

THE CONTINUOUS-TIME EHRENFEST PROCESS IN TERM STRUCTURE MODELLING

A. KAPLUN,* *TU Dortmund University*

Abstract

In this paper, a finite-state mean-reverting model for the short rate, based on the continuous-time Ehrenfest process, will be examined. Two explicit pricing formulae for zero-coupon bonds will be derived in the general and special symmetric cases. Its limiting relationship to the Vasicek model will be examined with some numerical results.

Keywords: Ehrenfest model; interest rate derivative; short rate; term structure; Vasicek model; zero-coupon bond

2010 Mathematics Subject Classification: Primary 91G30

Secondary 60J28; 91G20; 33C52

1. Introduction

One of the fundamental approaches to term structure modelling is based on the specification of the short-term interest rate—the short rate. Vasicek [19] first introduced a mean-reverting short-rate model with Gaussian distribution and derived a closed-form representation for the zero-coupon bond (ZCB) price. Since then, a variety of short-rate models have become established, each having its advantages and disadvantages.

Albeit the earliest, Vasicek's model is still very popular among practitioners owing to its analytical tractability with regard to ZCB prices and European options thereof. Unfortunately, the model has some shortcomings. The most prominent of these is the possibility that the interest rates become negative—a fact relating to all models with Gaussian distribution. Even though the probability of negative rates is rather small, not only is the realism of the model questionable, but also problems may appear while valuing ZCBs with a long time to maturity and a low interest rate level.

The idea of using both the discrete- and continuous-time versions of the Ehrenfest process in finance is well known. The discrete-time approach was used, for example, by Okunev and Tippett [13] in modelling accumulated cashflows, by Takahashi [18] in exploring changes in stock prices and exchange rates for currencies, and by Bühlmann [5] in modelling interest rates. With regard to the modelling of interest rates, it seems that the discrete-time approach leads in general only to a recursively computable term structure. Sumita *et al.* [16] studied the passage times and the historical maximum of the Ornstein–Uhlenbeck process via an approximation by means of a special case of the continuous-time Ehrenfest process.

In this paper we propose a finite-state mean-reverting model for the short rate related to the continuous-time Ehrenfest process. By choosing arbitrary lower and upper bounds for the rate, the respective short-rate process can be seen as a suitably linearly transformed birth-and-death process on $\{0, 1, \dots, N\}$, $N \in \mathbb{N}$. By choosing the lower bound to be nonnegative, the problem

Received 5 October 2009; revision received 5 May 2010.

* Postal address: Department of Mathematics, TU Dortmund University, Vogelpothsweg 87, D-44221 Dortmund, Germany. Email address: alexander.kaplun@math.uni-dortmund.de

of negative interest rates can be avoided. Furthermore, the model allows an explicit evaluation of ZCB prices. In this way, the model aims at realism and analytical tractability.

The main outcome of the paper is the derivation of pricing formulae for ZCBs in the general and the special symmetric cases of the model. In both cases the arbitrage-free ZCB price at time t and maturity T is given as follows:

$$P(t, T) = C P_1(t, T)^k P_0(t, T)^{N-k},$$

where C is a constant, $k \in \{0, 1, \dots, N\}$, and P_1 and P_0 can be expressed in terms of the ${}_1F_1$ hypergeometric functions of a matrix argument given in Section 2 (see also [7]). In the general case the model is governed by five parameters—a valuable fact considering the fitting of the model to the market data. The special case provides four parameters and is characterized by the symmetry of the underlying distribution with respect to the mean-reverting value. The advantage here is that we have more tractable expressions for P_1 and P_0 from the computational point of view. Moreover, a suitably transformed symmetric case of the model yields the Vasicek model in the limit as N tends to ∞ .

The paper is organized as follows. In Section 2 we give a short review of the special functions we will encounter throughout this paper. Section 3 deals with the Ehrenfest process in continuous time. The main results are given in Section 4, where the *Ehrenfest short-rate model* is defined and the ZCB pricing formulae are derived. Section 5 gives an overview of the Vasicek model and its limiting relationship to the Ehrenfest short-rate model. Section 6 illustrates the advantages of the Ehrenfest short-rate model.

2. Special functions and orthogonal polynomials

Throughout this paper we make use of some well-known facts concerning the *Krawtchouk polynomials* (see [8], [17, Section 2.82], and [20]) and ${}_1F_1$ functions (see [7]), as well as some of their practical implications. In the interests of clarity we give in this section an overview of these special functions.

2.1. Hypergeometric functions of matrix argument

Definition 1. (a) A *partition* $m := (m_1, m_2, \dots, m_n)$ is an n -tuple ($n \in \mathbb{N}$) of nonnegative integers such that $m_1 \geq m_2 \geq \dots \geq m_n$.

(b) For a partition m , the *generalized Pochhammer symbol* is defined by

$$[a]_m := \prod_{j=1}^n (a - j + 1)_{m_j},$$

where $(a)_k := a(a + 1) \cdots (a + k - 1)$ denotes the usual *Pochhammer symbol*.

(c) For a partition m , the *normalized Schur function of index m* is defined by

$$Z_m(z) := |m|! \frac{\prod_{1 \leq j < k \leq n} (m_j - m_k - j + k)}{\prod_{j=1}^n (m_j + n - j)!} \frac{\det(z_i^{m_j + n - j})}{\prod_{1 \leq i < j \leq n} (z_i - z_j)}.$$

(d) The *hypergeometric function of a matrix argument* ${}_pF_q$ is defined as a real-analytic function on the space S_n ($n \in \mathbb{N}$) of $n \times n$ Hermitian matrices with eigenvalues $z := (z_1, \dots, z_n)^T \in \mathbb{R}^n$,

and is given by the series

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; z) := \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{|m|=j} \frac{[a_1]_m \cdots [a_p]_m}{[b_1]_m \cdots [b_q]_m} Z_m(z), \tag{1}$$

where, for $1 \leq i \leq p$ and $1 \leq j \leq q$, $a_i \in \mathbb{C}$ and none of the numbers $-b_i + j - 1$ is a nonnegative integer.

We will be particularly concerned with the ${}_1F_1$ function, which is also known as the *confluent hypergeometric function of a matrix argument*. From Theorem 4.1 of [7] we know that it converges absolutely for all $z \in S_n$. An important result, which will be crucial later on, is given without proof in the following remark (see [7, p. 25]).

Remark 1. Let Δ_n denote the *standard simplex* in \mathbb{R}^n ($n \in \mathbb{N}$), defined by

$$\Delta_n := \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0, i = 1, \dots, n, \sum_{i=1}^n x_i \leq 1 \right\}. \tag{2}$$

Then, for $a > 0$, the following equation holds:

$${}_1F_1(1; a + n; z) = (a)_n \int_{\Delta_n} \left(1 - \sum_{i=1}^n x_i \right)^{a-1} \exp\left(\sum_{i=1}^n z_i x_i \right) dx_1 \cdots dx_n. \tag{3}$$

In order to compute the ${}_pF_q$ function numerically, we truncate the series in (1) at $m \leq H$ as follows:

$${}_pF_q^H(a_1, \dots, a_p; b_1, \dots, b_q; z) := \sum_{j=0}^H \frac{1}{j!} \sum_{|m|=j} \frac{[a_1]_m \cdots [a_p]_m}{[b_1]_m \cdots [b_q]_m} Z_m(z). \tag{4}$$

Koev and Edelman [10] provided an effective algorithm for computing the ${}_pF_q^H$ function. For $z \in S_n$, the complexity of their algorithm is linear in n and subexponential in H , which is acceptable if we consider the fast convergence of the power series (4).

2.2. Krawtchouk polynomials

For given $N \in \mathbb{N}$ and $0 < p = 1 - q < 1$, the Krawtchouk polynomials $K_l(x) := K_l(x; N; p)$ are the orthogonal polynomials that relate to the binomial distribution $B_{N,p}$ and the probability mass function $\omega(x) := \binom{N}{x} p^x q^{N-x}$ at the points $x = 0, 1, \dots, N$. They can be defined in two different, but equivalent ways.

