

UNIVERSITY OF PADOVA

DEPARTMENT OF MATHEMATICS
“Tullio Levi-Civita”

Master’s degree in Mathematical Sciences

An affine jump-diffusion model in the negative rates environment

Supervisor
Prof. Claudio Fontana

Candidate
Chiara Rossato
1203102

25 September 2020

Academic Year 2019-2020

Contents

Contents	3
Introduction	5
1 Challenges in the post-crisis framework	7
1.1 Post-crisis circumstances	8
1.1.1 Measures of risk-free rates in the euro area	9
1.2 Literature review	11
1.2.1 The SABR model	11
1.2.2 The shadow-rate models	12
1.3 An affine term structure model with stochastic lower bound	15
1.3.1 The dynamics under the physical probability measure P	16
2 The model	19
2.1 The characterization of the affine short rate model	19
2.2 The affine transform formula	27
2.2.1 From Riccati ODEs system to the affine transform formula	35
2.3 The probability of future jumps	41
2.4 Equivalent changes of measures	44
2.4.1 The exponentially affine change of measure	45
2.4.2 A more general change of measure	51
2.4.3 Comparison between the two density processes	55
2.5 Fitting the initial term structure	55
3 Simulations	59
3.1 The discretization schemes for the affine diffusion process	59
3.2 The Ogata's algorithm	61
3.3 Running the simulations	64
3.3.1 Simulation of the lower bound	64
3.3.2 Model's simulation	66
Conclusions	69

A Mathematical tools	71
A.1 Point processes	71
A.1.1 Stochastic intensity	72
A.1.2 Hawkes processes	74
A.2 Jump-diffusions	77
A.3 Affine jump-diffusions	79
A.3.1 Affine diffusions	82
A.3.2 Transforms	83
A.3.3 Change of measure	87
B Stochastic integration	93
B.1 Brownian integration	93
B.2 Extension of stochastic integration theory	94
C Matlab Implementation	97
Bibliography	103

Introduction

This Thesis develops a new affine term structure model providing a short rate that is allowed to take negative values and is bounded from below by a randomly varying level. With this approach we conveniently represent many of the most relevant empirical features of the financial market that raised after the financial crisis of 2007-08, when the spread of negative rates became not only realistic but real.

To achieve analytical and computational tractability we define a model described by an affine jump-diffusion process. Under suitable technical conditions, its conditional Laplace transform can be explicitly specified in terms of the solutions of a family of ordinary differential equations (ODEs) that are the essence of our affine short rate model.

This work is structured as follows. In the first Chapter, we describe the main challenges originated from the global financial crisis, in particular the beginning of the negative rates environment and the consequent need of formulation of new mathematical models. To this aim we present the available time series of the EONIA and ECB deposit facility rates (observed from January 1999 to May 2020). In the euro area these two rates are the proxies for the risk-free short-term rate and its lower bound, respectively. We analyse their behaviour in order to correctly define a new short rate model which is able to capture these empirical features. We also present a brief overview of the works which deal with negative rates highlighting their main characteristics. In particular, we describe the affine term structure model with stochastic lower bound introduced by Monfort et al. in [52]. We consider their model as a starting point to develop a more sophisticated model in continuous time.

The second Chapter concerns the formulation of our affine short rate model. The short rate process $(r_t)_{t \geq 0}$ is specified as a linear combination of a deterministic function used to achieve a fit to the initially observed term structure, a \mathbb{R}_+^d -valued affine diffusion process $(X_t)_{t \geq 0}$ of macroeconomic factors and a stochastic lower bound $(SLB_t)_{t \geq 0}$ defined by a bivariate counting process arising from Hawkes processes' framework. This corresponds to setting r_t as

$$r_t := h(t) + \langle \ell, X_t \rangle + SLB_t \quad \text{for } t \geq 0,$$

for some $\ell \in \mathbb{R}_+^d$. We prove that this model belongs to the class of affine short rate models since it is completely described by a $(d + 4)$ -dimensional affine jump-diffusion process $(Z_t)_{t \geq 0}$. This feature appears to be particularly useful in providing both flexibility to capture many of the empirical features as well as computational tractability. Indeed, under suitable technical conditions the affine property allows to derive a closed-form

expression of the following discounted transform

$$\mathbb{E} \left[e^{-\int_t^T r(Z_s, s) ds} e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t \right] = e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle} \quad \text{for } u \in \mathbb{C}^{d+4}. \quad (1)$$

The functions $\phi : [t, T] \rightarrow \mathbb{C}$ and $\psi : [t, T] \rightarrow \mathbb{C}^{d+4}$ defining this relationship are the solutions of a generalised Riccati ODEs system which therefore uniquely characterizes the model. We investigate the validity of the affine transform formula (1) in terms of the existence and uniqueness of the solutions ϕ and ψ . In addition, we prove that the conditional distribution of the stochastic lower bound process $(SLB_t)_{t \geq 0}$ can be computed directly from the affine transform formula (1).

We have so far introduced the affine short rate model for the process $(r_t)_{t \geq 0}$ with respect to the physical probability measure P . However, the analytical tractability deriving from the affine property of the model may be lost under a change of measure. Therefore, we characterize the risk-neutral measures Q which preserve the affine structure of the model. Firstly, we suppose that the density process defining the change of measure is an exponential-affine form in $(Z_t)_{t \geq 0}$ and then generalize this approach. We investigate these changes of measures and give sufficient conditions under which the affine structure of the model is preserved, from the physical P to an equivalent probability measure Q .

In the third Chapter we focus on the empirical behavior of our short rate model simulating the process $(r_t)_{t \geq 0}$ by means of the software Matlab. For simplicity we set the short rate process as a linear combination of the affine diffusion $(X_t)_{t \geq 0}$ and the stochastic lower bound $(SLB_t)_{t \geq 0}$, only. In order to generate paths of the affine diffusion process $(X_t)_{t \geq 0}$ we analyse different approximation methods that may be seen as modifications of the Euler or the Milstein schemes. Then, we focus on the simulation of the bivariate counting process that determines the stochastic lower $(SLB_t)_{t \geq 0}$. To this aim we apply the Ogata's algorithm.

Finally, in Appendix A we set a common notation and terminology giving a brief description of point processes, analyzing in detail Hawkes processes. We introduce the main concepts for the study of affine jump-diffusion processes and prove general results needed for our analysis. Some notions on stochastic integration are summarized in Appendix B. In Appendix C we report the Matlab code used to simulate the short rate process $(r_t)_{t \geq 0}$ and its stochastic lower bound $(SLB_t)_{t \geq 0}$.

Chapter 1

Challenges in the post-crisis framework

In this Chapter we aim at introducing the main challenges arisen from the global financial crisis of 2007-08. What we call global financial crisis is a crisis of credit markets whose roots can be traced back to the fallout from the housing market in the United States in early 2007. Unfortunately it spread to economies and financial markets all round the world with significant effects, such as slow growth, fall in the inflation rate and increasing financial uncertainty. In this sluggish and severe scenario banks took too much risk becoming weak and consequently they reported persistent large-scale losses. Then, some of these financial institutions became insolvent and had to be taken over or saved by their governments.

Not only national governments but also the main central banks took numerous actions to prevent the crisis from spreading further and lowering their official policy rates was one of the main operations. Reference rates reached negative values and this was contrary to the theoretical assumption of mathematical models that interest rates could not drop below zero. Consequently, there was a need to create new mathematical foundations able to provide adequate models for negative interest rates.

In this Chapter, first of all we will give a brief description of the introduction of increasingly negative central bank policy rates and analyze its impact on the study of theoretical models. We will investigate whether the no-longer valid assumption of non-negativity of interest rates was true and to this aim we will follow the ideas introduced in [39] which provides a theoretical economy approximating the reality in which the zero lower bound on interest rates is inadequate. This crucial change has led to the necessity to provide new mathematical solutions on the formulation of models and for this purpose we will describe some models which deal with negative rates, in particular the affine term structure model with stochastic lower bound introduced by Monfort et al. in [52]. Their discrete-time model is characterized by the presence of a time-varying lower bound on interest rates which takes value in \mathbb{R} and therefore it allows to provide negative rates. In Chapter 2 we will take this model as a reference to develop a continuous-time one.

1.1 *Post-crisis circumstances*

For most of history interest rates have been positive and indeed economists have traditionally assumed they cannot fall below zero. However, based on the empirical evidence, this assumption of a zero lower bound has become inadequate: in fact, after the financial crisis of 2007-08 several central banks around the world have reacted to economic changes with the introduction of unconventional policies keeping their policy rates below zero to provide monetary stimulus with the aim of encouraging economic growth and stabilizing inflation expectations, as it is richly explained in [4] (Section 2.1).

In June 2014 the ECB (European Central Bank) became the first major central bank to lower one of its policy rates to negative territory and it was from that moment that negative interest rates came to global attention as a major phenomenon. In the ECB's press release dated 5 June 2014 there was communicated the decision to cut the interest rate on the deposit facility to -0.10% . The introduction of negative rates was part of a policies' package aimed at fending off deflationary risks, as it is studied in [4] (Section 3.1). In other words, in the euro area in those years inflation was below its target rate of 2% and the ECB intended to stimulate economic activity by encouraging bank lending because, with the introduction of negative deposit rate, commercial banks were charged for their deposits at the central bank and therefore were penalized for holding overnight deposits.

Subsequently, in December 2014 the SNB (Swiss National Bank) announced the introduction of negative interest rate on sight deposit account balances which was imposed at -0.25% . Bech et al. in [12] (Section *Context for negative policy rates*) explain that the goal was to take the three-month Libor into negative territory replying to foreign capital inflow pressures and discouraging them. These reasons are also written in the SNB's press release dated 18 December 2014, where it is explained that the SNB's aim was the minimum exchange rate of 1.20 Swiss francs per euro. Furthermore, the adoption of negative interest rates was accompanied by the introduction of a tiered reserve system where negative rates only apply to reserves above a certain exemption threshold (for details see [12] (Box 2)).

These previously thought as unconventional policies not only have been adopted by the ECB and the SNB but most recently also by the BoJ (Bank of Japan). On January 2016 it announced that it would apply a negative rate of -0.10% to part of the balances in current accounts with the aim of achieving price stability and providing more accommodative financial conditions. These justifications are reported in the press release dated 29 January 2016 and in this document the BoJ illustrated that it would be adopted a remuneration schedule that divides deposits at the central bank into three tiers, to each of which a positive interest rate, a zero interest rate, or a negative interest rate would be applied, respectively, allowing that negative interest rate only applies to a fraction of bank reserves, as it is deeper analyzed in [4] (Section 3.7). It is similar to what happened in Switzerland, the difference is that the SNB has fixed only two tiers for the banks' reserves.

However, negative policy rates were not entirely new. The SNB sporadically intro-

duced negative interest rates on foreign deposits during the 1970s in order to reduce capital inflows and excessive appreciation of Swiss franc, as Arteta et al. call to mind in [7] (Section 1). Moreover, yields on some Japanese government bonds were negative for a brief period in late 1990s. Despite these events, economists thought that these were minor occurrences and they paid no attention to negative rates until the financial crisis of 2007-2008.

Summing up, over the past few years in order to stimulate economic recovery or reduce foreign capital inflows central banks have resorted to low rate policies more and more overcoming the zero lower bound. Roughly speaking, central banks made financial institutions pay to hold overnight deposits at them, nevertheless some banks have chosen to hold higher balances at central banks because of uncertain financial environment. This is in contradiction with what happened before the global crisis, when financial institutions were used to minimize holdings in excess reserves because central banks deposit rates are below money market rates.

1.1.1 Measures of risk-free rates in the euro area

In the euro area the global financial crisis of 2007-2008 has another substantial consequence because it has posed a number of challenges for measuring risk-free rates. The deposit rate cuts to negative territory have been transmitted directly to other interest rates, in particular government bond yields which fell below zero. Additionally, as a consequence of sluggish growth and unsure financial market, it followed credit rating downgrades that have mechanically shrunk the pool of AAA government bonds, in turn making the AAA curve less representative of the euro area as a whole. As it is analyzed in [26] (Section 4.3) the development of the overnight index swap OIS market provided an alternative way of measuring euro area risk-free rates taking a crucial role in this framework. Therefore, the yield measures chosen were no longer based on government bonds, but instead on OIS rates based on EONIA, which is the overnight unsecured interbank rate in the euro area, making them particularly informative from a monetary policy perspective.

In light of what has been said, we may consider the EONIA rate as the proxy for the risk-free short-term rate in the euro area. Therefore, we want to analyse its behaviour with the aim of outlining its fundamental and characterizing features in order to be able to define a short rate model as likely as possible. We will define this model in Chapter 2. To this purpose we need to study Figure [1.1] which depicts the current and historical values assumed by the EONIA and ECB deposit facility rates.

In Figure [1.1] we recognize that the ECB deposit facility rate (marked by the blue dotted line) provides a lower bound to the EONIA rate (marked by the yellow line). Indeed, the overnight money market rate does not fall below the deposit facility rate because it would not make sense for a bank to park its reserves overnight with another bank that pays an interest rate below the deposit facility rate. This sentence is the consequence of the fact that commercial banks generally hold deposits at the ECB that pays an interest on banks' excess reserves, that is the deposit facility rate. Therefore, the deposit facility rate is the overnight interest rate on reserves above the minimum

level paid by the ECB to a bank that deposits money with it, whereas EONIA expresses the weighted average of unsecured overnight interbank lending. This is the reason why EONIA does not fall below the ECB deposit facility rate that is its lower bound, therefore. The distance between these two rates depends on the monetary policy implementation and other macroeconomic factors. It was almost constant around 100 basis points until the onset of the financial crisis, whereas after 2008 the difference between the EONIA and ECB deposit facility rates has shrunk, and nowadays their values are very close with an average distance of 10 basis points.

Another empirical observation is that the lower bound (the ECB deposit facility rate) has a step behavior. We notice correlated jumps' successions: there is a persistence of growing phases, in which the deposit facility rate is moving upward in progression and similarly periods of negative decrease, as we experienced in recent years. The Figure 1 shows that since 2012 the ECB deposit rate has lowered to zero, and subsequently from June 2014 to negative values (it crossed the red line that represents the zero level). Even now the ECB deposit rate is negative.

Moreover, unlike the piecewise constant behaviour of its lower bound, we can observe that the EONIA rate is more variable and this is because it is an average decided every day by the main banks and therefore it is more susceptible to the market. Although EONIA is not directly influenced by the ECB, the change in the reference policy rates alters its trend. The Figure 1 shows frequent spikes in the EONIA rate but since 2012 its behaviour has changed quite substantially and the rate turned less volatile and stuck to its lower bound.

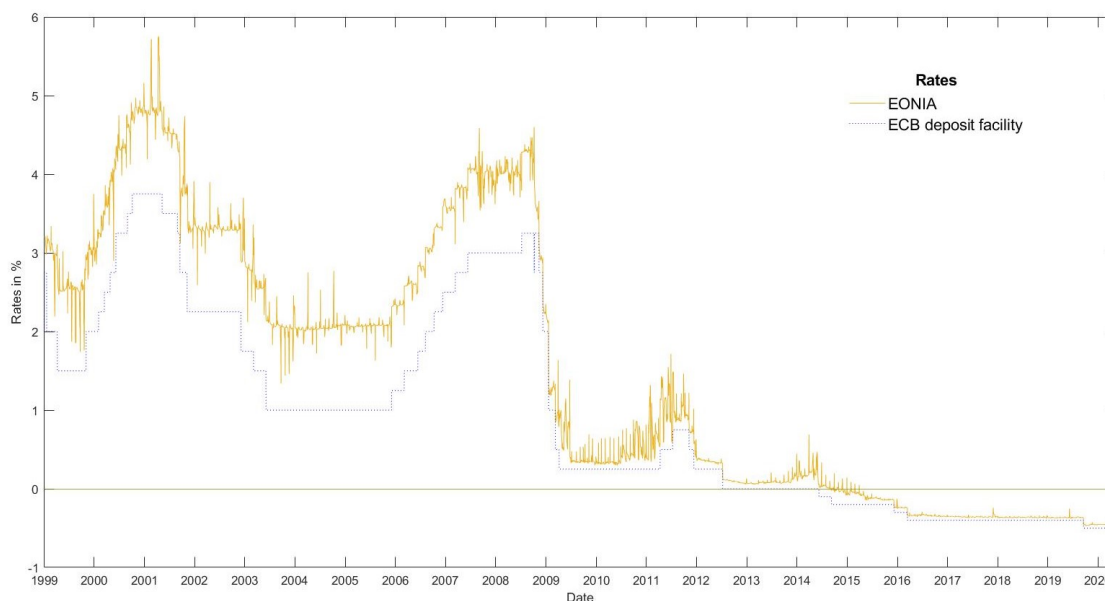


Figure 1.1: *Current and historical EONIA and ECB deposit facility rates.*
Historical data 04/01/1999 - 13/05/2020 from the website <http://sdw.ecb.europa.eu>

1.2 Literature review

Before 2008 interest rates were usually positive, or at most zero, and therefore economists viewed the zero lower bound on interest rates as unlikely to be relevant and thus it was not considered to be a constraint to develop term structure models. However, as analyzed in the previous section, the recent empirical evidence shows that this is not the case.

In [39] Jarrow studies this matter and concludes that the belief that there exists a zero lower bound on interest rates is wrong and consequently such a bound interferes negatively with the estimation of term structure of interest rates. The object of his paper is to explore the possible non-existence of a zero lower bound. To argue this he extends the Heath-Jarrow-Morton model to include the cash trading, i.e. cash is a traded asset which can be exchanged, developing a theoretical economy composed by consumers, firms, non-bank financial institutions and banks. He provides realistic constraints on the trading activity and proves that negative interest rates are consistent with an arbitrage-free environment.

Justified by the presence of this model and the recent empirical evidence reported in the previous section, we say that the zero lower bound does not necessarily reflect the reality because interest rates can take negative values. Then, we proceed by presenting the existing term structure models which may perform negative rates. We analyse the literature works that concern the European area and highlight their main features.

1.2.1 The SABR model

The SABR model was first introduced by Hagan et al. in [34] for a rate $(r_t)_{t \geq 0}$ and its volatility $(v_t)_{t \geq 0}$ satisfying the following stochastic differential equations (SDEs) under the risk-neutral probability measure:

$$\begin{aligned} dr_t &= r_t^\beta v_t dW_t^1, & r_0 &= \bar{r} \\ dv_t &= \gamma v_t dW_t^2, & v_0 &= \bar{v} \end{aligned}$$

with $(W_t^1)_{t \geq 0}$ and $(W_t^2)_{t \geq 0}$ Brownian motions with correlation $\mathbb{E}[dW_t^1 dW_t^2] = \rho dt$ and β, \bar{r}, γ and \bar{v} constants such that the power parameter satisfies $0 \leq \beta < 1$. In this model rates are assumed to be positive and the simplest way to take into account negative rates is to shift the SABR process with a constant strictly positive shift s :

$$\begin{aligned} dr_t &= (r_t + s)^\beta v_t dW_t^1 \\ dv_t &= \gamma v_t dW_t^2. \end{aligned}$$

This shifted model is characterized by a constant lower bound $-s$ which cannot change over time. Another limitation is represented by the absorbing property of the lower bound: upon reaching the value $-s$ the rate should stay there forever. However, these features are not realistic.

A different extension of the SABR model which does not require determining a shift

is presented by Antonov et al. in [5]. It is the free boundary SABR model:

$$\begin{aligned} dr_t &= |r_t|^\beta v_t dW_t^1, & r_0 &= \bar{r} \\ dv_t &= \gamma v_t dW_t^2, & v_0 &= \bar{v} \end{aligned}$$

for $0 \leq \beta < \frac{1}{2}$. This model can handle negative rates. A great advantage is the existence of a closed-form formula for the price of the European interest-rate options, the caplets as an example, in the special case of zero correlation between the Brownian motions ($\rho = 0$). For the general case an exact analytical solution to the pricing problem does not exist but an efficient approximation can be presented in terms of simple integrals.

1.2.2 The shadow-rate models

After the financial crisis economists have reconsidered models based on Gaussian processes that lead to the possibility of negative interest rates. Then, to set an appropriate lower bound for the rates these models started to be based on the shadow-rate concept, first introduced by Black in [13]. In these shadow-rate models the Gaussian dynamics drives the shadow short rate s_t while the short rate is defined as

$$r_t := \max\{s_t, LB_t\} \tag{1.1}$$

for a specific lower bound LB . That is, r_t equals the shadow rate if this is above the lower bound, while the short rate remains at the lower bound if the shadow rate is below the bound. This ensures that the short rate does not fall below the given lower bound.

In [48] Lemke & Vladu develop a discrete-time shadow-rate model to analyse the euro-area yield curve from 1999 to mid-2015. Under the risk-neutral probability measure three latent factors $(X_t)_{t \in \mathbb{N}} = ((X_t^1, X_t^2, X_t^3)^\top)_{t \in \mathbb{N}}$ follow a first-order Gaussian vector autoregressive process:

$$X_t = K_0 + K_1 X_{t-1} + \Sigma \epsilon_t \tag{1.2}$$

where $K_0 = (K_{01}, 0, 0)^\top$, K_1 is diagonal, Σ is lower triangular and $\epsilon_t \sim \mathcal{N}(0, 1_3)$. We mention that the superscript $^\top$ indicates transposition. The shadow short rate $(s_t)_{t \in \mathbb{N}}$ is specified as an affine function of factors:

$$s_t := X_t^1 + X_t^2 + X_t^3. \tag{1.3}$$

Then, the shadow rate can reach any positive or negative level. The short rate $(r_t)_{t \in \mathbb{N}}$ is given by (1.1) for a lower bound LB defined as a step function that assumes three different values: it is zero until the 2014 and then changes at two points in time, i.e. in May 2014 at -10 basis points and in September 2014 at -20 basis points.

If we assume that \mathbb{Q} is a risk-neutral probability measure, then we can define the price at time t of a zero-coupon bond that pays one unit at time $t + n$ as

$$P_{n,t} := \mathbb{E}^{\mathbb{Q}} \left[e^{-\sum_{i=0}^{n-1} r_{t+i}} \middle| \mathcal{F}_t \right]$$

and the bond yields for maturity n as

$$y_{n,t} := -\frac{1}{n} \log(P_{n,t}).$$

With the lower bound LB restriction for the short rate it is not possible to find explicit functions which map factors into bond yields for each maturity n . However, to allow the bond pricing Lemke & Vladu exploit the analytical approximation for the forward rate $f_{n,t} := (n+1)y_{n+1,t} - ny_{n,t}$ proposed by Wu & Xia in [60].

In [62] Wu & Xia develop a new shadow-rate model for which the short-term interest rate is defined by (1.1). Similar to the previous model their shadow rate is an affine function of three latent factors (1.3) whose risk-neutral dynamics follow a first-order Gaussian vector autoregressive process (1.2) for $K_0 = (0, 0, 0)^\top$:

$$X_t = K_1 X_{t-1} + \Sigma \epsilon_t.$$

Compared to the Lemke & Vladu's model the difference is represented by the stochastically time-varying lower bound $(LB_t)_{t \in \mathbb{N}}$. Since it is referred to the euro area, is assumed to take multiple of 10 basis points, its maximum value to be zero and its minimum -100 .

It is computed that the forward rate can be approximated by

$$f_{n,t} \approx \int g_n(l, X_t) \mathbb{Q}(LB_{t+n} = l | \mathcal{F}_t) dl \quad (1.4)$$

for a non-negative function g_n . We highlight that the formula prices taking account of the uncertainty of the future dynamics of the lower bound. In order to specify the risk-neutral probability $\mathbb{Q}(LB_{t+n} = l | \mathcal{F}_t)$ the authors introduce a sequence of random variables $(\Delta_t)_{t \in \mathbb{N}}$ that describes the direction in which the lower bound is moving, that is $\Delta_t = +1$ is the up state and the down state is represented by $\Delta_t = -1$. It follows that

$$\mathbb{Q}(LB_{t+n} = l | \mathcal{F}_t) = \mathbb{Q}(LB_{t+n} = l, \Delta_{t+n} = +1 | \mathcal{F}_t) + \mathbb{Q}(LB_{t+n} = l, \Delta_{t+n} = -1 | \mathcal{F}_t)$$

and therefore, with some computations, in the pricing formula (1.4) the integral is replaced by a finite sum allowing an analytical approximation for the bond prices.

In [61] Wu & Xia modify the previous shadow-rate model such that the lower bound $(LB_t)_{t \in \mathbb{N}}$ either stays where it is or moves down by 10 basis points, i.e.

$$\begin{aligned} \mathbb{Q}(LB_{t+1} = LB_t - 0.1 | \mathcal{F}_t) &= \alpha_{1,t} \\ \mathbb{Q}(LB_{t+1} = LB_t | \mathcal{F}_t) &= 1 - \alpha_{1,t}. \end{aligned}$$

To capture the forecast about the next moves of the lower bound, the authors introduce two binary random variables Δ_t and Δ_t^l referred to the immediate monetary policy and to the future one, respectively. If $\Delta_t = 1$ there is a high probability of a cut in an immediately following period, whereas $\Delta_t = 0$ implies that it is more likely the lower bound to be unchanged. On the other hand, $\Delta_t^l = 1$ indicates a high probability of a cut in longer horizons, whereas $\Delta_t^l = 0$ implies a lower possibility for future cuts. It follows

that in the pricing formula (1.4) the probability $\mathbb{Q}(LB_{t+n} = l | \mathcal{F}_t)$ can be replaced by the joint distributions of the lower bound LB and the Δ , Δ^l states:

$$\sum_{\substack{d=\pm 1 \\ dl=\pm 1}} \mathbb{Q}(LB_{t+n} = l, \Delta_{t+n} = d, \Delta_{t+n}^l = dl | \mathcal{F}_t).$$

Another discrete-time shadow-rate model is developed by Geiger & Schupp [31]. Similar to [48], it is assumed that the pricing factors $(X_t)_t = ((X_t^1, X_t^2, X_t^3)^\top)$ follow a first-order Gaussian vector autoregressive process (1.2). Then, the shadow rate $(s_t)_{t \in \mathbb{N}}$ is specified by (1.3) and the short rate $(r_t)_{t \in \mathbb{N}}$ by (1.1). The lower bound $(LB_t)_{t \in \mathbb{N}}$ is time-varying and equal to zero until July 2012 and afterward is defined by

$$LB_t := \gamma_t DFR_t + (1 - \gamma_t) DFR_{t+1} + sp_t$$

for the Deposit Facility rate $(DFR_t)_{t \in \mathbb{N}}$. The model is characterized by month-end observations and γ_t is the fraction of the month from t to $t + 1$ that stands for the calendar effect of the ECB's meets. The ECB does not meet at the end of the month and at most once a month, then γ_t represents the number of days between the end of the current month t and the next meeting date as a fraction of the month from t to $t + 1$. Lastly, sp_t is the spread between the EONIA and the Deposit Facility rate.

Analogously to this set-up, in [46] Kortela focuses on a continuous-time shadow-rate model that shapes the euro-area data from January 1999 to March 2016. Under the risk-neutral probability measure three latent factors $(X_t)_{t \geq 0}$ evolve as an Ornstein-Uhlenbeck process:

$$d \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} = - \begin{pmatrix} 0 & 0 & 0 \\ 0 & \phi & -\phi \\ 0 & 0 & \phi \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} + \Sigma dW_t,$$

where ϕ is a positive constant and $(W_t)_{t \geq 0}$ is a three-dimensional Brownian motion. The shadow rate is affine in the first two latent factors

$$s_t := X_t^1 + X_t^2,$$

and the short rate is defined by (1.1) where the lower bound LB is supposed to be exogenous. The authors introduce four different types of time-varying lower bounds. The first discussed is the minimum of the observed interest rates across maturities, that is $LB_t = \min\{R_t(\tau), 0\}$ where $R_t(\tau)$ is the observed OIS interest rate at time t for maturity τ . Another possibility is to choose the most negative yield in the sample up to t , that is $LB_t = \min\{\{R_u(\tau)\}_{u \leq t}, 0\}$. This lower bound is not increasing, i.e. $LB_t \leq LB_s$ for $s \leq t$. The third representation for the lower bound is an estimated sequence of lower bounds and the last specification is zero or the negative value given by the deposit facility rate, i.e. $LB_t = \min\{DFR_t, 0\}$.

Remark 1.1. The substantial disadvantage of the shadow-rate models is the lack of explicit formulas for bond prices, only analytical approximations are available. This Thesis's aim is to propose an extremely tractable model for which closed-form solutions for bond prices are computed.

1.3 An affine term structure model with stochastic lower bound

The purpose of this Thesis is to provide a term structure model that is able to accommodate negative short-term rates and then, following the ideas developed by Monfort et al. in [52], we want to define a stochastic lower bound *SLB* on interest rates which takes values in \mathbb{R} , and not only \mathbb{R}_+ . We do not define a possibly negative static lower bound but incorporate stochastic variations in the level of the interest rate bound to develop a better approximation of what happens in reality.

In order to get a better understanding of our continuous-time model described in detail in Chapter 2, we now provide a brief description of the term structure model of Monfort et al. [52]. We highlight that the authors of this paper are concerned with a discrete-time model, whereas one of the main objectives of this Thesis is its continuous-time generalization. The reference affine model discussed in [52] is characterized by a time-varying lower bound which can take negative values but, to study it we have first to refer to [51]. This article introduces the uni-variate gamma-zero distribution characterizing its features and extends it to the dynamic case with an affine process called autoregressive gamma-zero (ARG-Zero) and its multi-variate affine counterpart (VARG).

We have to give some important definitions and recall that a gamma distribution is a two-parameter family of positive probability distributions. The distribution $\gamma_\nu(\mu)$ is defined by a shape parameter $\nu > 0$ and a scale parameter $\mu > 0$. Its probability density function is given by

$$f(x; \nu, \mu) = \frac{e^{-\frac{x}{\mu}} x^{\nu-1}}{\Gamma(\nu) \mu^\nu} 1_{\{x>0\}}.$$

Because $\gamma_\nu(\mu)$ converges to the Dirac distribution at zero when ν goes to zero, the gamma distribution can be extended to the case $\nu = 0$ if $\gamma_0(\mu)$ is considered as the Dirac measure δ_0 . We also recall that the Poisson distribution is characterized by an intensity parameter $\lambda > 0$ and its probability density function is given by

$$f(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} 1_{\{x \in \mathbb{N}\}}.$$

If we consider a Poisson random variable Z , that is $Z \sim \mathcal{P}o(\lambda)$ for $\lambda > 0$, the non-centered gamma distribution $\gamma_\nu(\lambda, \mu)$ is a mixture of $\gamma_{\nu+Z}(\mu)$ distributions for $\nu > 0$ and $\mu > 0$.

Definition 1.2. *Let X be a non-negative random variable. We say that $X \sim \gamma_0(\lambda, \mu)$, that is X follows a gamma-zero distribution with parameters $\lambda > 0$, $\mu > 0$, if its conditional distribution given a random variable $Z \sim \mathcal{P}o(\lambda)$ is*

$$(X|Z) \sim \gamma_Z(\mu).$$

This definition is taken from [51]. We highlight that if $X \sim \gamma_0(\lambda, \mu)$, then $X = 0$ if and only if $Z = 0$, and $P(Z = 0) = e^{-\lambda} > 0$. Therefore, X equals zero with a strictly positive probability because the Poisson random variable Z equals zero with a strictly

positive probability and when $Z = 0$ the gamma distribution $\gamma_Z(\mu) = \gamma_0(\mu)$ is the Dirac distribution at zero. This is the key feature of the gamma-zero distribution, i.e. it has a point-mass at zero.

We now turn to the dynamic case where $(X_t)_{t \in \mathbb{N}}$ is a discrete-time random process that Monfort et al. call autoregressive gamma, denoted by $\text{ARG}_\nu(\alpha, \beta, \mu)$ for parameters α, β, ν and μ (see [51]).

Definition 1.3. $(X_t)_{t \in \mathbb{N}}$ is an $\text{ARG}_\nu(\alpha, \beta, \mu)$ process if for all $t \in \mathbb{N}$ the conditional distribution of X_{t+1} given $\underline{X}_t = (X_t, X_{t-1}, \dots)$ is a non-centered gamma distribution, that is

$$(X_{t+1} | \underline{X}_t) \sim \gamma_\nu(\alpha + \beta X_t, \mu),$$

for $\alpha \geq 0, \beta > 0, \nu \geq 0$ and $\mu > 0$.

In the reference article it is proved that the conditional Fourier transform of the $\text{ARG}_\nu(\alpha, \beta, \mu)$ process is exponential-affine in X_t for $t \in \mathbb{N}$. The process is called autoregressive gamma-zero if the parameter $\nu = 0$, and in this case it follows that, given a Poisson random variable $Z \sim \mathcal{P}o(\alpha + \beta X_t)$, the conditional distribution $((X_{t+1} | \underline{X}_t) | Z) \sim \gamma_Z(\mu)$ and therefore, as for the static gamma-zero distribution, the ARG_0 process has a zero-point mass. The previous definition can be generalized to the multi-variate case as follows (see [52]).

Definition 1.4. Fixed $d \in \mathbb{N}$, the \mathbb{R}^d -valued process $(X_t)_{t \in \mathbb{N}}$ is a $\text{VARG}_\nu(\alpha, \beta, \mu)$ process if for all $t \in \mathbb{N}$ its scalar components $(X_{t+1}^i | \underline{X}_t)$ are independent conditionally on \underline{X}_t and their conditionally distribution is the gamma distribution, that is

$$(X_{t+1}^i | \underline{X}_t) \sim \gamma_{\nu^i}(\alpha^i + \langle \beta^i, X_t \rangle, \mu^i),$$

for $i = 1, \dots, d$ and parameters $\alpha^i \geq 0, \nu^i \geq 0, \mu^i > 0$ and $\beta^i > \underline{0}$, where $\underline{0}$ is the d -dimensional null vector and the last inequality is componentwise.

We remark that in the notation of $\text{VARG}_\nu(\alpha, \beta, \mu)$ definition $\alpha := (\alpha^1, \dots, \alpha^d)$, $\nu := (\nu^1, \dots, \nu^d)$ and $\mu := (\mu^1, \dots, \mu^d)$ denote d -dimensional row vectors, whereas $\beta := (\beta^1, \dots, \beta^d)$ is a $(d \times d)$ matrix. Moreover, with $\langle \cdot, \cdot \rangle$ we denote the inner product in \mathbb{R}^d .

In Definition [1.4] the assumption of conditional independent components and the affine property that is deduced from $\text{ARG}_\nu(\alpha, \beta, \mu)$ definition make the conditional Fourier transform of the $\text{VARG}_\nu(\alpha, \beta, \mu)$ process an exponential-affine function in X_t for $t \in \mathbb{N}$, and therefore the model of Monfort et al. analytically tractable.

1.3.1 The dynamics under the physical probability measure P

We have defined the main mathematical ingredients that we are going to use in the construction of the affine term structure model proposed by Monfort et al. in [52], where the authors introduce the model specifying the involved processes' dynamics under the physical probability measure P .

First of all, let us to define an intensity process $(\lambda_t)_{t \in \mathbb{N}}$ that is an $\text{ARG}_0(\alpha, \beta, \mu)$ process, for parameters $\alpha \geq 0$, $\beta > 0$ and $\mu > 0$ that will be determined in the calibration phase. We remark that from this definition λ_t is non-negative for every $t \in \mathbb{N}$. Then, we introduce a stochastic process $(Y_t)_{t \in \mathbb{N}}$ such that Y_t is determined by the difference $\Delta Y_t = Y_t - Y_{t-1}$ that is in turn defined by the sum of two discrete processes

$$\Delta Y_t := \Delta N_t^+ - \Delta N_t^- \in \mathbb{N},$$

where $(\Delta N_t^+ | \lambda_t, \Delta N_{t-1}^+, \Delta N_{t-1}^-) \sim \mathcal{P}o(\lambda_t)$ whereas, after fixing a positive constant η , $(\Delta N_t^- | \lambda_t, \Delta N_{t-1}^+, \Delta N_{t-1}^-) \sim \mathcal{P}o(\eta Y_{t-1})$. We remark that it depends on ΔN_t^+ , and not only on ΔN_{t-1}^+ . Now, we can determine the stochastic lower bound as

$$SLB_t := -cY_t$$

where the constant c is a scale factor that is multiple of 10 basis points and employed in order to specify the SLB as a percentage quantity. We remark that from this construction it follows that $(\Delta N_t^+)_{t \in \mathbb{N}}$ is the process responsible of SLB decrease and, on the other hand, the force responsible of SLB increase is $(\Delta N_t^-)_{t \in \mathbb{N}}$.

We adopt the same notation of the reference article [52] and identify the process driving the SLB as $(\tilde{X}_t)_{t \in \mathbb{N}} = ((\lambda_t, Y_t)^\top)_{t \in \mathbb{N}}$. Then, we introduce a new process $(X_t)_{t \in \mathbb{N}}$ independent of the already defined $(\tilde{X}_t)_{t \in \mathbb{N}}$ and impose that $(X_t)_{t \in \mathbb{N}}$ is a 4-dimensional VARG process, i.e. for parameters $\alpha^i \geq 0$, $\beta^i > 0$, $\mu^i > 0$ and $\nu^i \geq 0$

$$\begin{aligned} (X_{t+1}^i | X_t) &\sim \gamma_0(\alpha^i + \langle \beta^i, X_t \rangle, \mu^i) \quad \text{for } i = 1, 2, \\ (X_{t+1}^i | X_t) &\sim \gamma_{\nu^i}(\alpha^i + \langle \beta^i, X_t \rangle, \mu^i) \quad \text{for } i = 3, 4. \end{aligned}$$

In other words, $(X_t)_{t \in \mathbb{N}}$ is composed by conditionally independent factors with gamma-zero and gamma distributions. Therefore, X_t^1 and X_t^2 can stay at zero for every $t \in \mathbb{N}$.

We denote the global process

$$(Z_t)_{t \in \mathbb{N}} := (X_t^\top, \tilde{X}_t^\top)_{t \in \mathbb{N}}^\top.$$

In [52] the risk-free rate between times t and $t+1$, and known at t , is denoted by r_t and it is assumed that it is given by

$$r_t := SLB_t + X_t^1 + X_t^2. \quad (1.5)$$

If we want to rewrite (1.5) in terms of the global process we obtain that

$$r_t = \langle \delta, Z_t \rangle \quad (1.6)$$

if $\delta = (1, 1, 0, 0, 0, -c)^\top$. We see that the interest rate process $(r_t)_{t \in \mathbb{N}}$ is a linear combination of $(Z_t)_{t \in \mathbb{N}}$.

We can conclude that the model proposed by Monfort et al. allows analytical tractability and, in the meantime, is flexible enough to match relevant empirical facts.

Indeed, recalling Figure [1.1](#) we can state that this model is appropriate to fit the observed features characterizing the EONIA and ECB deposit facility rates.

The model is described by the SLB which is proportional to $(Y_t)_{t \in \mathbb{N}}$ that is in turn the difference of two non-negative processes. Therefore, the SLB can handle negative values and is determined by a sequence of jumps that simulates the step behaviour of the ECB deposit facility rate. Furthermore, the interest rate defined by [\(1.5\)](#), or equivalently by [\(1.6\)](#), is a linear combination of the SLB and the first two components of $(X_t)_{t \in \mathbb{N}}$ which are determined by the gamma-zero distribution. It follows that the interest rate can never fall below the SLB.

