



**HAL**  
open science

## Computation of Expected Shortfall by fast detection of worst scenarios

Bruno Bouchard, Adil Reghai, Benjamin Virrion

► **To cite this version:**

Bruno Bouchard, Adil Reghai, Benjamin Virrion. Computation of Expected Shortfall by fast detection of worst scenarios. *Quantitative Finance*, 2021, 21 (7), pp.1087-1108. 10.1080/14697688.2021.1880618 . hal-02619589

**HAL Id: hal-02619589**

**<https://hal.science/hal-02619589>**

Submitted on 25 May 2020

**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.

# Computation of Expected Shortfall by fast detection of worst scenarios

Bruno Bouchard\*, Adil Reghai†, Benjamin Virrion‡,§

May 25, 2020

## Abstract

We consider a multi-step algorithm for the computation of the historical expected shortfall such as defined by the Basel *Minimum Capital Requirements for Market Risk*. At each step of the algorithm, we use Monte Carlo simulations to reduce the number of historical scenarios that potentially belong to the set of worst scenarios. The number of simulations increases as the number of candidate scenarios is reduced and the distance between them diminishes. For the most naive scheme, we show that the  $\mathbb{L}^p$ -error of the estimator of the Expected Shortfall is bounded by a linear combination of the probabilities of inversion of favorable and unfavorable scenarios at each step, and of the last step Monte Carlo error associated to each scenario. By using concentration inequalities, we then show that, for sub-gamma pricing errors, the probabilities of inversion converge at an exponential rate in the number of simulated paths. We then propose an adaptative version in which the algorithm improves step by step its knowledge on the unknown parameters of interest: mean and variance of the Monte Carlo estimators of the different scenarios. Both schemes can be optimized by using dynamic programming algorithms that can be solved off-line. To our knowledge, these are the first non-asymptotic bounds for such estimators. Our hypotheses are weak enough to allow for the use of estimators for the different scenarios and steps based on the same random variables, which, in practice, reduces considerably the computational effort. First numerical tests are performed.

**Keywords:** Expected Shortfall, ranking and selection, sequential design, Bayesian filter.

## 1 Introduction

The Basel *Minimum Capital Requirements for Market Risk* [4] has brought two main changes in the way that investment banks need to compute their capital requirements. Expected Shortfall (ES) replaces Value at Risk (VaR) as the main risk indicator for the computation of capital requirements. The advantages of ES over VaR have been brought forward in Artzner et al. [2], and Expected Shortfall is now considered by most researchers and practitioners as superior to VaR as a risk measure, because it respects the sub-additivity axiom, see [1, 2, 23]. The second main change is that the number of required daily computations of ES has been multiplied by a factor of up to 90. Where banks used to need to compute one VaR per day, they now need to compute up to three ES per liquidity horizon and risk class, as well as three ES per liquidity horizon for all risk classes combined. The above has triggered several works on the fast computation of ES.

Mathematically, if  $V$  is a random variable modeling the level of loss<sup>1</sup> of a portfolio that will be known at a future time, and  $0 < \alpha < 1$ , the expected shortfall of level  $\alpha \in (0, 1)$  is defined by

$$ES_\alpha := \frac{1}{\alpha} \int_0^\alpha VaR_\gamma(V) d\gamma, \quad (1)$$

---

\*CEREMADE, CNRS, Université Paris Dauphine, PSL University.

†Natixis.

‡Natixis and CEREMADE, CNRS, Université Paris Dauphine, PSL University.

§The authors would like to thank Nicolas Baradel for helping with the code, Rida Mahi and Mathieu Bernardo from the Natixis Quantitative Research Teams for providing the first results and ideas on the Fast Detection Algorithm, and finally William Leduc for providing all the necessary data to obtain the different book parameters.

<sup>1</sup>All over this paper, we measure the performances in terms of losses. A positive number is a loss, a negative number is a gain.

where  $\text{VaR}_\gamma$  is the Value at Risk at level  $\gamma$ , i.e.

$$\text{VaR}_\gamma(V) := \max \{x \in \mathbb{R} : \mathbb{P}[V \geq x] > \gamma\}. \quad (2)$$

Nearly all of the literature concentrates on studying the ES by using parametric, non-parametric or semi-parametric approaches to approximate the distribution of  $V$  based on historical data. See in particular [9, 11, 12, 15, 16, 19, 20, 21, 24, 25, 26, 27]. Another approach consists in using the fact that  $V$  is the risk neutral value of a book, and therefore of the form  $\mathbb{E}[P|S]$  in which  $S$  is a random variable associated to market parameters and  $P$  represents the future (discounted) payoffs of the book. This suggests using a nested Monte Carlo approach : simulate a set of values in the distribution of  $S$  (outer scenarios), and, for each simulation of  $S$ , compute a Monte Carlo estimator of  $\mathbb{E}[P|S]$  by using simulations in the conditional distribution (inner scenarios). This is for instance the approach of [8, 13].

But, as defined in the regulatory document of Basel [4], the expected shortfall is based on  $n_s = 253$  scenarios of market parameters  $s = (s^i)_{i \leq n_s}$  that are generated in a deterministic way. Therefore,  $S$  is just uniformly distributed in the sequence  $s$  and there is no need for simulating outer scenarios. Since  $V$  is defined by a pricing formula  $\mathbb{E}[P|S]$  that is fully described by the value of  $S$ , there is also no room for approximating the law of  $V$  based on historical data, if we are only interested by the requirements of [4]. The only issue is to compute in an efficient way the loss impacts  $(\mu^i)_{i \leq n_s}$  of the book,

$$\mu^i := (\mathbb{E}[P|S = s^i] - \mathbb{E}[P|S = s^0]), \quad i = 1 \dots, n_s,$$

in which  $s^0$  is the current value of the market parameters, and then compute the average over the  $n_w = 6$  worst impacts, say

$$\text{ES} = \frac{1}{n_w} \sum_{i=1}^{n_w} \mu^i, \quad (3)$$

if, for ease of notations, we assume that

$$\mu^1 \geq \mu^2 \geq \dots \geq \mu^{n_s-1} \geq \mu^{n_s}. \quad (4)$$

Methods that are in line with the above have also been studied, in particular in [17, 22] in which the authors define a distance on the space of scenarios induced by the distance between their risk factors. Starting with the original outer-level scenarios (called “prediction points”), they determine “design points” that are included in their convex hull. Inner-level paths are simulated in order to evaluate the portfolio value at the design points. These values are then used to establish a metamodel of the portfolio price with respect to the risk factors, and this metamodel is then used to select among the prediction points those that are most likely to be part of the worst scenarios set. They are then added to the design points, and evaluated by using inner-level simulations, after which the metamodel is updated.

These methods are very smart but neglect one important point for practitioners: the cost of launching a pricer is high, as it typically entails instantiating thousands of objects at initialization, as well as volatility surface calibrations and sometimes even graphical interfaces. Furthermore, these pricers usually do not have the flexibility to add dynamically, at each inner-level pricing, new paths to a given scenario. Therefore, we do not allow ourselves to adapt our strategies at such a level of granularity.

Instead, we will consider strategies that only entail  $L$ -levels of sets of simulations, where  $L$  is typically quite low, so as not to pay too many times the overhead of launching the pricer and/or calibrating the required volatility surfaces. We also do not use any concept of distance between scenarios induced by their risk factors. Although this enables [17] and [22] to obtain better empirical convergence rates, we see at least one problem with this approach: at the scale of a bank, the space of risk factors is both of a very high dimension (a few thousands) and with a very complex geometry (the payoffs of the portfolio’s derivative products are usually non-convex, and path-dependent), so that it is very difficult to establish a model describing the proximity of scenarios in a robust way.

We thus study a relatively simple procedure that also has the advantage of allowing us to establish non-asymptotic bounds on the  $L^p$ -error of our estimator, in the spirit of the simplest ranking by mean procedures, see e.g. [3, 5, 6, 14]. It consists in using a first set of simulated paths to provide a crude estimation of the impact factors  $\mu^i$ . These first estimators are ranked to select the  $q_1 < n_s$  outer-level scenarios with the highest estimated impact values. Then, only the impact values of these  $q_1$  pre-selected scenarios are estimated again by using the previous estimators together with a new set of simulated

paths. Among these new estimators we select the scenarios with the  $q_2 < q_1$  highest estimated impact factors. And so on. After  $L \geq 2$  steps,  $L$  being small in practice, we just keep the mean of the six highest estimated impacts.

The rationale behind this is that a first crude estimation should be sufficient to rule out a large part of the scenarios from the candidates of being in the 6 worst ones, because the corresponding values should be far enough. While the number of candidates reduces, one can expect that the differences between the corresponding impacts diminish as well and that more Monte Carlo simulations are needed to differentiate them. Under an indifference zone hypothesis, similar to the one used in the above mentioned paper, and a sub-gamma distribution assumption, the convergence is exponential in the number of simulations used at the different steps and of order 1/2 in the total number of simulations. See Proposition 2.2 and Corollary 2.3 below.

The optimal number of additional paths that should be used at each step to minimize the strong estimation error, given a maximal computational cost, can be determined by a simple dynamic programming algorithm, that can be solved off-line, see Section 2.4. In theory, this requires the a priori knowledge of the means and covariances of our estimators, which are obviously not known in practice. However, one can easily define a version based on a robust specification of the error. One can also take advantage of the different simulation sets to improve our prior on the true hidden parameters. This leads to a natural adaptative algorithm, see Section 3, for which convergence is also proved, see Proposition 3.3. Estimating the optimal policy associated to this adaptative algorithm is costly but can be done off-line by using a neural network approximation combined with a backward dynamic programming algorithm. We explain how this can be done in Section 3.3 (further details are in Appendix B).

The rest of the paper is organized as follows. Section 2 is dedicated to the most naive deterministic algorithm. In particular, Section 2.5 gives a very easy to use two levels algorithm for the case where the impacts decrease linearly in the scenarios' rank order. The adaptative version of the algorithm is presented in Section 3. Finally, we perform first numerical tests in Section 4.

## 2 Algorithm with a deterministic effort allocation

In this section, we describe the simplest version of the algorithm. It uses a pre-defined deterministic number of simulations. We establish a strong  $L^p$ -error bound and discuss several ways of choosing the optimal strategy for minimizing this error.

### 2.1 The algorithm

From now on, we assume that  $\mathbb{E}[P|S = s^0]$  is known perfectly and set it to 0 for ease of notations. As explained above, the algorithm relies on the idea of selecting progressively the scenarios that will be used for the computation of the Expected Shortfall. Namely, let  $P_{|s} := (P_{|s^i})_{i \leq n_s}$  be a  $n_s$ -dimensional random variable such that each  $P_{|s^i}$  has the law of  $P$  given  $S = s^i$ . We first simulate independent copies  $(P_j^1, \dots, P_j^{n_s})_{j \geq 1}$  of  $P_{|s}$  and compute the Monte Carlo estimators of  $\mathbb{E}[P|S = s^i]$ ,  $i \leq n_s$ :

$$\hat{\mu}_1^i := \frac{1}{N_1} \sum_{j=1}^{N_1} P_j^i \quad \text{for } i \leq n_s,$$

for some  $N_1 \geq 1$ . Among these random variables, we then select the ones that are the most likely to coincide with the worst scenarios  $s^1, \dots, s^{n_w}$ , for some  $1 \leq n_w < n_s$ . To do this, one considers the (random) permutation  $\mathbf{m}_1$  on  $\llbracket 1, n_s \rrbracket$  such that the components of  $(\hat{\mu}_1^{\mathbf{m}_1(i)})_{i \leq n_s}$  are in decreasing order:

$$\begin{cases} \hat{\mu}_1^{\mathbf{m}_1(1)} \geq \hat{\mu}_1^{\mathbf{m}_1(2)} \geq \dots \geq \hat{\mu}_1^{\mathbf{m}_1(n_s)}, \\ \mathbf{m}_1(i) < \mathbf{m}_1(i') \text{ if } \hat{\mu}_1^{\mathbf{m}_1(i)} = \hat{\mu}_1^{\mathbf{m}_1(i')} \text{ for } 1 \leq i < i' \leq n_s, \end{cases}$$

and only keep the indexes  $(\mathbf{m}_1(\ell))_{\ell \leq q_1}$  of the corresponding  $q_1 \geq n_w$  highest values, i.e. the indexes belonging to

$$\mathfrak{J}_1 := \mathfrak{J}_0 \cap \mathbf{m}_1(\llbracket 1, q_1 \rrbracket) \quad \text{in which } \mathfrak{J}_0 := \llbracket 1, n_s \rrbracket.$$

We then iterate the above procedure on the scenarios in  $\mathfrak{J}_1$  and so on. Namely, we fix  $L \geq 1$  different thresholds  $(q_\ell)_{\ell=0,\dots,L-1}$  such that

$$n_w =: q_{L-1} \leq \dots \leq q_0 := n_s. \quad (5)$$

Assuming that  $\mathfrak{J}_{\ell-1}$  is given, for some  $1 \leq \ell - 1 \leq L - 1$ , we compute the estimators<sup>2</sup>

$$\hat{\mu}_\ell^i := \frac{1}{N_\ell} \sum_{j=1}^{N_\ell} P_j^i \quad \text{for } i \leq n_s, \quad (6)$$

for some  $N_\ell \geq N_{\ell-1}$ . If  $\ell \leq L - 1$ , we consider the (random) permutation  $\mathbf{m}_\ell : \llbracket 1, q_{\ell-1} \rrbracket \mapsto \mathfrak{J}_{\ell-1}$  such that the components of  $(\hat{\mu}_\ell^i)_{i \in \mathfrak{J}_{\ell-1}}$  are in decreasing order

$$\begin{cases} \hat{\mu}_\ell^{\mathbf{m}_\ell(1)} \geq \hat{\mu}_\ell^{\mathbf{m}_\ell(2)} \geq \dots \geq \hat{\mu}_\ell^{\mathbf{m}_\ell(q_{\ell-1})}, \\ \mathbf{m}_\ell(i) < \mathbf{m}_\ell(i') \text{ if } \hat{\mu}_\ell^{\mathbf{m}_\ell(i)} = \hat{\mu}_\ell^{\mathbf{m}_\ell(i')} \text{ for } 1 \leq i < i' \leq n_s, \end{cases} \quad (7)$$

and only keep the elements in

$$\mathfrak{J}_\ell := \mathfrak{J}_{\ell-1} \cap \mathbf{m}_\ell(\llbracket 1, q_\ell \rrbracket)$$

for the next step. If  $\ell = L$ , we just compute the final estimator of the ES given by

$$\widehat{\text{ES}} := \frac{1}{n_w} \sum_{i=1}^{n_w} \hat{\mu}_L^{\mathbf{m}_{L-1}(i)} = \frac{1}{n_w} \sum_{i \in \mathfrak{J}_{L-1}} \hat{\mu}_L^i.$$

Note that only the  $L - 1$ -first steps are used to select the worth scenarios, the step  $L$  is a pure Monte Carlo step. Again, the general idea is to reduce little by little the number of candidate scenarios to be part of the worst ones. As the number of candidates diminishes, one increases the number of simulated paths so as to reduce the variance of our Monte Carlo estimators and be able to differentiate between potentially closer true values of the associated conditional expectations.

**Remark 2.1.** *Note that, given  $j$ , we do not assume that the  $P_j^i, i \leq n_s$ , are independent. The simulations associated to different scenarios are in general not independent. Moreover, the  $\hat{\mu}_\ell^i, \ell \leq L$ , use the same simulated paths, only the number of used simulations changes. Both permit to reduce the computational cost, by allowing the use of the same simulations of the underlying processes across scenarios and steps.*

## 2.2 General a-priori bound on the $\mathbb{L}^p$ error

In this section, we first provide a general  $\mathbb{L}^p$  estimate of the error. A more tractable formulation will be provided in Corollary 2.3 under an additional sub-gamma distribution assumption.

From now on, we assume that  $P_{|s} \in \mathbb{L}^p$  for all  $p \geq 1$ , and we use the notations<sup>3</sup>

$$\begin{aligned} q &:= (q_0, q_1, \dots, q_L), \quad N = (N_0, N_1, \dots, N_L) \\ \delta q_\ell &:= q_{\ell-1} - q_\ell \quad \text{and} \quad \delta N_\ell := N_\ell - N_{\ell-1}, \quad \text{for } 1 \leq \ell \leq L, \end{aligned} \quad (8)$$

$$\delta \hat{\mu}_\ell^i := \frac{\sum_{j=N_{\ell-1}+1}^{N_\ell} P_j^i}{\delta N_\ell} = \frac{N_\ell \hat{\mu}_\ell^i - N_{\ell-1} \hat{\mu}_{\ell-1}^i}{\delta N_\ell}, \quad \text{for } 1 \leq i \leq n_s, \quad (9)$$

with the convention  $0/0 = 0$ .

<sup>2</sup>Note from the considerations below that only the elements  $(\hat{\mu}_\ell^i)_{i \in \mathfrak{J}_{\ell-1}}$  are needed in practice, the others are only defined here because they will be used in our proofs.

<sup>3</sup>The element  $q_L$  and  $N_0$  are defined for notational convenience, they never appear in our algorithm. To fix ideas, they can be set to  $q_L = n_w$  and  $N_0 = 0$  all over this paper.

**Proposition 2.2.** For all  $p \geq 1$ ,

$$\begin{aligned} \mathbb{E} \left[ \left| \text{ES} - \widehat{\text{ES}} \right|^p \right]^{\frac{1}{p}} &\leq \sum_{\ell=1}^{L-1} (\delta q_\ell)^{\frac{1}{p}} \max_{(i,k) \in \llbracket 1, n_w \rrbracket \times \llbracket q_\ell+1, n_s \rrbracket} (\mu^i - \mu^k) \mathbb{P}[\hat{\mu}_\ell^k > \hat{\mu}_\ell^i]^{\frac{1}{p}} \\ &+ \frac{1}{n_w} \frac{\delta N_L}{N_L} \max_{1 \leq i_1 < \dots < i_{n_w} \leq n_s} \left( \sum_{j=1}^{n_w} \mathbb{E} \left[ \left| \delta \hat{\mu}_L^{i_j} - \mu^{i_j} \right|^p \right]^{\frac{1}{p}} \right) \\ &+ \frac{1}{n_w} \frac{N_{L-1}}{N_L} \sum_{i=1}^{n_s} \mathbb{E} \left[ \left| \hat{\mu}_{L-1}^i - \mu^i \right|^p \right]^{\frac{1}{p}}. \end{aligned} \quad (10)$$

Before providing the proof of this general estimate, let us make some comments. The last two terms in (10) are natural as they are due to the Monte Carlo error made on the estimation of the various conditional expectations that can enter, after the  $(L-1)$ -levels selection procedure, in the estimation of ES. Note that it corresponds to the estimation errors using the cumulated number of Monte Carlo simulations  $N_{L-1}$  of step  $L-1$  and the number  $N_L - N_{L-1}$  of simulations used only for the last step. In practice, these numbers should be sufficiently large. The first term involves the quantities  $\max_{(i,k) \in \llbracket 1, n_w \rrbracket \times \llbracket q_\ell+1, n_s \rrbracket} (\mu^i - \mu^k) \mathbb{P}[\hat{\mu}_\ell^k > \hat{\mu}_\ell^i]^{\frac{1}{p}}$  with  $\ell = 1, \dots, L-1$ . Each term corresponds to the situation in which an element  $i \in \llbracket 1, n_w \rrbracket$  gets out the set of selected indexes  $\mathcal{J}_\ell$  exactly at the  $\ell$ -th step. In the worst situation, it is replaced by an element of index  $k$  larger than  $q_\ell$  and this can happen only if  $\hat{\mu}_\ell^k > \hat{\mu}_\ell^i$ . The probability of this event is controlled by the number of Monte Carlo simulations  $N_\ell$  used at the step  $\ell$  but also by the *distance* between the two scenarios. More specifically, for  $\ell$  small, one expects that  $\mathbb{P}[\hat{\mu}_\ell^k > \hat{\mu}_\ell^i]$  is small because the law of  $P_{|s_k}$  is concentrated far away from where the law of  $P_{|s_i}$  is. This quantity potentially increases with  $\ell$ , as we reduce the number of selected indexes. This should be compensated by an increase in the number of used Monte Carlo simulations. Otherwise stated, we expect to balance the various terms of (10) by considering a suitable increasing sequence  $(N_\ell)_{\ell \leq L}$ .

