WP 2019/5

Cheikh Mbaye and Frédéric Vrins

Affine term-structure models: A time-changed approach with perfect fit to market curves

LOUVAIN FINANCE

LFIN

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/lfin/working-papers.html

Affine term-structure models : a time-changed approach with perfect fit to market curves

Cheikh Mbaye[1] & Frédéric Vrins[1,2*]

1. Louvain Finance Center (LFIN)

2. Center for Operations Research and Econometrics (CORE)

UC Louvain, Belgium

Abstract

We address the so-called *calibration problem* which consists of fitting in a tractable way a given model to a specified term structure like, e.g., yield, prepayment or default probability curves. Time-homogeneous jump-diffusions like Vasicek or Cox-Ingersoll-Ross (possibly coupled with compound Poisson jumps, JCIR, a.k.a. SRJD), are tractable processes but have limited flexibility; they fail to replicate actual market curves. The deterministic shift extension of the latter, Hull-White or JCIR++ (a.k.a. SSRJD) is a simple but yet efficient solution that is widely used by both academics and practitioners. However, the shift approach may not be appropriate when positivity is required, a common constraint when dealing with credit spreads or default intensities. In this paper, we tackle this problem by adopting a time change approach, leading to the TC-JCIR model. On the top of providing an elegant solution to the calibration problem under positivity constraint, our model features additional interesting properties in terms of variance. It is compared to the shift extension on various credit risk applications such as credit default swap, credit default swaption and credit valuation adjustment under wrong-way risk. The TC-JCIR model is able to generate much larger implied volatilities and covariance effects than JCIR++ under positivity constraint, and therefore offers an appealing alternative to the shift extension in such cases.

Keywords: model calibration, credit risk, stochastic intensity, jump-diffusions, term-structure models, time-change techniques

*Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium. E-mail: frederic.vrins@uclouvain.be. The research of Cheikh Mbaye is funded by the National Bank of Belgium and an FSR grant. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the National Bank of Belgium. The work of F. Vrins was supported by the Fonds de la Recherche Scientifique F.S.R.-FNRS, Grant J.0037.18.

1 Introduction

Model calibration is a standard problem in many areas of finance Brigo and Mercurio (2006); Joshi (2003); Veronesi (2010). It consists of tuning a model such that it "best fits" market quotes at a given time. As an example, financial markets provide a set of prices associated with liquid instruments, that openly trade on the market. Alongside with risk management (hedging), the main purpose of a model here is to act as an "interpolation/extrapolation" tool, i.e. to obtain the value of products at a given time t for which the market does not disclose prices in a transparent way. This could happen because either the product to be priced is "exotic" (i.e. is too "special", it does not quote openly on a platform, only on a bilateral basis) or because its cashflow schedule is not in line with the products that currently trade openly at t (a situation that commonly happens since products that were "standard" at inception, may have time-to-expiry or moneyness levels that are no longer "standard" afterwards).

Mathematically, model calibration is nothing but an optimization problem. Starting from a set of prices quoted on the market (called "market prices") for a set of specific financial products (called "calibration instruments"), model calibration consists of computing the model parameters such that the prices generated by the model (called "model prices") best fit to the market prices, according to some error function. Model calibration is crucial in finance; it is strongly related to arbitrage opportunities. In practice, only models that are able to reproduce the market prices of "simple instruments" (either in a perfect way, or at least up to the bid-ask spread) are trustworthy enough when it comes to pricing other instruments. For instance, one can price exotic derivatives (like barrier options) using stochastic volatility model like Heston in a semi-analytical way Carr and Madan (1999); Heston (1993). The parameters of the Heston model will be obtained by "calibration" to a volatility surface, i.e. to a set of liquid ("plain vanilla") options, like European calls and puts of various strikes and maturities. The justification behind this is that in a no-arbitrage, complete market setup, the price of an option can be obtained by computing the cost of setting up a self-financing hedging strategy. This cost depends on the prevailing prices of the hedging instruments. If the model fails to correctly price the latter, there is no chance it can correctly price the option.

In this work, we focus on financial calibration problems arising in other asset classes: interest-rates and credit Brigo and Mercurio (2006); Duffie and Singleton (2003). When specifying an interest rate model to price a derivative on, say, the Libor 3M index, one needs to make sure that the model generates a discount curve that is in line with that extracted from market quotes

of simpler Libor 3M-indexed products. In this case, the set of calibration instruments could be forward rate agreements (FRA), interest rate swaps (IRS), as well as vanilla cap/floors or swaptions. Similarly, adjusting the value of derivatives for counterparty risk (a problem known as credit valuation adjustment, or CVA) generally involves a stochastic model to represent the default of the counterparty with whom the trade is executed. The default probability of the counterparty can be extracted from a set of calibration instruments, which prices are driven by the default likelihood of the counterparty, e.g., corporate bonds of credit default swaps (CDS). In this context, the default model must be "calibrated" in such a way that the default probability curve generated by the stochastic model agrees with that implied from the prices of the corresponding instruments (see, e.g., Gregory (2010) and Stein and Pong (2011) for a general overview of CVA and Brigo et al. (2014) for a discussion of bilateral CVA in presence of collateralization agreements).

The calibration constraint rises practical issues. Indeed, the models that are actually used in the industry must have a tractability that is compatible with real-time pricing but, as explained above, must be flexible enough to match the information conveyed by the calibration instruments. Affine term structure models (ATSM) have been extensively used in fixed income modeling because of their analytical tractability. See, e.g., Duffie et al. (2003) and Duffie and Kan (1996) for an excellent review and a mathematical analysis of this class of processes. In practice, homogeneous affine jump diffusion (HAJD) models are extremely popular. The Vasicek (Ornstein-Uhlenbeck) model Vasicek (1977) is a short-rate model being widely used in both industry and academia. It is a time-homogeneous affine diffusion model that postulates Gaussian dynamics. If negative rates need to be ruled out, positive dynamics like the CIR (Cox-Ingersoll-Ross, also known as square-root diffusion, SRD) Cox et al. (1985) can be preferred, possibly with independent compounded Poisson jumps (JCIR or SRJD). However, it is in general impossible with either models (even in a multi-factor setup) to achieve a perfect fit: the flexibility of HAJD is limited, they are in general unable to generate a given discount curve. The same problem arises when dealing with credit derivatives: it is generally impossible to make sure that the default intensity process, modeled with HAJD dynamics, will generate a default probability curve that is in line with the corresponding curve, exogenously given by the market via the calibration instruments.

Several routes can be followed to deal with this issue. The first one consists of disregarding this lack of flexibility. Nevertheless, working with a model that fails to yield a perfect fit to

the market is often unacceptable in practice. Indeed, as explained above, the models are used to value derivatives positions, and a mismatch with the market can introduce a tremendous bias in the valuation of the book of companies or financial institutions. Another possibility is to significantly increase the complexity of the models. This is often to be avoided in practice, for computational, identification or over-fitting issues. A trade-off consists of extending the "simple models" in such a way that they can fit to the market. In fact, several authors show that a great flexibility can be obtained by shifting HAJD models in a deterministic way. Dybvig (1997) for instance show that the term structure of interest rates can be reproduced by adding a deterministic shift to Ho and Lee (1986) or Vasicek processes. Later, Brigo and Merccurio (2001) extended this idea to a broader class of models, thereby providing a simple but very clever solution to the calibration problem. Instead of considering a HAJD, one could simply adjust it with a deterministic function φ: the resulting process will have the required flexibility. When shifted this way, the Vasicek, CIR and JCIR models respectively correspond to the Hull-White, the CIR++ or the JCIR++ (a.k.a. SSRJD) models Brigo and Mercurio (2006). This trick is actually very powerful: it solves the calibration problem at no cost, since the model's dynamics remain affine. Moreover, the shift function is known analytically, as a function of the parameters of the underlying HAJD and the market curve to be fitted by the model.

Yet, this approach suffers from an important limitation. Because of the shift, there is no reason that the range of the shifted process agrees with that of the underlying HAJD process. For instance, shifting a positive process with a deterministic function may result in a process that could take on negative values. It all depends on the mismatch between the information conveyed by the calibration instruments on the one hand, and the parameters of the underlying HAJD on the other hand. In general, there is no reason to believe that the implied shift function will preserve the range of the HAJD model. This is problematic in many cases, and in credit risk modeling in particular: negative default intensities, for instance, make no sense. To circumvent this issue, one could think of adding a non-negativity constraint on the shift in the calibration step. But, as we will show, this drastically restricts the parameters of the underlying HAJD, hence the randomness embedded in the model. This explains why this solution is often not considered by practitioners: the shift approach (without positivity constraint) remains the standard approach, even if positivity is required, theoretically speaking. It seems that in absence of a valid alternative, one actually prefers to rely on a model providing a perfect fit, even though the latter suffers from theoretical inconsistencies.

In this paper we introduce an alternative to the deterministic shift. Using an equally simple – but intrinsically different – technique, we adjust a HAJD so as to allow for a perfect fit to a given market curve, without affecting the model's tractability, but also without introducing the aforementioned inconsistencies. More specifically, instead of shifting a HAJD, we time-change it. Time change techniques were first studied in 1965 Dambis (1965); Dubins and Schwartz (1965). The first application to finance dates back from the early 2000. Geman et al. used Lévy processes and interpreted the new time scale as the business time, in contrast with the calendar time Geman et al. (2001). This was then applied to stochastic volatility models Carr et al. (2003). Thanks to subordinated Lévy models, the authors introduced the leverage effect, as well as a long-term skew. Many other financial applications of time change techniques can be found in the review Swishchuk (2016). More recently, Mendoza-Arriaga and Linetsky used stochastic time change processes to introduce two-side jumps in positive processes. The analytical tractability of the resulting model is preserved to some extend. This model has been recently applied to counterparty credit risk Mbaye and Vrins (2018). In this work, we exploit the time change idea in yet another way, to solve a completely different problem. Our purpose is to time-change HAJDs so as to obtain models with the desired calibration flexibility, without affecting tractability and preserving the range of the original process. The intuition is that by slowing down or speeding up the time of the latent HAJD, at the appropriate rate, one would obtain a model that could fit most discount curves, and actually every default probability curve. Moreover, the time change function is easily found using simple numerical methods (namely, inversion of easy functions or ordinary differential equation). Eventually, our time-changed HAJD is proven to feature larger implied volatilities compared to the corresponding valid (i.e. non-negative) shifted HAJD. To illustrate the power of our approach, we provide two applications taken from credit: pricing of CDS options and computation of derivatives pricing accounting for counterparty risk under exposure-credit dependence (wrong-way risk, WWR). In either cases, all the considered default models perfectly fit the risk-neutral default probability curve extracted from market quotes associated to the CDS of the reference entity. The obtained results illustrate the nice feature of large implied volatilities : they are able to generate larger option prices compared to the shift approach calibrated on a same probability curve under non-negativity constraint.

Eventually, observe that although we focus on examples featuring reduced-form models when pricing of credit-sensitive instruments, our approach is of potential use for other models, in many areas of finance and insurance. Ongoing work suggests that it can be applied to other

default models, including the firm-value (structural) models Merton (1974), but also to linear-rational (polynomial) models Filipovic et al. (2017). Other models could be considered as well, like Jeanblanc and Vrins (2018) or Crépey et al. (2012).In terms of applications, the proposed method can be used in life insurance, to calibrate mortality rate to mortality tables. Our time-changed process could also be used to model prepayment rates in mortgage-backed securities (MBS). These products naturally exhibit a negative convexity due to a the negative relationship between interest and prepayment rates: householders tend to refinance their loans when interest rates drop. This calls for a stochastic prepayment (i.e. positive) rate, that will be negatively correlated with interest rates, and which parameters could be calibrated so as to agree with the averaged values given in the PSA measure, the indicator attached to MBS securities that characterizes the prepayment speed in MBS Veronesi (2010). Eventually, the proposed method could be applied in many other applications, including the modeling of performance degradation of devices or materials through time, which average outstanding performances evolution are given according to some quality standards.

The paper is organized as follows. In Section 2 the calibration problem is introduced and two specific cases (cashflow discounting and probability curves) are discussed. We then recall in Section 3 how a shifted version of time-homogeneous affine jump diffusions can fit every discount curve. We pay specific attention to the case where the resulting process needs to meet a positivity constraint. We then introduce in Section 4 our alternative model, specifically devoted to this case, focusing on the most common HAJD, namely the Vasicek and JCIR (generalizing the CIR) models. Eventually, we compare in Section 5 our model's performance to that of the shift approach on three different pricing problems taken from credit risk: CDS curve calibration, pricing of CDS options and pricing of credit valuation adjustment under wrong-way risk.

2 The calibration problem

Consider a given time-s *market curve* $P_s^{market}(t)$, $t \geq s$. The calibration problem consists of finding, for a given model, the (set of) parameter(s) $\Xi = \Xi^\star$ such that the corresponding *model curve* $P_s^{model}(t) = P_s^{model}(t;\Xi)$ "best fits" the market curve, according to some criterion. Mathematically speaking, this is an optimization problem that consists of finding a set of parameters that minimizes an error function between model and market values,

$$\Xi^\star := \arg\min_{\Xi} \left\| P_s^{model}(\cdot;\Xi) - P_s^{market}(\cdot) \right\| , \tag{1}$$

where $\|f(\cdot) - g(\cdot)\|$ represents a divergence measure between two functions f, g. In practice, one often computes the mean-square error (MSE) between f and g on a set of maturities $\mathcal{T} := \{T_1, \ldots, T_n\}$:

$$\|f(\cdot) - g(\cdot)\| := \frac{1}{n} \sum_{i=1}^{n} \left(f(T_i) - g(T_i)\right)^2 \ . \tag{2}$$

A model with parameter Ξ is said to *perfectly fit* the market up to horizon T whenever $P_s^{model}(t; \Xi) = P_s^{market}(t)$ for all $s \leq t \leq T$ or using a shorthand notation, $P_s^{model} \equiv P_s^{market}$.

2.1 Setup

A model can be either static or dynamic. For instance, the Nelsen-Siegel model Nelson and Siegel (1987) postulates a parametric form for the yield curve, but is a static model: the resulting curve does not correspond to the yield curve generated by the dynamics of a stochastic model. We focus on continuous-time dynamic models in the sequel.

We consider a frictionless market free of arbitrage opportunities in which trading takes place continuously over the time interval $[0, T]$, where T is a fixed time horizon. Uncertainty in the market is modelled through a filtered probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{Q})$. In this setup, $\mathbb{G} = (\mathcal{G}_t, t \in [0, T])$ represents the information flow and corresponds to the filtration generated by the stochastic market variables (risk factors, prices, interest rates, default intensities, default event, etc), $\mathcal{G} := \mathcal{G}_T$, and $\mathbb{Q}$ denotes the risk-neutral probability measure referred to as the *pricing measure*. In the sequel, we shall focus on a specific class of P^{model} and P^{market} functions: we assume they are *discount curves*, a set of functions that we now define.

Definition 1 (Discount curve)**.** *A time-s discount curve is any differentiable function of the form*

$$P_s : [s, \infty) \to \mathbb{R}_0^+, \quad t \mapsto P_s(t)$$

satisfying $P_s(s) = 1$.

In the specific $s = 0$ case, a *time-0 discount curve* $P_0(t)$ is simply called a *discount curve* and is noted $P(t)$, assuming implicitly that $t \geq 0$. Any time-s discount curve admits an exponential-integral from:

Lemma 1. *Every time-s discount curve* P_s *admits a representation in terms of time-s instantaneous forward rate curve* f_s*:*

$$P_s(t) = e^{-\int_s^t f_s(u)du} \ , \quad t \geq s.$$

If moreover $P_s(t)$ is strictly decreasing on (s, ∞), then $f_s(t)$ is strictly positive for all $t > s$.

Proof. Since $P_s(t) > 0$ for all $t \geq s$ and $P_s(t)$ is differentiable with respect to t on (s, ∞), then one can define the time-s instantaneous forward rate function as $f_s(t) := \frac{-1}{P_s(t)} \frac{d}{dt} P_s(t) = -\frac{d}{dt} \ln P_s(t)$ for all $t > s$; the value $f_s(s)$ is not identified but can be defined by, e.g., the limit as $t \downarrow s$. Moreover, if P_s is strictly decreasing on (s, ∞) then $f_s(t) > 0$ for all $t > s$. □

Discount curves are of paramount importance in finance. As suggested by the name, $P_s(t)$ allows one to compute the time-s value of a cashflow paid at time $t \geq s$, both in a credit risk-free and credit risky setup. As a special case of the second framework, they encompass survival probability curves, defined as one minus cumulative distribution functions. This is elaborated in the next two subsections.

2.1.1 Discounting in a default-free market

In this application, $P_s(t)$ stands for the time-s price of a risk-free zero-coupon bond (ZCB) with maturity t and face value 1, denominated in a given currency. In particular, $P_s^{market}(t)$ and $P_s^{model}(t; \Xi)$ respectively give the market and the model prices of that instrument.

Indeed, in a no-arbitrage setup, the price of a financial instrument paying a single cashflow (payoff) at a given maturity is given by the risk-neutral conditional expectation of the payoff, discounted at the risk-free rate from the payment date (maturity) back to the valuation date. Adopting a short-rate model, $P_s^{model}(t)$ corresponds to the $\mathbb{Q}$-expectation of the stochastic discount factor $D_s(t) := e^{-\int_s^t r_u du}$, the negative exponential of the risk-free short-rate process r, integrated from the valuation time s up to the payment time t, conditional upon the information prevailing at the pricing time Brigo and Mercurio (2006):

$$P_s^{model}(t) = \mathbb{E}\left[1 D_s(t) | \mathcal{G}_s\right] = \mathbb{E}\left[e^{-\int_s^t r_u du} \middle| \mathcal{G}_s\right] =: P_s^r(t) \ .$$

In this context, we aim at finding a model x to depict the risk-free short rate dynamics r that would be tractable enough, and provide a perfect fit to any yield curve, the curve that gives the set of prices of ZCBs with increasing maturities.

2.1.2 Discounting in a defaultable market

Adopting the same framework as before, the time-s price of a zero-coupon bond paying one unit of currency at time $t \geq s$ *contingent on the fact that the issuer doesn't default prior to*

the payment date is given by a similar expression as before. It suffices to replace the risk-free payoff 1 by the risky one, namely $\mathbb{1}_{\{\tau>t\}}$, where the random variable τ represents the default time of the issuer and $\mathbb{1}_A$ is the indicator function defined as 1 if A is true and zero otherwise. Mathematically,

$$P_s^{model}(t) = \mathbb{E}\left[\mathbb{1}_{\{\tau>t\}} D_s(t) \middle| \mathcal{G}_s\right] =: \bar{P}_s^r(t) \ .$$

In such a context, $\bar{P}_s^r$ corresponds to a risky discounting, where the term *risk* is referring to the possibility for the issuer not to meet her financial obligations.

To proceed, we need to model the default event. To that end, we consider a *reduced-form* (a.k.a. *intensity*) default model. We refer the reader to Duffie and Singleton (1999) and Lando (2004) for an extensive exposition of this class of models. In this framework, the default time $\tau := \tau(\lambda)$ is defined as the passage time of the process $\Lambda := (\Lambda_t, t \in [0,T])$ defined as $\Lambda_t := \int_0^t \lambda_s ds$ above a unit-mean exponential random variable $\mathcal{E}$ independent from every other processes. The process λ is an *intensity*, i.e. it is positive, so that Λ is almost surely increasing. In this model, the default event $\{\tau \leq t\}$ is modeled as $\{\Lambda_t \geq \mathcal{E}\}$ and the survival probability is given by

$$\mathbb{Q}(\tau > t) = \mathbb{Q}(\Lambda_t < \mathcal{E}) = \mathbb{Q}\left(U < e^{-\Lambda_t}\right) = \mathbb{E}\left[e^{-\Lambda_t}\right] \ ,$$

where $U := e^{-\mathcal{E}}$ is a random variable uniformly distributed on $[0,1]$.

