# An application to credit risk of a hybrid Monte Carlo-Optimal quantization method

GIORGIA CALLEGARO [*]

*Scuola Normale Superiore di Pisa,*
*Université d'Évry Val d'Essonne &*
*ENSIIE*

ABASS SAGNA [†]

*Université Pierre et Marie Curie, Paris VI*

June 20, 2011

---

[*]Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, I-56100 Pise, Italy. Université d'Évry Val d'Essonne, Département de Mathématiques, Bd. F. Mitterrand, F-91025 Évry Cedex. Ecole Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise, 1 square de la Résistance, 91025 Evry Cedex. Phone +33 (0)1 69 47 02 05, e-mail: `g.callegaro@sns.it`. The author's research is supported by an AMaMeF exchange grant and the "Chaire Risque de Crédit" of the French Banking Federation.

[†]Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, Université Pierre et Marie Curie (Paris VI), 175 Rue du Chevaleret, 75013 Paris. Phone +33 (0)1 44 27 72 22, e-mail: `abass.sagna@gmail.com`.

**Abstract**

In this paper we use a hybrid Monte Carlo-Optimal quantization method to approximate the conditional survival probabilities of a firm, given a structural model for its credit default, under partial information.

We consider the case when the firm's value is a non-observable stochastic process $(V_t)_{t \geq 0}$ and investors in the market have access to a process $(S_t)_{t \geq 0}$, whose value at each time $t$ is related to $(V_s, 0 \leq s \leq t)$. We are interested in the computation of the conditional survival probabilities of the firm given the "investor's information".

As an application, we analyze the shape of the credit spread curve for zero coupon bonds in two examples. Calibration to available market data is analyzed.

# Introduction

In this paper we compute the conditional survival probabilities of a firm, in a market that is not transparent to bond investors, by using both Monte Carlo and optimal quantization methods. This allows us to analyze the credit spread curve under partial information in some examples, in order to investigate the degree of transparency and riskiness of a firm, as viewed by bond-market participants.

To introduce the problem, recall that most of the bonds traded in the market are corporate bonds and treasury bonds, that are consequently subject to many kinds of risks, such as market risk (due for example to changes in the interest rate), counterparty risk and liquidity risk. One of the main challenges in credit risk modeling is, then, to quantify the risk associated to these financial instruments.

The methodology for modeling a credit event can be split into two main approaches: the structural approach, introduced by Merton in 1974 and the reduced form approach (or "intensity based"), originally developed by Jarrow and Turnbull in 1992.

The structural approach consists in modeling the credit event as the first hitting time of a barrier by the firm value process.

In reduced form models the default intensity is directly modeled and it is given by a function of latent state variables or predictors of default.

The first approach, in which we are interested, is intuitive by the economic point of view, but it presents some drawbacks: the firm value process can not be easily observed in practice, since it is not a tradeable security, and a continuous firm's value process implies a predictable credit event, leading to unnatural and undesirable features, such as null spreads for surviving firms for short maturities.

Despite the apparent difference between the two models (see, e.g., Jarrow and Protter, 2004), some recent results, starting from the seminal paper Duffie and Lando (2001), have unified the two approaches by means of information reduction. We can also cite Cetin, Jarrow, Protter and Yildirim (2004), where they consider an alternative method with respect to Duffie and Lando (2001), namely, a reduction of the manager's information set, to pass from structural to reduced form models; Giesecke (2006), where the role of the investor's information in a first passage model is investigated and Giesecke and Goldberg (2004), where a structural model with unobservable barrier is studied.

We consider a structural model under partial information, in which investors can not observe the firm value process, but they have access to another process whose value is related to the firm value process. We show in two examples that yield spreads for surviving firms are strictly positive at zero maturity, since investors are uncertain about the nearness of the current firm value to the trigger level at which the firm would declare default. The shape of the term structure of credit spreads may be useful, then, in practice to estimate the degree of transparency and of riskiness of a firm, from the investors' point of view.

We show that the computation of the conditional survival probabilities under partial information leads to a nonlinear filtering problem involving the conditional survival probabilities under full information. These former quantities are approximated (when no closed formula is available) by a Monte Carlo procedure. As concerns the (non)linear filtering problem, in continuous and discrete time, several

computational techniques are known. An overview of some existing methods can be found in Bain and Crişan (2009). These techniques include e.g. particle filtering, the extended Kalman filter, etc. Optimal quantization is an alternative method in discrete time. One of the advantages of this method with respect to the others existing is that, once an optimal quantization of the signal process has been obtained, it can be kept off-line and used instantaneously to estimate the filter. This is the main reason why we use optimal quantization to estimate the discrete time filter distribution. For a comparison between particle filtering and optimal quantization see e.g. Sellami (2008).

The paper is organized as follows. In the first section, we present the market model and we decompose our problem into two problems (**P1**) and (**P2**), that are, respectively, the computation of conditional survival probability in a full information setting and the approximation of the filter distribution. Section 2 and Section 3 are devoted to the solution of the previous two problems. We provide error estimates in Section 4 and, finally, in Section 5 we present two numerical examples concerning the application to credit risk and we calibrate the given model by minimizing the quadratic error between the predicted credit spreads and market credit spreads.

# 1    Market model and problem definition

Let us consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, representing all the randomness of our economic context. In this paper we concentrate our attention on a single firm model, in which the company is subject to default risk and we use a structural approach to characterize the default time.

The process representing the value of the firm, given for example by its value of financial statement, is denoted by $(V_t)_{t \geq 0}$ and we suppose that it can be modeled as the solution to the following stochastic differential equation

$$\begin{cases} dV_t &= b(t, V_t)dt + \sigma(t, V_t)dW_t, \\ V_0 &= v_0, \end{cases} \tag{1.1}$$

where the functions $b : [0, +\infty) \times \mathbb{R} \to \mathbb{R}$ and $\sigma : [0, +\infty) \times \mathbb{R} \to \mathbb{R}$ are Lipschitz in $x$ uniformly in $t$ and $W$ is a standard one-dimensional Brownian motion. We suppose that $\sigma(t, x) > 0$ for every $(t, x) \in [0, +\infty) \times \mathbb{R}$.

In our setting the process $V$ is non observable (it is also known as *state* or *signal*), but investors have access to the values of another stochastic process $S$, providing noisy information about the value of the firm, that can be thought, for example, as the price of an asset issued by the firm.

This observation process follows a diffusion of the type

$$\begin{cases} dS_t &= S_t \left[ \psi(V_t)dt + \nu(t)dW_t + \delta(t)d\bar{W}_t \right], \\ S_0 &= s_0, \end{cases} \tag{1.2}$$

where $\psi$ is locally bounded and Lipschitz, $\nu$ and $\delta$ are bounded deterministic continuous functions and $\bar{W}$ is a one-dimensional Brownian motion independent of $W$. Note that in this model the return on $S$ is a (nonlinear) function of $V$ affected by a noise. A key observation here concerns the volatility of $S$, that cannot be a function of $V$. Otherwise we would be able, under suitable regularity properties of this

4

function to obtain estimations of the firm value from the market observations of the quadratic variation of $S$.

We will deal with two different filtrations, representing different levels of information available to agents in the market and we suppose that they satisfy the usual hypotheses: a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfies the usual hypotheses if $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets and if the filtration is right-continuous.

The first and basic information set is the "default-free" filtration, the one generated by the observation process $S$, which we will denote, for each $t \geq 0$,

$$\mathcal{F}_t^S := \sigma(S_s, 0 \leq s \leq t)$$

and the second one is the full information filtration $(\mathcal{G}_t)_{t\geq 0}$, i.e., the information available for example to a small number of stock holders of the company, who have access to $S$ and $V$ at each time $t$. In our case, the full information filtration is the one generated by the stochastic pair process $(W, \bar{W})$. In conclusion we have

$$\mathcal{F}_t^S \subsetneq \mathcal{G}_t, \quad \forall\, t \geq 0,$$

and we observe that the following *immersion* property holds (see Coculescu, Geman and Jeanblanc, 2008, Proposition 3.1, for an analogous analysis):

**Lemma 1.1.** *Any $(\mathcal{F}_t^S)_t$-local martingale is a $(\mathcal{G}_t)_t$-local martingale. We will say that filtration $(\mathcal{F}_t^S)_t$ is* immersed *in the full filtration $(\mathcal{G}_t)_t$.*

Suppose now that a finite time horizon $T$ is fixed. For a given $s$, $0 \leq s < T$, we observe the process $S$ from 0 to $s$. At time $s$ if the firm has already defaulted we do nothing. Otherwise, we invest in derivatives issued by the firm. Then, in practice, following a structural approach, we define the default time as

$$\tau := \inf \{u \geq s : V_u \leq a\}, \tag{1.3}$$

where as usual $\inf \emptyset = +\infty$ and $a \in \mathbb{R}, 0 < a < v_0$ (notice that in numerical examples we will consider models where $V_t \in (0, +\infty)$, eventually by stopping the process $V$ at the default time $\tau$ by considering the process $(V_{t\wedge\tau})_t$). We are then interested in computing the following quantity, for a given $t$, $s < t < T$,

$$\mathbb{P}\left( \inf_{s\leq u\leq t} V_u > a \middle| \mathcal{F}_s^S \right), \tag{1.4}$$

that is the conditional survival probability of the firm up to time $t$, given the collected information on $S$ up to time $s$. We will see in Section 5 how this quantity plays a fundamental role (if computed under a pricing measure) in the computation of credit spreads for zero coupon bonds.

**Remark 1.2.** *In our setting, introducing the filtration $\mathcal{G}_t^{(s)} := \mathcal{F}_t^S \vee \sigma(t \wedge \tau), t \geq 0$, we have $\mathcal{G}_s^{(s)} = \mathcal{F}_s^S$, so that*

$$\mathbb{P}\left( \inf_{s\leq u\leq t} V_u > a \middle| \mathcal{F}_s^S \right) = \mathbb{P}\left( \inf_{s\leq u\leq t} V_u > a \middle| \mathcal{G}_s^{(s)} \right).$$

## 1.1 Reduction to a nonlinear filtering problem

Using the law of iterated conditional expectations, the Markov property of $V$ and the independence between $W$ and $\bar{W}$, we find, for each $(s,t) \in \mathbb{R}^+ \times \mathbb{R}^+, s < t$,

$$
\begin{aligned}
\mathbb{P}\left(\inf_{s \leq u \leq t} V_u > a \Big| \mathcal{F}_s^S\right) &= \mathbb{E}\left[\mathbb{P}\left(\inf_{s \leq u \leq t} V_u > a \Big| \mathcal{G}_s\right) \Big| \mathcal{F}_s^S\right] \\
&= \mathbb{E}\left[\mathbb{P}\left(\inf_{s \leq u \leq t} V_u > a \Big| V_s\right) \Big| \mathcal{F}_s^S\right] \\
&= \mathbb{E}\left[F(s,t,V_s) | \mathcal{F}_s^S\right], \qquad \mathbb{P}-\text{a.s.} \qquad (1.5)
\end{aligned}
$$

where, for every $x \in \mathbb{R}$,

$$
F(s,t,x) := \mathbb{P}\left(\inf_{s \leq u \leq t} V_u > a \Big| V_s = x\right). \qquad (1.6)
$$

Finally,

(**P1**) *if* we compute $F(s,t,x)$ for every $x \in \mathbb{R}$, which is now a conditional survival probability given the full information filtration, and

(**P2**) *if* we obtain the *filter distribution* at time $s$, i.e., the conditional distribution of $V_s$ given $\mathcal{F}_s^S$, $\Pi_{V_s | \mathcal{F}_s^S}$,

then we are done, since it suffices to compute the integral

$$
\begin{aligned}
\mathbb{E}\left[F(s,t,V_s) | \mathcal{F}_s^S\right] &= \int_{-\infty}^{\infty} F(s,t,x) \Pi_{V_s | \mathcal{F}_s^S}(\mathrm{d}x) \\
&= \int_a^{\infty} F(s,t,x) \Pi_{V_s | \mathcal{F}_s^S}(\mathrm{d}x).
\end{aligned}
$$

It remains to solve the two "intermediate problems" (**P1**) and (**P2**). Let us consider first problem (**P2**).

