Article

On the Deterministic-Shift Extended CIR Model in a Negative Interest Rate Framework

Marco Di Francesco [1,†] and Kevin Kamm [2,*,†]

1   UnipolSai Assicurazioni, Via Stalingrado 45, 40128 Bologna, Italy; marco.difrancesco@unipolsai.it
2   Dipartimento di Matematica, Università di Bologna, 40126 Bologna, Italy
*   Correspondence: kevin.kamm@unibo.it
†   These authors contributed equally to this work.

Abstract: In this paper, we propose a new exogenous model to address the problem of negative interest rates that preserves the analytical tractability of the original Cox–Ingersoll–Ross (CIR) model with a perfect fit to the observed term-structure. We use the difference between two independent CIR processes and apply the deterministic-shift extension technique. To allow for a fast calibration to the market swaption surface, we apply the Gram–Charlier expansion to calculate the swaption prices in our model. We run several numerical tests to demonstrate the strengths of this model by using Monte-Carlo techniques. In particular, the model produces close Bermudan swaption prices compared to Bloomberg's Hull–White one-factor model. Moreover, it finds constant maturity swap (CMS) rates very close to Bloomberg's CMS rates.

Keywords: CIR model; negative interest rates; calibration; Riccati equations; swaptions; Bermudan swaptions

1. Introduction

This paper is the natural extension of Di Francesco and Kamm (2021), where we modeled interest rates by means of a short-rate model defined as the difference of two independent Cox–Ingersoll–Ross (CIR) processes in a negative interest rate framework.

We are extending the previous short-rate model by adding a deterministic function to allow for a perfect fit to the observed market term-structure while preserving the analytical tractability of an affine model and its features.

Let us briefly recall our findings in Di Francesco and Kamm (2021). We derived an analytical formula for the zero-coupon price of the non-extended model (see Theorem A1) by solving the associated Riccati equations in Lemma A1 explicitly and calibrated it to the market term-structure. Such short-rate models, where the observed term-structure is an output depending on the model parameters, are called *endogenous*. We performed several numerical experiments on two different dates obtaining good results in the sense that the model reproduced the market term-structures with negative interest rates very well, and it generates more realistic distributions of interest rates with a slight skewness and fatter tail with respect to the normal distribution. However, as reported in the numerical tests, the model failed to capture the full swaption surface due to the fact that the model parameters were constant and the Brownian motions were independent.

To improve the fit to the swaption surface, we suggest transforming the endogenous model into an exogenous one, in which the observed term-structure is an input.

A basic strategy to transform an endogenous model into an exogenous one is the inclusion of time-dependent parameters to exactly reproduce the observed term-structure. In fact, matching the term-structure exactly is equivalent to solving a system with an infinite number of equations. However, this is only possible by introducing an infinite number of parameters or, equivalently, a deterministic function of time. We follow the method illustrated in (Brigo and Mercurio 2006, pp. 95 ff. Section 3.8 A General Deterministic-Shift

Extension) to extend any time-homogeneous short-rate model, so as to exactly reproduce any observed term-structure of interest rates while preserving the possible analytical tractability of the original model.

To be more precise, we consider the CIR dynamics for $z \in \{x, y\}$

$$dz(t) = k_z(\theta_z - z(t))dt + \sigma_z \sqrt{z(t)}dW_z(t), \quad z(0) = z_0 \tag{1}$$

under a martingale measure $\mathbb{Q}$ with $k_z, \theta_z, \sigma_z \in \mathbb{R}_{>0}$ and define the short-rate as

$$r(t) := x(t) - y(t) + \psi(t), \tag{2}$$

where W_y and W_x are two independent standard Brownian motions on a stochastic basis $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{Q} \right)$ and $\psi(t) := f^M(0, t) - f(0, t)$ is a deterministic function defined as the difference of the market and model instantaneous forward rate.

Since the market term-structure is now an input, we can calibrate the model parameters to the swaption surface. However, simple Monte-Carlo techniques are in general very slow and memory demanding. Therefore, we resort to an approximation formula known as the Gram–Charlier expansion (cf. Tanaka et al. 2010) in our model. This allows for a fast and accurate calibration procedure.

1.1. Description of the Main Results

In this paper, we will first of all extend the results of Di Francesco and Kamm (2021) by using a deterministic shift extension. The zero-coupon price in the extended model (2) is given in the next Lemma.

Lemma 1. *Let* $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{Q} \right)$ *be a stochastic basis, where* $\mathbb{Q}$ *is a martingale measure,* $T > 0$ *a finite time horizon and let the* σ*-algebra* $(\mathcal{F}_t)_{t \in [0,T]}$ *fulfill the usual conditions and support two independent standard Brownian motions* W_x *and* W_y*. The price of a zero-coupon bond in the model* $r(t) := x(t) - y(t) + \psi(t)$ *is given by*

$$P(t, T) = \frac{P^M(0, T)}{P^M(0, t)} \frac{P^{CIR^-}(0, t)}{P^{CIR^-}(0, T)} P^{CIR^-}(t, T),$$

where $P^{CIR^-}(t, T)$ *is the zero-coupon price from Theorem A1 and* $P^M(0, T)$ *the market zero-curve.*

The derivation of this result is straightforward and is referred to in Section 2 alongside a recollection of basic results on swaps and swaptions.

We will see that it is necessary to study the so-called *swap moments* to derive the Gram–Charlier expansion. In our model, we will find explicit formulas allowing for fast swaption pricing and it is part of the next technical Lemma.

Lemma 2. *Let everything be as in Lemma 1. The so-called swap moments at time* $t < T_0$ *of order* $m \in \mathbb{N}$ *are given by*

$$M^m(t) := \mathbb{E}^{\mathbb{Q}^{T_0}} \left[\left(\mathrm{Swap}_{T_0}^{T_N}(T_0; K, \zeta) \right)^m \Big| \mathcal{F}_t \right] = \left(\frac{P^{CIR^-}(0, T_0)}{P^M(0, T_0)} \right)^m \frac{1}{P^{CIR^-}(t, T_0)}$$

$$\sum_{\substack{0 \le k_0, \dots, k_N \le N \\ k_0 + \dots + k_N = m}} \frac{m!}{k_0! \cdots k_N!} \widetilde{a}_0^{k_0} \cdots \widetilde{a}_N^{k_N} \left(M_x(t, T_0) e^{-N_x(t, T_0)x(t)} M_y(t, T_0) e^{N_y(t, T_0)y(t)} \right)$$

where we suppress the dependency of N_z, M_z *on* k_i *for readability. The coefficients* $\widetilde{a}_i$ *are given by*

$$\tilde{a}_0 := \zeta \frac{P^M(0, T_0)}{P^{CIR-}(0, T_0)}, \quad \tilde{a}_N := -\zeta(1 + K\alpha_N)\frac{P^M(0, T_N)}{P^{CIR-}(0, T_N)}, \quad \tilde{a}_i := -\zeta K\alpha_i \frac{P^M(0, T_i)}{P^{CIR-}(0, T_i)},$$

for $i = 1, \ldots, N-1$, year fractions α_i, fixed swap rate K and swap type $\zeta = 1$ for a payer swap and $\zeta = -1$ for a receiver swap.

Moreover, the functions $M_z, N_z, z \in \{x, y\}$ are defined as

$$M_z(t, T_0) = a_z \left(\frac{\phi_1^z \exp(\phi_2^z(T_0 - t))}{\phi_1^z + \phi_2^z (\exp(\phi_1^z(T_0 - t)) - 1)(1 + b_z(\phi_1^z - \phi_2^z))} \right)^{\phi_3^z}, \quad a_z = \prod_{j=0}^N A_z(T_0, T_j)^{k_j}$$

$$N_z(t, T_0) = \frac{b_z \phi_1^z + (\exp(\phi_1^z(T_0 - t)) - 1)(1 + b_z(\phi_1^z - \phi_2^z))}{\phi_1^z + \phi_2^z (\exp(\phi_1^z(T_0 - t)) - 1)(1 + b_z(\phi_1^z - \phi_2^z))}, \quad b_z = \sum_{j=0}^N k_j B_z(T_0, T_j),$$

where A_z, B_z are the functions defined in Theorem A1. The swap cumulants $c_l(t)$ at time t are now given by the formulas in Appendix D by setting $\mu_i := M^l(t), l = 1, \ldots, m$.

For the proof of this Lemma, we follow Tanaka et al. (2010) closely, which is referred to in Section 3.1.

The main result of this paper is the approximation of swaption prices by the Gram–Charlier expansion with short-rate (2), which follows immediately from Lemma 2 by using Proposition 2 and is referred to in Section 3.2.

Theorem 1. *Let everything be as in Lemma 2.*
The time t price of a $T_0 \times (T_N - T_0)$ payer ($\zeta = 1$) and receiver ($\zeta = -1$) swaption is given by

$$\text{Swaption}_{T_0}^{T_N}(t; K, \zeta) = P(t, T_0) \left(C_1 \mathcal{N}\left(\frac{C_1}{\sqrt{C_2}}\right) + \sqrt{C_2}\, \varphi\left(\frac{C_1}{\sqrt{C_2}}\right) \left(1 + \sum_{l=3}^{\infty} (-1)^l q_l H_{l-2}\right) \right),$$

where $\mathcal{N}$ denotes the cdf of the normal distribution, φ is the pdf or the normal distribution and H_l are the probabilist's Hermite polynomials (see Appendix C). The coefficients $q_0 = 1, q_1 = q_2 = 0$, and for $n \geq 3$

$$q_n = \sum_{m=1}^{\lfloor \frac{n}{3} \rfloor} \sum_{\substack{k_1 + \cdots + k_m = n \\ k_i \geq 3}} \frac{C_{k_1} \cdots C_{k_m}}{m! k_1! \cdots k_m!} \left(\frac{1}{\sqrt{C_2}}\right)^n$$

for $C_l := c_l(t)P(t, T_n)^l$ with $c_l(t)$ being the swap cumulants from Lemma 2 for fixed $t \geq 0$.

This formula will provide the necessary ingredient for the numerical experiments in Section 4, making it possible to calibrate the model to the swaption surface very efficiently. After successfully calibrating the model, we apply it to find constant maturity swap rates in Section 4.4 and Bermudan swaption pricing in Section 4.5 using the Least-Square Monte-Carlo technique. We will see a good performance of this model compared to the reference data downloaded from Bloomberg.

1.2. Review of the Literature and Comparison

Historically, the theory of interest-rate modeling started on the assumption of specific one-dimensional dynamics for the instantaneous spot rate process r. These models are convenient for defining all fundamental quantities (rates and bonds) by no-arbitrage arguments as it is the expectation of a functional of the process r. Indeed, the price at time $t > 0$ of a contingent claim with payoff H_T, $T > t$, under the risk-neutral measure Q is given by (cf. Pascucci 2011)

$$H_t = E_t^{\mathbb{Q}}\left[e^{-\int_t^T r(s)ds} H_T\right], \tag{3}$$

where $E_t^{\mathbb{Q}}$ denotes the conditional expectation with respect to some filtration $\mathcal{F}_t$ under measure $\mathbb{Q}$. In particular, choosing $H_T := P(T,T) = 1$, where $P(t,T)$ denotes a zero-coupon bond.

The literature on interest rate modeling is very vast, and our short literature review is by no means exhaustive. We refer to (Björk 2004; Brigo and Mercurio 2006; Hull 2006) for a comprehensive review and description of these models.

