A TRACTABLE TERM STRUCTURE MODEL WITH ENDOGENOUS INTERPOLATION AND POSITIVE INTEREST RATES

ERIK SCHLÖGL AND LUTZ SCHLÖGL

ABSTRACT. This paper presents the one– and the multifactor versions of a term structure model in which the factor dynamics are given by Cox/Ingersoll/Ross (CIR) type "square root" diffusions with piecewise constant parameters. The model is fitted to initial term structures given by a finite number of data points, interpolating endogenously. Closed form and near–closed form solutions for a large class of fixed income contingent claims are derived in terms of a noncentral chi-square distribution whose noncentrality parameter is in turn noncentral chi-square distributed. Implementation details on this distribution are given in the appendix.

1. INTRODUCTION

Three often cited requirements for term structure models applied in practice are

(i) fit to the initial term structure observed in the market
(ii) analytical tractability for fast solutions for derivative pricing and hedging
(iii) non-negative interest rates.

Requirement (i) is often extended to fitting an initial term structure of interest rate volatility.

The most tractable class, the Gauss–Markov models, are precluded by requirement (iii). Taking this as given, we are faced with a tradeoff between requirements (i) and (ii). Of the factor models satisfying (iii), the Cox, Ingersoll and Ross (1985) (CIR) model is arguably the most tractable. Unfortunately, if one extends the original CIR model as in Hull and White (1990) and allows for a time dependent drift in order to calibrate to an observed initial term structure, the closed form solutions do not carry over from the constant parameter case, so we are faced with a tradeoff between requirements (i) and (ii). However, looking at a market for fixed income instruments, we observe interest rates or bond prices for only a finite number of maturities. If one considers the money market and/or swap market, this number is quite small. This presents a possibility to avoid the tradeoff between fitting initial term structure data and fast analytical solutions: The observed data points divide the time line into segments. On these segments we inductively construct short rate processes of the CIR type with constant parameters, chosen so as to give an exact fit of the observed term structure. The processes are pieced together to yield a continuous short rate process.

In taking this approach, we need to consider the CIR stochastic differential equation with nondeterministic initial conditions and show that the solution does not explode under the

---

Date. November 1997.

Financial support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged. Part of this work was completed while the first author was a guest of the Centre for Analytical Finance (CAF) at Aarhus University, Denmark. We would like to thank Jørgen Aase Nielsen, Daniel Sommer, Dieter Sondermann and Asbjørn Trolle Hansen for helpful discussions. All remaining errors are ours.

same restrictions on the parameters as in the deterministic case; thus interest rates are almost surely strictly positive. This result is derived in appendix A.

The inductive construction of the short rate process allows us to fit the model to an initial term structure consisting of a finite number of data points. By taking the relevant expectations, the model yields a complete initial term structure for the continuum of maturities, endogenously interpolating between the observed data. Thus it is more parsimonious in its assumptions in the sense that there is no need for an exogenous interpolation rule and the interpolation is consistent with the assumed short rate dynamics.

Herein also lies the difference to another possible resolution of the aforementioned tradeoff. In what he calls a "simple class of square root models", Jamshidian (1995) restricts the CIR model with time–varying coefficients to a class satisfying the condition that the ratio of the mean reversion level $\theta(t)$ and the square of the volatility parameter $\sigma^2(t)$ is identical for all t. The time–dependence of the parameters allows initial term structures of interest rates and of interest rate volatilities to be fitted. However, the way initial forward rates are interpolated determines how volatilities evolve over time. Our approach could thus be seen as an alternative to the "simple class", the choice between the two depending on what should be endogenous to the model.

Scott (1995) also addresses the tradeoff between analytical tractability and calibration to market data in CIR type models. He shifts the short rate realizations of a constant parameter model by a deterministic component in order to fit the initial term structure, thus avoiding the need to perform calculations with a time dependent drift parameter. The yield curve interpolation is exogenous to this model and the short rate volatility parameter is constant across time, precluding calibration to market volatilities.

In the next section we introduce the model and show how to fit it to initial zero coupon bond prices, as well as discussing how volatility structures can be input into the model, this being an additional requirement often put forth by practitioners in addition to the three already mentioned. The formulae for contingent claims pricing and hedging are derived in section 3, interpolated zero coupon bond prices being given as a special case. Section 4 shows how the results in the one-factor case can be extended to a multifactor model.

2. The Model

2.1. **The Short Rate Process.** We wish to specify the model so that the dynamics of the short rate process are given by a generalized CIR equation with piecewise constant coefficients. We call this the *segmented square root model.* To formalize, let points in time $0 = T_0 < T_1 < \ldots < T_N$ and constants $\theta_1, \ldots, \theta_N$, $a_1, \ldots, a_N$, $\sigma_1, \ldots, \sigma_N \in \mathbb{R}_{++}$ be given. We define a step function $\theta : \mathbb{R}_+ \to \mathbb{R}_{++}$ in the following manner:

$$\theta(t) := \theta_1 \; 1_{\{0\}}(t) + \sum_{i=1}^{N} \theta_i \; 1_{]T_{i-1},T_i]}(t) + \theta_N \; 1_{]T_N;+\infty[}(t)$$

Step functions $\sigma, a : \mathbb{R}_+ \to \mathbb{R}_{++}$ are defined analogously. Now let $(\Omega, \mathcal{F}, \mathbb{F}, Q)$ be a stochastic basis satisfying the usual hypotheses on which a standard one-dimensional Brownian motion W is defined. Note that since we are interested in arbitrage pricing of derivative securities, we are constructing the model immediately under the risk neutral measure Q.[1] The short rate r is a continuous $\mathbb{F}$-adapted stochastic process defined on Ω. We specify

[1] I.e. the equivalent measure under which all asset prices discounted by the savings account $\exp\{\int r(s)ds\}$ are martingales.

the dynamics of r by demanding that r is a solution of the following stochastic differential equation:

$$dr(t) = (\theta(t) - a(t)r(t))dt + \sigma(t)\sqrt{r(t)}dW(t) \tag{1}$$

We shorten notation by defining two functions $\alpha, \beta : \mathbb{R}_+ \times]0;+\infty[\to \mathbb{R}$ as follows:

$$\begin{aligned}
\alpha(t,x) &:= \sigma_1\sqrt{x}\, 1_{\{0\}}(t) + \sum_{i=1}^{N} \sigma_i\sqrt{x}\, 1_{]T_{i-1};T_i]}(t) + \sigma_N\sqrt{x}\, 1_{]T_N;+\infty[}(t) \\
\beta(t,x) &:= (\theta_1 - a_1x)1_{\{0\}}(t) + \sum_{i=1}^{N}(\theta_i - a_ix)1_{]T_{i-1};T_i]}(t) + (\theta_N - a_Nx)1_{]T_N;+\infty[}(t)
\end{aligned}$$

The dynamic equation now takes the form:

$$dr(t) = \beta(t, r(t))dt + \alpha(t, r(t))dW(t) \tag{2}$$

We stress the fact that a solution of this equation by definition only assumes values in $]0;+\infty[$, so that the short rate is automatically strictly positive at all times. We will show that solving equation (2) is equivalent to iteratively solving classical Cox/Ingersoll/Ross equations, i.e. equations of the type:

$$dr(t) = (\theta - ar(t))dt + \sigma\sqrt{r(t)}dW(t). \tag{3}$$

Here θ, a, σ are strictly positive constants. If these fulfill the inequality $2\theta \geq \sigma^2$, then using Feller's test for explosions (cf. Karatzas and Shreve (1988), Proposition 5.5.22 and Theorem 5.5.29) one can prove that solutions of (3) with nonrandom initial conditions cannot explode. Using the result shown in appendix A, it follows that solutions of (3) with random initial conditions do not explode either.

2.1.1. THEOREM. *For each $i \in \{1, \ldots, N\}$ let the constants $\theta_i, \sigma_i \in]0;+\infty[$ fulfill the inequality $2\theta_i \geq \sigma_i^2$. Let $(\Omega, \mathcal{F}, \mathbb{F}, Q)$ be a stochastic basis fulfilling the usual hypotheses, carrying a standard, one–dimensional Brownian motion W. Let $f : \Omega \to]0;+\infty[$ be an $\mathcal{F}_0$–measurable random variable. Then there exists a continuous, $\mathbb{F}$–adapted process r with values in $]0;+\infty[$, so that for each $t \in \mathbb{R}_+$, we have*

$$r(t) = f + \int_0^t \alpha(s, r(s))dW(s) + \int_0^t \beta(s, r(s))ds \quad Q\text{–a.s.}$$

PROOF: See appendix B.

2.1.2. REMARK. The coefficients of the classical CIR equation are locally Lipschitz, therefore pathwise uniqueness holds for the equation (3). By an iteration procedure analogous to the one used in the proof of the theorem above, it follows that pathwise uniqueness also holds for the equation (2). Therefore, in our model the short rate is uniquely determined up to indistinguishability by the coefficients in the dynamic equation and the initial interest rate.

2.2. **Fitting the Model to Zero Coupon Bond Prices.** In the segmented model, as in the original constant parameter CIR model, zero coupon bond prices can be expressed in closed form as functions of the state variable and the model parameters. Since we are constructing the model in such a manner as to have one segment for each zero coupon bond maturity, the drift parameters θ_j can be chosen in such a manner as to fit the bond prices.