# 2 Approximation of the filter by optimal quantization

We recall in what follows some facts about optimal vector quantization.

## 2.1 A brief overview on optimal quantization

Consider an $\mathbb{R}^d$-valued random variable $X$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with finite $r$-th moment and probability distribution $\mathbb{P}_X$. Quantizing $X$ on a given grid $\Gamma = \{x^1, \cdots, x^N\}$ consists in projecting $X$ on the grid $\Gamma$ following the closest neighbor rule. The induced mean $L^r$-error $(r > 0)$

$$
\|X - \mathrm{Proj}_\Gamma(X)\|_r = \|\min_{1 \leq i \leq N} |X - x^i|\|_r,
$$

where $\|X\|_r := [\mathbb{E}(|X|^r)]^{1/r}$, is called the $L^r$-*mean quantization error* and the projection of $X$ on $\Gamma$, $\mathrm{Proj}_\Gamma(X)$, is called *the quantization of $X$*. As a function of the grid $\Gamma$ the $L^r$-mean quantization error is continuous and reaches a minimum over all

the grids with size at most $N$. A grid $\Gamma^\star$ minimizing the $L^r$-mean quantization error over all the grids with size at most $N$ is called an $L^r$-*optimal quantizer*.

Moreover, the $L^r$-mean quantization error goes to 0 as the grid size $N \to +\infty$ and the convergence rate is ruled by Zador theorem:

$$\min_{\Gamma, \ |\Gamma|=N} \|X - \mathrm{Proj}_\Gamma(X)\|_r = Q_r(\mathbb{P}_X)N^{-1/d} + o(N^{-1/d})$$

where $Q_r(\mathbb{P}_X)$ is a nonnegative constant. We shall say no more about the basic results on optimal vector quantization. For a complete background on this field we refer to Graf and Luschgy (2000).

The first application of optimal quantization methods to numerical probability appears in Pagès (1997). It consists in estimating $\mathbb{E}f(X)$ (it may also be a conditional expectation) by

$$\mathbb{E}f\big(\mathrm{Proj}_{\Gamma^\star}(X)\big) = \sum_{i=1}^{N} f(x^{\star,i}) \ p_i \tag{2.1}$$

where $\Gamma^\star = \{x^{\star,1}, \cdots, x^{\star,N}\}$ is an $L^r$-optimal grid for $X$ and $p_i = \mathbb{P}\big(\mathrm{Proj}_{\Gamma^\star}(X) = x^{\star,i}\big)$. The induced quantization error estimate depends on the regularity of the function $f$.

• If $f : \mathbb{R}^d \mapsto \mathbb{R}$ is Lipschitz continuous and $r \geq 2$, introducing $[f]_{\mathrm{Lip}} := \sup_{x \neq y} \frac{|f(x)-f(y)|}{|x-y|}$, then

$$\begin{aligned}
\big|\mathbb{E}f(X) - \mathbb{E}f\big(\mathrm{Proj}_{\Gamma^\star}(X)\big)\big| \ &\leq \ \mathbb{E}\big|f(X) - f\big(\mathrm{Proj}_{\Gamma^\star}(X)\big)\big| \\
&\leq \ [f]_{\mathrm{Lip}}\|X - \mathrm{Proj}_{\Gamma^\star}(X)\|_1 \\
&\leq \ [f]_{\mathrm{Lip}}\|X - \mathrm{Proj}_{\Gamma^\star}(X)\|_2.
\end{aligned}$$

• If the derivative $Df$ of $f$ is Lipschitz and $r \geq 2$, then, for any optimal grid $\Gamma^\star$, we have

$$\big|\mathbb{E}f(X) - \mathbb{E}f\big(\mathrm{Proj}_{\Gamma^\star}(X)\big)\big| \leq [Df]_{\mathrm{Lip}}\|X - \mathrm{Proj}_{\Gamma^\star}(X)\|_2^2.$$

How to numerically compute the quadratic optimal quantizers or $L^r$-optimal (or stationary) quantizers in general, the associated weights and $L^r$-mean quantization errors is an important issue from the numerical point of view. Several algorithms are used in practice. In the one dimensional framework, the $L^r$-optimal quantizers are unique up to the grid size as soon as the density of $X$ is strictly log-concave. In this case the Newton algorithm is a commonly used algorithm to carry out the $L^r$-optimal quantizers when closed or semi-closed formulas are available for the gradient (and the hessian matrix).

When the dimension $d$ is greater than 2 the $L^r$-optimal grids are not uniquely determined and all $L^r$-optimal quantizers search algorithms are based on zero search recursive procedures like Lloyd's I algorithms (or generalized Lloyd's I algorithms which are the natural extension of the quadratic case), the Competitive Learning Vector Quantization (CLVQ) algorithm (see Gersho and Gray, 1992), stochastic algorithms (see Pagès, 2008, and Pagès and Printems, 2003), etc. From now on we consider quadratic optimal quantizers.

7

## 2.2 General results on discrete time nonlinear filtering

For an overview on nonlinear filtering problems in interest rate and credit risk models we refer to Frey and Runggaldier (2009) and references therein and, focusing on filtering theory in credit risk, we also have to mention the seminal papers Kusuoka (1999) and Nakagawa (2001).

We consider a general *discrete time* setting, in which we recall the relevant formulas and the desired approximation of the filter (see, e.g., Pagès and Pham, 2005 and Pham, Runggaldier and Sellami, 2005, for a detailed background). We introduce a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ (notice that $\mathbb{P}$ is not the same measure we considered in Section 1, but for simplicity we will use the same notation) and we suppose that:

- *the signal process* $(X_k)_{k \in \mathbb{N}}$ is a finite-state Markov chain with state space $E$, with known probability transition, from time $k-1$ to time $k$, $P_k(x_{k-1}, \mathrm{d}x_k), k \geq 1$, and given initial law $\mu$, and

- *the observation process* is an $\mathbb{R}^q$-valued process $(Y_k)_{k \in \mathbb{N}}$ such that $Y_0 = y_0$ and the pair $(X_k, Y_k)_{k \in \mathbb{N}}$ is a Markov chain.

Furthermore, we suppose that for all $k \geq 1$

(**H**) the law of $Y_k$ conditional on $(X_{k-1}, Y_{k-1}, X_k)$ admits a density

$$y_k \mapsto g_k(X_{k-1}, Y_{k-1}, X_k, y_k),$$

so that the probability transition of the Markov chain $(X_k, Y_k)_{k \in \mathbb{N}}$ is given by $P_k(x_{k-1}, \mathrm{d}x_k)g_k(x_{k-1}, y_{k-1}, x_k, y_k)\mathrm{d}y_k$, with initial law $\mu(\mathrm{d}x_0)\delta_0(\mathrm{d}y_0)$.

In this discrete time setting we are interested in computing conditional expectations of the form

$$\Pi_{Y,n}f := \mathbb{E}\left[f(X_n)|Y_1, \ldots, Y_n\right],$$

for suitable functions $f$ defined on $E$, i.e., we are interested in computing at some time $n$ the law $\Pi_{Y,n}$ of $X_n$ given the past observation $Y = (Y_1, \ldots, Y_n)$. Having fixed the observation $Y = (Y_1, \ldots, Y_n) = (y_1, \ldots, y_n) =: y$ we will write $\Pi_{y,n}$ instead of $\Pi_{Y,n}$.

It is evident that, in the case when the state space of the signal consists of a finite number of points, the filter is characterized by a finite-dimensional vector: if for example each $X_k$ takes values in a set $\{x_k^1, \ldots, x_k^{N_k}\}$ (as in the case where we quantize a process X at discrete times $t_k, k = 0, \cdots, n$ with grids of size $N_k$), then the discrete time filter distribution will be fully determined by the $N_k$-vector with components

$$\Pi_{Y,k}^i = \mathbb{P}\left(X_k = x_k^i | Y_1, \ldots, Y_k\right), \quad i = 1, \ldots, N_k.$$

It is for this reason that, following Pagès and Pham (2005), we apply optimal quantization results in order to obtain a spatial discretization, on a grid $\Gamma_k = \{x_k^1, \ldots, x_k^{N_k}\}$, of the state $X_k, k = 0, \ldots, n$, and we characterize the filter distribution by means of the finite number of points $\{x_0, x_1^1, \ldots, x_1^{N_1}, x_2^1, \ldots, x_2^{N_2}, \ldots, x_n^1, \ldots, x_n^{N_n}\}$ making up the grids $(\Gamma_k)_k$.

In what follows we recall the basic recursive filtering equation, that we will use in our numerics to approximate the filter. By applying the Markov property of $X$ and $(X, Y)$ and Bayes' formula, we find:

$$\Pi_{y,n} f = \frac{\pi_{y,n} f}{\pi_{y,n} \mathbb{1}}, \tag{2.2}$$

where $\pi_{y,n}$ is the *un-normalized filter*, defined by

$$\pi_{y,n} f = \int \cdots \int f(x_n) \mu(\mathrm{d}x_0) \prod_{k=1}^{n} g_k(x_{k-1}, y_{k-1}, x_k, y_k) P_k(x_{k-1}, \mathrm{d}x_k). \tag{2.3}$$

Equivalently, we recall the following recursive formula, that can be directly obtained as well by applying Bayes' formula and the Markov property:

$$\Pi_{y,k}(\mathrm{d}x_k) \propto \int g_k(x_{k-1}, y_{k-1}, x_k, y_k) P_k(x_{k-1}, \mathrm{d}x_k) \Pi_{y,k-1}(\mathrm{d}x_{k-1}),$$

where now $y$ in $\Pi_{y,k-1}$ represents the realization of the vector $(Y_1, \ldots, Y_{k-1})$ and we do not have equality because we need to re-normalize.

