^L \gamma_{N_\ell} \varepsilon(h_\ell))$ (Decomposition (2.28), Lemma 2.5(i) and Theorem 2.12)	$O(h_L^{2(1+\theta)} + \sum_{\ell=0}^L \tilde{\gamma}_{N_\ell}^\ell \varepsilon(h_\ell)^{1+\theta})$ (Decomposition (5.6), Lemma 2.5(i) and Theorem 5.4)
Number of levels	$L \sim \frac{ \ln \varepsilon }{\ln M}$ (Proposition 2.13(i))	$L \sim \frac{ \ln \varepsilon }{(1+\theta) \ln M}$ (Proposition 5.6)
Best complexity	$\text{Cost}_{\text{MLSA}} =$ $\begin{cases} O(\varepsilon^{-3+\frac{p_\star}{2(1+p_\star)}}) & \text{if (2.23) holds,} \\ O(\varepsilon^{-\frac{5}{2}}  \ln \varepsilon ^{\frac{1}{2}}) & \text{if (2.24) holds,} \\ O(\varepsilon^{-\frac{5}{2}}) & \text{if (2.25) holds.} \end{cases}$ (Proposition 2.13(iii))	$\text{Cost}_{\text{adMLSA}} =$ $\begin{cases} O(\varepsilon^{-2-\frac{3p_\star+2}{2p_\star(p_\star+1)}-\iota}), \forall \iota > 0, & \text{if (4.4) holds,} \\ O(\varepsilon^{-2}  \ln \varepsilon ^{\frac{7}{2}}) & \text{if (4.5) holds,} \\ O(\varepsilon^{-2}  \ln \varepsilon ^{\frac{5}{2}}) & \text{if (4.6) holds.} \end{cases}$ (Theorem 5.10)

**Table 2.** Comparison of MLSA and adMLSA.  $\varepsilon > 0$  designates a prescribed accuracy.

## 6 Heuristics

### 6.1 Refinement Parameters

In view of Theorem 4.12 and Theorem 5.10, to compute the VaR efficiently, we need to set  $\theta = \frac{p_\star-2}{p_\star+2}$  within the framework (4.4) and  $\theta = 1$  within the frameworks (4.5) and (4.6). Note that  $\theta \approx 1$  for  $p_\star$  large enough within the framework (4.4).

The choice  $r = 1 + \frac{1}{\theta}$  together with the assumption  $\theta \approx 1$  lead to thresholds in (3.11) of the form

$$\psi_n^{\ell,k} = h_{\ell+k}^{\frac{\theta}{1+\theta}} w_n \approx h_{\ell+k}^{\frac{1}{2}} w_n,$$

where  $w_n$  is the saturation factor. Complementing [27], under this choice of  $r$  and  $\theta$ , our strategy can be seen as conducting consecutive Student t-tests on the null hypotheses “ $X^{h_{\ell+k}} = \xi$ ”,  $0 \leq k \leq \lceil \theta \ell \rceil$ , until the first rejection.

### 6.2 Refinement Strategy

Let us fix some level  $\ell$  and iteration number  $n$ , and denote  $\xi^\ell$  and  $\xi^{\ell-1}$  the corresponding fine and coarse iterates. For simplicity, we omit  $n$  from our notation. To increase efficiency, we propose to only refine the fine sample  $X^{h_\ell}$  to  $\tilde{X}^{h_\ell} = X^{h_{\ell+\eta^\ell(\xi^\ell)}}$ , then discard the  $Z$  simulations indexed

from  $KM^{\ell-1+\eta^\ell(\xi^\ell)} + 1$  onward to recover the coarse sample  $\tilde{X}^{h_{\ell-1}}$ :

$$\begin{aligned}\tilde{X}^{h_{\ell-1}} &= \frac{1}{KM^{\ell-1+\eta^\ell(\xi^\ell)}} \sum_{k=1}^{KM^{\ell-1+\eta^\ell(\xi^\ell)}} \varphi(Y, Z^k), \\ \tilde{X}^{h_\ell} &= \frac{1}{M} \tilde{X}^{h_{\ell-1}} + \frac{1}{KM^{\ell+\eta^\ell(\xi^\ell)}} \sum_{k=KM^{\ell-1+\eta^\ell(\xi^\ell)+1}^{KM^{\ell+\eta^\ell(\xi^\ell)}} \varphi(Y, Z^k),\end{aligned}$$

where  $(Z^k) \stackrel{\text{i.i.d.}}{\sim} Z$  are independent of  $Y$ . This amounts to setting  $\eta^{\ell-1}(\xi^{\ell-1}) = (\eta^\ell(\xi^\ell) - 2)^+$ . While this strategy for  $\eta^{\ell-1}$  is not centered on  $\xi^{\ell-1}$ , numerical trials have shown it to produce better results.

### 6.3 Confidence Constant

In the following, we tune the confidence constant  $C_{\text{ad}}$  appearing in (3.11) on a grid. Remark 3.3(v) however suggests a different treatment for this constant.

Denote  $\sigma = \sqrt{\text{Var}(\varphi(Y, Z)|Y)}$  the sampling standard deviation, and  $\sigma^h$ ,  $h = \frac{1}{K} \in \mathcal{H}$ , its empirical approximation given by

$$\sigma^h = \left( \frac{1}{K} \sum_{k=1}^K \varphi(Y, Z^k)^2 - (X^h)^2 \right)^{\frac{1}{2}}. \quad (6.1)$$

An alternate refinement strategy consists in incorporating an estimation of  $C_{\text{ad}}$  in (3.11) with the quantity  $C_p \sigma^{h_{\ell+k}}$  at the refinement step  $k$ , where  $C_p$  is set to 3 to retrieve a 99%-confidence on the closeness between  $X^{h_{\ell+k}}$  and  $X^0$ .

We term  $\sigma$ -adNSA and  $\sigma$ -adMLSA the adNSA and adMLSA versions where  $C_{\text{ad}}$  is estimated by  $(C_p \sigma^{h_{\ell+k}})_{0 \leq k \leq \lceil \theta \ell \rceil}$ . For comparative purposes, these versions are run as well and their performances are reported and discussed in subsequent analyses.

### 6.4 Unsaturated Refinement

To further reduce complexity, we propose to look into a desaturated variant of the refinement strategy (3.10), whereby the saturation factor is set to 1:

$$\eta^\ell(\xi) = \lceil \theta \ell \rceil \wedge \min \{k \in \llbracket 0, \lceil \theta \ell \rceil \rrbracket : |X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}} \psi^{\ell, k}\}, \quad \ell \in \mathbb{N}^*, \quad \xi \in \mathbb{R}, \quad (6.2)$$

with the convention  $\min \emptyset = +\infty$ , where  $C_{\text{ad}} > 0$ ,  $r > 1$ ,  $0 < \theta \leq 1$ , and

$$\psi^{\ell, k} = h_{\theta^\ell(r-1)+k}^{\frac{1}{r}}, \quad k \in \llbracket 0, \lceil \theta \ell \rceil \rrbracket, \quad \ell \in \mathbb{N}^*.$$

To be able to incorporate this refinement strategy in an adaptive multilevel scheme, we need the corresponding optimal iteration amounts. These unfortunately cannot be easily obtained for this heuristic. Given that the strategy (6.2) resembles that of [31, Eq (3.2)] for adaptive MLMC, one could simply consider the iteration amounts derived in [31]. The complexity theorems [31, Theorem 3.10 and 3.11] content with pointing to the best complexities, without providing the corresponding iteration amounts. These however may be found in the proof of the earlier work of [24, Theorem 3.1]. But adopting these iteration amounts remains problematic for a couple of reasons. First of all, as pointed out in Section 3, the saturation factor incorporated in the refinement strategy (3.10) has been worked out from the bias and variance upper estimates derived in Lemma 4.2 and Proposition 4.4. Substituting this factor with one, which is tantamount to unsaturation, leads to upper estimates that do not asymptotically vanish with large iteration amounts. Besides, the frameworks used by [24] or [31] are different from ours in (4.4), (4.5) and

(4.6). Considering all of these concerns, we opt for the prudent choice of using our own optimal iteration amounts calculated in Theorem 5.10 as obtained by factoring in a saturation effect. In other words, the unsaturated refinement strategy (6.2) is simply treated as a heuristic in the pursuit of faster execution.

The adaptive NSA and MLSA algorithms corresponding to the heuristic (6.2), that we dub u-adNSA and u-adMLSA, will therefore be run with the same iteration amounts  $n$  and  $(N_\ell)_{0 \leq \ell \leq L}$  delineated in Theorems 4.12 and 5.10.

## 7 Financial Case Studies

In the ensuing numerical studies, we illustrate the closure of the performance gap between MLSA and SA, which is made possible by using our adaptive refinement strategy. To that end, we revisit the VaR use cases already handled by SA, NSA and MLSA in [14, Sections 5 & 6]. The implementations for the case studies can be found at [github.com/azarlouzi/ada\\_mlsa](https://github.com/azarlouzi/ada_mlsa).

### 7.1 European Option

We succinctly recall here the setting of [14, Section 5]. We refer to the developments therein for a rigorous derivation of the ensuing statements.

Consider a European option of maturity  $T = 1$  and payoff  $\mathbb{R} \ni x \mapsto -x^2$ , on an underlying asset following a standard Brownian motion dynamic  $(W_t)_{0 \leq t \leq 1}$ . The risk-free rate is null and the pricing is performed under  $\mathbb{P}$ . The option's value at time  $t \in [0, 1]$  is

$$V_t = \mathbb{E}[-W_1^2 | W_t],$$

and its associated loss at a horizon  $\tau \in (0, 1)$  is

$$X^0 = V_0 - V_\tau.$$

We are interested in retrieving the VaR  $\xi_\star^0$  of this loss at some confidence level  $\alpha \in (0, 1)$ .

#### 7.1.1 Analytical and Simulation Formulas

Let  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$\varphi(y, z) := -(\sqrt{\tau}y + \sqrt{1-\tau}z)^2, \quad y, z \in \mathbb{R}.$$

On the one hand,

$$X^0 \stackrel{\mathcal{L}}{=} -1 - \mathbb{E}[\varphi(Y, Z) | Y] = \tau(Y^2 - 1),$$

where  $Y$  and  $Z$  are independent and of law  $\mathcal{N}(0, 1)$ .  $X^0$  can thus be simulated exactly, hence the benchmarking unbiased SA scheme [4] is applicable to estimate the VaR.

On the other hand, for a bias parameter  $h = \frac{1}{K} \in \mathcal{H}$ ,  $X^0$  can be approximated by

$$X^h = -1 - \frac{1}{K} \sum_{k=1}^K \varphi(Y, Z^k),$$

where  $Y, Z^1, \dots, Z^K \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ . We can then apply NSA, MLSA, adNSA,  $\sigma$ -adNSA, u-adNSA, adMLSA,  $\sigma$ -adMLSA and u-adMLSA on this basis to approximate the VaR.

Finally, the VaR  $\xi_\star^0$  at level  $\alpha$  has an analytical form:

$$\xi_\star^0 = \tau \left\{ \left( F^{-1} \left( \frac{1-\alpha}{2} \right) \right)^2 - 1 \right\}, \quad (7.1)$$

where  $F$  is the standard Gaussian cdf. Its evaluation will help assess the estimation errors of the aforementioned SA schemes.

### 7.1.2 Numerical Results

We conduct below a performance comparison of the different SA schemes discussed in this paper. We set the confidence level to  $\alpha = 97.5\%$  and the time horizon to  $\tau = 0.5$ , which yields  $\xi_\star^0 \approx 2.012$  via (7.1).

To minimize complexity, all algorithms are run with  $\beta = 1$  and their respective theoretical optimal iteration amounts. SA, NSA, adNSA and  $\sigma$ -adNSA are run with  $\gamma_n = \frac{1}{100+n}$ . MLSA, adMLSA,  $\sigma$ -adMLSA and u-adMLSA are run with  $M = 2$  under the framework (4.4), with  $\delta = 0.05$  and  $p_\star = 11$ . The parameter  $p_\star$  has been fine-tuned on a grid of values to minimize RMSE and execution time. We set  $\theta = \frac{p_\star-2}{p_\star+2}$  and  $r = 1 + \frac{1}{\theta}$  for the adaptive schemes, as recommended in Section 6. Identical  $h_0$  parametrization was applied to adNSA,  $\sigma$ -adNSA, MLSA, adMLSA,  $\sigma$ -adMLSA and u-adMLSA, and identical  $\gamma_n$  parametrization was applied to MLSA, adMLSA,  $\sigma$ -adMLSA and u-adMLSA. Further parametrizations of the adaptive and multilevel schemes, obtained via a grid search, are described in Table 3. The subsequent level amounts computed via (2.29), (4.10) and (5.7) are reported in columns  $L$ ,  $\ell_{\text{ad}}$  and  $L_{\text{ad}}$ . For adNSA and adMLSA, the iteration amounts scaling factor  $C$  (e.g.  $C$  represents  $C_1^{1,\delta}$  for adMLSA as per (5.9)) and the confidence constant  $C_{\text{ad}}$  were jointly optimized by grid search, which led to the choices  $C = 2$  and  $C_{\text{ad}} = 0.5$  for adNSA and  $C = 2$  and  $C_{\text{ad}} = 12$  for adMLSA and u-adMLSA. As for  $\sigma$ -adNSA and  $\sigma$ -adMLSA, we use the critical value  $C_p = 3$  with scaling factor  $C = 2$ .

$\varepsilon$	$h_0$	$\ell_{\text{ad}}$
$\frac{1}{32}$	$\frac{1}{16}$	1
$\frac{1}{64}$	$\frac{1}{32}$	1
$\frac{1}{128}$	$\frac{1}{32}$	2
$\frac{1}{256}$	$\frac{1}{32}$	2
$\frac{1}{512}$	$\frac{1}{32}$	3

$\varepsilon$	$h_0$	$L$	$\gamma_n$
$\frac{1}{32}$	$\frac{1}{16}$	1	$\frac{2}{2.5 \times 10^3 + n}$
$\frac{1}{64}$	$\frac{1}{32}$	1	$\frac{2}{4 \times 10^3 + n}$
$\frac{1}{128}$	$\frac{1}{32}$	2	$\frac{0.75}{9 \times 10^3 + n}$
$\frac{1}{256}$	$\frac{1}{32}$	3	$\frac{0.25}{10^4 + n}$
$\frac{1}{512}$	$\frac{1}{32}$	4	$\frac{0.09}{10^4 + n}$

$\varepsilon$	$h_0$	$L_{\text{ad}}$	$\gamma_n$
$\frac{1}{32}$	$\frac{1}{16}$	1	$\frac{2}{2.5 \times 10^3 + n}$
$\frac{1}{64}$	$\frac{1}{32}$	1	$\frac{2}{4 \times 10^3 + n}$
$\frac{1}{128}$	$\frac{1}{32}$	2	$\frac{0.75}{9 \times 10^3 + n}$
$\frac{1}{256}$	$\frac{1}{32}$	2	$\frac{0.25}{10^4 + n}$
$\frac{1}{512}$	$\frac{1}{32}$	3	$\frac{0.09}{10^4 + n}$

**Table 3.** Parametrizations of adNSA and  $\sigma$ -adNSA (left), MLSA (center) and adMLSA,  $\sigma$ -adMLSA and u-adMLSA (right), by prescribed accuracy.

Root-mean-square errors (RMSEs) relative to the true  $\xi_\star^0$ , as well as corresponding average execution times over 200 runs, for a prescribed accuracy ranging in  $\{\frac{1}{7}, \frac{1}{13}, \frac{1}{26}, \frac{1}{52}, \frac{1}{103}\}$ , are graphed for each algorithm on a logarithmic scale in Figure 2. Recalling that  $\xi_\star^0 \approx 2.012$ , these prescribed accuracies correspond to accuracy rates ranging from  $\frac{1/7}{2.012} \approx 7.10\%$  down to  $\frac{1/103}{2.012} \approx 0.48\%$ . Before feeding them into the different SA algorithms, we adjust each accuracy  $\varepsilon$  by a scaling factor of  $C' = \frac{1}{5}$  into  $C'\varepsilon$  ( $C'$  represents the constant on the right hand side of (5.5)).

Figure 3 plots the average execution times against the prescribed accuracies themselves to illustrate the achieved complexities. The slopes fitted on these curves, shown in dashed lines in Figures 2 and 3, are reported in Table 4.

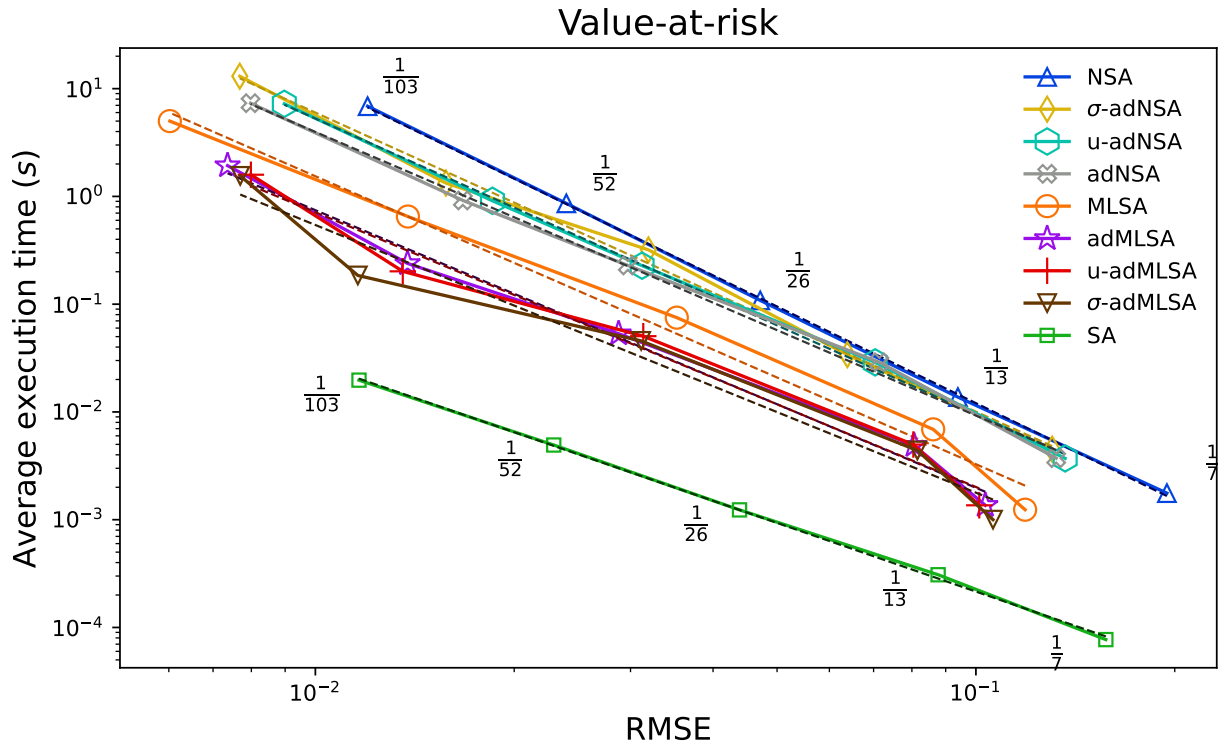


Figure 2. Performance comparison of the different SA schemes.

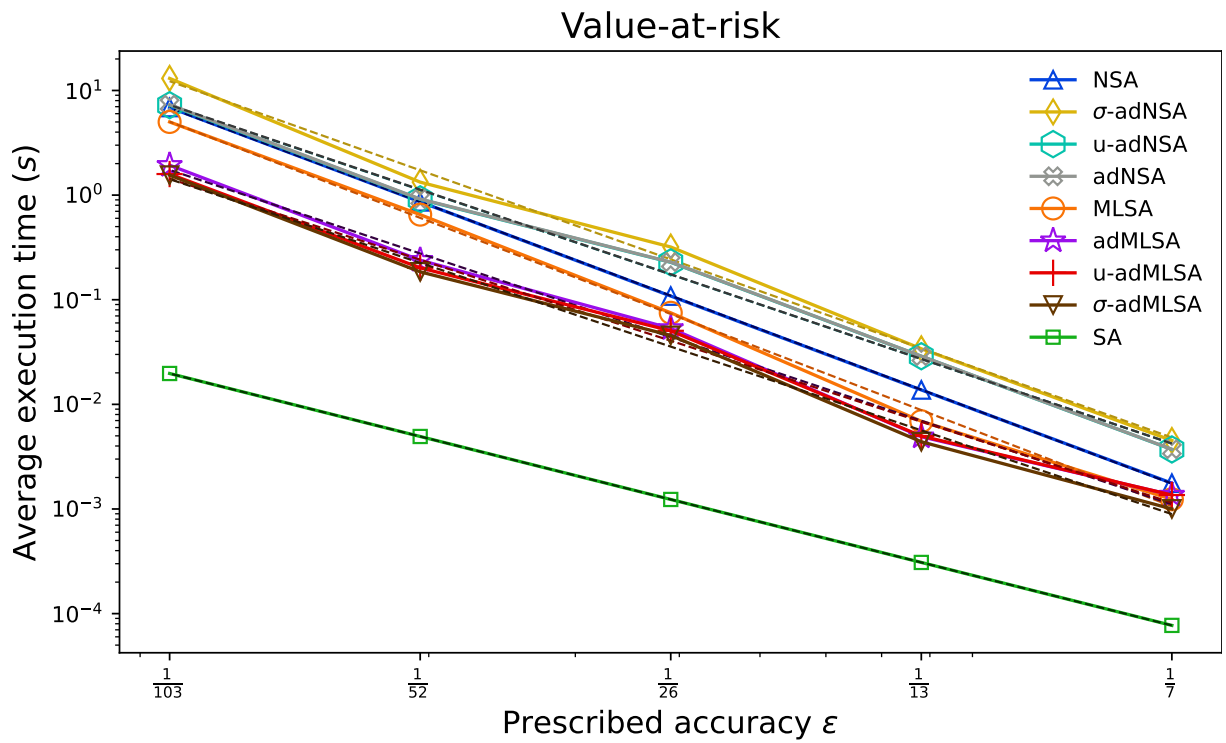


Figure 3. Complexity comparison of the different SA schemes.

SA scheme	NSA	$\sigma$ -adNSA	u-adNSA	adNSA	MLSA	adMLSA	u-adMLSA	$\sigma$ -adMLSA	SA
RMSE	-2.98	-2.78	-2.74	-2.63	-2.67	-2.58	-2.55	-2.48	-2.12
$\varepsilon$	-2.98	-2.83	-2.68	-2.69	-3.05	-2.65	-2.57	-2.66	-2.00

**Table 4.** Reported slopes from Figures 2 and 3.

[14, Section 5.3] already provides a thorough discussion on SA, NSA and MLSA. We retain here that MLSA achieves a partial gain on the gap of performance separating SA and NSA.

The novelties here are the adaptive schemes. On Figure 2, adNSA,  $\sigma$ -adNSA and u-adNSA achieve comparable performances, outperforming NSA by a margin that seems to widen for smaller accuracies. adNSA seems to be slightly outperforming both  $\sigma$ -adNSA and u-adNSA. The overperformance over  $\sigma$ -adNSA may be attributed to the overhead computation performed by the latter algorithm to recompute the confidence constant  $C_{\text{ad}}$ . The overperformance over u-adNSA can be explained by the desaturated refinement strategy used by the latter algorithm. adMLSA,  $\sigma$ -adMLSA and u-adMLSA show to be outperforming MLSA: for a target RMSE of order  $10^{-2}$ , adMLSA,  $\sigma$ -adMLSA and u-adMLSA run on average in 1s while MLSA runs in 2s, achieving a 2-fold speed-up over their non-adaptive counterpart.  $\sigma$ -adMLSA seems to be slightly outperforming adMLSA, which can be explained by the precision brought in by the recomputation of the standard deviation  $\sigma$ , in spite of the entailed computational overhead. The advantage of adMLSA over  $\sigma$ -adMLSA remains that, once calibrated on a few accuracies, the confidence constant  $C_{\text{ad}}$  for adMLSA needs not be readjusted, which eliminates all additional upstream fine-tuning. u-adMLSA seems also to be slightly outperforming adMLSA, which can be attributed to the computational gain achieved by the simpler unsaturated refinement strategy, despite the implied refinement imprecision.

Eventually, an examination of the fitted slopes highlights that adNSA achieves the theoretical complexity exponent predicted by Theorem 4.12. It also demonstrates that adMLSA approaches the quadratic complexity anticipated by Theorem 5.10(i) for  $p_\star$  large.

Note that, in a comparable numerical setting, [31] instead plots the quantity “Average execution time  $\times$  RMSE<sup>2</sup>” versus the RMSE on a logarithmic scale. This choice is intended to illustrate the underlying complexity results, recalling that the average execution time serves as a proxy for the computational cost, while the RMSE represents the prescribed accuracy. Under a quasi-quadratic complexity regime (Cost  $\propto \varepsilon^{-2-\delta}$ , with  $\delta > 0$  small), one expects a log-linear relationship between these quantities, with a mildly negative slope ( $\ln(\text{Cost } \varepsilon^2) \propto -\delta \ln \varepsilon$ ). The main drawback of this graphical representation is that the quantity “Average execution time  $\times$  RMSE<sup>2</sup>” lacks a clear interpretation, as it does not directly convey a tangible performance gain by directly comparing execution times for a given level of accuracy.

## 7.2 Interest Rate Swap

The following setting is recapitulated from [14, Section 6].

We consider a swap, of strike  $\bar{K}$ , maturity  $T$  and nominal  $\bar{N}$ , with each leg being worth 1 at inception, and issued at par on some underlying interest (or FX) rate  $(S_t)_{0 \leq t \leq T}$  that follows a Black-Scholes model with risk neutral drift  $\bar{\kappa}$  and constant volatility  $\bar{\sigma}$ . At the coupon dates  $0 < T_1 < \dots < T_d = T$ , the swap remunerates the cash flows  $\Delta T_i (S_{T_{i-1}} - \bar{K})$ , where  $\Delta T_i = T_i - T_{i-1}$ ,  $T_0 = 0$ . The risk-free rate is  $\bar{r}$  and the risk neutral probability measure is  $\mathbb{P}$ .

For  $t \in [0, T]$ , let  $i_t$  be the integer such that  $t \in [T_{i_t-1}, T_{i_t})$  if  $t \in [0, T)$ , and  $+\infty$  otherwise. The fair value of the swap at time  $t \in [0, T]$  is

$$P_t = \bar{N} \mathbb{E} \left[ \sum_{i=i_t}^d e^{-\bar{r}(T_i-t)} \Delta T_i (S_{T_{i-1}} - \bar{K}) \middle| S_t \right].$$

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

$$X^0 = e^{-\bar{r}\tau} P_\tau.$$

We are interested in computing the VaR  $\xi_\star^0$  of this loss at some confidence level  $\alpha \in (0, 1)$ .

### 7.2.1 Analytical and Simulation Formulas

On the one hand,

$$X^0 \stackrel{\mathcal{L}}{=} \bar{N}AS_0 \left( \exp \left( -\frac{\bar{\sigma}^2}{2}\tau + \bar{\sigma}\sqrt{\tau}U \right) - 1 \right), \quad \text{where } A := \sum_{i=2}^d e^{-\bar{r}T_i} \Delta T_i e^{\bar{\kappa}T_{i-1}}$$

and  $U \sim \mathcal{N}(0, 1)$ . This allows simulating  $X^0$  exactly, hence the availability of SA to approximate the VaR.

On the other hand,  $X^0$  satisfies

$$X^0 \stackrel{\mathcal{L}}{=} \mathbb{E}[\varphi(Y, Z)|Y], \tag{7.2}$$

where  $Y \in \mathbb{R}$  is independent of  $Z = (Z_1, \dots, Z_{d-1}) \in \mathbb{R}^{d-1}$ , with

$$\begin{aligned} Y &:= \exp \left( -\frac{\bar{\sigma}^2}{2}\tau + \bar{\sigma}\sqrt{\tau}U_0 \right), \\ Z_1 &:= \exp \left( -\frac{\bar{\sigma}^2}{2}(T_1 - \tau) + \bar{\sigma}\sqrt{T_1 - \tau}U_1 \right), \\ Z_i &:= \exp \left( -\frac{\bar{\sigma}^2}{2}\Delta T_i + \bar{\sigma}\sqrt{\Delta T_i}U_i \right), \quad 2 \leq i \leq d-1, \\ \varphi(y, z) &:= \bar{N}S_0 \sum_{i=2}^d e^{-\bar{r}T_i} \Delta T_i e^{\bar{\kappa}T_{i-1}} \left( y \prod_{j=1}^{i-1} z_j - 1 \right), \quad y \in \mathbb{R}, \quad z = (z_1, \dots, z_{d-1}) \in \mathbb{R}^{d-1}, \end{aligned}$$

and  $(U_i)_{0 \leq i \leq d-1} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ . The nested Monte Carlo averaging (2.7) is thus available to approximate (7.2), hence the applicability of NSA, adNSA,  $\sigma$ -adNSA, u-adNSA, MLSA, adMLSA,  $\sigma$ -adMLSA and u-adMLSA to approximate the VaR.

Finally, the VaR  $\xi_\star^0$  at level  $\alpha$  is available analytically:

$$\xi_\star^0 = \bar{N}AS_0 \left( \exp \left( F^{-1}(\alpha)\bar{\sigma}\sqrt{\tau} - \frac{\bar{\sigma}^2}{2}\tau \right) - 1 \right), \tag{7.3}$$

where  $F$  is the standard Gaussian cdf. The output of this formula will serve as a benchmark for the outcomes of the aforementioned algorithms.

### 7.2.2 Numerical Results

For the case study, we set  $S_0 = 1\%$ ,  $\bar{r} = 2\%$ ,  $\bar{\kappa} = 12\%$ ,  $\bar{\sigma} = 20\%$ ,  $T = 1$  year,  $\Delta T_i = 3$  months,  $\tau = 7$  days and  $\alpha = 85\%$ . We use the 30/360 day count fraction convention. (7.3) yields  $\xi_\star^0 \approx 219.64$ .

We run all algorithms at their theoretical optimums, with  $\beta = 1$  and the corresponding optimal iteration amounts. The learning rate  $\gamma_n = \frac{100}{n}$  is employed for SA and  $\gamma_n = \frac{50}{n}$  for NSA, adNSA,  $\sigma$ -adNSA and u-adNSA. For adNSA,  $\sigma$ -adNSA, u-adNSA, MLSA, adMLSA,  $\sigma$ -adMLSA and u-adMLSA, we adopt the framework (4.4) with the exponents  $p_\star = 8$  (fine-tuned on a grid to reduce RMSE and execution time) and  $\delta = 0.05$ , and the geometric step  $M = 2$ . As suggested in Section 6, we set  $\theta = \frac{p_\star - 2}{p_\star + 2}$  and  $r = 1 + \frac{1}{\theta}$ . For every prescribed accuracy  $\varepsilon \in \{\frac{1}{2}, 2, 4, 8, 15\}$ , we tune  $h_0$  (governing the number of levels  $L$  for MLSA,  $\ell_{\text{ad}}$  for adNSA,  $\sigma$ -adNSA and u-adNSA, and

$L_{\text{ad}}$  for adMLSA,  $\sigma$ -adMLSA and u-adMLSA) and the learning rate  $(\gamma_n)_{n \geq 1}$  for MLSA, adMLSA,  $\sigma$ -adMLSA and u-adMLSA on suitable grids. Table 5 lists these parametrizations by prescribed accuracy. The iteration amounts scaling factor  $C$  and adaptive refinement confidence constant  $C_{\text{ad}}$  are tuned on grids. We retain  $C = 2$  for adMLSA,  $\sigma$ -adMLSA and u-adMLSA,  $C_{\text{ad}} = 30$  for adMLSA and u-adMLSA,  $C = 2$  for adNSA,  $\sigma$ -adNSA and u-adNSA, and  $C_{\text{ad}} = 300$  for adNSA and u-adNSA. We eventually set the critical value  $C_p = 3$  for  $\sigma$ -adNSA and  $\sigma$ -adMLSA.