On each constant parameter segment, the bond price formula is as in the classical CIR case. To get the price of a bond whose time to maturity covers two segments $0 = T_0 < T_1 < T_2$, we apply the CIR formula at the segment boundary T_1 and take the expectation under the T_1 forward measure[2]. Iterating this procedure yields

2.2.1. PROPOSITION. *In the segmented square root model with time segments $[T_{j-1}; T_j]$, $j \in \{1; \dots ; N\}$, the time T_{j-1} prices of zero coupon bonds with maturity T_k, $j \leq k \leq N$, are exponential affine functions of the short rate realization $r(T_{j-1})$:*

$$B(r(T_{j-1}), T_{j-1}, T_k) = \mathcal{C}_{j-1,k} \exp\{-\mathcal{D}_{j-1,k}\, r(T_{j-1})\} \tag{4}$$

with $\mathcal{C}_{j-1,k}$ and $\mathcal{D}_{j-1,k}$ recursively defined as

$$\mathcal{C}_{j-1,k} := \mathcal{C}_{j,k} \cdot \mathcal{A}_j(T_{j-1}, T_j) \cdot \left(\frac{b_j}{b_j + 2\mathcal{D}_{j,k}} \right)^{\frac{1}{2}\nu_j}$$

$$\mathcal{D}_{j-1,k} := \mathcal{B}_j(T_{j-1}, T_j) + \eta_j \frac{\mathcal{D}_{j,k}}{b_j + 2\mathcal{D}_{j,k}}, \qquad \text{where } \mathcal{C}_{k,k} := 1 \text{ and } \mathcal{D}_{k,k} := 0,$$

and

$$\mathcal{A}_j(T_{j-1}, T_j) := \left(2c_j w_j(T_{j-1}, T_j) \exp\left\{ \frac{1}{2}(c_j + a_j)(T_j - T_{j-1}) \right\} \right)^{\frac{1}{2}\nu_j}$$

$$\mathcal{B}_j(T_{j-1}, T_j) := 2w_j(T_{j-1}, T_j) \left(\exp\{c_j(T_j - T_{j-1})\} - 1 \right)$$

$$w_j(T_{j-1}, T_j) := \left((c_j + a_j) \exp\{c_j(T_j - T_{j-1})\} + c_j - a_j \right)^{-1}$$

$$c_j := \sqrt{a_j^2 + 2\sigma_j^2}$$

$$\eta_j := \frac{16 w_j(T_{j-1}, T_j)^2 c_j^2 \exp\{c_j(T_j - T_{j-1})\}}{\sigma_j^2 \mathcal{B}_j(T_{j-1}, T_j)}$$

$$b_j := \frac{4}{\sigma_j^2} \mathcal{B}_j(T_{j-1}, T_j)^{-1}$$

$$\nu_j := \frac{4\theta_j}{\sigma_j^2}$$

PROOF: Let $k - j = 0$. Then (4) becomes

$$B(r(T_{j-1}), T_{j-1}, T_j) = \mathcal{A}_j(T_{j-1}, T_j) \exp\{-\mathcal{B}_j(T_{j-1}, T_j) r(T_{j-1})\}$$

which is the original CIR formula for the price of a zero coupon bond, valid since we have a constant parameter CIR process on $[T_{j-1}, T_j]$. Now consider some $k > j$. If (4) is valid over $(n-1)$ segments then (4) determines all zero coupon bond prices

$$B(r(T_j), T_j, T_k) \qquad \forall\, k - j < n.$$

If there are n segments there remains one bond price not determined by (4). Since under the time T_1 forward measure all assets valued with respect to the zero coupon bond maturing in T_1 are martingales, we can write this remaining bond price as

$$B(r(T_0), T_0, T_n) \;=\; B(r(T_0), T_0, T_1) E^{T_1} \left[\frac{B(r(T_1), T_1, T_n)}{B(r(T_1), T_1, T_1)} \right], \tag{5}$$

[2] I.e. the equivalent measure under which all asset prices discounted by the zero coupon bond maturing in T_1 are martingales.

where the bond price in T_1 is given by (4) with a suitable shift of indices. Setting $\tilde{r}(T_1) := b_1 r(T_1)$ we can then employ the result in Jamshidian (1987) that $\tilde{r}(T_1)$ conditioned on $r(T_0)$ is noncentral chi–square distributed under the T_1 forward measure, with ν_1 degrees of freedom and non–centrality parameter $\eta_1 r(T_0)$. Thus (5) becomes

$$\begin{aligned} B(r(T_0),T_0,T_n) &= B(r(T_0),T_0,T_1)\cdot \\ &\int_0^\infty \mathcal{C}_{1,n}\exp\{-\mathcal{D}_{1,n}\,r(T_1)\}q_{\chi^2}((b_1 r(T_1),\nu_1,\eta_1 r(T_0))d(b_1 r(T_1)). \end{aligned} \tag{6}$$

Applying lemma C.1 in appendix C to (6) we get

$$\begin{aligned} B(r(T_0),T_0,T_n) &= B(r(T_0),T_0,T_1)\mathcal{C}_{1,n}\exp\left\{-\frac{\mathcal{D}_{1,n}}{b_1+2\mathcal{D}_{1,n}}\eta_1 r(T_0)\right\}\left(\frac{b_1}{b_1+2\mathcal{D}_{1,n}}\right)^{\frac{1}{2}\nu_1} \\ &\quad\cdot\underbrace{\int_0^\infty q_{\chi^2}\left((b_1+2\mathcal{D}_{1,n})r(T_1),\nu_1,\frac{b_1}{b_1+2\mathcal{D}_{1,n}}\eta_1 r(T_0)\right)d((b_1+2\mathcal{D}_{1,n})r(T_1))}_{=1} \\ &= \mathcal{A}_1(T_0,T_1)\mathcal{C}_{1,n}\left(\frac{b_1}{b_1+2\mathcal{D}_{1,n}}\right)^{\frac{1}{2}\nu_1}\exp\left\{-\left(\mathcal{B}(T_0,T_1)+\frac{\mathcal{D}_{1,n}}{b_1+2\mathcal{D}_{1,n}}\eta_1\right)r(T_0)\right\} \\ &= \mathcal{C}_{0,n}\exp\{-\mathcal{D}_{0,n}\,r(T_0)\}. \end{aligned}$$

□

Starting with $j = 1$ and solving (4) for θ_j, we can thus successively calculate all θ_j for $j \in \{1;\ldots;N\}$. Note, however, that strongly downward sloping initial forward rate curves can lead to negative θ_j, and thus this model shares the disadvantage all CIR type models with a time dependent drift coefficient in that it cannot fit all possible initial term structures.

2.3. **Initial Volatility Term Structures.** The problem of calibrating a model to observed volatility structures has two dimensions, of which the term structure of volatilities for forward rates or zero coupon bonds of different maturities is most often cited. Besides this maturity dimension, however, there is the temporal dimension of how volatilities evolve. Historical estimates of deterministic volatility coefficients usually assume that these coefficients do not change over time. When calibrating a model to implied volatilities, the first dimension is given by for example prices of options on zero coupon bonds of different maturities, and the temporal dimension of the volatility structure is determined by prices of options on zero coupon bonds with the same time to maturity, but different option expiries.

In the "simple class of square root models" of Jamshidian (1995), the maturity dimension is covered by choosing the (time–dependent) speed of mean reversion to match input volatilities for forward rates. Instantaneous forward rates are given by

$$f(r(t),t,T) = -\frac{\partial}{\partial T}\ln B(r(t),t,T)$$

By Itô's Lemma, initial forward rate volatilities in our model are therefore $\sigma_1\sqrt{r(T_0)}\frac{\partial}{\partial T_k}\mathcal{D}_{0,k}$ and we can state

2.3.1. PROPOSITION. *Initial forward rate volatilities in the segmented square root model are* $\sigma_1\sqrt{r(T_0)}\frac{\partial}{\partial T_k}\mathcal{D}_{0,k}$, *with*

$$\frac{\partial}{\partial T_k}\mathcal{D}_{0,k} = \left(\prod_{i=1}^{k-1}\eta_i b_i(b_i + 2\mathcal{D}_{i,k})^{-2}\right)\frac{\partial}{\partial T_k}\mathcal{B}_k(T_{k-1}, T_k) \tag{7}$$

and $\mathcal{D}$, η, b *and* $\mathcal{B}$ *defined as in proposition 2.2.1.*

Note that forward rate curves in our model are continuous (see proposition 3.2.1).