Now for any $k \in \{1, \cdots, n\}$ note that

$$\pi_{y,k} f = \mathbb{E}\Big( f(X_k) \prod_{i=1}^{k} g_i(X_{i-1}, y_{i-1}, X_i, y_i) \Big).$$

Therefore, introducing the natural filtration of $X$, $(\mathcal{F}_k^X)_{k \in \mathbb{N}}$, we have

$$
\begin{aligned}
\pi_{y,k} f &= \mathbb{E}\left( \mathbb{E}\Big( f(X_k) \prod_{i=1}^{k} g_i(X_{i-1}, y_{i-1}, X_i, y_i) | \mathcal{F}_{k-1}^X \Big) \right) \\
&= \mathbb{E}\left( \mathbb{E}\big( f(X_k) g_k(X_{k-1}, y_{k-1}, X_k, y_k) | \mathcal{F}_{k-1}^X \big) \prod_{i=1}^{k-1} g_i(X_{i-1}, y_{i-1}, X_i, y_i) \right) \\
&= \mathbb{E}\left( H_{y,k}(f(X_{k-1})) \prod_{i=1}^{k-1} g_i(X_{i-1}, y_{i-1}, X_i, y_i) \right), \tag{2.4}
\end{aligned}
$$

where $H_{y,k}, k = 1, \ldots, n$, is a family of bounded transition kernels defined on bounded measurable functions $f : E \to \mathbb{R}$ by:

$$
\begin{aligned}
H_{y,k} f(x_{k-1}) &:= \mathbb{E}\left[ f(X_k) g_k(x_{k-1}, y_{k-1}, X_k, y_k) | X_{k-1} = x_{k-1} \right] \\
&= \int f(x_k) g_k(x_{k-1}, y_{k-1}, x_k, y_k) P_k(x_{k-1}, \mathrm{d}x_k), \tag{2.5}
\end{aligned}
$$

with $x_{k-1} \in E$. Furthermore, for every $x \in E$, we have

$$H_{y,0} f(x) := \pi_{y,0} f = \mathbb{E}\left[ f(X_0) \right] = \int f(x_0) \mu(\mathrm{d}x_0).$$

It follows, then, from (2.4) that

$$\pi_{y,k} f = \pi_{y,k-1} H_{y,k} f, \quad k = 1, \ldots, n, \tag{2.6}$$

so that we finally obtain the recursive expression

$$\pi_{y,n} = H_{y,0} \circ H_{y,1} \circ \cdots \circ H_{y,n}.$$

9

## 2.3 Estimation of the filter and related error

The estimation of the filter by optimal quantization is already studied in Pagès and Pham (2005) and in Sellami (2005). It consists first in quantizing for every time step $k$ the random variable $X_k$ by considering

$$\widehat{X}_k = \mathrm{Proj}_{\Gamma_k}(X_k), \quad k = 0, \cdots, n, \tag{2.7}$$

where $\Gamma_k$ is a grid of $N_k$ points $x_k^i$, $i = 1, \cdots, N_k$ to be optimally chosen and where $\mathrm{Proj}_{\Gamma_k}$ denotes the closest neighbor projection on the grid $\Gamma_k$.

Owing to Equation (2.6) our aim is to estimate the filter using an approximation of the probability transition $P_k(x_{k-1}, dx_k)$ of $X_k$ given $X_{k-1}$. These transition probabilities are approximated by the probability transition matrix $\hat{p}_k := (\hat{p}_k^{ij})$ of $\widehat{X}_k$ given $\widehat{X}_{k-1}$:

$$\hat{p}_k^{ij} = \mathbb{P}(\widehat{X}_k = x_k^j | \widehat{X}_{k-1} = x_{k-1}^i), \ i = 1, \cdots, N_{k-1}, \ j = 1, \cdots, N_k. \tag{2.8}$$

Then, following Equation (2.5), being the observation $y := (y_0, \cdots, y_k)$ fixed, the transition kernel matrix $H_{y,k}$ is estimated by the quantized transition kernel $\widehat{H}_{y,k}$

$$\widehat{H}_{y,k} = \sum_{j=1}^{N_k} \widehat{H}_{y,k}^{ij} \delta_{x_{k-1}^i}, \quad i = 1, \cdots, N_{k-1}, \quad k = 1, \cdots, n,$$

where

$$\widehat{H}_{y,k}^{ij} = g_k(x_{k-1}^i, y_{k-1}, x_k^j, y_k) \hat{p}_k^{ij}, \quad i = 1, \cdots, N_{k-1}, \ j = 1, \cdots, N_k$$

and where the $x_k^j$'s, $j = 1, \cdots, N_k$ are the (quadratic) optimal quantizers of $X_k$. The initial kernel matrix $H_{y,0}$ is estimated by

$$\widehat{H}_{y,0} = \sum_{i=1}^{N_0} \mathbb{P}(\widehat{X}_0 = x_0^i) \, \delta_{x_0^i}.$$

This leads to the following forward induction to approximate $\pi_{y,n}$:

$$\widehat{\pi}_{y,0} = \widehat{H}_{y,0}, \qquad \widehat{\pi}_{y,k} = \widehat{\pi}_{y,k-1} \widehat{H}_{y,k}, \qquad k = 1, \cdots, n, \tag{2.9}$$

or, equivalently,

$$\begin{cases} \widehat{\pi}_{y,0} = \widehat{H}_{y,0} \\ \widehat{\pi}_{y,k} = \left( \sum_{i=1}^{N_{k-1}} \widehat{H}_{y,k}^{ij} \widehat{\pi}_{y,k-1}^i \right)_{j=1,\cdots,N_k}, \quad k = 1, \cdots, n. \end{cases}$$

Finally, the filter approximation at time $t_n$ is

$$\widehat{\Pi}_{y,n} f = \frac{\widehat{\pi}_{y,n} f}{\widehat{\pi}_{y,n} \mathbb{1}}. \tag{2.10}$$

In order to have some upper bound of the quantization error estimate of $\Pi_{y,n} f$ by $\widehat{\Pi}_{y,n} f$ let us make the following assumptions.

**(A1)** The transition operators $P_k(x, dy)$ of $X_k$ given $X_{k-1}$, $k = 1, \cdots, n$ are Lipschitz.

Recall that a probability transition $P$ on $E$ is C-Lipschitz (with C > 0) if for any Lipschitz function $f$ on $E$ with ratio $[f]_{Lip}$, $Pf$ is Lipschitz with ratio $[Pf]_{Lip} \leq C[f]_{Lip}$. Then, one may define the Lipschitz ratio $[P]_{Lip}$ by

$$[P]_{Lip} = \sup \left\{ \frac{[Pf]_{Lip}}{[f]_{Lip}}, f \text{ a nonzero Lipschitz function} \right\} < +\infty.$$

If the transition operators $P_k(x, dy)$, $k = 1, \cdots, n$ are Lipschitz, it follows that

$$[P]_{Lip} := \max_{k=1,\cdots,n} [P_k]_{Lip} < +\infty.$$

**(A2)** (i) For every $k = 1, \cdots, n$, the functions $g_k$ (recall hypothesis **(H)**) are bounded on $E \times \mathbb{R}^q \times E \times \mathbb{R}^q$ and we set

$$\mathrm{K}_g^n := \max_{k=1,\cdots,n} \|g_k\|_\infty.$$

(ii) For every $k = 1, \cdots, n$, there exist two positive functions $[g_k^1]_{Lip}$ and $[g_k^2]_{Lip}$ defined on $\mathbb{R}^q \times \mathbb{R}^q$ so that for every $x, x', \widehat{x}, \widehat{x}' \in E$ and $y, y' \in \mathbb{R}^q$,

$$|g_k(x, y, x', y') - g_k(\widehat{x}, y, \widehat{x}', y')| \leq [g_k^1]_{Lip}(y, y') \, |x - \widehat{x}| + [g_k^2]_{Lip}(y, y') \, |x' - \widehat{x}'|.$$

The following result gives the error bound of the estimation of the filter (see Pagès and Pham, 2005, Theorem 3.1, for details of the proof).

**Theorem 2.1.** *Suppose that Assumptions **(A1)** and **(A2)** hold true. For every bounded Lipschitz function $f$ on $E$ and for every n-tuple of observations $y = (y_1, \cdots, y_n)$, we have for every $p \geq 1$,*

$$|\Pi_{y,n} f - \widehat{\Pi}_{y,n} f| \leq \frac{\mathrm{K}_g^n}{\phi_n(y) \vee \widehat{\phi}_n(y)} \sum_{k=0}^n \mathrm{B}_k^n(f, y, p) \, \|X_k - \widehat{X}_k\|_p \tag{2.11}$$

*with*

$$\phi_n(y) := \pi_{y,n} \mathbf{1}, \quad \widehat{\phi}_n(y) := \widehat{\pi}_{y,n} \mathbf{1}$$

*and*

$$\begin{aligned} \mathrm{B}_k^n(f, y, p) \;:=\; & (2 - \delta_{2,p})[P]_{Lip}^{n-k}[f]_{Lip} + 2\left( \frac{\|f\|_\infty}{K_g} \left([g_{k+1}^1]_{Lip}(y_k, y_{k+1}) + [g_k^2]_{Lip}(y_{k-1}, y_k)\right) \right.\\ + \; & (2 - \delta_{2,p})\frac{\|f\|_\infty}{K_g} \sum_{j=k+1}^n [P]_{Lip}^{j-k-1} \left.\left([g_j^1]_{Lip}(y_{j-1}, y_j) + [P]_{Lip}[g_j^2]_{Lip}(y_{j-1}, y_j)\right) \right). \end{aligned}$$

*(Convention: $g_0 = g_{n+1} \equiv 0$ and $\delta_{n,p}$ is the usual Kronecker symbol).*

**Remark 2.2.** Concerning the above $L^p$-error bounds, remark that in the quadratic case ($p = 2$) the coefficients $B_k^n$ are smaller than in the $L^1$ case, even if the $L^1$ quantization error is smaller than the quadratic quantization error.

## 2.4 Application to the estimation of $\Pi_{V_s | \mathcal{F}_s^S}$

We focus now on solving problem (**P2**) and, in order to obtain the discrete time approximation of the desired filter $\Pi_{V_s | \mathcal{F}_s^S}$, we fix a time discretization grid $t_0 = 0 < \cdots < t_n = s$ in the interval $[0, s]$ and we apply the results in the previous subsections by working with the corresponding quantized process $\widehat{V}$ (we identify $X$ with $V$ and $Y$ with $S$). From now on $(V_k)_{k=0,\cdots,n}$, will denote either the continuous time process $V$ taken at discrete times $t_k, k = 0, \cdots, n$, or the discrete time Euler scheme relative to $V$.

First of all, let us make the following remark concerning the conditional law of $S_t$ given $((V_u)_{u \in [s,t]}, S_s)$. This will ensure that in our case Hypothesis (**H**) is verified.