Among all possible classifications, we can divide these models into two major categories: the endogenous and exogenous models. In chronological order, the first short-rate models belong to the first group: the Vasicek model Vasicek (1977), the Dothan model and the Cox, Ingersoll and Ross (CIR) Cox et al. (1985). In particular, the CIR model has been regarded as the reference model in interest rate modeling by both practitioners and academics for several decades for several reasons. First of all, it was derived from a general equilibrium framework. Secondly, it generates more realistic interest rate distributions with skewness and a fatter tail with respect to normal distribution. Thirdly, it avoids negative interest rates. There is a rich literature on extensions to the classical CIR model in order to obtain more sophisticated models, which could fit the market data better, allowing to price interest rate derivatives more accurately. For example, Chen (1996), proposed a three-factor model; Brigo and Mercurio (2006), proposed a jump diffusion model (JCIR).

However, in the last decade, the financial industry encountered a paradigm shift by allowing the possibility of negative interest rates, making the classical CIR model unsuitable.

Recently, Orlando et al. suggest in several papers (cf. Orlando et al. 2019a, 2019b, 2020) a new framework that they call the CIR# model, which fits the market term-structure of interest rates. Additionally, it preserves the market volatility, as well as the analytical tractability of the original CIR model. Their new methodology consists of partitioning the entire available market data sample, which usually consists of a mixture of probability distributions of the same type. They use a technique to detect suitable sub-samples with normal or gamma distributions. In the next step, they calibrate the CIR parameters to shifted market interest rates, such that the interest rates are positive, and use a Monte Carlo scheme to simulate the expected value of interest rates.

In addition to historical reasons, endogenous models are important for their simplicity and analytical tractability, in particular for the possibility of pricing bonds and bond options analytically. However, there are some drawbacks. Since these models use only a few constant parameters, they are not able to simultaneously reproduce a given term-structure and volatility curve satisfactorily. Moreover, some shapes of the zero-coupon curve can never be reproduced (for example, an inverted shape curve with the Vasicek model). The need for an exact fit to the currently observed yield curve led some authors to introduce exogenous term-structure models. The first model was proposed by Ho and Lee (see HO and LEE 1986), but we believe the most popular among practitioners is the Hull and White extended Vasicek model (see Hull and White 1990). A generalization of this model with a good calibration to swaption market prices was found in Di Francesco (2012), while Mercurio and Pallavicini (2005) proposed a mixing Gaussian model coupled with parameter uncertainty. Moreover, in Russo and Torri (2019), the authors calibrate a one- and two-factor Hull–White model using swaptions under a market-consistent framework compatible with negative interest rates.

On the one hand, these models can handle negative interest rates with a very good analytical tractability. On the other hand, the distribution of continuously compounded interest rates shows all the undesirable features of the Gaussian distribution.

In this paper, we extend the endogenous model of Di Francesco and Kamm (2021) to an exogenous model by adding a deterministic shift and show how the Gram–Charlier expansion of Tanaka et al. (2010) can be utilized to calibrate our model to the swaption surface. We will see a good performance of the model with respect to determining constant maturity swap rates and pricing Bermudan swaptions.

We performed tests on two different dates 30 December 2019 and 30 November 2020. On the first date, the market zero rates were partially negative, and on the second date, they were completely negative. We saw similar numerical results on both dates and decided for the sake of brevity to only present the results on 30 December 2019. For the interested reader, we will make the data on 30 November 2020 as well as the code of the numerical implementation available online. The paper is organized as follows. In Section 2, we first introduce the deterministic shift extension and the corresponding zero-coupon price. This is followed by a reminder of the relevant features of swaps and swaptions in Section 2.1.

In Section 3, we will derive the Gram–Charlier expansion. This is achieved by first recalling how a probability density of a random variable can be approximated by Hermite polynomials. We will see that it is necessary to study the cumulants or, equivalently, the moments of this random variable. In our case, this will be the *swap moments*, and we will show how to derive them from the so-called *bond moments* by solving some Riccati equations, which will have explicit solutions in our model, making it possible to compute swaption prices very fast.

After that, in Section 4, we will conduct some numerical experiments. First, we calibrate our model to the market swaption surface at 30 December 2019 in Section 4.2. Subsequently, we simulate the model by using the Euler–Maruyama scheme in Section 4.3 and compute CMS rates in Section 4.4. We conclude our numerical tests by pricing Bermudan swaptions in Section 4.5. Finally, we summarize the results of the paper in Section 5 and discuss possible extensions for future research.

2. A Model for Negative Interest Rates with Perfect Fit to the Term-Structure

Let us define $\alpha := (\alpha_x, \alpha_y)$, $\alpha_z := (k_z, \theta_z, \sigma_z)$, $z \in \{x, y\}$. We want to use the general deterministic shift extension by (Brigo and Mercurio 2006, pp. 95 ff. Chapter 3.8 A General Deterministic-Shift Extension) or Brigo and Mercurio (2001) in the case of multifactor models. We note that contrary to the presented ideas in the aforementioned papers, we do not need to introduce another probability space for our purposes and will use the same risk-neutral measure for all dynamics.

Thus, we are interested in the following short rate model on $(\Omega, \mathcal{F}, \mathbb{Q})$

$$r(t; \alpha) := r^{\mathrm{CIR}\text{-}}(t; \alpha) + \psi(t; \alpha) = x(t; \alpha_x) - y(t; \alpha_y) + \psi(t; \alpha_x), \quad \psi(0; \alpha) = 0, \qquad (4)$$

where $r^{\mathrm{CIR}\text{-}}$ denotes the short-rate model without the deterministic shift extension. We will suppress the dependency on the parameters α for readability whenever there is no confusion.

Likewise, we recall from Theorem A1 in the Appendix A that the price of the zero-coupon bond for the non-extended model is given by

$$P^{\mathrm{CIR}\text{-}}(t, T) = A_x(t, T)e^{-B_x(t,T)x(t)}A_y(t, T)e^{B_y(t,T)y(t)}.$$

Analogous to (Brigo and Mercurio 2001, p. 5, Theorem 3.1), we easily derive the price of a zero-coupon bond in the short-rate model (4) for given parameters α

$$P(t, T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(s)ds\right)\bigg|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T x(s) - y(s) + \psi(s)ds\right)\bigg|\mathcal{F}_t\right]$$

$$= \exp\left(-\int_t^T \psi(s)ds\right)P^{\mathrm{CIR}\text{-}}(t, T)$$

because ψ is deterministic.

To ensure a perfect fit to the initial term-structure, we set as in (Brigo and Mercurio 2001, pp. 5–6, Corollary 3.2)

$$\psi(t; \alpha) = f^M(0, t) - f^\alpha(0, t),$$

where $f^M(0,t)$ is the instantaneous market forward rate and

$$f^\alpha(0,t) = -\frac{\partial_T(A_x(0,t))}{A_x(0,t)} + \partial_T(B_x(0,t))x(0) - \frac{\partial_T(A_y(0,t))}{A_y(0,t)} - \partial_T(B_y(0,t))y(0)$$

is the instantaneous market rate for $r^{\text{CIR-}}$ (see Appendix B).

More conveniently, we observe that this is equivalent to asking that the following equation holds

$$\exp\left(-\int_t^T \psi^\alpha(s)ds\right) = \frac{P^M(0,T)}{P^M(0,t)}\frac{P^{\text{CIR-}}(0,t)}{P^{\text{CIR-}}(0,T)},$$

where we used $P^z(t,T) = \exp\left(-\int_t^T f^z(t,s)ds\right), z \in \{M, \text{CIR-}\}$.

In total, this leads to the following formula for the zero-coupon price of the deterministic shift extended model

$$P(t,T) = \frac{P^M(0,T)}{P^M(0,t)}\frac{P^{\text{CIR-}}(0,t)}{P^{\text{CIR-}}(0,T)}P^{\text{CIR-}}(t,T) \qquad (5)$$

and $P(0,T) = P^M(0,T)$ is guaranteed.

2.1. Swaption Price Formula

In Di Francesco and Kamm (2021), we calibrated the short-rate model $r^{\text{CIR-}}$ to the initial term-structure. For the deterministic shift extended model, this is not possible since a perfect fit is guaranteed regardless of the parameters α. Therefore, we will calibrate to market swaption prices, for which we will recall all essentials in this section.

We are following (Björk 2004, pp. 428 ff. Chapter 27.7 Swaps); (Brigo and Mercurio 2006, pp. 19 ff.) and (Schrager and Pelsser 2006, pp. 3 ff.) in this section.

A swap is a financial contract between two counterparties with fixed resettlement dates $T_0, T_1, \ldots, T_N$, $N \in \mathbb{N}$. The contract itself contains two cashflows, one—called the floating leg—is payments of future interest rates and the other—called the fixed leg—is a fixed amount of payments. The receiver of a swap will receive at the fixed dates the amount of the fixed leg and pays the amounts of the floating leg to the other counterparty, giving it its name: the floating leg is swapped for the fixed leg. Additionally, a payer swap refers to the case, when the floating leg is received and the fixed rate paid. We will distinguish the different kinds by introducing the factor ζ, which will be equal to $+1$ in case of a payer swap and -1 in case of a receiver swap.

Such a contract with maturity T_0 and tenor $T_N - T_0$ and resettlements $T_0, \ldots, T_N$ is commonly called a $T_0 \times (T_N - T_0)$ swap.

The net value of a $T_0 \times (T_N - T_0)$ payer and receiver swap at time $t < T_0$ is given by

$$\text{Swap}_{T_0}^{T_N}(t; K, \zeta) := \zeta\left(P(t,T_0) - P(t,T_N) - K\sum_{i=1}^N \alpha_i P(t,T_i)\right) \qquad (6)$$

where $\alpha_i = T_i - T_{i-1}$ is the day-count convention and K the fixed rate, see for instance (Björk 2004, pp. 429 ff.). To ease notation, we will suppress the explicit dependency on the T_0 and T_N whenever there is no confusion.

A particular fixed rate K called *par or forward swap rate* is of special interest, which are usually quoted in the market. It is the one such that $\text{Swap}^N(t; K, \zeta) = 0$ (which is independent of ζ), and we will denote it a bit more generally by

$$R_n^N(t) := \frac{P(t,T_n) - P(t,T_N)}{\sum_{i=n+1}^N \alpha_i P(t,T_i)}, \quad n = 0, \ldots, N-1.$$

Moreover, we will denote the so-called *accrual factor* or *present value of a basis point* by

$$S_n^N(t) := \sum_{i=n+1}^{N} \alpha_i P(t, T_i), \quad n = 0, \ldots, N-1.$$

Now, we are able to discuss swaptions (cf. Björk 2004, pp. 430 ff.). A $T_0 \times (T_N - T_0)$ payer, receiver swaption with swaption strike K is a contract, at maturity T_0 gives the holder the right to enter into a $T_0 \times (T_N - T_0)$ payer, receiver swap with fixed rate K.

Its arbitrage free price at time $t < T_0$ is given by

$$\text{Swaption}_{T_0}^{T_N}(t; K, \zeta) = \mathbb{E}^{\mathbb{Q}} \left[\exp\left(- \int_t^{T_0} r(s) ds \right) \left(\zeta \left(R_0^N(T_0) - K \right) \right)^+ S_0^N(T_0) \Big| \mathcal{F}_t \right] \quad (7)$$

We will use this formulation for our Monte-Carlo calibration procedure together with (5).

