Finance and Stochastics manuscript No.
(will be inserted by the editor)

Robert A. Jarrow · **Philip Protter** · **A. Deniz Sezer**

Information Reduction via Level Crossings in a Credit Risk Model

Received: date / Accepted: date

Abstract This paper provides an alternative credit risk model based on information reduction where the market only observes the firm's asset value when it crosses certain levels, interpreted as changes significant enough for the firm's management to make a public announcement. For a class of diffusion processes we are able to provide explicit expressions for the firm's default intensity process and its zero-coupon bond prices.

Keywords Reduced form models · structural models · credit risk · information reduction · diffusion · level-crossings · Brownian motion with drift

JEL Classification G13 · D82

Mathematics Subject Classification (2000) 60G55 · 60G60

1 Introduction

The credit risk literature studies the pricing and hedging of financial securities that can default and pay off less than promised. For pricing and hedging purposes,

Supported in part by NSF grant DMS-0604020 and NSA grant H98230-06-1-0079

Robert A. Jarrow
Johnson Graduate School of Management, Cornell University, Ithaca, NY, 14853

Philip Protter
School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 14853

Deniz Sezer
Department of Mathematics and Statistics, York University, Toronto, ON, Canada M3J 1P3
Tel.: 1-647-200 95 83
Fax: 1-416-736 21 00 x33956
E-mail: dsezer@mathstat.yorku.ca

characterizing the time to default, a random variable, is essential. Two modeling approaches have arisen in this regard. The first, due to Merton [16], characterizes the default time as the first hitting time of the firm's asset value X_t to a fixed level (the default barrier) x. This approach requires the market (modeler) to have complete and continuous information about the firm's asset value process and default barrier. This is called the *structural* approach to credit risk. The second, due to Jarrow and Turnbull [10],[11], characterizes the default time as the first jump time of a doubly stochastic point process, usually with an intensity process. This approach requires the market to have information regarding the point process, usually observable histories of relevant state variables and the default process itself. This is called the *reduced form* approach to credit risk. Although seemingly distinct, Jarrow and Protter [12] have recently argued (following Duffie and Lando [6]) that these two approaches are related via information reduction. They state that one can view a reduced form model as a structural model where the information set is reduced from complete and continuous observations of the firm's asset value process and default barrier to a coarser information set. This coarser information set can transform the typical default time process of a structural model into that of a reduced form model.

Various examples of this information reduction transformation have already appeared in the literature. Duffie and Lando [6] use delayed information with discrete time steps. The market observes the firm's asset value X_t at discrete time steps $\{t_1,\ldots,t_n,\ldots\}$ for a class of diffusion processes. But, they add Gaussian noise to the observation process, the interpretation being that one only observes the firm's asset value X_t with noise. Dufresne, Goldstein and Helwege [5] use delayed information with discrete time steps, with X_t a geometric BM process, where the time delay is itself a simple binomial random variable. Guo, Jarrow, Zeng [8] generalize these two papers to the market receiving delayed information about X_t that can occur at either discrete or random time steps. These observation times are interpreted as occurring when the firm issues quarterly reports or when the credit rating (health of the firm) changes. GJZ study a diffusion model with jumps in either the drift, volatility, or the asset value. A related example is that of Çetin, Jarrow, Protter, Yildirim [4] whose information corresponds to knowing when X_t crosses a level 0 (i.e. the sign of X_t). The default time is more complicated in CJPY. It corresponds to the duration of an excursion of X_t below zero. Consequently, in CJPY, X_t is not interpreted as the firm's asset value but as the firm's cash flows.[1] Several other papers have extended the ideas of Duffie and Lando recently, in different directions. We mention Nagagawa [17], Jeanblanc and Valchev [13], and the master's thesis of M. Tolotti [18].

Our paper adds another and alternative type of information reduction in a structural credit risk model. We assume that the market only observes the firm's asset value X_t when it crosses various levels $x_1,\ldots,x_N$. The interpretation is that the firm makes public announcements about the firm's asset value only when there are significant changes in its economic standing (when the asset value reaches the crossings). As in the structural approach, default is the first time X_t hits the lowest

[1] There are two related papers. One by Giesecke and Goldberg [7] who have the default barrier being unknown, with X_t observed continuously. The second by Kusuoka [15] who has X_t observed continuously, but with noise. The default time is given exogenously, and not related to the first time X_t hits some level.

one of these levels. This alternative information set requires a different and more complex set of mathematics, recently developed by Sezer [19], to characterize the firm's default process. Using this newly developed mathematics, we study the characterization of the firm's default time and the pricing of a firm's risky debt. Analytic expressions for a firm's conditional default probability and risky zero-coupon bond prices are provided. For special cases of the general model, these expressions are easily computed.

Much of the underlying mathematics this paper is based on, was developed in the thesis (Cornell University, 2005) of A. Deniz Sezer, itself inspired by an old paper of Jacod and Mémin ([9]). The published work will appear in [19]. The main contribution of this paper is its application to Finance, and the main mathematical contribution is the new result, Theorem 4.

An outline for this paper is as follows. Section 2 presents the model structure, section 3 applies this model structure to characterizing a firm's credit risk, section 4 is the analysis of a particular model, section 5 concludes.

2 Model Setup

Let $X = (\Omega, \mathscr{G}, \mathscr{G}_t, X_t, \theta_t, P^x)$ be a non-singular diffusion with values in some interval $I \subset \mathbb{R}$. Here we are using the formulation of a Markov process as given in (for example) [2] or [20]; in particular, θ denotes the shift operator: One has $X_t \circ \theta_s = X_{t+s}$. For simplicity we take $I = \mathbb{R}$. We assume that the infinitesimal generator of X is given by

$$\mathscr{A} := \frac{1}{2}a(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx}$$

where $a(x)$ is strictly positive and continuous, and $b(x)$ is locally integrable on $(-\infty, \infty)$. We refer to $\mathbb{G} \doteq (\mathscr{G}_t)_{t\geq 0}$ as the complete information set.

We consider a finite collection of points $\mathscr{L} = \{x_1, x_2 \ldots, x_N\}$ in $\mathbb{R}$ with $x_1 < x_2 \ldots < x_N$. The region indicator function $R(x)$ associated with $\mathscr{L}$ is defined as

$$R(x) = i \text{ if } x \in R_i$$

where $R_i = (x_i, x_{i+1}], i = 1, \ldots N-1$, $R_0 = (-\infty, x_1]$ and $R_N = (x_N, \infty)$.

We let $\mathbb{F}^0 = (\mathscr{F}_t^0)_{t\geq 0}$ be the filtration generated by $(R(X_t))_{t\geq 0}$. We fix $P = P^{x_{i*}}$ for some $x_{i*} \in \mathscr{L}$ and let $\mathscr{F} = (\vee_{t\geq 0}\mathscr{F}_t^0) \vee \mathscr{N}$ where $\mathscr{N}$ is the set of null sets of P. We denote by $\mathbb{F}$ the filtration $(\mathscr{F}_t)_{t\geq 0}$ obtained by letting $\mathscr{F}_t = \mathscr{F}_t^0 \vee \mathscr{N}$. Then $(\Omega, \mathscr{F}, \mathbb{F}, P)$ is a complete stochastic basis [19].