**Remark 2.3.** Let $s \leq t$. Using the form of the solution of the SDE (1.2)

$$S_t = S_s \exp \left( \int_s^t \left( \psi(V_u) - \frac{1}{2}(\nu^2(u) + \delta^2(u)) \right) du + \int_s^t \nu(u) dW_u + \int_s^t \delta(u) d\bar{W}_u \right),$$

we notice that

$$\mathcal{L}\big(S_t | (V_u)_{s \leq u \leq t}, S_s\big) = \mathrm{LN}(m_{s,t}; \sigma_{s,t}^2), \tag{2.12}$$

where

$$m_{s,t} = \log(S_s) + \int_s^t \left( \psi(V_u) - \frac{1}{2}(\nu^2(u) + \delta^2(u)) - \nu(u) \frac{b(u, V_u)}{\sigma(u, V_u)} \right) du + \int_s^t \frac{\nu(u)}{\sigma(u, V_u)} dV_u$$

and

$$\sigma_{s,t}^2 = \int_s^t \delta^2(u) du.$$

$\mathrm{LN}(m; \sigma^2)$ stands for the lognormal distribution with mean $m$ and variance $\sigma^2$.

Now, suppose that we temporarily have a time discretization grid from 0 to $t$: $u_0 = 0 < u_1 < \cdots < u_m = t$. For $m$ large enough we can estimate the mean and the variance appearing in Equation (2.12) by using an Euler Scheme. When the estimations of the mean $m_{s,t}$ and variance $\sigma_{s,t}^2$ between two discretization steps are respectively denoted by $m_k$ and $\sigma_k^2$ and we have:

$$\mathcal{L}\big(S_k | V_{k-1}, S_{k-1}, V_k\big) = \mathrm{LN}(m_k; \sigma_k^2) \tag{2.13}$$

with

$$\begin{aligned} m_k &= \log S_{k-1} + \left( \psi(V_{k-1}) - \frac{1}{2}\left( \nu^2(u_{k-1}) + \delta^2(u_{k-1}) \right) - \nu(u_{k-1}) \frac{b(u_{k-1}, V_{k-1})}{\sigma(u_{k-1}, V_{k-1})} \right) \Delta_k \\ &\quad + \frac{\nu(u_{k-1})}{\sigma(u_{k-1}, V_{k-1})} \Delta V_k \end{aligned}$$

and

$$\sigma_k^2 = \delta^2(u_{k-1}) \Delta_k,$$

where $S_k := S_{u_k}$; $V_k := V_{u_k}$; $\Delta V_k = V_k - V_{k-1}$; $\Delta_k = u_k - u_{k-1}$. So, the law of $S_k$ conditional on $(V_{k-1}, S_{k-1}, V_k)$ admits the density (i.e., Hypothesis (**H**) is fulfilled)

$$g_k(V_{k-1}, S_{k-1}, V_k, x) = \frac{1}{\sigma_k x \sqrt{2\pi}} \exp \left( -\frac{1}{2\sigma_k^2} (\log x - m_k)^2 \right), \quad x \in (0, +\infty). \tag{2.14}$$

12

**Remark 2.4.** (a) In the case where

$$\begin{cases} dV_t = \mu V_t dt + \sigma V_t dW_t, & V_0 = v_0, \\ dS_t = rS_t dt + \nu S_t dW_t + \delta S_t d\bar{W}_t, & S_0 = s_0, \end{cases}$$

we directly deduce from Remark 2.3 that for every $s \le t$

$$\mathcal{L}\big(S_t | (V_u)_{s \le u \le t}, S_s\big) = \text{LN}\,\big(m_{s,t}; \sigma_{s,t}^2\big).$$

with

$$m_{s,t} = \log S_s + \left[r - \mu\frac{\nu}{\sigma} - \frac{1}{2}(\sigma^2 + \delta^2) + \frac{\nu\sigma}{2}\right](t-s) + \frac{\nu}{\sigma}\log\frac{V_t}{V_s} \quad \text{and} \quad \sigma_{s,t}^2 = \delta^2(t-s).$$

(b) (About the transition probabilities in Equation (2.8)) In a general setting the transition probabilities

$$\hat{p}_k^{ij} = \mathbb{P}(\widehat{V}_k = v_k^j | \widehat{V}_{k-1} = v_{k-1}^i), \ i = 1, \cdots, N_{k-1}, \ j = 1, \cdots, N_k$$

where $\{v_p^q, p = 0, \cdots, n; q = 1, \cdots, N_p\}$ are the quadratic optimal quantizers of the process $V$, can be estimated by Monte Carlo. However, in some specific cases the continuous time transition densities $p(s, t, x, dy) := \mathbb{P}(V_t \in dy | V_s = x), 0 \le s < t$, are explicitly obtained as solutions to the Kolmogorov equations. For example in the case of item (a) of the remark,

$$p(s, t, x, dy) = \frac{1}{\sigma y\sqrt{2\pi(t-s)}}\, e^{-\frac{1}{2\sigma^2(t-s)}\left[\log\left(\frac{y}{x}\right) - \left(\mu - \frac{\sigma^2}{2}\right)(t-s)\right]^2} dy. \tag{2.15}$$

This density can also be derived from the explicit form of $V$. In such situations, the $\hat{p}_k^{ij}$'s are estimated by

$$\hat{p}_k^{ij} \approx C(t_{k-1}, t_k, v_{k-1}^i, v_k^{j+}) - C(t_{k-1}, t_k, v_{k-1}^i, v_k^{j-}),$$

where $C(t_{k-1}, t_k, v_{k-1}^i, \cdot)$ is the cumulative distribution function associated with $p(t_{k-1}, t_k, v_{k-1}^i, dy)$ and where, for every $k = 0, \ldots, n$,

$$v_k^{j+} := \frac{v_k^j + v_k^{j+1}}{2}; \ v_k^{j-} := \frac{v_k^j + v_k^{j-1}}{2}; \ j = 2, \ldots, N_k - 1; \ v_k^{1-} = 0; \ v_k^{N_k+} = +\infty.$$

In both situations, when estimating the $\hat{p}_k^{ij}$'s by Monte Carlo or by optimal quantization, we commit an additional error.

Once problem **(P2)** solved, owing to Equation (1.5) we use optimal quantization to estimate the $\mathbb{P}\big(\inf_{s \le u \le t} V_u > a | \mathcal{F}_s^S\big)$ on the set $\{\tau > s\}$ by

$$\sum_{i=1}^{N_n} F(s, t, v_n^i)\, \widehat{\Pi}_{y,n}^i, \tag{2.16}$$

where $v_n^i, i = 1, \cdots, N_n$ is the quadratic optimal grid of the process $V$ at time $t_n = s$, $\widehat{\Pi}_{y,n}^i$ is the $i$-th coordinate of the optimal filter $\widehat{\Pi}_{y,n}$ given in (2.10) and, for every $i$, $F(s, t, v_n^i)$ is defined as in (1.6). Note that this last function has in general no explicit expression. In such case, we will estimate it by Monte Carlo as specified in the next section.

13

# 3 Approximation by Monte Carlo of survival probabilities under full information

The aim of this section is to solve problem (**P1**), i.e., to compute, for each pair of positive values $(s, t), s \leq t \leq T$,

$$\mathbb{P}\left(\inf_{s \leq u \leq t} V_u > a \Big| V_s\right) = \mathbb{E}\left(\mathbb{1}_{\{\inf_{s \leq u \leq t} V_u > a\}} | V_s\right), \tag{3.1}$$

where in our general setting the firm value $V$ follows *a priori* a diffusion of the type (1.1). Notice that in the specific case where $V$ is a geometric Brownian motion there exists a closed-formula, that we recall below.
If

$$dV_t = \mu V_t dt + \sigma V_t dW_t, \quad V_0 = v_0,$$

then

$$\mathbb{P}\left(\inf_{s \leq u \leq t} V_u > a | V_s\right) = \Phi(h_1(V_s, t - s)) - \left(\frac{a}{V_s}\right)^{\sigma^{-2}(\mu - \sigma^2/2)} \Phi(h_2(V_s, t - s)) \tag{3.2}$$

where

$$h_1(x, u) = \frac{1}{\sigma\sqrt{u}}\left(\log\left(\frac{x}{a}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)u\right),$$

$$h_2(x, u) = \frac{1}{\sigma\sqrt{u}}\left(\log\left(\frac{a}{x}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)u\right)$$

and where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-u^2/2} du$ is the cumulative distribution function of the standard Gaussian law. For an overview on the computation of boundary crossing probabilities see e.g. Chesney, Jeanblanc and Yor, 2009, Borodin and Salminen, 2002 or Revuz and Yor, 1999.

Since in general we cannot use directly the result in Equation (3.2), we have to resort to an approximation method. Several techniques can be used to estimate these probabilities, such as in Kahalé (2007), where the crossing probabilities are calculated via Schwartz distributions in the specific case of drifted Brownian motion and in Linetsky (2004a) and Linetsky (2004b), where the survival probabilities and hitting densities relative to the CIR, the CEV and to the OU diffusions are expressed as infinite series of exponential densities:

$$\mathbb{P}_{v_0}(\tau > t) = \sum_{n=1}^{\infty} c_n e^{-\lambda_n t}, \quad t > 0, \tag{3.3}$$

where $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n \to \infty$ as $n \to \infty$ and $(c_n)_n$ are explicitly given in terms of the solution of the Sturm-Liouville equation and the eigenvalues of the Sturm-Liouville problem. When the basic solutions to the Sturm-Liouville equation are known, this approach provides efficient estimates of the survival probabilities.

In this paper, we will adopt the "regular Brownian bridge method", originally introduced in Baldi (1995). From the numerical viewpoint, if the exact $c_n$ and $\lambda_n$ in Equation (3.3) can be exactly computed, Linetsky's procedure may be more efficient

than the "regular Brownian bridge method" (except in the Black-Scholes setting, see Section 5.1). Nevertheless, it will be more time consuming than the last one since obtaining e.g. the first one hundred exact $c_n$ and $\lambda_n$ takes 'several minutes' (see Linetsky (2004b)). In order to find an approximated solution to problem (**P1**) by means of the regular Brownian bridge method, we consider the interval $[s,t]$ and we discretize it by means of $u_0 = s < u_1 < \cdots < t = u_N$. We denote by $\bar{V}$ the continuous Euler scheme relative to $V$. This process is defined by

$$\bar{V}_u = \bar{V}_{\underline{u}} + b(\underline{u}, \bar{V}_{\underline{u}})(u - \underline{u}) + \sigma(\underline{u}, \bar{V}_{\underline{u}})(W_u - W_{\underline{u}}), \quad \bar{V}_s = v_s,$$

with $\underline{u} = u_k$ if $u \in [u_k, u_{k+1})$, for the given time discretization grid $u_k := s + \frac{k(t-s)}{N}, k = 0, \cdots, N$, on the set $[s,t]$.

The regular Brownian bridge method is connected to the knowledge of the distribution of the minimum (or the maximum) of the continuous Euler scheme $\bar{V}$ relative to the process $V$ over the time interval $[s,t]$, given its values at the discrete time observation points $s = u_0 < u_1 < \cdots < u_N = t$ (see, e.g., Glasserman, 2003).