Definition 2. For $x, l \in \{0, 1, \dots, N\}$, we set

$$K_l(x) := {}_2F_1\left(-l, -x; -N; \frac{1}{p}\right) := \sum_{k=0}^N \frac{(-l)_k (-x)_k}{(-N)_k k!} \left(\frac{1}{p}\right)^k, \tag{5}$$

or

$$K_l(x) := \binom{N}{l}^{-1} \sum_{k=0}^N (-1)^k \binom{N-x}{l-k} \binom{x}{k} \left(\frac{q}{p}\right)^k, \tag{6}$$

where ${}_2F_1$ is the classical Gauss hypergeometric function (see [17, Section 4.21]).

Definition 2 leads to the following basic well-known properties.

Lemma 1. (a) *Symmetry:*

$$K_l(x) = K_x(l) \tag{7}$$

for all $x, l \in \{0, 1, \dots, N\}$.

(b) $K_0(x) = K_l(0) = 1$ for all $x, l \in \{0, 1, \dots, N\}$.

(c) $K_1(x) = 1 - x/Np$.

(d) *Generating function:*

$$\left(1 - \frac{q}{p}s\right)^i (1 + s)^{N-i} = \sum_{k=0}^N \binom{N}{k} K_l(i) s^k \tag{8}$$

for all $x \in \{0, 1, \dots, N\}$.

(e) *Recurrence relation:*

$$-xK_l(x) = (N - l)pK_{l+1}(x) - [(N - l)p + lq]K_l(x) + lqK_{l-1}(x) \tag{9}$$

for all $x, l \in \{0, 1, \dots, N\}$.

(f) *Orthogonality relation:*

$$\sum_{x=0}^N K_l(x)K_m(x)\omega(x) = \frac{\delta_{l,m}}{\pi_l}, \tag{10}$$

where

$$\pi_l := \binom{N}{l} \left(\frac{p}{q}\right)^l = \omega(l)q^{-N}$$

for all $l, m \in \{0, 1, \dots, N\}$.

(g) For $l, m \in \{0, 1, \dots, N\}$, $m \leq l$,

$$B_{m,l} := \sum_{x=0}^N xK_l(x)K_m(x)\omega(x) = \begin{cases} 0 & \text{if } m \leq l - 2, \\ -lq/\pi_{l-1} & \text{if } m = l - 1, \\ ((N - l)p + lq)/\pi_l & \text{if } m = l. \end{cases} \tag{11}$$

Note that $B_{l,m} = B_{m,l}$.

Proof. Parts (a)–(e) follow from (5), part (f) follows from (6), and part (g) is a direct application of (9) and (10) (see, e.g. [17, Section 2.82]).

3. Original Ehrenfest model

The original Ehrenfest model describes the heat exchange between two isolated bodies, each of arbitrary temperature. The temperatures are symbolized by the number of fluctuating balls in two urns with a total of $N \in \mathbb{N}$ balls. For details of the continuous- and discrete-time versions of the model, we refer the reader to [2], [8], [11], and [15].

In this section we shall discuss the continuous-time Ehrenfest process. Primarily, its representation as a sum of independent simple processes will allow us to show the main result of this paper, which we state in the next section. Furthermore, we examine the transition semigroup of the Ehrenfest process and explore some of its basic properties.

3.1. Ehrenfest process

Let N balls, initially distributed between urns I and II in such a way that k balls are in the urn I and $N - k$ balls are in the urn II, fluctuate independently in continuous time between the two urns. We fix a fluctuation parameter $\lambda > 0$ and independent Poisson processes $(N_t^1)_{t \geq 0}, \dots, (N_t^N)_{t \geq 0}$ with intensity λ . Let $(\hat{Y}_n)_{n \in \mathbb{N}}$ be a Markov chain with the state space $\{0, 1\}$ and transition probability matrix

$$P := \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}. \tag{12}$$

Then, the subordinated Markov chain $(Y_t^l := \hat{Y}_{N_t^l})_{t \geq 0}$ describes the state of the l th ball at time t , where $Y_t^l = 1$ or 0 when the l th ball is in urn I or II, respectively. Let k balls be initially in urn I and $N - k$ balls in urn II. Hence,

$$\left(X_t := \sum_{l=1}^N Y_t^l \right)_{t \geq 0} \tag{13}$$

is a Markov process with the state space $E := \{0, 1, \dots, N\}$, denoting the number of balls in urn I at time t . We call $(X_t)_{t \geq 0}$ the (continuous-time) Ehrenfest process. A discrete-time version of (13) with arbitrary α and β was studied by Krafft and Schaefer [11].

Remark 2. (a) A special case of (13) with $\alpha = \beta = 1$, first suggested by Siegert [15] and also studied by Bingham [2], where the transitions become ‘deterministic’ in the sense of switching between states 0 and 1, will be important for us later on in Section 4. (Its discrete-time analogue leads to the original Ehrenfest chain.)

(b) Karlin and McGregor [8] provided an alternative but equivalent definition of (13) as a birth-and-death process with the state space E . Here, the time intervals between events are independently exponentially distributed with intensity γ , and, for $i \in E$, the birth and death rates are $\lambda_i := \gamma\alpha(N - i)/N$ and $\mu_i = \gamma\beta i/N$, respectively, where α and β are given as above. It can be verified that in the setting at hand $\gamma = \lambda N$.

Before computing the transition semigroup of $(X_t)_{t \geq 0}$ we need the following result.

Lemma 2. The transition semigroup of $(Y_t^l)_{t \geq 0}$ is given by

$$P(t) = \begin{pmatrix} q + pe^{-\lambda(\alpha+\beta)t} & p - pe^{-\lambda(\alpha+\beta)t} \\ q - qe^{-\lambda(\alpha+\beta)t} & p + qe^{-\lambda(\alpha+\beta)t} \end{pmatrix}, \tag{14}$$

where $p := \alpha/(\alpha + \beta)$ and $q := 1 - p$.

Proof. Since $(Y_t^l)_{t \geq 0}$ is a Markov chain subordinated by a Poisson process with index λ , the associated transition semigroup can be written as (see [3, p. 333])

$$P(t) = e^{-\lambda t} e^{\lambda t P} := e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} P^n. \tag{15}$$

We can avoid the computation of P^n by writing $P = \mu S + (1 - \mu)I$, where

$$S = \begin{pmatrix} q & p \\ q & p \end{pmatrix}$$

is a stochastic matrix with $S^2 = S$, I is the identity matrix, and $\mu := \alpha + \beta$. Then, (15) becomes

$$\begin{aligned} P(t) &= e^{-\lambda t} e^{\lambda t(\mu S + (1-\mu)I)} \\ &= e^{-\lambda t} e^{\lambda \mu t S} e^{\lambda(1-\mu)tI} \\ &= e^{-\lambda \mu t} e^{\lambda \mu t S} \\ &= e^{-\lambda(\alpha+\beta)t} \sum_{n=0}^{\infty} \frac{(\lambda(\alpha + \beta)t)^n}{n!} S^n. \end{aligned}$$

Since $S^n = S$ for all $n \in \mathbb{N}$, we can easily compute the above series, which completes the proof.

Remark 3. Analogously to the proof above, we see that a variation of α and β in (14) can be equivalently described as a suitable modification of p and λ . Thus, the distribution of $(Y_t)_{t \geq 0}$ depends only on the two parameters p and λ or, equivalently, α and β .

Theorem 1. (Properties of the Ehrenfest process.) *Let $(X_t)_{t \geq 0}$ be the Ehrenfest process given by (13).*

(a) *The transition probabilities $p_{ij}(t) := P(X_{t+s} = j \mid X_s = i)$ are given by*

$$p_{ij}(t) = \binom{N}{j} \left(\frac{p}{q}\right)^j \sum_{x=0}^N \binom{N}{x} p^x q^{N-x} K_i(x) K_j(x) e^{-\lambda(\alpha+\beta)xt}, \tag{16}$$

where $p := \alpha/(\alpha + \beta)$, $q := 1 - p$, and $K_i(\cdot) := K_i(\cdot; p; N)$ ($i \in E$) are the Krawtchouk polynomials as given in Definition 2.

