BERGISCHE UNIVERSITÄT WUPPERTAL

Dissertation

Ergodicity properties of affine term structure models and applications.

ausgeführt zum Zwecke der Erlangung des akademischen Grades

eines Doktors der Naturwissenschaften

eingereicht an der Bergische Universität Wuppertal,
Fakultät für Mathematik und Naturwissenschaften,
Angewandte Mathematik-Stochastik.

Von: **Chiraz Trabelsi**

unter der Leitung von

a.o. Univ. Prof. Barbara Rüdiger

Juni 2016

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20160624-084201-9
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20160624-084201-9]

Deutschsprachige Kurzfassung

In Dissertation werden die Ergodizitäts Eigenschaften der affinen Zinsstrukturmodelle sowie deren praktische Anwendung untersucht. Erstens betrachten wir die affine Zinsstrukturen des Cox-Ingersoll-Ross Modells (CIR-Modell genannt). Dieses Modell (1985) wurde von John C. Cox, Jonathan E. Ingersoll und Stephen A. Ross als ein alternatives Modell eingeführt, um den Nachteil des Vasicek-Modells, in dem der Zinssatz negativ werden kann zu überwinden. Wir zeigen die Harris's positiv-Rekurrenz des CIR-Prozesses, Daraus folgen Ergodizitätseigenschaften für einen Transformation der CIR Prozess. Dies wird in der Kalibrierung der Parameter eines Kredit Migration Modelles angewendet.

Ferner konzentrieren wir uns auf eine Verallgemeinerung des CIR-Modell, das durch Zugabe von Sprüngen erhalten wird, nämlich der Grund affine jump-Diffusion Modell (BAJD). Dieses Modell wurde von Duffie und Gârleanu als Erweiterung des CIR-Modell mit Sprüngen eingeführt. Wir leiten eine Formel für die Übergangsdichten des Prozesses. Beachten Sie, dass diese Tatsache bereits in einem speziellen Fall von Filipović entdeckt wurde [13]. Außerdem beweisen wir die Harris's positiv-Rekurrenz und die exponentielle Ergodizität des BAJD, und Kalibrieren einen Transformation davon.

Eine andere Erweiterung des klassischen CIR-Modelles mit lich Sprüngen sind die Sprung-Diffusion CIR-Verfahren (JCIR). Sie werden mit Hilfe eines reinen Sprung Lévy Prozess eingeführt . Wir finden eine untere Schranke für die Übergangsdichten und zeigen wir die Existenz eines Foster-Ljapunovfunktion, aus denen wir die exponentielle Ergodizität ableiten.

Schließlich untersuchen wir einige Eigenschaften von nicht affinen Zinsstrukturmodelle .

Abstract

The aim of this thesis is to study the ergodicity properties of affine term structure models as well as the practical applications. First, we consider the affine term structure model called the Cox- Ingersoll-Ross model (abbreviated CIR). This model was introduced in 1985 by John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross as an alternative model to overcome the disadvantage of Vasicek model, in which the interest rate can become negative. We show the positive Harris recurrence of the CIR process, from which we get an ergodicity results for a transformation of the CIR process. This is applied in the calibration of the parameters of a credit migration model.

Later we focus on an extension of the CIR model that is obtained by adding jumps, namely the basic affine jump-diffusion (BAJD). This model has been introduced by Duffie and Gârleanu as an extension of the CIR model with jumps. We derive a closed formula for the transition densities of the BAJD. Note that this fact has already been discovered in a special case by Filipović [13]. Further, we prove the positive Harris recurrence and the exponential ergodicity of the BAJD, and calibrate the transformation of it.

Another extension of the classical CIR model including jumps is the jump-diffusion CIR process (shorted as JCIR). This is introduced with the help of a pure-jump Lévy process. We find a lower bound on the transition densities and we show the existence of a Foster-Lyapunov function from which we derive the exponential ergodicity.

Finally, we investigate some properties of non affine term structure models.

Acknowledgement

First of all, I would like to thank my supervisor Professor Barbara Rüdiger, who guided my progresses with her excellent support during my doctoral studies. I am grateful to her for the proposal of an interesting subject within the field of ergodic theory for affine term structure models. I thank her, for her numerous advice and suggestions which considerably improved this thesis. I am extremely thankful to her contribution to my scientific education as well as others aspects of my professional development.

I want to express my deeply-felt gratitude to my Professor Habib Ouerdiane, who is the head of Laboratory of Stochastic Analysis and Applications to Finance at University of Tunis ElManar, Tunisia, for introducing me into the world of stochastic analysis through his Master's courses and his very active research team. He has given me the opportunity to get in contact with various scientists from different countries.

Special thanks are also due to Dr. Peng Jin, I have found working with him really stimulating and rewarding, the discussions with him have been very enriching. He was always full of ideas, making connections between different subjects with startling creativity, providing me with high motivation.

I am also thankful to Professor Vidyadhar Mandrekar, from Michigan State University USA, for providing me with helpful remarks and interesting discussions. I had a fortune to write and publish my first article in collaboration with him.

I appreciate the financial support of University of Wuppertal. It is been a pleasure and an honor to study here, and I have learnt a lot since I came here. My visit has resulted in fruitful scientific collaboration, which I hope will continue in the future.

The referees Professor Padmanabhan Sundar, Professor Barbara Rüdiger and Professor Vidyadhar Mandrekar offered their precious time to read this thesis and write the reports.

I want to thank Professor Balint Farkas for agreeing to be in the Ph.D. committee for the evaluation of my thesis.

The financial support of this work was provided by the DAAD grant, Faculty of Sciences of Tunis, the Ministry of Higher Education of Tunisia, University of Wuppertal and the insurance company DeBeKa in Koblenz, to which I am thankful.

Special thanks are also due to my colleagues for the good times at FST and at BUW. Thanks for the nice work environment and the interesting discussions. Moreover, I would like to thank my friends Dr. Hafida Laasri and Barun Sarkar for the pleasant time they provided me in Wuppertal. I am pleased to thanks Mrs. Kerstin Scheibler for her help.

I would like to take this opportunity to thank my mother Latifa, my father Fathi and my brothers Mohamed and Ghaith, for their patience, confidence and encouragement over the years leading up to this work. I would like to thank all my family Trabelsi, Ksaier and Hemdani.

I want also to thank my husband Elyes, who in the last years has become a part of my life, thanks for his love, support and kindness.

Finally, I want to thank my son Mohamed Wissem for bringing so much joy to my life. This thesis is dedicated to him.

Contents

1 Short term Interest rate models **15**
1.1 Cox-Ingersoll-Ross model 16
1.1.1 Mean reversion of the CIR model 17
1.1.2 Transition density function of the CIR process 18
1.2 Basic affine jump diffusion process 20
1.3 Jump-diffusion CIR process 22

2 Ergodic results on transformation of the CIR and application **24**
2.1 Affine and regularity properties 24
2.1.1 Affine property 24
2.1.2 Regularity property 28
2.2 Positive Harris recurrence 32
2.3 Ergodicity results 34
2.4 Application in one credit migration model 38

3 Positive Harris recurrence, exponential ergodicity and calibration of the BAJD **44**
3.1 Characteristic function of the BAJD 44
3.2 Mixtures of Bessel distributions 47
3.3 Transition density of the BAJD 50
3.3.1 Case i): $\Delta > 0$ 51
3.3.2 Case ii): $\Delta < 0$ 53
3.3.3 Case iii): $\Delta = 0$ 53
3.4 Positive Harris recurrence of the BAJD 55
3.5 Exponential ergodicity of the BAJD 61
3.6 Calibration for the BAJD-process 67

4 Exponential Ergodicity of the Jump-Diffusion CIR Process **71**
4.1 Characteristic function of the JCIR 72

4.1.1 Special Case i): $\nu = 0$, No Jumps 74
4.1.2 Special Case ii): $\theta = 0$ and $x = 0$ 75
4.2 Lower bound for the transition densities of JCIR 75
4.3 Exponential ergodicity of JCIR 81

5 Non affine term structure models **86**
5.1 Connection between OU-process and CIR-process in the jumps case . . . 86
5.2 Some investigations 91

Introduction

This thesis is devoted to study the Harris recurrence and the ergodicity of affine term structure models. The term structure models have been the focus of many studies over a century now. They can mainly be put in three categories: short rate models, forward rate models and market models. In this thesis we consider only the short rate models. In the literature of financial mathematics, the short rate models were the first studied dynamic term structure models. These mathematical models describe the future evolution of interest rates by describing the future evolution of short rates. Interest rate modeling has gained special attention during the last few decades which has resulted in reliable models. In order to understand better the evolution of interest rates, researchers have attempted to identify processes and rational investor's behaviors. The short rates are typically described as a diffusion process. The diffusion models have become one of the core areas in statistical sciences and financial modeling.

An often referenced short rate model is the celebrated Cox-Ingersoll-Ross model (CIR). It was introduced in 1985 by John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross (see [7]) and is one of an interesting process which became quite popular in finance. This model was done to illustrate the workings of a general equilibrium model and was proposed as an extension of the Vasicek model (see [51]). The bad property of possible negativity in the Vasicek model is removed in the Cox-Ingersoll-Ross model under the so-called Feller condition and hence ensuring that the origin is inaccessible to the process. The nice mean reversion property in the Vasicek model is preserved in the Cox-Ingersoll-Ross model. Mean reversion means that prices and returns eventually move back towards the mean or the average. In finance, mean reversion is the assumption that a stock's price will tend to move to the average price over time. After that Cox et al. proposed the CIR process for modeling short term interest rates, it is also used in the valuation of interest rate derivatives and for modeling stochastic volatility in the Heston model [21]. The popularity of the CIR process in all main branches of financial modeling stems from its desirable property of positivity, its richness of behaviour and its mathematical properties. In the literature the CIR is also known as the square root diffusion or Feller process.

The financial crisis of $2008 - 2009$ has once again made it clear that the extreme behavior of financial assets cannot be described using only the traditional models based on Gaussian processes but also we have to consider the Lévy processes. Therefore we are interested by another type of short term interest rate model which is an extension of the CIR model including positive jumps. The instantaneous interest rate is modeled as a mixture of CIR process and a compound Poisson process. This model is called basic affine jump diffusion process (BAJD) and in which the Lévy process takes the form of a compound Poisson process with exponentially distributed jumps. The BAJD was introduced by Duffie and Gârleanu [10] to describe the dynamics of default intensity. It was also used by Filipović and Keller-Ressel et al. as a short rate model. Motivated by some applications in finance, the long-time behavior of the BAJD has been well studied. Keller-Ressel et al. proved that the BAJD possesses a unique invariant probability measure. The existence and the approximations of the transition densities of the BAJD can be found in [14].

A more general extension of the CIR model including jumps is the so-called Jump-diffusion CIR process (abbreviated as JCIR). In this model the jumps are introduced with the help of a pure-jump Lévy process. The BAJD and JCIR models belong to the class of affine processes. A complete characterization of the class of regular affine processes was given in [9]. During the last decades affine processes have became very popular due to their tractability and their flexibility, since often there are explicit solution for bond prices.

Many other aspects of affine term structure models are still under current investigation, mainly their long-term behavior and ergodicity properties. Initially, the ergodic theory was introduced in statistical mechanics by Boltzmann. The word ergodic is a mixture of two Greek words: "ergon" (work) and "odos" (path). It is not easy to give a simple definition of ergodic theory because it uses techniques from many fields, mainly is the study of the long-term average behavior of systems evolving in time.

Our main focus is to study the ergodicity properties of affine term structure models such as CIR, BAJD and JCIR processes, as well as the positive Harris recurrence for the CIR and BAJD . We investigate some properties of non-affine term structure models in this work.

To attain our major objective, we give a brief outline of how we intend to proceed and what each chapter contains:

The first chapter covers the basic definitions and properties of short term interest rate models: CIR model and its extensions, which are the BAJD and JCIR models.

In the second Chapter, we managed to show that the CIR process is positive Harris recurrent. Harris recurrence was first introduced by Harris [20] for discrete Markov chains and then was extended in [1] to a general continuous time Markov process. The applications of Harris recurrence have been found in queueing theory and stochastic control. A recent application for interest rate models was given in [2], where Harris recurrence was used as a principal assumption to enable the authors to prove consistency of some estimators of jump-diffusion models for interest rate. In the first Section 2.1, we give the affine and regularity properties of the CIR process. The main results are stated in Section 2.2, we show that the CIR process is positive Harris recurrent. Then in Section 2.3, we establish the ergodicity results on the transformation of the CIR process. In the last Section 2.4, we take advantage of this study to apply the ergodicity results in the calibration of one credit migration model. We should remark that the results presented in the last section have been derived in [39] with a different method. Most of the results of Section 2.2-2.4 are taken from our paper [26]. This paper is coauthored with Peng Jin, Vidyadhar Mandrekar and Barbara Rüdiger.

Recently, the long-term behavior of affine processes with the state space $\mathbb{R}_+$ has been studied by in [38] (see also [35]), motivated by some financial applications in affine term structure models of interest rates. In particular, they have found some sufficient conditions such that the affine process converges weakly to a limit distribution. This limit distribution was later shown by Keller-Ressel in [32] as the unique invariant probability measure of the process. Under further sharper assumptions it was even shown in [42] that the convergence of the law of the process to its invariant probability measure under the total variation norm is exponentially fast, which is called the exponential ergodicity in the literature. The method used in [42] to show the exponential ergodicity is based on some coupling techniques.

In the third Chapter, we investigate the long-time behavior of the BAJD model. More precisely, we show that the BAJD process is also positive Harris recurrent. As a well-known fact, Harris recurrence implies the existence of (up to the multiplication by a positive constant) unique invariant measure. Therefore, our result on the positive Harris recurrence of the BAJD provides another way of proving the existence and uniqueness of invariant measures for the BAJD. Another consequence of the positive Harris recurrence is the limit theorem for additive functional (see e.g. [29, Theorem20.21]). Some applications of Harris recurrence in statistics and calibrations of some financial models can be found in [2] and [26].

We give some preliminaries on the BAJD process in the Section 3.1. Next, we introduce the so-called Bessel distributions and some mixtures of Bessel-distributions in Section 3.2. In Section 3.3, we derive the transition densities of the BAJD. This formula indicates that the law of the BAJD process at any time is a convolution of a mixture of Gamma-distributions with a noncentral chi-square distribution. We should point out that this fact has already been discovered by Filipović [13] for a special case. In the main results of Section 3.4, we were successful in showing that the BAJD is positive Harris recurrent.

The second main result is described in Section 3.5, namely we show the exponential ergodicity of the BAJD. We should indicate that the BAJD does not satisfy the assumptions required in [42] in order to get the exponential ergodicity. Our method is different and is based on the existence of a Foster-Lyapunov function. In the last Section 3.6, we apply these results to show another consequence of Harris recurrence in the calibration of the BAJD. This calibration result motivated by one discussion with the member of the research group in DeBeKa insurance company in Koblenz (19 July 2012).

Most of the results of Section 3.2-3.5 are taken from the joint paper with Peng Jin and Barbara Rüdiger [27].

In the fourth Chapter, we compute explicitly the characteristic function of the JCIR in Section 4.1. Moreover, this enables us to represent the distribution of the JCIR as the convolution of two distributions. The first distribution coincides with the distribution of the CIR model. However, the second distribution is more complicated. We give a sufficient condition such that the second distribution is singular at zero. In this way we derive a lower bound estimate of the transition densities of the JCIR in Section 4.2. The problem that we consider in Section 4.3, is the exponential ergodicity of the JCIR. Namely, we show the existence of a Foster-Lyapunov function and then apply the general framework of Meyn and Tweedie. Most of the results of Section 4.2 and 4.3 are taken from the proceeding [28]. This proceeding is coauthored with Peng Jin and Barbara Rüdiger.

In the last Chapter, we show that CIR model driven by Lévy noise is equal to the square of Ornstein-Uhlenbeck process with jumps and a positive drift in Section 5.1.Unfortunally this model is non-affine term structures models, we will give some investigations on the non-affine term structure models in the last Section 5.2.

Chapter 1

Short term Interest rate models

Interest rates are of fundamental importance in the economy in general and in financial markets in particular. The movements of interest rate plays an important role in decision of investment and risk management in financial markets. One-factor models are a popular class of interest rate models which are used for these purposes, especially in the pricing of interest rate derivatives. In the literature one can find several references. The well known Birgo and Mercurio [6] and Lamberton and Lapeyre [41] are the most complete references about interest rate models for both theoretical and practical aspects. Interest rate modeling has gained special attention during the last few decades which has resulted in reliable models. An empirical observations suggest that the dynamics of the interest rate should be modelled by a stochastic process, since they are heavily varying over time.

First, we postulate the following general process for the short-term interest rate. Short rate models use the instantaneous spot rate $X(t)$ as the basic state variable. The stochastic differential equation describing the dynamics of $X(t)$ is usually stated under the spot measure. We further assume that the interest rate process is Markovian and its dynamic is described by the following first-order stochastic differential equation:

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW_t \tag{1.1}$$

where μ and σ are suitably chosen drift and diffusion coefficients respectively, and W is the standard Brownian motion driving the process. These models are referred as one-factor models, as there is only one stochastic drivers. The models with multiple stochastic drivers are called multi-factor models. The drift and diffusion coefficients need to satisfy some regularity requirements to guarantee the existence of unique solution of the SDE (1.1). The regularity requirements make sure that the solution does not explode (growth conditions) and its unique (Lipschitz conditions). These solutions are called strong solutions, which means that any other Itô process that solves (1.1) is equal to X almost everywhere. Various choices of the coefficients μ and σ lead to different dynamics of the instantaneous rate. We shall focus on the Cox-Ingersoll-Ross model and its extension including jumps. Throughout this thesis, we assume that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$ will always be a filtered probability space satisfying the usual conditions, i.e.,

1. $(\Omega, \mathcal{F}_t, P)$ is complete for all $t \in \mathbb{R}_+$, $\mathcal{F}_0$ contains all the P-null sets in $\mathcal{F}$ for all $t \in \mathbb{R}_+$.
2. $\mathcal{F}_t = \mathcal{F}_{t+}$ where $\mathcal{F}_{t+} = \cap_{s>t}\mathcal{F}_s$, for all $t \geq 0$, i.e. the filtration is right-continuous.

In this section, we will focus on the celebrated Cox-Ingersoll-Ross short rate model.

1.1 Cox-Ingersoll-Ross model

We consider the well-known Cox, Ingersoll and Ross model (shorted as CIR model), this model is a diffusion process suitable for modeling the term structure of interest rates. It was introduced in 1985 by John C. Cox, Jonathan A. Ingersoll and Stephen A. Ross [7] as an extension of the Vasicek model [51]. They suggest modelling the behavior of instantaneous interest rate by the following stochastic differential equation

$$dX_t = a(\theta - X_t)dt + \sigma\sqrt{|X_t|}dW_t, \quad X_0 \geq 0, \tag{1.2}$$

where $a, \sigma > 0$ and $\theta \geq 0$ are constants and W_t is a one-dimensional standard Brownian motion. The CIR model is one of the standard "short rate" model in financial mathematics.

Now, we recall some well-known properties of the solution of the CIR model. Note that we can not apply the theorem of existence and uniqueness for the SDE because the volatility term $\sigma\sqrt{x}$ does not satisfy the Lipschitz condition. However, from the Hölder property of the square root function, by a theorem due to Yamade and Watanabe (see [30, Proposition 5.2.13]), the strong uniqueness holds for the above SDE (1.2). Ikeda and Watanabe prove that there is a (pathwise) unique non-negative strong solution $(X_t, t \geq 0)$ of (1.2) with only positive initial value X_0 (see [23, Example 8.2], [25, Theorem 1.5.5.1]).

If $\theta = 0$ and $X_0 = 0$, the solution of the SDE (1.2) is $X_t \equiv 0$, and from the comparison theorem for one-dimensional diffusion processes ([25, Theorem 1.5.5.9]), it follows that $X_t \geq 0$ if $X_0 \geq 0$. In that case, we omit the absolute value and we consider the positive solution of the following SDE

$$\begin{cases} dX_t = a(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad t \geq 0 \\ X_0 = x \geq 0, \end{cases} \tag{1.3}$$

This solution is often called a Cox-Ingersoll-Ross (CIR) process or a square-root process see [11].
By applying Itô formula for the process $(X_t, t \geq 0)$ one can get for all $t \geq 0$

$$\begin{aligned} d(e^{at}X_t) &= ae^{at}X_t dt + e^{at}dX_t \\ &= a\theta e^{at}dt + \sigma e^{at}\sqrt{|X_t|}dW_t. \end{aligned}$$

Then we have

$$X_t = e^{-at}\big(x + a\theta \int_0^t e^{as}ds + \sigma \int_0^t e^{as}\sqrt{|X_s|}dW_s\big)$$

by taking expectation on both sides

$$E(X_t) = e^{-at}E(x) + a\theta \int_0^t e^{-a(t-s)}ds$$

Hence

$$\lim_{t \to +\infty} E(X_t) = E(X_\infty) = \theta,$$

then θ is called the long-term value.

1.1.1 Mean reversion of the CIR model

The most important feature which this model exhibits is the mean reversion properties, which means that if the interest rate is bigger than the long-term mean $(X > \theta)$, then the coefficient $a > 0$ makes the drift become negative so that the rate will be pulled down in the direction of θ. Similarly, if the interest rate is smaller than the long-term mean $X < \theta$, then the coefficient $a > 0$ makes the drift term become positive so that the rate will be pulled up in the direction of θ. Therefore the parameter a is called the speed of mean reversion, it gives the speed of adjustment and has to be positive in order to maintain stability around the long-term value θ. The parameter σ is the volatility coefficient.
Now, we give some properties of the CIR process. We denote $(X_t^x, t \geq 0)$ the CIR process started from an initial point x and τ_0^x the stopping time, is the first time when the process hit 0, and defined by

$$\tau_0^x = \inf\{t \geq 0; X_t^x = 0\}$$

with, as usual, $\inf \emptyset = \infty$.

Proposition 1.

1. If $2a\theta \geq \sigma^2$, we have $P(\tau_0^x = \infty) = 1,$ for all $x > 0$.

2. If $2a\theta < \sigma^2$, we have $P(\tau_0^x < \infty) = 1,$ for all $x > 0$.

For the proof of this proposition we refer to (Exercise 34 page 137, [41]). The above proposition give us that an examination of the boundary classification criteria shows that the rate can reach zero if $\sigma^2 > 2a\theta$. In this case zero is accessible but is not absorbing as explained intuitively below see proof ([41, Exercice 34]). If $2a\theta \geq \sigma^2$, the upward drift is sufficiently large to make the origin inaccessible, this condition is called the Feller condition [11]. In other words, the condition $2a\theta \geq \sigma^2$ makes sure that zero is never reached, so that we can grant that $X(t)$ remains always positive. In either case, the singularity of the diffusion coefficient at the origin implies that an initially non-negative interest rate can never subsequently become negative.

Intuitively, when the interest rate is at a low level (approaches zero), the volatility term $\sigma\sqrt{x}$ also becomes close to zero (cancelling the effect of the randomness). Consequently, when the rate gets close to zero, its evolution becomes dominated by the drift factor, which pushes the rate upwards (towards equilibrium). When the interest rate is high then the volatility is high and this is a desired property.

1.1.2 Transition density function of the CIR process

The SDE (1.3) has no general, explicit solution, although its transition density function can be characterized. The transition density of the CIR process is first found in [11] by Laplace transform methods. Duffie et al [9] exploited the affine structure of the CIR process to identify the Fourier transform of the law of the $(X_t, t \geq 0)$. A more probabilistic method to get the transition density was mentioned in Yor et al [17].
In this section we briefly explain how to compute the transition density function of the CIR process via squared Bessel processes this method used in [17, 25]. For full details the readers are referred to [17, 25, 49].

Definition 1. *For every $\delta \geq 0$ and $x \geq 0$ the unique strong solution to the equation*

$$R_t = x + \delta t + 2\int_0^t \sqrt{R_s} dB_s \tag{1.4}$$

is called the square of a δ-dimensional Bessel process started at x and is denoted by $BESQ_x^\delta$.

Remark 1. The number δ is called the *dimension* of BESQ$^\delta$, since a BESQ$^\delta$ process R_t can be represented by the square of the Euclidean norm of δ-dimensional Brownian motion B_t if $\delta \in \mathbb{N}$: $R_t = |B_t|^2$.

Definition 2. *The square root of BESQ$^\delta$, $\delta \geq 0, y = \sqrt{x} \geq 0$ is called the Bessel process of dimension δ started at y and is denoted by BES$^\delta$.*

The CIR process (1.3) can be represented as a time-changed squared Bessel process

$$X_t = e^{-at} R\bigg(\frac{\sigma^2}{4a}\big(e^{at} - 1\big)\bigg),$$

where R is a squared Bessel process with dimension $\delta = \frac{4a\theta}{\sigma^2}$ started at x. This relation is used by Delbaen and Shirakawa [8] and Szatzschneider [50]. For $\delta > 0$, the transition density for BESQ$^\delta$ is equal to

$$q_t^\delta(x,y) = \frac{1}{2t}\Big(\frac{y}{x}\Big)^{\frac{\nu}{2}} \exp\bigg\{-\frac{x+y}{2t}\bigg\} I_\nu\bigg(\frac{\sqrt{xy}}{t}\bigg),$$

where $t > 0, x > 0, \nu \equiv \frac{\delta}{2} - 1$ and I_ν is the modified Bessel function of the first kind of index ν, see e.g. [49]. Using the transition density of the squared Bessel, it is easy to obtain the transition density of CIR process

$$p(t,x,y) = \rho e^{-u-v}\Big(\frac{v}{u}\Big)^{\frac{q}{2}} I_q\big(2(uv)^{\frac{1}{2}}\big) \tag{1.5}$$

for $t > 0, x > 0$ and $y \geq 0$, where

$$\rho \equiv \frac{2a}{\sigma^2\Big(1 - e^{-at}\Big)}, \qquad u \equiv \rho x e^{-at},$$

$$v \equiv \rho y, \qquad q \equiv \frac{2a\theta}{\sigma^2} - 1,$$

and $I_q(\cdot)$ is the first-order modified Bessel function with index q, defined by:

$$I_q(x) = \Big(\frac{x}{2}\Big)^q \sum_{k=0}^{\infty} \frac{\big(\frac{x}{2}\big)^{2k}}{k!\Gamma(q+k+1)}$$

We should remark that for $x = 0$ the formula of the density function $p(t,x,y)$ given in (1.5) is no more valid. In this case the density function is given by

$$p(t,0,y) = \frac{\rho}{\Gamma(q+1)} v^q e^{-v} \tag{1.6}$$

for $t > 0$ and $y \geq 0$.
The distribution density is the non-central chi-square, $\chi^2[2v; 2q+2, 2u]$, with $2q+2$ degrees of freedom, and non-centrality parameter $2u$.