We refer to $\mathbb{F}$ as the reduced information set. $R = (R(X_t))_{t\geq 0}$ is an optional process of $\mathbb{F}$. Let $g_t = \sup\{s \leq t : X_s \in \mathscr{L}\}$ and $U_t = t - g_t$ respectively the last exit time (before t) and the time since the last exit time from $\mathscr{L}$. Here we take $\sup\{\emptyset\} = 0$. Both g_t and U_t are càdlàg and adapted to $\mathbb{F}$. $\mathbb{F}$ is also generated by the following pair of càdlàg processes, $(X_{g_t}, \text{sign}(X_t - X_{g_t})U_t)_{t\geq 0}$.

A key structure for our analysis is an excursion interval. Since X is continuous, P a.s. for all ω the sets $M_i(\omega) = \{t : X_t(\omega) = x_i\}$ are closed, and so is their union. We can write

$$(\cup_{i=1}^N M_i)^c = \cup_{k=1}^\infty (a_k, b_k)$$

where the (a_k, b_k) are disjoint open intervals. We refer to these intervals as the excursion intervals away from $\mathscr{L}$. During an excursion interval (a,b), X lies entirely in one of the regions R_i with both X_a and $X_b \in \mathscr{L}$. The path $f(s) = X(a+s)$ $0 \le s \le b-a$ is called an excursion. The starting point of an excursion can be any of the points x_i. An excursion starting from x_i can be either in the region R_i (an upward excursion) or R_{i-1} (a downward excursion). If it is in R_i (resp. R_{i-1}) then the ending point can be either x_i or x_{i+1} (resp. x_{i-1}).

Loosely speaking these are the four different types of excursions (with starting point x_i) observable with the information contained in $\mathbb{F}$. For each type of excursion there corresponds a Lévy measure on $(0,\infty]$, which we denote by $F_i^{j\pm}$, where a $(+)$ (resp. $(-)$) is for an upward (resp. downward) excursion, and $j = 0$ (resp. 1) is for an excursion ending at x_i (resp. $x_{i\pm1}$). These measures are constructed using the excursion measure n_i of X at x_i (see [19]). As we do not assume that the reader is familiar with the excursion theory of Markov processes, we are not going to show this construction here. Rather, for the purposes of this paper, it is enough to know that these measures exist, and that they are related to the excursions, precisely, as follows: We define the stopping times

$$T_1^x = \inf\{t : U_t > x\}$$

$$S_1^x = \inf\{t > T_1^x : U_t = 0\},$$

and then define recursively

$$T_n^x = \inf\{t > S_{n-1}^x : U_t > x\}$$

$$S_n^x = \inf\{t > T_n^x : U_t > x\}.$$

S_n^x is the right end point of the nth excursion interval whose length is greater than x. Then $T_n^x - x$ corresponds to the left end point of that excursion interval. Let $A_{i,\pm}^n = \{X_{T_n^x - x} = x_i, \pm X_{T_n^x} > \pm X_{T_n^x - x}\}$, noting that these sets give a partition of $\{T_n^x < \infty\}$. On $A_{i,\pm}^n$, the joint conditional distribution of $S_n^x - T_n^x + x$ (note that this is equal to the length of the excursion) and $X_{S_n^x}$ (which can be either x_i or $x_{i\pm1}$) given $\mathscr{F}_{T_n^x}$ can be given in terms of the measures $F_i^{1\pm}$ and $F_i^{0\pm}$. In particular, on $A_{i,\pm}^n$,

$$P(S_n^x - T_n^x + x \in dy, X_{S_n^x} = x_{i\pm1} | \mathscr{F}_{T_n^x}) = \frac{F_i^{1\pm}(dy)}{F_i^{\pm}(x,\infty]} 1_{\{y>x\}} \tag{1}$$

and

$$P(S_n^x - T_n^x + x \in dy, X_{S_n^x} = x_i | \mathscr{F}_{T_n^x}) = \frac{F_i^{0\pm}(dy)}{F_i^{\pm}(x,\infty]} 1_{\{y>x\}} \tag{2}$$

where $F_i^{\pm} = F_i^{1\pm} + F_i^{0\pm}$.

Note that when we write $F_i^{1\pm}$, we are excluding F_1^{1-} and F_N^{1+}, as these are not defined because there are no levels below x_1 or above x_N. Occasionally, however, these will appear in our notation, in which case they should always be taken as 0. Similarly, we take $x_0 = -\infty$ and $x_{N+1} = \infty$.

Some properties of $F_i^{j\pm}$ that we borrow from excursion theory are as follows. $F_i^{j\pm}$, except F_1^{0-} and F_N^{0+}, do not have mass at ∞. This corresponds to the fact that X can not have an excursion of infinite length in any bounded region. However,

F_1^{0-} or F_N^{1+} can have mass at ∞ depending on whether or not X is transient. $F_i^{0\pm}$ have infinite total mass, this corresponds to the regularity of x_i. On the other hand $F_i^{1\pm}$ have finite total mass, and this corresponds to the fact that X can not have infinitely many excursions from x_i to $x_{i\pm 1}$ in finite time intervals. Lastly, none of the measures $F_i^{j\pm}$ charges singletons $\{x\}$ for $0 < x < \infty$ (see, [19]).

Let us assume that $x_{i*} \neq x_1$ and let τ be the hitting time of X to $\{x_1\}$. τ is a stopping time of $\mathbb{F}$ since it is also the hitting time of $(X_{g_t})_{t\geq 0}$ to $\{x_1\}$. This stopping time will be fundamental to the economic analysis subsequently discussed.

Let A be the cumulative intensity process of τ, i.e. the compensator of the one point process $1_{\{\tau \leq t\}}$.