The affine term structure model described in [\[52\]](#) is flexible and particularly efficient in matching the remarkable empirical features observed from the onset of the financial crisis that the already existing models could not explain. This justifies our interest in extending it to continuous time, and we will do it in Chapter 2.

Chapter 2

The model

In Chapter 1 we have described the properties that appeared from the description of the EONIA and ECB deposit facility rates and presented a discrete-time model which can provide negative rates and exhibit the other analysed features. Our aim is to extend it to a continuous-time framework proposing a mathematically appealing affine term structure model. Specifically, we want to be able to simultaneously match the following characteristics

- short rate can take negative values;
- time-varying (stochastically) lower bound.

We will develop a continuous-time generalization of the model of Monfort et al. [52] recalled at the end of the previous Chapter introducing an affine short rate model for the process $(r_t)_{t \geq 0}$ with respect to the physical probability measure P . Precisely, we will specify the short rate process as a linear combination of the stochastic lower bound $(SLB_t)_{t \geq 0}$, a process of macroeconomic factors $(X_t)_{t \geq 0}$ and a deterministic function. We will illustrate the analytical tractability of our model and following the results in [44] we will investigate the validity of the affine transform formula in terms of the solutions of the generalized Riccati ODEs which characterize our model. At the end of the Chapter we will study its structure under a different probability measure Q because we are interested in characterizing a parametrized family of probability measures Q which preserve its affine structure.

Regarding the notations used in this Chapter we refer to Appendix A.

2.1 The characterization of the affine short rate model

In this section we are going to introduce an affine term structure model characterized by two fundamental ingredients: a multivariate affine diffusion process $(X_t)_{t \geq 0}$ which represents the vector of macroeconomic variables and a bivariate point process $(N_t)_{t \geq 0}$ which models the stochastic lower bound SLB .

Let (Ω, \mathcal{F}, P) be a given probability space endowed with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ satisfying the *usual conditions*:

- right-continuity, i.e. $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s > t} \mathcal{F}_s$ for $t \geq 0$;
- completeness, i.e. \mathcal{F}_0 contains all P -null sets of \mathcal{F} .

In this presentation P denotes the physical probability measure. We suppose that all stochastic processes introduced in the following are defined on (Ω, \mathcal{F}, P) and adapted to the filtration \mathbb{F} .

Let $(W_t)_{t \geq 0}$ be a d -dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{F}, P)$, with $d \in \mathbb{N}$. We consider a stochastic process $(X_t)_{t \geq 0}$ taking values in \mathbb{R}_+^d whose dynamics is expressed by

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x_0. \quad (2.1)$$

We shall assume the given process may be very general representing a combination of observed and latent variables which capture salient macroeconomic features. The observed variables can be directly pointed out from market data, whereas the latent factors cannot be directly measured. They do not have a direct economic meaning but are related to the general economic trend, expectations, monetary policy shocks and others directly non-measurable concepts that play a meaningful role in describing the state of the economy and the determination of the term structure movements.

We assume that $(X_t)_{t \geq 0}$ is an affine diffusion process in the sense of Definition [A.10](#). This means that $(X_t)_{t \geq 0}$ is a Markov process and $\mu : \mathbb{R}_+^d \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}_+^d \rightarrow \mathbb{R}^{d \times d}$ are measurable functions such that for all $x \in \mathbb{R}_+^d$

- $\mu(x) = K_0 + K_1 x$ for $K_0 \in \mathbb{R}^d$ and $K_1 \in \mathbb{R}^{d \times d}$,
- $\sigma(x)\sigma(x)^\top = H(x)$ that is a $(d \times d)$ matrix with elements for $i, j = 1, \dots, d$
 $(\sigma(x)\sigma(x)^\top)_{ij} = H(x)_{ij} = \langle (H_1)_{ij}, x \rangle$ for $H_1 \in \mathbb{R}^{d \times d \times d}$.

We notice that the given definition of $H(x)$ differs from [\(A.14\)](#) for the H_0 term; in our framework $H_0 = \mathbf{0}$, where $\mathbf{0}$ denotes the $(d \times d)$ null matrix.

Assumption 2.1. *The parameters K_0 , K_1 and H_1 satisfy the following conditions:*

- $K_0 \in \mathbb{R}_+^d$;
- $K_1 \in \mathbb{R}^{d \times d}$ with $(K_1)_{ij} \geq 0$ for all $i, j = 1, \dots, d$ with $i \neq j$;
- H_1^k is a $(d \times d)$ symmetric and positive semi-definite matrix such that $(H_1^k)_{ij} = (H_1^k)_{ji} = 0$ for all $i, j = 1, \dots, d$ with $i \neq k$, for $k = 1, \dots, d$.

Remark 2.2. Assumption [2.1](#) concerns the admissibility conditions explained in Assumption [A.12](#). As a consequence, it is guaranteed that X_t is \mathbb{R}_+^d -valued P -a.s..

We can rewrite the SDE (2.1) as

$$dX_t = (K_0 + K_1 X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x_0.$$

As a direct consequence of Theorem A.14 there is no loss of generality in assuming that the volatility matrix is of the following simple form

$$H(x) = \begin{pmatrix} x_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & x_d \end{pmatrix}$$

for $x \in \mathbb{R}_+^d$. In other words, $H(x) = \text{diag}(x_1, \dots, x_d)$ that is the $(d \times d)$ diagonal matrix with diagonal elements x_1, \dots, x_d .

It follows that $(X_t)_{t \geq 0}$ is specified by the SDE

$$dX_t = (K_0 + K_1 X_t)dt + \sqrt{\text{diag}(X_t)}dW_t, \quad X_0 = x_0. \quad (2.2)$$

The notation $\sqrt{\text{diag}(x)}$ is to be understood componentwise, i.e. it is a diagonal matrix with elements $(\sqrt{\text{diag}(x)})_{ii} = \sqrt{x_i}$ for $i = 1, \dots, d$. We highlight that the matrix $\sqrt{\text{diag}(X_t)}$ is well-defined for all $t \geq 0$ because $(X_t)_{t \geq 0}$ takes values in \mathbb{R}_+^d .

Similar to [52], we define the stochastic lower bound process $(SLB_t)_{t \geq 0}$ as a difference of two point processes responsible of its piecewise constant behaviour. To this aim we introduce $(N_t)_{t \geq 0} = ((N_t^1, N_t^2)^\top)_{t \geq 0}$ that is a bivariate counting process which models the sequence of jumps of the stochastic lower bound. From Definition A.3 it follows that $(N_t^a)_{t \geq 0}$ is a simple point process, for $a = 1, 2$, and the possibility of simultaneous occurrence of two jumps is ruled out, i.e.

$$\Delta N_t^1 \Delta N_t^2 = 0 \quad P\text{-a.s.} \quad \text{for every } t \geq 0.$$

We recall that $\Delta N_t^a = N_t^a - N_{t-}^a$ is the jump of the a -th point process at time $t \geq 0$ with the convention that $\Delta N_0^a = 0$, as it is defined in Appendix A.

We intend to uniquely determine the counting process $(N_t)_{t \geq 0}$ through its intensity vector $(\lambda_t)_{t \geq 0} = ((\lambda_t^1, \lambda_t^2)^\top)_{t \geq 0}$. We use the Hawkes processes setting as a starting point to characterize it because of the self-excitement property of Hawkes processes, meaning that each jump increases the likelihood of future jumps. For $a = 1, 2$ $(\lambda_t^a)_{t \geq 0}$ is defined by

$$\lambda_t^a := \lambda_0^a + \langle \Lambda_a, X_t \rangle + \delta_a \int_0^t e^{-\gamma_a(t-u)} dN_u^a \quad \text{for } t \geq 0, \quad (2.3)$$

for $\lambda_0^a \in \mathbb{R}_{++}$, $\delta_a, \gamma_a \in \mathbb{R}_+$ and $\Lambda_a \in \mathbb{R}_+^d$. This is well-defined according to (A.9).

The likelihood of future jumps directly depends on the occurrence of past jumps of the same component. In other words, (2.3) describes how jumps of the first process component $(N_t^1)_{t \geq 0}$ influence the intensity only of the first component, and analogously for the second component. This positive influence of the past jumps on the current value of intensity is the fundamental characteristic deriving from Hawkes processes' framework

but it occurs only if $\delta_a > 0$, for $a = 1, 2$. The jump intensity component $(\lambda_t^a)_{t \geq 0}$ is determined by jumps in the past with associated weights $\delta_a e^{-\gamma_a(\cdot)}$; the parameter δ_a allows jumps' clusters and, on the other hand, the coefficient γ_a accounts for a decay effect, in the sense that past jumps have a diminishing effect on the intensity over time.

This description reflects the historical trend of the ECB deposit facility rate and capture its persistence (see Figure 1.1). In fact, we have observed that when a broad downward trend occurs in the ECB deposit facility rate it is more likely to keep moving down than up, and vice-versa if it is moving upward.

The stochastic intensity vector defined by (2.3) depends also on the d -variate process $(X_t)_{t \geq 0}$ of macroeconomic factors, thus generalizing the classical intensity specification used in the context of Hawkes processes.

We assume that the two components of the process $(N_t)_{t \geq 0}$ have jumps of unit size. In other words, $\Delta N_t^a \in \{0, 1\}$ for all $t \geq 0$ and $a = 1, 2$.

From (2.3) we deduce that the intensity process $(\lambda_t^a)_{t \geq 0}$ satisfies the SDE

$$\begin{aligned} d\lambda_t^a &= \delta_a dN_t^a - \gamma_a \delta_a \left(\int_0^t e^{-\gamma_a(t-u)} dN_u^a \right) dt + \sum_{i=1}^d (\Lambda_a)_i dX_t^i \\ &= \delta_a dN_t^a + \gamma_a \left(\lambda_0^a + \langle \Lambda_a, X_t \rangle - \lambda_t^a \right) dt + \langle \Lambda_a, dX_t \rangle \\ &= \left[\gamma_a \left(\lambda_0^a + \langle \Lambda_a, X_t \rangle - \lambda_t^a \right) + \langle \Lambda_a, (K_0 + K_1 X_t) \rangle \right] dt \\ &\quad + \delta_a dN_t^a + \langle \sqrt{\text{diag}(X_t)} \Lambda_a, dW_t \rangle. \end{aligned}$$

Recalling Definition A.3 the counting process is associated with a simple point process $(T_n)_{n \geq 0}$ which represents times when $(N_t)_{t \geq 0}$ jumps and a sequence of $\{1, 2\}$ -valued random variables $(Z_n)_{n \geq 0}$ that indicates which component jumps. Now, to simplify the notation used in the following we denote the jump times of the counting component $(N_t^a)_{t \geq 0}$ as $(T_n^a)_{n \geq 0}$, for $a = 1, 2$.

We define the stochastic lower bound process as follows

$$SLB_t := c(N_t^1 - N_t^2) \quad \text{for all } t \geq 0, \quad (2.4)$$

where $c > 0$ is a fixed scale factor and it is used to specify the SLB as a percentage quantity. For example, if we take the ECB deposit facility rate as reference for the SLB then, we have to set $c = 0.1\%$ since the deposit facility rate can be changed by multiples of 10 basis points.

Adopting the definition of short rate process given in the model of [52] and analyzed in (1.5) we define $(r_t)_{t \geq 0}$ by

$$r_t := h(t) + SLB_t + \langle \ell, X_t \rangle \quad \text{for all } t \geq 0, \quad (2.5)$$

where $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ is an integrable function, i.e. $\int_0^t h(u) du < \infty$, and $\ell \in \mathbb{R}_+^d$. We observe that $(r_t)_{t \geq 0}$ linearly depends on the non-negative process $(X_t)_{t \geq 0}$ and the counting process $(N_t)_{t \geq 0}$ that models the SLB . The deterministic function h is introducing to reproduce any observed yield curve, proceeding similarly as in Brigo & Mercurio [17].

Market data have to be used in order to fit exactly the initial term structure of interest rates and, therefore, determine the function h . We denote the theoretical price at time t of a zero coupon bond with maturity date T_m by $p(t, T_m)$ and assume that $p(t, T_m) = \Pi(t, T_m, X_t, N_t^1, N_t^2)$ for a smooth function $\Pi : \mathbb{R}_+^{d+2} \times \mathbb{N}^2 \rightarrow \mathbb{R}_+$. We also suppose that the term structure of discount factors which is currently observed in the market is given by the smooth function $p^M(0, \cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that $t \rightarrow p^M(0, t)$. The initial term structure matches the term structure of interest rate observed in the market if and only if $p^M(0, t) = \Pi(0, t, X_0, 0, 0)$ for all $t \geq 0$ and from this equality we can determine the function h . We will analyse the expression taken by the function h in a next section after the study of change of probability measure, from the physical P to a martingale measure Q .

Remark 2.3. The function h can take negative values and therefore for some $t \geq 0$ the short rate r_t may fall below the lower bound SLB_t according to (2.5). We can enforce its positivity but this appears to be too restrictive since h is determined by market data.

If h is differentiable, then $(r_t)_{t \geq 0}$ satisfies the SDE

$$\begin{aligned} dr_t &= h'(t)dt + c(dN_t^1 - dN_t^2) + \langle \ell, dX_t \rangle \\ &= \left[h'(t) + \langle \ell, (K_0 + K_1 X_t) \rangle \right] dt + c(dN_t^1 - dN_t^2) + \langle \sqrt{\text{diag}(X_t)} \ell, dW_t \rangle, \end{aligned}$$

where $'$ represents the derivative with respect to time t .

Summing up, our short rate model is characterized by an affine diffusion process $(X_t)_{t \geq 0}$ and a stochastic lower bound $(SLB_t)_{t \geq 0}$, and therefore it can be completely described by the following system

$$\begin{cases} dX_t = (K_0 + K_1 X_t)dt + \sqrt{\text{diag}(X_t)}dW_t \\ SLB_t = c(N_t^1 - N_t^2) \\ r_t = h(t) + SLB_t + \langle \ell, X_t \rangle. \end{cases}$$

This model relies on several random factors to capture the potential variability of the market term structure and can generate negative rates which are bounded from below by the SLB that is not constant, but can vary over time.

We now prove that the model belongs to the class of affine jump-diffusion short rate models. The analytical tractability of the model will be ensured by this fundamental property.

Theorem 2.4. *Let $(X_t)_{t \geq 0}$ be a d -dimensional affine diffusion which takes values in \mathbb{R}_+^d whose dynamic is determined by (2.2) and $(N_t)_{t \geq 0}$ a bivariate point process uniquely determined by the intensity process $(\lambda_t)_{t \geq 0}$ satisfying (2.3). We specify the short rate process $(r_t)_{t \geq 0}$ by (2.5), that is*

$$r_t = h(t) + c(N_t^1 - N_t^2) + \langle \ell, X_t \rangle$$

for $c > 0$, $\ell \in \mathbb{R}_+^d$ and $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ a deterministic function such that $\int_0^t h(u)du < \infty$. Then, this continuous-time model is an affine jump-diffusion short rate model.

The proof of Theorem [2.4](#) is fundamental because it introduces the main concepts that we use in the whole Thesis.

Proof. To prove the statement we construct an affine jump-diffusion process $(Z_t)_{t \geq 0}$ in an enlarged state space following the idea developed in [\[27\]](#) (Proposition 4.4).

To this effect, for $a = 1, 2$ let us define the auxiliary process $(\tilde{\lambda}_t^a)_{t \geq 0}$ by

$$\tilde{\lambda}_t^a := \lambda_0^a + \delta_a \int_0^t e^{-\gamma_a(t-u)} dN_u^a \quad \text{for all } t \geq 0. \quad (2.6)$$

It holds that

$$\begin{aligned} d\tilde{\lambda}_t^a &= \delta_a dN_t^a - \gamma_a \delta_a \int_0^t e^{-\gamma_a(t-u)} dN_u^a dt \\ &= \delta_a dN_t^a - \gamma_a \left(\delta_a \int_0^t e^{-\gamma_a(t-u)} dN_u^a + \lambda_0^a \right) dt + \gamma_a \lambda_0^a dt \\ &= \gamma_a (\lambda_0^a - \tilde{\lambda}_t^a) dt + \delta_a dN_t^a. \end{aligned}$$

Now, we define the $(d + 4)$ -dimensional process $(Z_t)_{t \geq 0}$ by

$$Z_t := \begin{pmatrix} X_t \\ \tilde{\lambda}_t^1 \\ \tilde{\lambda}_t^2 \\ N_t^1 \\ N_t^2 \end{pmatrix} \quad \text{for all } t \geq 0. \quad (2.7)$$

We aim at verifying that $(Z_t)_{t \geq 0}$ is an affine jump-diffusion according to Definition [A.10](#). This is sufficient to prove the Theorem because, clearly, the short rate is a time-dependent affine function of $(Z_t)_{t \geq 0}$ since

$$r_t = h(t) + \left\langle \begin{pmatrix} \ell \\ 0 \\ 0 \\ c \\ -c \end{pmatrix}, \begin{pmatrix} X_t \\ \tilde{\lambda}_t^1 \\ \tilde{\lambda}_t^2 \\ N_t^1 \\ N_t^2 \end{pmatrix} \right\rangle.$$

Obviously $(Z_t)_{t \geq 0}$ is a Markov process by construction. In fact, the process $(X_t)_{t \geq 0}$ is Markov by definition and also $(N_t, \tilde{\lambda}_t)_{t \geq 0} = ((N_t^1, N_t^2, \tilde{\lambda}_t^1, \tilde{\lambda}_t^2)^\top)_{t \geq 0}$ is a Markov process (see [\[10\]](#) (Proposition 2)). From the independence of these two processes it follows that $(Z_t)_{t \geq 0}$ is Markov.

We can represent the process $(Z_t)_{t \geq 0}$ as a solution to the following SDE

$$dZ_t = \bar{\mu}(Z_t) dt + \bar{\sigma}(Z_t) d\bar{W}_t + \sum_{a=1}^2 \zeta_a dJ_t^a, \quad Z_0 = z_0 := (x_0^\top, \lambda_0^1, \lambda_0^2, 0, 0)^\top. \quad (2.8)$$

With $(\bar{W}_t)_{t \geq 0}$ we denote the \mathbb{R}^{d+4} -valued Brownian motion $((W_t^\top, W_t^1, W_t^2, W_t^3, W_t^4)^\top)_{t \geq 0}$ which is composed by $(W_t)_{t \geq 0}$, i.e. the d -dimensional Brownian motion that we have already introduced in (2.1), and four independent uni-dimensional Brownian motions $(W_t^i)_{t \geq 0}$ for $i = 1, \dots, 4$ (see [19] (Proposizione 2.30)). Then, with $(J_t^a)_{t \geq 0}$ we identify the \mathbb{N}^{d+4} -valued pure jump process composed by $(d+4)$ identical point processes $(N_t^a)_{t \geq 0}$, for $a = 1, 2$. We remark that this construction is coherent with Definition A.8

Now, we have to correctly determine the coefficients which appear in (2.8). We use similar symbols for the coefficients of $(Z_t)_{t \geq 0}$ and $(X_t)_{t \geq 0}$ to avoid further notation.

For all $z \in \mathbb{R}_+^{d+2} \times \mathbb{N}^2$ we recognize that the drift coefficient is

$$\bar{\mu}(z) = \bar{K}_0 + \bar{K}_1 z = \begin{pmatrix} K_0 \\ \gamma_1 \lambda_0^1 \\ \gamma_2 \lambda_0^2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} K_1 & 0 & 0 & 0 & 0 \\ 0 & -\gamma_1 & 0 & 0 & 0 \\ 0 & 0 & -\gamma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} z,$$

whereas the volatility matrix is defined by

$$\bar{\sigma}(z) \bar{\sigma}(z)^\top = \bar{H}(z) = \begin{pmatrix} \text{diag}((z_1, \dots, z_d)^\top) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We highlight that in the matrices' writing 0 indicates both the null vector (column or row) in \mathbb{R}^d and the scalar, but the difference is sufficiently clear from the held position. Following the notation of Definition A.10 we deduce that \bar{H}_0 is the $((d+4) \times (d+4))$ null matrix, i.e. $\bar{H}_0 = \mathbf{0}$, and the only non-null elements of the volatility tensor \bar{H}_1 are $(\bar{H}_1)_{i,i} = \epsilon(i)$ for $i = 1, \dots, d$, where $\epsilon(i)$ denotes the $(d+4)$ -dimensional vector whose i -th component is 1 and all the other are equal to 0.

Looking at (2.8) the jump matrices in $\mathbb{R}^{(d+4) \times (d+4)}$ are defined by

$$\zeta_1 = \begin{pmatrix} \mathbf{0} & 0 & 0 & 0 & 0 \\ 0 & \delta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \zeta_2 = \begin{pmatrix} \mathbf{0} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.9)$$

The intensities of the jump processes $(J_t^1)_{t \geq 0}$ and $(J_t^2)_{t \geq 0}$ are

$$\begin{aligned} \lambda^1(z) &= l_0^1 + \langle l_1^1, z \rangle = 0 + \left\langle \begin{pmatrix} \Lambda_1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, z \right\rangle, \\ \lambda^2(z) &= l_0^2 + \langle l_1^2, z \rangle = 0 + \left\langle \begin{pmatrix} \Lambda_2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, z \right\rangle. \end{aligned} \tag{2.10}$$

Moreover, their probability jump distributions are

$$\nu^1 = (\delta_{\{1\}}^\top, \delta_{\{1\}}, \delta_{\{1\}}, \delta_{\{1\}}, \delta_{\{1\}})^\top = \nu^2, \tag{2.11}$$

where $\delta_{\{1\}}$ represents the Dirac measure with point mass at 1.

To conclude the proof we need to prove that the coefficients $(\bar{K}_0, \bar{K}_1, \bar{H}_0, \bar{H}_1, l_0, l_1, \nu)$ satisfy the admissibility conditions of Assumption [A.12](#), where we identify $m = d + 4$ and consequently $I = \{1, \dots, d + 4\}$ and $J = \emptyset$. It is an easy check:

- $\bar{K}_0 \in \mathbb{R}_+^{d+4}$ because of $K_0 \in \mathbb{R}_+^d$, $\gamma_1, \gamma_2 \in \mathbb{R}_+$ and $\lambda_0^1, \lambda_0^2 \in \mathbb{R}_{++}$;
- $(\bar{K}_1)_{ij} \in \mathbb{R}_+$ for $i, j = 1, \dots, d + 4$ with $i \neq j$ because $(K_1)_{ij} \in \mathbb{R}_+$ for $i, j = 1, \dots, d$ with $i \neq j$ and the other non-diagonal elements are null;
- \bar{H}_0 is a symmetric and positive-semi-definite matrix because trivially $\bar{H}_0 = \mathbf{0}$;
- for every $k = 1, \dots, d$ \bar{H}_1^k is the $((d + 4) \times (d + 4))$ matrix such that the only non-null element is $(\bar{H}_1^k)_{kk} = 1$, it is symmetric and positive semi-definite because every principal minor is non-negative and evidently for $i, j = 1, \dots, d + 4$ with $i \neq k$ $(\bar{H}_1^k)_{ij} = (\bar{H}_1^k)_{ji} = 0$;
- for every $k = d + 1, \dots, d + 4$ $\bar{H}_1^k = \mathbf{0}$, it is trivially symmetric and positive semi-definite and $(\bar{H}_1^k)_{ij} = (\bar{H}_1^k)_{ji} = 0$ for $i, j = 1, \dots, d + 4$ with $i \neq k$;
- for every $a = 1, 2$ $l_0^a \in \mathbb{R}_+$ because it is simply null;
- for every $a = 1, 2$ $l_1^a \in \mathbb{R}_+^{d+4}$ because $\Lambda_a \in \mathbb{R}_+^d$ by definition;
- for every $a = 1, 2$ $(\zeta^a)_{i,i} \in \mathbb{R}_+$ for $i = 1, \dots, d + 4$ because the only non-null diagonal elements are $\delta_a \in \mathbb{R}_+$ and 1.

Therefore, we have proved that the constructed process $(Z_t)_{t \geq 0}$ is an affine jump-diffusion. We can conclude that our continuous-time model belongs to the class of affine jump-diffusion short rate models, as considered in Appendix A (Section A.3). \square

To complete the description of the modeling framework we have to compute the jump transforms and the parameters defining the short rate process. We start by computing the jump transforms for the probability jump distributions ν^1 and ν^2 , that is

$$\theta^a(c) = \int_{\mathbb{N}^{d+4}} e^{\langle c, u \rangle} d\nu^a(u) \quad \text{for } c \in \mathbb{C}^{d+4}, \quad a = 1, 2.$$

From (2.11) we identify $\nu^1 = \nu = \nu^2$ and it follows that for every $c \in \mathbb{C}^{d+4}$

$$\theta^1(c) = \theta^2(c) = \int_{\mathbb{N}^{d+4}} e^{\langle c, u \rangle} d\nu(u) = e^{\sum_{i=1}^{d+4} c_i}. \quad (2.12)$$

If we express the short rate process $(r_t)_{t \geq 0}$ defined in (2.5) as a function of $(Z_t)_{t \geq 0}$, then this relation is determined by the parameters

$$\rho_0(t) = h(t) \quad \text{and} \quad \rho_1(t) = \rho_1 = \begin{pmatrix} \ell \\ 0 \\ 0 \\ c \\ -c \end{pmatrix}. \quad (2.13)$$

Indeed, from (2.5) we obtain that

$$\begin{aligned} r_t &= h(t) + c(N_t^1 - N_t^2) + \langle \ell, X_t \rangle \\ &= h(t) + \left\langle \begin{pmatrix} \ell \\ 0 \\ 0 \\ c \\ -c \end{pmatrix}, \begin{pmatrix} X_t \\ \tilde{\lambda}_t^1 \\ \tilde{\lambda}_t^2 \\ N_t^1 \\ N_t^2 \end{pmatrix} \right\rangle = \rho_0(t) + \langle \rho_1, Z_t \rangle. \end{aligned} \quad (2.14)$$

2.2 The affine transform formula

In this section we aim at characterizing the discounted Laplace transform of the affine processes $(Z_t)_{t \geq 0}$. We follow the approach developed by Duffie et al. [24] which is reviewed in detail in Appendix A.

We recall that (see equation (A.24) in the appendix)

$$\Psi(u, Z_t, t, T) = \mathbb{E} \left[e^{-\int_t^T r(Z_s, s) ds} e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t \right] \quad \text{for } u \in \mathbb{C}^{d+4}. \quad (2.15)$$

In Appendix A we show that under integrability conditions this transform is easy to compute because it becomes an exponential affine expression on Z_t and it is explicitly known up to the solution of some ODEs, precisely a generalized Riccati system. In some cases the solution of this system is known exactly, whereas in other cases one has to resort to suitable numerical schemes, such as the Runge-Kutta method.

We state the next proposition to find the characterizing ODEs system related to the affine process $(Z_t)_{t \geq 0}$.

Proposition 2.5. *Let $(u, T) \in \mathbb{C}^{d+4} \times \mathbb{R}_+$ and suppose that $\chi^d = (\bar{K}, \bar{H}, l^1, l^2, \theta^1, \theta^2, \rho)$ are well-behaved at (u, T) (see Definition [A.15](#)). Then, the transform Ψ defined in [\(2.15\)](#) is given by*

$$\Psi(u, Z_t, t, T) = e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle},$$

where $\phi \in \mathbb{C}$ and $\psi \in \mathbb{C}^{d+4}$ solve the following generalized Riccati system:

$$\begin{aligned} \frac{d\phi}{dt}(t, T, u) &= h(t) - \langle (K_0^\top, \gamma_1 \lambda_0^1, \gamma_2 \lambda_0^2, 0, 0)^\top, \psi(t, T, u) \rangle \\ \phi(T, T, u) &= 0, \end{aligned} \quad (2.16)$$

$$\begin{aligned} \frac{d\psi}{dt}(t, T, u) &= \begin{pmatrix} \ell \\ 0 \\ 0 \\ c \\ -c \end{pmatrix} + \begin{pmatrix} -K_1^\top \psi|_d(t, T, u) \\ \gamma_1 \psi_{d+1}(t, T, u) \\ \gamma_2 \psi_{d+2}(t, T, u) \\ 0 \\ 0 \end{pmatrix} \\ &\quad - \frac{1}{2} \begin{pmatrix} (\psi_1^2(t, T, u), \dots, \psi_d^2(t, T, u))^\top \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ &\quad - \begin{pmatrix} \Lambda_1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \left(e^{\delta_1 \psi_{d+1}(t, T, u) + \psi_{d+3}(t, T, u)} - 1 \right) \\ &\quad - \begin{pmatrix} \Lambda_2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \left(e^{\delta_2 \psi_{d+2}(t, T, u) + \psi_{d+4}(t, T, u)} - 1 \right) \\ \psi(T, T, u) &= u. \end{aligned} \quad (2.17)$$

With the notation $\psi|_d(t, T, u)$ we denote the restriction of $\psi(t, T, u) \in \mathbb{C}^{d+4}$ to the first d components, that is $(\psi_1(t, T, u), \dots, \psi_d(t, T, u))^\top$.

We observe that [\(2.16\)](#) is a trivial ODE. Indeed, if ψ denotes the solution to [\(2.17\)](#) then the function ϕ can be obtain by direct integration, that is

$$\phi(t, T, u) = - \int_t^T \left(h(s) - \langle (K_0^\top, \gamma_1 \lambda_0^1, \gamma_2 \lambda_0^2, 0, 0)^\top, \psi(s, T, u) \rangle \right) ds. \quad (2.18)$$

Proof. This proposition is a consequence of Theorem [A.16](#). It suffices to verify that the system defined previously by [\(2.16\)](#) and [\(2.17\)](#) coincides with that identified by equations [\(A.26\)](#) and [\(A.27\)](#), where we have to replace the coefficients of the model.

For brevity of notation we omit the dependencies on T and u of the functions ϕ and ψ and, with some abuse of notation, we denote both the scalar and the null vector (row or column) of \mathbb{R}^d by $\mathbf{0}$.

In our modeling framework the ODE (A.26) becomes:

$$\begin{aligned} \frac{d\phi}{dt}(t) &= \rho_0(t) - \langle \bar{K}_0, \psi(t) \rangle - \frac{1}{2} \langle \psi(t), \bar{H}_0 \psi(t) \rangle - \sum_{a=1}^2 l_0^a [\theta^a(\zeta_a \psi(t)) - 1] \\ &= h(t) - \left\langle \begin{pmatrix} K_0 \\ \gamma_1 \lambda_0^1 \\ \gamma_2 \lambda_0^2 \\ 0 \\ 0 \end{pmatrix}, \psi(t) \right\rangle - \frac{1}{2} \langle \psi(t), \mathbf{0} \psi(t) \rangle - 0 \\ &= h(t) - \left\langle \begin{pmatrix} K_0 \\ \gamma_1 \lambda_0^1 \\ \gamma_2 \lambda_0^2 \\ 0 \\ 0 \end{pmatrix}, \psi(t) \right\rangle, \end{aligned}$$

where $\mathbf{0}$ denotes the null matrix of dimension $((d+4) \times (d+4))$, as usual.

Before analyzing directly (A.27) we have to do some work. First of all

$$\zeta_1 \psi(t) = \begin{pmatrix} \mathbf{0} & 0 & 0 & 0 & 0 \\ 0 & \delta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \psi|_d(t) \\ \psi_{d+1}(t) \\ \psi_{d+2}(t) \\ \psi_{d+3}(t) \\ \psi_{d+4}(t) \end{pmatrix} = \begin{pmatrix} \underline{\mathbf{0}} \\ \delta_1 \psi_{d+1}(t) \\ 0 \\ \psi_{d+3}(t) \\ 0 \end{pmatrix},$$

where $\underline{\mathbf{0}}$ is the d -dimensional null vector. From this equivalence we obtain that

$$\theta^1(\zeta_1 \psi(t)) = \theta^1((0, \delta_1 \psi_{d+1}(t), 0, \psi_{d+3}(t), 0)^\top) = e^{\delta_1 \psi_{d+1}(t) + \psi_{d+3}(t)}. \quad (2.19)$$

Analogously from

$$\zeta_2 \psi(t) = \begin{pmatrix} \mathbf{0} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \psi|_d(t) \\ \psi_{d+1}(t) \\ \psi_{d+2}(t) \\ \psi_{d+3}(t) \\ \psi_{d+4}(t) \end{pmatrix} = \begin{pmatrix} \underline{\mathbf{0}} \\ 0 \\ \delta_2 \psi_{d+2}(t) \\ 0 \\ \psi_{d+4}(t) \end{pmatrix},$$

it follows that

$$\theta^2(\zeta_2 \psi(t)) = \theta^2((0, 0, \delta_2 \psi_{d+2}(t), 0, \psi_{d+4}(t))^\top) = e^{\delta_2 \psi_{d+2}(t) + \psi_{d+4}(t)}. \quad (2.20)$$

Therefore we can compute

$$\begin{aligned}
\frac{d\psi}{dt}(t) &= \rho_1 - \bar{K}_1^\top \psi(t) - \frac{1}{2} \psi(t)^\top \bar{H}_1 \psi(t) - \sum_{a=1}^2 l_1^a [\theta^a(\zeta_a \psi(t)) - 1] \\
&= \begin{pmatrix} \ell \\ 0 \\ 0 \\ c \\ -c \end{pmatrix} - \begin{pmatrix} K_1^\top & 0 & 0 & 0 & 0 \\ 0 & -\gamma_1 & 0 & 0 & 0 \\ 0 & 0 & -\gamma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \psi|_d(t) \\ \psi_{d+1}(t) \\ \psi_{d+2}(t) \\ \psi_{d+3}(t) \\ \psi_{d+4}(t) \end{pmatrix} \\
&\quad - \frac{1}{2} \begin{pmatrix} \sum_{i,j=1}^{d+4} \psi_i(t) (\bar{H}_1^1)_{ij} \psi_j(t) \\ \dots \\ \sum_{i,j=1}^{d+4} \psi_i(t) (\bar{H}_1^d)_{ij} \psi_j(t) \\ \sum_{i,j=1}^{d+4} \psi_i(t) (\bar{H}_1^{d+1})_{ij} \psi_j(t) \\ \dots \\ \sum_{i,j=1}^{d+4} \psi_i(t) (\bar{H}_1^{d+4})_{ij} \psi_j(t) \end{pmatrix} \\
&\quad - \begin{pmatrix} \Lambda_1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} [\theta^1(\zeta_1 \psi(t)) - 1] - \begin{pmatrix} \Lambda_2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} [\theta^2(\zeta_2 \psi(t)) - 1].
\end{aligned}$$

From the proof of Theorem 2.4 we see that for $k = 1, \dots, d$ the only non-null elements of the $((d+4) \times (d+4))$ matrix \bar{H}_1^k are $(\bar{H}_1^k)_{kk} = 1$. While \bar{H}_1^k is the null matrix for $k = d+1, \dots, d+4$. Therefore

$$\begin{aligned}
\frac{d\psi}{dt}(t) &= \begin{pmatrix} \ell \\ 0 \\ 0 \\ c \\ -c \end{pmatrix} - \begin{pmatrix} K_1^\top \psi|_d(t) \\ -\gamma_1 \psi_{d+1}(t) \\ -\gamma_2 \psi_{d+2}(t) \\ 0 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \psi_1(t)^2 \\ \dots \\ \psi_d(t)^2 \\ 0 \\ \dots \\ 0 \end{pmatrix} \\
&\quad - \begin{pmatrix} \Lambda_1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} [e^{\delta_1 \psi_{d+1}(t) + \psi_{d+3}(t)} - 1] - \begin{pmatrix} \Lambda_2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} [e^{\delta_2 \psi_{d+2}(t) + \psi_{d+4}(t)} - 1],
\end{aligned}$$

thus proving (2.17). □

We are interested in investigating the meaning of *well-behaved* for the discounted characteristics $\chi^d = (\bar{K}, \bar{H}, l^1, l^2, \theta^1, \theta^2, \rho)$ at $(u, T) \in \mathbb{C}^{d+4} \times \mathbb{R}_+$. According to Definition A.15 χ^d are *well-behaved* if the generalized Riccati system defined by (2.16) and (2.17) is solved uniquely and if for $0 \leq t \leq T$ the following integrability conditions hold:

1. $\mathbb{E}[|\kappa_t|] < \infty$ for $\kappa_t = e^{-\int_0^t r(Z_s, s) ds} e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle}$;
2. $\mathbb{E}\left[\left(\int_0^t \langle \eta_s, \eta_s \rangle ds\right)^{\frac{1}{2}}\right] < \infty$ for $\eta_t = \kappa_t \psi(t, T, u)^\top \bar{\sigma}(Z_t)$;
3. $\mathbb{E}\left[\int_0^t |\gamma_s| ds\right] < \infty$ for $\gamma_t = \kappa_t \sum_{a=1}^2 \bar{\lambda}^a(Z_t) [\theta^a(\zeta_a \psi(t, T, u)) - 1]$.

To this end, we first have to determine explicitly the processes $(\kappa_t)_{t \geq 0}$, $(\eta_t)_{t \geq 0}$ and $(\gamma_t)_{t \geq 0}$ in terms of the model coefficients. It follows that

$$\kappa_t = e^{-\int_0^t (h(s) + c(N_s^1 - N_s^2) + \langle \ell, Z_s |_d \rangle) ds} e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle},$$

where $Z_t|_d$ coincides with X_t because the first d coordinates of the vector Z_t correspond to X_t . Then

$$\begin{aligned} \mathbb{E}[|\kappa_t|] &= e^{-\int_0^t h(s) ds} \mathbb{E}\left[\left|e^{-\int_0^t (cN_s^1 + \langle \ell, X_s \rangle) ds} e^{\int_0^t cN_s^2 ds} e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle}\right|\right] \\ &= e^{-\int_0^t h(s) ds} \mathbb{E}\left[e^{-\int_0^t (cN_s^1 + \langle \ell, X_s \rangle) ds} e^{\int_0^t cN_s^2 ds} \left|e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle}\right|\right] \\ &\leq e^{-\int_0^t h(s) ds} \mathbb{E}\left[e^{\int_0^t cN_s^2 ds} e^{Re(\phi(t, T, u)) + \langle Re(\psi(t, T, u)), Z_t \rangle}\right]. \end{aligned}$$

In the last line we have employed the fact that $\int_0^t (cN_s^1 + \langle \ell, X_s \rangle) ds \in \mathbb{R}_+$, therefore the exponential is less than 1. It follows that in order to verify the first integrability condition the last written expected value has to be finite. Considering that for all $z \in \mathbb{R}_+^{d+2} \times \mathbb{N}^2$ $\bar{\sigma}(z) = \text{diag}((\sqrt{z_1}, \dots, \sqrt{z_d}, 0, 0, 0, 0)^\top)$ we compute

$$\psi(t, T, u)^\top \bar{\sigma}(Z_t) = (\psi_1(t, T, u) \sqrt{(Z_t)_1}, \dots, \psi_d(t, T, u) \sqrt{(Z_t)_d}, 0, 0, 0, 0)$$

and therefore the second integrability condition becomes

$$\mathbb{E}\left[\left(\int_0^t \langle \eta_s, \eta_s \rangle ds\right)^{\frac{1}{2}}\right] = \mathbb{E}\left[\left(\int_0^t \kappa_s^2 \sum_{i=1}^d (\psi_i(s, T, u)^2 (X_s)_i) ds\right)^{\frac{1}{2}}\right] < \infty.$$

From (2.10) it follows that for all $z \in \mathbb{R}^{d+2} \times \mathbb{N}^2$

$$\bar{\lambda}^1(z) = \langle \Lambda_1, z|_d \rangle + z_{d+1} = \langle \Lambda_1, x \rangle + \tilde{\lambda}^1 \quad \text{and} \quad \bar{\lambda}^2(z) = \langle \Lambda_2, z|_d \rangle + z_{d+2} = \langle \Lambda_2, x \rangle + \tilde{\lambda}^2.$$

Therefore, regarding the last integrability condition from (2.19) and (2.20) it holds that

$$\begin{aligned} \gamma_t &= \kappa_t \left(\bar{\lambda}^1(Z_t) [\theta^1(\zeta_1 \psi(t, T, u)) - 1] + \bar{\lambda}^2(Z_t) [\theta^2(\zeta_2 \psi(t, T, u)) - 1] \right) \\ &= \kappa_t \left((\langle \Lambda_1, X_t \rangle + \tilde{\lambda}_t^1) [e^{\delta_1 \psi_{d+1}(t, T, u) + \psi_{d+3}(t, T, u)} - 1] \right. \\ &\quad \left. + (\langle \Lambda_2, X_t \rangle + \tilde{\lambda}_t^2) [e^{\delta_2 \psi_{d+2}(t, T, u) + \psi_{d+4}(t, T, u)} - 1] \right). \end{aligned}$$

We can not further analyze the three integrability conditions since they depend on the specific values of the model coefficients that are yet to be determined.