(b) *The conditional mean and variance of $(X_t)_{t \geq 0}$ are given by*

$$E[X_t \mid X_0 = i] = Np - (Np - i)e^{-\lambda(\alpha+\beta)t}, \tag{17}$$

$$\begin{aligned} \text{var}[X_t \mid X_0 = i] &= Np(1 - p) + (Np - i)(2p - 1)e^{-\lambda(\alpha+\beta)t} \\ &\quad - (Np - i)^2(2p - 1)e^{-2\lambda(\alpha+\beta)t}. \end{aligned} \tag{18}$$

(c) *The limiting distribution of $(X_t)_{t \geq 0}$ is a stationary distribution and is given by the binomial distribution $B_{N,p}$ on E with parameter p .*

Proof. (a) The proof here is similar to that in [2]. First, we compute the moment generating function of $(Y_t^l)_{t \geq 0}$, given Y_0^l . In the following we suppress the dependence of $(Y_t^l)_{t \geq 0}$ on l when it is clear from the context. From Lemma 2 we have

$$\begin{aligned} E[z^{Y_t} \mid Y_0 = 1] &= p_{10}(t) + p_{11}(t)z = q + pz - q(1 - z)e^{-\lambda(\alpha+\beta)t}, \\ E[z^{Y_t} \mid Y_0 = 0] &= p_{00}(t) + p_{01}(t)z = q + pz + p(1 - z)e^{-\lambda(\alpha+\beta)t}. \end{aligned}$$

Since $(Y_t^l)_{t \geq 0}$ are independent for all $l = 1, \dots, N$, we obtain from (13) the moment generating

function of $(X_t)_{t \geq 0}$, given $X_0 = i$, as follows:

$$\begin{aligned} \sum_{j=0}^N p_{ij}(t)z^j &= E\left[z^{\sum_{l=1}^N Y_l^i} \mid X_0 = i\right] \\ &= E\left[\prod_{l=1}^N z^{Y_l^i} \mid X_0 = i\right] \\ &= E[z^{Y_t} \mid Y_0 = 1]^i E[z^{Y_t} \mid Y_0 = 0]^{N-i} \\ &= [q + pz - q(1 - z)e^{-\lambda(\alpha+\beta)t}]^i [q + pz + p(1 - z)e^{-\lambda(\alpha+\beta)t}]^{N-i} \\ &= (q + pz)^N \left[1 - \frac{q}{p} \frac{p(1 - z)}{q + pz} e^{-\lambda(\alpha+\beta)t}\right]^i \left[1 + \frac{p(1 - z)}{q + pz} e^{-\lambda(\alpha+\beta)t}\right]^{N-i}. \end{aligned}$$

Applying (a) and (d) of Lemma 1, we obtain

$$\begin{aligned} \sum_{j=0}^N p_{ij}(t)z^j &= (q + pz)^N \sum_{x=0}^N \binom{N}{x} K_i(x) \left(\frac{p(1 - z)}{q + pz} e^{-\lambda(\alpha+\beta)t}\right)^x \\ &= \sum_{x=0}^N \binom{N}{x} K_i(x) (q + pz)^{N-x} p^x (1 - z)^x e^{-\lambda(\alpha+\beta)xt} \\ &= \sum_{x=0}^N \binom{N}{x} p^x q^{N-x} K_i(x) \left(1 - \frac{q}{p} \frac{pz}{q}\right)^x \left(1 + \frac{pz}{q}\right)^{N-x} e^{-\lambda(\alpha+\beta)xt}. \end{aligned}$$

Applying (8) once again, we obtain

$$\begin{aligned} \sum_{j=0}^N p_{ij}(t)z^j &= \sum_{x=0}^N \binom{N}{x} p^x q^{N-x} K_i(x) \left[\sum_{j=0}^N \binom{N}{j} K_j(x) \left(\frac{pz}{q}\right)^j\right] e^{-\lambda(\alpha+\beta)xt} \\ &= \sum_{j=0}^N \left[\binom{N}{j} \left(\frac{p}{q}\right)^j \sum_{x=0}^N \binom{N}{x} p^x q^{N-x} K_i(x) K_j(x) e^{-\lambda(\alpha+\beta)xt}\right] z^j. \end{aligned}$$

Equating coefficients of z^j completes the proof of the claim.

(b) Combining (16) with results (a), (b), and (g) of Lemma 1, we obtain

$$\begin{aligned} E[X_t \mid X_0 = i] &= \sum_{j=0}^N j p_{ij}(t) \\ &= \sum_{j=0}^N j \omega(j) \sum_{x=0}^N \pi_x K_i(x) K_j(x) e^{-\lambda(\alpha+\beta)xt} \\ &= \sum_{x=0}^N \pi_x K_i(x) e^{-\lambda(\alpha+\beta)xt} \sum_{j=0}^N j K_x(j) \omega(j) \\ &= \sum_{x=0}^N \pi_x B_{0,x} K_i(x) e^{-\lambda(\alpha+\beta)xt}, \end{aligned} \tag{19}$$

where $B_{0,x} \neq 0$ for $x \in \{0, 1\}$ as given in (11), which implies (17). Moreover,

$$\begin{aligned}
 & E[X_t^2 \mid X_0 = i] \\
 &= \sum_{j=0}^N j^2 p_{ij}(t) \\
 &= \sum_{j=0}^N j^2 \omega(j) \sum_{x=0}^N \pi_x K_i(x) K_j(x) e^{-\lambda(\alpha+\beta)xt} \\
 &= \sum_{x=0}^N \pi_x K_i(x) e^{-\lambda(\alpha+\beta)xt} \sum_{j=0}^N j^2 K_x(j) \omega(j) \\
 &= \sum_{x=0}^N \pi_x K_i(x) e^{-\lambda(\alpha+\beta)xt} \sum_{j=0}^N j \omega(j) [-(N-x)p K_{x+1}(j) \\
 &\qquad\qquad\qquad + [(N-x)p + xq] K_x(j) - xq K_{x-1}(j)] \\
 &= \sum_{x=0}^N \pi_x K_i(x) e^{-\lambda(\alpha+\beta)xt} [-(N-x)p B_{0,x+1} + [(N-x)p + xq] B_{0,x} - xq B_{0,x-1}],
 \end{aligned}$$

where $B_{0,x}$ is given by (11). A straightforward computation of

$$\text{var}[X_t \mid X_0 = i] = E[X_t^2 \mid X_0 = i] - E[X_t \mid X_0 = i]^2$$

leads immediately to (18).

(c) The proof is clear.

4. Ehrenfest’s short-rate model

In this section we introduce a finite-state mean-reverting short-rate model associated with the continuous-time Ehrenfest process (13) and give its basic properties. As a main result, we exploit the algebraic-combinatorial roots of the Ehrenfest process and derive explicit pricing formulae for ZCBs in the general and special cases of the process, both of which have their advantages.

4.1. Definition and properties

Let $[r_m, r_M] \subseteq \mathbb{R}$ be an interval on the real line. We decompose it into N equal pieces of length $h := (r_M - r_m)/N$ and consider the process

$$(R_t^{(N)} := hX_t^{(N)} + r_m)_{t \geq 0} \tag{20}$$

as a *short-rate process* with state space $E := \{r_k := hk + r_m, k = 0, \dots, N\}$, where $(X_t^{(N)})_{t \geq 0}$ is the Ehrenfest process given by (13) with $\alpha, \beta \in (0, 1]$. Considering Remark 2(b), we note that $(R_t^{(N)})_{t \geq 0}$ can be seen as an affine linearly transformed birth-and-death process on $\{0, 1, \dots, N\}$. In the case at hand, N can be interpreted as the state space discretization parameter. Clearly, $(R_t := R_t^{(N)})_{t \geq 0}$ also depends on N . We will suppress this dependence when it is clear from the context. Bearing in mind Remark 3, we call this short-rate model the $\mathcal{E}(p, \lambda)$ model.

From Theorem 1 we immediately obtain the conditional mean and variance of $(R_t)_{t \geq 0}$, given $R_0 := r_k \in E$, as follows:

$$\begin{aligned} E[R_t \mid R_0 = hk + r_m] &= h E[X_t \mid X_0 = k] + r_m \\ &= (r_M - r_m) \left(p - \left(p - \frac{i}{N} \right) e^{-\lambda(\alpha + \beta)t} \right) + r_m, \\ \text{var}[R_t \mid R_0 = hk + r_m] &= h^2 \text{var}[X_t \mid X_0 = k] \\ &= \frac{(r_M - r_m)^2}{N^2} (Np(1 - p) + (Np - i)(2p - 1)e^{-\lambda(\alpha + \beta)t} \\ &\quad - (Np - i)^2(2p - 1)e^{-2\lambda(\alpha + \beta)t}), \end{aligned}$$

where $p = \alpha / (\alpha + \beta)$. We also obtain the *mean reversion* of $(R_t)_{t \geq 0}$:

$$\begin{aligned} \lim_{t \rightarrow \infty} E[R_t] &= pr_M + (1 - p)r_m, \tag{21} \\ \lim_{t \rightarrow \infty} \text{var}[R_t] &= \frac{(r_M - r_m)^2}{N} p(1 - p) < \infty. \end{aligned}$$

Thus, we have a total of five parameters, r_m, r_M, p, λ , and N , to fit the model to the market data. Here, p governs the skewness of the underlying distribution, $r_M - r_m, N$, and p have an impact on its kurtosis, and λ influences the speed of reversion to the mean reverting value $pr_M + (1 - p)r_m$.