Swaption Prices under the Forward Measure

For the Gram–Charlier expansion, we will make use of the fact that the stochastic discount factor in (7) can be removed by a clever change of measure. For fixed T_0, the T_0-forward measure $\mathbb{Q}^{T_0}$ is defined as the martingale measure for the numeraire process $p(t, T_0)$ (cf. Björk 2004, pp. 403 ff. Chapter 26.4 Forward measure), and we have the following:

Proposition 1. *For any T_0-claim X, we have*

$$\mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_0} r(s) ds} X \Big| \mathcal{F}_t \right] = P(t, T_0) \mathbb{E}^{\mathbb{Q}^{T_0}} [X | \mathcal{F}_t].$$

Thus, the price at time $t < T_0$ of a payer ($\zeta = 1$), receiver ($\zeta = -1$) swaption under the T_0-forward measure is given by

$$\text{Swaption}(t; K, \zeta) = P(t, T_0) \mathbb{E}^{\mathbb{Q}^{T_0}} \left[\left(\zeta \left(R_0^N(T_0) - K \right) \right)^+ S_0^N(T_0) \Big| \mathcal{F}_t \right] \quad (8)$$

3. Gram–Charlier Expansion

We will use all the results available in (Tanaka et al. 2010, pp. 3 ff. Section 2.1 Gram–Charlier expansion) and apply them to our case.

Let us first of all make the following observation: The payer ($\zeta = 1$) and receiver ($\zeta = -1$) swap value (6) can both be rewritten as

$$\text{Swap}(t; K, \zeta) := \sum_{i=0}^{N} a_i^\zeta P(t, T_i),$$

where a_i^ζ are equal to

$$a_0^\zeta := \zeta, \qquad a_N^\zeta := -\zeta(1 + K\alpha_N), \qquad a_i^\zeta := -\zeta K\alpha_i, \quad i = 1, \ldots, N-1.$$

For the remainder of this section, we will drop the dependency on ζ for the coefficients a_i to ease the notation. Now, with this notation, we can rewrite the swaption prices (8) to obtain

$$\text{Swaption}(t; K, \zeta) = P(t, T_0) \mathbb{E}^{\mathbb{Q}^{T_0}} \left[\left(\text{Swap}^N(T_0; K, \zeta) \right)^+ \Big| \mathcal{F}_t \right]$$

$$\stackrel{!}{=} P(t, T_0) \int_0^\infty x f(x) dx,$$

for an unknown density function f. The idea of the Gram–Charlier expansion is to approximate this density function f by using the orthonormal basis of Hermite polynomials

(see Appendix C), which is the content of the next Proposition (cf. Tanaka et al. 2010, p. 3, Proposition 2.1; Cheng 2013, p. 5, Proposition 2.1.2).

Proposition 2. *Assume that a random variable Y has the continuous density function f and has finite cumulants c_k, $k \geq 1$. Then the following holds:*

1. *f can be expanded as*

$$f(x) = \sum_{n=0}^{\infty} \frac{q_n}{\sqrt{c_2}} H_n\left(\frac{x - c_1}{\sqrt{c_2}}\right) \varphi\left(\frac{x - c_1}{\sqrt{c_2}}\right),$$

where H_n are the probabilist's Hermite polynomials and φ the probability density function of the standard normal distribution, as well as $q_0 = 1$, $q_1 = q_2 = 0$, and for $n \geq 3$

$$q_n = \frac{1}{n!}\mathbb{E}\left[H_n\left(\frac{Y - c_1}{\sqrt{c_2}}\right)\right] = \sum_{m=1}^{\lfloor \frac{n}{3} \rfloor} \sum_{\substack{k_1 + \cdots + k_m = n \\ k_i \geq 3}} \frac{c_{k_1} \cdots c_{k_m}}{m! k_1! \cdots k_m!}\left(\frac{1}{\sqrt{c_2}}\right)^n.$$

2. *For any $a \in \mathbb{R}$*

$$\mathbb{E}[Y \mathbb{1}_{Y \geq a}] = c_1 \mathcal{N}\left(\frac{c_1 - a}{\sqrt{c_2}}\right) + \sqrt{c_2}\varphi\left(\frac{c_1 - a}{\sqrt{c_2}}\right)$$
$$+ \sum_{n=3}^{\infty} (-1)^{n-1} q_n \varphi\left(\frac{c_1 - a}{\sqrt{c_2}}\right)\left[a H_{n-1}\left(\frac{c_1 - a}{\sqrt{c_2}}\right) - \sqrt{c_2} H_{n-2}\left(\frac{c_1 - a}{\sqrt{c_2}}\right)\right],$$

where, furthermore, $\mathcal{N}$ denotes the cumulative distribution function of the standard normal distribution.

In particular, we have

$$q_3 = \frac{c_3}{3! c_2^{\frac{3}{2}}}, \qquad q_4 = \frac{c_4}{4! c_2^{\frac{5}{2}}}, \qquad q_5 = \frac{c_5}{5! c_2^{\frac{6}{2}}}, \qquad q_6 = \frac{c_6 + 10 c_3^2}{6! c_2^{\frac{6}{2}}}, \qquad q_7 = \frac{c_7 + 35 c_3 c_4}{7! c_2^{\frac{7}{2}}}$$

Therefore, all that is required is to determine the swap cumulants. This will be performed in several steps: First, we will use the fact that cumulants can be computed from moments, see Appendix D. Second, we compute the so-called swap moments in Equation (9), which in turn are computed from so-called bond moments. Last but not least, in order to compute the bond moments, we need to derive a new system of Riccati equations in Equation (11).

3.1. Bond and Swap Moments

Since cumulants can be expressed by moments and vice versa (see Appendix D), we will study the *Swap moments* in this section, which we denote by

$$M^m(t) := \mathbb{E}^{\mathbb{Q}^{T_0}}\left[(\text{Swap}(T_0))^m \big| \mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}^{T_0}}\left[\left(\sum_{i=0}^{N} a_i P(T_0, T_i)\right)^m \Bigg| \mathcal{F}_t\right].$$

It can be shown by induction that the m-th power can be rewritten as

$$\left(\sum_{i=0}^{N} a_i P(T_0, T_i)\right)^m = \sum_{0 \leq i_1, \ldots, i_m \leq N} a_{i_1} \cdots a_{i_m}\left(\prod_{k=1}^{m} P(T_0, T_{i_k})\right).$$

Now, notice that all a_i are $\mathcal{F}_t$ measurable, and therefore,

$$M^m(t) = \sum_{0 \le i_1,\dots,i_m \le N} a_{i_1} \cdots a_{i_m} \mathbb{E}^{\mathbb{Q}^{T_0}} \left[\prod_{k=1}^m P(T_0, T_{i_k}) \middle| \mathcal{F}_t \right] \tag{9}$$

and we will call $\mathbb{E}^{\mathbb{Q}^{T_0}} \left[\prod_{k=1}^m P(T_0, T_{i_k}) \middle| \mathcal{F}_t \right]$ the *bond moments*.

Similar to (Cheng 2013, pp. 44–46), we will reduce the problem to finding the bond moments for the short-rate model without a deterministic-shift extension by using (5)

$$
\begin{aligned}
M^m(t) &= \sum_{0 \le i_1,\dots,i_m \le N} a_{i_1} \cdots a_{i_m} \mathbb{E}^{\mathbb{Q}^{T_0}} \left[\prod_{k=1}^m P(T_0, T_{i_k}) \middle| \mathcal{F}_t \right] \\
&= \sum_{0 \le i_1,\dots,i_m \le N} a_{i_1} \cdots a_{i_m} \mathbb{E}^{\mathbb{Q}^{T_0}} \left[\prod_{k=1}^m \frac{P^M(0, T_{i_k})}{P^M(0, T_0)} \frac{P^{\text{CIR-}}(0, T_0)}{P^{\text{CIR-}}(0, T_{i_k})} P^{\text{CIR-}}(T_0, T_{i_k}) \middle| \mathcal{F}_t \right] \\
&= \left(\frac{P^{\text{CIR-}}(0, T_0)}{P^M(0, T_0)} \right)^m \sum_{0 \le i_1,\dots,i_m \le N} a_{i_1} \cdots a_{i_m} \prod_{k=1}^m \frac{P^M(0, T_{i_k})}{P^{\text{CIR-}}(0, T_{i_k})} \mathbb{E}^{\mathbb{Q}^{T_0}} \left[\prod_{k=1}^m P^{\text{CIR-}}(T_0, T_{i_k}) \middle| \mathcal{F}_t \right] \\
&= \left(\frac{P^{\text{CIR-}}(0, T_0)}{P^M(0, T_0)} \right)^m \sum_{0 \le i_1,\dots,i_m \le N} a_{i_1}^* \cdots a_{i_m}^* \mathbb{E}^{\mathbb{Q}^{T_0}} \left[\prod_{k=1}^m P^{\text{CIR-}}(T_0, T_{i_k}) \middle| \mathcal{F}_t \right],
\end{aligned}
$$

where $a_{i_k}^* = a_{i_k} \frac{P^M(0, T_{i_k})}{P^{\text{CIR-}}(0, T_{i_k})}$.

Thus, we only have to calculate the bond moments for the CIR- model.

For a numerical implementation, the m-fold sum over all permutations of i_k is unfavorable. Therefore, we rewrite it as follows: By definition, there will always be m coefficients a_{i_k} in the m-fold sum, but it is possible to obtain, e.g., a_{i_1} twice, etc. Hence, fixing indices for a_0 up to a_N, we can sum over the powers of all occurrences, which have to sum up to m. However, the individual products of the coefficients can appear multiple times as well, e.g., for $m = 2, N = 2$ summing over all permutations would lead to two times the term $a_0 a_1$ since we encounter $i_0 = 0$, $i_1 = 1$ and $i_1 = 1$, $i_0 = 0$. Finally, we derive, similar to (Cheng 2013, p. 28, Remark 4.2.1), the following expression

$$\sum_{0 \le i_1,\dots,i_m \le N} a_{i_1} \cdots a_{i_m} \left(\prod_{k=1}^m P(T_0, T_{i_k}) \right) = \sum_{\substack{0 \le k_0,\dots,k_N \le N \\ k_0 + \cdots + k_N = m}} \frac{m!}{k_0! \cdots k_N!} a_0^{k_0} \cdots a_N^{k_N} \left(\prod_{j=0}^N P(T_0, T_j)^{k_j} \right).$$

Finding this set of indices is known as the *subset sum problem*, which is NP-hard but can be solved by, e.g., dynamical programming. The interested reader is referred to Curtis and Sanches (2017) for recent developments using a GPU for large subset sum problems. In our case, m will be at most 7, and due to annual payments, N will be at most equal to the maximal tenor plus one, i.e., 11, which is considered as a small subset sum problem for which we will utilize a simpler implementation. Even with semi-annual payments, a simple implementation with dynamic programming is sufficient, since we will need to calculate the subset sum problems only once and pass it to the calibration procedure.

Now, let us derive the Riccati equation for the bond moments. First of all, notice that the affine structure of $P(t, T)$ is preserved

$$\prod_{j=0}^{N} P^{\text{CIR-}}(T_0, T_j)^{k_j}$$

$$= \prod_{j=0}^{N} \left(A_x(T_0, T_j) e^{-B_x(T_0,T_j)x(T_0)} A_y(T_0, T_j) e^{B_y(T_0,T_j)y(T_0)} \right)^{k_j}$$

$$= \left(\prod_{j=0}^{N} A_x(T_0, T_j)^{k_j} \right) e^{-\sum_{j=0}^{N} k_j B_x(T_0,T_j)x(T_0)} \left(\prod_{j=0}^{N} A_y(T_0, T_j)^{k_j} \right) e^{\sum_{j=0}^{N} k_j B_y(T_0,T_j)y(T_0)}$$

$$=: A_x(T_0, \{k_0, \ldots, k_N\}) e^{-B_x(T_0,\{k_0,\ldots,k_N\})x(T_0)} A_y(T_0, \{k_0, \ldots, k_N\}) e^{B_y(T_0,\{k_0,\ldots,k_N\})y(T_0)}$$

(10)

By Proposition 1, we have also for $t \leq T_0$

$$\mathbb{E}^{\mathbb{Q}^{T_0}} \left[\prod_{j=0}^{N} P^{\text{CIR-}}(T_0, T_j)^{k_j} \middle| \mathcal{F}_t \right]$$

$$= \frac{1}{P^{\text{CIR-}}(t, T_0)} \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_0} r^{\text{CIR-}}(s)ds} A_x(T_0, \{k_0, \ldots, k_N\}) e^{-B_x(T_0,\{k_0,\ldots,k_N\})x(T_0)} \right.$$

$$\left. A_y(T_0, \{k_0, \ldots, k_N\}) e^{B_y(T_0,\{k_0,\ldots,k_N\})y(T_0)} \middle| \mathcal{F}_t \right]$$

$$\overset{!}{=} \frac{1}{P^{\text{CIR-}}(t, T_0)} M_x(t, T_0) e^{-N_x(t,T_0)x(t)} M_y(t, T_0) e^{N_y(t,T_0)y(t)}.$$

We notice that by martingale pricing, the discounted price process $e^{-\int_0^t r^{\text{CIR-}}(s)ds} M_x(t, T_0)$ $e^{-N_x(t,T_0)} M_y(t, T_0) e^{N_y(t,T_0)}$ has to be a martingale. Since it has an affine structure as well, it places us exactly in the same situation as in the derivation of Lemma A1 seen in Di Francesco and Kamm (2021) with the difference of variable terminal conditions.