Theorem 1 *P almost surely, for all $0 < t < \tau$,*

$$A_t = \lim_{x\to 0} A_t^x$$

where for $x > 0$,

$$A_t^x = \begin{cases} 0 & \text{if } t \leq T_1^x \\ A_{S_{n-1}^x} + 1_{\{x_1 < X_{T_n^x} < x_2\}} \int_x^{(t\wedge S_n^x)-T_n^x+x} \frac{F_2^{1-}(ds)}{F_2^-[s,\infty)} & \text{if } T_n^x < t < T_{n+1}^x \end{cases}$$

The proof of Theorem 1 is omitted. It is obtained, however, as a direct consequence of Theorem 6 of [19]. Theorem 1 tells us that the compensator increases only during excursions in the region (x_1, x_2). If the measure F_2^{1-} is absolutely continuous with respect to Lebesgue measure with density f_2^{1-} then

$$\lambda(t) = \begin{cases} 0 & \text{if } X_t \geq x_2 \\ \frac{f_2^{1-}(U_t)}{F_2^-[U_t,\infty)} & \text{if } x_1 < X_t < x_2 \end{cases}$$

is the intensity process (conditional hazard rate), i.e. $A_t = \int_0^{\tau\wedge t} \lambda(s)ds$. This follows from the monotone convergence theorem, once we observe

$$A_t^x = \int_0^t 1_{\{U_s > x\}} \lambda(s) ds.$$

For any given $T > 0$, we would like to compute the càdlàg version of the martingale $Y_t = E[1_{\{\tau\leq T\}}|\mathscr{F}_t]$. For each i let p_i be the function defined on $[0,\infty)$ as

$$p_i(t) = P^{x_i}(\tau \leq T - t).$$

Each p_i is a continuous function of t, since the distribution function of τ under each P^{x_i} is continuous.

Theorem 2 *P a.s. for all t such that $t < \tau \wedge T$*

$$Y_t = \begin{cases} \sum_{i=2}^N p_i(t) 1_{\{X_t = x_i\}} \quad \textit{if} \quad U_t = 0 \\ \\ \frac{1}{F_i^{\pm}[U_t,\infty]} \left(\int_{U_t}^\infty F_i^{1\pm}(ds) p_{i\pm 1}(g_t + s) + \int_{U_t}^\infty F_i^{0\pm}(ds) p_i(g_t + s) \right) \\ \textit{if} \quad U_t > 0 \textit{ and } X_{g_t} = x_i, \pm X_t > \pm X_{g_t}. \end{cases} \tag{3}$$

Proof We first derive Y_t for the first case, i.e. for all t s.t. $U_t = 0$. Since $Y = (Y_t)_{t\geq 0}$ is a bounded martingale, by the optional sampling theorem

$$Y_{S_n^x} 1_{\{S_n^x < T\wedge\tau\}} = E[1_{\{\tau\leq T\}}|\mathscr{F}_{S_n^x}] 1_{\{S_n^x<T\wedge\tau\}}.$$

On $\{S_n^x < T\wedge\tau\}$, $X_{S_n^x} = x_i$ for some $i \neq 1$. Therefore, by the strong Markov property of X at S_n^x we have that P a.s.

$$E[1_{\{S_n^x<T\wedge\tau\}}1_{\{\tau\leq T\}}1_{\{X_{S_n^x}=x_i\}}|\mathscr{F}_{S_n^x}] = p_i(S_n^x)1_{\{X_{S_n^x}=x_i\}}1_{\{S_n^x<T\wedge\tau\}}.$$

Therefore P a.s.

$$Y_{S_n^x}1_{\{S_n^x<T\wedge\tau\}} = \sum_{i=2}^{N} p_i(S_n^x)1_{\{X_{S_n^x}=x_i\}}1_{\{S_n^x<T\wedge\tau\}}. \tag{4}$$

From this we can easily deduce that P a.s. (4) holds for all $n\geq 1$ and $x\in\mathbb{Q}$ (therefore for all $x>0$). The regularity of x_i implies that the set $D(\omega) = \{S_n^x(\omega), n\geq 1, x>0\}$ is dense in $\{t : U_t(\omega) = 0\}$. By the martingale representation theorem of [19], Y does not jump on the set $\{t : U_t(\omega) = 0\} - D(\omega)$, hence by continuity we have the proof for the first case.

It remains to find Y_t for t such that $U_t > 0$ and $t < T\wedge\tau$. If $U_t > 0$ then we must have $T_n^x < t < S_n^x$ for some $n\geq 1$ and $x > 0$. From the previous step P-a.s. on $A_{i,\pm}^n \cap \{S_n^x \leq \tau\}$,

$$\begin{aligned} Y_{S_n^x} = \; & p_i\left((T_n^x - x) + (S_n^x - T_n^x + x)\right)1_{\{X_{S_n^x}=x_i\}} \\ & + p_{i\pm1}((T_n^x - x) + (S_n^x - T_n^x + x))1_{\{X_{S_n^x}=x_{i\pm1}\}} \end{aligned}$$

(1), (2), and the continuity of Y_t imply that P a.s. for all $t < T\wedge\tau$ satisfying $T_n^x < t < S_n^x$,

$$\begin{aligned} Y_t = \; & \frac{1}{F_i^{\pm}[t - T_n^x + x, \infty]} \left(\int_{t-T_n^x+x}^{\infty} F_i^{0\pm}(ds) p_i(T_n^x - x + s) \right. \\ & \left. + \int_{t-T_n^x+x}^{\infty} F_i^{1\pm}(ds) p_{i\pm1}(T_n^x - x + s) \right). \end{aligned}$$

Since $U_t = t - T_n^x + x$ and $g_t = T_n^x - x$, we have also proved the second case.

Remark 1 We can interpret (2) as follows. Let us consider the function

$$p(x,t) = P^x(\tau \leq T - t).$$

By the strong Markov property and the continuity of X,

$$p(x,t) < p(y,t) \qquad \text{for } x_1 \leq y \leq x.$$

For $t < T\wedge\tau$,

$$\pm p_{i\pm1}(t) < \pm Y_t 1_{\{X_{g_t}=x_i, \pm X_t > \pm X_{g_t}\}} < \pm p_i(t) \tag{5}$$

Therefore for fixed t, Y_t is lowest when X is in the region R_1. $p_{i\pm1}(t)$ and $p_i(t)$ give strict bounds for Y_t during an excursion in the region R_i, and Y_t attains one of the bounds at the end of the excursion. The value during the excursion should be interpreted as a weighted average of the values at the end of the excursion with the joint conditional distribution of the ending point and the duration of the excursion.

3 Bond Pricing

This section presents the economic interpretation of the previous model setup. Considered is a continuous trading economy with time horizon $[0,\infty)$. Traded are default-free zero-coupon bonds of all maturities, with a corresponding spot rate r_t, and a risky zero-coupon bond. For simplicity, we assume that the spot rate is deterministic.

The risky zero-coupon bond is issued by a firm that promises to pay 1 dollar at time T. Of course, the promise may not be fullfilled. If the firm defaults prior to time T, then (for simplicity) we assume that the bond pays off $\delta < 1$ dollars. Other payoff structures are possible, see Bielecki and Rutkowski [1]. We let $v(t,T)$ denote the time t value of the risky zero-coupon bond maturing at time T (as viewed by the market).

For pricing, we assume the existence of an equivalent probability measure such that all traded bond prices, discounted by the spot rate of interest, are martingales. Of course, it is well known that this is equivalent to assuming our economy is arbitrage-free. We let the probability measure P as defined in the preceding section be an equivalent martingale probability measure.[2]

We let the firm's asset value X_t follow the diffusion process characterized in the preceding section. As is standard in structural models (see Bielecki and Rutkowski [1]) we assume that default occurs when the firm's asset value falls below the level x_1. Thus, τ (the first hitting time of x_1) is our default time.