PROOF: By induction, we show the validity of the more general version of (7)

$$\frac{\partial}{\partial T_k}\mathcal{D}_{j,k} = \left(\prod_{i=j+1}^{k-1}\eta_i b_i(b_i + 2\mathcal{D}_{i,k})^{-2}\right)\frac{\partial}{\partial T_k}\mathcal{B}_k(T_{k-1}, T_k) \tag{8}$$

For $k - j = 1$, (8) becomes

$$\frac{\partial}{\partial T_k}\mathcal{D}_{j,k} = \frac{\partial}{\partial T_k}\mathcal{B}_k(T_{k-1}, T_k)$$

which is obviously true since $\mathcal{D}_{j,j+1} = \mathcal{B}_{j+1}(T_j, T_{j+1})$. Now let (8) be valid for some $k - j \geq 1$. Then for $k + 1$ (or by a simple change of notation $j - 1$) we have

$$\begin{aligned}
\frac{\partial}{\partial T_{k+1}}\mathcal{D}_{j,k+1} &= \frac{\partial}{\partial T_{k+1}}(\mathcal{B}_{j+1}(T_j, T_{j+1}) + \eta_{j+1}(b_{j+1}\mathcal{D}^{-1}_{j+1,k+1} + 2)^{-1}) \\
&= \eta_{j+1}(b_{j+1}\mathcal{D}^{-1}_{j+1,k+1} + 2)^{-2} b_{j+1}\mathcal{D}^{-2}_{j+1,k+1}\frac{\partial}{\partial T_{k+1}}\mathcal{D}_{j+1,k+1} \\
&= \eta_{j+1}b_{j+1}(b_{j+1} + 2\mathcal{D}_{j+1,k+1})^{-2}\left(\prod_{i=j+2}^{k}\eta_i b_i(b_i + 2\mathcal{D}_{i,k+1})^{-2}\right)\frac{\partial}{\partial T_{k+1}}\mathcal{B}_{k+1}(T_k, T_{k+1}) \\
&= \left(\prod_{i=j+1}^{k}\eta_i b_i(b_i + 2\mathcal{D}_{i,k+1})^{-2}\right)\frac{\partial}{\partial T_{k+1}}\mathcal{B}_{k+1}(T_k, T_{k+1})
\end{aligned}$$

□

For given σ_k, we can thus use (7) to inductively calculate the speed of mean reversion parameters a_k to match initial forward rate volatilities for the maturities T_1 to T_n. Alternatively, one could calculate the σ_k for given parameters a_k.

Calibrating the models to an initial volatility structure along the maturity dimension, we therefore still retain a degree of freedom that is already taken in the "simple class". Thus our model differs from the simple class in how the temporal dimension of the volatility structure is specified: In the simple class, the way initial forward rates are interpolated determines how volatilities evolve over time, while in the segmented model the factor volatility, be it the short rate or some yield of forward rate[3], determines the endogenous interpolation.

[3]Note that affine models can be reparameterized in any yield or forward rate instead of the short rate (see Duffie and Kan (1992, 1996)).

3. Pricing and Hedging

3.1. **Contingent Claim Valuation.** Having calibrated the model, we can now proceed to price other assets relative to the initial term structure, given the assumptions of the model. We will consider contingent claims whose payoffs can be expressed as linear combinations of European (exchange) options on securities whose terminal function is a simple exponential in r, thus allowing us to apply lemma C.1. As will be discussed below, this covers a wide range of fixed income derivatives.

We begin by deriving the exchange option formula in the classical CIR case. Consider the value $V_j(t)$ of a security at time t given by

$$V_j(t, r(t)) \;=\; h_j(t)\exp\{-g_j(t)r(t)\}, \tag{9}$$

where h_j and g_j are deterministic functions of t. The payoff $C(t_m)$ at expiry t_m of a European exchange option on two such securities is defined as

$$C\left(t_m, r(t_m)\right) \;:=\; \left[V_1\left(t_m, r(t_m)\right) - V_2\left(t_m, r(t_m)\right)\right]^+. \tag{10}$$

Note that for $h_2(t_m) = K$ and $g_2(t_m) = 0$ we have a European call option on V_1. Alternatively, setting $h_1(t_m) = K$ and $g_1(t_m) = 0$ yields a European put option on V_2. Let

$$k \;:=\; \max\{n \in \{0; \dots ; N\} \mid T_n < t_m\}. \tag{11}$$

On $[T_k; T_{k+1}]$ we have a constant parameter CIR process, and following Jamshidian (1987) we know that $\tilde{r}(t_m) := \tilde{b}\cdot r(t_m)$ conditioned on $r(T_k)$ is noncentral chi–square distributed under the t_m forward measure, with ν_{k+1} degrees of freedom and noncentrality parameter $\tilde{\eta}\cdot r(T_k)$, where

$$\begin{aligned}\tilde{b} &:= \frac{4}{\sigma_{k+1}^2}\mathcal{B}_{k+1}(T_k, t_m)^{-1}\\ \tilde{\eta} &:= \frac{16 w_{k+1}(T_k, t_m)^2\, c_{k+1}^2 \exp\{c_{k+1}(t_m - T_k)\}}{\sigma_{k+1}^2 \mathcal{B}_{k+1}(T_k, t_m)}\end{aligned}$$

with $\mathcal{B}$, c and w defined as in proposition 2.2.1. Therefore

$$\begin{aligned} C(T_k, r(T_k)) =& B(r(T_k), T_k, t_m) E^{t_m}\left[\left[V_1(t_m, r(t_m)) - V_2(t_m, r(t_m))\right]^+ | \mathcal{F}_{T_k}\right] \\ =& B(r(T_k), T_k, t_m)\left(\int_{\mathbb{Z}} h_1(t_m)\exp\{-g_1(t_m)r(t_m)\} q_{\chi^2}\left(\tilde{b}r(t_m), \nu_{k+1}, \tilde{\eta}r(T_k)\right) d(\tilde{b}r(t_m))\right.\\ &\left. - \int_{\mathbb{Z}} h_2(t_m)\exp\{-g_2(t_m)r(t_m)\} q_{\chi^2}\left(\tilde{b}r(t_m), \nu_{k+1}, \tilde{\eta}r(T_k)\right) d(\tilde{b}r(t_m))\right)\end{aligned} \tag{12}$$

with

$$\mathbb{Z} := \{r(t_m) > 0 \mid V_1(t_m, r(t_m)) > V_2(t_m, r(t_m))\}.$$

Given the functional form of V_1 and V_2, we have either $\mathbb{Z}=]0; r^*[$ or $\mathbb{Z}=]r^*; \infty[$ for some deterministic $r^* > 0$. We consider $\mathbb{Z}=]0; r^*[$; the calculations for $\mathbb{Z}=]r^*; \infty[$ are analogous. Applying lemma C.1, (12) becomes

$$
\begin{aligned}
&C(T_k, r(T_k))\\
=&B(r(T_k), T_k, t_m)\left(h_1(t_m)\exp\left\{-\frac{g_1(t_m)}{\tilde{b}+2g_1(t_m)}\,\tilde{\eta}\,r(T_k)\right\}\left(\frac{\tilde{b}}{\tilde{b}+2g_1(t_m)}\right)^{\frac{1}{2}\nu_{k+1}}\right.\\
&\qquad\int_0^{(\tilde{b}+2g_1(t_m))r^*} q_{\chi^2}\left((\tilde{b}+2g_1(t_m))r(t_m),\ \nu_{k+1},\ \frac{\tilde{b}}{\tilde{b}+2g_1(t_m)}\,\tilde{\eta}\,r(T_k)\right) d\left((\tilde{b}+2g_1(t_m))r(t_m)\right)\\
&\quad - h_2(t_m)\exp\left\{-\frac{g_2(t_m)}{\tilde{b}+2g_2(t_m)}\,\tilde{\eta}\,r(T_k)\right\}\left(\frac{\tilde{b}}{\tilde{b}+2g_2(t_m)}\right)^{\frac{1}{2}\nu_{k+1}}\\
&\qquad\left.\int_0^{(\tilde{b}+2g_2(t_m))r^*} q_{\chi^2}\left((\tilde{b}+2g_2(t_m))r(t_m),\ \nu_{k+1},\ \frac{\tilde{b}}{\tilde{b}+2g_2(t_m)}\,\tilde{\eta}\,r(T_k)\right) d\left((\tilde{b}+2g_2(t_m))r(t_m)\right)\right)
\end{aligned}
$$

which we can write as

$$
\begin{aligned}
(13)\qquad C(T_k, r(T_k)) = \quad &\hat{h}_1\exp\{-\hat{g}_1 r(T_k)\}\int_0^{\hat{b}_1 r^*} q_{\chi^2}\left(\hat{b}_1 r(t_m), \nu_{k+1}, \hat{\eta}_1 r(T_k)\right) d(\hat{b}_1 r(t_m))\\
&-\hat{h}_2\exp\{-\hat{g}_2 r(T_k)\}\int_0^{\hat{b}_2 r^*} q_{\chi^2}\left(\hat{b}_2 r(t_m), \nu_{k+1}, \hat{\eta}_2 r(T_k)\right) d(\hat{b}_2 r(t_m))
\end{aligned}
$$

with

$$
\begin{aligned}
\hat{h}_j &:= \mathcal{A}_{k+1}(T_k, t_m)h_j(t_m)\left(\frac{\tilde{b}}{\tilde{b}+2g_j(t_m)}\right)^{\frac{1}{2}\nu_{k+1}}\\
\hat{g}_j &:= \mathcal{B}_{k+1}(T_k, t_m) + \frac{g_j(t_m)\tilde{\eta}}{\tilde{b}+2g_j(t_m)}\\
\hat{b}_j &:= \tilde{b}+2g_j(t_m)\\
\hat{\eta}_j &:= \frac{\tilde{b}}{\tilde{b}+2g_j(t_m)}\tilde{\eta}
\end{aligned}
$$

and

$$
\int_0^{\hat{b}_j r^*} q_{\chi^2}\left(\hat{b}_j r(t_m), \nu_{k+1}, \hat{\eta}_j r(T_k)\right) d\left(\hat{b}_j r(t_m)\right) = \chi^2_{\nu_{k+1}, \hat{\eta}_j r(T_k)}(\hat{b}_j r^*)
$$

the value of the noncentral chi–square distribution function.