Obviously, (10) implies that the algorithm converges as  $N_\ell \rightarrow \infty$  for all  $\ell \leq L$ , see Proposition 3.3 below for a proof in a more general framework.

*Proof of Proposition 2.2.* We split the error into a permutation and a Monte Carlo error:

$$\mathbb{E} \left[ \left| \text{ES} - \widehat{\text{ES}} \right|^p \right]^{\frac{1}{p}} \leq \mathbb{E} \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \mu^i - \mu^{\mathfrak{m}_{L-1}(i)} \right|^p \right]^{\frac{1}{p}} + \mathbb{E} \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \hat{\mu}_L^{\mathfrak{m}_{L-1}(i)} - \mu^{\mathfrak{m}_{L-1}(i)} \right|^p \right]^{\frac{1}{p}}. \quad (11)$$

Let us first look at the second term which corresponds to a Monte Carlo error. We have

$$\begin{aligned} \mathbb{E} \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \hat{\mu}_L^{\mathfrak{m}_{L-1}(i)} - \mu^{\mathfrak{m}_{L-1}(i)} \right|^p \right]^{\frac{1}{p}} &\leq \frac{N_{L-1}}{N_L} \frac{1}{n_w} \mathbb{E} \left[ \left| \sum_{i \leq n_w} \hat{\mu}_{L-1}^{\mathfrak{m}_{L-1}(i)} - \mu^{\mathfrak{m}_{L-1}(i)} \right|^p \right]^{\frac{1}{p}} \\ &+ \frac{N_L - N_{L-1}}{N_L} \frac{1}{n_w} \mathbb{E} \left[ \left| \sum_{i \leq n_w} \frac{\sum_{j=N_{L-1}+1}^{N_L} \hat{P}_j^{\mathfrak{m}_{L-1}(i)}}{N_L - N_{L-1}} - \mu^{\mathfrak{m}_{L-1}(i)} \right|^p \right]^{\frac{1}{p}} \end{aligned}$$

in which

$$\mathbb{E} \left[ \left| \sum_{i \leq n_w} \hat{\mu}_{L-1}^{\mathfrak{m}_{L-1}(i)} - \mu^{\mathfrak{m}_{L-1}(i)} \right|^p \right]^{\frac{1}{p}} \leq \sum_{i \leq n_s} \mathbb{E} \left[ \left| \hat{\mu}_{L-1}^i - \mu^i \right|^p \right]^{\frac{1}{p}},$$

and

$$\begin{aligned}
\mathbb{E} \left[ \left| \sum_{i \leq n_w} \frac{\sum_{j=N_{L-1}+1}^{N_L} \hat{P}_j^{\mathbf{m}_{L-1}(i)}}{N_L - N_{L-1}} - \mu^{\mathbf{m}_{L-1}(i)} \right|^p \right]^{\frac{1}{p}} &= \mathbb{E} \left[ \mathbb{E} \left[ \left| \sum_{i \leq n_w} \delta \hat{\mu}_L^{\mathbf{m}_{L-1}(i)} - \mu^{\mathbf{m}_{L-1}(i)} \right|^p \middle| \mathbf{m}_{L-1} \right] \right]^{\frac{1}{p}} \\
&\leq \left( \max_{1 \leq i_1 < \dots < i_{n_w} \leq n_s} \mathbb{E} \left[ \left| \sum_{j=1}^{n_w} \delta \hat{\mu}_L^{i_j} - \mu^{i_j} \right|^p \right] \right)^{\frac{1}{p}} \\
&\leq \max_{1 \leq i_1 < \dots < i_{n_w} \leq n_s} \sum_{j=1}^{n_w} \mathbb{E} \left[ \left| \delta \hat{\mu}_L^{i_j} - \mu^{i_j} \right|^p \right]^{\frac{1}{p}}.
\end{aligned}$$

To discuss the first term in the right-hand side of (11), the permutation error, let us first define  $S_q[A]$  as the collection of the  $q$  smallest elements of a set  $A \subset \mathbb{N}$ . If  $i \in \llbracket 1, n_w \rrbracket \cap \mathcal{J}_{\ell-1} \setminus \mathcal{J}_\ell$ , then  $i \in S_{q_\ell}[\mathcal{J}_{\ell-1}] \setminus \mathcal{J}_\ell$  and therefore there exists  $k_i \in \mathcal{R}_\ell := \mathcal{J}_\ell \setminus S_{q_\ell}[\mathcal{J}_{\ell-1}]$ . Thus, on the set  $\{\{i_1, \dots, i_J\} = (\mathcal{J}_{\ell-1} \setminus \mathcal{J}_\ell) \cap \llbracket 1, n_w \rrbracket\}$ , one can define  $\mathfrak{k}_\ell(i_1) := \max \mathcal{R}_\ell$  and  $\mathfrak{k}_\ell(i_{j+1}) := \max\{k < \mathfrak{k}_\ell(i_j) : k \in \mathcal{R}_\ell\}$  for  $j+1 \leq J$ . Note that

$$\{i \in \mathcal{J}_{\ell-1} \setminus \mathcal{J}_\ell\} \subset \{\hat{\mu}_\ell^{\mathfrak{k}_\ell(i)} > \hat{\mu}_\ell^i\} \quad \text{and} \quad |\mathcal{R}_\ell| \leq q_{\ell-1} - q_\ell, \quad (12)$$

since  $\mathcal{R}_\ell \subset \mathcal{J}_{\ell-1} \setminus S_{q_\ell}[\mathcal{J}_{\ell-1}]$  and  $|\mathcal{J}_{\ell-1}| = q_{\ell-1}$ . Let  $\mathbf{A}_{q,q'}$  denote the collection of subsets  $A$  of  $\llbracket q+1, n_s \rrbracket$  such that  $|A| = q'$ . Then, it follows from (4), Hölder's inequality and (12) that

$$\begin{aligned}
\mathbb{E} \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \mu^i - \mu^{\mathbf{m}_{L-1}(i)} \right|^p \right]^{\frac{1}{p}} &\leq \frac{1}{n_w} \sum_{i \leq n_w} \sum_{\ell=1}^{L-1} \mathbb{E} \left[ \left| (\mu^i - \mu^{\mathfrak{k}_\ell(i)}) \mathbf{1}_{\{i \in \mathcal{J}_{\ell-1} \setminus \mathcal{J}_\ell\}} \right|^p \right]^{\frac{1}{p}} \\
&\leq \max_{i \leq n_w} \sum_{\ell=1}^{L-1} \mathbb{E} \left[ \left| (\mu^i - \mu^{\mathfrak{k}_\ell(i)}) \mathbf{1}_{\{i \in \mathcal{J}_{\ell-1} \setminus \mathcal{J}_\ell\}} \right|^p \right]^{\frac{1}{p}} \\
&\leq \max_{i \leq n_w} \sum_{\ell=1}^{L-1} \left( \max_{A \subset \mathbf{A}_{q_\ell, \delta q_\ell}} \sum_{k \in A} \mathbb{E} \left[ \left| (\mu^i - \mu^k) \right|^p \mathbf{1}_{\{i \in \mathcal{J}_{\ell-1} \setminus \mathcal{J}_\ell, \mathfrak{k}_\ell(i) = k\}} \right] \right)^{\frac{1}{p}} \\
&\leq \sum_{\ell=1}^{L-1} (\delta q_\ell)^{\frac{1}{p}} \max_{(i,k) \in \llbracket 1, n_w \rrbracket \times \llbracket q_\ell+1, n_s \rrbracket} (\mu^i - \mu^k) \mathbb{P}[\hat{\mu}_\ell^k > \hat{\mu}_\ell^i]^{\frac{1}{p}}.
\end{aligned}$$

□

### 2.3 Error bound for Sub-Gamma distributions

To illustrate how the general error bound of Proposition 2.2 can be used in practice to decide of the sequence  $(q_\ell, N_\ell)_\ell$ , we now consider the case where the components of  $P_{\mathbf{1}_s}$  have sub-gamma distributions, and apply Bernstein's inequality in (10), see e.g. [7, Chapter 2]. This requires the following assumption.

**Assumption 1.** *There exists  $c \in \mathbb{R}_+$  such that the random variables  $Z[i, k] := (P_{\mathbf{1}_s^i} - \mu^i) - (P_{\mathbf{1}_s^k} - \mu^k)$ ,  $i, k \leq n_s$ , satisfy Bernstein's condition :*

$$\mathbb{E} [|Z[i, k]|^p] \leq \frac{p! c^{p-2}}{2} \mathbb{E} [Z[i, k]^2], \quad i, k \leq n_s, \quad \text{for all } p \geq 3.$$

From now on, we shall assume that the constant  $c$  is known. It can usually be estimated in practice.

**Corollary 2.3.** *Assume that Assumption 1 holds. Then, for all  $p \geq 1$ ,*

$$\mathbb{E} \left[ \left| \text{ES} - \widehat{\text{ES}} \right|^p \right]^{\frac{1}{p}} \leq F_p(q, N) \quad (13)$$

in which

$$\begin{aligned}
F_p(q, N) &:= \sum_{\ell=1}^{L-1} (\delta q_\ell)^{\frac{1}{p}} \max_{(i,k) \in \llbracket 1, n_w \rrbracket \times \llbracket q_\ell+1, n_s \rrbracket} (\mu^i - \mu^k) e^{-\frac{N_\ell(\mu^i - \mu^k)^2}{2p(\sigma_{ik}^2 + c(\mu^i - \mu^k))}} \\
&+ \frac{1}{n_w} \frac{\delta N_L}{N_L} \max_{1 \leq i_1 < \dots < i_{n_w} \leq n_s} \sum_{j=1}^{n_w} \left( C_{p,\sigma} \frac{p\sigma_{i_j}^p}{(\delta N_L)^{\frac{p}{2}}} + C_{p,c} \frac{pc^p}{(\delta N_L)^p} \right)^{\frac{1}{p}} \\
&+ \frac{1}{n_w} \frac{N_{L-1}}{N_L} \sum_{i=1}^{n_s} \left( C_{p,\sigma} \frac{p\sigma_i^p}{(N_{L-1})^{\frac{p}{2}}} + C_{p,c} \frac{pc^p}{(N_{L-1})^p} \right)^{\frac{1}{p}}
\end{aligned} \tag{14}$$

with

$$\begin{cases} \sigma_{ik}^2 := \text{Var}[P_{|s^i} - P_{|s^k}] \text{ and } \sigma_i^2 := \text{Var}[P_{|s^i}], \quad i, k \leq n_s \\ C_{p,\sigma} := 2^{p-1} \Gamma\left(\frac{p}{2}\right) \text{ and } C_{p,c} := 4^p \Gamma(p) \end{cases} \tag{15}$$

where  $\Gamma$  is the Gamma function defined by

$$\Gamma(y) = \int_0^{+\infty} x^{y-1} e^{-x} dx, \quad y > 0.$$

The upper-bound of Corollary 2.3 has two advantages on Proposition 2.2. First, the dependence on  $(q_\ell, N_\ell)_{\ell \geq 0}$  is more explicit. It depends on unknown quantities, but we can estimate (at least rough) confidence intervals for them, see e.g. Section 2.4 below. Second, as we will see in the next section, it allows one to define a tractable deterministic optimal control problem satisfying a dynamic programming principle, or even simple heuristics (see Section 2.5), to select an appropriate sequence  $(q_\ell, N_\ell)_{\ell \geq 0}$ .

*Proof of Corollary 2.3.* The first term in (14) is an upper-bound for the first term in the right-hand side of (10), see [7, Theorem 2.1]. As for the two other terms in (10), we use the usual argument, for  $i \leq n_s$ ,

$$\mathbb{E} \left[ |\delta \hat{\mu}_L^i - \mu^i|^p \right] = \int_0^\infty px^{p-1} \mathbb{P}[|\delta \hat{\mu}_L^i - \mu^i| \geq x] dx$$

and

$$\mathbb{E} \left[ |\hat{\mu}_{L-1}^i - \mu^i|^p \right] = \int_0^\infty px^{p-1} \mathbb{P}[|\hat{\mu}_{L-1}^i - \mu^i| \geq x] dx,$$

and then appeal to [7, Theorem 2.1] again to deduce that

$$\begin{aligned}
\mathbb{E} \left[ |\delta \hat{\mu}_L^i - \mu^i|^p \right] &\leq \int_0^\infty px^{p-1} e^{-\frac{\delta N_L x^2}{2(\sigma_i^2 + cx)}} dx \\
&\leq \int_0^\infty px^{p-1} e^{-\frac{\delta N_L x^2}{4\sigma_i^2}} \mathbf{1}_{\{x \leq \frac{\sigma_i^2}{c}\}} dx + \int_0^\infty px^{p-1} e^{-\frac{\delta N_L x}{4c}} \mathbf{1}_{\{x > \frac{\sigma_i^2}{c}\}} dx \\
&\leq \frac{p(\sigma_i^2)^{\frac{p}{2}}}{(\delta N_L)^{\frac{p}{2}}} \int_0^\infty y^{p-1} e^{-\frac{y^2}{4}} dy + \frac{pc^p}{(\delta N_L)^p} \int_0^\infty y^{p-1} e^{-\frac{y}{4}} dy, \\
&\leq \frac{p\sigma_i^p}{(\delta N_L)^{\frac{p}{2}}} 2^{p-1} \Gamma\left(\frac{p}{2}\right) + \frac{pc^p}{(\delta N_L)^p} 4^p \Gamma(p),
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E} \left[ |\hat{\mu}_{L-1}^i - \mu^i|^p \right] &\leq \int_0^\infty px^{p-1} e^{-\frac{N_{L-1} x^2}{2(\sigma_i^2 + cx)}} dx \\
&\leq \int_0^\infty px^{p-1} e^{-\frac{N_{L-1} x^2}{4\sigma_i^2}} \mathbf{1}_{\{x \leq \frac{\sigma_i^2}{c}\}} dx + \int_0^\infty px^{p-1} e^{-\frac{N_{L-1} x}{4c}} \mathbf{1}_{\{x > \frac{\sigma_i^2}{c}\}} dx \\
&\leq \frac{p(\sigma_i^2)^{\frac{p}{2}}}{(N_{L-1})^{\frac{p}{2}}} \int_0^\infty y^{p-1} e^{-\frac{y^2}{4}} dy + \frac{pc^p}{(N_{L-1})^p} \int_0^\infty y^{p-1} e^{-\frac{y}{4}} dy, \\
&\leq \frac{p\sigma_i^p}{(N_{L-1})^{\frac{p}{2}}} 2^{p-1} \Gamma\left(\frac{p}{2}\right) + \frac{pc^p}{(N_{L-1})^p} 4^p \Gamma(p).
\end{aligned}$$

□



**Remark 2.4.** If the  $(\hat{\mu}_\ell^i)_{i \leq n_s}$  and  $(\delta \hat{\mu}_\ell^i)_{i \leq n_s}$  are Gaussian, which is the case asymptotically, then the bound of Corollary 2.3 remains valid with  $c = 0$ . This fact will be used later on for simplifying our numerical algorithms.

## 2.4 Optimal a-priori allocation by deterministic dynamic programming based on fixed a-priori bounds

Given  $N := (N_\ell)_{0 \leq \ell \leq L}$  and  $q = (q_\ell)_{0 \leq \ell \leq L-1}$ , the total computation cost is

$$C(q, N) := \sum_{\ell=0}^{L-1} q_\ell (N_{\ell+1} - N_\ell)$$

with the convention  $N_0 := 0$ . Let  $\mathcal{N}$  denote the collection of non-decreasing sequences  $N := (N_\ell)_{0 \leq \ell \leq L}$  with values in  $\mathbb{N}$  such that  $N_0 = 0$ , and let  $\mathcal{Q}$  denote the collections of non-increasing sequences<sup>4</sup>  $q = (q_\ell)_{0 \leq \ell \leq L}$  with values in  $\llbracket n_w, n_s \rrbracket$  satisfying (5). In this section, we fix a total effort  $K > 0$  and recall how  $F_p(q, N)$ , as defined in (14), can be minimized over the collection  $\mathcal{A}$  of sequences  $(N, q) \in \mathcal{N} \times \mathcal{Q}$  satisfying  $C(N, q) \leq K$  by using a standard dynamic programming approach.