The function $\bar{P}_s^r(t)$ can be proven to be a time-s discount curve in many cases. To show this, we first define a sub-filtration $\mathbb{F}$ such that all processes are $\mathbb{F}$-adapted except those featuring τ (i.e. those featuring $\mathcal{E}$ or U, which are independent from $\mathcal{F}_T$). We then define a second filtration $\mathbb{H} = (\mathcal{H}_t, t \in [0,T])$, the filtration generated by the default process $\mathcal{H}_t = \sigma(\mathbb{1}_{\{\tau<u\}}, u < t)$. Eventually, the total filtration $\mathbb{G}$ is recovered by progressively enlarging $\mathbb{F}$ with $\mathbb{H}$: $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$. Hence, τ is a $\mathbb{G}$-stopping time, but not an $\mathbb{F}$-stopping time. In such a case, one can replace $\mathcal{G}_s$ by $\mathcal{F}_s$ in the expression providing the time-s price of the risk-free ZCB:

$$P_s^r(t) = \mathbb{E}\left[D_s(t) | \mathcal{G}_s\right] = \mathbb{E}\left[D_s(t) | \mathcal{F}_s \vee \mathcal{H}_s\right] = \mathbb{E}\left[D_s(t) | \mathcal{F}_s\right] \ .$$

A central result in stochastic calculus is the so-called *Key lemma*. This fundamental theorem allows one to write the $\mathcal{G}_s$-conditional expectation of $X\mathbb{1}_{\{\tau>t\}}$ as the $\mathcal{F}_s$-conditional expectation of $Xe^{-\int_s^t \lambda_u du}$, rescaled by $\mathbb{1}_{\{\tau>s\}}$, for every integrable and $\mathcal{F}_t$-measurable random variable X. It is originally due to Dellacherie and Meyer Dellacherie and Meyer (1980), although its use

in financial applications have been put forward by Bielecki, Jeanblanc and Rutkowski Bielecki and Rutkowski (2002) (see, e.g., Bielecki et al. (2011) for numerous examples in credit risk and Brigo and Vrins (2018) for a specific application in counterparty credit risk). Applying the Key lemma to the risky ZCB formula above yields, with $X \leftarrow D_s(t)$,

$$\bar{P}_s^r(t) = \mathbb{1}_{\{\tau>s\}} \mathbb{E}\left[e^{-\int_s^t \lambda_u du} D_s(t) \middle| \mathcal{F}_s\right] = \mathbb{1}_{\{\tau>s\}} \mathbb{E}\left[e^{-\int_s^t (\lambda_u + r_u) du} \middle| \mathcal{F}_s\right] =: \mathbb{1}_{\{\tau>s\}} P_s^{\lambda+r}(t) \ .$$

Eventually, in the special case where $r \equiv 0$, $D \equiv 1$ so that $\bar{P}_s^r(t)$ collapses to

$$\bar{P}_s^r(t) = \mathbb{E}\left[\mathbb{1}_{\{\tau>t\}} | \mathcal{G}_s\right] = \mathbb{Q}\left(\tau > t | \mathcal{G}_s\right) = \mathbb{1}_{\{\tau>s\}} P_s^{\lambda}(t) \ .$$

Hence, on the event $\{\tau > s\}$, $\bar{P}_s^r(t) = P_s^{\lambda}(t)$ agrees with the survival probability function associated with τ, conditional upon $\mathcal{G}_s$.

In this specific context, we are interested in a model x to depict the dynamics of the intensity process λ that would be tractable enough, and provide a perfect fit to any valid survival probability curve extracted from the prices of defaultable instruments like corporate bonds or credit default swaps (CDS).

Remark 1. *Notice that in contrast to rates, that can – and some of them currently do – take negative value, non-negativity is a formal requirement when x represents an intensity process λ. A default model featuring "negative intensities" is theoretically flawed, and is problematic. Indeed, modelling the event $\{\tau > t\}$ as $\{\Lambda_t < \mathcal{E}\}$ yields a survival indicator process $\mathbb{1}_{\{\tau>t\}}$ that might jump both up and down, i.e. the reference entity could be "brought back to life". One could of course think of replacing the default event using a first passage time, thereby revisiting the default time definition as $\tau := \inf\{t \geq 0 : \Lambda_t \geq \mathcal{E}\}$. However, one looses the analytical tractability for the survival probability since in this case, $\mathbb{Q}(\tau > t)$ does no longer agree with $\mathbb{Q}(\Lambda_t < \mathcal{E}) = \mathbb{E}\left[e^{-\int_0^t \lambda_u du}\right] = P_0^{\lambda}(t)$. The $x \geq 0$ constraint is also a natural requirement when it represents a credit spread.*

2.2 The perfect fit problems

Equation (1) suggests that the calibration problem consists of finding the parameters of a given model to minimize the discrepancies between market and model curves up to a time horizon T. However, it is clear that the choice of the model class will also have a substantial impact. Indeed, depending on the model chosen, the minimum of the error function could be large, small or even zero, in which case the perfect fit is obtained.

Inspired by the financial problems mentioned in Section 2.1, we consider the following problems directly related to the perfect fit constraint (up to a given time horizon T, that is implicit in the sequel). The first one does not impose any constraint on the process x to consider.

Problem 1. *Find a tractable process x satisfying*

$$P_s^x(t) := \mathbb{E}\left[e^{-\int_s^t x_u du}\middle| \mathcal{F}_s\right] = P_s(t)$$

for $t \in [0,T]$ and every given discount curve P_s.

Depending on the application at hand, one may need to impose additional constraints on x. As suggested by the risky discounting example, non-negativity is a crucial one. This leads us to consider a second (constrained) problem.

Problem 2. *Find a tractable positive process x (i.e. such that $\mathbb{Q}(x_t \geq 0) = 1$ and $\mathbb{Q}(x_t > 0) > 0$ for all $t \in [0,T]$) satisfying*

$$P_s^x(t) := \mathbb{E}\left[e^{-\int_s^t x_u du}\middle| \mathcal{F}_s\right] = P_s(t)$$

for $t \in [0,T]$ and every strictly decreasing discount curve P_s.

In either problems, *tractability* refers to the fact that model calibration (1) – that features an optimization over the parameter space – is not too cumbersome, computationally. Solving this optimization problem typically requires many iterations, hence numerous evaluations of the objective function. This suggests that a highly desirable feature of the model is to admit a closed form expression for P^{model} or, at least, that the latter can be computed without having to rely on time-consuming numerical methods like, e.g., Monte Carlo simulations.

3 Shifted homogeneous affine models

In order to solve these two problems, we consider what is probably the most tractable family of models, namely *affine processes* and, more specifically *time-homogeneous affine processes.* Indeed, for a one-factor affine model $y := (y_t, t \in [0,T])$, many expressions are available analytically, as well as for its integrated version $Y := (Y_t, t \in [0,T])$, $Y_t := \int_0^t y_u du$. In particular, $P_s^y(t) := \mathbb{E}\left[e^{-\int_s^t y_u du}\middle| \mathcal{F}_s\right]$ is merely the conditional moment generating function of $Y_t - Y_s$, $t \geq s$.

3.1 Affine processes and affine jump-diffusions

As recalled in the introduction, ATSM models are widely used in finance because they offer an appealing modeling framework : they are scarce, and empirical evidences suggest that they depict relatively well the market dynamics. Affine models are characterized as follows Filipovic (2005).

Definition 2 (Affine process)**.** *An affine process is any process y satisfying*

$$P_s^y(t) := \mathbb{E}\left[e^{-\int_s^t y_u du}\middle| \mathcal{F}_s\right] = e^{A_s^y(t;\Xi)-B_s^y(t;\Xi)y_s} =: P_s^y(t;\Xi) \tag{3}$$

where Ξ is the (set of) parameter(s) governing y and A_s^y, B_s^y are differentiable functions satisfying $A_s^y(s;\Xi) = B_s^y(s;\Xi) = 0$.

Provided that the A, B functions are known, the analytical form (3) facilitates in a tremendous way calibration procedures such as (1) when the considered P^{model} function takes the form of the conditional expectation in (3), as illustrated on the risk-free and risky discounting applications. This explains why such models are so popular in term-structure modeling.

For such processes, the function P_s^y is thus well-defined for every s, is positive, and satisfies $P_s^y(s) = 1$. It is therefore a time-s discount curve in the sense of Definition 1 since it is obviously differentiable on (s, ∞). For instance, P_s^r and P_s^λ in the above two examples are time-s discount curves whenever r and λ are affine processes, respectively.

It is known (see, e.g., Brigo and Mercurio (2006) and Duffie and Kan (1996)) that every diffusion with affine drift and diffusion coefficients, regular enough so that a solution exists, is an affine process. Similarly, every jump-diffusion with such types of drift and variance coefficients and independent compounded Poisson jumps (i.e. exponentially-distributed jumps arriving according to a Poisson process) is also affine.

Definition 3 (Affine jump-diffusions, AJD)**.** *A stochastic process y is called an affine jump-diffusion if its dynamics take the form*

$$dy_t = (a(t) + b(t)y_t)dt + \sqrt{c(t) + d(t)y_t}dW_t + dJ_t \tag{4}$$

with W an $\mathbb{F}$-Brownian motion and J a compound Poisson process independent from W, defined according to $J_t := \sum_{j=1}^{N_t} \zeta_i$ where N is a Poisson process with instantaneous jump rate $\omega(t) \geq 0$ and ζ_i's are i.i.d. exponentially distributed random variables with mean $\alpha \geq 0$. In the special case where the parameters $(a, b, c, d, \alpha, \omega)$ are constant, y is said time-homogeneous, or simply homogeneous, or HAJD.

As explained above, affine models are specifically relevant in our context when A, B are known in closed form. This is the case for HAJD. Three important homogeneous cases are the Ornstein-Uhlenbeck, the square-root diffusion and the square-root jump-diffusion. The first one, widely known as the Vasicek (VAS) model, corresponds to the special case where $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\beta, -\kappa, \eta^2, 0, 0, 0)$ and $y_0 \in \mathbb{R}$. The second model is the Cox-Ingersoll-Ross with (CIR), and is associated to $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\beta, -\kappa, 0, \delta^2, 0, 0)$, with $y_0, \beta > 0$. Eventually, the JCIR is an extension of the CIR, associated with parameters $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\beta, -\kappa, 0, \delta^2, \alpha, \omega)$. The speed of mean-reversion κ is assumed to be positive in all models. When the initial value y_0 is part of the parameters, we note the parameter set Ξ_0. In contrast to VAS which is a Gaussian model, the CIR and JCIR models are non-negative. We recall (and derive) some properties of these processes in the Appendix (Section 7.1) for further references.

Observe that the sum of two affine processes x, y is, generally speaking, not an affine process. Hence, it is not clear whether the risky discounting curve $P_s^{\lambda+r}$ is a time-s discount curve, even in the simple case where both r, λ are affine processes. Some special cases are discussed in the Appendix (Section 7.2). In the sequel, we consider a specific pricing time, say $s = 0$ without loss of generality, and drop the observation time subscript for conciseness.

HAJD models like VAS, CIR and JCIR seem appropriate to solve problems 1 and 2. Unfortunately, they do not allow for a perfect fit to a given discount curve P, except in very special cases. Indeed, it is not possible in general, for such type of processes x, to find Ξ (or Ξ_0) such that $P^{model} := P^x(\cdot; \Xi) \equiv P$, even up to a finite horizon T.

3.2 A deterministic shift extension

The starting point is to notice that the limited capacities of homogeneous models result from their rigid parametric form. Therefore, an interesting route is to consider a family of models x defined as time-dependent transform of a base HAJD model y in such a way that the model's tractability is not affected. In this section, we recall the general deterministic shift extension approach. The latter has been introduced in the seminal paper Brigo and Merccurio (2001) in order, precisely, to address calibration issues such as Problem 1. In this model, $x := x^\varphi$ is defined as a HAJD (y) that is shifted in a time-dependent way using a deterministic function φ:

$$x_t^\varphi := y_t + \varphi(t) \ . \tag{5}$$

Interestingly, $P^{model}(t) := P^{x^\varphi}(t;\Xi)$ where x^φ remains affine (although no longer homogeneous) and is hence analytically tractable in terms of calibration since :

$$P^{x^\varphi}(t;\Xi) = e^{A^{x^\varphi}(t;\Xi) - B^{x^\varphi}(t;\Xi)x_0}$$

with

$$\begin{aligned} A^{x^\varphi}(t;\Xi) &= A^y(t;\Xi) - \int_0^t \varphi(u)du + B^y(t;\Xi)\varphi(0) \ , \\ B^{x^\varphi}(t;\Xi) &= B^y(t;\Xi) \ . \end{aligned}$$

Clearly, the dynamics of x^φ are easily obtained from that of y. Indeed, assuming

$$dy_t = \mu(t, y_t)dt + \sigma(t, y_t)dW_t + dJ_t \ , \tag{6}$$

the dynamics of x^φ read, when φ is differentiable, as

$$dx_t^\varphi = dy_t + \varphi'(t)dt = (\mu(t, x_t^\varphi - \varphi(t)) + \varphi'(t))dt + \sigma(t, x_t^\varphi - \varphi(t))dW_t + dJ_t \ , \quad x_0^\varphi = y_0 + \varphi(0) \ .$$

It can be shown that in the particular case where y is a HAJD, then x^φ remains an AJD, even though no longer homogeneous, unless $\varphi(t)$ is constant. For instance, if the dynamics of y obey (4), then x^φ is governed by the same type of dynamics since

$$dx_t^\varphi = (a^\varphi(t) + b(t)x_t^\varphi)\, dt + \sqrt{c^\varphi(t) + d(t)x_t^\varphi}dW_t + dJ_t \ . \tag{7}$$

where $a^\varphi(t) := a(t) + \varphi'(t) - b(t)\varphi(t)$ and $c^\varphi(t) := c(t) - d(t)\varphi(t)$. As already noticed in Brigo and Merccurio (2001), whatever the base model y, the parameter Ξ and the discount curve P^{market}, there always exists a shift function $\varphi(t) = \varphi^\star(t;\Xi)$ that provides a perfect fit between the x^φ-model and the market. This is summarized in the next lemma.

Remark 2. *The shift approach may look suspicious: adding a deterministic function to a stochastic process is arguably a somewhat artificial way to fix the model's limitations in terms of calibration. However, as clear from (7), shifting the model in a deterministic way actually amounts to consider an inhomogeneous model, i.e. with time-dependent long-term mean. For instance, the Vasicek model* $(a(t), b(t), c(t), d(t)) = (0, -\kappa, \eta, 0)$ *shifted with* $\varphi(t) \leftarrow \int_0^t \beta(s)e^{-\kappa(t-s)}ds$ *yields a HAJD with* $(a(t), b(t), c(t), d(t)) = (\kappa\beta(t), -\kappa, \eta, 0)$*, which is known as the Hull-White (HW) model Hull and White (1990). Moreover, the later is itself a particular case of the Heath-Jarrow-Morton (HJM) model Heath et al. (1992) which consists of modeling the entire instantaneous forward curve* $f_s(t)$ *with* $df_s(t) = \mu(s,t)dt + \eta e^{-\kappa(t-s)}dW_s$ *where the drift* $\mu(s,t)$ *is given by no-arbitrage, and provided that the initial discount curve and the long-term mean*

obey the relationship $\beta(t) = \frac{d}{dt}f_0(t) + \kappa f_0(t) + \frac{\eta^2}{2\kappa}(1 - e^{-2\kappa t})$. *Therefore, any instantaneous forward curve* f^{market} *(hence discount curve* P^{market}*) can be fitted with either models provided that one takes* $f_0(t) \leftarrow f^{market}(t)$ *as initial curve (HJM), the corresponding long-term mean* $\beta(t)$ *(HW), or the associated shift* $\varphi(t)$ *(shifted Vasicek). These models became very popular among practitioners, essentially because of their ability to replicate market curves, i.e. to solve Problem 1.*

Lemma 2. *The* x*-model defined according to* (5) *where* y *is a HAJD solves Problem 1 provided that*

$$\varphi(t) \leftarrow \varphi^\star(t;\Xi) := \frac{d}{dt}\ln\frac{P^y(t;\Xi)}{P^{market}(t)} = f^{market}(t) - f^y(t;\Xi)\ . \tag{8}$$

where f^{market} *and* f^y *are the instantaneous forward rate functions associated with* P^{market} *and* P^y*, respectively.*

Proof. Indeed, because y is a HAJD, P^y is a discount curve and from Lemma 1, it admits a representation in terms of forward rates f^y. By assumption, same holds true for P^{market}. Eventually,

$$P^{x^\varphi}(t;\Xi) = \mathbb{E}\left[e^{-\int_0^t x_u du}\right] = e^{-\int_0^t \varphi^\star(u;\Xi)du}\,\mathbb{E}\left[e^{-\int_0^t y_u du}\right] = e^{-\int_0^t f^{market}(u)du} = P^{market}(t)\ .$$

The model is tractable since $f^y(t;\Xi) = -\frac{d}{dt}\ln P^y(t;\Xi)$ can be computed in closed form. □

It is worth noting that, for a given model y, the perfect fit can be attained for every parameters Ξ. This suggests that the calibration problem (1) is ill-posed. Indeed, the choice of Ξ is completely arbitrary since the error between P^{x^φ} and P^{market} can be set to zero for any Ξ, provided that one chooses $\varphi(t) \leftarrow \varphi^\star(t;\Xi)$. In particular, one could take the null process for y and $\varphi(t) = f^{market}(t)$. This trivial choice rends x^φ deterministic, which is most likely not the desired result. A common practice to circumvent this indeterminacy is thus either (i) to extend the set of calibration instruments, incorporating products that are sensitive to volatility (like interest-rate or credit options in the above asset classes), or (ii) to require the y-model to fit the market "as best as possible" (to get $\Xi^\star$) and then take $\varphi(t;\Xi^\star)$ as shift function:

$$P^{model}(t) := P^{x^\varphi}(t;\Xi^\star) \text{ where } \Xi^\star := \arg\min_{\Xi}\|P^y(\cdot;\Xi) - P^{market}(\cdot)\|\ ,\ \ \varphi(t) \leftarrow \varphi^\star(t) := \varphi^\star(t;\Xi^\star)\ . \tag{9}$$

This approach is particularly relevant when no or little "volatility-sensitive" instruments are quoted on the market. The role of the shift is thus merely to compensate the remaining discrepancies between the market curve P^{market} and the one generated by the "best" parametric

model y, $P^y(\cdot;\Xi^\star)$. Adding a shift to the VAS, CIR or JCIR models yield the Hull-White, CIR++ or JCIR++, respectively Brigo and Mercurio (2006).[1]

Remark 3. *On the top of the appealing affine structure, the shifted model is highly tractable because many statistical properties of the process are available in closed form. Indeed, as recalled in Section 7.1, the k-th moment* $m^y(k,t) := \mathbb{E}[y_t^k]$ *and the moment generating function (MGF)* $\psi^y(u,t) := \mathbb{E}[e^{uy_t}]$ *of a time-homogeneous affine model* y *are known analytically, as well as those of their time-integrals* $Y_t := \int_0^t y_u du$, $m^Y(k,t)$ *and* $\psi^Y(u,t)$. *Due to the simple shift structure, the corresponding expressions for* x^φ, *the shifted model, are readily available. For instance, the k-th moment of* x_t^φ *and* X_t^φ *are given by Newton's binomial formula applied to* $(y_t+\varphi(t))^k$ *and the MGFs simply collapse to* $\psi^{x^\varphi}(u,t) = e^{u\varphi(t)}\psi^y(u,t)$ *and* $\psi^{X^\varphi}(u,t) = e^{u\int_0^t \varphi(s)ds}\psi^Y(u,t)$.