Continuing our analysis of the definition of *well-behaved* characteristics we are interested in studying the existence and uniqueness of solutions of (2.16) and (2.17). Clearly ϕ is uniquely determined by ψ , as we showed in (2.18), and consequently we have to study only the Riccati ODEs system (2.17) that can be written as

$$\begin{aligned}\frac{d\psi}{dt}(t, T, u) &= f(t, \psi(t, T, u)) \\ \psi(T, T, u) &= u,\end{aligned}\tag{2.21}$$

for a C^∞ -function $f : [0, T] \times \mathbb{C}^{d+4} \rightarrow \mathbb{C}^{d+4}$ (we remark that the smoothness of f is deduced from its definition because f is a sum of C^∞ -functions). We refer to [45] (Theorems 8.1, 8.2) and apply these uniqueness and existence theorems to the real and the complex part of (2.21) and then it follows that it is locally solved uniquely. In other words, in a neighbourhood of (T, u) there exists a unique function ψ that solves (2.21) and consequently (2.17).

We are interested in discussing the existence of a closed-form solution for (2.17) in the real domain, that is for every final condition $u \in \mathbb{R}^{d+4}$ we aim at determining an explicit function $\psi : [0, T] \rightarrow \mathbb{R}^{d+4}$ which solves the Riccati system. From the structure of the system it follows that the final two components of ψ verify an easily solvable system composed by two independent ODEs. In other words, ψ_{d+3} satisfies the ODE

$$\frac{d\psi_{d+3}}{dt}(t, T, u) = c \quad \text{for } 0 \leq t \leq T \quad \text{with } \psi_{d+3}(T, T, u) = u_{d+3},$$

and ψ_{d+4} the following one

$$\frac{d\psi_{d+4}}{dt}(t, T, u) = -c \quad \text{for } 0 \leq t \leq T \quad \text{with } \psi_{d+4}(T, T, u) = u_{d+4}.$$

It is easily deduced that $\psi_{d+3}(t, T, u) = c(t-T) + u_{d+3}$ and $\psi_{d+4}(t, T, u) = c(T-t) + u_{d+4}$. If we replace these solutions in the equations that define the system (2.17), then ψ_{d+1} solves the ODE

$$\begin{aligned}\frac{d\psi_{d+1}}{dt}(t, T, u) &= \gamma_1 \psi_{d+1}(t, T, u) - e^{\delta_1 \psi_{d+1}(t, T, u) + c(t-T) + u_{d+3}} + 1 \\ \psi_{d+1}(T, T, u) &= u_{d+1},\end{aligned}\tag{2.22}$$

whereas ψ_{d+2} verifies the following one

$$\begin{aligned}\frac{d\psi_{d+2}}{dt}(t, T, u) &= \gamma_2 \psi_{d+2}(t, T, u) - e^{\delta_2 \psi_{d+2}(t, T, u) + c(T-t) + u_{d+4}} + 1 \\ \psi_{d+2}(T, T, u) &= u_{d+2}.\end{aligned}\tag{2.23}$$

Now, we aim at analysing the previous ODEs in order to find exact solutions ψ_{d+1} and ψ_{d+2} . We notice that (2.22) and (2.23) are non-linear first-order ODEs but we are not able to find closed-form solutions if γ_1 and δ_1 are different from zero simultaneously, as well as γ_2 and δ_2 .

If $\delta_1 = 0$ then (2.22) can be rewritten in

$$\frac{d\psi_{d+1}}{dt}(t, T, u) = \gamma_1 \psi_{d+1}(t, T, u) - e^{c(t-T)+u_{d+3}} + 1.$$

It is a linear ODE and, for $\gamma_1 \neq 0$ and $\gamma_1 \neq c$ its solution is

$$\begin{aligned} \psi_{d+1}(t, T, u) &= e^{\int_T^t \gamma_1 ds} \left(u_{d+1} + \int_T^t \left(1 - e^{c(s-T)+u_{d+3}} \right) e^{-\int_T^s \gamma_1 du} ds \right) \\ &= e^{\gamma_1(t-T)} \left(u_{d+1} - \frac{e^{\gamma_1 T}}{\gamma_1} \left(e^{-\gamma_1 t} - e^{-\gamma_1 T} \right) - \frac{e^{(\gamma_1-c)T+u_{d+3}}}{c-\gamma_1} \left(e^{(c-\gamma_1)t} - e^{(c-\gamma_1)T} \right) \right) \\ &= u_{d+1} e^{-\gamma_1(T-t)} + \frac{e^{-\gamma_1(T-t)}}{\gamma_1} - \frac{1}{\gamma_1} + \frac{e^{u_{d+3}}}{c-\gamma_1} e^{-\gamma_1(T-t)} - \frac{e^{u_{d+3}}}{c-\gamma_1} e^{-c(T-t)} \\ &= \left(u_{d+1} + \frac{1}{\gamma_1} + \frac{e^{u_{d+3}}}{c-\gamma_1} \right) e^{-\gamma_1(T-t)} - \frac{1}{\gamma_1} - \frac{e^{u_{d+3}}}{c-\gamma_1} e^{-c(T-t)}. \end{aligned}$$

If $\delta_1 = \gamma_1 = 0$ and $c \neq 0$ the function ψ_{d+1} is determined by

$$\begin{aligned} \psi_{d+1}(t, T, u) &= u_{d+1} - \int_T^t e^{c(s-T)+u_{d+3}} ds + t - T \\ &= u_{d+1} + \frac{e^{u_{d+3}}}{c} \left(1 - e^{-c(T-t)} \right) + t - T. \end{aligned}$$

If $\delta_1 = \gamma_1 = c = 0$, then ψ_{d+1} is specified by

$$\psi_{d+1}(t, T, u) = u_{d+1} + (e^{u_{d+3}} - 1)(T - t).$$

Lastly, if $\delta_1 = 0$ and $\gamma_1 = c \neq 0$ the function ψ_{d+1} is

$$\psi_{d+1}(t, T, u) = \left(u_{d+1} + \frac{1}{\gamma_1} + e^{u_{d+3}}(T - t) \right) e^{-\gamma_1(T-t)} - \frac{1}{\gamma_1}.$$

Remark 2.6. If $\delta_1 = 0$, then the jump intensity component $(\lambda_t^1)_{t \geq 0}$ is simply defined by $\lambda_t^1 = \lambda_1^0 + \langle \Lambda_1, X_t \rangle$ and therefore we can assume that $\gamma_1 = 0$ without loss of generality. We have analyzed also the case $\gamma_1 \neq 0$ for completeness.

By means of analogous computations, if $\delta_2 = 0$ and $\gamma_2 \neq 0$ then (2.23) becomes a linear ODE with explicit solution defined by

$$\psi_{d+2}(t, T, u) = \left(u_{d+2} - \frac{1}{\gamma_2} - \frac{e^{u_{d+4}}}{c + \gamma_2} \right) e^{-\gamma_2(T-t)} - \frac{1}{\gamma_2} + \frac{e^{u_{d+4}}}{c + \gamma_2} e^{c(T-t)}.$$

If $\delta_2 = \gamma_2 = 0$ and $c \neq 0$ the function ψ_{d+2} is specified by

$$\psi_{d+2}(t, T, u) = u_{d+2} + \frac{e^{u_{d+4}}}{c} \left(e^{c(T-t)} - 1 \right) + t - T.$$

Whereas, if $\delta_2 = \gamma_2 = c = 0$

$$\psi_{d+2}(t, T, u) = u_{d+2} + (e^{u_{d+4}} - 1)(T - t).$$

On the other hand, if $\delta_1 \neq 0$ but $\gamma_1 = 0$ then (2.22) becomes

$$\frac{d\psi_{d+1}}{dt}(t, T, u) = -e^{\delta_1\psi_{d+1}(t, T, u) + c(t-T) + u_{d+3}} + 1. \quad (2.24)$$

Remark 2.7. The case $\gamma_1 = 0$ represents the absence of the decay effect of past jumps, that is the intensity process $(\lambda_t^1)_{t \geq 0}$ is defined by

$$\lambda_t^1 = \lambda_0^1 + \langle \Lambda_1, X_t \rangle + \delta_1 \int_0^t dN_u^1 = \lambda_0^1 + \langle \Lambda_1, X_t \rangle + \delta_1 N_t^1.$$

This is well-defined since the L^1 -property (A.7) is satisfied. It follows that $(N_t^1)_{t \geq 0}$ is non-explosive.

By means of replacement $\psi_{d+1}(t, T, u) \rightarrow w_{d+1}(t, T, u) := e^{-\delta_1\psi_{d+1}(t, T, u)}$ the non-linear ODE (2.24) becomes a linear ODE such that

$$\begin{aligned} \frac{dw_{d+1}}{dt}(t, T, u) &= \delta_1 e^{c(t-T) + u_{d+3}} - \delta_1 w_{d+1}(t, T, u) \\ w_{d+1}(T, T, u) &= e^{-\delta_1 u_{d+1}}. \end{aligned}$$

Its solution is

$$\begin{aligned} w_{d+1}(t, T, u) &= e^{-\int_T^t \delta_1 ds} \left(e^{-\delta_1 u_{d+1}} + \int_T^t \delta_1 e^{c(s-T) + u_{d+3}} e^{\int_T^s \delta_1 du} ds \right) \\ &= \left(e^{-\delta_1 u_{d+1}} - \frac{\delta_1}{c + \delta_1} e^{u_{d+3}} \right) e^{\delta_1(T-t)} + \frac{\delta_1}{c + \delta_1} e^{u_{d+3}} e^{-c(T-t)}. \end{aligned}$$

Then, we can revert the change and deduce that $\psi_{d+1}(t, T, u) = -\frac{\log w_{d+1}(t, T, u)}{\delta_1}$. We remark that this last step is permitted only if $w_{d+1}(t, T, u) > 0$, that is if

$$\log \left(\frac{\delta_1}{c + \delta_1} \right) + \delta_1 u_{d+1} + u_{d+3} \leq 0, \quad (2.25)$$

or if the following conditions hold

$$\begin{aligned} \log \left(\frac{\delta_1}{c + \delta_1} \right) + \delta_1 u_{d+1} + u_{d+3} &> 0; \\ t > T + \frac{1}{c + \delta_1} \left(\log \left(\frac{\delta_1}{c + \delta_1} e^{u_{d+3}} - e^{-\delta_1 u_{d+1}} \right) - \log \left(\frac{\delta_1}{c + \delta_1} \right) - u_{d+3} \right). \end{aligned}$$

In the same spirit, we use a similar change of variable in case of $\delta_2 \neq 0$ and $\gamma_2 = 0$ in (2.23), i.e. $\psi_{d+2}(t, T, u) \rightarrow w_{d+2}(t, T, u) := e^{-\delta_2\psi_{d+2}(t, T, u)}$. Therefore, the non-linear ODE (2.23) becomes a linear ODE of the following form

$$\begin{aligned} \frac{dw_{d+2}}{dt}(t, T, u) &= \delta_2 e^{c(T-t) + u_{d+4}} - \delta_2 w_{d+2}(t, T, u) \\ w_{d+2}(T, T, u) &= e^{-\delta_2 u_{d+2}}. \end{aligned}$$

For $\delta_2 \neq c$ its explicit solution is

$$w_{d+2}(t, T, u) = \left(e^{-\delta_2 u_{d+2}} - \frac{\delta_2}{\delta_2 - c} e^{u_{d+4}} \right) e^{\delta_2(T-t)} + \frac{\delta_2}{\delta_2 - c} e^{u_{d+4}} e^{c(T-t)}.$$

Consequently, $\psi_{d+2}(t, T, u) = -\frac{\log w_{d+2}(t, T, u)}{\delta_2}$ by reverting the change of variable if

$$\begin{aligned} \delta_2 &> c; \\ \log \left(\frac{\delta_2}{\delta_2 - c} \right) + \delta_2 u_{d+2} + u_{d+4} &\leq 0, \end{aligned} \tag{2.26}$$

or if

$$\begin{aligned} \delta_2 &> c; \\ \log \left(\frac{\delta_2}{\delta_2 - c} \right) + \delta_2 u_{d+2} + u_{d+4} &> 0; \\ t &> T + \frac{1}{\delta_2 - c} \left(\log \left(\frac{\delta_2}{\delta_2 - c} e^{u_{d+4}} - e^{-\delta_2 u_{d+2}} \right) - \log \left(\frac{\delta_2}{\delta_2 - c} \right) - u_{d+4} \right), \end{aligned}$$

or if

$$\begin{aligned} \delta_2 &< c; \\ t &> T - \frac{1}{c - \delta_2} \left(\log \left(\frac{\delta_2}{c - \delta_2} e^{u_{d+4}} + e^{-\delta_2 u_{d+2}} \right) - \log \left(\frac{\delta_2}{c - \delta_2} \right) - u_{d+4} \right). \end{aligned}$$

If $\gamma_2 = 0$ and $\delta_2 = c \neq 0$ it follows that

$$w_{d+2}(t, T, u) = \left(e^{-\delta_2 u_{d+2}} - \delta_2 e^{u_{d+4}} (T - t) \right) e^{\delta_2(T-t)}$$

and, therefore, $\psi_{d+2}(t, T, u) = -\frac{\log w_{d+2}(t, T, u)}{\delta_2}$ if

$$t > T - \frac{1}{\delta_2} e^{-\delta_2 u_{d+2} - u_{d+4}}.$$

In conclusion, if $\delta_1 = \delta_2 = 0$ we are able to find explicit closed-form solutions ψ_{d+1} and ψ_{d+2} to (2.22) and (2.23), respectively. However, in this particular case for $a = 1, 2$ the intensity process $(\lambda_t^a)_{t \geq 0}$ defined by (2.3) loses its particular characteristic deriving from Hawkes processes' framework. In other words, $\lambda_t^a = \lambda_0^a + \langle \Lambda_a, X_t \rangle$. The intensity is only a linear combination of the process $(X_t)_{t \geq 0}$ and, as a consequence, the counting process $(N_t^a)_{t \geq 0}$ does not influence its value. On the other hand, in the previous analysis we have achieved solutions ψ_{d+1} and ψ_{d+2} also if $\gamma_1 = \gamma_2 = 0$ and $\delta_1 \neq 0 \neq \delta_2$. In this case $\lambda_t^a = \lambda_0^a + \langle \Lambda_a, X_t \rangle + \delta_a N_t^a$ and it follows that the counting process $(N_t^a)_{t \geq 0}$ influences the intensity value but without a decay effect over time.

2.2.1 From Riccati ODEs system to the affine transform formula

In Proposition 2.5 we have showed that if the characteristics $\chi^d = (\bar{K}, \bar{H}, l^1, l^2, \theta^1, \theta^2, \rho)$ are *well-behaved* at $(u, T) \in \mathbb{C}^{d+4} \times \mathbb{R}_+$, then

$$\mathbb{E} \left[e^{-\int_t^T r(Z_s, s) ds} e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t \right] = e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle}. \tag{2.27}$$

Nevertheless, verifying that χ^d are *well-behaved* is onerous and for this reason we aim at simplifying this condition. The article [44] of Keller-Ressel & Mayerhofer addresses this problem; the authors first show that the affine transform formula (2.27) holds for $u \in \mathbb{R}^{d+4}$ up to the maximal lifetime T of the solutions ϕ and ψ of the generalized Riccati system (2.16) and (2.17) and finally perform the extension for complex exponents. We highlight that Keller-Ressel & Mayerhofer use the notion of minimal solution since their article deals with general affine process and therefore the solution of the related ODEs system may not be unique. In our case, the solution's uniqueness is guaranteed by the smoothness of the functions which determine the derivatives of ϕ and ψ .

We first consider the real case and state the following Lemma which is a direct consequence of the main result of [44] (Theorem 2.14).

Lemma 2.8. *Let $u \in \mathbb{R}^{d+4}$ and suppose that the ODEs system (2.16) and (2.17) with terminal condition u has solutions ϕ and ψ up to time T . Then*

$$\mathbb{E} \left[e^{-\int_t^T r(Z_s, s) ds} e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t \right] = e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle} \quad \text{for } 0 \leq t \leq T.$$

It follows that in order to verify the validity of the affine transform formula in the real case we have only to analyse the existence of solutions ψ and ϕ of the generalized Riccati system. We have already observed that (2.16) is easy to solve once we know the function ψ and its solution ϕ is determined by (2.18). Therefore, we can focus our analysis on (2.17), i.e. the ODEs which characterize the function ψ .

Let $T > 0$ be fixed. To simplify the notation we introduce the function y such that $y(t) := \psi(T - t)$ for all $t \geq 0$. Consequently, we write the ODEs satisfied by y as follows:

$$\begin{aligned} \frac{dy}{dt}(t) &= g(t, y(t)) \\ y(0) &= u, \end{aligned} \tag{2.28}$$

for $u \in \mathbb{R}^{d+4}$ and the C^∞ -function $g : [0, t(u)) \times \mathbb{R}^{d+4} \rightarrow \mathbb{R}^{d+4}$. With $t(u)$ we denote the lifetime of the solution y , i.e. either $t(u) = +\infty$ or $\lim_{t \rightarrow t(u)} \|y(t)\| = +\infty$.

In the previous section we have already observed that in a neighbourhood of $(0, u)$ there exists a unique function y that solves (2.28). By analysing these Riccati ODEs we aim at identifying every initial value $u \in \mathbb{R}^{d+4}$ such that there exists a unique global solution y that is not in general guaranteed due to the presence of the quadratic term. In other words, we want to study if the function y blows up in finite time, i.e. $t(u) < +\infty$. We remark that this is equivalent to study the existence of a global solution ψ that satisfies (2.17), or equivalently (2.21), for any final time T .

Firstly, in the previous section we have computed the closed-form expressions of ψ_{d+3} and ψ_{d+4} which are linear in the time variable and therefore we derive that y_{d+3} and y_{d+4} do not explode at any finite time. Then, to study the existence of a global solution y to (2.28) we replace the found expressions of y_{d+3} and y_{d+4} and investigate this system. We note that the ODEs which describe the derivative of y_{d+1} and y_{d+2} are independent from $y|_d$. It follows that we can investigate them and find sufficient conditions in order

to guarantee global solutions y_{d+1} and y_{d+2} . Then, we study the first d ODEs and finally analyse if their solution $y|_d$ is global, in turn.

Before continuing the analysis of the existence of a global solution y to (2.28) we enunciate a result stated in [28] (Corollary 10.5).

Lemma 2.9. *We define $\mathcal{H} := \{x \in \mathbb{R} | x \geq 0\}$ the half space in \mathbb{R} and consider the continuous map $b : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ satisfying for $t \geq 0$*

$$\begin{aligned} b(t, x) &\geq 0 \quad \text{for all } x \in \partial\mathcal{H} \\ b(t, x) &= b(t, x^+) \quad \text{for all } x \in \mathbb{R}, \end{aligned}$$

where $+$ denotes the positive part. Then, any solution f of the ODE $\frac{df}{dt}(t) = b(t, f(t))$ with $f(0) \in \mathbb{R}_+$ verifies $f(t) \in \mathbb{R}_+$ for all $t \geq 0$.

Recalling that $y(t) = \psi(T - t)$ for all $t \geq 0$, from (2.22) and (2.23) we deduce the ODEs which describe y_{d+1} and y_{d+2} in order to analyze if they are global solutions:

$$\begin{aligned} \frac{dy_{d+1}}{dt}(t) &= -\gamma_1 y_{d+1}(t) + e^{\delta_1 y_{d+1}(t) - ct + u_{d+3}} - 1 =: g_{d+1}(t, y_{d+1}(t)) \\ y_{d+1}(0) &= u_{d+1}, \\ \frac{dy_{d+2}}{dt}(t) &= -\gamma_2 y_{d+2}(t) + e^{\delta_2 y_{d+2}(t) + ct + u_{d+4}} - 1 =: g_{d+2}(t, y_{d+2}(t)) \\ y_{d+2}(0) &= u_{d+2}. \end{aligned} \tag{2.29}$$

Proposition 2.10. *The solutions $y_{d+1}, y_{d+2} : [0, +\infty) \rightarrow \mathbb{R}$ of (2.29) are global if one of the following conditions are verified*

1. $\delta_1 = 0$;
2. $\gamma_1 = 0$, $\delta_1 \neq 0$ and $\log\left(\frac{\delta_1}{c + \delta_1}\right) + \delta_1 u_{d+1} + u_{d+3} \in \mathbb{R}_-$;
3. $u_{d+1}, u_{d+3} \in \mathbb{R}_-$,

and another one among

1. $\delta_2 = 0$;
2. $\gamma_2 = 0$, $\delta_2 > c$ and $\log\left(\frac{\delta_2}{\delta_2 - c}\right) + \delta_2 u_{d+2} + u_{d+4} \in \mathbb{R}_-$;
3. $c = 0$ and $u_{d+2}, u_{d+4} \in \mathbb{R}_-$.

In order to prove this Proposition we exploit the closed-form expressions that we have computed in the previous section for the functions ψ_{d+1} and ψ_{d+2} .

Proof. To study the ODE characterizing y_{d+1} we have to distinguish different cases. For $\delta_1 = 0$ we have computed an explicit global solution, whereas for $\gamma_1 = 0 \neq \delta_1$ there exists an unique global solution only if

$$\log\left(\frac{\delta_1}{c + \delta_1}\right) + \delta_1 u_{d+1} + u_{d+3} \leq 0. \tag{2.30}$$

This is a direct consequence of (2.25). A sufficient condition for (2.30) to be satisfied is $u_{d+1}, u_{d+3} \in \mathbb{R}_-$. Lastly, if $\delta_1 \neq 0 \neq \gamma_1$ we can not compute a closed-form solution y_{d+1} . Nevertheless, if $u_{d+1}, u_{d+3} \in \mathbb{R}_-$ we consider $b(t, x) := -\gamma_1 x^+ - e^{-\delta_1 x^+ - ct + u_{d+3}} + 1$ for $x \in \mathbb{R}$ and analyse the function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ satisfying

$$\begin{aligned} \frac{df}{dt}(t) &= b(t, f(t)) \\ f(0) &= -u_{d+1}. \end{aligned}$$

From Lemma 2.9 we deduce that the function f takes value in \mathbb{R}_+ . Then,

$$\frac{df}{dt}(t) = -\gamma_1 f(t)^+ - e^{-\delta_1 f(t)^+ - ct + u_{d+3}} + 1 = -\gamma_1 f(t) - e^{-\delta_1 f(t) - ct + u_{d+3}} + 1.$$

Setting $f(t) = -y_{d+1}(t)$ for all $t \geq 0$ it follows that y_{d+1} has to take values in \mathbb{R}_- for every initial point $u_{d+1} \in \mathbb{R}_-$. We deduce that

$$|g_{d+1}(t, y_{d+1}(t))| = |-\gamma_1 y_{d+1}(t) + e^{\delta_1 y_{d+1}(t) - ct + u_{d+3}} - 1| \leq \gamma_1 |y_{d+1}(t)| + 1$$

and therefore g_{d+1} is linearly bounded with respect to y_{d+1} . From [45] (Proposition 7.8) we conclude that there exists a global negative solution y_{d+1} for every $u_{d+1}, u_{d+3} \in \mathbb{R}_-$.

Accordingly, if $\delta_2 = 0$ an explicit global solution was computed in the previous section, whereas if $\gamma_2 = 0 \neq \delta_2$ an unique global solution y_{d+2} exists whether (2.26) holds, i.e.

$$\begin{aligned} \delta_2 &> c; \\ \log\left(\frac{\delta_2}{\delta_2 - c}\right) + \delta_2 u_{d+2} + u_{d+4} &\leq 0. \end{aligned}$$

If $c = 0$ and $u_{d+4} \in \mathbb{R}_-$, then y_{d+2} takes negative values for every $u_{d+2} \in \mathbb{R}_-$ as a consequence of Lemma 2.9. Therefore,

$$|g_{d+2}(t, y_{d+2}(t))| = |-\gamma_2 y_{d+2}(t) + e^{\delta_2 y_{d+2}(t) + u_{d+4}} - 1| \leq \gamma_2 |y_{d+2}(t)| + 1$$

and we can conclude that a global negative solution y_{d+2} exists. \square

Corollary 2.11. *If $c = 0$ and $u_{d+1}, u_{d+2}, u_{d+3}, u_{d+4} \in \mathbb{R}_-$, then y_{d+1} and y_{d+2} are negative global solutions of (2.29).*

Up to this point we have imposed some sufficient conditions for guaranteeing global solutions $y_{d+1}, y_{d+2}, y_{d+3}$ and y_{d+4} . Now, we investigate the blow-up behaviour of the first d components of y which satisfy the following Riccati ODEs system:

$$\begin{aligned} \frac{dy|_d}{dt}(t) &= -\ell + K_1^\top y|_d(t) + \frac{1}{2}(y_1^2(t), \dots, y_d^2(t))^\top \\ &\quad + \Lambda_1(e^{\delta_1 y_{d+1}(t) - ct + u_{d+3}} - 1) + \Lambda_2(e^{\delta_2 y_{d+2}(t) + ct + u_{d+4}} - 1) \\ &=: -\ell + K_1^\top y|_d(t) + \frac{1}{2}(y_1^2(t), \dots, y_d^2(t))^\top + \Lambda_1 \alpha(t) + \Lambda_2 \beta(t) \\ y|_d(0) &= u|_d. \end{aligned} \tag{2.31}$$

Proposition 2.12. *Suppose that K_1 is an upper triangular matrix. If $c = 0$, $u_{d+1}, u_{d+2}, u_{d+3}, u_{d+4} \in \mathbb{R}_-$ and $u \in \mathbb{R}_^d$, then there exists a unique global negative solution $y|_d$ to (2.31).*

Proof. We remark that if these hypotheses are satisfied, then from Corollary 2.11 we deduce that y_{d+1} and y_{d+2} are negative global solutions of (2.29). Consequently, the functions α and β defined in (2.31) are \mathbb{R}_- -valued.

To prove the Proposition we begin with the observation that $y|_d$ has to take negative values. We demonstrate this statement by induction on the index i employing similar arguments of the proof of Proposition 2.10. For $i = 1$, because of the upper triangular shape of the matrix K_1 the function y_1 satisfies the ODE

$$\frac{dy_1}{dt}(t) = -\ell_1 + (K_1)_{11}y_1(t) + \frac{1}{2}y_1^2(t) + (\Lambda_1)_1\alpha(t) + (\Lambda_2)_1\beta(t).$$

From Lemma 2.9 we deduce that y_1 has to take negative values. Now, suppose that y_k is negative-valued for $k = 1, \dots, i-1$. Then, the i -th component y_i verifies

$$\frac{dy_i}{dt}(t) = -\ell_i + \sum_{j=1}^{i-1} (K_1)_{ji}y_j(t) + (K_1)_{ii}y_i(t) + \frac{1}{2}y_i^2(t) + (\Lambda_1)_i\alpha(t) + (\Lambda_2)_i\beta(t). \quad (2.32)$$

Setting $b(t, x) = \ell_i - \sum_{j=1}^{i-1} (K_1)_{ji}y_j(t) + (K_1)_{ii}x^+ - \frac{1}{2}(x^+)^2 - (\Lambda_1)_i\alpha(t) - (\Lambda_2)_i\beta(t)$ and $f(t) = -y_i(t)$ in the statement of Lemma 2.9 it follows that the solution y_i of (2.32) takes value in \mathbb{R}_- for every initial point $u_i \in \mathbb{R}_-$. By induction we can conclude that $y|_d$ is $\mathbb{R}_^d$ -valued. From this property we deduce that

$$\begin{aligned} \frac{d\|y|_d(t)\|^2}{dt} &= 2 \left[-\langle y|_d(t), \ell \rangle + \langle y|_d(t), K_1^\top y|_d(t) \rangle + \frac{1}{2}(y_1^3(t) + \dots + y_d^3(t)) \right. \\ &\quad \left. + \langle y|_d(t), \Lambda_1 \rangle \alpha(t) + \langle y|_d(t), \Lambda_2 \rangle \beta(t) \right] \\ &\leq C(t)(1 + \|y|_d(t)\|^2), \end{aligned}$$

for a positive function C . Consequently,

$$\|y|_d(t)\|^2 \leq \|u|_d\|^2 + \int_0^t C(s)(1 + \|y|_d(s)\|^2) ds.$$

Then, Gronwall's Lemma (see [11] Lemma 1.1) implies that $\|y|_d(t)\|$ is finite for every $t \geq 0$, that is $y|_d$ is a global negative solution. \square

Remark 2.13. From Corollary 2.11 and Proposition 2.12 we can conclude that if K_1 is an upper triangular matrix, $u \in \mathbb{R}_^{d+4}$ and $c = 0$, then there exists an unique solution to the Riccati system (2.28) which does not explode in finite time. Therefore, the exponential transform (2.27) holds for every final time $T > 0$ as a consequence of Lemma 2.8.

Now, we focus on the problem of the existence of the exponential transform formula without discounting, that is we analyse the following conditional expectation

$$\mathbb{E}\left[e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t\right] \text{ for } u \in \mathbb{R}^{d+4}. \quad (2.33)$$

We observe that (2.33) is equivalent to the transform (2.27) if we consider the short rate process identically null, that is $h(t) = 0$ for $t \geq 0$, $c = 0$ and $\ell = 0$ because of its definition $r_t = h(t) + c(N_t^1 - N_t^2) + \langle \ell, X_t \rangle$. Then, from Proposition 2.12 we can deduce that if K_1 is an upper triangular matrix then the exponential transform formula without discounting holds for $u \in \mathbb{R}_-^{d+4}$.

Corollary 2.14. *If K_1 is an upper triangular matrix, then for every $u \in \mathbb{R}_-^{d+4}$ and $T > 0$ it holds that*

$$\mathbb{E}\left[e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t\right] = e^{\tilde{\phi}(t, T, u) + \langle \tilde{\psi}(t, T, u), u_{d+3}, u_{d+4} \rangle^\top, Z_t} \text{ for } 0 \leq t \leq T, \quad (2.34)$$

where, omitting T and u dependencies, $\tilde{\phi} : [0, T] \rightarrow \mathbb{R}_-$ is defined by

$$\tilde{\phi}(t) = \int_t^T \left(\langle K_0, \tilde{\psi}|_d(s) \rangle + \gamma_1 \lambda_0^1 \tilde{\psi}|_{d+1}(s) + \gamma_2 \lambda_0^2 \tilde{\psi}|_{d+2}(s) \right) ds \quad (2.35)$$

and $\tilde{\psi} : [0, T] \rightarrow \mathbb{R}_-^{d+2}$ is the unique solution of the following ODEs

$$\begin{aligned} \frac{d\tilde{\psi}}{dt}(t) &= \begin{pmatrix} -K_1^\top \tilde{\psi}|_d(t) \\ \gamma_1 \tilde{\psi}|_{d+1}(t) \\ \gamma_2 \tilde{\psi}|_{d+2}(t) \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \tilde{\psi}^2|_d(t) \\ 0 \\ 0 \end{pmatrix} \\ &\quad - \begin{pmatrix} \Lambda_1 \\ 1 \\ 0 \end{pmatrix} (e^{\delta_1 \tilde{\psi}|_{d+1}(t) + u_{d+3}} - 1) - \begin{pmatrix} \Lambda_2 \\ 0 \\ 1 \end{pmatrix} (e^{\delta_2 \tilde{\psi}|_{d+2}(t) + u_{d+4}} - 1) \end{aligned} \quad (2.36)$$

with final condition $\tilde{\psi}(T) = u|_{d+2}$. With the notation $\tilde{\psi}^2|_d(t)$ we mean the square of each component of $\tilde{\psi}|_d(t)$, that is the \mathbb{R}^d -vector defined by $(\tilde{\psi}_1^2(t), \dots, \tilde{\psi}_d^2(t))^\top$.

Remark 2.15. We highlight that the previous ODEs system (2.36) is exactly (2.17) in which we have fixed $\ell = 0$ and $c = 0$. Analogously, (2.35) is similar to (2.18) considering h as the null function.

To complete the treatment of exponential moments we want to extend the validity of the affine transform formula to the complex case giving an analogue of (2.34). This extension is permitted under the hypothesis that the Riccati system (2.36) with real-valued final condition is solvable until time T . We state a result proved in [44] (Theorem 2.26).

Lemma 2.16. *Let $u \in \mathbb{C}^{d+4}$ and suppose that the ODEs system (2.36) admits a unique real-valued solution up to time T with terminal value $Re(u|_{d+2})$. Then the system (2.36)*

has a unique complex solution $\tilde{\psi}$ that exists up to time T and satisfying the terminal condition $\tilde{\psi}(T) = u|_{d+2}$. It follows that

$$\mathbb{E}\left[e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t\right] = e^{\tilde{\phi}(t, T, u) + \langle (\tilde{\psi}(t, T, u), u_{d+3}, u_{d+4})^\top, Z_t \rangle} \quad \text{for } 0 \leq t \leq T.$$

The function $\tilde{\phi}$ is determined by (2.35).

2.3 The probability of future jumps

We aim at finding the explicit expression for the conditional probability of future jumps of the stochastic lower bound SLB . That is, the purpose of this section is computing closed-form formulas that express the probability that the lower bound rises or falls by a certain amount on a certain future horizon.

Let t and T two time horizons such that $0 \leq t \leq T$. For all $s \in c\mathbb{Z}$ we investigate

$$\mathbb{P}(\Delta SLB_T = s | \mathcal{F}_t) = \mathbb{P}(SLB_T - SLB_t = s | \mathcal{F}_t).$$

Recalling the SLB definition in (2.4)

$$\begin{aligned} \mathbb{P}(\Delta SLB_T = s | \mathcal{F}_t) &= \sum_{n=\frac{s}{c}}^{\infty} \mathbb{P}\left(\Delta N_T^1 = n, \Delta N_T^2 = n - \frac{s}{c} \middle| \mathcal{F}_t\right) \mathbf{1}_{\{s > 0\}} \\ &\quad + \sum_{n=0}^{\infty} \mathbb{P}\left(\Delta N_T^1 = n, \Delta N_T^2 = n - \frac{s}{c} \middle| \mathcal{F}_t\right) \mathbf{1}_{\{s \leq 0\}}. \end{aligned} \quad (2.37)$$

To simplify the notation used in the following we denote

$$p_t(n, s, T) := \mathbb{P}\left(\Delta N_T^1 = n, \Delta N_T^2 = n - \frac{s}{c} \middle| \mathcal{F}_t\right). \quad (2.38)$$

Proposition 2.17. *If K_1 is an upper triangular matrix, then it holds that*

$$p_t(n, s, T) = \frac{1}{n!} \frac{1}{(n - \frac{s}{c})!} \frac{\partial^{2n-s/c}}{\partial v_1^n \partial v_2^{n-s/c}} \left(e^{\tilde{\phi}_{v_1 v_2}(t, T, \underline{0}) + \langle \tilde{\psi}_{v_1 v_2}(t, T, \underline{0}), (X_t^\top, \tilde{\lambda}_t^1, \tilde{\lambda}_t^2)^\top \rangle} \right) \Big|_{\substack{v_1=0 \\ v_2=0}}. \quad (2.39)$$

We recall that $\tilde{\lambda}_t^a = \lambda_0^a + \delta_a \int_0^t e^{-\gamma_a(t-u)} dN_u^a$, for $a = 1, 2$. The \mathbb{R} -valued function $\tilde{\phi}_{v_1 v_2}$ is determined by

$$\begin{aligned} \tilde{\phi}_{v_1 v_2}(t, T, \underline{0}) &= \int_t^T \langle K_0, (\tilde{\psi}_{v_1 v_2})|_d(s, T, \underline{0}) \rangle ds \\ &\quad + \int_t^T \left(\gamma_1 \lambda_0^1 (\tilde{\psi}_{v_1 v_2})_{d+1}(s, T, \underline{0}) + \gamma_2 \lambda_0^2 (\tilde{\psi}_{v_1 v_2})_{d+2}(s, T, \underline{0}) \right) ds \end{aligned} \quad (2.40)$$

and the \mathbb{R}^{d+2} -valued function $\tilde{\psi}_{v_1 v_2}$ satisfy the following ODEs system

$$\begin{aligned} \frac{d\tilde{\psi}_{v_1 v_2}}{dt}(t, T, \underline{0}) &= \begin{pmatrix} -K_1^\top(\tilde{\psi}_{v_1 v_2})|_d(t, T, \underline{0}) \\ \gamma_1(\tilde{\psi}_{v_1 v_2})_{d+1}(t, T, \underline{0}) \\ \gamma_2(\tilde{\psi}_{v_1 v_2})_{d+2}(t, T, \underline{0}) \end{pmatrix} - \frac{1}{2} \begin{pmatrix} (\tilde{\psi}_{v_1 v_2}^2)|_d(t, T, \underline{0}) \\ 0 \\ 0 \end{pmatrix} \\ &\quad - \begin{pmatrix} \Lambda_1 \\ 1 \\ 0 \end{pmatrix} (e^{\delta_1(\tilde{\psi}_{v_1 v_2})_{d+1}(t, T, \underline{0}) + \log(v_1)} - 1) \\ &\quad - \begin{pmatrix} \Lambda_2 \\ 0 \\ 1 \end{pmatrix} (e^{\delta_2(\tilde{\psi}_{v_1 v_2})_{d+2}(t, T, \underline{0}) + \log(v_2)} - 1) \end{aligned} \quad (2.41)$$

with terminal condition $\tilde{\psi}_{v_1 v_2}(T, T, \underline{0}) = \underline{0}$. We highlight that $\underline{0}$ denotes the null vector of \mathbb{R}^{d+2} .

Remark 2.18. In the statement of Proposition 2.17 we require that K_1 is an upper triangular matrix in order to the ODEs system (2.41) admits a unique global solution (see Proposition 2.12).