The conditional expectation and conditional variance of the short rate $(X_t, t \geq 0)$ can be calculated explicitly which can be useful for calibrating the parameters, for $s > t$ we get

$$\begin{aligned} E[X_s|X_t] &= X_t e^{-a(s-t)} + \theta\big(1 - e^{-a(s-t)}\big) \\ Var[X_s|X_t] &= X_t(\frac{\sigma^2}{a})\big(e^{-a(s-t)} - e^{-2a(s-t)}\big) + \theta\big(\frac{\sigma^2}{a}\big)(1 - e^{-a(s-t)}\big)^2 \end{aligned}$$

The properties of the distribution of the future interest rates are those expected values. As a approaches infinity, the mean goes to θ and the variance to zero, and when a approaches zero, the conditional mean goes to the current interest rate and the variance to $\sigma^2 X(t)(s-t)$. Another important issue concerning the CIR process is its long-term behavior. In the original paper [7] by John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross, the steady state of the CIR model is shown to be a Gamma distribution. In other words, as $t \to \infty$, $X_t \to X_\infty$ where X_∞ follows a Gamma distribution with shape parameter $\frac{2a\theta}{\sigma^2}$ and scale parameter $\frac{\sigma^2}{2a}$. The steady state mean and variance are θ and $\frac{\sigma^2\theta}{2a}$, respectively. The corresponding Gamma distribution is an invariant measure for the CIR process. It is also well known that this invariant measure is ergodic (see [3]). [4] investigated the recurrent properties of the CIR process and proved that $\big[0, \frac{\sqrt{(2\sigma^2+4a\theta)(3a\theta+\sigma^2)}}{2a} + 1\big]$ is a recurrent region for the CIR process. As well known, the CIR process is an affine process in $\mathbb{R}_+$, the laplace transform of the value of the process at time t is the exponential of an affine function of its initial value. General affine processes and their applications in finance have been investigated in great detail in [9]. Among other things, it is proved in [9] that any stochastically continuous affine process is a Feller process. In particular, the CIR process is a Feller process in $\mathbb{R}_+$.

1.2 Basic affine jump diffusion process

The CIR model captures many features of the real world interest rates. In particular, the interest rate in the CIR model is non-negative and mean-reverting. Because of its vast applications in mathematical finance, some extensions of the CIR model have been introduced and studied, see e.g. [10, 13, 42].

Here, we propose to analyze a stochastic process, which is a basic affine jump diffusion (shorted as BAJD). It can be seen as a generalization of the classical Cox-Ingersoll-Ross process including jumps. The BAJD process is given as the unique strong solution $X := (X_t)_{t\geq 0}$ to the following stochastic differential equation

$$dX_t = a(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t + dJ_t, \quad X_0 \geq 0, \tag{1.7}$$

where a, θ, σ are positive constants, $(W_t)_{t\geq 0}$ is a one-dimensional Brownian motion and $(J_t)_{t\geq 0}$ is a compound Poisson process, i.e.

$$J(t) = \sum_{i=1}^{N(t)} Y_i,$$

where N_t is a Poisson process with constant jump intensity c, the $(Y_i)_{i\in\mathbb{N}}$ are independent and exponentially distributed with parameter d, which are also independent of $(N_t)_{t\geq 0}$.

- In this model only positive jumps are allowed.
- The jump size and inter-arrival times are exponentially distributed with parameter d and c.
- $(J_t)_{t\geq 0}$ is a pure-jump Lévy process with Lévy measure

  $$\nu(dy) = \begin{cases} cde^{-dy}dy, & y \geq 0, \\ 0, & y < 0, \end{cases}$$

  for some constants $c > 0$ and $d > 0$.

We assume that all the above processes are defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$.

The BAJD process $X = (X_t)_{t\geq 0}$ given by (1.7), has been introduced in 2001 by Duffie and Gârleanu [10] and is attractive for modelling default times τ in credit risk applications, i.e.

$$\mathbb{P}(\tau > t + s|\mathcal{F}_t) = E\Big[\exp\Big(\int_t^{t+s} -X_u du\Big)|\mathcal{F}_t\Big],$$

since both the moment generating function and the characteristic function are known in closed form. Notice that the BAJD process can be seen as a special case of CBI-process (Continuous-state Branching process with Immigration), if we choose the parameters according to [13, Thm. 5.3], as follows

$$\alpha = \frac{1}{2}\sigma^2, \quad b = a\theta, \quad \beta = -a,$$
$$m(dy) := \nu(dy) = cde^{-dy}dy, \qquad \mu = 0,$$

for some constants $c > 0$ and $d > 0$. A special case is the no-jump, i.e. $c = 0$, just yields the classical CIR model.

It was also used in [13] and [38] as a short-rate model. Due to its simple structure, it is later referred as the basic affine jump-diffusion. The existence and uniqueness of strong

solutions to the SDE (1.7) follow from the main results of [15]. At the same time, the BAJD process $X = (X_t)_{t\geq 0}$ in (1.7) stays non-negative due to vanishing volatility and positive drift near the origin. This fact can be shown rigorously with the help of comparison theorems for SDEs, for more details we refer the reader to [15].
If the coefficient of the linear term in the drift is negative, and the constant term is positive, then BAJD process is mean-reverting, which is an important empirical feature observed in credit markets.

As its name implies, the BAJD belongs to the class of affine processes. Roughly speaking, affine processes are Markov processes for which the logarithm of the characteristic function of the process is affine with respect to the initial state. Affine processes on the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$ have been thoroughly investigated by Duffie et al [9], as well as in [36]. In particular, it was shown in [9] (see also [36]) that any stochastic continuous affine process on $\mathbb{R}_+^m \times \mathbb{R}^n$ is a Feller process and a complete characterization of its generator has been derived. Results on affine processes with the state space $\mathbb{R}_+$ can also be found in [13]. Affine processes have found vast applications in mathematical finance, because of their complexity and computational tractability. As mentioned in [9], these applications include the affine term structure models of interest rates, affine stochastic volatility models, and many others.

1.3 Jump-diffusion CIR process

A more general extension of the classical CIR model including jumps is the so-called jump-diffusion CIR process (shorted as JCIR). The jumps of the JCIR are introduced with the help of a pure-jump Lévy process. The JCIR process is defined as the unique strong solution $X := (X_t, t \geq 0)$ to the following stochastic differential equation

$$dX_t = a(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t + dJ_t, \quad X_0 \geq 0, \tag{1.8}$$

where $a, \sigma > 0$, $\theta \geq 0$ are constants, $(W_t, t \geq 0)$ is a one-dimensional Brownian motion and $(J_t, t \geq 0)$ is a pure-jump Lévy process with its Lévy measure ν concentrated on $(0, \infty)$ and satisfying

$$\int_{(0,\infty)} (\xi \wedge 1)\nu(d\xi) < \infty, \tag{1.9}$$

independent of the Brownian motion $(W_t, t \geq 0)$.

The initial value X_0 is assumed to be independent of $(W_t, t \geq 0)$ and $(J_t, t \geq 0)$. We assume that all the above processes are defined on some filtered probability space

$(\Omega, \mathcal{F}, (\mathcal{F})_{t\geq 0}, P)$. We remark that if we choose the parameters according to [13, Thm. 5.3], as follows

$$\alpha = \frac{1}{2}\sigma^2, \quad b = a\theta, \quad \beta = -a,$$
$$m(dy) := \nu(d\xi), \qquad \mu = 0,$$

then we get the JCIR process as a special case of CBI-process (see [13]).

The existence and uniqueness of strong solutions to (1.8) are guaranteed by [15, Thm. 5.1].

The JCIR process preserves the mean-reverting and non-negative properties of the classical CIR process (1.3), more precisely the term $a(\theta - X_t)$ in (1.8) defines a mean reverting drift pulling the process towards its long-term value θ with a speed of adjustment equal to a. Since the diffusion coefficient in the SDE (1.8) degenerate at 0 and only positive jumps are allowed, the JCIR process $(X_t, t \geq 0)$ stays non-negative if $X_0 \geq 0$. This fact can be shown rigorously with the help of comparison theorems for SDEs, for more details we refer the readers to [15].

Clearly,the JCIR defined in (1.8) includes the classical CIR as well as the basic affine jump-diffusion (or BAJD) as a special case, in which the Lévy process $(J_t, t \geq 0)$ takes the form of a compound Poisson process with exponentially distributed jumps. The BAJD was introduced by Duffie and Gârleanu [10] to describe the dynamics of default intensity.

Chapter 2

Ergodic results on transformation of the CIR and application

As a main result of this chapter, we will prove the positive Harris recurrence of the classical CIR process. Ergodic results on transformations of the CIR process will be given. We will also show that if $g : \mathbb{R}_+ \to \mathbb{R}$ is continuous, $f : \mathbb{R} \to \mathbb{R}_+$ is measurable and $(X_t, t \geq 0)$ is the CIR process, then $\frac{1}{N}\sum_{j=0}^{N-1} f\big(\int_j^{j+1} g\big(X_s\big)ds\big)$ converges almost surely to a constant. An application of the ergodic results in one credit migration model will be presented too.

2.1 Affine and regularity properties

Recall that the CIR process $(X_t, t \geq 0)$ is given as the unique strong solution of the following stochastic differential equation

$$dX_t = a(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad X_0 = x \geq 0, \tag{2.1}$$

where $a, \theta, \sigma > 0$ are constants and $(W_t, t \geq 0)$ is a one-dimensional Brownian motion defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ with $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions.

In this section we give the affine and regularity properties of CIR model.

2.1.1 Affine property

The process X given by (2.1) is a special affine process. The set of affine processes contains a large class of important Markov processes such as continuous state branching processes and Ornstein-Uhlenbeck processes. Further, a lot of models in financial mathematics are affine such as the Vasicek, CIR and Heston model, but also extensions of these models that are obtained by adding jumps. A precise mathematical formulation and a complete characterization of regular affine processes are due to Duffie et al. in 2003 [9]. Later several authors have contributed to the theory of general affine processes as Filipović, Mayerhofer and Keller-Ressel.

The class of affine processes introduced by Duffie et al. consists of all continuous-time Markov processes taking values in $\mathbb{R}_+^m \times \mathbb{R}^n$ for integers $m \geq 0$ and $n \geq 0$, whose log-characteristic function depends in an affine way on the initial state vector of the process. Stochastic processes of this type have been studied also where $D = \mathbb{R}_+$ or $\mathbb{R}$, they have been obtained as continuous-time limits of classic Galton-Watson branching processes with and without immigration. First let us recall the definition of affine process.

Definition 3. *(One-dimensional affine process) A time-homogenous Markov process $(X_t)_{t\geq 0}$ taking values in $D = \mathbb{R}_{\geq 0}$ or $\mathbb{R}$ is called affine if the characteristic function is exponentially affine in x. More precisely, this means that there exist $\mathbb{C}$-valued functions $\phi(t,u)$ and $\psi(t,u)$ defined on $D \times \mathcal{U}$, where*

$$\mathcal{U} := \begin{cases} \{u \in \mathbf{C} : \Re u \leq 0\}, & \text{if } D = \mathbb{R}_{\geq 0} \\ \{u \in \mathbf{C} : \Re u = 0\} & \text{if } D = \mathbb{R}. \end{cases}$$

and $\Re u$ denotes the real part of u, such that

$$E_x\left[e^{uX_t}\right] = exp\big(\phi(t,u) + x\psi(t,u)\big), \tag{2.2}$$

for all $x \in D$ and $(t,u) \in \mathbb{R}_+ \times \mathcal{U}$.

Duffie, Filipović and Schachermayer [9] introduce the following regularity conditions:

Definition 4. *An affine process is said to be regular, if the derivatives*

$$F(u) = \partial_t \phi(t,u)|_{t=0} \ \ and \ \ R(u) = \partial_t \psi(t,u)|_{t=0}$$

exist, and are continuous at $u = 0$.

Under this regularity assumptions, the functions F and R completely characterize the process $(X_t)_{t\geq 0}$. Moreover, the functions $\phi(t,u)$ and $\psi(t,u)$ defined in (2.2) are the solutions of the following generalized Riccati equations

$$\begin{cases} \partial_t \phi(t,u) = F\big(\psi(t,u)\big), & \phi(0,u) = 0, \\ \partial_t \psi(t,u) = R\big(\psi(t,u)\big), & \psi(0,u) = u, \end{cases} \tag{2.3}$$

For the readers we refer to, Duffie et al. [9, Theorem 2.7].

It is well-known that the CIR process belongs to the class of affine process in $\mathbb{R}_+$. More precisely, if we can find functions $\phi(t,u)$ and $\psi(t,u)$ with the initial conditions $\phi(0,u) = 0$ and $\psi(0,u) = u$, such that (2.2). For $T > 0$ and $u \in \mathcal{U} := \{u \in \mathbf{C} : \Re u \leq 0\}$ define the complex-valued Itô process

$$M(t) := f(t, X_t) = exp\big(\phi(T-t,u) + X_t\psi(T-t,u)\big)$$

Now, we assume that the functions ϕ and ψ are sufficiently differentiable then we can apply Itô formula and obtain

$$\begin{aligned}
\frac{df(t,X_t)}{f(t,X_t)} &= -\Big(\partial_t\phi(T-t,u) + \partial_t\psi(T-t,u)X_t\Big)dt + \psi(T-t,u)dX_t + \frac{1}{2}\sigma^2\psi(T-t,u)^2X_tdt \\
&= -\Big(\partial_t\phi(T-t,u) + \partial_t\psi(T-t,u)X_t - \psi(T-t,u)a(\theta - X_t) - \frac{1}{2}\sigma^2\psi(T-t,u)^2X_t\Big)dt \\
&+ \sigma\psi(T-t,u)\sqrt{X_t}dW_t, \quad t \leq T.
\end{aligned}$$

We denote

$$I(t) := \partial_t\phi(T-t,u) + \partial_t\psi(T-t,u)X_t - \psi(T-t,u)a(\theta - X_t) - \frac{1}{2}\sigma^2\psi(T-t,u)^2X_t,$$

we can write

$$\frac{dM(t)}{M(t)} = -I(t)dt + \sigma\psi(T-t,u)\sqrt{X_t}dW_t. \tag{2.4}$$

Since M is a martingale, we have

$$I(t) = 0, \text{for all } t \leq T \; a.s.$$

Letting $t \to 0$ by continuity of the parameters, we thus obtain

$$\partial_t\phi(T,u) + \partial_t\psi(T,u)x = \psi(T,u)a(\theta - x) + \frac{1}{2}\sigma^2\psi(T,u)^2x$$

for all $x \in \mathbb{R}_+$, $T > 0$ and $u \in \mathcal{U}$.

Note that both sides are affine in x therefore we can collect the coefficients and we get

$$\begin{aligned}
\partial_t\phi(t,u) &= a\theta\psi(t,u), \\
\partial_t\psi(t,u) &= -a\psi(t,u) + \frac{1}{2}\sigma^2\psi^2(t,u).
\end{aligned}$$

We also know the initial conditions

$$\phi(0,u)=0 \qquad and \qquad \psi(0,u)=u.$$

We derived that the functions $\phi(t,u)$ and $\psi(t,u)$ satisfy ordinary differential equations of the form

$$\begin{cases}\partial_t\phi(t,u)=F\big(\psi(t,u)\big), & \phi(0,u)=0,\\ \partial_t\psi(t,u)=R\big(\psi(t,u)\big), & \psi(0,u)=u,\end{cases} \tag{2.5}$$

with F and R satisfies

$$F(u)=a\theta u \tag{2.6}$$

$$R(u)=\frac{\sigma^2u^2}{2}-au. \tag{2.7}$$

The ordinary differential equations (2.5) are called the generalized Riccati equations. Solving the system 2.5 gave $\phi(t,u)$ and $\psi(t,u)$ in explicit form. One can remark that the ordinary differential equation for ψ, the second equation of the generalized Riccati equations, is a Bernoulli differential equation with parameter 2 and it can be represented as follows

$$\begin{aligned}\psi(t,u) &= \psi(0,u)e^{-\int_0^t ads}\big(1-u\int_0^t\frac{\sigma^2}{2}e^{-as}ds\big)^{-1}\\ &= \frac{ue^{-at}}{1-\frac{\sigma^2}{2a}u(1-e^{-at})}.\end{aligned} \tag{2.8}$$

Note that the first Riccati equation is just an integral, and ϕ may be written explicitly as:

$$\begin{aligned}\phi(t,u)-\phi(0,u) &= a\theta\int_0^t\psi(s,u)ds\\ &= a\theta\int_0^t\frac{ue^{-as}}{1-\frac{\sigma^2}{2a}u(1-e^{-as})}ds,\ let\ \ y=1-\frac{\sigma^2}{2a}u(1-e^{-as})\\ &= a\theta\int_1^{y(t)}-\frac{2}{\sigma^2}\frac{dy}{y}\end{aligned}$$

$$\phi(t,u)=-\frac{2a\theta}{\sigma^2}\log\big(1-\frac{\sigma^2}{2a}u(1-e^{-at})\big) \tag{2.9}$$

According to (2.8) and (2.9) the characteristic function of X_t is given by

$$\begin{aligned} E_x[e^{uX_t}] =& \int_{\mathbb{R}_+} p(t,x,y)e^{uy}dy \\ =& \big(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\big)^{-\frac{2a\theta}{\sigma^2}} \cdot \exp\Big(\frac{xue^{-at}}{1-\frac{\sigma^2}{2a}u(1-e^{-at})}\Big) \end{aligned}$$

2.1.2 Regularity property

So far the CIR process is defined as a solution of a stochastic differential equation. This setting is broadly used in the literature, especially in the area of financial mathematics. Another somewhat different setting, initiated by Duffie et al [9], is to construct the CIR process as a Markov process on the canonical path space. This approach has some advantage when we have to deal with the laws of CIR process from different starting points and it is also applicable for other affine models.

Since later we need to apply the ergodic theory of Feller processes, we adopt the approach of Duffie et al [9] in this section. To be precise, we first establish the connection between these two settings.

Let $\mathbb{R}_+ := [0, \infty)$. Consider the CIR process $(X_t, t \geq 0)$ starting from $x \in \mathbb{R}_+$, namely $(X_t, t \geq 0)$ is the unique strong solution to (2.1).

We denote $E_x(\cdot)$ and $P_x(\cdot)$ as the expectation and probability respectively given the initial condition $X_0 = x$, with $x \geq 0$ being a constant. The semigroup (T_t) associated with the CIR process is defined as

$$T_t f(x) := E_x[f(X_t)] = \int_{\mathbb{R}_+} p(t,x,y)f(y)dy, \tag{2.10}$$

where $f : \mathbb{R}_+ \to \mathbb{R}$ is bounded and continuous and we recall that $p(t, x, y)$ is the transition density function of the CIR process started at x and is given by

$$p(t,x,y) = \rho e^{-u-v}\Big(\frac{v}{u}\Big)^{\frac{q}{2}} I_q\big(2(uv)^{\frac{1}{2}}\big) \tag{2.11}$$

for $t > 0, x > 0$ and $y \geq 0$, where

$$\begin{aligned} \rho \equiv& \frac{2a}{\sigma^2\Big(1-e^{-at}\Big)}, \qquad & u \equiv \rho x e^{-at}, \\ v \equiv& \rho y, \qquad & q \equiv \frac{2a\theta}{\sigma^2} - 1, \end{aligned}$$

and $I_q(\cdot)$ is the first-order modified Bessel function with index q, defined by:

$$I_q(x) = \left(\frac{x}{2}\right)^q \sum_{k=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{2k}}{k!\Gamma(q+k+1)}.$$

For $x = 0$ the density function is given by

$$p(t,0,y) = \frac{\rho}{\Gamma(q+1)} v^q e^{-v} \tag{2.12}$$

for $t > 0$ and $y \geq 0$.

We write $C_0 = C_0(\mathbb{R}_+)$ for the class of continuous functions which vanish at infinity.

It is well-known that CIR process is an affine process (see [9]). It is already shown in [9, Section 8] (see also [36]) that the semigroup of every stochastic continuous affine process is a Feller semigroup. Since CIR process is a diffusion process, it is obviously stochastic continuous. Thus we know that $(T_t)_{t\geq 0}$ defined in (2.10) is a Feller semigroup.

We denote the canonical path space by $\hat{\Omega}$, namely $\hat{\Omega} = C\big([0,\infty);\mathbb{R}_+\big)$, and let $(\hat{X}_t, t \geq 0)$ be the canonical process on $\hat{\Omega}$. Let $(\hat{\mathcal{F}}_t)_{t\geq 0}$ be the filtration generated by the canonical process $(\hat{X}_t, t \geq 0)$, namely $\hat{\mathcal{F}}_t := \sigma(\hat{X}_s, 0 \leq s \leq t)$ and $\hat{\mathcal{F}} := \sigma(\hat{X}_s, s \geq 0)$. The map

$$X : (\Omega, \mathcal{F}) \to (\hat{\Omega}, \hat{\mathcal{F}})$$

induces a measure $\hat{P}_x$ on $(\hat{\Omega}, \hat{\mathcal{F}})$, which is the law of the CIR process starting from x on the canonical path space. Since $(T_t)_{t\geq 0}$ is Feller semigroup, the Markov process $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$ is a Feller process.

Following [29, Chapter 20] we give the definition of a regular Markov process on $\mathbb{R}_+$.

Definition 5. *Consider a continuous-time Markov process Z with state space $\big(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)\big)$ and distributions P_x. The process is said to be regular if there exist a locally finite measure μ on $\mathbb{R}_+$ and a continuous function $(t,x,y) \mapsto p(t,x,y) > 0$ on $(0,\infty) \times \mathbb{R}_+^2$ such that*

$$P_x\{Z_t \in B\} = \int_B p(t,x,y)\mu(dy), \quad x \in \mathbb{R}_+,\ B \in \mathcal{B}(\mathbb{R}_+),\ t > 0.$$

The measure μ is called the “supporting measure" of the process. It is unique up to an equivalence (see [29, page 399]).

Proposition 2. *The CIR process is a regular Feller process on $\mathbb{R}_+$.*

We will explain in the proof that the supporting measure can not be a Lebesgue measure.

Proof. As we have mentioned it before, the Feller property is already proved in [9, Section 8]. We only need to prove the regularity property.

The modified Bessel functions of the first kind can be expanded as

$$I_q(r) = \left(\frac{r}{2}\right)^q \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}r^2\right)^k}{k!\Gamma(q+k+1)}$$

thus has the following asymptotic forms

$$I_q(r) = \frac{1}{\Gamma(q+1)}\left(\frac{r}{2}\right)^q + O(r^{q+2}) \tag{2.13}$$

for small arguments $0 < r \ll \sqrt{q+1}$. If $\frac{2a\theta}{\sigma^2} < 1$, it follows that

$$p(t,x,0) := \lim_{y\to 0} p(t,x,y) = \infty, \quad \forall x \in \mathbb{R}_+.$$

Thus $(t,x,y) \mapsto p(t,x,y)$ is not continuous on $(0,\infty)\times\mathbb{R}_+^2$. On the other hand, if $\frac{2a\theta}{\sigma^2} > 1$, then we have

$$p(t,x,0) := \lim_{y\to 0} p(t,x,y) = 0$$

Therefore in both cases the behavior of $p(t,x,y)$ at point $y = 0$ violates the regularity condition. To overcome this difficulty, we define a measure η on $\big(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)\big)$ as

$$\eta(dx) := h(x)dx, \tag{2.14}$$

where

$$h(x) = \begin{cases} x^{\frac{2a\theta}{\sigma^2}-1}, & 0 \le x \le 1, \\ 1, & x > 1. \end{cases}$$

Then the transition density of the CIR process with respect to the new measure η is given by

$$\tilde{p}(t,x,y) = \frac{p(t,x,y)}{h(y)}, \quad t > 0,\ x \ge 0,\ y > 0. \tag{2.15}$$

Recall that

$$\rho \equiv \frac{2a}{\sigma^2\Big(1-e^{-at}\Big)}, \qquad u \equiv \rho x e^{-at},$$

$$v \equiv \rho y, \qquad q \equiv \frac{2a\theta}{\sigma^2} - 1.$$

At the point $y = 0$ we define

$$\tilde{p}(t,x,0) := \lim_{y\to 0} \tilde{p}(t,x,y) = \frac{1}{\Gamma(q+1)}\rho^{q+1}e^{-u} \in (0,\infty) \tag{2.16}$$

From (2.15) we get

$$0 < \tilde{p}(t,x,y) < \infty, \quad t > 0,\ x \geq 0,\ y > 0, \tag{2.17}$$

since $h(y)$ and $p(t,x,y)$ are positive and finite if $y > 0$. It follows from (2.16) and (2.17) that $0 < \tilde{p}(t,x,y) < \infty$ for all $(t,x,y) \in (0,\infty) \times \mathbb{R}^2_+$.

Moreover, the function $\tilde{p}(t,x,y)$ is continuous on $(0,\infty)\times(0,\infty)\times(0,\infty)$, which follows from the continuity and positivity of the functions $h(y)$ and $p(t,x,y)$ with $y > 0$. Next we prove the continuity of $\tilde{p}(t,x,y)$ at the point $(0,0,t_0)$.

Let $\delta > 0$ be sufficiently small. Then for $|t - t_0| \leq \delta$ and $0 \leq x, y \leq \delta$ we have

$$\begin{aligned}
&|\tilde{p}(t,x,y) - \tilde{p}(t_0,0,0)| \\
\leq&|\tilde{p}(t,x,y) - \tilde{p}(t,0,y)| + |\tilde{p}(t,0,y) - \tilde{p}(t,0,0)| + |\tilde{p}(t,0,0) - \tilde{p}(t_0,0,0)| \\
\leq&\left|\frac{p(t,x,y) - p(t,0,y)}{h(y)}\right| + \left|\frac{p(t,0,y)}{h(y)} - \tilde{p}(t,0,0)\right| + |\tilde{p}(t,0,0) - \tilde{p}(t_0,0,0)|.
\end{aligned} \tag{2.18}$$

By (2.11) and (2.13) we get

$$\begin{aligned}
&\left|\frac{p(t,x,y) - p(t,0,y)}{h(y)}\right| \\
=&\frac{1}{|y^q|}\left|\rho e^{-u-v}\Big(\frac{v}{u}\Big)^{\frac{q}{2}}\Big(\frac{1}{\Gamma(q+1)}(uv)^{\frac{q}{2}} + O\big((uv)^{\frac{q}{2}+1}\big)\Big) - \frac{\rho}{\Gamma(q+1)}v^q e^{-v}\right| \\
=&\frac{1}{|y^q|}\left|\frac{\rho}{\Gamma(q+1)}e^{-v}\big(e^{-u}-1\big)v^q + O\big(uv^{q+1}\big)\right| \\
\leq&\left|\frac{\rho^{q+1}}{\Gamma(q+1)}e^{-v}\big(e^{-u}-1\big)\right| + O(uv).
\end{aligned} \tag{2.19}$$

By (2.12) and (2.16) we have

$$\begin{aligned}
&\left|\frac{p(t,0,y)}{h(y)} - \tilde{p}(t,0,0)\right| \\
=&\left|\frac{\rho v^q e^{-v}}{\Gamma(q+1)y^q} - \frac{1}{\Gamma(q+1)}\rho^{q+1}\right| = \left|\frac{\rho^{q+1}}{\Gamma(q+1)}\big(e^{-v}-1\big)\right|.
\end{aligned} \tag{2.20}$$

Since ρ is a continuous function with respect to the variable t, it follows from (2.16) that

$$\lim_{t\to t_0} |\tilde{p}(t,0,0) - \tilde{p}(t_0,0,0)| = 0. \tag{2.21}$$

It follows from (2.18), (2.19), (2.20) and (2.21) that

$$\lim_{(t,x,y)\to(t_0,0,0)} |\tilde{p}(t,x,y) - \tilde{p}(t_0,0,0)| = 0.$$

The continuity at other remaining points can be proved with a similar argument. Thus $(t,x,y) \mapsto \tilde{p}(t,x,y)$ is a positive continuous function on $(0,\infty)\times\mathbb{R}_+^2$. Therefore the CIR process is regular with respect to the measure η. □

2.2 Positive Harris recurrence

In this section we prove the main result of this chapter, namely we show that the CIR process, as a Feller process on $\mathbb{R}_+$, is positive Harris recurrent. Recall that $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$ is the CIR process realized on the canonical path space. By (2.11) we know that

$$\hat{P}_x(\hat{X}_t \in A) = \int_A p(t,x,y)dy, \quad \forall\, A \in \mathcal{B}(\mathbb{R}_+).$$

According to Proposition (2), we know that $(\hat{\Omega}, \hat{\mathcal{F}}_t, \hat{P}_x, x \in \mathbb{R}_+)$ is a regular Feller process with respect to the measure η defined in (2.14).