(13) is the formula for pricing the exchange option defined by (10) and (9) in a constant parameter CIR model. Given this price at T_k, the next lower segment boundary to option expiry t_m, in analogy to proposition 2.2.1 we now state the pricing formula for the earlier segment boundaries T_n, $n < k$:

3.1.1. Proposition. *In the segmented square root model with time segments $[T_{n-1}; T_n]$, $n \in \{1; \ldots; N\}$, consider an exchange option defined by (10) and (9), k defined by (11).*

For $n \leq k$, the time T_n price of the option is given by

$$
\begin{aligned}
(14)\qquad C(T_n, r(T_n)) &= \mathcal{C}^{(1)}_{n,k} \exp\left\{-\mathcal{D}^{(1)}_{n,k} r(T_n)\right\} P\left(b^{(1)}_{n,k+1} r(t_m) \leq z^*_m\right) \\
&\quad -\mathcal{C}^{(2)}_{n,k} \exp\left\{-\mathcal{D}^{(2)}_{n,k} r(T_n)\right\} P\left(b^{(2)}_{n,k+1} r(t_m) \leq z^*_m\right) \\
\text{if}\quad V_1(t_m, r(t_m)) &> V_2(t_m, r(t_m)) \quad \text{for} \quad r(t_m) \in]0; r^*[\quad \text{and} \\
(15)\qquad C(T_n, r(T_n)) &= \mathcal{C}^{(1)}_{n,k} \exp\left\{-\mathcal{D}^{(1)}_{n,k} r(T_n)\right\} \left(1 - P\left(b^{(1)}_{n,k+1} r(t_m) \leq z^*_m\right)\right) \\
&\quad -\mathcal{C}^{(2)}_{n,k} \exp\left\{-\mathcal{D}^{(2)}_{n,k} r(T_n)\right\} \left(1 - P\left(b^{(2)}_{n,k+1} r(t_m) \leq z^*_m\right)\right) \\
\text{if}\quad V_1(t_m, r(t_m)) &> V_2(t_m, r(t_m)) \quad \text{for} \quad r(t_m) \in]r^*; \infty[,
\end{aligned}
$$

with $\mathcal{C}^{(j)}_{n,k}$ and $\mathcal{D}^{(j)}_{n,k}$ recursively defined as in proposition 2.2.1, however with $\mathcal{C}^{(j)}_{k,k}$ and $\mathcal{D}^{(j)}_{k,k}$ given by (13):

$$
\mathcal{C}^{(j)}_{k,k} := \hat{h}_j \qquad \mathcal{D}^{(j)}_{k,k} := \hat{g}_j.
$$

$P\left(b^{(j)}_{n,k+1} r(t_m) \leq z^*_m\right) = \int_0^{z^*_m} p\left(b^{(j)}_{n,k+1} r(t_m) = z_m\right) dz_m$ *is the distribution function of a $(k-n+1)$–times multiple compound*[4] *noncentral χ^2 distributed random variable with degrees of freedom ν_{n+s}, $s \in \{1; \ldots ; k-n+1\}$, noncentrality parameters*

$$
\lambda^{(j)}_s := \begin{cases} \hat{\eta}_j r(T_k) & s = k-n+1 \\ \frac{\eta_{n+s}\, b_{n+s}}{b_{n+s} + 2\mathcal{D}^{(j)}_{n+s,k}} r(T_{n+s-1}) & s \in \{1; \ldots ; k-n\} \end{cases}
$$

and transformation coefficients

$$
b^{(j)}_{n,n+s} := \begin{cases} \hat{b}_j & s = k-n+1 \\ b_{n+s} + 2\mathcal{D}^{(j)}_{n+s,k} & s \in \{1; \ldots ; k-n\} \end{cases}
$$

and z^*_m is given by

$$
z^*_m = b^{(j)}_{n,k+1} (g_1 - g_2)^{-1} \ln \frac{h_1}{h_2}.
$$

PROOF: Again, we carry out the proof by induction, showing the first of the two analogous cases (14) and (15): For $n = k$, (14) is identical to (13). Let (14) be valid for some $0 < n \leq k$. Then we have for $n-1$:

$$
\begin{aligned}
& C\left(T_{n-1}, r(T_{n-1})\right) \\
=\ & B\left(r(T_{n-1}), T_{n-1}, T_n\right) \int_0^\infty C\left(T_n, r(T_n)\right) q_{\chi^2}\left(b_n r(T_n), \nu_n, \eta_n r(T_{n-1})\right) d(b_n r(T_n))
\end{aligned}
$$

$$
\begin{aligned}
=\ & \mathcal{A}_n(T_{n-1}, T_n) \exp\left\{-\mathcal{B}_n(T_{n-1}, T_n) r(T_{n-1})\right\} \\
& \cdot \Bigg(\mathcal{C}^{(1)}_{n,k} \int_0^\infty \exp\{-\mathcal{D}^{(1)}_{n,k} r(T_n)\} \int_0^{z^*_m} p\left(b^{(1)}_{n,k+1} r(t_m) = z_m\right) dz_m q_{\chi^2}\left(b_n r(T_n), \nu_n, \eta_n r(T_{n-1})\right) d\left(b_n r(T_n)\right) \\
& -\mathcal{C}^{(2)}_{n,k} \int_0^\infty \exp\{-\mathcal{D}^{(2)}_{n,k} r(T_n)\} \int_0^{z^*_m} p\left(b^{(2)}_{n,k+1} r(t_m) = z_m\right) dz_m q_{\chi^2}\left(b_n r(T_n), \nu_n, \eta_n r(T_{n-1})\right) d\left(b_n r(T_n)\right)\Bigg).
\end{aligned}
$$

[4] see appendix D

By interchanging the order of integration and applying lemma C.1, we get

$$\begin{aligned}
&\mathcal{A}_n(T_{n-1},T_n)\exp\{-\mathcal{B}_n(T_{n-1},T_n)r(T_{n-1})\}\\
&\cdot\mathcal{C}^{(j)}_{n,k}\int_0^\infty \exp\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\}\int_0^{z^*_m} p\left(b^{(j)}_{n,k+1}r(t_m)=z_m\right)dz_m q_{\chi^2}(b_n r(T_n),\nu_n,\eta_n r(T_{n-1}))d(b_n r(T_n))\\
=\;&\mathcal{A}_n(T_{n-1},T_n)\exp\{-\mathcal{B}_n(T_{n-1},T_n)r(T_{n-1})\}\\
&\cdot\mathcal{C}^{(j)}_{n,k}\int_0^{z^*_m}\int_0^\infty p\left(b^{(j)}_{n,k+1}r(t_m)=z_m\right)\exp\left\{-\frac{\mathcal{D}^{(j)}_{n,k}}{b_n+2\mathcal{D}^{(j)}_{n,k}}\eta_n r(T_{n-1})\right\}\left(\frac{b_n}{b_n+2\mathcal{D}^{(j)}_{n,k}}\right)^{\frac{1}{2}\nu_n}\\
&\cdot q_{\chi^2}\left((b_n+2\mathcal{D}^{(j)}_{n,k})r(T_n),\,\nu_n,\,\frac{\eta_n b_n}{b_n+2\mathcal{D}^{(j)}_{n,k}}r(T_{n-1})\right)d\left((b_n+2\mathcal{D}^{(j)}_{n,k})\cdot r(T_n)\right)dz_m\\
=\;&\mathcal{C}^{(j)}_{n-1,k}\exp\left\{-\mathcal{D}^{(j)}_{n-1,k}\,r(T_{n-1})\right\}\int_0^{z^*_m} p\left(b^{(j)}_{n-1,k+1}\cdot r(t_m)=z_m\right)dz_m.
\end{aligned}$$

□

3.1.2. Remark. Note that the $\mathcal{C}^{(j)}_{n,k}\exp\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\}$ are simply the time T_n values of the underlying assets, so (14) can also be written as

$$C(T_n,r(T_n))=V_1(T_n,r(T_n))P\left(b^{(1)}_{n,k+1}r(t_m)\le z^*_m\right)-V_2(T_n,r(T_n))P\left(b^{(2)}_{n,k+1}r(t_m)\le z^*_m\right)$$

3.1.3. Remark. We can use proposition 3.1.1 to calculate the option price at any time t as a function of $r(t)$ by first setting

$$n:=\max\{i\in\{0;\ldots;N\}|t>T_i\}$$

and then $T_n:=t$.

Proposition 3.1.1 allows us to price a wide class of contingent claims. For one, all claims which can be represented as portfolios of (European) options on zero coupon bonds are covered. This includes caps and floors, and also swaptions and options on coupon bonds, since in the one–factor model considered here bond prices are monotonic in the short rate, thus the argument of Jamshidian (1989) is applicable.

Spread options on forward LIBOR can also be priced. More generally, consider a spread option on two forward rates with actuarial compounding. Such a time T forward rate with compounding period α is given at time t by

$$f_n(t,T,T+\alpha)=\frac{1}{\alpha}\left(\frac{B(r(t),t,T)}{B(r(t),t,T+\alpha)}-1\right)$$

and thus the payoff on an option on the spread between two such rates is

$$C(t_m,r(t_m)):=\frac{1}{\alpha}\left[\frac{B(r(t_m),t_m,T_1)}{B(r(t_m),t_m,T_1+\alpha)}-\frac{B(r(t_m),t_m,T_2)}{B(r(t_m),t_m,T_2+\alpha)}\right]^+$$

Since the quotients of zero coupon bond prices are simple exponentials of $r(t_m)$, proposition 3.1.1 applies. Similarly, futures on simple exponentials of r remain simple exponentials in the segmented square root model, as they do in the original CIR, so proposition 3.1.1 can also be used to price options on futures on zero coupon bonds. Finally, proposition 3.1.1 can easily be extended to options on linear combinations of simple exponentials of r, in order to valuate options on nominal yield spreads, for example.

