A SPARSE GRID APPROACH TO BALANCE SHEET RISK MEASUREMENT

CYRIL BÉNÉZET[1], JÉRÉMIE BONNEFOY[2], JEAN-FRANÇOIS CHASSAGNEUX[1], SHUOQING DENG[3], CAMILO GARCIA TRILLOS[4] AND LIONEL LENÔTRE[5]

Abstract. In this work, we present a numerical method based on a sparse grid approximation to compute the loss distribution of the balance sheet of a financial or an insurance company. We first describe, in a stylised way, the assets and liabilities dynamics that are used for the numerical estimation of the balance sheet distribution. For the pricing and hedging model, we chose a classical Black & Scholes model with a stochastic interest rate following a Hull & White model. The risk management model describing the evolution of the parameters of the pricing and hedging model is a Gaussian model. The new numerical method is compared with the traditional nested simulation approach. We review the convergence of both methods to estimate the risk indicators under consideration. Finally, we provide numerical results showing that the sparse grid approach is extremely competitive for models with moderate dimension.

1. INTRODUCTION

The goal of this paper is to present a robust and efficient method to numerically assess risks on the balance sheet distribution of, say, an insurance company, at a given horizon. In practice, it is chosen to be one year, consistently with the Solvency 2 regulation, the prudential framework for assessing the required solvency capital for an European insurance company.

On a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$, the balance sheet of the company is a random process summarised, at any time $t \geq 0$, by the value of the assets of the company $(A_t)_{t \geq 0}$ and the value of the liabilities $(L_t)_{t \geq 0}$. The quantity of interest is the Profit and Loss (PnL in the sequel) associated to the balance sheet, which is given by

$$P_t = L_t - A_t , \quad t \geq 0 .$$

By convention, and adopting the point of view of risk management, we measure the loss as a positive quantity.

On the Liability side, the insurance company has sold a structured financial product which depends on the evolution of a one-dimensional stock price (S_t) and the risk-free interest rate (r_t). Several insurance products could be of this type, in particular Unit-Linked (with or without financial guarantees) and Variable Annuity contracts. For those contracts, client's money is invested in equity and bond markets while the insurance company might also provide with financial guarantees similar to long-term put options. The long maturity of

---

[1] Paris Diderot University, LPSM.

[2] Group Risk Management, GIE AXA.

[3] Paris Dauphine University, CEREMADE.

[4] University College London.

[5] Ecole Polytechnique, CMAP.

those contracts requires the introduction of a model for interest rate as they are very sensitive to Interest Rate curve movements. The value L_1 is just the price of this product taking into account the value of some risk factors $\mathcal{X}_1$ (stock price, interest rate curve etc.) at time $t = 1$ used to calibrate the pricing model.

On the Asset side, the insurance company manages some assets to hedge the risk associated to the product sale. The pricing actually includes a margin which is secured through hedging. The hedging assets are the stock and swaps of several maturities, in practice mostly concentrated on the long term. In practice, bond futures are also included sometimes. The hedging portfolio is typically rebalanced on a weekly basis and the hedging quantities are determined by a financial model, taken to be the same as the liability pricing model, whose inputs are the risk factors $\mathcal{X}_t$ at the time t when the hedge is computed.

We describe precisely in Section 2, the pricing and hedging model, the dynamics of the risk factor $\mathcal{X}$ and the value of the asset and liability side of the balance sheet. Let us stress that the risk factor model is given under the so-called *real-world* probability measure $\mathbb{P}$, which might be objectively calibrated using time series of financial markets or represents the management view. This *real-world* model may be –and most of the time is– completely different from the pricing and hedging model which might be simplified for runtime/trackability purposes, prudent (pricing and hedging include a margin) or being constrained by regulation.

Our goal is then to compute various risk indicator for the loss distribution of the balance sheet at one year namely the distribution of P_1 under the *real-world* probability measure $\mathbb{P}$, that we denote hereafter η.

Precisely, we measure the risk associated to η using a (law invariant) risk measure defined over the class of square integrable measures $\varrho : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$. First, we consider for ϱ the so called Value-at-Risk ($V@R$), which is defined by the left-side quantile:

$$V@R_p(\eta) = \inf \left\{ q \in \mathbb{R} \mid \eta \left((-\infty, q] \right) \geq p \right\} . \tag{1}$$

We will also work with the class of *spectral risk measures*: a spectral risk measure is defined as

$$\varrho_h(\eta) = \int_0^1 V@R_p(\eta) h(p) \mathrm{d}p , \tag{2}$$

where h is a non-decreasing probability density on $[0, 1]$. In the numerics, we will focus on the Average Value-at-Risk ($AV@R$) which is given by

$$AV@R_\alpha(\eta) = \frac{1}{1-\alpha} \int_\alpha^1 V@R_p(\eta) \mathrm{d}p , \tag{3}$$

and is a special case of a spectral risk measure.

For a law invariant risk measure ϱ, we denote by $\Re$ its "lift" on $L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}) =: L^2$, namely $\Re[X] = \varrho([X])$ for any $X \in L^2$, where $[X]$ denotes the law of X. The lift $\Re_h$ from a spectral risk measure ϱ_h satisfies the following properties:

(1) *Monotonicity*: $\Re_h[X] \leq \Re_h[Y]$, for $X \leq Y \in L^2$;
(2) *Cash invariance*: $\Re_h[X + c] = \Re_h[X] + c$ for $X \in L^2$ and $c \in \mathbb{R}$;
(3) *Positive homogeneity*: $\Re[tX] = t\Re[X]$, $t \geq 0$ and $X \in L^2$.
(4) *Convexity*: $\Re[tX + (1-t)Y] \leq t\Re[X] + (1-t)\Re[Y]$, whenever $0 \leq t \leq 1$, for $X, Y \in L^2$;

Let us stress the fact that $V@R$ only satisfies 1-3. We refer to [7] and the references therein for more insights on risk measures and spectral risk measures.

In our setting, the loss distribution η of the balance sheet PnL is obtained through the following expression:

$$\eta = p_1 \sharp \nu ,$$

where $\sharp$ denotes the push-forward operator, $p_1 : \mathbb{R}^\theta \to \mathbb{R}$ is the *function* describing the PnL in terms of the risk factors, and ν stands for the distribution of the risk factors $\mathcal{X}$. In practice the estimation of $\varrho(\eta)$ requires to

sample from η. In turn, this demands for a sample of the model parameter distribution ν and for a numerical approximation of p_1. In this note, we compare two main approaches to form the sample of η given one of ν.

The first one is known as the *nested simulation* approach: It is a two-step method. First, a set of "outer simulation", describing the random values of the risk factors, is drawn. Then, for each value of the risk factors, a sample of "inner simulation" is drawn to compute the various hedge and prices. In this approach, all computations are realised "online". The main advantage of this approach is its simplicity to implement in practice, described in the first paragraph of Subsection 3.1. However, it is well known that this approach is quite greedy, even if optimised as in [6]. We also want to stress the fact that when computing the η-sample, no information about p_1 is stored for future work: for example if ν is modified, due to time or a model change, a full recalculation would be required.

The other approach we chose to adopt and would like to promote is a *grid approach* where the approximation of p_1 is made "offline", by a Monte Carlo approach, and then stored. The numerical computation is then done through a (multi-linear) interpolation on a grid. The main drawback of this approach is that the size of the grid, in high dimension, can become untractable, especially if one uses regular grid. To partially circumconvent this difficulty, we introduce a *sparse grid* [2] which reduces drastically the number of point to be used (equivalently, values to be stored) with only small reduction of the accuracy of the method.

We prove that, for a spectral risk measure, the two approaches give an estimation of $\varrho(\eta)$ which converges to the true value, see Theorem 3.8.

Furthermore, we show in the numerical Section 4 that using the grid approach together with a sparse grid of low level allows to get a good approximation of the loss distributions η, and of some related risk measures, while reducing drastically the computational time and allowing to keep information about the balance sheet function p_1. Last, this permits to numerically quantify uncertainty. Indeed, since the computations on the grid are stored, the computation of the distribution of the PnL under other distributions for the parameters is almost instantaneous and can be compared with the results obtained with the initial one. An application to uncertainty estimation is given in the last numerical application.

The rest of the paper is organised as follows. In the Section 2, we first describe the mathematical models that are used to describe the evolution of the prices under the risk-neutral measure $\mathbb{Q}$. We then describe precisely how A and L are specified. In Section 3, we describe the two numerical methods used to compute L_t and A_t at any given time $t \geq 0$. In particular, we show how to efficiently compute, at time t, the quantities to hold in the hedging portfolio, which are expressed in term of the derivatives of the claim's price. We also explain how to compute the price of the product and of the assets used to construct the hedging porfolio, leading to the computation of L_t and A_t. We show how to obtain an approximation of the distribution of P_1 under the physical measure $\mathbb{P}$, and we prove an upper bound for the mean square error of the overall procedure. Finally, in Section 4, we present our numerical results, comparing the two methods.

2. Financial Model

In this section, we give the precise specification of the asset and liability sides of the balance sheet. We also present the *risk-neutral* model and the *real-world* model that are used.

2.1. **Description of the sold product**

Let us assume that a company sells a contingent claim at time $t = 0$ which is a (discretely) path-dependent option with a payoff function G paid at the maturity $T > 0$, depending upon the evolution of a one-dimensional risky asset's price S. We focus here on:
A put lookback option, that is a discretely path-dependent option whose strike at maturity T is given by the maximum of the asset's price S over the times $t \in \{\tau_0 = 0, \tau_1, \cdots, \tau_\kappa = T\}$ where $\kappa \geq 1$:

$$G(S_{\tau_0}, \ldots, S_{\tau_\kappa}) = \left(\max_{0 \leq \ell \leq \kappa} S_{\tau_\ell}\right) - S_T. \tag{4}$$

Remark 2.1. The proxy provided above is close to financial guarantees offered in Variable Annuity contracts. Those contracts are structured insurance products composed of a fund investment on top of which both insurance and financial protection are added. In our case, the contract is a *Guaranteed Minimum Accumulation Benefit including a ratchet mechanism.* At time $t = 0$, the customer invests his/her money in the underlying fund and will receive at a given maturity the maximum between the terminal fund value and its *terminal benefit base* in case she is still alive. The terminal benefit base is equal to the maximum of the underlying fund values observed at each anniversary date of the contract (ratchet mechanism). We do not consider the modeling of death/survival in this proceedings, neither the possibility that client can surrender at any time during the life of the contract.

2.2. **Market model under the risk-neutral measure**

We assume that all pricing and hedging is done with a market risk-neutral measure $\mathbb{Q}$.

The derivative with a payoff function G as above depends upon a one-dimensional stock's price $S = (S_t)_{t \in [0,T]}$. We assume here that the dynamics of the asset under $\mathbb{Q}$ are of the Black & Scholes type as described in Section 2.2.2 with a stochastic interest rate $r = (r_t)_{t \in [0,T]}$ which follows a Hull & White model.

As the payoff G is a proxy of Variable Annuity guarantee which is a long term Savings product (in practice maturity ranges from 10 to 30 years depending on product type), the modeling of interest rate is essential as the product and therefore the overall balance sheet of the company is very sensitive to this risk.

2.2.1. *The short rate model*

Let $\Theta \in \mathbb{R}^d$ ($d := 3$ in the sequel) be a set of parameters representing some market observations. The short rate evolution is governed by the Hull & White dynamics

$$r_s^{t,\Theta} = r_t^{t,\Theta} + \int_t^s a \left(\mu_u^{t,\Theta} - r_u^{t,\Theta} \right) du + b \left(B_s - B_t \right), \ s \in [t,T], \tag{5}$$

where B is a $\mathbb{Q}$-Brownian motion, a and b are real constants and $\mu^{t,\Theta} : [t,T] \to \mathbb{R}$ is a function. We refer to [1] for a more complete analysis of the Hull & White short rate model.

The parameter $\mu^{t,\Theta}$ is calibrated using the market observations Θ, so that the model reproduces the interest rate curve observed on the market. It is given by

$$\mu_s^{t,\Theta} = f^{\Theta}(t,s) + \frac{1}{a} \frac{\partial f^{\Theta}(t,s)}{\partial s} + \frac{b^2}{2a^2} \left(1 - e^{-2a(s-t)} \right), s \in [t,T]. \tag{6}$$

We refer to the Appendix for a derivation of (6).

As a consequence, the Θ parameter must be chosen in order to represent adequately the forward rate curve observed on the market.