Proof. Following the approach of [27] (Section 2.6), to prove the statement it suffices to compute the conditional probability generating function of the bivariate counting process $(N_t)_{t \geq 0} = ((N_t^1, N_t^2)^\top)_{t \geq 0}$, that is

$$\mathbb{E} \left[v_1^{N_T^1 - N_t^1} v_2^{N_T^2 - N_t^2} \middle| \mathcal{F}_t \right] = \sum_{\substack{n_1=0 \\ n_2=0}}^{\infty} v_1^{n_1} v_2^{n_2} \mathbb{P}(N_T^1 - N_t^1 = n_1, N_T^2 - N_t^2 = n_2 | \mathcal{F}_t) \quad (2.42)$$

for $v_1, v_2 \in (0, 1)$; in this case the sum of the right-hand side of (2.42) is guaranteed to converge and, therefore, the conditional expected value exists (see [37] (Chapter 7)).

We have that

$$\begin{aligned} \mathbb{E} \left[v_1^{N_T^1 - N_t^1} v_2^{N_T^2 - N_t^2} \middle| \mathcal{F}_t \right] &= \frac{1}{v_1^{N_t^1} v_2^{N_t^2}} \mathbb{E} \left[v_1^{N_T^1} v_2^{N_T^2} \middle| \mathcal{F}_t \right] \\ &= \frac{1}{v_1^{N_t^1} v_2^{N_t^2}} \mathbb{E} \left[e^{\log(v_1) N_T^1} e^{\log(v_2) N_T^2} \middle| \mathcal{F}_t \right] \\ &= \frac{1}{v_1^{N_t^1} v_2^{N_t^2}} \mathbb{E} \left[e^{\langle (\underline{0}^\top, \log(v_1), \log(v_2))^\top, Z_T \rangle} \middle| \mathcal{F}_t \right], \end{aligned} \quad (2.43)$$

where we recall that $Z_T = (X_T, \lambda_T^1, \lambda_T^2, N_T^1, N_T^2)^\top$. Then, the conditional probability generating function can be obtained as a direct consequence of Corollary 2.14. In fact, the expected value (2.43) is exactly the left-hand side of (2.34) if we consider $u = (\underline{0}^\top, \log(v_1), \log(v_2))^\top \in \mathbb{R}_-^{d+4}$.

It follows that

$$\begin{aligned} \mathbb{E} \left[v_1^{N_{t+h}^1 - N_t^1} v_2^{N_{t+h}^2 - N_t^2} \middle| \mathcal{F}_t \right] &= \frac{1}{v_1^{N_t^1} v_2^{N_t^2}} e^{\tilde{\phi}_{v_1 v_2}(t, T, \underline{0}) + \langle (\tilde{\psi}_{v_1 v_2}(t, T, \underline{0}), \log(v_1), \log(v_2))^\top, Z_t \rangle} \\ &= e^{\tilde{\phi}_{v_1 v_2}(t, T, \underline{0}) + \langle \tilde{\psi}_{v_1 v_2}(t, T, \underline{0}), (X_t^\top, \tilde{\lambda}_t^1, \tilde{\lambda}_t^2)^\top \rangle} \end{aligned} \quad (2.44)$$

for the functions $\tilde{\phi}_{v_1 v_2}$ and $\tilde{\psi}_{v_1 v_2}$ which satisfy (2.40) and (2.41), respectively.

Differentiating the probability generating function we can recover all probabilities $p_t(n, s, T)$ for every $n \in \mathbb{N}$ (see [56] (Section 3.3, Remark 2)). In fact, $p_t(n, s, T)$ is represented by the coefficient of $v_1^n v_2^{n-s/c}$ in the expansion of (2.44) into a power series. \square

Then, the conditional probability of SLB 's future jumps is determined by

$$\mathbb{P}(\Delta SLB_T = s | \mathcal{F}_t) = \sum_{n=\frac{s}{c}}^{\infty} p_t(n, s, T) \mathbf{1}_{\{s>0\}} + \sum_{n=0}^{\infty} p_t(n, s, T) \mathbf{1}_{\{s \leq 0\}}$$

for $p_t(n, s, T)$ given in (2.39).

Corollary 2.19. *If K_1 is an upper triangular matrix, then it holds that*

$$\mathbb{P}(SLB_s = SLB_t \text{ for all } t \leq s \leq T | \mathcal{F}_t) = e^{\tilde{\phi}(t, T, \underline{0}) + \langle \tilde{\psi}(t, T, \underline{0}), (X_t^\top, \tilde{\lambda}_t^1, \tilde{\lambda}_t^2)^\top \rangle}, \quad (2.45)$$

where the function $\tilde{\phi}$ is determined by

$$\begin{aligned} \tilde{\phi}(t, T, \underline{0}) &= \int_t^T \langle K_0, \tilde{\psi}|_d(s, T, \underline{0}) \rangle ds \\ &\quad + \int_t^T \left(\gamma_1 \lambda_0^1 \tilde{\psi}_{d+1}(s, T, \underline{0}) + \gamma_2 \lambda_0^2 \tilde{\psi}_{d+2}(s, T, \underline{0}) \right) ds \end{aligned} \quad (2.46)$$

and $\tilde{\psi}$ satisfy the following system:

$$\frac{d\tilde{\psi}}{dt}(t, T, \underline{0}) = \begin{pmatrix} -K_1^\top \tilde{\psi}|_d(t, T, \underline{0}) \\ \gamma_1 \tilde{\psi}_{d+1}(t, T, \underline{0}) \\ \gamma_2 \tilde{\psi}_{d+2}(t, T, \underline{0}) \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \tilde{\psi}^2|_d(t, T, \underline{0}) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \Lambda_1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \Lambda_2 \\ 0 \\ 1 \end{pmatrix}. \quad (2.47)$$

Remark 2.20. From the system (2.47) it follows that the last two components of $\tilde{\psi}$ verify two easily solvable independent ODEs and therefore they are determined by

$$\begin{aligned} \tilde{\psi}_{d+1}(t, T, \underline{0}) &= \frac{1}{\gamma_1} (e^{-\gamma_1(T-t)} - 1), \\ \tilde{\psi}_{d+2}(t, T, \underline{0}) &= \frac{1}{\gamma_2} (e^{-\gamma_2(T-t)} - 1). \end{aligned}$$

Proof. The result is a direct consequence of Proposition [2.17](#) since the following events' equality holds:

$$\mathcal{S} := \{SLB_s = SLB_t \text{ for all } t \leq s \leq T\} = \{\Delta N_T^1 = 0, \Delta N_T^2 = 0\} =: \mathcal{N}.$$

Clearly, $\mathcal{S} \supseteq \mathcal{N}$ due to the SLB 's definition. We show the opposite inclusion by contradiction: if $\Delta N_T^1 \neq 0$ (or equivalently $\Delta N_T^2 \neq 0$), then there exists $t \leq \bar{s} \leq T$ such that $\Delta N_{\bar{s}}^1 = N_{\bar{s}}^1 - N_t^1 \neq 0$ ($\Delta N_{\bar{s}}^2 \neq 0$). It necessarily follows that $SLB_{\bar{s}} \neq SLB_t$ since the possibility of simultaneous occurrence of two jumps is ruled out.

In view of this equality we can compute the conditional probability that the stochastic lower bound SLB remains unchanged in the time interval $[t, T]$ since

$$\mathbb{P}(SLB_s = SLB_t \text{ for all } t \leq s \leq T | \mathcal{F}_t) = p_t(0, 0, T). \quad (2.48)$$

If K_1 is an upper triangular matrix, from [\(2.39\)](#) it results that

$$\mathbb{P}(SLB_s = SLB_t \text{ for all } t \leq s \leq T | \mathcal{F}_t) = \lim_{\substack{v_1 \rightarrow 0 \\ v_2 \rightarrow 0}} e^{\tilde{\phi}_{v_1 v_2}(t, T, \mathbf{0}) + \langle \tilde{\psi}_{v_1 v_2}(t, T, \mathbf{0}), (X_t^\top, \tilde{\lambda}_t^1, \tilde{\lambda}_t^2)^\top \rangle},$$

for the C^∞ -functions $\tilde{\phi}_{v_1 v_2}$ and $\tilde{\psi}_{v_1 v_2}$ which verify [\(2.40\)](#) and [\(2.41\)](#), respectively. Since

$$\begin{aligned} \lim_{v_1 \rightarrow 0} e^{\delta_1(\tilde{\psi}_{v_1 v_2})_{d+1}(t, T, \mathbf{0}) + \log(v_1)} &= 0, \\ \lim_{v_2 \rightarrow 0} e^{\delta_2(\tilde{\psi}_{v_1 v_2})_{d+2}(t, T, \mathbf{0}) + \log(v_2)} &= 0 \end{aligned}$$

we derive that the conditional probability [\(2.48\)](#) is determined by

$$\mathbb{P}(SLB_s = SLB_t \text{ for all } t \leq s \leq T | \mathcal{F}_t) = e^{\tilde{\phi}(t, T, \mathbf{0}) + \langle \tilde{\psi}(t, T, \mathbf{0}), (X_t^\top, \tilde{\lambda}_t^1, \tilde{\lambda}_t^2)^\top \rangle},$$

for the functions $\tilde{\phi}$ and $\tilde{\psi}$ which satisfy [\(2.46\)](#) and [\(2.47\)](#), respectively. \square

2.4 Equivalent changes of measures

The short rate model introduced so far has been completely described under the physical probability measure P . We point out that the properties of our framework described in Section 2.2 may be lost under a change of measure. Therefore, we need to study the behavior of the model under a change of probability measure, from the physical P to an equivalent probability measure Q . In particular, we aim at studying sufficient conditions under which the change of measure preserves the affine structure of the short rate model. In other words, our purpose is to investigate the structure of the process $(Z_t)_{t \geq 0} = (X_t^\top, \tilde{\lambda}_t^1, \tilde{\lambda}_t^2, N_t^1, N_t^2)_{t \geq 0}$ under an equivalent change of the probability measure, as a consequence of Theorem [2.4](#).

Since throughout this section we deal with more than one probability measure defined on the same measurable space (Ω, \mathcal{F}) , we want to make the notation clear. In order to distinguish we denote by $\mathbb{E}[\cdot]$ the expectation with respect to the physical probability measure P and by $\mathbb{E}^Q[\cdot]$ the expectation with respect to any other probability measure Q . Similarly, $(W_t)_{t \geq 0}$ is the P -Brownian motion, while $(W_t^Q)_{t \geq 0}$ denotes a Brownian motion with respect to Q .

2.4.1 The exponentially affine change of measure

We focus our attention on the driving element of the short rate model, i.e. the jump-diffusion process $(Z_t)_{t \geq 0}$ which satisfies the SDE (2.8) with respect to the physical measure P . Our goal consists in characterizing an equivalent probability measures Q on (Ω, \mathcal{F}) which ensures that $(Z_t)_{t \geq 0}$ maintains its affine structure. It follows that $(Z_t)_{t \geq 0}$ will satisfy an SDE of the type (A.12) on $(\Omega, \mathcal{F}, \mathbb{F}, Q)$ with respect to an \mathbb{R}^{d+4} -valued Brownian motion $(W_t^Q)_{t \geq 0}$ and for some parameters which fulfill the admissibility conditions of Assumption A.12. Then, as a consequence of Theorem 2.4 our model is an affine short rate model also under the probability measure Q .

We fix a time horizon $T > 0$. Similar to the density process introduced in [24] (Appendix C) and recalled in Appendix A (Section A.3.3), we define a positive process of the form

$$L_t := \frac{1}{C} e^{k \int_0^t r(Z_s, s) ds} e^{\alpha(t) + \langle \beta(t), Z_t \rangle} \quad \text{for } 0 \leq t \leq T, \quad (2.49)$$

where $C := e^{\alpha(0) + \langle \beta(0), Z_0 \rangle}$ is a normalizing constant, $k \in \mathbb{R}$ and $\alpha : [0, T] \rightarrow \mathbb{R}$ and $\beta : [0, T] \rightarrow \mathbb{R}^{d+4}$ are C^1 -functions.

Theorem 2.21. *Let $(L_t)_{t \in [0, T]}$ be the process defined by (2.49) satisfying for $0 \leq t \leq T$ the following integrability conditions:*

1. $\mathbb{E} \left[e^{k \int_0^t (\langle \ell, X_s \rangle + cN_s^1 - cN_s^2) ds} e^{\langle \beta(t), Z_t \rangle} \right] < \infty;$
2. $\mathbb{E} \left[\int_0^t L_s (e^{\delta_1 \beta_{d+1}(s) + \beta_{d+3}(s)} - 1) (\Lambda_1 X_s + \tilde{\lambda}_s^1) ds \right] + \mathbb{E} \left[\int_0^t L_s (e^{\delta_2 \beta_{d+2}(s) + \beta_{d+4}(s)} - 1) (\Lambda_2 X_s + \tilde{\lambda}_s^2) ds \right] < \infty;$
3. $\mathbb{E} \left[\left(\sum_{i=1}^d \int_0^t L_s^2 \beta_i(s)^2 (X_s)_i ds \right)^{\frac{1}{2}} \right] < \infty,$

for $\alpha : [0, T] \rightarrow \mathbb{R}$ and $\beta : [0, T] \rightarrow \mathbb{R}^{d+4}$ which solve the ODEs system:

$$\begin{aligned} \frac{d\alpha}{dt}(t) &= -k\rho_0(t) - \langle \bar{K}_0, \beta(t) \rangle \\ \frac{d\beta}{dt}(t) &= -k\rho_1 - \bar{K}_1^\top \beta(t) - \frac{1}{2} \beta(t)^\top \bar{H}_1 \beta(t) \\ &\quad - \left(e^{\delta_1 \beta_{d+1}(t) + \beta_{d+3}(t)} - 1 \right) l_1^1 - \left(e^{\delta_2 \beta_{d+2}(t) + \beta_{d+4}(t)} - 1 \right) l_1^2, \end{aligned} \quad (2.50)$$

where $\beta(t)^\top \bar{H}_1 \beta(t)$ is the \mathbb{R}^{d+4} -valued vector with k -element $\sum_{i,j=1}^{d+4} \beta_i(t) (\bar{H}_1)_{ijk} \beta_j(t)$. We recall that $l_1^1 = (\Lambda_1^\top, 1, 0, 0, 0)^\top$ and $l_1^2 = (\Lambda_2^\top, 0, 1, 0, 0)^\top$.

Let Q be a probability measure on (Ω, \mathcal{F}) such that $\frac{dQ}{dP} = L_T$. Then, the process $(Z_t)_{t \in [0, T]}$ is an affine jump-diffusion process under the probability measure Q with char-

acteristics $\chi^Q = (\bar{K}^Q, \bar{H}^Q, l^Q, \theta^Q)$ uniquely specified by

$$\begin{aligned}
\bar{K}_0^Q(t) &= \bar{K}_0^Q = \bar{K}_0; \\
\bar{K}_1^Q(t) &= \bar{K}_1 + \bar{H}_1\beta(t); \\
\bar{H}_0^Q(t) &= \mathbf{0}; \\
\bar{H}_1^Q(t) &= \bar{H}_1^Q = \bar{H}_1; \\
l_0^a(t) &= 0 \text{ for } a = 1, 2; \\
l_1^a(t) &= l_1^a\theta^a(\zeta_a\beta(t)) \text{ for } a = 1, 2; \\
\theta^a(c, t) &= \frac{\theta^a(c + \zeta_a\beta(t))}{\theta^a(\zeta_a\beta(t))} \text{ for } c \in \mathbb{C}^{d+4}, a = 1, 2.
\end{aligned} \tag{2.51}$$

Remark 2.22. We have made the structure of the process $(Z_t)_{t \in [0, T]}$ more flexible accommodating time dependence in its coefficients. In fact, under the probability measure Q the process is affine but it is described by time-dependent characteristics, whereas under the physical measure P the characteristics are time-independent.

Remark 2.23. We highlight that the parameter \bar{K}_0 remains the same even after the change of measure, i.e. $\bar{K}_0^Q = \bar{K}_0$, and it is a consequence of the fact that $\bar{H}_0 = \mathbf{0}$ which is the null matrix of dimension $((d+4) \times (d+4))$ (see (A.31)).

We rewrite the previous ODEs system (2.50) componentwise:

$$\begin{aligned}
\frac{d\alpha}{dt}(t) &= -kh(t) - \langle K_0, \beta|_d(t) \rangle - \gamma_1\lambda_0^1\beta_{d+1}(t) - \gamma_2\lambda_0^2\beta_{d+2}(t) \\
\frac{d\beta_i}{dt}(t) &= -kl_i - \sum_{j=1}^d (K_1)_{ji}\beta_j(t) - \frac{1}{2}\beta_i^2(t) \quad \text{for } i = 1, \dots, d \\
&\quad - \left(e^{\delta_1\beta_{d+1}(t)+\beta_{d+3}(t)} - 1 \right) (\Lambda_1)_i - \left(e^{\delta_2\beta_{d+2}(t)+\beta_{d+4}(t)} - 1 \right) (\Lambda_2)_i \\
\frac{d\beta_{d+1}}{dt}(t) &= \gamma_1\beta_{d+1}(t) - e^{\delta_1\beta_{d+1}(t)+\beta_{d+3}(t)} + 1 \\
\frac{d\beta_{d+2}}{dt}(t) &= \gamma_2\beta_{d+2}(t) - e^{\delta_2\beta_{d+2}(t)+\beta_{d+4}(t)} + 1 \\
\frac{d\beta_{d+3}}{dt}(t) &= -kc \\
\frac{d\beta_{d+4}}{dt}(t) &= kc.
\end{aligned} \tag{2.52}$$

From this we notice that the ODEs system defined by (2.50) is similar to that describing the discounted Laplace transform in Proposition 2.5 (see (2.16) and (2.17)). It is a consequence of the fact that the choice of the density process (2.49) is analogous to that made in (A.30); we have only to change k with -1 and the functions α and β with $\phi(t, T, u)$ and $\psi(t, T, u)$, respectively, for $(u, T) \in \mathbb{C}^{d+4} \times \mathbb{R}_+$ such that the discounted characteristics $\chi^d = (\bar{K}, \bar{H}, l^1, l^2, \theta^1, \theta^2, \rho)$ are *well-behaved* at (u, T) . Then, since the

density process defined by (A.30) is employed in the proof of Theorem A.16 (and consequently of Proposition 2.5) the equivalence between the two ODEs systems follows, up to the stated change.

Proof of Theorem 2.21. We aim at proving that $(L_t)_{t \in [0, T]}$ is a positive P -martingale under the required integrability conditions, and therefore may define an equivalent probability measure Q uniquely determined by the Radon-Nikodym derivative $\frac{dQ}{dP} = L_T$.

For all $0 \leq t \leq T$

$$\begin{aligned} \mathbb{E}[|L_t|] &= \mathbb{E}[L_t] = \frac{1}{C} \mathbb{E} \left[e^{k \int_0^t r(Z_s, s) ds} e^{\alpha(t) + \langle \beta(t), Z_t \rangle} \right] \\ &= \frac{1}{C} \mathbb{E} \left[e^{k \int_0^t (\rho_0(s) + \langle \rho_1, Z_s \rangle) ds} e^{\alpha(t) + \langle \beta(t), Z_t \rangle} \right] \\ &= \frac{1}{C} e^{k \int_0^t \rho_0(s) ds + \alpha(t)} \mathbb{E} \left[e^{k \int_0^t \langle \rho_1, Z_s \rangle ds} e^{\langle \beta(t), Z_t \rangle} \right]. \end{aligned}$$

If we replace the expressions of $\rho_0(s)$ and ρ_1 made explicit by (2.13), then we obtain

$$\mathbb{E}[L_t] = \frac{1}{C} e^{k \int_0^t h(s) ds + \alpha(t)} \mathbb{E} \left[e^{k \int_0^t (\langle \ell, Z_s \rangle + cN_s^1 - cN_s^2) ds} e^{\langle \beta(t), Z_t \rangle} \right]$$

Since $Z_t|_d = X_t$, in order that $\mathbb{E}[L_t] < \infty$ we have to verify that

$$\mathbb{E} \left[e^{k \int_0^t (\langle \ell, X_s \rangle + cN_s^1 - cN_s^2) ds} e^{\langle \beta(t), Z_t \rangle} \right] < \infty \quad (2.53)$$

which is exactly the first integrability condition of the Theorem. Then, $(L_t)_{t \in [0, T]}$ is a P -martingale if $\mathbb{E}[L_t | \mathcal{F}_s] = L_s$ for $0 \leq s \leq t \leq T$. By applying Itô's formula we derive that

$$\begin{aligned} L_t - L_0 &= \int_0^t L_s \left[kr(Z_s, s) + \alpha'(s) + \langle \beta'(s), Z_s \rangle \right] ds \\ &\quad + \int_0^t L_s \beta(s)^\top dZ_s^c \\ &\quad + \frac{1}{2} \int_0^t L_s \langle \beta(s), \bar{\sigma}(Z_s) \bar{\sigma}(Z_s)^\top \beta(s) \rangle ds + \sum_{0 < s \leq t} \Delta L_s. \end{aligned}$$

Recalling (2.8) it follows that

$$\begin{aligned} L_t - L_0 &= \int_0^t L_s \left[kr(Z_s, s) + \alpha'(s) + \langle \beta'(s), Z_s \rangle \right. \\ &\quad \left. + \langle \bar{\mu}(Z_s), \beta(s) \rangle + \frac{1}{2} \left\| \bar{\sigma}(Z_s)^\top \beta(s) \right\|^2 \right] ds \\ &\quad + \int_0^t L_s \beta(s)^\top \bar{\sigma}(Z_s) dW_s + \sum_{a=1}^2 \sum_{0 < T_n^a \leq t} (L_{T_n^a} - L_{T_n^a-}), \end{aligned} \quad (2.54)$$

where we remind that T_n^a denotes the n -th jump time of the counting process $(N_t^a)_{t \in [0, T]}$. It follows that

$$\begin{aligned} \mathbb{E} \left[\sum_{a=1}^2 \sum_{t < T_n^a \leq s} (L_{T_n^a} - L_{T_n^{a-}}) \middle| \mathcal{F}_t \right] &= \\ &= \mathbb{E} \left[\sum_{a=1}^2 \sum_{t < T_n^a \leq s} \mathbb{E} \left[L_{T_n^a} - L_{T_n^{a-}} \middle| N_{T_n^a}^a, T_n^a \right] \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[\sum_{a=1}^2 \sum_{t < T_n^a \leq s} \mathbb{E} \left[L_{T_n^a} - \left(e^{\langle \beta(T_n^a), \zeta_a \Delta J_{T_n^a}^a \rangle} - 1 \right) \middle| N_{T_n^a}^a, T_n^a \right] \middle| \mathcal{F}_t \right]. \end{aligned}$$

The first equality is a consequence of the fact that $L_{T_n^a} - L_{T_n^{a-}}$ is generated only by the jump of the process $(N_t^a)_{t \in [0, T]}$ at time T_n^a and, therefore, it does not depend on the whole history \mathcal{F}_t . In the third line we have made the difference $L_{T_n^a} - L_{T_n^{a-}}$ more explicit. We use the definition of jump transform:

$$\begin{aligned} \mathbb{E} \left[\sum_{a=1}^2 \sum_{t < T_n^a \leq s} (L_{T_n^a} - L_{T_n^{a-}}) \middle| \mathcal{F}_t \right] &= \\ &= \mathbb{E} \left[\sum_{a=1}^2 \sum_{t < T_n^a \leq s} L_{T_n^a} - (\theta^a(\zeta_a \beta(T_n^a)) - 1) \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[\sum_{a=1}^2 \int_t^s L_{v-} (\theta^a(\zeta_a \beta(v)) - 1) dN_v^a \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[\sum_{a=1}^2 \int_t^s L_{v-} (\theta^a(\zeta_a \beta(v)) - 1) \lambda^a(Z_v) dv \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[\sum_{a=1}^2 \int_t^s L_v (\theta^a(\zeta_a \beta(v)) - 1) \lambda^a(Z_v) dv \middle| \mathcal{F}_t \right]. \end{aligned}$$

The third equality follows from [\[15\]](#) (Theorem T8) because the a -th counting process has intensity $(\bar{\lambda}^a(Z_t))_{t \in [0, T]}$ and $(\kappa_{t-}(\theta(\zeta_a \psi(t, T, u)) - 1))_{t \in [0, T]}$ is an \mathcal{F}_t -predictable process. However, in order to apply that theorem for $0 \leq t \leq T$ we have to check that

$$\mathbb{E} \left[\sum_{a=1}^2 \int_0^t L_s |\theta^a(\zeta_a \beta(s)) - 1| \lambda^a(Z_s) ds \right] < \infty. \quad (2.55)$$

If we insert [\(2.9\)](#) in [\(2.55\)](#) it derives that

$$\begin{aligned} \mathbb{E} \left[\int_0^t L_s \left(\left| \theta^1((\underline{0}^\top, \delta_1 \beta_{d+1}(s), 0, \beta_{d+3}(s), 0)^\top) - 1 \right| \lambda^1(Z_s) \right. \right. \\ \left. \left. + \left| \theta^2((\underline{0}^\top, 0, \delta_2 \beta_{d+2}(s), 0, \beta_{d+4}(s))^\top) - 1 \right| \lambda^2(Z_s) \right) ds \right] < \infty, \end{aligned}$$

where $\underline{0}$ is the null vector of \mathbb{R}^d . Then, using (2.12) for the definitions of the jump transforms θ^1 and θ^2 and (2.10) for the intensities' parameters it follows that we have to verify that

$$\mathbb{E} \left[\int_0^t L_s \left(\left| e^{\delta_1 \beta_{d+1}(s) + \beta_{d+3}(s)} - 1 \right| (\Lambda_1 X_s + \tilde{\lambda}_s^1) + \left| e^{\delta_2 \beta_{d+2}(s) + \beta_{d+4}(s)} - 1 \right| (\Lambda_2 X_s + \tilde{\lambda}_s^2) \right) ds \right] < \infty, \quad (2.56)$$

corresponding to the second integrability condition in the statement of the theorem. Going back to (2.54), we add and subtract $\sum_{a=1}^2 \int_0^t L_s (\theta^a(\zeta_a \beta(s)) - 1) \lambda^a(Z_s) ds$:

$$\begin{aligned} L_t - L_0 &= \int_0^t L_s \left[kr(Z_s, s) + \alpha'(s) + \langle \beta'(s), Z_s \rangle + \langle \bar{\mu}(Z_s), \beta(s) \rangle \right. \\ &\quad \left. + \frac{1}{2} \left\| \bar{\sigma}(Z_s)^\top \beta(s) \right\|^2 + \sum_{a=1}^2 (\theta^a(\zeta_a \beta(s)) - 1) \lambda^a(Z_s) \right] ds \\ &\quad + \int_0^t L_s \beta(s)^\top \bar{\sigma}(Z_s) dW_s \\ &\quad + \sum_{a=1}^2 \left[\sum_{0 < T_n^a \leq t} (L_{T_n^a} - L_{T_n^a-}) - \int_0^t L_s (\theta^a(\zeta_a \beta(s)) - 1) \lambda^a(Z_s) ds \right]. \end{aligned}$$

We specify all the affine dependence on $(Z_t)_{t \in [0, T]}$ of the drift parameters and therefore obtain that

$$\begin{aligned} L_t - L_0 &= \int_0^t L_s \left[k\rho_0(s) + \alpha'(s) + \langle \bar{K}_0, \beta(s) \rangle \right. \\ &\quad \left. + \left\langle \left\{ k\rho_1 + \beta'(s) + \bar{K}_1^\top \beta(s) + \frac{1}{2} \beta(s)^\top \bar{H}_1 \beta(s) \right. \right. \right. \\ &\quad \left. \left. \left. + \left(e^{\delta_1 \beta_{d+1}(s) + \beta_{d+3}(s)} - 1 \right) l_1^1 + \left(e^{\delta_2 \beta_{d+2}(s) + \beta_{d+4}(s)} - 1 \right) l_1^2 \right\}, Z_s \right\rangle \right] ds \\ &\quad + \int_0^t L_s \beta(s)^\top \bar{\sigma}(Z_s) dW_s \\ &\quad + \sum_{a=1}^2 \left[\sum_{0 < T_n^a \leq t} (L_{T_n^a} - L_{T_n^a-}) - \int_0^t L_s (\theta^a(\zeta_a \beta(s)) - 1) \lambda^a(Z_s) ds \right]. \end{aligned}$$

It follows that $(L_t)_{t \in [0, T]}$ is a martingale because

$$\mathbb{E} \left[\left(\int_0^t L_s^2 \langle \beta(s)^\top \bar{\sigma}(Z_s), \beta(s)^\top \bar{\sigma}(Z_s) \rangle ds \right)^{\frac{1}{2}} \right] = \mathbb{E} \left[\left(\sum_{i=1}^d \int_0^t L_s^2 \beta_i(s)^2 (X_s)_i ds \right)^{\frac{1}{2}} \right] < \infty$$

and the ODEs system (2.50) is verified.

Then, from the martingale property of the density process $(L_t)_{t \in [0, T]}$ we can conclude that $\frac{dQ}{dP} = L_T$ is well-defined and, consequently, an equivalent probability measure Q on the measurable space (Ω, \mathcal{F}) can be determined.

Under the probability measure Q the affine structure of the process $(Z_t)_{t \in [0, T]}$ is preserved. It also follows that its characteristics satisfy (2.51) under Q and, consequently, the admissibility conditions of Assumption A.12 are fulfilled. We can omit these computations because they are analogous to those made in the proof of Theorem A.17. It is a direct consequence of the fact that the Radon-Nikodym derivative (2.49) is similar to that defined in Appendix A, as we have already noticed. \square

Through the density process $(L_t)_{t \in [0, T]}$ defined by (2.49) we have considered a family of probability measures parameterized by the constant k and functions α and β which preserves the affine structure of the process $(Z_t)_{t \in [0, T]}$ and, therefore, of the short rate model.

After the change of measure the parameters that characterize the affine structure of $(Z_t)_{t \in [0, T]}$ are modified and our purpose is to analyze them in details. Let

$$\bar{W}_t^Q := \bar{W}_t - \int_0^t \bar{\sigma}(Z_s)^\top \beta(s) ds \quad \text{for } 0 \leq t \leq T. \quad (2.57)$$

By computations analogous to those made in the proof of Theorem A.17 it follows that (2.57) defines an \mathbb{R}^{d+4} -valued Brownian motion under Q . Then, from (2.8) we can write the SDE satisfied by $(Z_t)_{t \in [0, T]}$ under the new probability measure Q :

$$dZ_t = \bar{\mu}^Q(Z_t, t) dt + \bar{\sigma}(Z_t) d\bar{W}_t^Q + \sum_{a=1}^2 \zeta_a dJ_t^a, \quad Z_0 = z_0 := (x_0^\top, \lambda_0^1, \lambda_0^2, 0, 0)^\top, \quad (2.58)$$

where

$$\begin{aligned} \mu^Q(z, t) &= \bar{K}_0^Q(t) + \bar{K}_1^Q(t)z = \bar{K}_0 + (\bar{K}_1 + \bar{H}_1\beta(t))z \\ &= \begin{pmatrix} K_0 \\ \gamma_1 \lambda_0^1 \\ \gamma_2 \lambda_0^2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} K_1 + \text{diag}((\beta_1(t), \dots, \beta_d(t))^\top) & 0 & 0 & 0 & 0 \\ 0 & -\gamma_1 & 0 & 0 & 0 \\ 0 & 0 & -\gamma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} z, \end{aligned}$$

because for the definition of \bar{H}_1 the non-null elements of the $((d+4) \times (d+4))$ matrix $\bar{H}_1\beta(t)$ are $\langle (\bar{H}_1)_{ii}, \beta(t) \rangle = \beta_i(t)$ for $i = 1, \dots, d$.

From (2.51) we can compute the Q -intensity process of the jump process $(J_t^a)_{t \in [0, T]}$, for $a = 1, 2$. It is determined by

$$\lambda^{aQ}(z, t) = \theta^a(\zeta_a \beta(t)) \lambda^a(z).$$

For $a = 1$ we obtain that

$$\begin{aligned} \lambda^{1Q}(Z_t, t) &= e^{\delta_1 \beta_{d+1}(t) + \beta_{d+3}(t)} \langle l_1^1, Z_t \rangle \\ &= e^{\delta_1 \beta_{d+1}(t) + \beta_{d+3}(t)} (\langle \Lambda_1, X_t \rangle + \tilde{\lambda}_t^1), \end{aligned}$$

meanwhile, for $a = 2$

$$\begin{aligned}\lambda^{2Q}(Z_t, t) &= e^{\delta_2\beta_{a+2}(t)+\beta_{a+4}(t)} \langle l_1^2, Z_t \rangle \\ &= e^{\delta_2\beta_{a+2}(t)+\beta_{a+4}(t)} (\langle \Lambda_2, X_t \rangle + \tilde{\lambda}_t^2).\end{aligned}$$

We recall that $\tilde{\lambda}_t^a = \lambda_0^a + \delta_a \int_0^t e^{-\gamma_a(t-u)} dN_u^a$, for $a = 1, 2$ (see (2.6)).

We conclude that the affine structure of the model is preserved under the change of probability measure if the hypotheses of Theorem 2.21 are satisfied. Therefore, the properties of our framework described in Section 2.2 are unchanged under Q and, as a consequence, we can obtain an explicit expression for the conditional characteristic function of the joint process $(Z_t)_{t \in [0, T]}$ under both the physical P and another equivalent probability measure Q (see Proposition 2.5).

2.4.2 A more general change of measure

To determine a new probability measure equivalent to the physical P we have chosen the process $(L_t)_{t \in [0, T]}$ defined in (2.49) as density process since it arises naturally when we price interest rate derivatives. Now our goal is to generalize that approach. To this purpose we fix a time horizon $T > 0$ and introduce the process $(\xi_t)_{t \in [0, T]}$ which satisfies the following SDE

$$\frac{d\xi_t}{\xi_{t-}} = \theta_t^\top dW_t + \gamma_t^1 dM_t^1 + \gamma_t^2 dM_t^2, \quad \xi_0 = 1, \quad (2.59)$$

where $M_t^a := N_t^a - \int_0^t \lambda_s^a ds$ denotes the compensated counting process, for $a = 1, 2$, while $(\theta_t)_{t \in [0, T]}$, $(\gamma_t^1)_{t \in [0, T]}$ and $(\gamma_t^2)_{t \in [0, T]}$ are integrable processes which take values in \mathbb{R}^d , $(-1, +\infty)$ and $(-1, +\infty)$, respectively. We characterize them in the following Theorem (see [29] (Theorem 3.2) as a reference)).

Theorem 2.24. *Let $(\theta_t)_{t \in [0, T]}$, $(\gamma_t^1)_{t \in [0, T]}$ and $(\gamma_t^2)_{t \in [0, T]}$ be the processes introduced in (2.59). We suppose that $\theta_t = \theta(X_t, t)$ and $\gamma_t^a = \gamma^a(Z_t, t)$, for $a = 1, 2$, such that*

1. $\int_0^t \langle \theta(X_s, s), \theta(X_s, s) \rangle ds < \infty$ P -a.s.;
2. $\int_0^t |\gamma^a(Z_s, s)| \lambda_s^a ds < \infty$ P -a.s..

Furthermore, we require that

$$\sqrt{\text{diag}(X_t)} \theta(X_t, t) = (K_0^Q(t) - K_0) + (K_1^Q(t) - K_1) X_t, \quad (2.60)$$

for functions $K_0^Q : [0, T] \rightarrow \mathbb{R}_+^d$, $K_1^Q : [0, T] \rightarrow \mathbb{R}^{d \times d}$ such that $(K_1^Q)_{ij} : [0, T] \rightarrow \mathbb{R}_+$ for $i, j = 1, \dots, d$ with $i \neq j$, and

$$\gamma^a(t, Z_t) := \frac{(l_0^{aQ}(t) - l_0^a) + \langle (l_1^{aQ}(t) - l_1^a), Z_t \rangle}{l_0^a + \langle l_1^a, Z_t \rangle} \quad (2.61)$$

for $l_0^{aQ} : [0, T] \rightarrow \mathbb{R}_+$ and $l_1^{aQ} : [0, T] \rightarrow \mathbb{R}_+^{d+4}$ such that for $0 \leq t \leq T$ it is verified that $l_0^{aQ}(t) + \langle l_1^{aQ}(t), Z_t \rangle > 0$ P -a.s..

If $\mathbb{E}[\xi_T] = 1$, then we define the probability measure Q via the Radon-Nikodym derivative $\frac{dQ}{dP} = \xi_T$. It is a probability measure equivalent to P such that it preserves the affine structure of the short rate model.

Remark 2.25. If X_t does not have null components, then the $(d \times d)$ matrix $\sqrt{\text{diag}(X_t)}$ is non-singular. Consequently, $\theta(t, X_t)$ is completely determined by

$$\theta(X_t, t) := (\sqrt{\text{diag}(X_t)})^{-1}((K_0^Q(t) - K_0) + (K_1^Q(t) - K_1)X_t),$$

where $^{-1}$ indicates the inverse matrix.

Remark 2.26. We point out that (2.61) is well-defined because the denominator is strictly positive. Indeed, from the definition of Z_t given in (2.7) and the intensities' parameters (2.10) we deduce that

$$l_0^a + \langle l_1^a, Z_t \rangle \geq \tilde{\lambda}_t^a \geq \lambda_0^a > 0.$$

In addition, rearranging terms in (2.61) we notice that $\gamma^a(t, Z_t) > -1$:

$$\gamma^a(Z_t, t) = \frac{l_0^{aQ}(t) + \langle l_1^{aQ}(t), Z_t \rangle}{l_0^a + \langle l_1^a, Z_t \rangle} - 1 > -1$$

because $l_0^{aQ}(t) + \langle l_1^{aQ}(t), Z_t \rangle > 0$. In order to get this condition, it is sufficient to assume that $l_0^{aQ} : [0, T] \rightarrow \mathbb{R}_{++}$ or $(l_1^{aQ})_{d+a} : [0, T] \rightarrow \mathbb{R}_{++}$. If l_0^{aQ} is a \mathbb{R}_{++} -valued function, then the assertion is obviously verified. Otherwise, if $(l_1^{aQ})_{d+a} : [0, T] \rightarrow \mathbb{R}_{++}$ then

$$\langle l_1^{aQ}(t), Z_t \rangle \geq (l_1^{aQ})_{d+a}(t)\tilde{\lambda}_t^a \geq (l_1^{aQ})_{d+a}(t)\lambda_0^a > 0.$$

In the statement of Theorem 2.24 we have required that $(\xi_t)_{t \in [0, T]}$ is a true martingale under the probability measure P . To have this hypothesis satisfied it is sufficient that

$$\mathbb{E} \left[\int_0^T \xi_{s-}^2 \langle \theta(X_s, s), \theta(X_s, s) \rangle ds + \sum_{a=1}^2 \sum_{0 \leq s \leq T} \xi_{s-}^2 (\gamma^a(Z_s, s))^2 \Delta N_s^a \right] < \infty. \quad (2.62)$$

Indeed,

$$\begin{aligned} [\xi, \xi]_T &= [\xi, \xi]_T^c + \sum_{0 \leq s \leq T} (\Delta \xi_s)^2 \\ &= \int_0^T \xi_{s-}^2 \langle \theta(X_s, s), \theta(X_s, s) \rangle ds + \sum_{0 \leq s \leq T} \xi_{s-}^2 (\gamma^1(Z_s, s) \Delta N_s^1 + \gamma^2(Z_s, s) \Delta N_s^2)^2 \\ &= \int_0^T \xi_{s-}^2 \langle \theta(X_s, s), \theta(X_s, s) \rangle ds \\ &\quad + \sum_{0 \leq s \leq T} \xi_{s-}^2 (\gamma^1(Z_s, s))^2 \Delta N_s^1 + \sum_{0 \leq s \leq T} \xi_{s-}^2 (\gamma^2(Z_s, s))^2 \Delta N_s^2. \end{aligned}$$

The last equality is a direct consequence of the fact that the point processes $(N_t^1)_{t \in [0, T]}$ and $(N_t^2)_{t \in [0, T]}$ do not have common jumps by definition. If (2.62) is verified, then $\mathbb{E}[[\xi, \xi]_T] < \infty$ and, consequently, $(\xi_t)_{t \in [0, T]}$ is a square-integrable P -martingale (see [55] (Chapter II, Corollary 3)).