We assume that firm's management knows the asset value process filtration $\mathbb{G}$, but the market (the entity that determines the risky zero-coupon bond's price) only has the reduced information set represented by the filtration $\mathbb{F}$ above, corresponding to knowledge of the level crosses of the firm's asset value process X_t. The interpretation is that the firm makes public announcements about the firm's asset value only when there are significant changes in its economic standing (when the asset value reaches the crossings).

Given this structure, management values the firm's zero-coupon bond according to the formula

$$\begin{aligned} v^{mgmt}(t,T) &= E[\{\delta 1_{\{\tau\le T\}} + (1-1_{\{\tau\le T\}})\}e^{-\int_t^T r_s ds}|\mathscr{G}_t] \\ &= 1-[(1-\delta)E[1_{\{\tau\le T\}}|\mathscr{G}_t]]e^{-\int_t^T r_s ds} \\ &= 1-[(1-\delta)p(X_t,t)]e^{-\int_t^T r_s ds} \end{aligned}$$

for $t < T\wedge\tau$, where the last equality follows from the Markov property.

In contrast, the market values the firm's risky zero-coupon bond according to the formula:

$$v(t,T) = [1-(1-\delta)]E[1_{\{\tau\le T\}}|\mathscr{F}_t]e^{-\int_t^T r_s ds} = [1-(1-\delta)]Y_t e^{-\int_t^T r_s ds}$$

where Y_t is given in the previous section.

[2] There is an issue of uniqueness of the equivalent martingale probability measure. In general, if the market is incomplete, the equivalent martingale probability measure is not unique. In this circumstance, we select from the set of equivalent martingale measures, the one that is uniquely determined by an economic equilibrium in the (observed) economy as implied by the standard setting involving a specification of preferences and endowments.

In contrast to the management's using only X_t and $T-t$ to determine the price, the market evaluates the price using the following variables: X_{g_t}, U_t, $R(X_t)$, and $T-t$. The interdependence between these variables, which in turn determines the price, is complex and described in terms of the Lévy measures $F_i^{j\pm}$. However, a number of observations can be made without much difficulty.

First, as a consequence of Remark 1, for fixed t, the bond price is lower when the asset value is in a region closer to the default level x_1. When the asset value is at the boundaries of any region, the information of which is available to both the market and the management, the market price is equal to the management's price. When the asset value is in the interior of any region, i.e. during an excursion, the market does not know the true asset value, hence has to make an estimate of it. Consequently in this case the market's price is going to differ from the management's price.

In particular we are going to have the following relationship between the management's price and the market's price: Suppose at a given time t, $X_{g_t} = x_i$ and $\pm X_t > \pm X_{g_t}$. Let x_t solve the following equation:

$$p(x_t, t) = Y_t. \tag{6}$$

(6) has a unique solution in (x_i, x_{i+1}) (resp. (x_{i-1}, x_i)) due to Remark 1. Then

$$v^{mgmt}(t,T) - v(t,T) > 0 \text{ if } X_t > x_t$$

$$v^{mgmt}(t,T) - v(t,T) < 0 \text{ if } X_t < x_t.$$

The interpretation of this relationship is straightforward. When the true asset value is sufficiently far from the default level, the market underestimates the price, and the opposite happens when the asset value is sufficiently close to the default level.

Key objects which determine the price are the functions $p_i(t)$, and the measures $F_i^{j\pm}$, which are both related to the differential equation $\mathscr{A}u = \lambda u$. The following is well known. Let $\tilde{\Phi}_\lambda^+$ (resp. $\tilde{\Phi}_\lambda^-$) denote the decreasing (resp. increasing) solution of $\mathscr{A}\Phi = \lambda \Phi$ on $(-\infty, \infty)$. If τ^y is the hitting time of y then for $x \neq y$

$$E^x(e^{-\lambda \tau^y}) = \begin{cases} \frac{\tilde{\Phi}_\lambda^+(x)}{\tilde{\Phi}_\lambda^+(y)} & \text{if } x > y \\ \frac{\tilde{\Phi}_\lambda^-(x)}{\tilde{\Phi}_\lambda^-(y)} & \text{if } x < y \end{cases}. \tag{7}$$

Note that the right side is the Laplace transform of the density g_x^y of τ^y under P^x. We can compute $p(x,t)$ by integrating $g_x^{x_1}$. Equations (7) leads us to the Laplace transform of $g_x^{x_1}$, and inverting this we can find $g_x^{x_1}$.

D. Sezer [19] gives formulae to obtain the Laplace transforms of the functions $F_i^{j\pm}[x,\infty]$. Let $\Phi_{i,\lambda}^+$ (resp. $\Phi_{i,\lambda}^-$), (excluding $\Phi_{1,\lambda}^-$ and $\Phi_{N,\lambda}^+$, which are not defined), denote the decreasing (resp. increasing) solution of $\mathscr{A}\Phi = \lambda\Phi$ on $(-\infty, x_{i+1}]$ (resp. $[x_{i-1}, \infty)$) with boundary condition $\Phi(x_{i+1}) = 0$ (resp. $\Phi(x_{i-1}) = 0$) normalized to have $\Phi_{i,\lambda}^\pm(x_i) = 1$. For $\lambda > 0$ let

$$\psi_i^\pm(\lambda) = -\pm \frac{1}{2} \frac{(\Phi_{i,\lambda}^\pm)'(x_i)}{\Phi_{i,\lambda}^\pm(x_i)},$$

and similarly

$$\tilde{\psi}_i^{\pm}(\lambda) = -\pm \frac{1}{2}\frac{(\tilde{\Phi}_\lambda^{\pm})'(x_i)}{\tilde{\Phi}_\lambda^{\pm}(x_i)}.$$

We put $\psi_i^{\pm}(0) = \lim_{\lambda\to 0}\psi_i^{\pm}(\lambda)$ and $\tilde{\psi}_i^{\pm}(0) = \lim_{\lambda\to 0}\tilde{\psi}_i^{\pm}(\lambda)$. Finally we let

$$\psi_{i,i+1}(\lambda) = \psi_i^{+}(\lambda) + \tilde{\psi}_i^{-}(\lambda)$$

and

$$\psi_{i,i-1}(\lambda) = \psi_i^{-}(\lambda) + \tilde{\psi}_i^{+}(\lambda).$$

Theorem 3 *[19] Let*

$$\psi_i^{j\pm}(\lambda) = \int_{(0,\infty]}(1-e^{-\lambda x})F_i^{j\pm}(dx) = \lambda\int_0^\infty e^{-\lambda x}F_i^{j\pm}[x,\infty]dx.$$

Then

i) $\psi_i^{1\pm}(\lambda) = P^{x_i}(\tau^{x_{i\pm 1}} < \infty)\psi_{i,i+1}(0) - P^{x_i}(e^{-\lambda\tau^{x_{i\pm 1}}})\psi_{i,i+1}(\lambda)$,

ii) Excluding ψ_1^{0-} *and* ψ_N^{0+}, $\psi_i^{0\pm}(\lambda) = -\psi_i^{\pm}(0) + \psi_i^{\pm}(\lambda)$,

iii) $\psi_N^{0+}(\lambda) = \tilde{\psi}_N^{+}(\lambda)$ *and* $\psi_1^{0-}(\lambda) = \tilde{\psi}_1^{-}(\lambda)$.