4.2. Zero-coupon bond

We assume that an equivalent martingale measure (or risk-neutral measure) exists (see [12, Proposition 4.2]), and that the underlying probability measure P is this equivalent martingale measure. Let $(\mathcal{F}_t)_{t \geq 0}$ be the natural filtration of $(R_t)_{t \geq 0}$. Then, the arbitrage-free ZCB price at time t with a face value of one monetary unit and maturity at T is given by (see [4, p. 51])

$$P^{(N)}(t, T) = E \left[\exp \left(- \int_t^T R_s ds \right) \mid \mathcal{F}_t \right], \quad r \in E. \tag{22}$$

In the following we also omit explicitly writing out the dependence of $P(t, T)$ on N when it is clear from the context.

The calculation of (22) within the $\mathcal{E}(p, \lambda)$ model with arbitrary $\alpha, \beta \in (0, 1]$ is inspired by the proof of Theorem 3.1 of [6]. There, Delbaen and Shirakawa represented the transition probabilities of the underlying short-rate process as a weighted series of Jacobi polynomials. Using orthogonality relations of the Jacobi polynomials, they obtained a pricing formula for ZCBs in the associated model. However, this formula is only semi-explicit, since it contains multiple integrals that have to be calculated iteratively. We will avoid this problem by representing such integrals in terms of ${}_1F_1$ functions.

Theorem 2. (The ZCB price in the $\mathcal{E}(p, \lambda)$ model.) *Let $(R_t)_{t \geq 0}$ be given by definition (20) with $\alpha, \beta \in (0, 1]$, $p = \alpha / (\alpha + \beta)$, and $\lambda > 0$. The price at time $t \geq 0$ of a ZCB with maturity at T , conditional on $R_t = r \in E$, is given by*

$$P(t, T) = e^{-r_m(T-t)} P_1(t, T)^k P_0(t, T)^{N-k},$$

where $k := (r - r_m)/h \in \{0, \dots, N\}$ and, for $m \in \{0, 1\}$,

$$P_m(t, T) := 1 + \sum_{n=1}^{\infty} (-1)^n \frac{(h(T-t))^n}{n!} \times \sum_{i_1=0}^1 \cdots \sum_{i_n=0}^1 K_m(i_n) \left[\prod_{j=1}^n \left(\frac{p}{q}\right)^{i_j} B_{i_{j-1}, i_j} \right] {}_1F_1(1; n+1; z^{(n)}). \tag{23}$$

Here, K_m is given according to (b) and (c) of Lemma 1, ${}_1F_1$ is defined by (1), $z^{(n)} := -\lambda(\alpha + \beta)(T-t)(i_1, \dots, i_n)^T \in \mathbb{R}^n$, $i_0 := 0$, and

$$B_{i_{j-1}, i_j} := \begin{cases} p-1 & \text{if } |i_{j-1} - i_j| = 1, \\ i_j(1-2p) + p & \text{if } i_{j-1} = i_j. \end{cases}$$

Proof. Let $r_k = hk + r_m$, $k \in \{0, \dots, N\}$, be the state of $(R_t)_{t \geq 0}$ at time t . Bearing in mind the fact that $(R_t)_{t \geq 0}$ is a Markov process, and using definitions (13) and (20), from (22), we obtain

$$\begin{aligned} P(t, T) &= E \left[\exp \left(- \int_t^T \left(h \sum_{l=1}^N Y_s^l + r_m \right) ds \right) \middle| X_t = k \right] \\ &= e^{-r_m(T-t)} E \left[\prod_{l=1}^N \exp \left(- \int_t^T h Y_s^l ds \right) \middle| X_t = k \right] \\ &= e^{-r_m(T-t)} E_{1,t} \left[\exp \left(- \int_t^T h Y_s ds \right) \right]^k E_{0,t} \left[\exp \left(- \int_t^T h Y_s ds \right) \right]^{N-k}, \end{aligned} \tag{24}$$

where $E_{m,t}[\cdot] := E[\cdot | X_t = m]$ for $m \in \{0, 1\}$. The last equality holds because of the independence of $(Y_t^l)_{t \geq 0}$ for all $l = 1, \dots, N$. In the following we omit writing out the dependence on particular l , and set

$$P_m(t, T) := E_{m,t} \left[\exp \left(- \int_t^T h Y_s ds \right) \right], \quad m \in \{0, 1\}. \tag{25}$$

Using (25), we rewrite (24) as

$$P(t, T) = e^{-r_m(T-t)} P_1(t, T)^k P_0(t, T)^{N-k}. \tag{26}$$

From the power series representation of the exponential function, we obtain

$$\begin{aligned} P_m(t, T) &= E_{m,t} \left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(- \int_t^T h Y_s ds \right)^n \right] \\ &= 1 + \sum_{n=1}^{\infty} (-1)^n h^n \int_t^T \int_{s_1}^T \cdots \int_{s_{n-1}}^T E_{m,t} [Y_{s_n} \cdots Y_{s_1}] ds_n \cdots ds_2 ds_1, \end{aligned} \tag{27}$$

where $t =: s_0 < s_1 < \dots < s_n < T$. The last equality follows from

$$\left(\int_t^T Y_s ds \right)^n = n! \int_t^T \int_{s_1}^T \cdots \int_{s_{n-1}}^T Y_{s_n} \cdots Y_{s_1} ds_n \cdots ds_2 ds_1$$

and the dominated convergence theorem.

For the given $t < s_1 < \dots < s_n < T$ and $t_j := s_j - s_{j-1}$, we have

$$E_{m,t}[Y_{s_n} \dots Y_{s_1}] = \sum_{m_1=0}^1 \dots \sum_{m_n=0}^1 \prod_{j=1}^n m_j p_{m_{j-1}, m_j}(t_j),$$

where $m_0 := m$. Using the symmetry relation (7), we write the transition probabilities $p_{m_{j-1}, m_j}(t_j)$ given in Theorem 1 as

$$p_{m_{j-1}, m_j}(t_j) = w(m_j) \sum_{i_j=0}^1 \pi_{i_j} K_{m_{j-1}}(i_j) K_{m_j}(i_j) e^{-\lambda_{i_j} t_j},$$

where $\lambda_{i_j} := \lambda(\alpha + \beta) i_j$. Analogously to the calculation of the expected value (19) in the proof of Theorem 1, we obtain

$$E_{m_{n-1}, s_{n-1}}[Y_{s_n}] = \sum_{i_1=0}^1 \pi_{i_1} B_{i_0, i_1} K_{m_{n-1}}(i_1) e^{-\lambda_{i_1} t_n},$$

where B_{i_0, i_1} is defined by (11) and $i_0 := 0$. Iteratively, we obtain

$$E_{m,t}[Y_{s_n} \dots Y_{s_1}] = \sum_{i_1=0}^1 \dots \sum_{i_n=0}^1 K_m(i_n) \left[\prod_{j=1}^n \pi_{i_j} B_{i_{j-1}, i_j} e^{-\lambda_{i_j} t_{n-j+1}} \right].$$

Hence, (27) becomes

$$P_m(t, T) = 1 + \sum_{n=1}^{\infty} (-1)^n h^n \sum_{i_1=0}^1 \dots \sum_{i_n=0}^1 K_m(i_n) \left[\prod_{j=1}^n \pi_{i_j} B_{i_{j-1}, i_j} \right] \times \int_t^T \int_{s_1}^T \dots \int_{s_{n-1}}^T \exp\left(-\sum_{k=1}^n \lambda_{n-k+1}(s_k - s_{k-1})\right) ds_n \dots ds_2 ds_1. \tag{28}$$

In order to evaluate the multiple integrals above, we transform the integration domain to the standard simplex Δ_n defined by (2) via the following mapping:

$$J: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \mapsto \begin{pmatrix} (T-t)s_1 + t \\ (T-t)(s_1 + s_2) + t \\ \vdots \\ (T-t)(s_1 + \dots + s_n) + t \end{pmatrix}. \tag{29}$$

Using (28), we rewrite (29) as

$$P_m(t, T) = 1 + \sum_{n=1}^{\infty} (-1)^n h^n \sum_{i_1=0}^1 \dots \sum_{i_n=0}^1 K_m(i_n) \left[\prod_{j=1}^n \pi_{i_j} B_{i_{j-1}, i_j} \right] \times (T-t)^n \int_{\Delta_n} e^{\langle z^{(n)}, x \rangle} dx, \tag{30}$$

where $z^{(n)} := -(T-t)(\lambda_{i_1}, \dots, \lambda_{i_n})^T \in \mathbb{R}^n$ and $\langle \cdot, \cdot \rangle$ denotes the standard inner product.