**Lemma 3.1.**

$$\mathcal{L}\left(\min_{u \in [s,t]} \bar{V}_u \big| \bar{V}_{u_k} = v_k, k = 0, \cdots, N\right) = \mathcal{L}\left(\min_{k=0,\cdots,N-1} F^{-1}_{v_k,v_{k+1}}(U_k)\right) \qquad (3.4)$$

where $(U_k)_{k=0,\cdots,N-1}$ are i.i.d random variables uniformly distributed over the unit interval and $F^{-1}_{v_k,v_{k+1}}$ is the inverse function of the conditional cumulative function $F_{v_k,v_{k+1}}$, defined by

$$F_{v_k,v_{k+1}}(u) := \begin{cases} \exp\left(-\frac{2N}{(t-s)\sigma^2(u_k,v_k)}(u - v_k)(u - v_{k+1})\right) & \text{if } u \le \min(v_k, v_{k+1}) \\ 1 & \text{otherwise.} \end{cases}$$

We deduce from the previous lemma the following result.

**Proposition 3.2.** *We have:*

$$\mathbb{P}\left(\min_{s \le u \le t} \bar{V}_u > a \big| \bar{V}_s\right) = \mathbb{E}\left(\prod_{k=0}^{N-1} G_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a) \big| \bar{V}_s\right),$$

*with*

$$G_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a) = 1 - F_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a).$$

**Proof.** We have (recall that $\bar{V}_s = \bar{V}_{u_0}$)

$$\mathbb{P}\left(\min_{s \le u \le t} \bar{V}_u > a \big| \bar{V}_s\right) = \mathbb{E}\left(\mathbb{P}\left(\min_{s \le u \le t} \bar{V}_u > a \big| \bar{V}_{u_k}, k = 0, \ldots, N\right) \big| \bar{V}_s\right)$$

$$= \mathbb{E}\left(\mathbb{P}\left(\min_{k=0,\cdots,N-1} F^{-1}_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(U_k) > a\right) \big| \bar{V}_s\right).$$

Since the function $F_{x,y}(\cdot)$ is non-decreasing and the $U_k$'s are i.i.d uniformly distributed random variables, we have

$$\mathbb{P}\left(\min_{s \le u \le t} \bar{V}_u > a \big| \bar{V}_s\right) = \mathbb{E}\left(\prod_{k=0}^{N-1} \mathbb{P}\left(U_k \ge F_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a)\right) \big| \bar{V}_s\right)$$

$$= \mathbb{E}\left(\prod_{k=0}^{N-1} \left(1 - F_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a)\right) \big| \bar{V}_s\right),$$

15

which gives the announced result. $\qquad\square$

By using Proposition 3.2, we estimate the survival probability under full information

$$\mathbb{P}\left(\inf_{s\leq u\leq t}\bar{V}_u > a\Big|\bar{V}_s = v\right)$$

by the following Monte Carlo procedure:

- *Time grid specification.* Fix $u_0 = s < u_1 < \cdots < t = u_N$, the set of $N+1$ points for the (discrete time) Euler scheme in the interval $[s,t]$;

- *Trajectories simulation.* Starting from $v$ and having fixed $M$ (number of Monte Carlo simulations), for $j = 1, \ldots, M$, simulate the discrete path $(\bar{V}^j_{u_k})_{k=0,\ldots,N}$;

- *Computation of the survival probability.* For $j = 1, \ldots, M$, compute (recall that, for every $j$, $\bar{V}^j_{u_0} = v$)

$$p^j_{s,t}(v;a) := \prod_{k=0}^{N-1} G_{\bar{V}^j_{u_k}, \bar{V}^j_{u_{k+1}}}(a). \tag{3.5}$$

- *Monte Carlo procedure.* Finally, apply the Monte Carlo paradigm and get the following approximating value

$$\mathbb{P}\left(\inf_{s\leq u\leq t}\bar{V}_u > a\Big|\bar{V}_s = v\right) \approx \frac{\sum_{j=1}^M p^j_{s,t}(v;a)}{M}. \tag{3.6}$$

As a consequence, combining formulas (2.16) and (3.6) leads to the following hybrid Monte Carlo - optimal quantization formula on the set $\{\tau > s\}$

$$\mathbb{P}\left(\inf_{s\leq u\leq t} V_u > a\Big|\mathcal{F}^S_s\right) \approx \frac{1}{M}\sum_{j=1}^M\sum_{i=1}^{N_n} p^j_{s,t}(v^i_n;a)\,\widehat{\Pi}^i_{y,n} \tag{3.7}$$

where $p^j_{s,t}(\cdot\,;a)$ was introduced in (3.5).

# 4    The error analysis

We now focus on the analysis of the error induced by approximating $\mathbb{P}\left(\inf_{s\leq u\leq t} V_u > a\Big|\mathcal{F}^S_s\right)$ by

$$\frac{1}{M}\sum_{j=1}^M\sum_{i=0}^{N_n} p^j_{s,t}(v^i_n;a)\,\widehat{\Pi}^i_{y,n}.$$

We distinguish three types of error. The first error is induced by the approximation of the filter $\Pi_{y,n}$ appearing in Equation (2.2) by $\hat{\Pi}_{y,n}$, defined in (2.10). This error was already discussed in Section 2.3 in a general setting. The second one is the error deriving from the approximation of

$$\mathbb{P}\left(\inf_{s\leq u\leq t} V_u > a\Big|V_s = v\right) \quad \text{by} \quad \mathbb{P}\left(\inf_{s\leq u\leq t} \bar{V}_u > a\Big|\bar{V}_s = v\right),$$

16

where $\bar{V}$ is the (continuous) Euler scheme relative to the process $V$ (in the Black-Scholes model, there is no need to use an Euler scheme since Equation (1.1) admits an explicit solution). The last one is the error arising from the approximation of the survival probability under full information by means of Monte Carlo simulations.

We next discuss the second and third kinds of error.

▷ **Error induced by the Euler scheme**. We here refer to Gobet (1998), in which the author starts by investigating the case of a one-dimensional diffusion and to the successive related article Gobet (2000) for the multidimensional case. In the two papers the considered diffusion has homogeneous coefficients $b$ and $\sigma$. We start by recalling here some important convergence results we find therein, we will then adapt these results to our case.

Suppose that $X$ is a diffusion taking values in $\mathbb{R}$, with $X_0 = x$, and define $\tau'$ as the first exit time from an open set $D \subset \mathbb{R}$:

$$\tau' := \inf \{u \geq 0 : X_u \notin D\}.$$

Let $\tau'_c$ denotes the exit time from the domain $D$ of the continuous Euler process $\bar{X}$. In order to give the error bound in the approximation of $\mathbb{E}_x \left( \mathbb{1}_{\{\tau'>t\}} f(X_t) \right)$ by $\mathbb{E}_x \left( \mathbb{1}_{\{\tau'_c>t\}} f(\bar{X}_t) \right)$ the following hypotheses are needed:

**(H1)** $b$ is a $\mathcal{C}_b^\infty(\mathbb{R}, \mathbb{R})$ function and $\sigma$ is in $\mathcal{C}_b^\infty(\mathbb{R}, \mathbb{R})$,

**(H2)** there exists $\sigma_0 > 0$ such that $\forall x \in \mathbb{R}, \sigma(x)^2 \geq \sigma_0^2$ (*uniform ellipticity*),

**(H3)** $\mathbb{P}_x \left( \inf_{t \in [0,T]} X_t = a \right) = 0$.

The following proposition states that, under Hypothesis **(H3)**, the approximation error goes to zero as the number of time discretization steps goes to infinity.

**Proposition 4.1 (Convergence).** *Suppose that $b$ and $\sigma$ are Lipschitz, $D = (a, +\infty)$ and that (H3) holds. If $f \in \mathcal{C}_b^0(\bar{D}, \mathbb{R})$ then,*

$$\lim_{N \to +\infty} \left| \mathbb{E}_x[\mathbb{1}_{\{\tau'_c>T\}} f(\bar{X}_T)] - \mathbb{E}_x[\mathbb{1}_{\{\tau'>T\}} f(X_T)] \right| = 0.$$

Note that in the homogeneous case, when $D = (a, +\infty)$, a sufficient condition in order for **(H3)** to hold is (see Gobet, 1998, Prop. 2.3.2)

$$\sigma(a) \neq 0. \tag{4.1}$$

On the other hand, the rate of convergence is given by the following

**Proposition 4.2 (Rate of convergence).** *Under Hypotheses (H1) and (H2), if $f \in \mathcal{C}_b^1(\bar{D}, \mathbb{R})$, then there exists an increasing function $K(T)$ such that*

$$\left| \mathbb{E}_x[\mathbb{1}_{\{\tau'_c>T\}} f(\bar{X}_T)] - \mathbb{E}_x[\mathbb{1}_{\{\tau'>T\}} f(X_T)] \right| \leq \frac{1}{\sqrt{N}} K(T) ||f||_D^{(1)},$$

*where $||f||_D^{(1)} = \sum_{j=0}^1 \sup_{x \in D} |f^{(j)}(x)|$.*

**Remark 4.3.** One notes, by generalizing the proof of Propositions 2.3.1, 2.4.3 and 2.3.2 in Gobet (1998), that the two previous propositions and condition (4.1) still hold when the diffusion coefficients are in-homogeneous, as in our setting, by replacing Hypotheses (**H1**) and (**H2**) by (**I**) and (**J**) :

(**I**) $b$ and $\sigma$ are $\mathcal{C}_b^{\infty}$ functions with respect to both arguments $t$ and $v$, with uniformly bounded partial derivatives with respect to $v$,

(**J**) $\sigma$ is uniformly elliptic, i.e., $\exists \alpha > 0$ such that $\sigma^2(t,v) \geq \alpha, \forall (t,v) \in [0,T] \times \mathbb{R}$.

$\triangleright$ **Error induced by Monte Carlo approximation**. This error comes from the estimation of $\mathbb{P}\left(\min_{s \leq u \leq t} \bar{V}_u > a \middle| \bar{V}_s = v_s^i\right) = \mathbb{E}\left(\prod_{k=0}^{N-1} G_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a) \middle| \bar{V}_s = v_s^i\right)$, for every $i = 1, \cdots, N_s$, by

$$\frac{\sum_{j=1}^{M} p_{s,t}^j(v_s^i; a)}{M},$$

where $p_{s,t}^j(\cdot\,; a)$ is defined in (3.5). We have for every $i = 1, \cdots, N_s$,

$$\left\| \mathbb{E}\left( \prod_{k=0}^{N-1} G_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a) \middle| \bar{V}_s = v_s^i \right) - \frac{\sum_{j=1}^{M} p_{s,t}^j(v_s^i; a)}{M} \right\|_2 = \mathcal{O}\left(\frac{1}{\sqrt{M}}\right). \tag{4.2}$$

By adapting the previous results to our case, namely by identifying $V$ with $X$ and $S$ with $Y$, one deduces an error bound for the estimation of $\Pi_{y,n}F(s,t,\cdot)$ by $\widehat{\Pi}_{y,n}F_{MN}(s,t,x)$, where $n$ is the dimension of the observation vector $y$ (or, equivalently, $n+1$ is the number of points in the time discretization grid of the interval $[0,s]$) and where $F_{MN}(s,t,x)$ is a Monte Carlo estimation of $F(s,t,\cdot)$ of size $M$, based on a time discretization grid, between $s$ and $t$, of size $N+1$. We state, then, the main result of this section.