The stability and ergodic theory of continuous-time Markov processes has a large literature which includes many different approaches. For the readers we refer to [44, 45, 46]. Recurrence theory permits us to establish stability even for models for which the stationary equations can not be explicitly solved (like the CIR process).

Let us first recall some basic definitions

Definition 6. *A continuous-time Markov process Y on the state space $\big(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)\big)$ is said to be Harris recurrent if for some σ-finite measure μ*

$$P_x\Big(\int_0^\infty \mathbf{1}_A(Y_s)ds = \infty\Big) = 1, \tag{2.22}$$

for any $x \in \mathbb{R}_+$ and $A \in \mathcal{B}(\mathbb{R}_+)$ with $\mu(A) > 0$.

Definition 7. *A continuous-time Markov process Y with state space $\big(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)\big)$ is said to be uniformly transient if*

$$\sup_x E_x\Big[\int_0^\infty \mathbf{1}_K(Y_s)ds\Big] < \infty \tag{2.23}$$

for every compact $K \subset \mathbb{R}_+$.

Harris recurrence means that the Markov process $(Y_t, t \geq 0)$ visit the Borel set A with $\mu(A) > 0$ infinitely often. It was shown in [29, Theorem 20.17] that any Feller process is either Harris recurrent or uniformly transient. We are now ready to prove the positive recurrence in the sense of Harris for the CIR process $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$.

Lemma 1. *The CIR process* $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$ *is not uniformly transient.*

Proof. We take $K = [0, M]$ with $M > 0$. Then for any fixed $x \in (0, \infty)$

$$\begin{aligned} \hat{E}_x\Big[\int_0^\infty \mathbf{1}_{[0,M]}(\hat{X}_t)dt\Big] &= \int_0^\infty \hat{E}_x\big[\mathbf{1}_{[0,M]}(\hat{X}_t)\big]dt \\ &= \int_0^\infty \int_0^M p(t,x,y)dydt \\ &= \int_0^M dy \int_0^\infty p(t,x,y)dt. \end{aligned}$$

The modified Bessel functions of the first kind have the following asymptotic forms, for small arguments $0 < r \ll \sqrt{q+1}$, one obtains

$$I_q(r) \approx \frac{1}{\Gamma(q+1)}(\frac{r}{2})^q$$

where Γ denotes the Gamma function. Let $\epsilon > 0$ be small enough. For any $y \in [\epsilon, M]$ and large enough t we have

$$p(t,x,y) \approx \frac{\rho}{\Gamma(q+1)}e^{-\rho y}(\rho y)^q$$

and thus for $y \in [\epsilon, M]$

$$\int_0^\infty p(t,x,y)dt = \infty.$$

It follows that

$$\hat{E}_x\Big[\int_0^\infty \mathbf{1}_{[0,M]}(\hat{X}_t)dt\Big] = \int_0^M dy \int_0^\infty p(t,x,y)dt = \infty.$$

This proves that the CIR process $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$ is not uniformly transient. □

Harris recurrence guarantees the existence of a unique (up to multiplication by a constant) invariant measure for the Markov process (see e.g. [31]), but not necessarily finite. If this invariant measure is finite, then the process is called positive Harris recurrent.

Theorem 2. *The CIR process* $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$ *is positive Harris recurrent.*

Proof. In Proposition (2) we have shown that the CIR process is a regular Feller process with respect to the measure ρ defined in (2.14). It follows from Lemma 1 and [29, Theorem 20.17] that the CIR process is Harris recurrent and ρ can be taken as a possible reference measure in place of μ in the Definition 6. Due to [29, Theorem 20.18] (see also [40, Theorem 1.3.5]) it is possible to construct a locally finite invariant measure μ for the CIR process. Furthermore μ is equivalent to ρ and every σ-finite, invariant measure for the CIR process agrees with μ up to a normalization. It was shown in [7] that μ is a Gamma distribution and has the form

$$\mu(dy) := \frac{\omega^\nu}{\Gamma(\nu)} y^{\nu-1} e^{-\omega y} dy, \quad y \geq 0, \tag{2.24}$$

where $\omega \equiv \frac{2a}{\sigma^2}$ and $\nu \equiv \frac{2a\theta}{\sigma^2}$. Thus the CIR process is positive Harris recurrent. □

2.3 Ergodicity results

As a consequence of positive Harris recurrence we are able to prove the strong ergodicity of the CIR process. Based on Birkhoff's ergodic theorem and the strong ergodicity we give the ergodic results on transformation of the CIR.

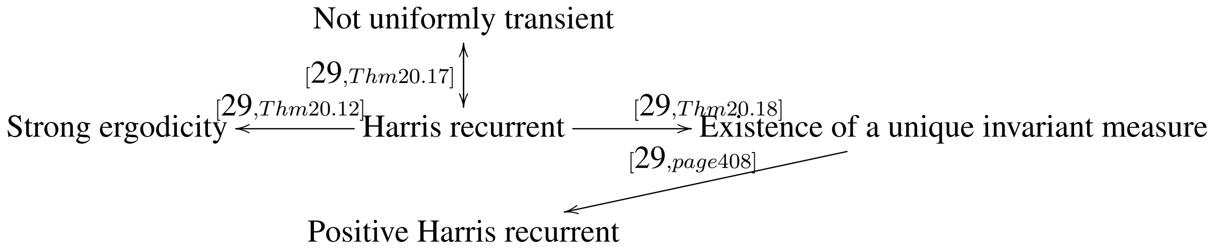
Definition 8. *(i) The tail σ-field on $\hat{\Omega}$ is defined as $\hat{\mathcal{T}} = \cap_{t\geq 0}\hat{\mathcal{T}}_t$, where $\hat{\mathcal{T}}_t = \sigma\{\hat{X}_s : s \geq t\}$.*
(ii) A σ-field $\mathcal{G} \subset \hat{\mathcal{F}}$ on $\hat{\Omega}$ is said to be $\hat{P}_\nu$-trivial if $\hat{P}_\nu(A) = 0$ or $\hat{P}_\nu(A) = 1$ for every $A \in \mathcal{G}$, where $\hat{P}_\nu(\cdot) := \int_{\mathbb{R}_+} \hat{P}_x(\cdot)\nu(dx)$ denotes the distribution of the CIR process with initial distribution ν.

From the positive Harris recurrence of the CIR process we reproduce the following well-known fact.

Corollary 1. *The CIR process $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$ is strongly ergodic, meaning that the tail σ-field $\hat{\mathcal{T}}$ of the CIR process is $\hat{P}_\mu$-trivial for every μ.*

Proof. According to [29, Theorem 20.12] any Harris recurrent Feller process is strongly ergodic. From Theorem 2 we know that the CIR process $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$ is a strongly ergodic Markov process which means the tail σ-field $\hat{\mathcal{T}}$ of the CIR process is $\hat{P}_\mu$-trivial for every μ, see [29, Theorem 20.10]. □

Now we can summarize our results with the following diagram



Based on Corollary 1, we now apply the key result of stationarity theory which is Birkhoff's ergodic theorem to get an ergodicity result for a transformation of the CIR process, which will be applied the next section to calibrate parameters of a credit migration model.

Let $(S, \mathcal{S})$ be an arbitrary measurable space. Given a measure μ and a measurable transformation T on S, we say that T be a μ-preserving map on S if $\mu \circ T^{-1} = \mu$. Thus, if ξ be a random element in S with distribution μ, then T is μ-preserving if and only if $T\xi \overset{d}{=} \xi$.

Now we recall Birkhoff's ergodic Theorem (see e.g. [29, Theorem 10.6]). Let $(S, \mathcal{S})$ be a measurable space and ξ be a random element in S with distribution μ, and let T be a μ-preserving map on S with invariant σ-field $\mathcal{I} := \{A \in \mathcal{S} : T^{-1}A = A\}$ and let $\mathcal{I}_\xi := \{\xi^{-1}A : A \in \mathcal{I}\}$. Then for any measurable function $f \geq 0$ on S,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k \xi) \to E[f(\xi)|\mathcal{I}_\xi] \quad \text{a.s.}$$

The same convergence holds in L^p for some $p \geq 1$ when $f \in L^p(\mu)$.

We now consider the canonical CIR process $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_x, x \in \mathbb{R}_+)$ which is previous constructed in this section. Recall that the measure μ defined in (2.24) is the unique invariant probability measure for the CIR process. Suppose that $g : \mathbb{R}_+ \to \mathbb{R}$ is a continuous function. We define a random sequence

$$\xi : (\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}_\mu) \to \mathbb{R}^\infty$$

by

$$\xi(\hat{\omega}) := \left(\int_0^1 g\big(\hat{X}_s\big)(\hat{\omega}) ds, \int_1^2 g\big(\hat{X}_s\big)(\hat{\omega}) ds, \cdots \right)$$

The shift operator $\theta : \mathbb{R}^\infty \to \mathbb{R}^\infty$ is defined as

$$\theta(x_0, x_1, x_2, x_3, \cdots) := (x_1, x_2, x_3, \cdots)$$

for $x = (x_0, x_1, x_2, \cdots)$ and the invariant σ-field $\mathcal{I}$ on $\mathbb{R}^\infty$ generated by the shift operator θ is given by

$$\mathcal{I} := \{A \in \mathcal{B}(\mathbb{R}^\infty) : \theta^{-1}A = A\}.$$

Lemma 2. *Suppose that $g : \mathbb{R}_+ \to \mathbb{R}$ is continuous and $f : \mathbb{R} \to \mathbb{R}_+$ is measurable. Then we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} f\Big(\int_j^{j+1} g\big(\hat{X}_s\big)(\hat{\omega})ds\Big) = \hat{E}_\mu\Big[f\Big(\int_0^1 g\big(\hat{X}_s\big)ds\Big)\Big|\mathcal{I}_\xi\Big]$$

for $\hat{P}_\mu$-almost all $\hat{\omega} \in \hat{\Omega}$, where $\mathcal{I}_\xi := \{\xi^{-1}A : A \in \mathcal{I}\}$.

Proof. Since the initial distribution μ is an invariant measure for the CIR process

$$\hat{X}_1 \sim \mu.$$

It follows now from homogeneous Markov property that the random sequence ξ is stationary, i.e.

$$\theta\xi \overset{d}{=} \xi.$$

According to Birkhoff's ergodic Theorem, for any measurable function $h \geq 0$ on $\big(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty)\big)$,

$$\frac{1}{N}\sum_{i=0}^{N-1} h(\theta^i\xi) \to \hat{E}_\mu[h(\xi)|\mathcal{I}_\xi] \quad \text{a.s.}$$

where $\mathcal{I}_\xi := \xi^{-1}\mathcal{I}$ and $\mathcal{I}$ is the invariant σ-field on $\mathbb{R}^\infty$ generated by the shift operator θ. Especially, if we take

$$h(x) := f\big(x_0\big)$$

for $x = (x_0, x_1, x_2, \cdots) \in \mathbb{R}^\infty$, then we get

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} f\Big(\int_j^{j+1} g\big(\hat{X}_s\big)(\hat{\omega})ds\Big) = \hat{E}_\mu\Big[f\Big(\int_0^1 g\big(\hat{X}_s\big)ds\Big)\Big|\mathcal{I}_\xi\Big]$$

for $\hat{P}_\mu$-almost all $\hat{\omega} \in \hat{\Omega}$. □

Lemma 3. *The invariant σ-field $\mathcal{I}_\xi$ of ξ is $\hat{P}_\mu$-trivial, namely*

$$\hat{P}_\mu(A) = 0 \text{ or } \hat{P}_\mu(A) = 1 \text{ for every } A \in \mathcal{I}_\xi.$$

Proof. Since $\hat{X}_t$ is continuous and $g : \mathbb{R}_+ \to \mathbb{R}$ is continuous, thus

$$\sigma(\xi_j) = \sigma\bigg(\int_j^{j+1} g\big(\hat{X}_s\big)ds\bigg) \subset \sigma\{\hat{X}(s) : j \leq s \leq j+1\}.$$

Thus the tail σ-field of the random sequence ξ is contained in the tail σ-field $\mathcal{T}$ of the CIR process $\hat{X}_t$, where $\mathcal{T} = \cap_{t\geq 0}\mathcal{T}_t$ and $\mathcal{T}_t = \sigma\{\hat{X}_s : s \geq t\}$. Because the invariant σ-field of the random sequence ξ is contained in the tail σ-field of ξ, we get $\mathcal{I}_\xi \subset \mathcal{T}$. According to Corollary 1, the tail σ-field $\mathcal{T}$ of the CIR process $\hat{X}_t$ is $\hat{P}_\mu$-trivial, it follows that $\mathcal{I}_\xi$ is also $\hat{P}_\mu$-trivial. □

Theorem 3. *Suppose that $g : \mathbb{R}_+ \to \mathbb{R}$ is continuous and $f : \mathbb{R} \to \mathbb{R}_+$ is measurable. Then for any $x \in \mathbb{R}_+$ we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{j=0}^{N-1} f\Big(\int_j^{j+1} g\big(\hat{X}_s\big)(\hat{\omega})ds\Big) = \hat{E}_\mu\Big[f\Big(\int_0^1 g\big(\hat{X}_s\big)(\hat{\omega})ds\Big)\Big]$$

for $\hat{P}_x$-almost all $\hat{\omega} \in \hat{\Omega}$, where μ is given (2.24)and is the unique invariant probability measure for the CIR process.

Proof. Since $\mathcal{I}_\xi$ is $\hat{P}_\mu$-trivial, the conditional expectation

$$\hat{E}_\mu\Big[f\Big(\int_0^1 g\big(\hat{X}_s\big)ds\Big)\Big|\mathcal{I}_\xi\Big]$$

is a constant and equals

$$\hat{E}_\mu\Big[f\Big(\int_0^1 g\big(\hat{X}_s\big)ds\Big)\Big].$$

From Lemma 2 we get

$$\lim_{N\to\infty} \frac{1}{N}\sum_{j=0}^{N-1} f\Big(\int_j^{j+1} g\big(\hat{X}_s\big)(\hat{\omega})ds\Big) = \hat{E}_\mu\Big[f\Big(\int_0^1 g\big(\hat{X}_s\big)ds\Big)\Big] \tag{2.25}$$

where the convergence in (2.25) holds for $\hat{P}_\mu$-almost all $\hat{\omega} \in \hat{\Omega}$. If we set

$$N := \big\{\hat{\omega} \in \hat{\Omega} : \text{ the convergence in (2.25) fails for } \hat{\omega}\big\},$$

then

$$\hat{P}_\mu(N) = \int_{\mathbb{R}_+} \hat{P}_x(N)\mu(dx) = 0,$$

which implies $\hat{P}_x(N) = 0$ for μ-almost all $x \in \mathbb{R}_+$. For any $x \in \mathbb{R}_+$, by the Markov property, it holds

$$\begin{aligned}\hat{P}_x(N) =& \hat{E}_x[\mathbf{1}_N] = \hat{E}_x\big[\hat{E}_x[\mathbf{1}_N|\hat{\mathcal{F}}_1]\big]\\ =& \hat{E}_x\big[\hat{P}_{\hat{X}_1}[\theta_1^{-1}(N)]\big] = \int_0^\infty \hat{P}_y(N)p(1,x,y)dy\\ =& 0,\end{aligned}$$

where $p(t, x, y)$ denotes the transition density function of the CIR process. In the above calculation we have used the fact that $\theta_1^{-1}(N) = N$, namely the pre-image of N under the shift operator θ is still N. Thus we have proved that the convergence in (2.25) holds for $\hat{P}_x$-almost all $\hat{\omega} \in \hat{\Omega}$. □

2.4 Application in one credit migration model

In this section we show a simple application of Theorem 3 in calibration of the parameters in one credit migration model. We should remark that the results presented in this section have been derived in [39] with a different method. Their method is very analytical and relies very much on the affine structure of the CIR process. In contrast to [39] our method is more probabilistic and can be extended to more general models.

The idea of Hurd and Kuznetsov [22] was to generalize the $0-1$ process to a finite state Markov chain on $\{1, 2, \cdots, k\}$ where each state represents credit rating or distance to default of the firm. Here we consider a simpler version of the credit migration model of Hurd and Kuznetsov where $k = 8$. Consider the finite state space $\{1, 2, \cdots, 8\}$, which can be identified with Moody's rating classes via the mapping:

$$\{1, 2, \cdots, 8\} \leftrightarrow \{\text{AAA, AA, A, BBB, BB, B, CCC, default}\}.$$

The credit migration matrix $P(s, t), 0 \le s \le t$, is a stochastic 8×8 matrix and describes all possible transition probabilities between rating classes from time s to time t, namely

$$P(s,t) = \Big(p_{ij}(s,t)\Big)_{1\le i,j\le 8} \tag{2.26}$$

where each $p_{ij}(s, t)$ in (2.26) represents the transition probability from state i to state j from time s to time t. The last column of the migration matrix $P(s, t)$ represents the absorbing state of default, i.e. the probability of leaving the default state equals zero. It was assumed in [22] that the migration matrix $P(s, t)$ is given by

$$P(s,t) = \exp\left(\left(\int_s^t X_r dr\right)\cdot \hat{P}\right) \tag{2.27}$$

where $\hat{P}$ is a 8×8 constant matrix and X_t is a CIR process with long-term average 1, namely X_t satisfies

$$X_t = X_0 + \int_0^t a(1 - X_s)ds + \int_0^t \sigma\sqrt{X_s}dW_s, \quad t \geq 0$$

The matrix $\hat{P}$ is called the generator matrix. Thus the dynamics of the migration matrix is determined by two factors: the generator $\hat{P}$ and the CIR process X_t.

A natural question is how to calibrate the parameters of the above credit migration model. More precisely, how can one determine the generator matrix $\hat{P}$ and the parameters a, θ, σ of the CIR process?

Among many other things, this problem was considered by [39] and they presented the following way to calibrate the parameters of the above model, with extra assumptions on the generator matrix $\hat{P}$. The starting point is the Moody-matrix P_{Moody}, which is derived by Moody as the historical average of one year migration matrix, based on the historical data from 1920 to 1996. The logarithm matrix of P_{Moody} is given by

$$\hat{P}_{Moody} := \log(P_{Moody}) = \log(id - (id - P_{Moody})) = -\sum_{j=1}^{\infty} \frac{1}{j}(id - P_{Moody})^j \tag{2.28}$$

According to [24, Theorem 2.2], the right-hand side in (2.28) converges and thus $\hat{P}_{Moody}$ is well-defined and

$$\exp(\hat{P}_{Moody}) = P_{Moody}$$

It was indicated by [39] that $\hat{P}_{Moody}$ is diagonalizable and thus can be written as

$$\hat{P}_{Moody} = G \cdot (-\hat{D}_{Moody,ii}) \cdot G^{-1} \tag{2.29}$$

where $\hat{D}_{Moody,ii}$ is a diagonal matrix and $G = (g_{ik})_{1 \leq j,k \leq 8}$ is a matrix whose columns are the corresponding eigenvectors of $\hat{P}_{Moody}$. Thus we get

$$P_{Moody} = G \cdot (e^{-\hat{D}_{Moody,ii}}) \cdot G^{-1}$$

Instead of finding a full 8×8 generator matrix, it was proposed in [39] to seek the generator matrix in the form

$$\hat{P} = G \cdot (-\hat{D}) \cdot G^{-1}$$

where $\hat{D} = (\hat{D}_{ii})$ is a diagonal matrix with diagonal elements $\hat{D}_{ii} \geq 0$ and G is given in (2.29).

According to (2.27) the migration matrix from time j to $j+1$, $j=0,1,2,\cdots$ is given by

$$
\begin{aligned}
P(j,j+1) &= \exp\left(\left(\int_j^{j+1} X_s ds\right)\cdot \hat{P}\right) \\
&= G\cdot\left(e^{-\hat{D}_{ii}\int_j^{j+1} X_s ds}\right)\cdot G^{-1} \\
&= G\cdot\begin{pmatrix} e^{-\hat{D}_{11}\int_j^{j+1} X_s ds} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & e^{-\hat{D}_{88}\int_j^{j+1} X_s ds} \end{pmatrix}\cdot G^{-1}
\end{aligned}
$$

Thus the Cesàro average $\frac{1}{N}\sum_{j=0}^{N-1} P(j,j+1)$ equals

$$
G\cdot\begin{pmatrix} \frac{1}{N}\sum_{j=0}^{N-1} e^{-\hat{D}_{11}\int_j^{j+1} X_s ds} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & \frac{1}{N}\sum_{j=0}^{N-1} e^{-\hat{D}_{88}\int_j^{j+1} X_s ds} \end{pmatrix}\cdot G^{-1} \tag{2.30}
$$

For each $1\le i\le 8$ by taking $g(x)=\hat{D}_{ii}x$ and $f(x)=e^{-x}$ in Theorem 3 of last section we get

$$
\lim_{N\to\infty}\frac{1}{N}\sum_{j=0}^{N-1} e^{-\hat{D}_{ii}\int_j^{j+1} X_s ds} = \mathbf{E}_\mu\left[e^{-\hat{D}_{ii}\int_0^1 X_s ds}\right] \tag{2.31}
$$

and the convergence in (2.31) holds almost surely. Since each term under the limit sign on the left hand side of (2.31) is bounded, by dominated convergence theorem, the convergence in (2.31) holds also in L^p for any $p\ge 1$. We should remark that the L^2 convergence in (2.31) was obtained in [39] with a different method. Our method used here is more probabilistic and provides us with a stronger convergence in (2.31).

Since the Laplace transform of $\int_0^1 X_s ds$ is well-known (for example, see [3, Lemma 2]), we get

$$
E_x\left[e^{-\hat{D}_{ii}\int_0^1 X_s ds}\right] = e^{aA(0,\hat{D}_{ii},1)+x\cdot B(0,\hat{D}_{ii},1)}
$$

with deterministic functions

$$B(\lambda, u, t) := -\frac{\lambda(h - a + (h+a)e^{-ht}) + 2u(1 - e^{-ht})}{\sigma^2\lambda(1 - e^{-ht}) + h + a + (h-a)e^{-ht}}$$

$$A(\lambda, u, t) := \frac{2}{\sigma^2}\log\left(\frac{2he^{\frac{1}{2}(a-h)t}}{\sigma^2\lambda(1 - e^{-ht}) + h + a + (h-a)e^{-ht}}\right)$$

and $h := \sqrt{a^2 + 2u\sigma^2}$. Therefore

$$\begin{aligned} E_\mu\left[e^{-\hat{D}_{ii}\int_0^1 X_s ds}\right] &= \int_0^\infty E_x\left[e^{-\hat{D}_{ii}\int_0^1 X_s ds}\right]\mu(dx) \\ &= \int_0^\infty e^{aA(0,\hat{D}_{ii},1) + x\cdot B(0,\hat{D}_{ii},1)}\frac{\omega^\nu}{\Gamma(\nu)}x^{\nu-1}e^{-\omega x}dx \\ &= e^{a\hat{D}_{ii}A(0,\hat{D}_{ii},1)}\left(\frac{2a}{2a - \sigma^2\hat{D}_{ii}B(0,\hat{D}_{ii},1)}\right)^{\frac{2a}{\sigma^2}}. \end{aligned} \tag{2.32}$$

On the other hand P_{Moody} is the historical average of one year migration matrix and thus it is reasonable to assume that

$$P_{Moody} = \lim_{N\to\infty}\frac{1}{N}\sum_{j=0}^{N-1}P(j, j+1), \tag{2.33}$$

if the limit on the right-hand side exists.

It now follows from (2.30), (2.31), (2.32) and (2.33) that

$$-\hat{D}_{Moody,ii} = a\hat{D}_{ii}A(0,\hat{D}_{ii},1) + \frac{2a}{\sigma^2}\log\left(\frac{2a}{2a - \sigma^2\hat{D}_{ii}B(0,\hat{D}_{ii},1)}\right). \tag{2.34}$$

for each $1 \le i \le 8$.

To get more equations for the unknown parameters, we consider the probability of rating downgrade from level A to the level BBB within time j und $j+1$, which is given by:

$$p_{3,4}(j, j+1) := \sum_{k=1}^{8} g_{3k}\cdot e^{-\hat{D}_{kk}\int_j^{j+1}X_s ds}\cdot g^{k4} \tag{2.35}$$

where $(g^{ik})_{1\le i,k\le 8}$ is the inverse of the matrix G. Similar to (2.31) we know that as $N \to \infty$

$$\frac{1}{N}\sum_{j=0}^{N-1}p_{3,4}(j, j+1)$$

converges almost surely to a constant, denoted by $E_{3,4,\infty}$, and we have

$$E_{3,4,\infty} = \sum_{k=1}^{8} g_{3k} e^{a\hat{D}_{kk}A(0,\hat{D}_{kk},1)} \left(\frac{2a}{2a - \sigma^2 \hat{D}_{kk} B(0,\hat{D}_{kk},1)} \right)^{\frac{2a}{\sigma^2}} g^{k4}$$

Define the ergodic variation

$$V_{3,4,\infty} := \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} \left(p_{3,4}(j,j+1) - E_{3,4,\infty} \right)^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} p_{3,4}^2(j,j+1) - E_{3,4,\infty}^2 \tag{2.36}$$

Since

$$p_{3,4}^2(j,j+1) = \sum_{i,k=1}^{8} g_{3i} g_{3k} e^{-(\hat{D}_{ii}+\hat{D}_{kk}) \int_j^{j+1} X_s ds} \cdot g^{i4} g^{k4}$$

it follows again from Theorem 3 that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} p_{3,4}^2(j,j+1) = \sum_{i,k=1}^{8} g_{3i} g_{3k} e^{(\hat{D}_{ii}+\hat{D}_{kk})aA(0,\hat{D}_{ii}+\hat{D}_{kk},1)} \cdot \left(\frac{2a}{2a-\sigma^2(\hat{D}_{ii}+\hat{D}_{kk})B(0,\hat{D}_{ii}+\hat{D}_{kk},1)} \right)^{\frac{2a}{\sigma^2}} g^{i4} g^{k4}$$

Thus we get

$$V_{P,3,4,\infty} = \sum_{i,k=1}^{8} g_{3i} g_{3k} e^{(\hat{D}_{ii}+\hat{D}_{kk})a\theta A(0,\hat{D}_{ii}+\hat{D}_{kk},1)} \cdot \left(\frac{2a}{2a - \sigma^2(\hat{D}_{ii}+\hat{D}_{kk}) B(0,\hat{D}_{ii}+\hat{D}_{kk},1)} \right)^{\frac{2b}{\sigma^2}} g^{i4} g^{k4} - \sum_{i,k=1}^{8} g_{3i} g_{3k} e^{\hat{D}_{ii}a\theta A(0,\hat{D}_{ii},1) + \hat{D}_{kk}a\theta A(0,\hat{D}_{kk},1)} \cdot \left(\frac{2a}{2a - \sigma^2 \hat{D}_{ii} B(0,\hat{D}_{ii},1)} \right)^{\frac{2b}{\sigma^2}} \left(\frac{2a}{2a - \sigma^2 \hat{D}_{kk} B(0,\hat{D}_{kk},1)} \right)^{\frac{2b}{\sigma^2}} g^{i4} g^{k4} \tag{2.37}$$

Based on the historical data from 1920 to 1996 Moody also gave the standard variation of one year transition probabilities between different rating classes. For example the standard variation of the one year transition probability from rating class A to BBB is 0.053. We could approximately assume that this value coincides with the square of the ergodic variation defined in (2.36), namely

$$V_{3,4,\infty} = (0.053)^2. \tag{2.38}$$

Summarizing (2.34), (2.37) and (2.38) we get 9 equations for 10 unknown parameters. By fitting Moody's standard variation of one year transition probability of another rating transition we will get an extra equation. Thus all parameters of this migration matrix model can be uniquely determined by solving the 10 equations we have derived.