Therefore, we have the same Riccati equation but different terminal values dependent on $k_0, \ldots, k_N$.

For generic terminal values $a_z, b_z \in \mathbb{R}_{\geq 0}$, the explicit solution is given by

$$M_z(t, T_0) = a_z \left(\frac{\phi_1^z \exp(\phi_2^z(T_0 - t))}{\phi_1^z + \phi_2^z \left(\exp(\phi_1^z(T_0 - t)) - 1 \right) \left(1 + b_z(\phi_1^z - \phi_2^z) \right)} \right)^{\phi_3^z}, \qquad M_z(T_0, T_0) = a_z$$

$$N_z(t, T_0) = \frac{b_z \phi_1^z + \left(\exp(\phi_1^z(T_0 - t)) - 1 \right) \left(1 + b_z(\phi_1^z - \phi_2^z) \right)}{\phi_1^z + \phi_2^z \left(\exp(\phi_1^z(T_0 - t)) - 1 \right) \left(1 + b_z(\phi_1^z - \phi_2^z) \right)}, \qquad N_z(T_0, T_0) = b_z.$$

(11)

As seen from our derivation in Equation (10), the terminal values $t = T_0$ are equal to

$$a_z = A_z(T_0, \{k_0, \ldots, k_N\}), \quad b_z = B_z(T_0, \{k_0, \ldots, k_N\}), \quad z \in \{x, y\},$$

and we can now compute the bond moments and, therefore, the swap moments for the Gram–Charlier expansion. Thus, using the one-to-one relationship between moments and cumulants in Appendix D, we have an explicit formula for the swap cumulants and we can apply Proposition 2, which is part of the next subsection.

3.2. Expansion Formula

As described in Tanaka et al. (2010), we can now use Proposition 2 to formulate the Gram–Charlier expansion formula:

$$\text{Swaption}(t; K, \zeta) = P(t, T_0) \left(C_1 \mathcal{N} \left(\frac{C_1}{\sqrt{C_2}} \right) + \sqrt{C_2} \varphi \left(\frac{C_1}{\sqrt{C_2}} \right) \left(1 + \sum_{l=3}^{\infty} (-1)^l q_l H_{l-2} \right) \right)$$

where we replace the c_n in Proposition 2 by $C_n := c_n(t)P(t, T_0)^n$ for $n \geq 1$ and the swap cumulants $c_n(t)$ are derived from the swap moments $M^m(t)$ using their one-to-one relationship shown in Appendix D.

In the following, we will denote by

$$\text{GC}(L; K, \zeta) := P(t, T_0) \left(C_1 \mathcal{N} \left(\frac{C_1}{\sqrt{C_2}} \right) + \sqrt{C_2} \varphi \left(\frac{C_1}{\sqrt{C_2}} \right) \left(1 + \sum_{l=3}^{L} (-1)^l q_l H_{l-2} \right) \right)$$

the L-th order of the Gram–Charlier expansion of the $T_0 \times (T_N - T_0)$ swaption with strikes K and swaption type ζ with annual payment dates.

4. Numerical Tests

We will now perform some numerical experiments in our model. In Section 4.1, we will briefly discuss the market data, which we will use to perform all numerical tests in the subsequent sections. Afterward, we will describe the calibration procedure of our model in Section 4.2. This is followed by a short subsection on simulating the model with the Euler–Maruyama scheme in Section 4.3, and in Section 4.4, we investigate the par rates of constant maturity swaps (CMS). Last but not least, we compare the model Bermudan swaption prices to Bloomberg's Hull–White one factor model prices in Section 4.5.

For the calculations, we used Matlab 2021a with the (Global) Optimization Toolbox running on Windows 10 Pro, on a machine with the following specifications: processor Intel(R) Core(TM) i7-8750H CPU @ 2.20 GHz and 2x32 GB (Dual Channel) Samsung SODIMM DDR4 RAM @ 2667 MHz. All calculations were sped-up by multiprocessing on a single CPU whenever possible.

4.1. Market Data

To obtain the market zero-coupon bond term-structure, we first build the EUR Euribor-swap curve, which was created from the most liquid interest rate instruments available in the market and constructed as follows: We consider deposit rates and Euribor rates with maturity from one day to one year and par-swap rates versus six-month Euribor rates with maturity from two years to thirty years. Then, the zero interest curve and the zero-coupon bond curve are calculated using a standard "bootstrapping" technique in conjunction with cubic spline interpolation of the continuously compounded rate (cf. Miron and Swannell 1991 for more details).

We tested the model on two different dates 30 December 2019 and 30 November 2020. As mentioned in the introduction, we chose those particular dates because on 30 December 2019 the market zero rates were partially negative (up to year six) and on 30 November 2020 they were completely negative. This enables us to test the compatibility of the model in the most relevant different scenarios. However, since the results on 30 December 2019 and 30 November 2020 are very similar, we decided to present only the results on 30 December 2019 and make all the data on 30 November 2020 available online to shorten the presentation.

As aforementioned, we will calibrate the model to swaption prices (Table A3). They are computed by Bachelier's formula from normal volatilities quoted in the market (Table A1), and the swaption strikes can be found in Table A2.

After the calibration, we will assess the performance of the model by comparing its prediction of par CMS rates to Bloomberg's CMS rates in Section 4.4 and pricing Bermudan swaptions. The benchmark for Bermudan swaption prices will be Bloomberg's Hull–White one factor model alongside the corresponding strikes. The values are displayed in Tables A4–A6, respectively.

All data were downloaded from Bloomberg and used in the following subsections for our numerical experiments. We start in the next subsection by calibrating our model to the swaption surface.

4.2. Calibration

In this subsection, we will discuss how we use the Gram–Charlier expansion to calibrate our model to parts of the swaption surface in Table A3. Since we are using a deterministic shift extension, a perfect fit to the market zero-coupon curve (see Table A7) is always guaranteed. Let us denote the parameter vector by $\Pi := \left[\phi_1^x, \phi_2^x, \phi_3^x, \phi_1^y, \phi_2^y, \phi_3^y, x_0, y_0\right]^T \in \mathbb{R}_{>0}^8$. We will formulate the calibration procedure as a constraint minimization problem in $\mathbb{R}_{>0}^8$ for the parameters Π with the objective function

$$f(\Pi) := \sum_{l \in \mathcal{L}} \sum_{T_0 \in \mathcal{M}} \sum_{T_N \in \mathcal{T}} \left(\frac{\text{MarketSwaption}_{T_0}^{T_N}(K, \zeta)}{\text{GC}_{T_0}^{T_N}(l, \Pi; K, \zeta)} - 1 \right)^2, \tag{12}$$

where $\mathcal{L} \subset \mathbb{N}$ is a set of natural numbers containing the orders of the Gram–Charlier expansion, $\mathcal{M}$ is a set of maturities and $\mathcal{T}$ a set of final times. We will go into further details on how to choose these sets in Remark 1.

The objective function describes the relative square difference between the market swaption prices and the theoretical prices derived by the Gram–Charlier expansion using the short-rate model (4). We would like to note that one could also think of different objective functions, which might lead to slightly different results. For instance, we tested absolute squared errors, but the results were similar.

The set of admissible parameters $\mathcal{A}$ will consist of the following constraints arising from the well-definedness of the Formulas (A3):

1. First of all, let us note that there is a one-to-one correspondence between the parameters Π and k_z, σ_z and θ_z if one is looking for positive real solutions only. We have

$$k_x = 2\phi_2^x - \phi_1^x, \qquad\qquad k_y = 2\phi_2^y - \phi_1^y,$$
$$\sigma_x = \sqrt{2\left(\phi_2^x \phi_1^x - \left(\phi_2^x\right)^2\right)}, \qquad \sigma_y = \sqrt{-2\left(\phi_2^y \phi_1^y - \left(\phi_2^y\right)^2\right)}, \tag{13}$$
$$\theta_x = -\frac{\phi_x^2 \phi_x^3 (\phi_x^1 - \phi_x^2)}{\phi_x^1 - 2\phi_x^2}, \qquad \theta_y = \frac{\phi_y^2 \phi_y^3 (\phi_y^1 - \phi_y^2)}{\phi_y^1 - 2\phi_y^2}.$$

2. We require $\sigma_z \in \mathbb{R}_{\geq 0}$, $z \in \{x, y\}$. By rearranging (13), these conditions are equivalent to $\phi_1^x \geq \phi_2^x$ and $\phi_2^y \geq \phi_1^y$;
3. A positive mean-reversion speed, i.e., $k_z \geq 0$, is equivalent to $2\phi_2^z \geq \phi_1^z$, $z \in \{a, b\}$;
4. The Feller condition $2k_z\theta_z \geq \sigma_z^2$ is equivalent to $\phi_3^z \geq 1$, $z \in \{a, b\}$;
5. A positive mean for each CIR process, i.e., $\theta_z \geq 0$, is by positivity of σ_z^2 and k_z equivalent to $\phi_3^z \geq 0$, which is already satisfied by the Feller condition;
6. The parameter ϕ_1^z, assuming that it is real-valued, is positive by definition, meaning that by the positivity of the mean reversion speed, ϕ_2^z will be as well. Therefore, all ϕ are positive;
7. As both CIR processes x_t and y_t, individually, are positive processes, we additionally require $x_0 \geq 0$ and $y_0 \geq 0$.

The advantage of using the parameters Π instead of k_z, σ_z and θ_z is that we can rewrite these conditions as a system of linear inequality constraints in matrix notation $A \cdot \Pi \leq 0$, where

$$A := \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 0 \end{bmatrix}$$

with boundary conditions $\Pi_i \geq 0$, $i = 1, \ldots, 8$, and $\Pi_3 = \phi_3^x \geq 1$, as well as $\Pi_6 = \phi_3^y \geq 1$.

In total, the set of admissible parameters is given by

$$\mathcal{A} := \left\{ \Pi \in \mathbb{R}^8_{\geq 0}, \Pi_3, \Pi_6 \geq 1 : A \cdot \Pi \leq 0 \right\}. \tag{14}$$

Finally, a solution Π^* to the calibration problem is a minimizer of

$$\min_{\Pi \in \mathcal{A}} f(\Pi). \tag{15}$$

Before we present some results, we would like to make the following remark on the choices of $\mathcal{L}$, $\mathcal{M}$ and $\mathcal{T}$.