The equation $\mathscr{A}u = \lambda u$ for positive λ is known to have explicit solutions for a number of choices for the coefficients of $\mathscr{A}$. The simplest case is when the coefficients are constant, i.e. when X is Brownian motion with drift. Restricting our attention to this case we can get an intuitive understanding of the relationship between bond prices and the model's parameters.

For the rest of the discussion we assume that X is Brownian motion with drift, i.e. X is given by

$$X_t = \mu t + \sigma B_t,$$

where B is a standard Brownian motion, $\mu \in \mathbb{R}$, and $\sigma > 0$. The corresponding infinitesimal generator is given by

$$\mathscr{A} = \frac{1}{2}\sigma^2\frac{d^2}{dx^2} + \mu\frac{d}{dx}.$$

We remark that X is an $\mathbb{R}$-valued process, therefore we will assume that X models the logarithmic transform of the asset value, which is equivalent to assuming that the asset value follows a geometric Brownian motion.[3] The levels x_i correspond to the logarithmic transform of the actual thresholds. The information structure does not change since the logarithm is a monotone transformation.

In this case, a formula for $p_i(t)$ is standard (see e.g. [14], p.10):

$$p_i(t) = N\left(\frac{-(x_i - x_1) - \mu(T-t)}{\sigma\sqrt{T-t}}\right) + e^{\frac{-2(x_i - x_1)\mu}{\sigma^2}}N\left(\frac{-(x_i - x_1) + \mu(T-t)}{\sigma\sqrt{T-t}}\right) \quad (8)$$

where $N(.)$ is the cumulative distribution function of the standard normal distribution.

[3] If the market observed X_t continuously, then the equivalent martingale probability measure would have $\mu = r - \sigma^2/2$.

We can also explicitly compute the measures $F_i^{j\pm}$, by inversion after computing their Laplace transforms using Theorem 3. We find that each $F_i^{j\pm}$ (excluding F_N^{0+} and F_1^{0-}) has a density, $f_i^{j\pm}$, with respect to Lebesgue measure, and

$$f_i^{1\pm}(x) = e^{\pm\frac{\mu|x_i - x_{i\pm1}|}{\sigma^2}} e^{-\frac{\mu^2}{2\sigma^2}x} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{n^2\pi^2\sigma^2}{|x_i - x_{i\pm1}|^2} e^{-\frac{n^2\pi^2\sigma^2 x}{|x_i - x_{i\pm1}|^2}}, \tag{9}$$

$$f_i^{0\pm}(x) = e^{-\frac{\mu^2}{2\sigma^2}x} \sum_{n=1}^{\infty} \frac{n^2\pi^2\sigma^2}{|x_i - x_{i\pm1}|^2} e^{-\frac{n^2\pi^2\sigma^2 x}{|x_i - x_{i\pm1}|^2}}. \tag{10}$$

F_N^{0+} (resp. F_1^{0-}) has mass at ∞ if and only if $\mu > 0$ (respectively $\mu < 0$). We have

$$F_N^{0+}(\infty) = \frac{|\mu| + \mu}{2\sigma}, \qquad F_1^{0-}(\infty) = \frac{|\mu| - \mu}{2\sigma}.$$

The restriction on $(0,\infty)$ of both F_N^{0+}, and F_1^{0-} have densities, with respect to Lebesgue measure, which we denote by f_N^{0+} and f_1^{0-}. These densities are equal and given by

$$f_N^{0+}(x) = f_1^{0-}(x) = e^{-\frac{\mu^2}{2\sigma^2}x} \frac{1}{\sigma\sqrt{2\pi x}}.$$

We remark that neither A_t nor Y_t should change as we scale the underlying process X, as long as we scale the points x_i with it. This is because if $R(.)$ is the region indicator function associated with the level set $\{x_1, \dots, x_n\}$, and $\tilde{R}(.)$ is the region indicator function associated with the level set $\{\sigma x_1, \dots, \sigma x_n\}$, then the processes $R(X_t)$ and $\tilde{R}(\sigma X_t)$ are identical. This is reflected in the formulae for $F_i^{j\pm}$. $\sigma F_i^{j\pm}$ depends only on μ/σ and $|x_i - x_{i\pm1}|/\sigma$, and Y_t does not depend on the scaling of $F_i^{j\pm}$. The functions $p_i(t)$, also, depend only on μ/σ and $|x_i - x_{i\pm1}|/\sigma$. Hence the parameters of the model are μ/σ and $|x_i - x_{i\pm1}|/\sigma$. For the rest of the discussion we assume $\sigma = 1$.

From Remark 1, we know that Y_t moves within the boundaries $0 < p_2(t) \dots < p_N(t)$ during $[0, \tau \wedge T]$. The parameters affect Y_t in two ways, by first changing the boundaries $p_i(t)$ and within these boundaries, by changing the relative position of Y_t. We first look at how the boundaries change. Each $p_i(\cdot)$ is a continuous monotone function in t with $p_i(T) = 0$. Clearly as μ increases, $p_i(\cdot)$ increases point-wise, because an increase in the drift makes it less likely to hit the default level. Also, the longer the distance $x_i - x_1$, the smaller p_i. The distances $x_i - x_{i\pm1}$ determine how far apart the boundaries, the p_i, are from each other.

We know that Y_t is lowest if $X_t > x_N$, in which case μ changes the rate of convergence of Y_t to 0. The higher μ, the faster the convergence. During an excursion in $[x_i, x_{i\pm1}]$ that began at x_i, (here we assume $x_{i\pm1} \neq x_N$) we know that Y_t is between $p_i(t)$ and $p_{i\pm1}(t)$. We can think of $F_i^{0\pm}$ and $F_i^{1\pm}$ as two forces acting on Y_t, the first one pulling Y_t towards $p_i(t)$, the latter towards $p_{i\pm1}(t)$. So the position of Y_t between $p_i(t)$ and $p_{i\pm1}(t)$ is determined by the relative strength of $F_i^{1\pm}$ versus $F_i^{0\pm}$. (9), (10) indicate that the ratio

$$\frac{f_i^{1\pm}}{f_i^{0\pm}}(x) \tag{11}$$

is monotone in x and converges to $\exp(\pm\mu|x_i - x_{i\pm1}|/\sigma^2)$ as $x \to \infty$. Therefore although initially dominated by $F_i^{0\pm}$, $F_i^{1\pm}$ gets stronger as the excursion continues, so, Y_t gets closer to $p_{i\pm1}(t)$. Of course, this should be interpreted in a relative sense, because $p_{i\pm1}(t)$ and $p_i(t)$ are also getting closer to each other as they both converge to 0. The limit

$$e^{\pm\frac{\mu|x_i - x_{i\pm1}|}{\sigma^2}}$$

gives us an indication on the role of μ. The drift μ changes the relative strength of $F_i^{1\pm}$ versus $F_i^{0\pm}$. In particular, for non-negative μ, f_i^{1+} is always dominated by f_i^{0+} whereas for positive μ, f_i^{1+} dominates f_i^{0+} for large x.