Given  $(\bar{q}, \bar{N}) \in \mathcal{Q} \times \mathcal{N}$  and  $0 \leq \ell \leq L-1$ , we write

$$\begin{aligned} F_p(\ell, \bar{q}, \bar{N}) := & \frac{1}{n_w} \frac{\delta \bar{N}_L}{\bar{N}_L} \max_{1 \leq i_1 < \dots < i_{n_w} \leq n_s} \sum_{j=1}^{n_w} \left( C_{p,\sigma} \frac{p \sigma_{i_j}^p}{(\delta \bar{N}_L)^{\frac{p}{2}}} + C_{p,c} \frac{p c^p}{(\delta \bar{N}_L)^p} \right)^{\frac{1}{p}} \\ & + \frac{1}{n_w} \frac{\bar{N}_{L-1}}{\bar{N}_L} \sum_{i=1}^{n_s} \left( C_{p,\sigma} \frac{p \sigma_i^p}{(\bar{N}_{L-1})^{\frac{p}{2}}} + C_{p,c} \frac{p c^p}{(\bar{N}_{L-1})^p} \right)^{\frac{1}{p}} + \mathbf{1}_{\{\ell < L-1\}} \sum_{\ell'=\ell+1}^{L-1} f_p(\bar{q}_{\ell'}, \bar{q}_{\ell'-1}, \bar{N}_{\ell'}), \end{aligned} \quad (16)$$

where

$$f_p(\bar{q}_{\ell'}, \bar{q}_{\ell'-1}, \bar{N}_{\ell'}) := (\delta \bar{q}_{\ell'})^{\frac{1}{p}} \max_{(i,k) \in \llbracket 1, n_w \rrbracket \times \llbracket \bar{q}_{\ell'}+1, n_s \rrbracket} (\mu^i - \mu^k) e^{-\frac{\bar{N}_{\ell'} (\mu^i - \mu^k)^2}{2p(\sigma_{ik}^2 + c(\mu^i - \mu^k))}},$$

and define

$$\hat{F}_p(\ell, \bar{q}, \bar{N}) = \min_{(\bar{q}', \bar{N}') \in \mathcal{A}(\ell, \bar{q}, \bar{N})} F_p(\ell, \bar{q}', \bar{N}')$$

where<sup>5</sup>

$$\mathcal{A}(\ell, \bar{q}, \bar{N}) := \{(\bar{q}', \bar{N}') \in \mathcal{Q} \times \mathcal{N} : (\bar{q}'_l, \bar{N}'_l)_{0 \leq l \leq \ell} = (\bar{q}_l, \bar{N}_l)_{0 \leq l \leq \ell} \text{ and } C(\bar{q}', \bar{N}') \leq K\}, \ell \geq 0.$$

Then, the dynamic programming principle implies that

$$\hat{F}_p(\ell, \bar{q}, \bar{N}) = \min_{(\bar{q}', \bar{N}') \in \mathcal{A}(\ell, \bar{q}, \bar{N})} \left[ \hat{F}_p(\ell+1, \bar{q}', \bar{N}') + f_p(\bar{q}'_{\ell+1}, \bar{q}_\ell, \bar{N}'_{\ell+1}) \right], \text{ for } 0 \leq \ell < L-1.$$

This reduces the search for an optimal selection of  $(\bar{N}, \bar{q})$  to  $L-1$  one-step optimization problems, which is much simpler to solve than the optimization problem associated to the left-hand side of (13).

In practice, the exact values of  $(\mu^i, \sigma_i^2)_{i \leq n_s}$  and  $(\sigma_{ik}^2)_{i,k \leq n_s}$  are not known. However, one can consider robust versions of the above. For instance, if we know that there exists some  $(\underline{\delta}_q, \bar{\delta}_q)_{q \leq n_s}$  and  $\bar{\sigma}^2$  such that

$$\begin{cases} 0 \leq \underline{\delta}_q \leq \mu^i - \mu^k \leq \bar{\delta}_q, (i, k) \in \llbracket 1, n_w \rrbracket \times \llbracket q+1, n_s \rrbracket \\ \sigma_i^2 \vee \sigma_k^2 \vee \sigma_{ik}^2 \leq \bar{\sigma}^2, (i, k) \in \llbracket 1, n_w \rrbracket \times \llbracket n_w+1, n_s \rrbracket, \end{cases} \quad (17)$$

<sup>4</sup>We write  $(q_\ell)_{0 \leq \ell \leq L}$  for convenience also  $q_L$  will never play any role.

<sup>5</sup>In the following, we only write  $\mathcal{A}(0)$  for  $\ell = 0$  as it does not depend on  $(\bar{q}, \bar{N})$ .

then one can similarly minimize the upper-bound of  $F_p$  defined as

$$\begin{aligned} & \frac{\delta \bar{N}_L}{\bar{N}_L} \left( C_{p,\sigma} \frac{p\bar{\sigma}^p}{(\delta \bar{N}_L)^{\frac{p}{2}}} + C_{p,c} \frac{pc^p}{(\delta \bar{N}_L)^p} \right) + \frac{n_s}{n_w} \frac{\bar{N}_{L-1}}{\bar{N}_L} \left( C_{p,\sigma} \frac{p\bar{\sigma}^p}{(\bar{N}_{L-1})^{\frac{p}{2}}} + C_{p,c} \frac{pc^p}{(\bar{N}_{L-1})^p} \right) \\ & + \mathbf{1}_{\{\ell < L-1\}} \sum_{\ell'=\ell}^{L-1} \tilde{f}_p(\bar{q}_{\ell'}, \bar{q}_{\ell'-1}, \bar{N}_{\ell'}), \end{aligned}$$

with

$$\tilde{f}_p(\bar{q}_{\ell'}, \bar{q}_{\ell'-1}, \bar{N}_{\ell'}) := (\delta \bar{q}_{\ell'})^{\frac{1}{p}} \max_{\delta_{\bar{q}_{\ell'}} \leq \delta \leq \bar{\delta}_{\bar{q}_{\ell'}}} \delta e^{-\frac{\bar{N}_{\ell'} \delta^2}{2p(\bar{\sigma}^2 + c\delta)}}.$$

This corresponds to a worst case scenario, when only the a priori bounds  $(\delta_q, \bar{\delta}_q)_{q \leq n_s}$  and  $\bar{\sigma}^2$  are known. In the above, one can also impose that  $q$  takes values in a given subset of  $\bar{Q}$  of  $\mathcal{Q}$ . In this case, we will only need to know  $(\delta_q, \bar{\delta}_q)_{q \in \bar{Q}}$ .

We refer to Section 4 below for numerical tests that show that such an algorithm seems to perform pretty well. Note that the optimization can be done off-line.

## 2.5 Simplified 2-levels algorithm for a linear indifference zone's size

Inspired by [3, 5, 6, 14], we assume here that we know the value of a constant  $\delta_0 > 0$  such that the impacts of the  $n_w$  worst scenarios have values that are separated by at least  $(k - n_w)\delta_0$  from the  $k$ -th worst scenario, for  $k > n_w$ :

$$\mu^{n_w} - \mu^k \geq (k - n_w) \delta_0, \quad \forall k \in \llbracket n_w + 1, n_s \rrbracket. \quad (18)$$

To illustrate this, we plot on Figures 1-4 the curves  $k \mapsto |\mu^{n_w} - \mu^k|$  for different formerly used test books of Natixis. We see that they are more flat on the interval  $[100, 120]$ , so that a rather conservative value would be the minimum (over the different books) of  $(\mu^{100} - \mu^{120})/20$ . Another choice in practice could be to take the ratio  $(\mu^{n_w} - \mu^{100})/(100 - n_w)$  which amounts to considering only the first part of the curve, and neglecting points that are anyway far from the worst scenarios.

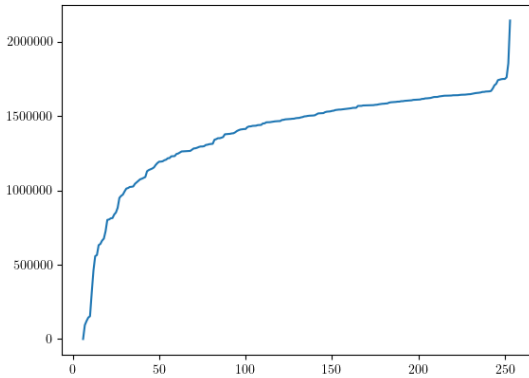


Figure 1:  $k \mapsto |\mu^{n_w} - \mu^k|$  for Book #1

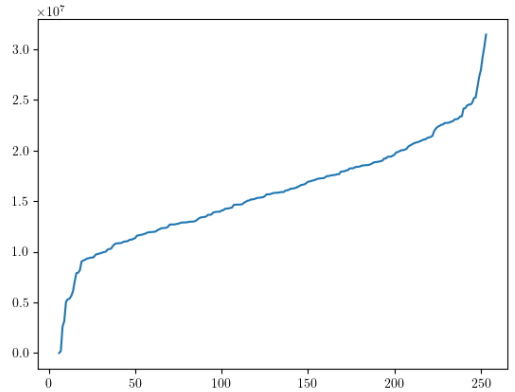


Figure 2:  $k \mapsto |\mu^{n_w} - \mu^k|$  for Book #2

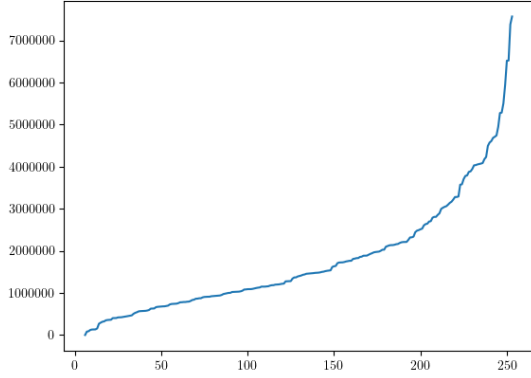


Figure 3:  $k \mapsto |\mu^{n_w} - \mu^k|$  for Book #3

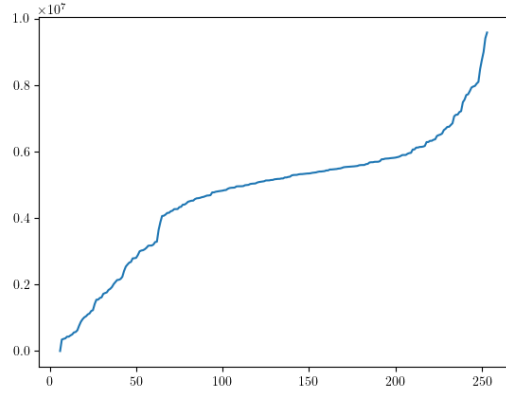


Figure 4:  $k \mapsto |\mu^{n_w} - \mu^k|$  for Book #4

We now consider a simplified version of the algorithm of Section 2.1 where we only do one intermediate “fast pricing” (meaning  $N_1$  rather small) and one final “full pricing” (meaning  $N_2$  large). In theory, this corresponds to  $L = 3$  with  $q_2 = n_w$ ,  $\delta N_3 = 0$  and  $\delta N_2 \rightarrow \infty$ . As  $\delta N_2 \rightarrow \infty$ , the second and third terms in (14) vanish, as well as the component of the first term corresponding to  $\ell = 2$ . We therefore neglect them. In practice, we only take  $N_2$  large enough (and given) from the point of view of the bank, and minimize over  $(q_1, N_1)$  the remaining term in (14):

$$F_1^\infty(q_1) := (n_s - q_1)^{\frac{1}{p}} \max_{(i,k) \in \llbracket 1, n_w \rrbracket \times \llbracket q_1 + 1, n_s \rrbracket} (\mu^i - \mu^k) e^{-\frac{N_1(\mu^i - \mu^k)^2}{2p(\bar{\sigma}^2 + c(\mu^i - \mu^k))}},$$

in which  $\bar{\sigma}$  is estimated to be as in (17), under the computation cost constraint

$$C(N_1, q_1) = q_1(N_2 - N_1) + n_s N_1 \leq K$$

for some given maximal cost  $K \in \mathbb{N}^*$ .

For  $N_1$  (or  $K$ ) large enough, the condition (18) leads to minimizing over  $q_1 \in \llbracket n_w, n_s \rrbracket \cap [1, K/N_2]$  the upper-bound

$$h_0^p(q_1) := (n_s - q_1)^{\frac{1}{p}} \times (q_1 + 1 - n_w) \delta_0 \exp\left(-\frac{(K - q_1 N_2)(q_1 + 1 - n_w)^2 \delta_0^2}{2p(n_s - q_1)(\bar{\sigma}^2 + c(q_1 + 1 - n_w) \delta_0)}\right). \quad (19)$$

The optimal  $q_1^*$  can then be found easily by performing a one-dimensional numerical minimization. Upon replacing  $n_s - q_1$  by  $n_s$  in the denominator of the exponential term, which provides a further upper-bound, the optimum can even be computed explicitly, see Appendix A. This provides a very easy to use algorithm.

Considering the case  $p = 1$ , let us now perform first numerical tests to see if the proxy based on  $h_0^1$  is far from  $F_1^\infty$ . We use the parameters of Tables 1 and 2 below and  $\mu^i = -i\delta_0$ ,  $i \leq n_s$ , where  $\delta_0 := (\mu^{n_w} - \mu^{100})/(100 - n_w)$  for the  $\mu^i$ 's of Figure 6. In particular, we take  $c = 0$ , see Remark 2.4.

In Figure 5, the two increasing curves show the optimum  $q_1^*$  (right axis) as found when applying the deterministic dynamic programming algorithm (dashed line) of Section 2.4 associated to the real sample book curve of Figure 6, and the heuristic (solid line) based on (19). The two decreasing curves show the corresponding  $F_1^\infty(q_1^*)$  (left axis) found when applying the deterministic dynamic programming algorithm (dotted line) and the heuristic (dashdot line). We see that the heuristic and the real minimizer are extremely close. The noise in the lines associated to the dynamic programming algorithm are due to grid effects.

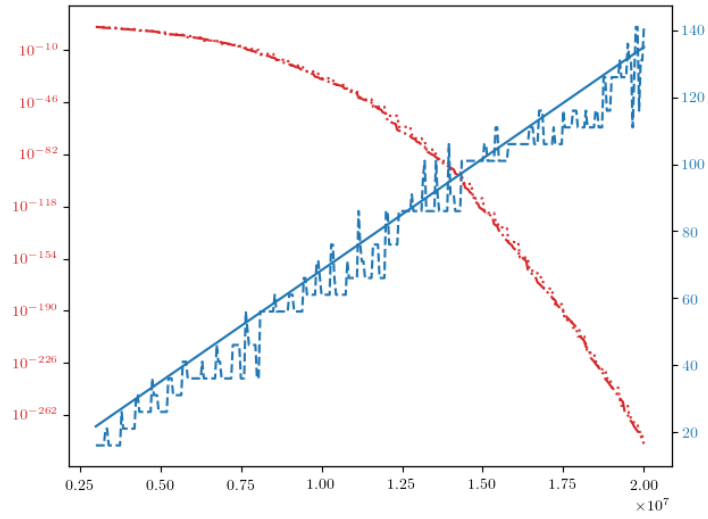


Figure 5:  $q_1^*$  vs  $K$  for the distribution of Figure 6

$\delta_0$	2 766
$c$	0
$\bar{\sigma}$	$\sqrt{2(1-\rho)} \times 2\,200\,000$
$\rho$	0.6
$n_s$	253
$n_w$	6

Table 1: Sample Book Parameters

$K$	$10^7$
$N_2$	$10^5$

Table 2: Computing Power

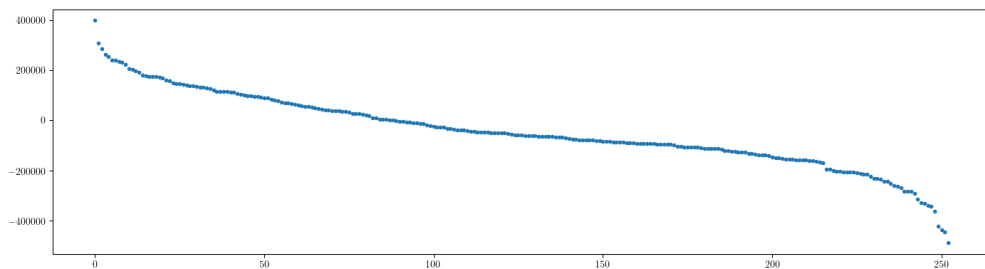


Figure 6: Sample book distribution :  $i \mapsto \mu^i$  for  $i \leq n_s$

### 3 Adaptative algorithm

Although the true value  $\theta_o = (\mu_o, \Sigma_o)$  of the vector of means and of the covariance matrix of  $P_{|s}$  are unknown, we can set on it a prior distribution, e.g. based on previous Monte Carlo experiments, rather than just working on robust bounds as in the end of Section 2.4. Since the estimation of ES uses Monte Carlo simulations of  $P_{|s}$ , the knowledge of these quantities can be improved along the different steps  $\ell$  of our estimation procedure. This suggests an adaptative algorithm for the optimization of the numerical effort allocation, in which we learn progressively the true value of these parameters, or part of them. From now on, we therefore view the true value of the parameters as a random variable  $\tilde{\theta} := (\tilde{\mu}, \tilde{\Sigma})$  on which a prior law  $\nu_0$  is set. At each step  $\ell$ , new Monte Carlo simulations will allow us to update this prior, and our strategy for the next steps accordingly.

#### 3.1 Error bounds and convergence for predictable strategies

Let us first adapt the proof of Proposition 2.2 and Corollary 2.3 to the case where controls are not deterministic but stochastic processes. Given a stochastic process  $\alpha$  with values in  $\mathcal{Q} \times \mathcal{N}$ , we set  $(q^\alpha, N^\alpha) := \alpha$  where  $q^\alpha$  and  $N^\alpha$  are respectively  $\mathcal{Q}$  and  $\mathcal{N}$ -valued. We then define  $\hat{\mu}^\alpha = (\hat{\mu}_\ell^\alpha)_{\ell \leq L}$ ,  $(\mathcal{J}_\ell^\alpha, \mathfrak{m}_\ell^\alpha)_{\ell \leq L}$  as in Section 2.1 except that we see  $\hat{\mu}_\ell^\alpha$  as a  $q_\ell^\alpha$ -dimensional random variables with entries given by  $(\hat{\mu}_\ell^{\alpha, i})_{i \in \mathcal{J}_\ell^\alpha}$ . We use the same convention for  $\delta \hat{\mu}_\ell^\alpha$ , recall (9). We say that  $\alpha$  is admissible if it is predictable with respect to  $(\mathcal{F}_\ell^\alpha)_{\ell \leq L}$  in which  $\mathcal{F}_0^\alpha$  is trivial and  $\mathcal{F}_\ell^\alpha = \mathcal{F}_{\ell-1}^\alpha \vee \sigma(P_j^i, (i, j) \in \mathcal{J}_\ell^\alpha \times \llbracket 1, N_\ell^\alpha \rrbracket)$ . We call  $\mathcal{A}^{\text{ad}}$  the collection of such processes. Then, one defines

$$\widehat{\text{ES}}^\alpha := \frac{1}{n_w} \sum_{i=1}^{n_w} \hat{\mu}_L^{\alpha, \mathfrak{m}_{L-1}^\alpha(i)}, \quad \alpha \in \mathcal{A}^{\text{ad}}.$$

The true value of the expected shortfall is now also written as a random variable

$$\widetilde{\text{ES}} := \frac{1}{n_w} \sum_{i=1}^{n_w} \tilde{\mu}^{\tilde{\mathfrak{m}}(i)},$$

in which  $\tilde{\mathfrak{m}}$  is the random permutation such that

$$\begin{cases} \tilde{\mu}^{\tilde{\mathfrak{m}}(1)} \geq \tilde{\mu}^{\tilde{\mathfrak{m}}(2)} \geq \dots \geq \tilde{\mu}^{\tilde{\mathfrak{m}}(n_s)}, \\ \tilde{\mathfrak{m}}(i) < \tilde{\mathfrak{m}}(i') \text{ if } \tilde{\mu}^{\tilde{\mathfrak{m}}(i)} = \tilde{\mu}^{\tilde{\mathfrak{m}}(i')} \text{ for } 1 \leq i < i' \leq n_s. \end{cases}$$

We let  $\mathcal{M}$  be a collection of laws on  $\mathbb{R}^{n_s} \times \mathbb{S}^{n_s}$ , where  $\mathbb{S}^{n_s}$  denotes the collection of covariance matrices of size  $n_s$ . Given  $\nu \in \mathcal{M}$ , we denote by  $\mathbb{E}^\nu$  the expectation operator given that  $\tilde{\theta}$  admits the law  $\nu$ . When  $\nu$  is a Dirac mass, we retrieve the situation of Section 2 (up to re-ordering in a deterministic way the components of  $\mu$ ).