Proof of Theorem 2.24. From [55] (Section II, Theorem 37) we deduce that the unique solution of the SDE (2.59) is

$$\begin{aligned} \xi_t &= e^{\int_0^t \theta(X_s, s)^\top dW_s + \int_0^t \gamma^1(Z_s, s) dM_s^1 + \int_0^t \gamma^2(Z_s, s) dM_s^2 - \frac{1}{2} \int_0^t \langle \theta(X_s, s), \theta(X_s, s) \rangle ds} \\ &\quad \prod_{0 < s \leq t} (1 + \gamma^1(Z_s, s) \Delta N_s^1 + \gamma^2(Z_s, s) \Delta N_s^2) e^{-\gamma^1(Z_s, s) \Delta N_s^1 - \gamma^2(Z_s, s) \Delta N_s^2} \\ &= e^{\int_0^t \theta(X_s, s)^\top dW_s - \int_0^t \gamma^1(Z_s, s) \lambda_s^1 ds - \int_0^t \gamma^2(Z_s, s) \lambda_s^2 ds - \frac{1}{2} \int_0^t \langle \theta(X_s, s), \theta(X_s, s) \rangle ds} \\ &\quad \prod_{0 < s \leq t} (1 + \gamma^1(Z_s, s) \Delta N_s^1 + \gamma^2(Z_s, s) \Delta N_s^2). \end{aligned} \tag{2.63}$$

Hence, we infer that $\xi_t > 0$ because $\gamma^a(Z_t, t) > -1$ and $\Delta N_t^a \in \{0, 1\}$, for $0 \leq t \leq T$ and $a = 1, 2$. We conclude that $(\xi_t)_{t \in [0, T]}$ is a positive P -local martingale and, as a consequence of Fatou's lemma, it is a supermartingale. Therefore, $(\xi_t)_{t \in [0, T]}$ is a true martingale because $\mathbb{E}[\xi_T] = 1$. It follows that $\frac{dQ}{dP} = \xi_T$ is well-defined and a new probability measure Q on the measurable space (Ω, \mathcal{F}) can be determined.

Now, our purpose is to study how this change of measure reflects on the parameters which characterize the model. Let

$$\bar{W}_t^Q := W_t - \int_0^t \theta(X_s, s) ds \quad \text{for } 0 \leq t \leq T. \tag{2.64}$$

For Girsanov's theorem it is a Q -Brownian motion on \mathbb{R}^d . Then, from (2.2) it follows that the SDE satisfied by the process $(X_t)_{t \in [0, T]}$ under the probability measure Q is

$$\begin{aligned} dX_t &= (\mu(X_t) + \sqrt{\text{diag}(X_t)} \theta(X_t, t)) dt + \sqrt{\text{diag}(X_t)} dW_t^Q \\ &= (K_0^Q(t) + K_1^Q(t) X_t) dt + \sqrt{\text{diag}(X_t)} dW_t^Q, \end{aligned} \tag{2.65}$$

where the second equality is a consequence of (2.60). We can conclude that $(X_t)_{t \in [0, T]}$ is an affine diffusion under Q where the time-dependent drift is represented by

$$\mu^Q(x, t) = K_0^Q(t) + K_1^Q(t)x.$$

The hypotheses in the statement of the Theorem on the functions $K_0^Q : [0, T] \rightarrow \mathbb{R}_+^d$ and $K_1^Q : [0, T] \rightarrow \mathbb{R}^{d \times d}$ guarantee that the admissibility conditions of Assumption A.12 are fulfilled.

To analyse the Q -intensity of the counting process $(N_t^a)_{t \in [0, T]}$ for $a = 1, 2$ we refer to [38] (Section III, Theorem 3.11). It follows that

$$\begin{aligned} M_t^{aQ} &= M_t^a - \int_0^t \frac{1}{\xi_{s-}} d\langle \xi, M^a \rangle_s \\ &= \left(N_t^a - \int_0^t \lambda_s^a ds \right) - \int_0^t \gamma_s^a \lambda_s^a ds \\ &= N_t^a - \int_0^t \lambda_s^a (1 + \gamma^a(Z_s, s)) ds \end{aligned}$$

is a Q -local martingale. Then, by the martingale characterization of intensity (see [15] (Chapter II, Theorem T9)) we derive that

$$\begin{aligned} \lambda_t^{aQ} &= (1 + \gamma^a(Z_t, t)) \lambda_t^a \\ &= \left(\frac{l_0^{aQ}(t) + \langle l_1^{aQ}(t), Z_t \rangle}{\lambda_t^a} \right) \lambda_t^a = l_0^{aQ}(t) + \langle l_1^{aQ}(t), Z_t \rangle \end{aligned}$$

is the intensity associated to $(N_t^a)_{t \in [0, T]}$ under the probability measure Q . We observe that the positivity of the Q -intensity process, i.e. $\lambda_t^{aQ} > 0$ for $0 \leq t \leq T$ and $a = 1, 2$, is guaranteed by the hypotheses of the Theorem.

Computations similar to those in the proof of Theorem 2.4 show that Q is a probability measure which preserves the affine structure of the short rate model. This means that under the measure Q the $(d+4)$ -dimensional process $(Z_t)_{t \in [0, T]}$ is an affine jump-diffusion process satisfying the SDE

$$dZ_t = \bar{\mu}^Q(Z_t, t) dt + \bar{\sigma}(Z_t) d\bar{W}_t^Q + \sum_{a=1}^2 \zeta_a dJ_t^a, \quad Z_0 = z_0 := (x_0^\top, \lambda_0^1, \lambda_0^2, 0, 0)^\top, \quad (2.66)$$

where $(\bar{W}_t^Q)_{t \in [0, T]}$ denotes the Brownian motion on \mathbb{R}^{d+4} for which $\bar{W}_t^Q|_d = W_t^Q$ that is defined in (2.64), and the remaining components are identified with independent unidimensional Brownian motions $(W_t^i)_{t \in [0, T]}$, for $i = 1, \dots, 4$.

For every $z \in \mathbb{R}_+^{d+2} \times \mathbb{N}^2$ we identify

$$\mu^Q(z, t) = \bar{K}_0^Q(t) + \bar{K}_1^Q(t)z = \begin{pmatrix} K_0^Q(t) \\ \gamma_1 \lambda_0^1 \\ \gamma_2 \lambda_0^2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} K_1^Q(t) & 0 & 0 & 0 & 0 \\ 0 & -\gamma_1 & 0 & 0 & 0 \\ 0 & 0 & -\gamma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} z$$

and for $a = 1, 2$ the intensity of the jump process $(J_t^a)_{t \in [0, T]}$ is determined by

$$\lambda^{aQ}(z, t) = l_0^{aQ}(t) + \langle l_1^{aQ}(t), z \rangle.$$

□

2.4.3 Comparison between the two density processes

Since we have discussed about two changes of measure, we aim at comparing the two proposed Radon-Nikodym derivatives through which we have defined the probability measures Q equivalent to the physical one P . We recall that the process $(\xi_t)_{t \in [0, T]}$ satisfies

$$\frac{d\xi_t}{\xi_{t-}} = \theta(X_t, t)^\top dW_t + \gamma^1(Z_t, t) dM_t^1 + \gamma^2(Z_t, t) dM_t^2, \quad \xi_0 = 1.$$

On the other hand, from the definition of $(L_t)_{t \in [0, T]}$ given in (2.49) using Itô's formula we obtain that

$$\begin{aligned} \frac{dL_t}{L_{t-}} &= \left(kr_t + \alpha'(t) + \langle \beta'(t), Z_t \rangle + \beta(t)^\top \bar{K}_0 + \beta(t)^\top \bar{K}_1 Z_t + \frac{1}{2} \langle \beta(t), \bar{\sigma}(Z_t) \bar{\sigma}(Z_t)^\top \beta(t) \rangle \right) dt \\ &\quad + \beta(t)^\top \bar{\sigma}(Z_t) d\bar{W}_t + (e^{\langle \beta(t), \Delta Z_t \rangle} - 1) dN_t^1 + (e^{\langle \beta(t), \Delta Z_t \rangle} - 1) dN_t^2 \\ &= \left(-(\theta^1(\zeta_1 \beta(t)) - 1) \langle l_1^1, Z_t \rangle - (\theta^2(\zeta_2 \beta(t)) - 1) \langle l_1^2, Z_t \rangle \right) dt \\ &\quad + \beta(t)^\top \bar{\sigma}(Z_t) d\bar{W}_t + (e^{\delta_1 \beta_{d+1}(t) + \beta_{d+3}(t)} - 1) dN_t^1 + (e^{\delta_2 \beta_{d+2}(t) + \beta_{d+4}(t)} - 1) dN_t^2, \end{aligned}$$

where the second equality is a consequence of the ODEs system (2.50). Computing the jump transforms θ^1 and θ^2 it follows that

$$\frac{dL_t}{L_{t-}} = \beta(t)^\top \bar{\sigma}(Z_t) d\bar{W}_t + (e^{\delta_1 \beta_{d+1}(t) + \beta_{d+3}(t)} - 1) dM_t^1 + (e^{\delta_2 \beta_{d+2}(t) + \beta_{d+4}(t)} - 1) dM_t^2, \quad L_0 = 1.$$

To connect the two Radon-Nikodym derivatives it is sufficient to compare the SDEs that define them since they have the same initial value and satisfy the conditions of Theorem 117 of [57] which guarantees the pathwise uniqueness of SDEs' solutions. Consequently, if for $0 \leq t \leq T$ we consider the following equivalences

$$\begin{aligned} \theta(X_t, t) &= \begin{pmatrix} \beta_1(t) \sqrt{(X_t)_1} \\ \dots \\ \beta_d(t) \sqrt{(X_t)_d} \end{pmatrix} \\ \gamma^1(Z_t, t) &= \gamma^1(t) = e^{\delta_1 \beta_{d+1}(t) + \beta_{d+3}(t)} - 1 \\ \gamma^2(Z_t, t) &= \gamma^2(t) = e^{\delta_2 \beta_{d+2}(t) + \beta_{d+4}(t)} - 1, \end{aligned}$$

then we relate the change of measure characterized by the density process $(\xi_t)_{t \in [0, T]}$ to the exponentially affine one defined by the Radon-Nikodym derivative $\frac{dQ}{dP} = L_T$.

2.5 Fitting the initial term structure

In this section we will define the deterministic function $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ introduced in equation (2.5). Proceeding similarly as in Brigo & Mercurio [17] we will find the exact expression of h in order to fit the term structure of interest rates observed at time $t = 0$.

We suppose that Q is a risk-neutral measure equivalent to P such that $\frac{dP}{dQ} = \xi_T$ for the density process $(\xi_t)_{t \in [0, T]}$ characterized by the dynamics (2.59) for a fixed time

horizon $T > 0$. The measure Q is chosen in such a way that the affine structure of the model is preserved.

We denote the theoretical price at time t of a zero coupon bond with maturity date $T_m \leq T$ and unit face value by $p(t, T_m)$ so that

$$p(t, T_m) = \mathbb{E}^Q \left[e^{-\int_t^{T_m} r(Z_s, s) ds} \right].$$

If we suppose that the discounted characteristics $\chi^{dQ} = (\bar{K}^Q, \bar{H}^Q, l^Q, \theta^Q, \rho)$ are *well-behaved* at $(\underline{0}, T_m) \in \mathbb{R}^{d+4} \times [0, T]$, then from Theorem [A.16](#) it follows that

$$p(t, T_m) = \mathbb{E}^Q \left[e^{-\int_t^{T_m} r(Z_s, s) ds} \right] = e^{\phi(t, T_m, \underline{0}) + \langle \psi(t, T_m, \underline{0}), Z_t \rangle} =: \Pi(t, T_m, Z_t),$$

where $\Pi : \mathbb{R}_+^{d+4} \times \mathbb{N}^2 \rightarrow \mathbb{R}_+$ is a C^1 -function because so are ϕ and ψ . We recall that the function ϕ is defined by

$$\begin{aligned} \phi(t, T_m, \underline{0}) &= - \int_t^{T_m} h(s) ds + \int_t^{T_m} \langle \bar{K}_0^Q(s), \psi(s, T_m, \underline{0}) \rangle ds \\ &\quad + \frac{1}{2} \int_t^{T_m} \langle \psi(s, T_m, \underline{0}), \bar{H}_0^Q(s) \psi(s, T_m, \underline{0}) \rangle ds \\ &\quad + \sum_{a=1}^2 \int_t^{T_m} l_0^{aQ}(s) \left(\theta^{aQ}(\zeta_a \psi(s, T_m, \underline{0}), s) - 1 \right) ds, \end{aligned} \quad (2.67)$$

whereas ψ verifies

$$\begin{aligned} \frac{d\psi}{dt}(t, T_m, \underline{0}) &= (\ell^\top, 0, 0, c, -c)^\top - (\bar{K}_1^Q(t))^\top \psi(t, T_m, \underline{0}) \\ &\quad - \frac{1}{2} \psi(t, T_m, \underline{0})^\top \bar{H}_1^Q(t) \psi(t, T_m, \underline{0}) - \sum_{a=1}^2 l_1^{aQ}(t) (\theta^{aQ}(\zeta_a \psi(t, T_m, \underline{0}), t) - 1), \\ \psi(T) &= \underline{0}. \end{aligned}$$

We assume that the term structure which is currently observed in the market is given by the C^1 -function $p^M(0, \cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that $t \rightarrow p^M(0, t)$. In order to identify h we require that the initial term structure matches the term structure of interest rate observed in the market, therefore the following equality is required to hold:

$$p^M(0, t) = \Pi(0, t, Z_0) \quad \text{for } 0 \leq t \leq T. \quad (2.68)$$

We take into account that a new short rate model is defined considering $(\tilde{r}_t)_{t \in [0, T]}$ in place of $(r_t)_{t \in [0, T]}$, where $r_t - \tilde{r}_t = h(t)$. It follows that

$$\tilde{p}(t, T_m) = \mathbb{E}^Q \left[e^{-\int_t^{T_m} \tilde{r}(Z_s, s) ds} \right] = e^{\tilde{\phi}(t, T_m, \underline{0}) + \langle \tilde{\psi}(t, T_m, \underline{0}), Z_t \rangle} = \tilde{\Pi}(t, T_m, Z_t)$$

for $\tilde{\psi} : [0, T_m] \rightarrow \mathbb{R}^{d+4}$ such that $\tilde{\psi}(t, T_m, \underline{0}) = \psi(t, T_m, \underline{0})$ and $\tilde{\phi} : [0, T_m] \rightarrow \mathbb{R}$ which is determined by

$$\begin{aligned} \tilde{\phi}(t, T_m, \underline{0}) &= \int_t^{T_m} \langle \bar{K}_0^Q(s), \psi(s, T_m, \underline{0}) \rangle ds + \frac{1}{2} \int_t^{T_m} \langle \psi(s, T_m, \underline{0}), \bar{H}_0^Q(s) \psi(s, T_m, \underline{0}) \rangle ds \\ &\quad + \sum_{a=1}^2 \int_t^{T_m} l_0^{aQ}(s) \left(\theta^{aQ}(\zeta_a \psi(s, T_m, \underline{0}), s) - 1 \right) ds. \end{aligned}$$

Looking at (2.67) it follows that $\tilde{\phi}(t, T_m, \underline{0}) - \phi(t, T_m, \underline{0}) = \int_t^{T_m} h(s) ds$ for all $0 \leq t \leq T_m$.

In order to verify (2.68) we derive that

$$\begin{aligned} p^M(0, t) &= \Pi(0, t, Z_0) = \mathbb{E}^Q \left[e^{-\int_0^t r(Z_s, s) ds} \right] \\ &= e^{-\int_0^t h(s) ds} \mathbb{E}^Q \left[e^{-\int_0^t \tilde{r}(Z_s, s) ds} \right] = e^{-\int_0^t h(s) ds} \tilde{\Pi}(0, t, Z_0). \end{aligned}$$

By taking the logarithm of both members the previous equality becomes

$$\begin{aligned} \log(p^M(0, t)) &= - \int_0^t h(s) ds + \log(\tilde{\Pi}(0, t, Z_0)) \\ &= - \int_0^t h(s) ds + \tilde{\phi}(0, t, Z_0) + \langle \psi(0, t, Z_0), Z_0 \rangle. \end{aligned} \tag{2.69}$$

Lastly, adopting the notation of [17] (Section 2) for a maturity $0 \leq t \leq T$ we define by $\tilde{f}(0, t)$ and $f^M(0, t)$ the instantaneous forward rates at time 0 associated respectively to the bond prices $\tilde{p}(0, t)$ and $p^M(0, t)$, that is

$$\begin{aligned} \tilde{f}(0, t) &:= - \frac{d(\log \tilde{p})}{dt}(0, t) = - \frac{d(\log \tilde{\Pi})}{dt}(0, t, Z_0) = -\tilde{\phi}'(0, t, \underline{0}) - \langle \psi'(0, t, \underline{0}), Z_0 \rangle \\ f^M(0, t) &:= - \frac{d(\log p^M)}{dt}(0, t) \end{aligned}$$

If we differentiate (2.69) and employ previous definitions, then

$$-f^M(0, t) = -h(t) - \tilde{f}(0, t).$$

We conclude that the model fits the term structure observed at time $t = 0$ if and only if it holds that

$$\begin{aligned} h(t) &= f^M(0, t) - \tilde{f}(0, t) \\ &= f^M(0, t) + \tilde{\phi}'(0, t, \underline{0}) + \langle \psi'(0, t, \underline{0}), Z_0 \rangle \end{aligned}$$

Substituting the exact expression of the derivatives of $\tilde{\phi}$ and ψ :

$$\begin{aligned}
h(t) = & f^M(0, t) - \langle \bar{K}_0^Q(0), \psi(0, t, \underline{0}) \rangle - \frac{1}{2} \langle \psi(0, t, \underline{0}), \bar{H}_0^Q(0) \psi(0, t, \underline{0}) \rangle \\
& - \sum_{a=1}^2 l_0^{aQ}(0) (\theta^{aQ}(\zeta_a \psi(0, t, \underline{0}), 0) - 1) \\
& + \langle \ell, X_0 \rangle - \langle (\bar{K}_1^Q(0))^\top \psi(0, t, \underline{0}), Z_0 \rangle - \frac{1}{2} \langle \psi(0, t, \underline{0}) \bar{H}_1^Q(0) \psi(0, t, \underline{0}), Z_0 \rangle \\
& - \sum_{a=1}^2 l_1^{aQ}(0) (\theta^a(\zeta_a \psi(0, t, \underline{0}), 0) - 1).
\end{aligned}$$

Chapter 3

Simulations

We recall that the short rate process is defined by (2.5):

$$r_t = h(t) + SLB_t + \langle \ell, X_t \rangle \quad \text{for all } t \geq 0.$$

We intend to simulate this process by means of the software Matlab in order to investigate the empirical behavior of our short rate model. To reduce the computation time we consider a one-dimensional diffusion process $(X_t)_{t \geq 0}$ and set $h(t) = 0$ for all $t \geq 0$ for convenience. It follows that

$$r_t = c(N_t^1 - N_t^2) + \ell X_t.$$

Let $T > 0$ be a fixed time horizon. We intend to replicate the short rate process over the time interval $[0, T]$ simulating the diffusion process $(X_t)_{t \in [0, T]}$ by means of a simple recursive method and the bi-dimensional point process $(N_t)_{t \in [0, T]}$ characterized by the intensity function $(\lambda_t)_{t \in [0, T]}$ determined by (2.3), i.e.

$$\lambda_t^a = \lambda_0^a + \langle \Lambda_a, X_t \rangle + \delta_a \int_0^t e^{-\gamma_a(t-u)} dN_u^a, \quad \text{for } a = 1, 2.$$

3.1 The discretization schemes for the affine diffusion process

The SDE which defines the dynamics of the \mathbb{R}_+ -valued affine diffusion $(X_t)_{t \in [0, T]}$ is represented by (2.2), that is

$$dX_t = (K_0 + K_1 X_t)dt + \sqrt{X_t}dW_t, \quad X_0 = x_0.$$

In the following we will simulate the paths of this process discretizing the time interval $[0, T]$ and simulating its dynamics. In other words, we will only generate the process for a finite number of times $0 = t_0 < t_1 < \dots < t_n = T$ that is the partition of $[0, T]$ into segments of the same length Δt , i.e. $t_i = \frac{iT}{n}$ for each $i = 0, \dots, n$. Doing

so, we approximate the continuous-time process $(X_t)_{t \in [0, T]}$ by a discrete-time process introducing errors into the simulation result. It follows that a large number of time steps may be needed to reduce errors to an acceptable level, thus increasing the computation time. Even if its convergence rate is slow we use a discretization scheme because it is simple, direct and easy to implement.

The natural way to simulate the affine diffusion is the Euler method which is determined by the following recursive scheme:

$$\begin{aligned} X_{t_i} &= X_{t_{i-1}} + (K_0 + K_1 X_{t_{i-1}}) \Delta t + \sqrt{X_{t_{i-1}}} (W_{t_i} - W_{t_{i-1}}) \quad \text{for } i = 1, \dots, n \\ X_{t_0} &= x_0. \end{aligned}$$

Nevertheless, the Euler scheme is not well-defined because it can lead to negative values for which the square root is not defined. To correct this problem Delbaen & Deelstra [22] consider the following scheme:

$$\begin{aligned} X_{t_i} &= X_{t_{i-1}} + (K_0 + K_1 X_{t_{i-1}}) \Delta t + \sqrt{X_{t_{i-1}}^+} (W_{t_i} - W_{t_{i-1}}) \quad \text{for } i = 1, \dots, n \\ X_{t_0} &= x_0, \end{aligned}$$

while Lord et al. in [50] propose the full truncation scheme:

$$\begin{aligned} X_{t_i} &= X_{t_{i-1}} + (K_0 + K_1 X_{t_{i-1}}^+) \Delta t + \sqrt{X_{t_{i-1}}^+} (W_{t_i} - W_{t_{i-1}}) \quad \text{for } i = 1, \dots, n \\ X_{t_0} &= x_0. \end{aligned}$$

Doing so, the schemes may take negative values but they are still well-defined because of the positive part. Another modification of the Euler method is the implicit Euler scheme proposed by Brigo & Alfonsi in [16]:

$$\begin{aligned} X_{t_i} &= X_{t_{i-1}} + (K_0 + K_1 X_{t_i} - \frac{1}{2}) \Delta t + \sqrt{X_{t_i}} (W_{t_i} - W_{t_{i-1}}) \quad \text{for } i = 1, \dots, n \\ X_{t_0} &= x_0. \end{aligned}$$

We see that $\sqrt{X_{t_i}}$ appears as a root of a second degree polynomial. The scheme is well-defined when n is large enough and $K_0 \geq \frac{1}{2}$. For $i = 1, \dots, n$ it holds that

$$X_{t_i} = \left(\frac{(W_{t_i} - W_{t_{i-1}}) + \sqrt{(W_{t_i} - W_{t_{i-1}})^2 + 4(X_{t_{i-1}} + (K_0 - \frac{1}{2})\Delta t)(1 - K_1\Delta t)}}{2(1 - K_1\Delta t)} \right)^2. \quad (3.1)$$

In [2] Alfonsi has proposed another implicit scheme that is obtained from

$$\begin{aligned} \sqrt{X_{t_i}} &= \sqrt{X_{t_{i-1}}} + \left(\frac{K_0 - \frac{1}{4}}{2\sqrt{X_{t_i}}} + \frac{K_1}{2} \sqrt{X_{t_i}} \right) \Delta t + \frac{1}{2} (W_{t_i} - W_{t_{i-1}}) \quad \text{for } i = 1, \dots, n \\ X_{t_0} &= x_0. \end{aligned}$$

Multiplying by $\sqrt{X_{t_i}}$ we observe that $\sqrt{X_{t_i}}$ is a root of a second degree polynomial function. If $K_0 \geq \frac{1}{4}$ and n is large enough we get the following value

$$X_{t_i} = \left(\frac{\frac{1}{2}(W_{t_i} - W_{t_{i-1}}) + \sqrt{X_{t_{i-1}}} + \sqrt{\left(\frac{1}{2}(W_{t_i} - W_{t_{i-1}}) + \sqrt{X_{t_{i-1}}}\right)^2 + 4\left(1 - \frac{K_1}{2}\Delta t\right)\frac{K_0 - 1/4}{2}\Delta t}}{2\left(1 - \frac{K_1}{2}\Delta t\right)} \right)^2.$$

Another discretization method is represented by the Milstein scheme:

$$X_{t_i} = \left(\sqrt{X_{t_{i-1}}} + \frac{1}{2}(W_{t_i} - W_{t_{i-1}}) \right)^2 + \left(K_0 - \frac{1}{4} + K_1 X_{t_{i-1}} \right) \Delta t \quad \text{for } i = 1, \dots, n$$

$$X_{t_0} = x_0.$$

This scheme is well-defined only for $K_1 \geq 0$ and $K_0 \geq \frac{1}{4}$. The following scheme studied in [2] can be seen as a correction to the Milstein scheme:

$$X_{t_i} = \left(\left(1 + \frac{K_1}{2}\Delta t\right) \sqrt{X_{t_{i-1}}} + \frac{(W_{t_i} - W_{t_{i-1}})}{2\left(1 + \frac{K_1}{2}\Delta t\right)} \right)^2 \quad \text{for } i = 1, \dots, n$$

$$X_{t_0} = x_0.$$

It is defined for $K_0 \geq \frac{1}{4}$ and for n large enough.

In all the above proposed schemes the Brownian increments $(W_{t_i} - W_{t_{i-1}})$ are independent, normally distributed with zero mean and standard deviation equal to $\sqrt{\Delta t}$, for every $i = 1, \dots, n$. Therefore, in the Matlab code each of them will be generated by a random variable $U \sim \mathcal{N}(0, \sqrt{\Delta t})$.

3.2 The Ogata's algorithm

To simulate the bi-dimensional counting process $(N_t)_{t \in [0, T]}$ we consider an algorithm given for past-dependent point processes and introduced by Ogata in [53]. This method is based on the thinning algorithm which was developed by Lewis and Shedler in [49] for the simulation of non-homogeneous Poisson processes and relies on the following fundamental result (see [53] (Proposition 1)).

Lemma 3.1. *Consider the bi-dimensional point process $(N_t)_{t \in [0, T]} = ((N_t^1, N_t^2)^\top)_{t \in [0, T]}$ on the time interval $[0, T]$ with intensity $(\lambda_t)_{t \in [0, T]}$. Suppose we can find a one-dimensional \mathcal{F}_t -predictable process $(\lambda_t^*)_{t \in [0, T]}$ satisfying*

$$\lambda_t^1 + \lambda_t^2 \leq \lambda_t^* \quad P\text{-a.s. for } 0 \leq t \leq T.$$

We introduce the one-dimensional point process $(N_t^)_{t \geq 0}$ characterized by the intensity $(\lambda_t^*)_{t \in [0, T]}$ and define*

$$\lambda_t^0 := \lambda_t^* - \lambda_t^1 - \lambda_t^2 \quad \text{for } 0 \leq t \leq T.$$

Let $t_1^, \dots, t_{N_{T^*}}^* \in (0, T]$ be the jump times of the process $(N_t^*)_{t \geq 0}$. To each of them we give a mark $a = 0, 1, 2$ with probability $\frac{\lambda_{t_j^*}^a}{\lambda_{t_j^*}^*}$, $j = 1, \dots, N_{T^*}$. Then the times with marks*

$a = 1, 2$ provide a bivariate point process $(\tilde{N}_t)_{t \in [0, T]}$ such that

$$\tilde{N}_t = N_t \quad P\text{-a.s. for } 0 \leq t \leq T.$$

We are going to construct an algorithm based on this Lemma. The intuition is to generate a faster Poisson process and remove jump times probabilistically so that the remaining times are characterized by the intensity $(\lambda_t)_{t \geq 0} = ((\lambda_t^1, \lambda_t^2)^\top)_{t \geq 0}$ determined by (2.3).

Algorithm 3.1. *Simulation of the bivariate point process $(N_t)_{t \in [0, T]}$.*

<p>Data: <i>model's parameters, x_{max} and T</i></p> <p>Result: <i>jumps times $\tau^1 = \{T_1^1, \dots, T_{n^1}^1\}$ and $\tau^2 = \{T_1^2, \dots, T_{n^2}^2\}$</i></p> <p><i>Initialize $n^1 = n^2 = m = 0$, $s_0 = 0$, $\tau^1 = \tau^2 = \emptyset$;</i></p> <p>while $s_m < T$ do</p> <div style="padding-left: 20px;"> <p>$\lambda_{s_m}^{1*} = \lambda_0^1 + \langle \Lambda_1, x_{max} \rangle + \delta_1 \sum_{s \in \tau^1} e^{-\gamma_1(s_m - s)} + \delta_1$;</p> <p>$\lambda_{s_m}^{2*} = \lambda_0^2 + \langle \Lambda_2, x_{max} \rangle + \delta_2 \sum_{s \in \tau^2} e^{-\gamma_2(s_m - s)} + \delta_2$;</p> <p>$\lambda_{s_m}^* = \lambda_{s_m}^{1*} + \lambda_{s_m}^{2*}$;</p> <p>$U \sim Unif[0, 1]$;</p> <p>$w = -\frac{\log(U)}{\lambda_{s_m}^*}$;</p> <p>$s_{m+1} = s_m + w$;</p> <p>$D \sim Unif[0, 1]$;</p> <p>if $D \leq \frac{\lambda_{s_{m+1}}^1 + \lambda_{s_{m+1}}^2}{\lambda_{s_m}^*}$ then</p> <div style="padding-left: 20px;"> <p>if $\lambda_{s_{m+1}}^1 \geq D \lambda_{s_m}^*$ then</p> <div style="padding-left: 20px;"> <p>$n^1 = n^1 + 1$;</p> <p>$\tau^1 = \tau^1 \cup \{s_{m+1}\}$;</p> </div> <p>else</p> <div style="padding-left: 20px;"> <p>$n^2 = n^2 + 1$;</p> <p>$\tau^2 = \tau^2 \cup \{s_{m+1}\}$;</p> </div> <p>end</p> </div> <p>else</p> <div style="padding-left: 20px;"> <p><i>nothing</i>;</p> </div> <p>end</p> <p>$m = m + 1$;</p> </div> <p>end</p>

The algorithm provides two times' lists $\tau^1 = \{T_1^1, \dots, T_{n^1}^1\}$ and $\tau^2 = \{T_1^2, \dots, T_{n^2}^2\}$ composed by the jump times of the components of the simulated counting process $(N_t)_{t \in [0, T]}$. As input data we supply the time interval $[0, T]$ that we are considering for the simulation and the model's parameters that uniquely identify the point process. In addition, we have to supply the \mathbb{R}_+^d -valued constant x_{max} that represents the maximum value assumed by the affine process $(X_t)_{t \in [0, T]}$. It is the output of a previous simulation in which the Euler discretization method is used.

Remark 3.2. We compute x_{max} in a previous simulation because of the definition of the

intensity process $(\lambda_t^a)_{t \in [0, T]}$. In other words, for $a = 1, 2$

$$\lambda_t^a = \lambda_0^a + \langle \Lambda_a, X_t \rangle + \delta_a \sum_{0 < T_n^a \leq t} e^{-(t - T_n^a)},$$

where the sum is over the jump times in the interval $(0, t]$. Let $0 < \bar{T}_n^a < \bar{T}_{n+1}^a$ be two consecutive jump times of the a -component process. Then, on $[\bar{T}_n^a, \bar{T}_{n+1}^a)$ the intensity λ_t^a is composed by a constant, a random scalar product and a monotonically decreasing time function:

$$\max_{t \in [\bar{T}_n^a, \bar{T}_{n+1}^a)} \lambda_t^a = \lambda_0^a + \max_{t \in [\bar{T}_n^a, \bar{T}_{n+1}^a)} \langle \Lambda_a, X_t \rangle + \delta_a \sum_{0 < T_n^a \leq \bar{T}_n^a} e^{-(\bar{T}_n^a - T_n^a)}.$$

It follows that we have to compute x_{max} at an earlier time in order to correctly define $(\lambda_t^*)_{t \geq 0}$ as in Lemma 3.1.

For the initialization we have to look at the counters n^1 and n^2 for the number of jumps made by $(N_t^1)_{t \in [0, T]}$ and $(N_t^2)_{t \in [0, T]}$ and the counter m which describes the number of simulated jumps. Let s_0 be the first jump time. It will symbolize the candidate jump time. The sets τ^1 and τ^2 in which the jumps of $(N_t)_{t \in [0, T]}$ are stored are empty at the beginning since there have not occurred any jumps at the start of the simulation.

The algorithm generates new candidate jumps s_m as long as $s_m < T$, so as long as the m -th candidate jump time is not bigger than the final time of the considered interval $[0, T]$. At each cycle to generate a jump time the algorithm simulates an homogeneous Poisson process characterized by the intensity $\lambda_{s_m}^*$. Then, the new inter-arrival time w is identified by the first jump time of that Poisson process and therefore the new candidate jump time is equal to the old value plus the generated inter-arrival time, i.e. $s_{m+1} = s_m + w$. Then, by means of the result stated in Lemma 3.1 we can verify if the candidate time s_{m+1} has to be rejected or accepted. If so, it represents the new jump times of the point process $(N_t)_{t \in [0, T]}$ and therefore has to be included in τ^1 or τ^2 .

Generating an uniform variable D we check if the jump time s_{m+1} corresponds to a jump of $(N_t)_{t \in [0, T]}$. If

$$D \leq \frac{\lambda_{s_{m+1}}^1 + \lambda_{s_{m+1}}^2}{\lambda_{s_m}^*},$$

then the generated time belongs to the point process $(N_t)_{t \in [0, T]}$. In other words, employing the notation of Lemma 3.1 the mark $a = 1$ or $a = 2$ is assigned to s_{m+1} . In order to understand which mark belongs to the time s_{m+1} , that is to which component process it belongs, we have to find $\bar{a} \in \{1, 2\}$ such that

$$\sum_{a=1}^{\bar{a}-1} \lambda_{s_{m+1}}^a < D \lambda_{s_m}^* \leq \sum_{a=1}^{\bar{a}} \lambda_{s_{m+1}}^a$$

and therefore assign the jump time s_{m+1} to the component process $(N_t^{\bar{a}})_{t \in [0, T]}$. We raise its counter $n^{\bar{a}}$. We also have to raise the general counter m which numbers the candidate jump times that we have analysed.

We start the cycle over again as long as the last candidate time $s_m < T$. If $s_m \geq T$ the simulation is finished. The only thing we have to do is to check whether the last found jump time is bigger than T .

Remark 3.3. We note that in the first cycle we generate the first point of an homogeneous Poisson process characterized by the intensity $\lambda_{s_0}^*$. Then, in the second one the intensity value is $\lambda_{s_1}^*$ and a new homogeneous Poisson process is generated, and so on. In this way we do not simulate an homogeneous Poisson process on the whole time interval $[0, T]$ but in each cycle we generate a new homogeneous Poisson process which is independent from the previous Poisson process. Therefore, we simulate a non-homogeneous Poisson process on $[0, T]$ (see [15] (Chapter II, Theorem T5)). This counting process is characterized by the intensity process $(\lambda_t^*)_{t \in [0, T]}$ defined by

$$\lambda_t^* = \lambda_{s_m}^* \text{ if } s_m \leq t < s_{m+1}.$$

It satisfies

$$\lambda_t^1 + \lambda_t^2 \leq \lambda_t^* \text{ P-a.s. for } 0 \leq t \leq T,$$

where $((\lambda_t^1, \lambda_t^2)^\top)_{t \in [0, T]}$ is the intensity which defines $((N_t^1, N_t^2)^\top)_{t \in [0, T]}$.

3.3 Running the simulations

Based on the Algorithm 3.1 a Matlab code has been written to simulate the stochastic lower bound $(SLB_t)_{t \in [0, T]}$ and then the short rate process $(r_t)_{t \in [0, T]}$ in order to give an intuition on how the parameters influence the model's structure. The code can be found in Appendix C.

3.3.1 Simulation of the lower bound

We fix a small time interval $[0, 10]$ in which simulations are performed and consider that the counting process $(N_t)_{t \in [0, T]}$ is a pure bivariate Hawkes process. In other words, the parameters $\Lambda_1 = \Lambda_2 = 0$ and therefore the affine diffusion process do not influence the intensity vector which is simply defined by $\lambda_t^a = \lambda_0^a + \delta_a \int_0^t e^{-\gamma_a(t-s)} dN_s^a$, for $a = 1, 2$.

We recall that

$$SLB_t = c(N_t^1 - N_t^2).$$

For the parameters' choice $\lambda_0^1 = \lambda_0^2 = 0.1$, $\delta_1 = 0.2$, $\delta_2 = 0.8$, $\gamma_1 = \gamma_2 = 0.6$ and $c = 0.1$ a path of $(SLB_t)_{t \in [0, 10]}$ is shown in Figure 3.1.

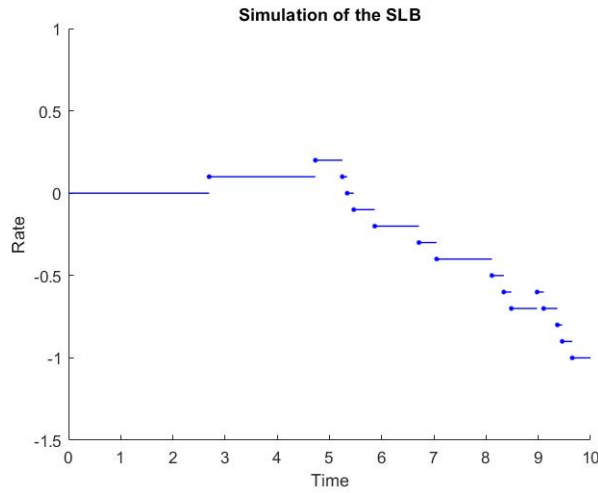


Figure 3.1: Path of $(SLB_t)_{t \in [0,10]}$ for the intensities $\lambda_t^1 = 0.1 + 0.2 \int_0^t e^{-0.6(t-s)} dN_s^1$, $\lambda_t^2 = 0.1 + 0.8 \int_0^t e^{-0.6(t-s)} dN_s^2$.

Although Figure 3.1 performs only a simulation, we can observe that as soon as a jump occurs in the second component process $(N_t^2)_{t \in [0,10]}$ representing downward jumps, then a cluster of downward jumps comes in succession. This feature is not so clear in the first component $(N_t^1)_{t \in [0,10]}$ and this difference is due to the parameter δ which determines the scale of the positive influence of past jumps. Since $\delta_1 < \delta_2$, then a jump of the second component process has more impact on the intensity $(\lambda_t^2)_{t \in [0,T]}$ with respect to a jump of the first component on $(\lambda_t^1)_{t \in [0,T]}$.