**Theorem 4.4.** *Suppose that the transition operators of $V_k$ given $V_{k-1}$, $k = 1, \ldots, n$, satisfy Assumption (**A1**) and that the conditional law of $S_k$ given $(V_{k-1}, S_{k-1}, V_k)$ admits a density satisfying (**A2**). Suppose, furthermore, that the coefficients $b$ and $\sigma$ of $V$ fulfill Hypotheses (**H1**)-(**H2**). Then, for every $p \geq 1$,*

$$|\Pi_{y,n}F(s,t,\cdot) - \widehat{\Pi}_{y,n}F_{MN}(s,t,\cdot)| \leq \frac{K_g^n}{\phi_n(y) \vee \hat{\phi}_n(y)} \sum_{k=0}^{n} B_k^n(F(s,t,\cdot), y, p) \, \|V_k - \widehat{V}_k\|_p$$

$$+ \; \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right),$$

*where $n$ is the dimension of the observation vector $y$, $N$ stands for the size of the time discretization grid for the Euler scheme from $s$ to $t$ and $M$ is the number of Monte Carlo trials. Furthermore, $K_g^n, \phi_n(y), \hat{\phi}_n(y)$ and $B_k^n, k = 0, \ldots, n$, are introduced in Theorem 2.1.*

**Remark 4.5.** (About the hypotheses of Theorem 4.4) We consider the case when $V$ is a time homogeneous diffusion.

▷ Concerning Assumption **(A2)** (i), the conditional density functions $g_k$ given in Equation (2.14) are bounded on $\mathbb{R} \times (0, +\infty) \times \mathbb{R} \times (\varepsilon, +\infty)$ for every $\varepsilon > 0$. The Lipschitz condition **(A2)** (ii) holds.

▷ If we suppose that the coefficients $b$ and $\sigma$ of the diffusion $V$ are Lipschitz, we show, by using the Euler scheme relative to $V$, that the transition operators $P_k$ defined by $P_k f(x) := \mathbb{E}(f(V_k)|V_{k-1} = x)$, satisfy

$$|P_k f(x) - P_k f(x')| \le C[f]_{\text{Lip}}|x - x'|$$

for every Lipschitz function $f$ with Lipschitz constant $[f]_{\text{Lip}}$. Then Hypothesis **(A1)** holds true.

▷ As concerns the Lipschitz property of the function $F(s, t, \cdot)$, it follows from Proposition 2.2.1 in Gobet (1998), in the case when the coefficients of the diffusion satisfy Hypotheses **(H1)** and **(H2)** and for $t > s$.

***Proof (of Theorem 4.4).*** We have

$$
\begin{aligned}
|\Pi_{y,n} F(s, t, \cdot) - \widehat{\Pi}_{y,n} F_{MN}(s, t, \cdot)| &\le |\Pi_{y,n} F(s, t, \cdot) - \widehat{\Pi}_{y,n} F(s, t, \cdot)| \\
&+ |\widehat{\Pi}_{y,n} F(s, t, \cdot) - \widehat{\Pi}_{y,n} F_{MN}(s, t, \cdot)|.
\end{aligned}
$$

The error bound of the first term on the right-hand side of the above inequality is given by Theorem 2.1. As concerns the second term, we have

$$
\begin{aligned}
|\widehat{\Pi}_{y,n} F(s, t, \cdot) - \widehat{\Pi}_{y,n} F_{MN}(s, t, \cdot)| &= \left| \sum_{i=1}^{N_s} \widehat{\Pi}_{y,n}^i (F(s, t, v_s^i) - F_{MN}(s, t, v_s^i)) \right| \\
&\le \sup_{v \in \mathbb{R}} |F(s, t, v) - F_{MN}(s, t, v)| \sum_{i=1}^{N_s} \widehat{\Pi}_{y,n}^i \\
&= \sup_{v \in \mathbb{R}} |F(s, t, v) - F_{MN}(s, t, v)|.
\end{aligned}
$$

On the other hand, we have for every $v \in \mathbb{R}$

$$
\begin{aligned}
|F(s, t, v) - F_{MN}(s, t, v)| &\le \left| \mathbb{P}_v(\tau > t) - \mathbb{E}_v\left( \prod_{k=0}^{N-1} G_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a) \right) \right| \\
&+ \left\| \mathbb{E}_v\left( \prod_{k=0}^{N-1} G_{\bar{V}_{u_k}, \bar{V}_{u_{k+1}}}(a) \right) - \frac{\sum_{j=1}^{M} p_{s,t}^j(v; a)}{M} \right\|_2.
\end{aligned}
$$

We then deduce from Proposition 4.2 and from Equation (4.2) that

$$|F(s, t, v) - F_{MN}(s, t, v)| \le \mathcal{O}\left( \frac{1}{\sqrt{N}} \right) + \mathcal{O}\left( \frac{1}{\sqrt{M}} \right),$$

which completes the proof since the error bounds do not depend on $v$. $\qquad\square$

# 5  Numerical results

In the numerical experiments we deal with the estimation of the credit spread for zero coupon bonds. For simplicity we suppose that investors are risk neutral so that

here we directly work under a risk neutral probability $\mathbb{Q}$. We also suppose that the market is complete (remark that $V$ is not a traded asset then it will be necessary to complete the market), so that $\mathbb{Q}$ is unique.

In this section $S$ represents the stock price of an asset issued by the firm. We fix $s$ and, given the observations of $S$ from 0 to $s$, we estimate the spread curve for different maturities $t$ ($t > s$). The credit spread is the difference in yield between a corporate bond and a risk-less bond (Treasury bond) with the same characteristics. It can be seen as a measure of the riskiness relative to a corporate bond, with respect to a risk-free bond. If we suppose for simplicity that the face value is equal to 1 and the recovery rate is zero, the credit spread under partial information from time $s$ to maturity $t$, $\mathbb{S}(s, t)$, equals (see e.g. Bielecki and Rutkowski, 2004 and Coculescu, Geman and Jeanblanc, 2008)

$$\mathbb{S}(s, t) = -\frac{\log\left(\mathbb{Q}(\inf_{s < u \leq t} V_u > a | \mathcal{F}_s^S)\right)}{t - s}.$$

This section is divided into two parts. We first focus on simulations, namely, having arbitrarily fixed the model parameters, we simulate different trajectories of $S$ and we compute, in two examples, the credit spreads for zero coupon bonds. The second part is devoted to calibration.

## 5.1 Simulations

We consider two models for the dynamics of the firm value $V$: a Black-Scholes one and a CEV (Constant Elasticity of Variance) model. In both cases we fix $s = 1$ and, given the simulated trajectory of $S$ from 0 to $s$, we estimate the spreads $\mathbb{S}(1, t)$ for different maturities $t$ varying 0.1 by 0.1 from 1.1 to 11 (the time unit is expressed in years).

▷ *The Black-Scholes model.* We consider the following model for the firm value's and the observed process' dynamics:

$$\begin{cases} dV_t = V_t(\mu dt + \sigma dW_t), & V_0 = v_0, \\ dS_t = S_t(r dt + \nu dW_t + \delta d\bar{W}_t), & S_0 = v_0, \end{cases} \tag{5.1}$$

so that

$$\frac{dS_t}{S_t} = \frac{\nu}{\sigma}\frac{dV_t}{V_t} + \left(r - \mu\frac{\nu}{\sigma}\right)dt + \delta d\bar{W}_t. \tag{5.2}$$

For simplicity, we set $r = \mu = 0.03$ and $\sigma = \nu$, meaning that the return on $S$ is the return on $V$ affected by a noise. The other parameters values are $\sigma = 0.05$, $\delta = 0.1$ and $v_0 = 86.3$. The barrier $a$ is fixed to 76. It is important to note that since $V$ is not traded in the market, the return on $V$ is not necessarily equal to the interest rate $r$.

Notice that when $V$ evolves following a Black-Scholes dynamics, the quantization grids of the firm value process can be derived instantaneously from optimal quadratic functional quantization grids of the Brownian motion, which can be downloaded on the website `www.quantize.math-fi.com` (for more information about functional quantization for numerics see e.g. Pagès and Printems (2005)). This drastically cuts down the computational cost and allows working with grids of higher size. Furthermore, the transition probabilities are estimated using Equation (2.15) and

the survival probabilities $F(s, t, v_n^i), i = 1, \cdots, N_n$ (under $\mathbb{Q}$) in Equation (2.16) are computed via Equation (3.2). We then obtain a single spread estimate in one second.

We set the number $n$ of discretization points over $[0, s]$ equal to 50 and for every $k = 1, \cdots, n$, the quantization grid size $N_k$ is set to 966, with $N_0 = 1$.

Numerical results are presented in Figures 2 and 3. Figure 2 is relative to the partial information case, where three simulated trajectories of the observable process $S$ and the corresponding credit spreads are depicted. Figure 3 treats the full information case, where we suppose that we directly observe $V$. Then, in three examples, corresponding to three different trajectories of $V$ (left hand side of Figure 2), we compute the corresponding credit spreads (right hand side of Figure 2).

We deduce from (5.1) (with $\mu = r$) that

$$S_t = V_t e^{-\frac{1}{2}\delta^2 t + \delta \bar{W}_t}.$$

The correlation coefficient is, then, given for every $t$ by

$$\rho(t) := \sqrt{\frac{e^{\sigma^2 t} - 1}{e^{(\sigma^2 + \delta^2)t} - 1}},$$

meaning that the firm value $V$ is positively correlated to the observation process $S$. Observe that when $\sigma < \delta$, $\rho(t)$ is a strictly decreasing function and goes to 0 as $t$ goes to infinity. This tells us that the a posteriori information on $V$ given $S$ decreases as the maturity $t$ increases. This is what we observe in the spreads curves from Figures 2 and 3, since for large maturities the spreads values almost coincide for analogous trajectories (e.g. for trajectories SU and VU).
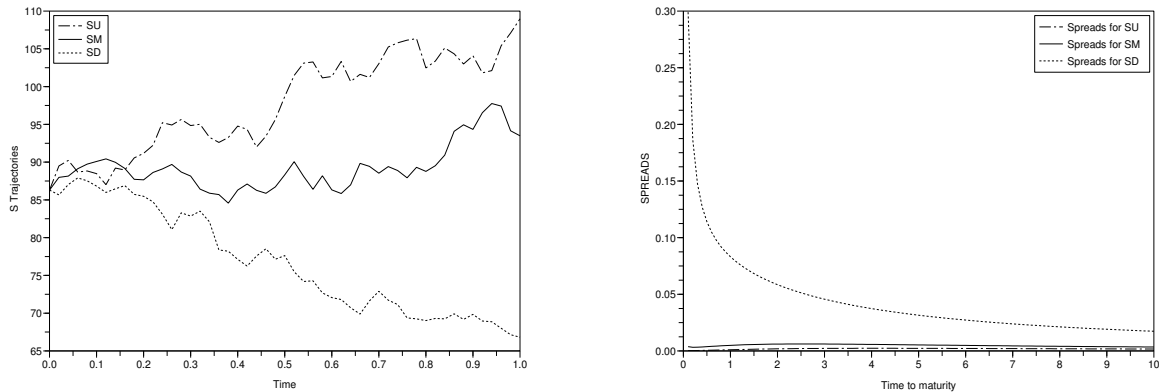


Figure 1: Three trajectories of the observed process $S$ (on the left) and the corresponding spreads (on the right).