Remark 1. *As always, in calibration procedures with parametrized models, there is the notion of over- and underfitting the data. Overfitting usually occurs when there are more parameters than independent values to calibrate to. For example, we saw a very good fit for a single swaption price. Underfitting, on the other hand, occurs when the model is not able to fit to the whole data, e.g., fitting this model to the entire swaption surface.*

In our experiments, we determined that four to six values performed best with regards to the Bermudan swaption pricing (Section 4.5) and finding the CMS par rates (Section 4.4). This is not very surprising since the model has in total eight parameters, but since the two CIR processes are independent and subtracted to deal with the negative interest rates, it has essentially four parameters to model the data.

Therefore, we decided to perform tests on columns of the swaption surface and excluded short maturities. Additionally, removing the last maturity in the column from the calibration increased the speed of the optimization with usually the same accuracy. Additionally, we performed tests on several diagonals of the swaption surface with similar results and, therefore, decided to focus only on columns in this paper.

Another aspect of this calibration procedure is the question of which orders to use the Gram–Charlier expansion. Since it is an orthogonal expansion, there is no a priori error estimate of the truncated expansion formula. This also means that increasing the order might not be beneficial for the accuracy. By comparing the Gram–Charlier swaption prices with Monte-Carlo swaption prices (see Section 4.3) using the same parameters, we found both prices to be closer to each other if we were using the orders three, five and seven in the calibration procedure. A non-rigorous and heuristic idea behind this reasoning is that if the three orders are close to each other, then the expansion "converges" to the correct price of the swaption in a loose sense.

To conclude, to avoid over- and underfitting, we will calibrate to columns of the swaption surface starting with maturity five and ending with maturity 15. Moreover, to have a "stable" Gram–Charlier swaption price, we will use the orders three, five and seven in all experiments.

To solve (15) numerically, we would like to use `Matlab`'s function fmincon in the (Global) Optimization Toolbox. In order to use this function, we need an initial guess of the parameter Π, and the computational time will depend on that choice.

Our experiments showed that initial guesses with small admissible values worked best for fmincon. Therefore, we use the following hand-made parameters as initial points for fmincon

$$I_1 := [0.1, 0.095, 0.3, 0.095, 0.1, 0.3, 0.01, 0.01]^T, \qquad I_2 := \frac{1}{2}I_1$$

and compare the performance to parameters found by `Matlab`'s function ga. For the algorithms, e.g., the interior point algorithm for `fmincon`, used by `Matlab`, we refer the reader to Gilli et al. (2011), in the context of financial mathematics.

In Table 1, we show the value of (12) after the calibration procedure and its computational time in seconds in the case of a payer swaption on 30 December 2019. We display four different choices of initial points, first of all only using ga, second ga as an initial point for fmincon, third I_1 as initial point for fmincon and last but not least I_2 for fmincon. We

can see that the model fits the swaption values best using columns with larger tenor but the computational time increases as well for all methods. Furthermore, we can see that our choices I_1 and I_2 in conjunction with `fmincon` outperforms `ga` with respect to accuracy and it is significantly faster than the combination of `ga` and `fmincon`. Therefore, we will use in the following experiments only `fmincon` with I_1 or I_2 to present the results. In Table 2, we show the results of (15) with initial point I_1 using `fmincon` for reproducibility.

Table 1. Computational times and values of (12) using different initial points and different swaption columns in Table A3 and corresponding strikes in Table A2 in the case of payer swaptions and maturities ranging from 5 to 15.

Tenor Method	1	2	5	7	10
ga	$f(\Pi)=$ 3.94×10^{-2} in 76.2 s	$f(\Pi)=$ 7.12×10^{-2} in 85.8 s	$f(\Pi)=$ 5.75×10^{-2} in 100 s	$f(\Pi)=$ 2.27×10^{-2} in 168 s	$f(\Pi)=$ 1.79×10^{-2} in 891 s
ga & fmincon	$f(\Pi)=$ 3.94×10^{-2} in 76.6 s	$f(\Pi)=$ 7.92×10^{-2} in 87.6 s	$f(\Pi)=$ 6.61×10^{-3} in 118.6 s	$f(\Pi)=$ 1.12×10^{-3} in 206.6 s	$f(\Pi)=$ 8.04×10^{-4} in 945.9 s
I_1 & fmincon	$f(\Pi)=$ 7.90×10^{-2} in 0.9 s	$f(\Pi)=$ 4.78×10^{-2} in 0.8 s	$f(\Pi)=$ 6.62×10^{-3} in 2.47 s	$f(\Pi)=$ 1.10×10^{-3} in 52 s	$f(\Pi)=$ 3.00×10^{-4} in 181 s
I_2 & fmincon	$f(\Pi)=$ 8.62×10^{-1} in 0.3 s	$f(\Pi)=$ 5.80×10^{-1} in 1.35 s	$f(\Pi)=$ 6.55×10^{-3} in 33.3 s	$f(\Pi)=$ 1.12×10^{-3} in 49.9 s	$f(\Pi)=$ 6.95×10^{-4} in 93.9 s

Table 2. Calibrated parameters Π^* using I_1 with `fmincon` and different swaption columns in Table A3 and corresponding strikes in Table A2 in the case of payer swaptions and maturities ranging from 5 to 15.

Tenor Π^*	1	2	5	7	10
ϕ_1^x	0.082	0.114	0.109	0.113	0.118
ϕ_2^x	0.0477	0.0947	0.0846	0.0899	0.092
ϕ_3^x	1.05	1.13	1.99	2	2
ϕ_1^y	0.155	0.0241	0.584	0.00192	0.00741
ϕ_2^y	0.165	0.0521	0.597	0.00851	0.00151
ϕ_3^y	1.33	1.19	1.26	1.78	1.73
x_0	0.000126	0.00147	0.00017	0.000107	0.00151
y_0	0.000128	0.0024	0.0021	0.0991	0.0988

4.3. Euler–Monte-Carlo Simulation

In order to forecast the future expected interest rate for, e.g., pricing Bermudan swaptions in Section 4.5, we use the Euler–Maruyama scheme to simulate the instantaneous spot rate r (2). We refer to Dereich et al. (2012) and the references therein for a list of different Euler-type methods to simulate a CIR process. In our experiments, we simulate the processes $x(t)$ and $y(t)$ by the truncated Euler scheme defined as follows:

First of all, we fix a homogeneous time grid $0 = t_0 \leq t_1 \leq \cdots \leq t_N = T$ for the interval $[0, T]$ with $N + 1$ time points and mesh $\Delta t_i := t_{i+1} - t_i \equiv \Delta := \frac{T}{N}$ for all $i = 0, \ldots, N-1$. Secondly, we simulate the two independent Brownian motions W_z, $z \in \{x, y\}$, and define their time increment as $\Delta W_z(t_i) := W_z(t_{i+1}) - W_z(t_i)$. In total, we compute $r(t_{i+1}) := x(t_{i+1}) - y(t_{i+1})$ for $i = 0, \ldots, N-1$, where

$$x(t_{i+1}) = x(t_i) + k_x(\theta_x - x(t_i))\Delta t_i + \sigma_x \sqrt{\max(x(t_i), 0)} \Delta W_x(t_i)$$

$$y(t_{i+1}) = y(t_i) + k_y(\theta_y - y(t_i))\Delta t_i + \sigma_y \sqrt{\max(y(t_i), 0)} \Delta W_y(t_i).$$

(16)

We choose the max inside the square-root to ensure that the square-root remains real because, due to discretization effects, the positivity of $x(t_i)$ and $y(t_i)$ might be violated.

In all of our experiments, we use $M = 10,000$ simulations and mesh size $\Delta = \frac{1}{256}$. On the one hand, looking at the fast calibration times using the Gram–Charlier approximation in Section 4.2, it is clear that Monte-Carlo methods cannot compete with respect to speed. On the other hand, since the Gram–Charlier expansion has no a priori error bound, let us now validate the calibration results by computing the Monte-Carlo prices with the parameters obtained by the Gram–Charlier expansion in Table 2. In Table 3, we compare the swaption prices obtained by selected orders of the Gram–Charlier expansion to the Monte-Carlo prices and also the Monte-Carlo prices to the market prices. To compare the prices, we will use an average absolute error, i.e., for $X, Y \in \mathbb{R}^{d_1, d_2}$

$$\|X - Y\| := \frac{1}{d_1 d_2} \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} |X_{ij} - Y_{ij}|.$$

Table 3. Average absolute errors of Monte-Carlo prices compared to Gram–Charlier prices and market prices using the parameters shown in Table 2.

Tenor Methods	1	2	5	7	10
MC − GC3	4.96×10^{-4}	1.17×10^{-3}	5.48×10^{-4}	6.23×10^{-4}	6.18×10^{-4}
MC − GC5	2.90×10^{-4}	1.07×10^{-4}	9.54×10^{-4}	2.93×10^{-4}	3.57×10^{-4}
MC − GC7	7.65×10^{-4}	1.05×10^{-3}	2.14×10^{-4}	1.97×10^{-4}	2.86×10^{-4}
MC − Market	3.93×10^{-4}	8.99×10^{-4}	4.58×10^{-4}	3.85×10^{-4}	3.73×10^{-4}

The average absolute error between the Gram–Charlier orders and the Monte-Carlo prices are usually of order 10^{-4}, and the Monte-Carlo prices compared to the market prices usually of order 10^{-4}, as well. It is important to note while reading this table that the prices themselves are usually of order 10^{-2}; therefore, the accuracy is usually up to two significant orders. Hence, this validates the parameters obtained by the calibration with the Gram–Charlier expansion, and we can proceed with finding CMS rates in the next subsection using Monte-Carlo techniques.

4.4. Pricing Constant Maturity Swaps (CMS)

In this section, we want to use the calibrated model to compute the par rates of constant maturity swaps (CMS) using Monte-Carlo simulation. We refer the reader to (Brigo and Mercurio 2006, pp. 557 ff. Section 13.7 Constant-Maturity-Swaps; Tanaka et al. 2010, pp. 7 ff.) for more details.

Let us recall the definition of a CMS:

Definition 1. *A constant maturity swap (CMS) is a variant of an interest rate swap between two parties, such that at each payment date starting at T_0 and ending at T_N a fixed rate K is swapped with a c-year swap rate.*

Analogously, we distinguish between payer and receiver CMS. In receiver CMS, the fixed rate is received and the floating rate paid, and vice versa for payer CMS.

Furthermore, as before, we will assume annual settlements between the effective date T_0 and maturity T_N and denote the payment dates by $\mathcal{T} := \{T_0, T_1, \ldots, T_N\}$. The net value of a $T_0 \times T_N + c$ CMS with fixed rate K and index c at time 0 under the risk-neutral measure is

$$\text{CMS}_{T_0}^{T_N}(0; K, c, \zeta) := \mathbb{E}^{\mathbb{Q}} \left[\sum_{i=1}^{N} \exp\left(- \int_0^{T_{i-1}} r(s) ds \right) \zeta \alpha_i \left(R_{i-1}^{i-1+c}(T_{i-1}) - K \right) \right]. \tag{17}$$

By rearranging (17), we can compute the par CMS rates by setting it to zero and solve for K, i.e.,

$$K = \frac{\mathbb{E}^{\mathbb{Q}}\left[\sum_{i=1}^{N} \alpha_i \exp\left(-\int_0^{T_{i-1}} r(s)ds\right) R_{i-1}^{i-1+c}(T_{i-1})\right]}{\mathbb{E}^{\mathbb{Q}}\left[\sum_{i=1}^{N} \alpha_i \exp\left(-\int_0^{T_{i-1}} r(s)ds\right)\right]}$$

$$= \frac{\mathbb{E}^{\mathbb{Q}}\left[\sum_{i=1}^{N} \alpha_i \exp\left(-\int_0^{T_{i-1}} r(s)ds\right) R_{i-1}^{i-1+c}(T_{i-1})\right]}{\sum_{i=1}^{N} \alpha_i P(0, T_{i-1})}.$$

Remember that by the deterministic shift extension, we have $P(0, T) = P^M(0, T)$ in our model.