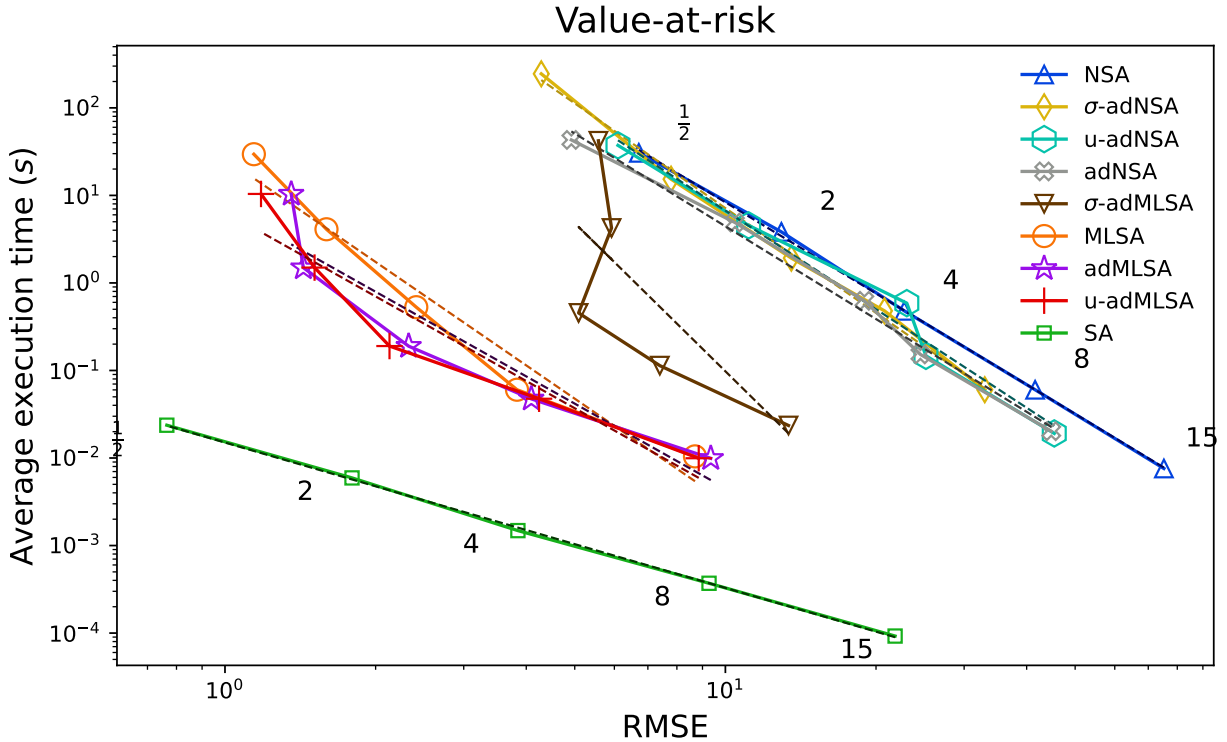
$\varepsilon$	$h_0$	$\ell_{\text{ad}}$
$\frac{1}{32}$	$\frac{1}{8}$	2
$\frac{1}{64}$	$\frac{1}{16}$	2
$\frac{1}{128}$	$\frac{1}{16}$	2
$\frac{1}{256}$	$\frac{1}{16}$	3
$\frac{1}{512}$	$\frac{1}{16}$	4

$\varepsilon$	$h_0$	$L$	$\gamma_n$
$\frac{1}{32}$	$\frac{1}{8}$	2	$\frac{6}{10+n}$
$\frac{1}{64}$	$\frac{1}{16}$	2	$\frac{20}{500+n}$
$\frac{1}{128}$	$\frac{1}{16}$	3	$\frac{21}{10^3+n}$
$\frac{1}{256}$	$\frac{1}{16}$	4	$\frac{20}{2 \times 10^3+n}$
$\frac{1}{512}$	$\frac{1}{16}$	5	$\frac{21}{3 \times 10^3+n}$

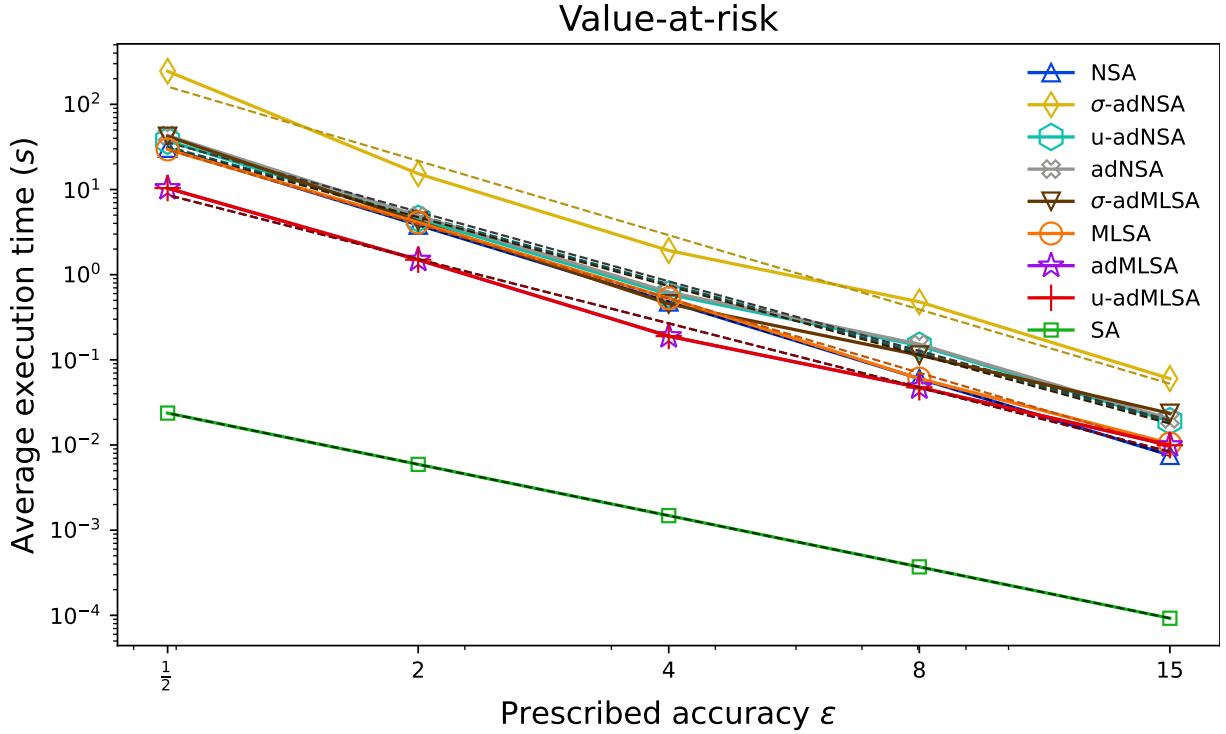
$\varepsilon$	$h_0$	$L_{\text{ad}}$	$\gamma_n$
$\frac{1}{32}$	$\frac{1}{8}$	2	$\frac{6}{10+n}$
$\frac{1}{64}$	$\frac{1}{16}$	2	$\frac{20}{500+n}$
$\frac{1}{128}$	$\frac{1}{16}$	2	$\frac{21}{10^3+n}$
$\frac{1}{256}$	$\frac{1}{16}$	3	$\frac{20}{2 \times 10^3+n}$
$\frac{1}{512}$	$\frac{1}{16}$	4	$\frac{21}{3 \times 10^3+n}$

**Table 5.** Parametrizations of adNSA and  $\sigma$ -adNSA (left), MLSA (center) and adMLSA and  $\sigma$ -adMLSA (right), by prescribed accuracy.

The joint evolution of the RMSE and average execution time over 200 runs for each SA scheme, for an accuracy  $\varepsilon$  looping through  $\{\frac{1}{2}, 2, 4, 8, 15\}$ , are plotted on a logarithmic scale on Figure 4. Since  $\xi_*^0 \approx 219.64$ , these accuracies coincide with accuracy rates in the tune of  $\frac{1/2}{219.64} \approx 6.83\%$  through  $\frac{15}{219.64} \approx 0.23\%$ . The prescribed accuracies  $\varepsilon$  have been adjusted by a factor of  $C' = \frac{1}{450}$  into  $C'\varepsilon$  before using them in the different SA schemes. Figure 5 showcases the average execution times against the prescribed accuracies. Table 6 reports the regressed slopes on these curves as depicted in dashed lines on Figures 4 and 5.



**Figure 4.** Performance comparison of the different SA schemes.



**Figure 5.** Complexity comparison of the different SA schemes.

SA scheme	NSA	$\sigma$ -adNSA	u-adNSA	adNSA	$\sigma$ -adMLSA	MLSA	adMLSA	u-adMLSA	SA
RMSE	-3.46	-3.98	-3.74	-3.52	-5.62	-3.92	-3.21	-3.21	-1.66
$\varepsilon$	-3.00	-2.90	-2.68	-2.71	-2.68	-2.90	-2.50	-2.50	-2.00

**Table 6.** Reported slopes from Figures 4 and 5.

We refer to [14, Section 6.3] for extended comments on SA, NSA and MLSA.

Some similar comments to the previous case study are applicable here. Concerning the nested schemes, adNSA,  $\sigma$ -adNSA and u-adNSA score a slight speed-up over NSA. As for the multilevel schemes, only adMLSA and u-adMLSA score a speed-up with respect to MLSA.  $\sigma$ -adMLSA seems to score worse than MLSA. This may be attributed to the computational overhead coming from the recomputation of the constant  $C_{\text{ad}}$  at each iteration. The performance margins between adNSA,  $\sigma$ -adNSA and u-adNSA on the one hand and between adMLSA and u-adMLSA on the other hand remain slim.

Although further fine-tuning may be required, the fitted slopes already indicate that the adaptive multilevel schemes adMLSA and u-adMLSA are closer than MLSA to the desired quadratic complexity.

## 8 Conclusion

For a prescribed accuracy  $\varepsilon > 0$ , the canonical  $\varepsilon^{-2}$  multilevel complexity order is retrieved, up to a logarithmic factor, for an MLSA scheme with a Heaviside-type update function. This is made possible by adopting an adaptive refinement strategy on the samples driving the multilevel algorithm's inner nested schemes. Our strategy allows us to close of the performance gap existing between the nested MLSA and the unbiased Robbins-Monro schemes for Heaviside-type update functions. The performance gain achieved by adaptive MLSA compared to its regular counterpart

is significant in practice, attaining a two-fold speed-up in certain cases and expected to increase exponentially with smaller prescribed accuracies. Potential future lines of research could involve applying a Polyak-Ruppert averaging [15] to adMLSA to promote greater numerical stability, or extending our findings on adaptive refinement to triply nested X-value adjustment estimations [27], or exploring the effect of sampling the risk factors  $Y$  and  $Z$  from an SDE trajectory [35].

**Acknowledgement.** The research of S. Crépey has benefited from the support of (by alphabetical order): *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; *Chair Futures of Quantitative Finance*, a partnership between LPSM at Université Paris Cité, CERMICS at École nationale des ponts et chaussées, and BNP Paribas Global Markets; *Chair Stress Test, RISK Management and Financial Steering*, led by the French Ecole polytechnique and its Foundation and sponsored by BNP Paribas.

The research of N. Frikha has benefited from the support of the Institut Europlace de Finance.

## A Auxiliary Results

Similarly to [13, Sect. 2.1] and [14, Appendix C], we define, for  $\mu \in \mathbb{R}_+$ ,  $h \in \overline{\mathcal{H}}$  and  $q \in \mathbb{N}^*$ , the Lyapunov function  $\mathcal{L}_{h,q}^\mu : \mathbb{R} \rightarrow \mathbb{R}_+$  given by

$$\mathcal{L}_{h,q}^\mu(\xi) = (V_h(\xi) - V_h(\xi_\star^h))^q \exp\left(\mu(V_h(\xi) - V_h(\xi_\star^h))\right), \quad \xi \in \mathbb{R}. \quad (\text{A.1})$$

The following lemma states some important properties of  $\mathcal{L}_{h,q}^\mu$ .

**Lemma A.1** ([14, Lemma C.1]). *Denote  $k_\alpha = 1 \vee \frac{\alpha}{1-\alpha}$ ,  $\mu_{h,q} = q\|V_h''\|_\infty$ , and  $\bar{\mathcal{L}}_{h,q} = \mathcal{L}_{h,q}^{\mu_{h,q}}$ ,  $h \in \overline{\mathcal{H}}$ ,  $q \in \mathbb{N}^*$ . Under Assumption 2.3, for all  $\mu \in \mathbb{R}_+$ ,  $h \in \overline{\mathcal{H}}$  and  $q \in \mathbb{N}^*$ ,*

(i)  $\mathcal{L}_{h,q}^\mu$  is twice continuously differentiable on  $\mathbb{R}$  and

$$(\mathcal{L}_{h,q}^\mu)'(\xi) = qV_h'(\xi)\mathcal{L}_{h,q-1}^\mu(\xi) + \mu V_h'(\xi)\mathcal{L}_{h,q}^\mu(\xi), \quad \xi \in \mathbb{R}. \quad (\text{A.2})$$

(ii) for all  $\xi \in \mathbb{R}$ ,

$$\bar{\mathcal{L}}_{h,q}(\xi) \leq k_\alpha^q |\xi - \xi_\star^h|^q \exp\left(\frac{qk_\alpha}{1-\alpha} \sup_{h' \in \overline{\mathcal{H}}} \|f_{X^{h'}}\|_\infty |\xi - \xi_\star^h|\right).$$

(iii) for all  $\xi \in \mathbb{R}$ ,

$$V_h'(\xi)(\mathcal{L}_{h,q}^\mu)'(\xi) \geq \lambda_{h,q}^\mu \mathcal{L}_{h,q}^\mu(\xi), \quad \text{where} \quad \lambda_{h,q}^\mu := \frac{3}{8}qV_h''(\xi_\star^h) \wedge \mu \frac{V_h''(\xi_\star^h)^4}{4[V_h'']_{\text{Lip}}^2}.$$

Furthermore, for all  $q' \geq q$ ,  $\inf_{h \in \overline{\mathcal{H}}} \lambda_{h,q}^{\mu_{h,q'}} > 0$ . In particular, letting  $\bar{\lambda}_{h,q} := \lambda_{h,q}^{\mu_{h,q}}$ , one has  $\inf_{h \in \overline{\mathcal{H}}} \bar{\lambda}_{h,q} > 0$ .

(iv) for all  $\xi \in \mathbb{R}$ ,

$$|(\mathcal{L}_{h,q}^\mu)''(\xi)| \leq \chi_{h,q}^\mu (\mathcal{L}_{h,q}^\mu(\xi) + \mathcal{L}_{h,q-1}^\mu(\xi)),$$

where

$$\chi_{h,q}^\mu := (q \vee \mu)\|V_h''\|_\infty + k_\alpha^2 \mu(\mu \vee 2) + q(2\mu \vee (q-1)) \left( \frac{3k_\alpha^2 [V_h'']_{\text{Lip}}^2}{V_h''(\xi_\star^h)^3} \vee \frac{3\|V_h''\|_\infty^2}{V_h''(\xi_\star^h)} \right).$$

Besides, for all  $q' \geq q$ ,  $|\lambda_{h,q}^{\mu_{h,q'}}|^2 \leq \chi_{h,q}^{\mu_{h,q'}}$ , and  $\sup_{h \in \overline{\mathcal{H}}} \chi_{h,q}^{\mu_{h,q'}} < \infty$ . In particular, introducing  $\bar{\chi}_{h,q} := \chi_{h,q}^{\mu_{h,q}}$ , then  $|\bar{\lambda}_{h,q}|^2 \leq \bar{\chi}_{h,q}$ , and  $\sup_{h \in \overline{\mathcal{H}}} \bar{\chi}_{h,q} < \infty$ .

(v) for all  $\mu' \in \mathbb{R}_+$  and all  $\xi \in \mathbb{R}$ ,

$$(\xi - \xi_*^h)^{2q} \leq \kappa_{h,q} (\mathcal{L}_{h,q}^\mu(\xi) + \mathcal{L}_{h,2q}^{\mu'}(\xi)), \quad \text{where} \quad \kappa_{h,q} := \frac{3^q}{V_h''(\xi_*^h)^q} \vee \frac{3^{2q} [V_h'']_{\text{Lip}}^{2q}}{V_h''(\xi_*^h)^{4q}}.$$

Moreover,  $\sup_{h \in \overline{\mathcal{H}}} \kappa_{h,q} < \infty$ .

For  $\omega \in \mathbb{R}$ , let

$$\Phi_\omega(t) := \begin{cases} \omega^{-1}(t^\omega - 1), & \text{if } \omega \neq 0, \\ \ln t + \gamma^\mathcal{E}, & \text{if } \omega = 0, \end{cases} \quad t \in \mathbb{R}_+, \quad (\text{A.3})$$

where  $\gamma^\mathcal{E}$  is the Euler-Mascheroni constant. The next result generalizes [14, Lemma C.2].

**Lemma A.2.** *Let  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\gamma_1 > 0$  and  $\beta \in (0, 1]$ . Using the convention  $\prod_\emptyset = 1$ , define*

$$\Pi_{k+1:n}^\delta := \prod_{j=k+1}^n (1 - \lambda \gamma_j + \zeta \gamma_j^{1+\delta}), \quad k \in \llbracket 0, n \rrbracket, \quad n \in \mathbb{N}^*,$$

where  $\delta, \lambda > 0$  and  $\zeta \geq 0$  are constants independent of  $k$  and  $n$ . Assume that  $1 - \lambda \gamma_n + \zeta \gamma_n^{1+\delta} > 0$ ,  $n \in \mathbb{N}^*$ . Then, for all  $n \in \mathbb{N}^*$ ,

(i) for all  $k \in \llbracket 0, n-1 \rrbracket$ ,

$$\Pi_{k+1:n}^\delta \leq \begin{cases} \exp\{-\lambda \gamma_1 (\Phi_{1-\beta}(n+1) - \Phi_{1-\beta}(k+1))\} \\ \quad \times \exp\{2^{\beta(1+\delta)} \zeta \gamma_1^{1+\delta} (\Phi_{1-\beta(1+\delta)}(n+1) - \Phi_{1-\beta(1+\delta)}(k+1))\} & \text{if } \beta \in (0, 1), \\ \exp(\zeta(1 + \frac{1}{\delta}) \gamma_1^{1+\delta} + \frac{\lambda \gamma_1}{2}) \frac{(k+1)^{\lambda \gamma_1}}{(n+1)^{\lambda \gamma_1}} & \text{if } \beta = 1. \end{cases}$$

(ii) for all  $p \geq 1$ ,

$$\sum_{k=1}^n \gamma_k^p \Pi_{k+1:n}^\delta \leq \begin{cases} \zeta^{-1} (\gamma_1^{p-(1+\delta)} \vee (\frac{\lambda}{2\zeta})^{\frac{p-(1+\delta)}{\delta}}) \exp(2^{1+\beta(1+\delta)} \zeta \gamma_1^{1+\delta} \Phi_{1-\beta(1+\delta)}(n+1)) \\ \quad \times \exp(-\frac{\lambda \gamma_1}{2} \Phi_{1-\beta}(n+1)) & \text{if } \beta \in (0, 1), \\ \gamma_1^p \exp\{\zeta(1 + \frac{1}{\delta}) \gamma_1^{1+\delta} + \frac{\lambda \gamma_1}{2} + ((\lambda \gamma_1) \vee p) \ln(2)\} \frac{\Phi_{\lambda \gamma_1 - (p-1)}(n+1)}{(n+1)^{\lambda \gamma_1}} & \text{if } \beta = 1. \end{cases}$$

*Proof.* We follow closely the strategy employed in the proof of [14, Lemma C.2].

Let  $n \in \mathbb{N}^*$ ,  $k \in \llbracket 0, n \rrbracket$  and  $p \geq 1$ . Using that  $1 + x \leq e^x$ ,  $x \in \mathbb{R}$ ,

$$\Pi_{k+1:n}^\delta \leq \exp\left(-\lambda \sum_{j=k+1}^n \gamma_j\right) \exp\left(\zeta \sum_{j=k+1}^n \gamma_j^{1+\delta}\right), \quad (\text{A.4})$$

with the convention  $\sum_\emptyset = 0$ .

*Step 1. Case  $\beta = 1$ .*

*Step 1.1. Inequality on  $\Pi_{k+1:n}^\delta$ .*

Denoting  $\psi$  is the digamma function (c.f. Abramowitz et al. [1, Section 6.3]), we have

$$\sum_{j=k+1}^n \gamma_j = \gamma_1 (\psi(n+1) - \psi(k+1)).$$

We recall that  $\psi$  satisfies (see e.g. Alzer [3])

$$\ln x - \frac{1}{x} \leq \psi(x) \leq \ln x - \frac{1}{2x}, \quad x > 0.$$

Thus

$$\sum_{j=k+1}^n \gamma_j \geq \gamma_1 \ln \left( \frac{n+1}{k+1} \right) - \frac{\gamma_1}{n+1} + \frac{\gamma_1}{2(k+1)} \geq \gamma_1 \ln \left( \frac{n+1}{k+1} \right) - \frac{\gamma_1}{2}.$$

Besides, via a series-integral comparison,

$$\sum_{j=1}^{\infty} \gamma_j^{1+\delta} \leq \gamma_1^{1+\delta} \left( 1 + \frac{1}{\delta} \right).$$

Thus, recalling (A.4),

$$\Pi_{k+1:n}^{\delta} \leq \exp \left( \zeta \left( 1 + \frac{1}{\delta} \right) \gamma_1^{1+\delta} + \frac{\lambda \gamma_1}{2} \right) \frac{(k+1)^{\lambda \gamma_1}}{(n+1)^{\lambda \gamma_1}}. \quad (\text{A.5})$$

*Step 1.2. Inequality on  $\sum_{k=1}^n \gamma_k^p \Pi_{k+1:n}^{\delta}$ .*

Via the upper bound in (A.5),

$$\sum_{k=1}^n \gamma_k^p \Pi_{k+1:n}^{\delta} \leq \frac{\gamma_1^p}{(n+1)^{\lambda \gamma_1}} \exp \left( \zeta \left( 1 + \frac{1}{\delta} \right) \gamma_1^{1+\delta} + \frac{\lambda \gamma_1}{2} \right) \sum_{k=1}^n \frac{(k+1)^{\lambda \gamma_1}}{k^p}.$$

If  $\lambda \gamma_1 \leq p$ , then

$$\sum_{k=1}^n \frac{(k+1)^{\lambda \gamma_1}}{k^p} \leq 2^p \sum_{k=1}^n \frac{1}{(k+1)^{p-\lambda \gamma_1}} \leq 2^p \Phi_{\lambda \gamma_1 - p + 1}(n+1).$$

Otherwise, if  $\lambda \gamma_1 > p$ , then

$$\sum_{k=1}^n \frac{(k+1)^{\lambda \gamma_1}}{k^p} \leq 2^{\lambda \gamma_1} \sum_{k=1}^n k^{\lambda \gamma_1 - p} \leq 2^{\lambda \gamma_1} \Phi_{\lambda \gamma_1 - p + 1}(n+1).$$

All in all,

$$\sum_{k=1}^n \gamma_k^p \Pi_{k+1:n}^{\delta} \leq \gamma_1^p \exp \left( \zeta \left( 1 + \frac{1}{\delta} \right) \gamma_1^{1+\delta} + \frac{\lambda \gamma_1}{2} + ((\lambda \gamma_1) \vee p) \ln(2) \right) \frac{\Phi_{\lambda \gamma_1 - p + 1}(n+1)}{(n+1)^{\lambda \gamma_1}}.$$

*Step 2. Case  $\beta \in (0, 1)$ .*

*Step 2.1. Inequality on  $\Pi_{k+1:n}^{\delta}$ .*

Via a series-integral comparison,

$$\sum_{j=k+1}^n \gamma_j \geq \gamma_1 (\Phi_{1-\beta}(n+1) - \Phi_{1-\beta}(k+1)), \quad (\text{A.6})$$

and also

$$\sum_{j=k+1}^n \gamma_j^{1+\delta} \leq 2^{\beta(1+\delta)} \gamma_1^{1+\delta} (\Phi_{1-\beta(1+\delta)}(n+1) - \Phi_{1-\beta(1+\delta)}(k+1)). \quad (\text{A.7})$$

Therefore, recalling (A.4),

$$\begin{aligned} \Pi_{k+1:n}^{\delta} &\leq \exp \left( -\lambda \gamma_1 (\Phi_{1-\beta}(n+1) - \Phi_{1-\beta}(k+1)) \right) \\ &\quad \times \exp \left( 2^{\beta(1+\delta)} \zeta \gamma_1^{1+\delta} (\Phi_{1-\beta(1+\delta)}(n+1) - \Phi_{1-\beta(1+\delta)}(k+1)) \right). \end{aligned}$$

Step 2.2. Inequality on  $\sum_{k=1}^n \gamma_k^p \Pi_{k+1:n}^\delta$ .

Let

$$n_0 := \inf \left\{ n \in \mathbb{N}^* : \gamma_n \leq \left( \frac{\lambda}{2\zeta} \right)^{\frac{1}{\delta}} \right\} - 1.$$

Note that

$$1 - \lambda\gamma_n + \zeta\gamma_n^{1+\delta} \leq 1 - \frac{\lambda}{2}\gamma_n, \quad n \geq n_0 + 1,$$

and that

$$n_0 = \left\lceil \left( \gamma_1 \left( \frac{2\zeta}{\lambda} \right)^{\frac{1}{\delta}} \right)^{\frac{1}{\beta}} \right\rceil - 1 < \left( \gamma_1 \left( \frac{2\zeta}{\lambda} \right)^{\frac{1}{\delta}} \right)^{\frac{1}{\beta}}. \quad (\text{A.8})$$

Now,

$$\begin{aligned} \sum_{k=1}^n \gamma_k^p \Pi_{k+1:n}^\delta &= \sum_{k=1}^{n_0 \wedge n} \gamma_k^p \Pi_{k+1:n_0 \wedge n}^\delta \Pi_{n_0 \wedge n+1:n}^\delta + \sum_{k=n_0 \wedge n+1}^n \gamma_k^p \Pi_{k+1:n}^\delta \\ &\leq \left( \sum_{k=1}^{n_0 \wedge n} \gamma_k^p \prod_{j=k+1}^{n_0 \wedge n} (1 + \zeta\gamma_j^{1+\delta}) \right) \prod_{j=n_0 \wedge n+1}^n \left( 1 - \frac{\lambda}{2}\gamma_j \right) \\ &\quad + \sum_{k=1}^n \gamma_k^p \prod_{j=k+1}^n \left( 1 - \frac{\lambda}{2}\gamma_j \right). \end{aligned} \quad (\text{A.9})$$

Let us study each term in the upper bound separately.

On the one hand,

$$\begin{aligned} &\sum_{k=1}^{n_0 \wedge n} \gamma_k^p \prod_{j=k+1}^{n_0 \wedge n} (1 + \zeta\gamma_j^{1+\delta}) \\ &\leq (\gamma_1^{p-(1+\delta)} \vee \gamma_{n_0}^{p-(1+\delta)}) \sum_{k=1}^{n_0 \wedge n} \gamma_k^{1+\delta} \prod_{j=k+1}^{n_0 \wedge n} (1 + \zeta\gamma_j^{1+\delta}). \end{aligned} \quad (\text{A.10})$$

If  $p \geq 1 + \delta$ , then, given the decreasing monotonicity of  $(\gamma_n)_{n \geq 1}$ ,

$$\gamma_1^{p-(1+\delta)} \vee \gamma_{n_0}^{p-(1+\delta)} \leq \gamma_1^{p-(1+\delta)}. \quad (\text{A.11})$$

Otherwise, if  $p < 1 + \delta$ , then, taking into account (A.8),

$$\gamma_1^{p-(1+\delta)} \vee \gamma_{n_0}^{p-(1+\delta)} \leq \left( \frac{n_0^\beta}{\gamma_1} \right)^{(1+\delta)-p} < \left( \frac{2\zeta}{\lambda} \right)^{\frac{(1+\delta)-p}{\delta}}. \quad (\text{A.12})$$

Besides, note that, by telescopic summation,

$$\begin{aligned} &\sum_{k=1}^{n_0 \wedge n} \gamma_k^{1+\delta} \prod_{j=k+1}^{n_0 \wedge n} (1 + \zeta\gamma_j^{1+\delta}) \\ &= \frac{1}{\zeta} \sum_{k=1}^{n_0 \wedge n} \left( \prod_{j=k}^{n_0 \wedge n} (1 + \zeta\gamma_j^{1+\delta}) - \prod_{j=k+1}^{n_0 \wedge n} (1 + \zeta\gamma_j^{1+\delta}) \right), \\ &\leq \frac{1}{\zeta} \prod_{j=1}^{n_0 \wedge n} (1 + \zeta\gamma_j^{1+\delta}). \end{aligned} \quad (\text{A.13})$$

Thus, coming back to (A.10), using (A.11), (A.12), (A.13), telescopic summation and the inequality  $1 + x \leq e^x$ ,  $x \in \mathbb{R}$ ,

$$\begin{aligned} & \sum_{k=1}^{n_0 \wedge n} \gamma_k^p \prod_{j=k+1}^{n_0 \wedge n} (1 + \zeta \gamma_j^{1+\delta}) \\ & \leq \zeta^{-1} \left( \gamma_1^{p-(1+\delta)} \vee \left( \frac{\lambda}{2\zeta} \right)^{\frac{p-(1+\delta)}{\delta}} \right) \prod_{j=1}^{n_0 \wedge n} (1 + \zeta \gamma_j^{1+\delta}) \\ & \leq \zeta^{-1} \left( \gamma_1^{p-(1+\delta)} \vee \left( \frac{\lambda}{2\zeta} \right)^{\frac{p-(1+\delta)}{\delta}} \right) \exp \left( \zeta \sum_{j=1}^{n_0 \wedge n} \gamma_j^{1+\delta} \right). \end{aligned}$$

Moreover, note that, for  $j \leq n_0$ ,  $\gamma_j \geq (\frac{\lambda}{2\zeta})^{\frac{1}{\delta}}$ , so that  $\frac{\lambda}{2} \gamma_j \leq \zeta \gamma_j^{1+\delta}$ . Hence, using the inequality  $1 + x \leq e^x$ ,  $x \in \mathbb{R}$ ,

$$\begin{aligned} \prod_{j=n_0 \wedge n+1}^n \left( 1 - \frac{\lambda}{2} \gamma_j \right) & \leq \exp \left( -\frac{\lambda}{2} \sum_{j=1}^n \gamma_j \right) \exp \left( \frac{\lambda}{2} \sum_{j=1}^{n_0 \wedge n} \gamma_j \right) \\ & \leq \exp \left( -\frac{\lambda}{2} \sum_{j=1}^n \gamma_j \right) \exp \left( \zeta \sum_{j=1}^{n_0 \wedge n} \gamma_j^{1+\delta} \right). \end{aligned}$$

Combining the two preceding upper bounds and using (A.6) and (A.7) with  $k = 0$ ,

$$\begin{aligned} & \left( \sum_{k=1}^{n_0 \wedge n} \gamma_k^p \prod_{j=k+1}^{n_0 \wedge n} (1 + \zeta \gamma_j^{1+\delta}) \right) \prod_{j=n_0 \wedge n+1}^n \left( 1 - \frac{\lambda}{2} \gamma_j \right) \\ & \leq \zeta^{-1} \left( \gamma_1^{p-(1+\delta)} \vee \left( \frac{\lambda}{2\zeta} \right)^{\frac{p-(1+\delta)}{\delta}} \right) \exp \left( 2\zeta \sum_{j=1}^{n_0 \wedge n} \gamma_j^{1+\delta} \right) \exp \left( -\frac{\lambda}{2} \sum_{j=1}^n \gamma_j \right) \\ & \leq \zeta^{-1} \left( \gamma_1^{p-(1+\delta)} \vee \left( \frac{\lambda}{2\zeta} \right)^{\frac{p-(1+\delta)}{\delta}} \right) \exp \left( 2^{1+\beta(1+\delta)} \zeta \gamma_1^{1+\delta} \Phi_{1-\beta(1+\delta)}(n+1) \right) \\ & \quad \times \exp \left( -\frac{\lambda \gamma_1}{2} \Phi_{1-\beta}(n+1) \right). \end{aligned} \tag{A.14}$$

On the other hand, for  $m \in \llbracket 1, n \rrbracket$ ,

$$\sum_{k=1}^n \gamma_k^p \prod_{j=k+1}^n \left( 1 - \frac{\lambda}{2} \gamma_j \right) = \sum_{k=1}^m \gamma_k^p \prod_{j=k+1}^n \left( 1 - \frac{\lambda}{2} \gamma_j \right) + \sum_{k=m+1}^n \gamma_k^p \prod_{j=k+1}^n \left( 1 - \frac{\lambda}{2} \gamma_j \right). \tag{A.15}$$

By the inequality  $1 + x \leq e^x$ ,  $x \in \mathbb{R}$  and the decreasing monotony of  $(\gamma_n)_{n \geq 1}$ ,

$$\begin{aligned} \sum_{k=1}^m \gamma_k^p \prod_{j=k+1}^n \left( 1 - \frac{\lambda}{2} \gamma_j \right) & \leq \prod_{j=m+1}^n \left( 1 - \frac{\lambda}{2} \gamma_j \right) \sum_{k=1}^m \gamma_k^p \\ & \leq \exp \left( -\frac{\lambda}{2} \sum_{j=m+1}^n \gamma_j \right) \gamma_1^{p-1} \sum_{k=1}^m \gamma_k. \end{aligned} \tag{A.16}$$

Using series-integral comparisons,

$$2^{-\beta} \gamma_1 (\Phi_{1-\beta}(n) - \Phi_{1-\beta}(m)) \leq \sum_{j=m+1}^n \gamma_j \leq \sum_{j=1}^n \gamma_j \leq 2^\beta \gamma_1 \Phi_{1-\beta}(n+1). \tag{A.17}$$

Hence, combining (A.16) and (A.17),

$$\sum_{k=1}^m \gamma_k^p \prod_{j=k+1}^n \left(1 - \frac{\lambda}{2} \gamma_j\right) \leq 2^\beta \gamma_1^p \exp\left(-2^{-(\beta+1)} \lambda \gamma_1 (\Phi_{1-\beta}(n) - \Phi_{1-\beta}(m))\right) \Phi_{1-\beta}(n+1). \quad (\text{A.18})$$

Besides, since  $(\gamma_n)_{n \geq 1}$  is decreasing,

$$\sum_{k=m+1}^n \gamma_k^p \prod_{j=k+1}^n \left(1 - \frac{\lambda}{2} \gamma_j\right) \leq \gamma_{m+1}^{p-1} \sum_{k=m+1}^n \gamma_k \prod_{j=k+1}^n \left(1 - \frac{\lambda}{2} \gamma_j\right). \quad (\text{A.19})$$

By telescopic summation,

$$\sum_{k=m+1}^n \gamma_k \prod_{j=k+1}^n \left(1 - \frac{\lambda}{2} \gamma_j\right) = \frac{2}{\lambda} \sum_{k=m+1}^n \left( \prod_{j=k+1}^n \left(1 - \frac{\lambda}{2} \gamma_j\right) - \prod_{j=k}^n \left(1 - \frac{\lambda}{2} \gamma_j\right) \right) \leq \frac{2}{\lambda}. \quad (\text{A.20})$$

Thus, gathering (A.19) and (A.20),

$$\sum_{k=m+1}^n \gamma_k^p \prod_{j=k+1}^n \left(1 - \frac{\lambda}{2} \gamma_j\right) \leq \frac{2\gamma_{m+1}^{p-1}}{\lambda} = \frac{2\gamma_1^{p-1}}{\lambda(m+1)^{(p-1)\beta}}. \quad (\text{A.21})$$

Collecting the upper bounds from (A.15), (A.18) and (A.21),

$$\begin{aligned} & \sum_{k=1}^n \gamma_k^p \prod_{j=k+1}^n \left(1 - \frac{\lambda}{2} \gamma_j\right) \\ & \leq 2^\beta \gamma_1^p \exp\left(-2^{-(\beta+1)} \lambda \gamma_1 (\Phi_{1-\beta}(n) - \Phi_{1-\beta}(m))\right) \Phi_{1-\beta}(n+1) + \frac{2\gamma_1^{p-1}}{\lambda(m+1)^{(p-1)\beta}}. \end{aligned}$$

Taking  $m = \lfloor \frac{n}{2} \rfloor$ , and using that  $2^\beta \leq 1 + \beta$ ,  $\beta \in (0, 1]$ , one has

$$\Phi_{1-\beta}(n) - \Phi_{1-\beta}\left(\left\lfloor \frac{n}{2} \right\rfloor\right) \geq \frac{n^{1-\beta}}{2}.$$

Thus

$$\begin{aligned} & \sum_{k=1}^n \gamma_k^p \prod_{j=k+1}^n \left(1 - \frac{\lambda}{2} \gamma_j\right) \\ & \leq 2^\beta \gamma_1^p \exp(-2^{-(\beta+2)} \lambda \gamma_1 n^{1-\beta}) \Phi_{1-\beta}(n+1) + \frac{2^{1+(p-1)\beta} \gamma_1^{p-1}}{\lambda n^{(p-1)\beta}}. \end{aligned} \quad (\text{A.22})$$

Finally, injecting the upper bounds of (A.14) and (A.22) into (A.9) concludes the proof.  $\square$

The proof of the following result is included within the preliminary steps of the proof of [14, Proposition 5.2], and is therefore omitted.