We suppose here that the forward rate curve $f^{\Theta}(t, \cdot)$ is directly observed and is a linear combination of three elementary functions $h^{t,1}, h^{t,2}, h^{t,3}$ from $[t,T]$ to $\mathbb{R}$, given by

$$h^{t,1}(s) := h^1(s-t), \ h^{t,2}(s) := h^2(s-t), \ \text{and} \ h^{t,3}(s) := h^3(s-t), \ s \in [t,T],$$

where, for $u \in [0,T]$:

$$h^1(u) = \begin{cases} 1 & \text{if } u \le \frac{t_1+t_2}{2}, \\ 2\frac{\frac{t_2+t_3}{2}-u}{t_3-t_1} & \text{if } u \in [\frac{t_1+t_2}{2}, \frac{t_2+t_3}{2}], \\ 0 & \text{otherwise}, \end{cases} \qquad h^3(u) = \begin{cases} 0 & \text{if } u \le \frac{t_2+t_3}{2} \\ 2\frac{u-\frac{t_2+t_3}{2}}{t_4-t_2} & \text{if } u \in [\frac{t_2+t_3}{2}, \frac{t_3+t_4}{2}], \\ 1 & \text{otherwise}, \end{cases}$$

$$\text{and } h^2(u) = 1 - h^1(u) - h^3(u),$$

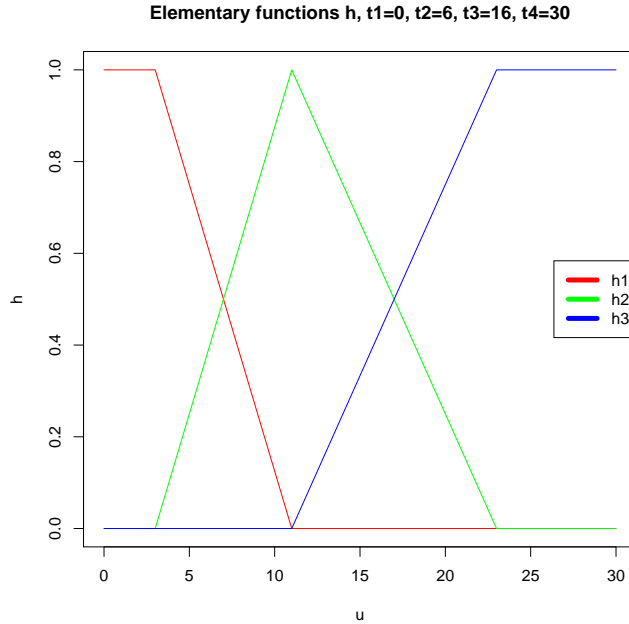where $0 \leq t_1 < t_2 < t_3 < t_4 \leq T$ are four fixed real numbers.

FIGURE 1. Building blocks for the forward interest rate curve.

The function $h^{t,1}$ (resp. $h^{t,2}, h^{t,3}$) model the short (resp. middle, long) term structure of the interest rates curve.

In a nutshell, the short rate model is determined by:

(1) the time of observation $t \in [0, T]$,
(2) the three-dimensional parameter $\Theta := \{\theta_1, \theta_2, \theta_3\} \in \mathbb{R}^3$, where $\theta_1, \theta_2, \theta_3$ are such that

$$f^{\Theta}(t, \cdot) = \theta_1 h^{t,1} + \theta_2 h^{t,2} + \theta_3 h^{t,3}, \tag{7}$$

$f^{\Theta}(t, \cdot)$ being the observed forward rates curve.

As a result, given an observation (t, Θ) as above, the short rate process $r^{t,\Theta}$ has the dynamics (5), where $\mu^{t,\Theta}$ is computed using (6).

Remark 2.2. In practice, the parameters a, b appearing in (5) can be calibrated so that the model reproduces the prices, observed on the market, of some contracts such as swaps or swaptions. We could more generally allow the parameters a, b of the Hull & White model to depend upon the market observations Θ. The parameter Θ should live in a higher-dimensional space to take into account the observed swap(tion)s prices. Regular recalibration of parameters is largely performed by practitioners, in particular when they perform dynamic hedging.

There are several reasons explaining the choice of this model. First of all, it is quite simple to calibrate using the data. In fact, the function μ is directly given as a function of the forward rate curve. We should note again that the choice of keeping a, b fixed through time simplifies the calibration. Secondly, we will see later in Proposition 3.2 that this short rate model, associated to the stock model described below, leads to an exact simulation under the risk-neutral measure. Lastly, closed and easily tractable formulas can be obtained for the

prices of the zero-coupon bonds and swaps which are the products used to construct the hedging portfolio and then to compute the value of the company's assets A.

These prices are as follows.

Proposition 2.3. *Let $(t, \Theta) \in [0, 1] \times \mathbb{R}^3$ be a market observation, and consider the process $\left(r_s^{t,\Theta}\right)_{s \in [t,T]}$ given by (5), where the parameter $\mu^{t,\Theta}$ is defined with (6) and (7).*

(1) *The price at time t of a zero-coupon maturing at time $u \in [t, T]$ is given by:*

$$P^{t,\Theta,u} = \exp\left(-\int_t^u f^\Theta(t, s)\mathrm{d}s\right), \tag{8}$$

and its derivatives with respect to $\Theta := (\theta_1, \theta_2, \theta_3)$ are given by:

$$\frac{\partial P^{t,\Theta,u}}{\partial \theta_i} = -P^{t,\Theta,u}\int_t^u h^{t,i}(s)\mathrm{d}s. \tag{9}$$

(2) *Let $(0, \Theta_0)$ be the observation made at time 0. Consider a swap contract issued in $s = 0$, with maturity $M > 0$, rate $R > 0$, with coupons versed at every time $i \in \{1, \ldots, M\}$. Then, the price of this contract at time t is given by:*

$$SW^{t,\Theta,M,R} = \frac{P^{t,\Theta,1}}{P^{0,\Theta_0,1}} - P^{t,\Theta,M} - R\sum_{i=1}^M P^{t,\Theta,i}, \tag{10}$$

and its derivatives with respect to Θ are given by:

$$\frac{\partial SW^{t,\Theta,M,R}}{\partial \theta_j} = -\frac{P^{t,\Theta,1}}{P^{0,\Theta_0,1}}\int_t^1 h^{t,j}(s)\mathrm{d}s + P^{t,\Theta,M}\int_t^M h^{t,j}(r)\mathrm{d}r + R\sum_{i=1}^M \left(P^{t,\Theta,i}\int_t^{t+i} h^{t,j}(s)\mathrm{d}s\right). \tag{11}$$

2.2.2. *The stock model*

Given the observations Θ of the interest rate factors and the risky asset's price $x \in (0, \infty)$, the evolution of the price under the neutral-risk measure $\mathbb{Q}$ is given by

$$S_s^{t,x,\Theta} = x + \int_t^s r_u^{t,\Theta} S_u^{t,x,\Theta}\mathrm{d}u + \int_t^s \sigma S_u^{t,x,\Theta}\mathrm{d}\tilde{W}_u, s \in [t, T], \tag{12}$$

where $\sigma > 0$, $\tilde{W}$ is another $\mathbb{Q}$-Brownian motion, whose quadratic covariation with B is given by

$$\langle B, \tilde{W}\rangle_t := \rho\, t, t \in [0, T],$$

where $\rho \in [-1, 1]$. Equivalently, $S^{t,x,\Theta}$ can be written as:

$$S_s^{t,x,\Theta} = x + \int_t^s r_u^{t,\Theta} S_u^{t,x,\Theta}\mathrm{d}u + \int_t^s \rho\sigma S_u^{t,x,\Theta}\mathrm{d}B_u + \int_t^s \sqrt{1-\rho^2}\sigma S_u^{t,x,\Theta}\mathrm{d}W_u, s \in [t, T] \tag{13}$$

where W is a $\mathbb{Q}$-Brownian motion, independent of B.

Remark 2.4. In practice, the parameter σ appearing in (12) can be calibrated so that the model reproduces the prices of some derivatives over the risky asset. This can be taken into account by increasing the dimension of the space where Θ lives, and by adding this calibration procedure.

Remark 2.5. Naturally, the general sparse grid approach can be applied to different models and functional representations. We made the choice of using a Black & Scholes model for the stock value and a Hull & White model with the given functional representation in terms of h^1, h^2, h^3 for the short rate, since they are convenient to obtain explicit pricing and sensitivities formulae as we show in the following.

2.3. **Modeling the Balance Sheet**

The key point for us is to approximate the distribution of the balance sheet of an insurance company at time $t = 1$ (here a year) given the market observations at $t = 0$. As mentionned in the introduction, the PnL is a process P which can be decomposed as

$$P_t = L_t - A_t, \; t \in [0, 1], \tag{14}$$

where L is the value of the liabilities of the company and A is the value of the assets.
We assume that at time $t = 0$, the balance sheet is clear, meaning that the company has no asset nor liability, that is $L_0 = A_0 = 0$.

We describe precisely in the two subsequent sections how these quantities are defined. Importantly, we denote by $\bar{\mathcal{X}}_t := (\bar{S}_t, \bar{\Theta}_t)$, $0 \le t \le 1$, the stochastic process representing the random evolution of the market parameter under the *real-world* measure $\mathbb{P}$. Namely, $\bar{S}$ is the stock price and $\bar{\Theta}$ the interest rate curve parameters as described above in Section 2.2.1. It is important to have in mind that the model chosen for the stock price S (under $\mathbb{Q}$) and $\bar{S}$ (under $\mathbb{P}$) will be completely different as they do not serve the same purpose (pricing-hedging on one hand, risk management of the Balance Sheet or regulatory assessment of required capital on the other hand).

2.3.1. *Liability side*

For any market observation $\bar{\mathcal{X}}_t := (\bar{S}_t, \bar{\Theta}_t)$, the value $L_t =: \ell(t, \bar{S}_t, \bar{\Theta}_t)$ of the liabilities has to be computed, especially at time $t = 1$ in our application. The company's liabilities are reduced to one derivative product sold at $t = 0$. In our setting, the contingent claim's price is simply given by:

$$\ell(t, x, \Theta) = \mathbb{E}^{\mathbb{Q}} \left[e^{- \int_t^T r_s^{t,\Theta} \mathrm{d}s} G(S^{t,x,\Theta}) \right], \tag{15}$$

where $(r^{t,\Theta}, S^{t,x,\Theta})_{t \le s \le T}$ are the risk neutral dynamics of the short rate and stock price, see Section 2.2.1 and Section 2.2.2, calibrated from the observed market parameter (x, Θ) at time t.

We recall that the payoff G depends on $S^{t,x,\Theta}$ only through the values $S_\tau^{t,x,\Theta}$, $\tau \in \Gamma_G := \{\tau_0, \ldots, \tau_\kappa\}$ ($\kappa \ge 0$), see (4).

As explained in more detail below, the computation of P_1 first requires to approximate $\ell(t, x, \Theta)$ for (t, x, Θ) on a (possibly stochastic) discrete grid of $[0, 1] \times (0, \infty) \times \mathbb{R}^3$. This approximation L at any point $(t, x, \Theta) \in [0, 1] \times (0, \infty) \times \mathbb{R}^3$ basically follows from the simulation of the processes $r^{t,\Theta}$ and $S^{t,x,\Theta}$ under the risk-neutral measure $\mathbb{Q}$ and a Monte Carlo procedure. We will see in Section 3 that the simulation can be done in an exact manner in our model.

Remark 2.6. A classical approach to compute ℓ would be to use a dynamic programming principle. Step by step, it requires

(1) To numerically obtain $\ell(1, x, \Theta)$ for all (x, Θ) on the grid with (15),
(2) Then to iteratively compute $\ell(t_{k+1}, x, \Theta)$ for all x, Θ on the grid using

$$\ell(t_k, x, \Theta) = \mathbb{E}^{\mathbb{Q}} \left[e^{- \int_{t_k}^{t_{k+1}} r_s^{t_k,\Theta} \mathrm{d}s} \mathbb{E}^{\mathbb{Q}} \left[e^{- \int_{t_{k+1}}^T r_s^{t_k,\Theta} \mathrm{d}s} G(S^{t_k,x,\Theta}) | \mathcal{F}_{t_{k+1}} \right] \right]. \tag{16}$$

However, the inner conditional expectation is *not* of the form $\ell(t_{k+1}, x, \Theta)$ for $x > 0$ and $\Theta \in \mathbb{R}^3$. In fact, the time of the market observation Θ is still t_k, while the discount factor goes only from t_{k+1} to T, in contrast with

(15). Therefore, using the dynamic programming equation (16) would require to introduce some additional and artificial parameters, namely the time of calibration and the value of $r^{t_k,\Theta}$ at time t_{k+1}. This would made the overall procedure heavier that is why we compute ℓ using simply (15).

2.3.2. *Asset side*

The company wants to replicate the product with payoff G[1]. The classical theory of mathematical finance ensures that it is equivalent, in theory, to possess a portfolio which negates the variations of the price of the product with respect to the evolution of the underlying parameters.

In our context, the insurer wants to be protected against the variations with respect to the stock price S_t and the interest rate curve, which is modeled through the parameter Θ.

The dynamic hedging portfolio is constructed and rebalanced in discrete time, on the time grid $\Gamma := \{t_0 = 0 < t_1 < \cdots < t_n = 1\}$ (in practice, the porfolio will be rebalanced up to the maturity of the product, but in our setting, we are only interested in the portfolio's value up to $t = 1$). At each time $t \in \Gamma$, the insurer computes the derivatives of the price with respect to S_t and Θ_t, and then buys some financial assets (the stock and swaps) in order to construct a portfolio whose derivatives match those computed.