3.3 Dealing with the positivity constraint

As discussed above, the deterministic shift extension nicely solves Problem 1. In order to solve Problem 2 however, one first considers a non-negative base process y. Yet, there is no reason that the shifted process x^φ would remain non-negative. For instance, taking CIR dynamics for y, x^φ is non-negative on $[s,t]$ if and only if $\min_{u\in[s,t]}\varphi(u) \geq 0$. From (8), the shift function depends both on the y model (and its parameters Ξ) and on the market curve.

Remark 4. *Observe that the optimization problem (9) is contradictory with non-negative shift functions. Indeed, by construction of* $\Xi^\star$, $P^y(\cdot;\Xi^\star)$ *passes through* P^{market}. *Consequently, the shift* $\varphi(t) \leftarrow \varphi^\star(t;\Xi^\star)$ *will lead to a perfect fit, but will correct for both negative and positive errors. In other words,* φ *will change of sign. Therefore, this strategy does not provide a valid solution to Problem 2. This will be illustrated on a real example in Section 5.1.*

In order to satisfy the non-negativity constraint mentioned in Problem 1, one needs to force the non-negativity constraint on the shift at the optimal parameters. The shift function under positivity constraint is referred to with the notation $\varphi^{\star,+}(t)$ to stress the difference with the unconstrained counterpart, $\varphi^\star(t)$.

Lemma 3. *Let* y *be a HAJD that is non-negative on* $[0,T]$ *with parameters* $\Xi^\star$ *given by*

$$\Xi^{\star,+} := \arg\min_{\Xi} \|P^y(\cdot;\Xi) - P^{market}(\cdot)\| \quad \textit{subject to} \quad f^y(t;\Xi) \leq f^{market}(t), \quad \forall\, 0 \leq t \leq T\,. \tag{10}$$

Then, the x^φ*-model (5) with* $\varphi(t) \leftarrow \varphi^{\star,+}(t) := \varphi^\star(t;\Xi^{\star,+})$ *solves Problem 2.*

[1]Notice that the Hull-White model is a Vasicek model where the long-term mean parameter is replaced by a deterministic function of time.

Proof. The condition on the instantaneous forward rates ensures that shift function $\varphi^{\star,+}$ will be non-negative on $[0,T]$; this is obvious from (8). Hence, since y is assumed to be non-negative, so is the shifted process x^{φ}. Moreover, taking $\varphi(s) \leftarrow \varphi^{\star}(s;\Xi)$ yields a perfect fit for every Ξ, by construction, including $\Xi = \Xi^{\star,+}$. □

Notice that there always exists a set of parameters Ξ such that the constraint is met. Indeed, all parameters Ξ associated to the deterministic case $y \equiv 0$ yield $f^{y}(\cdot;\Xi) \equiv 0$. Clearly, the constraint is met since $f^{market}(t)$ is strictly positive given that P^{market} is strictly decreasing, by assumption. The shift is simply given by the market forward rate $\varphi(t) \leftarrow \varphi^{\star,+}(t) = f^{market}(t)$. However, the trivial process parameter is likely not to be satosfactory.

In order to deal with Problem 2, we need to consider a non-negative base model y. Given that we focus on HAJDs, we consider the CIR and JCIR models. To make the distinction between the two shifted models, we call S-(J)CIR the (J)CIR process shifted with $\varphi(t) \leftarrow \varphi^{\star}(t) = \varphi^{\star}(t;\Xi^{\star})$ (i.e. without positivity constraint, and parameter $\Xi^{\star}$ given by (9)) and PS-(J)CIR the (J)CIR process shifted with $\varphi(t) \leftarrow \varphi^{\star,+}(t) = \varphi^{\star}(t;\Xi^{\star,+})$ (i.e. under positivity constraint, and parameter $\Xi^{\star}$ given by (10)). Although the PS-(J)CIR allows both for a perfect fit and the non-negativity constraint, one may argue that it is not as tractable as the (J)CIR. Indeed, the optimization problem (10) is more difficult than (9) due to the constraint on the instantaneous forwards, even if some sufficient conditions on the parameters can be found. Second, and probably more importantly, this constraint is binding, in the sense that it often deeply impacts the optimal parameter $\Xi^{\star,+}$. Even if it is unlikely that the optimal solution corresponds to the deterministic case, it often yields dynamics associated to rates that feature "little randomness". This will be illustrated in Section 5, first by comparing the variance of the integrated S-CIR and PS-CIR processes, as well as the impact when dealing with financial applications. These two points are discussed in (Brigo and Mercurio, 2006, sec. 3.9.3, p.107-109). To circumvent this issue in an interest rate framework, the authors suggest to relax the strict positivity constraint. By working in a setup where positivity is expected but not guaranteed, they obtain a process that yields much more realistic results in terms of implied volatility levels. This is perfectly fine in such a context as positivity of rates might be desirable (in some cases), but zero is by no means a strict lower bound (neither theoretically nor practically). Yet, this is more problematic when it comes to model such things as default intensities, because this kind of applications requires both strict positivity and, typically, large volatility. Increasing the variance of the CIR++

process without breaking Feller's constraint[2] can be achieved by incorporating compounded Poisson jumps (JCIR++) but, unfortunately, increasing the jump activity while maintaining the calibration to a given market curve f^{market} is difficult under the positivity constraint. Indeed, the minimum of the implied shift function is driven down when increasing the jump activity because the difference $f^{\text{JCIR}}(t) - f^{\text{CIR}}(t)$ is non-negative and increases with ω, α for $\alpha, \omega > 0$ (see Appendix, Section 7.1.3). This observation combined with (8) leads to a lower shift function φ for the JCIR++ than for the corresponding CIR++. For this reason, there is a need for an alternative to CIR++ and JCIR++ that would combine (i) tractability, (ii) the prefect fit feature, (iii) the large implied volatility and (iv) positivity.

4 The deterministic time-changed extension

In order to circumvent the drawbacks of the deterministic shift extension with regards to Problem 2, we propose a different approach. In the same spirit as the shift, we aim at finding a model x by adjusting a time-homogeneous affine model y, that would benefit from a set of desirable properties.

4.1 Model setting

The x-model is obtained by time-changing a HAJD y using a specific (but deterministic) *clock* Θ that may differ from the calendar clock. A clock is a time change function that can differ from identity, but having specific properties.

Definition 4 (Clock)**.** *A clock is an application*

$$\Theta : \mathbb{R}^+ \to \mathbb{R}^+, \; t \mapsto \Theta(t)$$

that is a grounded, increasing and differentiable. In other words, a clock is any function Θ *of the form*

$$\Theta(t) := \int_0^t \theta(u)du \quad \textit{where} \quad \theta(u) > 0 \; , \quad \forall \, u \geq 0 \; .$$

Clearly, $\Theta(t) = t$ is the calendar clock, and any function of the form $\Theta(t) = kt$, $k > 0$, is again a clock, corresponding to a constant rescaling of the calendar time.

[2]Increasing the volatility of the CIR++ process by increasing the diffusion paramter δ just breaks the Feller's condition ($2\kappa\beta \geq \delta^2$) and leads to an intensity process that almost surely equals to zero at a given time interval.

Similarly to (5), we define our model as $x = x^\theta$, obtained from the following transform of the base process y:

$$x_t^\theta := \theta(t) y_{\Theta(t)} \ . \tag{11}$$

The dynamics of x^θ are given by Ito's product rule. Defining the process $y^\theta := \left(y_{\Theta(t)}, t \in [0,T]\right)$, one gets

$$dx_t^\theta = y_t^\theta d\theta(t) + \theta(t) dy_t^\theta \ , \tag{12}$$

where the dynamics of y^θ are given in the below lemma.

Lemma 4. *Let Θ be a clock and consider a base model y with dynamics (6). Then, the dynamics of y^θ take the form*

$$dy_t^\theta = \mu\left(\Theta(t), y_t^\theta\right)\theta(t)dt + \sigma\left(\Theta(t), y_t^\theta\right)\sqrt{\theta(t)}dB_t + dJ_t^\theta, \quad y_0^\theta = y_0 \tag{13}$$

where B is an $\mathbb{F}^\theta$-Brownian motion, $\mathbb{F}^\theta := (\mathcal{F}_{\Theta(t)}, t \in [0,T])$ and J^θ an inhomogeneous compounded Poisson process with jump size mean α and time-t intensity $\omega\theta(t)$.

Proof. By definition, we have

$$y_t^\theta := y_{\Theta(t)} = y_0 + \int_0^{\Theta(t)} \mu(u, y_u)du + \int_0^{\Theta(t)} \sigma(u, y_u)dW_u + \int_0^{\Theta(t)} dJ_u \ .$$

Hence,

$$\int_0^{\Theta(t)} \mu(u, y_u)du = \int_0^t \mu\left(\Theta(u), y_{\Theta(u)}\right)\theta(u)du = \int_0^t \mu\left(\Theta(u), y_u^\theta\right)\theta(u)du \ ,$$

and

$$\int_0^{\Theta(t)} \sigma(u, y_u)dW_u = \int_0^t \sigma\left(\Theta(u), y_{\Theta(u)}\right)dW_{\Theta(u)} = \int_0^t \sigma\left(\Theta(u), y_u^\theta\right)\sqrt{\theta(u)}dB_u \ .$$

Indeed, Θ is a clock, hence $\theta > 0$ and the process $W^\theta := (W_{\Theta(t)}, t \in [0,T])$ is a local martingale with quadratic variation $\langle W^\theta, W^\theta\rangle_t = \Theta(t)$. From Jeanblanc et al. (2009), the process $B := (B_t, t \in [0,T])$ defined as

$$B_t := \int_0^t \frac{1}{\sqrt{\theta(u)}} dW_{\Theta(u)} \tag{14}$$

is then a Brownian motion. Differentiating y_t^θ leads to (13). With regards to the compounded Poisson process, notice that $dJ_t = \zeta_{N_t} dN_t$ and $dJ_t^\theta = dJ_{\Theta(t)} = \zeta_{N_\Theta(t)} dN_{\Theta(t)}$. The process N^θ defined as $N_t^\theta := N_{\Theta(t)}$ is a Poisson process with instantaneous intensity $\omega\theta(t)$. Hence, the dynamics of J^θ are given by $J_0^\theta = 0$ and $\zeta_{N_t^\theta} dN_t^\theta$, so that J^θ is a compounded Poisson process with jump size mean α and time-t instantaneous rate of jumps arrival, $\omega\theta(t)$. □

This model looks appealing for several reasons. First, just as the shift extension, it is a deterministic adjustment of a base model and is hence expected to be tractable when the latter is, say, a HAJD. Second, because x_t^θ is a positive rescaling of the process y sampled at time $\Theta(t)$, the range of x^θ is linked to that of y. In particular, if the range of y is $\mathbb{R}$, as for the Vasicek model, then so is the range of x^θ. However, if y is non-negative as in the (J)CIR case, then so is x^θ. Hence, this solves the drawback of the shift approach related to Problem 2. Eventually, the time-dependent feature of the clock rate θ is expected to provide additional flexibility in the calibration properties of x^θ with respect to that of the homogeneous model y. Two questions remain open in this respect. First, we need to clarify the circumstances under which the model provides a perfect fit. Second, in the case where the perfect fit can be achieved, we need to provide an efficient procedure to compute the resulting "optimal clock", $\Theta^\star$. The price to pay is that, in contrast with the shift extension, the time-changed model is not fully flexible. Indeed, starting with a given model y, the x^θ model can only generate specific shapes for discount curves. We are thus more dependent on the initial choice of the base model y. Fortunately, it turns out that a perfect fit is achievable for a wide set of market curves, including all decreasing discount curves, considered in Problem 2. This is clearly the most important case since (i) it corresponds to the case where the shift approach fails to provide a convincing solution and (ii) it is probably the most common case in practice, since it encompasses the class of discount curves with non-negative rates (or, more generally, with non-negative instantaneous forward rates), as well as the set of all continuous survival probability curves. Moreover, even if the mathematical expression of the clock $\Theta^\star$ is not available in closed form, its numerical computation turns out to be easy. This leads us to the first fundamental result of the paper.[3]

Theorem 1. *Let P^{market} be a discount curve and y a model such that P^y is a discount curve. Define the x^θ-model as in (11). Then, $P^{x^\theta} \equiv P^{market}$ provided that $\Theta \leftarrow \Theta^\star$ where $\Theta^\star$ satisfies the first-order ODE*

$$\theta^\star(t) := \frac{d}{dt}\Theta^\star(t) = \frac{f^{market}(t)}{f^y(\Theta^\star(t))}\,, \tag{15}$$

with f^{market}, f^y the corresponding instantaneous forward curves. Moreover, if P^{market} and P^y are strictly decreasing, the solution to (15) exists, is a clock, and is given by

$$\Theta^\star(t) := Q^y\left(P^{market}(t)\right)\,, \tag{16}$$

where Q^y is the inverse of the base-model discount curve, P^y.

[3]When no confusion is possible, the explicit reference to the model parameters Ξ is avoided to ease the notations.

Proof. See Section 7.3. □

Observe that the optimal clock $\Theta^\star$ actually depends from the y-model parameters Ξ. Just like for the shift, we actually have $\Theta^\star(t) = \Theta^\star(t;\Xi)$. Although other frameworks are possible, we set $\Xi = \Xi^\star$ as in (9). Similar to the function φ in the shift approach, the purpose of the clock Θ is then to absorb the remaining errors between $P^y(\cdot;\Xi^\star)$ and P^{market}.

4.2 Time-changed homogeneous affine diffusions

As in the shift extension, a time-changed model x^θ enjoys a similar tractability level to that of the base model y. Indeed the k-th moment is $m^{x^\theta}(k,t) = \theta(t)^k m^y(k,\Theta(t))$ and moment generating function is $\psi^{x^\theta}(u,t) = e^{u\theta(t)}\psi^y(u,\Theta(t))$, whereas those of X_t^θ coincide with those of $Y_{\Theta(t)}$. Hence, a tractable model x^θ can be obtained by considering HAJD processes as base model y. We illustrate our method by analyzing two calibration problems that can be solved by considering the Vasicek and the JCIR processes.

It is clear from Lemma 4 that in the particular case where y is a HAJD, then x^θ is a scaled version of an inhomogeneous affine jump diffusion (AJD), unless $\theta(t)$ is a positive constant, in which case it remains a HAJD. To see this, suppose that the dynamics of y obey (4). From Lemma 4, y^θ is governed by

$$dy_t^\theta = \left(a(\Theta(t)) + b(\Theta(t))y_t^\theta\right)\theta(t)dt + \sqrt{\left(c(\Theta(t)) + d(\Theta(t))y_t^\theta\right)\theta(t)}dB_t + dJ_t^\theta\,. \tag{17}$$

Interestingly, y^θ is still an AJD. In the sequel, we focus on the special case where the base model y is a HAJD, i.e. takes the form (4) with constant parameters $(a(t),b(t),c(t),d(t),\alpha,\omega(t)) = (\kappa\beta,-\kappa,\eta^2,\delta^2,\alpha,\omega)$. To simplify the notation, we specify the model parameters using the vector $\Xi = (\kappa,\beta,\eta,\delta,\alpha,\omega)$.

4.2.1 Time-changed Vasicek

Our time change approach can be easily used to solve Problem 1 in the most common case where the forward curve f^{market} is arbitrary (monotonic, humped, etc) provided that it is positive. As there is no constraint on the range of the process x^θ, let us postulate Vasicek dynamics for the base process with parameters $\Xi = (\kappa,\beta,\eta,0,0,0)$:

$$dy_t = \kappa(\beta - y_t)dt + \eta dW_t, \quad y_0 \in \mathbb{R}\,.$$

The forward curve associated to this model is given by $f^y(t) = f^{\mathrm{VAS}}(t) := f_0^{\mathrm{VAS}}(t)$ in (24):

$$f^{\mathrm{VAS}}(t) = (1-\mathrm{e}^{-\kappa t})\frac{\kappa^2\beta - \eta^2/2}{\kappa^2} + \frac{\eta^2}{2\kappa^2}\,\mathrm{e}^{-\kappa t}(1-\mathrm{e}^{-\kappa t}) + y_0\,\mathrm{e}^{-\kappa t}\,. \tag{18}$$

It can thus be used to select an appropriate Vasicek model. The next corollary provides guidelines to generate decreasing discount curve, associated with the most common case of positive instantaneous forwards.

Corollary 1. *Let P^{market} be a strictly decreasing market curve. Then, for every Vasicek model with parameters satisfying $y_0 \geq 0$ and $2\kappa^2\beta > \eta^2$, there exists a clock $\Theta^\star$ such that $P^{x^\theta} \equiv P^{market}$.*

Proof. Because y is a Vasicek process, P^y is a discount curve. Moreover, the conditions $y_0 \geq 0$ and $2\kappa^2\beta > \eta^2$ guarantee that the forward curve (18) is strictly positive, hence P^y is strictly decreasing. From Theorem 1, the clock $\Theta^\star$ exists and is given by (15) with f^y given in (18). □

Notice that the dynamics of the time-changed Vasicek model x_t^θ are given by (12) with

$$dy_t^\theta = \kappa(\beta - y_t^\theta)\theta(t)dt + \eta\sqrt{\theta(t)}dB_t, \quad y_0^\theta = y_0 \ ,$$

showing that y^θ remains a Gaussian process. Fitting perfectly a strictly decreasing discount curve (without further constraints on the process) is a special case of Problem 1, that can also be solved using the shift approach (5) by taking $x \leftarrow x^\varphi$ where y is a Vasicek with arbitrary parameters Ξ and $\varphi(t) \leftarrow \varphi^\star(t;\Xi) = f^{market}(t) - f^{\text{VAS}}(t)$. The main interest of the time-changed approach is actually when considering Problem 2.

4.2.2 Time-changed (J)CIR

The following result is the second main contribution of the paper. It shows that the time change approach $x \leftarrow x^\theta$ provides a solution to Problem 2.

Corollary 2. *Let y be an almost-surely positive HAJD with parameters Ξ. Then, the model x^θ defined in (11) with $\Theta \leftarrow \Theta^\star(t;\Xi)$ solves Problem 2.*

Proof. Because y is a HAJD, P^y is a discount curve and is tractable analytically. Moreover, the latter is strictly decreasing since y is almost-surely positive. We conclude the proof by relying on Theorem 1. □

Let us now consider the JCIR model, i.e. the HAJD with $\Xi = (\kappa, \beta, 0, \delta, \alpha, \omega)$. The CIR is recovered as a special case by choosing (α, ω) such that $\alpha\omega = 0$.

Then,

$$dy_t = \kappa(\beta - y_t)dt + \delta\sqrt{y_t}dW_t + dJ_t, \quad y_0 > 0 \ , \tag{19}$$

where κ, β, δ are strictly positive constants and ω, α are non-negative. The optimal clock $\Theta^\star$ leading to the perfect fit to a given strictly decreasing curve P^{market} is given by (15) where the forward curve associated to this model is given by $f^y(t) = f^{\text{JCIR}}(t) := f_0^{\text{JCIR}}(t)$ in (28) :

$$f^{\text{JCIR}}(t) = \frac{2\kappa\beta(e^{t\gamma}-1)}{2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)} + y_0 \frac{4\gamma^2 e^{t\gamma}}{[2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)]^2} + \frac{2\omega\alpha(e^{t\gamma}-1)}{2\gamma+(\kappa+\gamma+2\alpha)(e^{t\gamma}-1)}, \tag{20}$$

where $\gamma := \sqrt{\kappa^2+2\delta^2}$. The dynamics of the time-changed process $x_t^\theta = \theta(t) y_t^\theta$ are given by (12) with

$$dy_t^\theta = \kappa(\beta - y_t^\theta)\theta(t)dt + \delta\sqrt{\theta(t)y_t^\theta}dB_t + dJ_t^\theta, \quad y_0^\theta = y_0 \, .$$

where B is an $\mathbb{F}^\theta$-Brownian motion and J^θ is an inhomogeneous compound Poisson process with jump size mean α and time-t instantaneous rate of arrival $\omega\theta(t)$.