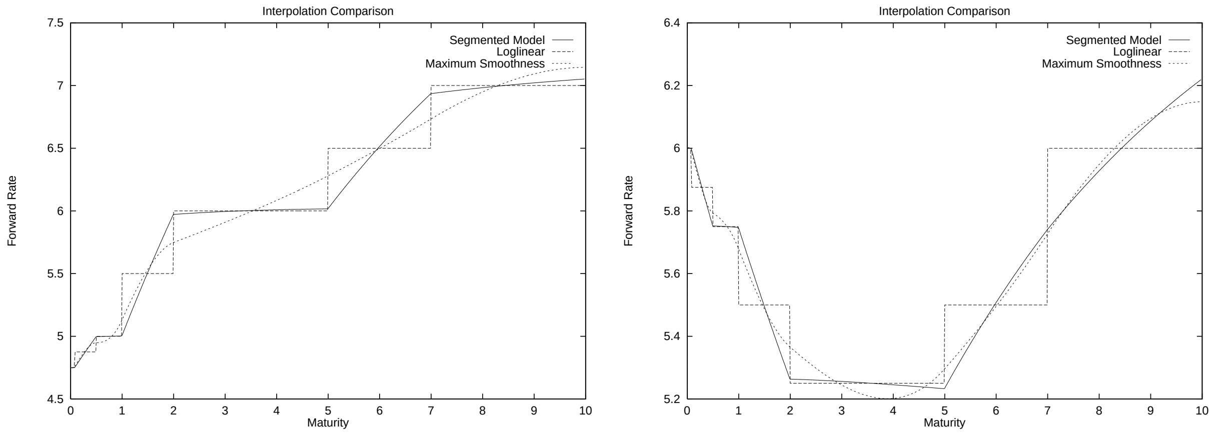

FIGURE 1. Interpolated term structures

The pricing formulae (14) and (15) call for the evaluation of multiple compound noncentral χ^2 distribution functions, which at first glance appears to be a task of high numerical complexity. However, as discussed in appendix D, the number of operations needed to calculate the value of an n–times multiple compound noncentral χ^2 distribution function only grows linearly in n. Furthermore, note that this n is determined by the number of time segments up to the expiry of the option only, since the value of the underlyings at option expiry is known explicitly as a function of r.

3.2. **Term Structure Interpolation.** Proposition 3.1.1 also provides interpolated zero coupon bond prices: If one sets $h_1(t_m) = 1$ and $g_1(t_m) = g_2(t_m) = h_2(t_m) = 0$, then $C(T_n, r(T_n)) = \mathcal{C}^{(1)}_{n,k} \exp\{-\mathcal{D}^{(1)}_{n,k} r(T_n)\}$ is the time T_n price of a zero coupon bond maturing in t_m.

3.2.1. PROPOSITION. *Interpolated forward rate curves in the segmented square root model are continuous.*

PROOF: At any future point in time t, we can without loss of generality set $T_j := t$, where T_{j+1} is the earliest segment boundary greater than t in the original segmentation. We have

$$f(r(T_j), T_j, t_m) = -\frac{\partial}{\partial t_m} \ln \mathcal{C}^{(1)}_{j,k} \exp\{-\mathcal{D}^{(1)}_{j,k} r(T_j)\} \tag{16}$$

Since $\mathcal{C}^{(1)}_{j,k}$ and $\mathcal{D}^{(1)}_{j,k}$ are smooth functions of t_m for $t_m \in]T_k; T_{k+1}[$, we only need to show continuity of (16) on the segment boundaries, i.e. for

$$f(r(T_j), T_j, T_k) = -\frac{\partial}{\partial T_k} \ln \mathcal{C}_{j,k} + r(T_j) \frac{\partial}{\partial T_k} \mathcal{D}_{j,k} \tag{17}$$

with $\mathcal{C}_{j,k}$ and $\mathcal{D}_{j,k}$ defined as in proposition 2.2.1. It is sufficient to show the continuity of the two terms in (17) separately. $\frac{\partial}{\partial T_k}\mathcal{D}_{j,k}$ is already given in (8). Consider

$$\begin{aligned}
\frac{\partial}{\partial T_k}\mathcal{B}_k(T_{k-1}, T_k) &= -2w_k(T_{k-1}, T_k)^2 (c_k + a_k) c_k \exp\{c_k(T_k - T_{k-1})\}(\exp\{c_k(T_k - T_{k-1})\} - 1) \\
&\quad +2w_k(T_{k-1}, T_k) c_k \exp\{c_k(T_k - T_{k-1})\} \\
&= 2w_k(T_{k-1}, T_k)^2((c_k + a_k) c_k \exp\{c_k(T_k - T_{k-1})\}(1 - \exp\{c_k(T_k - T_{k-1})\}) \\
&\quad +c_k \exp\{c_k(T_k - T_{k-1})\}((c_k + a_k)\exp\{c_k(T_k - T_{k-1})\} + c_k - a_k)) \\
&= 4c_k^2 w_k(T_{k-1}, T_k)^2 \exp\{c_k(T_k - T_{k-1})\} = \eta_k b_k^{-1}
\end{aligned}$$

Inserting this into (8), we get

$$\begin{aligned}
\left.\frac{\partial}{\partial T_k}\mathcal{D}_{j,k}\right|_{T_k=T_{k-1}} &= \prod_{i=j+1}^{k-1}\eta_i b_i(b_i+2\mathcal{D}_{i,k})^{-2} \\
&= \left(\prod_{i=j+1}^{k-2}\eta_i b_i(b_i+2\mathcal{D}_{i,k})^{-2}\right)\eta_{k-1}b_{k-1}^{-1} \\
&= \frac{\partial}{\partial T_{k-1}}\mathcal{D}_{j,k-1}
\end{aligned}$$

thus showing the continuity of the second term on the segment boundary. The first term is

$$-\frac{\partial}{\partial T_k}\ln\mathcal{C}_{j,k} = -\frac{\partial}{\partial T_k}\ln\mathcal{A}_k(T_{k-1}.T_k) + \sum_{i=j+1}^{k-1}\nu_i(b_i+2\mathcal{D}_{i,k})^{-1}\frac{\partial}{\partial T_k}\mathcal{D}_{i,k} \tag{18}$$

where

$$\begin{aligned}
-\frac{\partial}{\partial T_k}\ln\mathcal{A}_k(T_{k-1}.T_k) &= -\frac{1}{2}\nu_k\left(2c_k w_k(T_{k-1},T_k)\exp\left\{\frac{1}{2}(c_k+a_k)(T_k-T_{k-1})\right\}\right)^{-1} \\
&\quad\cdot\Bigg(2c_k\frac{1}{2}(c_k+a_k)w_k(T_{k-1},T_k)\exp\left\{\frac{1}{2}(c_k+a_k)(T_k-T_{k-1})\right\} \\
&\qquad -2c_k w_k(T_{k-1},T_k)^2(c_k+a_k)c_k\exp\{c_k(T_k-T_{k-1})\} \\
&\qquad \exp\left\{\frac{1}{2}(c_k+a_k)(T_k-T_{k-1})\right\}\Bigg) \\
&= -\frac{1}{2}\nu_k\left(\frac{1}{2}(c_k+a_k)-c_k w_k(T_{k-1},T_k)(c_k+a_k)\exp\{c_k(T_k-T_{k-1})\}\right) \\
&= \frac{1}{4}\nu_k w_k(T_{k-1},T_k)(c_k+a_k)(c_k-a_k)(\exp\{c_k(T_k-T_{k-1})\}-1) \\
&= \frac{1}{4}\nu_k\sigma_k^2\mathcal{B}_k(T_{k-1},T_k) = \nu_k b_k^{-1}
\end{aligned}$$

Inserting this into (18), by the continuity of the $\frac{\partial}{\partial T_k}\mathcal{D}_{i,k}$ we get

$$\begin{aligned}
\left.-\frac{\partial}{\partial T_k}\ln\mathcal{C}_{j,k}\right|_{T_k=T_{k-1}} &= \sum_{i=j+1}^{k-1}\nu_i(b_i+2\mathcal{D}_{i,k})^{-1}\frac{\partial}{\partial T_k}\mathcal{D}_{i,k} \\
&= \left(\sum_{i=j+1}^{k-2}\nu_i(b_i+2\mathcal{D}_{i,k-1})^{-1}\frac{\partial}{\partial T_{k-1}}\mathcal{D}_{i,k-1}\right)+\nu_{k-1}b_{k-1}^{-1} \\
&= \frac{\partial}{\partial T_{k-1}}\ln\mathcal{C}_{j,k-1}
\end{aligned}$$

□

Figure 1 shows examples of how the segmented square root model interpolates initial term structures, as compared to loglinear interpolation of zero coupon bond prices and the "maximum smoothness" approach of Adams and van Deventer (1994).

3.3. **Hedging.** Although in the literature on CIR type models hedging strategies are generally not discussed at length, we do so here. For one because the existence of a self-financing duplicating strategy is the justification for pricing derivative instruments by arbitrage, and secondly because the hedge ratios in the underlying securities are *not* as readily apparent from the pricing formulae as in the Black and Scholes (1973) model.