Figure 1 gives a simulation of how Y_t changes during an excursion. The excursion starts at $t = 0.2$ and goes on until $t = 0.3$. There are four levels, $x_1 = 0$, $x_2 = 0.25$, $x_3 = 0.75$, $x_4 = 1$. X has no drift, $\sigma = 1$, and $T = 1$. Given these data, Y_t is computed under all possible scenarios. These scenarios are (i) $X_t > x_4$, (ii) $x_3 < X_t < x_4$, $X_{g_t} = x_4$, (iii) $x_3 < X_t < x_4$, $X_{g_t} = x_3$, (iv) $x_2 < X_t < x_3$, $X_{g_t} = x_3$, (v) $x_2 < X_t < x_3$, $X_{g_t} = x_2$, (vi) $X_t < x_2$, $X_{g_t} = x_3$. We observe that scenario (i) gives the lowest Y_t, whereas (vi) gives the highest. The boundaries, p_2, p_3, p_4 are also given. We note that the values of Y_t under (ii) and (iii), start from p_4 and p_3, respectively, and converge to each other as the excursion continues. This is because $F_4^{1-} = F_3^{1+}$ (as $\mu = 0$) converges to $F_4^{0-} = F_3^{0+}$. The rate of convergence is solely determined by $x_4 - x_3$ (see the discussion on jump rates). Under ((iv) and (v)), we see a similar picture, except with a slower rate of convergence since $x_3 - x_2 > x_4 - x_3$.

We remark that under (ii) , Y_t is always lower than it is under (iii), and this is always the case whenever $\mu \geq 0$. Indeed, when $\mu > 0$, for $u > 0$, and $y > u$,

$$\frac{f_{i+1}^{0-}(y)}{F_i^{-}[u,\infty)} \geq \frac{f_i^{1+}(y)}{F_i^{+}[u,\infty)}$$

and when $x_i < X_t < x_{i+1}$ these are the weight distributions for the lower boundary, p_{i+1}, under the two possible scenarios, $X_{g_t} = x_i$ and $X_{g_t} = x_{i+1}$.

Last, we look at the jump rates

$$\lambda_i^{j\pm}(x) = \frac{f_i^{j\pm}(x)}{F_i^{\pm}[x,\infty]}.$$

$\lambda_i^{j\pm}(x)$ corresponds to the instantaneous likelihood of an excursion ending at the point $i \pm 1$, when its age has reached x. We note that λ_2^{1-} is related to the intensity process, λ_t, of the default time as follows:

$$\lambda_t = \begin{cases} 0 \text{ if } X_t \geq x_2 \\ \lambda_2^{1-}(U_t) \text{ if } x_1 < X_t < x_2. \end{cases}$$

Theorem 4 *i) $\lambda_i^{0\pm}$ is monotone decreasing with*

$$\lim_{x\to 0} \lambda_i^{0\pm}(x) = \infty,$$

$$\lim_{x\to\infty} \lambda_i^{0\pm}(x) = \left(\frac{\pi^2\sigma^2}{(x_i - x_{i\pm1})^2} + \frac{\mu^2}{2\sigma^2}\right)(1 + e^{\pm\frac{\mu|x_i - x_{i\pm1}|}{\sigma^2}})^{-1}.$$

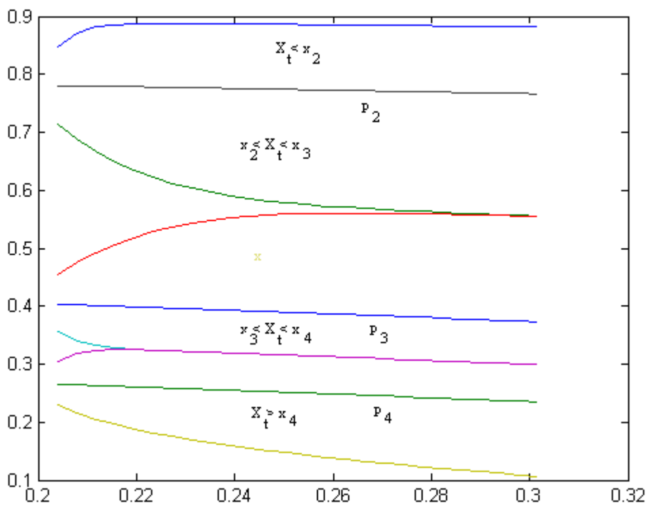

Fig. 1 *Y_t, during an excursion. Excursion starts at $t = 0.2$, goes on till $t = 0.3$.*

ii) $\lambda_i^{1\pm}$ is monotone increasing with

$$\lim_{x\to 0} \lambda_i^{1\pm}(x) = 0,$$

$$\lim_{x\to\infty} \lambda_i^{1\pm}(x) = \left(\frac{\pi^2\sigma^2}{(x_i - x_{i\pm 1})^2} + \frac{\mu^2}{2\sigma^2}\right)(1 + e^{-\pm\frac{\mu|x_i - x_{i\pm 1}|}{\sigma^2}})^{-1}.$$

Proof We first prove (i) for λ_1^{0-} and λ_N^{0+}. It is enough to show $F_1^{0-}[x,\infty)(= F_N^{0+}[x,\infty))$ is log-convex. This is true because f_1^{0-} is log-convex, and the integral of a log-convex function over the interval $[x,\infty)$ is log-convex as a function of x. The assertions involving the limits hold with $x_0 = -\infty$ and $x_{N+1} = \infty$, and can be shown by an application of L'Hôpital's rule.

The rest of the proof is concerned with $\lambda_i^{j\pm}$ excluding λ_1^{0-} and λ_N^{0+}. We observe that all the terms after the first term inside the sums in (10) and (9) are negligible for large t. Hence the limits as $t \to \infty$ are found by direct computation using the first terms.