The time change technique applied to a JCIR (TC-JCIR) therefore solves Problem 2. In particular, in contrast to the S-JCIR (that focuses on parameters such that φ is positive), the positivity constraint on x^θ is automatically satisfied for every (strictly decreasing) market curve and every Ξ (such that y is not trivially equal to 0). However, we have shown that it is possible to ensure positivity by considering the PS-JCIR, $x^{\varphi,+}$. Working with $\Xi^{\star,+}$ instead of $\Xi^\star$ can make the job, but at the expenses of having a process $x^{\varphi,+}$ that is, to a large extend, deterministic (i.e. $x_t^{\varphi,+}$ varies in a small neighborhood around $f^{market}(t)$). Consequently, TC-JCIR model are expected to feature a higher volatility compared to the corresponding PS-JCIR, at least up to some time horizon. This is summarized in the next theorem, which is the third main result of the paper.

Theorem 2. *Let P^{market} be a strictly decreasing discount curve and y be a JCIR++ process with parameter Ξ such that the perfect fit JCIR++ model $x_t^{\varphi^\star}$ is positive. Then, the ODE (15) with $f^y(t) = f^{\text{JCIR}}(t;\Xi)$ given by (20) admits a solution that satisfies $\Theta^\star(t) = \Theta^\star(t;\Xi) \geq t$. Moreover, the variance of the corresponding perfect fit TC-JCIR model $x_t^{\theta^\star}$ satisfies:*

1) $\mathbb{V}\left[X_t^{\theta^\star}\right] \geq \mathbb{V}\left[X_t^{\varphi^\star}\right]$, $\forall\, t \geq 0$,

2) $\mathbb{V}\left[x_t^{\theta^\star}\right] \geq \mathbb{V}\left[x_t^{\varphi^\star}\right]$ *if one of the following holds:*

- *i)* $y_0 = \beta + \frac{\omega\alpha}{\kappa}$,
- *ii)* f^{market} *constant and* $y_0 \leq \beta + \frac{\omega\alpha}{\kappa}$,
- *iii)* $y_0 > \beta + \frac{\omega\alpha}{\kappa}$ *and* $t < \Theta^{\star -1}(t_1)$,
- *iv)* $(\kappa\beta + \omega\alpha)/\gamma < y_0 < \beta + \frac{\omega\alpha}{\kappa}$ *and* $t > \Theta^{\star -1}(t_2)$

where

$$t_1 := \frac{1}{\kappa}\ln\left(1+\frac{y_0+2\omega\alpha^2/\gamma^2}{y_0-\beta-\omega\alpha/\kappa}\right) \quad \textit{and} \quad t_2 := \frac{1}{\gamma}\ln\frac{(\gamma-\kappa)(\kappa\beta+y_0\gamma+\omega\alpha)-2\omega\alpha^2}{(\kappa+\gamma)(y_0\gamma-\kappa\beta-\omega\alpha)-2\omega\alpha^2}.$$

Proof. See Section 7.4. □

To sum up, the TC-JCIR model (including the TC-CIR) provides an elegant solution to Problem 2: the process $x^{\theta^\star}$ is non-negative (in contrast with the S-JCIR $x^{\varphi^\star}$), is almost as tractable as the simple JCIR diffusion (in contrast with the PS-JCIR $x^{\varphi^{\star,+}}$), provides a perfect fit to every strictly decreasing discount curve (as both JCIR++ models) and features, to some extend, a larger variance (compared to the PS-JCIR $x^{\varphi^{\star,+}}$). In particular, it is observed, empirically, that the variance of the integral of the TC-JCIR remains similar to that of the unconstrained (i.e. flawed, but high-volatility) S-JCIR model $x^{\varphi^\star}$. Therefore, when a positivity constraint is required, the TC-JCIR avoids the drawbacks of the JCIR++ models. The only price to pay is that the clock is not available in closed form, but requires a (simple) numerical inversion. The properties of the model, namely the perfect fit and high-variance features, are illustrated in the next section on various applications taken from credit risk modeling.

5 Application to Credit Risk Modelling

We consider a reduced-form default model as in Section 2.1.2 by using a CIR base model y (i.e. (19) with $J \equiv 0$). The default intensity λ is modelled either as a CIR++ ($\lambda \leftarrow \lambda_t^\varphi := y_t + \varphi(t)$) or using the TC-CIR ($\lambda \leftarrow \lambda_t^\theta := \theta(t) y_{\Theta(t)}$). Observe that depending on the pair (P^{market}, Ξ), the CIR++ process can feature negative values. This will be the case when taking $\Xi \leftarrow \Xi^\star$ given using the MSE approach (9), unless there is an explicit constraint as in (10), leading to take $\Xi \leftarrow \Xi^{\star,+}$. Bear in mind that when λ represents an intensity process, the S-CIR model (λ^φ) is actually flawed as there is a non-zero probability to observe negative intensities, and $P^{\lambda^\varphi}(t)$ cannot be interpreted as a survival probability associated to a Cox model. Yet, we give the results of the model as a benchmark since, as explained in the introduction, it is a very standard approach.

We compare the CIR++ (S-CIR and PS-CIR) to the TC-CIR on several aspects related to a real case example where the reference entity is Ford Inc. We also discuss the TC-JCIR case when relevant. We first analyze the perfect fit feature of both types of models, as well as the non-negativity property of λ. We then compare the variance of the integrated processes Λ. We then analyse their behaviors in two different applications, namely the pricing of various credit

default swaptions (a.k.a. CDS options, or CDSO) with Ford as reference entity, or on the credit valuation adjustment (CVA) of prototypical FRA and IRS exposures where Ford is the trade counterparty.

It it well-admitted that "pure credit instruments" like CDS or CDSO are quite insensitive to the stochasticity of the interest rates in realistic conditions. This has been discussed explicitly for the CIR base model in Brigo and Alfonsi (2005) and Brigo and Cousot (2006). Hence, we consider a deterministic short rate process, which is stressed by the notation $r_u = r(u)$. In this case, one simply gets $P_s^r(t) = D_s(t) = e^{-\int_s^t r(u)du}$.[4]

In the sequel, we first illustrate the perfect fit feature of S-CIR, PS-CIR and TC-CIR when the default model is calibrated on the survival probability curve of Ford Inc. We then use the model to price CDSO and compute CVA figures.

5.1 Perfect fit of CDS term-structure

We consider the CDS term-structure of Ford Inc, and show that considering a set of parameter Ξ, there exist φ and Θ that yield a perfect fit. In the sequel, we drop the star superscript on the shift and clock functions. Hence, $\Xi^\star$ corresponds to the CIR parameter optimized without constraint to a given P^{market} curve, and φ and Θ refer to the corresponding optimal shift and clock functions. The corresponding parameters found under a non-negativity constraint are noted $\Xi^{\star+}, \varphi^+$ and Θ^+, respectively.

A credit default swap (CDS) is a financial instrument used by two parties – called the protection buyer and the protection seller – to transfer to the protection seller the financial loss that the protection buyer would suffer if a particular default event happened to a third party called the *reference entity*. Typically, we set τ as the default time of the latter. In a default swap contracted at time t, started at time T_a with maturity T_b, the protection buyer pays a coupon (of spread) k at a set of payment dates $T_a, \ldots, T_b$ as long as the reference entity does not default. The protection seller agrees to make a single payment LGD to the protection buyer if the default occurs between T_a and T_b. When applicable, the protection buyer makes a final payment corresponding to the spread accrued since the last payment date before default. For more details about the mechanics of this product, we refer to Brigo and Alfonsi (2005) and Brigo and El-Bachir (2010).[5]

[4]Given that the interest rates have little impact on the figure and that our main objective is to discuss the impact of the default model, we considered zero risk-free rate in the numerical applications below.

[5]For more details about the actual market conventions, we refer the interested reader to Markit (2004) and

The CDS term-structure consists of a set of par spreads associated with CDS of various maturities. The time-t par spread $s_t(T_i)$ of a CDS contract of maturity T_i is defined as the contract spread k that sets the value of the CDS contract to 0 at time t. The par spreads have been taken from Bloomberg on November 12, 2018 and are shown on the table below.

Maturity (years)	1	3	5	7	10
Spread (bps)	18.3	136.6	191.9	267.6	280.6

Table 1: CDS spread term structure of Ford Inc. on November 12, 2018. Source: Bloomberg.

In this context, the market curve P^{market} to be fitted is the risk-neutral survival probability curve, defined as $G(t) := \mathbb{Q}(\tau > t)$ associated with the default time τ of a given reference entity (here, Ford Inc.). It can be extracted from CDS quotes by inverting the no-arbitrage pricing formulae of the corresponding financial instruments. In practice, one only has a couple of calibration equations, say n, given by the number of market quotes (here, $n = 5$). It is therefore not possible to estimate the full (i.e. infinite-dimensional) market curve G without further assumptions. It is common market practice to consider the CDS model from the International Swap and Derivative Association (ISDA) – a.k.a the JP Morgan model – Markit (2004), that provides a slightly simplified version of the actual no-arbitrage pricing formula applying to CDSs. In this approach, the curve G is parametrized via a positive hazard rate function h, playing a similar role as the instantaneous forward rate f^{market},

$$G(t) := e^{-\int_0^t h(s)ds} ,$$

where h is itself parametrized by n constants $h_1, h_2, \ldots, h_n$ bootstrapped from the spreads $s_1, s_2, \ldots, s_n$ associated with the maturities $T_1, T_2, \ldots, T_n$. Let us focus on the horizon $T = T_n$. It is market practice to assume that h is piecewise constant between the maturities, i.e. to postulate the parametric form:

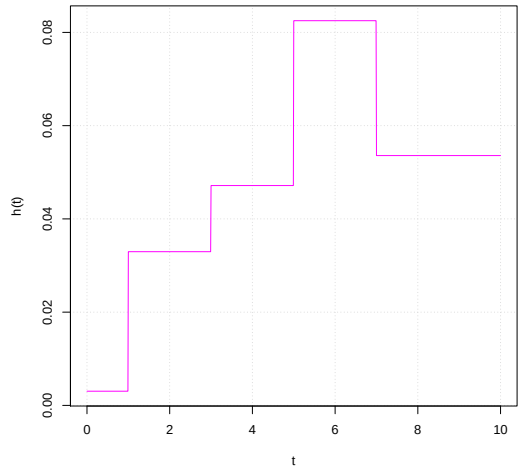
$$h(t) = \sum_{i=1}^{n} \mathbb{1}_{\{T_{i-1} \leq t < T_i\}} h_{i-1} ,$$

where $T_0 := 0$, $h_0 := \frac{s_1}{1-R}$ with $LGD := 1 - R$, $R = 40\%$ the assumed recovery rate of the firm and h_i's are positive constants. Even if less standard, another specifications like, e.g., a piecewise linear parametrization could be preferred:

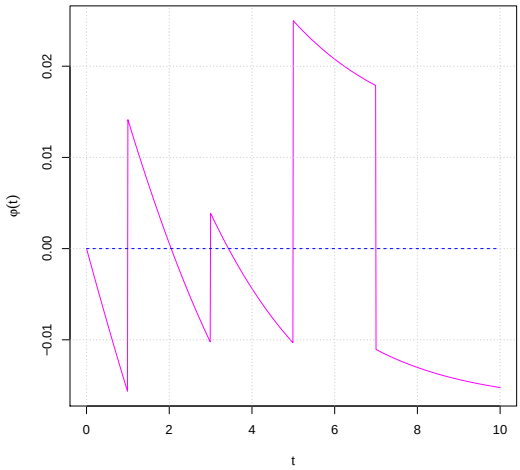
$$h(t) = \sum_{i=1}^{n} \mathbb{1}_{\{T_{i-1} \leq t < T_i\}} \left[\frac{h_i - h_{i-1}}{T_i - T_{i-1}} (t - T_{i-1}) + h_{i-1} \right] .$$

Markit (March 13, 2009).

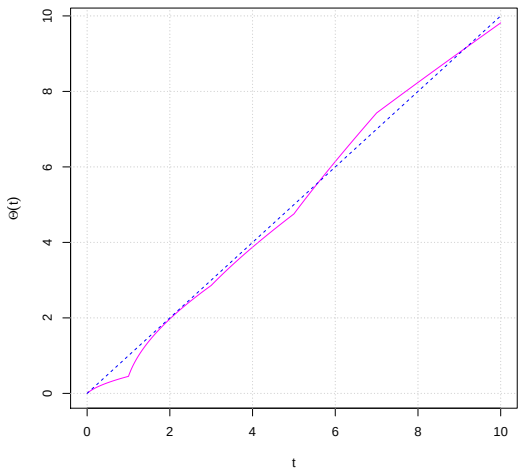
These two different specifications of the hazard rate function are considered on panels (a) of Figure 1 and 2, respectively. These frameworks yield similar (yet, slightly different) market curves $G(t)$ (green curves on panels (d)). For each of them, we start by computing the "best" base CIR model y. In line with market practice, we take $\Xi \leftarrow \Xi^\star$ using (1) with $P^{model} \leftarrow P^y$ considering (2) as error function and $\mathcal{T}$ the set of available liquid CDS maturities available. In each case, we consider the two adjusted intensity models associated to the optimal shift (φ, given by (9)) or optimal clock (Θ, given by (16)). The latter are shown on panels (b) and (c), respectively. The model curves P^{λ^φ} (S-CIR) and P^{λ^θ} (TC-CIR) are shown in magenta on panels (d); they agree with each other, and collapse to $G(t)$ due to the perfect fit. Notice that the parametrization of the hazard rate function has little importance: the survival probability curves G, P^y and P^λ are very similar in either cases. Similarly, the clock functions Θ look very similar in both panels (c) of Fig. 1 and 2.



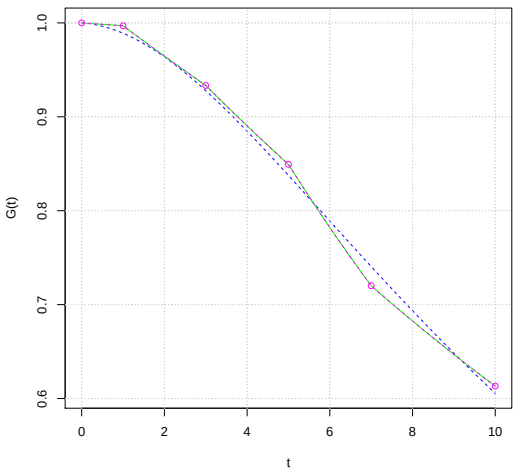
(a) Piecewise constant hazard rate (h)



(b) Optimal shift ($\varphi = \varphi^\star$)

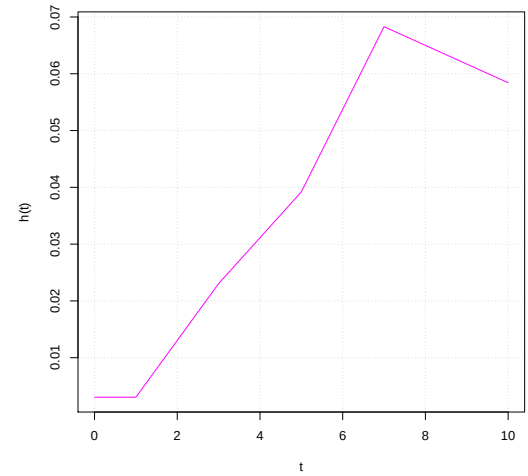


(c) Optimal clock ($\Theta = \Theta^\star$)

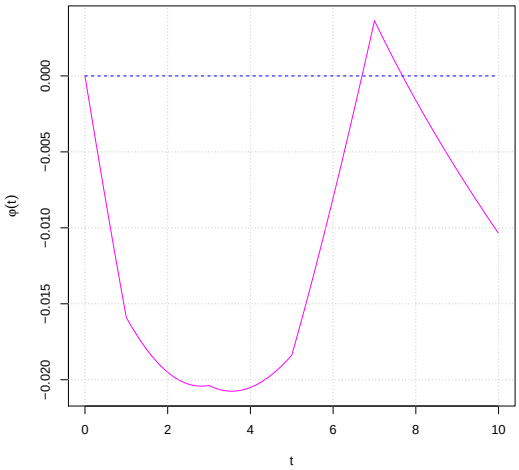


(d) Survival probability ($P^y, G = P^{\lambda^\varphi} = P^{\lambda^\theta}$)

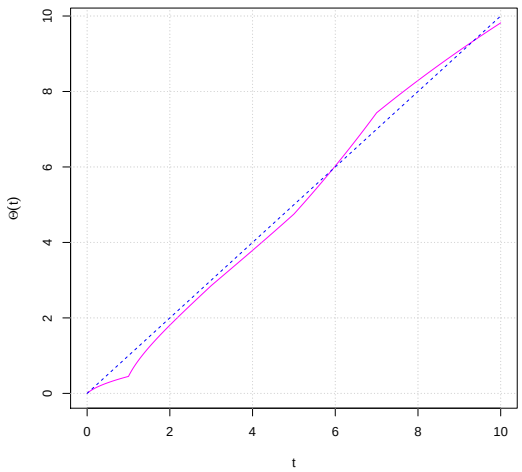
Figure 1: Fitting Ford Inc. CDS term-structure with adjusted CIR models. The survival probability curve $G(t)$ is parametrized with a piecewise constant hazard rate function $h(t)$ extracted from market prices taken from Bloomberg on November 12, 2018, panel (a). The base model y is a CIR with parameters $\Xi = \Xi^\star$ where $\Xi^\star = (\kappa, \beta, \eta, \delta, \alpha, \omega) = (0.0555, 0.3018, 0, 0.2939, 0, 0)$ is obtained from (9) and $y_0 = h_0$. The shift function $\varphi(t) \leftarrow \varphi^\star(t, \Xi^\star)$ is shown in panel (b). Panel (c) gives the clock $\Theta(t) \leftarrow \Theta^\star(t; \Xi^\star)$. Eventually, panel (d) yields the survival probability curves given by the market ($G(t)$, green), or associated to $\mathbb{Q}(\tau(\lambda) > t)$ for various intensity models λ : the best base model $\lambda \leftarrow y$ (leading to $\mathbb{Q}(\tau(y) > t) = P^y(t, \Xi^\star)$, dashed blue), $\lambda \leftarrow \lambda^\varphi$ (S-CIR) and $\lambda \leftarrow \lambda^\theta$ (TC-CIR model). By construction of φ and Θ, the last two curves coincide (magenta) and agree with $G(t)$.