Chapter 3

Positive Harris recurrence, exponential ergodicity and calibration of the BAJD

In this chapter, we find the transition densities of the basic affine jump-diffusion (BAJD), which has been introduced by Duffie and Gârleanu as an extension of the classical CIR model with jumps. We prove the positive Harris recurrence and exponential ergodicity of the BAJD. Furthermore, we prove that the unique invariant probability measure π of the BAJD is absolutely continuous with respect to the Lebesgue measure and we also derive a closed form formula for the density function of π.

3.1 Characteristic function of the BAJD

We may now recall some preliminaries on the BAJD process. As defined in (1.7), it is the unique strong solution $X = (X_t)_{t\geq 0}$ to the following stochastic differential equation

$$dX_t = a(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t + dJ_t, \quad X_0 \geq 0. \tag{3.1}$$

where a, θ, σ are positive constants, $(W_t)_{t\geq 0}$ is a one-dimensional Brownian motion and $(J_t)_{t\geq 0}$ is a one-dimensional non-decreasing pure jump Lévy process with identical independent increments, characterized by its jump measure ν which is supported on $(0, \infty)$. The Lévy measure ν is given by

$$\nu(dy) = \begin{cases} cde^{-dy}dy, & y \geq 0, \\ 0, & y < 0, \end{cases}$$

for some constants $c > 0$ and $d > 0$. Throughout this chapter we denote $P_x(\cdot)$ and $E_x(\cdot)$ as the probability and expectation respectively given the initial condition $X_0 = x$, with $x \geq 0$ being a constant.

By the affine structure of the BAJD process X, the characteristic function of X_t (given that $X_0 = x$) is of the form

$$E_x\left[e^{uX_t}\right] = \exp\left(\phi(t,u) + x\psi(t,u)\right), \quad u \in \mathcal{U} := \{u \in \mathbb{C} : \Re u \leq 0\}, \tag{3.2}$$

where the functions $\phi(t,u)$ and $\psi(t,u)$ solve the generalized Riccati equations

$$\begin{cases} \partial_t \phi(t,u) = F\big(\psi(t,u)\big), & \phi(0,u) = 0, \\ \partial_t \psi(t,u) = R\big(\psi(t,u)\big), & \psi(0,u) = u \in \mathcal{U}, \end{cases} \tag{3.3}$$

with

$$\begin{aligned} F(u) &= a\theta u + \frac{cu}{d-u}, \quad u \in \mathbb{C} \setminus \{d\}, \\ R(u) &= \frac{\sigma^2 u^2}{2} - au, \quad u \in \mathbb{C}. \end{aligned}$$

Remark that, it is not difficult to find the explicit form of the functions $\partial_t\phi$ and $\partial_t\psi$. We have to proceed as in the Section (2.1). More precisely, we can find the functions $\partial_t\phi$ and $\partial_t\psi$ under the initial conditions such that

$$M_t := f(t, X_t) = \exp\left(\phi(T-t,u) + X_t\psi(T-t,u)\right)$$

is a martingale.

By applying Itô formula, one can get

$$\frac{dM(t)}{M(t)} = -I(t)dt + \sigma\psi(T-t,u)\sqrt{X_t}dW_t + \int_{(0,\infty)} \left(e^{y\psi(T-t,u)} - 1\right)\nu(dy), \tag{3.4}$$

We can derive the generalized Riccati equation by collecting the coefficients.
One can remark that the function R does not depend on the parameters of the jumps c and d. For this reason, the solution of the second equation of the system (3.3), ψ is the same as in the case of classical CIR (2.8).

$$\psi(t,u) = \frac{ue^{-at}}{1 - \frac{\sigma^2}{2a}u(1-e^{-at})} \tag{3.5}$$

Note that the first equation in the generalized Riccati equation is just an integral, and ϕ may be written explicitly as:

$$\phi(t,u) = \begin{cases} -\frac{2a\theta}{\sigma^2}\log\left(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\right) \\ \qquad + \frac{c}{a-\frac{\sigma^2 d}{2}}\log\left(\frac{d - \frac{\sigma^2 du}{2a} + \left(\frac{\sigma^2 d}{2a} - 1\right)ue^{-at}}{d-u}\right), & \text{if } \Delta \neq 0, \\ -\frac{2a\theta}{\sigma^2}\log\left(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\right) + \frac{cu(1-e^{-at})}{a(d-u)}, & \text{if } \Delta = 0, \end{cases} \tag{3.6}$$

where $\Delta = a - \sigma^2 d/2$. Here the complex-valued logarithmic function $\log(\cdot)$ is to be understood as its main branch defined on $\mathbb{C} - \{0\}$.

According to (3.2), (3.5) and (3.6), the characteristic function of X_t is given by

$$E_x[e^{uX_t}] = \begin{cases} \left(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\right)^{-\frac{2a\theta}{\sigma^2}} \cdot \left(\frac{d - \frac{\sigma^2 du}{2a} + \left(\frac{\sigma^2 d}{2a} - 1\right)ue^{-at}}{d-u}\right)^{\frac{c}{a - \frac{\sigma^2 d}{2}}} \\ \quad \cdot \exp\left(\frac{xue^{-at}}{1 - \frac{\sigma^2}{2a}u(1-e^{-at})}\right), \qquad \text{if } \Delta \neq 0, \\ \left(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\right)^{-\frac{2a\theta}{\sigma^2}} \cdot \exp\left(\frac{cu(1-e^{-at})}{a(d-u)} + \frac{xue^{-at}}{1 - \frac{\sigma^2}{2a}u(1-e^{-at})}\right), \\ \qquad \text{if } \Delta = 0. \end{cases} \tag{3.7}$$

Here the complex-valued power functions $z^{-2a\theta/\sigma^2} := \exp(-(2a\theta/\sigma^2)\log z)$ and $z^{c/(a-\sigma^2 d/2)} := \exp((c \log z)/(a - \sigma^2 d/2))$ are to be understood as their main branches defined on $\mathbb{C} - \{0\}$. Obviously $E_x\big[\exp(uX_t)\big]$ is continuous in $t \geq 0$ and thus the BAJD process X is stochastically continuous.

We should point out that if we allow the parameter c to be 0, then the stochastic differential equation (3.1) turns into

$$dZ_t = a(\theta - Z_t)dt + \sigma\sqrt{Z_t}dW_t, \quad Z_0 = x \geq 0. \tag{3.8}$$

To avoid confusions we have used Z_t instead of X_t here. The unique solution $Z := (Z_t)_{t \geq 0}$ to (3.8) is the well-known Cox-Ingersoll-Ross (CIR) process and it holds

$$E_x[e^{uZ_t}] = \left(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\right)^{-\frac{2a\theta}{\sigma^2}} \cdot \exp\left(\frac{xue^{-at}}{1 - \frac{\sigma^2}{2a}u(1-e^{-at})}\right). \tag{3.9}$$

Later we will find a distribution ν_t on $\mathbb{R}_+$ such that

$$\int_{\mathbb{R}+} e^{uy}\nu_t(dy) = \begin{cases} \left(\frac{d - \frac{\sigma^2 du}{2a} + \left(\frac{\sigma^2 d}{2a} - 1\right)ue^{-at}}{d-u}\right)^{\frac{c}{a - \frac{\sigma^2 d}{2}}}, & \text{if } \Delta \neq 0, \\ \exp\left(\frac{cu(1-e^{-at})}{a(d-u)}\right), & \text{if } \Delta = 0. \end{cases} \tag{3.10}$$

Then it follows from (3.7), (3.9) and (3.10) that the distribution of the BAJD is the convolution of the distribution of the CIR process and ν_t. In light of this observation we can thus identify the transition probabilities $p(t, x, y)$ of the BAJD with

$$p(t, x, y) = \int_{\mathbb{R}_+} f(t, x, y - z)\nu_t(dz), \qquad x, y \geq 0, \ t > 0, \tag{3.11}$$

to avoid confusion we have used f instead of p, where $f(t, x, y)$ denotes the transition densities of the CIR process.

Remark 4. For a different way of representing the distribution of X_t as a convolution we refer the reader to [13]. In fact it was indicated in [13, Remark 4.8] that the distribution of any affine process on $\mathbb{R}_+$ can be represented as the convolution of two distributions on $\mathbb{R}_+$.

3.2 Mixtures of Bessel distributions

To find a distribution ν_t with the characteristic function of the form (3.10) and study the distributional properties of the BAJD, it is inevitable to encounter the Bessel distributions and mixtures of Bessel distributions.

We start with a slight variant of the Bessel distribution defined in [19, p.15]. Suppose that α and β are positive constants. A probability measure $\mu_{\alpha,\beta}$ on $\big(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)\big)$ is called a Bessel distribution with parameters α and β if

$$\mu_{\alpha,\beta}(dx) = e^{-\alpha}\delta_0(dx) + \beta e^{-\alpha-\beta x}\sqrt{\frac{\alpha}{\beta x}} \cdot I_1(2\sqrt{\alpha\beta x})dx, \tag{3.12}$$

where δ_0 is the Dirac measure at the origin and I_1 is the modified Bessel function of the first kind, namely,

$$I_1(r) = \frac{r}{2}\sum_{k=0}^{\infty}\frac{\left(\frac{1}{4}r^2\right)^k}{k!(k+1)!}, \qquad r \in \mathbb{R}.$$

Now we consider mixtures of Bessel distributions. Let $\gamma > 0$ be a constant and define a probability measure $m_{\alpha,\beta,\gamma}$ on $\mathbb{R}_+$ as follows:

$$m_{\alpha,\beta,\gamma}(dx) := \int_0^\infty \mu_{\alpha t,\beta}(dx)\frac{t^{\gamma-1}}{\Gamma(\gamma)}e^{-t}dt.$$

Similar to [19], we can easily calculate the characteristic function of $\mu_{\alpha,\beta}$ and $m_{\alpha,\beta,\gamma}$.

Lemma 4. *For $u \in \mathcal{U}$ we have:*

(*i*) $\displaystyle\int_0^\infty e^{ux}\mu_{\alpha,\beta}(dx) = e^{\frac{\alpha u}{\beta-u}}.$

(*ii*) $\displaystyle\int_0^\infty e^{ux}m_{\alpha,\beta,\gamma}(dx) = \Big(\frac{1}{\alpha+1} + \frac{\alpha}{\alpha+1}\cdot\frac{1}{1-\frac{\alpha+1}{\beta}\cdot u}\Big)^\gamma.$

Proof. (i) If $u \in \mathcal{U}$, then

$$\begin{aligned}\int_0^\infty e^{ux}\mu_{\alpha,\beta}(dx) &= \int_0^\infty e^{ux}e^{-\alpha}\delta_0(dx) + \int_0^\infty \beta e^{ux}e^{-\alpha-\beta x}\sqrt{\frac{\alpha}{\beta x}} \cdot I_1(2\sqrt{\alpha\beta x})dx \\ &= e^{-\alpha} + e^{-\alpha}\int_0^\infty \beta e^{-\beta x} \cdot e^{ux}\sqrt{\frac{\alpha}{\beta x}}\big(\sqrt{\alpha\beta x}\big) \cdot \sum_{k=0}^\infty \frac{(\alpha\beta x)^k}{k!(k+1)!}dx \\ &= e^{-\alpha} + e^{-\alpha}\int_0^\infty \alpha\beta e^{(u-\beta)x} \cdot \sum_{k=0}^\infty \frac{(\alpha\beta x)^k}{k!(k+1)!}dx \\ &= e^{-\alpha} + \alpha\beta e^{-\alpha}\sum_{k=0}^\infty \int_0^\infty e^{(u-\beta)x}\frac{(\alpha\beta x)^k}{k!(k+1)!}dx \\ &= e^{-\alpha} + e^{-\alpha}\sum_{k=0}^\infty \Big(\frac{\alpha\beta}{\beta-u}\Big)^{k+1} \cdot \frac{1}{(k+1)!} \\ &= e^{-\alpha}\sum_{k=0}^\infty \Big(\frac{\alpha\beta}{\beta-u}\Big)^{k} \cdot \frac{1}{k!} = e^{-\alpha} \cdot e^{\frac{\alpha\beta}{\beta-u}} = e^{\frac{\alpha u}{\beta-u}}.\end{aligned}$$

(ii) For $u \in \mathcal{U}$, we get

$$\begin{aligned}\int_0^\infty e^{ux}m_{\alpha,\beta,\gamma}(dx) &= \int_0^\infty \Big(\int_0^\infty e^{ux}\mu_{\alpha t,\beta}(dx)\Big)\frac{t^{\gamma-1}}{\Gamma(\gamma)}e^{-t}dt \\ &= \int_0^\infty e^{\alpha t\cdot\frac{u}{\beta-u}} \cdot \frac{t^{\gamma-1}}{\Gamma(\gamma)}e^{-t}dt = \Big(1+\alpha\cdot\frac{u}{u-\beta}\Big)^{-\gamma} \\ &= \Big(\frac{-\beta+(\alpha+1)u}{-\beta+u}\Big)^{-\gamma} = \Big(\frac{-\beta+u}{-\beta+(\alpha+1)u}\Big)^{\gamma} \\ &= \Big(\frac{\frac{1}{\alpha+1}((\alpha+1)u-\beta)+\frac{\beta}{\alpha+1}-\beta}{-\beta+(\alpha+1)u}\Big)^{-\gamma} \\ &= \Big(\frac{1}{\alpha+1}+\frac{\alpha}{\alpha+1}\cdot\frac{1}{1-\frac{\alpha+1}{\beta}\cdot u}\Big)^{\gamma}. \qquad (3.13)\end{aligned}$$

□

Lemma 5. *(i) The measure $m_{\alpha,\beta,\gamma}$ can be represented as follows:*

$$m_{\alpha,\beta,\gamma}(dx) = \Big(\frac{1}{1+\alpha}\Big)^{\gamma}\delta_0 + g_{\alpha,\beta,\gamma}(x)dx, \quad x \geq 0, \tag{3.14}$$

where

$$g_{\alpha,\beta,\gamma}(x) := \sum_{k=1}^\infty \frac{\alpha^k\Gamma(k+\gamma)}{(\alpha+1)^{k+\gamma}\Gamma(\gamma)k!}\Gamma(x;k,\beta), \quad x \geq 0, \tag{3.15}$$

and $\Gamma(x;k,\beta)$ *denotes the density function of the Gamma distribution with parameters* k *and* β.
(ii) The function $g_{\alpha,\beta,\gamma}(x)$ *defined in (3.15) is a continuous function with variables* $(\alpha,\beta,\gamma,x) \in D := (0,\infty)\times(0,\infty)\times(0,\infty)\times[0,\infty)$.

Proof. (i) We can write

$$\begin{aligned}
m_{\alpha,\beta,\gamma}(dx) &= \int_0^\infty \mu_{\alpha t,\beta}(dx)\frac{t^{\gamma-1}}{\Gamma(\gamma)}e^{-t}dt \\
&= \int_0^\infty \Big(e^{-\alpha t}\delta_0(dx) + \beta e^{-\alpha t-\beta x}\sqrt{\frac{\alpha t}{\beta x}}\cdot I_1(2\sqrt{\alpha t\beta x})dx\Big)\frac{t^{\gamma-1}}{\Gamma(\gamma)}e^{-t}dt \\
&= \Big(\frac{1}{1+\alpha}\Big)^\gamma\delta_0(dx) + \int_0^\infty \alpha\beta e^{-\alpha t-\beta x}\sum_{k=0}^\infty\frac{(\alpha t\beta x)^k}{k!(k+1)!}\cdot\frac{t^\gamma}{\Gamma(\gamma)}e^{-t}dtdx \\
&= \Big(\frac{1}{1+\alpha}\Big)^\gamma\delta_0(dx) + \alpha\beta e^{-\beta x}\cdot\sum_{k=0}^\infty\Big(\int_0^\infty(\alpha\beta x)^k\frac{e^{-(\alpha+1)t}t^{\gamma+k}}{\Gamma(\gamma)k!(k+1)!}dt\Big)dx \\
&= \Big(\frac{1}{1+\alpha}\Big)^\gamma\delta_0(dx) + \sum_{k=0}^\infty\frac{\alpha^{k+1}\Gamma(k+\gamma+1)}{(\alpha+1)^{k+\gamma+1}\Gamma(\gamma)(k+1)!}\Gamma(x;k+1,\beta)dx \\
&= \Big(\frac{1}{1+\alpha}\Big)^\gamma\delta_0(dx) + \sum_{k=1}^\infty\frac{\alpha^{k}\Gamma(k+\gamma)}{(\alpha+1)^{k+\gamma}\Gamma(\gamma)k!}\Gamma(x;k,\beta)dx.
\end{aligned}$$

(ii) By the definition of $g_{\alpha,\beta,\gamma}(x)$ we have

$$\begin{aligned}
g_{\alpha,\beta,\gamma}(x) &= \int_0^\infty \beta e^{-\alpha t-\beta x}\sqrt{\frac{\alpha t}{\beta x}}\cdot I_1(2\sqrt{\alpha t\beta x})\frac{t^{\gamma-1}}{\Gamma(\gamma)}e^{-t}dt \\
&= \int_0^\infty \alpha\beta e^{-\alpha t-\beta x}\Big(\sum_{k=0}^\infty\frac{(\alpha t\beta x)^k}{k!(k+1)!}\Big)\frac{t^\gamma}{\Gamma(\gamma)}e^{-t}dt \\
&= \int_0^\infty \frac{\alpha\beta t^\gamma}{\Gamma(\gamma)}e^{-(\alpha+1)t-\beta x}\Big(\sum_{k=0}^\infty\frac{(\alpha t\beta x)^k}{k!(k+1)!}\Big)dt
\end{aligned}$$

Suppose that $(\alpha_0,\beta_0,\gamma_0,x_0) \in D$ and $\delta > 0$ is small enough such that $\gamma_0-\delta>0$, $\alpha_0-\delta>0$ and $\beta_0-\delta>0$. Then for $(\alpha,\beta,\gamma,x)\in K_\delta$ with

$$K_\delta := \{(\alpha,\beta,\gamma,x)\in D : \max\{|\alpha-\alpha_0|,|\beta-\beta_0|,|\gamma-\gamma_0|,|x-x_0|\}\le\delta\}$$

we get

$$\begin{aligned}\frac{\alpha\beta t^{\gamma}}{\Gamma(\gamma)}e^{-(\alpha+1)t-\beta x}\Big(\sum_{k=0}^{\infty}\frac{(\alpha t\beta x)^k}{k!(k+1)!}\Big) \leq & \frac{\alpha\beta t^{\gamma}}{\Gamma(\gamma)}e^{-(\alpha+1)t-\beta x}\Big(\sum_{k=0}^{\infty}\frac{(\alpha t)^k(\beta x)^k}{(k!)^2}\Big) \\ \leq & \frac{\alpha\beta t^{\gamma}}{\Gamma(\gamma)}e^{-(\alpha+1)t-\beta x}\cdot e^{\alpha t}e^{\beta x} \leq \frac{\alpha\beta t^{\gamma}}{\Gamma(\gamma)}e^{-t} \\ \leq & c_\delta\big(t^{\gamma_0-\delta}1_{[0,1]}(t)+t^{\gamma_0+\delta}e^{-t}1_{(1,\infty)}(t)\big) \end{aligned} \tag{3.16}$$

for some constant $c_\delta > 0$, since $\frac{\alpha\beta}{\Gamma(\gamma)}$ is continuous and thus bounded for $(\alpha,\beta,\gamma,x) \in K_\delta$. If $(\alpha_n,\beta_n,\gamma_n,x_n) \to (\alpha_0,\beta_0,\gamma_0,x_0)$ as $n \to \infty$, then by dominated convergence we get

$$\lim_{n\to\infty} g_{\alpha_n,\beta_n,\gamma_n}(x_n) = g_{\alpha_0,\beta_0,\gamma_0}(x_0),$$

namely $g_{\alpha,\beta,\gamma}(x)$ is a continuous function on D. □

Remark 5. If we write $\delta_0 = \Gamma(0,\beta)$, namely considering the Dirac measure δ_0 as a degenerated Gamma distribution, then the representation in (3.14) shows that the measure $m_{\alpha,\beta,\gamma}$ is a mixture of Gamma distributions $\Gamma(k,\beta),\ k \in \mathbb{Z}_+$, namely

$$m_{\alpha,\beta,\gamma} = \Big(\frac{1}{1+\alpha}\Big)^{\gamma}\Gamma(0,\beta) + \sum_{k=1}^{\infty}\frac{\alpha^k\Gamma(k+\gamma)}{(\alpha+1)^{k+\gamma}\Gamma(\gamma)k!}\Gamma(k,\beta).$$

3.3 Transition density of the BAJD

In this section we shall derive a closed form expression for the transition density of the BAJD. We should mention that in [13, Chapter 7] the density functions of the pricing semigroup associated to the BAJD was derived for some special cases. Essentially, the method used in [13] could be used to derive the density functions of the BAJD in the case where $c/(a-\sigma^2 d/2) \in \mathbb{Z}$. Here we proceed like [13] but deal with more general parameters. In order to do this, we first find, by using the results of the previous section, a probability measure ν_t on $\mathbb{R}_+$ whose characteristic function satisfies (3.10).

We recall that the BAJD process $X = (X_t)_{t\geq 0}$ is given by (3.1). We distinguish between three cases according to the sign of $\Delta := a - \sigma^2 d/2$.

3.3.1 Case i): $\Delta > 0$

From (3.7) we know that

$$E_x[e^{uX_t}] = \big(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\big)^{-\frac{2a\theta}{\sigma^2}} \cdot \exp\Big(\frac{xue^{-at}}{1-\frac{\sigma^2}{2a}u(1-e^{-at})}\Big)$$
$$\cdot\Big(\frac{d - \frac{\sigma^2 du}{2a} + \big(\frac{\sigma^2 d}{2a} - 1\big)ue^{-at}}{d-u}\Big)^{\frac{c}{a-\frac{\sigma^2 d}{2}}} \tag{3.17}$$

The product of the first two terms on the right-hand side of (3.17) coincides with the characteristic function of the CIR process $Z = (Z_t)_{t\geq 0}$ defined in (1.3). It is well-known that the transition density function of the CIR process is given by

$$f(t,x,y) = \rho e^{-u-v}\Big(\frac{v}{u}\Big)^{\frac{q}{2}} I_q\big(2(uv)^{\frac{1}{2}}\big) \tag{3.18}$$

for $t > 0, x > 0$ and $y \geq 0$, where

$$\rho \equiv \frac{2a}{\sigma^2\Big(1-e^{-at}\Big)}, \qquad u \equiv \rho x e^{-at},$$
$$v \equiv \rho y, \qquad q \equiv \frac{2a\theta}{\sigma^2} - 1,$$

and $I_q(\cdot)$ is the modified Bessel function of the first kind of order q, namely

$$I_q(r) = \big(\frac{r}{2}\big)^q \sum_{k=0}^{\infty} \frac{\big(\frac{1}{4}r^2\big)^k}{k!\Gamma(q+k+1)}, \qquad r > 0.$$

We should remark that for $x = 0$ the formula of the density function $f(t,x,y)$ given in (3.18) is not valid any more. In this case we have

$$f(t,0,y) = \frac{\rho}{\Gamma(q+1)} v^q e^{-v} \tag{3.19}$$

for $t > 0$ and $y \geq 0$.

Thus

$$\int_{\mathbb{R}_+} f(t,x,y)e^{uy}dy = \big(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\big)^{-\frac{2a\theta}{\sigma^2}} \cdot \exp\Big(\frac{xue^{-at}}{1-\frac{\sigma^2}{2a}u(1-e^{-at})}\Big).$$

Now we want to find a probability measure ν_t with

$$\begin{aligned}\int_{\mathbb{R}_+} e^{uy}\nu_t(dy) =&\Big(\frac{d-\frac{\sigma^2 du}{2a}+\big(\frac{\sigma^2 d}{2a}-1\big)ue^{-at}}{d-u}\Big)^{\frac{c}{a-\frac{\sigma^2 d}{2}}}\\ =&\Big(\frac{d-uL_1(t)}{d-u}\Big)^{\frac{c}{a-\frac{\sigma^2 d}{2}}}\\ =&\Big(L_1(t)+\big(1-L_1(t)\big)\frac{1}{1-\frac{u}{d}}\Big)^{\frac{c}{a-\frac{\sigma^2 d}{2}}},\end{aligned} \tag{3.20}$$

where $L_1(t) := \exp(-at) + \sigma^2 d\big(1-\exp(-at)\big)/(2a)$. If such a measure ν_t exists, then the law of X_t can be written as the convolution of the law of Z_t and ν_t.

Comparing the characteristic functions (3.13) and (3.20), it is easy to see that we can seek the measure ν_t as a mixture of Bessel distributions. More precisely, we define

$$\nu_t := m_{\alpha_1(t),\beta_1(t),\gamma_1} \tag{3.21}$$

with

$$\begin{cases}\alpha_1(t) := \frac{1}{L_1(t)}-1\\ \beta_1(t) := \frac{d}{L_1(t)}\\ \gamma_1 := \frac{c}{a-\frac{\sigma^2 d}{2}}.\end{cases} \tag{3.22}$$

Then the characteristic function of ν_t coincides with (3.20). Since the probability measure $m_{\alpha_1(t),\beta_1(t),\gamma_1}$ is of the form (3.14), it follows now from (3.9), (3.17) and (3.20) that the law of X_t is absolutely continuous with respect to the Lesbegue measure and its density function $p(t,x,y)$ is given by

$$p(t,x,y) = \Big(\frac{1}{1+\alpha_1(t)}\Big)^{\gamma_1} f(t,x,y) + \int_0^y f(t,x,y-z)g_{\alpha_1(t),\beta_1(t),\gamma_1}(z)dz \tag{3.23}$$

for $t>0,\ x\geq 0$ and $y\geq 0$, where the function g is defined in (3.15).