**Lemma A.3.** (i) Assume that there exists  $p \geq 1$  such that

$$\mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^p] < \infty.$$

Then, for all  $h, h' \in \overline{\mathcal{H}}$ ,

$$\mathbb{E}[|X^h - X^{h'}|^p] \leq C|h - h'|^{\frac{p}{2}},$$

where

$$C = \sqrt{2} B_p \mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^p]^{\frac{1}{p}},$$

with  $B_p$  being a positive constant that depends only on  $p$ .

(ii) Assume that there exists  $\mathfrak{g} > 0$  such that, for all  $u \in \mathbb{R}$ ,

$$\mathbb{E}\left[\exp\left(u(\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y])\right)\middle|Y\right] \leq e^{\mathfrak{g}u^2} \quad \mathbb{P}\text{-as.}$$

Then, for all  $h, h' \in \overline{\mathcal{H}}$  and all  $u \in \mathbb{R}$ ,

$$\mathbb{E}\left[\exp\left(u(X^h - X^{h'})\right)\middle|Y\right] \leq \exp(\mathfrak{g}u^2|h - h'|) \quad \mathbb{P}\text{-as.}$$

We refer to Remarks 2.9(i) and (ii) for related comments.

## B Proofs of the Convergence Results

*Proof of Lemma 3.1.* The proof is similar to [14, Proposition 4.2(ii)]. Using the tower law, one has

$$\begin{aligned} & \mathbb{E}[|\mathbb{1}_{\{X^{h'} > \xi\}} - \mathbb{1}_{\{X^h > \xi\}}|] \\ &= \mathbb{P}(X^{h'} \leq \xi < X^h) + \mathbb{P}(X^h \leq \xi < X^{h'}) \\ &= \mathbb{P}(X^h + (h')^{\frac{1}{2}}G_h^{h'} \leq \xi < X^h) + \mathbb{P}(X^h \leq \xi < X^h + (h')^{\frac{1}{2}}G_h^{h'}) \\ &= \mathbb{P}(X^h - (h')^{\frac{1}{2}}(G_h^{h'})^- \leq \xi < X^h) + \mathbb{P}(X^h \leq \xi < X^h + (h')^{\frac{1}{2}}(G_h^{h'})^+) \\ &= \mathbb{E}[F_{X^h|G_h^{h'}}(\xi + (h')^{\frac{1}{2}}(G_h^{h'})^-) - F_{X^h|G_h^{h'}}(\xi)] \\ &\quad + \mathbb{E}[F_{X^h|G_h^{h'}}(\xi) - F_{X^h|G_h^{h'}}(\xi - (h')^{\frac{1}{2}}(G_h^{h'})^+)] \\ &= \mathbb{E}[F_{X^h|G_h^{h'}}(\xi + (h')^{\frac{1}{2}}(G_h^{h'})^-) - F_{X^h|G_h^{h'}}(\xi - (h')^{\frac{1}{2}}(G_h^{h'})^+)] \\ &\leq \left( \sup_{0 \leq h_1 < h_2 \in \overline{\mathcal{H}}} \mathbb{E}[K_{h_1}^{h_2}|G_{h_1}^{h_2}] \right) (h')^{\frac{1}{2}}. \end{aligned}$$

□

*Proof of Lemma 4.2.* Let  $\ell, n \in \mathbb{N}^*$  and  $\xi \in \mathbb{R}$ ,  $C_{\text{ad}} > 0$ ,  $r > 1$  and  $\theta \in (0, 1]$ .

By (3.10) and the law of total probability,

$$\begin{aligned} & \mathbb{P}(X^{h_{\ell+n}(\xi)} \leq \xi) - F_{X^{h_{\ell+n}(\xi)}}(\xi) \\ &= \sum_{k=0}^{\lceil \theta \ell \rceil} \mathbb{P}(X^{h_{\ell+k}} \leq \xi, \eta_n^\ell(\xi) = k) - \mathbb{P}(X^{h_{\ell+\lceil \theta \ell \rceil}} \leq \xi, \eta_n^\ell(\xi) = k) \\ &= \sum_{k=0}^{\lceil \theta \ell \rceil - 1} \mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil \theta \ell \rceil}} \leq \xi\}}) \mathbb{1}_{\{\eta_n^\ell(\xi) = k\}}]. \end{aligned} \quad (\text{B.1})$$

Let  $k \in \llbracket 0, \lceil \theta \ell \rceil - 1 \rrbracket$ . On the one hand,

$$\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil \theta \ell \rceil}} \leq \xi\}} = \mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi < X^{h_{\ell+\lceil \theta \ell \rceil}}\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil \theta \ell \rceil}} \leq \xi < X^{h_{\ell+k}}\}}. \quad (\text{B.2})$$

On the other hand, by (3.10),

$$\{\eta_n^\ell(\xi) = k\} \subset \{|X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}}\psi_n^{\ell, k}\}. \quad (\text{B.3})$$

Therefore, combining (B.2) and (B.3),

$$\begin{aligned} & \mathbb{E}[|\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil \theta \ell \rceil}} \leq \xi\}}| \mathbb{1}_{\{\eta_n^\ell(\xi) = k\}}|] \\ &\leq \mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi < X^{h_{\ell+\lceil \theta \ell \rceil}}\}} + \mathbb{1}_{\{X^{h_{\ell+\lceil \theta \ell \rceil}} \leq \xi < X^{h_{\ell+k}}\}}) \mathbb{1}_{\{|X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}}\psi_n^{\ell, k}\}}] \\ &\leq \mathbb{P}(|X^{h_{\ell+k}} - X^{h_{\ell+\lceil \theta \ell \rceil}}| \geq |X^{h_{\ell+k}} - \xi|, |X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}}\psi_n^{\ell, k}) \\ &\leq \mathbb{P}(|X^{h_{\ell+k}} - X^{h_{\ell+\lceil \theta \ell \rceil}}| \geq C_{\text{ad}}\psi_n^{\ell, k}), \end{aligned} \quad (\text{B.4})$$

where we used that

$$\{X^{h_{\ell+k}} \leq \xi < X^{h_{\ell+\lceil\theta\ell\rceil}\} \cup \{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi < X^{h_{\ell+k}}\} \subset \{|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}| \geq |X^{h_{\ell+k}} - \xi|\}.$$

(i)a. By (B.4), Markov's inequality and Lemma A.3(i),

$$\begin{aligned} & |\mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi\}}) \mathbb{1}_{\{\eta_n^\ell(\xi)=k\}}]| \\ & \leq \mathbb{P}(|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}| \geq C_{\text{ad}} u_n^{-\frac{1}{p_\star}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}}) \\ & \leq C_{\text{ad}}^{-p_\star} u_n h_{\theta\ell(r-1)+k}^{-\frac{p_\star}{r}} \mathbb{E}[|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}|^{p_\star}] \\ & \leq C_{11} u_n h_{\theta\ell(r-1)+k}^{-\frac{p_\star}{r}} h_{\ell+k}^{\frac{p_\star}{2}}, \end{aligned} \tag{B.5}$$

where

$$C_{11} = \sqrt{2} C_{\text{ad}}^{-p_\star} B_{p_\star} \mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{p_\star}]^{\frac{1}{p_\star}},$$

with  $B_{p_\star}$  being a positive constant that depends only on  $p_\star$ . Using (B.5), (B.1) and the condition (4.4) on  $r$  and on  $\theta$ , it follows

$$\begin{aligned} & |\mathbb{P}(X^{h_{\ell+\eta_n^\ell(\xi)}} \leq \xi) - F_{X^{h_{\ell+\lceil\theta\ell\rceil}}(\xi)| \\ & \leq C_{11} u_n \sum_{k=0}^{\lceil\theta\ell\rceil-1} h_{\theta\ell(r-1)+k}^{-\frac{p_\star}{r}} h_{\ell+k}^{\frac{p_\star}{2}} \\ & \leq \frac{C_{11} h_1^{-p_\star(\frac{1}{r}-\frac{1}{2})}}{M^{p_\star(\frac{1}{r}-\frac{1}{2})} - 1} u_n M^{-\frac{p_\star\ell}{2}(1-\theta)} \\ & \leq C_{12} u_n h_\ell^{1+\theta}, \end{aligned} \tag{B.6}$$

where

$$C_{12} = \frac{C_{11} h_0^{-\frac{2p_\star}{p_\star+2}} h_1^{-p_\star(\frac{1}{r}-\frac{1}{2})}}{M^{p_\star(\frac{1}{r}-\frac{1}{2})} - 1}.$$

(i)b. Let  $\lambda > 0$ . By (B.4), the definition (3.10) for the case (2.24), Markov's exponential inequality, the fact that  $e^{|x|} \leq e^x + e^{-x}$ ,  $x \in \mathbb{R}$ , and Lemma A.3(ii),

$$\begin{aligned} & |\mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi\}}) \mathbb{1}_{\{\eta_n^\ell(\xi)=k\}}]| \\ & \leq \mathbb{P}\left(|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}| \geq C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1}\gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \\ & \leq \exp\left(-\lambda C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1}\gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \mathbb{E}[\exp(\lambda|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}|)] \\ & \leq 2 \exp\left(-\lambda C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1}\gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right) + \mathfrak{g}\lambda^2 h_{\ell+k}\right). \end{aligned}$$

Minimizing the above upper bound with respect to  $\lambda$  yields

$$\begin{aligned} & |\mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi\}}) \mathbb{1}_{\{\eta_n^\ell(\xi)=k\}}]| \\ & \leq 2 \exp\left(-\frac{C_{\text{ad}}^2 h_{\theta\ell(r-1)+k}^{\frac{2}{r}} \ln\left((\gamma_1^{-1}\gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)}{4\mathfrak{g}h_{\ell+k}}\right) \\ & = 2\left(\frac{\gamma_n}{\gamma_1} h_{\ell+k}^{1+\theta}\right)^{\frac{C_{\text{ad}}^2 h_{\theta\ell(r-1)+k}^{\frac{2}{r}}}{8\mathfrak{g}h_{\ell+k}}}. \end{aligned}$$

Observe that  $\frac{\gamma_n}{\gamma_1} h_{\ell+k}^{1+\theta} \leq 1$ . Now, consider the condition (4.6) on  $h_0$ ,  $r$  and  $\theta$ . On the one hand, one has  $\theta \leq 1 \leq \frac{r}{2(r-1)}$ , so that  $-2\theta(1 - \frac{1}{r}) + 1 \geq 0$ , i.e.  $M^{-2\theta\ell(1-\frac{1}{r})+\ell} \geq 1$ . On the other hand, one also has  $\frac{C_{\text{ad}}^2 h_0^{\frac{2}{r}-1}}{8g} \geq 1$ . Thus

$$\begin{aligned} & |\mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi\}}) \mathbb{1}_{\{\eta_n^\ell(\xi)=k\}}]| \\ & \leq 2 \left( \frac{\gamma_n}{\gamma_1} h_{\ell+k}^{1+\theta} \right) \frac{C_{\text{ad}}^2 h_0^{\frac{2}{r}-1}}{8g} M^{-2\theta\ell(1-\frac{1}{r})+\ell} \\ & \leq C_{21} \gamma_n h_{\ell+k}^{1+\theta}, \end{aligned}$$

where

$$C_{21} = 2\gamma_1^{-1}.$$

Therefore, by (B.1),

$$|\mathbb{P}(X^{h_{\ell+\eta_n^\ell(\xi)}} \leq \xi) - F_{X^{h_{\ell+\lceil\theta\ell\rceil}}(\xi)}| \leq C_{22} \gamma_n h_\ell^{1+\theta}. \quad (\text{B.7})$$

where

$$C_{22} = \frac{C_{21}}{1 - M^{-1}}.$$

(ii) Via (B.4), the definition (3.10) for the case (3.9), the definition (3.8) and Markov's exponential inequality,

$$\begin{aligned} & |\mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi\}}) \mathbb{1}_{\{\eta_n^\ell(\xi)=k\}}]| \\ & \leq \mathbb{P}\left(|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}| \geq C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1} \gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \\ & = \mathbb{P}\left(|G_{h_{\ell+\lceil\theta\ell\rceil}}^{h_{\ell+k}}| \geq C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} h_{\ell+k}^{-\frac{1}{2}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1} \gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \\ & \leq \exp\left(-v_0 C_{\text{ad}}^2 h_{\theta\ell(r-1)+k}^{\frac{2}{r}} h_{\ell+k}^{-1} \ln\left((\gamma_1^{-1} \gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \mathbb{E}[\exp(v_0 |G_{h_{\ell+\lceil\theta\ell\rceil}}^{h_{\ell+k}}|^2)]. \end{aligned}$$

Noting that  $\frac{\gamma_n}{\gamma_1} h_{\ell+k}^{1+\theta} \leq 1$ , and using the condition (4.6) on  $h_0$ ,  $r$  and  $\theta$ , which implies in particular that  $\theta \leq 1 \leq \frac{r}{2(r-1)}$ , so that  $-2\theta(1 - \frac{1}{r}) + 1 \geq 0$ , hence  $M^{-2\theta\ell(1-\frac{1}{r})+\ell} \geq 1$ , and also that  $\frac{1}{2} v_0 C_{\text{ad}}^2 h_0^{\frac{2}{r}-1} \geq 1$ , one has

$$\begin{aligned} & |\mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi\}}) \mathbb{1}_{\{\eta_n^\ell(\xi)=k\}}]| \\ & \leq \left( \sup_{0 \leq h < h' \in \overline{\mathcal{H}}} \mathbb{E}[\exp(v_0 |G_h^{h'}|^2)] \right) \left( \frac{\gamma_n}{\gamma_1} h_{\ell+k}^{1+\theta} \right)^{\frac{1}{2} v_0 C_{\text{ad}}^2 h_0^{\frac{2}{r}-1} M^{-2\theta\ell(1-\frac{1}{r})+\ell}} \\ & \leq C_{31} \gamma_n h_{\ell+k}^{1+\theta}, \end{aligned}$$

where

$$C_{31} = \gamma_1^{-1} \left( \sup_{0 \leq h < h' \in \overline{\mathcal{H}}} \mathbb{E}[\exp(v_0 |G_h^{h'}|^2)] \right).$$

Thus, by (B.1),

$$|\mathbb{P}(X^{h_{\ell+\eta_n^\ell(\xi)}} \leq \xi) - F_{X^{h_{\ell+\lceil\theta\ell\rceil}}(\xi)}| \leq C_{32} \gamma_n h_\ell^{1+\theta}. \quad (\text{B.8})$$

where

$$C_{32} = \frac{C_{31}}{1 - M^{-1}}.$$

□

*Proof of Proposition 4.4.* In this proof, we omit the superscript  $\mu$  from the Lyapunov function  $\mathcal{L}_{h,q}^\mu$  and the parameters  $\lambda_{h,q}^\mu$  and  $\chi_{h,q}^\mu$ , as defined in Lemma A.1, and simply denote  $\mathcal{L}_{h,q}$ ,  $\lambda_{h,q}$  and  $\chi_{h,q}$ . For  $\ell \in \mathbb{N}$ , let  $(\tilde{\mathcal{F}}_n^{h_\ell})_{n \geq 0}$  be the filtration defined by  $\tilde{\mathcal{F}}_0^{h_\ell} = \sigma(\tilde{\xi}_0^{h_\ell})$  and  $\tilde{\mathcal{F}}_n^{h_\ell} = \sigma(\tilde{\xi}_0^{h_\ell}, X_1^{h_\ell + \lceil \theta \ell \rceil}, \dots, X_n^{h_\ell + \lceil \theta \ell \rceil})$ ,  $n \in \mathbb{N}^*$ .

*Step 1. Linearization of the adNSA dynamics.*

Let  $\ell, n \in \mathbb{N}^*$ . Define

$$v_n^\ell(\xi) := \mathbb{E}[H(\xi, X^{h_\ell + \eta_n^\ell(\xi)})] = 1 - \frac{1}{1 - \alpha} \mathbb{P}(X^{h_\ell + \eta_n^\ell(\xi)} \geq \xi), \quad \xi \in \mathbb{R}. \quad (\text{B.9})$$

The dynamics (4.1) can then be rewritten as

$$\tilde{\xi}_{n+1}^{h_\ell} = \tilde{\xi}_n^{h_\ell} - \gamma_{n+1} V'_{h_\ell + \lceil \theta \ell \rceil}(\tilde{\xi}_n^{h_\ell}) - \gamma_{n+1} r_{n+1}^{h_\ell} - \gamma_{n+1} e_{n+1}^{h_\ell}, \quad (\text{B.10})$$

where

$$r_{n+1}^{h_\ell} := v_{n+1}^\ell(\tilde{\xi}_n^{h_\ell}) - V'_{h_\ell + \lceil \theta \ell \rceil}(\tilde{\xi}_n^{h_\ell}), \quad (\text{B.11})$$

$$e_{n+1}^{h_\ell} := H(\tilde{\xi}_n^{h_\ell}, \tilde{X}_{n+1}^{h_\ell}) - v_{n+1}^\ell(\tilde{\xi}_n^{h_\ell}). \quad (\text{B.12})$$

By Lemma 4.2,

$$|r_{n+1}^{h_\ell}| = \frac{1}{1 - \alpha} |\mathbb{P}(X^{h_\ell + \eta_{n+1}^\ell(\xi)} \leq \xi)_{|\xi = \tilde{\xi}_n^{h_\ell}} - \mathbb{P}(X^{h_\ell + \lceil \theta \ell \rceil} \leq \xi)_{|\xi = \tilde{\xi}_n^{h_\ell}}| \leq CM h_{\ell + \lceil \theta \ell \rceil} \tilde{u}_{n+1}, \quad (\text{B.13})$$

where

$$\tilde{u}_{n+1} = \begin{cases} u_{n+1} & \text{if (4.4) holds,} \\ \gamma_{n+1} & \text{if (4.5) or (4.6) holds.} \end{cases} \quad (\text{B.14})$$

and

$$C = \begin{cases} C_1 & \text{if (4.4) holds,} \\ C_2 & \text{if (4.5) holds,} \\ C_3 & \text{if (4.6) holds,} \end{cases} \quad (\text{B.15})$$

with  $C_1$ ,  $C_2$  and  $C_3$  being the positive constants defined in Lemma 4.2.

*Step 2. Testing the Lyapunov function.*

Let  $\mu \in \mathbb{R}_+$  and  $q \in \mathbb{N}^*$ . In the spirit of [14, Theorem 3.7], we test the Lyapunov  $\mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q}$  along the dynamics (B.10). Using a second order Taylor expansion,

$$\begin{aligned} \mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q}(\tilde{\xi}_{n+1}^{h_\ell}) &= \mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q}(\tilde{\xi}_n^{h_\ell} - \gamma_{n+1} V'_{h_\ell + \lceil \theta \ell \rceil}(\tilde{\xi}_n^{h_\ell}) - \gamma_{n+1} r_{n+1}^{h_\ell} - \gamma_{n+1} e_{n+1}^{h_\ell}) \\ &= \mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q}(\tilde{\xi}_n^{h_\ell}) - \gamma_{n+1} \mathcal{L}'_{h_\ell + \lceil \theta \ell \rceil, q}(\tilde{\xi}_n^{h_\ell})(V'_{h_\ell + \lceil \theta \ell \rceil}(\tilde{\xi}_n^{h_\ell}) + r_{n+1}^{h_\ell} + e_{n+1}^{h_\ell}) \\ &\quad + \gamma_{n+1}^2 H(\tilde{\xi}_n^{h_\ell}, \tilde{X}_{n+1}^{h_\ell})^2 \int_0^1 (1-t) \mathcal{L}''_{h_\ell + \lceil \theta \ell \rceil, q}(t \tilde{\xi}_{n+1}^{h_\ell} + (1-t) \tilde{\xi}_n^{h_\ell}) dt. \end{aligned} \quad (\text{B.16})$$

It follows from Lemmas A.1(iii)-(iv), (A.2) and (B.13) that

$$\begin{aligned} \mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q}(\tilde{\xi}_{n+1}^{h_\ell}) &\leq \mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q}(\tilde{\xi}_n^{h_\ell}) \\ &\quad \times (1 - \lambda_{h_\ell + \lceil \theta \ell \rceil, q} \gamma_{n+1} + CM \mu \|V'_{h_\ell + \lceil \theta \ell \rceil}\|_\infty h_{\ell + \lceil \theta \ell \rceil} \gamma_{n+1} \tilde{u}_{n+1}) \\ &\quad - \gamma_{n+1} \mathcal{L}'_{h_\ell + \lceil \theta \ell \rceil, q}(\tilde{\xi}_n^{h_\ell}) e_{n+1}^{h_\ell} + CM q \|V'_{h_\ell + \lceil \theta \ell \rceil}\|_\infty h_{\ell + \lceil \theta \ell \rceil} \gamma_{n+1} \tilde{u}_{n+1} \mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q-1}(\tilde{\xi}_n^{h_\ell}) \\ &\quad + \chi_{h_\ell + \lceil \theta \ell \rceil, q} \gamma_{n+1}^2 H(\tilde{\xi}_n^{h_\ell}, \tilde{X}_{n+1}^{h_\ell})^2 \\ &\quad \times \int_0^1 (1-t) \left( \mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q}(t \tilde{\xi}_{n+1}^{h_\ell} + (1-t) \tilde{\xi}_n^{h_\ell}) + \mathcal{L}_{h_\ell + \lceil \theta \ell \rceil, q-1}(t \tilde{\xi}_{n+1}^{h_\ell} + (1-t) \tilde{\xi}_n^{h_\ell}) \right) dt. \end{aligned} \quad (\text{B.17})$$

The mean value theorem guarantees that, for all  $t \in [0, 1]$ , there exists  $a_n^{h_\ell}(t) \in \mathbb{R}$  such that

$$V_{h_{\ell+\lceil\theta\ell\rceil}}(t\tilde{\xi}_{n+1}^{h_\ell} + (1-t)\tilde{\xi}_n^{h_\ell}) = V_{h_{\ell+\lceil\theta\ell\rceil}}(\tilde{\xi}_n^{h_\ell}) + tV'_{h_{\ell+\lceil\theta\ell\rceil}}(a_n^{h_\ell}(t))(\tilde{\xi}_{n+1}^{h_\ell} - \tilde{\xi}_n^{h_\ell}). \quad (\text{B.18})$$

Besides, from (2.3),

$$|\tilde{\xi}_{n+1}^{h_\ell} - \tilde{\xi}_n^{h_\ell}| = \gamma_{n+1}|H(\tilde{\xi}_n^{h_\ell}, \tilde{X}_{n+1}^{h_\ell})| \leq k_\alpha \gamma_{n+1}, \quad (\text{B.19})$$

with  $k_\alpha = \frac{\alpha}{1-\alpha} \vee 1$ . From (B.18), (B.19) and the fact that  $\|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty \leq k_\alpha$ , it ensues that, for all  $t \in [0, 1]$ ,

$$\begin{aligned} & V_{h_{\ell+\lceil\theta\ell\rceil}}(t\tilde{\xi}_{n+1}^{h_\ell} + (1-t)\tilde{\xi}_n^{h_\ell}) - V_{h_{\ell+\lceil\theta\ell\rceil}}(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) \\ & \leq V_{h_{\ell+\lceil\theta\ell\rceil}}(\tilde{\xi}_n^{h_\ell}) - V_{h_{\ell+\lceil\theta\ell\rceil}}(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) + k_\alpha^2 \gamma_{n+1}. \end{aligned} \quad (\text{B.20})$$

Using the inequality  $e^x \leq e\mathbb{1}_{x \leq 1} + x^q e^x \mathbb{1}_{x > 1} \leq e(1 + x^q e^x)$ ,  $x \geq 0$ , and the very definition (A.1) of  $\mathcal{L}_{h,q}$ , it follows

$$\mathcal{L}_{h,0}^\mu(\xi) \leq e(1 + \mu^q \mathcal{L}_{h,q}^\mu(\xi)), \quad \xi \in \mathbb{R}, \quad \mu \in \mathbb{R}_+, \quad h \in \overline{\mathcal{H}}, \quad q \in \mathbb{N}^*. \quad (\text{B.21})$$

Thus, by (B.20) and (B.21), for all  $t \in [0, 1]$ ,

$$\begin{aligned} & \mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q}(t\tilde{\xi}_{n+1}^{h_\ell} + (1-t)\tilde{\xi}_n^{h_\ell}) \\ & \leq (V_{h_{\ell+\lceil\theta\ell\rceil}}(\tilde{\xi}_n^{h_\ell}) - V_{h_{\ell+\lceil\theta\ell\rceil}}(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) + k_\alpha^2 \gamma_{n+1})^q \\ & \quad \times \exp\left(\mu(V_{h_{\ell+\lceil\theta\ell\rceil}}(\tilde{\xi}_n^{h_\ell}) - V_{h_{\ell+\lceil\theta\ell\rceil}}(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) + k_\alpha^2 \gamma_{n+1})\right) \\ & \leq 2^{q-1} \exp(\mu k_\alpha^2 \gamma_{n+1}) (\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell}) + k_\alpha^{2q} \gamma_{n+1}^q \mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},0}(\tilde{\xi}_n^{h_\ell})) \\ & \leq \sigma_q^\mu (\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell}) + \gamma_{n+1}^q), \end{aligned} \quad (\text{B.22})$$

with

$$\sigma_q^\mu := 2^{q-1} \exp(\mu k_\alpha^2 \gamma_1) ((1 + e\mu^q k_\alpha^{2q} \gamma_1^q) \vee e k_\alpha^{2q}) \geq 2^{q-1}, \quad \mu \in \mathbb{R}_+, \quad h \in \overline{\mathcal{H}}, \quad q \in \mathbb{N}^*.$$

Plugging the upper bounds of (B.19) and (B.22) into (B.17) yields

$$\begin{aligned} & \mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_{n+1}^{h_\ell}) \leq \mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell}) \\ & \quad \times (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},q} \gamma_{n+1} + (CM\mu \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil} \gamma_{n+1} \tilde{u}_{n+1} + \nu_{h_{\ell+\lceil\theta\ell\rceil},q}^\mu \gamma_{n+1}^2)) \\ & \quad + (CMq \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty \gamma_{n+1} \tilde{u}_{n+1} h_{\ell+\lceil\theta\ell\rceil} + \nu_{h_{\ell+\lceil\theta\ell\rceil},q}^\mu \gamma_{n+1}^2) \mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q-1}(\tilde{\xi}_n^{h_\ell}) \\ & \quad - \gamma_{n+1} \mathcal{L}'_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell}) e_{n+1}^{h_\ell} + \nu_{h_{\ell+\lceil\theta\ell\rceil},q}^\mu \gamma_{n+1}^{q+1}, \end{aligned} \quad (\text{B.23})$$

where

$$\nu_{h,q}^\mu := \frac{1}{2} \chi_{h,q}^\mu k_\alpha^2 ((\gamma_1 \sigma_{h,q}^\mu + \sigma_{h,q-1}) \vee \sigma_{h,q}^\mu), \quad \mu \in \mathbb{R}_+, \quad h \in \overline{\mathcal{H}}, \quad q \in \mathbb{N}^*. \quad (\text{B.24})$$

Hereafter, to simplify notation, we will simply denote  $\nu_{h,q}$  instead of  $\nu_{h,q}^\mu$ .

*Step 3. Inequality on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell})]$ .*

Via the tower law,

$$\mathbb{E}[\mathcal{L}'_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell}) e_{n+1}^{h_\ell}] = \mathbb{E}[\mathcal{L}'_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell}) \mathbb{E}[e_{n+1}^{h_\ell} | \tilde{\mathcal{F}}_n^{h_\ell}]]. \quad (\text{B.25})$$

For  $n \in \mathbb{N}$ , by the independence of the  $(X_{n+1}^{h_\ell+k})_{1 \leq k \leq \lceil\theta\ell\rceil}$  and  $\tilde{\mathcal{F}}_n^{h_\ell}$ ,

$$\begin{aligned} & \mathbb{E}[e_{n+1}^{h_\ell} | \tilde{\mathcal{F}}_n^{h_\ell}] \\ & = \mathbb{E}[H(\tilde{\xi}_n^{h_\ell}, X_{n+1}^{h_\ell+\eta_{n+1}^\ell(\tilde{\xi}_n^{h_\ell})}) | \tilde{\mathcal{F}}_n^{h_\ell}] - \mathbb{E}[H(\xi, X_{n+1}^{h_\ell+\eta_{n+1}^\ell(\xi)})]_{|\xi=\tilde{\xi}_n^{h_\ell}} \\ & = \mathbb{E}\left[\sum_{k=1}^{\lceil\theta\ell\rceil} H(\tilde{\xi}_n^{h_\ell}, X_{n+1}^{h_\ell+k}) \mathbb{1}_{\{\eta_{n+1}^\ell(\tilde{\xi}_n^{h_\ell})=k\}} \middle| \tilde{\mathcal{F}}_n^{h_\ell}\right] - \sum_{k=1}^{\lceil\theta\ell\rceil} \mathbb{E}[H(\xi, X_{n+1}^{h_\ell+k}) \mathbb{1}_{\{\eta_{n+1}^\ell(\xi)=k\}}]_{|\xi=\tilde{\xi}_n^{h_\ell}} \\ & = 0. \end{aligned}$$

Hence  $(e_n^{h_\ell})_{n \geq 1}$  is a  $((\tilde{\mathcal{F}}_n^{h_\ell})_{n \geq 0}, \mathbb{P})$ -martingale increment sequence, and, by (B.25),

$$\mathbb{E}[\mathcal{L}'_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell})e_{n+1}^{h_\ell}] = 0.$$

Thus, by taking the expectation in both sides of the inequality (B.23),

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_{n+1}^{h_\ell})] &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q}(\tilde{\xi}_n^{h_\ell})] \\ &\quad \times (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},q}\gamma_{n+1} + (CM\mu\|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil}\gamma_{n+1}\tilde{u}_{n+1} + \nu_{h_{\ell+\lceil\theta\ell\rceil},q}\gamma_{n+1}^2)) \\ &\quad + (CMq\|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil}\gamma_{n+1}\tilde{u}_{n+1} + \nu_{h_{\ell+\lceil\theta\ell\rceil},q}\gamma_{n+1}^2)\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},q-1}(\tilde{\xi}_n^{h_\ell})] \\ &\quad + \nu_{h_{\ell+\lceil\theta\ell\rceil},q}\gamma_{n+1}^{q+1}. \end{aligned} \tag{B.26}$$

(i) We hereby aim to use Lemma A.1(v) with  $h = h_{\ell+\lceil\theta\ell\rceil}$  and  $q = 1$ . To that end, we provide below sharper controls for  $\mathbb{E}[\tilde{\mathcal{L}}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_n^{h_\ell})]$  and  $\mathbb{E}[\tilde{\mathcal{L}}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})]$ . According to Lemma A.1(ii), by assumption, for  $\mu \in \{\mu_{h_{\ell+\lceil\theta\ell\rceil},q'}, q' \in \{1, 2\}\}$ ,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_0^{h_\ell})] \vee \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_0^{h_\ell})] \\ \leq \mathbb{E}\left[ (1 + 2k_\alpha^2(|\tilde{\xi}_0^{h_\ell}|^2 + |\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}|^2)) \right. \\ \left. \times \exp\left(\frac{2k_\alpha}{1-\alpha}\|f_{X^{h_{\ell+\lceil\theta\ell\rceil}}}\|_\infty(|\tilde{\xi}_0^{h_\ell}| + |\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}|)\right) \right] < \infty. \end{aligned}$$

*Step 1. Inequality on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_n^{h_\ell})]$ .*

Taking  $q = 1$  in (B.26) and (B.21),

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_{n+1}^{h_\ell})] &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_n^{h_\ell})] \\ &\quad \times (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_{n+1} + (1 + e\mu)(CM\mu\|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil}\gamma_{n+1}\tilde{u}_{n+1} + \nu_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_{n+1}^2)) \\ &\quad + CM e\|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil}\gamma_{n+1}\tilde{u}_{n+1} + (1 + e\mu)\nu_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_{n+1}^2. \end{aligned}$$

Let

$$\delta' = \begin{cases} \delta & \text{if (4.4) holds,} \\ \beta & \text{if (4.5) or (4.6) holds.} \end{cases} \tag{B.27}$$

Note that, by (B.14),

$$\tilde{u}_n = \frac{\gamma_1}{n^{\delta'}} = \gamma_1^{1-\frac{\delta'}{\beta}} \gamma_n^{\frac{\delta'}{\beta}}, \quad n \in \mathbb{N}^*, \tag{B.28}$$

so that

$$\gamma_n \tilde{u}_n = \frac{\gamma_1^2}{n^{\beta+\delta'}} = \gamma_1^{1-\frac{\delta'}{\beta}} \gamma_n^{1+\frac{\delta'}{\beta}}, \quad n \in \mathbb{N}^*, \tag{B.29}$$

hence

$$\gamma_n \tilde{u}_n \vee \gamma_n^2 = \frac{\gamma_1^2}{n^{\beta+(\delta' \wedge \beta)}} = \gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_n^{1+\frac{\delta' \wedge \beta}{\beta}}, \quad n \in \mathbb{N}^*. \tag{B.30}$$

Therefore, using in particular that  $\gamma_n \tilde{u}_n \leq \gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_n^{1+\frac{\delta' \wedge \beta}{\beta}}$ ,  $n \in \mathbb{N}^*$ ,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_{n+1}^{h_\ell})] &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_n^{h_\ell})] \\ &\quad \times \left( 1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_{n+1} \right. \\ &\quad \left. + (1 + e\mu)(CM\mu\|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil} + \nu_{h_{\ell+\lceil\theta\ell\rceil},1})\gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_{n+1}^{1+\frac{\delta' \wedge \beta}{\beta}} \right) \\ &\quad + CM\gamma_1^{1-\frac{\delta'}{\beta}} e\|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil}\gamma_{n+1}^{1+\frac{\delta'}{\beta}} + (1 + e\mu)\nu_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_{n+1}^2. \end{aligned} \tag{B.31}$$

Recalling the definition of  $\chi_{h,q}^\mu$  from Lemma A.1(iv) and the definition of  $\nu_{h,q}^\mu$  given in (B.24), let

$$c_{h,q}^\mu := \frac{\nu_{h,q}^\mu}{\chi_{h,q}^\mu} = \frac{1}{2}(\sigma_q^\mu \vee (\gamma_1 \sigma_q^\mu + \sigma_{q-1}^\mu)) k_\alpha^2 > \frac{1}{2}, \quad \bar{c}_{h,q} := c_{h,q}^{\mu_{h,q}},$$

$$\mu \in \mathbb{R}_+, \quad h \in \bar{\mathcal{H}}, \quad q \in \mathbb{N}^*. \quad (\text{B.32})$$

Let  $q' \geq q$  and set  $\mu = \mu_{h,q'}$ . Recalling that  $|\lambda_{h,q}^{\mu_{h,q'}}|^2 \leq \chi_{h,q}^{\mu_{h,q'}}$  according to Lemma A.1(iv), and that  $\gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_n^{1+\frac{\delta' \wedge \beta}{\beta}} \geq \gamma_n^2$ ,  $n \in \mathbb{N}^*$ , according to (B.30), it ensues that, for all  $n \in \mathbb{N}^*$ ,

$$\begin{aligned} 1 - \lambda_{h,q}^{\mu_{h,q'}} \gamma_n + (1 + e\mu_{h,q'}^q)(CM\mu_{h,q'} \|V'_h\|_\infty h + \nu_{h,q}^{\mu_{h,q'}}) \gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_n^{1+\frac{\delta' \wedge \beta}{\beta}} \\ \geq 1 - \lambda_{h,q}^{\mu_{h,q'}} \gamma_n + \nu_{h,q}^{\mu_{h,q'}} \gamma_n^2 \\ \geq 1 - (\chi_{h,q}^{\mu_{h,q'}})^{\frac{1}{2}} \gamma_{n+1} + \chi_{h,q}^{\mu_{h,q'}} c_{h,q}^{\mu_{h,q'}} \gamma_n^2 \\ \geq 1 - \frac{1}{4c_{h,q}^{\mu_{h,q'}}} \\ \geq \frac{1}{2} > 0. \end{aligned} \quad (\text{B.33})$$

Therefore, for  $q' \geq 1$ , setting  $\mu = \mu_{h_{\ell+[\theta\ell]},q'}$  and iterating  $n$  times the inequality (B.31),

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_n^{h_\ell})] &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_0^{h_\ell})] \Pi_{1:n}^{h_{\ell+[\theta\ell]},1,q'} \\ &\quad + CM\gamma_1^{1-\frac{\delta'}{\beta}} e \|V'_{h_{\ell+[\theta\ell]}}\|_\infty \sum_{k=1}^n h_{\ell+[\theta\ell]} \gamma_k^{1+\frac{\delta'}{\beta}} \Pi_{k+1:n}^{h_{\ell+[\theta\ell]},1,q'} \\ &\quad + (1 + e\mu_{h_{\ell+[\theta\ell]},q'}) \nu_{h_{\ell+[\theta\ell]},1} \sum_{k=1}^n \gamma_k^2 \Pi_{k+1:n}^{h_{\ell+[\theta\ell]},1,q'}, \end{aligned} \quad (\text{B.34})$$

where

$$\Pi_{k:n}^{h,q,q'} := \prod_{j=k}^n \left( 1 - \lambda_{h,q}^{\mu_{h,q'}} \gamma_j + (1 + e\mu_{h,q'}^q)(CM\mu_{h,q'} \|V'_h\|_\infty h + \nu_{h,q}^{\mu_{h,q'}}) \gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_j^{1+\frac{\delta' \wedge \beta}{\beta}} \right),$$

$$h \in \bar{\mathcal{H}}, \quad q \in \mathbb{N}^*, \quad q' \geq q, \quad k, n \in \mathbb{N}^*,$$

with the convention  $\prod_{\emptyset} = 1$ .