First of all, we notice that the short term spreads under partial information, being the default time totally inaccessible, do not vanish, as it is the case in the full information model. Moreover, since $V_t$ and $S_t$ are positively correlated, it is expected that the more the trajectory of $S$ behaves "badly", the higher the short term spreads are, as shown in Figure 2.
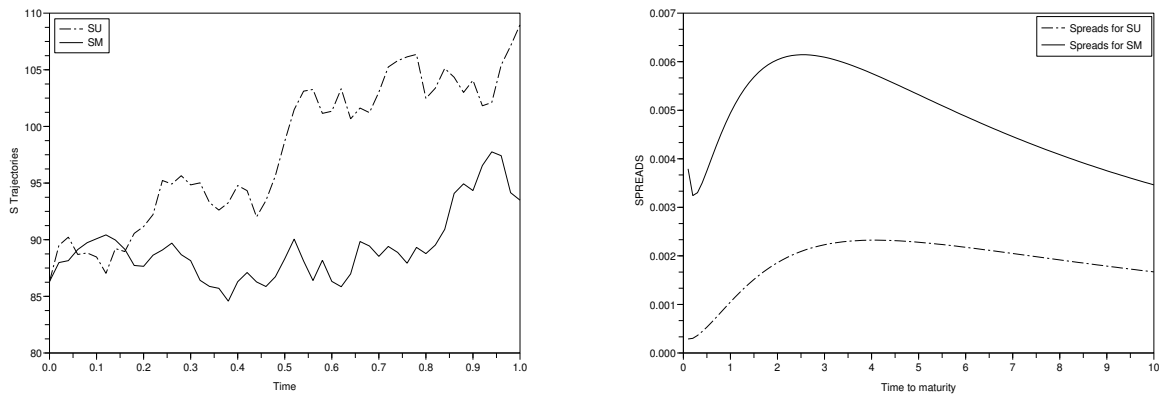
Figure 2: Three trajectories of the observed process $S$ (on the left) and the corresponding spreads (on the right).
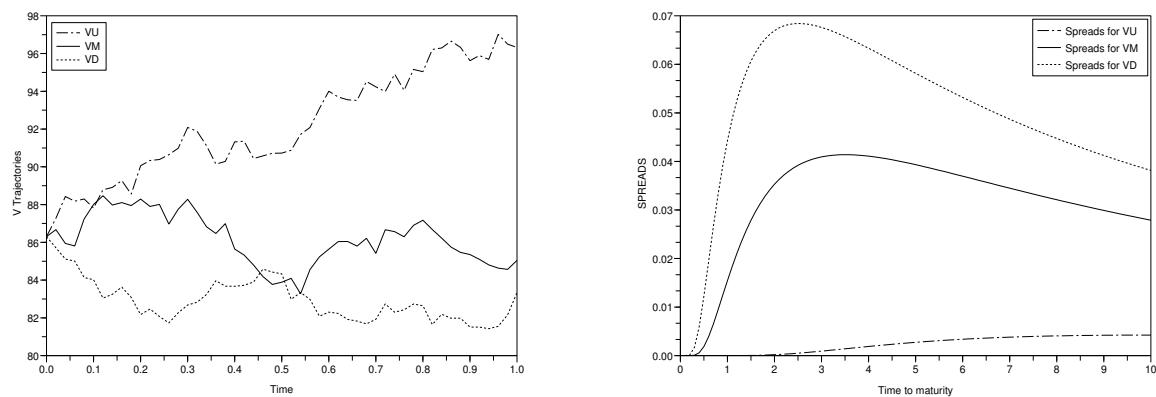


Figure 3: Three trajectories of the value process in the full information case (on the left) and the corresponding spreads (on the right).

22

In the full information setting, on the other hand, the short term spreads are always equal to zero, but in "bad" situations (for example in the case of trajectory VD on the left-hand side of Figure 3) the medium term spreads can be higher than in the partial information model.

▷ *The CEV model.* We suppose now that the firm value's and the observed process' dynamics are given by

$$\begin{cases} dV_t & = & V_t(\mu dt + \gamma V_t^\beta dW_t), & V_0 = v_0, \\ dS_t & = & S_t(rdt + \sigma dW_t + \delta d\bar{W}_t), & S_0 = v_0, \end{cases} \quad (5.3)$$

where $\mu = r = 0.03$, $\gamma = 744.7$ (it is chosen so that the initial volatility equals 0.10), $\beta = -2$ (notice that in this case one of the characteristics of the model is that the *leverage* effect holds: a firm value process increase implies a decrease in the variance of the price process return), $\sigma = 0.05, \delta = 0.1, v_0 = 86.3$. The barrier $a$ here is set to be equal to 79.

For numerics, the number $n$ of discretization points over $[0, s]$ equal to 50 and for every $k = 1, \cdots, n$, the quantization grid size $N_k$ is set to 60, with $N_0 = 1$ (here, since we cannot obtain the quantization grids from the optimal quadratic quantization grids of the Brownian motion, we use Lloyd's algorithm). The number of Euler discretization steps $N$ equals 50 for $t$ varying 0.1 by 0.1 from 1.1 to 3.0 and $N = 100$ for $t$ varying 0.1 by 0.1 from 3.1 to 11.0. The number of Monte Carlo trials $M$ is set to 100000. In the quantization phase we obtain the optimal grid by carrying out 80 Lloyd's I procedures.

Numerical results are presented in Figure 5, where three simulated trajectories of the observable process $S$ and the corresponding spreads are depicted. We first notice that the spreads in this example are higher than the ones in the previous example. This is due to the fact that in this case the observed process $S$ is more volatile, as it can be seen from Figure 5, compared to Figure 2.

Secondly, we remark, as in the previous example, that the more the trajectory of S behaves "badly", the higher the short term spreads are, as shown in Figure 5.

Moreover, notice that the spread curves corresponding to the two worst $S$ trajectories seem to cross, however a zoom in the graph shows that it is not the case and that the spreads curve for SD CEV is always above the one for SM CEV. This can be explained by noticing that the model we use keeps the memory of all the observed path and that the trajectory SD CEV is globally worse than the trajectory SM CEV.

**Remark 5.1.** a) The most important fact from the numerical point of view is that, as soon as the process $V$ is quantized over $[0, s]$, the survival probability $\mathbb{Q}(\inf_{s \leq u \leq t} V_u > a | \mathcal{F}_s^S)$ is estimated for every maturity $t > s$ without modifying the optimal quantization grid of $V$.
b) As expected, in both Black-Scholes and CEV models, numerical tests confirm that the spread increases as the barrier $a(a < v_0)$, tends to $v_0$.

## 5.2 Calibration issues

For calibration to real data, we consider the Black-Scholes model

$$\begin{cases} dV_t & = & V_t(\mu dt + \sigma dW_t), & V_0 = v_0, \\ dS_t & = & S_t(rdt + \sigma dW_t + \delta d\bar{W}_t), & S_0 = s_0, \end{cases} \quad (5.4)$$
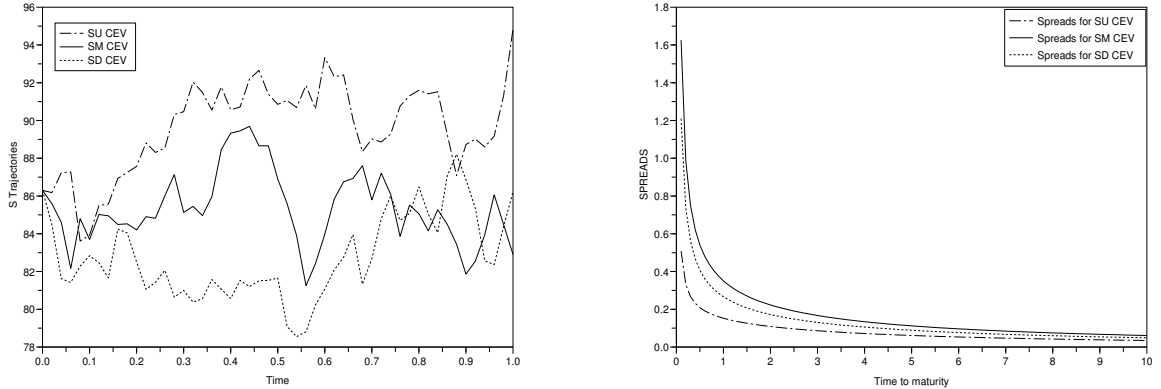
Figure 4: Three trajectories of the observed process $S$ in the CEV model (on the left) and the corresponding spreads (on the right).

even if the methodology presented below may be applied to other models. The calibration has been done in two steps. The first step, related to the learning phase, consists in calibrating the parameters of the stock price $S$ in the observation interval $[0, s]$. The remaining parameters are, then, calibrated from the market data for credit spreads. Recall that the quantization grids of the firm value process can be derived from the optimal quadratic functional quantization grids of the Brownian motion.

$\triangleright$ *Calibration of $S$'s parameters.* We work on JP Morgan weekly stock prices data (available on Yahoo finance website `www.finance.yahoo.com/`) for the period 03/22/2009 - 03/22/2010, corresponding in our setting to the observation time interval $[0, s]$ with $s = 1$. The data set is, then, of size 53 and each considered stock price $S_i, i = 0, \cdots, 52$, is computed as the average between the bid and ask prices (see Figure 6 on the left). The considered interest rate $r = 0.51\%$ is obtained as the average of the three-months U.S. Libor rates in the period March 2009 - March 2010. Given the above model for $S$, one can estimate the parameter $\theta := \sqrt{\sigma^2 + \delta^2}$ using elementary statistical theory. The obtained estimation $\hat{\theta}$ from real data is $\hat{\theta} = 0.2496$.