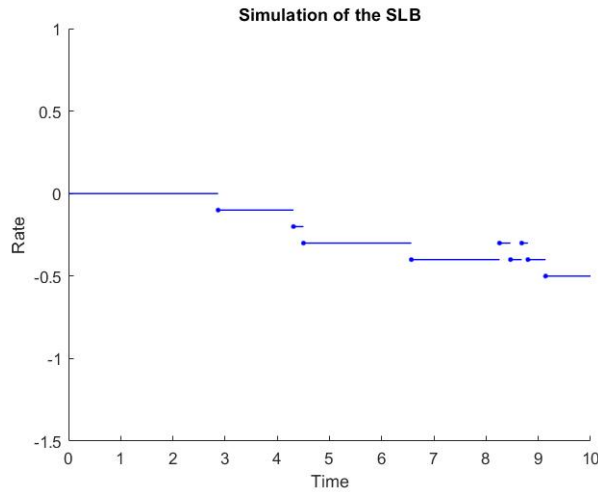


Figure 3.2: Path of $(SLB_t)_{t \in [0,10]}$ for the intensities $\lambda_t^1 = 0.1 + 0.2 \int_0^t e^{-0.6(t-s)} dN_s^1$, $\lambda_t^2 = 0.1 + 0.8 \int_0^t e^{-1.2(t-s)} dN_s^2$.

To investigate the meaning of the parameter γ we run another simulation considering $c = 0.1$, $\lambda_0^1 = \lambda_0^2 = 0.1$, $\delta_1 = 0.2$, $\delta_2 = 0.8$, $\gamma_1 = 0.6$ and $\gamma_2 = 1.2$. With respect to the first parameters' choice γ_2 is doubled. If we compare the second component process drawn in Figure 3.1 with respect to the one represented in Figure 3.2 then we distinctly notice that in the second figure the cluster effect of $(N_t^2)_{t \in [0,10]}$ is less evident. In fact, in the second simulation the decay effect controlled by the parameter γ_2 is amplified.

3.3.2 Model's simulation

The purpose of this section is to investigate how our short rate model behaves in comparison to the time series of the EONIA and ECB deposit facility rates that are shown in Figure 1.1. In order to do this we fix the parameter $c = 0.1$ since the deposit facility rates moves by multiples of 10 basis points. To lower the computation time we focus on a restricted time period, namely from 2 January 2014 to 31 December 2019 since this period is characterized by negative rates. In the following those dates will be identified by T_{min} and T_{max} , respectively.

Firstly, we set $\Lambda_1 = \Lambda_2 = 0$ and simulate the short rate process $(r_t)_{t \in [T_{min}, T_{max}]}$, i.e.

$$r_t = SLB_t + 0.9 \cdot X_t,$$

and $(SLB_t)_{t \in [T_{min}, T_{max}]}$ that is described by the bivariate counting process identified by the following intensities

$$\begin{aligned} \lambda_t^1 &= 0.0001 + 0.00001 \int_0^t e^{-20(t-s)} dN_s^1, \\ \lambda_t^2 &= 0.004 + 0.00001 \int_0^t e^{-20(t-s)} dN_s^2. \end{aligned}$$

We set $\delta_a = 0.00001$ and $\gamma_a = 20$, for $a = 1, 2$, in order to reduce the jumps' positive influence on the intensity value and increase the decay effect since in the reference time interval $[T_{min}, T_{max}]$ the ECB deposit facility rate does not change frequently. In addition, $\lambda_0^1 < \lambda_0^2$ because the ECB deposit facility rate is decreased in four points in time from 2 January 2014 to 31 December 2019, and never increased.

To reproduce the paths of the affine diffusion process $(X_t)_{t \in [T_{min}, T_{max}]}$ we employ the discretization scheme described in (3.1) with discretization step $\Delta t = 0.1$. We use it since in [2] Alfonsi shows that this scheme converges strongly, i.e.

$$\lim_{n \rightarrow 0} \mathbb{E} \left[\max_{0 \leq i \leq n} |\bar{X}_{t_i} - X_{t_i}| \right] = 0,$$

where \bar{X}_{t_i} defines the approximation of X_{t_i} by means of the scheme (3.1). We consider the following dynamics

$$dX_t = (0.5 - 0.01 \cdot X_t)dt + \sqrt{X_t}dW_t, \quad X_0 = 0.169.$$

We choose the parameter $K_0 = 0.5$ since the discretization scheme (3.1) is well-defined only for $K_0 \geq 0.5$ and the parameter $K_1 = -0.01$ in order to ensure mean reversion of

the affine diffusion. We highlight that from this coefficients' choice it follows that the zero value is precluded because $2 \cdot K_0 = 1$ (see for example [1] (Proposition 1.2.15)), and therefore $r_t > SLB_t$ for all $T_{min} \leq t \leq T_{max}$. This is because in the reference time interval the EONIA rate is always greater than the ECB deposit facility rate. Lastly, the initial value X_0 is set equal to 0.169 in order that

$$r_0 = SLB_0 + 0.9 \cdot X_0 = 0.9 \cdot 0.169 = 0.152.$$

That is, the initial value of the simulated short rate process $(r_t)_{t \in [T_{min}, T_{max}]}$ matches the EONIA rate observed at date 2 January 2014 which is the first point in time of our reference interval $[T_{min}, T_{max}]$.

The results of the simulation are shown in Figure 3.3.

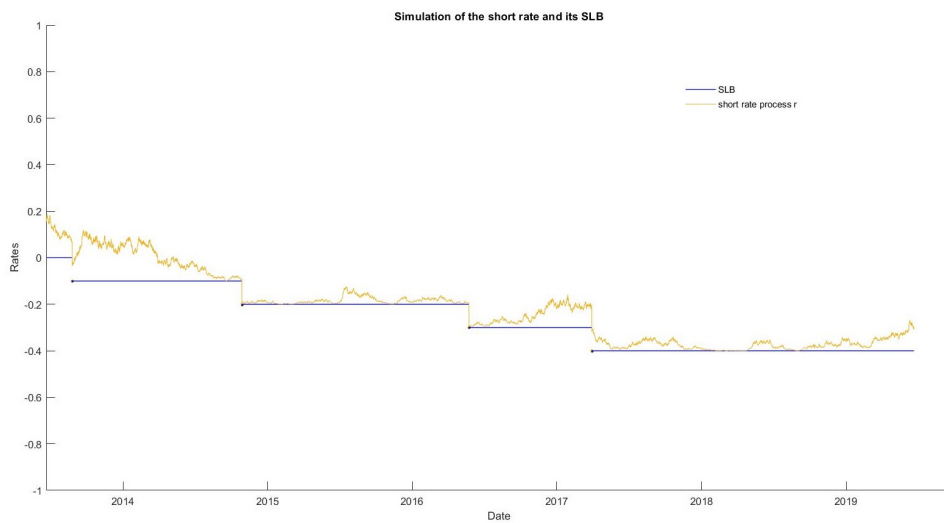


Figure 3.3: Paths of $(r_t)_{t \in [T_{min}, T_{max}]}$ and $(SLB_t)_{t \in [T_{min}, T_{max}]}$ for the intensities $\lambda_t^1 = 10^{-4} + 10^{-5} \int_0^t e^{-20(t-s)} dN_s^1$, $\lambda_t^2 = 4 \cdot 10^{-3} + 10^{-5} \int_0^t e^{-20(t-s)} dN_s^2$.

First of all, Figure 3.3 shows that the model works correctly since it can reproduce negative rates and the stochastic lower bound $(SLB_t)_{t \in [T_{min}, T_{max}]}$ provides a floor for the short rate $(r_t)_{t \in [T_{min}, T_{max}]}$. In addition, their distance is almost constant around 20 basis points in 2014, whereas from 2015 onward it shrinks, and this reflects the real behaviour of the EONIA and ECB deposit facility rates observed in Figure 1.1. Nevertheless, the model does not fit the market values correctly because we did not have calibrate the model's parameters since it is outside the scope of this Thesis. The aim of this simulation is simply to give some visual evidence and intuition of the features which characterize the model that we have introduced.

Now, we set $\Lambda_1 = 0.01$ and $\Lambda_2 = 0.1$. We choose these parameters so that the values assumed by the intensities $(\lambda_t^1)_{t \in [T_{min}, T_{max}]}$ and $(\lambda_t^2)_{t \in [T_{min}, T_{max}]}$ do not become too high. For the same reason we modify the parameters $\lambda_0^1 = 0.00001$ and $\lambda_0^2 = 0.0004$. Then,

the jump intensities become

$$\lambda_t^1 = 0.00001 + 0.01 \cdot X_t + 0.00001 \int_0^t e^{-20(t-s)} dN_s^1,$$

$$\lambda_t^2 = 0.0004 + 0.1 \cdot X_t + 0.00001 \int_0^t e^{-20(t-s)} dN_s^2.$$

We perform another simulation and the results are plotted in Figure [3.4](#).

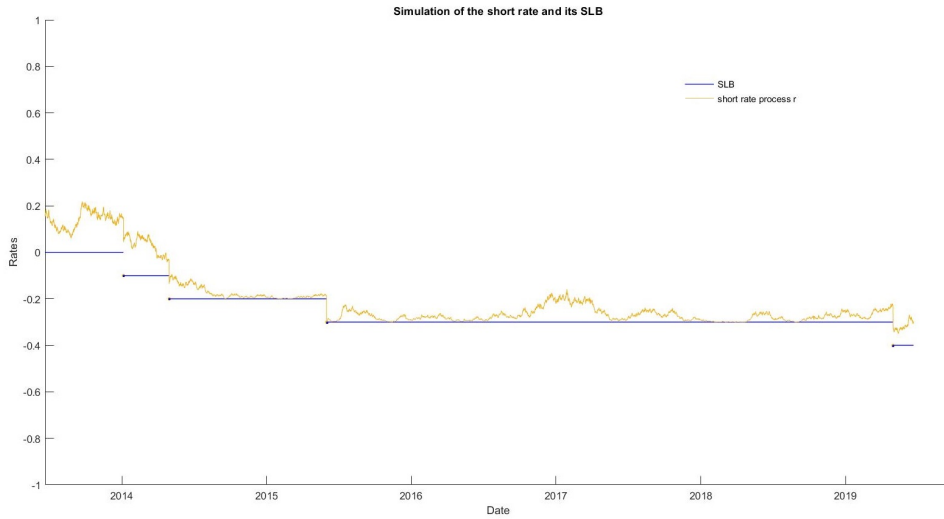


Figure 3.4: Paths of $(r_t)_{t \in [T_{min}, T_{max}]}$ and $(SLB_t)_{t \in [T_{min}, T_{max}]}$ for the intensities $\lambda_t^1 = 10^{-5} + 10^{-2} X_t + 10^{-5} \int_0^t e^{-20(t-s)} dN_s^1$, $\lambda_t^2 = 4 \cdot 10^{-4} + 10^{-1} X_t + 10^{-5} \int_0^t e^{-20(t-s)} dN_s^2$.

Conclusions

The aim of this Thesis was to propose a new affine term structure model able to provide a short rate bounded from below by a randomly varying level. This stochastic lower bound is allowed to take negative values since negative rates can be rightfully regarded as the most relevant feature of financial markets in the post-crisis environment.

Under the physical probability measure P we have introduced our short rate model and given suitable admissibility conditions in order to get an affine model, as it is stated in Theorem 2.4. We have analysed that the affine nature allows for a great analytical tractability proving the validity of the affine transform formula for the functions ϕ and ψ solving the generalized Riccati ODEs (2.16) and (2.17). Since the affine nature of our framework may be lost under a change of measure from the physical P to a risk-neutral probability measure Q , we studied some sufficient conditions under which the affine structure is preserved.

To investigate the empirical behavior of our model we have simulated the short rate process. We studied how to approximate the continuous-time affine diffusion $(X_t)_{t \geq 0}$ by a recursive scheme, whereas to simulate the bi-dimensional counting process $(N_t)_{t \geq 0}$ we considered the Ogata's algorithm.

In the results of the simulations plotted in Figure 3.3 and Figure 3.4 we have remarked that our short rate model does not correctly fit the real market values of the EONIA and ECB deposit facility rates. Precisely, the spiky jumps that characterizes the historical behaviour of the EONIA rate (see Figure 1.1) are not replicated by our model since we have specified the macroeconomic factors $(X_t)_{t \geq 0}$ as an affine diffusion on \mathbb{R}_+^d . In order to overcome this limit we can allow the presence of jumps on $(X_t)_{t \geq 0}$.

In the one-dimensional case we can generalize our affine short rate model considering that the macroeconomic factors follow an α -CIR process for the parameter $1 < \alpha \leq 2$, that is

$$dX_t = (K_0 + K_1 X_t)dt + \sqrt{X_t}dW_t + \sigma \sqrt[\alpha]{X_t}dZ_t, \quad X_0 = x_0,$$

where $K_0 \in \mathbb{R}_+$, $K_1 \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$ and $(Z_t)_{t \geq 0}$ is a compensated spectrally positive α -stable Lévy process, independent of the Brownian motion $(W_t)_{t \geq 0}$. The α -CIR process is the natural extension of the affine diffusion process employed in our model; adding a jump part driven by the α -stable Lévy process we introduce a jump behavior on $(X_t)_{t \geq 0}$. In [42] Jiao et al. study this process and its properties and prove that it is an affine process. Therefore with this generalization the affine structure of our short rate model is preserved and therefore the results discussed in the Thesis can be adapted in order to

study a more general setting presenting a good balance between analytical tractability and flexibility.

Appendix A

Mathematical tools

In this appendix we introduce basic definitions and results that we use in the whole dissertation for the study of affine jump-diffusion processes. First of all, we set a common notation and terminology giving a brief description of point processes, analyzing in detail Hawkes processes. Then we introduce jump-diffusions and discuss the generalized version of the Itô formula. We concentrate on affine jump-diffusions that open the way to term structure models consistent with availability of closed-form bond-pricing formulas, as analyzed by Duffie et al. in [24] and later by Errais et al. in [27].

Let (Ω, \mathcal{F}, P) be the probability space defined in Chapter 1 endowed with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ satisfying the *usual conditions*. We suppose that the stochastic processes introduced in the following are defined on this probability space and adapted to the filtration \mathbb{F} .

A.1 Point processes

A point process aims to describe events that occur randomly over time. On the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ a point process can be viewed in two different ways: a sequence of non-negative random variables representing jumps times or via its associated counting process which increases in unit steps and is constant between these times. They are two equivalent representations and therefore we use them interchangeably in the text. We follow the point processes' theory developed in [15], unless otherwise specified.

Definition A.1. *A simple point process $(T_n)_{n \geq 0}$ is a sequence of random variables such that*

$$\begin{aligned} T_0 &= 0; \\ T_n &\leq T_{n+1} \quad \text{and} \quad T_n < T_{n+1} \quad \text{if} \quad T_n < \infty. \end{aligned}$$

This sequence models the times when jumps occur. We define the explosion time as the random variable $T_\infty := \lim_{n \rightarrow \infty} T_n$ and say that if $T_\infty = +\infty$, then $(T_n)_{n \geq 0}$ is non-explosive.

The counting process $(N_t)_{t \geq 0}$ associated to $(T_n)_{n \geq 0}$ is defined as

$$N_t = \sum_{n \geq 1} 1_{\{T_n \leq t\}} \text{ for } t \geq 0 \text{ with } N_0 = 0.$$

Therefore, $(N_t)_{t \geq 0}$ increases by one unit at times $(T_n)_{n \geq 0}$ and between these times is constant. We note that N_t counts the number of jumps that occurred up to time t .

Since we have assumed that all processes are \mathcal{F}_t -adapted, then it follows that the random variables $(T_n)_{n \geq 0}$ are \mathcal{F}_t -stopping times. Therefore, since the counting process $(N_t)_{t \geq 0}$ is \mathcal{F}_t -adapted then for all $t \geq 0$ it is true that $\mathcal{F}_t \ni \{N_t \leq n\} = \{T_{n+1} > t\}$ for every $n \geq 0$.

We can give an alternative definition of counting process.

Definition A.2. A counting process $(N_t)_{t \geq 0}$ is defined by

$$N_t := \begin{cases} n & \text{if } t \in [T_n, T_{n+1}) \\ +\infty & \text{if } t \geq T_\infty. \end{cases}$$

From this definition we deduce that equivalently the point process is non-explosive if $N_t < \infty$ P -a.s. for $t \geq 0$. We want to remark that the counting process is locally integrable and even locally bounded because $N_{T_n} \leq n$ for every $n \geq 0$.

The standard Poisson process can be defined as a counting process without explosion whose increments are independent and stationary, as it is showed in [40] (Section 8.2.1).

Definition A.3. Let $(T_n)_{n \geq 0}$ be a simple point process and $(Z_n)_{n \geq 0}$ a sequence of $\{1, 2, \dots, A\}$ -valued random variables representing A different types of jumps. Then the double sequence $(T_n, Z_n)_{n \geq 0}$ is called an A -variate point process. Define

$$N_t^a := \sum_{n \geq 1} 1_{\{T_n \leq t\}} 1_{\{Z_n = a\}} \text{ for } t \geq 0 \text{ and } 1 \leq a \leq A.$$

The A -vector process $(N_t)_{t \geq 0} = ((N_t^1, \dots, N_t^A)^\top)_{t \geq 0}$ is the A -variate counting process associated with $(T_n, Z_n)_{n \geq 0}$.

From the above definition it follows that $(N_t^a)_{t \geq 0}$ is a simple point process for every $a = 1, \dots, A$. Furthermore, we emphasize that the possibility of the simultaneous occurrence of two events (of either the same or different types) is ruled out.

A.1.1 Stochastic intensity

Given a counting process $(N_t)_{t \geq 0}$ its probability structure can be uniquely characterized via its intensity. The intensity $(\lambda_t)_{t \geq 0}$ is a non-negative \mathcal{F}_t -progressively measurable process defined heuristically by Bacry et al. in [10] as

$$\lambda_t := \lim_{\Delta \rightarrow 0} \frac{\mathbb{E}[N_{t+\Delta} - N_t | \mathcal{F}_t]}{\Delta} \text{ } P\text{-a.s.},$$

provided that the limit exists. For a rigorous definition we refer to [15] (Chapter II, Definition D7).

Definition A.4. Let $(N_t)_{t \geq 0}$ be a simple point process and $(\lambda_t)_{t \geq 0}$ a non-negative \mathcal{F}_t -progressively measurable process such that for all $t \geq 0$

$$\int_0^t \lambda_s ds < \infty \quad P\text{-a.s.} \quad (\text{A.1})$$

The point process admits the intensity $(\lambda_t)_{t \geq 0}$ if for all non-negative predictable processes $(C_t)_{t \geq 0}$ it holds that

$$\mathbb{E} \left[\int_0^\infty C_s dN_s \right] = \mathbb{E} \left[\int_0^\infty C_s \lambda_s ds \right]. \quad (\text{A.2})$$

Following the result stated in [15] (Chapter II, Theorem T8) the subsequent integration theorem holds.

Theorem A.5. The counting process $(N_t)_{t \geq 0}$ with intensity $(\lambda_t)_{t \geq 0}$ is non-explosive and the process $(N_t - \int_0^t \lambda_s ds)_{t \geq 0}$ is a local martingale.

Proof. To prove the non-explosion property we define for all $n \geq 0$

$$\tau_n = \begin{cases} \inf \{t : \int_0^t \lambda_s ds \geq n\} & \text{if } \{t : \int_0^t \lambda_s ds \geq n\} \neq \emptyset \\ +\infty & \text{otherwise.} \end{cases}$$

$(\tau_n)_{n \geq 0}$ is an increasing sequence of stopping times such that $\lim_{n \rightarrow +\infty} \tau_n = +\infty$ because of (A.1). If we evaluate (A.2) with the predictable process $(1_{\{t \leq \tau_n\}})_{t \geq 0}$ for a fixed $n \geq 0$, then we conclude that $\mathbb{E}[N_{\tau_n}] = \mathbb{E}[\int_0^{\tau_n} dN_s] = \mathbb{E}[\int_0^{\tau_n} \lambda_s ds] \leq n < \infty$. Therefore, $N_{\tau_n} < \infty$ P -a.s. for $n \geq 0$ and this property implies the non-explosion property of the counting process since $\lim_{n \rightarrow +\infty} \tau_n = +\infty$.

To prove the local martingale property, for $0 \leq u \leq v$ and $B \in \mathcal{F}_u$ we consider the predictable process $(1_B 1_{\{t \leq T_n\}} 1_{\{u < t \leq v\}})_{t \geq 0}$ where $(T_n)_{n \geq 0}$ is the point process associated to the counting process. Consequently (A.2) becomes

$$\mathbb{E}[1_B (N_{v \wedge T_n} - N_{u \wedge T_n})] = \mathbb{E} \left[1_B \int_{u \wedge T_n}^{v \wedge T_n} dN_s \right] = \mathbb{E} \left[1_B \int_{u \wedge T_n}^{v \wedge T_n} \lambda_s ds \right]. \quad (\text{A.3})$$

In particular, for $u = 0$, $s = t$ and $B = \Omega$ it follows that $\mathbb{E}[N_{t \wedge T_n}] = \mathbb{E} \left[\int_0^{t \wedge T_n} \lambda_s ds \right]$. Hence, by definition of T_n it holds that $\mathbb{E} \left[\int_0^{t \wedge T_n} \lambda_s ds \right] < \infty$ since $N_{t \wedge T_n} \leq n$. Rearranging terms in (A.3) we obtain:

$$\mathbb{E} \left[1_B \left(N_{v \wedge T_n} - \int_0^{v \wedge T_n} \lambda_s ds \right) \right] = \mathbb{E} \left[1_B \left(N_{u \wedge T_n} - \int_0^{u \wedge T_n} \lambda_s ds \right) \right].$$

From the arbitrariness of times $0 \leq u \leq v$ and B with $B \in \mathcal{F}_u$ we deduce that the process $(N_{t \wedge T_n} - \int_0^{t \wedge T_n} \lambda_s ds)_{t \geq 0}$ is a martingale for every fixed $n \geq 0$. Then, $(T_n)_{n \geq 0}$ is a sequence of stopping times with $\lim_{n \rightarrow +\infty} T_n = +\infty$ and therefore $(N_t - \int_0^t \lambda_s ds)_{t \geq 0}$ is a local martingale. \square

Consequently, following the argument of the proof for all $n \geq 0$ and $0 \leq u \leq t$

$$\mathbb{E}[N_{t \wedge T_n} - N_{u \wedge T_n} | \mathcal{F}_u] = \mathbb{E}\left[\int_{u \wedge T_n}^{t \wedge T_n} \lambda_s ds \middle| \mathcal{F}_u\right] \quad P\text{-a.s.},$$

letting $n \rightarrow +\infty$ we argue that

$$\mathbb{E}[N_t - N_u | \mathcal{F}_u] = \mathbb{E}\left[\int_u^t \lambda_s ds \middle| \mathcal{F}_u\right] \quad P\text{-a.s.} \quad (\text{A.4})$$

This is a consequence of the fact that $\lim_{n \rightarrow +\infty} T_n = +\infty$ because $(N_t)_{t \geq 0}$ is a non-explosive point process.

A.1.2 Hawkes processes

A particular class of multivariate point processes consists of Hawkes processes. They are characterized by a stochastic intensity vector that is a simple linear function of past jumps and thus they are path-dependent point processes. Each jump excites the process increasing the likelihood of subsequent jumps (self-exciting property). In this subsection we give the definition of a Hawkes process and introduce the main properties adopting the approach of Bacry et al. [10].

Definition A.6. We consider a A -variate counting process $(N_t)_{t \geq 0} = ((N_t^1, \dots, N_t^A)^\top)_{t \geq 0}$ whose associated intensity vector is $(\lambda_t)_{t \geq 0} = ((\lambda_t^1, \dots, \lambda_t^A)^\top)_{t \geq 0}$. A Hawkes process is a counting process $(N_t)_{t \geq 0}$ such that the intensity vector is given by

$$\lambda_t^a := u^a + \sum_{b=1}^A \int_0^t g_{ab}(t-s) dN_s^b \quad \text{for } a = 1, \dots, A, \quad (\text{A.5})$$

where $u^a \in \mathbb{R}_+$ and $g_{ab} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a measurable L^1 -integrable function, for all $a, b = 1, \dots, A$.

In literature the parameter u^a is called the baseline intensity and g_{ab} the kernel function. In equation (A.5) the Lebesgue-Stieltjes integral with respect to the counting process $(N_t)_{t \geq 0}$ is defined pathwise as

$$\sum_{b=1}^A \sum_{n \geq 1} g_{ab}(t - T_n) \mathbf{1}_{\{T_n \leq t\}} \mathbf{1}_{\{Z_n = b\}}. \quad (\text{A.6})$$

Looking at the previous expression, or equivalently (A.5), we notice that between jumps the intensity process $(\lambda_t^a)_{t \geq 0}$ is constant if g_{ab} are constant functions for $a, b = 1, \dots, A$.

By Definition A.6 we highlight the self-exciting property that identifies Hawkes processes. In fact, the Hawkes process is characterized by the intensity (A.5) which in turn changes in response to jumps of the counting processes $(N_t^b)_{t \geq 0}$ for $b = 1, \dots, A$.

For all $a, b = 1, \dots, A$ we require that the function g_{ab} is L^1 -integrable, that is

$$\|g_{ab}\|_{L^1} := \int_0^{+\infty} g_{ab}(s) ds < \infty. \quad (\text{A.7})$$

From this property we derive that the Hawkes process $(N_t^a)_{t \geq 0}$ is non-explosive. To prove this statement we fix $t \geq 0$ and define $\bar{T}_n := t \wedge T_n$ for every $n \geq 0$, where we recall that $(T_n)_{n \geq 0}$ is the point process associated to $(N_t)_{t \geq 0}$. It follows that $N_{\bar{T}_n}^a \leq n$ and, therefore, from (A.5) we derive that

$$\begin{aligned} \mathbb{E}[N_{\bar{T}_n}^a] &= u^a \mathbb{E}[\bar{T}_n] + \sum_{b=1}^A \mathbb{E} \left[\int_0^{\bar{T}_n} \int_0^t g_{ab}(t-s) dN_s^b dt \right] \\ &= u^a \mathbb{E}[\bar{T}_n] + \sum_{b=1}^A \mathbb{E} \left[\int_0^{\bar{T}_n} \int_s^{\bar{T}_n} g_{ab}(t-s) dt dN_s^b \right] \\ &= u^a \mathbb{E}[\bar{T}_n] + \sum_{b=1}^A \mathbb{E} \left[\int_0^{\bar{T}_n} \int_0^{\bar{T}_n-s} g_{ab}(t) dt dN_s^b \right] \end{aligned} \quad (\text{A.8})$$

where in the second line we have used Fubini's theorem (see [25] (Theorem 3.29)). If we define the function $G_{ab}(u) := \int_0^u g_{ab}(t) dt$, then integrating by parts we deduce that

$$\begin{aligned} 0 &= G_{ab}(\bar{T}_n - \bar{T}_n) N_{\bar{T}_n}^b - G_{ab}(\bar{T}_n - 0) N_0^b \\ &= \int_0^{\bar{T}_n} \left(\frac{dG_{ab}}{ds}(\bar{T}_n - s) \right) N_s^b ds + \int_0^{\bar{T}_n} G_{ab}(\bar{T}_n - s) dN_s^b \\ &= - \int_0^{\bar{T}_n} g_{ab}(\bar{T}_n - s) N_s^b ds + \int_0^{\bar{T}_n} G_{ab}(\bar{T}_n - s) dN_s^b. \end{aligned}$$

It follows that $\int_0^{\bar{T}_n} \int_0^{\bar{T}_n-s} g_{ab}(t) dt dN_s^b = \int_0^{\bar{T}_n} g_{ab}(\bar{T}_n - s) N_s^b ds$ and inserting this last equality in (A.8) it holds that

$$\mathbb{E}[N_{\bar{T}_n}^a] = u^a \mathbb{E}[\bar{T}_n] + \sum_{b=1}^A \mathbb{E} \left[\int_0^{\bar{T}_n} g_{ab}(\bar{T}_n - s) N_s^b ds \right].$$

Since $\lim_{n \rightarrow +\infty} N_{\bar{T}_n} = N_t$ we can conclude that

$$\mathbb{E}[N_t^a] = u^a t + \sum_{b=1}^A \mathbb{E} \left[\int_0^t g_{ab}(t-s) N_s^b ds \right].$$

We remind that g_{ab} is an L^1 -integrable function for every $a, b = 1, \dots, A$ and thus $\mathbb{E}[N_t^a] < \infty$ for any $t \geq 0$. It follows that the intensity process (A.5) is well-defined because (A.1) is satisfied. In fact, by means of analogous computations to (A.8) and

definition of L^1 -norm we deduce that

$$\begin{aligned} \mathbb{E} \left[\int_0^t \lambda_s^a ds \right] &= u^a t + \sum_{b=1}^A \mathbb{E} \left[\int_0^t \int_0^{t-u} g_{ab}(s) ds dN_u^b \right] \\ &\leq u^a t + \sum_{b=1}^A \|g_{ab}\|_{L^1} \mathbb{E}[N_t^b] < \infty. \end{aligned}$$

A Hawkes process is stationary if its distribution does not change under time shift. Adopting the notation of [30] the point process $(N_t)_{t \geq 0}$ is stationary if the process $((\delta_s N)_t)_{t \geq 0}$ has the same distribution as $(N_t)_{t \geq 0}$ for any $s > 0$, where δ_s is the shift operator defined by $(\delta_s N)_t := N_{t+s}$ for $t \geq 0$, i.e. it is a shift by s units of time. This implies that a stationary Hawkes process $(N_t)_{t \geq 0}$ has stationary increments. Consequently, for every $t \geq 0$ and $a = 1, \dots, A$ the distribution of λ_t^a does not depend on the time variable. Then, from

$$\begin{aligned} \mathbb{E}[\lambda_t^a] &= u^a t + \mathbb{E} \left[\sum_{b=1}^A \int_0^t g_{ab}(t-s) dN_s^b \right] \\ &= u^a t + \mathbb{E} \left[\sum_{b=1}^A \int_0^t g_{ab}(t-s) \lambda_s^b ds \right] \end{aligned}$$

it follows that

$$\lambda^a = u^a + \sum_{b=1}^A \lambda^b \int_0^{+\infty} g_{ab}(t) dt.$$

For the univariate case, i.e. $A = 1$, if we assume the stationary property we deduce that $\lambda = u + \lambda \|g\|_{L^1}$. Rearranging terms we agree that

$$\lambda = \frac{u}{1 - \|g\|_{L^1}}.$$

The intensity must be positive and therefore $\|g\|_{L^1} < 1$. We state the following result provided by [36] (Theorem 1).

Lemma A.7. *If $u > 0$ and the kernel function satisfies the property $\|g\|_{L^1} < 1$, then there exists a stationary counting process $(N_t)_{t \geq 0}$ characterized by a finite intensity process which satisfies (A.5) for $A = 1$.*

When the kernel functions g_{ab} are set to be exponentially decaying, then (A.5) becomes

$$\lambda_t^a = u^a + \sum_{b=1}^a \delta_{ab} \int_0^t e^{-\gamma_{ab}(t-s)} dN_s^b \quad \text{for } a = 1, \dots, A, \quad (\text{A.9})$$

for some $\delta_{ab}, \gamma_{ab} \in \mathbb{R}_+$. Because of parameters' non-negativity g_{ab} is L^1 -integrable and thus the associated Hawkes process is well-defined. For $a = 1, \dots, A$ the jump intensity

λ_t^a is raised by each jump before time t , that is ΔN_s^b with $s < t$ and $b = 1, \dots, A$. The parameters δ_{ab} determine the scale of the positive influence of past jumps, whereas γ_{ab} determine its time decay. Hence, the response of previous jumps on the intensity function in t decays exponentially and is driven by the parameters γ_{ab} .

In this particular case the couple consisting of the jump process and its intensity $(N_t, \lambda_t)_{t \geq 0}$ is a Markov process, and also the intensity process itself is a Markov process, as is it stated in [10] (Proposition 2).

A.2 Jump-diffusions

We define the main family of stochastic processes used in this dissertation: jump-diffusion processes. Loosely speaking, they are processes whose dynamics are driven by a Brownian motion and a jump process. We recall the following fundamental definitions adopting the notation of [27, 40].

Definition A.8. A pure jump process $(J_t)_{t \geq 0}$ on \mathbb{R}^d is defined by

$$J_t := \sum_{n \geq 1} Z_n 1_{\{T_n \leq t\}} = \sum_{n=1}^{N_t} Z_n \quad \text{for } t \geq 0 \quad \text{with } J_0 = 0, \quad (\text{A.10})$$

where $(T_n)_{n \geq 0}$ is a non-explosive simple point process with associated counting process $(N_t)_{t \geq 0}$ and $(Z_n)_{n \geq 0}$ a sequence of independent identically distributed \mathbb{R}^d -valued random variables. By convention we set $Z_0 = 0$. The sequence $(Z_n)_{n \geq 0}$ is supposed to be independent from the point process.

We recognize that $(J_t)_{t \geq 0}$ is a discontinuous stochastic process which has a jump of size Z_n at time T_n and is constant between two consecutive jumps. Therefore, the point process $(T_n)_{n \geq 0}$ represents jump times and the variables $(Z_n)_{n \geq 0}$ jump sizes.

Denoting the left limit by $J_{t-} := \lim_{s \rightarrow t, s < t} J_s$ for $t > 0$ with the convention that $J_{0-} = J_0 = 0$, we describe the jumps of the process as $\Delta J_t := J_t - J_{t-}$. Consequently, ΔJ_t is non-null only if $t = T_n$ for some $n \geq 1$ and in this case $\Delta J_t = \Delta J_{T_n} = J_{T_n} - J_{T_n-}$. We emphasize that from the non-explosion property of the point process it follows that for every fixed time $t \geq 0$ the number of jumps before t is finite P -a.s..

Equivalently, we adopt the definition of [27] (Section 4) and define $(J_t)_{t \geq 0}$ as a temporally consistent \mathbb{R}^d -valued point process, i.e. all the components of the process share jump times $(T_n)_{n \geq 0}$ and differ only for jump sizes, that is

$$J_t = (J_t^1, \dots, J_t^d)^\top \quad \text{for } t \geq 0,$$

where $J_t^i := \sum_{n=1}^{N_t} Z_n^i$ for each $i = 1, \dots, d$ and $(Z_n^i)_{n \geq 0}$ is a sequence of random variables taking values in \mathbb{R} .

Otherwise specified from now on we refer to Definition [A.8] and uniquely characterize the jump process $(J_t)_{t \geq 0}$ through the intensity of the point process $(\lambda_t)_{t \geq 0}$ and the jump distribution ν on \mathbb{R}^d that specifies the probability distribution of $(Z_n)_{n \geq 0}$.

Definition A.9. An \mathbb{R}^d -valued jump diffusion process $(X_t)_{t \geq 0}$ is a strong solution to the SDE

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t + dJ_t, \quad (\text{A.11})$$

where $(W_t)_{t \geq 0}$ is a d -dimensional Brownian motion and two measurable functions are assigned, $\mu : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$.

It follows that the driving terms in (A.11) are the Brownian motion $(W_t)_{t \geq 0}$ and the jump process $(J_t)_{t \geq 0}$. Adopting the notation of [40] (Section 9.1.2) we observe that $(X_t)_{t \geq 0}$ is made up of a continuous process $(X_t^c)_{t \geq 0}$ and a purely discontinuous one $(J_t)_{t \geq 0}$. Indeed, we can write

$$\begin{aligned} X_t - X_0 &= \int_0^t \mu(X_s, s)ds + \int_0^t \sigma(X_s, s)dW_s + \int_0^t dJ_s \\ &= X_t^c + \sum_{0 < s \leq t} \Delta J_s \\ &= X_t^c + J_t, \end{aligned}$$

From the previous equation it can be easily deduced that $(J_t)_{t \geq 0}$ is the jump part of $(X_t)_{t \geq 0}$, and therefore the jump times of the process $(X_t)_{t \geq 0}$ are those of $(J_t)_{t \geq 0}$ and for all $t \geq 0$ the jump size is $\Delta X_t = X_t - X_{t-} = \Delta X_t^c + \Delta J_t = \Delta J_t$ P -a.s..

For $t \geq 0$ we define $Y_t := f(X_t, t)$ for a $C_{x,t}^{2,1}$ function¹ $f : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}$, thus the Itô's formula becomes (see [38] (Chapter I, Theorem 4.57))

$$\begin{aligned} f(X_t, t) - f(X_0, 0) &= \int_0^t \frac{\partial f}{\partial s}(X_s, s)ds + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s, s)dX_s^{c^i} \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s, s)d[X^{c^i}, X^{c^j}]_s \\ &\quad + \sum_{0 < s \leq t, \Delta J_s \neq 0} f(X_{s-} + \Delta J_s, s) - f(X_{s-}, s-), \end{aligned}$$

where the last addend is equal to

$$\sum_{0 < T_n \leq t} f(X_{T_n-} + \Delta J_{T_n}, T_n) - f(X_{T_n-}, T_n-).$$

The sum is finite P -a.s. for any $t \geq 0$. The enunciated Itô's formula is also valid when the function f is complex-valued and in this case we consider the real and purely imaginary parts separately.

¹ $C_{x,t}^{2,1}$ means C^2 with respect to x and C^1 with respect to t .

A.3 Affine jump-diffusions

This section introduces affine jump-diffusion processes. We provide the definition and further technical restrictions that the coefficients which characterize the process have to satisfy. This section is based on [24, 27].

Definition A.10. *The process $(Z_t)_{t \geq 0}$ taking value in some state space $D \subseteq \mathbb{R}^d$ is an affine jump-diffusion if it is a Markov process and a strong solution to the SDE*

$$dZ_t = \mu(Z_t)dt + \sigma(Z_t)dW_t + \sum_{a=1}^A \zeta_a dJ_t^a, \quad Z_0 = z, \quad (\text{A.12})$$

where the drift $\mu : D \rightarrow \mathbb{R}^d$ is supposed to be continuous and the diffusion function $\sigma : D \rightarrow \mathbb{R}^{d \times d}$ measurable such that $\sigma(z)\sigma(z)^\top =: H(z)$ is continuous in z . We impose the affine structure given by

$$\mu(z) = K_0 + K_1 z \text{ for } K_0 \in \mathbb{R}^d, K_1 \in \mathbb{R}^{d \times d} \quad K = (K_0, K_1); \quad (\text{A.13})$$

$$(H(z))_{ij} = (H_0)_{ij} + \langle (H_1)_{ij}, z \rangle \text{ for } H_0 \in \mathbb{R}^{d \times d}, H_1 \in \mathbb{R}^{d \times d \times d} \quad H = (H_0, H_1). \quad (\text{A.14})$$

Each ζ_a is a d -dimensional diagonal matrix and $(J_t^a)_{t \geq 0}$ a pure jump process on \mathbb{R}_+^d , for $a = 1, \dots, A$. Therefore, the components of each process N^a share event times and differ only in jump sizes. We determine $(J_t^a)_{t \geq 0}$ by means of a fixed probability distribution ν^a on \mathbb{R}_+^d independent of the state process and the intensity process $(\lambda^a(Z_t))_{t \geq 0}$ for some affine function $\lambda^a : D \rightarrow [0, \infty)$:

$$\lambda^a(z) = l_0^a + \langle l_1^a, z \rangle \text{ for } l_0^a \in \mathbb{R}, l_1^a \in \mathbb{R}^d \quad l^a = (l_0^a, l_1^a). \quad (\text{A.15})$$

Following the convention of Duffie et al. [24] H_1 is a tensor of dimension $d \times d \times d$. We remark that if we fix the first two indices i and j , then the tensor turns into a \mathbb{R}^d -valued vector determined by $(H_1)_{ij} = ((H_1^1)_{ij}, \dots, (H_1^d)_{ij})^\top$. Analogously, fixing the third index k the tensor is reduced to a $(d \times d)$ -matrix H_1^k with elements $(H_1^k)_{ij}$.