3.3.2 Case ii): $\Delta < 0$

Similar to the case (i), it suffices to find a probability measure ν_t with

$$\begin{aligned}\int_{\mathbb{R}_+} e^{uy}\nu_t(dy) &= \Big(\frac{d-\frac{\sigma^2 du}{2a}+\big(\frac{\sigma^2 d}{2a}-1\big)ue^{-at}}{d-u}\Big)^{\frac{c}{a-\frac{\sigma^2 d}{2}}}\\ &= \Big(\frac{d-u}{d-\frac{\sigma^2 du}{2a}+\big(\frac{\sigma^2 d}{2a}-1\big)ue^{-at}}\Big)^{\frac{-c}{a-\frac{\sigma^2 d}{2}}}\\ &= \Big(\frac{d-u}{d-L_1(t)u}\Big)^{\frac{-c}{a-\frac{\sigma^2 d}{2}}}\\ &= \Bigg(\frac{1}{L_1(t)}+(1-\frac{1}{L_1(t)})\cdot\frac{1}{1-\frac{L_1(t)u}{d}}\Bigg)^{\frac{-c}{a-\frac{\sigma^2 d}{2}}}. \end{aligned} \tag{3.24}$$

Since $\Delta = a - \sigma^2 d/2 < 0$, therefore $\sigma^2 d/2a > 1$ and

$$L_1(t) = e^{-at} + \frac{\sigma^2 d}{2a}\cdot\big(1-e^{-at}\big) > 1.$$

According to the formula (3.12), we can choose

$$\nu_t = m_{\alpha_2(t),\beta_2(t),\gamma_2}$$

with the parameters α_2, β_2 and γ_2 defined by

$$\begin{cases}\alpha_2(t) := L_1(t)-1\\ \beta_2 := d\\ \gamma_2 := \frac{-c}{a-\frac{\sigma^2 d}{2}}.\end{cases} \tag{3.25}$$

Similar to the case (i), the transition densities $p(t,x,y)$ of X is given by

$$p(t,x,y) = \Big(\frac{1}{1+\alpha_2(t)}\Big)^{\gamma_2} f(t,x,y) + \int_0^y f(t,x,y-z)g_{\alpha_2(t),\beta_2,\gamma_2}(z)dz \tag{3.26}$$

for $t>0,\ x\geq 0$ and $y \geq 0$, where the function g is defined in (3.15).

3.3.3 Case iii): $\Delta = 0$

In this case we need to find a probability measure ν_t with

$$\int_{\mathbb{R}_+} e^{uy}\nu_t(dy) = \exp\Big(\frac{cu(1-e^{-at})}{a(d-u)}\Big).$$

According to the formula (3.14) we can take ν_t as a Bessel distribution $\mu_{\alpha_3(t),\beta_3}$ with the parameters $\alpha_3(t)$ and β_3 defined by

$$\begin{cases} \alpha_3(t) := \frac{c}{a}(1 - e^{-at}) \\ \beta_3 := d. \end{cases} \tag{3.27}$$

Thus in this case the transition densities $p(t, x, y)$ of X is given by

$$\begin{aligned} p(t, x, y) = \int_0^y f(t, x,y - z)\beta_3 e^{-\alpha_3(t)-\beta_3 z}\sqrt{\frac{\alpha_3(t)}{\beta_3 z}} I_1(2\sqrt{\alpha_3(t)\beta_3 z})dz \\ +e^{-\alpha_3(t)} f(t, x, y) \end{aligned} \tag{3.28}$$

for $t > 0,\ x \geq 0$ and $y \geq 0$.

Summarizing the above three cases we get the following theorem.

Theorem 6. *Let $X = (X_t)_{t\geq 0}$ be the BAJD defined in (3.1). Then the law of X_t given that $X_0 = x \geq 0$ is absolutely continuous with respect to the Lesbegue measure and thus possesses a density function $p(t, x, y)$, namely*

$$P_x(X_t \in A) = \int_A p(t, x, y)dy, \quad t \geq 0,\ A \in \mathcal{B}(\mathbb{R}_+)$$

According to the sign of $\Delta = a - \sigma^2 d/2$, the density $p(t, x, y)$ is given by (3.23), (3.26) and (3.28) respectively.

Although the density functions in (3.23), (3.26) and (3.28) are essentially different, they do share some similarities. In the following corollary we give a unified representation of $p(t, x, y)$.

Corollary 2. *Irrelevant of the the sign of $\Delta = a - \sigma^2 d/2$, the transition densities $p(t, x, y)$ of X can be expressed in a unified form as*

$$p(t, x, y) = L(t)f(t, x, y) + \int_0^y f(t, x, y - z)h(t, z)dz, \tag{3.29}$$

where $L(t)$ is continuous function in $t > 0$ which satisfies $0 < L(t) < 1$ for $t > 0$, the function $h(t, z)$ is non-negative and continuous in $(t, z) \in (0, \infty) \times [0, \infty)$ and satisfies $\int_{\mathbb{R}_+} h(t, z)dz = 1 - L(t)$.

3.4 Positive Harris recurrence of the BAJD

It was shown in [9] (see also [36]) that the semigroup of any stochastically continuous affine process on the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$ is a Feller semigroup. Define the semigroup of the BAJD by

$$T_t f(x) := \int_{\mathbb{R}_+} p(t,x,y) f(y) dy, \tag{3.30}$$

where $f : \mathbb{R}_+ \to \mathbb{R}$ is bounded. Since the BAJD process X is stochastically continuous and affine, thus $(T_t)_{t\geq 0}$ is a Feller semigroup.

To show the positive Harris recurrence, we need first to prove the regularity property of BAJD. To this aim, we first analyse the continuity properties of the integral which appears on the right hand side of (3.29).

Lemma 6. *Let $f(t,x,y)$ be the transition density of the CIR process given in (3.18) and $h(t,z)$ be the same as in (3.29). Then the function $F(t,x,y)$ defined by*

$$F(t,x,y) := \int_0^y f(t,x,y-z) h(t,z) dz \tag{3.31}$$

is continuous with variables $(t,x,y) \in (0,\infty) \times [0,\infty) \times [0,\infty)$. Moreover if $M > 1$ is a constant, then

$$|F(t,x,y)| \leq C y^{\frac{2a\theta}{\sigma^2}} \tag{3.32}$$

for all

$$(t,x,y) \in K_M := \{(t,x,y) : \frac{1}{M} \leq t \leq M,\ 0 \leq x \leq M,\ 0 \leq y \leq \frac{1}{M}\},$$

where $C > 0$ is a constant which depends on M.

Proof. For simplicity we set $q := 2a\theta/\sigma^2 - 1$ as in (3.18). Since $h(t,z)$ is continuous in $(t,z) \in (0,\infty) \times [0,\infty)$, thus there exists a constant $c_1 > 0$ depending on M such that

$$|h(t,z)| \leq c_1 \quad \text{for} \quad \frac{1}{M} \leq t \leq M, \quad 0 \leq z \leq \frac{1}{M}. \tag{3.33}$$

Therefore if $(t,x,y) \in K_M$, we have

$$|F(t,x,y)| \leq c_1 \int_0^y f(t,x,y-z) dz. \tag{3.34}$$

According to (3.18) and (3.19) we have

$$|f(t,x,y-z)| \le c_2|y-z|^q \qquad \text{if} \quad (t,x,y-z) \in K_M \text{ and } y \neq z, \tag{3.35}$$

where $c_2 > 0$ is a constant depending on M.
It follows from (3.34) and (3.35) that

$$|F(t,x,y)| \le c_1c_2 \int_0^y |y-z|^q dz = \frac{c_1c_2}{q+1} y^{q+1} \le c_3 y^{q+1} \tag{3.36}$$

for $(t,x,y) \in K_M$, if we set $c_3 := c_1c_2/(q+1)$. Thus (3.32) is proved. Noting that $F(t,x,0) = 0$ for $t > 0$ and $x \ge 0$, the continuity of the function F at points $(t,x,0)$ is an immediate consequence of the estimate (3.36).
We now proceed to prove the continuity of F at other points. Suppose that $t_0 > 0$, $x_0 \ge 0$ and $y_0 > 0$ are fixed. Let $\epsilon > 0$ be arbitrary. We choose $\delta_1 > 0$ small enough such that $y_0 - 2\delta_1 > 0$ and $t_0 - \delta_1 > 0$. As in (3.33) and (3.35) there exist constants $c_4, c_5 > 0$, which depend on δ_1, such that

$$|h(t,z)| \le c_4 \quad \text{for} \quad t \in [t_0-\delta_1, t_0+\delta_1], \;\; z \in [0, y_0+\delta_1] \tag{3.37}$$

and

$$|f(t,x,y-z)| \le c_5|y-z|^q \tag{3.38}$$

$for\ t \in [t_0-\delta_1, t_0+\delta_1],\ x \in [0, x_0+\delta_1]$ and $0 < y-z \le y_0+\delta_1$. Set

$$K_{\delta_2} := [t_0-\delta_2, t_0+\delta_2] \times [0, x_0+\delta_2] \times [y_0-\delta_2, y_0+\delta_2].$$

We choose $\delta_2 > 0$ small enough such that $\delta_2 < \delta_1$ and $c_4c_5(3\delta_2)^{q+1}/(q+1) < \epsilon/3$. If $(t,x,y) \in K_{\delta_2}$ then it holds

$$\begin{aligned} \Big|\int_{y_0-2\delta_2}^{y} f(t,x,y-z)h(t,z)dz\Big| &\le c_4c_5 \int_{y_0-2\delta_2}^{y} (y-z)^q dz \\ &= \frac{c_4c_5}{q+1}(y-y_0+2\delta_2)^{q+1} \\ &\le \frac{c_4c_5}{q+1}(3\delta_2)^{q+1} < \frac{\epsilon}{3}. \end{aligned} \tag{3.39}$$

If $(t,x,y) \in K_{\delta_2}$ and $0 \le z \le y_0 - 2\delta_2$, then $\delta_2 \le y - z \le y_0 + \delta_2$ and by (3.37) and (3.38) we have

$$\begin{aligned} |f(t,x,y-z)h(t,z)| &\le c_4c_5|y-z|^q \\ &\le c_4c_5(|\delta_2|^q + |y_0+\delta_2|^q). \end{aligned} \tag{3.40}$$

Since for fixed $z \in [0, y_0 - 2\delta_2]$ the function $f(t, x, y-z)h(t, z)$ is continuous in $(t, x, y) \in K_{\delta_2}$, it follows from (3.40) and dominated convergence theorem that

$$F_2(t, x, y) := \int_0^{y_0 - 2\delta_2} f(t, x, y - z)h(t, z)dz$$

is a continuous function in $(t, x, y) \in K_{\delta_2}$. This implies the existence of a constant δ with $0 < \delta < \delta_2$ such that

$$|F_2(t, x, y) - F_2(t_0, x_0, y_0)| < \frac{\epsilon}{3}, \tag{3.41}$$

if $(t, x, y) \in K_\delta := [t_0 - \delta, t_0 + \delta] \times [0 \vee (x_0 - \delta), x_0 + \delta] \times [y_0 - \delta, y_0 + \delta]$. Thus it follows from (3.39) and (3.41) that

$$|F(t, x, y) - F(t_0, x_0, y_0)| < \epsilon$$

for $(t, x, y) \in K_\delta$. The continuity of the function F at (t_0, x_0, y_0) is proved. $\square$

According to (3.29) and Lemma 6, the boundary behavior of the transition densities $p(t, x, y)$ of X at $y = 0$ depends very much on the behavior of $f(t, x, y)$ at $y = 0$. As shown in the proof of proposition 2 of chapter 2, if $2a\theta/\sigma^2 < 1$, then

$$f(t, x, 0) := \lim_{y \to 0} f(t, x, y) = \infty, \quad \forall x \in \mathbb{R}_+,$$

which means that $(t, x, y) \mapsto f(t, x, y)$ is not continuous on $(0, \infty) \times \mathbb{R}_+^2$; if $2a\theta/\sigma^2 > 1$, then

$$f(t, x, 0) := \lim_{y \to 0} f(t, x, y) = 0.$$

Therefore, in both cases the behavior of $f(t, x, y)$ at the boundary $y = 0$ violates the regularity condition; as a consequence, the transition densities $p(t, x, y)$ of X are also not regular. To overcome this difficulty, we proceed as in the proof of proposition 2 of chapter 2 and define a new measure η on $\big(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)\big)$ as

$$\eta(dx) := \kappa(x)dx, \tag{3.42}$$

where

$$\kappa(x) = \begin{cases} x^{\frac{2a\theta}{\sigma^2} - 1}, & 0 \le x \le 1, \\ 1, & x > 1. \end{cases}$$

Then the transition densities of the BAJD process with respect to the new measure η is given by

$$\tilde{p}(t, x, y) = \frac{p(t, x, y)}{\kappa(y)}, \quad t > 0,\ x \ge 0,\ y > 0. \tag{3.43}$$

Theorem 7. *The transition densities $\tilde{p}(t,x,y)$ of the BAJD process with respect to the measure η satisfies*

$$0 < \tilde{p}(t,x,y) < \infty, \quad t > 0,\ x \geq 0,\ y \geq 0$$

and is continuous in $(t,x,y) \in (0,\infty) \times [0,\infty) \times [0,\infty)$. Thus the BAJD is a regular Feller process on the state space $\mathbb{R}_+$.

For the definition of regular continuous-time Markov process we refer to the definition (5) of Chapter 2.

Proof. From (3.29) we can write the transition density of the BAJD process with respect to the measure η as

$$\begin{aligned} \tilde{p}(t,x,y) &= L(t)\frac{f(t,x,y)}{\kappa(y)} + \frac{F(t,x,y)}{\kappa(y)} \\ &= L(t)\tilde{f}(t,x,y) + \frac{F(t,x,y)}{\kappa(y)} \end{aligned} \tag{3.44}$$

where $\tilde{f}(t,x,y)$ is the transition density of CIR with respect the new measure η and F is defined in (3.31). We have already shown in Proposition 2 of Chapter 2 that

$$0 < \tilde{f}(t,x,y) < \infty, \quad t > 0,\ x \geq 0,\ y \geq 0, \tag{3.45}$$

and $(t,x,y) \mapsto \tilde{f}(t,x,y)$ is a continuous function on $(0,\infty) \times \mathbb{R}_+^2$.
From Lemma 6, we know that the function $F(t,x,y)$ appearing in the second summand in (3.44) is continuous on $(0,\infty) \times \mathbb{R}_+^2$ and

$$|F(t,x,y)| \leq Cy^{\frac{2a\theta}{\sigma^2}}, \qquad \text{if } (t,x,y) \in K_M,$$

where $C > 0$ is a constant depending on M. Now it is clear that

$$0 \leq \lim_{y\to 0} \frac{F(t,x,y)}{\kappa(y)} \leq \lim_{y \to 0} C|y| = 0.$$

Since $F(t,x,0) = 0$ and $L(t)$ is continuous in $t > 0$, it follows that the function $\tilde{p}(t,x,y)$ is continuous at points $(t_0, x_0, 0)$, if $t_0 > 0$ and $x_0 \geq 0$. The continuity of the function $\tilde{p}(t,x,y)$ at other points is also clear, because all the functions appearing in (3.44) are continuous and $0 < \kappa(y) < \infty$ for $y > 0$. Noting that $0 < L(t) < 1$ for $t > 0$, we get

$$0 < \tilde{p}(t,x,y) < \infty \quad \text{for all } (t,x,y) \in (0,\infty) \times \mathbb{R}_+^2.$$

Therefore the BAJD process is a regular Feller process with η as the supporting measure. □

Let us recall the following definitions from Chapter 2, we refer to [45, p.490] and [29, p.405].

Definition 9. *Consider a time-homogeneous Markov process $Y = (Y_t)_{t\geq 0}$ with the state space $\mathbb{R}_+$ and distributions P_x, $x \in \mathbb{R}_+$.*
(i) Y is said to be Harris recurrent if for some σ-finite measure μ

$$P_x\Big(\int_0^\infty \mathbf{1}_A(Y_s)ds = \infty\Big) = 1, \tag{3.46}$$

for any $x \in \mathbb{R}_+$ and $A \in \mathcal{B}(\mathbb{R}_+)$ with $\mu(A) > 0$. It was shown in [16] that if Y is Harris recurrent then it possesses a unique (up to a renormalization) invariant measure. If the invariant measure is finite, then the process Y is called positive Harris recurrent.
(ii) Y is said to be uniformly transient if

$$\sup_x E_x\Big[\int_0^\infty \mathbf{1}_K(Y_s)ds\Big] < \infty \tag{3.47}$$

for every compact $K \subset \mathbb{R}_+$.

Lemma 7. *The BAJD is not uniformly transient.*

Proof. Let $m > 0$, $K := [0, m]$ and $x \in (0, \infty)$ be fixed. Then

$$\begin{aligned} E_x\Big[\int_0^\infty \mathbf{1}_{[0,m]}(X_t)dt\Big] &= \int_0^\infty E_x\big[\mathbf{1}_{[0,m]}(X_t)\big]dt \\ &= \int_0^\infty \int_0^m p(t,x,y)dydt \\ &= \int_0^m dy \int_0^\infty p(t,x,y)dt. \end{aligned}$$

From (3.29) and (3.31) we know that

$$p(t,x,y) = L(t)f(t,x,y) + F(t,x,y).$$

It follows from (3.23), (3.26) and (3.28) that

$$0 < \lim_{t\to\infty} L(t) = \begin{cases} \left(\frac{\sigma^2 d}{2a}\right)^{\frac{c}{a-\frac{\sigma^2 d}{2}}}, & \text{if } \Delta \neq 0, \\ e^{-\frac{c}{a}}, & \text{if } \Delta = 0. \end{cases} \tag{3.48}$$

Since $F(t,x,y)$ is non-negative, thus there exists large enough $T > 0$ such that

$$p(t,x,y) \geq \lambda f(t,x,y) \quad \text{for } t \geq T,$$

where $\lambda > 0$ is a constant. Let $\epsilon > 0$ be small enough. According to Lemma 1 of Chapter 2 we know that for any $x > 0$ and $y \in [\epsilon, m]$ it holds

$$\int_0^\infty f(t,x,y)dt = \infty.$$

Therefore

$$\begin{aligned} E_x\Big[\int_0^\infty \mathbf{1}_{[0,m]}(X_t)dt\Big] &= \int_0^m dy \int_0^\infty p(t,x,y)dt \\ &\geq \int_\epsilon^m dy \int_R^\infty \lambda f(t,x,y)dt = \infty. \end{aligned}$$

This proves that the BAJD is not uniformly transient. □

Theorem 8. *The BAJD is Harris recurrent.*

Proof. We have shown that the BAJD is a regular Feller process with the measure η, which is defined in (3.42), as a supporting measure. It follows from Lemma 7 and [29, Theorem 20.17] that the BAJD is Harris recurrent and the measure η satisfies (3.46). □

Remark 9. Since the BAJD is a Harris recurrent Feller process, it follows from [29, Theorem 20.18] that the BAJD has a locally finite invariant measure π which is equivalent to the supporting measure η. Moreover every σ -finite invariant measure of the BAJD agrees with π up to a renormalization. The existence and uniqueness of an invariant probability measure for the BAJD has already been proved in [32] (see also [38]). Thus we can assume π to be a probability measure. The characteristic function of π was given in [38] and has the form

$$(3.49)\qquad \int_{\mathbb{R}_+} e^{uz}\pi(dz) = \begin{cases} \big(1-\frac{\sigma^2}{2a}u\big)^{-\frac{2a\theta}{\sigma^2}} \cdot \Big(\frac{d-\frac{\sigma^2 du}{2a}}{d-u}\Big)^{\frac{c}{a-\frac{\sigma^2 d}{2}}}, & \text{if } \Delta \neq 0, \\ \big(1-\frac{\sigma^2}{2a}u\big)^{-\frac{2a\theta}{\sigma^2}} \cdot \exp\Big(\frac{cu}{a(d-u)}\Big), & \text{if } \Delta = 0. \end{cases}$$

Corollary 3. *The BAJD is positive Harris recurrent. Its unique invariant probability measure π is absolute continuous with respect to the Lebesgue measure and thus has a density function $l(\cdot)$, namely $\pi(dy) = l(y)dy,\ y \in \mathbb{R}_+$. If $\Delta \neq 0$, then we have*

$$l(y) = \Big(\frac{\sigma^2 d}{2a}\Big)^{\frac{c}{a-\frac{\sigma^2 d}{2}}}\Gamma\big(y;\frac{2a\theta}{\sigma^2},\frac{\sigma^2}{2a}\big) + \int_0^y \Gamma\big(y-z;\frac{2a\theta}{\sigma^2},\frac{\sigma^2}{2a}\big)h(z)dz, \quad y \geq 0,$$

where $\Gamma\big(y;2a\theta/\sigma^2,\sigma^2/(2a)\big)$ denotes the density function of the Gamma distribution with parameters $2a\theta/\sigma^2$ and $\sigma^2/(2a)$, and

$$h(z) = \begin{cases} g_{\frac{2a}{\sigma^2 d}-1,\frac{2a}{\sigma^2},\frac{c}{a-\frac{\sigma^2 d}{2}}}(z), & \textit{if } \Delta > 0, \\ g_{\frac{\sigma^2 d}{2a}-1,d,-\frac{c}{a-\frac{\sigma^2 d}{2}}}(z), & \textit{if } \Delta < 0, \end{cases}$$

with g defined in (3.15). If $\Delta = 0$, then we have

$$l(y) = e^{-\frac{c}{a}}\Gamma\big(y; \frac{2a\theta}{\sigma^2}, \frac{\sigma^2}{2a}\big) + \int_0^y \Gamma\big(y-z; \frac{2a\theta}{\sigma^2}, \frac{\sigma^2}{2a}\big) de^{-\frac{c}{a}-dz}\sqrt{\frac{c}{adz}} \cdot I_1\big(2\sqrt{\frac{cdz}{a}}\big)dz$$

for $y \geq 0$.

Proof. We apply the same method which we used in Section 3.3 to find the transition densities of the BAJD. Since the characteristic function of π is given by (3.49) and noting that the first term on the right hand side of (3.49) corresponds to the characteristic function of a Gamma distribution, we can represent the measure π as a convolution of a Gamma distribution with a probability measure ν. If $\Delta = 0$ the measure ν is a Bessel distribution, otherwise it is a mixture of Bessel distributions. By identifying the parameters of the Gamma distribution and the Bessel or mixture of Bessel distributions, we get an explicit formula of the density function of π. □

Corollary 4. *Let $X = (X_t)_{t\geq 0}$ be the BAJD defined by (3.1). Then for any $f \in \mathcal{B}_b(\mathbb{R}_+)$ we have*

$$\frac{1}{t}\int_0^t f(X_s)ds \;\to\; \int_{\mathbb{R}_+} f(x)\pi(dx) \quad a.s.$$

as $t \to \infty$, where π is the unique invariant probability measure of the BAJD.

Proof. The above convergence follows from Corollary 3 and [29, Theorem 20.21]. □

3.5 Exponential ergodicity of the BAJD

Let $\|\cdot\|_{TV}$ denote the total-variation norm for signed measures on $\mathbb{R}_+$, namely

$$\|\mu\|_{TV} = \sup_{A\in\mathcal{B}(\mathbb{R}_+)} \{|\mu(A)|\}.$$

The total variation norm for signed measures on $\mathbb{R}_+$ is a special case of the norm $\|\cdot\|_h$, which is defined by

$$\|\mu\|_h = \sup_{|g|\leq h} \Big| \int_{\mathbb{R}_+} g d\mu \Big|$$

for a function h on $\mathbb{R}_+$ with $h \geq 1$. Obviously it holds $\|\mu\|_{TV} \leq \|\mu\|_h$, given that $h \geq 1$.

Let $P^t(x,\cdot) := P_x(X_t \in \cdot)$ be the distribution of the BAJD process X at time t given that $X_0 = x$ with $x \geq 0$.

Definition 10. *A continuous-time Markov process X is said to be h-exponentially ergodic if there exist a constant $\beta \in (0,1)$ and a finite-valued function $B(\cdot)$ such that*

$$\|P^t(x,\cdot) - \pi\|_h \leq B(x)\beta^t, \quad \forall t > 0, \quad x \in \mathbb{R}_+,$$

where π is the unique invariant probability measure of X.

In this section we will find a function $h \geq 1$ such that the BAJD is h-exponentially ergodic, where π is the unique invariant probability measure of the BAJD.

We first show the existence of a Foster-Lyapunov function, which is essential for the exponential ergodicity to hold.

Let $\mathcal{A}$ denote the extended generator of the BAJD and $D(\mathcal{A})$ the domain of $\mathcal{A}$. For the definitions of extended generator of Markov processes and its domain, the reader is referred to [46, p.521]. For $\lambda > 0$ define a function $f_\lambda : \mathbb{R}_+ \to \mathbb{R}$ by $f_\lambda(x) := \exp(-\lambda x),\ x \in \mathbb{R}_+$. Let $\Lambda := \{f_\lambda :\ \lambda > 0\}$ and denote the linear hull of Λ by $\mathcal{L}(\Lambda)$. Then we know from [13, Theorem 4.3] and [13, Section 7] (see also [14, Section 6]) that $\mathcal{L}(\Lambda) \subset D(\mathcal{A})$ and

$$\mathcal{A}g(x) = \frac{1}{2}\sigma^2 x g''(x) + (a\theta - ax)g'(x) + cd\int_{\mathbb{R}_+}\big(g(x+y) - g(x)\big)e^{-dy}dy$$

for $g \in \mathcal{L}(\Lambda)$.