We first provide a natural extension of Proposition 2.2.

**Proposition 3.1.** *For all  $p \geq 1$ ,  $\nu \in \mathcal{M}$ , and  $\alpha \in \mathcal{A}^{\text{ad}}$ ,*

$$\begin{aligned} \mathbb{E}^\nu \left[ \left| \widetilde{\text{ES}} - \widehat{\text{ES}}^\alpha \right|^p \right]^{\frac{1}{p}} &\leq \frac{1}{n_w} \mathbb{E}^\nu \left[ \left| \sum_{i \in \mathcal{J}_{L-1}^\alpha} \hat{\mu}_L^{\alpha, i} - \tilde{\mu}^i \right|^p \right]^{\frac{1}{p}} \\ &+ \sum_{\ell=1}^{L-1} \mathbb{E}^\nu \left[ \delta q_\ell^\alpha \max_{(i, k) \in \tilde{\mathfrak{m}}_{\ell-1}^\alpha(\llbracket 1, n_w \rrbracket) \times \llbracket q_\ell^\alpha + 1, q_{\ell-1}^\alpha \rrbracket} (\tilde{\mu}^i - \tilde{\mu}^k)^p \mathbb{P}^\nu[\hat{\mu}_\ell^{\alpha, k} > \hat{\mu}_\ell^{\alpha, i} | \mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta})] \right]^{\frac{1}{p}}, \end{aligned} \quad (20)$$

with the convention  $\max_\emptyset = 0$  and in which  $\tilde{\mathfrak{m}}_{\ell-1}^\alpha$  is defined as  $\tilde{\mathfrak{m}}$  but on the subset  $\mathcal{J}_{\ell-1}^\alpha$  instead of  $\mathcal{J}_0^\alpha = \llbracket 1, n_s \rrbracket$ .

*Proof.* We proceed as in the proof of Proposition 2.2 to obtain that

$$\mathbb{E}^\nu \left[ \left| \widetilde{\text{ES}} - \widehat{\text{ES}}^\alpha \right|^p \right]^{\frac{1}{p}} \leq \mathbb{E}^\nu \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \hat{\mu}_L^{\alpha, \mathfrak{m}_{L-1}^\alpha(i)} - \tilde{\mu}^{\mathfrak{m}_{L-1}^\alpha(i)} \right|^p \right]^{\frac{1}{p}} + \mathbb{E}^\nu \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \tilde{\mu}^{\tilde{\mathfrak{m}}(i)} - \tilde{\mu}^{\mathfrak{m}_{L-1}^\alpha(i)} \right|^p \right]^{\frac{1}{p}},$$

where

$$\mathbb{E}^\nu \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \hat{\mu}_L^{\alpha, \mathfrak{m}_{L-1}^\alpha(i)} - \tilde{\mu}^{\mathfrak{m}_{L-1}^\alpha(i)} \right|^p \right]^{\frac{1}{p}} = \frac{1}{n_w} \mathbb{E}^\nu \left[ \left| \sum_{i \in \mathcal{J}_{L-1}^\alpha} \hat{\mu}_L^{\alpha, i} - \tilde{\mu}^i \right|^p \right]^{\frac{1}{p}}.$$

We define  $\mathfrak{k}_\ell^\alpha$  as  $\mathfrak{k}_\ell$  in the proof of Proposition 2.2 for the strategy  $\alpha$ , with  $\mathcal{R}_\ell$  replaced by  $\mathcal{R}_\ell^\alpha := \mathcal{J}_\ell^\alpha \setminus \tilde{\mathfrak{m}}(\mathcal{S}_{q_\ell^\alpha}[\tilde{\mathfrak{m}}^{-1}(\mathcal{J}_{\ell-1})])$ . Then,

$$\begin{aligned} & \mathbb{E}^\nu \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \tilde{\mu}^{\tilde{\mathfrak{m}}(i)} - \tilde{\mu}^{\mathfrak{m}_{L-1}^\alpha(i)} \right|^p \right]^{\frac{1}{p}} \\ & \leq \mathbb{E}^\nu \left[ \left| \frac{1}{n_w} \sum_{i \leq n_w} \sum_{\ell=1}^{L-1} (\tilde{\mu}^{\tilde{\mathfrak{m}}(i)} - \tilde{\mu}^{\mathfrak{k}_\ell^\alpha(\tilde{\mathfrak{m}}(i))}) \mathbf{1}_{\{\tilde{\mathfrak{m}}(i) \in \mathcal{J}_{\ell-1}^\alpha \setminus \mathcal{J}_\ell^\alpha\}} \right|^p \right]^{\frac{1}{p}} \\ & \leq \frac{1}{n_w} \sum_{\ell=1}^{L-1} \sum_{i \leq n_w} \mathbb{E}^\nu \left[ \left| (\tilde{\mu}^{\tilde{\mathfrak{m}}(i)} - \tilde{\mu}^{\mathfrak{k}_\ell^\alpha(\tilde{\mathfrak{m}}(i))}) \right|^p \mathbf{1}_{\{\tilde{\mathfrak{m}}(i) \in \mathcal{J}_{\ell-1}^\alpha \setminus \mathcal{J}_\ell^\alpha\}} \right]^{\frac{1}{p}} \\ & \leq \frac{1}{n_w} \sum_{\ell=1}^{L-1} \sum_{i \leq n_w} \mathbb{E}^\nu \left[ \sum_{k \in \tilde{\mathfrak{m}}_{\ell-1}^\alpha(\llbracket q_\ell^\alpha + 1, q_{\ell-1}^\alpha \rrbracket)} \left| (\tilde{\mu}^{\tilde{\mathfrak{m}}(i)} - \tilde{\mu}^k) \right|^p \mathbf{1}_{\{\tilde{\mathfrak{m}}(i) \in \mathcal{J}_{\ell-1}^\alpha \setminus \mathcal{J}_\ell^\alpha, \mathfrak{k}_\ell^\alpha(\tilde{\mathfrak{m}}(i)) = k\}} | \mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta}) \right]^{\frac{1}{p}} \\ & \leq \sum_{\ell=1}^{L-1} \mathbb{E}^\nu \left[ \delta q_\ell^\alpha \max_{(i,k) \in \tilde{\mathfrak{m}}_{\ell-1}^\alpha(\llbracket 1, n_w \rrbracket) \times \llbracket q_\ell^\alpha + 1, q_{\ell-1}^\alpha \rrbracket} (\tilde{\mu}^i - \tilde{\mu}^k)^p \mathbb{P}^\nu[\hat{\mu}_\ell^{\alpha, k} > \hat{\mu}_\ell^{\alpha, i} | \mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta})] \right]^{\frac{1}{p}}. \end{aligned}$$

□

**Remark 3.2.** Note that, when  $\alpha$  is deterministic and  $\nu$  is concentrated on a Dirac, the right-hand side of (20) is bounded from above by

$$\begin{aligned} & \frac{1}{n_w} \frac{\delta N_L^\alpha}{N_L^\alpha} \max_{1 \leq i_1 < \dots < i_{n_w} \leq n_s} \sum_{j=1}^{n_w} \mathbb{E}^\nu \left[ \left| \delta \hat{\mu}_L^{\alpha, i_j} - \tilde{\mu}^{i_j} \right|^p \right]^{\frac{1}{p}} + \frac{1}{n_w} \frac{N_{L-1}^\alpha}{N_L^\alpha} \sum_{i=1}^{n_s} \mathbb{E}^\nu \left[ \left| \hat{\mu}_{L-1}^{\alpha, i} - \tilde{\mu}^i \right|^p \right]^{\frac{1}{p}} \\ & + \sum_{i=1}^{n_s} \sum_{\ell=1}^{L-1} (\delta q_\ell^\alpha)^{\frac{1}{p}} \mathbb{E}^\nu \left[ \max_{(i,k) \in \tilde{\mathfrak{m}}\llbracket 1, n_w \rrbracket \times \llbracket q_\ell^\alpha + 1, n_s \rrbracket} (\tilde{\mu}^i - \tilde{\mu}^k)^p \mathbb{P}^\nu[\hat{\mu}_\ell^{\alpha, k} > \hat{\mu}_\ell^{\alpha, i} | \mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta})] \right]^{\frac{1}{p}}, \end{aligned}$$

which coincides with the bound of Proposition 2.2

The above guarantees the convergence of the algorithm.

**Proposition 3.3.** Let  $(K^n)_{n \geq 1} \subset \mathbb{N}^*$  be a sequence converging to infinity and let  $(\alpha^n)_{n \geq 1}$  be a sequence in  $\mathcal{A}^{\text{ad}}$  such that  $C(q^{\alpha^n}, N^{\alpha^n}) \leq K^n$  for each  $n \geq 1$ . Assume further that  $\min_{1 \leq \ell \leq L} N_\ell^{\alpha^n} \rightarrow \infty$  a.s. Let  $\nu$  be concentrated on the Dirac mass on  $\theta_\circ$ . Then,

$$\mathbb{E}^\nu \left[ \left| \widetilde{\text{ES}} - \widehat{\text{ES}}^{\alpha^n} \right|^p \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

*Proof.* It suffices to use the fact that, for some  $C_p > 0$ ,

$$\mathbb{E}^\nu \left[ \left| \hat{\mu}_\ell^{\alpha^n, i} - \tilde{\mu}^i \right|^p \right] \leq C_p \mathbb{E}^\nu \left[ \frac{\delta N_\ell^{\alpha^n}}{N_\ell^{\alpha^n}} \mathbb{E}^\nu \left[ \left| \delta \hat{\mu}_\ell^{\alpha^n, i} - \mu_\circ^i \right|^p | \mathcal{F}_{\ell-1}^{\alpha^n} \right] \right] + C_p \mathbb{E}^\nu \left[ \frac{N_{\ell-1}^{\alpha^n}}{N_\ell^{\alpha^n}} \left| \hat{\mu}_{\ell-1}^{\alpha^n, i} - \mu_\circ^i \right|^p \right],$$

in which

$$\frac{\delta N_\ell^{\alpha^n}}{N_\ell^{\alpha^n}} \mathbb{E}^\nu \left[ \left| \delta \hat{\mu}_\ell^{\alpha^n, i} - \mu_\circ^i \right|^p | \mathcal{F}_{\ell-1}^{\alpha^n} \right] \rightarrow 0, \quad \nu_\circ - \text{a.s.},$$

for all  $\ell > 1$  and  $i \leq n_s$ . By induction, this implies that

$$\mathbb{E}^\nu \left[ \left| \hat{\mu}_\ell^{\alpha^n, i} - \tilde{\mu}^i \right|^p \right] = \mathbb{E}^\nu \left[ \left| \hat{\mu}_\ell^{\alpha^n, i} - \mu_\circ^i \right|^p \right] \rightarrow 0$$

for all  $\ell \leq L$  and  $i \leq n_s$ . Moreover, for some  $C > 0$ ,

$$\mathbb{E}^\nu \left[ (\tilde{\mu}^i - \tilde{\mu}^k)^p \mathbb{P}^\nu [\hat{\mu}_\ell^{\alpha^n, k} > \hat{\mu}_\ell^{\alpha^n, i} | \mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta})] \right] \leq C \mathbf{1}_{\{\mu_\circ^i - \mu_\circ^k > 0\}} \frac{\mathbb{E}^\nu [|\hat{\mu}_\ell^{\alpha^n, i} - \hat{\mu}_\ell^{\alpha^n, k} - (\mu_\circ^i - \mu_\circ^k)|]}{\mu_\circ^i - \mu_\circ^k} \rightarrow 0$$

for all  $i < k$  and  $\ell \leq L - 1$ .  $\square$

Using the fact that a control  $\alpha \in \mathcal{A}^{\text{ad}}$  is predictable, one can then proceed as in the proof of Corollary 2.3 to derive a more tractable upper-bound. It appeals to the following version of Assumption 1.

**Assumption 2.** *There exists  $c > 0$  such that, for all  $\nu \in \mathcal{M}$ ,*

$$\mathbb{E}^\nu \left[ |Z[i, k]|^p | \sigma(\tilde{\theta}) \right] \leq \frac{p! c^{p-2}}{2} \mathbb{E}^\nu \left[ |Z[i, k]|^2 | \sigma(\tilde{\theta}) \right] \quad \nu - \text{a.s.}, \quad \text{for all } i, k \leq n_s, \quad p \geq 3.$$

**Corollary 3.4.** *Let Assumption 2 holds. Then, for all  $p \geq 1$ ,  $\alpha \in \mathcal{A}^{\text{ad}}$  and  $\nu \in \mathcal{M}$ ,*

$$\mathbb{E}^\nu \left[ \left| \widehat{\text{ES}} - \widetilde{\text{ES}}^\alpha \right|^p \right]^{\frac{1}{p}} \leq F_p^{\text{ad}}(\alpha, \nu)$$

in which

$$F_p^{\text{ad}}(\alpha, \nu) := \frac{1}{n_w} \mathbb{E}^\nu \left[ \left| \sum_{i \in \mathcal{I}_{L-1}^\alpha} \hat{\mu}_L^{\alpha, i} - \tilde{\mu}^i \right|^p \right]^{\frac{1}{p}} + \sum_{\ell=1}^{L-1} \mathbb{E}^\nu \left[ f_p^{\text{ad}}(\ell, \alpha, \tilde{\theta}) \right]^{\frac{1}{p}}$$

where

$$f_p^{\text{ad}}(\ell, \alpha, \tilde{\theta}) := \delta q_\ell^\alpha \max_{(i, k) \in \tilde{\mathfrak{m}}_{\ell-1}^\alpha(\llbracket 1, n_w \rrbracket \times \llbracket q_\ell^\alpha + 1, q_{\ell-1}^\alpha \rrbracket)} (\tilde{\mu}^i - \tilde{\mu}^k)^p \left( e^{-\frac{\delta N_\ell^\alpha (\rho_\ell^\alpha [i, k])^2}{2(\tilde{\sigma}_{ik}^2 + c\rho_\ell^\alpha [i, k])}} \mathbf{1}_{\{\rho_\ell^\alpha [i, k] \geq 0\}} + \mathbf{1}_{\{\rho_\ell^\alpha [i, k] < 0\}} \right) \quad (21)$$

with

$$\rho_\ell^\alpha [i, k] := \tilde{\mu}^i - \tilde{\mu}^k + \frac{N_{\ell-1}^\alpha}{\delta N_\ell^\alpha} (\hat{\mu}_{\ell-1}^{\alpha, i} - \hat{\mu}_{\ell-1}^{\alpha, k}) \quad \text{for } \ell \geq 1 \text{ and } i, k \leq n_s.$$

*Proof.* We use Bernstein's inequality, see [7, Theorem 2.1], conditionally to  $\mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta})$ , to deduce that

$$\begin{aligned} & \mathbb{P}^\nu [\hat{\mu}_\ell^{\alpha, k} > \hat{\mu}_\ell^{\alpha, i} | \mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta})] \\ &= \mathbb{P}^\nu [\delta \hat{\mu}_\ell^{\alpha, k} - \tilde{\mu}^k - (\delta \hat{\mu}_\ell^{\alpha, i} - \tilde{\mu}^i) > \frac{N_{\ell-1}^\alpha}{\delta N_\ell^\alpha} (\hat{\mu}_{\ell-1}^{\alpha, i} - \hat{\mu}_{\ell-1}^{\alpha, k}) - (\tilde{\mu}^k - \tilde{\mu}^i) | \mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta})] \\ &\leq e^{-\frac{\delta N_\ell^\alpha (\rho_\ell^\alpha [i, k])^2}{2(\tilde{\sigma}_{ik}^2 + c\rho_\ell^\alpha [i, k])}} \mathbf{1}_{\{\rho_\ell^\alpha [i, k] \geq 0\}} + \mathbf{1}_{\{\rho_\ell^\alpha [i, k] < 0\}}. \end{aligned}$$

$\square$

### 3.2 A generic progressive learning algorithm

Let us now describe how the result of Corollary 3.4 can be turned into a (stochastic) dynamic programming algorithm, in the spirit of Section 2.4, that can be implemented in practice.

By Jensen's inequality, the upper-bound of Corollary 3.4 can be rewritten as

$$\mathbb{E}^\nu \left[ \left| \widehat{\text{ES}} - \widetilde{\text{ES}}^\alpha \right|^p \right] \leq F_p^{\text{ad}}(0, \alpha, \nu)^p \quad (22)$$

where

$$F_p^{\text{ad}}(0, \alpha, \nu) := \mathbb{E}^\nu \left[ \left| \frac{1}{n_w} \sum_{i \in \mathfrak{J}_{L-1}^\alpha} \hat{\mu}_L^{\alpha, i} - \tilde{\mu}^i \right|^p + \sum_{\ell=1}^{L-1} f_p^{\text{ad}}(\ell, \alpha, \tilde{\theta}) \right],$$

to which we can associate the optimal control problem<sup>6</sup>

$$\hat{F}_p^{\text{ad}}(\ell, \alpha, \nu) = \text{ess inf}_{\alpha' \in \mathcal{A}^{\text{ad}}(\ell, \alpha)} F_p^{\text{ad}}(\ell, \alpha', \nu) \quad \text{for } 0 \leq \ell \leq L-1, \nu \in \mathcal{M} \text{ and } \alpha \in \mathcal{A}^{\text{ad}},$$

where

$$\mathcal{A}^{\text{ad}}(\ell, \alpha) := \{\alpha' = (q', N') \in \mathcal{A}^{\text{ad}} : (\alpha'_i)_{0 \leq i \leq \ell} = (\alpha_i)_{0 \leq i \leq \ell} \text{ and } C(q', N') \leq K\}$$

and

$$F_p^{\text{ad}}(\ell, \alpha', \nu) := \mathbb{E}^\nu \left[ \left| \frac{1}{n_w} \sum_{i \in \mathfrak{J}_{L-1}^{\alpha'}} \hat{\mu}_L^{\alpha', i} - \tilde{\mu}^i \right|^p + \mathbf{1}_{\{\ell < L-1\}} \sum_{l=\ell+1}^{L-1} f_p^{\text{ad}}(l, \alpha', \tilde{\theta}) \middle| \mathcal{F}_\ell^{\alpha'} \right].$$

It admits a dynamic programming principle that involves a Bayesian update of the prior law on  $\tilde{\theta}$  at each step of the algorithm, see e.g. [10].