To model this framework, we decompose the hedging portfolio's value A in two parts:

$$A_t = A_t^\Delta + A_t^\rho. \tag{17}$$

The process A^Δ is the value of the portfolio obtained to cancel the variations of the price with respect to S, while A^ρ is defined to deal with the variations with respect to Θ.

Remark 2.7. Obviously, since in practice the hedging is done in discrete time and some underlying parameter are not considered, the payoff G is not exactly replicated, nor super-replicated. Therefore the PnL of the company is not null, nor always positive, and the goal of this proceedings is precisely to propose a new numerical method to estimate the distribution of this quantity at time $t = 1$.

To construct the hedging portfolio, the insurer can buy the underlying stock, together with three swap contracts, defined by some rates $R_1, R_2, R_3 > 0$ and maturity dates $T_1, T_2, T_3 \in [1, T]$. Interest rate hedging is performed so that the portfolio is insensitive to the variations of the main maturities of the interest rate curve. For long-term products, this means building an hedging portfolio containing several different maturities from 1 year to 30 year. Here, only 3 maturities representing short, medium and long-term part of the curve are considered for simplicity. The formula for their price $\mathrm{SW}^{t,\Theta,T_i,R_i}$, $i = 1, 2, 3$ is given in Proposition 2.3 above. We now describe how to compute the quantities of assets and swaps to buy at a time $t \in \Gamma$, to rebalance the hedging portfolio. Denote by Δ (resp. $\rho_i, i = 1, 2, 3$) the quantities of stock (resp. swap with rate R_i and maturity date T_i, $i = 1, 2, 3$). Then the value of the portfolio of the company is given by:

$$\Pi_t = \Delta S_t + \sum_{i=1}^{3} \rho_i \mathrm{d}SW^{t,\Theta_t,T_i,R_i} - \ell(t, S_t, \Theta_t). \tag{18}$$

---

[1]In practice, the pricing embeds a margin and the objective of replicating the payoff G is to secure it.

By It's formula, assuming a semimartingale decomposition for the process Θ under $\mathbb{P}$, we get:

$$
\begin{aligned}
d\Pi_t = {} & (\Delta - \Delta(t, S_t, \Theta_t)) \, dS_t \\
& + \left(\rho_1 \frac{\partial SW^{t,\Theta,T_1,R_1}}{\partial \theta_1} + \rho_2 \frac{\partial SW^{t,\Theta,T_2,R_2}}{\partial \theta_1} + \rho_3 \frac{\partial SW^{t,\Theta,T_3,R_3}}{\partial \theta_1} - \frac{\partial \ell}{\partial \theta_1}(t, x, \Theta) \right) d\theta_1 \\
& + \left(\rho_1 \frac{\partial SW^{t,\Theta,T_1,R_1}}{\partial \theta_2} + \rho_2 \frac{\partial SW^{t,\Theta,T_2,R_2}}{\partial \theta_2} + \rho_3 \frac{\partial SW^{t,\Theta,T_3,R_3}}{\partial \theta_2} - \frac{\partial \ell}{\partial \theta_2}(t, x, \Theta) \right) d\theta_2 \\
& + \left(\rho_1 \frac{\partial SW^{t,\Theta,T_1,R_1}}{\partial \theta_3} + \rho_2 \frac{\partial SW^{t,\Theta,T_2,R_2}}{\partial \theta_3} + \rho_3 \frac{\partial SW^{t,\Theta,T_3,R_3}}{\partial \theta_3} - \frac{\partial \ell}{\partial \theta_3}(t, x, \Theta) \right) d\theta_3 \\
& + dt \text{ terms.}
\end{aligned}
$$

To cancel the risks induced by the variations of the stock price and the interest rate curve, it is needed that the four first terms in the previous equation cancel.

Those considerations lead to the following construction for the hedging portfolio $A = A^\Delta + A^\rho$:

Δ-hedging: The value of A^Δ at time 1 is

$$
A_1^\Delta = \sum_{i=0}^{n-1} \Delta(t_i, \bar{S}_{t_i}, \bar{\Theta}_{t_i}) \left(\bar{S}_{t_{i+1}} - \bar{S}_{t_i} \right) \quad \text{where} \quad \Delta(t, x, \Theta) := \frac{\partial L}{\partial x}(t, x, \Theta). \tag{19}
$$

Note that the function Δ has to be computed, at each time $t_i, i = 0, \ldots, n$, and for any market situation (x, Θ) at this time. A method leading to a numerical estimation of Δ is proposed in Section 3.

ρ-hedging: The value of A_1^ρ is

$$
A_1^\rho = \sum_{i=0}^{n-1} \sum_{j=1}^{3} \rho^j(t_i, \bar{S}_{t_i}, \bar{\Theta}_{t_i}) \left(SW^{t_{i+1}, \bar{\Theta}_{t_{i+1}}, T_j, R_j} - SW^{t_i, \bar{\Theta}_{t_i}, T_j, R_j} \right). \tag{20}
$$

At time $t \in [0, T]$, for a market at (x, Θ), the quantities $\rho^i(t, x, \Theta)$, $i = 1, 2, 3$ of each swap contract required for the hedging are given by the solution of the linear system

$$
\begin{cases}
\rho_1 \frac{\partial SW^{t,\Theta,T_1,R_1}}{\partial \theta_1} + \rho_2 \frac{\partial SW^{t,\Theta,T_2,R_2}}{\partial \theta_1} + \rho_3 \frac{\partial SW^{t,\Theta,T_3,R_3}}{\partial \theta_1} &= \frac{\partial \ell}{\partial \theta_1}(t, x, \Theta) \\
\rho_1 \frac{\partial SW^{t,\Theta,T_1,R_1}}{\partial \theta_2} + \rho_2 \frac{\partial SW^{t,\Theta,T_2,R_2}}{\partial \theta_2} + \rho_3 \frac{\partial SW^{t,\Theta,T_3,R_3}}{\partial \theta_2} &= \frac{\partial \ell}{\partial \theta_2}(t, x, \Theta) \\
\rho_1 \frac{\partial SW^{t,\Theta,T_1,R_1}}{\partial \theta_3} + \rho_2 \frac{\partial SW^{t,\Theta,T_2,R_2}}{\partial \theta_3} + \rho_3 \frac{\partial SW^{t,\Theta,T_3,R_3}}{\partial \theta_3} &= \frac{\partial \ell}{\partial \theta_3}(t, x, \Theta)
\end{cases} \tag{21}
$$

One key quantity to compute for us in this setting is thus the vector of sensitivities $\left(\frac{\partial \ell}{\partial \theta_i}(t, x, \Theta) \right)$, $i = 1, 2, 3$.

Remark 2.8. (i) We choose to always use the same swap contracts issued at $t = 0$ as hedging instruments. We could have decided to enter for free in swaps (at the swap rate) at each rebalancing time. However, this strategy requires to keep the memory of the swap rate in order to compute the swap price at the next rebalancing date.
(ii) The T_i, R_i should be chosen so that they represent some liquid contracts.

2.3.3. *The global PnL function*

From the previous two sections, we conclude that the PnL of the balance sheet at time 1 can be expressed as,

$$
P_1 = p_1 \left((t, \bar{S}_t, \bar{\Theta}_t)_{t \in \Gamma} \right)
$$

where $(\bar{S}, \bar{\Theta})$ are the market parameters (risk factors) and the PnL function $p_1 : \mathbb{R}^\gamma \to \mathbb{R}$, with $\gamma = 4 \times (n+1)$, is given by

$$\ell(t_n, x_n, \Theta_n) - \sum_{i=0}^{n-1} \Delta(t_i, x_i, \Theta_i)(x_{i+1} - x_i) - \sum_{i=0}^{n-1} \sum_{j=1}^{3} \rho^j(t_i, x_i, \Theta_i) \left(\mathrm{SW}^{t_{i+1}, \Theta_{i+1}, T_j, R_j} - \mathrm{SW}^{t_i, \Theta_i, T_j, R_j} \right) . \quad (22)$$

In the next section, we describe the model we will consider for the market parameter $(\bar{S}, \bar{\Theta})$.

2.4. **Market parameters under the *real-world* measure**

We describe here the model that will be used for the simulation of the market parameters in the numerical part. Let us insist that this *real-world* measure $\mathbb{P}$ might represent the view of the management on the evolution of the market parameter on the period $[0, 1]$. As already mentioned, it can be completely different from the model used for the *risk-neutral* pricing.

We assume that we know, or at least are able to estimate, the first two moments of the distribution of $(X_1 := \log(S_1), \Theta_1) = (X_1, (\theta_1)_1, (\theta_2)_1, (\theta_3)_1)$ under $\mathbb{P}$. More precisely, we assume that under $\mathbb{P}$, this random vector has mean and covariance matrix given by

$$\mu = (\mu_X, \mu_1, \mu_2, \mu_3), \quad (23)$$

$$V = (V_{ij})_{i,j=0,1,2,3}. \quad (24)$$

To model the random process $(X_t, (\theta_1)_t, (\theta_2)_t, (\theta_3)_t)_{t \in [0,1]}$ under $\mathbb{P}$, we assume that its dynamics are given by

$$X_t = X_0 + b_0 t + c_{00} W_t^0 + c_{01} W_t^1 + c_{02} W_t^2 + c_{03} W_t^3, \quad (25)$$

$$(\theta_1)_t = (\theta_1)_0 + b_1 t + c_{11} W_t^1 + c_{12} W_t^2 + c_{13} W_t^3 \quad (26)$$

$$(\theta_2)_t = (\theta_2)_0 + b_2 t + c_{22} W_t^2 + c_{23} W_t^3 \quad (27)$$

$$(\theta_3)_t = (\theta_3)_0 + b_3 t + c_{33} W_t^3, \quad (28)$$

where $W^i, i = 0, 1, 2, 3$ are independent $\mathbb{P}$-Brownian motions.

Proposition 2.9. *There exists at most one set of coefficients $b_i, c_{ij}, i, j = 0, 1, 2, 3$, such that the random vector $(X_1, (\theta_1)_1, (\theta_2)_2, (\theta_3)_2)$ has mean μ and covariance matrix V.*

Proof. We refer to the appendix for a proof, cf. Proposition 5.1. □

Remark 2.10. It is well-known that it is difficult to estimate accurately the drift parameter in a Black-Scholes model. This makes our computation subject to model risk. We leave it to a future research work to find a robust way to approximate the law of P_1 under $\mathbb{P}$. Nevertheless, let us point out that the grid approach allows us to compute, with minimal re-computation, risk measures for different approximations of the law of P_1. This is one of the advantages of this method with respect to using "nested simulations", as illustrated in Section 4.

3. Numerical methods

In this section, we describe the two numerical methods that we use to compute the risk indicator on the balance sheet PnL. The first one is known as *nested simulation* approach and is already used in the industry, see the seminal paper [6]. The second one is a *sparse grid* approach and is designed to be more efficient that the *nested simulation* approach in the case of moderate dimensions. In the next section, we present numerical simulations that confirms this fact for the model with moderate dimension that we consider here.

3.1. **Estimating the risk measure**

Given a risk measure ϱ and the loss distribution η of the balance sheet at one year, we estimate the quantity of interest $\varrho(\eta)$ by simulating a sample of N i.i.d random variables $(\Psi_j)_{1 \leq j \leq N}$ with law η and then computing simply $\varrho(\eta^N)$ using the formulae (1), (2) and (3) with η^N instead of η. Here, η^N stands for the empirical measure associated to the Ψ_j i.e.

$$\eta^N = \frac{1}{N} \sum_{j=1}^{N} \delta_{\Psi_j} \,,$$

where δ_x is the Dirac mass at the point x.

Recall, that in our work, the loss distribution η is obtained through the following expression:

$$\eta = p_1 \sharp \nu$$

p_1 being described in (22) and ν stands for the distribution of the market parameters. Namely, ν is the law of the random variable

$$\bar{\mathcal{X}} := (\bar{S}_t, \bar{\Theta}_t)_{t \in \Gamma} \tag{29}$$

under the *real world* probability measure $\mathbb{P}$.

In order to estimate $\varrho(\eta)$ for a chosen risk measure, we need to be able to sample from η which implies two steps in our setting. First, we need to be able to sample $\bar{\mathcal{X}}$ and then we use an approximation p_1^v of p_1, $v \in \{\mathcal{N}, \mathcal{S}\}$, namely:

- $p_1^{\mathcal{N}}$ if one chooses the *nested simulation* approach;
- $p_1^{\mathcal{S}}$ if one chooses the *sparse grid* approach.

Eventually, the estimator of $\varrho(\mu)$ is given by

$$R^v := \varrho(p_1^v \sharp \nu^N) \,, \quad \text{for } v \in \{\mathcal{N}, \mathcal{S}\} \,. \tag{30}$$

To summarise, the two numerical methods have the following steps.
Nested simulation approach.

(1) Outer step: Simulate the model parameters $(\mathcal{X}_j)_{j=1,\dots,N}$.
(2) Inner step: Simulate $\Psi_j = p^{\mathcal{N}}(\mathcal{X}_j)$ using MC simulations. This requires to compute the option prices with Monte Carlo estimates, the interest rate derivative prices, and the various sensitivities of the price, see subsection 3.2.2. All these computations are done using closed-form formulae that are derived below.
(3) Estimate the risk measure.