Definition 11. *A function V belongs to $D(\mathcal{A})$ is called Foster-Lyapunov function if there exist constants $k, M \in (0,\infty)$ such that*

$$\mathcal{A}V(x) \leq -kV(x) + M, \quad \forall x \in \mathbb{R}_+.$$

Lemma 8. *The function $V(x) := \exp(\gamma x)$, $x \in \mathbb{R}_+$, with small enough $\gamma > 0$ belongs to $D(\mathcal{A})$. Moreover, V is a Foster-Lyapunov function for the BAJD, with $k = 1$ and*

$$M := e^{\gamma x_0} + \gamma e^{\gamma x_0} \cdot \big(a\theta + \frac{c}{d-\gamma}\big) < \infty.$$

Proof. Firstly, it is easy to see that ψ and ϕ defined in (3.5) and (3.6), respectively, are still solutions of (3.3) if we let $u = \gamma$ in (3.3), (3.5) and (3.6) with a small enough $\gamma > 0$. It follows from [34, Theorem 2.14] that (3.7) holds also for $u = \gamma$. In particular,

$$E_x[\exp(\gamma X_t)] = \exp\big(\phi(t,\gamma) + x\psi(t,\gamma)\big) < \infty, \quad \forall x \in \mathbb{R}_+,\ t \geq 0. \tag{3.50}$$

By Itô's formula and using (3.50), we can easily verify that

$$V(X_t) - V(x) - \int_0^t \big(\mathcal{A}V(X_s)\big)ds, \quad t \geq 0,$$

is a P_x-square-integrable martingale, where

$$\begin{aligned}\mathcal{A}V(x) =& \frac{1}{2}\sigma^2\gamma^2 x e^{\gamma x} + (a\theta - ax)\gamma e^{\gamma x} + cd\int_{\mathbb{R}_+}\big(e^{\gamma(x+y)} - e^{\gamma x}\big)e^{-dy}dy\\ =& \frac{1}{2}\sigma^2\gamma^2 x e^{\gamma x} + (a\theta - ax)\gamma e^{\gamma x} + \frac{c\gamma}{d-\gamma}e^{\gamma x}\\ =& \gamma e^{\gamma x}\cdot\Big(\big(\frac{1}{2}\sigma^2\gamma - a\big)x + a\theta + \frac{c}{d-\gamma}\Big).\end{aligned}$$

Hence

$$E_x[V(X_t)] - V(x) = E_x\Big[\int_0^t \big(\mathcal{A}V(X_s)\big)ds\Big], \quad t \geq 0.$$

We can find $C > 0$ and small $\epsilon > 0$ such that

$$\Big|\big(\frac{1}{2}\sigma^2\gamma - a\big)x + a\theta + \frac{c}{d-\gamma}\Big| \leq Ce^{\epsilon x}, \quad x \in \mathbb{R}_+.$$

By (3.3) with $u = \gamma + \epsilon$, we obtain

$$\begin{aligned}\int_0^t E_x[|\mathcal{A}V(X_s)|]ds \leq& C\gamma\int_0^t E_x\big[\exp\big((\gamma+\epsilon)X_s\big)\big]ds\\ =& C\gamma\int_0^t \exp\big(\phi(s,\gamma+\epsilon) + x\psi(s,\gamma+\epsilon)\big)ds < \infty.\end{aligned}$$

This verifies $V \in D(\mathcal{A})$.

If $\gamma > 0$ is small enough, then $\sigma^2\gamma/2 - a < 0$ and there exists $x_0 > 0$ with

$$\big(\frac{1}{2}\sigma^2\gamma - a\big)x_0 + a\theta + \frac{c}{d-\gamma} = -\frac{1}{\gamma}.$$

Thus we have for $x \in [x_0, \infty)$

$$\mathcal{A}V(x) = \gamma e^{\gamma x}\cdot\Big(\big(\frac{1}{2}\sigma^2\gamma - a\big)x + a\theta + \frac{c}{d-\gamma}\Big) \leq -e^{\gamma x}$$

and for $x \in [0, x_0]$

$$\mathcal{A}V(x) = \gamma e^{\gamma x}\cdot\Big(\big(\frac{1}{2}\sigma^2\gamma - a\big)x + a\theta + \frac{c}{d-\gamma}\Big) \leq \gamma e^{\gamma x_0}\cdot(a\theta + \frac{c}{d-\gamma}).$$

It follows for all $x \in \mathbb{R}_+$

$$\mathcal{A}V(x) \leq -e^{\gamma x} + e^{\gamma x_0} + \gamma e^{\gamma x_0}\cdot(a\theta + \frac{c}{d-\gamma}) \leq -V(x) + M$$

with

$$M := e^{\gamma x_0} + \gamma e^{\gamma x_0}\cdot\big(a\theta + \frac{c}{d-\gamma}\big) < \infty.$$

□

Lemma 9. *Let the constants γ, k and M, as well as $V(x) = \exp(\gamma x)$ be the same as in Lemma 8. Then the BAJD satisfies*

$$E_x[V(X_t)] \leq e^{-kt}V(x) + \frac{M}{k},$$

or equivalently

$$\int_{\mathbb{R}_+} V(y)p(t,x,y)dy \leq e^{-kt}V(x) + \frac{M}{k}$$

for all $x \in \mathbb{R}_+,\ t > 0$.

Proof. Let $V(x) = \exp(\gamma x)$ and $g(x,t) := V(x) \cdot \exp(kt) = \exp(\gamma x + kt)$, where the constants γ and k are the same as in Lemma 8. Then $g_x = \gamma \exp(\gamma x + kt)$, $g_{xx} = \gamma^2 \exp(\gamma x + kt)$ and $g_t = k \exp(\gamma x + kt)$. By applying Itô's formula and then taking the expectation, we get for all $x \in \mathbb{R}_+,\ t > 0$,

$$\begin{aligned}
&e^{kt}E_x[V(X_t)] - V(x)\\
=&E_x[g(X_t,t)] - E_x[g(X_0,0)]\\
=&E_x\Big[\int_0^t \big(e^{ks} \cdot \mathcal{A}V(X_s) + ke^{ks} \cdot V(X_s)\big)ds\Big]\\
\leq&E_x\Big[\int_0^t \big(e^{ks} \cdot \big(-kV(X_s) + M\big) + ke^{ks} \cdot V(X_s)\big)ds\Big]\\
=&\mathbb{E}_x\Big[\int_0^t Me^{ks}ds\Big] = \frac{M}{k}e^{kt} - \frac{M}{k} \leq \frac{M}{k}e^{kt}.
\end{aligned}$$

Thus for $x \in \mathbb{R}_+,\ t > 0$,

$$E_x[V(X_t)] \leq e^{-kt}V(x) + \frac{M}{k}.$$

□

Applying the main results of [46] and Lemma 9, we get the following theorem.

Theorem 10. *Let $h(x) := 1 + \exp(\gamma x)$ with the constant $\gamma > 0$ small enough. Then the BAJD is h-exponentially ergodic, namely there exist constants $\beta \in (0,1)$ and $C \in (0,\infty)$ such that*

$$\|P^t(x,\cdot) - \pi\|_h \leq C\big(e^{\gamma x} + 1\big)\beta^t, \quad t > 0, \quad x \in \mathbb{R}_+. \tag{3.51}$$

Proof. Basically, we follow the proof of [46, Theorem 6.1]. Since the details of the proof may obscure the idea, we now outline the reasoning behind our analysis. Instead of showing (3.51) for all $t > 0$, we first show (3.51) for some specially chosen $t_n > 0$. More precisely, we find constants $\beta \in (0,1)$ and $C \in (0,\infty)$ such that

$$\|P^{\delta n}(x,\cdot) - \pi\|_h \leq C\big(e^{\gamma x} + 1\big)\beta^n, \quad n \in \mathbb{Z}_+, \quad x \in \mathbb{R}_+, \tag{3.52}$$

where $\delta > 0$ is a constant and $h := V + 1 = \exp(\gamma x) + 1$. Then there is a simple trick in [46] to conclude (3.51) from (3.52). Thus, it suffices to show (3.52), or equivalently, to show the exponential ergodicity for the δ-skeleton chain of the BAJD. The exponential ergodicity of Markov chains has been studied in [44]. In particular, to show that a Markov chain is exponentially ergodic, [44, Theorem 6.3] provides some sufficient conditions: the chain admits a Foster-Lyapunov function and is irreducible and aperiodic, and all compact subsets are petite. If we can show that these conditions are satisfied for the δ-skeleton chain of the BAJD, then we are done.

Now we proceed to show that the conditions of [44, Theorem 6.3] are satisfied for the δ-skeleton chain of X. For any $\delta > 0$ we consider the δ-skeleton chain $Y_n^\delta := X_{n\delta},\ n \in \mathbb{Z}_+$. Then $(Y_n^\delta)_{n\in\mathbb{Z}_+}$ is a Markov chain with transition kernel $p(\delta, x, y)$ on the state space $\mathbb{R}_+$ and the law of the Y_n (started from $Y_0 = x$) is given by $P^{\delta n}(x,\cdot)$. It is easy to see that invariant measures for the BAJD process $(X_t)_{t\geq 0}$ are also invariant measures for $(Y_n^\delta)_{n\in\mathbb{Z}_+}$. Thus the probability measure π in Corollary 3 is also an invariant probability measure for the chain $(Y_n^\delta)_{n\in\mathbb{Z}_+}$.

Let $V(x) = \exp(\gamma x)$ be the same as in Lemma 9. It follows from the Markov property and Lemma 9 that

$$E_x[V(Y_{n+1})|Y_0, Y_1, \cdots, Y_n] = \int_{\mathbb{R}_+} V(y)p(\delta, Y_n, y)dy \leq e^{-\delta k}V(Y_n) + \frac{M}{k},$$

where k and M are positive constants. If we set $V_0 := V$ and $V_n := V(Y_n)$, $n \in \mathbb{N}$, then

$$E_x[V_1] \leq e^{-\delta k}V_0(x) + \frac{M}{k}$$

and

$$E_x[V_{n+1}|Y_0, Y_1, \cdots, Y_n] \leq e^{-\delta k}V_n + \frac{M}{k}, \quad n \in \mathbb{N}.$$

Now we proceed to show that the chain $(Y_n^\delta)_{n\in\mathbb{Z}_+}$ is λ-irreducible, strong aperiodic, and all compact subsets of $\mathbb{R}_+$ are petite for the chain $(Y_n^\delta)_{n\in\mathbb{Z}_+}$.

"λ-irreducibility": We show that the Lebesgue measure λ on $\mathbb{R}_+$ is an irreducibility measure for $(Y_n^\delta)_{n\in\mathbb{Z}_+}$. Let $A \in \mathcal{B}(\mathbb{R}_+)$ and $\lambda(A) > 0$, then

$$P[Y_1 \in A|Y_0 = x] = P_x[X_\delta \in A] = \int_A p(\delta, x, y)dy > 0,$$

since $p(\delta, x, y) > 0$ for any $x \in \mathbb{R}_+$ and $y > 0$. This shows that the chain $(Y_n^\delta)_{n\in\mathbb{Z}_+}$ is irreducible with λ being an irreducibility measure.

"Strong aperiodicity" (see [44, p.561] for a definition): To show the strong aperiodicity of $(Y_n)_{n\in\mathbb{N}_0}$, we need to find a set $C \in \mathcal{B}(\mathbb{R}_+)$, a probability measure ν with $\nu(C) = 1$, and $\epsilon > 0$ such that

$$L(x, C) > 0, \qquad x \in \mathbb{R}_+ \tag{3.53}$$

and

$$P_x(Y_1 \in A) \geq \epsilon \cdot \nu(A), \quad x \in C, \quad A \in \mathcal{B}(\mathbb{R}_+), \tag{3.54}$$

where $L(x, C) := \mathbb{P}_x(Y_m \in C$ for some m $\in \mathbb{N})$. To this end set $C := [0, 1]$ and $g(y) := \inf_{x\in[0,1]} p(\delta, x, y)$, $y > 0$. Since for fixed $y > 0$ the function $p(\delta, x, y)$ is strictly positive and continuous in $x \in [0, 1]$, thus we have $g(y) > 0$ and $0 < \int_{(0,1]} g(y)dy \leq 1$. Define

$$\nu(A) := \frac{1}{\int_{(0,1]} g(y)dy} \int_{A\cap(0,1]} g(y)dy, \qquad A \in \mathcal{B}(\mathbb{R}_+).$$

Then for any $x \in [0, 1]$ and $A \in \mathcal{B}(\mathbb{R}_+)$ we get

$$P_x(Y_1 \in A) = \int_A p(\delta, x, y)dy \geq \int_{A\cap(0,1]} g(y)dy = \nu(A) \int_{(0,1]} g(y)dy,$$

so (3.54) holds with $\epsilon := \int_{(0,1]} g(y)dy$.

Obviously

$$L(x, [0, 1]) \geq P_x(Y_1 \in [0, 1]) = P_x(X_\delta \in [0, 1]) = \int_{[0,1]} p(\delta, x, y)dy > 0$$

for all $x \in \mathbb{R}_+$, which verifies (3.53).

"Compact subsets are petite": We have shown that λ is an irreducibility measure for $(Y_n^\delta)_{n\in\mathbb{Z}_+}$. According to [44, Theorem 3.4(ii)], to show that all compact sets are petite, it suffices to prove the Feller property of $(Y_n^\delta)_{n\in\mathbb{Z}_+}$, but this follows from the fact that $(Y_n^\delta)_{n\in\mathbb{Z}_+}$ is a skeleton chain of the BAJD process $(X_t)_{t\geq 0}$, which possesses the Feller property.

According to [44, Theorem 6.3] (see also the proof of [44, Theorem 6.1]), the probability measure π is the only invariant probability measure of the chain $(Y_n^\delta)_{n\in\mathbb{Z}_+}$ and there exist constants $\beta \in (0, 1)$ and $C \in (0, \infty)$ such that

$$\|P^{\delta n}(x, \cdot) - \pi\|_h \leq C\big(e^{\gamma x} + 1\big)\beta^n, \quad n \in \mathbb{Z}_+, \quad x \in \mathbb{R}_+,$$

where $h := V + 1 = \exp(\gamma x) + 1$.

Then we can proceed as in [46, p.536] and get the inequality (3.51). $\square$

Since $\|\mu\|_{TV} \leq \|\mu\|_h$, it follows immediately the following corollary.

Corollary 5. *The BAJD is exponentially ergodic, namely there exist constants* $\beta \in (0,\infty)$ *and* $C \in (0,\infty)$ *such that*

$$\|P^t(x,\cdot) - \pi\|_{TV} \leq C\big(e^{\gamma x} + 1\big)\beta^t, \quad \forall t > 0, \quad x \in \mathbb{R}_+. \tag{3.55}$$

Remark 11. In Chapter 2 we proved that the CIR process is positive Harris recurrent. If we allow the parameter $c = 0$, then all results of this section still hold and thus are also true for the CIR process. In particular Theorem 10 is also true for the CIR process. In this case the unique invariant probability measure of the CIR process is the Gamma distribution $\Gamma\big(2a\theta/\sigma^2, \sigma^2/(2a)\big)$ and has the characteristic function $\big(1 - \sigma^2 u/(2a)\big)^{-2a\theta/\sigma^2}$.

3.6 Calibration for the BAJD-process

Now, we will use the results of this chapter to provide another consequence of Harris recurrence property for the BAJD process. We have already shown in section (3.4) that our result on the positive Harris recurrence of the BAJD provides another way of proving the existence and uniqueness of an invariant probability measure for the BAJD. We denote it by π, this measure is the convolution of a Gamma distribution with Bessel or mixture of Bessel distributions.

From the ergodicity point of view, we also show in corollary (4) that another consequence of the positive Harris recurrence is the limit theorem for additive functionals, namely we show that for all $f \in \mathcal{B}(\mathbb{R}_+)$, $\frac{1}{t}\int_0^t f(X_s)ds$ converges almost surely to $\int_{\mathbb{R}_+} f(x)\pi(dx)$. Obviously the BAJD process is strongly ergodic. In fact, it follow from the exponential ergodicity result of the BAJD, which was proved in the last section. Note that, the exponential ergodicity implies strong ergodicity but not vice-versa. Also, strong ergodicity of the BAJD follow from [29, Theorem 20.12], which say that any Harris recurrent Feller process is strongly ergodic.

Let us first recall the definition of affine term structure (see e.g. [13, Definition 3.1]).

Definition 12. *An* $\mathbb{R}_+-$*valued homogeneous Markov process* X *with*

$$E_x\big[e^{-\int_0^t X_s ds}\big] = e^{p(t)+xq(t)} \tag{3.56}$$

for some functions p *and* q*, is said to provide an affine term structure (ATS).*

We want the term $e^{-\int_0^t X_s ds}$ to be well-defined $\mathbb{R}_+-$valued adapted process, for each P_x. This is equivalent to assume that the process X is progressively measurable and

$$\int_0^t X_s ds < \infty, \quad P_x - a.s. \quad \forall t, x \in \mathbb{R}_+.$$

This implies that X is conservative. Also X is a special Feller process and the function p and q satisfy a system of generalized Riccati equations. The functions $p(t)$ and $q(t)$ are nonnegative and non-decreasing functions with initial conditions $p(0) = q(0) = 0$.

Now we apply Theorem (3) and by taking $g(x) = x$ and $f(x) = e^{-x}$ we get

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} e^{-\int_j^{j+1} X_s ds} = E_\pi\left[e^{-\int_0^1 X_s ds}\right] \tag{3.57}$$

and this convergence holds almost surely. Now, we will calibrate this limit

$$E_\pi\left[e^{-\int_0^1 X_s ds}\right] = ?$$

Since the BAJD process belongs to the class of affine term structure, by the definition we know that

$$E_x\left[e^{-\int_0^t X_s ds}\right] = e^{p(t)+xq(t)} \tag{3.58}$$

where the functions p and q appearing in the term structure solve the following generalized Riccati equations

$$\begin{cases} \partial_t p(t) = F_1\big(q(t)\big), & p(0) = 0, \\ \partial_t q(t) = R_1\big(q(t)\big), & q(0) = 0, \end{cases} \tag{3.59}$$

and

$$F_1(u) = a\theta u + \frac{cu}{d-u}, \quad u \in \mathbb{C} \setminus \{d\}, \tag{3.60}$$

$$R_1(u) = \frac{\sigma^2 u^2}{2} - au - 1, \quad u \in \mathbb{C}. \tag{3.61}$$

(see [38, Proposition 3.4]).

The second equation in the above system (3.59) is well-known particular case of the Riccati equations because the coefficients considered are constants. In this case this equation can be reduced to a separable differential equation.

$$\frac{dq(t)}{dt} = \frac{\sigma^2}{2} q^2(t) - aq(t) - 1, \quad q(0) = 0.$$

The solution is described by the integral of rational function with quadratic function in the denominator.

$$\int_0^t \frac{dq(s)}{\frac{\sigma^2}{2}q^2(s) - aq(s) - 1} = \int_0^t ds = t.$$

Now

$$\begin{aligned}\int_0^t \frac{dq(s)}{\frac{\sigma^2}{2}q^2(s) - aq(s) - 1} &= \int_0^t \Big(\frac{1}{q(s) - \frac{a-\sqrt{a^2+2\sigma^2}}{\sigma^2}} + \frac{-1}{q(s) - \frac{a+\sqrt{a^2+2\sigma^2}}{\sigma^2}}\Big)dq(s) \\ &= -\sqrt{a^2+2\sigma^2}t.\end{aligned}$$

Finally one can get

$$q(t) = \frac{1 - e^{b_1 t}}{c_1 + d_1 e^{b_1 t}}$$

where $c_1 = -\frac{a+\sqrt{a^2+2\sigma^2}}{2}$, $d_1 = \frac{a-\sqrt{a^2+2\sigma^2}}{2}$ and $b_1 = -\sqrt{a^2+2\sigma^2}$. Now for the first equation of the system (3.59), we denote

$$I_1 := a\theta \int_0^t q(s)ds = a\theta \int_0^t \frac{1 - e^{b_1 s}}{c_1 + d_1 e^{b_1 s}} ds$$

and

$$I_2 := c\int_0^t \frac{q(s)}{d - q(s)} ds = c\int_0^t \frac{1 - e^{b_1 s}}{(dd^1 + 1)e^{b_1 s} + (dc_1 - 1)} ds$$

By change of variable $r = e^{b_1 s}$ in $I_1$0it follow that

$$I_1 = \frac{2a\theta}{\sigma^2} \log\Big(\frac{b_1 e^{d_1 t}}{c_1 + d_1 e^{b_1 t}}\Big) = \frac{2a\theta}{\sigma^2} \log\Big(\frac{b_1 e^{(a+b_1)t/2}}{c_1 + d_1 e^{b_1 t}}\Big)$$

For the second integral I_2, we distinguish between two cases. If $d \neq d_1$ then by change of variable is easy to get

$$I_2 = \frac{ct}{c_1 d - 1} - \frac{cd}{\frac{1}{2}\sigma^2 d^2 - ad - 1} \log\Big(\frac{(1 + dd_1)e^{b_1 t} + c_1 d - 1}{b_1 d}\Big).$$

If $d = d_1$ then we get

$$I_2 = \frac{ct}{c_1 d - 1} - \frac{cd}{\frac{1}{2}\sigma^2 d^2 - ad - 1} + \frac{d_1}{b_1^2}\big(e^{b_1 t} - 1\big)$$

Therefore the solution p can be represented explicitly in the following way:

$$p(t) = \begin{cases} \frac{2a\theta}{\sigma^2}\log\big(\frac{b_1 e^{(a+b_1)t/2}}{c_1 + d_1 e^{b_1 t}}\big) + \frac{ct}{c_1 d - 1} \\ \quad -\frac{cd}{\frac{1}{2}\sigma^2 d^2 - ad - 1}\log\Big(\frac{(1+dd_1)e^{b_1 t} + c_1 d - 1}{b_1 d}\Big), & \text{if } d \neq d_1, \\ \frac{2a\theta}{\sigma^2}\log\big(\frac{b_1 e^{(a+b_1)t/2}}{c_1 + d_1 e^{b_1 t}}\big) + \frac{ct}{c_1 d - 1} + \frac{d_1}{b_1^2}\big(e^{b_1 t} - 1\big), & \\ & \text{if } d = d_1. \end{cases}$$

Since the BAJD has an explicit form of invariant measure hence we calibrate this limit in the explicit form

$$
\begin{aligned}
E_\pi\left[e^{-\int_0^1 X_s ds}\right] &= \int_0^\infty E_x\left[e^{-\int_0^1 X_s ds}\right]\pi(dx) \\
&= \int_0^\infty e^{p(1)+xq(1)}\pi(dx) \\
&= e^{p(1)}\int_0^\infty e^{xq(1)}\pi(dx) \\
&= \begin{cases} e^{p(1)}\left(1-\frac{\sigma^2}{2a}q(1)\right)^{-\frac{2a\theta}{\sigma^2}}\cdot\left(\frac{d-\frac{\sigma^2 dq(1)}{2a}}{d-q(1)}\right)^{\frac{c}{a-\frac{\sigma^2 d}{2}}}, & \text{if } \Delta\neq 0, \\ e^{p(1)}\left(1-\frac{\sigma^2}{2a}q(1)\right)^{-\frac{2a\theta}{\sigma^2}}\cdot\exp\left(\frac{cq(1)}{a(d-q(1))}\right), & \text{if } \Delta = 0. \end{cases}
\end{aligned}
$$

where $\Delta = a - \frac{\sigma^2 d}{2}$.

Chapter 4

Exponential Ergodicity of the Jump-Diffusion CIR Process

Over the last few years, Lévy processes and other processes including jumps became quite popular in finance. Jumps processes have been playing increasingly important roles in various applications, for example in short term interest rate. A naturel generalization of the classical Cox-Ingersoll-Ross process takes into account the jumps have been studied by Duffie and Gârleanu in [10], Filipović in [13] and Li and Ma in [42].

In this chapter, we study the jump-diffusion CIR process (abbreviated as JCIR), which is an extension of the classical CIR model. The jumps of the JCIR are introduced with the help of a pure-jump Lévy process $(J_t, t \geq 0)$. The JCIR process is defined as the unique strong solution $X := (X_t, t \geq 0)$ to the following stochastic differential equation

$$dX_t = a(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t + dJ_t, \quad X_0 \geq 0, \tag{4.1}$$

where $a, \sigma > 0$, $\theta \geq 0$ are constants, $(W_t, t \geq 0)$ is a one-dimensional Brownian motion and $(J_t, t \geq 0)$ is a pure-jump Lévy process with its Lévy measure ν concentrated on $(0, \infty)$ and satisfying

$$\int_{(0,\infty)} (\xi \wedge 1)\nu(d\xi) < \infty, \tag{4.2}$$

independent of the Brownian motion $(W_t, t \geq 0)$. The initial value X_0 is assumed to be independent of $(W_t, t \geq 0)$ and $(J_t, t \geq 0)$. We assume that all the above processes are defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{t\geq 0}, P)$. We should remark that the existence and uniqueness of strong solutions to the SDE (4.1) are guaranteed by [15, Thm. 5.1].

The drift factor $a(\theta - X_t)$ in (4.1) is exactly the same as in the BAJD process (3.1), defines a mean reverting drift pulling the process towards its long-term value θ with a speed of adjustment equal to a. Since the diffusion coefficient $\sigma\sqrt{X_t}$ in the SDE (4.1) is degenerate at 0 and only positive jumps are allowed, the JCIR process $(X_t, t \geq 0)$ stays non-negative if $X_0 \geq 0$. This fact can be shown rigorously with the help of comparison theorems for SDEs, for more details we refer the readers to [15].

In this chapter we are interested in two problems concerning the JCIR defined in (4.1). The first one is to study the transition density estimates of the JCIR, namely we find a lower bound on the transition densities of the JCIR. Our idea to establish the lower bound of the transition densities is as follows. It well known that the JCIR is also an affine processes in $\mathbf{R}_+$. Based on the exponential-affine structure of the JCIR, we are able to compute its characteristic function explicitly. Moreover, this enables us to represent the distribution of the JCIR as the convolution of two distributions. The first distribution coincides with the distribution of the CIR model. However, the second distribution is more complicated. We will give a sufficient condition such that the second distribution is singular at the point 0. In this way we derive a lower bound estimate of the transition densities of the JCIR.

4.1 Characteristic function of the JCIR

In this section we use the exponential-affine structure of the JCIR process to derive its characteristic functions.

We recall that the JCIR process $(X_t, t \geq 0)$ is defined to be the solution to (4.1) and it depends obviously on its initial value X_0. From now on we denote by $(X_t^x, t \geq 0)$ the JCIR process started from an initial point $x \geq 0$, namely

$$dX_t^x = a(\theta - X_t^x)dt + \sigma\sqrt{X_t^x}dW_t + dJ_t, \quad X_0^x = x. \tag{4.3}$$

Since the JCIR process is an affine jump diffusion, the corresponding characteristic functions of $(X_t^x, t \geq 0)$ is of an exponential-affine form, i.e. there exist functions $\phi(t, u)$ and $\psi(t, u)$ such that

$$E\big[e^{uX_t^x}\big] = e^{\phi(t,u)+x\psi(t,u)}, \quad \text{for all} \quad t \geq 0,\ u \in \mathcal{U},\ x \geq 0, \tag{4.4}$$

where $\mathcal{U} := \{u \in \mathbf{C} : \Re u \leq 0\}$, $\Re u$ denotes the real part of u. Moreover, the functions $\phi(t, u)$ and $\psi(t, u)$ are the unique solutions of the generalized Riccati equations

$$\begin{cases} \partial_t\phi(t, u) = F\big(\psi(t, u)\big), & \phi(0, u) = 0, \\ \partial_t\psi(t, u) = R\big(\psi(t, u)\big) & \psi(0, u) = u \in \mathcal{U}, \end{cases} \tag{4.5}$$

with the functions F and R are of Lévy-Khitchine form

$$F(u) = a\theta u + \int_{(0,\infty)} (e^{u\xi} - 1)\nu(d\xi), \tag{4.6}$$

$$R(u) = \frac{\sigma^2 u^2}{2} - au. \tag{4.7}$$

For general case see [[13], Theorem 4.3].