Loosely speaking, an affine jump-diffusion process is a jump-diffusion for which the drift, the covariance matrix and the intensity of each jump process are determined by some affine functions of the process $(Z_t)_{t \geq 0}$.

The infinitesimal generator \mathbf{A} of $(Z_t)_{t \geq 0}$ is defined at a C^2 -function with compact support $f : D \rightarrow \mathbb{R}$ by

$$\begin{aligned} \mathbf{A}f(z) = & \langle \nabla f(z), \mu(z) \rangle + \frac{1}{2} \text{tr} \left[\mathbf{H}f(z) \sigma(z) \sigma(z)^\top \right] \\ & + \sum_{a=1}^A \lambda^a(z) \int_{\mathbb{R}_+^d} [f(z + \zeta_a u) - f(z)] d\nu^a(u), \end{aligned} \quad (\text{A.16})$$

where $\mathbf{H}f(z)$ denotes the Hessian matrix. The affine dependence of this generator on the value z emphasizes the fact that the process $(Z_t)_{t \geq 0}$ which satisfies (A.12) is an affine

jump-diffusion. In other words, we can identify the affine dependencies with respect to $(Z_t)_{t \geq 0}$ from the generator's structure. Indeed, rewriting (A.16) it follows that

$$\begin{aligned} \mathbf{A}f(z) &= \langle \nabla f(z), (K_0 + K_1 z) \rangle \\ &+ \frac{1}{2} \text{tr} \left[\mathbf{H}f(z) (H_0 + \langle H_1, z \rangle) \right] \\ &+ \sum_{a=1}^A (l_0^a + \langle l_1^a, z \rangle) \int_{\mathbb{R}_+^d} [f(z + \zeta_a u) - f(z)] d\nu^a(u), \end{aligned}$$

where with $\langle H_1, z \rangle$ we denote the $(d \times d)$ matrix with elements $(\langle H_1, z \rangle)_{ij} = \langle (H_1)_{ij}, z \rangle$.

For $c \in \mathbb{C}^d$ we define the jump transforms

$$\theta^a(c) := \int_{\mathbb{R}_+^d} e^{\langle c, u \rangle} d\nu^a(u) \quad (\text{A.17})$$

and suppose that these integrals are well-defined for all $a = 1, \dots, A$.

Until now we have introduced the fundamental components which describe the affine jump-diffusion process $(Z_t)_{t \geq 0}$ uniquely. We denote $l := (l^1, \dots, l^A)$, $\nu := (\nu^1, \dots, \nu^A)$ and $\theta := (\theta^1, \dots, \theta^A)$. The parameters (K, H, l, θ) determine the distribution of the process totally. We adopt the notation used in [24] (Section 2.2) and define χ .

Definition A.11. *The process $(Z_t)_{t \geq 0}$ is said to have the characteristics $\chi = (K, H, l, \theta)$.*

We follow the approach described in [28] (Section 10.2) and from now on we assume that the state space is of the canonical form $D = \mathbb{R}_+^m \times \mathbb{R}^n$ for integers $m, n \geq 0$ with $m+n = d$. We define the index sets $I := \{1, \dots, m\}$ and $J := \{m+1, \dots, d\}$. Intuitively, the set I collects the indices of the first m elements of the D -valued process $(Z_t)_{t \geq 0}$, that is the components taking values in \mathbb{R}_+ , while the set J collects the remaining ones taking values in \mathbb{R} . We have to ensure that $(Z_t)_{t \geq 0}$ does not leave the set D .

Assumption A.12. *The parameters $(K_0, K_1, H_0, H_1, l_0, l_1, \nu)$ are called admissible if*

- H_0, H_1^k for $k \in I \cup J$ are symmetric and positive semi-definite matrices
- $(H_0)_{ij} = 0$ for $i, j \in I$ (and so for $i \in I, j \in J$ and $i \in J, j \in I$)
- $H_1^k = \mathbf{0}$ for $k \in J$
- $(H_1^k)_{ij} = (H_1^k)_{ji} = 0$ for $k \in I, i \in I \setminus \{k\}, j \in I \cup J$
- $K_0 \in \mathbb{R}_+^m \times \mathbb{R}^n$
- $(K_1)_{ij} = 0$ for $i \in I, j \in J$
- $(K_1)_{ij} \in \mathbb{R}_+$ for $i \in I, j \in I \setminus \{i\}$
- $l_0^a, (l_1^a)_i \in \mathbb{R}_+$ for $i \in I$, for every $a = 1, \dots, A$

- $(l_1^a)_j = 0$ for $j \in J$, for every $a = 1, \dots, A$
- $(\zeta^a)_{i,i} \in \mathbb{R}_+$ for $i \in I$, for every $a = 1, \dots, A$

Before giving an intuitive interpretation of these conditions we have to state a result proved by Filipović in [28] (Lemma 10.11). It considers only the case of affine diffusions, that is $A = 0$.

Lemma A.13. *Suppose that $\mu : D \rightarrow \mathbb{R}^d$ and $\sigma : D \rightarrow \mathbb{R}^{d \times d}$ in (A.12) admit a continuous and measurable extension to \mathbb{R}^d , respectively, and such that $H : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ is continuous. Let $u \in \mathbb{R}^d \setminus \{0\}$ and define $\mathcal{S} := \{z \in \mathbb{R}^d \mid \langle u, z \rangle \geq 0\}$.*

Fix $z \in \partial\mathcal{S}$ and let $(Z_t)_{t \geq 0}$ be a solution of (A.12) with initial condition $Z_0 = z$. If $Z_t \in \mathcal{S}$ for all $t \geq 0$, then necessarily

$$\langle u, H(z)u \rangle = 0 \quad \text{and} \quad \langle u, \mu(z) \rangle \geq 0. \quad (\text{A.18})$$

Conversely, if (A.18) holds for all $z \in \mathbb{R}^d \setminus \mathring{\mathcal{S}}$, where $\mathring{\mathcal{S}}$ denotes the interior, then for any solution $(Z_t)_{t \geq 0}$ with $Z_0 \in \mathcal{S}$ it holds that $Z_t \in \mathcal{S}$ for all $t \geq 0$.

Now we can illustrate the admissibility conditions formulated in Assumption A.12. We suppose that $(Z_t)_{t \geq 0}$ is an affine jump-diffusion with characteristics $\chi = (K, H, l, \theta)$ and taking values in $D = \mathbb{R}_+^m \times \mathbb{R}^n$. We rewrite equations (A.13) and (A.14) obtaining

$$\begin{aligned} \mu(z) &= \begin{pmatrix} (K_0)_1 \\ \dots \\ (K_0)_d \end{pmatrix} + \sum_{h \in I} \begin{pmatrix} (K_1)_{1h} \\ \dots \\ (K_1)_{dh} \end{pmatrix} z_h + \sum_{h \in J} \begin{pmatrix} (K_1)_{1h} \\ \dots \\ (K_1)_{dh} \end{pmatrix} z_h, \\ H(z) &= H_0 + \sum_{k \in I} H_1^k z_k + \sum_{k \in J} H_1^k z_k. \end{aligned}$$

Obviously, $H(z)$ is a symmetric and positive semi-definite matrix for all $z \in D$ if and only if H_0, H_1^k are symmetric and positive semi-definite for $k \in I$ and $H_1^k = \mathbf{0}$ for all $k \in J$, where $\mathbf{0}$ defines the null $(d \times d)$ matrix.

Now, let z be a point at the boundary of D and so $z_l = 0$ for some $l \in I$. If we denote the l -th vector of the canonical basis of \mathbb{R}^d with e_l , then Lemma A.13 implies that

$$\begin{aligned} \left\langle e_l, \left(H_0 + \sum_{k \in I \setminus \{l\}} H_1^k z_k \right) e_l \right\rangle &= 0, \\ \left\langle e_l, \left\{ \begin{pmatrix} (K_0)_1 \\ \dots \\ (K_0)_d \end{pmatrix} + \sum_{k \in I \setminus \{l\}} \begin{pmatrix} (K_1)_{1k} \\ \dots \\ (K_1)_{dk} \end{pmatrix} z_k + \sum_{h \in J} \begin{pmatrix} (K_1)_{1h} \\ \dots \\ (K_1)_{dh} \end{pmatrix} z_h \right\} \right\rangle &\geq 0. \end{aligned}$$

The previously defined conditions ensure that $(Z_t)_{t \geq 0}$ does not cross the boundary of D and they have to hold for all $z_k \geq 0$ with $k \in I \setminus \{l\}$ and $z_h \in \mathbb{R}$ with $h \in J$. Then, for all $l \in I$ it follows that $H_0 e_l = \underline{0}$ and $H_1^k e_l = \underline{0}$ for all $k \in I \setminus \{l\}$, where $\underline{0}$ denotes the null

vector in \mathbb{R}^d . Furthermore, $K_0 \in \mathbb{R}_+^m \times \mathbb{R}^n$ and for all $l \in I$, $(K_1)_{lk} \geq 0$ for $k \in I \setminus \{l\}$ and $(K_1)_{ll} = 0$ for $k \in J$.

Lastly, we analyse the jump terms. For $a = 1, \dots, A$ the intensity of the a -th jump process is $\lambda^a(z) = l_0 + \langle l_1, z \rangle$. It has to be positive for all $z \in D$ and, consequently, $l_0^a \in \mathbb{R}_+$, $l_1^a \in \mathbb{R}_+^d$ and $(l_1^a)_j = 0$ for all $j \in J$. If the process is at z , a point at the boundary of D and so $z_l = 0$ for some $l \in I$, we have to force that it does not leave the set $\mathbb{R}_+^m \times \mathbb{R}^n$ with a jump. It is sufficient to require that the diagonal element $(\zeta^a)_{ll} \in \mathbb{R}_+$ for all $l \in I$. Therefore, up to a renaming of the indexes, the discussed conditions are equivalent to Assumption [A.12](#).

A.3.1 Affine diffusions

If the previously defined affine process $(Z_t)_{t \geq 0}$ is characterized by continuous paths, then it does not satisfy the SDE [\(A.12\)](#) but the following one

$$dZ_t = \mu(Z_t)dt + \sigma(Z_t)dW_t, \quad Z_0 = z, \quad (\text{A.19})$$

with the usual hypotheses, i.e. $\mu : D \rightarrow \mathbb{R}^d$ continuous and $\sigma : D \rightarrow \mathbb{R}^{d \times d}$ measurable such that $H(z) = \sigma(z)\sigma(z)^\top$ is continuous.

Making explicit the coefficients' affine dependence with respect to the process we rewrite equation [\(A.19\)](#) as

$$dZ_t = (K_0 + K_1 Z_t)dt + \sigma(Z_t)dW_t, \quad Z_0 = z, \quad (\text{A.20})$$

such that $(\sigma(z)\sigma(z)^\top)_{ij} = (H_0)_{ij} + \langle (H_1)_{ij}, z \rangle$ for $i, j = 1, \dots, d$. Our aim is to convert the above equation by a linear transformation into a canonical representation in which the matrix H is of special form. Specifically, developing the contents of [\[28\]](#) (Section 10.4) we use Itô's formula and obtain that for every invertible $(d \times d)$ matrix Λ , the linear transform $y = \Lambda z$ satisfies

$$dY_t = (\Lambda K_0 + \Lambda K_1 \Lambda^{-1} Y_t)dt + \Lambda \sigma(\Lambda^{-1} Y_t) dW_t, \quad Y_0 = \Lambda z. \quad (\text{A.21})$$

Therefore the drift and the diffusion matrix of $(Y_t)_{t \geq 0}$ are affine in y :

$$\begin{cases} \mu_y &= \Lambda K_0 + \Lambda K_1 \Lambda^{-1} y \\ H_y &= \Lambda \sigma(\Lambda^{-1} Y_t) (\Lambda \sigma(\Lambda^{-1} Y_t))^\top = \Lambda H(\Lambda^{-1} y) \Lambda^\top. \end{cases} \quad (\text{A.22})$$

Consequently, the existence and uniqueness of a strong solution to SDE described in [\(A.20\)](#) is invariant with respect to non-singular linear transformations. We will show that there exists a classification method ensuring a unique canonical representation. The diffusion matrix $H(z)$ can always be brought into block-diagonal form by a regular linear transform Λ with $\Lambda(\mathbb{R}_+^m \times \mathbb{R}^n) = \mathbb{R}_+^m \times \mathbb{R}^n$. This feature is showed in the next result based on [\[28\]](#) (Lemma 10.5). We recall that $I = \{1, \dots, m\}$ and $J = \{m+1, \dots, d\}$.

Theorem A.14. *There exists an invertible $(d \times d)$ matrix Λ with $\Lambda(\mathbb{R}_+^m \times \mathbb{R}^n) = \mathbb{R}_+^m \times \mathbb{R}^n$ such that $\Lambda H(\Lambda^{-1} y) \Lambda^\top$ is block-diagonal of the form*

$$\Lambda H(\Lambda^{-1} y) \Lambda^\top = \begin{pmatrix} \text{diag}(y_1, \dots, y_q, 0, \dots, 0) & 0 \\ 0 & p + \sum_{i \in I} y_i \pi_i \end{pmatrix}$$

for $0 \leq q \leq m$ and some symmetric positive semi-definite $(n \times n)$ matrices p, π_1, \dots, π_m . We notice that $\text{diag}(y_1, \dots, y_q, 0, \dots, 0)$ denotes the $(m \times m)$ matrix with diagonal elements $y_1, \dots, y_q, 0, \dots, 0$.

Proof. From the affine dependence on the state variable we know that $\Lambda H(\Lambda^{-1}y)\Lambda^\top$ is block-diagonal for every $z = \Lambda^{-1}y$ if and only if $\Lambda H_0 \Lambda^\top$ and $\Lambda H_1^k \Lambda^\top$ are block-diagonal for all $k \in I$. By permutation and scaling of the first m coordinate axes, we may assume that there exists an integer $0 \leq q \leq m$ such that $(H_1^1)_{11} = \dots = (H_1^q)_{qq} = 1$ and $(H_1^k)_{kk} = 0$ for $q < k \leq m$. Hence H_0 and H_1^k for $q < k \leq m$ are already block-diagonal of the form

$$H_0 = \begin{pmatrix} 0 & 0 \\ 0 & \{H_0\}_{i,j \in J} \end{pmatrix} \quad \text{and} \quad H_1^k = \begin{pmatrix} 0 & 0 \\ 0 & \{H_1^k\}_{i,j \in J} \end{pmatrix},$$

where $\{H_0\}_{i,j \in J}$ denotes the $(n \times n)$ matrix composed by the elements $(H_0)_{ij}$ for $i, j \in J$ of the $(d \times d)$ matrix H_0 , and similarly for H_1^k .

We notice that for $1 \leq k \leq q$ we may have non-zero off-diagonal elements in the k -th row $\{(H_1^k)_{kj}\}_{j \in J}$. We thus define the $(n \times m)$ matrix $\mathcal{D} := (\delta_1, \dots, \delta_m)$ with k -th column $\delta_k := -\{(H_1^k)_{kj}\}_{j \in J}$ and set

$$\Lambda := \begin{pmatrix} 1_m & 0 \\ \mathcal{D} & 1_n \end{pmatrix},$$

where 1_m and 1_n are identity matrices of dimension m and n , respectively. \mathcal{D} is invertible and maps $\mathbb{R}_+^m \times \mathbb{R}^n$ onto $\mathbb{R}_+^m \times \mathbb{R}^n$. Moreover, if we use the same notation that we defined above, $\mathcal{D}\{H_1^k\}_{i,j \in I} = -\{H_1^k\}_{i \in J, j \in I}$ for $k \in I$. From here it follows that

$$\begin{aligned} \Lambda H_1^k &= \begin{pmatrix} 1_m & 0 \\ \mathcal{D} & 1_n \end{pmatrix} \begin{pmatrix} \{H_1^k\}_{i,j \in I} & \{H_1^k\}_{i \in I, j \in J} \\ \{H_1^k\}_{i \in J, j \in I} & \{H_1^k\}_{i,j \in J} \end{pmatrix} \\ &= \begin{pmatrix} \{H_1^k\}_{i,j \in I} & \{H_1^k\}_{i \in I, j \in J} \\ 0 & \mathcal{D}\{H_1^k\}_{i \in I, j \in J} + \{H_1^k\}_{i,j \in J} \end{pmatrix}. \end{aligned}$$

And thus

$$\begin{aligned} \Lambda H_1^k \Lambda^\top &= \begin{pmatrix} \{H_1^k\}_{i,j \in I} & \{H_1^k\}_{i \in I, j \in J} \\ 0 & \mathcal{D}\{H_1^k\}_{i \in I, j \in J} + \{H_1^k\}_{i,j \in J} \end{pmatrix} \begin{pmatrix} 1_m & \mathcal{D} \\ 0 & 1_n \end{pmatrix} \\ &= \begin{pmatrix} \{H_1^k\}_{i,j \in I} & \mathcal{D}\{H_1^k\}_{i,j \in I} + \{H_1^k\}_{i \in I, j \in J} \\ 0 & \mathcal{D}\{H_1^k\}_{i \in I, j \in J} + \{H_1^k\}_{i,j \in J} \end{pmatrix} \\ &= \begin{pmatrix} \{H_1^k\}_{i,j \in I} & 0 \\ 0 & \mathcal{D}\{H_1^k\}_{i \in I, j \in J} + \{H_1^k\}_{i,j \in J} \end{pmatrix}, \end{aligned}$$

where in the last equality we use the fact that $\mathcal{D}\{H_1^k\}_{i,j \in I} = -\{H_1^k\}_{i \in J, j \in I}$ for $k \in I$ together with the feature of H_1^k , it is a symmetric matrix. Since $\Lambda H_0 \Lambda^\top = H_0$ the assertion is proved. \square

A.3.2 Transforms

The affine structure permits to reduce the calculation of the Laplace transform of the process $(Z_t)_{t \geq 0}$ and of certain related random variables to the solution of an ODEs

system that can sometimes be solved analytically. We will show that many transforms used to solve several important problems in asset pricing are computable explicitly as exponentially affine functions of $(Z_t)_{t \geq 0}$. We follow the article [24] (Section 2.2) and derive the ODEs system, and consequently the discounted Laplace transform of $(Z_t)_{t \geq 0}$.

Let the short interest rate process $(r(Z_t, t))_{t \geq 0}$ be an affine function

$$r(z, t) = \rho_0(t) + \langle \rho_1(t), z \rangle \quad \rho = (\rho_0, \rho_1), \quad (\text{A.23})$$

for some $\rho_0 : [0, \infty) \rightarrow \mathbb{R}$ and $\rho_1 : [0, \infty) \rightarrow \mathbb{R}^d$ L^1 -integrable functions. A short rate model of this form is called affine short rate model.

We underline the fact that the process characteristics $\chi = (K, H, l, \theta)$ and the parameter ρ capture the distribution of $(Z_t)_{t \geq 0}$ and the effects of the discounting at rate r , respectively. Including ρ in the process coefficients we define the discounted characteristics $\chi^d = (K, H, l, \theta, \rho)$. It is relevant for pricing because it determines the following discounted Laplace transform

$$\Psi(u, Z_t, t, T) := \mathbb{E} \left[e^{-\int_t^T r(Z_s, s) ds} e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t \right] \quad (\text{A.24})$$

for $u \in \mathbb{C}^d$. We will show that, under suitable conditions,

$$\Psi(u, z, t, T) = e^{\phi(t, T, u) + \langle \psi(t, T, u), z \rangle}, \quad (\text{A.25})$$

where $\phi(t, T, u)$ and $\psi(t, T, u)$ are \mathbb{C} and \mathbb{C}^d -valued functions, respectively, that solve the following backward ODEs system:

$$\frac{d\phi}{dt}(t) = \rho_0(t) - \langle K_0, \psi(t) \rangle - \frac{1}{2} \langle \psi(t), H_0 \psi(t) \rangle - \sum_{a=1}^A l_0^a [\theta^a(\zeta_a \psi(t)) - 1] \quad (\text{A.26})$$

$$\phi(T) = 0,$$

$$\frac{d\psi}{dt}(t) = \rho_1(t) - K_1^\top \psi(t) - \frac{1}{2} \psi(t)^\top H_1 \psi(t) - \sum_{a=1}^A l_1^a [\theta^a(\zeta_a \psi(t)) - 1] \quad (\text{A.27})$$

$$\psi(T) = u.$$

We have omitted the dependencies on T and u of the functions ϕ and ψ for simplicity of notation. In (A.27) with $\psi(t)^\top H_1 \psi(t)$ we have denoted the \mathbb{C}^d -valued vector with k -element $\sum_{i,j} \psi_i(t) (H_1)_{ijk} \psi_j(t)$.

Looking at (A.27) we note that the ODE for ψ is a system consisting of d one-dimensional ODEs, while the ODE (A.26) which characterizes ϕ is only one-dimensional. The closed-form solution for these ODEs is not always available, and in this case we have to resort to suitable numerical methods. If (A.27) is solvable, then ϕ is determined by ψ via simple integration

$$\begin{aligned} \phi(t, T, u) = & \int_t^T \left(-\rho_0(s) + \langle K_0, \psi(s, T, u) \rangle \right. \\ & \left. + \frac{1}{2} \langle \psi(s, T, u), H_0 \psi(s, T, u) \rangle + \sum_{a=1}^A l_0^a [\theta^a(\zeta_a \psi(s, T, u)) - 1] \right) ds. \end{aligned}$$

The discounted characteristics $\chi^d = (K, H, l, \theta, \rho)$ fully specify the process $(Z_t)_{t \geq 0}$ and the discounting effects of the associated rate. We give the following definition which is formulated by Duffie et al. [24] and state the fundamental result of this section (see [24], Proposition 1).

Definition A.15. *Suppose that (A.26) and (A.27) are solved uniquely by ϕ and ψ . The characteristics $\chi^d = (K, H, l, \theta, \rho)$ are well-behaved at $(u, T) \in \mathbb{C}^n \times \mathbb{R}_+$ if the following technical integrability conditions hold for $0 \leq t \leq T$:*

1. $\mathbb{E}[|\kappa_t|] < \infty$, where $\kappa_t := e^{-\int_0^t r(Z_s, s) ds} e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle}$;
2. $\mathbb{E}\left[\left(\int_0^t \langle \eta_s, \eta_s \rangle ds\right)^{\frac{1}{2}}\right] < \infty$, where $\eta_t := \kappa_t \psi(t, T, u)^\top \sigma(Z_t)$;
3. $\mathbb{E}\left[\int_0^t |\gamma_s| ds\right] < \infty$, where $\gamma_t := \kappa_t \sum_{a=1}^A \lambda^a(Z_t) [\theta^a(\zeta_a \psi(t, T, u)) - 1]$.

Theorem A.16. *Fix $(u, T) \in \mathbb{C}^d \times \mathbb{R}_+$ and suppose that the discounted characteristics $\chi^d = (K, H, l, \theta, \rho)$ are well-behaved at (u, T) . Then the transform Ψ of $(Z_t)_{t \geq 0}$ defined by (A.24) is given by (A.25), that is*

$$\mathbb{E}\left[e^{-\int_t^T r(Z_s, s) ds} e^{\langle u, Z_T \rangle} \Big| \mathcal{F}_t\right] = e^{\phi(t, T, u) + \langle \psi(t, T, u), z \rangle}.$$

Before giving the theorem's proof we want to observe that both expected value and characteristics $\chi = (K, H, l, \theta)$ are relative to the same probability measure, that is the coefficients are determined under P and the mean is calculated with respect to the same measure P .

Proof. It suffices to show that $(\kappa_t)_{t \in [0, T]}$ is a martingale. Consequently $\kappa_t = \mathbb{E}[\kappa_T | \mathcal{F}_t]$ for all $0 \leq t \leq T$ and thus

$$\begin{aligned} e^{-\int_0^t r(Z_s, s) ds} e^{\phi(t, T, u) + \langle \psi(t, T, u), Z_t \rangle} &= \mathbb{E}\left[e^{-\int_0^T r(Z_s, s) ds} e^{\phi(T, T, u) + \langle \psi(T, T, u), Z_T \rangle} \Big| \mathcal{F}_t\right] \\ &= \mathbb{E}\left[e^{-\int_0^T r(Z_s, s) ds} e^{\langle u, Z_T \rangle} \Big| \mathcal{F}_t\right], \end{aligned}$$

where in the last passage we have used the terminal conditions of the ODEs system. Multiplying by $e^{\int_0^t r(Z_s, s) ds}$ we get the result.

For $0 \leq t \leq T$ obviously κ_t is \mathcal{F}_t -adapted and from the first integrability condition we deduce that $\mathbb{E}[|\kappa_t|] < \infty$. It remains to be proved that for $0 \leq t \leq s \leq T$ it holds the property $\kappa_s = \mathbb{E}[\kappa_t | \mathcal{F}_s]$. To this aim we have to do some work. To begin we use Itô's

formula for $0 \leq t \leq T$:

$$\begin{aligned}
\kappa_t - \kappa_0 &= \int_0^t \kappa_s \left[-r(Z_s, s) + \phi'(s, T, u) + \langle \psi'(s, T, u), Z_s \rangle \right] ds \\
&\quad + \int_0^t \kappa_s \psi(s, T, u)^\top dZ_s^c \\
&\quad + \frac{1}{2} \int_0^t \kappa_s \langle \psi(s, T, u), \sigma(Z_s) \sigma(Z_s)^\top \psi(s, T, u) \rangle ds + \sum_{0 < s \leq t} \Delta \kappa_s \\
&= \int_0^t \kappa_s \left[-r(Z_s, s) + \phi'(s, T, u) + \langle \psi'(s, T, u), Z_s \rangle \right. \\
&\quad \left. + \langle \mu(Z_s), \psi(s, T, u) \rangle + \frac{1}{2} \left\| \sigma(Z_s)^\top \psi(s, T, u) \right\|^2 \right] ds \\
&\quad + \int_0^t \kappa_s \psi(s, T, u)^\top \sigma(Z_s) dW_s + \sum_{a=1}^A \sum_{0 < T_n^a \leq t} (\kappa_{T_n^a} - \kappa_{T_n^a-}),
\end{aligned}$$

where $T_n^a = \inf\{t : J_t^a = n\}$ denotes the n -th jump time of the jump process $(J_t^a)_{t \in [0, T]}$ for $a = 1, \dots, A$. Rearranging the terms it follows that

$$\kappa_t - \kappa_0 = \int_0^t \kappa_s \mu_s^\kappa ds + \int_0^t \eta_s dW_s + D_t, \quad (\text{A.28})$$

for

$$\begin{aligned}
\mu_t^\kappa &= \phi'(t, T, u) + \langle \psi'(t, T, u), Z_t \rangle - r(Z_t, t) + \langle \mu(Z_t), \psi(t, T, u) \rangle \\
&\quad + \frac{1}{2} \left\| \sigma(Z_t)^\top \psi(t, T, u) \right\|^2 + \sum_{a=1}^A \lambda^a(Z_t) [\theta^a(\zeta_a \psi(t, T, u)) - 1]
\end{aligned}$$

and

$$D_t = \sum_{a=1}^A \sum_{0 < T_n^a \leq t} (\kappa_{T_n^a} - \kappa_{T_n^a-}) - \int_0^t \gamma_s ds,$$

where $\gamma_t = \kappa_t \sum_{a=1}^A \lambda^a(Z_t) [\theta^a(\zeta_a \psi(t, T, u)) - 1]$ as defined in the third condition of Definition [A.15](#). From ODEs [\(A.26\)](#) and [\(A.27\)](#) it follows that μ_t^κ is null for $0 \leq t \leq T$. Indeed, omitting T and u dependencies of ψ and ϕ for brevity of notation it follows that

$$\begin{aligned}
\mu_t^\kappa &= \phi'(t) - \rho_0(t) + \langle K_0, \psi(t) \rangle + \frac{1}{2} \langle \psi(t), H_0 \psi(t) \rangle + \sum_{a=1}^A l_0^a [\theta^a(\zeta_a \psi(t)) - 1] \\
&\quad + \left\langle \left\{ \psi'(t) - \rho_1(t) + K_1^\top \psi(t) + \frac{1}{2} \psi(t)^\top H_1 \psi(t) + \sum_{a=1}^A l_1^a [\theta^a(\zeta_a \psi(t)) - 1] \right\}, Z_t \right\rangle.
\end{aligned}$$

Consequently [\(A.28\)](#) becomes

$$\kappa_t - \kappa_0 = \int_0^t \eta_s dW_s + D_t \quad \text{for } 0 \leq t \leq T. \quad (\text{A.29})$$

The second integrability condition implies that the process $(\eta_t)_{t \in [0, T]} \in L^2([0, T])$ and therefore $(\int_0^t \eta_s dW_s)_{t \in [0, T]}$ is a martingale (see [54] (Chapter I, Corollary 4.23)). By the law of iterated expectations for $0 \leq t \leq s \leq T$ we obtain that

$$\begin{aligned}
\mathbb{E} \left[\sum_{a=1}^A \sum_{t < T_n^a \leq s} (\kappa_{T_n^a} - \kappa_{T_n^a-}) \middle| \mathcal{F}_t \right] &= \\
&= \mathbb{E} \left[\sum_{a=1}^A \sum_{t < T_n^a \leq s} \mathbb{E} [\kappa_{T_n^a} - \kappa_{T_n^a-} | J_{T_n^a-}^a, T_n^a] \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[\sum_{a=1}^A \sum_{t < T_n^a \leq s} \mathbb{E} [\kappa_{T_n^a-} (e^{\langle \psi(T_n^a, T, u), \zeta_a \Delta J_{T_n^a-}^a \rangle} - 1) | J_{T_n^a-}^a, T_n^a] \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[\sum_{a=1}^A \sum_{t < T_n^a \leq s} \kappa_{T_n^a-} (\theta^a(\zeta_a \psi(T_n^a, T, u)) - 1) \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[\sum_{a=1}^A \int_t^s \kappa_{v-} (\theta^a(\zeta_a \psi(v, T, u)) - 1) dJ_v^a \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[\sum_{a=1}^A \int_t^s \kappa_{v-} (\theta^a(\zeta_a \psi(v, T, u)) - 1) \lambda^a(Z_v) dv \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[\sum_{a=1}^A \int_t^s \kappa_v (\theta^a(\zeta_a \psi(v, T, u)) - 1) \lambda^a(Z_v) dv \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[\int_t^s \gamma_v dv \middle| \mathcal{F}_t \right].
\end{aligned}$$

The second line is a consequence of the definition of the process $(\kappa_t)_{t \in [0, T]}$, in other words for all fixed $a = 1, \dots, A$ the jump $\Delta \kappa_{T_n^a} = \kappa_{T_n^a} - \kappa_{T_n^a-}$ is generated only by the jump of the process $(J_t^a)_{t \in [0, T]}$ at the stochastic time T_n^a . In the third line we have made the jump transform's definition explicit. While the fourth line follows from [15] (Theorem T8) and in the last line we have only written down the definition of the process $(\gamma_t)_{t \in [0, T]}$. In conclusion, $(D_t)_{t \in [0, T]}$ is a martingale and so is the process $(\kappa_t)_{t \in [0, T]}$. \square

A.3.3 Change of measure

Let $(Z_t)_{t \geq 0}$ be an affine jump-diffusion with characteristics $\chi = (K, H, l, \theta)$ under the objective measure P . Fix the current date t and a future date $T > t$. If we are interested in pricing derivatives with generalized terminal payoff function $e^{\langle u, Z_T \rangle}$ for $u \in \mathbb{R}^d$ we have to compute the expected value of the given payoff under an equivalent risk-neutral measure Q

$$\mathbb{E}^Q \left[e^{-\int_0^T r(Z_s, s) ds} e^{\langle u, Z_T \rangle} \middle| \mathcal{F}_t \right].$$

We underline the fact that the expected value is given under the measure Q . In the previous section we showed that the discounted Laplace transform of an affine jump-diffusion

process has an exponential structure under the measure P . One may ask the question of whether the affine structure is preserved under an equivalent change of probability. The answer is positive if the density process which defines the change of measure is of exponential affine form in the process $(Z_t)_{t \geq 0}$.

Following the ideas of Duffie et al. [24] (Appendix C) we fix $T > 0$ and assume that the characteristics of $(Z_t)_{t \geq 0}$ χ are *well-behaved* at (m, T) for some $m \in \mathbb{R}^d$. It follows that ODEs (A.26) and (A.27) have unique solutions ϕ and ψ .

In this case the process specified by

$$\xi_t := e^{-\int_0^t r(Z_s, s) ds} e^{\phi(t, T, m) + \langle \psi(t, T, m), Z_t \rangle} \quad \text{for } 0 \leq t \leq T \quad (\text{A.30})$$

is a positive martingale, as proved in Theorem A.16. Consequently, we can define an equivalent probability measure $Q(m)$ through the following Radon-Nikodym derivative

$$\frac{dQ(m)}{dP} = \frac{\xi_T}{\xi_0},$$

where ξ_0 is a scale factor such that $\mathbb{E}\left[\frac{dQ(m)}{dP}\right] = 1$. We remark that the new measure depends on the choice of the parameter m .

The dynamic of the process $(Z_t)_{t \geq 0}$ changes under the new measure $Q(m)$ and so does its characteristics but the density process defined above preserves its affine structure, as we will show in the next theorem. Under $Q(m)$ the process $(Z_t)_{t \geq 0}$ is an affine jump-diffusion with different characteristics χ .

Theorem A.17. *Under the new measure $Q(m)$ the process $(Z_t)_{t \geq 0}$ preserves its affine structure with characteristics $\chi^{Q(m)} = (K^{Q(m)}, H^{Q(m)}, l^{Q(m)}, \theta^{Q(m)})$ given by*

$$\begin{aligned} K_0^{Q(m)}(t) &= K_0 + H_0 \psi(t, T, m); \\ K_1^{Q(m)}(t) &= K_1 + H_1 \psi(t, T, m); \\ H^{Q(m)}(t) &= H; \\ l_0^{aQ(m)}(t) &= l_0^a \theta^a (\zeta_a \psi(t, T, m)) \quad \text{for } a = 1, \dots, A; \\ l_1^{aQ(m)}(t) &= l_1^a \theta^a (\zeta_a \psi(t, T, m)) \quad \text{for } a = 1, \dots, A; \\ \theta^{aQ(m)}(c, t) &= \frac{\theta^a (c + \zeta_a \psi(t, T, m))}{\theta^a (\zeta_a \psi(t, T, m))} \quad \text{for } c \in \mathbb{C}^d \text{ and } a = 1, \dots, A. \end{aligned} \quad (\text{A.31})$$

Proof. Let

$$W_t^{Q(m)} = W_t - \int_0^t \sigma(Z_s)^\top \psi(s, T, m) ds \quad \text{for } 0 \leq t \leq T. \quad (\text{A.32})$$

By integration by parts and [38] (Chapter I, Theorem 4.52) for $0 \leq s \leq t \leq T$ it follows

that

$$\begin{aligned}
\xi_t W_t^{Q(m)} &= \xi_s W_s^{Q(m)} + \int_s^t \xi_{u-} dW_u^{Q(m)} + \int_s^t W_{u-}^{Q(m)} d\xi_u \\
&\quad + \sum_{s < u \leq t} (\xi_u - \xi_{u-})(W_u^{Q(m)} - W_{u-}^{Q(m)}) + \int_s^t d[\xi, W^{Q(m)}]_u^c \\
&= \xi_s W_s^{Q(m)} + \int_s^t \xi_{u-} (dW_u - \sigma(Z_u)^\top \psi(u, T, m) du) + \int_s^t W_u^{Q(m)} d\xi_u \\
&\quad + 0 + \int_s^t \xi_u \sigma(Z_u)^\top \psi(u, T, m) du \\
&= \xi_s W_s^{Q(m)} + \int_s^t \xi_{u-} dW_u + \int_s^t W_u^{Q(m)} d\xi_u.
\end{aligned}$$

Since $(W_t)_{t \in [0, T]}$ and $(\xi_t)_{t \in [0, T]}$ are martingales under the measure P , $(\xi_{t-})_{t \in [0, T]}$ and $(W_{t-}^{Q(m)})_{t \in [0, T]}$ predictable processes (see [38] (Chapter I, Section 4.d)) it follows that both $(\int_0^t \xi_{u-} dW_u)_{t \geq 0}$ and $(\int_0^t W_u^{Q(m)} d\xi_u)_{t \geq 0}$ are P -local martingales. Hence, the process $(\xi_t W_t^{Q(m)})_{t \in [0, T]}$ is a P -local martingale since it is sum of P -local martingales. From [38] (Chapter III, Proposition 3.8) we deduce that $(W_t^{Q(m)})_{t \in [0, T]}$ is a $Q(m)$ -local martingale because $(\xi_t)_{t \in [0, T]}$ is the density process which defines the change of measure.

Since $(\int_0^t \sigma(Z_s)^\top \psi(s, T, m) ds)_{t \in [0, T]}$ is a continuous process of finite variation, from [38] (Chapter I, Proposition 4.49) we compute

$$[W_i^{Q(m)}, W_j^{Q(m)}]_t = [W_i, W_j]_t + 0 = \delta_{ij} t$$

for every $0 \leq t \leq T$ and $i, j = 1, \dots, d$. By Lévy's characterization of Brownian motion (see [38] (Chapter II, Theorem 4.4)) we derive that $(W_t^{Q(m)})_{t \in [0, T]}$ is a Brownian motion under the measure $Q(m)$.