Let us first observe that, from step  $\ell$  on, our bound only involves the components of  $\tilde{\theta}$  associated to the indexes in  $\mathfrak{J}_\ell^\alpha$ . We therefore set

$$\tilde{\theta}_\ell^\alpha = (\tilde{\mu}_\ell^\alpha, \tilde{\Sigma}_\ell^\alpha) := \mathcal{T}_{\mathfrak{J}_{\ell-1}^\alpha}^{\mathfrak{J}_\ell^\alpha}(\tilde{\theta}_{\ell-1}^\alpha), \ell \geq 1, \quad \text{with } \tilde{\theta}_0^\alpha := \tilde{\theta}$$

where, for two subsets  $A' \subset A \subset \llbracket 1, n_s \rrbracket$  and  $(\mu, \Sigma) = ((\mu^i)_{i \in A}, (\Sigma^{ij})_{i, j \in A})$ , we define

$$\mathcal{T}_{A'}^A(\mu, \Sigma) = ((\mu^i)_{i \in A'}, (\Sigma^{ij})_{i, j \in A'}).$$

This means that the update of the prior can be restricted to a reduced number of components of  $\tilde{\theta}$ . This explains why we will concentrate on minimizing this upper-bound rather than directly the left-hand side of (22), which would lead to a very high-dimensional optimal control problem, at each step  $\ell$ . This way, we expect to reduce very significantly the computation cost of the corresponding ‘‘optimal’’ strategy.

In order to make the updating rule explicit, we use the following assumption.

**Assumption 3.** *Given  $\nu_0 \in \mathcal{M}$ , there exists a measure  $m$ , such that, for all  $\alpha \in \mathcal{A}^{\text{ad}}$  and  $1 \leq \ell \leq L$ , the law of  $I_\ell^\alpha := (P_j^i, (i, j) \in \mathfrak{J}_{\ell-1}^\alpha \times \llbracket N_{\ell-1}^\alpha + 1, N_\ell^\alpha \rrbracket)$  given  $\mathcal{F}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta})$  admits  $\nu_0$ -a.s. the density  $g_\ell^\alpha(\cdot, \tilde{\theta}_{\ell-1}^\alpha) := g(\cdot, \mathfrak{J}_{\ell-1}^\alpha, N_{\ell-1}^\alpha, N_\ell^\alpha, \tilde{\theta}_{\ell-1}^\alpha)$  with respect to  $m$ , in which  $g$  is a bounded measurable map<sup>7</sup>, that is continuous in its first argument, uniformly in the other ones. Moreover, for all  $\alpha \in \mathcal{A}^{\text{ad}}$  and  $\ell \leq L$ , the law of  $\tilde{\theta}$  given  $\mathcal{F}_\ell^\alpha$  belongs to  $\mathcal{M}$   $\nu_0$ -a.s.*

Under this assumption, we can compute the law  $\nu_\ell^{\alpha, \ell-1}$  of  $\tilde{\theta}_{\ell-1}^\alpha = \mathcal{T}_{\ell-1}^\alpha(\tilde{\theta})$  given  $\mathcal{F}_\ell^\alpha$  in terms of its counterpart  $\nu_{\ell-1}^\alpha$  given  $\mathcal{F}_{\ell-1}^\alpha$ , in which  $\mathcal{T}_{\ell-1}^\alpha := \mathcal{T}_{\mathfrak{J}_{\ell-2}^\alpha}^{\mathfrak{J}_{\ell-1}^\alpha} \circ \dots \circ \mathcal{T}_{\mathfrak{J}_0^\alpha}^{\mathfrak{J}_1^\alpha}$ . It is given by

$$\nu_\ell^{\alpha, \ell-1} = \mathcal{U}^{\frac{1}{2}}(\ell, \alpha, \nu_{\ell-1}^\alpha)$$

with  $\nu_0^\alpha = \nu$  and

$$\mathcal{U}^{\frac{1}{2}}(\ell, \alpha, \nu_{\ell-1}^\alpha)(A) := \frac{\int_{\mathcal{D}_{\ell-1}^\alpha} g_\ell^\alpha(I_\ell^\alpha, \theta) \mathbf{1}_{\{\theta \in A\}} \nu_{\ell-1}^\alpha(d\theta)}{\int_{\mathcal{D}_{\ell-1}^\alpha} g_\ell^\alpha(I_\ell^\alpha, \theta) \nu_{\ell-1}^\alpha(d\theta)} \quad (23)$$

<sup>6</sup>Only the conditional law given  $\mathcal{F}_\ell^\alpha$  of the components of  $\tilde{\theta}$  corresponding to indexes in  $\mathfrak{J}_\ell^\alpha$  play a role in the definition of  $\hat{F}_p^{\text{ad}}(\ell, \alpha, \nu)$  and  $F_p^{\text{ad}}(\ell, \alpha, \nu)$  below. To avoid introducing new complex notations, we shall indifferently take  $\nu$  or only the conditional law of the corresponding components as an argument, depending on the context.

<sup>7</sup>As for measurability, we identify  $\mathfrak{J}_{\ell-1}^\alpha$  to the element of  $\mathbb{R}^{n_s}$  with  $i$ -th entry given by  $\mathbf{1}_{\{i \in \mathfrak{J}_{\ell-1}^\alpha\}}$ .



for a Borel set  $A$  of  $\mathcal{D}_{\ell-1}^\alpha := \mathcal{T}_{\ell-1}^\alpha(\mathbb{R}^{n_s} \times \mathbb{S}^{n_s})$ . From this, one can deduce the law  $\nu_\ell^\alpha$  of  $\tilde{\theta}_\ell^\alpha = \mathcal{T}_\ell^\alpha(\tilde{\theta})$  given  $\mathcal{F}_\ell^\alpha$ , in the form

$$\nu_\ell^\alpha = \mathcal{U}(\ell, \alpha, \nu_{\ell-1}^\alpha),$$

by simply integrating on the components corresponding to indexes that are not in  $\mathcal{J}_\ell^\alpha$  (meaning that  $\mathcal{U}$  is explicit in terms of  $\mathcal{U}^{\frac{1}{2}}$ ).

We are now in position to state our dynamic programming principle, see e.g. [10]. Again, note that the law of  $f_p^{\text{ad}}(\ell+1, \alpha', \tilde{\theta})$  given  $\mathcal{F}_\ell^{\alpha'}$  depends on  $\tilde{\theta}$  only through  $\tilde{\theta}_\ell^{\alpha'}$ . For ease of notations, we identify all measures to an element of  $\mathcal{M}$  (even if it supported by a space smaller than  $\mathbb{R}^{n_s} \times \mathbb{S}^{n_s}$ ).

**Proposition 3.5.** *Let Assumption 3 hold. Then, for all  $\alpha \in \mathcal{A}^{\text{ad}}$ ,  $0 \leq \ell \leq L-2$  and  $\nu \in \mathcal{M}$ ,*

$$\hat{F}_p^{\text{ad}}(\ell, \alpha, \nu) = \text{ess inf}_{\alpha' \in \mathcal{A}^{\text{ad}}(\ell, \alpha)} \mathbb{E}^\nu[\hat{F}_p^{\text{ad}}(\ell+1, \alpha', \mathcal{U}(\ell+1, \alpha', \nu)) + f_p^{\text{ad}}(\ell+1, \alpha', \tilde{\theta}) | \mathcal{F}_\ell^\alpha].$$

In principle, this dynamic programming algorithm allows one to estimate numerically the optimal policy  $\alpha^*$  in a feed-back form, off-line. Importantly, solving this problem given an initial prior  $\nu_0$  is very different from first estimating the parameter  $\tilde{\theta}$  and then solving the control problem as if  $\tilde{\theta}$  was given. In the first case, we take into account the risk due to the uncertainty on the true value of  $\tilde{\theta}$ , not in the second one.

**Remark 3.6.** *In practice, the algorithm requires estimating and manipulating the law of a high-dimensional parameter, at least at the first steps. But the above can be modified by changing the filtration  $(\mathcal{F}_\ell^\alpha)_{\ell \leq L}$  in  $(\bar{\mathcal{F}}_\ell^\alpha)_{\ell \leq L}$  with  $\bar{\mathcal{F}}_\ell^\alpha = \sigma(\mathbf{1}_{\ell \geq \tau^\alpha} P_j^i, (i, j) \in \mathcal{J}_\ell^\alpha \times \llbracket 1, N_\ell^\alpha \rrbracket)$  with  $\tau^\alpha := \inf\{l \leq L : q_l^\alpha \leq \rho\}$  for some  $\rho > 0$ . In this case, no additional information is considered up to step  $\tau^\alpha$ , the update of the prior only takes place from step  $\tau^\alpha$  on and it only concerns  $\tilde{\theta}_{\tau^\alpha}^\alpha$  whose dimension is controlled by  $\rho$ . As for the first steps of the algorithm, namely before  $\tau^\rho$ , one can replace  $f_p^{\text{ad}}$  by a robust version in the spirit of Section 2.4.*

**Remark 3.7.** *The algorithm also requires knowing the conditional density  $g_\ell^\alpha$ . Although,  $P_{\mathbb{S}}$  can be simulated, its conditional density is not known in general. However, one can use a proxy and/or again modify the flow of information to reduce to a more explicit situation. Let us consider the situation in which  $(\mathcal{F}_\ell^\alpha)_{\ell \leq L}$  is replaced by  $(\bar{\mathcal{F}}_\ell^\alpha)_{\ell \leq L}$  with  $\bar{\mathcal{F}}_\ell^\alpha = \bar{\mathcal{F}}_{\ell-1}^\alpha \vee \sigma(\delta \hat{\mu}_\ell^{\alpha, i}, i \in \mathcal{J}_\ell^\alpha)$  and  $\bar{\mathcal{F}}_0^\alpha$  is trivial. Then, conditionally to  $\bar{\mathcal{F}}_{\ell-1}^\alpha \vee \sigma(\tilde{\theta}_{\ell-1}^\alpha)$ ,  $\sqrt{\delta N_\ell^\alpha} (\tilde{\Sigma}_\ell^\alpha)^{-1} (\delta \hat{\mu}_\ell^\alpha - \tilde{\mu}_\ell^\alpha)$  is asymptotically Gaussian as  $\delta N_\ell^\alpha$  increases to infinity. In practice, we can do as if  $\sqrt{\delta N_\ell^\alpha} (\tilde{\Sigma}_\ell^\alpha)^{-1} (\delta \hat{\mu}_\ell^\alpha - \tilde{\mu}_\ell^\alpha)$  was actually following a standard Gaussian distribution, conditionally to  $\tilde{\theta}_\ell^\alpha$  and  $\mathcal{F}_{\ell-1}^\alpha$ , which provides an explicit formula for the conditional density  $\bar{g}_\ell^\alpha$  of  $\delta \hat{\mu}_\ell^\alpha$  given  $\tilde{\theta}_{\ell-1}^\alpha$  and  $\mathcal{F}_{\ell-1}^\alpha$ , to be plugged into (23). Namely, the updating procedure takes the form*

$$\nu_\ell^\alpha = \check{\mathcal{U}}(\ell, \alpha, \nu_{\ell-1}^\alpha)$$

where  $\check{\mathcal{U}}$  is explicit.

Then, if the initial prior  $\nu_0$  is such that  $(\tilde{\mu}, \tilde{\Sigma})$  is a Normal-inverse-Wishart distribution, all the posterior distribution  $\nu_\ell^\alpha$ ,  $\ell \leq L$ , are such that  $(\tilde{\mu}, \tilde{\Sigma})$  remains in the class of Normal-inverse-Wishart distributions with parameters that can be computed explicitly from our simulations. Namely, if, given  $\bar{\mathcal{F}}_\ell^\alpha$ ,  $\tilde{\Sigma}$  has the distribution<sup>8</sup>  $\mathcal{W}_{i_\ell^\alpha}^{-1}(\Sigma_\ell^\alpha)$  and  $\tilde{\mu}$  has the distribution  $\mathcal{N}(\mathbf{m}_\ell^\alpha, \tilde{\Sigma}/\mathbf{k}_\ell^\alpha)$  given  $\tilde{\Sigma}$ , then the coefficients corresponding to the law given  $\bar{\mathcal{F}}_{\ell+1}^\alpha$  are

$$\begin{cases} \mathbf{i}_{\ell+1}^\alpha = \mathbf{i}_\ell^\alpha + \delta N_{\ell+1}^\alpha, \mathbf{k}_{\ell+1}^\alpha = \mathbf{k}_\ell^\alpha + \delta N_{\ell+1}^\alpha, \mathbf{m}_{\ell+1}^\alpha = \frac{1}{\kappa_\ell^\alpha + \delta N_{\ell+1}^\alpha} \left[ \kappa_\ell^\alpha \mathcal{T}_{\mathcal{J}_\ell^\alpha}^{\mathcal{J}_{\ell+1}^\alpha}(\mathbf{m}_\ell^\alpha) + \delta N_{\ell+1}^\alpha \delta \hat{\mu}_{\ell+1}^\alpha \right] \\ \Sigma_{\ell+1}^\alpha = \mathcal{T}_{\mathcal{J}_\ell^\alpha}^{\mathcal{J}_{\ell+1}^\alpha}(\Sigma_\ell^\alpha) + \sum_{j=N_\ell^\alpha+1}^{N_{\ell+1}^\alpha} (\mathcal{T}_{\ell+1}^\alpha(P_j) - \delta \hat{\mu}_{\ell+1}^\alpha) (\mathcal{T}_{\ell+1}^\alpha(P_j) - \delta \hat{\mu}_{\ell+1}^\alpha)^\top \\ \quad + \frac{\kappa_\ell^\alpha \delta N_{\ell+1}^\alpha}{\kappa_\ell^\alpha + \delta N_{\ell+1}^\alpha} (\mathcal{T}_{\mathcal{J}_\ell^\alpha}^{\mathcal{J}_{\ell+1}^\alpha}(\mathbf{m}_\ell^\alpha) - \delta \hat{\mu}_{\ell+1}^\alpha) (\mathcal{T}_{\mathcal{J}_\ell^\alpha}^{\mathcal{J}_{\ell+1}^\alpha}(\mathbf{m}_\ell^\alpha) - \delta \hat{\mu}_{\ell+1}^\alpha)^\top, \end{cases} \quad (24)$$

see e.g. [18, Section 9]. Later on, we shall write the corresponding law as  $\mathcal{NW}^{-1}(\mathbf{p}_{\ell+1}^\alpha)$  with

$$\mathbf{p}_{\ell+1}^\alpha := (\mathbf{m}_{\ell+1}^\alpha, \mathbf{k}_{\ell+1}^\alpha, \mathbf{i}_{\ell+1}^\alpha, \Sigma_{\ell+1}^\alpha).$$

<sup>8</sup>Hereafter  $\mathcal{W}_i^{-1}(\Sigma)$  stands for the Inverse-Wishart distribution with degree of freedom  $i$  and scale matrix  $\Sigma$ , while  $\mathcal{N}(\mathbf{m}, \Sigma)$  is the Gaussian distribution with mean  $\mathbf{m}$  and covariance matrix  $\Sigma$ .

### 3.3 Example of numerical implementation using neural networks

In this section, we aim at solving the version of the dynamic programming equation of Proposition 3.5, using an initial Normal-inverse-Wishart prior and the approximate updating procedure suggested in Remark 3.7:

$$\check{F}_p^{\text{ad}}(\ell, \alpha, \nu) = \text{ess inf}_{\alpha' \in \mathcal{A}^{\text{ad}}(\ell, \alpha)} \mathbb{E}^\nu [\check{F}_p^{\text{ad}}(\ell + 1, \alpha', \check{\mathcal{U}}(\ell + 1, \alpha', \nu)) + f_p^{\text{ad}}(\ell + 1, \alpha', \tilde{\theta}) | \bar{\mathcal{F}}_\ell^\alpha],$$

with  $\check{\mathcal{U}}$  as in Remark 3.7 and

$$\check{F}_p^{\text{ad}}(L - 1, \alpha, \nu) := \mathbb{E}^\nu \left[ \left| \frac{1}{n_w} \sum_{i \in \mathcal{J}_{L-1}^{\alpha'}} \hat{\mu}_L^{\alpha', i} - \tilde{\mu}^i \right|^p \middle| \bar{\mathcal{F}}_{L-1}^\alpha \right].$$

It would be tempting to use a standard grid-based approximation. However, to turn this problem in a Markovian one, one needs to let the value function at step  $\ell$  depend on  $q_\ell^\alpha$ ,  $N_\ell^\alpha$ ,  $C_\ell^\alpha$ ,  $\hat{\mu}_\ell^\alpha$  and  $\mathbf{p}_\ell^\alpha$ , where  $C_\ell^\alpha$  is the running cost of strategy  $\alpha$  up to level  $\ell$ , defined for  $\ell \neq 0$  by  $C_\ell^\alpha = \sum_{l=0}^{\ell-1} q_l^\alpha \delta N_{l+1}$  and  $C_0^\alpha = 0$ . The dimension is then  $1 + 1 + 1 + q_\ell^\alpha + (1 + q_\ell^\alpha + 1 + (q_\ell^\alpha)^2)$ . Even for  $q_\ell^\alpha = 20$ , the corresponding space is already much too big to construct a reasonable grid on it. We therefore suggest using a neural network approximation. Let us consider a family of bounded continuous functions  $\{\phi_x, x \in X\}$ ,  $X$  being a compact subset of  $\mathbb{R}^{d_X}$  for some  $d_X \geq 1$ , such that, for all  $q, \delta q \leq n_s$  and  $N, \delta N \geq 1$ ,

$$\phi.(\delta q, \delta N, q, N, C, \cdot) : (x, \mu, \mathbf{p}) \in X \times \mathbb{R}^q \times \mathbb{R}^{3+q+q^2} \mapsto \phi_x(\delta q, \delta N, q, N, C, \mu, \mathbf{p}) \in \mathbb{R} \text{ is continuous.}$$