Note that, it is not difficult to find the explicit form of the functions $\partial_t \phi$ and $\partial_t \psi$. We have to proceed as in the Section (2.1). More precisely, we can find the functions $\partial_t \phi$ and $\partial_t \psi$ under the initial conditions such that

$$M_t := \exp\big(\phi(T-t,u) + X_t \psi(T-t,u)\big)$$

is a martingale. We can derive the generalized Riccati equation by applying Itô formula for the jumps diffusion and then we collect the coefficients.

Solving the system (4.5) gives $\phi(t,u)$ and $\psi(t,u)$ in their explicit forms. One can remark that the second equation of the generalized Riccati equations is a Bernoulli differential equation with $n = 2$. Therefore the solution ψ can be represented explicitly in the following way:

$$\begin{aligned} \psi(t,u) &= \psi(0,u) e^{-\int_0^t a ds}\big(1 - u\int_0^t \frac{\sigma^2}{2} e^{-as} ds\big)^{-1} \\ &= \frac{ue^{-at}}{1 - \frac{\sigma^2}{2a}u(1-e^{-at})}. \end{aligned} \tag{4.8}$$

Note that the first Riccati equation is just an integral, and ϕ may be written explicitly as:

$$\begin{aligned} \phi(t,u) &= \frac{-2a\theta}{\sigma^2}\int_0^t \frac{\frac{\sigma^2}{2}ue^{-as}}{1-\frac{\sigma^2}{2a}u(1-e^{-as})} ds + \int_0^t \int_{(0,\infty)} \big(e^{\xi\psi(s,u)} - 1\big)\nu(d\xi) ds \\ &= -\frac{2a\theta}{\sigma^2}\log\big(1 - \frac{\sigma^2}{2a}u(1-e^{-at})\big) + \int_0^t \int_{(0,\infty)} \Big(e^{\xi\psi(s,u)} - 1\Big)\nu(d\xi) ds. \end{aligned} \tag{4.9}$$

Here the complex-valued logarithmic function $\log(\cdot)$ is understood to be its main branch defined on $\mathbf{C} \setminus \{0\}$.

It follows from (4.4), (4.8) and (4.9) that the characteristic functions of $(X_t^x, t \geq 0)$ is given by

$$E[e^{uX_t^x}] = \varphi_1(t,u,x)\varphi_2(t,u), \qquad u \in \mathcal{U}, \tag{4.10}$$

where $\varphi_1(t,u,x)$ and $\varphi_2(t,u)$ are defined as follows, for $t \geq 0$ and $u \in \mathcal{U}$

$$\begin{aligned} \varphi_1(t,u,x) : &= \big(1 - \frac{\sigma^2}{2a} u(1-e^{-at})\big)^{-\frac{2a\theta}{\sigma^2}} \exp\Big(\frac{xue^{-at}}{1 - \frac{\sigma^2}{2a}u(1-e^{-at})}\Big), \\ \varphi_2(t,u) : &= \exp\Big(\int_0^t \int_0^\infty \Big(e^{\xi\psi(s,u)} - 1\Big)\nu(d\xi)ds\Big), \end{aligned} \tag{4.11}$$

where the complex-valued power function $z^{-2a\theta/\sigma^2} := \exp\big(-(2a\theta/\sigma^2)\log z\big)$ is also understood to be its main branch defined on $\mathbf{C} \setminus \{0\}$.

According to the parameters of the JCIR process we look at two special cases:

4.1.1 Special Case i): $\nu = 0$, No Jumps

Notice that the case $\nu = 0$ corresponds to the classical CIR model $(Y_t, t \geq 0)$ satisfying the following stochastic differential equation

$$dY_t^x = a(\theta - Y_t^x)dt + \sigma\sqrt{Y_t^x}dW_t, \quad Y_0^x = x \geq 0. \tag{4.12}$$

It follows from (4.10) that the characteristic function of $(Y_t^x, t \geq 0)$ coincides with $\varphi_1(t,u,x)$, namely, for $u \in \mathcal{U}$

$$E[e^{uY_t^x}] = \varphi_1(t,u,x).$$

It is well known that the classical CIR model $(Y_t^x, t \geq 0)$ has transition density functions $f(t,x,y)$ given by

$$f(t,x,y) = \kappa e^{-u-v}\Big(\frac{v}{u}\Big)^{\frac{q}{2}} I_q\big(2(uv)^{\frac{1}{2}}\big) \tag{4.13}$$

for $t > 0, x > 0$ and $y \geq 0$, where

$$\begin{aligned} \kappa &\equiv \frac{2a}{\sigma^2\Big(1-e^{-at}\Big)}, \qquad u \equiv \kappa x e^{-at}, \\ v &\equiv \kappa y, \qquad\qquad\qquad q \equiv \frac{2a\theta}{\sigma^2} - 1, \end{aligned}$$

and $I_q(\cdot)$ is the modified Bessel function of the first kind of order q. For $x = 0$ the formula of the density function $f(t,x,y)$ is given by

$$f(t,0,y) = \frac{c}{\Gamma(q+1)} v^q e^{-v} \tag{4.14}$$

for $t > 0$ and $y \geq 0$.

4.1.2 Special Case ii): $\theta = 0$ and $x = 0$

We denote by $(Z_t, t \geq 0)$ the JCIR process given by

$$dZ_t = -aZ_t dt + \sigma\sqrt{Z_t} dW_t + dJ_t, \quad Z_0 = 0. \tag{4.15}$$

In this particular case the characteristic functions of $(Z_t, t \geq 0)$ is equal to $\varphi_2(t, u)$, namely, for $u \in \mathcal{U}$

$$E[e^{uZ_t}] = \varphi_2(t, u). \tag{4.16}$$

One can notice that $\varphi_2(t, u)$ resembles the characteristic function of a compound Poisson distribution.

4.2 Lower bound for the transition densities of JCIR

In this section we will find some conditions on the Lévy measure ν of $(J_t, t \geq 0)$ such that an explicit lower bound for the transition densities of the JCIR process given in (4.3) can be derived. As a first step we show that the law of $X_t^x, t > 0$, in (4.3) is absolutely continuous with respect to the Lebesgue measure and thus possesses a density function.

Lemma 10. *Consider the JCIR process $(X_t^x, t \geq 0)$ (started from $x \geq 0$) that is defined in (4.3). Then for any $t > 0$ and $x \geq 0$ the law of X_t^x is absolutely continuous with respect to the Lebesgue measure and thus possesses a density function $p(t, x, y),\ y \geq 0$.*

Proof. As shown in the previous section, it holds

$$E[e^{uX_t^x}] = \varphi_1(t, u, x)\varphi_2(t, u) = E[e^{uY_t^x}]E[e^{uZ_t}],$$

therefore the law of X_t^x, denoted by $\mu_{X_t^x}$, is the convolution of the laws of $\mu_{Y_t^x}$ and μ_{Z_t}. Where $\mu_{Y_t^x}$ and μ_{Z_t} are the laws of Y_t^x and Z_t respectively. Since $(Y_t^x, t \geq 0)$ is the well-known classical CIR process and has transition density function $f(t, x, y), t > 0, x, y \geq 0$ with respect to the Lebesgue measure, thus $\mu_{X_t^x}$ is also absolutely continuous with respect to the Lebesgue measure and possesses a density function.

□

In order to get a lower bound for the transition densities of the JCIR process we need the following lemma.

Lemma 11. *Suppose that $\int_{(0,1)} \xi \ln(1/\xi)\nu(d\xi) < \infty$. Then φ_2 defined by (4.11) is the characteristic function of a compound Poisson distribution. In particular, $P(Z_t = 0) > 0$ for all $t > 0$, where $(Z_t, t \geq 0)$ is defined by (4.15).*

Proof. It follows from (4.8), (4.11) and (4.16) that

$$E[e^{uZ_t}] = \varphi_2(t,u) = \exp\bigg(\int_0^t \int_{(0,\infty)} \Big(\exp\big(\tfrac{\xi u e^{-as}}{1-(\sigma^2/2a)(1-e^{-as})u}\big) - 1\Big)\nu(d\xi)ds\bigg),$$

where $u \in \mathcal{U}$. Define

$$\Delta := \int_0^t \int_{(0,\infty)} \Big(\exp\Big(\frac{\xi u e^{-as}}{1-(\sigma^2/2a)(1-e^{-as})u}\Big) - 1\Big)\nu(d\xi)ds.$$

If we rewrite

$$\exp\Big(\frac{\xi e^{-as}u}{1-(\sigma^2/2a)(1-e^{-as})u}\Big) = \exp\Big(\frac{\alpha u}{\beta - u}\Big), \tag{4.17}$$

where

$$\begin{cases} \alpha := \dfrac{2a\xi}{\sigma^2(e^{as}-1)} > 0, \\ \beta := \dfrac{2ae^{as}}{\sigma^2(e^{as}-1)} > 0, \end{cases} \tag{4.18}$$

then we recognize that the right-hand side of (4.17) is the characteristic function of a Bessel distribution with parameters α and β. Recall that a probability measure $\mu_{\alpha,\beta}$ on $\big(\mathbf{R}_+, \mathcal{B}(\mathbf{R}_+)\big)$ is called a Bessel distribution with parameters α and β if

$$\mu_{\alpha,\beta}(dx) = e^{-\alpha}\delta_0(dx) + \beta e^{-\alpha-\beta x}\sqrt{\frac{\alpha}{\beta x}} I_1(2\sqrt{\alpha\beta x})dx, \tag{4.19}$$

where δ_0 is the Dirac measure at the origin and I_1 is the modified Bessel function of the first kind, namely,

$$I_1(r) = \frac{r}{2}\sum_{k=0}^{\infty} \frac{\big(\frac{1}{4}r^2\big)^k}{k!(k+1)!}, \qquad r \in \mathbf{R}.$$

For more properties of Bessel distributions we refer the readers to [19, Sect. 3] (see also [12, p.438] and [27, Sect. 3]). Let $\hat{\mu}_{\alpha,\beta}$ denote the characteristic function of the Bessel distribution $\mu_{\alpha,\beta}$ with parameters α and β which are defined in (4.18). It follows from [27, Lemma 3.1] that

$$\hat{\mu}_{\alpha,\beta}(u) = \exp\Big(\frac{\alpha u}{\beta - u}\Big) = \exp\Big(\frac{\xi e^{-as}u}{1-(\sigma^2/2a)(1-e^{-as})u}\Big).$$

Therefore

$$\begin{aligned}\Delta &= \int_0^t \int_{(0,\infty)} \big(\hat{\mu}_{\alpha,\beta}(u) - 1\big)\nu(d\xi)ds \\ &= \int_0^t \int_{(0,\infty)} \Big(e^{\frac{\alpha u}{\beta - u}} - e^{-\alpha} + e^{-\alpha} - 1\Big)\nu(d\xi)ds.\end{aligned}$$

Set

$$\begin{aligned}\lambda &:= \int_0^t \int_{(0,\infty)} \big(1 - e^{-\alpha}\big)\nu(d\xi)ds \\ &= \int_0^t \int_{(0,\infty)} \Big(1 - e^{-\frac{2a\xi}{\sigma^2(e^{as}-1)}}\Big)\nu(d\xi)ds. \qquad (4.20)\end{aligned}$$

If $\lambda < \infty$, then

$$\begin{aligned}\Delta &= \int_0^t \int_{(0,\infty)} \Big(e^{\frac{\alpha u}{\beta - u}} - e^{-\alpha}\Big)\nu(d\xi)ds - \lambda \\ &= \lambda\Big(\frac{1}{\lambda}\int_0^t \int_{(0,\infty)} \big(e^{\frac{\alpha u}{\beta - u}} - e^{-\alpha}\big)\nu(d\xi)ds - 1\Big).\end{aligned}$$

The fact that $\lambda < \infty$ will be shown later in this proof.
Next we show that the term $(1/\lambda)\int_0^t \int_{(0,\infty)} \big(\exp\big(\alpha u/(\beta - u)\big) - \exp(-\alpha)\big)\nu(d\xi)ds$ can be viewed as the characteristic function of a probability measure ρ. To define ρ, we first construct the following measures

$$m_{\alpha,\beta}(dx) := \beta e^{-\alpha-\beta x}\sqrt{\frac{\alpha}{\beta x}}I_1(2\sqrt{\alpha\beta x})dx, \quad x \geq 0,$$

where I_1 is the modified Bessel function of the first kind. Noticing that the measure $m_{\alpha,\beta}$ is the absolute continuous component of the measure $\mu_{\alpha,\beta}$ in (4.19), we easily get

$$\hat{m}_{\alpha,\beta}(u) = \hat{\mu}_{\alpha,\beta}(u) - e^{-\alpha} = e^{\frac{\alpha u}{\beta - u}} - e^{-\alpha},$$

where $\hat{m}_{\alpha,\beta}(u) := \int_0^\infty e^{ux} m_{\alpha,\beta}(dx)$ for $u \in \mathcal{U}$. Recall that the parameters α and β defined by (4.18) depend on the variables ξ and s. We can define a measure ρ on $\mathbf{R}_+$ as follows:

$$\rho := \frac{1}{\lambda}\int_0^t \int_{(0,\infty)} m_{\alpha,\beta}\,\nu(d\xi)ds.$$

By the definition of the constant λ in (4.20) we get

$$\begin{aligned}\rho(\mathbf{R}_+) &= \frac{1}{\lambda}\int_0^t\int_{(0,\infty)} m_{\alpha,\beta}(\mathbf{R}_+)\nu(d\xi)ds\\ &= \frac{1}{\lambda}\int_0^t\int_{(0,\infty)}(1-e^{-\alpha})\nu(d\xi)ds\\ &= 1,\end{aligned}$$

i.e. ρ is a probability measure on $\mathbf{R}_+$, and for $u \in \mathcal{U}$

$$\begin{aligned}\hat{\rho}(u) &= \int_{(0,\infty)} e^{ux}\rho(dx)\\ &= \frac{1}{\lambda}\int_0^t\int_{(0,\infty)}(e^{\frac{\alpha u}{\beta-u}}-e^{-\alpha})\nu(d\xi)ds.\end{aligned}$$

Thus $\Delta = \lambda(\hat{\rho}(u)-1)$ and $E[e^{uZ_t}] = \varphi_2(t,u) = e^{\lambda(\hat{\rho}(u)-1)}$ is the characteristic function of a compound Poisson distribution.

Now we verify that $\lambda < \infty$. Noticing that

$$\begin{aligned}\lambda &= \int_0^t\int_{(0,\infty)}\big(1-e^{-\alpha}\big)\nu(d\xi)ds\\ &= \int_0^t\int_{(0,\infty)}\Big(1-e^{-\frac{2a\xi}{\sigma^2(e^{as}-1)}}\Big)\nu(d\xi)ds\\ &= \int_{(0,\infty)}\int_0^t\Big(1-e^{-\frac{2a\xi}{\sigma^2(e^{as}-1)}}\Big)ds\nu(d\xi),\end{aligned}$$

we introduce the change of variables $\dfrac{2a\xi}{\sigma^2(e^{as}-1)} := y$ and then get

$$\begin{aligned}dy &= -\frac{2a\xi}{\sigma^2(e^{as}-1)^2}ae^{as}ds\\ &= -y^2\frac{\sigma^2}{2\xi}\Big(\frac{2a\xi}{\sigma^2 y}+1\Big)ds.\end{aligned}$$

Therefore

$$\begin{aligned}\lambda &= \int_{(0,\infty)}\nu(d\xi)\int_{\infty}^{\frac{2a\xi}{\sigma^2(e^{at}-1)}}(1-e^{-y})\frac{-2\xi}{2a\xi y+\sigma^2y^2}dy\\ &= \int_{(0,\infty)}\nu(d\xi)\int_{\frac{2a\xi}{\sigma^2(e^{at}-1)}}^{\infty}(1-e^{-y})\frac{2\xi}{2a\xi y+\sigma^2y^2}dy\\ &= \int_{(0,\infty)}\nu(d\xi)\int_{\frac{\xi}{\delta}}^{\infty}(1-e^{-y})\frac{2\xi}{2a\xi y+\sigma^2y^2}dy,\end{aligned}$$

where $\delta := \dfrac{\sigma^2(e^{at}-1)}{2a}$. Define

$$M(\xi) := \int_{\frac{\xi}{\delta}}^{\infty} (1-e^{-y}) \frac{2\xi}{2a\xi y + \sigma^2 y^2} dy.$$

Then $M(\xi)$ is continuous on $(0,\infty)$. As $\xi \to 0$ we get

$$\begin{aligned} M(\xi) &= \int_{\frac{\xi}{\delta}}^{1} (1-e^{-y}) \frac{2\xi}{2a\xi y + \sigma^2 y^2} dy + 2\xi \int_{1}^{\infty} (1-e^{-y}) \frac{dy}{2a\xi y + \sigma^2 y^2} \\ &\leq 2\xi \int_{\frac{\xi}{\delta}}^{1} \frac{y}{2a\xi y + \sigma^2 y^2} dy + 2\xi \int_{1}^{\infty} \frac{1}{2a\xi y + \sigma^2 y^2} dy \\ &\leq 2\xi \int_{\frac{\xi}{\delta}}^{1} \frac{1}{2a\xi + \sigma^2 y} dy + 2\xi \int_{1}^{\infty} \frac{1}{\sigma^2 y^2} dy. \end{aligned}$$

Since

$$\begin{aligned} 2\xi \int_{\frac{\xi}{\delta}}^{1} \frac{1}{2a\xi + \sigma^2 y} dy &= \frac{2\xi}{\sigma^2} \Big[\ln(2a\xi + \sigma^2 y) \Big]_{\frac{\xi}{\delta}}^{1} \\ &= \frac{2\xi}{\sigma^2} \ln(2a\xi + \sigma^2) - \frac{2\xi}{\sigma^2} \ln(2a\xi + \frac{\sigma^2 \xi}{\delta}) \\ &\leq c_1 \xi + c_2 \xi \ln(\frac{1}{\xi}) \leq c_3 \xi \ln(\frac{1}{\xi}) \end{aligned}$$

for sufficiently small ξ, we conclude that

$$M(\xi) \leq c_4 \xi \ln(\frac{1}{\xi}), \quad \text{as } \xi \to 0.$$

If $\xi \to \infty$, then

$$\begin{aligned} M(\xi) &\leq \int_{\frac{\xi}{\delta}}^{\infty} (1-e^{-y}) \frac{2\xi}{2a\xi y + \sigma^2 y^2} dy \\ &\leq \int_{\frac{\xi}{\delta}}^{\infty} \frac{2\xi}{2a\xi y + \sigma^2 y^2} dy \leq 2\xi \int_{\frac{\xi}{\delta}}^{\infty} \frac{1}{\sigma^2 y^2} dy \\ &= \frac{2\xi}{\sigma^2} \int_{\frac{\xi}{\delta}}^{\infty} d(-\frac{1}{y}) = \frac{2\xi}{\sigma^2} \Big[-\frac{1}{y} \Big]_{\frac{\xi}{\delta}}^{\infty} \\ &= \frac{2\xi}{\sigma^2} \frac{\delta}{\xi} = \frac{2\delta}{\sigma^2} := c_5 < \infty. \end{aligned}$$

Therefore,

$$\lambda \leq c_4 \int_0^1 \xi \ln(\frac{1}{\xi})\nu(d\xi) + c_5 \int_1^\infty 1\nu(d\xi) < \infty.$$

□

With the help of the above Lemma 11 we can easily prove the following proposition.

Proposition 3. *Let* $p(t,x,y),\ t > 0, x, y \geq 0$ *denote the transition density of the JCIR process* $(X_t^x, t \geq 0)$ *defined in* (4.3). *Suppose that* $\int_{(0,1)} \xi \ln(\frac{1}{\xi})\nu(d\xi) < \infty$. *Then for all* $t > 0, x, y \geq 0$ *we have*

$$p(t,x,y) \geq P(Z_t = 0)f(t,x,y),$$

where $P(Z_t = 0) > 0$ *for all* $t > 0$ *and* $f(t,x,y)$ *are transition densities of the classical CIR process.*

Proof. According to Lemma 11, we have $P(Z_t = 0) > 0$ for all $t > 0$. Since

$$E[e^{uX_t^x}] = \varphi_1(t,u,x)\varphi_2(t,u) = E[e^{uY_t^x}]E[e^{uZ_t}],$$

the law of X_t^x, denoted by $\mu_{X_t^x}$, is the convolution of the laws of $\mu_{Y_t^x}$ and μ_{Z_t}. Where $\mu_{Y_t^x}$ and μ_{Z_t} are the laws of Y_t^x and Z_t respectively. Thus for all $A \in \mathcal{B}(\mathbf{R}_+)$

$$\begin{aligned}
\mu_{X_t^x}(A) &= \int_{\mathbf{R}_+} \mu_{Y_t^x}(A - y)\mu_{Z_t}(dy) \\
&\geq \int_{\{0\}} \mu_{Y_t^x}(A - y)\mu_{Z_t}(dy) \\
&\geq \mu_{Y_t^x}(A)\mu_{Z_t}(\{0\}) \\
&\geq P(Z_t = 0)\mu_{Y_t^x}(A) \\
&\geq P(Z_t = 0)\int_A f(t,x,y)dy,
\end{aligned}$$

where $A - y = \{z - y, z \in A\}$ and $f(t,x,y)$ are the transition densities of the classical CIR process given in (4.13). Since $A \in \mathcal{B}(\mathbf{R}_+)$ is arbitrary, we get

$$p(t,x,y) \geq P(Z_t = 0)f(t,x,y)$$

for all $t > 0, x, y \geq 0$. □

4.3 Exponential ergodicity of JCIR

Now we consider the second problem of this chapter which is the exponential ergodicity of the JCIR. According to the main results of [32] (see also [38]), the JCIR has a unique invariant probability measure π, given that some integrability condition on the Lévy measure of $(J_t, t \geq 0)$ is satisfied. Under some sharper assumptions we show that the convergence of the law of the JCIR process to its invariant probability measure under the total variation norm is exponentially fast, which is called the exponential ergodicity. Our method is the same as in Chapter 3 Section (3.5), namely we show the existence of a Foster-Lyapunov function and then apply the general framework of [44, 45, 46] to get the exponential ergodicity.

In this section we find some sufficient conditions such that the JCIR process is exponentially ergodic. We have derived a lower bound for the transition densities of the JCIR process in the previous section. Next we show that the function $V(x) = x$, $x \geq 0$, is a Foster-Lyapunov function for the JCIR process.

Lemma 12. *Suppose that $\int_{(1,\infty)} \xi\nu(d\xi) < \infty$. Then the function $V(x) = x$, $x \geq 0$, is a Foster-Lyapunov function for the JCIR process defined in (4.3), in the sense that for all $t > 0$, $x \geq 0$,*

$$E[V(X_t^x)] \leq e^{-at}V(x) + M,$$

where $0 < M < \infty$ is a constant.

Proof. We know that $\mu_{X_t^x} = \mu_{Y_t^x} * \mu_{Z_t}$, therefore

$$E[X_t^x] = E[Y_t^x] + E[Z_t].$$

Since $(Y_t^x, t \geq 0)$ is the classical CIR process starting from x, it is known that $\mu_{Y_t^x}$ is a non-central Chi-squared distribution and thus $E[Y_t^x] < \infty$. Next we show that $E[Z_t] < \infty$.

Let $u \in (-\infty, 0)$. By using Fatou's Lemma we get

$$\begin{aligned} E[Z_t] &= E\Big[\lim_{u\to 0} \frac{e^{uZ_t}-1}{u}\Big] \\ &\leq \liminf_{u\to 0} E\Big[\frac{e^{uZ_t}-1}{u}\Big] = \liminf_{u\to 0} \frac{E[e^{uZ_t}]-1}{u}. \end{aligned}$$

Recall that

$$E[e^{uZ_t}] = \varphi_2(t,u) = \exp\Big(\int_0^t \int_{(0,\infty)} \Big(e^{\frac{\xi u e^{-as}}{1-(\sigma^2/2a)(1-e^{-as})u}} - 1\Big)\nu(d\xi)ds\Big) = e^{\Delta(u)}.$$

Then we have for all $u \leq 0$

$$\begin{aligned} & \frac{\partial}{\partial u}\Big(\exp\Big(\frac{\xi u e^{-as}}{1-(\sigma^2/2a)(1-e^{-as})u}\Big)-1\Big) \\ = \quad & \frac{\xi e^{-as}}{\big(1-(\sigma^2/2a)(1-e^{-as})u\big)^2}\exp\Big(\frac{\xi u e^{-as}}{1-(\sigma^2/2a)(1-e^{-as})u}\Big) \\ \leq \quad & \frac{\xi e^{-as}}{\big(1-(\sigma^2/2a)(1-e^{-as})u\big)^2} \leq \xi e^{-as} \end{aligned}$$

and further

$$\int_0^t \int_{(0,\infty)} \xi e^{-as} \nu(d\xi) ds < \infty.$$

Thus $\Delta(u)$ is differentiable in u and

$$\Delta'(0) = \int_0^t \int_{(0,\infty)} \xi e^{-as} \nu(d\xi) ds = \frac{1-e^{-at}}{a} \int_{(0,\infty)} \xi \nu(d\xi).$$

It follows that

$$\begin{aligned} E[Z_t] &\leq \liminf_{u \to 0} \frac{\varphi_2(t,u)-\varphi_2(t,0)}{u} \\ &= \frac{\partial \varphi_2(t,u)}{\partial u}\Big|_{u=0} = e^{\Delta(0)}\Delta'(0) \\ &= \frac{1-e^{-at}}{a}\int_{(0,\infty)} \xi \nu(d\xi). \end{aligned}$$

Therefore under the assumption $\int_{(0,\infty)} \xi \nu(d\xi) < \infty$ we have proved that $E[Z_t] < \infty$. Furthermore,

$$E[Z_t] = \frac{\partial}{\partial u}\Big(E[e^{uZ_t}]\Big)\Big|_{u=0} = \frac{1-e^{-at}}{a}\int_{(0,\infty)} \xi \nu(d\xi).$$

On the other hand,

$$E[e^{uY_t^x}] = \big(1-(\sigma^2/2a)u(1-e^{-at})\big)^{-2a\theta/\sigma^2} \exp\Big(\frac{xue^{-at}}{1-(\sigma^2/2a)u(1-e^{-at})}\Big).$$

With a similar argument as above we get

$$E[Y_t^x] = \frac{\partial}{\partial u}\Big(E[e^{uY_t^x}]\Big)\Big|_{u=0} = \theta(1-e^{-at}) + xe^{-at}.$$

Altogether we get

$$\begin{aligned} E[X_t^x] &= E[Y_t^x] + E[Z_t] \\ &= (1-e^{-at})\big(\theta + \frac{1-e^{-at}}{a}\big) + xe^{-at} \\ &\leq \theta + \frac{1}{a} + xe^{-at}, \end{aligned}$$

namely

$$E[V(X_t^x)] \leq \theta + \frac{1}{a} + e^{-at}V(x).$$

□

If $\int_{(1,\infty)} \xi\nu(d\xi) < \infty$, then there exists a unique invariant probability measure for the JCIR process. This fact follows from [38, Thm. 3.16] and [32, Prop. 3.1]. Let $\|\cdot\|_{TV}$ denote the total-variation norm for signed measures on $\mathbf{R}_+$, namely

$$\|\mu\|_{TV} = \sup_{A\in\mathcal{B}(\mathbf{R}_+)} \{|\mu(A)|\}.$$

Let $P^t(x,\cdot) := P(X_t^x \in \cdot)$ be the distribution of the JCIR process at time t started from the initial point $x \geq 0$. Now we prove the main result of this chapter.