All computations are made "online".
Sparse grid approach.

(1) Fix a (sparse) grid $\mathcal{V}$ and compute the approximation $p^{\mathcal{S}}$ at each required value on the grid by an MC simulation. This involves exactly the same computations as 2. above.
(2) Simulate the N model parameter samples $(\mathcal{X}_j)$ and evaluate $\Psi_j = p^{\mathcal{S}}(\mathcal{X}_j)$.
(3) Estimate the risk measure.

Computations at step 1. are done "offline". The next two steps are done "online".
We now describe precisely how to compute $p^{\mathcal{N}}$ and $p^{\mathcal{S}}$.

3.2. **Nested Simulation approach**

In the *nested simulation* approach, once the market parameters $\bar{\mathcal{X}}$ have been simulated, the function $p_1^{\mathcal{N}}$ has itself to be computed. Recalling (22), this requires to evaluate the function $\ell^{\mathcal{N}}, \Delta^{\mathcal{N}}, \rho^{\mathcal{N}}$ (approximations of ℓ, Δ, ρ) at the value of the market parameters. This computation is made using again a Monte Carlo approach.

In order to compute the risk-neutral expectations in the above formulae, one has to sample the short rate process and compute its integral over $[t, T]$, and also to simulate the stock price S at least at the times $\tau_\ell \in \Gamma_G \cap [t, T]$. Under the model described in section 2.2, the simulation is exact and is described in the two following statements whose proofs are postponed to the appendix.

Let $(t, x, \Theta) \in [0, T] \times (0, \infty) \times \mathbb{R}^3$ be a market observation, and consider the processes $(r_s^{t,\Theta})_{s \in [t,T]}$ and $(S_s^{t,x,\Theta})_{s \in [t,T]}$ defined by (5), (12). We first start with an easy and well known observation.

Lemma 3.1. *We have the following decomposition for the short rate process:*

$$r^{t,\Theta} = \xi^t + \alpha^{t,\Theta}, \tag{31}$$

where $(\alpha_s^{t,\Theta})_{s \in [t,T]}$ is the deterministic function of time

$$\alpha_s^{t,\Theta} = \frac{1}{a} f^\Theta(t, s) + \frac{b^2}{2a^3} \left[1 - e^{-a(s-t)} \right]^2, s \in [t, T], \tag{32}$$

and $(\xi_s^t)_{s \in [t,T]}$ is the solution of the SDE

$$d\xi_s^t = -a\xi_s^t ds + b dB_s, s \in [t, T] \tag{33}$$

$$\xi_t^t = 0. \tag{34}$$

The following proposition provides a recursive way to produce sample paths of the triplet $(\xi^t, A^t, X^{t,x,\Theta})$ on the discrete grid $\{t\} \cup (\Gamma_G \cap [t, 1])$. We are thus in a position to simulate the evolution of $(r^{t,\Theta}, S^{t,x,\Theta})$ and the discount factor $\beta_T^{t,\theta} := e^{-\int_t^T r_s^{t,\Theta} ds}$ under the risk-neutral measure $\mathbb{Q}$.

Proposition 3.2. *Fix $t \leq s \leq u \leq T$. Conditionally upon $\mathcal{F}_s$, the triplet*

$$\left(\xi_u^t, A_u^{t,s} := \int_s^u \xi_r^t dr, X_u^{t,x,\Theta} := \log\left(S_u^{t,x,\Theta}\right) \right)$$

is a Gaussian vector of dimension 3, with mean vector and covariance matrix respectively given by

$$\mathbb{E}_s\left[\xi_u^t\right] = \xi_s^t e^{-a(u-s)},$$

$$\mathbb{E}_s\left[A_u^{t,s}\right] = \frac{\xi_s^t}{a}(1 - e^{-a(u-s)}), \tag{35}$$

$$\mathbb{E}_s\left[X_u^{t,x,\Theta}\right] = X_s^{t,x,\Theta} + \int_s^u \alpha_r dr - \frac{\sigma^2}{2}(u-s) + \frac{\xi_s^t}{a}(1 - e^{-a(u-s)}),$$

and

$$\text{Var}_s(\xi_u^t) = \frac{b^2}{2a}(1 - e^{-2a(u-s)}),$$

$$\text{Cov}_s(\xi_u^t, A_u^{t,s}) = \frac{b^2}{a^2}(1 - e^{-a(u-s)}) - \frac{b^2}{2a^2}(1 - e^{-2a(u-s)}),$$

$$\text{Var}_s(A_u^{t,s}) = \frac{b^2}{a^2}(u - s) - \frac{2b^2}{a^3}(1 - e^{-a(u-s)}) + \frac{b^2}{2a^3}(1 - e^{-2a(u-s)}),$$

$$\text{Cov}_s(\xi_u^t, X_u^{t,x,\Theta}) = \frac{b}{a}(\frac{b}{a} + \rho\sigma)(1 - e^{-a(u-s)}) - \frac{b^2}{2a^2}(1 - e^{-2a(u-s)}),$$

$$\text{Cov}_s(A_u^{t,s}, X_u^{t,x,\Theta}) = -\frac{1}{a}\text{Cov}(\xi_u^t, X_u^{t,x,\Theta}) + \frac{b}{a}(\frac{b}{a} + \rho\sigma)(u - s) - \frac{b^2}{a^3}(1 - e^{-a(u-s)}),$$

$$\text{Var}_s(X^{t,x,\Theta}) = (\rho\sigma + \frac{b}{a})^2(u - s) + (1 - \rho^2)\sigma^2(u - s) - 2\frac{b}{a^2}(\rho\sigma + \frac{b}{a})(1 - e^{-a(u-s)}) + \frac{b^2}{2a^3}(1 - e^{-2a(u-s)}).$$

$$(36)$$

3.2.1. *Approximation of the Liability side*

With the above results, the approximation at any time t of the liability function ℓ, denoted $\ell^{\mathcal{N}}$, is straightforward. It is given by

$$\ell^{\mathcal{N}}(t, x, \Theta) = \frac{1}{M} \sum_{k=1}^{M} \beta_T^{t,\Theta,k} G(S^{t,x,\Theta,k}), \tag{37}$$

with $((\beta_\tau^{t,\Theta,k}, S_\tau^{t,x,\Theta,k})_{\tau \in \Gamma_G})_{1 \le k \le M}$ are i.i.d realisations of $(\beta_\tau^{t,\Theta}, S_\tau^{t,x,\Theta})_{\tau \in \Gamma_G}$, recall (4).

3.2.2. *Approximation of the Asset side*

The approximation of the asset side is slightly more involveld as it requires the computation of sensitivities with respect to the model parameters: $\frac{\partial \ell}{\partial x}$ and $\frac{\partial \ell}{\partial \theta_i}$, $i = 1, 2, 3$, see (19), (20) and (21). We choose to compute the sensitivities using a "weight" approach namely we express them as expectation of the discounted payoff times a well chosen random weight. Note that other techniques are available to compute these sensitivities e.g. *Automatic differentiation*. In our context, we have compared the two methods and observed that the weight approach is more efficient, see Section 5.4 in the Appendix.

We now describe how to obtain the sensitivities in a convenient form for Monte Carlo simulation. Recall that we consider a liability function ℓ of the form:

$$\ell(t, x, \Theta) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta} ds} G(S^{t,x,\Theta})\right], \tag{38}$$

where G depends upon $S^{t,x,\Theta}$ through its values on a finite set $\Gamma_G = \{\tau_0 = 0, \ldots, \tau_\kappa = T\} \subset [0, T]$. In the following, we assume for simplicity that $\tau_1 > t$. Otherwise, there are deterministic terms that are to be added, but the method remains the same.

The Delta: We want to compute:

$$\frac{\partial \ell}{\partial x}(t, x, \Theta) = \frac{\partial}{\partial x}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta} ds} G(S^{t,x,\Theta})\right] \tag{39}$$

and we have the following result.

Proposition 3.3. *For all* $(t, x, \Theta) \in [0, 1] \times \mathbb{R} \times \mathbb{R}^3$, *the following holds:*

$$\frac{\partial \ell}{\partial x}(t, x, \Theta) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta} ds} G(S^{t,x,\Theta}) H^{x,\Theta}\left(e^{-\int_t^{\tau_1} \xi_s^t ds}, S_{\tau_1}^{t,x,\Theta}\right)\right], \tag{40}$$

with

$$H^{x,\Theta}(a,y) = \frac{\Sigma_{1,2}^{-1}a + \Sigma_{2,2}^{-1} \times \left(\log(y/x) - \int_t^{\tau_1} \alpha_r^{t,\Theta}dr + \frac{\sigma^2}{2}(\tau_1 - t)\right)}{x}, \tag{41}$$

where Σ is the covariance matrix of $(A_{\tau_1}^t, X_{\tau_1}^{t,x,\Theta})$, see (46).

Proof. We write the expectation as an integral, remembering that we know the law of the couple $(\xi_u^t, A_u^t, X_u^{t,x,\Theta})$ conditionally upon $\mathcal{F}_s$:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta}ds} G(S^{t,x,\Theta})\right] \tag{42}$$

$$= e^{-\int_t^T \alpha_s^{t,\Theta}ds} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{\tau_1} \xi_s^t ds} \prod_{\ell=1}^{\kappa-1} e^{-\int_{\tau_\ell}^{\tau_{\ell+1}} \xi_s^t ds} G(S^{t,x,\Theta})\right]$$

$$= e^{-\int_t^T \alpha_s^{t,\Theta}ds} \int_{\mathbb{R}^{2\kappa}} e^{-a_1} \prod_{\ell=1}^{\kappa-1} e^{-a_{\ell+1}} G(e^{x_1}, \dots, e^{x_\kappa}) d\mathbb{Q}_{(A_u^t, X_u^{t,x,\Theta})_{u \in \Gamma_G}}(a_1, \cdots, a_\kappa, x_1, \cdots, x_\kappa)$$

$$= e^{-\int_t^T \alpha_s^{t,\Theta}ds} \int_{\mathbb{R}^{2\kappa}} e^{-a_1} \prod_{\ell=1}^{\kappa-1} e^{-a_{\ell+1}} G(e^{x_1}, \dots, e^{x_\kappa}) p^{\Theta}(t, 0, \log(x), 1, a_1, x_1) \times \cdots$$

$$\times p^{\Theta}(t_{\kappa-1}, 0, x_{\kappa-1}, t_\kappa, a_\kappa, x_\kappa) da_1 \cdots da_\kappa dx_1 \cdots dx_\kappa,$$

where $p^{\Theta}(s, a, x, u, ., .)$ is the density of the couple $(A_u^{t,a}, X_u^{t,x,\Theta})$, conditionally upon $\mathcal{F}_s$. We have previously seen that it is a Gaussian vector with explicit mean vector and covariance matrix. Thus, using Fubini's theorem, we get, since there is no dependence on x except in the first density:

$$\frac{\partial \ell(t, x, \Theta)}{\partial x} = e^{-\int_t^T \alpha_s^{t,\Theta}ds} \int_{\mathbb{R}^{2\kappa}} e^{-a_1} \prod_{\ell=1}^{\kappa-1} e^{-a_{\ell+1}} G(e^{x_1}, \dots, e^{x_\kappa}) \frac{\partial p^{\Theta}(t, 0, \log(x), 1, a_1, x_1)}{\partial x} \times \cdots$$

$$\times p^{\Theta}(t_{\kappa-1}, 0, x_{\kappa-1}, t_\kappa, a_\kappa, x_\kappa) da_1 \cdots da_\kappa dx_1 \cdots dx_\kappa. \tag{43}$$

Consequently, the sensibility of the discounted price with respect to the initial stock price is computed only by calculating the derivative of the density with respect to x.

We have

$$p^{\Theta}(t, 0, \log(x), s, a, y) = \frac{1}{\det(2\pi\Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left(((a, y) - \mu)\Sigma^{-1}((a, y) - \mu)^{\top}\right)\right), \tag{44}$$

where, thanks to (35)-(36),

$$\mu = \left(\mathbb{E}_t[A_{\tau_1}^t], \mathbb{E}_t[X_{\tau_1}^{t,x,\Theta}]\right) \tag{45}$$

$$\Sigma = \begin{pmatrix} \mathrm{Var}(A_{\tau_1}^t)_t & \mathrm{Cov}(A_{\tau_1}^t, X_{\tau_1}^{t,x,\Theta})_t \\ \mathrm{Cov}(A_{\tau_1}^t, X_{\tau_1}^{t,x,\Theta})_t & \mathrm{Var}(X_{\tau_1}^{t,x,\Theta})_t \end{pmatrix}. \tag{46}$$

Still by (35)-(36), we see that only $\mathbb{E}_t[X_{\tau_1}^{t,x,\Theta}]$ depends upon x. Thus we get:

$$\frac{\partial f(t, 0, \log(x), s, a, y)}{\partial x} = \frac{\Sigma_{1,2}^{-1}a + \Sigma_{2,2}^{-1} \times \left(\log(y/x) - \int_t^{\tau_1} \alpha_r^{t,\Theta}dr + \frac{\sigma^2}{2}(\tau_1 - t)\right)}{x} f(t, 0, \log(x), s, a, y). \tag{47}$$

Reinjecting this equality into (43) and rewriting the result in term of expectations, we finally get the result.
□

The function $\Delta(t_i, \cdot)$ is computed using the Monte Carlo estimator given in (40) as in (37) where we simulate in addition the weight H.