If  $\gamma_n = \gamma_1 n^{-1}$ ,  $n \in \mathbb{N}^*$ , then, applying Lemma A.2,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_n^{h_\ell})] &\leq \tilde{C}_{1,q'}^{\ell,1} \frac{\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_0^{h_\ell})]}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1}} + \tilde{C}_{1,q'}^{\ell,1} \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1 - \delta'}(n+1)}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1}} h_{\ell+[\theta\ell]} \\ &\quad + \tilde{C}_{1,q'}^{\ell,1} \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1 - 1}(n+1)}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1}}, \end{aligned}$$

where

$$\begin{aligned} \tilde{C}_{1,q'}^{\ell,1} &= \exp \left( \zeta_{h_{\ell+[\theta\ell]},1,q'}^{1,\delta'} \left( 1 + \frac{1}{\delta'} \right) \gamma_1^{1+\delta'} + \frac{\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1}{2} \right), \\ \bar{C}_{1,q'}^{\ell,1} &= 2^{(\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1) \vee (1+\delta')} C \tilde{C}_{1,q'}^{\ell,1} M e \|V'_{h_{\ell+[\theta\ell]}}\|_\infty \gamma_1^2, \\ \tilde{C}_{1,q'}^{\ell,1} &= 2^{(\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1) \vee 2} (1 + e\mu_{h_{\ell+[\theta\ell]},q'}) \nu_{h_{\ell+[\theta\ell]},1} \tilde{C}_{1,q'}^{\ell,1} \gamma_1^2, \end{aligned}$$

with

$$\zeta_{h,q,q'}^{\beta,\delta} := (1 + e\mu_{h,q'}^q)(CM\mu_{h,q'}\|V'_h\|_\infty h + \nu_{h,q}^{\mu_{h,q'}})\gamma_1^{1-\frac{\delta\wedge\beta}{\beta}},$$

$$h \in \overline{\mathcal{H}}, \quad q \in \mathbb{N}^*, \quad q' \geq q, \quad \beta, \delta \in (0, 1].$$

If  $\bar{\lambda}_{h_{\ell+[\theta\ell]},1}\gamma_1 > 1$ , so that, for  $q' \geq 1$ ,  $\lambda_{h_{\ell+[\theta\ell]},1}^{\mu_{h_{\ell+[\theta\ell]},q'}}\gamma_1 \geq \bar{\lambda}_{h_{\ell+[\theta\ell]},1}\gamma_1 > 1 \geq \delta'$ , then, via the definition (A.3),

$$\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_n^{h_\ell})] \leq C_{1,q'}^{\ell,1}((h_\ell^{1+\theta}\tilde{u}_n) \vee \gamma_n), \quad (\text{B.35})$$

where

$$C_{1,q'}^{\ell,1} = \gamma_1^{-1}(\bar{C}_{1,q'}^{\ell,1}\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_0^{h_\ell})] + h_0^{-1}\bar{C}_{1,q'}^{\ell,1} + \tilde{C}_{1,q'}^{\ell,1}).$$

Likewise, if  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\beta \in (0, 1)$ , then, by Lemma A.2,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_n^{h_\ell})] &\leq (\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_0^{h_\ell})] + \bar{C}_{1,q'}^{\ell,\beta} + \tilde{C}_{1,q'}^{\ell,\beta}) \\ &\times \exp\left(2^{1+\beta+\delta'\wedge\beta}\zeta_{h_{\ell+[\theta\ell]},1,q'}^{\beta,\delta'}\gamma_1^{1+\frac{\delta'\wedge\beta}{\beta}}\Phi_{1-\beta-\delta'\wedge\beta}(n+1) - \frac{\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1}{2}\Phi_{1-\beta}(n+1)\right) \\ &+ 2^\beta(\bar{C}_{1,q'}^{\ell,\beta}\gamma_1^{1+\frac{\delta'}{\beta}}h_{\ell+[\theta\ell]} + \tilde{C}_{1,q'}^{\ell,\beta}\gamma_1^2)\exp(-2^{-(\beta+2)}\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1 n^{1-\beta})\Phi_{1-\beta}(n+1) \\ &+ \bar{C}_{1,q'}^{\ell,\beta}\frac{2^{1+\delta'}\gamma_1^{\frac{\delta'}{\beta}}}{\lambda_{h_{\ell+[\theta\ell]},1}n^{\delta'}}h_{\ell+[\theta\ell]} + \tilde{C}_{1,q'}^{\ell,\beta}\frac{2^{1+\beta}\gamma_1}{\lambda_{h_{\ell+[\theta\ell]},1}n^\beta}, \end{aligned}$$

where

$$\begin{aligned} \bar{C}_{1,q'}^{\ell,\beta} &= (\zeta_{h_{\ell+[\theta\ell]},1,q'}^{\beta,\delta'})^{-1}\left(\gamma_1^{\frac{\delta'-(\delta'\wedge\beta)}{\beta}} \vee \left(\frac{\lambda_{h_{\ell+[\theta\ell]},1}}{2\zeta_{h_{\ell+[\theta\ell]},1,q'}^{\beta,\delta'}}\right)^{\frac{\delta'-(\delta'\wedge\beta)}{\delta'\wedge\beta}}\right)CM\gamma_1^{1-\frac{\delta'}{\beta}}e\|V'_{h_{\ell+[\theta\ell]}}\|_\infty, \\ \tilde{C}_{1,q'}^{\ell,\beta} &= (\zeta_{h_{\ell+[\theta\ell]},1,q'}^{\beta,\delta'})^{-1}\left(\gamma_1^{\frac{\beta-(\delta'\wedge\beta)}{\beta}} \vee \left(\frac{\lambda_{h_{\ell+[\theta\ell]},1}}{2\zeta_{h_{\ell+[\theta\ell]},1,q'}^{\beta,\delta'}}\right)^{\frac{\beta-(\delta'\wedge\beta)}{\delta'\wedge\beta}}\right)(1 + e\mu_{h_{\ell+[\theta\ell]},q'})\nu_{h_{\ell+[\theta\ell]},1}. \end{aligned}$$

Hence

$$\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_n^{h_\ell})] \leq C_{1,q'}^{\ell,\beta}((h_\ell^{1+\theta}\tilde{u}_n) \vee \gamma_n), \quad (\text{B.36})$$

where

$$\begin{aligned} C_{1,q'}^{\ell,\beta} &= \gamma_1^{-1}(\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},1}(\tilde{\xi}_0^{h_\ell})] + \bar{C}_{1,q'}^{\ell,\beta} + \tilde{C}_{1,q'}^{\ell,\beta}) \\ &\times \sup_{m \geq 1} \left\{ m^\beta \exp\left(2^{1+\beta+\delta'\wedge\beta}\zeta_{h_{\ell+[\theta\ell]},1,q'}^{\beta,\delta'}\gamma_1^{1+\frac{\delta'\wedge\beta}{\beta}}\Phi_{1-\beta-\delta'\wedge\beta}(m+1)\right) \right. \\ &\quad \left. \times \exp\left(-\frac{\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1}{2}\Phi_{1-\beta}(m+1)\right) \right\} \\ &+ 2^\beta(\bar{C}_{1,q'}^{\ell,\beta}\gamma_1^{\frac{\delta'}{\beta}}h_{\ell+[\theta\ell]} + \tilde{C}_{1,q'}^{\ell,\beta}\gamma_1) \\ &\quad \times \sup_{m \geq 1} \{m^\beta \exp(-2^{-(\beta+2)}\lambda_{h_{\ell+[\theta\ell]},1}\gamma_1 m^{1-\beta})\Phi_{1-\beta}(m+1)\} \\ &+ \bar{C}_{1,q'}^{\ell,\beta}\frac{2^{1+\delta'}\gamma_1^{-1+\frac{\delta'}{\beta}}}{h_0\lambda_{h_{\ell+[\theta\ell]},1}} + \tilde{C}_{1,q'}^{\ell,\beta}\frac{2^{1+\beta}}{\lambda_{h_{\ell+[\theta\ell]},1}}, \end{aligned}$$

Note that  $\sup_{\ell \geq 1} C_{1,q'}^{\ell,\beta} < \infty$ ,  $q' \in \{1, 2\}$ ,  $\beta \in (0, 1]$ .

Step 2. Inequality on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})]$ .

Via (B.35) and (B.36), assuming that  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_1 > 1$  if  $\beta = 1$ , for  $q' \geq 1$ ,

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_n^{h_\ell})] \leq C_{1,q'}^{\ell,\beta} (2^{\delta'} M h_{\ell+\lceil\theta\ell\rceil} \tilde{u}_{n+1} + 2^\beta \gamma_{n+1}). \quad (\text{B.37})$$

Thus, taking  $q = 2$ ,  $h = h_{\ell+\lceil\theta\ell\rceil}$  and  $\mu = \mu_{h_{\ell+\lceil\theta\ell\rceil},q'}$ , with  $q' \geq 2$ , in (B.26), and using (B.30) and (B.37),

$$\begin{aligned} & \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_{n+1}^{h_\ell})] \leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})] \\ & \times \left( 1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_{n+1} + (CM \mu_{h_{\ell+\lceil\theta\ell\rceil},q'} \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil} + \nu_{h_{\ell+\lceil\theta\ell\rceil},2}) \gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_{n+1}^{\frac{1+\delta' \wedge \beta}{\beta}} \right) \\ & + 2^{1+\delta'} CM^2 \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty C_{1,q'}^{\ell,\beta} h_{\ell+\lceil\theta\ell\rceil}^2 \gamma_{n+1} \tilde{u}_{n+1}^2 \\ & + C_{1,q'}^{\ell,\beta} M (2^{1+\beta} C \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty + 2^{\delta'} \nu_{h_{\ell+\lceil\theta\ell\rceil},2}) h_{\ell+\lceil\theta\ell\rceil} \gamma_{n+1}^2 \tilde{u}_{n+1} \\ & + (2^\beta C_{1,q'}^{\ell,\beta} + 1) \nu_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_{n+1}^3 \\ & \leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})] (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_{n+1} + \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \gamma_{n+1}^{\frac{1+\delta' \wedge \beta}{\beta}}) \\ & + d_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \gamma_{n+1} (h_{\ell+\lceil\theta\ell\rceil} \tilde{u}_{n+1} + \gamma_{n+1})^2 \\ & \leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})] (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_{n+1} + \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \gamma_{n+1}^{\frac{1+\delta' \wedge \beta}{\beta}}) \\ & + 2d_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \gamma_1^{2-2\frac{\delta'}{\beta}} h_{\ell+\lceil\theta\ell\rceil}^2 \gamma_{n+1}^{1+2\frac{\delta'}{\beta}} + 2d_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \gamma_{n+1}^3, \end{aligned} \quad (\text{B.38})$$

where we used

$$\begin{aligned} \bar{\zeta}_{h,q,q'}^{\beta,\delta} &= (CM \mu_{h,q'} \|V'_h\|_\infty h + \nu_{h,q}^{\mu_{h,q'}}) \gamma_1^{1-\frac{\delta \wedge \beta}{\beta}}, \\ d_{h,q,q'}^{\beta,\delta} &= (2^{1+\delta'} C C_{1,q'}^{\ell,\beta} M^2 \|V'_h\|_\infty) \vee \left( \frac{C_{1,q'}^{\ell,\beta}}{2} (2^{\delta'} M + 2^{1+\beta} CM \|V'_h\|_\infty) \right) \\ & \vee ((2^\beta C_{1,q'}^{\ell,\beta} + 1) \nu_{h,q}^{\mu_{h,q'}}) \\ & \quad h \in \bar{\mathcal{H}}, \quad q \in \mathbb{N}^*, \quad q' \geq q, \quad \beta, \delta \in (0, 1]. \end{aligned}$$

and the fact that

$$\gamma_n \tilde{u}_n^2 = \frac{\gamma_1^3}{n^{\beta+2\delta'}} = \gamma_1^{2-2\frac{\delta'}{\beta}} \gamma_n^{1+2\frac{\delta'}{\beta}}, \quad n \in \mathbb{N}^*. \quad (\text{B.39})$$

Recall that  $\gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_n^{\frac{1+\delta' \wedge \beta}{\beta}} \geq \gamma_n^2$ ,  $n \in \mathbb{N}^*$  by (B.30). Therefore, via (B.33), for all  $n \in \mathbb{N}^*$ ,  $h \in \bar{\mathcal{H}}$ ,  $q \in \mathbb{N}^*$  and  $q' \geq q$ ,

$$\begin{aligned} & 1 - \lambda_{h,q}^{\mu_{h,q'}} \gamma_n + (CM \mu_{h,q'} \|V'_h\|_\infty h + \nu_{h,q}^{\mu_{h,q'}}) \gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_n^{\frac{1+\delta' \wedge \beta}{\beta}} \\ & \geq 1 - \lambda_{h,q}^{\mu_{h,q'}} \gamma_n + \nu_{h,q}^{\mu_{h,q'}} \gamma_n^2 \\ & \geq \frac{1}{2} > 0. \end{aligned} \quad (\text{B.40})$$

Hence, iterating  $n$  times the inequality (B.38) yields, for  $q' \geq 2$ ,

$$\begin{aligned} & \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})] \leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_0^{h_\ell})] \tilde{\Pi}_{1:n}^{h_{\ell+\lceil\theta\ell\rceil},2,q'} \\ & + 2d_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \gamma_1^{2-2\frac{\delta'}{\beta}} \sum_{k=1}^n h_{\ell+\lceil\theta\ell\rceil}^2 \gamma_k^{1+2\frac{\delta'}{\beta}} \tilde{\Pi}_{k+1:n}^{h_{\ell+\lceil\theta\ell\rceil},2,q'} \\ & + 2d_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \sum_{k=1}^n \gamma_k^3 \tilde{\Pi}_{k+1:n}^{h_{\ell+\lceil\theta\ell\rceil},2,q'}, \end{aligned} \quad (\text{B.41})$$

where

$$\begin{aligned} \tilde{\Pi}_{k:n}^{h,q,q'} &:= \prod_{j=k}^n \left( 1 - \lambda_{h,q}^{\mu_{h,q'}} \gamma_j + (CM\mu_{h,q'} \|V'_h\|_\infty h + \nu_{h,q}^{\mu_{h,q'}}) \gamma_1^{1-\frac{\delta' \wedge \beta}{\beta}} \gamma_j^{1+\frac{\delta' \wedge \beta}{\beta}} \right), \\ &h \in \bar{\mathcal{H}}, \quad q \in \mathbb{N}^*, \quad q' \geq q, \quad k, n \in \mathbb{N}^*, \end{aligned} \quad (\text{B.42})$$

with the convention  $\prod_\emptyset = 1$ .

If  $\gamma_n = \gamma_1 n^{-1}$ ,  $n \in \mathbb{N}^*$ , then, by Lemma A.2,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},2}(\tilde{\xi}_n^{h_\ell})] &\leq \widehat{C}_{2,q'}^{\ell,1} \frac{\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},2}(\tilde{\xi}_0^{h_\ell})]}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1}} + \bar{C}_{2,q'}^{\ell,1} \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1 - 2\delta'}(n+1)}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1}} h_{\ell+[\theta\ell]}^2 \\ &+ \tilde{C}_{2,q'}^{\ell,1} \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1 - 2}(n+1)}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1}}, \end{aligned} \quad (\text{B.43})$$

where

$$\begin{aligned} \widehat{C}_{2,q'}^{\ell,1} &= \exp\left(\bar{\zeta}_{h_{\ell+[\theta\ell]},2,q'}^{1,\delta'} \left(1 + \frac{1}{\delta'}\right) \gamma_1^{1+\delta'} + \frac{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1}{2}\right), \\ \bar{C}_{2,q'}^{\ell,1} &= 2^{1+(\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1) \vee (1+\delta')} \widehat{C}_{2,q'}^{\ell,1} d_{h_{\ell+[\theta\ell]},2,q'}^{1,\delta'} \gamma_1^3, \\ \tilde{C}_{2,q'}^{\ell,1} &= 2^{1+(\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1) \vee 2} \widehat{C}_{2,q'}^{\ell,1} d_{h_{\ell+[\theta\ell]},2,q'}^{1,\delta'} \gamma_1^3. \end{aligned}$$

If  $\bar{\lambda}_{h_{\ell+[\theta\ell]},1} \gamma_1 > 1$ , so that  $\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1 \geq \bar{\lambda}_{h_{\ell+[\theta\ell]},2} 2\gamma_1 \geq \bar{\lambda}_{h_{\ell+[\theta\ell]},1} \gamma_1 > 1 \geq \delta'$ , then

$$\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},2}(\tilde{\xi}_n^{h_\ell})] \leq C_{2,q'}^{\ell,1} ((h_\ell^{1+\theta} \tilde{u}_n) \vee \gamma_n), \quad (\text{B.44})$$

where

$$\begin{aligned} C_{2,q'}^{\ell,1} &= \gamma_1^{-1} \left( \widehat{C}_{2,q'}^{\ell,1} \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},2}(\tilde{\xi}_0^{h_\ell})] \right. \\ &+ \frac{\bar{C}_{2,q'}^{\ell,1}}{M^{\ell+[\theta\ell]}} \sup_{m \geq 1} \left\{ m^{\delta'} \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1 - 2\delta'}(m+1)}{(m+1)^{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1}} \right\} \\ &+ \left. \tilde{C}_{2,q'}^{\ell,1} \sup_{m \geq 1} \left\{ m \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1 - 2}(m+1)}{(m+1)^{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1}} \right\} \right). \end{aligned}$$

If  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\beta \in (0, 1)$ , then, as a consequence of Lemma A.2,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},2}(\tilde{\xi}_n^{h_\ell})] &\leq (\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},2}(\tilde{\xi}_0^{h_\ell})] + \bar{C}_{2,q'}^{\ell,\beta} + \tilde{C}_{2,q'}^{\ell,\beta}) \\ &\times \exp\left(2^{1+\beta+\delta' \wedge \beta} \bar{\zeta}_{h_{\ell+[\theta\ell]},2,q'}^{\beta,\delta'} \gamma_1^{1+\frac{\delta' \wedge \beta}{\beta}} \Phi_{1-\beta-\delta' \wedge \beta}(n+1) - \frac{\lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1}{2} \Phi_{1-\beta}(n+1)\right) \\ &+ 2^\beta (\bar{C}_{2,q'}^{\ell,\beta} \gamma_1^{1+2\frac{\delta'}{\beta}} h_{\ell+[\theta\ell]}^2 + \tilde{C}_{2,q'}^{\ell,\beta} \gamma_1^3) \exp(-2^{-(\beta+2)} \lambda_{h_{\ell+[\theta\ell]},2} 2\gamma_1 n^{1-\beta}) \Phi_{1-\beta}(n+1) \\ &+ \bar{C}_{2,q'}^{\ell,\beta} \frac{2^{1+2\delta'} \gamma_1^{2\frac{\delta'}{\beta}}}{\lambda_{h_{\ell+[\theta\ell]},2} 2n^{2\delta'}} h_{\ell+[\theta\ell]}^2 + \tilde{C}_{2,q'}^{\ell,\beta} \frac{2^{1+2\beta} \gamma_1^2}{\lambda_{h_{\ell+[\theta\ell]},2} 2n^{2\beta}}, \end{aligned} \quad (\text{B.45})$$

where

$$\begin{aligned} \bar{C}_{2,q'}^{\ell,\beta} &= (\bar{\zeta}_{h_{\ell+[\theta\ell]},2,q'}^{\beta,\delta'})^{-1} \left( \gamma_1^{\frac{2\delta'-\delta' \wedge \beta}{\beta}} \vee \left( \frac{\lambda_{h_{\ell+[\theta\ell]},2}}{2 \bar{\zeta}_{h_{\ell+[\theta\ell]},2,q'}^{\beta,\delta'}} \right)^{\frac{2\delta'-\delta' \wedge \beta}{\delta' \wedge \beta}} \right), \\ \tilde{C}_{2,q'}^{\ell,\beta} &= (\bar{\zeta}_{h_{\ell+[\theta\ell]},2,q'}^{\beta,\delta'})^{-1} \left( \gamma_1^{\frac{2\beta-\delta' \wedge \beta}{\beta}} \vee \left( \frac{\lambda_{h_{\ell+[\theta\ell]},2}}{2 \bar{\zeta}_{h_{\ell+[\theta\ell]},2,q'}^{\beta,\delta'}} \right)^{\frac{2\beta-\delta' \wedge \beta}{\delta' \wedge \beta}} \right). \end{aligned}$$

Thus

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})] \leq C_{2,q'}^{\ell,\beta}((h_\ell^{1+\theta}\tilde{u}_n) \vee \gamma_n), \quad (\text{B.46})$$

where

$$\begin{aligned} C_{2,q'}^{\ell,\beta} &= \gamma_1^{-1}(\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_0^{h_\ell})] + \bar{C}_{2,q'}^{\ell,\beta} + \tilde{C}_{2,q'}^{\ell,\beta}) \\ &\quad \times \sup_{m \geq 1} \left\{ m^\beta \exp\left(2^{1+\beta+\delta' \wedge \beta} \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \gamma_1^{1+\frac{\delta' \wedge \beta}{\beta}} \Phi_{1-\beta-\delta' \wedge \beta}(m+1)\right) \right. \\ &\quad \left. \times \exp\left(-\frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},2}\gamma_1}{2} \Phi_{1-\beta}(m+1)\right)\right\} \\ &\quad + 2^\beta (\bar{C}_{2,q'}^{\ell,\beta} \gamma_1^{2\frac{\delta'}{\beta}} h_{\ell+\lceil\theta\ell\rceil}^2 + \tilde{C}_{2,q'}^{\ell,\beta} \gamma_1^2) \\ &\quad \times \sup_{m \geq 1} \{m^\beta \exp(-2^{-(\beta+2)} \lambda_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_1 m^{1-\beta}) \Phi_{1-\beta}(m+1)\} \\ &\quad + \bar{C}_{2,q'}^{\ell,\beta} \frac{2^{1+2\delta'} \gamma_1^{-1+2\frac{\delta'}{\beta}}}{h_0 M^{\ell+\lceil\theta\ell\rceil} \lambda_{h_{\ell+\lceil\theta\ell\rceil},2}} + \tilde{C}_{2,q'}^{\ell,\beta} \frac{2^{1+2\beta} \gamma_1}{\lambda_{h_{\ell+\lceil\theta\ell\rceil},2}}. \end{aligned}$$

Note that  $\sup_{\ell \geq 1} C_{2,2}^{\ell,\beta} < \infty$ ,  $\beta \in (0, 1]$ .

*Step 3. Conclusion.*

Applying Lemma A.1(v) with  $h = h_{\ell+\lceil\theta\ell\rceil}$ ,  $q = 1$  yields

$$\mathbb{E}[(\tilde{\xi}_n^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})^2] \leq \kappa_{h_{\ell+\lceil\theta\ell\rceil},1}(\mathbb{E}[\bar{\mathcal{L}}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_n^{h_\ell})] + \mathbb{E}[\bar{\mathcal{L}}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})]).$$

Thus, via (B.35), (B.36), (B.44) and (B.46), if  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_1 > 1$  when  $\beta = 1$ , so that  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2}\gamma_1 \geq \bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_1 > 1$  when  $\beta = 1$ , then

$$\mathbb{E}[(\tilde{\xi}_n^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})^2] \leq C^{\ell,\beta}((h_\ell^{1+\theta}\tilde{u}_n) \vee \gamma_n),$$

where

$$C^{\ell,\beta} = \kappa_{h_{\ell+\lceil\theta\ell\rceil},1}(C_{1,1}^{\ell,\beta} + C_{2,2}^{\ell,\beta}). \quad (\text{B.47})$$

Observe that  $\sup_{\ell \geq 1} C^{\ell,\beta} < \infty$ ,  $\beta \in (0, 1]$ .

(ii) Our goal is to apply Lemma A.1(v) with  $q = 2$  and  $h = h_{\ell+\lceil\theta\ell\rceil}$ . Via Lemma A.1(ii), by assumption, for  $\mu \in \{\mu_{h_{\ell+\lceil\theta\ell\rceil},q'}, q' \in \{1, 2, 3, 4\}\}$ ,

$$\begin{aligned} &\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},1}(\tilde{\xi}_0^{h_\ell})] \vee \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_0^{h_\ell})] \vee \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_0^{h_\ell})] \vee \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_0^{h_\ell})] \\ &\leq \mathbb{E}[(1 + 8k_\alpha^4(|\tilde{\xi}_0^{h_\ell}|^4 + |\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}|^4)) \exp(k_\alpha(|\tilde{\xi}_0^{h_\ell}| + |\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}|))] < \infty. \end{aligned}$$

*Step 1. Alternative inequality on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})]$ .*

If  $\gamma_n = \gamma_1 n^{-1}$ ,  $n \in \mathbb{N}^*$ , and  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2}\gamma_1 > 2$ , which implies in particular that  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},1}\gamma_1 = \frac{\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2}\gamma_1}{2} > 1$  and that, for  $q' \geq 2$ ,  $\lambda_{h_{\ell+\lceil\theta\ell\rceil},2}^{\mu_{h_{\ell+\lceil\theta\ell\rceil},q'}} \gamma_1 \geq \bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2}\gamma_1 > 2 \geq 2\delta'$ , then, recalling the control on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})]$  derived in (B.43), for  $q' \geq 2$ ,

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})] \leq \check{C}_{2,q'}^{\ell,1}((h_\ell^{2(1+\theta)}\tilde{u}_n^2) \vee \gamma_n^2), \quad (\text{B.48})$$

where

$$\check{C}_{2,q'}^{\ell,1} = \gamma_1^{-2}(\check{C}_{2,q'}^{\ell,1} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_0^{h_\ell})] + h_0^{-2} \bar{C}_{2,q'}^{\ell,1} + \tilde{C}_{2,q'}^{\ell,1}).$$

Reconsidering the inequality on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})]$  obtained in (B.45), if  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\beta \in (0, 1)$ , then

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})] \leq \check{C}_{2,q'}^{\ell,\beta} ((h_\ell^{2(1+\theta)} \tilde{u}_n^2) \vee \gamma_n^2), \quad (\text{B.49})$$

where

$$\begin{aligned} \check{C}_{2,q'}^{\ell,\beta} &= \gamma_1^{-2} (\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_0^{h_\ell})]) + \bar{C}_{2,q'}^{\ell,\beta} + \tilde{C}_{2,q'}^{\ell,\beta} \\ &\quad \times \sup_{m \geq 1} \left\{ m^{2\beta} \exp \left( 2^{1+\beta+\delta' \wedge \beta} \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},2,q'}^{\beta,\delta'} \gamma_1^{1+\frac{\delta' \wedge \beta}{\beta}} \Phi_{1-\beta-\delta' \wedge \beta}(m+1) \right) \right. \\ &\quad \left. \times \exp \left( - \frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_1}{2} \Phi_{1-\beta}(m+1) \right) \right\} \\ &\quad + 2^\beta (\bar{C}_{2,q'}^{\ell,\beta} \gamma_1^{-1+2\frac{\delta'}{\beta}} h_{\ell+\lceil\theta\ell\rceil}^2 + \tilde{C}_{2,q'}^{\ell,\beta} \gamma_1) \\ &\quad \times \sup_{m \geq 1} \{ m^{2\beta} \exp(-2^{-(\beta+2)} \lambda_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_1 m^{1-\beta}) \Phi_{1-\beta}(m+1) \} \\ &\quad + \bar{C}_{2,q'}^{\ell,\beta} \frac{2^{1+2\delta'} \gamma_1^{-2+2\frac{\delta'}{\beta}}}{h_0^2 \lambda_{h_{\ell+\lceil\theta\ell\rceil},2}} + \tilde{C}_{2,q'}^{\ell,\beta} \frac{2^{1+2\beta}}{\lambda_{h_{\ell+\lceil\theta\ell\rceil},2}}. \end{aligned}$$

Note that  $\sup_{\ell \geq 1} \check{C}_{2,q'}^{\ell,\beta} < \infty$ ,  $q' \in \{2, 3, 4\}$ ,  $\beta \in (0, 1]$ .