We then fix a family  $\{\alpha^k\}_{k \leq \bar{k}}$  of deterministic paths of  $\mathcal{A}(0)$  and simulate independent copies  $\{\tilde{\theta}^j\}_{j \leq \bar{j}}$  of  $\tilde{\theta}$  according to  $\nu_0$ , a Normal-inverse-Wishart distribution  $\mathcal{NW}^{-1}(\mathbf{p}_0)$ . For each  $j$ , we consider an i.i.d. sequence  $(P_{j'}^{j,1}, \dots, P_{j'}^{j,n_s})_{j' \geq 1}$  in the law  $\mathcal{N}(\tilde{\mu}^j, \tilde{\Sigma}^j)$  with  $\tilde{\theta}^j =: (\tilde{\mu}^j, \tilde{\Sigma}^j)$ . We take these sequences independent and independent of  $\tilde{\theta}$ . For each  $k$  and  $j$ , we denote by  $(\hat{\mu}_\ell^{k,j})_{\ell \leq L}$ ,  $(\hat{\mathbf{p}}_\ell^{k,j})_{\ell \leq L}$  and  $(\mathcal{J}_\ell^{k,j})_{\ell \leq L}$  the paths  $(\hat{\mu}_\ell^{\alpha^k})_{\ell \leq L}$ ,  $(\mathbf{p}_\ell^{\alpha^k})_{\ell \leq L}$  and  $(\mathcal{J}_\ell^{\alpha^k})_{\ell \leq L}$  associated to the  $j$ -th sequence  $(P_{j'}^{j,1}, \dots, P_{j'}^{j,n_s})_{j' \geq 1}$  and the control  $\alpha^k$ . Similarly, we write  $f_p^{\text{ad},k,j}(\ell, \cdot)$  to denote the function  $f_p^{\text{ad}}(\ell, \cdot)$  defined as in (21) but in terms of  $\mathcal{J}_{\ell-1}^{k,j}$  in place of  $\mathcal{J}_{\ell-1}^\alpha$ . Given an integer  $r \geq 1$ , we first compute  $\check{x}_{L-1}$  as the argmin over  $x \in X$  of

$$\sum_{k=1}^{\bar{k}} \sum_{j=1}^{\bar{j}} \left| \mathbb{E}_{L-1}^{\nu^{k,j}} \left[ \frac{1}{n_w} \sum_{i \in \mathcal{J}_{L-1}^{k,j}} (\hat{\mu}_L^{k,j})^i - \tilde{\mu}^i \right]^p - \phi_x(0, \delta N_{L-1}^{\alpha^k}, q_{L-1}^{\alpha^k}, N_{L-1}^{\alpha^k}, C_{L-1}^{\alpha^k}, \hat{\mu}_{L-1}^{k,j}, \mathbf{p}_{L-1}^{k,j}) \right|^r$$

in which  $\mathbb{E}_{L-1}^{\nu^{k,j}}$  means that the expectation is taken only over  $\tilde{\mu}$  according to the law  $\nu_{L-1}^{k,j}$ , i.e.  $\mathcal{NW}^{-1}(\mathbf{p}_{L-1}^{k,j})$ , and  $(\cdot)^i$  means that we take the  $i$ -th component of the vector in the brackets. Then, for any  $\alpha \in \mathcal{A}^{\text{ad}}$ , we set

$$\check{\phi}_{L-1}(q_{L-1}^\alpha, N_{L-1}^\alpha, C_{L-1}^\alpha, \cdot) := \min_{(0, \delta N) \in \mathcal{A}(L-1, \alpha)} \phi_{\check{x}_{L-1}}(0, \delta N, q_{L-1}^\alpha, N_{L-1}^\alpha, C_{L-1}^\alpha, \cdot),$$

where

$$\mathcal{A}(L-1, \alpha) := \{(\delta q, \delta N) \in \{0\} \times \mathbb{N} : C_{L-1}^\alpha + n_w \delta N \leq K\}.$$

Given  $\check{\phi}_{\ell+1}$  for some  $\ell \leq L-2$ , we then compute a minimizer  $\check{x}_\ell \in X$  of

$$\sum_{k=1}^{\bar{k}} \sum_{j=1}^{\bar{j}} \left| \mathbb{E}_\ell^{\nu_\ell^{k,j}} \left[ \check{\phi}_{\ell+1}(q_{\ell+1}^{\alpha^k}, N_{\ell+1}^{\alpha^k}, C_{\ell+1}^{\alpha^k}, \hat{\mu}_{\ell+1}^k, \mathbf{p}_{\ell+1}^k) + f_p^{\text{ad},k,j}(\ell + 1, \alpha^k, \tilde{\theta}) \right] - \phi_x(\delta q_{\ell+1}^{\alpha^k}, \delta N_{\ell+1}^{\alpha^k}, q_\ell^{\alpha^k}, N_\ell^{\alpha^k}, C_\ell^{\alpha^k}, \hat{\mu}_\ell^{k,j}, \mathbf{p}_\ell^{k,j}) \right|^r,$$

where  $\mathbb{E}_\ell^{\nu_\ell^{k,j}}$  means that the expectation is computed over  $(\hat{\mu}_{\ell+1}^{\alpha^k}, \mathbf{p}_{\ell+1}^{\alpha^k}, \tilde{\theta}, \tilde{\mathbf{m}}_\ell^{\alpha^k})$  given  $(\hat{\mu}_\ell^k, \mathbf{p}_\ell^k) = (\hat{\mu}_\ell^{k,j}, \mathbf{p}_\ell^{k,j})$  and using the prior  $\nu_\ell^{k,j}$  on  $\tilde{\theta}$  associated to  $\mathbf{p}_\ell^{k,j}$ . Then, we set

$$\check{\phi}_\ell(q_\ell^\alpha, N_\ell^\alpha, C_\ell^\alpha, \cdot) := \min_{(\delta q, \delta N) \in \mathcal{A}(\ell, \alpha)} \phi_{\check{x}_\ell}(\delta q, \delta N, q_\ell^\alpha, N_\ell^\alpha, C_\ell^\alpha, \cdot),$$

where

$$\begin{aligned} A(\ell, \alpha) &:= \{(\delta q, \delta N) \in \llbracket 0, q_\ell^\alpha - n_w \rrbracket \times \mathbb{N} : C_\ell^\alpha + (q_\ell^\alpha - \delta q)\delta N \leq K\}, \ell < L - 2, \\ A(L - 2, \alpha) &:= \{(\delta q, \delta N) \in \{q_\ell^\alpha - n_w\} \times \mathbb{N} : C_{L-1}^\alpha + (q_{L-2}^\alpha - \delta q)\delta N \leq K\}, \end{aligned}$$

and so on until obtaining  $\phi_0(n_s, 0, 0, 0, \mathbf{p}_0)$ . By continuity of  $\phi(\cdot)$  and compactness of  $X$  and  $A(\ell, \alpha)$  for  $\alpha$  given, the minimum is achieved in the above, possibly not unique, and one can choose a measurable map  $\mathbf{a}_\ell^*$  such that

$$\mathbf{a}_\ell^*(q_\ell^\alpha, N_\ell^\alpha, C_\ell^\alpha, \cdot) \in \arg \min_{(\delta q, \delta N) \in A(\ell, \alpha)} \phi_{\tilde{x}_\ell^N}(\delta q, \delta N, q_\ell^\alpha, N_\ell^\alpha, C_\ell^\alpha, \cdot)$$

for all  $\alpha \in \mathcal{A}^{\text{ad}}$ . Then, given the parameter  $\mathbf{p}_0$  of our initial prior  $\nu_0$ , our estimator of the optimal policy is given by  $\alpha^* = (q^*, N^*)$  defined by induction by

$$(\delta q_1^*, \delta N_1^*) = \mathbf{a}_0^*(n_s, 0, 0, 0, \mathbf{p}_0) \quad \text{and} \quad (\delta q_{\ell+1}^*, \delta N_{\ell+1}^*) = \mathbf{a}_\ell^*(q_\ell^*, N_\ell^*, C_\ell^*, \hat{\mu}_\ell^{\alpha^*}, \mathbf{p}_\ell^{\alpha^*}) \quad \text{for } 0 < \ell < L.$$

Note that the above algorithm for the estimation of the optimal control only requires off-line simulations according to the initial prior  $\nu_0$ . It is certainly costly but does not require to evaluate the real financial book, it can be trained on a proxy, and can be done off-line. It can be combined with the approach of Remark 3.6 to reduce the computation time. In order to prepare for the use of a different initial prior, one can also slightly adapt the above algorithm by considering different initial values of  $\mathbf{p}_0$  (e.g. drawn from another distribution around  $\mathbf{p}_0$ ), so as to estimate  $\check{\phi}_0$  not only at the point  $\mathbf{p}_0$ . When applied to the real book, the update of the prior according to (24) leads to an additional cost that is negligible with respect to the simulation of the book. It leads to the computation of new priors associated to the financial book at hand, that can be used for a new estimation of the optimal policy or simply as a new initial prior for the next computation of the ES.

An example of a simple practical implementation is detailed in Appendix B, while numerical tests are performed in Section 4.

## 4 Numerical Experiments

This section is dedicated to first numerical tests of the different algorithms presented in the previous sections. The settings of the experiments are as follows. We first choose a Normal-inverse-Wishart prior distribution  $\nu_0$  with parameters  $\mathbf{p}_0 := (\mathbf{m}_0, \mathbf{k}_0, \mathbf{i}_0, \Sigma_0)$ . The vector  $\mathbf{m}_0$  is represented on Figure 6 with  $\mathbf{m}_0^i = \mu^i$ ,  $i \leq n_s$ , and  $\Sigma_0 = (\mathbf{i}_0 - n_s - 1)\Sigma$  where  $\Sigma$  has entries

$$\begin{cases} \Sigma^{ii} = 4.84 \times 10^{12} \text{ if } i = j \\ \Sigma^{ij} = \rho \times 4.84 \times 10^{12} \text{ if } i \neq j, \end{cases} \quad (25)$$

with  $\rho = 0.6$  or  $\rho = 0$  depending on the experiments below. As for  $\mathbf{k}_0$  and  $\mathbf{i}_0$ , they are chosen equal to 300, meaning that we have a low confidence in our prior. The computing power is  $K = 10^7$ .

We apply the four different algorithms on 5000 runs (i.e. 5000 independent implementations of each algorithm). For each run, we

- first simulate a value for the real scenarios and covariance matrices  $(\tilde{\mu}, \tilde{\Sigma}) \sim \mathcal{NW}^{-1}(\mathbf{p}_0)$ ,
- apply each of the four algorithms, with simulated prices following  $P_{|s} \sim \mathcal{N}(\tilde{\mu}, \tilde{\Sigma})$ ,
- for each algorithm, we measure the relative error  $\frac{\widehat{ES} - \widetilde{ES}}{\widetilde{ES}}$  and the error  $\widehat{ES} - \widetilde{ES}$ , where  $\widetilde{ES} = \frac{1}{n_w} \sum_{i=1}^{n_w} \tilde{\mu}^{\tilde{m}(i)}$ .

The four algorithms that we compare are:

- A Uniform Pricing Algorithm: All the scenarios are priced with  $K/n_s$  Monte Carlo simulations, and the estimator  $\widehat{ES}$  is the average of the  $n_w = 6$  worst scenarios. This is the most naive method, with only one step and where all scenarios are priced with an equal number of Monte Carlo simulations. It serves as a benchmark.

- The Heuristic Algorithm: We use the 2-levels strategy described in Section 2.5 with the book sample parameters of Table 1 and the computation parameters of Table 2. We do not evaluate the constant  $c$  of Assumption 1 but simply set it to 0, see Remark 2.4. The optimal strategy is given by  $(q_0, q_1, N_1, N_2) = (253, 68, 17\,297, 100\,000)$ .
- The Deterministic Algorithm: We run the deterministic algorithm of Section 2.4 optimized with  $\mu = \mathbf{m}_0$  as the values of the scenarios,  $\Sigma$  with  $\rho = 0.6$  as the covariance matrix and  $L = 4$ . Note that using the real mean parameter as an entry for optimization is quite favorable for this algorithm, although the “true” parameter of each run will actually deviate from this mean value. This gives us the strategy  $(q_0, q_1, q_2, q_3, N_0, N_1, N_2, N_3, N_4) = (253, 35, 10, 6, 0, 6\,000, 44\,000, 44\,000, 1\,235\,666)$ , which we apply to each run.
- The Adaptative Algorithm: We do the training part of the adaptative algorithm using our prior  $\mathbf{p}_0 := (\mathbf{m}_0, \mathbf{k}_0, \mathbf{i}_0, \Sigma_0)$ , with  $\rho = 0.6$ , as parameters and  $L = 4$ . We use a very simple one hidden-layer neural network. It could certainly be improved by using a more sophisticated multi-layers neural network, but this version will be enough for our discussion. Details on the implementation are given in the Appendix B. Once this is done, we apply the optimal adaptative strategy on each run.

#### 4.1 Positively correlated scenarios $\rho = 0.6$

In this first experiment, the simulated runs use the values  $\rho = 0.6$  and  $\mathbf{i}_0 = \mathbf{k}_0 = 300$ .

To get an idea of how much noise is added to the average scenario values in our simulations, we plot in Figure 7 the prior value  $\mu_0^i$  for each scenario of index  $i \leq n_s$  (this is the line) and the first 20  $\tilde{\mu}_j^i$  out of the 5 000 runs for each scenario (these are the points).

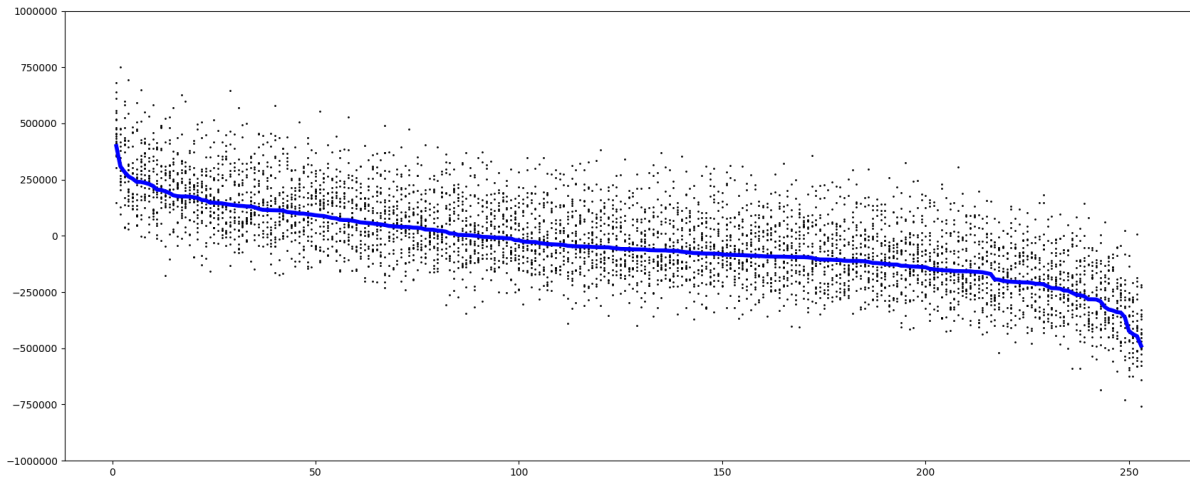


Figure 7: True value of  $\mu_0$  and simulations of  $\tilde{\mu}$

For the adaptative algorithm, the three mostly used strategies are:

- $(q_0, q_1, q_2, q_3, N_1, N_2, N_3, N_4) = (253, 40, 25, 6, 8\,399, 97\,995, 172\,504, 577\,252)$
- $(q_0, q_1, q_2, q_3, N_1, N_2, N_3, N_4) = (253, 40, 30, 6, 8\,399, 99\,733, 148\,560, 608\,040)$
- $(q_0, q_1, q_2, q_3, N_1, N_2, N_3, N_4) = (253, 40, 30, 6, 8\,399, 75\,033, 123\,860, 748\,007)$

Compared to the deterministic algorithm, we see that the adaptative one uses much less Monte Carlo simulations at the final steps and focuses more on the intermediate steps to select the worst scenarios. The deterministic algorithm is also more aggressive in the choice of  $q_1$  and  $q_2$ . This can be easily explained by the fact that the latter believes that the real distribution is not far from the solid curve on Figure 7 (up to standard deviation) while the adaptative one only knows a much more diffuse distribution corresponding

to the cloud of points of Figure 7 since his level of uncertainty is quite high for our choice  $i_0 = k_0 = 300$ .

On Figures 8-11, we plot the histograms of the relative errors. We see that the distribution is tightest for the deterministic algorithm, followed quite closely by the adaptative algorithm. Both of them perform very well. As expected, the uniform algorithm is very poor. Note that the heuristic one already very significantly improves the uniform algorithm, although it does not reach the precision of the two most sophisticated algorithms (without surprise). Because of the huge uncertainty mentioned above, the adaptative algorithm is rather conservative while the deterministic algorithm makes full profit of essentially knowing the correct distribution, and performs better. We will see in our second experiment that things will change when we will deviate from the parameters used for optimizing the deterministic algorithm (by simply passing from  $\rho = 0.6$  to  $\rho = 0$  in the simulated runs).

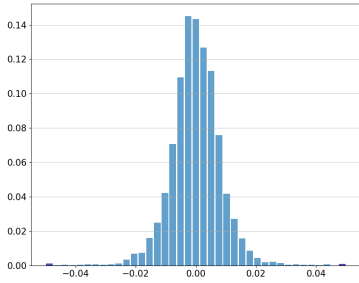


Figure 8: Relative Error for Adaptive Algorithm

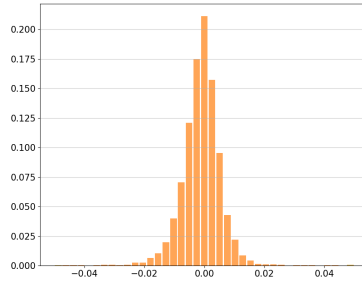


Figure 9: Relative Error for Determinist Algorithm

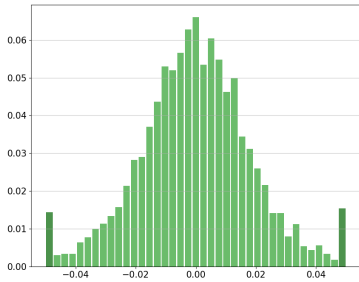


Figure 10: Relative Error for Heuristic Algorithm

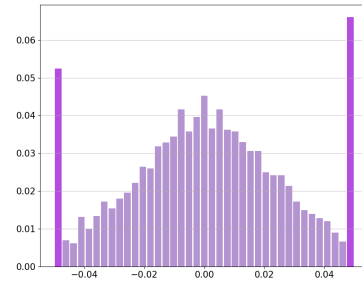


Figure 11: Relative Error for Uniform Algorithm

In Table 3, we provide the  $\mathbb{L}^1$  and relative errors (with standard deviations), the  $\mathbb{L}^2$  error and the number of correct selections, that is the number of runs for which a given algorithm has chosen the correct worst 6 scenarios. In terms of  $\mathbb{L}^1$  or  $\mathbb{L}^2$  error, the relative performance of the algorithms is as above. However, if we look at the number of correct selections, we see that the adaptive algorithm performs better than the other 3 algorithms. Again, by comparing the strategies of the deterministic and the adaptive algorithms, we see that those of the adaptive algorithm are more conservative on the ranking and filtering part versus the final pricing as it puts relatively more Monte Carlo simulations to detect the correct scenarios and relatively less for their estimation.

Algorithm	$\mathbb{L}^1$ Err.	$\mathbb{L}^1$ Err. Std	Rel. Err. (%)	Rel. Err. Std (%)	$\mathbb{L}^2$ Err.	Correct Selections
Ad. Alg.	1 891	20.4	0.623	0.00886	2377	4247
Det. Alg.	1 411	16.1	0.465	0.00693	1813	3499
Heur. Alg.	4 562	50.2	1.49	0.0234	5779	4054
Unif. Alg.	7 269	81.6	2.38	0.0348	9279	3500

Table 3: Errors for  $\rho = 0.6$

In Figures 12, we plot the function  $x \mapsto \mathbb{P}[X > 5000 - x]$  where  $X$  is the absolute error of the algorithm on a run.