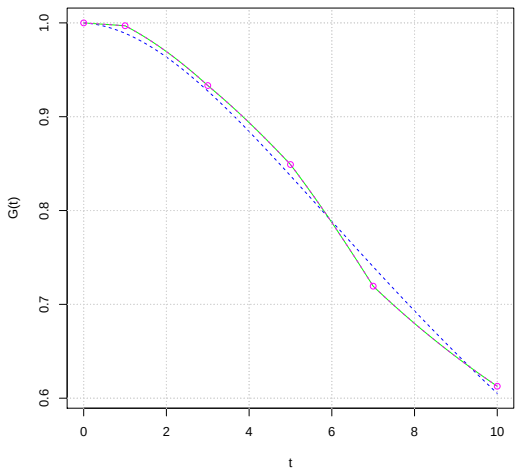
(a) Piecewise linear hazard rate (h)



(b) Optimal shift ($\varphi = \varphi^\star$)



(c) Optimal clock ($\Theta = \Theta^\star$)



(d) Survival probability ($P^y, G = P^{\lambda^\varphi} = P^{\lambda^\theta}$)

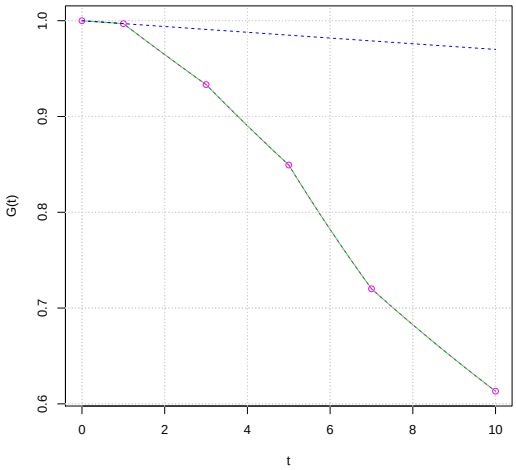
Figure 2: Fitting Ford Inc. CDS term-structure with adjusted CIR. The survival probability curve $G(t)$ is parametrized with a piecewise linear hazard rate function $h(t)$ extracted from market prices taken from Bloomberg on Novermber 12, 2018, panel (a). The base model y is a CIR with parameters $\Xi = \Xi^\star$ where $\Xi^\star = (\kappa, \beta, \eta, \delta, \alpha, \omega) = (0.0620, 0.2729, 0, 0.2926, 0, 0)$ is obtained from eq. (9) and $y_0 = h_0$. The shift function $\varphi(t) \leftarrow \varphi^\star(t, \Xi^\star)$ is shown in panel (b). Panel (c) gives the clock $\Theta(t) \leftarrow \Theta^\star(t; \Xi^\star)$. Eventually, panel (d) yields the survival probability curves given by the market ($G(t)$, green), or associated to $\mathbb{Q}(\tau(\lambda) > t)$ for various intensity models λ : the best base model $\lambda \leftarrow y$ (leading to $\mathbb{Q}(\tau(y) > t) = P^y(t, \Xi^\star)$, dashed blue), $\lambda \leftarrow \lambda^\varphi$ (S-CIR) and $\lambda \leftarrow \lambda^\theta$ (TC-CIR model). By construction of φ and Θ, the last two curves coincide (magenta) and agree with $G(t)$.

The parameters used in the numerical examples in the rest of the paper are given in Table 2.

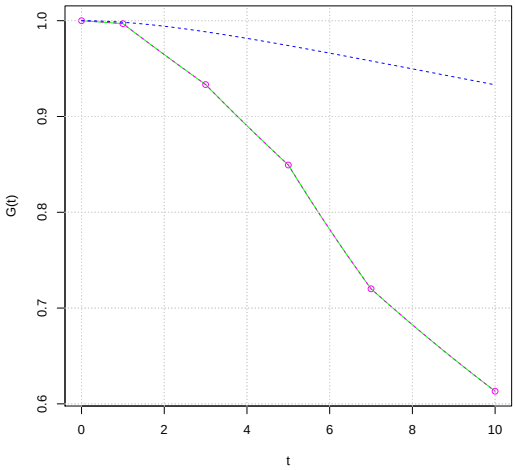
Ξ	κ	β	δ	y_0
$\Xi^\star$	0.0555	0.3018	0.2939	h_0
$\Xi^{\star,+}$	0.2118	0.0030	0.0006	h_0
$\Xi_0^\star$	0.0624	0.2975	0.3343	0.0000
$\Xi_0^{\star,+}$	$3.8252.10^{-01}$	$9.6881.10^{-03}$	$1.5195.10^{-01}$	$3.2093.10^{-10}$

Table 2: Calibration parameters using Ford piecewise constant hazard rate. Parameters $\Xi^\star$ and $\Xi^{\star,+}$ correspond to the parameters of the CIR model y with and without positivity constraint, where y_0 is set exogenously to the first level of the piecewise hazard rate function, $h_0 = 0.0030$. The other parameters, $\Xi_0^\star$ and $\Xi_0^{\star,+}$, correspond to the similar cases but where y_0 is a parameter that enters the optimization procedure. In all cases, we have taken $\alpha = \omega = 0$.

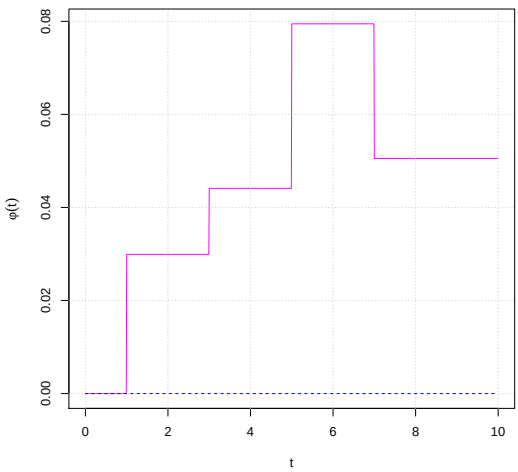
Notice that in both Figure 1 and 2, the shift function φ can take negative values. This means that the shift approach, S-CIR, yields negative default intensities λ^φ and, calibrated that way, is flawed. In particular, we cannot interpret λ^φ as a default intensity associated to a Cox process. This contrasts with the TC-CIR approach since λ^θ is a positive process if so is y. To fix this issue in a CIR++ framework, one needs to rely on PS-CIR. We note the corresponding processes y^+ and $\lambda^{\varphi,+}$. As illustrated on Figure 3 with our Ford example, this procedure is very restrictive: it leads to a curve P^y that is decreasing at a very low rate. In particular, the shape of $P^{\lambda^{\varphi,+}}$ essentially results from the shift, not from the base model y. This is problematic: it basically amounts to say that $h \approx \varphi$, i.e. that the PS-CIR process $\lambda^{\varphi,+}$ is essentially deterministic. This will put strong limitations on the resulting default model, and will be further discussed in the remaining subsections.



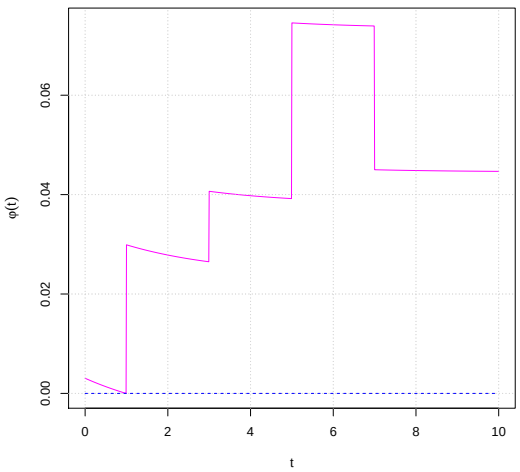
(a) Survival probability $(P^{y^+}, G = P^{\lambda^{\varphi,+}})$



(b) Survival probability $(P^{y^+}, G = P^{\lambda^{\varphi,+}})$



(c) Shift function $(\varphi = \varphi^{\star,+})$



(d) Shift function $(\varphi = \varphi^{\star,+})$

Figure 3: Fitting Ford Inc. CDS term-structure using $\lambda^{\varphi^\star,+}$ (PS-CIR). Panels (a) and (c) correspond to $y_0^+ = h_0$ whereas panels (b) and (d) correspond to the case where y_0^+ is one of the optimized parameters. The survival probability curve $G(t)$ is parametrized with a piecewise constant hazard rate function $h(t)$ extracted from market prices taken from Bloomberg on November 12 2018. The parameters $\Xi^{\star,+}$ are computed under the constraint $\varphi(t) \leftarrow \varphi^\star(t; \Xi^{\star,+}) \geq 0$. The base model y^+ is a CIR with parameters $\Xi = \Xi^{\star,+}$ (left) and $\Xi = \Xi_0^{\star,+}$ (right).

5.2 Variance analysis

Interesting observations can be made regarding the variance of the various integrated processes. As shown in the next two sections, they will have important consequences when considering financial applications, where Λ plays a central role in governing volatility and covariance effects.

First, observe that the integrated CIR process with optimal parameter $\Xi^\star$ is expected to feature a larger variance compared to the integrated CIR with parameter $\Xi^{\star,+}$. Because of the shift constraint, the discount curve P^y in the latter case rends to be much flatter than in the former case i.e. one expects to have, in general

$$P^y(t;\Xi^{\star,+}) \geq P^y(t;\Xi^\star) \ .$$

This can be observed from panels (a) and (b) of Figure 3. When working with $\Xi^{\star,+}$, a substantial part of the shape of $P^{market} = G$ comes from the deterministic shift. This amounts to limit the randomness of the process. Not surprisingly, this will impact the variance of the integrated process Y. Indeed, because the discount curve of the CIR process with parameter $\Xi^{\star,+}$ generally dominates that of the CIR process with parameter $\Xi^\star$, one intuitively expects the variance of the CIR with parameter $\Xi^\star$ to be larger than that of the CIR with parameter $\Xi^{\star,+}$, due to the zero lower bound. In other words, even if it seems difficult to provide a formal proof, one expects intuitively the following to hold, in general:

$$v^{\Lambda^\varphi}(t) = v^Y(t;\Xi^\star) \geq v^Y(t;\Xi^{\star,+}) = v^{\Lambda^{\varphi,+}}(t) \ .$$

This is indeed the case on Figure 4: v^{Λ^φ} (dotted blue) dominates $v^{\Lambda^{\varphi,+}}$ (solid blue).

Second, observe that for a given base process y, the variance of the integrated TC-CIR is always larger than that of the integrated PS-CIR. Indeed, when working under the positivity constraint (i.e. when y is driven by $\Xi^{\star,+}$), we necessarily have $\Theta^+(t) := \Theta(t;\Xi^{\star,+}) \geq t$, in agreement with Theorem 2. Because for any parameter, the variance of Y is an increasing function of time (Lemma 6 in the Appendix, Section 7.1.2) we have, for $\Xi \leftarrow \Xi^{\star,+}$ in particular,

$$v^{\Lambda^{\theta,+}}(t) = v^Y(\Theta^+(t);\Xi^{\star,+}) \geq v^Y(t;\Xi^{\star,+}) = v^{\Lambda^{\varphi,+}}(t) \ .$$

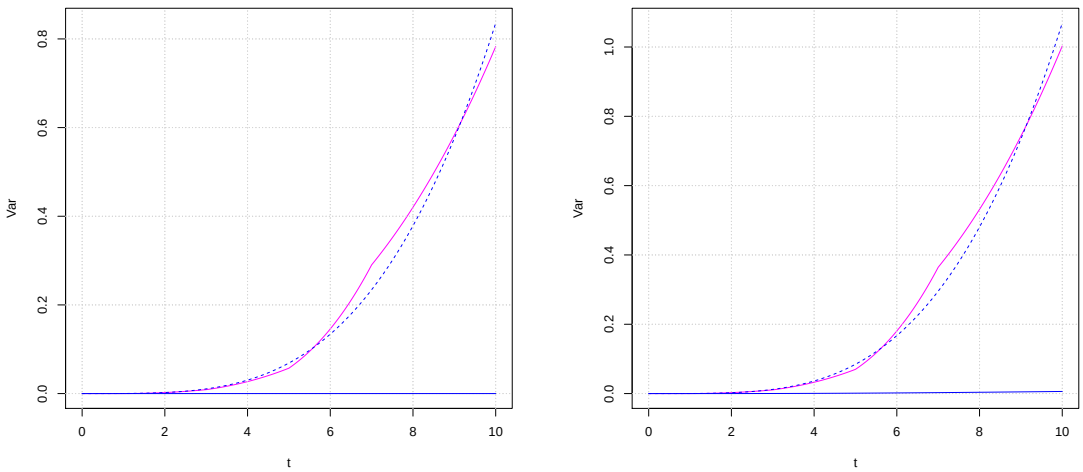
Third, we observe from Figure 4 that, in this example at least, the variance of the TC-CIR using $\Xi^\star$ is comparable to the variance of the S-CIR:

$$v^{\Lambda^\theta}(t) \approx v^Y(t;\Xi^\star) = v^{\Lambda^\varphi}(t) \ .$$

The fact that the variance of the S-CIR is expected to be close to that of the corresponding TC-CIR model can be understood intuitively as follows. As explained above the parameter

$\Xi \leftarrow \Xi^\star$ computed using (9) leads the HAJD y to best fits the market curve, and the clock is used to absorb the remaining discrepancies. Therefore, one expects the clock not to deviate much from the actual time, i.e $\theta(t) \approx 1$ and the two processes to behave similarly. In particular, the parameters of y^θ are those of y scaled by $\theta(t)$, and $x_t^\theta = \theta(t) y_t^\theta \approx y_t^\theta$, at least when the fit between P^{market} and the base HAJD model P^y is not too poor; see (17).

To sum up, we observe that when dealing with CIR++ under a positivity constraint, one has to choose between a valid (but low-volatility) PS-CIR process $\lambda^{\varphi,+}$, or a flawed (by high-volatility) S-CIR one λ^φ. By contrast, the TC-CIR model λ^θ is always valid (Corollary 2), always feature a variance that is larger than the PS-CIR counterpart (Theorem 2), and its variance is actually comparable to the large levels generated by the S-CIR. The TC-CIR thus proves to be a solid challenger to CIR++ models. In particular, its features are specifically interesting when dealing with actual credit risk applications, as we now point out based on two case studies.



(a) $y_0 = h_0$

(b) y_0 optimized

Figure 4: Variances of the integrated versions of $\lambda^{\varphi,+}$ (PS-CIR $\Xi = \Xi^{\star,+}$ (left) and $\Xi = \Xi_0^{\star,+}$ (right), solid blue), λ^φ (S-CIR $\Xi = \Xi^\star$ (left) and $\Xi = \Xi_0^\star$ (right), dashed blue), and λ^θ (TC-CIR with $\Xi = \Xi^\star$ (left) and $\Xi = \Xi_0^\star$ (right), magenta).

5.3 Pricing CDS options

We deal with the pricing of a CDS option (CDSO). Because CDSO is an option on CDS, we start by recalling the no-arbitrage pricing equation of a CDS. We note t the valuation time and

assume $\tau > t$ as it is pointless to price a CDS (or a CDSO) post-default. From the perspective of the protection buyer, the time-t value of a 1 dollar notional CDS $CDS_t(a,b,k)$ starting at time T_a with maturity T_b, $t \leq T_a < T_b$, a spread k and (known) loss given default $LGD = (1-R)$ is given by the difference of the conditional risk-neutral expectation of the protection and premium cashflows :

$$\begin{aligned} CDS_t(a,b,k) &= \mathbb{E}\left[(1-R)\mathbb{1}_{\{T_a \leq \tau \leq T_b\}} P_t^r(\tau) | \mathcal{G}_t\right] \\ &\quad -k\,\mathbb{E}\left[\sum_{i=a+1}^{b}\left(\mathbb{1}_{\{\tau \geq T_i\}}\alpha_i P_t^r(T_i) + \mathbb{1}_{\{T_{i-1} \leq \tau < T_i\}}\alpha_i \frac{\tau - T_{i-1}}{T_i - T_{i-1}} P_t^r(\tau)\right)\middle| \mathcal{G}_t\right] \end{aligned}$$

with α_i the day count fraction between dates T_{i-1} and T_i which, in a standard CDS, is around 0.25 (quarterly payment dates). In a reduced-form setup, when the default is triggered by the first jump of a Cox process with intensity λ, this expression can be developped explicitely thanks to the Key lemma:

$$CDS_t(a,b,k) = \mathbb{1}_{\{\tau > t\}}\left(-(1-R)\int_{T_a}^{T_b} P_t^r(u)\partial_u P_t^\lambda(u)du - k\ C_t(a,b)\right) , \tag{21}$$

where $C_t(a,b)$ is the risky duration, i.e. the time-t value of the CDS premia paid during the life of the contract when the spread is 1:

$$C_t(a,b) := \sum_{i=a+1}^{b} \alpha_i P_t^r(T_i) P_t^\lambda(T_i) - \int_{T_{i-1}}^{T_i} \frac{u - T_{i-1}}{T_i - T_{i-1}} \alpha_i P_t^r(u)\partial_u P_t^\lambda(u)du \ .$$

The spread which, at time t, sets the forward start CDS at 0, called par spread, is given by:

$$\mathbb{1}_{\{\tau > t\}} s_t(a,b) := \mathbb{1}_{\{\tau > t\}} \frac{-(1-R)\int_{T_a}^{T_b} P_t^r(u)\partial_u P_t^\lambda(u)du}{C_t(a,b)} \ . \tag{22}$$

The no-arbitrage price of a call option on such a contrat at time $t = 0$ becomes

$$\begin{aligned} PSO(a,b,k) &= \mathbb{E}\left[(CDS_{T_a}(a,b,k))^+ P^r(T_a)\right] \\ &= P^r(T_a)\mathbb{E}\left[e^{-\Lambda_{T_a}}\left((1-R) - \sum_{i=a+1}^{b}\int_{T_{i-1}}^{T_i} g_i(u) P_{T_a}^{r+\lambda}(u)du\right)^+\right] , \end{aligned}$$

where $g_i(u) := (1-R)(r(u) + \delta_{T_b}(u)) + k\frac{\alpha_i r(u)}{T_i - T_{i-1}}(1 - (u - T_{i-1}))$, with $\delta_s(u)$ the Dirac delta function centered at s.

Replacing the base intensity model (λ) by its shifted ($\lambda^\varphi, \lambda^{\varphi,+}$) or time-changed ($\lambda^\theta$) versions leads to model prices noted $PSO^\varphi(a,b,k), PSO^{\varphi,+}(a,b,k)$ and $PSO^\theta(a,b,k)$, respectively. Interestingly, these models are equally tractable as they feature similar expressions that can be

written in terms of the base process λ or its time integral, Λ. For instance, dropping the Ξ for short,

$$\Lambda^{\varphi}_{T_a} = \Lambda_{T_a} + \int_0^{T_a} \varphi(u)du \ , \quad \Lambda^{\theta}_{T_a} = \Lambda_{\Theta(T_a)} \ ,$$

and

$$\begin{aligned} P^{\lambda^{\varphi}}_{T_a}(u) &= P^{\lambda}_{T_a}(u) e^{\int_{T_a}^{u} \varphi(s)ds} = e^{A_{T_a}(u) - B_{T_A}(u)\lambda_{T_a} + \int_{T_a}^{u} \varphi(s)ds} \ , \\ P^{\lambda^{\theta}}_{T_a}(u) &= P^{\lambda}_{\Theta(T_a)}(u) = e^{A_{\Theta(T_a)}(u) - B_{\Theta(T_a)}(u)\lambda_{\Theta(T_a)}} \ , \end{aligned}$$

Recall that these expressions have a closed form when λ is a (J)CIR process.

Such kind of options has little liquidity. Models are then often compared in terms of their capabilities to generate large "implied volatilities". Indeed, empirical evidences show that this is a typical feature of CDS option quotes, when disclosed. Therefore, we compare the models in terms of their "Black volatilities": the volatility that one needs to plug in a "Black-Scholes" type of model to reproduce the model prices. Black model for PSO is recalled in the Appendix, Section 7.5. The Black volatility associated to a model price $PSO^{model}(a, b, k)$ is thus the volatility $\bar{\sigma}$ satisfying $PSO^{model}(a, b, k) = PSO^{Black}(a, b, k, \bar{\sigma})$. Recall that in all cases, the intensity process λ is calibrated to the market, i.e. $P^{\lambda}(t) = G(t)$. In other words, choosing, e.g., a CIR process for the base intensity process λ combined with the correct shift with (φ^+) or without (φ) positivity constraint, or eventually using the correct clock rate θ, all three models yield the same survival probability curve ($P^{\lambda^{\varphi}}(t) = P^{\lambda^{\varphi,+}}(t) = P^{\lambda^{\theta}}(t) = G(t)$). Hence, all these models agree on the par spread:

$$s_0(a, b) = \frac{(1 - R) \int_{T_a}^{T_b} P^r(u)h(u)G(u)du}{C_0(a, b)} \ .$$

We compare the S-CIR, the PS-CIR and the TC-CIR. The base HAJD process y in TC-CIR is taken to be the same as that of the S-CIR. One can see from Table 3 that the S-CIR features large implied volatilities. Recall however that it allows for negative intensities, hence is not appropriate. The PS-CIR model are not capable of generating large volatility levels, in line with the previous discussion. The TC-CIR fits in between: it rules out negative intensities, while maintaining substantial volatility levels.