Since (14) and (15) were determined using the no-arbitrage condition, there remain two conditions which a self-financing portfolio strategy duplicating the option must satisfy: The value of the portfolio must equal the value of the option at all times and the martingale part of the portfolio process must match the martingale part of the option process. For (14), the former yields[5]

$$\phi_1 = P\left(b^{(1)}_{n,k+1}r(t_m) \leq z^*_m\right) - \frac{\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}}{\mathcal{C}^{(1)}_{n,k}\exp\left\{-\mathcal{D}^{(1)}_{n,k}r(T_n)\right\}}\left(P\left(b^{(2)}_{n,k+1}r(t_m) \leq z^*_m\right) + \phi_2\right) \tag{19}$$

Let $(X)^M$ denote the (uniquely determined) martingale part of a continuous semimartingale X. By Itô's Lemma

$$d(C(T_n, r(T_n)))^M = \frac{\partial}{\partial r(T_n)}C(T_n, r(T_n))d(r(T_n))^M$$

with

$$\begin{aligned}\frac{\partial}{\partial r(T_n)}C(T_n, r(T_n)) = \sum_{j=1}^{2}(-1)^j \Bigg(&\mathcal{D}^{(j)}_{n,k}\mathcal{C}^{(j)}_{n,k}\exp\left\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\right\}P\left(b^{(j)}_{n,k+1}r(t_m) \leq z^*_m\right) \\ &- \mathcal{C}^{(j)}_{n,k}\exp\left\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\right\}\frac{\partial}{\partial r(T_n)}P\left(b^{(j)}_{n,k+1}r(t_m) \leq z^*_m\right)\Bigg)\end{aligned} \tag{20}$$

Similarly, the martingale parts of the processes of the underlyings are

$$d\left(\mathcal{C}^{(j)}_{n,k}\exp\left\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\right\}\right)^M = -\mathcal{D}^{(j)}_{n,k}\mathcal{C}^{(j)}_{n,k}\exp\left\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\right\}d(r(T_n))^M \tag{21}$$

In contrast to the case of options on lognormal assets (such as in the Black/Scholes case), the derivatives of the distribution function do not cancel out.

By (20) and (21), in order to match the martingale parts of the portfolio and the option processes, we must have

$$\begin{aligned}&-\mathcal{D}^{(1)}_{n,k}\mathcal{C}^{(1)}_{n,k}\exp\left\{-\mathcal{D}^{(1)}_{n,k}r(T_n)\right\}P\left(b^{(1)}_{n,k+1}r(t_m) \leq z^*_m\right) \\ &+\mathcal{C}^{(1)}_{n,k}\exp\left\{-\mathcal{D}^{(1)}_{n,k}r(T_n)\right\}\frac{\partial}{\partial r(T_n)}P\left(b^{(1)}_{n,k+1}r(t_m) \leq z^*_m\right) \\ &+\mathcal{D}^{(2)}_{n,k}\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}P\left(b^{(2)}_{n,k+1}r(t_m) \leq z^*_m\right) \\ &-\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}\frac{\partial}{\partial r(T_n)}P\left(b^{(2)}_{n,k+1}r(t_m) \leq z^*_m\right) \\ =&-\phi_1\mathcal{D}^{(1)}_{n,k}\mathcal{C}^{(1)}_{n,k}\exp\left\{-\mathcal{D}^{(1)}_{n,k}r(T_n)\right\} - \phi_2\mathcal{D}^{(2)}_{n,k}\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}\end{aligned}$$

[5]In order not to complicate the notation further, we apply remark 3.1.3 in what follows in this section.

and inserting (19)

$$\begin{aligned}
&\mathcal{C}^{(1)}_{n,k}\exp\left\{-\mathcal{D}^{(1)}_{n,k}r(T_n)\right\}\frac{\partial}{\partial r(T_n)}P\left(b^{(1)}_{n,k+1}r(t_m)\leq z^*_m\right)\\
&+\mathcal{D}^{(2)}_{n,k}\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}P\left(b^{(2)}_{n,k+1}r(t_m)\leq z^*_m\right)\\
&-\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}\frac{\partial}{\partial r(T_n)}P\left(b^{(2)}_{n,k+1}r(t_m)\leq z^*_m\right)\\
=&\mathcal{D}^{(1)}_{n,k}\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}P\left(b^{(2)}_{n,k+1}r(t_m)\leq z^*_m\right)+\phi_2(\mathcal{D}^{(1)}_{n,k}-\mathcal{D}^{(2)}_{n,k})\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}
\end{aligned}$$

Solving for ϕ_2,

$$\begin{aligned}
\phi_2=&(\mathcal{D}^{(1)}_{n,k}-\mathcal{D}^{(2)}_{n,k})^{-1}\left(\frac{\mathcal{C}^{(1)}_{n,k}\exp\left\{-\mathcal{D}^{(1)}_{n,k}r(T_n)\right\}}{\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}}\frac{\partial}{\partial r(T_n)}P\left(b^{(1)}_{n,k+1}r(t_m)\leq z^*_m\right)\right.\\
&\left.-\frac{\partial}{\partial r(T_n)}P\left(b^{(2)}_{n,k+1}r(t_m)\leq z^*_m\right)\right)-P\left(b^{(2)}_{n,k+1}r(t_m)\leq z^*_m\right)
\end{aligned}$$

and similarly

$$\begin{aligned}
\phi_1=&P\left(b^{(1)}_{n,k+1}r(t_m)\leq z^*_m\right)-(\mathcal{D}^{(1)}_{n,k}-\mathcal{D}^{(2)}_{n,k})^{-1}\left(\frac{\partial}{\partial r(T_n)}P\left(b^{(1)}_{n,k+1}r(t_m)\leq z^*_m\right)\right.\\
&\left.-\frac{\mathcal{C}^{(2)}_{n,k}\exp\left\{-\mathcal{D}^{(2)}_{n,k}r(T_n)\right\}}{\mathcal{C}^{(1)}_{n,k}\exp\left\{-\mathcal{D}^{(1)}_{n,k}r(T_n)\right\}}\frac{\partial}{\partial r(T_n)}P\left(b^{(2)}_{n,k+1}r(t_m)\leq z^*_m\right)\right)
\end{aligned}$$

The derivative of the multiple compound noncentral χ^2 distribution function with respect to $r(T_n)$ is given in lemma D.4.

Of course, the duplicating portfolio can also be constructed using two instruments different from the underlying securities, either directly as above or by first duplicating the underlyings: To duplicate the option using some zero coupon bond $B(r(T_n),T_n,T_x)$ and the savings account, we use the equations $(j=1;2)$:

$$\begin{aligned}
\mathcal{C}^{(j)}_{n,k}\exp\left\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\right\} &= \phi^{(j)}_0+\phi^{(j)}_xB(r(T_n),T_n,T_x)\\
-\mathcal{D}^{(j)}_{n,k}\mathcal{C}^{(j)}_{n,k}\exp\left\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\right\} &= -\phi^{(j)}_x\mathcal{D}_{n,x}B(r(T_n),T_n,T_x)
\end{aligned}$$

and thus

$$\begin{aligned}
\phi^{(j)}_x &= \frac{\mathcal{D}^{(j)}_{n,k}\mathcal{C}^{(j)}_{n,k}\exp\left\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\right\}}{\mathcal{D}_{n,x}B(r(T_n),T_n,T_x)}\\
\phi^{(j)}_0 &= \mathcal{C}^{(j)}_{n,k}\exp\left\{-\mathcal{D}^{(j)}_{n,k}r(T_n)\right\}\left(1-\frac{\mathcal{D}^{(j)}_{n,k}}{\mathcal{D}_{n,x}}\right)
\end{aligned}$$

Therefore

$$\begin{aligned}
&\phi^{(1)}_x\phi_1+\phi^{(2)}_x\phi_2 \qquad \text{of } B(r(T_n),T_n,T_x)\\
\text{and}\quad &\phi^{(1)}_0\phi_1+\phi^{(2)}_0\phi_2 \qquad \text{in the savings account}
\end{aligned}$$

duplicate the option.

4. Multifactor Extensions

For many applications, most prominently derivatives whose value depends on the shape of the yield curve, one-factor models are inadequate. This leads us to consider how the above techniques may be extended to the multifactor case.