Sensitivities with respect to the interest rates curve. We consider now the derivatives with respect to the interest rates curve. For $i = 1, 2, 3$, we want to compute:

$$\frac{\partial \ell}{\partial \theta_i}(t, x, \Theta) = \frac{\partial}{\partial \theta_i} \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s^{t,\Theta} \mathrm{d}s} G(S^{t,x,\Theta}) \right], \quad i = 1, 2, 3.$$

Proposition 3.4. *For all $(t, x, \Theta) \in [0, 1] \times (0, \infty) \times \mathbb{R}^3$ and all $i = 1, 2, 3$, we have the following identity (where we have set $\tau_0 = t$):*

$$\frac{\partial \ell}{\partial \theta_i}(t, x, \Theta) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s^{t,\Theta} \mathrm{d}s} G(S^{t,x,\Theta}) H_i^{x,\Theta} \left((\xi_{\tau_l}^t, e^{-\int_{\tau_{l-1}}^{\tau_l} \xi_u^t \mathrm{d}u}, S_{\tau_l}^{t,x,\Theta})_{l=1,\ldots,\kappa} \right) \right], \tag{48}$$

with

$$H_i^{t,x,\Theta} \left((r_\ell, a_\ell, s_\ell)_{\ell=1,\ldots,\kappa} \right) = -\int_t^{\tau_\kappa} h_s^{t,i} \mathrm{d}s + \sum_{\ell=1}^{\kappa} \left(\int_{\tau_{\ell-1}}^{\tau_\ell} h_s^{t,i} \mathrm{d}s \right) \left((\Sigma^{\tau_{\ell-1},\tau_\ell})_{1,3}^{-1}(r_\ell - \mu_1^{\tau_{\ell-1},\tau_\ell}) \right.$$
$$+ (\Sigma^{\tau_{\ell-1},\tau_\ell})_{2,3}^{-1}(a_\ell - \mu_2^{\tau_{\ell-1},\tau_\ell})$$
$$\left. + (\Sigma^{\tau_{\ell-1},\tau_\ell})_{3,3}^{-1}(\log(s_\ell) - \mu_3^{\tau_{\ell-1},\tau_\ell}) \right), \tag{49}$$

where $\mu^{s,u}$ and $\Sigma^{s,u}$ are the mean and the covariance matrix of the Gaussian vector $(\xi_u^t, A_u^t, X^{t,x,\Theta})$ conditionally upon $\mathcal{F}_s$.

Proof. Performing a similar analysis as the one above,

$$\frac{\partial \ell}{\partial \theta_i}(t, x, \Theta) = \frac{\partial e^{-\int_t^T \alpha_s^{t,\Theta} \mathrm{d}s}}{\partial \theta_i} \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T \xi_s^t \mathrm{d}s} G(S^{t,x,\Theta}) \right] + e^{-\int_t^T \alpha_s^{t,\Theta} \mathrm{d}s} \int_{\mathbb{R}^{2\kappa}} e^{-a_1} \prod_{\ell=1}^{\kappa-1} e^{-a_{\ell+1}} G(e^{x_1}, \ldots, e^{x_\kappa})$$
$$\frac{\partial \left(p^{\Theta}(t, 0, \log(x), 1, a_1, x_1) \ldots p^{\Theta}(t_{\kappa-1}, a_{\kappa-1}, x_{\kappa-1}, t_\kappa, a_\kappa, x_\kappa) \right)}{\partial \theta_i} \mathrm{d}a_1 \cdots \mathrm{d}a_\kappa \mathrm{d}x_1 \cdots \mathrm{d}x_\kappa. \tag{50}$$

Here, the computations are more involved since every density depends upon the $a_i, i = 1, 2, 3$, but the idea is the same as before.

The only difference is that, when we use (35)-(36) to differentiate, we see that the short term itself appears in the formulae. This is not a problem as we can rewrite the previous integral as an integral over $\mathbb{R}^{3\kappa}$, with the short rate process taken at times $\tau_l, l = 1, \ldots, \kappa$, as new variables to integrate against. □

The quantity $\frac{\partial \ell}{\partial \theta_i}$ $(i = 1, 2, 3)$ is computed using the Monte Carlo estimator of the formula (48). Then solving the system (21) allows to obtain the coefficients ρ_1, ρ_2, ρ_3.

This method to calculate derivatives allows us to compute the function $\ell(t, x, \Theta)$ and its four derivatives with only one Monte Carlo simulation. Furthermore, given a risk-neutral scenario, each quantity involved in formulae (41) and (49) can be exactly computed by integrating the elementary functions $h^{t,i}$ and by inverting real symmetric matrices of size 3×3. Thus, the weight functions $H^{t,x,\Theta}, H_i^{t,x,\Theta}$ are easily and accurately computed.

3.3. **Sparse Grid Approach**

The *nested simulation* approach requires the approximation of the function ℓ and its derivative $\frac{\partial \ell}{\partial x}$, $\frac{\partial \ell}{\partial \theta_i}$, $i = 1, 2, 3$ for each path $(\bar{S}_t^j, \bar{\Theta}_t^j)_{t \in \Gamma}$ of the market parameters. These values are computed on the fly which is quite time consuming.

We suggest here an alternative method which will pre-compute the quantities of interest (ℓ and its derivatives) and store them. The requested value for a given market parameter will then be obtained by an interpolation procedure.

A first simple approach is to consider an equidistant grid of the domain $A := \prod_{l=1}^{d}[m_p, M_p]$ which is a truncation of the support of $\mathcal{X}$ ($\mathbb{R}^d$, $d = 4$ in our setting). Then one can use a multi-linear interpolation to reconstruct the function in the whole space. If one sets 2^p points in one dimension, the total number of points will be 2^{dp} for one function at one given time and overall $(n+1)2^{dp}$ to store (here $d = 4$). This will become rapidly too big, especially if one allows the number of market parameters to grow. This is a typical example of the "curse of dimensionality" encountered in numerical analysis when dealing with problem of high dimension.

Instead of considering a regular grid, we shall rely on the use of *sparse grid*, which allows to lower the number of points required to store the numerical approximation of the function. We now present rapidly the main concepts linked to sparse grids, see [2] for a comprehensive survey. In our numerical examples, see Section 4, the sparse grid will be computed using the StOpt C++ library [5].

For each multi-index $\mathbf{k} \leq \mathbf{l}$, we define a grid mesh $h_{\mathbf{k}} = 2^{-\mathbf{k}}$ and grid points

$$\check{y}_{\mathbf{k},\mathbf{i}} = (m_1 + i_1(M_1 - m_1)h_{k_1}, \ldots, m_d + i_d(M_d - m_d)h_{k_d}), \ \mathbf{0} \leq \mathbf{i} \leq 2^{\mathbf{k}}.$$

Using the *hat* function,

$$y \in \mathbb{R} \mapsto \phi(y) := \begin{cases} 1 - |y| & \text{if } y \in [-1, 1] \\ 0 & \text{otherwise} \end{cases} \tag{51}$$

and we can associate to the previous grid a set of nodal basis function:

$$y \in \mathbb{R}^d \mapsto \phi_{\mathbf{k},\mathbf{i}}(y; A) = \prod_{l=1}^{d} \phi\left(\frac{y_l - \check{y}_{\mathbf{k},\mathbf{i}}^l}{2^{-i_l}}\right).$$

When using a "full" linear interpolation, the function is approximated using the whole set of nodal basis function at the finest level $\mathbf{l}$. Instead, we consider the sparse grid nodal space of order p defined by

$$\mathcal{V}_\kappa := \text{span}\{\phi_{\mathbf{l},\mathbf{j}}; (\mathbf{l}, \mathbf{j}) \in \mathcal{I}_p(A)\},$$

where

$$\mathcal{I}_\kappa := \Big\{(\mathbf{l}, \mathbf{j}): \quad 0 \leq \sum_{i=1}^{d} l_i \leq p; \quad \mathbf{0} \leq \mathbf{j} \leq 2^{\mathbf{l}};$$

$$(l_i > 0 \text{ and } j_i \text{ is odd}) \text{ or } (l_i = 0), \text{ for } i = 1, \ldots, d\Big\}. \tag{52}$$

For a function $\psi : A \to \mathbb{R}$ with support in A, we define its associated $\mathcal{V}_\kappa$-interpolator by

$$\pi_{\mathcal{V}_\kappa}^A[\psi](y) := \sum_{(\mathbf{l},\mathbf{j}) \in \mathcal{I}_\kappa(A)} \theta_{\mathbf{l},\mathbf{j}}(\psi; A)\phi_{\mathbf{l},\mathbf{j}}(y; A) \tag{53}$$

where the operator $\theta_{\mathbf{l},\mathbf{j}}$ can be defined recursively in terms of r, the dimension of $\mathbf{l}$, by:

$$\theta_{\mathbf{l},\mathbf{j}}(\psi; A) = \begin{cases} \psi(\breve{y}_{\mathbf{l},\mathbf{j}}); & r = 0 \\ \theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \breve{y}^r_{l_r,j_r}); A-); & l_r = 0 \\ \theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \breve{y}^r_{l_r,j_r}); A-) - \frac{1}{2}\theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \breve{y}^r_{l_r,j_r-1}); A-) \\ \qquad\qquad - \frac{1}{2}\theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \breve{y}^r_{l_r,j_r+1}); A-); & l_r > 0 \end{cases} \tag{54}$$

where, for a hypercube $A = \prod_{l=1}^d [m_l, M_l]$, $A- := \prod_{l=1}^{d-1} [m_l, M_l]$ and for a multi-index $\mathbf{k}$ with dimension $r \geq 1$, $\mathbf{k}- = (k_1, \ldots, k_{r-1})$.

Let us now introduce the approximation that we use to compute the loss distribution, namely

$$\ell^{\mathcal{S}}(\cdot) := \pi_{\mathcal{V}_\kappa}[\ell^{\mathcal{N}}](\cdot), \quad \left(\frac{\partial \ell}{\partial x}(t, \cdot)\right)^{\mathcal{S}} := \pi_{\mathcal{V}_\kappa}\left[\left(\frac{\partial \ell}{\partial x}(t, \cdot)\right)^{\mathcal{N}}\right](\cdot) \tag{55}$$

$$\text{and} \quad \left(\frac{\partial \ell}{\partial \theta_i}(t, \cdot)\right)^{\mathcal{S}} := \pi_{\mathcal{V}_\kappa}\left[\left(\frac{\partial \ell}{\partial \theta_i}(t, \cdot)\right)^{\mathcal{N}}\right](\cdot) \; i \in \{1, 2, 3\} \;, \; t \in \Gamma \;.$$

These functions are built by computing the coefficients appearing in (53), which are then stored in memory. For $\ell^{\mathcal{S}}$ say, this amounts to compute the function $\ell^{\mathcal{N}}$ on the sparse grid $\mathcal{V}_p$ and this is done by Monte Carlo simulation as in the previous approach, recall the definition of $\ell^{\mathcal{N}}$ in (37).

Complexity. The main limitation of this method is the memory usage and the time to pre-compute the functions. This is proportional to the number of point in the grid. This number can be estimated to be of $O(2^{\kappa-d+1} \frac{(\kappa-d+1)^{d-1}}{(d-1)!})$, see Proposition 4.1 in [2]. Let us insist on the fact that this is done "offline" compared to the *nested simulation* approach. For the "online" computations, the main effort is put in the evaluation of the function which is slightly more evolved than a linear interpolation and is of $O(\kappa)$ where κ is the maximum level that is chosen.

Remark 3.5. The computations at each point being independant, this Sparse Grid approach can be easily parallelised, hence improving further the gain of time observed in Subsection 4.1.

3.4. Convergence study

The goal is to obtain a reasonable approximation of the risk associated to the loss distribution of the balance sheet in an efficient way. In this section, we explain why the methods introduced above are indeed good approximations of the risk indicators. We also study theoretically the numerical complexity of both methods in terms of memory and time consumption.

3.4.1. *Error analysis*

For the risk estimation, we will investigate a root Mean Square Error (rMSE) of the following form

$$\epsilon^\upsilon := \mathbb{E}\big[|\varrho(p_1 \sharp \eta) - \varrho(p_1^\upsilon \sharp \eta^N)|^2\big]^{\frac{1}{2}} \;, \text{ for } \upsilon \in \{\mathcal{N}, \mathcal{S}\}.$$

The expectation operator $\mathbb{E}[\cdot]$ acts under $\mathbb{P}^{\otimes N} \otimes \mathbb{Q}^{\otimes M}$, namely it averages both on the simulation of the market parameters under the *real-world* measure used for calibration and the risk neutral evolution of the market model under the pricing measure.