Applying (3) with $a = 1$, we express the integrals in (30) as ${}_1F_1$ functions as follows:

$$n! \int_{\Delta_n} e^{(z^{(n)}, x)} dx = {}_1F_1(1; n + 1; z^{(n)}) \tag{31}$$

for all $z^{(n)} \in \mathbb{R}^n (n \in \mathbb{N}_0)$, setting ${}_1F_1(1; 1; z^{(0)}) := 1$. If we combine (26) and (30) with (31) and the fact that $\pi_{i_j} = (p/q)^{i_j}, i_j \in \{0, 1\}$, the theorem follows.

Now we consider the $\mathcal{E}(\frac{1}{2}, \lambda)$ model with $\alpha = \beta = 1$. On the one hand, we lose one of the fitting parameters, although the model is still well suited to model the term structure, and it yields the famous *Vasicek model* in the limit (see Section 5). On the other hand, we obtain a more tractable pricing formula for ZCBs, where, in contrast to the general case, no multiple sums need calculation, which improves the computational speed.

The calculation of the arbitrage-free ZCB price (22) in this setting is very intuitive and requires no knowledge of the transition probabilities of $(R_t)_{t \geq 0}$, since the only stochastic parameters are the arrival times of the underlying Poisson process.

Theorem 3. (The ZCB price in the $\mathcal{E}(\frac{1}{2}, \lambda)$ model.) *Let $(R_t)_{t \geq 0}$ be given by definition (20) with $\lambda > 0$ and $\alpha = \beta = 1$. The price at time $t \geq 0$ of a ZCB with maturity at T is given by*

$$P(t, T) = e^{-(r_m + \lambda N)(T-t)} P_1(t, T)^k P_0(t, T)^{N-k}, \tag{32}$$

where $k := (R_t - r_m)/h \in \{0, \dots, N\}$,

$$P_1(t, T) := \sum_{n=0}^{\infty} \frac{(\lambda(T-t))^{2n}}{(2n)!} \left\{ e^{-h(T-t)} {}_1F_1(1; 2n + 1; z^{(2n)}) + \frac{\lambda(T-t)}{2n+1} {}_1F_1(1; 2n + 2; -z^{(2n+1)}) \right\}, \tag{33}$$

$$P_0(t, T) := \sum_{n=0}^{\infty} \frac{(\lambda(T-t))^{2n}}{(2n)!} \left\{ {}_1F_1(1; 2n + 1; -z^{(2n)}) + \frac{\lambda(T-t)}{2n+1} e^{-h(T-t)} {}_1F_1(1; 2n + 2; z^{(2n+1)}) \right\}, \tag{34}$$

$z^{(2n)} := h(T-t)(0, 1, \dots, 0, 1)^T \in \mathbb{R}^{2n}, z^{(2n+1)} := h(T-t)(1, 0, 1, \dots, 0, 1)^T \in \mathbb{R}^{2n+1}$, and ${}_1F_1$ is defined by (1).

Proof. Let $r_k = hk + r_m, k \in \{0, \dots, N\}$, be the state of $(R_t)_{t \geq 0}$ at time t . Analogously to the derivation of expression (26) in the proof of Theorem 2, we obtain

$$P(t, T) = e^{-r_m(T-t)} \tilde{P}_1(t, T)^k \tilde{P}_0(t, T)^{N-k}, \tag{35}$$

where

$$\tilde{P}_y(t, T) := E_{y,t} \left[\exp \left(- \int_t^T h Y_s ds \right) \right], \quad y \in \{0, 1\}. \tag{36}$$

In order to evaluate $\tilde{P}_y(t, T)$, we count the number of jumps in the underlying Poisson process $(N_t)_{t \geq 0}$ within the time interval $(t, T]$, and, denoting the jump times by $(\tau_i)_{i \in \mathbb{N}}$ and setting $\tau_0 := t$, we split the integral on the right-hand side of (36), obtaining

$$\tilde{P}_y(t, T) = E_{y,t} \left[\sum_{n=0}^{\infty} \mathbf{1}_{\{N_{T-t}=n\}} \exp \left(- \sum_{i=0}^{n-1} \int_{\tau_i}^{\tau_{i+1}} h \hat{Y}_i ds - \int_{\tau_n}^T h \hat{Y}_n ds \right) \right].$$

At this point, we have to distinguish between even and odd numbers of jumps, since $(\hat{Y}_n)_{n \in \mathbb{N}}$ switches between 0 and 1 P-almost surely according to its transition probability matrix (12). Thus, conditional on $\{\hat{Y}_0 = 1\}$, the Markov chain $(\hat{Y}_n)_{n \in \mathbb{N}}$ stays at 1 after an even jump, whereas it stays at 0 after an odd jump. This consideration yields

$$\begin{aligned} \tilde{P}_1(t, T) &= \sum_{n=0}^{\infty} \left\{ \mathbb{E} \left[\mathbf{1}_{\{N_{T-t}=2n\}} \exp \left(- \sum_{i=0}^{n-1} \int_{\tau_{2i}}^{\tau_{2i+1}} h \, ds - \int_{\tau_{2n}}^T h \, ds \right) \right] \right. \\ &\quad \left. + \mathbb{E} \left[\mathbf{1}_{\{N_{T-t}=2n+1\}} \exp \left(- \sum_{i=0}^n \int_{\tau_{2i}}^{\tau_{2i+1}} h \, ds \right) \right] \right\} \\ &= \sum_{n=0}^{\infty} \left\{ \mathbb{E} \left[\exp \left(h \sum_{i=1}^{2n} (-1)^i \tau_i + t - T \right) \mid N_{T-t} = 2n \right] \mathbb{P}(N_{T-t} = 2n) \right. \\ &\quad \left. + \mathbb{E} \left[\exp \left(h \sum_{i=1}^{2n+1} (-1)^i \tau_i + t \right) \mid N_{T-t} = 2n + 1 \right] \mathbb{P}(N_{T-t} = 2n + 1) \right\}. \end{aligned}$$

Furthermore, from the order statistics property of the Poisson process (see, for example, [9, pp. 101–102]), we know that the joint density of the arrival times τ_1, \dots, τ_k ($k \in \mathbb{N}$) of $(N_t)_{t \geq 0}$ in $(t, T]$, conditional on $\{N_{T-t} = k\}$, is given by

$$\mathbb{P}(t < \tau_1 \leq \tau_2 \leq \dots \leq \tau_k \leq T \mid N_{T-t} = k) = \frac{n!}{(T-t)^k} \int_t^T \int_{t_1}^T \dots \int_{t_{k-1}}^T dt_k \dots dt_2 dt_1.$$

Hence, for $\tilde{P}_1(t, T)$, we obtain

$$\begin{aligned} \tilde{P}_1(t, T) &= \sum_{n=0}^{\infty} \left\{ \left[e^{-h(T-t)} \frac{(2n)!}{(T-t)^{2n}} \int_t^T \dots \int_{t_{2n-1}}^T \exp \left(h \sum_{i=1}^{2n} (-1)^i t_i \right) dt_{2n} \dots dt_1 \right] \right. \\ &\quad \times e^{-\lambda(T-t)} \frac{(\lambda(T-t))^{2n}}{(2n)!} \\ &\quad + \left[e^{-ht} \frac{(2n+1)!}{(T-t)^{2n+1}} \int_t^T \dots \int_{t_{2n}}^T \exp \left(h \sum_{i=1}^{2n+1} (-1)^i t_i \right) dt_{2n+1} \dots dt_1 \right] \\ &\quad \left. \times e^{-\lambda(T-t)} \frac{(\lambda(T-t))^{2n+1}}{(2n+1)!} \right\}. \tag{37} \end{aligned}$$