For $0 \leq t \leq T$ let

$$M_t^{aQ(m)} := J_t^a - \int_0^t \theta^a(\zeta_a \psi(s, T, m)) \lambda^a(Z_s) ds.$$

By Itô's formula and [38] (Chapter I, Theorem 4.52) with $0 \leq s \leq t \leq T$

$$\begin{aligned}
\xi_t M_t^{aQ(m)} &= \xi_s M_s^{aQ(m)} + \int_s^t \xi_{u-} dM_u^{aQ(m)} + \int_s^t M_{u-}^{aQ(m)} d\xi_u \\
&\quad + \sum_{s < u \leq t} (\xi_u - \xi_{u-})(M_u^{aQ(m)} - M_{u-}^{aQ(m)}) + \int_s^t d[\xi, M^{aQ(m)}]_u^c \\
&= \xi_s M_s^{aQ(m)} + \int_s^t \xi_{u-} \left(dJ_u^a - \theta^a(\zeta_a \psi(u, T, m)) \lambda^a(Z_u) du \right) \\
&\quad + \int_s^t M_{u-}^{aQ(m)} d\xi_u + \sum_{s < u \leq t} (\xi_u - \xi_{u-})(J_u^a - J_{u-}^a) + 0.
\end{aligned} \tag{A.33}$$

We define $(\Delta\xi_t^a)_{t \in [0, T]}$ as the jump induced by the a -th jump process in $(\xi_t)_{t \in [0, T]}$, that is

$$\Delta\xi_t^a := \begin{cases} \xi_{t-}(e^{\langle \psi(t, T, m), \Delta J_t^a \rangle} - 1) & \text{if } t = T_n^a \text{ for some } n \\ 0 & \text{otherwise,} \end{cases}$$

where we remind that T_n^a denotes the n -th jump time of the jump process $(J_t^a)_{t \in [0, T]}$. Consequently (A.33) becomes

$$\begin{aligned} \xi_t M_t^{aQ(m)} &= \xi_s M_s^{aQ(m)} + \int_s^t \xi_{u-} dJ_u^a - \int_s^t \xi_u \theta^a(\zeta_a \psi(u, T, m)) \lambda^a(Z_u) du \\ &\quad + \int_s^t M_{u-}^{aQ(m)} d\xi_u + \sum_{s < u \leq t} \Delta\xi_u^a \\ &= \xi_s M_s^{aQ(m)} + \int_s^t \xi_{u-} (dJ_u^a - \lambda^a(Z_u) du) + \int_s^t \xi_u \lambda^a(Z_u) du \\ &\quad - \int_s^t \xi_u \theta^a(\zeta_a \psi(u, T, m)) \lambda^a(Z_u) du + \int_s^t M_{u-}^{aQ(m)} d\xi_u + \sum_{s < u \leq t} \Delta\xi_u^a \\ &= \xi_s M_s^{aQ(m)} + \int_s^t \xi_{u-} dM_u^a \\ &\quad + \int_s^t \xi_u [\lambda^a(Z_u) - \theta^a(\zeta_a \psi(u, T, m)) \lambda^a(Z_u)] du \\ &\quad + \int_s^t M_{u-}^{aQ(m)} d\xi_u + \sum_{s < u \leq t} \Delta\xi_u^a \\ &= \xi_s M_s^{aQ(m)} + \int_s^t \xi_{u-} dM_u^a \\ &\quad - \int_s^t \xi_u [\theta^a(\zeta_a \psi(u, T, m)) \lambda^a(Z_u) - \lambda^a(Z_u)] du \\ &\quad + \int_s^t M_{u-}^{aQ(m)} d\xi_u + \sum_{s < u \leq t} \Delta\xi_u^a \\ &= \xi_s M_s^{aQ(m)} + \int_s^t \xi_{u-} dM_u^a + \int_s^t M_{u-}^{aQ(m)} d\xi_u + J_t^\xi. \end{aligned}$$

In the third equality we set $M_t^a := J_t^a - \int_0^t \lambda^a(Z_s) ds$ and in the last line

$$J_t^\xi := \sum_{s < u \leq t} \Delta\xi_u^a - \int_s^t \xi_u [\theta^a(\zeta_a \psi(u, T, m)) - 1] \lambda^a(Z_u) du.$$

Since $(\xi_t)_{t \in [0, T]}$ and $(M_t^a)_{t \in [0, T]}$ are P -martingales, referring to [38] (Chapter I, Section 4.d) we state that $(\int_0^t \xi_{u-} dM_u^a)_{t \in [0, T]}$ and $(\int_0^t M_{u-}^{aQ(m)} d\xi_u)_{t \in [0, T]}$ are P -local martingales. Consequently $(\xi_t M_t^{aQ(m)})_{t \in [0, T]}$ is a P -local martingale if and only if $(J_t^\xi)_{t \in [0, T]}$ is a P -local martingale. By the law of iterated expectations following a reasoning similar to

that used in the proof of Theorem [A.16](#), for $0 \leq s \leq t \leq T$

$$\begin{aligned}
\mathbb{E} \left[\sum_{s < u \leq t} \Delta \xi_u^a \middle| \mathcal{F}_s \right] &= \mathbb{E} \left[\sum_{s < T_n^a \leq t} \mathbb{E} \left[\Delta \xi_{T_n^a}^a \middle| J_{T_n^a-}^a, T_n^a \right] \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[\sum_{s < T_n^a \leq t} \mathbb{E} \left[\xi_{T_n^a-}^a \left(e^{\langle \psi(T_n^a, T, m), \Delta J_{T_n^a}^a \rangle} - 1 \right) \middle| J_{T_n^a-}^a, T_n^a \right] \middle| \mathcal{F}_s \right] \\
&= \mathbb{E} \left[\sum_{s < T_n^a \leq t} \xi_{T_n^a-}^a \left(\theta^a(\zeta_a \psi(T_n^a, T, m)) - 1 \right) \middle| \mathcal{F}_s \right] \\
&= \mathbb{E} \left[\int_s^t \xi_{u-}^a \left(\theta^a(\zeta_a \psi(u, T, m)) - 1 \right) dJ_u^a \middle| \mathcal{F}_s \right] \\
&= \mathbb{E} \left[\int_s^t \xi_u^a \left(\theta^a(\zeta_a \psi(u, T, m)) - 1 \right) \lambda^a(Z_u) du \middle| \mathcal{F}_s \right],
\end{aligned}$$

where the last equality is a consequence of [\[15\]](#) (Chapter II, Theorem T8) because the a -th jump process has intensity $(\lambda^a(Z_t))_{t \in [0, T]}$ under the measure P and the process $(\xi_{t-}^a (\theta^a(\zeta_a \psi(t, T, m)) - 1))_{t \in [0, T]}$ is \mathcal{F}_t -predictable.

It follows that $M^{aQ(m)}$ is a $Q(m)$ -local martingale. By the martingale characterization of intensity (see [\[15\]](#) (Chapter II, Theorem T9)) we conclude that $(J_t^a)_{t \in [0, T]}$ is a counting process with $Q(m)$ -intensity specified by $\lambda^{aQ(m)}(z, t) = l_0^{aQ(m)}(t) + \langle l_1^{aQ(m)}(t), z \rangle$.

If we rewrite the SDE [\(A.12\)](#) employing [\(A.32\)](#) the drift term becomes

$$\mu^{Q(m)}(z, t) = \mu(z) + H(z)\psi(t, T, m),$$

whereas the volatility coefficient H obviously remains unchanged after the change of probability measure.

Following the notation already used we employ $(\Delta J_t^a)_{t \in [0, T]}$ to denote the jump of the a -th jump process. Consequently, ΔJ_t^a is different from zero only if $t = T_n^a$ for some n , where T_n^a denotes the n -th jump time. For $c \in \mathbb{C}^d$ and $t = T_n^a$ the following holds:

$$\begin{aligned}
\lambda^{aQ(m)}(Z_t, t) \theta^{aQ(m)}(c, t) &= \mathbb{E}^{Q(m)} \left[e^{\langle c, \Delta J_t^a \rangle} \middle| \mathcal{F}_{t-} \right] \\
&= \mathbb{E}^P \left[\frac{\xi_t^a}{\xi_{t-}^a} e^{\langle c, \Delta J_t^a \rangle} \middle| \mathcal{F}_{t-} \right].
\end{aligned}$$

The first equality is deduced from the independence of the state process $(Z_t)_{t \in [0, T]}$ and the jump distribution $\nu = (\nu^1, \dots, \nu^A)^\top$. In the second line we have expressed the change of measure through the density process (see [\[38\]](#) (Chapter III, Proposition 3.8)). If we

make the ratio $\frac{\xi_t}{\xi_{t-}}$ more explicit remembering that $t = T_n^a$, then it follows that

$$\begin{aligned}
\lambda^{aQ(m)}(Z_t, t)\theta^{aQ(m)}(c, t) &= \mathbb{E}^P \left[e^{\langle \psi(t, T, m), \Delta Z_t^a \rangle} e^{\langle c, \Delta J_t^a \rangle} \middle| \mathcal{F}_{t-} \right] \\
&= \mathbb{E}^P \left[e^{\langle \zeta_a \psi(t, T, m), \Delta J_t^a \rangle} e^{\langle c, \Delta J_t^a \rangle} \middle| \mathcal{F}_{t-} \right] \\
&= \mathbb{E}^P \left[e^{\langle \zeta_a (\psi(t, T, m) + c), \Delta J_t^a \rangle} \middle| \mathcal{F}_{t-} \right] \\
&= \lambda^a(Z_t)\theta^a(\zeta_a \psi(t, T, m) + c, t) \\
&= \lambda^{aQ(m)}(Z_t, t) \frac{\theta^a(\zeta_a \psi(t, T, m) + c, t)}{\theta^a(\zeta_a \psi(t, T, m), t)}.
\end{aligned}$$

In the second line we used the relation $\Delta Z_t^a = \zeta_a \Delta J_t^a$ that is deduced from the SDE (A.12), while the following equality is a consequence of the independence of the state process $(Z_t)_{t \in [0, T]}$ and the jump distribution $\nu = (\nu^1, \dots, \nu^A)^\top$. In the last line we have substituted the definition of $\lambda^{aQ(m)}(t, Z_t) = \theta^a(\zeta_a \psi(t, T, m))\lambda^a(Z_t)$. \square

Appendix B

Stochastic integration

In this appendix we summarize some notions on stochastic integration that are needed in the study of affine jump-diffusions.

From now on our attention will be on 1-dimensional processes (the following theory can be easily extended to dimensions) in the time interval $[0, T]$ for a fixed $T > 0$. The present discussion is based on [54, 55].

B.1 Brownian integration

We work on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, P)$. The fact that the Brownian motion $(W_t)_{t \in [0, T]}$ does not have bounded variation prevents us to define pathwise the integral with respect to $(W_t)_{t \in [0, T]}$ in the Riemann-Stieltjes sense. On the other hand, it has finite quadratic variation and this property makes it possible to construct the stochastic integral for suitable classes of integrands.

Definition B.1. *The stochastic process $(u_t)_{t \in [0, T]}$ belongs to the class \mathbb{L}^2 if*

1. *u is progressively measurable;*
2. *$u \in L^2([0, T] \times \Omega)$, that is $\mathbb{E}[\int_0^T u_t^2 dt] < \infty$.*

If $(u_t)_{t \in [0, T]} \in \mathbb{L}^2$ is a simple process of the form

$$u_t = \sum_{n=1}^N u_n \mathbf{1}_{(t_{n-1}, t_n]}(t) \quad \text{for } 0 \leq t \leq T,$$

where $0 \leq t_0 < t_1 < \dots < t_N \leq T$ is a partition of the interval $[0, T]$ and for $n = 1, \dots, N$ u_n is a $\mathcal{F}_{t_{n-1}}$ -measurable random variable, then the Itô integral is defined in the following way (see [54] (Chapter 4, Definition 4.2)):

$$\int_0^T u_t dW_t := \sum_{n=1}^N u_n (W_{t_n} - W_{t_{n-1}}). \quad (\text{B.1})$$

We can extend the definition of stochastic integral to all stochastic processes in \mathbb{L}^2 . For $(u_t)_{t \in [0, T]} \in \mathbb{L}^2$ there exists a sequence $((u_t^n)_{t \in [0, T]})_{n \geq 0}$ of simple processes approximating the process $(u_t)_{t \in [0, T]}$ in $L^2([0, T] \times \Omega)$ (see [54] (Chapter 4, Lemma 4.8)). Thus we define

$$\int_0^T u_t dW_t := \lim_{n \rightarrow \infty} \int_0^T u_t^n dt \text{ in } L^2(\Omega). \quad (\text{B.2})$$

We use convergence in $L^2(\Omega)$ because sums in (B.1) do not converge pathwise because the paths of the Brownian motion are not of bounded variation P -a.s.. Now, we further extend the class of processes for which the stochastic Brownian integral is defined.

Definition B.2. *The stochastic process $(u_t)_{t \in [0, T]}$ belongs to the class \mathbb{L}_{loc}^2 if*

1. u is progressively measurable;
2. $\int_0^T u_t^2 dt < \infty$ P -a.s..

From [54] (Chapter 4, Section 4.4) it follows that for every $(u_t)_{t \in [0, T]} \in \mathbb{L}_{loc}^2$ there exists a sequence $((u_t^n)_{t \in [0, T]})_{n \geq 0}$ of simple processes approximating $(u_t)_{t \in [0, T]}$ in probability. We define the stochastic integral

$$\int_0^T u_t dW_t := \lim_{n \rightarrow \infty} \int_0^T u_t^n dW_t \text{ in probability.} \quad (\text{B.3})$$

B.2 Extension of stochastic integration theory

We follow the arguments presented in [55] (Chapter II, Section 4) and consider an increasing process $(C_t)_{t \in [0, T]}$, that is a càdlàg process whose paths are non-decreasing. Since it is an increasing process, we can easily use the Lebesgue-Stieltjes integral to define the stochastic integral with respect to $(C_t)_{t \in [0, T]}$. This procedure is based on the idea that any increasing right-continuous function f can be identified with a unique measure

$$\mu((u, s]) := f(s) - f(u),$$

for an arbitrary interval $(u, s] \subseteq [0, T]$. Then, a non-negative random measure ν_C is associated to $(C_t)_{t \in [0, T]}$ and consequently, if $(u_t)_{t \in [0, T]}$ is a bounded process, then the integral

$$\int_0^t u_s dC_s := \int_0^t u_s \nu_C(ds) \quad (\text{B.4})$$

is well-defined for each $0 \leq t \leq T$. It follows that the constructed integral is a càdlàg process.

We proceed analogously for a finite variation process $(C_t)_{t \in [0, T]}$, but in this case the induced measure $\nu_C(ds)$ is a signed measure. We can define the integral process as in (B.4), that is

$$\int_0^t u_s dC_s := \int_0^t u_s \nu_C(ds) \text{ for } 0 \leq t \leq T, \quad (\text{B.5})$$

where $(u_t)_{t \in [0, T]}$ is a bounded stochastic process.

According to [25] (Chapter 3, Section 3.2) we can replace boundedness of $(u_t)_{t \in [0, T]}$ by the weaker requirement of local boundedness. In such a case, (B.5) is well-defined and from [25] (Proposition 3.22) it follows that the constructed stochastic integral is a càdlàg process with finite variation. Furthermore, it holds that

$$\Delta \left(\int_0^t u_s dC_s \right) = \int_0^t u_s dC_s - \int_0^{t^-} u_s dC_s = u_t \Delta C_t.$$

If we consider the pure jump process $(J_t)_{t \in [0, T]}$ given in Definition A.8 as the integrator process, then the definition of stochastic integral is exactly

$$\int_0^t u_s dJ_s := \sum_{n \geq 1} u_{T_n} 1_{\{T_n \leq t\}} \quad \text{for } 0 \leq t \leq T, \quad (\text{B.6})$$

$(u_t)_{t \in [0, T]}$ is a locally bounded measurable stochastic process. Obviously $(T_n)_{n \geq 0}$ are the jump times related to $(J_t)_{t \in [0, T]}$. We highlight that (B.6) is reduced to be a finite sum from the non-explosion property of the point process $(T_n)_{n \geq 0}$ and therefore it has finitely many jumps in every finite time interval.

We point out that the Lebesgue-Stieltjes integral determined by (A.6) in the Hawkes intensity definition is exactly (B.6), with the only difference that in that case we are considering a multivariate counting process.

If the integrator does not have finite variation, then we must introduce suitable conditions such that the stochastic integral is well-defined.

Definition B.3. We let \mathbb{L} denote the space of càglàd adapted processes and \mathbb{D} indicate the space of càdlàg adapted processes. The predictable σ -algebra \mathcal{P} is the smallest σ -algebra on $[0, T] \times \Omega$ making all processes in \mathbb{L} measurable.

Denoting with \mathcal{A} the σ -algebra on $[0, T] \times \Omega$ generated by all progressively measurable processes it follows that $\mathcal{P} \subseteq \mathcal{A}$.

Definition B.4. A stochastic process $(u_t)_{t \in [0, T]}$ is predictable if it is \mathcal{P} -measurable.

Loosely speaking, we can think that $(u_t)_{t \in [0, T]}$ is predictable if u_t is \mathcal{F}_{t^-} -measurable, where \mathcal{F}_{t^-} stands for the information up to but excluding time t . Let $(u_t)_{t \in [0, T]} \in \mathbb{L}$ be a predictable process, then it is simple if it is of the form

$$u_t = \sum_{n=1}^N u_n 1_{(t_{n-1}, t_n]}(t) \quad \text{for } 0 \leq t \leq T,$$

where $0 \leq t_0 < \dots < t_N \leq T$ are stopping times and for fixed $n = 1, \dots, N$ u_n is a bounded random variable that is $\mathcal{F}_{t_{n-1}}$ -measurable. The stochastic integral of $(u_t)_{t \in [0, T]} \in \mathbb{L}$ with respect to a stochastic process $(C_t)_{t \in [0, T]} \in \mathbb{D}$ is defined as

$$\int_0^t u_s dC_s := \sum_{n=1}^N u_n (C_{t \wedge t_n} - C_{t \wedge t_{n-1}}) \quad \text{for } 0 \leq t \leq T. \quad (\text{B.7})$$

Because of the space of simple predictable processes is dense in \mathbb{L} under the uniform convergence in probability (see [54] (Theorem 14.14)), then for any $(u_t)_{t \in [0, T]} \in \mathbb{L}$ there exists a sequence $((u_t^n)_{t \in [0, T]})_{n \geq 0}$ of simple predictable processes approximating $(u_t)_{t \in [0, T]}$ uniformly in probability. Consequently we can define the stochastic integral of a predictable process $(u_t)_{t \in [0, T]} \in \mathbb{L}$ as

$$\int_0^t u_s dC_s := \lim_{n \rightarrow \infty} \int_0^t u_s^n dC_s \quad \text{uniformly in probability for } 0 \leq t \leq T. \quad (\text{B.8})$$

The definition of stochastic integral can be further extended but further generalizations are beyond the scope of this discussion.

Appendix C

Matlab Implementation

The Matlab code that have been used to simulate the short rate process $(r_t)_{t \geq 0}$ and its stochastic lower bound $(SLB_t)_{t \geq 0}$ is reported in this Appendix. It is based on the algorithm which is exhaustively explained in Chapter 3.

```
1 % Model's simulation
2 %
3 % Restriction: h=0,  $\mathbb{R}_+$ -valued X
4 %
5 % Inputs:
6 % model's parameters:
7 % X0 = initial value of X
8 % dX = (K0 + K1 X) dt + \sqrt{X} dW_t
9 % r = c (N_1 - N_2) + \langle ell, X \rangle
10 % N^1- > lam1 = lambda01 + \langle Lam1, X \rangle + delta1 \int_0^t e^{-\gamma_1(t-s)} dN^1_s
11 % N^2- > lam2 = lambda02 + \langle Lam2, X \rangle + delta2 \int_0^t e^{-\gamma_2(t-s)} dN^2_s
12 % simulation's parameters:
13 % T = final time (we work on [0, T])
14 % sim = number of time intervals for the discretization
15 %
16 % Output :
17 % number1 = number of jumps of N^1
18 % setT1 = jumps' times of N^1
19 % number2 = number of jumps of N^2
20 % setT2 = jumps' times of N^2
21 % plots for X, N_1, N_2, SLB and r
22
23 function [number1, setT1, number2, setT2] = simHw(X0, K0, K1, Lam1, Lam2,
           lambda01, lambda02, gamma1, gamma2, delta1, delta2, T, sim, ell, c)
24
25 % -> Code to simulate the diffusion process X (dim 1):
26
27 % Time increments
28 dt = T/sim; % length of time intervals
29 J = [1:sim+1]; % Matlab don't start from 0
30 x = [0:dt:T];
31
32 % Inizialization
```

```

33 X(1) = X0;
34
35 % Maximum of  $\langle \text{Lambda}^a, X_t \rangle$  on  $[0, T]$ 
36 max1=Lam1*X0;
37 max2=Lam2*X0;
38
39 % Simulate a trajectory
40 for j=2:sim+1 % because 0, dt, 2dt, ..., T=ndt (Matlab can't start from 0)
41     X(j) = (((dt^2)*normrnd(0,1)+sqrt(((dt^2)*normrnd(0,1))^2+4*(X(j-1)+(K0
42         -0.5)*dt)*(1+K1*dt))))^2)/(2*(1+K1*dt))^2;
43     max1=max(max1,Lam1*X(j));
44     max2=max(max2,Lam2*X(j));
45 end
46
47 % Visualize the trajectory of X
48 set(0, 'DefaultAxesColorOrder', [0.9290 0.6940 0.1250; 0 0 1]);
49 figure;
50 plot(x,X);
51 xlabel('Time t', 'FontSize', 16)
52 ylabel('X', 'FontSize', 16)
53 title('Simulation of a X trajectory')
54
55 % -> Code to simulate the 2-dim Hawkes process:
56
57 % Initialization
58 n1 = 1; % counter for the process N^1
59 n2 = 1; % counter for the process N^2
60 m = 1; % counter for the candidate points
61 s(1) = 0; % first candidate point of the Hawkes process
62 setT1 = []; % array of event times for N^1 (empty)
63 setT2 = []; % array of event times for N^2 (empty)
64
65 % Determination of jumps
66 while s(m) < T % as long as the candidate points are smaller than T we
67     generate new points
68     % Determine the value of lambda bar
69     if s(m) == 0
70         lambdabar1(m) = lambda01 + Lam1*X0;
71         lambdabar2(m) = lambda02 + Lam2*X0;
72     else
73         if n1 > 1 % check if there are already points in the Hawkes process N
74             ^1, if so calculate the right value of lambda bar
75             sum1 = 0;
76             for i = 1 : 1 : size(setT1, 2)
77                 sum1 = sum1 + delta1 * exp(-gamma1*(s(m) - setT1(i))); % if setT1
78                 isn't empty, lambabar is the sum of the intensity of the individual
79                 points and the maximum
76             end
77             lambdabar1(m) = lambda01 + sum1(end) + delta1 + max1; % the new
78             point increases the intensity with delta1
79         else % when there aren't any points in the Hawkes process N^1 yet
80             lambdabar1(m) = lambda01 + max1;

```

```

80     end
81     if n2 > 1 % check if there are already points in the Hawkes process N
      ^2
82         sum2 = 0;
83         for i = 1 : 1 : size(setT2, 2)
84             sum2 = sum2 + delta2 * exp(-gamma2*(s(m) - setT2(i)));
85         end
86         lambdabar2(m) = lambda02 + sum2(end) + delta2 + max2;
87     else % when there aren't any points in the Hawkes process N^2 yet
88         lambdabar2(m) = lambda02 + max2;
89     end
90 end
91
92 % Update the candidate point
93 w(m) = - log(rand)/(lambdabar1(m) + lambdabar2(m));
94 s(m+1) = s(m) + w(m); % generate the new candidate point
95
96 % Determine the value of the intensity for the new candidate point
97 if s(m+1) <= T
98     if s(m+1) == 0
99         lambdanew1(m) = lambda01 + Lam1*X0;
100        lambdanew2(m) = lambda02 + Lam2*X0;
101    else
102        cal = J(J.*dt >= s(m+1));
103        ind = cal(1);
104        if n1 > 1 % check if there are already point in the Hawkes process
          N^1, if so calculate the right value of lambda
105            sumnew1 = 0;
106            for i = 1 : 1 : size(setT1, 2)
107                sumnew1 = sumnew1 + delta1 * exp(-gamma1*(s(m+1) - setT1(i
          )))
108            end
109            lambdanew1(m) = lambda01 + sumnew1(end) + delta1 + Lam1*X(ind)
          ;
110        else
111            lambdanew1(m) = lambda01 + Lam1*X(ind); % there aren't any
          points in the Hawkes process N^1
112        end
113        if n2 > 1 % check if there are already point in the Hawkes process
          N^2, if so calculate the right value of lambda
114            sumnew2 = 0;
115            for i = 1 : 1 : size(setT2, 2)
116                sumnew2 = sumnew2 + delta2 * exp(-gamma2*(s(m+1) - setT2(i
          )))
117            end
118            lambdanew2(m) = lambda02 + sumnew2(end) + delta2 + Lam2*X(ind)
          ;
119        else
120            lambdanew2(m) = lambda02 + Lam2*X(ind); % there aren't any
          points in the Hawkes process N^2
121        end
122    end
123

```

```

124 % Check the new point
125 D(m+1) = rand; % generate a random variable
126 if D(m+1) * (lambdabar1(m) + lambdabar2(m)) <= lambdanew1(m) + lambdanew2(
    m) % determine if the point needs to be rejected
127 if D(m+1) * (lambdabar1(m) + lambdabar2(m)) <= lambdanew1(m)
128     setT1(n1) = s(m+1); % if not rejected the point is added to the Hawkes
        process N^1
129     n1 = n1 + 1; % the counter is raised by one
130 else
131     setT2(n2) = s(m+1); % if not rejected the point is added to the Hawkes
        process N^2
132     n2 = n2 + 1; % the counter is raised by one
133 end
134 end
135 end
136 m = m + 1; % otherwise we go on to the next candidate point
137 end
138
139 % Remove the first point from s since nothing occurred at time 0
140 s(1) = []; % remove starting point
141 m = m - 1; % one element of array is removed thus lower the array length
142 n1 = n1 - 1; % had to start at 1
143 n2 = n2 - 1; % had to start at 1
144
145 % Remove points that happened after T
146 if m > 0
147     if s(m) > T % check for the Poisson process
148         s(m) = []; % if the last value is bigger than T we remove it from the
            array
149         m = m - 1; % in this case the array size decreases by one since one point
            is removed
150     end
151 end
152 if n1 > 0
153     if setT1(n1) > T
154         setT1(n1) = [];
155         n1 = n1 - 1;
156     end
157 end
158 if n2 > 0
159     if setT2(n2) > T
160         setT2(n2) = [];
161         n2 = n2 - 1;
162     end
163 end
164
165 number1=length(setT1);
166 number2=length(setT2);
167
168
169 % -> Plot the graphs:
170 set(0, 'DefaultAxesColorOrder', [0 0.4470 0.7410]);
171 hold on

```

```

172
173 if isempty(setT1) == 0
174 figure;
175 y = @(e) 0*e;
176 plot([0 setT1(1)],y([0 setT1(1)]),'-')
177 for i=1:l:n1-1
178     x = linspace(setT1(i), setT1(i+1), 100); % plot between two event times
179     y = @(e) 1+y(e);
180     scatter(setT1(i), y(setT1(i)),80,'.')
181     plot(x, y(x), '-')
182 end
183 if setT1(n1) ~= T
184     x = linspace(setT1(n1), T, 100);
185     y = @(e) 1+y(e);
186     scatter(setT1(n1), y(setT1(n1)),80,'.')
187     plot(x, y(x), '-b')
188 else
189     y = @(e) 1+y(e);
190     scatter(setT1(n1), y(setT1(n1)),80,'.')
191 end
192 xlabel('Time')
193 ylabel('N^1')
194 title(['Simulation of the first component of N with parameters lambda_0^1 =
        ',num2str(lambda01), ', delta^1 = ', num2str(delta1), ' and gamma^1 =
        ',num2str(gamma1)])
195 end
196
197 if isempty(setT2) == 0
198 figure;
199 y = @(e) 0*e;
200 plot([0 setT2(1)],y([0 setT2(1)]),'-')
201 for i=1:l:n2-1
202     x = setT2(i) : 0.0001 : setT2(i+1); % plot between two event times
203     y = @(e) 1+y(e);
204     scatter(setT2(i), y(setT2(i)),80,'.')
205     plot(x, y(x), '-')
206 end
207 if setT2(n2) ~= T
208     x = linspace(setT2(n2),T,100);
209     y = @(e) 1+y(e);
210     scatter(setT2(n2), y(setT2(n2)),80,'.')
211     plot(x, y(x), '-')
212 end
213 xlabel('Time')
214 ylabel('N^2')
215 title(['Simulation of the second component of N with parameters lambda_0^2
        = ',num2str(lambda02), ', delta^2 = ', num2str(delta2), ' and gamma^2 =
        ',num2str(gamma2)])
216 end
217
218 T1T2 = sort([setT1(:)',setT2(:)']);
219 n12=length(T1T2);
220 if isempty(T1T2) == 0

```

```

221 figure;
222 y = @(e) 0*e;
223 plot([0 T1T2(1)],y([0 T1T2(1)]),'-')
224 for i=1:1:n12-1
225     x = linspace(T1T2(i), T1T2(i+1), 100); % plot between two event times
226     if ismember(T1T2(i),setT1) == 1
227         y = @(e) c+y(e);
228     else
229         y = @(e) -c+y(e);
230     end
231     scatter(T1T2(i), y(T1T2(i)),80,'.')
232     plot(x, y(x), '-b')
233 end
234 if T1T2(n12) ~= T
235     x = linspace(T1T2(n12), T, 100);
236     if ismember(T1T2(i),setT1) == 1
237         y = @(e) c+y(e);
238     else
239         y = @(e) -c+y(e);
240     end
241     scatter(T1T2(n12), y(T1T2(n12)),80,'.')
242     plot(x, y(x), '-b')
243 else
244     if ismember(T1T2(i),setT1) == 1
245         y = @(e) c+y(e);
246     else
247         y = @(e) -c+y(e);
248     end
249     scatter(T1T2(n1n2), y(T1T2(n1n2)),80,'.')
250 end
251 xlabel('Time')
252 ylabel('SLB')
253 title('Simulation of the SLB')
254 end
255
256 figure;
257 y=ell*X;
258 for i=1:n12
259     cal=J(J*dt>=T1T2(i));
260     if ismember(T1T2(i),setT1) == 1
261         y(cal)=y(cal)+c;
262     else
263         y(cal)=y(cal)-c;
264     end
265 end
266 x = [0:dt:T];
267 plot(x, y, '-b')
268 xlabel('Time')
269 ylabel('SLB')
270 title('Simulation of the short rate')
271
272 hold off
273 end

```

Bibliography

- [1] A. Alfonsi (2015), *Affine diffusions and related processes: simulation, theory and applications* (Vol. 6, p. 20), Cham, Switzerland: Springer.
- [2] A. Alfonsi (2005), On the discretization schemes for the CIR (and Bessel squared) processes, *Monte Carlo Methods and Applications*, 11(4), 355–384.
- [3] H. Amann (2011), *Ordinary differential equations: an introduction to nonlinear analysis* (Vol. 13), Walter de Gruyter.
- [4] S. Angrick, N. Nemoto (2017), Central banking below zero: the implementation of negative interest rates in Europe and Japan, *Asia Europe Journal*, 15(4): 417–443.
- [5] A. Antonov, M. Konikov, M. Spector (2015), The free boundary SABR: natural extension to negative rates (available at SSRN 2557046).
- [6] D. Applebaum (2009), *Lévy processes and stochastic calculus*, Cambridge university press.
- [7] C. Arteta, M. A. Kose, M. Stocker, T. Taskin (2016), *Negative interest rate policies: Sources and implications*, The World Bank.
- [8] M. Barczy, M. B. Alaya, A. Kebaier, G. Pap (2018), Asymptotic properties of maximum likelihood estimator for the growth rate for a jump–type CIR process based on continuous time observations, *Stochastic Processes and their Applications*, 128(4), 1135–1164.
- [9] E. Bacry, S. Delattre, M. Hoffmann, J. F. Muzy (2013), Some limit theorems for Hawkes processes and application to financial statistics, *Stochastic Processes and their Applications*, 123(7): 2475–2499.
- [10] E. Bacry, I. Mastromatteo, J. F. Muzy (2015), Hawkes processes in finance, *Market Microstructure and Liquidity*, 1(1): 1550005.
- [11] V. Barbu (2016), *Differential Equations*, Springer International Publishing.
- [12] M. L. Bech, A. Malkhozov (2016), How have central banks implemented negative policy rates?, *BIS Quarterly Review March*.

-
- [13] F. Black (1995), Interest rates as options, *The Journal of Finance*, 50(5), 1371–1376.
- [14] C. G. Bowsher (2007), Modelling security market events in continuous time: Intensity based, multivariate point process models, *Journal of Econometrics*, 141(2): 876–912.
- [15] P. Brémaud (1981), *Point Processes and Queues: Martingale Dynamics*, Springer, New York.
- [16] D. Brigo, A. Alfonsi (2005), Credit default swap calibration and derivatives pricing with the SSRD stochastic intensity model, *Finance and stochastics*, 9(1), 29–42.
- [17] D. Brigo, F. Mercurio (2001), A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models, *Finance and Stochastics*, 5(3): 369–387.
- [18] M. Broadie, Ö Kaya (2006), Exact simulation of stochastic volatility and other affine jump diffusion processes, *Operations research*, 54(2), 217–231.
- [19] F. Caravenna (2011), *Moto browniano e analisi stocastica* (available at <http://staff.matapp.unimib.it/fcaraven/download/other/dispense-3.3.pdf>).
- [20] Y. Chen (2016), Thinning Algorithms for Simulating Point Processes.
- [21] Y. Chen (2016), Multivariate Hawkes Processes and Their Simulations.
- [22] G. Deelstra, F. Delbaen (1998), Convergence of discretized stochastic (interest rate) processes with stochastic drift term, *Applied stochastic models and data analysis*, 14(1), 77–84.
- [23] D. Duffie, D. Filipović, W. Schachermayer (2003), Affine processes and applications in finance, *Annals of Applied Probability*, 13(3): 984–1053.
- [24] D. Duffie, J. Pan, K. Singleton (2000), Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, 68(6): 1343–1376.
- [25] E. Eberlein, J. Kallsen (2019), *Mathematical Finance*, Springer International Publishing.
- [26] ECB (2014), Euro area risk-free interest rates: Measurement issues, recent developments and relevance to monetary policy, *ECB Monthly Bulletin*, 2014(07): 63–77.
- [27] E. Errais, K. Giesecke, L. R. Goldberg (2010), Affine point processes and portfolio credit risk, *SIAM Journal on Financial Mathematics*, 1(1): 642–665.
- [28] D. Filipović (2009), *Term Structure Models: A Graduate Course*, Springer, Berlin - Heidelberg.

- [29] C. Fontana, J. M. A. Montes (2014). A unified approach to pricing and risk management of equity and credit risk, *Journal of Computational and Applied Mathematics*, 259(B): 350–361.
- [30] X. Gao, L. Zhu (2018), Functional central limit theorems for stationary Hawkes processes and application to infinite-server queues, *Queueing Systems*, 90(1-2): 161–206.
- [31] F. Geiger, F. Schupp (2018), With a little help from my friends: Survey-based derivation of euro area short rate expectations at the effective lower bound.
- [32] P. Glasserman, K. K. Kim (2010), Moment explosions and stationary distributions in affine diffusion models, *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 20(1), 1–33.
- [33] G. A. Grigoryan (2017), Global solvability tests for a scalar Riccati equation with complex coefficients, *Differential equations*, 53(4), 450–456.
- [34] P. S. Hagan, D. Kumar, A. S. Lesniewski, D. E. Woodward (2002), Managing smile risk, *The Best of Wilmott*, 1, 249–296.
- [35] A. G. Hawkes (1971), Point spectra of some mutually exciting point processes, *Journal of the Royal Statistical Society: Series B (Methodological)*, 33(3): 438–443.
- [36] A. G. Hawkes, D. Oakes (1974), A cluster process representation of a self-exciting process, *Journal of Applied Probability*, 493–503.
- [37] O. Ibe (2014), *Fundamentals of applied probability and random processes*, Academic Press.
- [38] J. Jacod, A. Shiryaev (2013), *Limit theorems for stochastic processes* (Vol. 288), Springer Science Business Media.
- [39] R. A. Jarrow (2013), *The zero-lower bound on interest rates: Myth or reality?*, *Finance Research Letters*, 10(4): 151–156.
- [40] M. Jeanblanc, M. Yor, M. Chesney (2009), *Mathematical methods for financial markets*, Springer Science Business Media.
- [41] R. P. Jena, K. K. Kim, H. Xing (2012), Long-term and blow-up behaviors of exponential moments in multi-dimensional affine diffusions, *Stochastic Processes and their Applications*, 122(8), 2961–2993.
- [42] Y. Jiao, C. Ma, S. Scotti (2017), Alpha–CIR model with branching processes in sovereign interest rate modeling, *Finance and Stochastics*, 21(3), 789–813.
- [43] A. Karr (1991), *Point Processes and Their Statistical Inference*, CRC press, second edition.

- [44] M. Keller-Ressel, E. Mayerhofer (2015), Exponential moments of affine processes, *The Annals of Applied Probability*, 25(2), 714–752.
- [45] A. C. King, J. Billingham, S. R. Otto (2003), *Differential equations: linear, nonlinear, ordinary, partial*, Cambridge University Press.
- [46] T. Kortela (2016), A shadow rate model with time-varying lower bound of interest rates, *Bank of Finland Research Discussion Paper*, (19).
- [47] P. J. Laub, T. Taimre, and P. K. Pollett (2015), Hawkes processes, Preprint (available at <https://arxiv.org/abs/1507.02822>).
- [48] W. Lemke, A. Vladu (2016), Below the zero lower bound: A shadow-rate term structure model for the euro area, Bundesbank Discussion Paper (available at <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1991.en.pdf>).
- [49] P. W. Lewis, G. S. Shedler (1979), Simulation of nonhomogeneous Poisson processes by thinning, *Naval research logistics quarterly*, 26(3), 403–413.
- [50] R. Lord, R. Koekoek, D. V. Dijk (2010), A comparison of biased simulation schemes for stochastic volatility models, *Quantitative Finance*, 10(2), 177–194.
- [51] A. Monfort, F. Pegoraro, J. P. Renne, G. Roussellet (2017), Staying at zero with affine processes: An application to term structure modelling. *Journal of Econometrics*, 201(2): 348–366.
- [52] A. Monfort, F. Pegoraro, J. P. Renne, D. Sabes (2019), Affine term structure models with stochastic lower bound: an application to Euro-area OIS rates, Presentation at the 12th Financial Risks International Forum, Paris, 18-19 March 2019.
- [53] Y. Ogata (1981), On Lewis’ simulation method for point processes, *IEEE transactions on information theory*, 27(1), 23–31.
- [54] A. Pascucci (2011), *PDE and martingale methods in option pricing*, Springer Science & Business Media.
- [55] P. E. Protter (2004), *Stochastic Integration and stochastic differential equations*, Springer, Berlin - Heidelberg, second edition, version 2.1.
- [56] V. K. Rohatgi, A. M. E. Saleh (2015), *An introduction to probability and statistics*, John Wiley & Sons.
- [57] S. I. T. U. Rong (2006), *Theory of stochastic differential equations with jumps and applications: mathematical and analytical techniques with applications to engineering*, Springer Science & Business Media.
- [58] E. J. van de Winkel (2019), *Parametric Hawkes Processes: Theory, Simulation and Application* (Bachelor’s thesis).

-
- [59] U. Vollmer, R. Bebenroth (2012), The financial crisis in Japan: causes and policy reactions by the Bank of Japan, *The European Journal of Comparative Economics*, 9(1): 51.
- [60] J. C. Wu, F. D. Xia (2016), Measuring the macroeconomic impact of monetary policy at the zero lower bound, *Journal of Money, Credit and Banking*, 48(2–3), 253–291.
- [61] J. C. Wu, F. D. Xia (2019), Negative Interest Rate Policy and the Yield Curve, *Journal of Applied Econometrics*.
- [62] J. C. Wu, F. D. Xia (2017), Time-varying lower bound of interest rates in Europe, *Chicago Booth Research Paper*, (17–06).
- [63] V. F. Zaitsev, A. D. Polyanin (2002), *Handbook of exact solutions for ordinary differential equations*, CRC press, second edition.