One might be concerned by the fact that the implied volatilities of the TC-CIR remain relatively small. This can be addressed in two ways. First, one can play with the parameter Ξ. However, the Feller constraint is often required to hold, which sets limits on the process' volatility. Another approach consists of considering a JCIR model as HAJD. Indeed, JCIR

is often considered when large volatilities are required. However, as explained in Section 3.3, increasing the volatility by boosting the jump activity while maintaining the calibration to a given market curve G reinforces the positivity issue. Fortunately, we do not have this problem in the TC-JCIR. One can drastically increase the jump activity without impacting the positivity of the TC-JCIR. As a consequence, the TC-JCIR seems very much appropriate when one needs a positive but yet high-volatility process. This is illustrated on Table 4 using the same jump parameters as those given in Brigo and Mercurio (2006). We keep the same parameter Ξ as before for the diffusion part, and play with the jump rate (ω) and jump size (α) in the compound Poisson process J. In every case, the clock Θ is chosen such that the model perfectly fits Ford's survival probability curve. Interestingly, the (positive) TC-JCIR model can feature much larger implied volatility levels than the PS-CIR. The results for the S-JCIR are also larger than the PS-CIR, but they are not shown because the negative intensity problem is magnified.

T_a	T_b	CIR++		TC-CIR
		λ^{φ}	$\lambda^{\varphi,+}$	λ^{θ}
1	3	67.12%	1.03%	43.68%
1	5	45.10%	1.25%	26.92%
1	7	27.72%	1.64%	16.86%
1	10	21.52%	1.65%	12.86%
3	5	61.30%	0.85%	57.16%
3	7	34.88%	1.15%	36.33%
3	10	27.65%	1.08%	27.17%
5	7	34.81%	1.16%	42.60%
5	10	30.96%	0.93%	31.19%
7	10	45.37%	0.63%	38.93%

T_a	T_b	CIR++		TC-CIR
		λ^{φ}	$\lambda^{\varphi,+}$	λ^{θ}
1	3	65.21%	9.48%	44.00%
1	5	42.35%	5.88%	26.56%
1	7	25.30%	4.87%	16.30%
1	10	19.37%	2.94%	12.27%
3	5	63.87%	8.02%	58.67%
3	7	35.20%	4.50%	36.10%
3	10	27.02%	3.44%	26.62%
5	7	35.77%	4.59%	43.32%
5	10	30.35%	3.70%	30.61%
7	10	45.05%	5.16%	39.40%

Table 3: Black volatilities for at-the-money ($k = s_0(a,b)$) CDS options implied by CIR++ models (S-CIR and PS-CIR) and the TC-CIR model with $y_0 = h_0$ (left) and y_0 optimized (right) using Monte Carlo simulation (500K paths with time step 0.01). In all the considered cases, the CIR++ model without shift is not valid since $\inf\{\lambda_t^{\varphi},\ t \in [0, T_a]\} < 0$. Among the two valid intensity models ($\lambda_t^{\varphi,+}$ and λ_t^{θ}), the latter exhibits a much higher implied volatility.

T_a	T_b	TC-JCIR (ω, α)		
		$(0,0)$	$(0.1, 0.1)$	$(0.15, 0.15)$
1	3	43.68%	79.04%	100.17%
1	5	26.92%	48.07%	69.69%
1	7	16.86%	30.43%	44.00%
1	10	12.86%	23.33%	33.75%
3	5	57.16%	65.50%	82.60%
3	7	36.33%	42.85%	53.17%
3	10	27.17%	33.17%	41.36%
5	7	42.60%	49.13%	60.11%
5	10	31.19%	37.34%	45.78%
7	10	38.93%	40.59%	46.02%

Table 4: Black volatilities for at-the-money ($k = s_0(a, b)$) CDS options implied by the TC-JCIR model (jump arrival rate ω and jump size α) using Monte Carlo simulation (10^6 paths with time step 0.01) and paramter set $\Xi = \Xi^\star$ but for various jump parameters (α, ω).

5.4 Wrong-way risk impact in credit valuation adjustments

A major concern of the post-crisis regulation is the modeling of the capital requirement of firms tacking into account some credit adjustment to the valuation under credit risk. Counterparty credit risk is defined as the risk that the counterparty of an over-the-counter (OTC) deal will default before the maturity of the contract. The latter can be seen as an option given to the counterparty, and can be priced in a risk-neutral setup by adjusting the OTC derivative, leading to CVA. The latter is nothing but the expected losses due to the missed payments associated to the OTC portfolio. In a risk-neutral specification and assuming $\tau > 0$, the current ($t = 0$) value of the CVA is expressed as:

$$\text{CVA} = \mathbb{E}\left[(1 - R)V_\tau^+ \mathbb{1}_{\{\tau \leq T\}}\right]$$

where V stands for the discounted exposure (i.e. the exposure process rescaled by the stochastic discount factor D). A straightforward application of the Key lemma (under some technical conditions that are valid here) yields

$$\text{CVA} = \mathbb{E}\left[(1 - R)\int_0^T V_u^+ \lambda_u e^{-\Lambda_u} du\right] . \tag{23}$$

The CVA of the shifted and the time-changed models, CVA^{φ} and CVA^{θ}, correspond to above expression, replacing (λ, Λ) by $(\lambda^{\varphi}, \Lambda^{\varphi})$ and $(\lambda^{\theta}, \Lambda^{\theta})$, respectively. The purpose of this section is to illustrate the order of magnitude of CVA figures that can be obtained with either models. In particular, we do not aim at representing a specific exposure. Instead, we simplify the analysis by considering two prototypical dynamics:

$$\begin{aligned} dV_t &= \nu dW_t^V \ , \\ dV_t &= \left(\gamma(T-t) - \frac{V_t}{T-t}\right) dt + \nu dW_t^V \ . \end{aligned}$$

where W^V is an $\mathbb{F}$-Brownian motion. The first SDE is that of a martingale, and can depict the evolution of the discounted price of a forward contract prior to its cashflow date. The second SDE corresponds to a Brownian bridge with drift, and mimics the dynamics of the discounted price of an asset paying continuous dividends. These two models have been previously used in Vrins (2017) and Brigo and Vrins (2018) to describe, in a schematic way, exposures of FRA and IRS. Calibration to actual exposures give indicative value for the parameters.

In general, there is no reason to assume that the Brownian motion driving the default intensity (W) would be independent of the Brownian motion driving the exposure (W^V): it depends on the problem at hand. Usually, we consider the general case of *wrong-way risk* (WWR) effect, obtained by introducing a correlation between the Brownian drivers. For the CIR++ we assume $dW_t dW_t^V = \rho dt$, whereas for the TC-CIR, we correlate W^V with the "synchronized" Brownian motion B in (14), i.e. $dB_t dW_t^V = \rho dt$ (see Mbaye and Vrins (2018) for more details). In the special case where the default time of the counterparty is independent from the discounted exposure (i.e. $\rho = 0$, that is no wrong-way risk) one can easily deduce from (23) the *independent* CVA formula

$$\text{CVA}^{\perp} = -(1-R) \int_0^T \mathbb{E}\left[V_u^+\right] d\,\mathbb{E}\left[e^{-\Lambda_u}\right] = (1-R) \int_0^T f^{\lambda}(u)\, \mathbb{E}\left[V_u^+\right] P^{\lambda}(u) du \ .$$

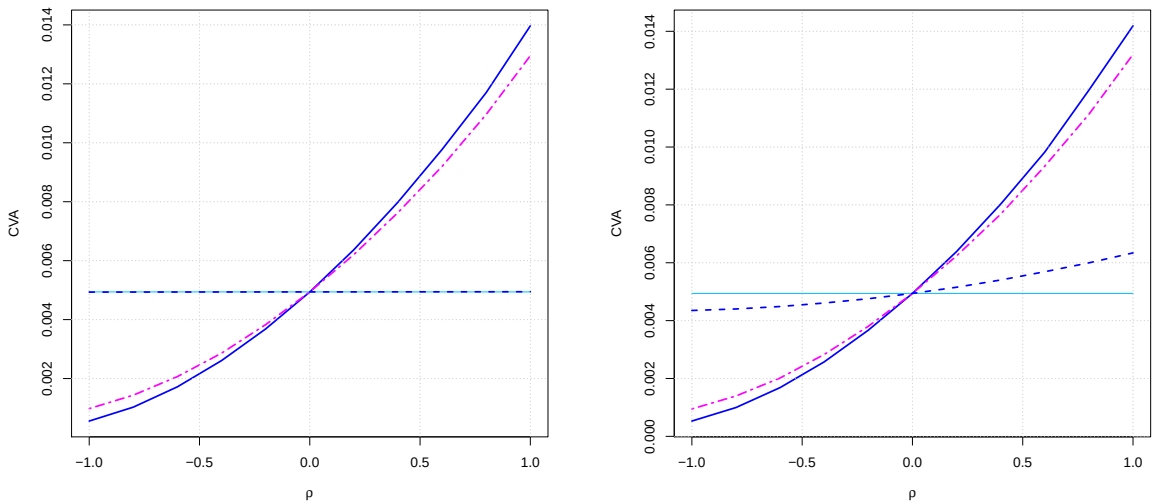
Recall that whatever the chosen model, it is assumed to be calibrated to the survival probability curve G, extracted from CDS prices. This leads to $P^{\lambda}(u) = G(t)$, and to the optimal shift and clock functions, namely φ or φ^+ in the S-CIR and PS-CIR cases, and Θ for the TC-CIR. In this case, $\text{CVA}^{\perp}$ does not depend on the default model:

$$\text{CVA}^{\perp} = -(1-R) \int_0^T \mathbb{E}\left[V_u^+\right] dG(u) = (1-R) \int_0^T h(u) \mathbb{E}\left[V_u^+\right] G(u) du \ .$$

However, the independent case $\rho = 0$ is unrealistic, and may lead to severe over or underestimations of CVA Kim and Leung (2016); Brigo and Vrins (2018); Breton and Marzouk (2018).

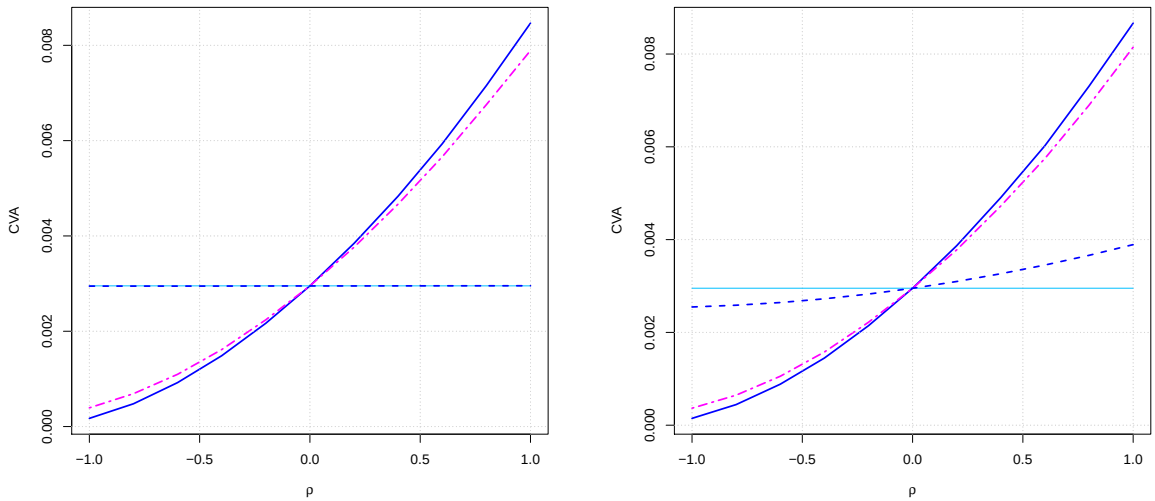
Under WWR, CVA becomes model-dependent. Figure 5 shows the evolution of CVA with respect to ρ for three different models: λ^{φ} (CIR++ without constraint, solid blue), $\lambda^{\varphi,+}$ (CIR++ with constraint, dashed blue) and λ^{θ} (TC-CIR, dashed magenta), all calibrated to Ford's survival probability curve G as before. Under no-WWR, the CVA is equal to the independent CVA (cyan): it is flat, model-free and can be computed using a simple integration. Under WWR, the CVAs are computed using Monte Carlo simulations (100K paths, time step of 0.01) and adaptive control variate [6]. The TC-CIR and S-CIR models exhibit the largest WWR effects and seem therefore appropriate to deal with high WWR applications. Recall that only TC-CIR is valid here as S-CIR gives room to negative intensities. The PS-CIR however is almost flat, equal to the independent CVA. This can be understood from the fact that WWR is essentially a covariance effect between V and $e^{-\Lambda}$. Hence, the models featuring large variance for Λ exhibit larger WWR effects at any (non-zero) fixed correlation level ρ. Eventually, TC-CIR provides an appealing trade-off: on the one hand, as the PS-CIR, it rules out the negative intensity problem inherent to the S-CIR model. But on the other hand, it preserves, to some extend, the variance of the S-CIR model, and therefore exhibits a much larger variance compared to PS-CIR.

[6] see Mbaye and Vrins (2018) for the implementation of the adaptive control variate applied on CVA computation.



(a) $dV_t = \nu dW_t^V$, $y_0 = h_0$

(b) $dV_t = \nu dW_t^V$, y_0 optimized



(c) $dV_t = \left(\gamma(T-t) - \frac{V_t}{T-t}\right) dt + \nu dW_t^V$, $y_0 = h_0$

(d) $dV_t = \left(\gamma(T-t) - \frac{V_t}{T-t}\right) dt + \nu dW_t^V$, y_0 optimized

Figure 5: Impact of the exposure-credit correlation ρ on CVA levels for prototypical 5Y Forward (top) and Swap exposure (bottom) with $y_0 = h_0$ (left) or y_0 optimized (right), $\nu = 8\%$ and $\gamma = 0.1\%$. The curves correspond to different intensity models: $\lambda^{\varphi,+}$ (PS-CIR $\Xi = \Xi^{\star,+}$ (left) and $\Xi = \Xi_0^{\star,+}$ (right), dotted blue), λ^{φ} (S-CIR $\Xi = \Xi^{\star}$ (left) and $\Xi = \Xi_0^{\star}$ (right), solid blue), and λ^{θ} (TC-CIR $\Xi = \Xi^{\star}$ (left) and $\Xi = \Xi_0^{\star}$ (right), dotted magenta). The case without wrong-way risk corresponds to the flat (cyan) line.

6 Conclusion

The calibration problem consists of finding the parameters of a model x so as to perfectly fit a given market curve. The perfect fit is an important feature in a pricing context, that is connected to no-arbitrage opportunities and corrects valuation of trading positions. This calls for two important features: the model x must be (i) flexible enough (to be able to generate various shapes) and (ii) tractable enough (to facilitate the parameters' optimization procedure). Time-homogeneous affine models like Vasicek, CIR or JCIR are very good candidates in this respect, and are widely used in interest rates and credit risk modeling. However, as such, they only feature a couple of constants and hence lack calibration flexibility. The deterministic shift extension offers an appealing solution. It consists of starting with a tractable base model y, that is shifted in a deterministic way with a function φ. The resulting process $x_t = y_t + \varphi(t)$ becomes fully flexible. Indeed, any discount or survival probability curve can be generated by such a model. Moreover, it has a tractability level that is very similar to that of y because φ is deterministic. Eventually, for every market curve, the shift $\varphi^\star$ that leads to the perfect fit is known in closed form, as a function of the y parameters and the market curve. However, this method is less appealing when the model x needs to fulfill some range constraints. Among those, non-negativity is of primary importance when modeling interest rates (depending on the type of economy at hand), mortality rate, prepayment rate or default intensities. In the deterministic shift approach indeed, starting with a non-negative base process y is not enough to guarantee that so will be x, without additional constraint on φ. Furthermore, this constraint becomes more and more severe when increasing the process volatility, due to the zero lower bound.

It seems obvious to rule out models allowing for "negative volatilities". However, surprisingly, the same does not seem to apply when it comes to "negative intensities". Yet, both are equally flawed. We believe the reason is twofold: first, negative intensities do not directly generate numerical problems (in contrast with volatilities that often appear in square-roots), so that the issue is less "obvious", second, there is a lack of a sound alternative. The positivity constraint can be dealt with by including a non-negativity constraint on φ. However, this again raises two problems. First the parameter optimization problem becomes more difficult and second, the resulting process x then features a much lower variance than without the constraint, which contradicts empirical evidences. Therefore, one often prefers to disregard the "negative intensities" issue, giving the priority to stochasticity and perfect fit.

In this paper, we develop such an alternative. It simply consists of time-changing a positive

homogeneous affine jump-diffusion. The model remains tractable, positive, the optimal clock is found by simple inversion and features larger implied volatility compared to the shift approach. Moreover, the perfect fit is achievable for a broad class of discount curves, including all decreasing discount curves. The features of the model have been illustrated on topical examples taken from credit risk, but other applications could be considered as well. This method thus proves to be a competitive challenger to the shift approach, at least under the (very common) positivity constraint, and when large volatility levels are needed.

7 Appendix

7.1 Properties of some Homogeneous Affine Jump-Diffusions

Let y be a $\mathbb{F}$-adapted jump-diffusion introduced in Definition 3 and $Y_t := \int_0^t y_u du$ its integrated version. We denote $v^x(t) = v^x(t;\Xi) := \mathbb{V}[x_t]$ the variance of a stochastic process x parametrized by Ξ at time t. Without explicit mention, all the results below that are given without proofs can be found in, e.g., Brigo and Mercurio (2006). New results are given in lemmas for further reference.