Analogously to the proof of Theorem 2, we consider the mapping J given in (29) and integrate (37) by substitution, which yields

$$\begin{aligned} \tilde{P}_1^{(N)}(t, T) &= e^{-\lambda(T-t)} \sum_{n=0}^{\infty} \frac{(\lambda(T-t))^{2n}}{(2n)!} \left\{ e^{-h(T-t)} (2n)! \int_{\Delta_{2n}} e^{(z^{(2n)}, x)} \, dx \right. \\ &\quad \left. + (2n+1)! \int_{\Delta_{2n+1}} e^{(-z^{(2n+1)}, x)} \, dx \right\}, \tag{38} \end{aligned}$$

where

$$z^{(2n)} := h(T-t)(0, 1, \dots, 0, 1)^\top \in \mathbb{R}^{2n}, \quad z^{(2n+1)} := h(T-t)(1, 0, 1, \dots, 0, 1)^\top \in \mathbb{R}^{2n+1},$$

and $\langle \cdot, \cdot \rangle$ denotes the standard inner product. Applying relation (31), we rewrite (38) as

$$\begin{aligned} \tilde{P}_1(t, T) &= e^{-\lambda(T-t)} \sum_{n=0}^{\infty} \frac{(\lambda(T-t))^{2n}}{(2n)!} \left\{ e^{-h(T-t)} {}_1F_1(1; 2n+1; z^{(2n)}) \right. \\ &\quad \left. + \frac{\lambda(T-t)}{2n+1} {}_1F_1(1; 2n+2; -z^{(2n+1)}) \right\} \\ &=: e^{-\lambda(T-t)} P_1(t, T). \end{aligned} \tag{39}$$

In similar fashion, we obtain

$$\begin{aligned} \tilde{P}_0(t, T) &= e^{-\lambda(T-t)} \sum_{n=0}^{\infty} \frac{(\lambda(T-t))^{2n}}{(2n)!} \left\{ {}_1F_1(1; 2n+1; -z^{(2n)}) \right. \\ &\quad \left. + \frac{\lambda(T-t)}{2n+1} e^{-h(T-t)} {}_1F_1(1; 2n+2; z^{(2n+1)}) \right\} \\ &=: e^{-\lambda(T-t)} P_0(t, T). \end{aligned} \tag{40}$$

If we combine (39) and (40) with (35), the theorem follows.

Remark 4. Note that Theorems 2 and 3 are based on two completely different approaches and yield different representations of the ZCB prices. However, both formulae involve the confluent hypergeometric function ${}_1F_1$ defined by (1).

4.3. Practical implementation

From Theorems 2 and 3, the ZCB prices can be computed approximately by truncating the series in the according formulae. We also use the truncated ${}_1F_1^H$ function defined by (4) as an approximation for the ${}_1F_1$ function.

In the setting of Theorem 2, we truncate the sum of series (23), obtaining

$$P(t, T; M, H) := e^{-r_m(T-t)} P_1(t, T; M, H)^k P_0(t, T; M, H)^{N-k}, \tag{41}$$

$$\begin{aligned} P_y(t, T; M, H) &:= 1 + \sum_{n=1}^M (-1)^n \frac{(h(T-t))^n}{n!} \\ &\quad \times \sum_{i_1=0}^1 \cdots \sum_{i_n=0}^1 K_y(i_n) \left[\prod_{j=1}^n \left(\frac{p}{q} \right)^{i_j} B_{i_{j-1}, i_j} \right] {}_1F_1^H(1; n+1; z^{(n)}) \end{aligned}$$

for $y \in \{0, 1\}$.

In the setting of Theorem 3, we truncate series (33) and (34), obtaining

$$P(t, T; M, H) := e^{-(r_m+a)(T-t)} P_1(t, T, M, H)^k P_0(t, T, M, H)^{N-k}, \tag{42}$$

$$\begin{aligned} P_1(t, T; M, H) &:= \sum_{n=0}^M \frac{(\lambda(T-t))^{2n}}{(2n)!} \left\{ e^{-h(T-t)} {}_1F_1^H(1; 2n+1; h(T-t)z^{(2n)}) \right. \\ &\quad \left. + \frac{\lambda(T-t)}{2n+1} {}_1F_1^H(1; 2n+2; -h(T-t)z^{(2n+1)}) \right\}, \end{aligned}$$

$$P_0(t, T; M, H) := \sum_{n=0}^M \frac{(\lambda(T-t))^{2n}}{(2n)!} \times \left\{ {}_1F_1^H(1; 2n+1; -h(T-t)z^{(2n)}) + \frac{\lambda(T-t)}{2n+1} e^{-h(T-t)} {}_1F_1^H(1; 2n+2; h(T-t)z^{(2n+1)}) \right\}.$$

The choice of the truncation parameters M and H is left to the practitioner and should be made with a view to maintaining a balance between the accuracy of the results and the computational speed. Some numerical examples, which provide numerical accuracy and computational speed for (41) and (42), will be given at the end of the next sections. In the following we omit explicitly writing out the dependence on M and H when it is clear from the context.

5. Connection to the Vasicek model

The Vasicek model [19] is one of the most popular short-rate models. Closed-form expressions of ZCB prices and European options thereof make the model highly appealing to practitioners. However, it also has some shortcomings.

In this section we give a short description of the Vasicek model. We point out its advantages and disadvantages. Here we follow Section 3.2.1 of [4]. We provide a convergence result, which shows that after a linear rescaling the $\mathcal{E}(\frac{1}{2}, \lambda)$ model converges weakly to the Vasicek model. We show the convergence of the respective ZCB prices and provide some numerical examples.

5.1. Vasicek model

The formulation of the Vasicek model under the risk-neutral measure P is

$$dr(t) = k[\theta - r(t)]dt + \sigma dW_t, \quad r(0) = r_0, \tag{43}$$

where k, θ, σ , and r_0 are positive constants, and $(W_t)_{t \geq 0}$ is the standard Wiener process on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ with the natural filtration $(\mathcal{F}_t)_{t \geq 0}$. Integration of (43) yields, for $t \geq s$,

$$r(t) = r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}) + \sigma \int_s^t e^{-k(t-u)} dW_u. \tag{44}$$

Thus, conditional on $\mathcal{F}_s(t \geq s)$, $r(t)$ is normally distributed with mean and variance

$$E[r(t) | \mathcal{F}_s] = r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}), \quad \text{var}[r(t) | \mathcal{F}_s] = \frac{\sigma^2}{2k}(1 - e^{-2k(t-s)}).$$

Hence, the short-rate process $(r_t)_{t \geq 0}$ tends to the mean-reverting value θ for $t \rightarrow \infty$. The drawbacks of the model are the possible negativity of the interest rates, implied by the Gaussian distribution, and the fact that it is driven by only three parameters, which makes the calibration an ill-posed problem and often yields poor results.

The price at time $t \geq 0$ of a ZCB with maturity at T , conditional on $r(t) = r$, is given by

$$P(t, T) = A(t, T)e^{-B(t, T)r}, \tag{45}$$

where

$$A(t, T) = \exp\left(\left(\theta - \frac{\sigma^2}{2k^2}\right)[B(t, T) - T + t] - \frac{\sigma^2}{4k}B(t, T)^2\right),$$

$$B(t, T) = \frac{1}{k}(1 - e^{-k(T-t)}).$$

5.2. Convergence results

It is well known that the Ehrenfest process converges weakly to the Ornstein–Uhlenbeck process (see, for instance, [9, pp. 168–173] or [16]). The following theorem shows that the $\mathcal{E}(\frac{1}{2}, \lambda)$ model also converges weakly to the Vasicek model.

Theorem 4. *Let $(r(t))_{t \geq 0}$ be given as in (43). Consider $(R_t^{(N)})_{t \geq 0}$ as defined in (20) with $\alpha = \beta \in (0, 1]$, $\lambda := \alpha/(\alpha + \beta)$, $r_m := \theta - \sigma\sqrt{N/2k}$, and $r_M := \theta + \sigma\sqrt{N/2k}$. Then,*

$$(R_t^{(N)})_{t \in [0, T]} \Rightarrow (r(t))_{t \in [0, T]} \quad \text{as } N \rightarrow \infty,$$

where ‘ \Rightarrow ’ denotes weak convergence.