4.1. Model Construction.
As in Chen and Scott (1995), let the short rate be given by the sum of independent state variables

$$r(t) = \sum_{j=1}^{J} z_j(t)$$

where the state variable dynamics are of the CIR type

$$dz_j(t) \;=\; \left(\theta^{(j)} - a^{(j)} z_j(t)\right) dt + \sigma^{(j)} \sqrt{z_j(t)} dW_t^{(j)}. \tag{22}$$

In our segmented version of the multifactor model, we replace (22) with the dynamics given in equation (2). For each segment $[T_{n-1}; T_n]$ we now have parameter vectors $\theta_n^{(\cdot)}, a_n^{(\cdot)}, \sigma_n^{(\cdot)}$, and the vector valued functions $\mathcal{A}_n^{(\cdot)}(T_{n-1}; T_n), \mathcal{B}_n^{(\cdot)}(T_{n-1}; T_n), \eta_n^{(\cdot)}, b_n^{(\cdot)}$ and $\nu_n^{(\cdot)}$ are defined element–wise as in the one–factor case. On a constant parameter segment, zero coupon bond prices are[6]

$$B(Z(T_{n-1}), T_{n-1}, T_n) = \left(\prod_{j=1}^{J} \mathcal{A}_n^{(j)}(T_{n-1}, T_n) \right) \exp \left\{ -\sum_{j=1}^{J} \mathcal{B}_n^{(j)}(T_{n-1}, T_n) z_j(T_{n-1}) \right\}.$$

The transformed factors $b_n^{(j)} z_j(T_n)$ are noncentral χ^2 distributed (conditioned on $z_j(T_{n-1})$) under the T_n forward measure, with $\nu_n^{(j)}$ degrees of freedom and noncentrality parameter $\eta_n^{(j)} z_j(T_{n-1})$. Since the factors are independently distributed, we can apply lemma C.1 for each factor separately and carry out the same induction as in the proof of proposition 2.2.1 to yield

4.1.1. Proposition. *In the multifactor segmented square root model with time segments $[T_{n-1}; T_n]$, $n \in \{1; \ldots; N\}$, the prices of zero coupon bonds at time T_{n-1} with maturity T_k, $n \le k \le N$ are given by*

$$B(Z(T_{n-1}), T_{n-1}, T_k) = \left(\prod_{j=1}^{J} \mathcal{C}_{n-1,k}^{(j)} \right) \exp \left\{ -\sum_{j=1}^{J} \mathcal{D}_{n-1,k}^{(j)} z_j(T_{n-1}) \right\}$$

with $\mathcal{C}_{n-1,k}^{(j)}$ and $\mathcal{D}_{n-1,k}^{(j)}$ recursively defined as

$$\mathcal{C}_{n-1,k}^{(j)} := \mathcal{C}_{n,k}^{(j)} \mathcal{A}_n^{(j)}(T_{n-1}, T_n) \left(\frac{b_n^{(j)}}{b_n^{(j)} + 2\mathcal{D}_{n,k}^{(j)}} \right)^{\frac{1}{2}\nu_n^{(j)}}$$

$$\mathcal{D}_{n-1,k}^{(j)} := \mathcal{B}_n^{(j)}(T_{n-1}, T_n) + \eta_n^{(j)} \frac{\mathcal{D}_{n,k}^{(j)}}{b_n^{(j)} + 2\mathcal{D}_{n,k}^{(j)}}$$

where

$$\mathcal{C}_{k,k}^{(j)} := 1 \qquad \textit{and} \qquad \mathcal{D}_{k,k}^{(j)} := 0.$$

[6] See Chen and Scott (1995).

Note that when fitting the model to an initial term structure in the multifactor case, there are more parameters which can be adjusted. Thus in a two–factor model one could choose to fit two initial zero coupon bond prices on each segment, reducing the number of segments.

4.2. **Option Pricing.** As in section 3, consider a European exchange option on two securities whose terminal values are exponential affine functions of the factors. Since the factors are independent, we can write the time T_k price of the option analogously to equation (12) as[7]

(23)

$$\begin{aligned}C(T_k, Z(T_k)) &= B(Z(T_k), T_k, t_m)\\ &\cdot \Bigg(\int_{\mathbb{Z}} h_1(t_m) \prod_{j=1}^{J} \exp\left\{-g_1^{(j)}(t_m) z_j(t_m)\right\} q_{\chi^2}\left(\tilde{b}^{(j)} z_j(t_m), \nu_{k+1}^{(j)}, \tilde{\eta}^{(j)} z_j(T_k)\right) d\left(\tilde{b}^{(j)} z_j(t_m)\right)\\ &- \int_{\mathbb{Z}} h_2(t_m) \prod_{j=1}^{J} \exp\left\{-g_2^{(j)}(t_m) z_j(t_m)\right\} q_{\chi^2}\left(\tilde{b}^{(j)} z_j(t_m), \nu_{k+1}^{(j)}, \tilde{\eta}^{(j)} z_j(T_k)\right) d\left(\tilde{b}^{(j)} z_j(t_m)\right)\Bigg)\end{aligned}$$

with

$$\begin{aligned}\mathbb{Z} &:= \left\{ Z(t_m) \in \mathbb{R}_{++}^J \,\middle|\, h_1(t_m) \exp\left\{-\sum_{j=1}^{J} g_1^{(j)}(t_m) z_j(t_m)\right\} > h_2(t_m) \exp\left\{-\sum_{j=1}^{J} g_2^{(j)}(t_m) z_j(t_m)\right\}\right\}\\ &= \left\{ Z(t_m) \in \mathbb{R}_{++}^J \,\middle|\, \sum_{j=1}^{J} \left(g_2^{(j)}(t_m) - g_1^{(j)}(t_m)\right) z_j(t_m) > \ln \frac{h_2(t_m)}{h_1(t_m)} \right\}.\end{aligned}$$

Again we can apply lemma C.1 and write (23) as

$$\begin{aligned}&C(T_k, Z(T_k)) = B(Z(T_k), T_k, t_m)\\ &\cdot \left(h_1(t_m) \left(\prod_{j=1}^{J} \exp\left\{ -\frac{g_1^{(j)}(t_m)}{\tilde{b}^{(j)} + 2 g_1^{(j)}(t_m)} \tilde{\eta}^{(j)} z_j(T_k) \right\} \left(\frac{\tilde{b}^{(j)}}{\tilde{b}^{(j)} + 2 g_1^{(j)}(t_m)} \right)^{\frac{1}{2}\nu_{k+1}^{(j)}} \right)\right.\\ &\cdot \int_{\mathbb{Z}} \left(\prod_{j=1}^{J} q_{\chi^2} \left(\left(\tilde{b}^{(j)} + g_1^{(j)}(t_m)\right) z_j(t_m), \nu_{k+1}^{(j)}, \frac{\tilde{b}^{(j)}}{\tilde{b}^{(j)} + 2 g_1^{(j)}(t_m)} \tilde{\eta}^{(j)} z_j(T_k) \right) d\left(\left(\tilde{b}^{(j)} + 2 g_1^{(j)}(t_m) \right) z_j(t_m) \right) \right)\\ &- h_2(t_m) \left(\prod_{j=1}^{J} \exp\left\{ -\frac{g_2^{(j)}(t_m)}{\tilde{b}^{(j)} + 2 g_2^{(j)}(t_m)} \tilde{\eta}^{(j)} z_j(T_k) \right\} \left(\frac{\tilde{b}^{(j)}}{\tilde{b}^{(j)} + 2 g_2^{(j)}(t_m)} \right)^{\frac{1}{2}\nu_{k+1}^{(j)}} \right)\\ &\left.\cdot \int_{\mathbb{Z}} \left(\prod_{j=1}^{J} q_{\chi^2} \left(\left(\tilde{b}^{(j)} + g_2^{(j)}(t_m)\right) z_j(t_m), \nu_{k+1}^{(j)}, \frac{\tilde{b}^{(j)}}{\tilde{b}^{(j)} + 2 g_2^{(j)}(t_m)} \tilde{\eta}^{(j)} z_j(T_k) \right) d\left(\left(\tilde{b}^{(j)} + 2 g_2^{(j)}(t_m) \right) z_j(t_m) \right) \right) \right).\end{aligned}$$

This is the exchange option formula for a multifactor CIR model with independent factors and constant parameters. In order to calculate the option price for some arbitrary time T_n in the segmented model and derive the multifactor version of proposition 3.1.1, we carry out the same induction steps as in the proof of 3.1.1. Given the option price $C(T_n, Z(T_n))$ at the segment boundary T_n as a function of the factor realizations $Z(T_n)$, the time T_{n-1}

[7]Unless otherwise stated, the notation in this section is defined as in section 3.

price of the option can be calculated as the discounted expectation under the time T_n forward measure Q^{T_n}:

$$C(T_{n-1}, Z(T_{n-1})) = B(Z(T_{n-1}), T_{n-1}, T_n) \int_{\mathbb{R}^J_{++}} C(T_n, Z(T_n)) Q^{T_n}(dZ(T_n)).$$

The joint density of the factors $Z(T_n)$ is given by the product of the factor densities because of independence. Noting the multiplicative structure of (23), which is retained in each induction step, we see that we can carry out the induction for each factor separately, yielding

4.2.1. PROPOSITION. *In the multifactor segmented square root model with time segments $[T_{n-1}; T_n]$, $n \in \{1; \ldots; N\}$ and J factors z_j, consider an exchange option on two assets whose values at option expiry t_m are exponential affine functions of the factors:*

$$V_{1,2}(t_m, Z(t_m)) = h_{1,2}(t_m) \exp\left\{ -\sum_{j=1}^{J} g^{(j)}_{1,2}(t_m) z_j(t_m) \right\}.$$