Before dealing with the second step of the calibration we study the impact of the noise parameter $\delta \in (0, \hat{\theta})$ on the credit spread $\left(\text{once } \delta \text{ is fixed, } \sigma = \sqrt{\hat{\theta}^2 - \delta^2}\right)$. For this purpose, we set $\mu = r$ to have

$$\frac{dS_t}{S_t} = \frac{dV_t}{V_t} + \delta d\bar{W}_t. \tag{5.5}$$

We plot in Figure **??** the term structure of credit spread $\mathbb{S}(1, t)$ for $t$ varying 0.1 by 0.1 from 1.1 to 6 and for $\delta = \{0.05, 0.10, 0.15, 0.20\}$. The considered values for $v_0$ and $a$ are $v_0 = 2,079,188,000\$$ and $a = 1,908,994,000\$$. They represent, respectively, the total assets value and the total liabilities balance sheet value of the firm at the end of March 2009 (both available on Yahoo finance website). In this numerical implementations we have set the number of discretization points over $[0, 1]$ to 53 and

24

the quantization grid size $N_k = 966$, for $k = 1, \cdots, 53$ and $N_0 = 1$. Numerical results show that the spreads increase as the noise parameter $\delta$ increases. This naturally comes from Equation (5.5), since the more $\delta$ is large, the more the information on $S$ is noisy and so the higher is the risk perception of the investor. Moreover, for small values of $\delta$ (as, for example, for $\delta = 0.05$), the term structure of credit spread has a form similar to the one we found in the complete information case (see Figure 3). Then, varying $\delta$ may allow us to obtain a rich set of different forms of the credit spread term structure.
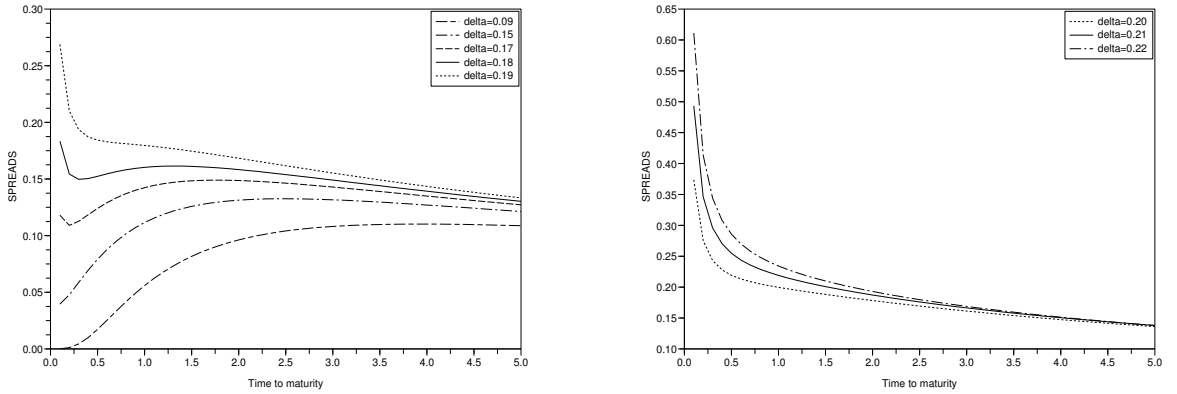
We now focus on the calibration to real data.



Figure 5: Spreads computed with different values of $\delta$.

▷ *Calibration.* As previously remarked, the parameters values $v_0$ and $a$ are known and they correspond to the total assets value and to the total liabilities value of the firm at the end of March 2009, namely $v_0 = 2,079,188,000\$$ and $a = 1,908,994,000\$$. Furthermore, we set the initial stock price value and the interest rate to, respectively, $s_0 = 27,365$ and $r = 0.51\%$.

We calibrate $\mu$ and $\delta$ on the credit spreads (for zero coupon bonds) market data, that is, given a set of credit spreads data $\{s_{t_i}, i = 1, \cdots, 4\}$, at time $s = 1$ and for different maturities $t_1 = 7/12; t_2 = 11/12; t_3 = 1; t_4 = 13/12$, we find $(\mu^\star, \delta^\star)$ that minimize the quadratic error

$$\sum_{i=1}^{4} \left( \mathbb{S}(1, t_i) - s_{t_i} \right)^2.$$

The market data $\{s_{t_i}, i = 1, \cdots, 4\}$ are obtained as the difference between riskless Treasury bond yields and JP Morgan zero coupon bonds (Medium Term Note zero coupon SER E principal protected bond) yields. Since there is a mismatch between the maturities of corporate and Treasury bonds in the sample, we interpolate the riskless yields in order to have a continuum of maturities and we compute the spreads for all the $t_i$'s. For the calibration we restricted our search procedure on the domain $[-0.1, 0.1] \times [0.01, 0.1]$. The optimal values obtained are $(\mu^\star, \delta^\star) = (0.03, 0.075)$ and the corresponding credit spread term structure over three years is depicted in Figure 6, right-hand side. The quadratic error equals $3.5 \times 10^{-3}$. Notice that the most challenging task in the calibration phase is the collection of real data, because zero

25

coupon corporate bond prices at a fixed time $s$, issued by the same firm and with identical features, are only given for a small number of different maturities. This is why the used set of data is of small size.
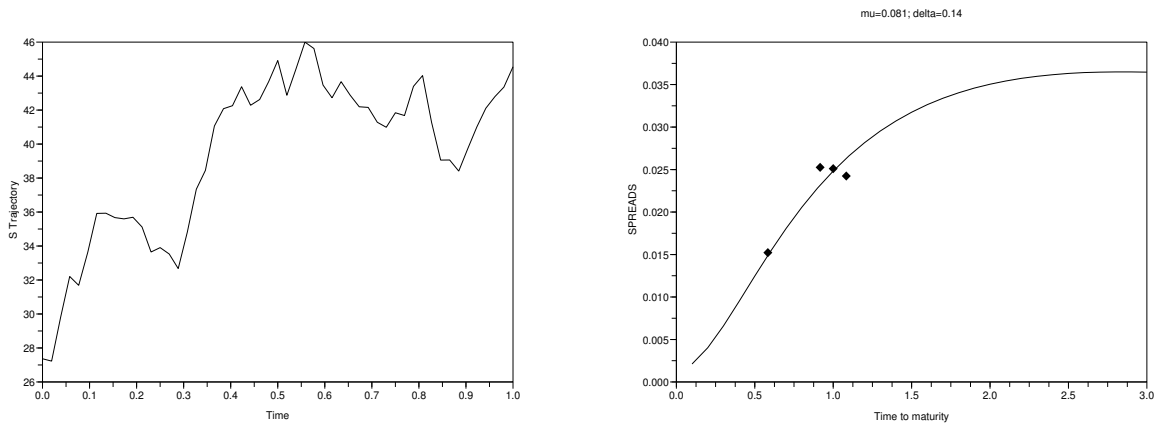


Figure 6: JP Morgan weekly stock prices over the period 03/22/2009 - 03/22/2010 (on the left) and corresponding credit spreads curve over three years obtained for $(\mu^\star, \delta^\star) = (0.03, 0.075)$ calibrated to market data (black square dots).

# References

Bain, A., & Crişan, D. (2009). *Fundamentals of stochastic filtering.* Springer.

Baldi, P. (1995). Exact asymptotics for the probability of exit from a domain and applications to simulations. *The Annals of Applied Probability*, 23(4), pp. 1644-1670.

Bielecki, T.R., & Rutkowski, M. (2004). *Credit Risk: Modeling, Valuation and Hedging.* Berlin: Springer (Springer Finance).

Borodin, A.N., & Salminen, P. (2002). *Handbook of Brownian motion - Facts and Formulae.* 2nd ed. Basel: Birkhäuser.

Cetin, U., Jarrow, R., Protter, P., & Yildirim, Y. (2004). Modeling credit risk with partial information. *The Annals of Applied Probability*, 14(3), pp. 1167-1178.

Chesney, M., Jeanblanc, M. & Yor, M. (2009). *Mathematical Methods for Financial Markets.* Springer Finance.

Coculescu, D., Geman, H., & Jeanblanc, M. (2008). Valuation of Default Sensitive Claims under Imperfect Information. *Finance and Stochastics*, 12(2), pp. 195-218.

Duffie, D., & Lando, D. (2001). Term Structures of Credit Spreads with Incomplete Accounting Information. *Econometrica*, 69(3), pp. 633-664.

Frey, R., & Runggaldier, W.J. (2009). Nonlinear filtering in models for interest rate and credit risk. To appear in *Handbook on nonlinear filtering.* Oxford University Press.

Gersho, A., & Gray, R. (1992). *Vector Quantization and Signal Compression.* Boston: Kluwer Academic Press.

Giesecke, K. (2006). Default and information. *Journal of Economic Dynamics and Control*, 30(11), pp. 2281-2303.

Giesecke, K., & Goldberg, L. (2004). Forecasting default in the face of uncertainty. *Journal of Derivatives*, 12(1), pp. 14-25.

Glasserman P. (2003). *Monte Carlo Methods in Financial Engineering.* Springer-Verlag: New York.

Gobet, E. (1998). *Schémas d'Euler pour diffusion tuée. Application aux options barrière.* PhD thesis. Paris: Université Paris VII.

Gobet, E. (2000). Weak approximation of killed diffusion using Euler schemes. *Stochastic Processes and their Applications*, 87, pp. 167-197.

Graf, S., & Luschgy, H. (2000). *Foundations of Quantization for Probability Distributions.* Volume 1730 of Lecture Notes in Mathematics. Springer - Verlag.

Jarrow, R.A., & Protter, P. (2004). Structural versus reduced form models: a new information based perspective. *Journal of investment management.* 2(2), pp. 1-10.

Kahalé, N. (2007). Analytic crossing probabilities for certain barriers by Brownian motion. Preprint.

Kusuoka, S. (1999). A remark on default risk models. *Advances in Mathematical Economics*, 1, pp. 69-82.

Linestky, V. (2004a). The spectral decomposition of the option value. *International Journal of Theoretical and Applied Finance*, 7(3), pp. 337-384.

Linestky, V. (2004b). Computing hitting time densities for CIR and OU diffusions: applications to mean-reverting models. *Journal of Computational Finance*, 7, pp. 1-22.

Nakagawa, H. (2001). A Filtering Model on Default Risk. *J. Math. Sci. Univ. Tokyo*, 8(1), pp. 107-142.

Pagès, G. (1997). A space quantization method for numerical integration. *Journal of Computational and Applied Mathematics*, 89(1), pp. 1-38.

Pagès, G., & Pham, H. (2005). Optimal quantization methods for nonlinear filtering with discrete time observations. *Bernoulli*, 11(5), pp. 893-932.

Pagès, G., & Printems, J. (2003). Optimal quadratic quantization for numerics: the Gaussian case. *Monte Carlo Methods and Appl.*, 9(2), pp. 135-165.

Pagès, G., & Printems, J. (2005). Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods and Appl.*, 11(4), pp. 407-446.

Pham, H., Runggaldier, W.J., & Sellami, A. (2005). Approximation by quantization of the filter process and applications to optimal stopping problems under partial observation. *Monte Carlo Methods and Applications*, 11(1), pp. 57-81.

Revuz, D., & Yor, M. (1999). *Continuous martingales and Brownian motion.* 3rd ed., volume 293 of Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Berlin: Springer-Verlag.

Sellami, A. (2005). *Méthodes de quantification optimale pour le filtrage et applications à la finance.* PhD thesis. Paris: Université Paris Dauphine.

Sellami, A. (2008). Comparative survey on nonlinear filtering methods: the quantization and the particle filtering approaches. *Journal of Statistical Computation and Simulation*, 78(2), pp. 93-113.