7.1.1 Vasicek model

The Vasicek model corresponds to the special HAJD case $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\beta, -\kappa, \eta^2, 0, 0, 0)$. The A^y, B^y functions in (3) are given by:

$$\begin{aligned} A_s^y(t;\Xi) &= \left(\beta - \frac{\eta^2}{2\kappa^2}\right)(B_s^y(t;\Xi) + s - t) - \frac{\eta^2}{4\kappa} B_s^y(t;\Xi)^2 \ , \\ B_s^y(t;\Xi) &= \frac{1}{\kappa}\left(1 - \mathrm{e}^{-\kappa(t-s)}\right) \ . \end{aligned}$$

The forward curve associated to this model is proven to be

$$f_s^{\mathrm{VAS}}(t) := (1 - \mathrm{e}^{-\kappa(t-s)})\frac{\kappa^2\beta - \eta^2/2}{\kappa^2} + \frac{\eta^2}{2\kappa^2}\,\mathrm{e}^{-\kappa(t-s)}(1 - \mathrm{e}^{-\kappa(t-s)}) + y_t\,\mathrm{e}^{-\kappa(t-s)} \ . \tag{24}$$

Moreover, both y and Y are Normally distributed at all times, with

$$\begin{aligned} \mathbb{E}[y_t] &= y_0^{-\kappa t} + \beta\kappa B_0^y(t;\Xi) \ , \\ \mathbb{E}[Y_t] &= y_0 B_0^y(t;\Xi) + \beta\left(t - B_0^y(t;\Xi)\right) \ , \\ v^y(t) &= \frac{\eta^2}{2\kappa}\left(1 - e^{-2\kappa t}\right) \ , \\ v^Y(t) &= \frac{\eta^2}{\kappa^2}\left[t + \frac{1 - e^{-2\kappa t}}{2\kappa} - 2B_0^y(t;\Xi)\right] \ . \end{aligned}$$

Lemma 5. *Let y be a Vasicek process and Y its time integral. The functions $v^y(t)$ and $v^Y(t)$ are increasing with respect to t.*

Proof. It is obvious for $v^y(t)$, and for $v^Y(t)$, a few manipulations lead to

$$\frac{d}{dt}v^Y(t) = \left(\frac{\eta}{\kappa}(1-e^{-\kappa t})\right)^2 \geq 0\,.$$

□

7.1.2 CIR model

The CIR model corresponds to the special HAJD case $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\beta, -\kappa, 0, \delta^2, 0, 0)$. The A^y, B^y functions in eq. (3) are given by:

$$\begin{aligned} A^y_s(t;\Xi) &= \frac{2\kappa\beta}{\delta^2}\ln\frac{2\gamma\exp\{(\kappa+\gamma)(t-s)/2)\}}{2\gamma+(\kappa+\gamma)(\exp\{(t-s)\gamma\}-1)}, \\ B^y_s(t;\Xi) &= \frac{2(\exp\{(t-s)\gamma\}-1)}{2\gamma+(\kappa+\gamma)(\exp\{(t-s)\gamma\}-1)}\,. \end{aligned}$$

where $\gamma := \sqrt{\kappa^2+2\delta^2}$. The forward curve associated to this model is given by Brigo and Mercurio (2006)

$$f^{\mathrm{CIR}}_s(t) := \frac{2\kappa\beta(e^{(t-s)\gamma}-1)}{2\gamma+(\kappa+\gamma)(e^{(t-s)\gamma}-1)} + y_t\frac{4\gamma^2 e^{(t-s)\gamma}}{[2\gamma+(\kappa+\gamma)(e^{(t-s)\gamma}-1)]^2}\,. \tag{25}$$

Important characteristics of the CIR processes can be computed explicitly, (see, e.g., Dufresne (2001)). For instance, y is distributed as a non-central chi-squared. The two first order moments of y and Y are respectively given by

$$\begin{aligned} \mathbb{E}[y_t] &= y_0 e^{-\kappa t} + \beta\left(1-e^{-\kappa t}\right)\,, \\ \mathbb{E}[y_t^2] &= y_0^2 e^{-2\kappa t} + \left(\frac{\delta^2}{2\kappa}+\beta\right)\left[2y_0\left(e^{-\kappa t}-e^{-2\kappa t}\right)+\beta\left(1-e^{-\kappa t}\right)^2\right]\,, \\ \mathbb{E}[Y_t] &= t\beta + \frac{(y_0-\beta)}{\kappa}(1-e^{-\kappa t})\,, \\ \mathbb{E}[Y_t^2] &= \left(\frac{y_0-\beta}{\kappa}\right)^2 + \delta^2\left(\frac{y_0}{\kappa^3}-\frac{5\beta}{2\kappa^3}\right) + t\beta\left(\frac{2y_0}{\kappa}-\frac{2\beta}{\kappa}+\frac{\delta^2}{\kappa^2}\right) + t^2\beta^2 \\ &\quad + e^{-\kappa t}\left[-2\left(\frac{y_0-\beta}{\kappa}\right)^2 + \frac{2\beta\delta^2}{\kappa^3} + t\left(-\frac{2y_0\beta}{\kappa}+\frac{2\beta^2}{\kappa}-\frac{2y_0\delta^2}{\kappa^2}+\frac{2\beta\delta^2}{\kappa^2}\right)\right] \\ &\quad + e^{-2\kappa t}\left[\left(\frac{y_0-\beta}{\kappa}\right)^2 - \frac{y_0\delta^2}{\kappa^3} + \frac{\beta\delta^2}{2\kappa^3}\right]\,. \end{aligned}$$

In contrast with the Vasicek model, the variance of the CIR is not always increasing monotonously with time; it depends on the parameters. However, the variance of the integrated CIR is increasing. These properties are proven in the next lemma, and will be central in the proof of Theorem 2.

Lemma 6. *Let y be a CIR process and Y its time integral. Then,*

$$v^y(t) = \frac{\delta^2}{\kappa}\left[y_0(e^{-\kappa t} - e^{-2\kappa t}) + \frac{\beta}{2\kappa}(1-e^{-\kappa t})^2\right] , \tag{26}$$

$$v^Y(t) = \frac{\delta^2}{\kappa^3}\left[\left(2\beta - 2\kappa t(y_0-\beta) - (y_0-\beta/2)\,e^{-\kappa t}\right)e^{-\kappa t} + t\kappa\beta + (y_0 - 5\beta/2)\right] . \tag{27}$$

The function $v^y(t)$ is increasing if $\beta \geq y_0$. Otherwise, it is first increasing up to a time $t^\star$, and then decreasing on $(t^\star, \infty)$. By contrast, $v^Y(t)$ is always increasing.

Proof. The computation of the variances is trivial from the first two moments recalled above. The derivative of the variance of the CIR is given by

$$\begin{aligned}\frac{d}{dt}v^y(t) &= y_0\delta^2(-e^{-\kappa t} + 2e^{-2\kappa t}) + \beta\delta^2 e^{-\kappa t}(1-e^{-\kappa t}) \\ &= \delta^2 e^{-\kappa t}(\beta - y_0) - \delta^2 e^{-2\kappa t}(\beta - 2y_0) .\end{aligned}$$

This expression has a root on the positive half-line at

$$t^\star = \frac{1}{\kappa}\ln\left(1 + \frac{y_0}{y_0-\beta}\right)$$

only if $y_0 > \beta$. Otherwise, $v^y(t)$ is always increasing in t. The derivative of the variance of the integrated CIR with respect of time can be written, after some manipulations, as

$$\frac{d}{dt}v^Y(t) = \frac{\delta^2}{\kappa^2}\left[2y_0 e^{-\kappa t}(e^{-\kappa t} - (1-\kappa t)) + \beta(1 - (2\kappa t + e^{-\kappa t})e^{-\kappa t})\right] .$$

Because $e^{-x} \geq 1 - x$ for all $x \geq 0$ and y_0, κ are positive constants, the first term is positive for all $t \geq 0$. On the other hand, $\beta > 0$, and it is enough to check that $1 - g(\kappa t) \geq 0$ for all $t \geq 0$, with $g(x) := (2x + e^{-x})e^{-x}$. Clearly, $g(0) = 1$ and $g'(x) = 2e^{-x}(1 - x - e^{-x}) \leq 0$ for all $x \geq 0$. Hence, $1 - g(\kappa t) \geq 0$ for all $t \geq 0$. □

7.1.3 JCIR model

The characteristics of the JCIR can be obtained by adjusting those of the corresponding CIR, i.e. with same initial value and diffusion parameters. We note z the former and y the latter, and similarly for their integrated versions (Z and Y, respectively). Hence, if the parameter set for the CIR (y, Y) is $\Xi_0 = (\kappa, \beta, \delta, 0, 0, y_0)$, the parameter set of the corresponding JCIR (z, Z) is $\Xi_0 = (\kappa, \beta, \delta, \alpha, \omega, z_0)$ with $z_0 = y_0$ and $\alpha, \omega \geq 0$. The functions associated to the discount curve are given by

$$\begin{aligned}A_s^z(t;\Xi) &= A_s^y(t;\Xi) + \frac{\alpha\omega}{\delta^2/2 - \kappa\alpha - \alpha^2}\ln\frac{2\gamma e^{\frac{\gamma+\kappa+2\alpha}{2}(t-s)}}{2\gamma + (\kappa+\gamma+2\alpha)(e^{(t-s)\gamma}-1)} , \\ B_s^z(t;\Xi) &= B_s^y(t;\Xi) .\end{aligned}$$

From the above functions, it is easy to see that the forward curve associated to this model reads as

$$f_s^{\mathrm{JCIR}}(t) := f_s^{\mathrm{CIR}}(t) + \frac{2\omega\alpha(e^{(t-s)\gamma}-1)}{2\gamma + (\kappa+\gamma+2\alpha)(e^{(t-s)\gamma}-1)} \ , \tag{28}$$

where $f_s^{\mathrm{CIR}}(t)$ is given in (25). For every valid parameters, $f_s^{\mathrm{JCIR}}(t) \geq f_s^{\mathrm{CIR}}(t)$ for all $t \geq s$. Regarding the moments, we have the following result.

Lemma 7. *Let y (resp. Y) be a CIR (resp. integrated CIR) and z (resp. Z) be a JCIR (resp. integrated JCIR) with same initial value, same diffusion parameters but with jumps governed by (ω, α). Then,*

$$\begin{aligned}
\mathbb{E}[z_t] &= \mathbb{E}[y_t] + \frac{\omega\alpha}{\kappa}(1-e^{-\kappa t}) \ , \\
\mathbb{E}[Z_t] &= \mathbb{E}[Y_t] + \frac{\omega\alpha}{\kappa^2}\left(\kappa t - (1-e^{-\kappa t})\right) \ . \\
v^z(t) &= v^y(t) + \omega\alpha\left[\frac{\delta^2}{2}\left(\frac{1-e^{-\kappa t}}{\kappa}\right)^2 + \alpha\frac{1-e^{-2\kappa t}}{\kappa}\right] \ , \\
v^Z(t) &= v^Y(t) + \frac{\alpha\omega}{\kappa^3}\left[\frac{1-e^{-\kappa t}}{\kappa}\left(\xi(3-e^{-\kappa t}) - 4\delta^2\right) + 2\delta^2 t e^{-\kappa t} + t\left(2\alpha\kappa + \delta^2\right)\right] \ ,
\end{aligned}$$

where $\xi := \delta^2/2 - \alpha\kappa$. The function $v^z(t)$ is increasing with respect to t unless $y_0 > \beta + \omega\alpha/\kappa$, in which case it is first increasing up to a time t_1, and then decreasing on (t_1, ∞). Moreover, $v^z(t) \geq v^y(t)$, $v^Z(t) \geq v^Y(t)$ and $v^Z(t)$ is always increasing.

Proof. Applying Ito's lemma we can solve the JCIR SDE (19) by

$$z_t = z_0 e^{-\kappa t} + \beta(1-e^{-\kappa t}) + \delta\int_0^t e^{-\kappa(t-s)}\sqrt{z_s}dW_s + \int_0^t e^{-\kappa(t-s)}dJ_s \ , \tag{29}$$

and find the SDE governing the integrated JCIR process

$$Z_t = t\beta + \frac{(z_0-\beta)}{\kappa}(1-e^{-\kappa t}) + \delta\int_0^t\int_0^s e^{-\kappa(s-u)}\sqrt{z_u}dW_u ds + \int_0^t\int_0^s e^{-\kappa(s-u)}dJ_u ds \ . \tag{30}$$

From (29), we can write

$$\begin{aligned}
\mathbb{E}[z_t] &= z_0 e^{-\kappa t} + \beta(1-e^{-\kappa t}) + \mathbb{E}\left[\int_0^t e^{-\kappa(t-s)}dJ_s\right] \\
&= \mathbb{E}[y_t] + \omega\alpha\int_0^t e^{-\kappa(t-s)}ds \\
&= \mathbb{E}[y_t] + \frac{\omega\alpha}{\kappa}(1-e^{-\kappa t}) \ .
\end{aligned}$$

Using Ito isometry,

$$\begin{aligned}
v^z(t) &= \mathbb{E}[(z_t - \mathbb{E}[z_t])^2] \\
&= \mathbb{E}\left[\left(\delta \int_0^t e^{-\kappa(t-s)}\sqrt{z_s}dW_s + \int_0^t e^{-\kappa(t-s)}dJ_s - \omega\alpha \int_0^t e^{-\kappa(t-s)}ds\right)^2\right] \\
&= \mathbb{E}\left[\left(\delta \int_0^t e^{-\kappa(t-s)}\sqrt{z_s}dW_s\right)^2\right] + \mathbb{E}\left[\left(\int_0^t e^{-\kappa(t-s)}(dJ_s - \omega\alpha ds)\right)^2\right] \\
&= v^y(t) + \frac{\delta^2\omega\alpha}{\kappa}\int_0^t \mathrm{e}^{-2\kappa(t-s)}(1-e^{-\kappa s})ds + 2\omega\alpha^2\int_0^t e^{-2\kappa(t-s)}ds \\
&= v^y(t) + \omega\alpha\left[\frac{\delta^2}{2}\left(\frac{1-e^{-\kappa t}}{\kappa}\right)^2 + \alpha\frac{1-e^{-2\kappa t}}{\kappa}\right].
\end{aligned}$$

Using a similar procedure applied to (30) combined with Fubini's theorem, one can derive the expectation and variance of the integrated JCIR :

$$\begin{aligned}
\mathbb{E}[Z_t] &= t\beta + \frac{(z_0-\beta)}{\kappa}(1-e^{-\kappa t}) + \mathbb{E}\left[\int_0^t\int_0^s e^{-\kappa(s-u)}dJ_u ds\right] \\
&= \mathbb{E}[Y_t] + \mathbb{E}\left[\int_0^t\int_s^t e^{-\kappa u}due^{\kappa s}dJ_s\right] \\
&= \mathbb{E}[Y_t] + \frac{\omega\alpha}{\kappa}\int_0^t(1-e^{-\kappa(t-s)})ds \\
&= \mathbb{E}[Y_t] + \frac{\omega\alpha}{\kappa^2}\left(\kappa t - (1-e^{-\kappa t})\right)
\end{aligned}$$

and

$$\begin{aligned}
v^Z(t) &= \mathbb{E}[(Z_t - \mathbb{E}[Z_t])^2] \\
&= \mathbb{E}\left[\left(\delta\int_0^t\int_0^s e^{-\kappa(s-u)}\sqrt{z_u}dW_u ds + \int_0^t\int_0^s e^{-\kappa(s-u)}dJ_u ds - \mathbb{E}\left[\int_0^t\int_0^s e^{-\kappa(s-u)}dJ_u ds\right]\right)^2\right] \\
&= \mathbb{E}\left[\left(\frac{\delta}{\kappa}\int_0^t(1-e^{-\kappa(t-s)})\sqrt{z_s}dW_s + \frac{1}{\kappa}\int_0^t(1-e^{-\kappa(t-s)})dJ_s - \frac{\omega\alpha}{\kappa}\int_0^t(1-e^{-\kappa(t-s)})ds\right)^2\right] \\
&= \mathbb{E}\left[\left(\frac{\delta}{\kappa}\int_0^t(1-e^{-\kappa(t-s)})\sqrt{z_s}dW_s\right)^2\right] + \mathbb{E}\left[\left(\frac{1}{\kappa}\int_0^t(1-e^{-\kappa(t-s)})(dJ_s - \omega\alpha ds)\right)^2\right] \\
&= v^Y(t) + \frac{\delta^2\omega\alpha}{\kappa^2}\int_0^t(1-e^{-\kappa(t-s)})^2(1-e^{-\kappa s})ds + \frac{2\omega\alpha^2}{\kappa^2}\int_0^t(1-e^{-\kappa(t-s)})^2ds \\
&= v^Y(t) + \frac{\alpha\omega}{\kappa^3}\left[\frac{1-e^{-\kappa t}}{\kappa}\left((\delta^2/2-\alpha\kappa)(3-e^{-\kappa t}) - 4\delta^2\right) + 2\delta^2 te^{-\kappa t} + t\left(2\alpha\kappa + \delta^2\right)\right].
\end{aligned}$$

Notice that the above results can be obtained using another procedure, namely by deriving once or twice the characteristic function $\Psi_t(u,v) = \mathbb{E}[e^{uz_t + vZ_t}]$ of (z_t, Z_t) which can be recovered from eq. (A.1) in Duffie and Gârleanu (2001). This procedure is however much heavier. [7]

[7] Be aware that there are typos in this formula. The correct expression can be found in eq. (B.9) in the draft version of Duffie and Gârleanu's paper, that is available for download on the authors' webpage.

The derivative of the variance of the JCIR is given by

$$\begin{aligned}\frac{d}{dt}v^z(t) &= y_0\delta^2(-e^{-\kappa t}+2e^{-2\kappa t})+\beta\delta^2 e^{-\kappa t}(1-e^{-\kappa t})+\frac{\delta^2\omega\alpha}{\kappa}e^{-\kappa t}(1-e^{-\kappa t})+2\omega\alpha^2 e^{-2\kappa t}\\ &= \delta^2 e^{-\kappa t}(\beta-y_0+\omega\alpha/\kappa)-\delta^2 e^{-2\kappa t}(\beta-2y_0+\omega\alpha/\kappa-2\omega\alpha^2/\delta^2)\ .\end{aligned}$$

This expression has a root on the positive half-line at

$$t_1 := \frac{1}{\kappa}\ln\left(1+\frac{y_0+2\omega\alpha^2/\delta^2}{y_0-\beta-\omega\alpha/\kappa}\right) \tag{31}$$

only if $y_0 > \beta+\omega\alpha/\kappa$. Otherwise, $v^y(t)$ is always increasing in t.

It seems intuitive that the variance of the JCIR cannot be smaller than that of the CIR, and similarly for the integrated versions. Because of the mean reverting effect however, this needs to be confirmed. It is obvious that $v^z(t)-v^y(t)\geq 0$. The term associated to $v^Z(t)-v^Y(t)$ starts at zero (since obviously $v^Z(0)=v^Y(0)=0$). This difference is increasing:

$$\frac{d}{dt}(v^Z(t)-v^Y(t)) = \frac{\delta^2\omega\alpha}{\kappa^3}\left(1-(2\kappa t+e^{-\kappa t})e^{-\kappa t}\right)+\frac{2\omega\alpha^2}{\kappa^2}\left(1-e^{-\kappa t}\right)^2\geq 0\ .$$

Indeed, the second term is obviously positive and the first term takes the form $\frac{\delta^2\omega\alpha}{\kappa^3}(1-g(\kappa t))$ where the function $g(x)=(2x+e^{-x})e^{-x}$ is shown to be bounded by 1 for $x\geq 0$ in the proof of Lemma 6. This shows that $v^Z(t)\geq v^Y(t)$. Because both $v^Y(t)$ (from Lemma 6) and $v^Z(t)-v^Y(t)$ are increasing; $v^Z(t)$ is itself increasing. □

7.2 Some special cases where P^{x+y} is a discount curve

Observe first that P^{x+y} is a discount curve whenever P^x, P^y are in the case where x, y are independent since then $P_s^{x+y}=P_s^x P_s^y$, and the product of two time-s discount curves is itself a time-s discount curve. The next lemma provides sufficient conditions on x, y for P^{x+y} to be a discount curve in the general case.

Lemma 8. *Let x, y be two squared-integrable stochastic processes. Suppose further that the functions*

$$P^z(t) := \mathbb{E}\left[e^{-\int_0^t z_u du}\right]\ , \qquad z\in\{4x,4y\}$$

are discount curves in the sense of Definition (1). Then, $P^{x+y}(t)$ is a discount curve.

Proof. We start with the lemma giving sufficient conditions to swap the expectation and derivative operators which can be found in, e.g., Pagès (2018).