Next, we show the monotonicity of $\lambda_i^{j\pm}$. Since

$$\lambda_i^{j\pm}(x) = \frac{f_i^{j\pm}(x)}{F_i^{j\pm}[x,\infty)} \frac{F_i^{j\pm}[x,\infty)}{F_i^{\pm}[x,\infty)} \tag{12}$$

and by (11), it is enough to show (a) $F_i^{0\pm}[x,\infty)$ is log-convex, and (b) $F_i^{1\pm}[x,\infty)$ is log-concave. We put

$$a_n = \frac{n^2\pi^2\sigma^2}{(x_i - x_{i\pm 1})^2},$$

$$b_n = \frac{n^2\pi^2\sigma^2}{(x_i - x_{i\pm 1})^2} + \frac{\mu^2}{2\sigma^2}.$$

Then

$$F_i^{0\pm}[x,\infty) = \sum_{n=1}^{\infty} \frac{a_n}{b_n} e^{-b_n x}. \tag{13}$$

(a) follows from Hölder's inequality. Indeed, for $0 < x_1 < x_2$ and $\lambda > 0$, let

$$x_n = e^{-b_n \lambda x_1}$$

and

$$y_n = e^{-b_n(1-\lambda)x_2}.$$

Letting

$$p = \frac{1}{\lambda}, \qquad q = \frac{1}{1-\lambda},$$

Hölder's inequality gives

$$\begin{aligned} F_i^{0\pm}[\lambda x_1 + (1-\lambda)x_2, \infty) &= \sum_{n=1}^{\infty} \frac{a_n}{b_n} x_n y_n \\ &\leq (\sum_{n=1}^{\infty} \frac{a_n}{b_n} x_n^p)^{1/p} (\sum_{n=1}^{\infty} \frac{a_n}{b_n} y_n^q)^{1/q} \\ &= F_i^{0\pm}[x_1,\infty)^{\lambda} F_i^{0\pm}[x_2,\infty)^{(1-\lambda)}. \end{aligned}$$

To prove (b), we show that the second derivative of $\log F_i^{1\pm}[x,\infty)$ is negative. Since

$$(\log F_i^{1\pm}[x,\infty))'' = -\frac{(f_i^{1\pm}(x))' F_i^{1\pm}[x,\infty) + (f_i^{1\pm}(x))^2}{F_i^{1\pm}[x,\infty)^2},$$

it is enough to show

$$(f_i^{1\pm}(x))' F_i^{1\pm}[x,\infty) + (f_i^{1\pm}(x))^2 > 0.$$

Note

$$\begin{aligned} &(f_i^{1\pm}(x))' F_i^{1\pm}[x,\infty) + (f_i^{1\pm}(x))^2 \\ &= \sum_{n=1}^{\infty} (-1)^n a_n b_n e^{-b_n t} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{a_n}{b_n} e^{-b_n t} \\ &\qquad - \sum_{n=1}^{\infty} (-1)^{n+1} a_n e^{-b_n t} \sum_{n=1}^{\infty} (-1)^n a_n e^{-b_n t} \\ &= \sum_{n=1}^{\infty} \sum_{m>n} (-1)^{n+m+1} (a_m b_m \frac{a_n}{b_n} + a_n b_n \frac{a_m}{b_m} - 2a_m a_n) e^{-(b_m+b_n)t} \end{aligned}$$

To show that the inner sum above is positive we borrow an inequality from the literature on inequalities in analysis. This inequality ([3], p.245) is a version of an inequality known as Szegö's inequality, is due to Brunk, and says the following:

Let f be convex on $[0,b]$. If $b \geq a_1 \geq \ldots \geq a_n \geq 0$, and $1 \geq h_1 \geq \ldots \geq h_n \geq 0$ then for odd n,

$$\sum_{i=1}^{n}(-1)^{i-1}h_i f(a_i) \geq f(\sum_{i=1}^{n}(-1)^{i-1}h_i a_i).$$

Note

$$\sum_{m>n}(-1)^{n+m+1}(a_m b_m \frac{a_m}{b_m} + a_n b_n \frac{a_m}{b_m} - 2a_m a_n)e^{-(b_m+b_n)t}$$

$$= \sum_{k=1}^{\infty}(-1)^{k-1}a_{n+k}a_n(\frac{b_{n+k}}{b_n} + \frac{b_n}{b_{n+k}} - 2)e^{-(b_{n+k}+b_n)t}.$$

Since $f(x) = e^{-x}$ is convex, and a_{n+k} is non-decreasing as k increases, non-negativity follows from Brunk's inequality once we show that

$$c_k = \frac{b_{n+k}}{b_n} + \frac{b_n}{b_{n+k}} - 2$$

is non-decreasing, too. To see this we first observe that $\frac{b_{n+k}}{b_n}$ is non-decreasing and greater than 1. This implies that c_k is non-decreasing, because the function $x + 1/x$ is non-decreasing for $x > 1$. This completes the proof of (b).

Last we compute the limits of $\lambda_i^{j\pm}$ as $x \to 0$. By a simple ratio test of the terms of the series in (10) and (13),

$$\lim_{x\to 0}\frac{f_i^{0\pm}(x)}{F_i^{0\pm}[x,\infty)} = \infty. \tag{14}$$

Since $F_i^{1\pm}(0,\infty) < \infty$ and $F_i^{0\pm}(0,\infty) = \infty$,

$$\lim_{x\to 0}\frac{F_i^{1\pm}[x,\infty)}{F_i^{\pm}[x,\infty)} = 0. \tag{15}$$

(12),(14) and (15) imply that $\lambda_i^{0\pm}(x) \to \infty$. Also since $F_i^{1\pm}[x,\infty)$ is log concave,

$$\lim_{x\to 0}\lambda_i^{1\pm}(x) = c\lim_{x\to 0}\frac{F_i^{1\pm}[x,\infty)}{F_i^{\pm}[x,\infty)}$$

for some $c < \infty$. Hence $\lim_{x\to 0}\lambda_i^{1\pm}(x) = 0$ by (15).

In Figure 2-3 we plot λ_i^{0+} versus λ_i^{1+} for different values of μ. We note as μ increases, λ_i^{0+}decreases and λ_i^{1+} increases, both point-wise. The limit of λ_i^{0+} is greater than the limit of λ_i^{1+} for $\mu > 0$ and this is reversed when $\mu < 0$. Both limits are the same when $\mu = 0$. The rate of convergence to the limit is determined by $x_i - x_{i+1}$, note that this is the only parameter attached to n^2 in the exponent of the summands in (10) and (9). (See Figure 4). The convergence of the jumping rates to a constant tells us that as the excursion continues, the dependence on the past (i.e. the age and the beginning point of the excursion) disappears, and the jump distribution starts to look like the distribution of a jump of a simple Markov process.