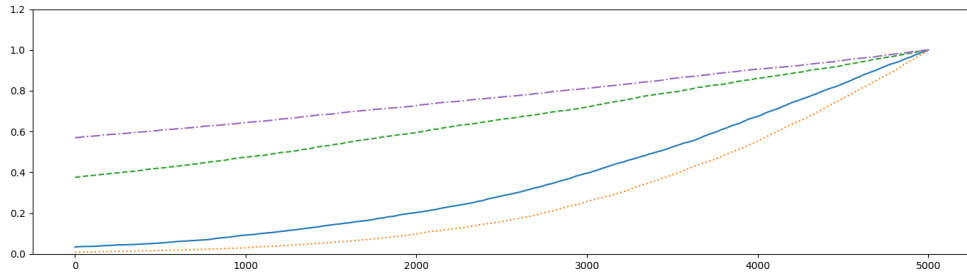


Figure 12: Tail Distribution of the errors. First top lines: Uniform and Heuristic algorithms, respectively. Solid line: Adaptive algorithm. Dotted line: Deterministic algorithm.

In Figure 13, we provide, for the first 4 runs, the values and real ranks of the 6 worst scenarios selected by each algorithm. The numbers displayed are the true ranks of the selected scenarios given by  $\tilde{\mu}$  and their y-coordinate is the value obtained when running the algorithm. “Real” is the real values as sampled.

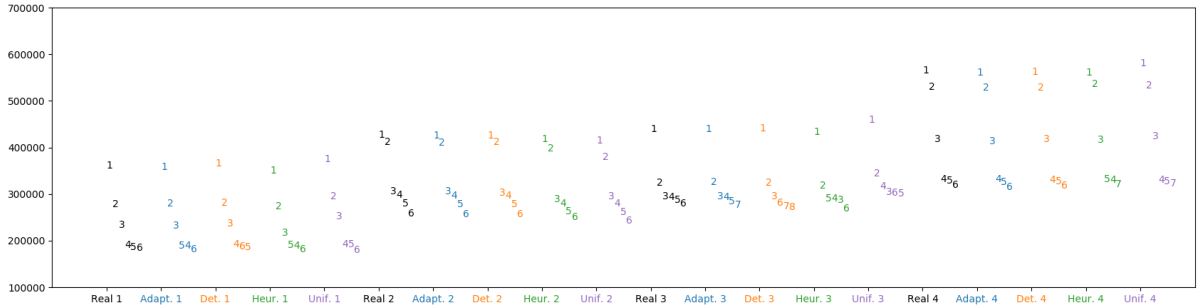


Figure 13: Worst Scenarios Ranks and Values

## 4.2 Uncorrelated scenarios $\rho = 0$

We now do the numerical test with  $\rho = 0$  as the true correlation. The deterministic and adaptive algorithm are still trained with  $\rho = 0.6$ , but  $P|_s$  is simulated using  $\rho = 0$ .

On Figures 14-17, we show the histograms of the relative errors. We see that the distribution of the relative errors is now tightest for the adaptive method, followed by the deterministic method, then by the heuristic and the uniform methods. Furthermore, we see that the distribution corresponding to the deterministic method is significantly biased to the left. This is actually true for all algorithms, but at a less significant level. This suggests that we now have a large part of the error that does not come from the final pricing error, but from errors in the selection of scenarios.

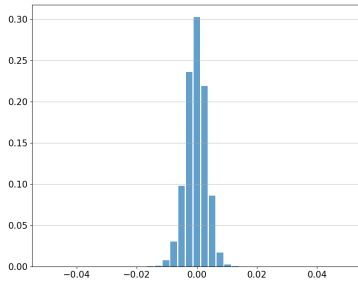


Figure 14: Relative Error for Adaptive Algorithm

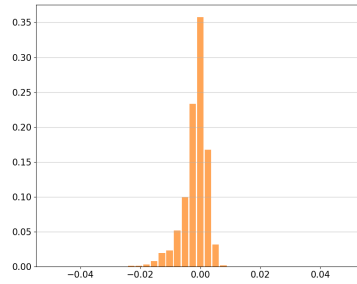


Figure 15: Relative Error for Determinist Algorithm

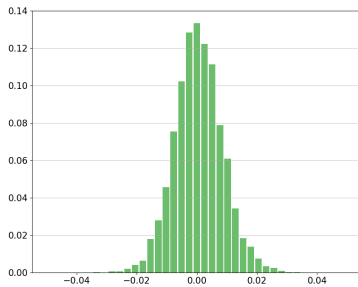


Figure 16: Relative Error for Heuristic Algorithm

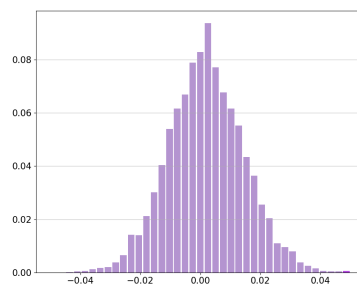


Figure 17: Relative Error for Uniform Algorithm

In Table 4, we provide the  $\mathbb{L}^1$  and relative errors (with standard deviations), the  $\mathbb{L}^2$  error and the number of correct selections for the 4 algorithms. For all algorithms, compared to the case  $\rho = 0.6$ , we see that we have simultaneously a lower number of correct selections of scenarios (which we could expect to increase the errors) and a lower  $\mathbb{L}^1$  error. This surprising result is explained by the fact that lowering the correlation has two effects. The filtering and ranking part of the algorithm becomes harder, as can be seen from Corollary 2.3. This explains why the number of correct selections becomes lower. However, we compute at the end an average over the  $n_w$  worst scenarios and the error on this average is lower when the pricings are uncorrelated compared to the case where they exhibit a positive correlation.

The adaptive algorithm has now simultaneously the lowest  $\mathbb{L}^1$  and  $\mathbb{L}^2$  errors, as well as the highest number of correct selections. We see that it is especially good in  $\mathbb{L}^2$  error, so we expect it to present a very low number of large errors. As, by construction, it has been trained to detect misspecifications of the parameters, it now has a clear advantage on the deterministic algorithm which does not see it. This results in an improvement of almost 20% of the  $\mathbb{L}^2$  error.

Following the above reasoning, we understand that, compared to the previous experiment, the final pricing error now plays a smaller role and the ranking and selection error a bigger role, which explains why the histogram of the errors for the deterministic algorithm is strongly biased to the left, as it now incorrectly selects scenarios more often.

Algorithm	$\mathbb{L}^1$ Err.	$\mathbb{L}^1$ Err. Std	Rel. Err. (%)	Rel. Err. Std (%)	$\mathbb{L}^2$ Err.	Correct Selections
Ad. Alg.	1 083	11.8	0.27	0.00294	1 366	3 930
Det. Alg.	1 175	17.5	0.293	0.00448	1 705	3 202
Heur. Alg.	2 547	28.33	0.628	0.00700	3 240	3 753
Unif. Alg.	4 062	44.7	1.00	0.0111	5 147	3 102

Table 4: Errors for  $\rho = 0$

In Figures 18, we plot the function  $x \mapsto \mathbb{P}[X > 5000 - x]$  where  $X$  is the absolute error of the algorithm on a run. As was suggested by the  $\mathbb{L}^2$  errors of Table 4, we see that the tail distribution of errors is lowest for the adaptative algorithm, followed by the deterministic algorithm (for big errors), and then by the heuristic and uniform algorithms.

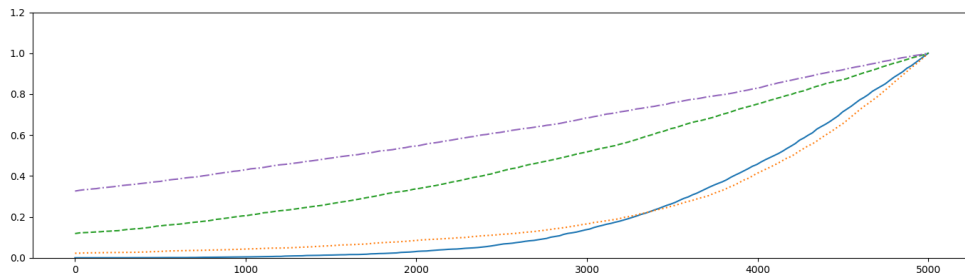


Figure 18: Tail Distribution of the errors. First top lines: Uniform and Heuristic algorithms, respectively. Solid line: Adaptative algorithm. Dotted line: Determinist algorithm

## 5 Conclusion

We propose in this paper different algorithms for the computation of the expected shortfall based on given historical scenarios. All are multi-steps algorithms that use Monte Carlo simulations to reduce the number of historical scenarios that potentially belong to the set of worst scenarios. We provide explicit error bounds and we test them on simulated data deviating from the true values of the historical impacts used for computing the associated optimal strategies. The first algorithm is a very easy to implement 2-steps algorithm that already provides relatively small errors on our numerical tests. A four step deterministic dynamic programming algorithm performs very well when real datas are not far from the parameters used in the optimization procedure. It seems even to be quite robust, as shown by our numerical test in the case where the true correlation parameter is not the one used for computing the optimal policy. Finally, we propose an adaptative algorithm that aims at learning the true value of the parameters at the different steps of the algorithm. Our first numerical tests suggest that it is more conservative than the deterministic one, but probably more robust to parameters misspecifications, as expected. The version we use is built on a very simple one hidden layer neural network and can certainly be considerably improved for industrial purposes.

## References

- [1] Carlo Acerbi and Dirk Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.
- [2] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.
- [3] Raghu Raj Bahadur and Herbert Robbins. The problem of the greater mean. *The Annals of Mathematical Statistics*, pages 469–487, 1950.



- [4] Basel Committee on Banking Supervision. *Minimum capital requirements for market risk*. 2016.
- [5] Robert E Bechhofer. A single-sample multiple decision procedure for ranking means of normal populations with known variances. *The Annals of Mathematical Statistics*, pages 16–39, 1954.
- [6] Robert E Bechhofer, Charles W Dunnett, and Milton Sobel. A tow-sample multiple decision procedure for ranking means of normal populations with a common unknown variance. *Biometrika*, 41(1-2):170–176, 1954.
- [7] Bernard Bercu, Bernard Delyon, and Emmanuel Rio. *Concentration inequalities for sums and martingales*. Springer, 2015.
- [8] Mark Broadie, Yiping Du, and Ciamac C Moallemi. Efficient risk estimation via nested sequential simulation. *Management Science*, 57(6):1172–1194, 2011.
- [9] Simon A Broda, Jochen Krause, and Marc S Paoletta. Approximating expected shortfall for heavy-tailed distributions. *Econometrics and statistics*, 8:184–203, 2018.
- [10] David Easley and Nicholas M Kiefer. Controlling a stochastic process with unknown parameters. *Econometrica: Journal of the Econometric Society*, pages 1045–1064, 1988.
- [11] Robert J Elliott and Hong Miao. Var and expected shortfall: a non-normal regime switching framework. *Quantitative Finance*, 9(6):747–755, 2009.
- [12] Christian Francq and Jean-Michel Zakoïan. Multi-level conditional var estimation in dynamic models. In *Modeling Dependence in Econometrics*, pages 3–19. Springer, 2014.
- [13] Michael B Gordy and Sandeep Juneja. Nested simulation in portfolio risk measurement. *Management Science*, 56(10):1833–1848, 2010.
- [14] Shanti S Gupta and S Panchapakesan. Sequential ranking and selection procedures. *Handbook of sequential analysis*, pages 363–380, 1991.
- [15] Lennart Hoogerheide and Herman K van Dijk. Bayesian forecasting of value at risk and expected shortfall using adaptive importance sampling. *International Journal of Forecasting*, 26(2):231–247, 2010.
- [16] Jochen Krause and Marc S Paoletta. A fast, accurate method for value-at-risk and expected shortfall. *Econometrics*, 2(2):98–122, 2014.
- [17] Ming Liu and Jeremy Staum. Stochastic kriging for efficient nested simulation of expected shortfall. *Journal of Risk*, 12(3):3, 2010.
- [18] Kevin P Murphy. Conjugate bayesian analysis of the gaussian distribution. [cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf](http://cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf).
- [19] Saralees Nadarajah, Bo Zhang, and Stephen Chan. Estimation methods for expected shortfall. *Quantitative Finance*, 14(2):271–291, 2014.
- [20] Luis Ortiz-Gracia and Cornelis W Oosterlee. Efficient var and expected shortfall computations for nonlinear portfolios within the delta-gamma approach. *Applied Mathematics and Computation*, 244:16–31, 2014.
- [21] Franco Peracchi and Andrei V Tanase. On estimating the conditional expected shortfall. *Applied Stochastic Models in Business and Industry*, 24(5):471–493, 2008.
- [22] Jimmy Risk and Michael Ludkovski. Sequential design and spatial modeling for portfolio tail risk measurement. *SIAM Journal on Financial Mathematics*, 9(4):1137–1174, 2018.
- [23] R Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471, 2002.

- [24] Jules Sadefo Kamdem. Value-at-risk and expected shortfall for linear portfolios with elliptically distributed risk factors. *International Journal of Theoretical and Applied Finance*, 8(05):537–551, 2005.
- [25] Jean-Guy Simonato. The performance of johnson distributions for computing value at risk and expected shortfall. *The Journal of Derivatives*, 19(1):7–24, 2011.
- [26] Keming Yu, A Allay, Shanchao Yang, and D Hand. Kernel quantile-based estimation of expected shortfall. 2010.
- [27] Meng-Lan Yueh and Mark CW Wong. Analytical var and expected shortfall for quadratic portfolios. *The Journal of Derivatives*, 17(3):33–44, 2010.

## A Proxy of the optimal strategy for the heuristic (19)

In the case  $p = 1$ , (19) can even be further simplified by using the upper-bound

$$\tilde{h}_0^1(q_1) \leq \max\{\tilde{h}_1(q_1); \tilde{h}_2(q_1)\} \quad (26)$$

where

$$\begin{aligned} \tilde{h}_1(q_1) &:= n_s (q_1 + 1 - n_w) \delta_0 \exp\left(-\frac{(K - q_1 N_2)(q_1 + 1 - n_w)\delta_0}{4n_s c}\right) \\ \tilde{h}_2(q_1) &:= n_s (n_s - n_w) \delta_0 \exp\left(-\frac{(K - q_1 N_2)((q_1 + 1 - n_w)\delta_0)^2}{4n_s \sigma^2}\right). \end{aligned}$$

The right-hand side of (26) is now tractable for minimization. Given,

$$\begin{cases} \Delta := (K - (n_w - 1)N_2)^2 - \frac{32n_s N_2 c}{\delta_0} \\ B := \frac{\sigma^2}{c\delta_0} + n_w - 1 \\ q_1^{2,*} := \max\left(\frac{n_w - 1}{3} + \frac{2K}{3N_2}, n_w\right) \\ q_1^{1,1,*} := \max\left(\frac{3(n_w - 1)}{4} + \frac{K - \sqrt{\Delta}}{4N_2}, n_w\right) \\ q_1^{1,2,*} := \max\left(\frac{3(n_w - 1)}{4} + \frac{K + \sqrt{\Delta}}{4N_2}, n_w\right) \end{cases} \quad (27)$$

the optimal policy  $q_1^h$  is defined by the following table<sup>9</sup>:

Cond. on $B$	Cond. $\Delta$	Cond. $q_1^{2,*}$	Cond. $q_1^{1,1,*}$	Cond. $q_1^{1,2,*}$	Choice of $q_1^h$
$\geq n_s$					$q_1^h := q_1^{2,*}$
$\leq n_w$	$> 0$				$q_1^h := \operatorname{argmin}_{q_1 \in \{n_w, q_1^{1,2,*}\}} h_0^1(q_1)$
$\leq n_w$	$\leq 0$				$q_1^h := n_w$
$n_w < \cdot < n_s$	$> 0$	$\leq B$	$\leq B$	$\leq B$	$q_1^h := \operatorname{argmin}_{q_1 \in \{q_1^{2,*}, B\}} h_0^1(q_1)$
$n_w < \cdot < n_s$	$> 0$	$\leq B$	$\leq B$	$\geq B$	$q_1^h := \operatorname{argmin}_{q_1 \in \{q_1^{2,*}, q_1^{1,2,*}\}} h_0^1(q_1)$
$n_w < \cdot < n_s$	$> 0$	$\leq B$	$\geq B$	$\geq B$	$q_1^h := \operatorname{argmin}_{q_1 \in \{q_1^{2,*}, B, q_1^{1,2,*}\}} h_0^1(q_1)$
$n_w < \cdot < n_s$	$> 0$	$\geq B$	$\leq B$	$\leq B$	$q_1^h := B$
$n_w < \cdot < n_s$	$> 0$	$\geq B$		$\geq B$	$q_1^h := \operatorname{argmin}_{q_1 \in \{B, q_1^{1,2,*}\}} h_0^1(q_1)$
$n_w < \cdot < n_s$	$\leq 0$	$\leq B$			$q_1^h := \operatorname{argmin}_{q_1 \in \{q_1^{2,*}, B\}} h_0^1(q_1)$
$n_w < \cdot < n_s$	$\leq 0$	$\geq B$			$q_1^h := B$

Table 5: Optimal  $q_1^h$  for  $\tilde{h}_0^1$ .

For simplicity, let us consider the case  $c = 0$ , see Remark 2.4.

On Figure 19, the square is  $q_1^{1,2,*} = 52.41$ , the circle is  $q_1^{2,*} = 68.33$  and the cross is the real optimum  $q_1^* = 71$  of  $h_0^1$ , for the parameters of Tables 1 and 2. We see that we actually almost reach the correct minimum. It corresponds (up to rounding) to  $N_1^{1,*} = 23723$ ,  $N_1^{2,*} = 17148$ ,  $N_1^* = 15934$ .

<sup>9</sup>We optimize here over real positive numbers.

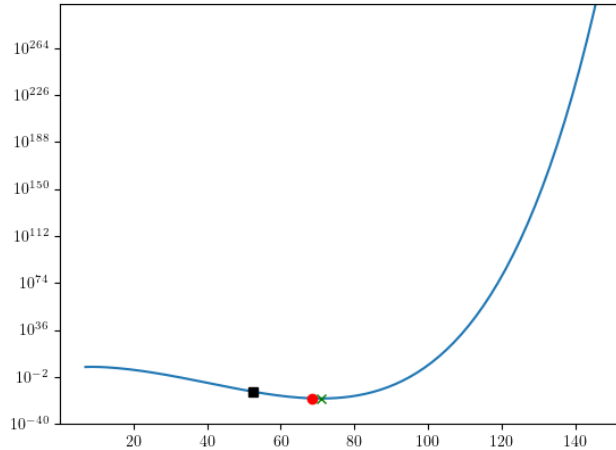


Figure 19: Square:  $h_0^1(q_1^{1,2,*})$ . Circle:  $h_0^1(q_1^{2,*})$ . Cross:  $h_0^1(q_1^*)$ .