The first observation is that under reasonable assumptions on the risk measure used in the risk indicator, the error performed in the numerical simulation can be separated in two main contribution: the error due to the sampling of the loss distribution coming from the sampling of the market parameters and the error made when approximating the different pricing and hedging functions.

Lemma 3.6. *Assume that ϱ has a* Monotonic *and* Cash Invariant *lift $\Re$, then*

$$\epsilon^v \leq \mathbb{E}\big[|\varrho(p_1 \sharp \eta) - \varrho(p_1 \sharp \eta^N)|^2\big]^{\frac{1}{2}} + \mathbb{E}\bigg[\sup_{1 \leq j \leq N} |p_1(\mathcal{X}^j) - p_1^v(\mathcal{X}^j)|^2\bigg]^{\frac{1}{2}} .$$

Proof. We denote by $\widehat{\mathcal{X}}^N(\omega)$ the random variable with distribution $\eta^N(\omega)$. Note that

$$p_1(\widehat{\mathcal{X}}^N(\omega)) \leq p_1^v(\widehat{\mathcal{X}}^N(\omega)) + \sup_{1 \leq j \leq N} |p_1(\mathcal{X}^j(\omega)) - p_1^v(\mathcal{X}^j(\omega))| \tag{56}$$

which leads to

$$\Re(p_1(\widehat{\mathcal{X}}^N(\omega))) \leq \Re(p_1^v(\widehat{\mathcal{X}}^N(\omega))) + \sup_{1 \leq j \leq N} |p_1(\mathcal{X}^j(\omega)) - p_1^v(\mathcal{X}^j(\omega))| .$$

By symmetry, we easily get

$$|\varrho(p_1 \sharp \eta) - \varrho(p_1^v \sharp \eta^N(\omega))| \leq |\varrho(p_1 \sharp \eta) - \varrho(p_1 \sharp \eta^N(\omega))| + \sup_{1 \leq j \leq N} |p_1(\mathcal{X}^j(\omega)) - p_1^v(\mathcal{X}^j(\omega))| \tag{57}$$

and then the proof is concluded using Minkowski inequality. $\qquad\square$

The error due to the approximation of the function p_1 is well understood when the function is smooth enough. Note that the asset side of the function is quite involved and we will not attempt to obtain the condition for smoothness of the overall function p_1. We will now simply review the error done on the liability part $\ell(1, \cdot)$ assuming that the mapping G is bounded and

$$(x, \Theta) \to \beta_T^{1,\Theta} G(S^{1,x,\Theta}) \in \mathcal{C}_b^2. \tag{58}$$

Even though, this cannot be almost surely true in the model presented above, we will assume that $\beta_T^{1,\Theta}$ is bounded in the discussion below. A more precise analysis should take care of these extreme events arising with small probability. Another possibility would be to force the interest rate to be non-negative, by truncation or by considering a CIR type of model for (5).

Lemma 3.7. *Assume that* (58) *holds true. Recall the definition of ℓ^N in* (37)*, then*

$$\mathbb{E}\bigg[\max_{1 \leq j \leq N} |\ell_1(\mathcal{X}^j) - \ell_1^N(\mathcal{X}^j)|^2\bigg]^{\frac{1}{2}} \leq C\sqrt{\frac{\log(N)}{M}} . \tag{59}$$

Proof. We denote by c the bound on the mapping $(x, \Theta) \to \beta_T^{1,\Theta} G(S^{1,x,\Theta})$ (recall the discussion after equation (58)) and thus the bound on ℓ_1. For the reader's convenience, we introduce

$$\Sigma_M^j := \sum_{k=1}^M \ell_1(\mathcal{X}^j) - \beta_T^{t,\mathcal{X}^j,k} G(S_T^{t,\mathcal{X}^j,k}),$$

and observe that $\mathbb{E}\big[\Sigma_M^j\big] = 0$ and recall that the (Σ_M^j) are i.i.d. We have, using Hoeffding Inequality,

$$\mathbb{E}\big[\mathbf{1}_{\{|\Sigma_M^j|^2 > z\}}\big] \leq 2\exp\left(-\frac{z}{cM}\right) .$$

Using the independence property, we obtain

$$\mathbb{E}\Big[\mathbf{1}_{\{\max_j |\Sigma_M^j|^2 \leq z\}}\Big] \geq \Big(1 - 2\exp\Big(-\frac{z}{cM}\Big)\Big)^N . \tag{60}$$

Now set $\xi := cM\log(N)$ and compute

$$\mathbb{E}\Big[\max_{1 \leq j \leq N} |\Sigma_M^j|^2\Big] = \int_0^\infty \mathbb{E}\Big[\mathbf{1}_{\{\max_j |\Sigma_M^j|^2 > z\}}\Big]\,\mathrm{d}z$$

$$\leq \xi + \int_\xi^\infty \mathbb{E}\Big[\mathbf{1}_{\{\max_j |\Sigma_M^j|^2 > z\}}\Big]\,\mathrm{d}z$$

$$\leq \xi + \int_\xi^\infty \Big\{1 - \Big(1 - 2\exp\Big(-\frac{z}{cM}\Big)\Big)^N\Big\}\,\mathrm{d}z.$$

Now, we observe that for $N \geq 2$, $2\exp\Big(-\frac{z}{cM}\Big) \leq 1$ for $z \geq \xi$. Using the fact that $1 - (1-u)^N \leq Nu$, for $u \in [0,1]$, we get

$$\mathbb{E}\Big[\max_{1 \leq j \leq N} |\Sigma_M^j|^2\Big] \leq \xi + 2N\int_\xi^\infty \exp\Big(-\frac{z}{cM}\Big)\,\mathrm{d}z$$

which leads to

$$\mathbb{E}\Big[\max_{1 \leq j \leq N} |\Sigma_M^j|^2\Big] \leq \xi + 2cM,$$

and concludes the proof.

$\square$

We conclude this section by giving the overall estimation error induced by the numerical procedure above. We will admit that the upper bound for the error given for $\ell(1, \cdot)$ in Lemma 3.7 holds true for the PnL function $p(1, \cdot)$ with a scaling by n coming from the number of rebalancing date.

Theorem 3.8. *Assume that ϱ_h is a spectral risk measure. Then, the following holds, for some $\alpha > 0$,*

(1) *for the* Nested Simulation *approach*

$$\epsilon^{\mathcal{N}} \leq C\left(\frac{1}{N^\alpha} + n\sqrt{\frac{\log(N)}{M}}\right) ; \tag{61}$$

(2) *for the* Sparse Grid *approach with maximum level κ*

$$\epsilon^{\mathcal{S}} \leq C\left(\frac{1}{N^\alpha} + n\left\{\sqrt{\frac{\log(N)}{M}} + 2^{-2\kappa}(\kappa - d + 1)^{(d-1)}\right\}\right) . \tag{62}$$

Proof. 1. We first show that

$$\mathbb{E}\Big[\max_{1 \leq j \leq N} |\ell_1(\mathcal{X}^j) - \ell_1^{\mathcal{S}}(\mathcal{X}^j)|^2\Big]^{\frac{1}{2}} \leq C\left(\sqrt{\frac{\log(N)}{M}} + 2^{-2\kappa}\kappa^{(d-1)}\right) . \tag{63}$$

Indeed, we have that

$$\ell_1^{\mathcal{S}} = \pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}]$$

$$= \ell_1^{\mathcal{N}} + \pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] - \ell_1^{\mathcal{N}} .$$

And we observe that

$$\pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] - \ell_1^{\mathcal{N}} = \frac{1}{M} \sum_{j=1}^{M} \pi_{\mathcal{V}_\kappa}[e^{-\int_1^T r_s^{1,\cdot,k} ds} G(S^{1,\cdot,k})] - e^{-\int_1^T r_s^{1,\cdot,k} ds} G(S^{1,\cdot,k})$$

Let us denote by $(x,\Theta) \mapsto \phi^k(x,\Theta) = e^{-\int_1^T r_s^{1,\Theta,k} ds} G(S^{1,x,\Theta,k})$ which is a random function as it depends on the random realisation of the (r,S) process. Under (58), ϕ^k is smooth enough to apply the results in Proposition 4.1 in [2] and we obtain

$$|\pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] - \ell_1^{\mathcal{N}}|_\infty \leq \frac{1}{M} \sum_{j=1}^{M} |\pi_{\mathcal{V}_\kappa}[\phi^k] - \phi^k|_\infty \leq C 2^{-2\kappa} \kappa^{d-1} . \tag{64}$$

We then observe that

$$\mathbb{E}\left[\max_{1 \leq j \leq N} |\ell_1(\mathcal{X}^j) - \ell_1^{\mathcal{S}}(\mathcal{X}^j)|^2 \right]^{\frac{1}{2}} \leq C \left(\mathbb{E}\left[\max_{1 \leq j \leq N} |\ell_1(\mathcal{X}^j) - \ell_1^{\mathcal{N}}(\mathcal{X}^j)|^2 \right]^{\frac{1}{2}} + |\pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] - \ell_1^{\mathcal{N}}|_\infty \right) .$$

The proof of (63) is concluded by combining the above inequality with (64) and Lemma 3.7.
2. We now prove (61). Applying Lemma 3.6, we obtain

$$\epsilon^{\mathcal{N}} \leq \mathbb{E}\big[|\varrho(p_1 \sharp \eta) - \varrho(p_1 \sharp \eta^N)|^2\big]^{\frac{1}{2}} + \mathbb{E}\left[\max_{1 \leq j \leq N} |\ell_1(\mathcal{X}^j) - \ell_1^{\mathcal{N}}(\mathcal{X}^j)|^2 \right]^{\frac{1}{2}} . \tag{65}$$

The second term in the right-hand side of the above inequality is controlled by using Lemma 3.7. We now study the first term in the right-hand side, which is the error introduced by the sampling of the loss distribution. Applying Corollary 11 in [7] to the spectral risk measure, we first get

$$\mathbb{E}\big[|\varrho(p_1 \sharp \eta) - \varrho(p_1 \sharp \eta^N)|^2\big]^{\frac{1}{2}} \leq C \mathbb{E}\big[\mathcal{W}_2(\eta, \eta^N)^2\big]^{\frac{1}{2}} .$$

We then use Theorem 1 in [4] to bound the Wasserstein distance, which concludes the proof for this step.
3. To prove (62), we follow similar arguments as in step 2. but using (63) instead of invoking Lemma 3.7. $\square$

Remark 3.9. We can compare the bound obtained for the nested simulation with the ones in [6]. Using a different approach, the authors prove a very nice bound on the overall error given by

$$C \left(\frac{1}{\sqrt{N}} + \frac{1}{M} \right) ,$$

for the $V@R$ (which is not a spectral risk measure) and $AV@R$. Note that the term $\frac{1}{M}$ is obtained by cancellation of the first order term through an error expansion. It would be interesting to understand under which assumptions their bound can be retrieved in our setting of general spectral risk measure. This topic is left for further research.

We conclude this Section by a short account on the numerical complexity of the two methods.

The *Nested Simulation* approach is a pure "online" method which is very simple to implement but has a huge drawback in term of running time. Each time an estimation is requested the numerical complexity is overall of nNM, where recall n is the number of rebalancing date, M the number of sample for the risk neutral simulation and N the number of sample for the real-world simulation. The memory requirements comes only from the estimation of the loss distribution and are of order N.

As already mentioned, the *Sparse Grid* approach is both an "online" and "offline" method. In terms of memory requirement, it is thus greedier than the *nested simulation* approach. On top of the memory needed to store the sample distribution (of order N), memory is also needed to store the sparse grid approximation $p^{\mathcal{S}}$, the requirement are of order $O(n2^{\kappa-d+1}\frac{(\kappa-d+1)^{d-1}}{(d-1)!})$. In term of running time, the gain is important as the complexity of evaluating $p^{\mathcal{S}}$ is of $O(\kappa)$ only, where κ is the maximum level used.

4. Numerics

In the numerical applications below, we will compare the loss distribution obtained via our two numerical procedures by computing the Wasserstein distance between the two empirical distributions. Since the loss distribution is one-dimensional, we use the following formula [8]: for two probability distribution on $\mathbb{R}$, η and $\tilde{\eta}$,

$$W_2(\eta, \tilde{\eta}) = (\int_0^1 |F_\eta^{-1}(u) - F_{\tilde{\eta}}^{-1}(u)|^2 \mathrm{d}u)^{\frac{1}{2}} . \tag{66}$$

In the setting of empirical distributions, the above distance is easily computed. Suppose $\eta = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ and $\tilde{\eta} = \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$.
We straightforwardly compute

$$W_2(\eta, \tilde{\eta})^2 = \sum_{i=1}^N \int_{\frac{i-1}{N}}^{\frac{i}{N}} |F_\eta^{-1}(u) - F_{\tilde{\eta}}^{-1}(u)|^2 \mathrm{d}u = \frac{1}{N}\sum_{i=1}^N |x_{(i)} - y_{(i)}|^2, \tag{67}$$

where the subscript (i) refers to the i-th order statistic of the distribution, since $x_{(i)}$ (resp. $y_{(i)}$) is simply the $\frac{i}{N}$-th quantile of η (resp. $\tilde{\eta}$).