Theorem 12. *Assume that*

$$\int_{(1,\infty)} \xi\,\nu(d\xi) < \infty \quad \textit{and} \quad \int_{(0,1)} \xi \ln(\frac{1}{\xi})\nu(d\xi) < \infty.$$

Let π be the unique invariant probability measure for the JCIR process. Then the JCIR process is exponentially ergodic, namely there exist constants $0 < \beta < 1$ and $0 < B < \infty$ such that

$$\|P^t(x,\cdot) - \pi\|_{TV} \leq B\big(x+1\big)\beta^t, \quad t \geq 0, \quad x \in \mathbf{R}_+. \tag{4.21}$$

Proof. Basically, we follow the proof of [46, Thm. 6.1]. For any $\delta > 0$ we consider the δ-skeleton chain $\eta_n^x := X_{n\delta}^x,\ n \in \mathbf{Z}_+$, where $\mathbf{Z}_+$ denotes the set of all non-negative integers. Then $(\eta_n^x)_{n\in\mathbf{Z}_+}$ is a Markov chain on the state space $\mathbf{R}_+$ with transition kernel $P^\delta(x,\cdot)$ and starting point $\eta_0^x = x$. It is easy to see that the measure π is also an invariant probability measure for the chain $(\eta_n^x)_{n\in\mathbf{Z}_+}$, $x \geq 0$.

Let $V(x) = x$, $x \geq 0$. It follows from the Markov property and Lemma 12 that

$$E[V(\eta_{n+1}^x)|\eta_0^x, \eta_1^x, \cdots, \eta_n^x] = \int_{\mathbf{R}_+} V(y)P^\delta(\eta_n^x, dy) \leq e^{-a\delta}V(\eta_n^x) + M,$$

where M is a positive constant. If we set $V_0 := V$ and $V_n := V(\eta_n^x)$, $n \in \mathbf{N}$, then

$$E[V_1] \le e^{-a\delta} V_0(x) + M$$

and

$$E[V_{n+1}|\eta_0^x, \eta_1^x, \cdots, \eta_n^x] \le e^{-a\delta} V_n + M, \quad n \in \mathbf{N}.$$

Now we proceed to show that the chain $(\eta_n^x)_{n\in\mathbf{Z}_+}$, $x \ge 0$, is λ-irreducible, strong aperiodic, and all compact subsets of $\mathbf{R}_+$ are petite for the chain $(\eta_n^x)_{n\in\mathbf{Z}_+}$.

"*λ-irreducibility*": We show that the Lebesgue measure λ on $\mathbf{R}_+$ is an irreducibility measure for $(\eta_n^x)_{n\in\mathbf{Z}_+}$. Let $A \in \mathcal{B}(\mathbf{R}_+)$ and $\lambda(A) > 0$. Then it follows from Prop. 3 that

$$P[\eta_1^x \in A | \eta_0^x = x] = P(X_\delta^x \in A) \ge P(Z_\delta = 0) \int_A f(\delta, x, y) dy > 0,$$

since $f(\delta, x, y) > 0$ for any $x \in \mathbf{R}_+$ and $y > 0$. This shows that the chain $(\eta_n^x)_{n\in\mathbf{Z}_+}$ is irreducible with λ being an irreducibility measure.

"*Strong aperiodicity*" (see [44, p.561] for a definition): To show the strong aperiodicity of $(\eta_n^x)_{n\in\mathbf{Z}_0}$, we need to find a set $B \in \mathcal{B}(\mathbf{R}_+)$, a probability measure m with $m(B) = 1$, and $\epsilon > 0$ such that

$$L(x, B) > 0, \qquad x \in \mathbf{R}_+, \tag{4.22}$$

and

$$P(\eta_1^x \in A) \ge \epsilon m(A), \quad x \in C, \quad A \in \mathcal{B}(\mathbf{R}_+), \tag{4.23}$$

where $L(x, B) := P(\eta_n^x \in B \text{ for some n} \in \mathbf{N})$. To this end set $B := [0, 1]$ and $g(y) := \inf_{x\in[0,1]} f(\delta, x, y)$, $y > 0$. Since for fixed $y > 0$ the function $f(\delta, x, y)$ strictly positive and continuous in $x \in [0, 1]$, thus we have $g(y) > 0$ and $0 < \int_{(0,1]} g(y)dy \le 1$. Define

$$m(A) := \frac{1}{\int_{(0,1]} g(y)dy} \int_{A\cap(0,1]} g(y)dy, \qquad A \in \mathcal{B}(\mathbf{R}_+).$$

Then for any $x \in [0, 1]$ and $A \in \mathcal{B}(\mathbf{R}_+)$ we get

$$\begin{aligned} P(\eta_1^x \in A) &= P(X_\delta^x \in A) \\ &\ge P(Z_\delta = 0) \int_A f(\delta, x, y) dy \\ &\ge P(Z_\delta = 0) \int_{A\cap(0,1]} g(y) dy \\ &\ge P(Z_\delta = 0) m(A) \int_{(0,1]} g(y) dy, \end{aligned}$$

so (4.23) holds with $\epsilon := P(Z_\delta = 0) \int_{(0,1]} g(y)dy$.

Obviously

$$L(x, [0,1]) \geq P(\eta_1^x \in [0,1]) = P(X_\delta^x \in [0,1]) \geq P(Z_\delta = 0) \int_{[0,1]} f(\delta, x, y)dy > 0$$

for all $x \in \mathbf{R}_+$, which verifies (4.22).

"*Compact subsets are petite*": We have shown that λ is an irreducibility measure for $(\eta_n^x)_{n \in \mathbf{Z}_+}$. According to [44, Thm. 3.4(ii)], to show that all compact sets are petite, it suffices to prove the Feller property of $(\eta_n^x)_{n \in \mathbf{Z}_+}$, $x \geq 0$. But this follows from the fact that $(\eta_n^x)_{n \in \mathbf{Z}_+}$ is a skeleton chain of the JCIR process, which is an affine process and possess the Feller property.

According to [44, Thm. 6.3] (see also the proof of [44, Thm. 6.1]), the probability measure π is the only invariant probability measure of the chain $(\eta_n^x)_{n \in \mathbf{Z}_+}$, $x \geq 0$, and there exist constants $\beta \in (0,1)$ and $C \in (0, \infty)$ such that

$$\|P^{\delta n}(x, \cdot) - \pi\|_{TV} \leq C\big(x+1\big)\beta^n, \quad n \in \mathbf{Z}_+, \quad x \in \mathbf{R}_+.$$

Then for the rest of the proof we can proceed as in [46, p.536] and get the inequality (4.21). $\square$

Chapter 5

Non affine term structure models

Our purpose in this Chapter is to study the ergodicity properties of interest rate models driven by the Lévy process. These processes are more realistic models for the term structure. The insurance company DeBeKa are interested to use this model to describe the movement of interest rate during financial crisis.

5.1 Connection between OU-process and CIR-process in the jumps case

The model described in this section is also an extension of the CIR model, is extended to include Lévy noise. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$ be a filtered probability space satisfying the usual conditions. Let $(X_t)_{t\geq 0}$ satisfy the following stochastic differential equation

$$dX_t = (b - aX_t)dt + \sigma\sqrt{X_t}dL_t \tag{5.1}$$

with starting point x_0, where a, b and σ are positive constants and $(L_t)_{t\geq 0}$ is one dimensional Lévy process denotes the jumps in the process and is given by

$$dL_t = dW_t + \int_{0<|u|<1} uq(dt, du) \tag{5.2}$$

where $\{W_t, t \geq 0\}$ is a standard Brownian motion and $q(dt, du)$ be the compensated Poisson random measure (cPrm for short) of $\{N(dt, du)\}$, given by

$$q(dt, du) = N(dt, du) - dt\nu(du),$$

where $\{N(dt, du)\}$ is a Poisson random measure with intensity $dt\nu(du)$ where $\nu(du)$ is a σ-finite measure on $(0, \infty)$ satisfying

$$\int_0^\infty (1 \wedge u^2)\nu(du) < \infty.$$

We suppose also that $\{W_t, t \geq 0\}$ and $\{N(dt, du)\}$ are independent.

Let us first recall the assumptions which can guarantee the existence and uniqueness of the solutions to the non-Lipschitz stochastic differential equations driven by Lévy noise, for more details we refer to [52].

Xu and Pei only focus on the study of an equation driven by continuous noise interspersed with small jumps, the non-Lipschitz SDE driven by Lévy noise has the following form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t + \int_{|u|<c} F(t, X_t, u)q(dt, du), X_0 = \xi. \tag{5.3}$$

The mapping b, σ and F are all assumed to be measurable and ξ a random variable and $E|\xi^2| < \infty$.

Now, we consider the following assumptions on the coefficients of SDE (5.3) :

- For each fixed $t \in [0, T]$, let $b(t, x)$, $\sigma(t, x)$ and $F(t, x, u)$ be continuous in x and for all $x, y \in \mathbb{R}$,

  $$\begin{aligned} |b(t,x) - b(t,y)|^2 + |\sigma(t,x) - \sigma(t,y)|^2 &+ \int_{|u|<c} |F(t,x,u) - F(t,y,u)|^2 \nu(du) \\ &\leq \lambda(t) K\big(|x-y|^2\big), \end{aligned}$$

  where $\lambda : [0, \infty) \to \mathbb{R}_+$ is an integrable function and $K(q)$ or $\frac{K(q)^2}{q}$ is concave function , $K(0) = 0$ and $\int \frac{1}{K(q)} dq = \infty$

- $b(t, 0)$, $\sigma(t, 0)$ and $F(t, 0, u)$ are integrable.

Under these assumptions Xu and Pei prove that the SDE (5.3) has a unique solution. The existence and uniqueness of solution to the SDE (5.1) follow from the main results of Xu and Pei [52, Thm 1,2].

Remark 13. The singularity of the noise term $\sqrt{X_t} dL_t$ at the origin does not guarantee that the solution of the above SDE 5.1 will always be non negative, since the process $(X_t)_{t \geq 0}$ can simply jump over the origin and to the negative part of the real line.

To overcome this disadvantage we establish the connection between this model and the square of Ornstein-Uhlenbeck process with jumps. Let $(Y_t)_{t\geq 0}$ be an Ornstein-Uhlenbeck process satisfying the following stochastic differential equation

$$dY_t = -\frac{a}{2}Y_t dt + \frac{\sigma}{2}dL_t \tag{5.4}$$

where a and σ are positive constants and $(L_t)_{t\geq 0}$ is a Lévy process of the form 5.2. This process is the classical Ornstein-Uhlenbeck SDE with the Browian motion replaced by a Lévy process. This is exist because of growth and Lipschitz conditions.

Lemma 13. *Now, we set* $Z_t = Y_t^2$*, the process* $(Z_t)_{t\geq 0}$ *is a solution to the following SDE*

$$\begin{aligned} dZ_t &= (\frac{\sigma^2}{4} - aZ_t)dt + \sigma\sqrt{Z_t}d\tilde{L}_t \\ &+ \int_{0<|u|<1} \frac{\sigma^2}{4}u^2\nu(du) + \int_{0<|u|<1} \frac{\sigma^2}{4}u^2 q(dt,du). \end{aligned}$$

where

$$\tilde{L}_t := \int_0^t sign(Y_s)dL_s = \int_0^t sign(Y_s)dW_s + \int_0^t \int_{0<|u|<1} sign(Y_s)uq(ds,du)$$

and $sign(x) := \begin{cases} +1, & for\ x \geq 0, \\ -1, & for\ x < 0. \end{cases}$

To prove this Lemma, let us first recall the Itô formula with respect to the cPrm. Let $(X_t)_{t\geq 0}$ be a one-dimensional Lévy process given by:

$$dX_t = G(X_t)dt + F(X_t)dW_t + \int_{0<|u|<1} H(t,u)q(dt,du), \qquad X_0 = x.$$

For $f \in \mathcal{C}^2(\mathbb{R})$ we have

$$\begin{aligned} df(X_t) &= \partial_x f(X_t)G(X_t)dt + \partial_x f(X_t)F(X_t)dW_t + \frac{1}{2}\partial_{xx}f(X_t)F^2(X_t)dt \\ &+ \int_{0<|u|<1} \big(f(X_t + H(t,u)) - f(X_t) - H(t,u)\partial_x f(X_t)\big)\nu(du)dt \\ &+ \int_{0<|u|<1} \big(f(X_t + H(t,u)) - f(X_t)\big)q(dt,du), \end{aligned}$$

we refer to ([43, Section 3.7]).

Proof. We denote $f(Y_t) = Y_t^2$ by applying the previous Itô formula one can get

$$\begin{aligned} df(Y_t) &= -aY_t^2 dt + +\sigma Y_t dW_t + \frac{\sigma^2}{4} dt \\ &+ \int_{0<|u|<1} \big((Y_t + \frac{\sigma}{2}u)^2 - Y_t^2 - 2Y_t \frac{\sigma}{2} u\big) dt\nu(du) \\ &+ \int_{0<|u|<1} \big((Y_t + \frac{\sigma}{2}u)^2 - Y_t^2\big) q(dt, du) \\ &= (\frac{\sigma^2}{4} - aY_t^2)dt + \sigma Y_t dW_t + \int_{0<|u|<1} \sigma Y_t u q(dt, du) \\ &+ \int_{0<|u|<1} \frac{\sigma^2}{4} u^2 dt\nu(du) + \int_{0<|u|<1} \frac{\sigma^2}{4} u^2 q(dt, du) \end{aligned}$$

then we get

$$\begin{aligned} dZ_t &= (\frac{\sigma^2}{4} - aZ_t)dt + \sigma\sqrt{Z_t} d\tilde{L}_t \\ &+ \int_{0<|u|<1} \frac{1}{4}\sigma^2 u^2 dt\nu(du) + \int_{0<|u|<1} \frac{1}{4}\sigma^2 u^2 q(dt, du) \end{aligned}$$

where

$$\tilde{L}_t := \int_0^t sign(Y_s) dL_s = \int_0^t sign(Y_s) dW_s + \int_0^t \int_{0<|u|<1} sign(Y_s) u q(ds, du)$$

One can remark that $\tilde{L}_t$ and $\int_{0<|u|<1} u^2 q(dt, du)$ are independent. The only difficulty is how to show that $\{\tilde{L}_t, t \geq 0\}$ is a Lévy process.

First, we will show that $\tilde{W}_t = \int_0^t sign(Y_s) dW_s$ is a Brownian motion. Our idea is to use the time change for the Brownian motion.

Let us first recall time change of Brownian motion. Let $c(t, \omega) \geq 0$ be an $\mathcal{F}_t$-adapted process. Define

$$\beta_t = \beta(t, \omega) := \int_0^t c(s, \omega) ds. \tag{5.5}$$

Note that β_t is also an $\mathcal{F}_t$-adapted process and for each ω the map $t \to \beta_t(\omega)$ is non-decreasing. Define

$$\alpha_t = \alpha(t, \omega) := \inf\{s : \beta_s > t\} \tag{5.6}$$

Following [48, Chapter 8] we recall the Theorem $8.5.2$:

Theorem 14. *Let* $d\tilde{B}_t = v(t,\omega)dB_t$, $v \in \mathbb{R}^{n\times m}$, $B_t \in \mathbb{R}^n$ *be an Itô integral in* $\mathbb{R}^n$, $Y(0) = 0$ *and assume that*

$$vv^T(t,\omega) = c(t,\omega)I_n,$$

for some process $c(t,\omega)$. *Let* α_t *and* β_t *as in (5.6) and (5.5) respectively then*

$$\tilde{B}_{\alpha_t} \quad \text{is an n-dimensional Brownian motion.}$$

By using time change we will prove that $\tilde{W}_t$ we define

$$\tilde{W}_t := \int_0^t sign(Y_s)dW_s.$$

In this case, we have $v(t,\omega) = sign(Y(t,\omega))$, $c(t,\omega) = sign(Y(t,\omega))^2 = 1$ and $\beta_t := \int_0^t \big(sign(Y_s)\big)^2 ds = t$.
Let

$$\alpha_t := \inf\{s : \beta_s > t\} = t.$$

After the time change, we know from theorem (14) that $\tilde{W}_{\alpha_t}$ is a Brownian motion. Since $\alpha_t = t$, we conclude that $\tilde{W}_t$ is also a Brownian motion. For full details for the time change the readers are referred to [48, Section 8.5].

Next step, is how to show that

$$\int_0^t \int_{0<|u|<1} sign(Y_s)uq(ds,du)$$

is still a Lévy process?

Note that this term is in terms of stochastic integral. If we assume that N(dt,du) is symmetric. The driving process of OU process is independent increment symmetric process.

If we succeed to prove this fact then we can write the CIR process with jumps 5.1 as a sum of the square of Ornstein-Uhlenbeck driven by jumps 5.4 and a positive drifts as follows

$$X_t = Y_t^2 + (b - \frac{\sigma^2}{4})t = \int_0^t \int_{0<|u|<1} \frac{1}{4}\sigma^2 u^2 \nu(du)ds + \int_0^t \int_{0<|u|<1} \frac{1}{4}\sigma^2 u^2 q(ds,du).$$

□

5.2 Some investigations

For term structure models driven by a Gaussian noise, Bhan and Mandrekar introduced a simple method for testing the stability of various important financial models like CIR and Vasicek by explicitly calculating their moments. They investigated the recurrence properties and ultimate boundedness of these models and hence the existence of invariant measure (see [4]). Motivated by the work of Govidan and Abreu [18], Bhan et al. in [5] proved the weak positive recurrence of the interest rate models driven by Lévy process and the existence of the invariant measure which follows from the exponential ultimate boundedness. We already derived in the last section, the equation for the square of Ornstein-Uhlenbeck process driven by jump process which is a special case of the processes considered in [5]. Hence, we have the weak recurrence and the existence of the invariant measure of this process.

Unfortunately these models are non affine term structure models, so we need a new mathematical tools to analyze Harris recurrence and exponential ergodicity.

First of all, we will consider the case when the Lévy process $(L_t)_{t\geq 0}$ is α-stable. Then we check in the literature how to analyse the Harris recurrence and the ergodicity for this type of process. Hence, we generalize the application in one credit migration model shown in the Chapter 2. More precisely, how one can calibrate the parameters of the non affine term structure model based on the ergodicity results.

Bibliography

[1] Azéma, J., Kaplan-Duflo, M. and Revuz, D.: Mesure invariante sur les classes récurrentes des processus de Markov, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete,* **8** (1967), 157–181.

[2] Bandi, F.M. and Nguyen, T.H.: On the functional estimation of jump-diffusion models, *J. Econometrics,* **116** (2003), 293–328.

[3] Ben Alaya, M. and Kebaier, A.: Parameter Estimation for the Square-root Diffusions: Ergodic and Nonergodic Cases, *Preprint,* (2011).

[4] Bhan, C. and Mandrekar, V.: Recurrence properties of term structure models, *Int. J. Contemp. Math. Sci.,* **3**(28)(2010), 1645–1652.

[5] Bhan, C., Chakraborty, P. and Mandrekar, V.: Invariant Measure and Stability of the Solution of a Stochastic Differential Equation Driven by a Jump Lévy Proces, *Int. J. Contemp. Math. Sci.,* **7**(1) (2012),33-52.

[6] Brigo, D. and Mercurio, F.: Interest Rate Models - Theory and Practice with Smile, Inflation and Credit, *S*pringer Verlag, (2001).

[7] Cox, J.C., Ingersoll, J.J.E., and Ross, S.A.: A theory of the term structure of interest rates, *Econometrica,* **53**(2) (1985), 385–407.

[8] Delbaen, F. and Shirakawa, H.: A Note of Option pricing for constant elasticity of variance model, *Preprint* (1996).

[9] Duffie, D. and Filipović, D. and Schachermayer, W.: Affine processes and applications in finance, *Ann. Appl. Probab.,* **13**(3) (2003), 984–1053.

[10] Duffie, D., Gârleanu, N.: Risk and valuation of collateralized debt obligations, *F*inancial Analysts Journal, **57**(1) (2001), 41–59.

[11] Feller, W.: Two singular diffusion problems, *Ann. of Math.,* **54**(1) (1951), 173–182.

[12] Feller, W.: An Introduction to Probability Theory and Its Applications, Vol. II. Second edition, John Wiley & Sons, Inc., New York-London-Sydney (1971)

[13] Filipović, D.: A general characterization of one factor affine term structure models, *Finance Stoch.,* 5(3) (2001), 389–412.

[14] Filipović, D., Mayerhofer, E., and Schneider, P.: Density approximations for multivariate affine jump-diffusion processes, *J. Econometrics,* 176(2) (2013), 93–111.

[15] Fu, Z.F., and Li, Z.H.: Stochastic equations of non-negative processes with jumps, *Stochastic Process. Appl.,* 120(3)(2010), 306–330.

[16] Getoor, R.K. 1980.: *Transience and recurrence of Markov processes*, Seminar on Probability, XIV (French) Lecture Notes in Math., vol. 784, Springer, Berlin, (Paris, 1978-1979), 397–409.

[17] Göing-Jaeschke, A. and Yor, M.: A survey and some generalizations of Bessel processes, *Bernoulli,* **9**(2) (2003), 313–349.

[18] Govindan, T.E., and Abreu, R.S.: Stability Behavior of Some Well-known Stochastic Financial Models, *Int. Jour. Contemp. Math. Sciences,* **3**(28) (2008), 1367–1378.

[19] Grigelionis, B.: Thorin classes of Lévy processes and their transforms, *Lith. Math. J.,* **48**(3) (2008), 294–315.

[20] Harris, T.E.: The existence of stationary measures for certain Markov processes, in: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* **II** (1956), 113–124, University of California Press, Berkeley and Los Angeles.

[21] Heston, S.L.: A closed-form solution for options with stochastic volatility, with applications to bond and currency options, *Review of Financial studies,* **6**(2) (1993), 327–343.

[22] Hurd T.R. and Kuznetsov A.: Affine Markov chain models of multifirm credit migration, *J. of credit risk,* **3** (2007), 3–29.

[23] Ikeda, N. and Watanabe, S.: *Stochastic differential equations and diffusion processes*, North-Holland Publishing Co., Amsterdam, (1989).

[24] Israel, R.B., Rosenthal, J.S. and Wei, J.Z.: Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings, *Math. Finance,* **11**(2) (2001), 245–265.

[25] Jeanblanc, M., Yor, M. and Chesney, M.: Mathematical Methods for Financial Markets, Springer, London, (2009).

[26] Jin, P., Mandrekar, V., Rüdiger, B., and Trabelsi, C.: Positive Harris recurrence of the CIR process and its applications, *Commun. Stoch. Anal.,* **7**(3) (2013), 409–424.

[27] Jin, P., Rüdiger, B., Trabelsi, C.: Positive Harris recurrence and exponential ergodicity of the basic affine jump-diffusion, *Stochastic Analysis and Appl.* **34**(1) (2015), 75–95.

[28] Jin, P., Rüdiger, B., Trabelsi, C.: Exponetial Ergodicity of the Jump-Diffusion CIR Process, *Springer Proceedings in Mathematics and statistics, Stochastics of Environmental and Financial Economics,* CAS, Oslo, Norway, (2014- 2015), 285–300.

[29] Kallenberg, O.: *Foundations of modern probability*, Springer-Verlag., New York, (2002).

[30] Karatzas, I. and Shreve, S. E.: *Brownian Motion and Stochastic Calculus*, Springer-Verlag, 2nd edition, (1991).

[31] Kaspi, H. and Mandelbaum, A.: On Harris recurrence in continuous time, *Math. Oper. Res.,* **19**(1) (1994), 211–222.

[32] Keller-Ressel, M.: Moment explosions and long-term behavior of affine stochastic volatility models, *M*ath. Finance, **21**(1) (2011), 73–98.

[33] Keller-Ressel, M., Schachermayer, W. and Teichmann, J.: Affine processes are regular, *Probab. Theory Related Fields,* **151** (2011), 591–611.

[34] Keller-Ressel, M., and Mayerhofer, E.: Exponential moments of affine processes, *Ann. Appl. Probab.,* **25**(2), (2015), 714–752.

[35] Keller-Ressel, M., and Mijatović, A.: On the limit distributions of continuous-state branching processes with immigration, *Stochastic Process. Appl.,* **122**(6), (2012), 2329–2345.

[36] Keller-Ressel, M., Schachermayer, W., and Teichmann, J.: Affine processes are regular, *Probab. Theory Related Fields,* **151**(3-4), (2011), 591–611.

[37] Keller-Ressel, M., Schachermayer, W., Teichmann, J.: Regularity of affine processes on general state spaces, *E*lectron. J. Probab., **18**(43) (2013).

[38] Keller-Ressel, M., and Steiner, T.: Yield curve shapes and the asymptotic short rate distribution in affine one-factor models, *Finance Stoch.,* **12**(2), (2008),149–172.

[39] Koehler, B. and Krafft,V.: Ein Kreditmodell basierend auf Migrationsmatrizen, *Preprint* (2011).

[40] Kunita, Hiroshi: *Stochastic flows and stochastic differential equations*, Cambridge University Press, Cambridge, (1997).

[41] Lamberton, D. and Lapeyre, B.: Intoduction to stochastic calculus applied to finance, *C*hapman and Hall, London, (1996).

[42] Li, Z.H., and Ma, C.H.: Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model, *Stochastic Process. Appl.,* **125**(8) (2015), 3196–3233.

[43] Mandrekar, V. and Rüdiger, B.: Stochastic Integration in Banach space, Theory and Applications, *Springer Cham Heidelberg New York Dordrecht London*, (2015).

[44] Meyn, S.P., and Tweedie, R.L.: Stability of Markovian processes. I. Criteria for discrete-time chains, *Adv. in Appl. Probab.,* **24**(3) (1992), 542–574.

[45] Meyn, S.P., and Tweedie, R.L.: Stability of Markovian processes. II. Continuous-time processes and sampled chains, *Adv. in Appl. Probab.,* **25**(3) (1993), 487–517.

[46] Meyn, S.P., and Tweedie, R.L.: Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes, *Adv. in Appl. Probab.,* **25**(3) (1993), 518–548.

[47] Meyn, S.P., and Tweedie, R.L.: Markov Chains and Stochastic Stability, *Cambridge University*, (2009).

[48] Øksendal, B.: Stochastic Differential Equations, *Springer-Verlag*, (2007).

[49] Revuz, D. and Yor, M.: *Continuous martingales and Brownian motion*, Springer-Verlag, Berlin, (1999).

[50] Szatzschneider, W: Extended Cox, Ingersoll & Ross Model, *Revista Mexicana de Economía y Finanzas,* **1**(4) (2012).

[51] Vasicek, O.A.: *An equilibrium characterization of the term structure*, *J*ournal of Financial Economics, **5** , 177-188, (1977).

[52] Xu, Y. and Pei, B.: *Existence and Stability of Solutions to Non-Lipschitz Stochastic Differential Equations Driven by Lévy Noise, A*pplied Mathematics and computation**263** (C), 398-409, (2015).