Using the same set of parameters, we plot on Figure 20 the two functions  $h_0^1$  and  $\tilde{h}_1$ . Although these two functions have values of different orders of magnitude, their shapes are quite close, which explains why we manage to obtain a relatively good approximation for the minimizer.

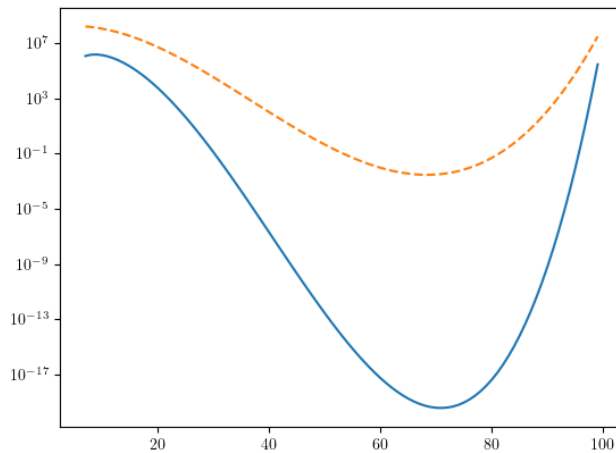


Figure 20: Solid line :  $h_0^1$ . Dashed line :  $\tilde{h}_1$ .

## B Precise implementation of the neural network algorithm

In this Appendix, we describe in more details how the neural network approximation of the optimal policy of the adaptative algorithm is constructed. All the parameters values are given in Tables 6, 7 and 8 below.

### B.1 Initialization

- In practice, the neural network input's size depends on the window size  $q$ . Therefore, we need to train different neural networks for each window size. In order to get enough points to train each of

these neural networks, we have chosen the grid

$$q_g = [6, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 253]$$

of possible values for  $q$ .

- We simulate independent copies  $\{\tilde{\theta}^j\}_{j \leq \text{j\_bar}} = \{(\tilde{\mu}^j, \tilde{\Sigma}^j)\}_{j \leq \text{j\_bar}}$  of  $\tilde{\theta}$ , where  $\text{j\_bar}$  is given in Table 8. For each  $1 \leq j \leq \text{j\_bar}$ ,  $\tilde{\Sigma}^j$  is an inverse-Wishart of parameters  $\mathbf{i}_0, \Sigma_0$ , and  $\tilde{\mu}^j$  is a Gaussian random vector of mean  $\mathbf{m}_0$  and covariance matrix  $\tilde{\Sigma}^j/\mathbf{k}_0$ . The parameters  $\mathbf{i}_0, \mathbf{k}_0$  and  $\Sigma_0$  are defined in Table 8 and (25), while  $\mathbf{m}_0^i = \mu^i$ ,  $i \leq n_s$ , with the  $\mu^i$ 's of Figure 6.

## B.2 Strategy Generation

To generate the deterministic strategies  $(\alpha^k)_{k \leq \text{k\_bar}}$ , where  $\text{k\_bar}$  is given in Table 8, we proceed as follows.

- For each  $1 \leq k \leq \text{k\_bar}$ , we simulate  $L + 1$  uniform random variables  $(U_n)_{n=0}^L$  between 0 and 1. We sort them in increasing order  $(U_{s(n)})_{n=0}^L$  and define a cost  $K_\ell := K(U_{s(\ell)} - U_{s(\ell-1)})$  when  $1 \leq \ell \leq L - 1$ , and  $K_L = K(U_{s(0)} + 1 - U_{s(L-1)})$ . The idea is that we select  $L + 1$  points randomly on a circle of total length  $K$ : we choose one of these points, and starting from it, the computational power that we will use at each level  $1 \leq \ell \leq L - 1$  is the length of the arc between the previous and the next point. For the last step, we take  $K$  times the length between the points  $L - 1$  and 0, so as to put, in average, twice more computational power on this last step.
- Once we have the computational cost for each step, we can choose the  $q_\ell$  for each strategy, so that we can deduce  $\delta N_{\ell+1} := K_\ell/q_\ell$ . For  $\ell = 0$ , we choose  $\text{q\_index}_0 = 18$ , where 18 is the number of terms in the grid  $q_g$ , which therefore gives  $q_0 = q_g[\text{q\_index}_0] = n_s$ . For  $\ell = L - 1$ , we choose  $\text{q\_index}_{L-1} = 0$ , that is,  $q_{L-1} = n_w$ . For  $1 \leq \ell \leq L - 2$ , we choose  $\text{q\_index}_\ell$  as a random integer between  $[L - \ell, \text{q\_index}_{\ell-1} - 1]$ . The choice of  $q_\ell$  is then  $q_\ell = q_g[\text{q\_index}_\ell]$ . We check that the sequence  $(N_\ell)_{1 \leq \ell \leq L}$  is non-decreasing. If this is the case, we keep it, if not, we reject it and do another run.

## B.3 Forward Pass

The next step is to generate all prices and execute for each  $k$  and each  $j$  the strategy  $k$ .

- For  $1 \leq j \leq \text{j\_bar}$ ,  $1 \leq k \leq \text{k\_bar}$  and  $1 \leq \ell \leq L$ , we simulate  $\delta N_\ell^k$  Gaussian variables  $(P_{j'}^{j,1}, \dots, P_{j'}^{j,n_s})_{j'=N_{\ell-1}^k}^{N_\ell^k}$  of mean  $\tilde{\mu}^j$  and covariance matrix  $\tilde{\Sigma}^j$  (independently across  $j$  and  $k$ ).
- We then update  $\hat{\mu}_\ell^{k,j}, \mathbf{m}_\ell^{k,j}, \mathbf{i}_\ell^{k,j}, \mathbf{k}_\ell^{k,j}, \Sigma_\ell^{k,j}$  accordingly, recall (24).
- Updating  $\Sigma_\ell^{k,j}$  from level  $\ell - 1$  to level  $\ell$  can use a lot of memory. Indeed,  $\sum_{j=N_{\ell-1}^k+1}^{N_\ell^k+1} (\mathcal{T}_{\ell+1}^\alpha(P_j) - \delta \hat{\mu}_{\ell+1}^\alpha)(\mathcal{T}_{\ell+1}^\alpha(P_j) - \delta \hat{\mu}_{\ell+1}^\alpha)^\top$  consists in  $\delta N_{\ell+1}^\alpha \times |q_{\ell+1}|^2$  terms, which can quickly exceed memory limits. Therefore, we do the sum with only `N_memory_new_pricings_opt` terms at a time, see Table 8 below.

## B.4 Computation of `f_precompute`, `running_costs` and `admissible_sets`

- In order to speed up the computation time, we now precompute several values that will be used many times afterwards. First, we compute `f_precompute`( $\ell, k, j$ ) as  $f_1^{\text{ad}}(\ell, \cdot)$  at the point corresponding to  $(k, j)$  except that, in the definition of  $f_1^{\text{ad}}(\ell, \cdot)$ , we replace the random permutation  $\tilde{\mathbf{m}}$  by its estimation from the previous step,  $\tilde{\mu}$  by its average under the posterior distribution at  $\ell + 1$ , and  $\tilde{\sigma}$  by its estimation at step  $\ell + 1$ .
- We compute `running_cost`( $\ell, k$ ) :=  $C_\ell^{\alpha^k}$  of each  $k$  at step  $\ell$ .

- We restrict the set of possible actions at step  $\ell$ , given that we have followed the strategy  $k$  so far, to `admissible_sets( $\ell, k$ )` defined as the collection of  $\{(\delta q_{\ell+1}^{k'}, \delta N_{\ell+1}^{k'})\}$ ,  $k' \leq \text{k\_bar}$ , such that

$$q_{\ell}^k + \delta q_{\ell+1}^{k'} \in q_g, N_{\ell+1}^{k'} > N_{\ell}^k, \text{running\_cost}(\ell, k) + q_{\ell}^k \delta N_{\ell+1}^{k'} \leq \max_{1 \leq k'' \leq \bar{k}} \text{running\_cost}(\ell + 1, k'').$$

The last condition avoids inducing a strategy with a running cost that is not present in our data set, when doing the one step optimization.

## B.5 Computation of the final expectations

We first pre-compute the quantities

$$\mathbb{E}_L^{\nu^{k,j}} \left[ \left| \frac{1}{n_w} \sum_{i \in \mathcal{J}_{L-1}^{k,j}} (\hat{\mu}_L^{k,j})^i - \tilde{\mu}^i \right| \right]$$

by Monte Carlo using  $N_e$  simulations. As the simulation of an inverse-Wishart random variable is significantly slower than the simulation of a Gaussian random variable, we only simulate 1 inverse-Wishart for  $N_p$  Gaussians. The values of  $N_e$  and  $N_p$  are given by `N_mu_tildes_simulated` and `N_wishart_proportion` of Table 8. The estimation is called `expectation $_L^{k,j}$` .

## B.6 Training of the neural network at level $L$

- We use a neural network with one inner layer with 256 neurons and 1 output layer with 1 neuron to fit  $(\text{expectation}_L^{k,j})_{j \leq \text{j\_bar}, k \leq \text{k\_bar}}$ . The neurons of the inner layer consist of the composition of the softplus function with an affine transformation of the inputs.
- We initialize the neural network parameters using a Xavier initialization. We then train the neural network by selecting a random new batch every `N_batch_change_proportion`. This random new batch is composed of the samples indexed by  $1 \leq \mathbf{m}_a(j) \leq \text{j\_batch}$  and strategies indexed by  $1 \leq \mathbf{m}_b(k) \leq \text{k\_batch}$ , where  $\mathbf{m}_a$  and  $\mathbf{m}_b$  are uniform random permutations of  $\llbracket 1, \text{j\_bar} \rrbracket$  and  $\llbracket 1, \text{k\_bar} \rrbracket$ . For each batch, the algorithm used for the training is the standard gradient descent of Tensorflow. We do `N_iter` training steps in total. The learning rate used is given in Table 7. In order to bring the input values of the parameters close to 0 and 1, we renormalize them according to the values in Table 6.

## B.7 Computation of the expectations at level $L - 1$

We now estimate

$$\mathbb{E}_{L-1}^{\nu^{k,j}} \left[ \check{\phi}_L(q_L^{\alpha^k}, N_L^{\alpha^k}, C_L^{\alpha^k}, \hat{\mu}_L^k, \mathbf{P}_L^k) \right]$$

where  $\check{\phi}_L$  is the fit of  $(\text{expectation}_L^{k,j})_{j \leq \text{j\_bar}, k \leq \text{k\_bar}}$  from the previous step. The most cpu demanding part is no more the simulation of the inverse-Wisharts, but the updates of the parameters of the inverse-Wishart. Therefore, we simulate as many Gaussian random variables as inverse-Wishart random variables, with  $N_e$  given by `N_mu_tildes_simulated_non_final_level` of Table 8.

For our computations, we need to update  $\Sigma_{L-1}^{k,j}$  to the corresponding posterior parameter according to (24). This can however lead to an enormous amount of multiplications and additions. Therefore, instead of updating the whole matrix, we only update the diagonal terms according to (24) and estimate non diagonal terms by keeping the correlation terms equal to the ones of  $\Sigma_{L-1}^{k,j}$ . This enables us to approximately gain a factor of  $q_L^k$  in speed in this critical step.

## B.8 Training of the neural network at level $L - 1$

- To fit the expectation of the previous step, we use a neural network with the same structure as in level  $L$ , with the same cost function.

- The initialization, choice of batches, and training of the neural network are the same as for the level  $L$ . The number of iteration, learning rate, and renormalization constants are given in Tables 6, 8 and 7.
- We take  $\text{j\_batch} = \min(\text{j\_batch\_size}, \text{j\_bar})$  and  $\text{k\_batch} = \min(\text{k\_batch\_size}, \text{k\_bar})$ , where  $\text{j\_batch\_size}$  and  $\text{k\_batch\_size}$  are defined in Table 8.

### B.9 Computation of the expectations at levels $0 \leq \ell \leq L - 2$

- The expectations at step  $\ell$  are computed by Monte Carlo after replacing the value function at step  $\ell + 1$  by its neural network approximation, and  $f_1^{\text{ad}}(\ell, \cdot)$  by  $\text{f\_precompute}(\ell, \cdot)$ .
- We simulate as many Gaussian random variables as inverse-Wishart random variables, with  $N_e$  given by  $\text{N\_mu\_tildes\_simulated\_non\_final\_level}$  of Table 8.
- We not not fully update  $\Sigma_\ell^{k,j}$  to the corresponding posterior parameter but proceed as in level  $L - 1$ .

### B.10 Training of neural networks at levels $0 \leq \ell \leq L - 2$

- We now have to optimize over  $q_\ell^k \in q_g$ . Therefore, we must now train up to  $|q_g|$  different neural networks (with different inputs' sizes). In practice, we only train neural networks indexed by  $q \in (q_\ell^k)_{1 \leq k \leq \text{k\_bar}} \subset q_g$ , that is, for all the choices of  $q$  that are obtained by at least one strategy at level  $\ell$ .
- We must also choose a  $\delta N$  that should be added as an entry of the neural network before optimizing. Furthermore, to help the neural networks converge, we decided to add  $\text{f\_precompute}(\ell, j, k)$  as an input.
- The loss function and the structure of the neural network is as above, and we still use Xavier initialization, and bring the inputs of the neural networks to reasonable values close to 0 and 1 by renormalizing them using the constants of Table 6.
- Compared to levels  $L$  and  $L - 1$ , the choice of batches is slightly different. Indeed, to train a neural network associated to  $q \in q_g$ , we only use strategies such that  $q_\ell^k = q$ . To do so, we first define  $S_q = \{k \in \llbracket 1, \text{k\_bar} \rrbracket : q_\ell^k = q\}$ . We then define  $\text{k\_batch} = \min(\text{k\_batch\_size}, |S_q|)$  and  $\text{j\_batch} = \min(\text{j\_batch\_size}, \text{j\_bar})$ . We then proceed nearly identically as for levels  $L$  and  $L - 1$ . We select a new batch every  $\text{N\_batch\_change\_proportion}$ , composed of indices  $1 \leq \mathbf{m}_a(j) \leq \text{j\_batch}$ ,  $1 \leq \mathbf{m}_b(k) \leq \text{k\_batch}$ , where  $\mathbf{m}_a$  and  $\mathbf{m}_b$  are uniform random permutations of  $\llbracket 1, \text{j\_bar} \rrbracket$  and  $S_q$ . For each batch, the algorithm used for the training is again the standard gradient descent of Tensorflow.
- Compared to levels  $L$  and  $L - 1$ , we found that making the neural networks converge was much harder. In particular, the learning rate had to be really fine tuned. In order to automatize the process, for each  $q$ , we proceed as follows. We do not intanciate one, but  $\text{number\_of\_neural\_networks\_for\_learning\_rate\_test}$  neural networks. For each of these neural networks, we do  $\text{N\_Iter\_learning\_rate\_test}$  training steps, but use different learning rates for each. For the first neural network, we use  $\text{base\_learning\_rate}$  as the learning rate, for the second,  $\text{base\_learning\_rate}/10$ , and for the  $k$ -th,  $\text{base\_learning\_rate}/10^{k-1}$ . For each of these neural networks, we store at each iteration step the log error. Once the  $\text{N\_Iter\_learning\_rate\_test}$  training steps have been done for each of these neural networks, we keep the neural network instance that has the lowest average log error. If it is the  $k$ -th neural network, we then train it again for  $\text{N\_Iter}$  training steps, using as learning rate  $\text{base\_learning\_rate}/10^k$ .

### B.11 Parallelization

In practice, we parallelize the forward pass according to the strategy indices  $k$ . We run  $\text{thread\_batch\_size}$  processes in parallel, where  $\text{thread\_batch\_size}$  is defined in Table 8.

At a given level  $\ell$ , the computation of expectation  $\mathbb{E}_\ell^{k,j}$  can be parallelized according to the sample indices  $j$ . In practice, we run  $\text{number\_of\_threads\_for\_level\_expectations}$  number of processes in parallel, where  $\text{number\_of\_threads\_for\_level\_expectations}$  is defined in Table 8.

For a given level, the training of each neural network corresponding to a given  $q \in q_g$  can be done independently. Therefore, at a given level, we multiprocessed our code in order to train all the neural networks in parallel.

## B.12 Normalization constants, implementation parameters, and learning rates

Level	q	m	$\Sigma$	N	running_cost	f_precompute
1	6	$10^6$	$10^{12}$	$10^4$	$10^7$	$10^5$
2	6	$10^6$	$10^{12}$	$10^4$	$10^7$	$10^5$
3	6	$10^6$	$10^{12}$	$10^4$	$10^5$	$10^6$
4	6	$10^6$	$10^{11}$	$10^4$	$10^5$	$10^6$

Table 6: Inputs' renormalization constants by Level

Level	q	base_learning_rate	Level	q	base_learning_rate
1	6	$10^{-9}$	2	6	$10^{-9}$
1	10	$10^{-9}$	2	10	$10^{-9}$
1	15	$10^{-9}$	2	15	$10^{-9}$
1	20	$10^{-9}$	2	20	$10^{-9}$
1	25	$10^{-9}$	2	25	$10^{-9}$
1	30	$10^{-9}$	2	30	$10^{-9}$
1	35	$10^{-9}$	2	35	$10^{-9}$
1	40	$10^{-9}$	2	40	$10^{-9}$
1	45	$10^{-9}$	2	45	$10^{-9}$
1	50	$10^{-9}$	2	50	$10^{-9}$
1	60	$10^{-9}$	2	60	$10^{-9}$
1	70	$10^{-9}$	2	70	$10^{-9}$
1	80	$10^{-9}$	2	80	$10^{-9}$
1	90	$10^{-9}$	2	90	$10^{-9}$
1	100	$10^{-9}$	2	100	$10^{-9}$
1	150	$10^{-9}$	2	150	$10^{-9}$
1	200	$10^{-10}$	2	200	$10^{-10}$
1	253	$10^{-10}$	2	253	$10^{-10}$
3	6	$10^{-7}$	4	6	$10^{-7}$

Table 7: Neural network base learning rates



Parameter	Value
j_batch_size	4
k_batch_size	4
N_batch_change_proportion	1 000
N_iter_show_proportion	100
smaller_learning_rate_proportion	10
N_Iter_smaller_learning_rate	10 000
L	4
n_s	253
n_w	6
k_bar	200
j_bar	40
i_0	300
k_0	300
$\Sigma_0$	$(300 - 253 - 1)\Sigma$
N_wishart_proportion	1 000
N_mu_tildes_simulated	1 000 000
thread_batch_size	4
number_of_threads_for_level_expectations	4
thread_batch_size_for_level_expectations	4
p	1
r	2
c	0
N_Iter	1 000 000
N_Iter_learning_rate_test	100 000
number_of_neural_networks_for_learning_rate_test	4
K	10 000 000
N_mu_tildes_simulated_non_final_level	1 000
N_memory_new_pricings_opt	100

Table 8: Implementation parameters