Lemma 9. *Let I be a nontrivial interval of $\mathbb{R}$, $\mathcal{B}(I)$ the Borel set of I and $\Psi : I\times\Omega\to\mathbb{R}$, $(x,\omega)\mapsto\Psi(x,\omega)$ be a $\mathcal{B}(I)\otimes\mathcal{G}$-measurable function. If the function Ψ satisfies:*

(i) For every $x \in I$, the random variable $\Psi(x, \omega) \in L^1$,

(ii) $\Psi_x(x,\omega) := \frac{\partial \Psi(x,\omega)}{\partial x}$ exists for all $x \in I$ a.s.,

(iii) There exists $Z \in L^1$ such that for every $x \in I$,

$$|\Psi_x(x,\omega)| \leq Z(\omega) \text{ a.s. .}$$

Then the function $\psi(x) := \mathbb{E}[\Psi(x, \omega)]$ is defined and differentiable at every $x \in I$ with derivative

$$\frac{d\psi(x)}{dx} = \mathbb{E}\left[\Psi_x(x,\omega)\right] \ .$$

From Cauchy-Schwarz inequality,

$$P^{2(x+y)}(t) = \mathbb{E}\left[e^{-2\int_0^t x_u du} e^{-2\int_0^t y_u du}\right] \leq \sqrt{\mathbb{E}\left[e^{-4\int_0^t x_u du}\right] \mathbb{E}\left[e^{-4\int_0^t y_u du}\right]} = \sqrt{P^{4x}(t) P^{4y}(t)}$$

where the RHS is finite for all $t \geq 0$ by assumption. This provides an upperbound to $P^{x+y}(t)$ since

$$P^{x+y}(t) \leq \sqrt{P^{2(x+y)}(t)}$$

from Jensen's inequality. Swapping the derivative and expectation operators yields

$$\frac{d}{dt} P^{x+y}(t) = \mathbb{E}\left[\frac{d}{dt} e^{-\int_0^t (x_u+y_u)du}\right] = -\mathbb{E}\left[(x_t + y_t) e^{-\int_t^t (x_u+y_u)du}\right] \ . \tag{32}$$

Applying Lemma 9 with

$$\Psi(t,\omega) = e^{-\int_0^t (x_u(\omega)+y_u(\omega))du}$$

shows that one can swap the expectation and derivative operators in (32) provided that

$$|\Psi_x(t,\omega)| = |-(x_t(\omega) + y_t(\omega)) e^{-\int_0^t (x_u(\omega)+y_u(\omega))du}| \leq Z(\omega)$$

for some L^1-random variable Z. To that end, obverve first that both expectations $\mathbb{E}\left[-x_t e^{-\int_0^t (x_u+y_u)du}\right]$ and $\mathbb{E}\left[-y_t e^{-\int_0^t (x_u+y_u)du}\right]$ are bounded by $Z_1 := |x_t| e^{-\int_0^t (x_u+y_u)du}$ and $Z_2 := |y_t| e^{-\int_0^t (x_u+y_u)du}$, respectively. Notice now that $Z_1, Z_2 \in L^1$ since

$$\mathbb{E}\left[|Z_1|\right] \leq \sqrt{\mathbb{E}\left[|x_t|^2\right] \mathbb{E}\left[e^{-2\int_0^t (x_u+y_u)du}\right]} = \sqrt{\mathbb{E}\left[x_t^2\right] P^{2(x+y)}(t)} \ ,$$

and similarly for Z_2. From the triangular inequality, $Z := Z_1 + Z_2 \in L^1$. Eventually,

$$|\Psi_x(t,\cdot)| = |-(x_t + y_t) e^{-\int_0^t (x_u+y_u)du}| \leq Z_1 + Z_2 = Z \ ,$$

and the swap in (32) is valid. Eventually, one can show that the expectations $\mathbb{E}\left[-x_t e^{-\int_0^t (x_u+y_u)du}\right]$ and $\mathbb{E}\left[-y_t e^{-\int_0^t (x_u+y_u)du}\right]$ are bounded below. Indeed, for all random variables X, Y with finite first two moments, $\mathbb{E}[XY] \geq \mathbb{E}[X]\,\mathbb{E}[Y] - \sqrt{\mathbb{V}[X]\mathbb{V}[Y]}$ provided that the variances exist. Hence,

$$\mathbb{E}\left[x_t e^{-\int_0^t (x_u+y_u)du}\right] \geq \mathbb{E}\left[x_s\right] P^{x+y}(t) - \sqrt{\left(\mathbb{E}[x_t^2] - (\mathbb{E}[x_t])^2\right)\left(P^{2(x+y)}(t) - (P^{x+y}(t))^2\right)} \,,$$

$\mathbb{E}\left[y_t e^{-\int_0^t (x_u+y_u)du}\right]$. This concludes the proof.

One can check that P^{x+y} is a discount curve when x, y are HAJD, possibly driven by correlated Brownian motions. Indeed, they have well-defined second moments. Moreover, ky is a HAJD for all $k > 0$ if so is y. Hence, $4x$ and $4y$ are HAJDs and P^{4x} and P^{4y} are discount curves, showing that all assumptions of Lemma 8 are met. □

Not surprisingly, one can weaken the assumptions on x, y to ensure that P^x, P^y and P^{x+y} are discount curves when x, y are bounded below.

Lemma 10. *P^y is a discount curve whenever y is bounded below. Similarly, P^{x+y} is a discount curve when x, y are integrable and bounded below.*

Proof. Let us fix $t > 0$. If there exists $k \geq 0$ such that $y > k$, then $e^{-\int_0^t y_u du} \leq e^{-kt} =: K$. Hence,

$$\left|\frac{d}{dt} e^{-\int_0^t y_u du}\right| = |y_t| e^{-\int_0^t y_u du} \leq K|y_t| \,.$$

Noting that y is integrable, one can use Lemma 9 with $\Psi(t, w) \leftarrow e^{-\int_0^t y_u(w)du}$ and $Z(\omega) \leftarrow K|y_t(\omega)|$, justifying the swap betweeen the derivative and expectation operators:

$$\frac{d}{dt} P^y(t) = \frac{d}{dt}\,\mathbb{E}\left[e^{-\int_0^t y_u du}\right] = \mathbb{E}\left[\frac{d}{dt} e^{-\int_0^t y_u du}\right] = \mathbb{E}\left[\frac{d}{dt} e^{-\int_0^t y_u du}\right] \,,$$

where the right-hand side is bounded by the conditional expectation of Z, which is integrable. The derivative of P^{x+y} is proven to exist similarly, since $x + y$ is obviously bounded below. □

7.3 Proof of Theorem 1

Observe first that for every θ and every t, one gets

$$\int_0^t x_u^\theta du = \int_0^t \theta(u) y_{\theta(u)} du = \int_0^{\Theta(t)} y_u du \,.$$

Hence, the expectation of their negative exponentials agree as well:

$$P^{x^\theta}(t) = P^y(\Theta(t)) \,.$$

The specific clock rate $\theta^\star$ given by the calibration equation thus satisfies, for all t,

$$P^{market}(t) = P^{x^{\theta^\star}}(t) = P^y(\Theta^\star(t)) \ . \tag{33}$$

Turning this equality in terms of instantaneous forward rates yields

$$\int_0^t f^{market}(u)du = \int_0^{\Theta^\star(t)} f^y(u)du \ .$$

Eq. (15) is just the differential form of the latter.

It is not clear, in general, to determine when this ODE admits a solution. However, a simple case is when P^y is a strictly decreasing discount curve. In this case indeed, P^y admits an inverse on the positive half line, noted Q^y. Apply Q^y to (33) yields $\Theta^\star = Q^y(P^{market}(t))$. Furthermore, the inverse of a decreasing function is decreasing, and the combination of two decreasing functions is itself increasing. Hence, if P^{market} is decreasing, $\Theta^\star(t)$ is continuous and strictly increasing. Moreover, $\Theta^\star(0) = Q^y\left(P^{market}(0)\right) = Q^y(1) = 0$. Hence, $\Theta^\star$ exists, and is a clock.

7.4 Proof of Theorem 2

It is known from Corollary 2 that for any (non-trivial) (J)CIR process y with parameter Ξ, there exits a clock $\Theta^\star(t) = \Theta^\star(t;\Xi)$ that yields a perfect fit between the curves P^x generated y the TC-JCIR $x_t^{\theta^\star} := \theta^\star(t)y_{\Theta^\star(t)}$. Same holds true for the JCIR++, $x_t^{\varphi^\star}$. This means:

$$P^{x^{\varphi^\star}}(t;\Xi) = P^{market}(t) = P^{x^{\theta^\star}}(t;\Xi) \ ,$$

or equivalently,

$$e^{-\int_0^t \varphi^\star(u)du} P^y(t;\Xi) = P^{market}(t) = P^y(\Theta^\star(t);\Xi) \ .$$

Because y is a JCIR, it can be arbitrarily close to 0 at any time t, hence the calibration constraint amounts to force $\varphi^\star(t) \geq 0$ (or equivalently, $f^{\text{JCIR}}(t) \leq f^{\text{market}}(t)) \ \forall\, t \geq 0$. This implies that $P^y(\Theta^\star(t);\Xi) \leq P^y(t;\Xi)$. Because $P^y(.;\Xi)$ is a decreasing function, the last inequality is equivalent to $\Theta^\star(t) \geq t$.

To prove 1), we start from the increasingness of $v^Y(t)$ (Lemma 7). Hence, $v^Y(\Theta^\star(t)) \geq v^Y(t)$ since $\Theta^\star(t) \geq t$.

From (28), we have, after some computations,

$$\begin{aligned}
\frac{d}{dt} f^{\text{JCIR}}(t) &= \frac{4\gamma^2 e^{t\gamma}\left[\kappa\beta(\gamma-\kappa+(\kappa+\gamma)e^{t\gamma}) + y_0\gamma(\gamma-\kappa-(\kappa+\gamma)e^{t\gamma}) + \omega\alpha(2\gamma+(\kappa+\gamma)(e^{t\gamma}-1))\right]}{[2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)]^3} \\
&= \frac{4\gamma^2 e^{t\gamma}\left[(\gamma-\kappa)(\kappa\beta+y_0\gamma+\omega\alpha) - 2\omega\alpha^2\right] + 4\gamma^2 e^{2t\gamma}\left[(\gamma+\kappa)(\kappa\beta-y_0\gamma+\omega\alpha) + 2\omega\alpha^2\right]}{[2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)]^3} \ .
\end{aligned}$$

From this expression, one can check that f^{JCIR} is strictly increasing if $y_0 < \beta + \omega\alpha/\kappa$ and $y_0\gamma \leq \kappa\beta + \omega\alpha$. It is strictly decreasing if $y_0 \geq \beta + \omega\alpha/\kappa$. Otherwise, i.e. if $y_0 < \beta + \omega\alpha/\kappa$ and $y_0\gamma > \kappa\beta + \omega\alpha$, the derivative has a root at

$$t_2 := \frac{1}{\gamma} \ln \frac{(\gamma - \kappa)(\kappa\beta + y_0\gamma + \omega\alpha) - 2\omega\alpha^2}{(\kappa + \gamma)(y_0\gamma - \kappa\beta - \omega\alpha) - 2\omega\alpha^2} \ .$$

i.e. f^{JCIR} is first increasing, then decreasing.

The constraint $\varphi^\star(t) \geq 0$ for all t, simply means that $f^{market}(t) \geq f^{\text{JCIR}}(t)$ and so $\theta^\star(t) \geq \frac{f^{\text{JCIR}}(t)}{f^{\text{JCIR}}(\Theta^\star(t))}$. Observe that the condition $y_0 = \beta + \omega\alpha/\kappa$ corresponds to the case where $v^y(t)$ is increasing and $f^{\text{JCIR}}(t)$ is decreasing, hence (i) holds.

If f^{market} is constant, we have that $f^{market}(t) \geq f^{\text{JCIR}}(t)$ which implies that $f^{market}(t) \geq f^{\text{JCIR}}(\Theta^\star(t))$. Clearly, if $f^{market}(t)$ is constant or f^{JCIR} is decreasing, then $\theta^\star(t) \geq 1$. And if $v^y(t)$ is increasing, $v^y(\Theta^\star(t)) \geq v^y(t)$ since $\Theta^\star(t) \geq t$. From the fact that $\mathbb{V}[x_t^{\theta^\star}] := \theta^\star(t)^2 v^y(\Theta^\star(t))$ and the variation of $v^y(t)$ (Lemma 7), (ii), (iii) and (iv) follow.

In particular, taking $\omega\alpha = 0$, we recover the CIR case which corresponds to the TC-CIR model.

7.5 Black model for PSO

In this context, the Black-Scholes model works as follows. We start by noting that the forward start CDS can be written in terms of the difference between the fair and the agreed premium cashflows. Indeed, the former corresponds to the protection leg. Inserting (22) in (21) yields

$$CDS_t(a,b,k) = \mathbb{1}_{\{\tau > t\}}(s_t(a,b) - k)\ C_t(a,b) \ .$$

The Payer default swaption becomes

$$PSO(a,b,k) = \mathbb{E}\left[(s_{T_a}(a,b) - k)^+\ C_{T_a}(a,b) D(T_a)\right] = C_0(a,b)\mathbb{E}^{(a,b)}\left[(s_{T_a}(a,b) - k)^+\right] \tag{34}$$

where $\mathbb{E}^{(a,b)}$ stands for the expectation under the equivalent measure $\mathbb{Q}^{(a,b)}$, associated with the numéraire $C(a,b)$. Interestingly, it is clear from (22) that the par spread $s(a,b)$ is a $\mathbb{Q}^{(a,b)}$-martingale on $[0, T_a]$. Hence, the Black-Scholes model for CDSO naturally postulates $\mathbb{Q}^{(a,b)}$-martingale dynamics for the par spread

$$ds_t(a,b) = \bar{\sigma} s_t(a,b) dW_t^s, \quad t \leq T_a$$

where W^s is a $\mathbb{Q}^{(a,b)}$-Brownian motion. Eventually, the expectation in (34) is given by the standard Black-Scholes formula by setting $r \leftarrow 0$. Hence, the Black-Scholes price of the PSO is given by

$$PSO^{Black}(a,b,k,\bar{\sigma}) = C_0(a,b)\left[s_0(a,b)\Phi(d_1) - k\Phi(d_2)\right]$$

where

$$d_1 = \frac{\ln \frac{s_0(a,b)}{k} + \frac{1}{2}\bar{\sigma}^2 T_a}{\bar{\sigma}\sqrt{T_a}}, \qquad d_2 = d_1 - \bar{\sigma}\sqrt{T_a}$$

and Φ is the distribution function of a standard Normal random variable.

References

T. Bielecki and M. Rutkowski. *Credit risk : modeling, valuation and hedging.* Springer finance. Springer, 2002.

T. Bielecki, M. Jeanblanc, and M. Rutkowski. Credit risk modeling. Technical report, Center for the Study of Finance and Insurance, Osaka University, Osaka (Japan), 2011.

M. Breton and O. Marzouk. Evaluation of counterparty risk for derivatives with early-exercise features. *Journal of Economic Dynamics and Control*, 88:1–20, 2018.

D. Brigo and A. Alfonsi. Credit default swaps calibration and option pricing with the SSRD stochastic intensity and interest rate model. *Finance and Stochastics*, 9:29–42, 2005.

D. Brigo and L. Cousot. A comparison between the SSRD model and the market model for cds options pricing. *International Journal of Theoretical and Applied Finance*, 9(3), 2006.

D. Brigo and N. El-Bachir. An exact formula for default swaptions pricing in the SSRJD stochastic intensity model. *Mathematical Finance*, 20(3):365–382, 2010.

D. Brigo and F. Merccurio. A deterministic-shift extension to analytically-tractable and time-homogeneous short-rate models. *Finance and Stochastics*, 5:369–388, 2001.

D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice.* Springer, 2006.

D. Brigo and F. Vrins. Disentangling Wrong-Way Risk: Pricing CVA via change of measures and drift adjustment. *European Journal of Operational Research*, 269:254–264, 2018.

D. Brigo, A. Capponi, and A. Pallavicini. Arbitrage-free bilateral counterparty risk valuation under collateralization and application to credit default swaps. *Mathematical Finance*, 24(1): 125–146, 2014.

P. Carr and D. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999.

P. Carr, H. Geman, D. Madan, and M. Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13:345–382, 2003.

JC. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

S. Crépey, M. Jeanblanc, and D. Wu. Informationally dynamized gaussian copula. Technical report, 2012. URL https://www.maths.univ-evry.fr/prepubli/363.pdf.

K.E. Dambis. On the decomposition of continuous submartingales. *Theory of Probability and its Applications*, 10:40914100, 1965.

B. Dellacherie and P.-A. Meyer. *Probabilités et Potentiel - Théorie des martingales*. Hermann, 1980.

L. Dubins and G. Schwartz. On continuous martingales. *Proceedings of the National Academy of Sciences*, 53:913916, 1965.

D. Duffie and N. Gârleanu. Risk and valuation of collateralized debt obligations. *Financial Analyst Journal*, 57:41–59, 2001.

D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379406, 1996.

D. Duffie and K. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12:687–720, 1999.

D. Duffie and K. Singleton. Credit risk: pricing, measurement, and management. *Princeton University Press*, 2003.

D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1056, 2003.

D. Dufresne. The integrated square-root process. Technical report, Research Paper 90, Center for Actuarial Studies, Department of Economics, University of Melbourne, 2001.

P.H. Dybvig. Bond and bond option pricing based on the current term structure. In M. Dempster and S. Pliska, editors, *Mathematics of Derivatives Securities (pp. 271-293)*. Cambridge University Press, 1997.

D. Filipovic. Time-inhomogeneous affines processes. *Stochastic processes and their application*, pages 639–659, 2005.

D. Filipovic, M. Larsson, and A.B. Trolle. Linear-rational term structure models. *Journal of Finance*, 72(2):655–704, 2017.

H. Geman, D. Madan, and M. Yor. Time changes for Lévy processes. *Mathematical Finance*, 11:79–96, 2001.

J. Gregory. *Counterparty Credit Risk*. Wiley Finance, 2010.

D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates. *Econometrica*, 60:77–105, 1992.

S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327343, 1993.

T. Ho and S.-B. Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41(5):10111029, 1986.

J. Hull and A. White. Pricing interest rate derivative securities. *Review of Financial Studies*, 3:573–592, 1990.

M. Jeanblanc and F. Vrins. Conic martingales from stochastic integrals. *Mathematical Finance*, 28(2):516–535, 2018.

M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets*. Springer Verlag, Berlin, 2009.

M. Joshi. *The Concepts and Practice of Mathematical Finance*. Cambridge University Press, 2003.

J. Kim and T. Leung. Pricing derivatives with counterparty risk and collateralization: A fixed point approach. *European Journal of Operational Research*, 249(2):525–539, 2016.

D. Lando. *Credit Risk Modeling: Theory and Applications*. Pinceton Series in Finance. Pinceton University Press, 2004.

Markit. ISDA CDS Standard Model. Technical report, Markit, 2004. http://www.cdsmodel.com/cdsmodel/.

Markit. The CDS Big-Bang: understanding the changes to the Global CDS Contract and North American conventions. Technical report, Markit, March 13, 2009. http://www.markit.com/cds/announcements/resource/cds_big_bang.pdf.

C. Mbaye and F. Vrins. A subordinated CIR intensity model with application to wrong-way risk CVA. *International Journal of Theoretical and Applied Finance*, 21(7):1850045, 2018.

R. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29:449470, 1974.

C. Nelson and A. Siegel. Parsimonious modelling of yield curves. *Journal of Business*, 60, 1987.

G. Pagès. *Numerical Probability: An Introduction with Applications to Finance.* Springer, Cham, 2018.

H. Stein and K. Pong. Counterparty valuation adjustments. In Brigo Bielecki and Patras, editors, *Credit Risk Frontiers: Subprime crisis, Pricing and Hedging, CVA, MBS, Ratings and Liquidity*, chapter 15. Wiley/Bloomberg Press, 2011.

A. Swishchuk. *Change of Time Methods in Quantitative Finance.* SpringerBriefs in Mathematics. Springer, 2016.

O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

P. Veronesi. *Fixed Income Analysis: Valuation, Risk, and Risk Management.* Wiley, 2010.

F. Vrins. Wrong-way risk CVA models with analytical EPE profiles under Gaussian exposure dynamics. *International Journal of Theoretical and Applied Finance*, 20(7):1750045, 2017.