Besides the Wasserstein distance between the two empirical distributions, we will also compare the estimated V@R and AV@R, which are computed in a similar way. Indeed, for $\alpha \in (0,1]$, we have:

$$V@R_\alpha(\eta) = F_\eta^{-1}(\alpha) = x_{(i_\alpha)}, \tag{68}$$

where $\frac{i_\alpha - 1}{N} < \alpha \leq \frac{i_\alpha}{N}$, $i_\alpha \in \{1, \dots, N\}$.
For a given $\alpha \in (0,1]$, we observe that

$$AV@R_\alpha(\eta) = \frac{1}{1-\alpha}\int_\alpha^1 V@R_p(\eta)\mathrm{d}p = \frac{1}{1-\alpha}\left(\int_\alpha^{\frac{i_\alpha}{N}} V@R_p(\eta)\mathrm{d}p + \sum_{i=i_\alpha}^{N-1}\int_{\frac{i}{N}}^{\frac{i+1}{N}} V@R_p(\eta)\mathrm{d}p\right)$$

which leads to

$$AV@R_{\bar{\alpha}}(\eta) = \frac{1}{1-\alpha}\left(\{\frac{i_\alpha}{N} - \alpha\}x_{(i_\alpha)} + \frac{1}{N}\sum_{i=i_\alpha}^{N-1} x_{(i+1)}\right). \tag{69}$$

Using the formulae (67), (68) and (69), we will now present numerical results showing the efficiency and usefulness of the *sparse grid* approach. We first start with a comparison with the classically used *nested simulation* approach.

4.1. **Sparse grid approach versus nested simulations approach**

We computed the empirical distribution of the PnL at horizon 1 year using the *nested simulations* approach, recall Section 3.2, and the *sparse grid* approach, recall Section 3.3.
For both methods, we used a sample of size $N = 11000$ describing the *real-world* evolution of S and Θ, recall

M	100	500	1000	2000	10000
Wasserstein distance	0.73	0.20	0.09	0.05	0
V@R/AV@R $\alpha = 0.005$	7.03 / 8.75	4.26 / 5.08	3.95 / 4.33	3.72 / 4.01	3.56 / 3.85
V@R/AV@R $\alpha = 0.01$	5.82 / 7.52	3.88 / 4.57	3.68 / 4.07	3.48 / 3.80	3.37 / 3.65
V@R/AV@R $\alpha = 0.05$	4.01 / 5.22	3.08 / 3.62	2.93 / 3.36	2.84 / 3.23	2.77 / 3.13
V@R/AV@R $\alpha = 0.1$	3.31 / 4.42	2.72 / 3.24	2.61 / 3.06	2.54 / 2.96	2.49 / 2.88

TABLE 1. Comparison of the metrics obtained with Nested simulations for varying risk-neutral simulation sample size.

Section 2.4.

For the nested simulations approach, using the overall error bound given in [6] we want to approximate the risk-neutral expectations with Monte Carlo simulations with samples of size $M \simeq \sqrt{N}$, that is $M = 100$. In practice however, we observe that convergence has not occured yet and we observe non-neglectible changes in the risk measures taken into account, see Table 1. In this table, we compute the Wasserstein distance with respect to the distribution obtained for $M = 10000$. Operational constraints do not allow us to change the size of the sample used for the Monte Carlo simulations under $\mathbb{P}$, We thus consider a sample of size $M = 2000$ for the *risk-neutral* Monte Carlo simulations. Following Proposition 2.9, we calibrated a Gaussian model such that $(X_1, (\theta_1)_1, (\theta_2)_1, (\theta_3)_1)$ has mean and covariance matrix given by:

$$\mu = (4.1 \times 10^{-5}, 0.01, 0.03, 0.01), \tag{70}$$

$$V = \begin{pmatrix} 0.004 & 3.2 \times 10^{-5} & 6.76 \times 10^{-6} & 0.000008 \\ 3.2 \times 10^{-5} & 3.1 \times 10^{-5} & 1.82 \times 10^{-5} & 1.5 \times 10^{-5} \\ 6.76 \times 10^{-6} & 1.82 \times 10^{-5} & 7.5 \times 10^{-5} & 8.1 \times 10^{-6} \\ 0.000008 & 1.5 \times 10^{-5} & 8.1 \times 10^{-6} & 2.7 \times 10^{-5} \end{pmatrix} \tag{71}$$

The risk-neutral simulations were computed by a Monte Carlo procedure, computed with the exact formulae in the Hull & White and Black & Scholes setting we used, recall Proposition 3.2. The volatility parameter used in the Black & Scholes model is set to $\sigma = 0.3$ while the parameters defining the Hull & White model are set to $a = 0.05$ and $b = 0.01$. Last, the covariation parameter between the two Brownian motions is set to $\rho = 0$. In this setting, the nested simulations method was tested with the Put Lookback option described in 2.1, with maturity $T = 30$ years. Figure 2 shows the outcome PnL's distribution.

We next looked at the grid method. Figure 3 shows the outcome PnL's distributions for sparse grids of level $1, 2, 3$, which respectively have cardinal 81, 297, 945. For each level, we chose the number of risk-neutral simulations M so that the error induced by the Monte Carlo estimation is small compared with the sparse interpolation error, meaning that we can run the program multiple times without changing significantly the outcome. Furthermore, as in the *nested simulations* case, we choose M so that increasing M has no effect on the distribution. Empirically, we chose $M = 20000$ for the sparse grid of level 1, 2 or 3. Figure 4 compares the distribution obtained with nested simulations with the distribution obtained with the sparse grid of level 3. Table 2 shows computational times comparison, and Table 3 shows V@R and AV@R comparison for the empirical distributions obtained in each case.

We observe that the computational time on the Sparse grid of level 3 with $M = 20000$ is similar to the one for the Nested simulations with only $M = 2000$. Moreover, a significant gain in time is obtained by the use of the sparse grid of level 2 only, which already gives good results, see Table 2. As already observed in Remark 3.5, this gain in time can be further improved by parallelisation of the computations. In addition, we observe that, once the computations on the grid are done, then the PnL distribution is almost straightforwardly obtained. This is a key feature of the method since the computations on the grid are to be kept. Indeed, if one needs to change the distribution of (S, Θ) under $\mathbb{P}$, say because the view of the risk management on the evolution of the
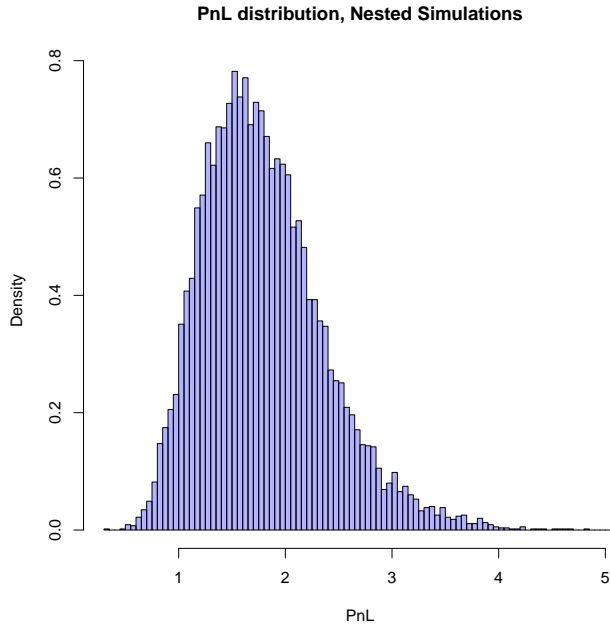
FIGURE 2. PnL distribution, nested simulations

market parameters has changed, then they can be re-used easily. In the next section, we give an application in this direction.

Level of the sparse grid	$l = 1$	$l = 2$	$l = 3$	Nested simulations
Computations on the grid	9 min 30 sec	35 min 10 sec	1h 58 min	
Computation of the PnL distribution	2 sec	4 sec	8 sec	2h 15min

TABLE 2. Computational times

4.2. Model risk

As mentioned above, an interesting feature of the *sparse grid* approach developed in this paper is the ability to change the distribution of the processes $X = log(S)$ and Θ under the *real-world* probability $\mathbb{P}$. In this Section, we use the model described in Section 2.4 to simulate a first sample. Then, we consider some uncertainty over the estimated moments μ, V of (X_1, Θ_1) used to calibrate the gaussian model: we only assume that the

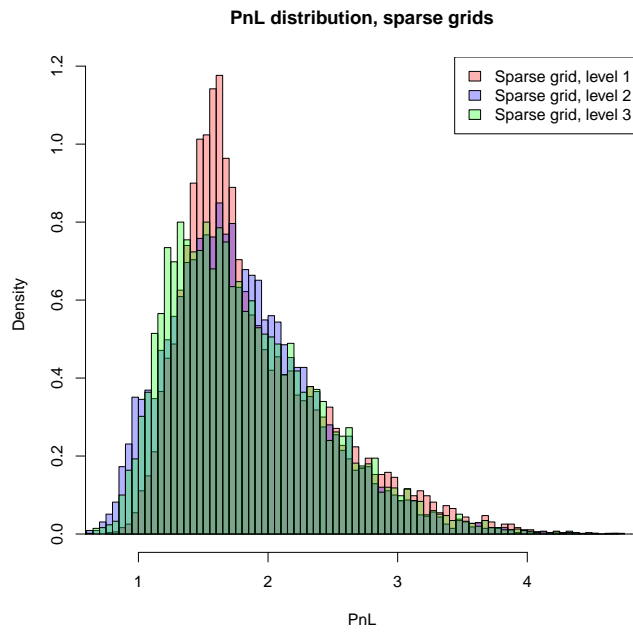| Level of the sparse grid	$l = 1$	$l = 2$	$l = 3$	Nested simulations
Wasserstein distance	0.13	0.06	0.05	0
V@R/AV@R $\alpha = 0.005$	3.87 / 4.09	3.64 / 3.94	3.72 / 3.96	3.72 / 4.01
V@R/AV@R $\alpha = 0.01$	3.68 / 3.93	3.43 / 3.73	3.54 / 3.79	3.48 / 3.80
V@R/AV@R $\alpha = 0.05$	3.08 / 3.43	2.81 / 3.17	2.95 / 3.30	2.84 / 3.23
V@R/AV@R $\alpha = 0.1$	2.73 / 3.16	2.54 / 2.92	2.64 / 3.04	2.54 / 2.96

TABLE 3. Comparison of the empirical distributions

PnL distribution, sparse grids

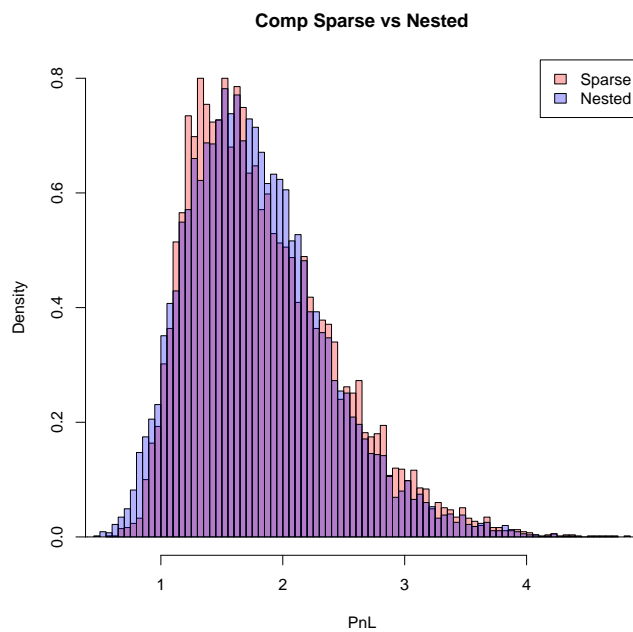FIGURE 3. PnL distribution, sparse grids

Comp Sparse vs Nested

FIGURE 4. PnL distribution, nested simulations versus sparse grid method

true moments lie in centered intervals around the estimations. In practice, we consider intervals of the form $[m \times 0.95, m \times 1.05]$, where m is the estimated moment under consideration.

To better understand the risk associated with this uncertainty under $\mathbb{P}$, we simulate two "extreme" new samples of (X, Θ), where every moment taken into account to calibrate the model are multiplied by 0.95 (resp. 1.05), and, thanks to the grid computations done before with the initial model, we are in position to compute almost instantaneously the empirical PnL distributions associated with these two new samples.

Table 4 shows the Wasserstein distance between the initial distributions and the two obtained for the shifted parameters, and the V@R and AV@R obtained at different quantile levels. We observe that with these small change the distribution are quite close to each other. The main discrepancy are obtained for the diminished moments.