In our experiment, we will use Monte-Carlo simulation for the short-rate (2) and display the results in Table 4 using the initial parameters I_2 for `fmincon` in the case of payer swaptions. In the first column, we see the effective date T_0, in the second the tenor T, such that $T_N = T_0 + T$ and in the third column the index c for the CMS. The next column shows Bloomberg's CMS rates, which is followed by the model CMS rates. In the last column, we can see the absolute error of market and model rates. We can observe that the majority of CMS rates are very close to each other, telling us that the model performs well on average using just one column of the swaption data for the calibration. Using different columns in the calibration for all different CMS rates would improve the results further.

Table 4. CMS rates computed with a calibration using I_2 and `fmincon` to the column with tenor 7 of the payer swaption surface with maturities ranging from 5 to 15.

Effective Date	Tenor	Index	Bloomberg's CMS Rate	Model CMS Rate	Abs Error
0	5	5	0.00145	0.00154	8.91×10^{-5}
0	10	5	0.00472	0.00499	0.000273
0	5	10	0.00465	0.0047	4.67×10^{-5}
0	10	10	0.00732	0.00738	6.06×10^{-5}
3	5	5	0.00562	0.00584	0.000226
3	5	10	0.00824	0.00825	8.32×10^{-6}
5	10	5	0.00999	0.01	3.64×10^{-5}
5	5	5	0.00958	0.00847	0.00112
5	5	10	0.011	0.0101	0.000912

4.5. Pricing Bermudan Swaptions

In this section, we want to use the calibrated model to compute the prices of Bermudan swaptions. A popular choice of literature on this subject is, e.g., (Brigo and Mercurio 2006, pp. 588 ff. Section 13.15 LFM: Pricing Bermudan Swaptions; Glasserman 2004, pp. 421 ff. Chapter 8 Pricing American Options) or more recently (Gatarek and Jabłecki 2021; Oosterlee and Grzelak 2019, pp. 422 ff. Section 13.3.2 European and Bermudan option example).

Now, let us define which type of Bermudan swaptions we are interested in.

Definition 2. *A T_N no-call T_0 or T_N nc T_0 Bermudan swaption with annual exercise dates gives its holder the right but not the obligation to enter at any time $\mathcal{T}_E^N := \{T_0, T_1, \ldots, T_{N-1}\}$ into an interest rate swap with first reset $T \in \mathcal{T}_E$, last payment T_N and fixed rate K.*

Let us give a quick example of a 10 nc 2 Bermudan swaption with annually spaced exercise dates. The holder can exercise this option starting from year two and afterward at the beginning of each consecutive year but not later than year nine. After exercising the option, the holder enters into a swap contract—for simplicity with annual settlements—ending at year 10.

Accordingly, the price at time t of a T_N nc T_0 Bermudan swaption is the solution to the following optimal stopping problem

$$\text{BSwaption}_{T_0}^{T_N}(t; K, \zeta) := \sup_{\substack{\tau \in \mathcal{T}_E^N \\ \tau \text{ stopping time}}} \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^\tau r(s)ds} S_\tau^N(\tau)\left(\zeta\left(K - R_\tau^N(\tau)\right)\right)^+ \right],$$

where the filtration is generated by the forward swap rate, i.e., $\mathcal{F}_t := \sigma\left(R_s^N(s) : s \leq t\right)$ augmented such that it satisfies the usual hypothesis.

For the implementation, we are interested in the special case of today's price, i.e., $t = 0$. We will use backward induction to compare the exercise value to the continuation value and compute the conditional expectations by the least square Monte-Carlo (LSMC) method (cf. Longstaff and Schwartz 2001). Let us be more precise:

We know that the price at time T_{N-1} is given by

$$\text{BSwaption}_{T_0}^{T_N}(T_{N-1}; K, \zeta) = \mathbb{E}_{T_{N-1}}^{\mathbb{Q}}\left[e^{-\int_{T_{N-1}}^{T_{N-1}} r(s)ds} S_{N-1}^N(T_{N-1})\left(\zeta\left(K - R_{N-1}^N(T_{N-1})\right)\right)^+ \right]$$

$$= P(T_{N-1}, T_N)\left(\zeta\left(K - R_{N-1}^N(T_{N-1})\right)\right)^+$$

by definition and measurability as well as the fact that the stopping time can only be equal to T_{N-1} in this case. This gives us the opportunity to inductively calculate the Bermudan swaption price backward. Thus, let us now assume that $\text{BSwaption}_{T_0}^{T_N}(T_{i+1}; K, \zeta)$ for $i = N - 2, \ldots, 0$ is known.

We would like to compare the so-called *continuation value*, which is the expected future payoff if the option is not exercised to the exercise value at all times $\mathcal{T}_E^N$, and is defined as

$$c(T_i) := \mathbb{E}_{T_i}^{\mathbb{Q}}\left[e^{-\int_{T_i}^{T_{i+1}} r(s)ds} \text{BSwaption}_{T_0}^{T_N}(T_{i+1}; K, \zeta) \right].$$

Since the optimal stopping time will pathwise choose the maximum of continuing the option or exercising it, we have a dynamic programming principle

$$\text{BSwaption}_{T_0}^{T_N}(T_i; K, \zeta) = \begin{cases} P(T_{N-1}, T_N)\left(\zeta\left(K - R_{N-1}^N(T_{N-1})\right)\right)^+, & i = N - 1 \\ \max\left(c(T_i), S_i^N(T_i)\left(\zeta\left(K - R_i^N(T_i)\right)\right)^+\right), & i = 0, \ldots, N - 2. \end{cases}$$

The price at time $t = 0$ is then given by

$$\text{BSwaption}_{T_0}^{T_N}(0; K, \zeta) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^{T_0} r(s)ds} \text{BSwaption}_{T_0}^{T_N}(T_0; K, \zeta) \right].$$

For completeness, we explain how to approximate the conditional expectation with the LSMC method in Appendix E. For the numerical implementation, we choose the polynomial basis.

In Table 5, we can see the average absolute error of the Bermudan swaption prices in our model compared to Bloomberg's prices. We used I_1 and I_2 s initial points for `fmincon` in the case of receiver and payer swaptions with different tenors. We can see that the average errors are very sensitive with respect to the calibrated parameters by looking at the results of I_1 and I_2 for a fixed tenor. Additionally, we notice that usually the results are better if we choose I_2 as an initial point. The best results, on average, are found while calibrating to the columns of the swaption surface with tenor 5 or 7. In Table 6, we show the absolute errors for the individual payer Bermudan swaptions using I_1 as the initial point calibrated to the column with tenor 5 and see an overall good match. Particularly, the column with tenor 7 in Table 5 is very accurate.

We focused in this experiment on the average errors only and not on specific Bermundan swaptions. If one desires to do so, there might be better choices as to which swaption prices to use for the calibration. Usually, the so-called co-terminal swaption prices are used to achieve better results for a specific Bermudan swaption. Since we are satisfied with the average performance of the model, we will not perform these individual tests for the sake of brevity.

Table 5. Average absolute errors of Monte-Carlo Bermudan swaption prices and Bloomberg's HW1 Bermudan swaption prices using the I_1 and I_2 as initial points for `fmincon`.

Methods \ Tenor	1	2	5	7	10
I_1 & `fmincon` (Payer)	0.0254	0.0129	0.0014	0.0073	0.203
I_2 & `fmincon` (Payer)	0.00196	0.00991	0.00269	0.00279	0.0088
I_1 & `fmincon` (Receiver)	0.948	0.0642	0.0036	0.0021	0.0102
I_2 & `fmincon` (Receiver)	0.00615	0.0149	0.0033	0.0021	0.0071

Table 6. Absolute errors of Monte-Carlo Bermudan payer swaption prices and Bloomberg's HW1 Bermudan swaption prices using the I_1 as initial points for `fmincon` calibrated to the column with tenor equal to 5.

Maturity \ Tenor	2	5	7	10
1	1.295×10^{-3}	5.452×10^{-4}	6.337×10^{-4}	2.488×10^{-3}
3	1.026×10^{-3}	6.628×10^{-4}	9.348×10^{-4}	2.931×10^{-3}
5	1.284×10^{-3}	1.404×10^{-3}	1.629×10^{-4}	3.605×10^{-3}
7	9.416×10^{-4}	1.191×10^{-3}	4.314×10^{-6}	2.271×10^{-3}
10	1.267×10^{-3}	1.305×10^{-3}	1.603×10^{-3}	2.470×10^{-3}

5. Conclusions

In this paper, we extended the short-rate of Di Francesco and Kamm (2021) by applying the deterministic-shift extension. We derived the swaption prices by using the Gram–Charlier expansion in this model and calibrated it to columns of the market swaption surface. The calibration is fast and accurate. Using Monte-Carlo techniques, we obtained close CMS rates compared to Bloomberg's rates. Furthermore, compared to Bloomberg's Bermudan swaption prices via the HW1 model, our model performed very well.

Let us discuss some limitations and potentials for future research on this model and methodology. Introducing correlations between the Brownian motions might impact the ability to solve the Riccati equations analytically, such that numerical solutions would be necessary. This would clearly impact the computational effort. Considering piecewise constant coefficients in the CIR processes might lead to similar problems but would make the model more flexible and could lead to a better fit the entire swaption surface. These problems are subject to future research.

Motivated by Russo and Torri (2019), we would like to extend the model by adding multiple risk factors. This enables the model to generate more yield-curve shapes. In this line of thought, it is possible to add an uneven number of risk factors, e.g., two risk factors for the positive CIR process but keep only one for the negative CIR process and vice versa. The impact of such choices are subject to future research. Again adding correlation structures for multiple risk factors is also a possible extension.

Furthermore, in this paper, we only considered the CIR framework and especially only Brownian motions as risk-factors. Another possible extension could be the JCIR framework by adding a jump process to (2).

Thanks to an anonymous referee, we were made aware that a discussion of how realistic this model is is in order. For a more in-depth discussion, we point the reader to Grabinski and Klinkova (2019) and the references therein. As aforementioned, we are using a Brownian framework for reasons of analytical and numerical tractability. However, this is far from sufficient to capture the full complexity of the real world. In this sense, the choice of constant parameters in (2) can also be criticized. Together with (12), they can be interpreted as regression parameters or averages with respect to the observed market data. Of course, it would be more realistic to choose time-dependent parameters to counteract some of these averaging effects. Nonetheless, even after adding time dependence, the model would not be able to perfectly describe the real world, but it would certainly improve it, and this is left to future research.

Appendix A. Results on the CIR-Model

The following results are taken from Di Francesco and Kamm (2021).