*Step 2. Inequality on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})]$ .*

To be able to apply Lemma A.1(v) with  $q = 2$  and  $h = h_{\ell+\lceil\theta\ell\rceil}$ , one needs a control on  $\mathbb{E}[\bar{\mathcal{L}}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_n^{h_\ell})]$ , which, ostensibly from the recursive relation (B.26), requires in turn a control on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})]$ . For  $q' \geq 3$ , assuming that  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_1 > 2$  if  $\beta = 1$ , then, by (B.48) and (B.49),

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},2}(\tilde{\xi}_n^{h_\ell})] \leq \check{C}_{2,q'}^{\ell,\beta} (2^{2\delta'} M^2 h_{\ell+\lceil\theta\ell\rceil}^2 \tilde{u}_{n+1}^2 + 2^{2\beta} \gamma_{n+1}^2).$$

Thus, by taking  $q = 3$  and  $h = h_{\ell+\lceil\theta\ell\rceil}$  in (B.26), for  $q' \geq 3$ ,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_{n+1}^{h_\ell})] &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},3} \gamma_{n+1} + \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_{n+1}^{1+\frac{\delta' \wedge \beta}{\beta}}) \\ &\quad + 2^{2\delta'} 3C \check{C}_{2,q'}^{\ell,\beta} M^3 \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil}^3 \tilde{u}_{n+1}^3 \gamma_{n+1} \\ &\quad + 2^{2\beta} 3C \check{C}_{2,q'}^{\ell,\beta} M \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil} \tilde{u}_{n+1} \gamma_{n+1}^3 \\ &\quad + 2^{2\delta'} \check{C}_{2,q'}^{\ell,\beta} M^2 \nu_{h_{\ell+\lceil\theta\ell\rceil},3} h_{\ell+\lceil\theta\ell\rceil}^2 \tilde{u}_{n+1}^2 \gamma_{n+1}^2 \\ &\quad + (2^{2\beta} \check{C}_{2,q'}^{\ell,\beta} + 1) \nu_{h_{\ell+\lceil\theta\ell\rceil},3} \gamma_{n+1}^4 \\ &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},3} \gamma_{n+1} + \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_{n+1}^{1+\frac{\delta' \wedge \beta}{\beta}}) \\ &\quad + \bar{d}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_{n+1} (h_{\ell+\lceil\theta\ell\rceil} \tilde{u}_{n+1} + \gamma_{n+1})^3 \\ &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},3} \gamma_{n+1} + \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_{n+1}^{1+\frac{\delta' \wedge \beta}{\beta}}) \\ &\quad + 4\bar{d}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_1^{3-3\frac{\delta'}{\beta}} h_{\ell+\lceil\theta\ell\rceil}^3 \gamma_{n+1}^{1+3\frac{\delta'}{\beta}} + 4\bar{d}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_{n+1}^4, \end{aligned} \quad (\text{B.50})$$

where we introduced

$$\begin{aligned} \bar{d}_{h,q,q'}^{\beta,\delta} &= (2^{2\delta'} 3C \check{C}_{2,q'}^{\ell,\beta} M^3 \|V'_h\|_\infty) \vee (2^{2\beta} C \check{C}_{2,q'}^{\ell,\beta} M \|V'_h\|_\infty) \\ &\quad \vee \left( \frac{2^{2\delta'}}{3} \check{C}_{2,q'}^{\ell,\beta} M^2 \nu_{h,q}^{\mu_{h,q'}} \right) \vee ((2^{2\beta} \check{C}_{2,q'}^{\ell,\beta} + 1) \nu_{h,q}^{\mu_{h,q'}}), \\ &\quad h \in \bar{\mathcal{H}}, \quad q \in \mathbb{N}^*, \quad q' \geq q, \quad \beta, \delta \in (0, 1], \end{aligned}$$

and used that

$$\gamma_n \tilde{u}_n^3 = \frac{\gamma_1^4}{n^{\beta+3\delta'}} = \gamma_1^{3-3\frac{\delta'}{\beta}} \gamma_n^{1+3\frac{\delta'}{\beta}}, \quad n \in \mathbb{N}^*. \quad (\text{B.51})$$

Considering (B.40) and recalling (B.42), for  $q' \geq 3$ , by iterating  $n$  times the inequality (B.50),

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_0^{h_\ell})] \tilde{\Pi}_{1:n}^{h_{\ell+\lceil\theta\ell\rceil},3,q'} \\ &\quad + 4\bar{d}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_1^{3-3\frac{\delta'}{\beta}} \sum_{k=1}^n h_{\ell+\lceil\theta\ell\rceil}^3 \gamma_k^{1+3\frac{\delta'}{\beta}} \tilde{\Pi}_{k+1:n}^{h_{\ell+\lceil\theta\ell\rceil},3,q'} \\ &\quad + 4\bar{d}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \sum_{k=1}^n \gamma_k^4 \tilde{\Pi}_{k+1:n}^{h_{\ell+\lceil\theta\ell\rceil},3,q'}. \end{aligned}$$

If  $\gamma_n = \gamma_1 n^{-1}$ ,  $n \in \mathbb{N}^*$ , then, by Lemma A.2,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] &\leq \hat{C}_{3,q'}^{\ell,1} \frac{\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_0^{h_\ell})]}{(n+1)^{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}}} + \bar{C}_{3,q'}^{\ell,1} \frac{\Phi_{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1-3\delta'}(n+1)}}{(n+1)^{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}}} h_{\ell+\lceil\theta\ell\rceil}^3 \\ &\quad + \tilde{C}_{3,q'}^{\ell,1} \frac{\Phi_{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1-3}(n+1)}}{(n+1)^{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}}}, \end{aligned}$$

where

$$\begin{aligned} \hat{C}_{3,q'}^{\ell,1} &= \exp\left(\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{1,\delta'} \left(1 + \frac{1}{\delta'}\right) \gamma_1^{1+\delta'} + \frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}}{2}\right), \\ \bar{C}_{3,q'}^{\ell,1} &= 2^{2+(\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}) \vee (1+3\delta')} \hat{C}_{3,q'}^{\ell,1} \bar{d}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{1,\delta'} \gamma_1^4, \\ \tilde{C}_{3,q'}^{\ell,1} &= 2^{2+(\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}) \vee 4} \hat{C}_{3,q'}^{\ell,1} \bar{d}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{1,\delta'} \gamma_1^4. \end{aligned}$$

In particular, if  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2} > 2$ , so that  $\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1} \geq \lambda_{h_{\ell+\lceil\theta\ell\rceil},2\gamma_1} \geq \bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2\gamma_1} > 2 \geq 2\delta'$ , then

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] \leq C_{3,q'}^{\ell,1} ((h_\ell^{2(1+\theta)} \tilde{u}_n^2) \vee \gamma_n^2), \quad (\text{B.52})$$

where

$$\begin{aligned} C_{3,q'}^{\ell,1} &= \gamma_1^{-2} \left( \hat{C}_{3,q'}^{\ell,1} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_0^{h_\ell})] \right. \\ &\quad \left. + \frac{\bar{C}_{3,q'}^{\ell,1}}{h_0 M^{\ell+\lceil\theta\ell\rceil}} \sup_{m \geq 1} \left\{ m^{2\delta'} \frac{\Phi_{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1-3\delta'}(m+1)}}{(m+1)^{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}}} \right\} \right. \\ &\quad \left. + \tilde{C}_{3,q'}^{\ell,1} \sup_{m \geq 1} \left\{ m^2 \frac{\Phi_{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1-3}(m+1)}}{(m+1)^{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}}} \right\} \right). \end{aligned}$$

If  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\beta \in (0, 1)$ , then, by Lemma A.2,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] &\leq (\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_0^{h_\ell})] + \bar{C}_{3,q'}^{\ell,\beta} + \tilde{C}_{3,q'}^{\ell,\beta}) \\ &\quad \times \exp\left(2^{1+\beta+\delta' \wedge \beta} \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_1^{1+\frac{\delta' \wedge \beta}{\beta}} \Phi_{1-\beta-\delta' \wedge \beta}(n+1) - \frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1}}{2} \Phi_{1-\beta}(n+1)\right) \\ &\quad + 2^\beta (\bar{C}_{3,q'}^{\ell,\beta} \gamma_1^{1+3\frac{\delta'}{\beta}} h_{\ell+\lceil\theta\ell\rceil}^3 + \tilde{C}_{3,q'}^{\ell,\beta} \gamma_1^4) \exp(-2^{-(\beta+2)} \lambda_{h_{\ell+\lceil\theta\ell\rceil},3\gamma_1} n^{1-\beta}) \Phi_{1-\beta}(n+1) \\ &\quad + \bar{C}_{3,q'}^{\ell,\beta} \frac{2^{1+3\delta'} \gamma_1^{3\frac{\delta'}{\beta}}}{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3n^{3\delta'}} h_{\ell+\lceil\theta\ell\rceil}^3 + \tilde{C}_{3,q'}^{\ell,\beta} \frac{2^{1+3\beta} \gamma_1^3}{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3n^{3\beta}}}, \end{aligned}$$

where

$$\begin{aligned}\bar{C}_{3,q'}^{\ell,\beta} &= (\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'})^{-1} \left( \gamma_1^{\frac{3\delta'-\delta'\wedge\beta}{\beta}} \vee \left( \frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3}}{2\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'}} \right)^{\frac{3\delta'-\delta'\wedge\beta}{\delta'\wedge\beta}} \right), \\ \tilde{C}_{3,q'}^{\ell,\beta} &= (\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'})^{-1} \left( \gamma_1^{\frac{3\beta-\delta'\wedge\beta}{\beta}} \vee \left( \frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3}}{2\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'}} \right)^{\frac{3\beta-\delta'\wedge\beta}{\delta'\wedge\beta}} \right).\end{aligned}$$

Thus

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] \leq C_{3,q'}^{\ell,\beta} ((h_\ell^{2(1+\theta)} \tilde{u}_n^2) \vee \gamma_n^2), \quad (\text{B.53})$$

where

$$\begin{aligned}C_{3,q'}^{\ell,\beta} &= \gamma_1^{-2} (\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_0^{h_\ell})] + \bar{C}_{3,q'}^{\ell,\beta} + \tilde{C}_{3,q'}^{\ell,\beta}) \\ &\quad \times \sup_{m \geq 1} \left\{ m^{2\beta} \exp \left( 2^{1+\beta+\delta'\wedge\beta} \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},3,q'}^{\beta,\delta'} \gamma_1^{1+\frac{\delta'\wedge\beta}{\beta}} \Phi_{1-\beta-\delta'\wedge\beta}(m+1) \right) \right. \\ &\quad \left. \times \exp \left( - \frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3} \gamma_1}{2} \Phi_{1-\beta}(m+1) \right) \right\} \\ &\quad + 2^\beta (\bar{C}_{3,q'}^{\ell,\beta} \gamma_1^{-1+3\frac{\delta'}{\beta}} h_{\ell+\lceil\theta\ell\rceil}^3 + \tilde{C}_{3,q'}^{\ell,\beta} \gamma_1^2) \\ &\quad \times \sup_{m \geq 1} \{ m^{2\beta} \exp(-2^{-(\beta+2)} \lambda_{h_{\ell+\lceil\theta\ell\rceil},3} \gamma_1 m^{1-\beta}) \Phi_{1-\beta}(m+1) \} \\ &\quad + \bar{C}_{3,q'}^{\ell,\beta} \frac{2^{1+3\delta'} \gamma_1^{-2+3\frac{\delta'}{\beta}}}{h_0 M^{\ell+\lceil\theta\ell\rceil} \lambda_{h_{\ell+\lceil\theta\ell\rceil},3}} + \tilde{C}_{3,q'}^{\ell,\beta} \frac{2^{1+3\beta} \gamma_1}{\lambda_{h_{\ell+\lceil\theta\ell\rceil},3}}.\end{aligned}$$

Note that  $\sup_{\ell \geq 1} C_{3,q'}^{\ell,\beta} < \infty$ ,  $q' \in \{3, 4\}$ ,  $\beta \in (0, 1]$ .

*Step 3. Inequality on  $\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_n^{h_\ell})]$ .*

For  $q' \geq 4$ , assuming that  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_1 > 2$  if  $\beta = 1$ , following from (B.52) and (B.53),

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},3}(\tilde{\xi}_n^{h_\ell})] \leq C_{3,q'}^{\ell,\beta} (2^{2\delta'} M^2 h_{\ell+\lceil\theta\ell\rceil}^2 \tilde{u}_{n+1}^2 + 2^{2\beta} \gamma_{n+1}^2).$$

Thus, the relation (B.26), with  $q = 4$  and  $h = h_{\ell+\lceil\theta\ell\rceil}$ , together with the decreasing monotonicity of  $(\gamma_n)_{n \geq 1}$  and the relation (B.51), give

$$\begin{aligned}\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_{n+1}^{h_\ell})] &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_n^{h_\ell})] (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},4} \gamma_{n+1} + \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'} \gamma_{n+1}^{1+\frac{\delta'\wedge\beta}{\beta}}) \\ &\quad + 2^{2+2\delta'} C C_{3,q'}^{\ell,\beta} M^3 \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil}^3 \tilde{u}_{n+1}^3 \gamma_{n+1} \\ &\quad + 2^{2+2\beta} C C_{3,q'}^{\ell,\beta} M \|V'_{h_{\ell+\lceil\theta\ell\rceil}}\|_\infty h_{\ell+\lceil\theta\ell\rceil} \tilde{u}_{n+1}^3 \gamma_{n+1}^3 \\ &\quad + 2^{2\delta'} C_{3,q'}^{\ell,\beta} M^2 \nu_{h_{\ell+\lceil\theta\ell\rceil},4} h_{\ell+\lceil\theta\ell\rceil}^2 \tilde{u}_{n+1}^2 \gamma_{n+1}^2 \\ &\quad + (2^{2\beta} C_{3,q'}^{\ell,\beta} + \gamma_1) \nu_{h_{\ell+\lceil\theta\ell\rceil},4} \gamma_{n+1}^4 \\ &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_n^{h_\ell})] (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},4} \gamma_{n+1} + \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'} \gamma_{n+1}^{1+\frac{\delta'\wedge\beta}{\beta}}) \\ &\quad + \hat{d}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'} \gamma_{n+1} (h_{\ell+\lceil\theta\ell\rceil} \tilde{u}_{n+1} + \gamma_{n+1})^3 \\ &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_n^{h_\ell})] (1 - \lambda_{h_{\ell+\lceil\theta\ell\rceil},4} \gamma_{n+1} + \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'} \gamma_{n+1}^{1+\frac{\delta'\wedge\beta}{\beta}}) \\ &\quad + 4 \hat{d}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'} \gamma_1^{3-3\frac{\delta'}{\beta}} h_{\ell+\lceil\theta\ell\rceil}^3 \gamma_{n+1}^{1+3\frac{\delta'}{\beta}} + 4 \hat{d}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'} \gamma_{n+1}^4,\end{aligned} \quad (\text{B.54})$$

where we introduced

$$\begin{aligned}\hat{d}_{h,q,q'}^{\beta,\delta} &= (2^{2+2\delta'} C C_{3,q'}^{\ell,\beta} M^3 \|V_h'\|_\infty) \vee \left( \frac{2^{2+2\beta}}{3} C C_{3,q'}^{\ell,\beta} M \|V_h'\|_\infty \right) \\ &\quad \vee \left( \frac{2^{2\delta'}}{3} C_{3,q'}^{\ell,\beta} M^2 \nu_{h,q}^{\mu_{h,q'}} \right) \vee ((2^{2\beta} C_{3,q'}^{\ell,\beta} + \gamma_1) \nu_{h,q}^{\mu_{h,q'}}), \\ &\quad h \in \bar{\mathcal{H}}, \quad q \in \mathbb{N}^*, \quad q' \geq q, \quad \beta, \delta \in (0, 1].\end{aligned}$$

Hence, iterating  $n$  times the inequality (B.54),

$$\begin{aligned}\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},4}(\tilde{\xi}_n^{h_\ell})] &\leq \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},4}(\tilde{\xi}_0^{h_\ell})] \tilde{\Pi}_{1:n}^{h_{\ell+[\theta\ell]},4,q'} \\ &\quad + 4\hat{d}_{h_{\ell+[\theta\ell]},4,q'}^{\beta,\delta'} \gamma_1^{3-3\frac{\delta'}{\beta}} \sum_{k=1}^n h_{\ell+[\theta\ell]}^3 \gamma_k^{1+3\frac{\delta'}{\beta}} \tilde{\Pi}_{k+1:n}^{h_{\ell+[\theta\ell]},4,q'} \\ &\quad + 4\hat{d}_{h_{\ell+[\theta\ell]},4,q'}^{\beta,\delta'} \sum_{k=1}^n \gamma_k^4 \tilde{\Pi}_{k+1:n}^{h_{\ell+[\theta\ell]},4,q'}.\end{aligned}$$

If  $\gamma_n = \gamma_1 n^{-1}$ ,  $n \in \mathbb{N}^*$ , then, by Lemma A.2,

$$\begin{aligned}\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},4}(\tilde{\xi}_n^{h_\ell})] &\leq \bar{C}_{4,q'}^{\ell,1} \frac{\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},4}(\tilde{\xi}_0^{h_\ell})]}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1}} + \bar{C}_{4,q'}^{\ell,1} \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1-3\delta'}(n+1)}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1}} h_{\ell+[\theta\ell]}^3 \\ &\quad + \tilde{C}_{4,q'}^{\ell,1} \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1-3}(n+1)}{(n+1)^{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1}},\end{aligned}$$

where

$$\begin{aligned}\bar{C}_{4,q'}^{\ell,1} &= \exp\left(\bar{\zeta}_{h_{\ell+[\theta\ell]},4,q'}^{1,\delta'} \left(1 + \frac{1}{\delta'}\right) \gamma_1^{1+\delta'} + \frac{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1}{2}\right), \\ \bar{C}_{4,q'}^{\ell,1} &= 2^{2+(\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1)\vee(1+3\delta')} \bar{C}_{4,q'}^{\ell,1} \hat{d}_{h_{\ell+[\theta\ell]},4,q'}^{1,\delta'} \gamma_1^4, \\ \tilde{C}_{4,q'}^{\ell,1} &= 2^{2+(\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1)\vee 4} \bar{C}_{4,q'}^{\ell,1} \hat{d}_{h_{\ell+[\theta\ell]},4,q'}^{1,\delta'} \gamma_1^4.\end{aligned}$$

In particular, if  $\bar{\lambda}_{h_{\ell+[\theta\ell]},2} > 2$ , so that  $\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1 \geq \lambda_{h_{\ell+[\theta\ell]},2}\gamma_1 \geq \bar{\lambda}_{h_{\ell+[\theta\ell]},2}\gamma_1 > 2 \geq 2\delta'$ , then

$$\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},4}(\tilde{\xi}_n^{h_\ell})] \leq C_{4,q'}^{\ell,1} ((h_\ell^{2(1+\theta)} \tilde{u}_n^2) \vee \gamma_n^2), \quad (\text{B.55})$$

where

$$\begin{aligned}C_{4,q'}^{\ell,1} &= \gamma_1^{-2} \left( \bar{C}_{4,q'}^{\ell,1} \mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},4}(\tilde{\xi}_0^{h_\ell})] \right. \\ &\quad \left. + \frac{\bar{C}_{4,q'}^{\ell,1}}{h_0 M^{\ell+[\theta\ell]}} \sup_{m \geq 1} \left\{ m^{2\delta'} \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1-3\delta'}(m+1)}{(m+1)^{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1}} \right\} \right. \\ &\quad \left. + \tilde{C}_{4,q'}^{\ell,1} \sup_{m \geq 1} \left\{ m^2 \frac{\Phi_{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1-3}(m+1)}{(m+1)^{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1}} \right\} \right).\end{aligned}$$

If  $\gamma_n = \gamma_1 n^{-\beta}$ ,  $n \in \mathbb{N}^*$ , with  $\beta \in (0, 1)$ , then, by Lemma A.2,

$$\begin{aligned}\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},4}(\tilde{\xi}_n^{h_\ell})] &\leq (\mathbb{E}[\mathcal{L}_{h_{\ell+[\theta\ell]},4}(\tilde{\xi}_0^{h_\ell})]) + \bar{C}_{4,q'}^{\ell,\beta} + \tilde{C}_{4,q'}^{\ell,\beta} \\ &\quad \times \exp\left(2^{1+\beta+\delta'\wedge\beta} \bar{\zeta}_{h_{\ell+[\theta\ell]},4,q'}^{\beta,\delta'} \gamma_1^{1+\frac{\delta'\wedge\beta}{\beta}} \Phi_{1-\beta-\delta'\wedge\beta}(n+1) - \frac{\lambda_{h_{\ell+[\theta\ell]},4}\gamma_1}{2} \Phi_{1-\beta}(n+1)\right) \\ &\quad + 2^\beta (\bar{C}_{4,q'}^{\ell,\beta} \gamma_1^{1+3\frac{\delta'}{\beta}} h_{\ell+[\theta\ell]}^3 + \tilde{C}_{4,q'}^{\ell,\beta} \gamma_1^4) \exp(-2^{-(\beta+2)} \lambda_{h_{\ell+[\theta\ell]},4}\gamma_1 n^{1-\beta}) \Phi_{1-\beta}(n+1) \\ &\quad + \bar{C}_{4,q'}^{\ell,\beta} \frac{2^{1+3\delta'} \gamma_1^{3\frac{\delta'}{\beta}}}{\lambda_{h_{\ell+[\theta\ell]},4} n^{3\delta'}} h_{\ell+[\theta\ell]}^3 + \tilde{C}_{4,q'}^{\ell,\beta} \frac{2^{1+3\beta} \gamma_1^3}{\lambda_{h_{\ell+[\theta\ell]},4} n^{3\beta}},\end{aligned}$$

where

$$\begin{aligned}\bar{C}_{4,q'}^{\ell,\beta} &= (\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'})^{-1} \left( \gamma_1^{\frac{3\delta'-\delta'\wedge\beta}{\beta}} \vee \left( \frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},4}}{2\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'}} \right)^{\frac{3\delta'-\delta'\wedge\beta}{\delta'\wedge\beta}} \right), \\ \tilde{C}_{4,q'}^{\ell,\beta} &= (\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'})^{-1} \left( \gamma_1^{\frac{3\beta-\delta'\wedge\beta}{\beta}} \vee \left( \frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},4}}{2\bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'}} \right)^{\frac{3\beta-\delta'\wedge\beta}{\delta'\wedge\beta}} \right).\end{aligned}$$

Thus

$$\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_n^{h_\ell})] \leq C_{4,q'}^{\ell,\beta} ((h_\ell^{2(1+\theta)} \tilde{u}_n^2) \vee \gamma_n^2), \quad (\text{B.56})$$

where

$$\begin{aligned}C_{4,q'}^{\ell,\beta} &= \gamma_1^{-2} (\mathbb{E}[\mathcal{L}_{h_{\ell+\lceil\theta\ell\rceil},4}(\tilde{\xi}_0^{h_\ell})]) + \bar{C}_{4,q'}^{\ell,\beta} + \tilde{C}_{4,q'}^{\ell,\beta} \\ &\quad \times \sup_{m \geq 1} \left\{ m^{2\beta} \exp \left( 2^{1+\beta+\delta'\wedge\beta} \bar{\zeta}_{h_{\ell+\lceil\theta\ell\rceil},4,q'}^{\beta,\delta'} \gamma_1^{1+\frac{\delta'\wedge\beta}{\beta}} \Phi_{1-\beta-\delta'\wedge\beta}(m+1) \right) \right. \\ &\quad \left. \times \exp \left( -\frac{\lambda_{h_{\ell+\lceil\theta\ell\rceil},4} \gamma_1}{2} \Phi_{1-\beta}(m+1) \right) \right\} \\ &\quad + 2^\beta (\bar{C}_{4,q'}^{\ell,\beta} \gamma_1^{-1+3\frac{\delta'}{\beta}} h_{\ell+\lceil\theta\ell\rceil}^3 + \tilde{C}_{4,q'}^{\ell,\beta} \gamma_1^2) \\ &\quad \times \sup_{m \geq 1} \{ m^{2\beta} \exp(-2^{-(\beta+2)} \lambda_{h_{\ell+\lceil\theta\ell\rceil},4} \gamma_1 m^{1-\beta}) \Phi_{1-\beta}(m+1) \} \\ &\quad + \bar{C}_{4,q'}^{\ell,\beta} \frac{2^{1+3\delta'} \gamma_1^{-2+3\frac{\delta'}{\beta}}}{h_0 M^{\ell+\lceil\theta\ell\rceil} \lambda_{h_{\ell+\lceil\theta\ell\rceil},4}} + \tilde{C}_{4,q'}^{\ell,\beta} \frac{2^{1+3\beta} \gamma_1}{\lambda_{h_{\ell+\lceil\theta\ell\rceil},4}}.\end{aligned}$$

Note that  $\sup_{\ell \geq 1} C_{4,4}^{\ell,\beta} < \infty$ ,  $\beta \in (0, 1]$ .

*Step 4. Conclusion.*

All in all, if  $\bar{\lambda}_{h_{\ell+\lceil\theta\ell\rceil},2} \gamma_1 > 2$  when  $\beta = 1$ , then, considering Lemma A.1(v) with  $q = 2$ ,  $h = h_{\ell+\lceil\theta\ell\rceil}$ ,  $\mu = \mu_{h_{\ell+\lceil\theta\ell\rceil},2}$  and  $\mu' = \mu_{h_{\ell+\lceil\theta\ell\rceil},4}$ , the inequalities (B.48), (B.49), (B.55) and (B.56) yield

$$\mathbb{E}[|\tilde{\xi}_n^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}|^4] \leq \bar{C}^{\ell,\beta} ((h_\ell^{2(1+\theta)} \tilde{u}_n^2) \vee \gamma_n^2),$$

where the

$$\bar{C}^{\ell,\beta} = \kappa_{h_{\ell+\lceil\theta\ell\rceil},2} (\check{C}_{2,2}^{\ell,\beta} + C_{4,4}^{\ell,\beta}) \quad (\text{B.57})$$

satisfy  $\sup_{\ell \geq 1} \bar{C}^{\ell,\beta} < \infty$ ,  $\beta \in (0, 1]$ .  $\square$

*Proof of Lemma 5.2.* Let  $\ell, n \in \mathbb{N}^*$  and  $\xi \in \mathbb{R}$ ,  $C_{\text{ad}} > 0$ ,  $r > 1$  and  $\theta \in (0, 1]$ .

Via (3.10) and the law of total probability,

$$\mathbb{E}[|\mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} > \xi\}} - \mathbb{1}_{\{X^0 > \xi\}}|] = \sum_{k=0}^{\lceil\theta\ell\rceil} \mathbb{E}[|\mathbb{1}_{\{X^{h_{\ell+k}} > \xi\}} - \mathbb{1}_{\{X^0 > \xi\}}| \mathbb{1}_{\{\eta_n^\ell(\xi) = k\}}] \leq A^\ell + B^\ell,$$

where

$$\begin{aligned}A^\ell &:= \mathbb{E}[|\mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} > \xi\}} - \mathbb{1}_{\{X^0 > \xi\}}|], \\ B^\ell &:= \sum_{k=0}^{\lceil\theta\ell\rceil-1} \mathbb{E}[|\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^0 \leq \xi\}}| \mathbb{1}_{\{\eta_n^\ell(\xi) = k\}}].\end{aligned} \quad (\text{B.58})$$

Let  $k \in \llbracket 0, \lceil\theta\ell\rceil - 1 \rrbracket$ . Observe that

$$|\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi\}}| = \mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi < X^{h_{\ell+\lceil\theta\ell\rceil}}\}} + \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi < X^{h_{\ell+k}}\}},$$

that

$$\{X^{h_{\ell+k}} \leq \xi < X^{h_{\ell+\lceil\theta\ell\rceil}}\} \cup \{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi < X^{h_{\ell+k}}\} \subset \{|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}| \geq |X^{h_{\ell+k}} - \xi|\},$$

and that, by (3.10),

$$\{\eta_n^\ell(\xi) = k\} \subset \{|X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}}\psi_n^{\ell,k}\}.$$

Hence

$$\begin{aligned} & \mathbb{E}[\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi\}} | \mathbb{1}_{\{\eta_n^\ell(\xi)=k\}}] \\ & \leq \mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi < X^{h_{\ell+\lceil\theta\ell\rceil}}\}} + \mathbb{1}_{\{X^{h_{\ell+\lceil\theta\ell\rceil}} \leq \xi < X^{h_{\ell+k}}\}}) \mathbb{1}_{\{|X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}}\psi_n^{\ell,k}\}}] \\ & \leq \mathbb{P}(|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}| \geq |X^{h_{\ell+k}} - \xi|, |X^{h_{\ell+k}} - \xi| \geq C_{\text{ad}}\psi_n^{\ell,k}) \\ & \leq \mathbb{P}(|X^{h_{\ell+k}} - X^{h_{\ell+\lceil\theta\ell\rceil}}| \geq C_{\text{ad}}\psi_n^{\ell,k}). \end{aligned} \tag{B.59}$$

In the following, we separately study  $A^\ell$  and  $B^\ell$  within each framework.

(i)a. *Step 1. Study of  $A^\ell$ .*

Using Lemma 2.8(i)a,

$$A^\ell \leq C_{1A} h_\ell^{(1+\theta)\frac{p_\star}{2(p_\star+1)}},$$

where

$$C_{1A} = h_0^{-\frac{p_\star}{2(p_\star+1)}} B_{p_\star} \mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{p_\star}]^{\frac{1}{p_\star+1}} \left( \sup_{h \in \overline{\mathcal{H}}} \|f_{X^h}\|_\infty \right)^{\frac{p_\star}{p_\star+1}},$$

with  $B_{p_\star}$  being a positive constant that depends only on  $p_\star$ .

*Step 2. Study of  $B^\ell$ .*

Via (B.59), the definition (3.10), Markov's inequality and Lemma A.3(i),

$$\begin{aligned} & \mathbb{E}[\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^0 \leq \xi\}} | \mathbb{1}_{\{\eta_n^\ell(\xi)=k\}}] \\ & \leq \mathbb{P}(|X^{h_{\ell+k}} - X^0| \geq C_{\text{ad}} u_n^{-\frac{1}{p_\star}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}}) \\ & \leq C_{\text{ad}}^{-p_\star} u_n h_{\theta\ell(r-1)+k}^{-\frac{p_\star}{r}} \mathbb{E}[|X^{h_{\ell+k}} - X^0|^{p_\star}] \\ & \leq C_{1b} h_{\theta\ell(r-1)+k}^{-\frac{p_\star}{r}} h_{\ell+k}^{\frac{p_\star}{2}}, \end{aligned}$$

where

$$C_{1b} = \gamma_1 \sqrt{2} C_{\text{ad}}^{-p_\star} B_{p_\star} \mathbb{E}[|\varphi(Y, Z) - \mathbb{E}[\varphi(Y, Z)|Y]|^{p_\star}]^{\frac{1}{p_\star}}.$$

Thus, by (B.58) and using the condition (5.2) on  $r$  and  $\theta$ ,

$$\begin{aligned} B^\ell & \leq C_{1b} \sum_{k=0}^{\lceil\theta\ell\rceil-1} h_{\theta\ell(r-1)+k}^{-\frac{p_\star}{r}} h_{\ell+k}^{\frac{p_\star}{2}} \\ & \leq \frac{C_{1b} h_1^{-p_\star(\frac{1}{r}-\frac{1}{2})}}{M^{p_\star(\frac{1}{r}-\frac{1}{2})} - 1} M^{-\frac{p_\star\ell}{2}(1-\theta)} \\ & \leq C_{1B} h_\ell^{(1+\theta)\frac{p_\star}{2(p_\star+1)}}, \end{aligned}$$

where

$$C_{1B} = \frac{C_{1b} h_0^{-\frac{2(p_\star+1)}{p_\star+2}} h_1^{-p_\star(\frac{1}{r}-\frac{1}{2})}}{M^{p_\star(\frac{1}{r}-\frac{1}{2})} - 1}.$$

(i)b. *Step 1. Study of  $A^\ell$ .*

Lemma 2.8(i)b entails

$$A^\ell \leq 2\sqrt{\mathfrak{g}h_{\ell+\lceil\theta\ell\rceil}} \left( 1 + \left( \sup_{h \in \overline{\mathcal{H}}} \|f_{X^h}\|_\infty \right) \sqrt{2|\ln(\mathfrak{g}h_{\ell+\lceil\theta\ell\rceil})|} \right).$$

On the one hand,

$$2\sqrt{\mathfrak{g}h_{\ell+\lceil\theta\ell\rceil}} \leq 2\sqrt{\mathfrak{g}h_0^{-1}h_\ell^{\frac{1+\theta}{2}}}.$$

On the other hand, using that  $\theta \leq 1$  under the condition (5.3) and the inequality  $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$ ,  $x, y \geq 0$ ,

$$\sqrt{2|\ln(\mathfrak{g}h_{\ell+\lceil\theta\ell\rceil})|} \leq \sqrt{2|\ln(\mathfrak{g}h_0^{-1}) + (1+\theta)\ln h_\ell|} \leq \sqrt{2|\ln(\mathfrak{g}h_0^{-1})|} + 2\sqrt{|\ln h_\ell|}.$$

Thus

$$A^\ell \leq C_{2A} h_\ell^{\frac{1+\theta}{2}} |\ln h_\ell|^{\frac{1}{2}},$$

where

$$C_{2A} = 4\mathfrak{g}^{\frac{1}{2}} h_0^{-\frac{1}{2}} \left( 1 + \left( \sup_{h \in \overline{\mathcal{H}}} \|f_{X^h}\|_\infty \right) \left( 2(2 \vee |\ln(\mathfrak{g}h_0^{-1})|) \right)^{\frac{1}{2}} \right).$$

*Step 2. Study of  $B^\ell$ .*

Let  $\lambda > 0$ . By the inequality (B.59), the definition (3.10), Markov's exponential inequality and Lemma A.3(ii),

$$\begin{aligned} & \mathbb{E}[\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^0 \leq \xi\}} | \mathbb{1}_{\{\eta_n^\ell(\xi) = k\}}] \\ & \leq \mathbb{P}\left(|X^{h_{\ell+k}} - X^0| \geq C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1}\gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \\ & \leq \exp\left(-\lambda C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1}\gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \mathbb{E}[\exp(\lambda|X^{h_{\ell+k}} - X^0|)] \\ & \leq 2 \exp\left(-\lambda C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1}\gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right) + \mathfrak{g}\lambda^2 h_{\ell+k}\right), \end{aligned}$$

where we used the inequality  $e^{|x|} \leq e^x + e^{-x}$ ,  $x \in \mathbb{R}$ . Minimizing the above upper bound with respect to  $\lambda$  yields

$$\begin{aligned} & \mathbb{E}[\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^0 \leq \xi\}} | \mathbb{1}_{\{\eta_n^\ell(\xi) = k\}}] \\ & \leq 2 \exp\left(-\frac{C_{\text{ad}}^2 h_{\theta\ell(r-1)+k}^{\frac{2}{r}} \ln\left((\gamma_1^{-1}\gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)}{4\mathfrak{g}h_{\ell+k}}\right) \\ & = 2 \left(\frac{\gamma_n}{\gamma_1} h_{\ell+k}^{1+\theta}\right)^{\frac{C_{\text{ad}}^2 h_{\theta\ell(r-1)+k}^{\frac{2}{r}}}{8\mathfrak{g}h_{\ell+k}}}. \end{aligned}$$

Note that  $\gamma_n \leq \gamma_1$  and  $h_{\ell+k}^{1+\theta} \leq 1$ . The condition (5.3) on  $h_0$ ,  $r$  and  $\theta$  implies in particular that  $\theta \leq 1 \leq \frac{r}{2(r-1)}$ , so that  $-2\theta(1 - \frac{1}{r}) + 1 \geq 0$ , hence  $M^{-2\theta\ell(1-\frac{1}{r})+\ell} \geq 1$ . It also entails that  $\frac{C_{\text{ad}}^2 h_0^{\frac{2}{r}-1}}{8\mathfrak{g}} \geq \frac{1}{2}$ . Hence

$$\begin{aligned} & \mathbb{E}[\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^0 \leq \xi\}} | \mathbb{1}_{\{\eta_n^\ell(\xi) = k\}}] \\ & \leq 2h_{\ell+k}^{(1+\theta)\frac{C_{\text{ad}}^2 h_0^{\frac{2}{r}-1}}{8\mathfrak{g}} M^{-2\theta\ell(1-\frac{1}{r})+\ell}} \\ & \leq 2h_{\ell+k}^{\frac{1+\theta}{2}}. \end{aligned}$$

Eventually, (B.58) yields

$$B^\ell \leq C_{2B} h_\ell^{\frac{1+\theta}{2}} \leq C_{2B} h_\ell^{\frac{1+\theta}{2}} |\ln h_\ell|^{\frac{1}{2}},$$

where

$$C_{2B} = \frac{2}{1 - M^{-\frac{1}{2}}}.$$

(ii) *Step 1. Study of  $A^\ell$ .*

Consequently from Lemma 3.1,

$$A^\ell \leq C_{3A} h_\ell^{\frac{1+\theta}{2}},$$

where

$$C_{3A} = h_0^{-\frac{1}{2}} \sup_{0 \leq h_1 < h_2 \in \bar{\mathcal{H}}} \mathbb{E}[K_{h_1}^{h_2} | G_{h_1}^{h_2}|].$$

*Step 2. Study of  $B^\ell$ .*

Via (B.59), the definition (3.10) and Markov's exponential inequality, and recalling the condition (5.4) on  $h_0$ ,  $r$  and  $\theta$ ,

$$\begin{aligned} & |\mathbb{E}[(\mathbb{1}_{\{X^{h_{\ell+k}} \leq \xi\}} - \mathbb{1}_{\{X^0 \leq \xi\}}) \mathbb{1}_{\{\eta_n^\ell(\xi) = k\}}]| \\ & \leq \mathbb{P}\left(|X^{h_{\ell+k}} - X^0| \geq C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1} \gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \\ & = \mathbb{P}\left(|G_0^{h_{\ell+k}}| \geq C_{\text{ad}} h_{\theta\ell(r-1)+k}^{\frac{1}{r}} h_{\ell+k}^{-\frac{1}{2}} \ln^{\frac{1}{2}}\left((\gamma_1^{-1} \gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \\ & \leq \exp\left(-v_0 C_{\text{ad}}^2 h_{\theta\ell(r-1)+k}^{\frac{2}{r}} h_{\ell+k}^{-1} \ln\left((\gamma_1^{-1} \gamma_n)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}}\right)\right) \mathbb{E}[\exp(v_0 |G_0^{h_{\ell+k}}|^2)] \\ & \leq \left(\sup_{h \in \mathcal{H}} \mathbb{E}[\exp(v_0 |G_0^h|^2)]\right) \left(\frac{\gamma_n}{\gamma_1} h_{\ell+k}^{1+\theta}\right)^{\frac{1}{2} v_0 C_{\text{ad}}^2 h_0^{\frac{2}{r}-1} M^{-2\theta\ell(1-\frac{1}{r})+\ell}} \\ & \leq C_{3b} h_{\ell+k}^{\frac{1+\theta}{2}}, \end{aligned}$$

with

$$C_{3b} = \sup_{h \in \mathcal{H}} \mathbb{E}[\exp(v_0 |G_0^h|^2)],$$

where we used that  $\gamma_n \leq \gamma_1$ , that  $\theta \leq 1 \leq \frac{r}{2(r-1)}$ , so that  $M^{-2\theta\ell(1-\frac{1}{r})+\ell} \geq 1$ , and that  $v_0 C_{\text{ad}}^2 h_0^{\frac{2}{r}-1} \geq 1$ . All in all, via (B.58),

$$B^\ell \leq C_{3B} h_\ell^{\frac{1+\theta}{2}},$$

where

$$C_{3B} = \frac{C_{3b}}{1 - M^{-\frac{1}{2}}}.$$

□

## C Proof of Theorem 5.4

For all  $\ell \in \mathbb{N}$ , let  $(\tilde{\mathcal{F}}_n^{h_\ell})_{n \geq 0}$  be the filtration defined by  $\tilde{\mathcal{F}}_0^{h_\ell} = \sigma(\tilde{\xi}_0^{h_\ell})$  and  $\tilde{\mathcal{F}}_n^{h_\ell} = \sigma(\tilde{\xi}_0^{h_\ell}, X_1^{h_{\ell+\lceil\theta\ell\rceil}}, \dots, X_n^{h_{\ell+\lceil\theta\ell\rceil}})$ ,  $n \geq 1$ .