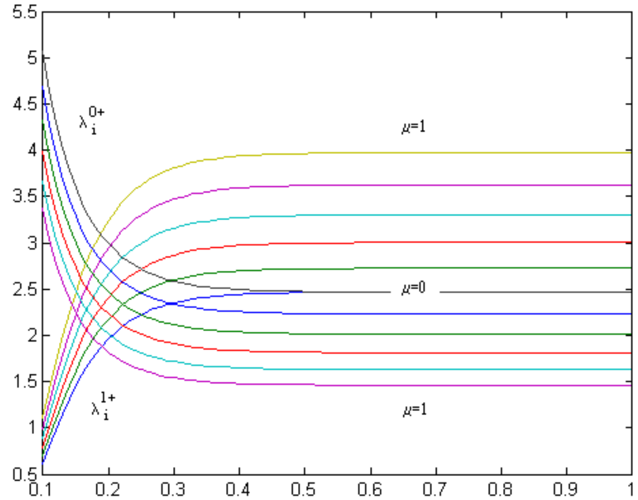

Fig. 2 λ_i^{1+} *versus* λ_i^{0+}. μ *varies from* 0 *to* 1 *with* 0.2 *increments.* $\sigma = 1$, $x_{i+1} - x_i = 1$.

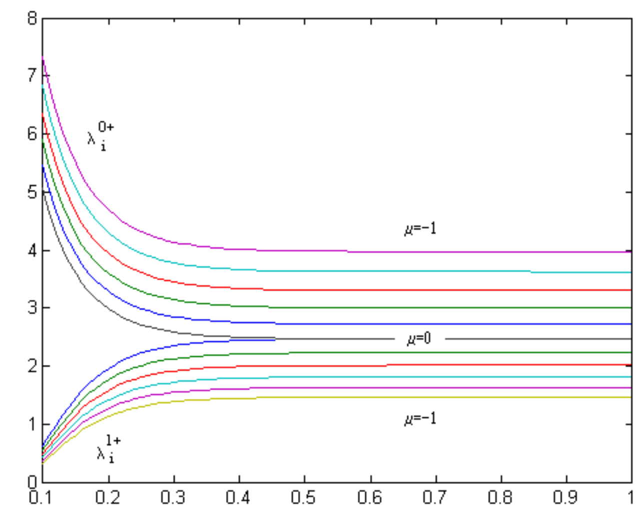

Fig. 3 λ_i^{1+} *versus* λ_i^{0+}, μ *varies from* 0 *to* 1 *with* 0.2 *increments.* $\sigma = 1$ $x_{i+1} - x_i = 1$.

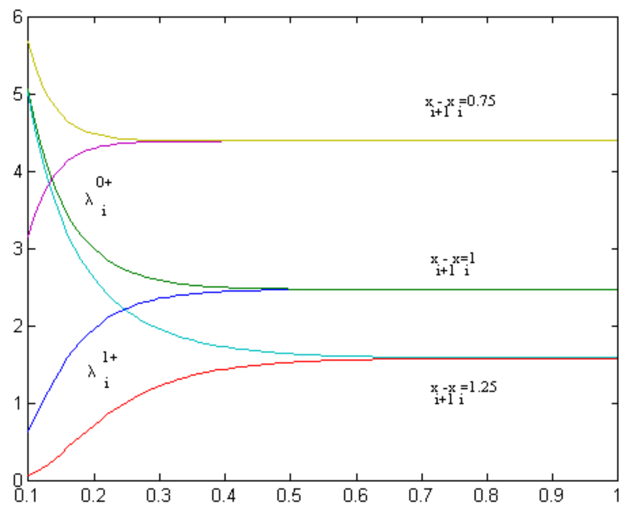

Fig. 4 λ_i^{1+} *versus* λ_i^{0+}, $x_{i+1} - x_i$ *varies from* 0.75 *to* 1.25. $\mu = 0$, $\sigma = 1$

4 Conclusion

This paper provides an alternative credit risk model based on information reduction where the market only observes the firm's asset value when it crosses certain levels, interpreted as changes significant enough for the firm's management to make a public announcement. For a class of diffusion processes we are able to provide explicit expressions for the firm's default intensity process and its zero-coupon bond prices.

5 Acknowledgements

The authors wish to thank the referee for helpful comments, and the Associate Editor for pointing out some important references we were unaware of.

References

1. T. Bielecki and M. Rutkowski, 2002, *Credit Risk: Modeling, Valuation and Hedging*, Springer-Verlag: New York.
2. R. Blumenthal and R. Getoor, 1968, *Markov Processes and Potential Theory*, Academic Press: New York
3. P.S. Bullen, 1998, *A Dictionary of Inequalities*, Addison Wesley Longman Limited.
4. U. Çetin, R. Jarrow, P. Protter and Y. Yildirim, 2004, "Modeling Credit Risk with Partial Information," *The Annals of Applied Probability* 14, 1167-1178.

5. P. Collin-Dufresne, R. Goldstein and J. Helwege, 2003, Is credit event risk priced? Modeling contagion via the updating of beliefs, *Working paper*, Carnegie Mellon University.
6. Duffie, J. D., and D. Lando, 2001, “Term Structure of Credit Spreads with Incomplete Accounting Information,” *Econometrica*, 69, 633-664.
7. K. Giesecke and L. Goldberg, 2004, “Forecasting Default in the Face of Uncertainty,” *Journal of Derivatives*, 12, 14-25
8. X. Guo, R. Jarrow, Y. Zeng, 2005, “Information Reduction in Credit Risk Models,” working paper, Cornell University.
9. J. Jacod and J. Mémin, 1976, “Un théorème de représentation des martingales pour les ensembles régénératif,” *Séminaire de Probabilités* 1003, 2439
10. R. Jarrow and S. Turnbull, 1992, “Credit Risk: Drawing the Analogy,” *Risk Magazine*, 5 (9).
11. R. Jarrow and S. Turnbull 1995, “Pricing Derivatives on Financial Securities Subject to Credit Risk.” *Journal of Finance*, 50, 1, 53-85.
12. R. Jarrow and P. Protter, 2004, “Structural versus Reduced Form Models: A New Information Based Perspective,”*Journal of Investment Management*, 2 (2), 1-10.
13. M. Jeanblanc and S. Valchev, 2005, “Partial Information and Hazard Process, ”*International Journal of Theoretical and Applied Finance*, 8, 807-838.
14. D. Lando, 2004, *Credit Risk Modeling* Princeton University Press, Princeton, New Jersey.
15. S. Kusuoka, 1999, “A remark on default risk models,” *Advances in Mathematical Economics*, 1, 69-82.
16. R. C. Merton, 1974, “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates,” *Journal of Finance*, 29, 449-470.
17. H. Nakagawa, 2001, “A filtering model on default risk, ” *J. Math. Sci. Univ. Tokyo*, 8, 107142.
18. P. Schönbucher and M. Tolotti, “Credit risk under incomplete accounting information: A discretized approach in filtering language, ”Preprint (Masters’ thesis of Tolotti, under the supervision of Schönbucher, at ETH Zürich).
19. D. Sezer, 2005, “Filtration shrinkage by level crossings of a diffusion”. Annals of Probability. Forthcoming.
20. M. Sharpe, 1988, *General Theory of Markov Processes* Academic Press, New York.