Theorem A1. *Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{Q}\right)$ be a stochastic basis, where $\mathbb{Q}$ is a martingale measure, $T > 0$ a finite time horizon and let the σ-algebra $(\mathcal{F}_t)_{t \in [0,T]}$ fulfill the usual conditions and support two independent standard Brownian motions W_x and W_y.*

The price of a zero-coupon bond in the model $r(t) = x(t) - y(t)$ with x and y being two independent CIR processes as in (1) is given by

$$P^{CIR\text{-}}(t, T) = A_x(t, T) e^{-B_x(t,T)x(t)} A_y(t, T) e^{B_y(t,T)y(t)}, \tag{A1}$$

where $t \leq T$ and for $z \in \{x, y\}$

$$
\begin{aligned}
A_z(t, T) &= \left(\frac{\phi_1^z e^{\phi_2^z(T-t)}}{\phi_2^z \left(e^{\phi_1^z(T-t)} - 1 \right) + \phi_1^z} \right)^{\phi_3^z} \\
B_z(t, T) &= \frac{e^{\phi_1^z(T-t)} - 1}{\phi_2^z \left(e^{\phi_1^z(T-t)} - 1 \right) + \phi_1^z}
\end{aligned}
\tag{A2}
$$

with $\phi_i^z \geq 0$, $i = 1, 2, 3$, $z \in \{x, y\}$, such that the Feller condition $2k_z\theta_z \geq \sigma_z^2$ is satisfied and

$$\begin{aligned}
\phi_1^x &= \sqrt{k_x^2 + 2\sigma_x^2}, & \phi_2^x &= \frac{k_x + \phi_1^x}{2}, & \phi_3^x &= \frac{2k_x\theta_x}{\sigma_x^2} \\
\phi_1^y &= \sqrt{k_y^2 - 2\sigma_y^2}, & \phi_2^y &= \frac{k_y + \phi_1^y}{2}, & \phi_3^y &= \frac{2k_y\theta_y}{\sigma_y^2}.
\end{aligned} \tag{A3}$$

Lemma A1. *Let everything be as in Theorem A1 but let $x(t)$ and $y(t)$ follow the general affine dynamics*

$$\begin{cases} dx(t) = (\lambda_x(t)x(t) + \eta_x(t))dt + \sqrt{\gamma_x(t)x(t) + \delta_x(t)}dW_x(t) \\ x(0) = x_0, \end{cases} \tag{A4}$$

$$\begin{cases} dy(t) = (\lambda_y(t)y(t) + \eta_y(t))dt + \sqrt{\gamma_y(t)y(t) + \delta_y(t)}dW_y(t) \\ y(0) = y_0, \end{cases} \tag{A5}$$

The initial values x_0, $y_0 \in \mathbb{R}$ are real-valued constants, and the coefficients $\lambda_z, \eta_z, \gamma_z, \delta_z, z \in \{x, y\}$, are all real-valued deterministic functions, such that (A4) and (A5) are well-defined.

Then, the price of a Zero-coupon bond is given by

$$P(t, T) = E_t^{\mathbb{Q}}\left[e^{-\int_t^T r(s)ds}\right] = A_x(t, T)e^{-B_x(t,T)x(t)}A_y(t, T)e^{B_y(t,T)y(t)}, \tag{A6}$$

where A_z and B_z, $z \in \{x, y\}$ are deterministic functions and are a classical solution to the following system of Riccati equations

$$\begin{cases} -1 - B_x(t, T)\lambda_x(t) - (\partial_t B_x)(t, T) + \dfrac{1}{2}B_x^2(t, T)\gamma_x(t) = 0, & B_x(T, T) = 0 \\[2mm] -B_x(t, T)\eta_x(t) + \dfrac{1}{2}B_x^2(t, T)\delta_x(t) + \partial_t(\log A_x)(t, T) = 0, & A_x(T, T) = 1 \\[2mm] 1 + B_y(t, T)\lambda_y(t) + (\partial_t B_y)(t, T) + \dfrac{1}{2}B_y^2(t, T)\gamma_y(t) = 0, & B_y(T, T) = 0 \\[2mm] B_y(t, T)\eta_y(t) + \dfrac{1}{2}B_y^2(t, T)\delta_y(t) + \partial_t(\log A_y)(t, T) = 0, & A_y(T, T) = 1. \end{cases} \tag{A7}$$

The Riccati equations for the CIR processes are given by defining $\lambda_z(t) \equiv -k_z, \eta_z(t) \equiv k_z\theta_z, \gamma_z(t) \equiv \sigma_z^2, \delta_z(t) \equiv 0$.

Appendix B. Instantaneous Forward Rate

The definition of the instantaneous forward rate (cf. Brigo and Mercurio 2006, p. 13, Equation (1.23)) is given by

$$f(t, T) := -\partial_T \log(P(t, T)).$$

By (A1), we therefore have

$$\begin{aligned}
f(t, T) &= -\partial_T\left(\log\left(A_x(t, T)e^{-B_x(t,T)x(t)}A_y(t, T)e^{B_y(t,T)y(t)}\right)\right) \\
&= -\partial_T(\log(A_x(t, T)) - B_x(t, T)x(t)) - \partial_T(\log(A_y(t, T)) + B_y(t, T)y(t)) \\
&= -\frac{\partial_T(A_x(t, T))}{A_x(t, T)} + \partial_T(B_x(t, T))x(t) - \frac{\partial_T(A_y(t, T))}{A_y(t, T)} - \partial_T(B_y(t, T))y(t).
\end{aligned}$$

Let $z \in \{x, y\}$ and consider the case of the CIR model (1). Then, those derivatives are given by the following expressions: Let us calculate the derivative of A_z first

$$\partial_T(A_z(t,T))$$

$$= \phi_z^3 \left(\frac{\phi_z^1 \phi_z^2 e^{\phi_z^2(T-t)}}{\phi_z^1 + \phi_z^2 \left(e^{\phi_z^1(T-t)} - 1 \right)} - \frac{\left(\phi_z^1 \right)^2 \phi_z^2 e^{\phi_z^1(T-t)} e^{\phi_z^2(T-t)}}{\left(\phi_z^1 + \phi_z^2 \left(e^{\phi_z^1(T-t)} - 1 \right) \right)^2} \right) \left(\frac{\phi_z^1 e^{\phi_z^2(T-t)}}{\phi_z^1 + \phi_z^2 \left(e^{\phi_z^1(T-t)} - 1 \right)} \right)^{\phi_z^3 - 1}.$$

Hence, we obtain

$$-\frac{\partial_T(A_z(t,T))}{A_z(t,T)} = \frac{\phi_z^2 \phi_z^3 \left(\phi_z^1 - \phi_z^2 \right) \left(e^{(T-t)\phi_z^1} - 1 \right)}{\phi_z^1 + \phi_z^2 \left(e^{(T-t)\phi_z^1} - 1 \right)}.$$

Now, we compute the derivative of B_z

$$\partial_T(B_z(t,T)) = \frac{\left(\phi_z^1 \right)^2 e^{(T-t)\phi_z^1}}{\left(\phi_z^1 + \phi_z^2 \left(e^{(T-t)\phi_z^1} - 1 \right) \right)^2}.$$

Appendix C. Hermite Polynomials

In this short section, we briefly recall the probabilist's Hermite polynomials, which are key to the Gram–Charlier expansion.

Definition A1. *The (probabilist's) Hermite polynomials $H_n(x)$ are defined as $H_0(x) \equiv 1$ and for $n \geq 1$*

$$(-1)^n (\varphi(x))^{-1} \left(\frac{d^n}{dx^n} \varphi \right)(x),$$

where $\varphi(x) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} \right)$.
Notice, that they are orthogonal with respect to the Gaussian measure, i.e.,

$$\int_{\mathbb{R}} H_m(x) H_n(x) \varphi(x) dx = \delta_{nm} n!.$$

In particular,

$$H_1(x) = x, H_2(x) = x^2 - 1, H_3(x) = x^3 - 3x, H_4(x) = x^4 - 6x^2 + 3,$$
$$H_5(x) = x^5 - 10x^3 + 15x, H_6(x) = x^6 - 15x^4 + 45x^2 - 15, H_7(x) = x^7 - 21x^5 + 105x^3 - 105x.$$

Appendix D. Cumulants and Moments

Let us denote by μ_i the moments and by c_i the cumulants. Their relationship towards each other is determined by the moment generating function (cf. Smith 1995) as follows

$$M(t) = 1 + \sum_{i=1}^{\infty} \mu_i \frac{t^i}{i!} = \exp\left(\sum_{i=1}^{\infty} c_i \frac{t^i}{i!} \right) = \exp(K(t)).$$

Therefore, assuming that the moments μ_i are known, we can compute the cumulants c_i by differentiating the formula from above

$$c_i = \frac{d^i}{dt^i} \log(M(t)) \bigg|_{t=0}.$$

Since we only need a few of them, we can compute the formulas and implement them directly, leading to

$$c_1 = \mu_1, \quad c_2 = \mu_2 - \mu_1^2, \quad c_3 = 2\mu_1^3 - 3\mu_2\mu_1 + \mu_3, \quad c_4 = -6\mu_1^4 + 12\mu_2\mu_1^2 - 4\mu_3\mu_1 - 3\mu_2^2 + \mu_4,$$

$$c_5 = 24\mu_1^5 - 60\mu_2\mu_1^3 + 20\mu_3\mu_1^2 + 30\mu_2^2\mu_1 - 5\mu_4\mu_1 - 10\mu_2\mu_3 + \mu_5$$

$$c_6 = -120\mu_1^6 + 360\mu_2\mu_1^4 - 120\mu_3\mu_1^3 - 270\mu_2^2\mu_1^2 + 30\mu_4\mu_1^2 + 120\mu_2\mu_3\mu_1 - 6\mu_5\mu_1 + 30\mu_2^3 - 10\mu_3^2$$
$$\quad - 15\mu_2\mu_4 + \mu_6$$

$$c_7 = 720\mu_1^7 - 2520\mu_2\mu_1^5 + 840\mu_3\mu_1^4 + 2520\mu_2^2\mu_1^3 - 210\mu_4\mu_1^3 - 1260\mu_2\mu_3\mu_1^2 + 42\mu_5\mu_1^2 - 630\mu_2^3\mu_1$$
$$\quad + 140\mu_3^2\mu_1 + 210\mu_2\mu_4\mu_1 - 7\mu_6\mu_1 + 210\mu_2^2\mu_3 - 35\mu_3\mu_4 - 21\mu_2\mu_5 + \mu_7.$$

Appendix E. Least Square Monte Carlo Method (LSMC)

In this section, we will demonstrate how to approximate the conditional expectation via LSMC. Let us first of all recall the following facts about the conditional expectation (cf. Pascucci 2011, pp. 654 ff.):

Let $X \in L^2(\Omega, \mathcal{F}, \mathbb{Q})$ and $\mathcal{A} \subseteq \mathcal{F}$ be a sub-σ-algebra.

1. Then, the conditional expectation is the L^2-best approximation, i.e.

$$\mathbb{E}^{\mathbb{Q}}\left[\left(X - \mathbb{E}^{\mathbb{Q}}[X|\mathcal{A}]\right)^2\right] \leq \mathbb{E}^{\mathbb{Q}}\left[(X - Y)^2\right]$$

for all $Y \in L^2(\Omega, \mathcal{A}, \mathbb{Q})$.

2. Furthermore, the factorization Lemma tells us that there exists a function u, such that

$$\mathbb{E}^{\mathbb{Q}}[Y|R] = u(R)$$

and combined with the argument above

$$u(R) = \arg\min_{v(\cdot)} \mathbb{E}^{\mathbb{Q}}\left[|v(R) - Y|^2\right]$$

where $v(\cdot)$ runs over all measurable functions.

The idea is now to approximate the function $u(x)$. Therefore, fix a basis $(b_i(x))_{i=1,\dots,n}$ and set $b^n(\cdot) := [b_1(\cdot), \dots, b_n(\cdot)]$. Then, we approximate u by $u(x) \approx \lambda^T b^n(x)$ where λ solves the least square problem

$$\lambda = \arg\min_{\alpha \in \mathbb{R}^n} \mathbb{E}^{\mathbb{Q}}\left[\left|\alpha^T b^n(R) - Y\right|^2\right].$$

The problem we encounter is that in this least square problem, we have random variables. Therefore, we can numerically deal with this problem by simulating those random variables, if it is possible, and view this least square problem as finding a linear regression for data points introduced by the realizations of the random variables.