Following (B.10), the dynamics (4.1) can be decomposed into

$$\tilde{\xi}_n^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}} = (1 - \gamma_n V_0''(\xi_\star^0)) (\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) - \gamma_n g_n^{h_\ell} - \gamma_n \rho_n^{h_\ell} - \gamma_n r_n^{h_\ell} - \gamma_n e_n^{h_\ell}, \quad (\text{C.1})$$

where  $(e_n^{h_\ell})_{n \geq 1}$  and  $(r_n^{h_\ell})_{n \geq 1}$  are defined in (B.11) and (B.12), and

$$g_n^{h_\ell} := (V_{h_{\ell+[\theta\ell]}}''(\xi_\star^{h_{\ell+[\theta\ell]}}) - V_0''(\xi_\star^0))(\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+[\theta\ell]}}), \quad (\text{C.2})$$

$$\rho_n^{h_\ell} := V_{h_{\ell+[\theta\ell]}}'(\tilde{\xi}_{n-1}^{h_\ell}) - V_{h_{\ell+[\theta\ell]}}''(\xi_\star^{h_{\ell+[\theta\ell]}})(\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+[\theta\ell]}}). \quad (\text{C.3})$$

By iterating (C.1),

$$\begin{aligned} \tilde{\xi}_n^{h_\ell} - \xi_\star^{h_{\ell+[\theta\ell]}} &= (\tilde{\xi}_0^{h_\ell} - \xi_\star^{h_{\ell+[\theta\ell]}})\Pi_{1:n} - \sum_{k=1}^n \gamma_k \Pi_{k+1:n} g_k^{h_\ell} - \sum_{k=1}^n \gamma_k \Pi_{k+1:n} \rho_k^{h_\ell} \\ &\quad - \sum_{k=1}^n \gamma_k \Pi_{k+1:n} r_k^{h_\ell} - \sum_{k=1}^n \gamma_k \Pi_{k+1:n} e_k^{h_\ell}, \end{aligned} \quad (\text{C.4})$$

where

$$\Pi_{k:n} = \prod_{j=k}^n (1 - \gamma_j V_0''(\xi_\star^0)) \quad (\text{C.5})$$

with the convention  $\prod_{\emptyset} = 1$ .

Given that  $\gamma_k \downarrow 0$  as  $k \uparrow \infty$ , there exists  $k_0 \geq 0$  such that for  $j \geq k_0$ ,  $(1 - \gamma_j V_0''(\xi_\star^0)) > 0$ . Hence, using the inequality  $1 + x \leq e^x$ ,  $x \in \mathbb{R}$ , for all  $n \in \mathbb{N}^*$ ,

$$\begin{aligned} |\Pi_{k:n}| &= |\Pi_{k:k_0-1}| \prod_{j=k_0 \vee k}^n (1 - \gamma_j V_0''(\xi_\star^0)) \\ &\leq |\Pi_{k:k_0-1}| \exp\left(-V_0''(\xi_\star^0) \sum_{j=k_0 \vee k}^n \gamma_j\right) \\ &\leq \bar{C} \exp\left(-V_0''(\xi_\star^0) \sum_{j=k}^n \gamma_j\right), \end{aligned} \quad (\text{C.6})$$

where

$$\bar{C} = 1 \vee \max_{1 \leq k \leq k_0} \left\{ |\Pi_{k:k_0-1}| \exp\left(V_0''(\xi_\star^0) \sum_{j=k_0 \wedge k}^{k-1} \gamma_j\right) \right\},$$

with the convention  $\sum_{\emptyset} = 0$ .

According to (5.1) and the decomposition (C.4),

$$\begin{aligned} \tilde{\xi}_{\mathbf{N}}^{\text{ML}} - \xi_\star^{h_{L+[\theta L]}} &= \tilde{\xi}_{N_0}^{h_0} - \xi_\star^{h_0} + \sum_{\ell=1}^L (\tilde{\xi}_{N_\ell}^{h_\ell} - \xi_\star^{h_{\ell+[\theta\ell]}} - (\tilde{\xi}_{N_\ell}^{h_{\ell-1}} - \xi_\star^{h_{\ell-1+[\theta(\ell-1)]}})) \\ &= \tilde{\xi}_{N_0}^{h_0} - \xi_\star^{h_0} + \sum_{\ell=1}^L (\tilde{\xi}_0^{h_\ell} - \xi_\star^{h_{\ell+[\theta\ell]}} - (\tilde{\xi}_0^{h_{\ell-1}} - \xi_\star^{h_{\ell-1+[\theta(\ell-1)]}}))\Pi_{1:N_\ell} \\ &\quad - \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (g_k^{h_\ell} - g_k^{h_{\ell-1}}) - \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}) \\ &\quad - \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (r_k^{h_\ell} - r_k^{h_{\ell-1}}) - \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}). \end{aligned} \quad (\text{C.7})$$

We study each term in the above decomposition separately.

*Step 1. Study of  $\tilde{\xi}_{N_0}^{h_0} - \xi_\star^{h_0}$ .*

Following Remark 3.3(ii), no refinement is applied at level 0, so that by Lemma 2.5(ii),

$$\mathbb{E}[(\tilde{\xi}_{N_0}^{h_0} - \xi_\star^{h_0})^2] \leq C_{-1} \gamma_{N_0},$$

where  $C_{-1} > 0$  is the constant described in (2.16).

*Step 2. Study of  $\sum_{\ell=1}^L (\tilde{\xi}_0^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}} - (\tilde{\xi}_0^{h_{\ell-1}} - \xi_\star^{h_{\ell-1+\lceil\theta(\ell-1)\rceil}})) \Pi_{1:N_\ell}$ .*

By assumption,  $\sup_{\ell \geq 0} \mathbb{E}[|\tilde{\xi}_0^{h_\ell}|^2] < \infty$ , and according to Lemma 2.5(i),  $(\xi_\star^h)_{h \in \bar{\mathcal{H}}}$  is bounded. Hence  $\sup_{\ell \geq 0} \mathbb{E}[(\tilde{\xi}_0^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})^2]^{\frac{1}{2}} \leq \sup_{\ell \geq 0} \mathbb{E}[|\tilde{\xi}_0^{h_\ell}|^2]^{\frac{1}{2}} + \sup_{\ell \geq 0} |\xi_\star^{h_\ell}| < \infty$ . Besides, by (C.6) and [15, Lemma A.1(ii)],  $\limsup_{n \uparrow \infty} \gamma_n^{-1} \exp(-V_0''(\xi_\star^0) \sum_{k=1}^n \gamma_k) = 0$  under the condition  $\bar{\lambda}_2 \gamma_1 > 2$ , since the latter entails that  $\gamma_1 V_0''(\xi_\star^0) > \bar{\lambda}_2 \gamma_1 > 2 > 0$ . Hence,

$$\begin{aligned} & \mathbb{E} \left[ \left( \sum_{\ell=1}^L (\tilde{\xi}_0^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}} - (\tilde{\xi}_0^{h_{\ell-1}} - \xi_\star^{h_{\ell-1+\lceil\theta(\ell-1)\rceil}})) \Pi_{1:N_\ell} \right)^2 \right]^{\frac{1}{2}} \\ & \leq 2 \sup_{\ell' \geq 0} \mathbb{E}[(\tilde{\xi}_0^{h_{\ell'}} - \xi_\star^{h_{\ell'+\lceil\theta\ell'\rceil}})^2]^{\frac{1}{2}} \sum_{\ell=1}^L |\Pi_{1:N_\ell}| \\ & \leq C_0 \sum_{\ell=1}^L \gamma_{N_\ell}, \end{aligned}$$

where

$$C_0 = 2\bar{C} \left( \sup_{\ell' \geq 0} \mathbb{E}[|\tilde{\xi}_0^{h_{\ell'}}|^2]^{\frac{1}{2}} + \sup_{\ell' \geq 0} |\xi_\star^{h_{\ell'}}| \right) \sup_{m \geq 1} \left\{ \gamma_m^{-1} \exp \left( -V_0''(\xi_\star^0) \sum_{k=1}^m \gamma_k \right) \right\}.$$

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

Recalling that, by Lemma 2.5(i),  $(\xi_\star^h)_{h \in \bar{\mathcal{H}}}$  is bounded, there exists a compact set  $\mathcal{K} \subset \mathbb{R}$  such that  $\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}} \in \mathcal{K}$ ,  $\ell \in \mathbb{N}$ . From Lemma 2.5(i) and Assumptions 2.3(iv) and 2.10,

$$\begin{aligned} & |V_0''(\xi_\star^0) - V_{h_{\ell+\lceil\theta\ell\rceil}}''(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})| \\ & \leq |V_0''(\xi_\star^0) - V_0''(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})| + |V_0''(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) - V_{h_{\ell+\lceil\theta\ell\rceil}}''(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})| \\ & \leq \frac{1}{1-\alpha} \left( [f_{X^0}]_{\text{Lip}} |\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}} - \xi_\star^0| + \sup_{\xi \in \mathcal{K}} |f_{X^0}(\xi) - f_{X^{h_{\ell+\lceil\theta\ell\rceil}}}(\xi)| \right) \quad (\text{C.8}) \\ & \leq \frac{[f_{X^0}]_{\text{Lip}} c' + c}{1-\alpha} (h_{\ell+\lceil\theta\ell\rceil} \vee h_{\ell+\lceil\theta\ell\rceil}^{\frac{1}{4}+\delta_0}) \\ & \leq C_g^1 h_\ell^{((\frac{1}{4}+\delta_0) \wedge 1)(1+\theta)}, \end{aligned}$$

where  $c' \geq 0$  is the constant described in (2.14), and

$$C_g^1 = M^{(\frac{1}{4}+\delta_0) \wedge 1} \frac{[f_{X^0}]_{\text{Lip}} c' + c}{1-\alpha}.$$

Recalling the definitions (4.7), (B.14) and (B.27), using Proposition 4.4(ii), (B.28) and the fact that

$$\tilde{u}_n^2 = \gamma_1^{\frac{2\beta-\delta'}{\beta}} \gamma_n^{\frac{2\delta'}{\beta}}, \quad n \in \mathbb{N}^*, \quad (\text{C.9})$$

one gets

$$\begin{aligned} & \mathbb{E}[(\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})^4] \\ & \leq \begin{cases} \left( \sup_{\ell' \geq 1} \mathbb{E}[|\tilde{\xi}_0^{h_{\ell'}}|^4] + \sup_{\ell' \geq 1} |\xi_\star^{h_{\ell'}}|^4 \right) & \text{if } n = 1, \\ C^\beta (\tilde{\gamma}_{n-1}^\ell)^2 & \text{if } n \geq 2, \end{cases} \\ & \leq \left\{ \left( 8\gamma_1^{-2} \left( \sup_{\ell' \geq 1} \mathbb{E}[|\tilde{\xi}_0^{h_{\ell'}}|^4] + \sup_{\ell' \geq 1} |\xi_\star^{h_{\ell'}}|^4 \right) \right) \vee (2^{2(\beta \vee \delta')} C^\beta) \right\} (h_\ell^{2(1+\theta)} \tilde{u}_n^2 + \gamma_n^2) \\ & \leq C_g^2 (h_\ell^{2(1+\theta)} \gamma_n^{\frac{2\delta'}{\beta}} + \gamma_n^2), \quad n \in \mathbb{N}^*, \end{aligned} \quad (\text{C.10})$$

where

$$C_g^2 = \left\{ \left( 8\gamma_1^{-2} \left( \sup_{\ell' \geq 1} \mathbb{E}[|\tilde{\xi}_0^{h_{\ell'}}|^4] + \sup_{\ell' \geq 1} |\xi_\star^{h_{\ell'}}|^4 \right) \vee (2^{2(\beta \vee \delta')} C^\beta) \right\} (\gamma_1^{2\frac{\beta-\delta'}{\beta}} \vee 1),$$

with  $C^\beta > 0$  being the constant described in (4.9). Consequently, using (C.2), (C.8) and the inequality  $(x+y)^{\frac{1}{4}} \leq x^{\frac{1}{4}} + y^{\frac{1}{4}}$ ,  $x, y \geq 0$ , one gets

$$\begin{aligned} \mathbb{E}[|g_n^{h_\ell}|^2]^{\frac{1}{2}} &\leq |V_0''(\xi_\star^0) - V_{h_{\ell+\lceil\theta\ell\rceil}}''(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})| \mathbb{E}[(\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})^4]^{\frac{1}{4}} \\ &\leq C_g^3 h_\ell^{((\frac{1}{4}+\delta_0)\wedge 1)(1+\theta)} (h_\ell^{\frac{1+\theta}{2}} \gamma_n^{\frac{\delta'}{2\beta}} + \gamma_n^{\frac{1}{2}}), \end{aligned} \quad (\text{C.11})$$

where

$$C_g^3 = C_g^1 (C_g^2)^{\frac{1}{4}}.$$

Recalling (C.6), under the condition  $\bar{\lambda}_2 \gamma_1 > 2$  if  $\beta = 1$ , which entails that  $\gamma_1 V_0''(\xi_\star^0) > \bar{\lambda}_2 \gamma_1 > 2 \geq \frac{1}{2} \vee \frac{\delta'}{2}$  if  $\beta = 1$ , by [15, Lemma A.1(i)],

$$\begin{aligned} &\mathbb{E} \left[ \left( \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (g_k^{h_\ell} - g_k^{h_{\ell-1}}) \right)^2 \right]^{\frac{1}{2}} \\ &\leq \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k |\Pi_{k+1:N_\ell}| (\mathbb{E}[|g_k^{h_\ell}|^2]^{\frac{1}{2}} + \mathbb{E}[|g_k^{h_{\ell-1}}|^2]^{\frac{1}{2}}) \\ &\leq C_g^3 (M+1) \sum_{\ell=1}^L h_\ell^{((\frac{1}{4}+\delta_0)\wedge 1)(1+\theta)} \left( \sum_{k=1}^{N_\ell} h_\ell^{\frac{1+\theta}{2}} \gamma_k^{1+\frac{\delta'}{2\beta}} |\Pi_{k+1:N_\ell}| + \sum_{k=1}^{N_\ell} \gamma_k^{\frac{3}{2}} |\Pi_{k+1:N_\ell}| \right) \\ &\leq C_g^4 \sum_{\ell=1}^L h_\ell^{((\frac{1}{4}+\delta_0)\wedge 1)(1+\theta)} (h_\ell^{\frac{1+\theta}{2}} \gamma_{N_\ell}^{\frac{\delta'}{2\beta}} + \gamma_{N_\ell}^{\frac{1}{2}}) \\ &\leq C_g^5 \sum_{\ell=1}^L h_\ell^{(\frac{1}{4}+(\delta_0\wedge\frac{3}{4}))(1+\theta)} (\tilde{\gamma}_{N_\ell}^\ell)^{\frac{1}{2}}, \end{aligned}$$

where

$$\begin{aligned} C_g^4 &= C_g^3 \bar{C} (M+1) \left( \sup_{m \geq 1} \left\{ \gamma_m^{-\frac{\delta'}{2\beta}} \sum_{k=1}^m \gamma_k^{1+\frac{\delta'}{2\beta}} \exp \left( -V_0''(\xi_\star^0) \sum_{j=k+1}^m \gamma_j \right) \right\} \right. \\ &\quad \left. \vee \sup_{m \geq 1} \left\{ \gamma_m^{-\frac{1}{2}} \sum_{k=1}^m \gamma_k^{\frac{3}{2}} \exp \left( -V_0''(\xi_\star^0) \sum_{j=k+1}^m \gamma_j \right) \right\} \right), \\ C_g^5 &= \sqrt{2} C_g^4 (\gamma_1^{-1+\frac{\delta'}{2\beta}} \vee 1). \end{aligned}$$

Hence, using the Cauchy-Schwarz inequality and the fact that, recalling (2.27),  $h^{\frac{1}{2}} \leq \epsilon(h)$ ,  $h \in \mathcal{H} \setminus \{1\}$ , one gets

$$\begin{aligned} &\mathbb{E} \left[ \left( \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (g_k^{h_\ell} - g_k^{h_{\ell-1}}) \right)^2 \right]^{\frac{1}{2}} \\ &\leq C_g^5 \left( \sum_{\ell=1}^L \tilde{\gamma}_{N_\ell}^\ell h_\ell^{\frac{1}{2}(1+\theta)} \right)^{\frac{1}{2}} \left( \sum_{\ell=1}^L h_\ell^{2(1+\theta)(\delta_0\wedge\frac{3}{4})} \right)^{\frac{1}{2}} \\ &\leq C_g \left( \sum_{\ell=1}^L \tilde{\gamma}_{N_\ell}^\ell \epsilon(h_\ell)^{1+\theta} \right)^{\frac{1}{2}}, \end{aligned} \quad (\text{C.12})$$

where

$$C_g = C_g^5 \left( \sum_{\ell=1}^{\infty} h_\ell^{2(1+\theta)(\delta_0 \wedge \frac{3}{4})} \right)^{\frac{1}{2}} = \frac{h_0^{(1+\theta)(\delta_0 \wedge \frac{3}{4})} C_g^5}{(1 - M^{-2(1+\theta)(\delta_0 \wedge \frac{3}{4})})^{\frac{1}{2}}}.$$

*Step 4. Study of  $\sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}})$ .*

Taking into account the fact that  $V'_{h_{\ell+\lceil\theta\ell\rceil}}(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) = 0$  and Assumption 2.3(iv), a first order Taylor expansion yields

$$\begin{aligned} & |V'_{h_{\ell+\lceil\theta\ell\rceil}}(\tilde{\xi}_{n-1}^{h_\ell}) - V''_{h_{\ell+\lceil\theta\ell\rceil}}(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})(\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})| \\ &= \left| (\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) \int_0^1 \left( V''_{h_{\ell+\lceil\theta\ell\rceil}}(t\tilde{\xi}_{n-1}^{h_\ell} + (1-t)\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) - V''_{h_{\ell+\lceil\theta\ell\rceil}}(\xi_\star^{h_{\ell+\lceil\theta\ell\rceil}}) \right) dt \right| \\ &\leq \frac{[f_{X^{h_\ell}}]_{\text{Lip}}}{2(1-\alpha)} (\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})^2. \end{aligned}$$

Hence, by (C.3), (C.10) and the inequality  $(x+y)^{\frac{1}{2}} \leq x^{\frac{1}{2}} + y^{\frac{1}{2}}$ ,  $x, y \geq 0$ ,

$$\mathbb{E}[|\rho_n^{h_\ell}|^2]^{\frac{1}{2}} \leq \frac{[f_{X^{h_\ell}}]_{\text{Lip}}}{2(1-\alpha)} \mathbb{E}[(\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_{\ell+\lceil\theta\ell\rceil}})^4]^{\frac{1}{2}} \leq C_\rho^1 (h_\ell^{1+\theta} \gamma_n^{\frac{\delta'}{\beta}} + \gamma_n),$$

where

$$C_\rho^1 = \frac{(C_g^2)^{\frac{1}{2}} \sup_{\ell' \geq 0} [f_{X^{h_{\ell'}}}]_{\text{Lip}}}{2(1-\alpha)}.$$

Recalling (C.6) and that, by assumption,  $\bar{\lambda}_2 \gamma_1 > 2$  if  $\beta = 1$ , i.e.  $\gamma_1 V_0''(\xi_\star^0) > \bar{\lambda}_2 \gamma_1 > 2 > 1 \vee \delta'$  if  $\beta = 1$ , via [15, Lemma A.1(i)]

$$\begin{aligned} & \mathbb{E} \left[ \left( \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (\rho_k^{h_\ell} - \rho_k^{h_{\ell-1}}) \right)^2 \right]^{\frac{1}{2}} \\ & \leq \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k |\Pi_{k+1:N_\ell}| (\mathbb{E}[|\rho_k^{h_\ell}|^2]^{\frac{1}{2}} + \mathbb{E}[|\rho_k^{h_{\ell-1}}|^2]^{\frac{1}{2}}) \\ & \leq C_\rho^1 (M+1) \sum_{\ell=1}^L \left( \sum_{k=1}^{N_\ell} h_\ell^{1+\theta} \gamma_k^{1+\frac{\delta'}{\beta}} |\Pi_{k+1:N_\ell}| + \sum_{k=1}^{N_\ell} \gamma_k^2 |\Pi_{k+1:N_\ell}| \right) \\ & \leq C_\rho^2 \sum_{\ell=1}^L (h_\ell^{1+\theta} \gamma_{N_\ell}^{\frac{\delta'}{\beta}} + \gamma_{N_\ell}) \\ & \leq C_\rho \sum_{\ell=1}^L \tilde{\gamma}_{N_\ell}^\ell, \end{aligned}$$

where

$$\begin{aligned} C_\rho^2 &= C_\rho^1 \bar{C} (M+1) \left( \sup_{m \geq 1} \left\{ \gamma_m^{-\frac{\delta'}{\beta}} \sum_{k=1}^m \gamma_k^{1+\frac{\delta'}{\beta}} \exp \left( -V_0''(\xi_\star^0) \sum_{j=k+1}^m \gamma_j \right) \right\} \right. \\ & \quad \left. \vee \sup_{m \geq 1} \left\{ \gamma_m^{-1} \sum_{k=1}^m \gamma_k^2 \exp \left( -V_0''(\xi_\star^0) \sum_{j=k+1}^m \gamma_j \right) \right\} \right), \\ C_\rho &= 2C_\rho^2 (\gamma_1^{-1+\frac{\delta'}{\beta}} \vee 1). \end{aligned}$$

*Step 5. Study of  $\sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (r_k^{h_\ell} - r_k^{h_{\ell-1}})$ .*

Recalling the definition (B.15), it follows from the inequality (B.13) that

$$|r_n^{h_\ell}| \leq C_r^1 h_\ell^{1+\theta} \gamma_n^{\frac{\delta'}{\beta}}.$$

where

$$C_r^1 = Ch_1^{-1} \gamma_1^{1-\frac{\delta'}{\beta}}.$$

Thus, using (C.6) and recalling that  $\bar{\lambda}_2 \gamma_1 > 2$  if  $\beta = 1$ , hence  $\gamma_1 V_0''(\xi_\star^0) > \bar{\lambda}_2 \gamma_1 > 2 > \delta'$  if  $\beta = 1$ , by [15, Lemma A.1(i)],

$$\begin{aligned} & \mathbb{E} \left[ \left( \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (r_k^{h_\ell} - r_k^{h_{\ell-1}}) \right)^2 \right]^{\frac{1}{2}} \\ & \leq \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k |\Pi_{k+1:N_\ell}| (\mathbb{E}[|r_k^{h_\ell}|^2]^{\frac{1}{2}} + \mathbb{E}[|r_k^{h_{\ell-1}}|^2]^{\frac{1}{2}}) \\ & \leq C_r^1 (M+1) \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} h_\ell^{1+\theta} \gamma_k^{1+\frac{\delta'}{\beta}} |\Pi_{k+1:N_\ell}| \\ & \leq C_r \sum_{\ell=1}^L h_\ell^{1+\theta} \gamma_{N_\ell}^{\frac{\delta'}{\beta}} \\ & \leq C_r \gamma_1^{-1+\frac{\delta'}{\beta}} \sum_{\ell=1}^L \tilde{\gamma}_{N_\ell}^\ell, \end{aligned}$$

where

$$C_r = C_r^1 (M+1) \sup_{m \geq 1} \left\{ \gamma_m^{-\frac{\delta'}{\beta}} \sum_{k=1}^m \gamma_k^{1+\frac{\delta'}{\beta}} \exp \left( -V_0''(\xi_\star^0) \sum_{j=k+1}^m \gamma_j \right) \right\}.$$

*Step 6. Study of  $\sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}})$ .*

Note that the random variables  $(\sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}))_{\ell \geq 1}$  are independent with zero mean and that, at each level  $\ell \in \mathbb{N}^*$ ,  $(e_k^{h_\ell} - e_k^{h_{\ell-1}})_{k \geq 1}$  are  $(\tilde{\mathcal{F}}_k^{h_\ell})_{k \geq 1}$ -martingale increments. Therefore

$$\mathbb{E} \left[ \left( \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:N_\ell} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) \right)^2 \right] = \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k^2 |\Pi_{k+1:n}|^2 \mathbb{E}[(e_k^{h_\ell} - e_k^{h_{\ell-1}})^2].$$

From (B.12), (B.9) and (2.3),

$$\begin{aligned} \mathbb{E}[(e_n^{h_\ell} - e_n^{h_{\ell-1}})^2] &= \mathbb{E}[\text{Var}(H(\tilde{\xi}_{n-1}^{h_\ell}, \tilde{X}_n^{h_\ell}) - H(\tilde{\xi}_{n-1}^{h_{\ell-1}}, \tilde{X}_n^{h_{\ell-1}}) | \tilde{\mathcal{F}}_{n-1}^{h_\ell})] \\ &\leq \mathbb{E}[(H(\tilde{\xi}_{n-1}^{h_\ell}, \tilde{X}_n^{h_\ell}) - H(\tilde{\xi}_{n-1}^{h_{\ell-1}}, \tilde{X}_n^{h_{\ell-1}}))^2] \\ &\leq \frac{1}{(1-\alpha)^2} (\mathbb{E}[|\mathbb{1}_{\{\tilde{X}_n^{h_\ell} > \tilde{\xi}_{n-1}^{h_\ell}\}} - \mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_\ell}\}}|] \\ &\quad + \mathbb{E}[|\mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_\ell}\}} - \mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_{\ell-1}}\}}|] \\ &\quad + \mathbb{E}[|\mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_{\ell-1}}\}} - \mathbb{1}_{\{\tilde{X}_n^{h_{\ell-1}} > \tilde{\xi}_{n-1}^{h_{\ell-1}}\}}|]). \end{aligned}$$

On the one hand, for  $j \in \{\ell-1, \ell\}$ , recalling the definition (2.27), by the tower law and Lemma 5.2,

$$\begin{aligned} \mathbb{E}[|\mathbb{1}_{\{\tilde{X}_n^{h_j} > \tilde{\xi}_{n-1}^{h_j}\}} - \mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_j}\}}|] &= \mathbb{E}[\mathbb{E}[|\mathbb{1}_{\{\tilde{X}_n^{h_j} > \tilde{\xi}_{n-1}^{h_j}\}} - \mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_j}\}}| | \tilde{\mathcal{F}}_{n-1}^{h_j}]] \\ &= \mathbb{E}[\mathbb{E}[|\mathbb{1}_{\{\tilde{X}_n^{h_j} > \xi\}} - \mathbb{1}_{\{X^0 > \xi\}}| | \xi = \tilde{\xi}_{n-1}^{h_j}]] \\ &\leq C'_e \epsilon(h_j)^{1+\theta} \\ &\leq C_e^1 \epsilon(h_\ell)^{1+\theta}, \end{aligned}$$

where

$$C_e^1 = C'_e M^{\frac{1}{2}} (\ln M)^{\frac{1}{2}},$$

and

$$C' = \begin{cases} C_1 & \text{if (5.2) holds,} \\ C_2 & \text{if (5.3) holds,} \\ C_3 & \text{if (5.4) holds,} \end{cases}$$

with  $C_1$ ,  $C_2$  and  $C_3$  being the positive constants defined in Lemma 5.2.