Proof. The proof follows analogously to [9, pp. 168–173], when we compute the conditional moments of $\Delta R_t := R_{t+\Delta t} - R_t$.

Remark 5. Karlin and McGregor showed in a rigorous way (see [8, pp. 371–373]) that the transition probability function (16) of the Ehrenfest process converges locally uniformly to the transition probability function of the Ornstein–Uhlenbeck process as $N \rightarrow \infty$, which sharpens the above result after a linear transformation of the underlying processes.

A direct consequence of Theorem 4 is the convergence of the respective ZCB prices.

Corollary 1. *Consider $(r(t))_{t \in [0, T]}$ and $(R_t^{(N)})_{t \in [0, T]}$ as in Theorem 4, and let $P(t, T)$ and $P^{(N)}(t, T)$ denote the associated ZCB prices at time t with maturity at T given by (45) and (32), respectively. Then, $P^{(N)}(t, T) \rightarrow P(t, T)$ as $N \rightarrow \infty$.*

Proof. Without loss of generality, let $t = 0$. We denote by $\mathbb{D} := \mathbb{D}[0, T]$ the space of real-valued functions on $[0, T]$ that are *right continuous and have left-hand limits*. From [1] (see page 123), we know that a metric exists that makes \mathbb{D} a *Polish space*, i.e. a metric, separable, and complete space. Clearly, $R_N := (R_t^{(N)})_{t \in [0, T]}$ and $r := (r(t))_{t \in [0, T]}$ both lie in \mathbb{D} . With Theorem 4, it follows that $R_N \Rightarrow r$ in \mathbb{D} .

Consider a linear operator \tilde{S} on \mathbb{D} , defined by

$$(\tilde{S}f)(t) := \int_0^t f(s) \, ds \quad \text{for } f \in \mathbb{D} \text{ and } t \in [0, T].$$

Clearly, \tilde{S} is a continuous operator on \mathbb{D} . Then, the operator S , defined by

$$(Sf)(t) := \exp(-(\tilde{S}R_N)(t)) = \exp\left(-\int_0^t f(s) \, ds\right) \quad \text{for } f \in \mathbb{D} \text{ and } t \in [0, T],$$

is a continuous operator on \mathbb{D} . Let $Y_N := (SR^{(N)}(t))_{t \in [0, T]}$ and $Y := (Sr(t))_{t \in [0, T]}$. Then, Theorem 5.1 of [1] yields $Y_N \Rightarrow Y$. Since Y_N is *uniformly integrable*, it follows from Theorem 5.4 of [1] that $E[Y_N] \rightarrow E[Y]$ as $N \rightarrow \infty$, which completes the proof.

Figure 1 illustrates the convergence result of Corollary 1. All computations were made on an INTEL® Core™ 2 Duo 2400MHz machine. We consider two scenarios: (a) we have a favourable set of parameters for the ZCB valuation; and (b) we choose an unrealistic high of 20% for the interest rate market volatility and a time to maturity of 10 years. Within the Vasicek model the valuation is done according to the pricing formula (45). The approximate values $P(0, T; M, H)$ of the ZCB prices in the $\mathcal{E}(\frac{1}{2}, \lambda)$ model are computed according to Corollary 1

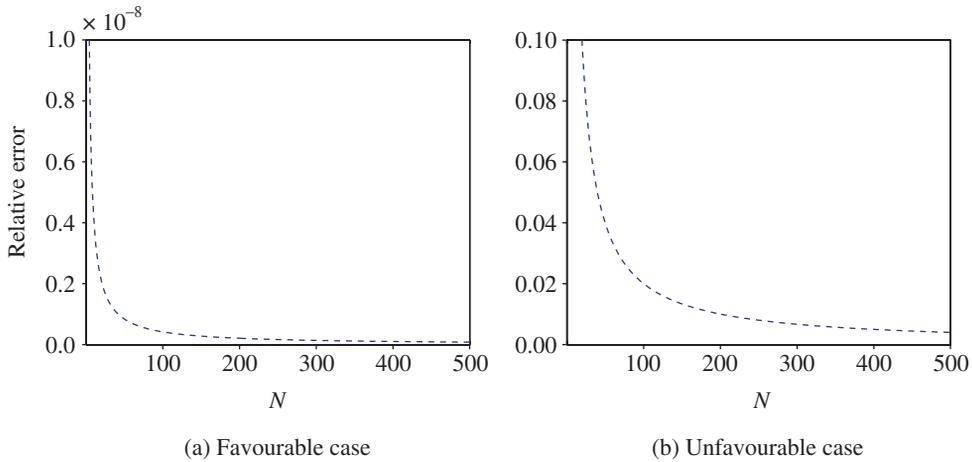


FIGURE 1: Relative price errors against N , approximating ZCB prices $P(0, T)$ in the Vasicek model by $P^{(N)}(0, T; 10, 30)$ in the $\mathcal{E}(\frac{1}{2}, \lambda)$ model via (42) in two scenarios: (a) Vasicek model with $k = 0.2$, $\sigma = 0.05$, $T = 1$ year, $\theta = 8\%$, and $r_0 = 5\%$; (b) Vasicek model with $k = 0.2$, $\sigma = 0.2$, $T = 10$ years, $\theta = 8\%$, and $r_0 = 5\%$.

via (42) with $M = 10$ and $H = 30$. In both cases we observe fast convergence of the respective prices.

Furthermore, we see that the choice of the truncating parameters $M = 10$ and $H = 30$ is satisfactory for our purpose. The computation time of one ZCB price via (42) is less than 0.1 seconds. Similar computation in the $\mathcal{E}(p, \lambda)$ model according to (41) takes 2.67 seconds.

6. Discussion

In this section we discuss the advantages of the $\mathcal{E}(p, \lambda)$ model with respect to the positivity of the interest rates. We use the case study of a ZCB valuation, showing that the $\mathcal{E}(p, \lambda)$ model can still be used when the Vasicek model reaches its limits.

The main shortcoming of all models with Gaussian distribution, including the Vasicek model, is the positive probability of the interest rates becoming negative. Although this probability is rather small, some problems may appear while valuing ZCBs with long residual maturity. For instance, Rogers [14] illustrated how an attempt to keep the probability of negative interest rates negligible by choosing suitable parameters of the Vasicek model in the limiting case $t \rightarrow \infty$ leads to an exponential growth in t of the ZCB prices. Conversely, the $\mathcal{E}(p, \lambda)$ model allows the choice of the lower and upper bounds, r_m and r_M , for the interest rate, and excludes the possibility of negative as well as unrealistically high positive interest rates.

Times of financial crisis are often accompanied by interest rates near 0%, as we see at present. The following example of pricing ZCBs in a respective scenario illustrates the advantage of the $\mathcal{E}(p, \lambda)$ model over the Vasicek model. First, we assume the Vasicek model given according to (43) with $\theta = 4\%$, $\sigma = 0.05$, $k = 0.1$, and $r(0) = 1\%$. Figure 2(a) shows three sample paths of the underlying process $(r_t)_{t \geq 0}$ over a period of 30 years simulated according to (44). We see that every path of the simulated process spends some time below the zero mark. Figure 2(b) demonstrates the weakness of the model in the case at hand, as we observe that the ZCB prices

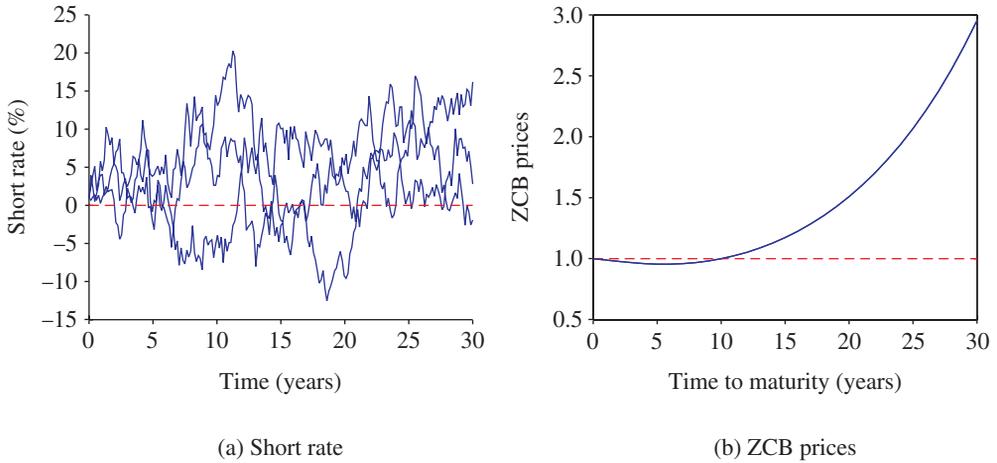


FIGURE 2: (a) Three sample paths of the short-rate process (43) in the Vasicek model with $k = 0.1$, $\theta = 4\%$, $\sigma = 0.05$, and $r(0) = 1\%$. (b) ZCB prices in the Vasicek model with the given parameters and residual maturities from 1 to 30 years.