Thus, let y_i be realizations of Y and set $y = [y_1, \dots, y_m]^T$. Additionally, let $b_{ij} = b_i(r_j)$, where r_j is a realization of R, and define the matrix $b = [b_{ij}]_{i=1,\dots,n;j=1,\dots,m}$.

Then, the above least square problem reads

$$\lambda = \arg\min_{\alpha \in \mathbb{R}^n} |b\alpha - y|^2.$$

This is known as the ordinary least square problem, and the optimal solution is given by

$$\lambda = \left(b^T b\right)^{-1} b^T y.$$

This tells us how to approximate the conditional expectation via a Monte Carlo linear regression approach.

Appendix F. Market Data

Table A1. Market data containing the volatility surface for the swaption pricing on 30 December 2019 in bps.

Maturity \ Tenor	1	2	3	4	5	7	10
1	17.5	21.8	26.8	31.4	35.2	40.2	45.6
2	25.4	29.3	33.5	36.4	39.5	43.5	47.5
3	34	36.7	39.2	41.1	43.2	46.2	49.3
4	40	41.5	43.4	44.8	46.2	48.4	50.9
5	43.7	44.6	45.8	47	48.4	50.1	52.3
7	49.7	49.8	50.5	51.4	52.1	53.1	54.4
10	54.6	54.4	54.7	54.9	55.1	55.2	55.6
15	54.8	54.4	54.5	54.4	54.2	54.2	54.4
20	53.6	53.2	53.4	53	52.9	52.8	52.5

Table A2. Market data containing the swaption strikes on 30 December 2019.

Maturity \ Tenor	1	2	5	7	10
1	−0.260793%	−0.195187%	−0.011405%	0.140129%	0.330514%
2	−0.129665%	−0.0782444%	0.139932%	0.273273%	0.449172%
5	0.268095%	0.38307%	0.556996%	0.655339%	0.757978%
7	0.547079%	0.611571%	0.76683%	0.830788%	0.891069%
10	0.880582%	0.907944%	0.967521%	0.988131%	0.992003%
15	1.04232%	1.04153%	1.01776%	0.985317%	0.924744%
20	0.925377%	0.901441%	0.827386%	0.778437%	0.721445%

Table A3. Market data containing the swaption prices on 30 December 2019.

Maturity \ Tenor	1	2	5	7	10
1	0.000702236	0.00175071	0.00706456	0.0112631	0.0181169
2	0.0014433	0.00333027	0.0111956	0.017189	0.0265694
5	0.00391314	0.00796766	0.0214221	0.0308074	0.0453508
7	0.00521117	0.0104082	0.0268942	0.0380283	0.0548627
10	0.00668368	0.0132567	0.0330802	0.045932	0.0651091
15	0.00781681	0.0154396	0.0378811	0.0525334	0.0743464
20	0.00840243	0.0166069	0.0407885	0.0565876	0.0795953

Table A4. Bloomberg's Hull–White one factor prices of receiver Bermudan swaptions on 30 December 2019.

Maturity \ Tenor	2	5	7	10
1	0.21%	1.06%	1.85%	3.28%
3	0.57%	1.83%	2.86%	4.63%
5	0.87%	2.48%	3.71%	5.72%
7	1.11%	3.03%	4.43%	6.65%
10	1.4%	3.62%	5.2%	7.59%

Table A5. Bloomberg's Hull–White one factor prices of payer Bermudan swaptions on 30 December 2019.

Maturity \ Tenor	2	5	7	10
1	0.25%	1.4%	2.55%	4.76%
3	0.6%	2.08%	3.42%	5.74%
5	0.9%	2.7%	4.16%	6.59%
7	1.13%	3.2%	4.75%	7.18%
10	1.41%	3.72%	5.33%	7.91%

Table A6. Market data containing the Bermudan swaption strikes on 30 December 2019.

Maturity \ Tenor	2	5	7	10
1	−0.194%	0.00912%	0.14%	0.33%
3	0.0789%	0.274%	0.432%	0.561%
5	0.335%	0.534%	0.644%	0.767%
7	0.612%	0.771%	0.84%	0.894%
10	0.926%	1.01%	0.994%	1.02%

Table A7. Market data containing the zero rate curve and zero coupon curve on 30 December 2019.

Maturity (in Years)	Zero Rate (in %)	Zero-Coupon Price
0.0833333333333333	−0.469999993219972	1.0004001991529
0.25	−0.388000020757318	1.00096969387991
0.5	−0.324999983422458	1.00163343819125
0.75	−0.314333918504417	1.00237481461989
1	−0.322000007145107	1.00323926670136
1.25	−0.323286440253412	1.00405360258242
1.5	−0.316161320131414	1.00476558980205
1.75	−0.303842297803669	1.00535001652119
2	−0.289547047577798	1.00582418019158
2.25	−0.275860329135469	1.00623288634409
2.5	−0.262835313503729	1.006604855007
2.75	−0.249892233800608	1.00691299093433
3	−0.236451346427202	1.00713375064174
3.25	−0.222084053437044	1.00725039326453
3.5	−0.20696636298112	1.00728054250496
3.75	−0.191425434683623	1.00721781901104
4	−0.175788428168744	1.00706740209126
4.25	−0.160311330630236	1.00684531811395
4.5	−0.144965462482105	1.00655553463348
4.75	−0.129650957156002	1.00618948972951
5	−0.114267959725112	1.00573933685071
5.25	−0.0987154224631581	1.00520062530541
5.5	−0.0828875612342017	1.00457454544122
5.75	−0.0666773874613114	1.00384671986489
6	−0.0499779242090881	1.00300667524933
6.25	−0.0327643402378897	1.00205088034181
6.5	−0.0153403983915723	1.00099833086134
6.75	0.00190798987986796	0.999871102605028
7	0.0185949131264351	0.998698306220564
7.25	0.0344518735623467	0.997505079039002

Table A7. *Cont.*

Maturity (in Years)	Zero Rate (in %)	Zero-Coupon Price
7.5	0.0496800311054812	0.996279818846146
7.75	0.0645979575189415	0.995003816465917
8	0.0795242260210216	0.993656440330286
8.25	0.0947347900819295	0.992214008696662
8.5	0.110335148849572	0.990662992494919
8.75	0.126388167535652	0.988997743889118
9	0.142956722993404	0.987213788328959
9.25	0.160050573928316	0.985308478446392
9.5	0.177466994199449	0.983284710270437
9.75	0.194950156980411	0.981173005874126
10	0.212244223803282	0.979004189945635
15	0.473523046821356	0.931543316237289
20	0.611338950693607	0.885166902653398
25	0.652327481657267	0.849865688031976
30	0.640345783904195	0.825611308910539

References

Björk, Tomas. 2004. *Arbitrage Theory in Continuous Time*. Oxford Finance Series. Incorporated: Oxford University Press.

Brigo, Damiano, and Fabio Mercurio. 2001. A deterministic–shift extension of analytically–tractable and time–homogeneous short–rate models. *Finance and Stochastics* 5: 369–87. [CrossRef]

Brigo, Damiano, and Fabio Mercurio. 2006. *Interest Rate Models: Theory and Practice: With Smile, Inflation and Credit*. Berlin: Springer.

Chen, Lin. 1996. A Three-Factor Model of the Term Structure of Interest Rates. In *Interest Rate Dynamics, Derivatives Pricing, and Risk Management*. Berlin: Springer. https://doi.org/10.1007/978-3-642-46825-4_1.

Cheng, Yin Hei. 2013. *Pricing Derivatives by Gram–Charlier Expansions*. Waterloo: UWSpace.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross. 1985. A theory of the term structure of interest rates. *Econometrica* 53: 385–407. [CrossRef]

Curtis, Vitor Venceslau, and Carlos Alberto Alonso Sanches. 2017. A low-space algorithm for the subset-sum problem on gpu. *Computers & Operations Research* 83: 120–24. [CrossRef]

Dereich, Steffen, Andreas Neuenkirch, and Lukasz Szpruch. 2012. An euler-type method for the strong approximation of the cox-ingersoll-ross process. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 468: 1105–15. [CrossRef]

Di Francesco, Marco. 2012. A general gaussian interest rate model consistent with the current term structure. *ISRN Probability and Statistics* 2012: 673607. [CrossRef]

Di Francesco, Marco, and Kevin Kamm. 2021. How to handle negative interest rates in a cir framework. *SeMA Journal* 1–26. [CrossRef]

Gatarek, Dariusz, and Juliusz Jabłecki. 2021. Between scylla and charybdis: The bermudan swaptions pricing odyssey. *Mathematics* 9: 112. [CrossRef]

Gilli, Manfred, Dietmar Maringer, and Enrico Schumann. 2011. *Numerical Methods and Optimization in Finance*, 1st ed. Amsterdam: Elsevier.

Glasserman, Paul. 2004. *Monte Carlo Methods in Financial Engineering*. New York: Springer.

Grabinski, Michael, and Galiya Klinkova. 2019. Wrong use of averages implies wrong results from many heuristic models. *Applied Mathematics* 10: 605–18. [CrossRef]

Ho, Thomas S. Y., and Sang-Bin Lee. 1986. Term structure movements and pricing interest rate contingent claims. *The Journal of Finance* 41: 1011–29. [CrossRef]

Hull, John C. 2006. *Options, Futures, and Other Derivatives*, 8th ed. Pearson Internat. ed. Hoboken: Pearson Prentice Hall.

Hull, John, and Alan White. 1990. Pricing interest-rate-derivative securities. *Review of Financial Studies* 3: 573–92. [CrossRef]

Longstaff, Francis A., and Eduardo S. Schwartz. 2001. Valuing american options by simulation: A simple least-squares approach. *The Finance* 14: 113–47. [CrossRef]

Mercurio, Fabio, and Andrea Pallavicini. 2005. Mixing gaussian models to price cms derivatives. *SSRN Electronic Journal*. [CrossRef]

Miron, Paul, and Philip Swannell. 1991. *Pricing and Hedging Swaps*. London: Euromoney Books.

Oosterlee, Cornelis W., and Lech A. Grzelak. 2019. *Mathematical Modeling and Computation In Finance: With Exercises and Python and Matlab Computer Codes*. London: World Scientific Publishing Europe Ltd.

Orlando, Giuseppe, Rosa Maria Mininni, and Michele Bufalo. 2019a. A new approach to forecast market interest rates through the CIR model. *Studies in Economics and Finance* 37: 267–92. [CrossRef]

Orlando, Giuseppe, Rosa Maria Mininni, and Michele Bufalo. 2019b. Interest rates calibration with a cir model. *Journal of Risk Finance* 20: 370–87. [CrossRef]

Orlando, Giuseppe, Rosa Maria Mininni, and Michele Bufalo. 2020. Forecasting interest rates through vasicek and cir models: A partitioning approach. *Journal of Forecasting* 39: 569–79. [CrossRef]

Pascucci, Andrea. 2011. *PDE and Martingale Methods in Option Pricing, Volume 2 of Bocconi & Springer Series.* Milan: Springer/Bocconi University Press. [CrossRef]

Russo, Vincenzo, and Gabriele Torri. 2019. Calibration of one-factor and two-factor hull–white models using swaptions. *Computational Management Science* 16: 275–95. [CrossRef]

Schrager, David F., and Antoon A. J. Pelsser. 2006. Pricing swaptions and coupon bond options in affine term structure models. *Capital Markets: Asset Pricing & Valuation eJournal* 16: 673–94.

Smith, Peter J. 1995. A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *The American Statistician* 49: 217–18.

Tanaka, Keiichi, Takeshi Yamada, and Toshiaki Watanabe. 2010. Applications of gram–charlier expansion and bond moments for pricing of interest rates and credit risk. *Quantitative Finance* 10: 645–62. [CrossRef]

Vasicek, Oldrich. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5: 177–88. [CrossRef]