Model	Initial model	Diminished moments	Augmented moments
Wasserstein distance	0	0.030	0.0070
V@R/AV@R $\alpha = 0.005$	3.37 / 3.63	3.43 / 3.81	3.33 / 3.67
V@R/AV@R $\alpha = 0.01$	3.18 / 3.45	3.17 / 3.54	3.16 / 3.45
V@R/AV@R $\alpha = 0.05$	2.61 / 2.95	2.63 / 2.97	2.60 / 2.93
V@R/AV@R $\alpha = 0.1$	2.34 / 2.71	2.37 / 2.73	2.32 / 2.69

TABLE 4. Comparison of the empirical distributions

5. Appendix

5.1. **Proof of** (6)

In this subsection, we shall give the proof of Proposition 2.1 for completeness. We remind that in the Hull-White model, the dynamics of the short rate is given by the following:

$$dr_s^{t,\Theta} = a(\mu_s^{t,\Theta} - r_s^{t,\Theta})ds + bdB_s \tag{72}$$

with $a, b \in \mathbb{R}$. We will prove that the mean-reverting θ_s can be calibrated by forward interest rate curve $f^\Theta(t,s)$ by:

$$\mu_s^{t,\Theta} = f^\Theta(t,s) + \frac{1}{a}\frac{\partial f^\Theta(t,s)}{\partial s} + \frac{b^2}{2a^2}(1 - e^{-2a(s-t)}) \tag{73}$$

The method is to express the price of the zero-coupon bond $P(t,s)$ in the following way:

$$\mathbb{E}[\exp(-\int_t^s r_u^{t,\Theta}du)] = P(t,s) = \exp(-\int_t^s f^\Theta(t,u)du) \tag{74}$$

Then by comparing both sides, we can determine $f^\Theta(t,s)$. First, it is easy to find out that the solution to (72) is

$$r_s^{t,\Theta} = r_t^{t,\Theta}e^{-a(s-t)} + a\int_t^s \mu_u^{t,\Theta}e^{-a(s-u)}du + b\int_t^s e^{-a(s-u)}dB_u$$

Then by straightforward calculation we have that

$$\int_t^s r_u^{t,\Theta}du = \frac{r_t^{t,\Theta}}{a}(1 - e^{-a(s-t)}) + \int_t^s \mu_u^{t,\Theta}(1 - e^{-a(s-u)})du + \frac{b}{a}\int_t^s (1 - e^{-a(s-u)})dB_u$$

So $\int_t^s r_u^{t,\Theta}du$ follows a normal distribution with mean

$$\mathbb{E}[\int_t^s r_u^{t,\Theta}du] = \frac{r_t^{t,\Theta}}{a}(1 - e^{-a(s-t)}) + \int_t^s \mu_u^{t,\Theta}(1 - e^{-a(s-u)})du$$

and variance

$$\mathbb{V}[\int_t^s r_u^{t,\Theta}du] = \frac{b^2}{a^2}\int_t^s (1 - e^{-a(s-u)})^2du$$

Now comparing the both sides of (74), we have that

$$\int_t^s f^\Theta(t,u)du = \mathbb{E}[\int_t^s r_u^{t,\Theta}du] - \frac{1}{2}\mathbb{V}[\int_t^s r_u^{t,\Theta}du]$$

Thus

$$\begin{aligned}
f^\Theta(t,s) &= \frac{\partial}{\partial s}\mathbb{E}[\int_t^s r_u^{t,\Theta}du] - \frac{1}{2}\frac{\partial}{\partial s}\mathbb{V}[\int_t^s r_u^{t,\Theta}du] \\
&= r_t^{t,\Theta}e^{-a(s-t)} + a\int_t^s \mu_u^{t,\Theta}e^{-a(s-u)}du - \frac{b^2}{2a^2}2a\int_t^s (e^{-a(s-u)} - e^{-2a(s-u)})du \\
&= r_t^{t,\Theta}e^{-a(s-t)} + a\int_t^s \mu_u^{t,\Theta}e^{-a(s-u)}du - \frac{b^2}{2a^2}(1 - e^{-a(s-t)})^2 \tag{75}
\end{aligned}$$

By straightforward differentiation, we have

$$\frac{\partial}{\partial s}f^\Theta(t,s) = -ar_t^{t,\Theta}e^{-a(s-t)} - a^2\int_t^s \mu_u^{t,\Theta}e^{-a(s-u)}du + a\mu_s^{t,\Theta} - \frac{b^2}{a}(e^{-a(s-t)} - e^{-2a(s-t)}) \tag{76}$$

Now by (75) and (76), we can easily verify that (73) is valid.

5.2. **Proof of Proposition 2.9**

We provide a recursive proof of Proposition 2.9, which allows to compute effectively the coefficients defining the processes.

Suppose more generally that a vector $\mu \in \mathbb{R}^n$ and a covariance matrix $V \in \mathbb{R}^{n \times n}$ is given. We look for n processes $X^i (i = 0, \ldots, n)$ defined by:

$$X_t^i = X_0^i + b_i t + \sum_{j=1}^n c_{ij} W_t^j, \tag{77}$$

where $W_t^j (j = 1, \ldots, n)$ are n independant Brownian motions, and $b \in \mathbb{R}^n, C = (c_{ij}) \in \mathbb{R}^{n \times n}$.

Proposition 5.1. *There is at most one $(b, C) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}$ such that:*

- *$c_{ij} = 0$ whenever $i > j$,*
- *$\mathbb{E}[X_1^i] = \mu_i (i = 1, \ldots, n)$,*
- *$Cov(X_1^i, X_1^j) = V_{ij} (i, j = 1, \ldots, n)$.*

Proof. We have $\mathbb{E}[X_1^i] = X_0^i + b_i$, so $b_i := \mu_i - X_0^i$ ensures $\mathbb{E}[X_1^i] = \mu_i$ for all i.

We next determine the matrix C thanks to a recursive algorithm:

Ascending step: Let $i, l \in \{1, \ldots, n\}$, and assume $c_{ik}, k > l$ and $c_{lk}, k \geq l$ are determined. Then we can determine c_{il}.

Indeed, if $i > l$, we set $c_{il} = 0$. If $i < l$, we have:

$$V_{il} = Cov(X_1^i, X_1^l) \tag{78}$$

$$= \sum_{j=1}^n c_{ij} c_{lj} \tag{79}$$

$$= \sum_{j=l}^n c_{ij} c_{lj} \tag{80}$$

$$= c_{il} c_{ll} + \sum_{j>l} c_{ij} c_{lj}. \tag{81}$$

Thus we set:

$$c_{il} = \frac{1}{c_{ll}} \left(V_{il} - \sum_{j>l} c_{ij} c_{lj} \right). \tag{82}$$

Back step: Let $l \in \{1, \ldots, n\}$ and assume c_{lj} is determined, for $k > l$. Then we can determine c_{ll}.

Indeed:

$$V_{ll} = \mathbb{V}(X_1^l) = \sum_{j=1}^n c_{lj}^2 = \sum_{j=l}^n c_{lj}^2 = c_{ll}^2 + \sum_{j>l} c_{lj}^2. \tag{83}$$

Thus we set:

$$c_{ll} = \sqrt{V_{ll} - \sum_{j>l} c_{lj}^2}. \tag{84}$$

$\square$

5.3. **Proofs of Lemma 3.1 and Proposition 3.2**

We prove here the Lemma 3.1 and the Proposition 3.2, which give a recursive procedure to simulate exactly under $\mathbb{Q}$.

Proof of Lemma 3.1. Let $(t, \Theta) \in [0, T] \times \mathbb{R}^3$ and consider the process $r^{t,\Theta} = (r_s^{t,\Theta})_{s \in [t,T]}$ defined by (5)-(6).
Let $s \in [t, T]$. An application of It's formula gives:

$$e^{as} r_s^{t,\Theta} = e^{at} r_t^{t,\Theta} + a \int_t^s e^{au} \mu_u^{t,\Theta} \mathrm{d}u + b \int_t^s e^{au} \mathrm{d}B_u, \tag{85}$$

and an easy computation using equality (6) shows that:

$$a \int_t^s e^{-a(s-u)} \mu_u^{t,\Theta} \mathrm{d}u = \alpha_s^{t,\Theta} - \alpha_t^{t,\Theta} e^{-a(s-t)}, \tag{86}$$

where $\alpha^{t,\Theta}$ is the defined by (32).
In addition, if ξ^t is defined by (33), applying It's formula again gives:

$$\xi_s^t = b \int_t^s e^{-a(s-u)} \mathrm{d}B_u. \tag{87}$$

Thus:

$$r_s^{t,\Theta} = e^{-a(s-t)} r_t^{t,\Theta} + \alpha_s^{t,\Theta} - e^{-a(s-t)} \alpha_t^{t,\Theta} + \xi_s^t, \tag{88}$$

which ends the proof as $r_t^{t,\Theta} = \alpha_t^{t,\Theta}$, by (75). $\square$

We now turn to the proof of Proposition 3.2.

Proof of Proposition 3.2. Let $t \leq s \leq u \leq T$. It's formula implies that the triplet $(\xi_r^t, A_r^{t,s}, X_r^{t,x,\Theta})_{r \in [s,u]}$ is the solution of the following linear stochastic differential equation:

$$d \begin{pmatrix} \xi_r \\ A_r \\ X_r \end{pmatrix} = \left[\begin{pmatrix} -a & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \xi_r \\ A_r \\ X_r \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \alpha_r^{t,\Theta} - \frac{\sigma^2}{2} \end{pmatrix} \right] \mathrm{d}t + \begin{pmatrix} b & 0 \\ 0 & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} \mathrm{d}B_r \\ \mathrm{d}W_r \end{pmatrix}, r \in [s, u], \tag{89}$$

with the initial conditions $\xi_s = \xi_s^t, A_s = 0, X_s = X_s^{t,x,\Theta}$. This linear equation has a closed form solution, and we find:

$$\xi_u^t = e^{-a(u-s)} \xi_s^t + b \int_s^u e^{-a(u-r)} \mathrm{d}B_r, \tag{90}$$

$$A_u^{t,s} = \frac{\xi_s^t}{a}(1 - e^{-a(u-s)}) + \frac{b}{a} \int_s^u (1 - e^{-a(u-r)}) \mathrm{d}B_r, \tag{91}$$

$$X_u^{t,x,\Theta} = X_s^{t,x,\Theta} + \int_s^u \alpha_r^{t,\Theta} \mathrm{d}r - \frac{\sigma^2}{2}(u-s) + \frac{\xi_s^t}{a}(1 - e^{-a(u-s)}) + \int_s^u \sqrt{1-\rho^2} \sigma \mathrm{d}W_r \tag{92}$$

$$+ \frac{b}{a} \int_s^u (1 - e^{-a(u-r)}) \mathrm{d}B_r + \int_s^u \rho\sigma \mathrm{d}B_r. \tag{93}$$

Conditionaly upon $\mathcal{F}_s$, the vector $(\xi_u^t, A_u^{t,s}, X_u^{t,x,\Theta})$ is Gaussian and the expectations and covariations given in the Proposition are easily computed thanks to the above formulae. $\square$

5.4. **Comparison with Automatic Differentiation.**

We use the stan math C++ library [3] which allows to easily implement a (Reverse Mode) Automatic Differentiation procedure in order to deduce the derivatives directly from the Monte-Carlo computation of the function L. We compare the results obtained with the weights method developed here with the results obtained by Automatic Differentiation. We also provide a comparison about the computational times.

Precisely, we computed the derivatives of ℓ with respect to the 4 variables $(x, \theta_1, \theta_2, \theta_3)$ at 256 points $(x^i, \theta_1^i, \theta_2^i, \theta_3^i)_{i=1,\ldots,256}$. In the case of the automatic differentiation, we only take 1000 risk-neutral simulations to compute ℓ, while for the approach involving the computation of weights, we took 10000 simulations to compute ℓ and its four derivatives.

Table 5 sums up the time taken for the computations. Clearly, the gain in time resulting by using the weights algorithm is really significant. Additionally, Figure 5 shows the accuracy in the computation using the weights derivatives in comparison with the Automatic Differentiation.

Algorithm - Option	Put Lookback
Automatic Differentiation	179 sec
Weights	97 sec

TABLE 5. Computational times

FIGURE 5. Results from the grid method versus results from Automatic Differentiation

Acknowledgements

References

[1] Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Springer Science & Business Media, 2007.

[2] Hans-Joachim Bungartz and Michael Griebel. Sparse grids. *Acta Numerica*, 13(May):147, 2004.

[3] Bob Carpenter, Matthew D Hoffman, Marcus Brubaker, Daniel Lee, Peter Li, and Michael Betancourt. The stan math library: Reverse-mode automatic differentiation in c++. *arXiv preprint arXiv:1509.07164*, 2015.

[4] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738, 2015.

[5] Hugo Gevret, Nicolas Langrené, Jerome Lelong, Xavier Warin, and Aditya Maheshwari. STochastic OPTimization library in C++. Research report, EDF Lab, May 2018.

[6] Michael B Gordy and Sandeep Juneja. Nested simulation in portfolio risk measurement. *Management Science*, 56(10):1833–1848, 2010.

[7] Alois Pichler. Evaluations of risk measures for different probability measures. *SIAM Journal on Optimization*, 23(1):530–551, 2013.

[8] Yu V Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory of Probability & Its Applications*, 1(2):157–214, 1956.