On the other hand, using Assumption 2.3(iv),

$$\begin{aligned} & \mathbb{E}[|\mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_\ell}\}} - \mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_{\ell-1}}\}}|] \\ &= \mathbb{E}[\mathbb{1}_{\{\tilde{\xi}_{n-1}^{h_\ell} < X^0 \leq \tilde{\xi}_{n-1}^{h_{\ell-1}}\}} + \mathbb{1}_{\{\tilde{\xi}_{n-1}^{h_{\ell-1}} < X^0 \leq \tilde{\xi}_{n-1}^{h_\ell}\}}] \\ &= \mathbb{E}[|F_{X^0}(\tilde{\xi}_{n-1}^{h_\ell}) - F_{X^0}(\tilde{\xi}_{n-1}^{h_{\ell-1}})|] \\ &\leq \|f_{X^0}\|_\infty \mathbb{E}[(\tilde{\xi}_{n-1}^{h_\ell} - \tilde{\xi}_{n-1}^{h_{\ell-1}})^4]^{\frac{1}{4}} \\ &\leq \|f_{X^0}\|_\infty (\mathbb{E}[(\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_\ell + \lceil \theta \ell \rceil})^4]^{\frac{1}{4}} + |\xi_\star^{h_\ell + \lceil \theta \ell \rceil} - \xi_\star^0| \\ &\quad + |\xi_\star^0 - \xi_\star^{h_{\ell-1} + \lceil \theta(\ell-1) \rceil}| + \mathbb{E}[(\xi_\star^{h_{\ell-1} + \lceil \theta(\ell-1) \rceil} - \tilde{\xi}_{n-1}^{h_{\ell-1}})^4]^{\frac{1}{4}}). \end{aligned}$$

Recalling (C.10), by using the inequality  $(x + y)^{\frac{1}{4}} \leq x^{\frac{1}{4}} + y^{\frac{1}{4}}$ ,  $x, y \geq 0$ ,

$$\mathbb{E}[(\tilde{\xi}_{n-1}^{h_\ell} - \xi_\star^{h_\ell + \lceil \theta \ell \rceil})^4]^{\frac{1}{4}} \leq (C_g^2)^{\frac{1}{4}} (h_\ell^{\frac{1+\theta}{2}} \gamma_n^{\frac{\delta'}{2\beta}} + \gamma_n^{\frac{1}{2}}).$$

Besides, by Lemma 2.5(i)

$$|\xi_\star^{h_\ell + \lceil \theta \ell \rceil} - \xi_\star^0| \leq c' h_0^{-1} h_\ell^{1+\theta},$$

recalling that  $c' \geq 0$  is the constant described in (2.14). Thus

$$\begin{aligned} & \mathbb{E}[|\mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_\ell}\}} - \mathbb{1}_{\{X^0 > \tilde{\xi}_{n-1}^{h_{\ell-1}}\}}|] \\ &\leq \|f_{X^0}\|_\infty ((C_g^2)^{\frac{1}{4}} (M+1) (h_\ell^{\frac{1+\theta}{2}} \gamma_n^{\frac{\delta'}{2\beta}} + \gamma_n^{\frac{1}{2}}) + c' h_0^{-1} (M^2 + 1) h_\ell^{1+\theta}) \\ &\leq C_e^2 (h_\ell^{1+\theta} + h_\ell^{\frac{1+\theta}{2}} \gamma_n^{\frac{\delta'}{2\beta}} + \gamma_n^{\frac{1}{2}}), \end{aligned}$$

where

$$C_e^2 = 2 \|f_{X^0}\|_\infty \left( ((C_g^2)^{\frac{1}{4}} (M+1)) \vee (c' h_0^{-1} (M^2 + 1)) \right).$$

Combining the previous results and invoking [15, Lemma A.1(i)], recalling the decreasing monotonicity of  $(\gamma_n)_{n \geq 1}$  and the fact that  $\bar{\lambda}_2 \gamma_1 > 2$  if  $\beta = 1$ , i.e.  $\gamma_1 V_0''(\xi_\star^0) > \bar{\lambda}_2 \gamma_1 > 2 > (1 + \frac{\delta'}{2}) \vee \frac{3}{2}$  if  $\beta = 1$ ,

$$\begin{aligned} & \mathbb{E} \left[ \left( \sum_{\ell=1}^L \sum_{k=1}^{N_\ell} \gamma_k \Pi_{k+1:n} (e_k^{h_\ell} - e_k^{h_{\ell-1}}) \right)^2 \right] \\ &\leq \frac{1}{(1-\alpha)^2} \sum_{\ell=1}^L \left( C_e^1 \sum_{k=1}^{N_\ell} \epsilon (h_\ell)^{1+\theta} \gamma_k^2 |\Pi_{k+1:N_\ell}|^2 + C_e^2 \sum_{k=1}^{N_\ell} h_\ell^{1+\theta} \gamma_k^2 |\Pi_{k+1:N_\ell}|^2 \right. \\ &\quad \left. + C_e^2 \sum_{k=1}^{N_\ell} h_\ell^{\frac{1+\theta}{2}} \gamma_k^{2+\frac{\delta'}{2\beta}} |\Pi_{k+1:N_\ell}|^2 + C_e^2 \sum_{k=1}^{N_\ell} \gamma_k^{\frac{5}{2}} |\Pi_{k+1:N_\ell}|^2 \right) \\ &\leq C_e \sum_{\ell=1}^L (\gamma_{N_\ell} \epsilon (h_\ell)^{1+\theta} + \gamma_{N_\ell} (\tilde{\gamma}_{N_\ell}^\ell)^{\frac{1}{2}}), \end{aligned}$$

where

$$C_e = \frac{1}{(1-\alpha)^2} \bar{C}^2 \left( (C_e^1 + C_e^2) \sup_{m \geq 1} \left\{ \gamma_m^{-1} \sum_{k=1}^m \gamma_k^2 \exp \left( -2V_0''(\xi_\star^0) \sum_{j=k+1}^m \gamma_j \right) \right\} \right. \\ \vee C_e^2 \left( \sup_{m \geq 1} \left\{ \gamma_m^{-1-\frac{\delta'}{2\beta}} \sum_{k=1}^m \gamma_k^{2+\frac{\delta'}{2\beta}} \exp \left( -2V_0''(\xi_\star^0) \sum_{j=k+1}^m \gamma_j \right) \right\} \right. \\ \left. \left. + \sup_{m \geq 1} \left\{ \gamma_m^{-\frac{3}{2}} \sum_{k=1}^m \gamma_k^{\frac{5}{2}} \exp \left( -2V_0''(\xi_\star^0) \sum_{j=k+1}^m \gamma_j \right) \right\} \right) \right).$$

*Step 7. Conclusion.*

Combining the previous upper bounds, we obtain

$$\mathbb{E}[(\tilde{\xi}_{\mathbf{N}}^{\text{ML}} - \xi_\star^{h_{L+[\theta L]}})^2] \leq \hat{C} \left( \gamma_{N_0} + \left( \sum_{\ell=1}^L \tilde{\gamma}_{N_\ell}^\ell \right)^2 + \sum_{\ell=1}^L \gamma_{N_\ell} (\tilde{\gamma}_{N_\ell}^\ell)^{\frac{1}{2}} + \sum_{\ell=1}^L \gamma_{N_\ell} \epsilon(h_\ell)^{1+\theta} \right),$$

where

$$\hat{C} = 6 \left( C_{-1} \vee (C_0)^2 \vee (C_g)^2 \vee (C_p)^2 \vee ((C_r)^2 \gamma_1^{-2+2\frac{\delta'}{\beta}}) \vee C_e \right). \quad (\text{C.13})$$

## D Proofs of the Complexity Results

*Proof of Proposition 4.6.* By Lemma 2.5(i), the bias error of adNSA satisfies

$$|\xi_\star^{h_{\ell+[\theta \ell]}} - \xi_\star^0| \leq C h_{\ell+[\theta \ell]} \leq \frac{C h_0}{M^{\ell(1+\theta)}}, \quad (\text{D.1})$$

where  $C > 0$  is the constant described in (2.14). Hence, assuming that  $h_0 > C^{-1}\varepsilon$ , in order to achieve a bias error of order  $\varepsilon > 0$ , it suffices to take  $\ell$  minimal such that

$$\frac{C h_0}{M^{\ell(1+\theta)}} \leq \varepsilon, \quad \text{i.e.} \quad \ell = \left\lceil \frac{\ln(C h_0 \varepsilon^{-1})}{(1+\theta) \ln M} \right\rceil \geq 1. \quad (\text{D.2})$$

□

*Proof of Lemma 4.8.* Let  $\ell, n \in \mathbb{N}^*$ . By the law of total probability,

$$\mathbb{E}[h_{\ell+\tilde{\eta}_n^\ell}^{-1}] = \sum_{k=0}^{[\theta \ell]} h_{\ell+k}^{-1} \mathbb{P}(\tilde{\eta}_n^\ell = k) \leq h_\ell^{-1} + \sum_{k=1}^{[\theta \ell]} h_{\ell+k}^{-1} \mathbb{P}(\tilde{\eta}_n^\ell = k).$$

For  $k \in \llbracket 1, [\theta \ell] \rrbracket$ , by the definition (3.10), the tower law and Assumption 2.3(iv),

$$\mathbb{P}(\tilde{\eta}_n^\ell = k) \leq \mathbb{P}(|X_n^{h_{\ell+k-1}} - \tilde{\xi}_{n-1}^{h_\ell}| < C_{\text{ad}} \psi_n^{\ell, k-1}) \\ = \mathbb{E}[\mathbb{P}(|X_n^{h_{\ell+k-1}} - \xi| < C_{\text{ad}} \psi_n^{\ell, k-1})_{|\xi = \tilde{\xi}_{n-1}^{h_\ell}}] \\ \leq 2C_{\text{ad}} \left( \sup_{\ell' \geq 0} \|f_{X^{h_{\ell'}}}\|_\infty \right) \psi_n^{\ell, k-1}.$$

Thus

$$\mathbb{E}[h_{\ell+\tilde{\eta}_n^\ell}^{-1}] \leq C_0 \left( h_\ell^{-1} + \sum_{k=1}^{[\theta \ell]} h_{\ell+k}^{-1} \psi_n^{\ell, k-1} \right),$$

where

$$C_0 = \left( 2C_{\text{ad}} \sup_{\ell' \geq 0} \|f_{X^{h_{\ell'}}}\|_\infty \right) \vee 1.$$

Recall now the definition (3.11) and the fact that  $r > 1$ . If (2.23) holds, then

$$\begin{aligned}
\mathbb{E}[h_{\ell+\tilde{\eta}_n^\ell}^{-1}] &\leq C_0 \left( h_\ell^{-1} + u_n^{-\frac{1}{p^*}} \sum_{k=1}^{\lceil \theta \ell \rceil} h_{\ell+k}^{-1} h_{\theta \ell(r-1)+k}^{\frac{1}{r}} \right) \\
&= C_0 \left( h_\ell^{-1} + \frac{h_0^{-1+\frac{1}{r}} n^{\frac{\delta}{p^*}}}{\gamma_1^{\frac{1}{p^*}} M^{-\ell+\theta \ell(1-\frac{1}{r})}} \sum_{k=1}^{\lceil \theta \ell \rceil} M^{(1-\frac{1}{r})k} \right) \\
&\leq C_0 n^{\frac{\delta}{p^*}} \left( h_\ell^{-1} + \frac{h_0^{-1+\frac{1}{r}}}{\gamma_1^{\frac{1}{p^*}} M^{-\ell+\theta \ell(1-\frac{1}{r})}} M^{1-\frac{1}{r}} \frac{M^{(2+\theta \ell)(1-\frac{1}{r})}}{M^{1-\frac{1}{r}} - 1} \right) \\
&= C_1 n^{\frac{\delta}{p^*}} h_\ell^{-1},
\end{aligned}$$

where

$$C_1 = C_0 \left( 1 + \frac{h_0^{\frac{1}{r}} M^{3(1-\frac{1}{r})}}{\gamma_1^{\frac{1}{p^*}} (M^{1-\frac{1}{r}} - 1)} \right).$$

Next, recalling that  $h_0 = \frac{1}{K}$ , note that, for  $k \in \llbracket 1, \lceil \theta \ell \rceil \rrbracket$ ,

$$\begin{aligned}
\ln \left( \left( \frac{\gamma_n}{\gamma_1} \right)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}} \right) &= \frac{\beta}{2} \ln n + \frac{1+\theta}{2} (\ell+k) \ln M + \frac{1+\theta}{2} \ln K \\
&\leq \frac{\beta}{2} \vee \left( \frac{1}{2} (1+\theta)(2+\theta) \ln M \right) \vee \left( \frac{1+\theta}{2} \ln K \right) (1 + \ell + \ln n) \\
&\leq C'_2 (\ell + \ln n),
\end{aligned}$$

where

$$C'_2 = \beta \vee ((1+\theta)(2+\theta) \ln M) \vee ((1+\theta) \ln K).$$

Hence, if (2.24) or (3.9) holds, then

$$\begin{aligned}
\mathbb{E}[h_{\ell+\tilde{\eta}_n^\ell}^{-1}] &\leq C_0 \left( h_\ell^{-1} + \sum_{k=1}^{\lceil \theta \ell \rceil} \ln^{\frac{1}{2}} \left( \left( \frac{\gamma_n}{\gamma_1} \right)^{-\frac{1}{2}} h_{\ell+k}^{-\frac{1+\theta}{2}} \right) h_{\ell+k}^{-1} h_{\theta \ell(r-1)+k}^{\frac{1}{r}} \right) \\
&\leq C_0 (\ell + \ln n)^{\frac{1}{2}} \left( h_\ell^{-1} + C'_2 \frac{h_0^{-1+\frac{1}{r}}}{M^{-\ell+\theta \ell(1-\frac{1}{r})}} \sum_{k=1}^{\lceil \theta \ell \rceil} M^{(1-\frac{1}{r})k} \right) \\
&\leq C_0 (\ell + \ln n)^{\frac{1}{2}} \left( h_\ell^{-1} + C'_2 \frac{h_0^{-1+\frac{1}{r}}}{M^{-\ell+\theta \ell(1-\frac{1}{r})}} M^{1-\frac{1}{r}} \frac{M^{(2+\theta \ell)(1-\frac{1}{r})}}{M^{1-\frac{1}{r}} - 1} \right) \\
&= C_2 (\ell + \ln n)^{\frac{1}{2}} h_\ell^{-1},
\end{aligned}$$

where

$$C_2 = C_0 \left( 1 + C'_2 \frac{h_0^{\frac{1}{r}} M^{3(1-\frac{1}{r})}}{M^{1-\frac{1}{r}} - 1} \right).$$

All in all,

$$\mathbb{E}[h_{\ell+\tilde{\eta}_n^\ell}^{-1}] \leq C \begin{cases} n^{\frac{\delta}{p^*}} h_\ell^{-1} & \text{if (2.23) holds,} \\ (\ell + \ln n)^{\frac{1}{2}} h_\ell^{-1} & \text{if (2.24) or (3.9) holds,} \end{cases}$$

where

$$C = \begin{cases} C_1 & \text{if (2.23) holds,} \\ C_2 & \text{if (2.24) or (3.9) holds.} \end{cases} \quad (\text{D.3})$$

□

*Proof of Proposition 4.10.* Let  $\ell, n \in \mathbb{N}^*$ . Based on (4.1), adNSA performs  $n$  iterations, where, at each iteration  $k \in \llbracket 1, n \rrbracket$ , it simulates one copy of  $Y$  of cost  $c_0$ , and simulates  $h_{\ell + \tilde{\eta}_k}^{-1}$  i.i.d. copies of  $Z$ , each copy being of cost  $c_0$ . Overall, the average computational cost of adNSA satisfies

$$\text{Cost}_{\text{adNSA}} \leq c_0 \sum_{k=1}^n (1 + \mathbb{E}[h_{\ell + \tilde{\eta}_k}^{-1}]) \leq 2c_0 \sum_{k=1}^n \mathbb{E}[h_{\ell + \tilde{\eta}_k}^{-1}].$$

By Lemma 4.8, if (2.23) holds,

$$\text{Cost}_{\text{adNSA}} \leq 2c_0 C' \sum_{k=1}^n h_{\ell}^{-1} k^{\frac{\delta}{p_{\star}}},$$

where  $C' > 0$  is the constant described in (D.3). Via a series-integral comparison,

$$\text{Cost}_{\text{adNSA}} \leq C_1 n^{1 + \frac{\delta}{p_{\star}}} h_{\ell}^{-1}, \quad \text{where} \quad C_1 = 2c_0 C' \left( 1 \vee \frac{2^{\frac{\delta}{p_{\star}}} p_{\star}}{p_{\star} + \delta} \right).$$

If (2.24) or (3.9) holds, using Lemma 4.8 and the inequality  $(x + y)^{\frac{1}{2}} \leq x^{\frac{1}{2}} + y^{\frac{1}{2}}$ ,  $x, y \geq 0$ ,

$$\text{Cost}_{\text{adNSA}} \leq 2c_0 C' \sum_{k=1}^n h_{\ell}^{-1} (\ell + \ln k)^{\frac{1}{2}} \leq 2c_0 C' h_{\ell}^{-1} \left( n \ell^{\frac{1}{2}} + \sum_{k=1}^n \ln^{\frac{1}{2}} k \right).$$

If  $n \geq 2$ , then, via a series-integral comparison, the inequality  $(x + y)^{\frac{1}{2}} \leq x^{\frac{1}{2}} + y^{\frac{1}{2}}$ ,  $x, y \geq 0$ , and the increasing monotony of  $\mathbb{R}_+^* \ni x \mapsto \ln^{\frac{1}{2}} x$ ,

$$\sum_{k=1}^n \ln^{\frac{1}{2}} k \leq n \ln^{\frac{1}{2}} 2 + \sum_{k=2}^n \ln^{\frac{1}{2}} (k - 1) \leq n \ln^{\frac{1}{2}} 2 + \int_1^n \ln^{\frac{1}{2}} s ds \leq 2n \ln^{\frac{1}{2}} n.$$

The above upper bound can also be obtained by integration by parts. Thus, for all  $n \in \mathbb{N}^*$ ,

$$\text{Cost}_{\text{adNSA}} \leq C_2 n (\ell^{\frac{1}{2}} + \ln^{\frac{1}{2}} n) h_{\ell}^{-1}, \quad \text{where} \quad C_2 = 4c_0 C'.$$

Overall,

$$\text{Cost}_{\text{adNSA}} \leq C \begin{cases} n^{1 + \frac{\delta}{p_{\star}}} h_{\ell}^{-1} & \text{if (2.23) holds,} \\ n (\ell^{\frac{1}{2}} + \ln^{\frac{1}{2}} n) h_{\ell}^{-1} & \text{if (2.24) or (3.9) holds,} \end{cases}$$

where

$$C = \begin{cases} C_1 & \text{if (2.23) holds,} \\ C_2 & \text{if (2.24) or (3.9) holds.} \end{cases} \quad (\text{D.4})$$

□

*Proof of Theorem 4.12.* Let  $n \in \mathbb{N}^*$  and  $\varepsilon > 0$ , and set  $\ell \in \mathbb{N}^*$  as in (4.10). According to (D.2), one has  $h_{\ell}^{1+\theta} \leq \frac{\varepsilon}{C'}$ , where  $C' > 0$  is the constant described in (2.14). Thus, consequently from Proposition 4.4(i), the statistical error of adNSA satisfies

$$\begin{aligned} \mathbb{E}[(\tilde{\xi}_n^{h_{\ell}} - \xi_{\star}^{h_{\ell} + \lceil \theta \ell \rceil})^2]^{\frac{1}{2}} &\leq (C^{\beta})^{\frac{1}{2}} (\tilde{\gamma}_n^{\ell})^{\frac{1}{2}} \\ &\leq C \begin{cases} (n^{-\frac{\delta}{2}} \varepsilon^{\frac{1}{2}}) \vee n^{-\frac{\beta}{2}} & \text{if (4.4) holds,} \\ n^{-\frac{\beta}{2}} & \text{if (4.5) or (4.6) holds,} \end{cases} \end{aligned}$$

where

$$C = \left( C^{\beta} ((C')^{-1} \vee 1) \gamma_1 \right)^{\frac{1}{2}}, \quad (\text{D.5})$$

with  $C^\beta > 0$  denoting the constant described in (4.8).

Assume (4.4) holds, with  $(n^{-\frac{\delta}{2}}\varepsilon^{\frac{1}{2}}) \vee n^{-\frac{\beta}{2}} = n^{-\frac{\delta}{2}}\varepsilon^{\frac{1}{2}}$ . To achieve a statistical error of  $\varepsilon$ , it suffices to take  $n$  minimal such that

$$Cn^{-\frac{\delta}{2}}\varepsilon^{\frac{1}{2}} \leq \varepsilon, \quad \text{i.e.} \quad n = \lceil C^{\frac{2}{\delta}}\varepsilon^{-\frac{1}{\delta}} \rceil. \quad (\text{D.6})$$

If (4.4) holds, with  $(n^{-\frac{\delta}{2}}\varepsilon^{\frac{1}{2}}) \vee n^{-\frac{\beta}{2}} = n^{-\frac{\beta}{2}}$ , or if (4.5) or (4.6) holds, then, to achieve a statistical error of order  $\varepsilon$ , we must take  $n$  minimal such that

$$Cn^{-\frac{\beta}{2}} \leq \varepsilon, \quad \text{i.e.} \quad n = \lceil C^{\frac{2}{\beta}}\varepsilon^{-\frac{2}{\beta}} \rceil. \quad (\text{D.7})$$

We now go back to the framework (4.4). If  $n$  is set as in (D.6), then

$$n^{-\frac{\delta}{2}}\varepsilon^{\frac{1}{2}} \leq C^{-1}\varepsilon, \quad \text{and} \quad n^{-\frac{\beta}{2}} > \frac{\varepsilon^{\frac{\beta}{2\delta}}}{(C^{\frac{2}{\delta}} + \varepsilon^{\frac{1}{\delta}})^{\frac{\beta}{2}}}. \quad (\text{D.8})$$

Thus, if  $n^{-\frac{\beta}{2}} \leq n^{-\frac{\delta}{2}}\varepsilon^{\frac{1}{2}}$  for all  $\varepsilon > 0$ , then, necessarily,  $\delta \leq \frac{\beta}{2}$ . Conversely, if  $\delta \leq \frac{\beta}{2}$ , then, according to (D.8), by taking  $\varepsilon$  small enough, it holds that  $n^{-\frac{\beta}{2}} \leq n^{-\frac{\delta}{2}}\varepsilon^{\frac{1}{2}}$ . Applying a similar reasoning to (D.7) shows that  $n^{-\frac{\delta}{2}}\varepsilon^{\frac{1}{2}} \leq n^{-\frac{\beta}{2}}$  if and only if  $\delta \geq \frac{\beta}{2}$ , provided that  $\varepsilon$  is taken small enough.

Let  $\widehat{C} > 0$  denote the constant defined in (D.4), set  $r > 1$ , and  $\ell$  as in (4.10), and take  $\varepsilon > 0$  small enough. Following Proposition 4.10, by setting  $n$  as in (D.6) if (4.4) holds with  $\delta \leq \frac{\beta}{2}$ , and as in (D.6) otherwise, it holds

$$\text{Cost}_{\text{adNSA}} \leq \bar{C} \begin{cases} \varepsilon^{-\frac{p_\star + \delta}{p_\star \delta} - \frac{1}{1+\theta}} & \text{if (4.4) holds with } \delta \leq \frac{\beta}{2}, \\ \varepsilon^{-\frac{2(\delta + p_\star)}{\beta p_\star} - \frac{1}{1+\theta}} & \text{if (4.4) holds with } \delta \geq \frac{\beta}{2}, \\ \varepsilon^{-\frac{2}{\beta} - \frac{1}{1+\theta}} |\ln \varepsilon|^{\frac{1}{2}} & \text{if (4.5) or (4.6) holds,} \end{cases}$$

where

$$\bar{C} = \widehat{C}(C')^{\frac{1}{1+\theta}} h_0^{\frac{1}{1+\theta}} h_1^{-1} \begin{cases} (C^{\frac{2}{\delta}} + 1)^{\frac{p_\star + \delta}{p_\star}} & \text{if (4.4) holds with } \delta \leq \frac{\beta}{2}, \\ (C^{\frac{2}{\beta}} + 1)^{\frac{p_\star + \delta}{p_\star}} & \text{if (4.4) holds with } \delta \geq \frac{\beta}{2}, \\ (1 + C^{\frac{2}{\beta}}) & \\ \times (\ln(1 + C^{\frac{2}{\beta}}) + \frac{2}{\beta} + 1 + \frac{1 + |\ln(Ch_0^{-1})|}{(1+\theta)\ln M}) & \text{if (4.5) or (4.6) holds.} \end{cases} \quad (\text{D.9})$$

The best complexity rates are then easily obtained by optimizing the upper bounds with respect to  $\theta$ ,  $\delta$  and  $\beta$ . □

*Proof of Proposition 5.6.* Lemma 2.5(i) guarantees that the bias error of adMLSA satisfies

$$|\xi_\star^{h_{L+\lceil \theta L \rceil}} - \xi_\star^0| \leq Ch_{L+\lceil \theta L \rceil} \leq \frac{Ch_0}{M^{L(1+\theta)}},$$

where  $C > 0$  is the constant described in (2.14). Thus, provided that  $h_0 > C^{-1}\varepsilon$ , it suffices to take  $L$  minimal so that

$$\frac{Ch_0}{M^{L(1+\theta)}} \leq \varepsilon, \quad \text{i.e.} \quad L = \left\lceil \frac{\ln(Ch_0\varepsilon^{-1})}{(1+\theta)\ln M} \right\rceil \geq 1.$$

□

*Proof of Proposition 5.8.* Recalling that no adaptation applies at level 0 (c.f. Remark 3.3(ii)), the average complexity of the adaptive multilevel SA algorithm satisfies

$$\begin{aligned} \text{Cost}_{\text{adMLSA}} &\leq c_0 \left( \sum_{n=1}^{N_0} (1 + h_0^{-1}) + \sum_{\ell=1}^L \sum_{n=1}^{N_\ell} (1 + \mathbb{E}[h_{\ell+\tilde{\eta}_n}^{-1} \vee h_{\ell-1+\tilde{\eta}_n}^{-1}]) \right) \\ &\leq 3c_0 \left( N_0 h_0^{-1} + \sum_{\ell=1}^L \sum_{n=1}^{N_\ell} (\mathbb{E}[h_{\ell+\tilde{\eta}_n}^{-1}] + \mathbb{E}[h_{\ell-1+\tilde{\eta}_n}^{-1}]) \right) \end{aligned}$$

Using a similar reasoning to the proof of Proposition 4.10, for  $\ell \in \mathbb{N}^*$ ,  $N_\ell \in \mathbb{N}^*$ , and  $j \in \{\ell-1, \ell\}$ ,

$$\sum_{n=1}^{N_\ell} \mathbb{E}[h_{j+\tilde{\eta}_n}^{-1}] \leq (C' \vee 1) \begin{cases} N_\ell^{1+\frac{\delta}{p^*}} h_\ell^{-1} & \text{if (2.23) holds,} \\ N_\ell(L^{\frac{1}{2}} + \ln^{\frac{1}{2}} N_\ell) h_\ell^{-1} & \text{if (2.24) or (3.9) holds,} \end{cases}$$

where  $C' > 0$  is the constant described in (D.4). Note that the previous inequality takes into account the case where  $\ell = 1$  and  $j = 0$ , for which no adaptation occurs. Therefore

$$\text{Cost}_{\text{adMLSA}} \leq C \begin{cases} \sum_{\ell=0}^L N_\ell^{1+\frac{\delta}{p^*}} h_\ell^{-1} & \text{if (2.23) holds,} \\ \sum_{\ell=0}^L N_\ell(L^{\frac{1}{2}} + \ln^{\frac{1}{2}} N_\ell) h_\ell^{-1} & \text{if (2.24) or (3.9) holds.} \end{cases}$$

where

$$C = 6c_0(C' \vee 1). \quad (\text{D.10})$$

□

*Proof of Theorem 5.10.* Define

$$\phi_L(N_\ell) := \begin{cases} N_\ell^{1+\frac{\delta}{p^*}} & \text{if (4.4) holds,} \\ N_\ell(L^{\frac{1}{2}} + \ln^{\frac{1}{2}} N_\ell) & \text{if (4.5) or (4.6) holds.} \end{cases}$$

Following [26, 14], a heuristic proxy for the upper estimate in (5.5) is  $\bar{C} \sum_{\ell=1}^L \gamma_{N_\ell} \epsilon(h_\ell)^{1+\theta}$ , where  $\bar{C}$  is the constant given in (C.13). To determine the optimal number of iterations  $N_0, \dots, N_L \in \mathbb{N}^*$ , we relax the problem to seeking the optimal values of  $N_0, \dots, N_L \in \mathbb{R}$  that minimize the complexity while constraining the proxy  $L^2(\mathbb{P})$ -error control to  $\epsilon^2$ :

$$\begin{cases} \min_{N_0, \dots, N_L \in \mathbb{R}} & C \sum_{\ell=0}^L \phi_L(N_\ell) h_\ell^{-1}, \\ \text{s.t.} & \bar{C} \sum_{\ell=0}^L \gamma_{N_\ell} \epsilon(h_\ell)^{1+\theta} = \epsilon^2. \end{cases}$$

where  $C > 0$  is the constant given in (D.10). We can check easily that the obtained solutions are positive and that, through them, the remaining terms in the upper estimate (5.5) are of order  $\epsilon^2$ .

(i) *Step 1. Iteration amounts.*

The problem to solve is

$$\begin{cases} \min_{N_0, \dots, N_L \in \mathbb{R}} & C \sum_{\ell=0}^L N_\ell^{\frac{p^*+\delta}{p^*}} h_\ell^{-1}, \\ \text{s.t.} & \bar{C} \gamma_1 \sum_{\ell=0}^L N_\ell^{-\beta} h_\ell^{\frac{(1+\theta)p^*}{2(p+1)}} = \epsilon^2. \end{cases}$$

The associated Lagrangian can be written as

$$\mathfrak{L} = C \sum_{\ell=0}^L \frac{N_\ell^{\frac{p^*+\delta}{p^*}}}{h_\ell} + \lambda \left( \bar{C} \gamma_1 \sum_{\ell=0}^L \frac{h_\ell^{\frac{(1+\theta)p^*}{2(p+1)}}}{N_\ell^\beta} - \epsilon^2 \right),$$

where  $\lambda \in \mathbb{R}$  is the Lagrangian multiplier. Optimizing with respect to  $N_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$ , leads to

$$N_\ell = \lambda^{\frac{p_\star}{(1+\beta)p_\star+\delta}} \left( \frac{\bar{C}\gamma_1\beta p_\star}{C(p_\star+\delta)} \right)^{\frac{p_\star}{(1+\beta)p_\star+\delta}} h_\ell^{\frac{(3+\theta)p_\star+2}{2(p_\star+1)((1+\beta)p_\star+\delta)}p_\star}. \quad (\text{D.11})$$

Injecting this value back into the Lagrangian gives

$$\begin{aligned} \mathfrak{L} &= C^{\frac{\beta p_\star}{(1+\beta)p_\star+\delta}} (\bar{C}\gamma_1)^{\frac{p_\star+\delta}{(1+\beta)p_\star+\delta}} \\ &\times \left( \left( \frac{\beta p_\star}{(1+\beta)p_\star+\delta} \right)^{\frac{p_\star+\delta}{(1+\beta)p_\star+\delta}} + \left( \frac{\beta p_\star}{(1+\beta)p_\star+\delta} \right)^{-\frac{\beta p_\star}{(1+\beta)p_\star+\delta}} \right) \\ &\times \lambda^{\frac{p_\star+\delta}{(1+\beta)p_\star+\delta}} \sum_{\ell=0}^L h_\ell^{-\frac{(2\beta-(1+\theta))p_\star+(2\beta-(1+\theta)\delta)}{2(p_\star+1)((1+\beta)p_\star+\delta)}p_\star} - \lambda \varepsilon^2. \end{aligned}$$

The optimal Lagrangian multiplier  $\lambda$  satisfies

$$\begin{aligned} \lambda^{\frac{p_\star}{(1+\beta)p_\star+\delta}} &= \left\{ C^{\frac{\beta p_\star}{(1+\beta)p_\star+\delta}} (\bar{C}\gamma_1)^{\frac{p_\star+\delta}{(1+\beta)p_\star+\delta}} \left( \frac{p_\star+\delta}{(1+\beta)p_\star+\delta} \right) \right. \\ &\times \left. \left( \left( \frac{\beta p_\star}{(1+\beta)p_\star+\delta} \right)^{\frac{p_\star+\delta}{(1+\beta)p_\star+\delta}} + \left( \frac{\beta p_\star}{(1+\beta)p_\star+\delta} \right)^{-\frac{\beta p_\star}{(1+\beta)p_\star+\delta}} \right) \right\}^{\frac{1}{\beta}} \\ &\times \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell=0}^L h_\ell^{-\frac{(2\beta-(1+\theta))p_\star+(2\beta-(1+\theta)\delta)}{2(p_\star+1)((1+\beta)p_\star+\delta)}p_\star} \right)^{\frac{1}{\beta}}. \end{aligned}$$

The optimal iteration amounts  $N_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$  are then obtained by injecting the previous value of the Lagrangian multiplier into (D.11):

$$N_\ell = \widehat{C}_1^{\beta,\delta} \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{(2\beta-(1+\theta))p_\star+(2\beta-(1+\theta)\delta)}{2(p_\star+1)((1+\beta)p_\star+\delta)}p_\star} \right)^{\frac{1}{\beta}} h_\ell^{\frac{(3+\theta)p_\star+2}{2(p_\star+1)((1+\beta)p_\star+\delta)}p_\star}, \quad (\text{D.12})$$

where

$$\begin{aligned} \widehat{C}_1^{\beta,\delta} &= C^{\frac{\beta p_\star}{(1+\beta)p_\star+\delta}} (\bar{C}\gamma_1)^{\frac{(p_\star+\delta)p_\star}{(1+\beta)p_\star+\delta}} \\ &\times \left( \left( \frac{\beta p_\star}{p_\star+\delta} \right)^{\frac{p_\star+\delta}{(1+\beta)p_\star+\delta}} + \left( \frac{\beta p_\star}{p_\star+\delta} \right)^{-\frac{\beta p_\star}{(1+\beta)p_\star+\delta}} \right)^{\frac{1}{\beta}} \left( \frac{p_\star+\delta}{(1+\beta)p_\star+\delta} \right)^{\frac{1}{\beta}}. \end{aligned}$$

It is easy to see that  $N_\ell > 0$ ,  $\ell \in \llbracket 0, L \rrbracket$ . Since we expect such values to be in  $\mathbb{N}^*$ , it suffices to take the ceiling thereof.