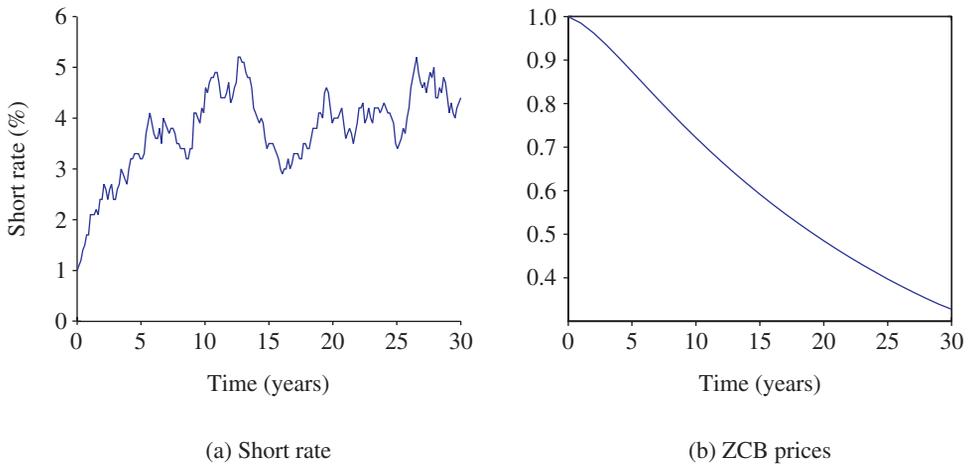


FIGURE 3: (a) A sample path of the short-rate process (20) in the $\mathcal{E}(p, \lambda)$ model with $\alpha = 0.1$, $\beta = 0.3$, $\lambda = 1$, $r_m = 0\%$, $r_M = 16\%$, $N = 160$, and $R_0 = 1\%$. (b) ZCB prices in the $\mathcal{E}(p, \lambda)$ model with the given parameters and residual maturities from 1 to 30 years.

are not monotone falling in the time to maturity and even exceed the upper bound of one monetary unit, which is contradictory to no-arbitrage principles.

Now we consider the $\mathcal{E}(p, \lambda)$ model in a similar hypothetical setting. We set the lower and upper bounds at $r_m = 0\%$ and $r_M = 16\%$, and the state space discretization parameter $N = 160$. We choose $\lambda = 1$, $\alpha = 0.1$, and $\beta = 0.3$, in that we have, with (21), a mean-reverting value of 4% as in the case above. Here, we set $R_0 = 1\%$ as well. Figure 3(a) demonstrates a

possible trajectory for the short-rate process $(R_t)_{t \geq 0}$ over 30 years, simulated on the basis of the underlying distribution. In Figure 3(b) we see the strictly monotone decreasing character of the respective ZCB prices as a function of the time to maturity, which is highly plausible.

7. Conclusions

In this paper we explored a finite-state mean-reverting short-rate model based on the Ehrenfest process. The respective short-rate process can be seen as an affine linearly transformed birth-and-death process on $\{0, 1, \dots, N\}$, $N \in \mathbb{N}$. The model provides a certain degree of analytical tractability, since it allows explicit pricing of ZCBs and solves the problem of negative interest rates characteristic of Gaussian models. The pricing formulae for ZCBs have been derived for both the general case and the special case, in which the underlying distribution is symmetric with respect to the mean-reverting value. The key to both approaches has turned out to be the representation of the underlying Ehrenfest process as a sum of independent binary processes, which has been possible only in continuous time. We also used the hypergeometric functions of a matrix argument and the Krawtchouk polynomials. The special case also benefits from a more tractable pricing formula for ZCBs.

We have seen that the Ehrenfest short-rate model is a good approximation to the Vasicek model under normal conditions and a better alternative to it in extreme cases, where the interest rates are low and the volatility is high, providing solely positive interest rates. A further advantage of the model is the availability of five fitting parameters in the general case.

Our conclusion is that especially the general case of the Ehrenfest short-rate model is an interesting enrichment in the field of term structure modelling, combining analytical tractability with the desired property of interest rates remaining positive.

Problems that remain open for the short-rate model that we have examined here are the derivation of an explicit pricing formula for European options on ZCBs, parameter estimates for the model under the objective measure. An extension of the model according to the three urn Ehrenfest model (see [8, pp. 363–368]) and a generalisation of the short-rate model, where the interest rate is allowed to jump more than one unit, are interesting open problems.

Acknowledgement

This paper comprises part of the author's PhD project at the Technische Universität Dortmund. I am deeply indebted to Professor Michael Voit for his patience and for being a wonderful research mentor.

References

- [1] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.
- [2] BINGHAM, N. H. (1991). Fluctuation theory for the Ehrenfest urn. *Adv. Appl. Prob.* **23**, 598–611.
- [3] BRÉMAUD, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer, New York.
- [4] BRIGO, D. AND MERCURIO, F. (2005). *Interest Rate Models - Theory and Practice*. Springer, Berlin.
- [5] BÜHLMANN, H. (1994). Continuous and discrete models in finance, in particular for stochastic interest rates. *Decisions Econom. Finance* **17**, 3–20.
- [6] DELBAEN, F. AND SHIRAKAWA, H. (2002). An interest rate model with upper and lower bounds. *Asia-Pacific Financial Markets* **9**, 191–209.
- [7] GROSS, K. I. AND RICHARDS, D. ST. P. (1989). Total positivity, spherical series, and hypergeometric functions of matrix argument. *J. Approx. Theory* **59**, 224–246.
- [8] KARLIN, S. AND MCGREGOR, J. (1965). Ehrenfest urn models. *J. Appl. Prob.* **2**, 352–376.
- [9] KARLIN, S. AND TAYLOR, H. M. (1981). *A Second Course in Stochastic Processes*. Academic Press, New York.
- [10] KOEV, P. AND EDELMAN, A. (2006). The efficient evaluation of the hypergeometric function of a matrix argument. *Math. Comput.* **75**, 833–846.

- [11] KRAFFT, O. AND SCHAEFER, M. (1993). Mean passage times for tridiagonal transition matrices and a two-parameter Ehrenfest urn model. *J. Appl. Prob.* **30**, 964–970.
- [12] LAKNER, P. (2006). Martingale measures for a class of right-continuous processes. *Math. Finance* **3**, 43–53.
- [13] OKUNEV, J. AND TIPPETT, M. (1989). Continuous-time stochastic calculus: a survey of applications in finance and accountancy. *IMA J. Manag. Math.* **2**, 157–171.
- [14] ROGERS, L. C. G. (1995). Which model for term-structure of interest rates should one use? In *Mathematical Finance* (IMA Vol. Math. Appl. **65**), eds M. H. A. Davis *et al.*, Springer, New York, pp. 93–116.
- [15] SIEGERT, A. J. F. (1949). On the approach to statistical equilibrium. *Phys. Rev.* **76**, 1708–1714.
- [16] SUMITA, U., GOTOH, J. AND JIN, H. (2006). Numerical exploration of dynamic behavior of Ornstein–Uhlenbeck processes via Ehrenfest process approximation. *J. Operat. Res. Soc. Japan* **49**, 256–278.
- [17] SZEGÖ, G. (1983). *Orthogonal Polynomials* (Amer. Math. Soc. Coll. Publ. **23**). American Mathematical Society, Providence, RI.
- [18] TAKAHASHI, H. (2004). Ehrenfest model with large jumps in finance. *Physica D* **189**, 61–69.
- [19] VASICEK, O. (1977). An equilibrium characterisation of the term structure. *J. Financial Econom.* **5**, 177–188.
- [20] VOIT, M. (1996). Asymptotic distributions for the Ehrenfest urn and related random walks. *J. Appl. Prob.* **33**, 340–356.