*Step 2. Complexity.*

Denote the rational fractions

$$A_{p_\star} := \frac{(3+\theta)p_\star+2}{2(p_\star+1)((1+\beta)p_\star+\delta)}p_\star, \quad (\text{D.13})$$

$$B_{p_\star} := \frac{(2\beta-(1+\theta))p_\star+(2\beta-(1+\theta)\delta)}{2(p_\star+1)((1+\beta)p_\star+\delta)}p_\star. \quad (\text{D.14})$$

Observe that

$$\left( 1 + \frac{\delta}{p_\star} \right) A_{p_\star} - 1 = -B_{p_\star}.$$

Thus, recalling that the  $N_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$ , are set as the ceilings of the value indicated in (D.12), by considering Proposition 5.8,

$$\text{Cost}_{\text{adMLSA}} \leq \bar{C}_1^{\beta,\delta} \left( \sum_{\ell=0}^L h_\ell^{-1} + \varepsilon^{-\frac{2(p_\star+\delta)}{\beta p_\star}} \left( \sum_{\ell=0}^L h_\ell^{-B_{p_\star}} \right)^{\frac{(1+\beta)p_\star+\delta}{\beta p_\star}} \right). \quad (\text{D.15})$$

where

$$\bar{C}_1^{\beta,\delta} = 2^{\frac{\delta}{p_\star}} \bar{C} (1 \vee (\hat{C}_1^{\beta,\delta})^{\frac{p_\star+\delta}{p_\star}}).$$

In view of (5.7),

$$\sum_{\ell=0}^L h_\ell^{-1} \leq \frac{h_L^{-1}}{1-M^{-1}} \leq \frac{(Ch_0)^{\frac{1}{1+\theta}} h_1^{-1}}{1-M^{-1}} \varepsilon^{-\frac{1}{1+\theta}}. \quad (\text{D.16})$$

Moreover,

$$\sum_{\ell=0}^L h_\ell^{-B_{p_\star}} \leq \begin{cases} \frac{h_0^{-B_{p_\star}}}{1-M^{B_{p_\star}}} & \text{if } B_{p_\star} < 0, \\ L+1 & \text{if } B_{p_\star} = 0, \\ \frac{h_L^{-B_{p_\star}}}{1-M^{-B_{p_\star}}} & \text{if } B_{p_\star} > 0. \end{cases}$$

Take  $\varepsilon$  small enough so that  $|\ln \varepsilon| \geq 1$ . Note that, following (5.7),

$$L+1 \leq \left( 2 + \frac{|\ln(Ch_0)| + 1}{(1+\theta) \ln M} \right) |\ln \varepsilon|. \quad (\text{D.17})$$

Besides, if  $B_{p_\star} > 0$ , then, it ensues from (5.7) that

$$h_L^{-B_{p_\star}} \leq (Ch_0)^{\frac{B_{p_\star}}{1+\theta}} h_1^{-B_{p_\star}} \varepsilon^{-\frac{B_{p_\star}}{1+\theta}}.$$

Thus

$$\sum_{\ell=0}^L h_\ell^{-B_{p_\star}} \leq \tilde{C}_1^{\beta,\delta,\theta} \times \begin{cases} 1 & \text{if } B_{p_\star} < 0, \\ |\ln \varepsilon| & \text{if } B_{p_\star} = 0, \\ \varepsilon^{-\frac{B_{p_\star}}{1+\theta}} & \text{if } B_{p_\star} > 0, \end{cases} \quad (\text{D.18})$$

where

$$\tilde{C}_1^{\beta,\delta,\theta} = \begin{cases} \frac{h_0^{-B_{p_\star}}}{1-M^{B_{p_\star}}} & \text{if } B_{p_\star} < 0, \\ 2 + \frac{|\ln(C'h_0)| + 1}{(1+\theta) \ln M} & \text{if } B_{p_\star} = 0, \\ \frac{(Ch_0)^{\frac{B_{p_\star}}{1+\theta}} h_1^{-B_{p_\star}}}{1-M^{-B_{p_\star}}} & \text{if } B_{p_\star} > 0. \end{cases}$$

Therefore, via (D.15), (D.16) and (D.18),

$$\text{Cost}_{\text{adMLSA}} \leq \check{C}_1 \varepsilon^{-\frac{1}{1+\theta}} + \check{C}_1 \begin{cases} \varepsilon^{-\frac{2(p_\star+\delta)}{\beta p_\star}} & \text{if } B_{p_\star} < 0, \\ \varepsilon^{-\frac{2(p_\star+\delta)}{\beta p_\star}} |\ln \varepsilon|^{\frac{(1+\beta)p_\star+\delta}{\beta p_\star}} & \text{if } B_{p_\star} = 0, \\ \varepsilon^{-\frac{2(p_\star+\delta)}{\beta p_\star} - \frac{(2\beta-(1+\theta)p_\star+2\beta-(1+\theta)\delta)}{(1+\theta)\beta(p_\star+1)}} & \text{if } B_{p_\star} > 0, \end{cases}$$

where

$$\check{C}_1 = \sup_{\beta,\delta,\theta} \check{C}_1^{\beta,\delta,\theta}, \quad \text{with} \quad \check{C}_1^{\beta,\delta,\theta} = \bar{C}_1^{\beta,\delta} \left( \frac{(Ch_0)^{\frac{1}{1+\theta}} h_1^{-1}}{1-M^{-1}} \vee (\tilde{C}_1^{\beta,\delta,\theta})^{\frac{p_\star+\delta}{p_\star}} \right). \quad (\text{D.19})$$

Hence

$$\text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \begin{cases} \varepsilon^{-\frac{2(p_\star+\delta)}{\beta p_\star}} & \text{if } B_{p_\star} < 0, \\ \varepsilon^{-\frac{2(p_\star+\delta)}{\beta p_\star}} |\ln \varepsilon|^{\frac{(1+\beta)p_\star+\delta}{\beta p_\star}} & \text{if } B_{p_\star} = 0, \\ \varepsilon^{-\frac{2(p_\star+\delta)}{\beta p_\star} - \frac{(2\beta-(1+\theta)p_\star+2\beta-(1+\theta)\delta)}{(1+\theta)\beta(p_\star+1)}} & \text{if } B_{p_\star} > 0. \end{cases} \quad (\text{D.20})$$

*Step 3. Best complexity.*

*Step 3.1. Case where  $B_{p_\star} < 0$ .*

If  $B_{p_\star} < 0$ , then, by the definition (D.14),

$$(2\beta - (1+\theta))p_\star + (2\beta - (1+\theta)\delta) < 0, \quad \text{i.e.} \quad \theta > \frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta}.$$

Given that  $\theta \in (0, \frac{p_\star - 2}{p_\star + 2}]$ , then, necessarily,

$$\frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta} < \frac{p_\star - 2}{p_\star + 2}, \quad \text{i.e.} \quad \beta < \frac{p_\star(p_\star + \delta)}{(p_\star + 1)(p_\star + 2)} < 1. \quad (\text{D.21})$$

Reconsidering (D.20), by using (D.21), and letting  $\theta = \frac{p_\star - 2}{p_\star + 2}$  and  $\beta \uparrow \frac{p_\star(p_\star + \delta)}{(p_\star + 1)(p_\star + 2)}$ ,

$$\inf_{\beta, \theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-2 - \frac{6}{p_\star} - \frac{4}{p_\star^2} - \iota},$$

for any  $\iota > 0$ . This bound is independent of the choice of  $\delta \in (0, 1]$ .

*Step 3.2. Case where  $B_{p_\star} = 0$ .*

If  $B_{p_\star} = 0$ , then, by the definition (D.14),

$$(2\beta - (1 + \theta))p_\star + (2\beta - (1 + \theta)\delta) = 0, \quad \text{i.e.} \quad \theta = \frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta}.$$

In particular,

$$\frac{1}{1 + \theta} = \frac{p_\star + \delta}{2\beta(p_\star + 1)}. \quad (\text{D.22})$$

Since  $\theta \in (0, \frac{p_\star - 2}{p_\star + 2}]$ , then, necessarily,

$$\frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta} \leq \frac{p_\star - 2}{p_\star + 2}, \quad \text{i.e.} \quad \beta \leq \frac{p_\star(p_\star + \delta)}{(p_\star + 1)(p_\star + 2)} < 1. \quad (\text{D.23})$$

Coming back to (D.20), via (D.22) and (D.23), by letting  $\beta = \frac{p_\star(p_\star + \delta)}{(p_\star + 1)(p_\star + 2)}$ ,

$$\inf_{\beta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-2 - \frac{6}{p_\star} - \frac{4}{p_\star^2}} |\ln \varepsilon|^{\frac{2p_\star^3 + (3+2\delta)p_\star^2 + (2+3\delta)p_\star + 2\delta}{\delta p_\star^2 + \delta^2 p_\star}}.$$

Note in particular that, for such choice of  $\beta$ ,

$$\theta = \frac{p_\star - 2}{p_\star + 2}.$$

This upper bound is minimal for  $\delta = 1$ , in which case,

$$\inf_{\beta, \delta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-2 - \frac{6}{p_\star} - \frac{4}{p_\star^2}} |\ln \varepsilon|^{\frac{2p_\star^2 + 3p_\star + 2}{p_\star}}.$$

*Step 3.3. Case where  $B_{p_\star} > 0$ .*

If  $B_{p_\star} > 0$ , then, by the definition (D.14),

$$(2\beta - (1 + \theta))p_\star + (2\beta - (1 + \theta)\delta) > 0, \quad \text{i.e.} \quad \theta < \frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta}.$$

Note that, following (D.14) and (D.20),

$$\text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-\frac{2(p_\star + \delta)}{\beta p_\star} - \frac{(2p_\star + \delta)((2\beta - (1 + \theta))p_\star + (2\beta - (1 + \theta)\delta))}{2(1 + \theta)(p_\star + 1)((1 + \beta)p_\star + \delta)}}. \quad (\text{D.24})$$

First, assume

$$\frac{p_\star - 2}{p_\star + 2} < \frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta}, \quad \text{i.e.} \quad \beta > \frac{p_\star(p_\star + \delta)}{(p_\star + 1)(p_\star + 2)}.$$

Per (D.24), taking  $\theta = \frac{p_\star - 2}{p_\star + 2}$ ,

$$\inf_{\theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-\frac{2(p_\star + \delta)}{\beta p_\star} - \frac{(2p_\star + \delta)((p_\star + 1)(p_\star + 2)\beta - p_\star(p_\star + \delta))}{2p_\star(p_\star + 1)((1 + \beta)p_\star + \delta)}}.$$

Thus, taking  $\beta = 1$ ,

$$\inf_{\beta, \theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-\frac{2(p_\star + \delta)}{p_\star} - \frac{(3 - \delta)p_\star + 2}{2p_\star(p_\star + 1)}}.$$

Hence, letting  $\delta \downarrow 0$ ,

$$\inf_{\beta, \delta, \theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-2 - \frac{3p_\star + 2}{2p_\star(p_\star + 1)} - \iota},$$

for any  $\iota > 0$ .

Assume now that

$$\frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta} \leq \frac{p_\star - 2}{p_\star + 2}, \quad \text{i.e.} \quad \beta \leq \frac{p_\star(p_\star + \delta)}{(p_\star + 1)(p_\star + 2)} < 1.$$

Via (D.24), letting  $\theta \uparrow \frac{(2\beta - 1)p_\star + 2\beta - \delta}{p_\star + \delta}$ ,

$$\inf_{\theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-\frac{2(p_\star + \delta)}{\beta p_\star} - \iota},$$

for any  $\iota > 0$ . Hence, setting  $\beta = \frac{p_\star(p_\star + \delta)}{(p_\star + 1)(p_\star + 2)}$ ,

$$\inf_{\beta, \theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-2 - \frac{6}{p_\star} - \frac{4}{p_\star^2} - \iota},$$

for any  $\iota > 0$ , independently from the choice of  $\delta \in (0, 1]$ .

*Step 3.4. Conclusion.*

All in all, the best complexity is attained when  $\beta = 1$ ,  $\delta \downarrow 0$  and  $\theta = \frac{p_\star - 2}{p_\star + 2}$ . The ensuing complexity satisfies

$$\boxed{\inf_{\beta, \delta, \theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_1 \varepsilon^{-2 - \frac{3p_\star + 2}{2p_\star(p_\star + 1)} - \iota}.}$$

(ii) *Step 0. Preliminaries.*

Take  $\varepsilon$  sufficiently small to ensure that  $|\ln \varepsilon| \geq 1$ . Note that, consequently from (5.7),

$$L \leq \tilde{C}_1 |\ln \varepsilon|, \quad \text{where} \quad \tilde{C}_1 = 1 + \frac{1 + |\ln(C h_0)|}{(1 + \theta) \ln M}. \quad (\text{D.25})$$

We postulate that, for some  $\tilde{C}_2 > 0$ ,

$$\ln N_\ell \leq \tilde{C}_2 |\ln \varepsilon|, \quad \ell \in \llbracket 0, L \rrbracket, \quad (\text{D.26})$$

a claim that we check a posteriori by explicitating  $\tilde{C}_2$ . Then, one has

$$C \sum_{\ell=0}^L \frac{N_\ell}{h_\ell} (L^{\frac{1}{2}} + \ln^{\frac{1}{2}} N_\ell) \leq C \tilde{C} |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L \frac{N_\ell}{h_\ell}, \quad \text{where} \quad \tilde{C} = (\tilde{C}_1)^{\frac{1}{2}} + (\tilde{C}_2)^{\frac{1}{2}}. \quad (\text{D.27})$$

*Step 1. Iteration amounts.*

We now solve the proxy problem

$$\begin{cases} \min_{N_0, \dots, N_L \in \mathbb{R}} & C \tilde{C} |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L N_\ell h_\ell^{-1}, \\ \text{s.t.} & \bar{C} \gamma_1 \sum_{\ell=0}^L N_\ell^{-\beta} h_\ell^{\frac{1+\theta}{2}} |\ln h_\ell|^{\frac{1+\theta}{2}} = \varepsilon^2. \end{cases}$$

The associated Lagrangian is

$$\mathfrak{L} = C\tilde{C} |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L \frac{N_\ell}{h_\ell} + \left( \bar{C}\gamma_1 \sum_{\ell=0}^L \frac{h_\ell^{\frac{1+\theta}{2}} |\ln h_\ell|^{\frac{1+\theta}{2}}}{N_\ell^\beta} - \varepsilon^2 \right),$$

where  $\lambda \in \mathbb{R}$  is the Lagrangian multiplier. The optimal  $N_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$  satisfy

$$N_\ell = \left( \frac{\bar{C}\gamma_1 \beta}{C\tilde{C}} \right)^{\frac{1}{1+\beta}} \lambda^{\frac{1}{1+\beta}} |\ln \varepsilon|^{-\frac{1}{2(1+\beta)}} h_\ell^{\frac{3+\theta}{2(1+\beta)}} |\ln h_\ell|^{\frac{1+\theta}{2(1+\beta)}}. \quad (\text{D.28})$$

Coming back to the Lagrangian, using the previous value for the  $N_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$ , one has

$$\begin{aligned} \mathfrak{L} &= (C\tilde{C})^{\frac{\beta}{1+\beta}} (\bar{C}\gamma_1)^{\frac{1}{1+\beta}} (\beta^{\frac{1}{\beta+1}} + \beta^{-\frac{\beta}{1+\beta}}) \\ &\quad \times \lambda^{\frac{1}{1+\beta}} |\ln \varepsilon|^{\frac{\beta}{2(1+\beta)}} \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} |\ln h_\ell|^{\frac{1+\theta}{2(1+\beta)}} - \lambda \varepsilon^2. \end{aligned}$$

The optimal Lagrangian multiplier verifies

$$\begin{aligned} \lambda^{\frac{1}{1+\beta}} &= \left( \frac{(C\tilde{C})^{\frac{\beta}{1+\beta}} (\bar{C}\gamma_1)^{\frac{1}{1+\beta}} (\beta^{\frac{1}{\beta+1}} + \beta^{-\frac{\beta}{1+\beta}})}{1 + \beta} \right)^{\frac{1}{\beta}} \\ &\quad \times \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1}{2(1+\beta)}} \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} |\ln h_\ell|^{\frac{1+\theta}{2(1+\beta)}} \right)^{\frac{1}{\beta}}. \end{aligned}$$

Injecting this value back into (D.28) yields

$$N_\ell = (\bar{C}\gamma_1)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} |\ln h_{\ell'}|^{\frac{1+\theta}{2(1+\beta)}} \right)^{\frac{1}{\beta}} h_\ell^{\frac{3+\theta}{2(1+\beta)}} |\ln h_\ell|^{\frac{1+\theta}{2(1+\beta)}}. \quad (\text{D.29})$$

Recalling that  $h_0 = \frac{1}{K}$ , note that

$$|\ln h_\ell| \leq (\ln K + \ln M)L. \quad (\text{D.30})$$

Besides,

$$\sum_{\ell=0}^L h_\ell^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} \leq \begin{cases} \frac{h_0^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}}}{1 - M^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}}} & \text{if } 2\beta < 1 + \theta, \\ L + 1 & \text{if } 2\beta = 1 + \theta, \\ \frac{h_L^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}}}{1 - M^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}}} & \text{if } 2\beta > 1 + \theta. \end{cases} \quad (\text{D.31})$$

Hence

$$\begin{aligned} &\ln \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} \right) \\ &\leq \begin{cases} -\ln \left( 1 - M^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} \right) & \text{if } 2\beta < 1 + \theta, \\ \ln(L + 1) & \text{if } 2\beta = 1 + \theta, \\ \frac{2\beta-(1+\theta)}{2(1+\beta)} (\ln K + L \ln M) - \ln \left( 1 - M^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} \right) & \text{if } 2\beta > 1 + \theta, \end{cases} \\ &\leq C'_2 L, \end{aligned} \quad (\text{D.32})$$

where

$$C'_2 = \left( \left( \frac{2\beta - (1+\theta)}{2(1+\beta)} \right)^+ \ln(MK) - \ln(1 - M^{-|\frac{2\beta - (1+\theta)}{2(1+\beta)}|}) \mathbf{1}_{\{2\beta \neq 1+\theta\}} \right) \vee 1. \quad (\text{D.33})$$

Hence, considering (D.25), the inequality (D.26) holds with

$$\tilde{C}_2 = \frac{2 + |\ln(\bar{C}\gamma_1)|}{\beta} + \tilde{C}_1 \left( \frac{3+\theta}{2(1+\beta)} \ln(MK) + \frac{1+\theta}{2\beta} \left( 1 + |\ln(\ln(MK))| \right) + \frac{C'_2}{\beta} \right).$$

*Step 2. Complexity.*

We set each  $N_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$ , as the ceiling of the value in (D.29) in order to retrieve iteration amounts in  $\mathbb{N}^*$ . Via Proposition 5.8, (D.27) and (D.30),

$$\begin{aligned} \text{Cost}_{\text{adMLSA}} &\leq C\tilde{C} |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L \frac{N_\ell}{h_\ell} \\ &\leq \hat{C}_2^{\beta, \theta} \left( |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L h_\ell^{-1} + \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta - (1+\theta)}{2(1+\beta)}} \right)^{\frac{1+\beta}{\beta}} \right), \end{aligned} \quad (\text{D.34})$$

where

$$\hat{C}_2^{\beta, \theta} = C\tilde{C} \left( 1 \vee (\bar{C}\gamma_1)^{\frac{1}{\beta}} (\tilde{C}_1 \ln(KM))^{\frac{1+\theta}{2\beta}} \right).$$

If  $2\beta > 1 + \theta$ , it follows from (5.7) that

$$h_L^{-\frac{2\beta - (1+\theta)}{2(1+\beta)}} \leq (Ch_0)^{\frac{2\beta - (1+\theta)}{2(1+\theta)(1+\beta)}} h_1^{-\frac{2\beta - (1+\theta)}{2(1+\theta)(1+\beta)}} \varepsilon^{-\frac{2\beta - (1+\theta)}{2(1+\theta)(1+\beta)}}.$$

Hence, via (D.31) and (D.17),

$$\sum_{\ell=0}^L h_\ell^{-\frac{2\beta - (1+\theta)}{2(1+\beta)}} \leq \tilde{C}_2^{\beta, \theta} \times \begin{cases} 1 & \text{if } 2\beta < 1 + \theta, \\ |\ln \varepsilon| & \text{if } 2\beta = 1 + \theta, \\ \varepsilon^{-\frac{2\beta - (1+\theta)}{2(1+\theta)(1+\beta)}} & \text{if } 2\beta > 1 + \theta, \end{cases} \quad (\text{D.35})$$

where

$$\tilde{C}_2^{\beta, \theta} = \begin{cases} (1 - M^{\frac{2\beta - (1+\theta)}{2(1+\beta)}})^{-1} h_0^{-\frac{2\beta - (1+\theta)}{2(1+\beta)}} & \text{if } 2\beta < 1 + \theta, \\ 2 + \frac{|\ln(Ch_0)| + 1}{(1+\theta) \ln M} & \text{if } 2\beta = 1 + \theta, \\ (1 - M^{-\frac{2\beta - (1+\theta)}{2(1+\beta)}})^{-1} (Ch_0)^{\frac{2\beta - (1+\theta)}{2(1+\theta)(1+\beta)}} h_1^{-\frac{2\beta - (1+\theta)}{2(1+\theta)(1+\beta)}} & \text{if } 2\beta > 1 + \theta. \end{cases} \quad (\text{D.36})$$

Thus, by gathering the results of (D.16) and (D.35) and coming back to (D.34),

$$\text{Cost}_{\text{adMLSA}} \leq \check{C}_2 \varepsilon^{-\frac{1}{1+\theta}} |\ln \varepsilon|^{\frac{1}{2}} + \check{C}_2 \begin{cases} \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} & \text{if } 2\beta < 1 + \theta, \\ \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{3(1+\beta)+\theta}{2\beta}} & \text{if } 2\beta = 1 + \theta, \\ \varepsilon^{-\frac{3(1+\theta)+2\beta}{2(1+\theta)\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} & \text{if } 2\beta > 1 + \theta, \end{cases}$$

where

$$\check{C}_2 = \sup_{\beta, \theta} \check{C}_2^{\beta, \theta}, \quad \text{with} \quad \check{C}_2^{\beta, \theta} = \hat{C}_2^{\beta, \theta} \left( \frac{(Ch_0)^{\frac{1}{1+\theta}} h_1^{-1}}{1 - M^{-1}} \vee (\tilde{C}_2^{\beta, \theta})^{\frac{1+\beta}{\beta}} \right). \quad (\text{D.37})$$

Thus

$$\text{Cost}_{\text{adMLSA}} \leq 2\check{C}_2 \begin{cases} \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} & \text{if } 2\beta < 1 + \theta, \\ \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{3(1+\beta)+\theta}{2\beta}} & \text{if } 2\beta = 1 + \theta, \\ \varepsilon^{-\frac{3(1+\theta)+2\beta}{2(1+\theta)\beta}} |\ln \varepsilon|^{\frac{1+\theta+\beta}{2\beta}} & \text{if } 2\beta > 1 + \theta. \end{cases} \quad (\text{D.38})$$

*Step 3. Best complexity.*

If  $2\beta < 1 + \theta$ , then, the complexity bound is minimal if  $\theta = 1$  and  $\beta \uparrow 1$ , in which case

$$\inf_{\beta, \theta} \text{Cost}_{\text{adMLSA}} \leq \check{C}_2 \varepsilon^{-2-\iota} |\ln \varepsilon|^{\frac{3}{2}+\iota},$$

for any  $\iota > 0$ .

If  $2\beta = 1 + \theta$ , then, the complexity bound is minimal if  $\beta = \theta = 1$ , in which case

$$\inf_{\beta, \theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_2 \varepsilon^{-2} |\ln \varepsilon|^{\frac{7}{2}}.$$

If  $2\beta > 1 + \theta$ , then, the complexity bound is minimal if  $\beta = 1$  and  $\theta \uparrow 1$ , in which case

$$\inf_{\beta, \theta} \text{Cost}_{\text{adMLSA}} \leq \check{C}_2 \varepsilon^{-2-\iota} |\ln \varepsilon|^{\frac{3}{2}+\iota},$$

for any  $\iota > 0$ .

Overall, the best complexity is attained for  $\beta = \theta = 1$ , and it satisfies

$$\boxed{\inf_{\beta, \theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_2 \varepsilon^{-2} |\ln \varepsilon|^{\frac{7}{2}}.}$$

(iii) *Step 0. Preliminaries.*

We proceed similarly to the previous framework by assuming that  $\varepsilon$  is small enough to ensure that  $|\ln \varepsilon| \geq 1$ , and that (D.26) holds.

*Step 1. Iteration amounts.*

Thus, taking into account (D.25) and (D.27), we look into the proxy problem

$$\begin{cases} \min_{N_0, \dots, N_L \in \mathbb{R}} & C\tilde{C} |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L N_\ell h_\ell^{-1}, \\ \text{s.t.} & \bar{C}\gamma_1 \sum_{\ell=0}^L N_\ell^{-\beta} h_\ell^{\frac{1+\theta}{2}} = \varepsilon^2. \end{cases}$$

The Lagrangian of the previous problem is

$$\mathfrak{L} = C\tilde{C} |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L \frac{N_\ell}{h_\ell} + \lambda \left( \bar{C}\gamma_1 \sum_{\ell=0}^L \frac{h_\ell^{\frac{1+\theta}{2}}}{N_\ell^\beta} - \varepsilon^2 \right),$$

with  $\lambda \in \mathbb{R}$  being the Lagrangian multiplier. The optimal  $N_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$ , are given by

$$N_\ell = \left( \frac{\bar{C}\gamma_1 \beta}{C\tilde{C}} \right)^{\frac{1}{1+\beta}} \lambda^{\frac{1}{1+\beta}} |\ln \varepsilon|^{-\frac{1}{2(1+\beta)}} h_\ell^{\frac{3+\theta}{2(1+\beta)}}.$$

Injecting these values back into the Lagrangian yields

$$\mathfrak{L} = (C\tilde{C})^{\frac{\beta}{1+\beta}} (\bar{C}\gamma_1)^{\frac{1}{1+\beta}} (\beta^{\frac{1}{\beta+1}} + \beta^{-\frac{\beta}{1+\beta}}) \lambda^{\frac{1}{1+\beta}} |\ln \varepsilon|^{\frac{\beta}{2(1+\beta)}} \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} - \lambda \varepsilon^2.$$

The optimal Lagrangian multiplier  $\lambda$  satisfies

$$\lambda^{\frac{1}{1+\beta}} = \left( \frac{(C\tilde{C})^{\frac{\beta}{1+\beta}} (\bar{C}\gamma_1)^{\frac{1}{1+\beta}} (\beta^{\frac{1}{\beta+1}} + \beta^{-\frac{\beta}{1+\beta}})}{1+\beta} \right)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1}{2(1+\beta)}} \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} \right)^{\frac{1}{\beta}}.$$

Thus

$$\boxed{N_\ell = (\bar{C}\gamma_1)^{\frac{1}{\beta}} \varepsilon^{-\frac{2}{\beta}} \left( \sum_{\ell'=0}^L h_{\ell'}^{-\frac{2\beta-(1+\theta)}{2(1+\beta)}} \right)^{\frac{1}{\beta}} h_\ell^{\frac{3+\theta}{2(1+\beta)}}.}$$

Using (D.32) and (D.25), we obtain that (D.26) holds, with

$$\tilde{C}_2 = \frac{2 + |\ln(\bar{C}\gamma_1)|}{\beta} + \tilde{C}_1 \left( \frac{3 + \theta}{2(1 + \beta)} \ln(MK) + \frac{C'_2}{\beta} \right),$$

where  $C'_2 > 0$  is the constant defined in (D.33).

*Step 2. Complexity.*

We set each  $N_\ell$ ,  $\ell \in \llbracket 0, L \rrbracket$ , as the ceiling of the value in (D.29) in order to retrieve iteration amounts in  $\mathbb{N}^*$ . Via Proposition 5.8, (D.27) and (D.30),

$$\begin{aligned} \text{Cost}_{\text{adMLSA}} &\leq C\tilde{C} |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L \frac{N_\ell}{h_\ell} \\ &\leq \hat{C}_3^{\beta, \theta} \left( |\ln \varepsilon|^{\frac{1}{2}} \sum_{\ell=0}^L h_\ell^{-1} + \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1}{2}} \left( \sum_{\ell=0}^L h_\ell^{-\frac{2\beta - (1+\theta)}{2(1+\beta)}} \right)^{\frac{1+\beta}{\beta}} \right), \end{aligned}$$

where

$$\hat{C}_3^{\beta, \theta} = C\tilde{C}(1 \vee (\bar{C}\gamma_1)^{\frac{1}{\beta}}).$$

Using (D.16) and (D.35),

$$\text{Cost}_{\text{adMLSA}} \leq \check{C}_3 \varepsilon^{-\frac{1}{1+\theta}} |\ln \varepsilon|^{\frac{1}{2}} + \check{C}_3 \begin{cases} \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1}{2}} & \text{if } 2\beta < 1 + \theta, \\ \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1+2\beta}{\beta}} & \text{if } 2\beta = 1 + \theta, \\ \varepsilon^{-\frac{3(1+\theta)+2\beta}{2(1+\theta)\beta}} |\ln \varepsilon|^{\frac{1}{2}} & \text{if } 2\beta > 1 + \theta, \end{cases}$$

where

$$\check{C}_3 = \sup_{\beta, \theta} \check{C}_3^{\beta, \theta}, \quad \text{with} \quad \check{C}_3^{\beta, \theta} = \hat{C}_3^{\beta, \theta} \left( \frac{(Ch_0)^{\frac{1}{1+\theta}} h_1^{-1}}{1 - M^{-1}} \vee (\tilde{C}_2^{\beta, \theta})^{\frac{1+\beta}{\beta}} \right), \quad (\text{D.39})$$

with  $\tilde{C}_2^{\beta, \theta} > 0$  being defined in (D.36). Hence

$$\text{Cost}_{\text{adMLSA}} \leq 2\check{C}_3 \begin{cases} \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1}{2}} & \text{if } 2\beta < 1 + \theta, \\ \varepsilon^{-\frac{2}{\beta}} |\ln \varepsilon|^{\frac{1+2\beta}{\beta}} & \text{if } 2\beta = 1 + \theta, \\ \varepsilon^{-\frac{3(1+\theta)+2\beta}{2(1+\theta)\beta}} |\ln \varepsilon|^{\frac{1}{2}} & \text{if } 2\beta > 1 + \theta. \end{cases} \quad (\text{D.40})$$

*Step 3. Best complexity.*

Similarly to the previous framework, the best complexity is attained for  $\beta = \theta = 1$ , and it satisfies

$$\inf_{\beta, \theta} \text{Cost}_{\text{adMLSA}} \leq 2\check{C}_3 \varepsilon^{-2} |\ln \varepsilon|^{\frac{5}{2}}.$$

□

## References

- [1] ABRAMOWITZ, M., STEGUN, I. A. and ROMER, R. H. (1988). Handbook of mathematical functions with formulas, graphs, and mathematical tables. *American Journal of Physics* **56**.
- [2] ACERBI, C. and TASCHE, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance* **26** 1487–1503.

- [3] ALZER, H. (1997). On some inequalities for the gamma and psi functions. *Mathematics of Computation* **66** 373–389.
- [4] BARDOU, O., FRIKHA, N. and PAGÈS, G. (2009). *Recursive computation of value-at-risk and conditional value-at-risk using MC and QMC*. In *Monte Carlo and Quasi-Monte Carlo Methods 2008* 193–208. Springer.
- [5] BARDOU, O., FRIKHA, N. and PAGÈS, G. (2009). Computing VaR and CVaR using stochastic approximation and adaptive unconstrained importance sampling. *Monte Carlo Methods and Applications* **15** 173–210.
- [6] BARDOU, O., FRIKHA, N. and PAGÈS, G. (2016). CVaR hedging using quantization-based stochastic approximation algorithm. *Mathematical Finance* **26** 184–229.
- [7] BARRERA, D., CRÉPEY, S., DIALLO, B., FORT, G., GOBET, E. and STAZHYNSKI, U. (2019). Stochastic approximation schemes for economic capital and risk margin computations. *ESAIM: Proceedings and Surveys* **65** 182–218.
- [8] BASEL COMMITTEE ON BANKING SUPERVISION. (2013). Consultative document: Fundamental Review of the Trading Book: a revised market risk framework. [bis.org/publ/bcbs265.pdf](https://www.bis.org/publ/bcbs265.pdf)
- [9] BEN ALAYA, M., HAJJI, K. and KEBAIER, A. (2023). Adaptive importance sampling for multilevel Monte Carlo Euler method. *Stochastics* **95** 303–327.
- [10] BEN-TAL, A. and TEBoulLE, M. (2007). An old-new concept of convex risk measures: the optimized certainty equivalent. *Mathematical Finance* **17** 449–476.
- [11] BERCU, B., COSTA, M. and GADAT, S. (2021). Stochastic approximation algorithms for superquantiles estimation. *Electronic Journal of Probability* **26** 1–29.
- [12] BROADIE, M., DU, Y. and MOALLEMI, C. C. (2011). Efficient risk estimation via nested sequential simulation. *Management Science* **57** 1172–1194.
- [13] COSTA, M. and GADAT, S. (2021). Non asymptotic controls on a recursive superquantile approximation. *Electronic Journal of Statistics* **15** 4718–4769.
- [14] CRÉPEY, S., FRIKHA, N. and LOUZI, A. (2025). A multilevel stochastic approximation algorithm for value-at-risk and expected shortfall estimation. *Finance and Stochastics* **29** 1015–1074.
- [15] CRÉPEY, S., FRIKHA, N., LOUZI, A. and PAGÈS, G. (2024). Asymptotic error analysis of multilevel stochastic approximations for the value-at-risk and expected shortfall. *Electronic Journal of Probability* **29** 1–56.
- [16] DEREICH, S. (2011). Multilevel Monte Carlo algorithms for Lévy-driven SDEs with Gaussian correction. *The Annals of Applied Probability* **21** 283–311.
- [17] DEREICH, S. and MÜLLER-GRONBACH, T. (2019). General multilevel adaptations for stochastic approximation algorithms of Robbins-Monro and Polyak-Ruppert type. *Numerische Mathematik* **142** 279–328.
- [18] DUFLO, M. (1996). *Algorithmes Stochastiques*. Springer.
- [19] ELFVerson, D., HELLMAN, F. and MÄLQVIST, A. (2016). A multilevel Monte Carlo method for computing failure probabilities. *SIAM/ASA Journal on Uncertainty Quantification* **4** 312–330.

- [20] FATHI, M. and FRIKHA, N. (2013). Transport-entropy inequalities and deviation estimates for stochastic approximation schemes. *Electronic Journal of Probability* **18** 1–36.
- [21] FRIKHA, N. (2016). Multi-level stochastic approximation algorithms *The Annals of Applied Probability* **26** 933–985.
- [22] FRIKHA, N. and MENOZZI, S. (2012). Concentration bounds for stochastic approximations. *Electronic Communications in Probability* **17** 1–15.
- [23] FÖLLMER, H. and SCHIED, A. (2010). *Convex risk measures* In *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Ltd.
- [24] GILES, M. B. (2008). Multilevel Monte Carlo path simulation. *Operations Research* **56** 607–617.
- [25] GILES, M. B. (2024). *MLMC techniques for discontinuous functions*. In *Monte Carlo and Quasi-Monte Carlo Methods 2022* 33–47. Springer International Publishing.
- [26] GILES, M. B. and HAJI-ALI, A.-L. (2019). Multilevel nested simulation for efficient risk estimation. *SIAM/ASA Journal on Uncertainty Quantification* **7** 497–525.
- [27] GILES, M. B., HAJI-ALI, A.-L. and SPENCE, J. (2023). Efficient risk estimation for the credit valuation adjustment. arXiv:2301.05886
- [28] GILES, M. B. and SZPRUCH, L. (2014). Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation. *The Annals of Applied Probability* **24** 1585–1620.
- [29] GIORGI, D., LEMAIRE, V. and PAGÈS, G. (2020). Weak error for nested multilevel Monte Carlo. *Methodology and Computing in Applied Probability* **22** 1325–1348.
- [30] GORDY, M. B. and JUNEJA, S. (2010). Nested simulation in portfolio risk measurement. *Management Science* **56** 1833–1848.
- [31] HAJI-ALI, A.-L., SPENCE, J. and TECKENTRUP, A. L. (2022). Adaptive multilevel Monte Carlo for probabilities. *SIAM Journal on Numerical Analysis* **60** 2125–2149.
- [32] HEINRICH, S. (1998). Monte Carlo complexity of global solution of integral equations. *Journal of Complexity* **14** 151–175.
- [33] HEINRICH, S. (2001). *Multilevel Monte Carlo methods*. In *Large-Scale Scientific Computing* 58–67. Springer.
- [34] HOEL, H., VON SCHWERIN, E., SZEPESSY, A. and TEMPONE, R. (2014). Implementation and analysis of an adaptive multilevel Monte Carlo algorithm. *Monte Carlo Methods and Applications* **20** 1–41.
- [35] RHEE, C.-H. and GLYNN, P. W. (2012). *A new approach to unbiased estimation for SDE's*. In *Proceedings of the Winter Simulation Conference*. WSC '12 1–2. Winter Simulation Conference.
- [36] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*. *Grundlehren der mathematischen Wissenschaften* **293**. Springer.
- [37] ROCKAFELLAR, R. T. and URYASEV, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk* **2** 21–41.