Define k as

$$k := \max\left\{ n \in \{0; \ldots; N\} | T_n < t_m \right\}.$$

For $n \leq k$, the time T_n price of the option is given by

$$\begin{aligned} C(T_n, Z(T_n)) = &\left(\prod_{j=1}^{J} \mathcal{C}^{(1,j)}_{n,k} \exp\left\{ -\mathcal{D}^{(1,j)}_{n,k} z_j(T_n) \right\} \right) P_1 \left(\sum_{j=1}^{J} \left(g^{(j)}_1(t_m) - g^{(j)}_2(t_m) \right) z_j(t_m) < \ln \frac{h_1(t_m)}{h_2(t_m)} \right) \\ &- \left(\prod_{j=1}^{J} \mathcal{C}^{(2,j)}_{n,k} \exp\left\{ -\mathcal{D}^{(2,j)}_{n,k} z_j(T_n) \right\} \right) P_2 \left(\sum_{j=1}^{J} \left(g^{(j)}_1(t_m) - g^{(j)}_2(t_m) \right) z_j(t_m) < \ln \frac{h_1(t_m)}{h_2(t_m)} \right) \end{aligned}$$

with $\mathcal{C}^{(i,j)}_{n,k}$ and $\mathcal{D}^{(i,j)}_{n,k}$ recursively defined as in proposition (4.1.1), however with

$$\begin{aligned} \mathcal{C}^{(i,j)}_{k,k} &:= \mathcal{A}^{(j)}_{k+1}(T_k, t_m) h_i(t_m) \left(\frac{\tilde{b}^{(j)}}{\tilde{b}^{(j)} + 2g^{(j)}_i(t_m)} \right)^{\frac{1}{2}\nu^{(j)}_{k+1}} \\ \mathcal{D}^{(i,j)}_{k,k} &:= \mathcal{B}^{(j)}_{k+1}(T_k, t_m) + \frac{g^{(j)}_i(t_m) \tilde{\eta}^{(j)}}{\tilde{b}^{(j)} + 2g^{(j)}_i(t_m)}. \end{aligned}$$

$P_i \left(\sum_{j=1}^{J} \left(g^{(j)}_1(t_m) - g^{(j)}_2(t_m) \right) z_j(t_m) < \ln \frac{h_1(t_m)}{h_2(t_m)} \right)$ is the distribution function of a weighted sum of independent $(k-n+1)$–times multiple compound noncentral χ^2 distributed factors with degrees of freedom $\nu^{(j)}_{n+s}$, $s \in \{1; \ldots; k-n+1\}$, noncentrality parameters

$$\lambda^{(i,j)}_s := \begin{cases} \frac{\tilde{b}^{(j)}}{\tilde{b}^{(j)} + 2g^{(j)}_i(t_m)} \tilde{\eta}^{(j)} z_j(T_k) & s = k-n+1 \\ \frac{\eta^{(j)}_{n+s} b^{(j)}_{n+s}}{b^{(j)}_{n+s} + 2\mathcal{D}^{(i,j)}_{n+s,k}} z_j(T_{n+s-1}) & s \in \{1; \ldots; k-n\} \end{cases}$$

and transformation coefficients

$$b^{(i,j)}_{n,n+s} := \begin{cases} \tilde{b}^{(j)} + 2g^{(j)}_i(t_m) & s = k-n+1 \\ b^{(j)}_{n+s} + 2\mathcal{D}^{(i,j)}_{n+s,k} & s \in \{1; \ldots; k-n\} \end{cases}$$

The distribution function can be evaluated using the technique described by Chen and Scott (1995); the characteristic function of the multiple compound noncentral χ^2 distribution is given by proposition D.5 in the appendix. Note that by representing the value of the distribution function as an integral of a product of the characteristic functions of independent factors, this technique reduces the dimension of the numerical integration to one for any number of factors and any number of segments. Thus for a large number of segments before option expiry it may be efficient to employ this technique even in the one–factor case.

5. Closing Remarks

We have constructed the one– and the multifactor versions of a term structure model with non-negative interest rates which fits an initial yield curve while retaining analytical tractability for fast solutions for derivative pricing and hedging. The factor stochastic differential equations are Cox/Ingersoll/Ross (CIR) type "square root" diffusions with piecewise constant parameters, where the constant parameter segments are determined by the initial term structure data, i.e. by the maturities for which zero coupon bond prices are given. Prices of European options on linear combinations of securities whose value at option expiry is an exponential affine function of the model factors can be expressed in terms of a "multiple compound" noncentral chi-square distribution function, i.e. a noncentral chi-square distribution whose noncentrality parameter is again (multiple compound) chi-square distributed.

In the one-factor case the number of segments n up to option expiry determines the numerical complexity of the problem of calculating this distribution function; the number of operations necessary grows only linearly in n. In the multifactor case and for a large number of segments in the one-factor case, the (explicitly derived) characteristic function can be used to calculate the value of the multiple compound noncentral chi-square distribution function by a one-dimensional numerical integration.

Thus we have arguably closed form solutions for a large class of fixed income derivatives, including caps, floors, yield spreads, options on interest rate futures and, in the one-factor case, swaptions.

Our approach to fitting an initial term structure does not require that we exogenously specify zero coupon bond prices for the continuum of maturities. Instead, the model interpolates endogenously in a manner consistent with the short rate dynamics. However, exogenous interpolation schemes such as splines can be approximated by a sufficiently large number of segments should one choose to do so.

Appendix A. On the Non-Explosion of Solutions of an SDE

Let $\mathcal{U} \subset \mathbb{R}^d$ be open. We denote by $\widehat{\mathcal{U}} := \mathcal{U} \cup \{\Delta\}$ the Alexandrov–Compactification of $\mathcal{U}$. Then $\widehat{\mathcal{U}}$ is compact and its topology has a countable basis, therefore $\widehat{\mathcal{U}}$ is Polish. Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a stochastic basis and $X = (X_t)_{t \in \mathbb{R}_+}$ be a continuous $\mathbb{F}$–adapted process with values in $\widehat{\mathcal{U}}$. We define the explosion time e_X of X as follows:

$$e_X := \inf\{t \in \mathbb{R}_+ \mid X_t = \Delta\}.$$

Then e_X is an $\mathbb{F}$–stopping time. If X_0 only assumes values in $\mathcal{U}$, then we have $e_X > 0$ and furthermore, e_X is predictable.
We denote by $M_{d,n}(\mathbb{R})$ the set of all $d \times n$ matrices with real entries. Let $\alpha : \mathbb{R}_+ \times \mathcal{U} \to M_{d,n}(\mathbb{R})$ and $\beta : \mathbb{R}_+ \times \mathcal{U} \to \mathbb{R}^d$ be two continuous functions; these will be the coefficients

of the SDE we wish to consider. All stochastic bases $(\Omega, \mathcal{F}, \mathbb{F}, P)$ will fulfill the usual hypotheses.

A.1. Definition. **(Solution)**: Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a stochastic basis on which an n–dimensional, standard Brownian motion $B = (B_t)_{t \in \mathbb{R}_+}$ is defined. A continuous, $\mathbb{F}$–adapted process X taking values in $\widehat{\mathcal{U}}$ is a solution of

$$dZ_t = \alpha(t, Z_t) dB_t + \beta(t, Z_t) dt \tag{A.1}$$

iff the following conditions are met:

1. X_0 only assumes values in $\mathcal{U}$.
2. For P–almost all $\omega \in \{e_X < +\infty\}$ we have
$$\forall t \geq e_X(\omega) : X_t(\omega) = \Delta.$$
(This condition is obviously fulfilled iff the two processes X and X^{ex} are indistinguishable.)
3. If τ is an $\mathbb{F}$–stopping time with $[\![0, \tau]\!] \subset [\![0, e_X[\![$, then for every $i \in \{1, \dots, d\}$ and every $t \in \mathbb{R}_+$ we have the following:
$$X_t^{(i)^\tau} = X_0^{(i)} + \sum_{j=1}^n \int_0^t \alpha_{ij}\left(s, X_s^\tau\right) dB_s^{(j)^\tau} + \int_0^{t\wedge\tau} \beta_i\left(s, X_s^\tau\right) ds \quad P\text{–a.s.}$$

The solution X does not explode iff $P[e_X = +\infty] = 1$.

A.2. Lemma. *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a stochastic basis carrying a normal n–dimensional Brownian motion W and let X be a continuous, $\mathbb{F}$–adapted process in $\widehat{\mathcal{U}}$ so that X_0 only assumes values in $\mathcal{U}$ and the two processes X and X^{ex} are indistinguishable. Let $(\tau_k)_{k \in \mathbb{N}}$ be an announcing sequence for e_X. Then the following two statements are equivalent:*

1. *The process X is a solution of*
$$dZ_t = \alpha(t, Z_t) dW_t + \beta(t, Z_t) dt.$$
2. *For each $i \in \{1, \dots, d\}$, $k \in \mathbb{N}$ and $t \in \mathbb{R}_+$ we have*
$$X_t^{(i)^{\tau_k}} = X_0^{(i)} + \sum_{j=1}^n \int_0^t \alpha_{ij}\left(s, X_s^{\tau_k}\right) dW_s^{(j)^{\tau_k}} + \int_0^{t\wedge\tau_k} \beta_i\left(s, X_s^{\tau_k}\right) ds \quad P\text{–a.s.}$$

Proof: Obviously, we only need to show that 2. implies 1. Let τ be any stopping time with $[\![0, \tau]\!] \subset [\![0, e_X[\![$. Fixing $i \in \{1, \dots, d\}$ and $t \in \mathbb{R}_+$ we must show:

$$X_t^{(i)^\tau} = X_0^{(i)} + \sum_{j=1}^n \int_0^t \alpha_{ij}\left(s, X_s^\tau\right) dW_s^{(j)^\tau} + \int_0^{t\wedge\tau} \beta_i\left(s, X_s^\tau\right) ds \quad P\text{–a.s.}$$

By the stopping rules for stochastic integrals and the assumption, we have for every $k \in \mathbb{N}$,

$$\begin{aligned} X_t^{(i)^{\tau_k \wedge \tau}} &= X_0^{(i)} + \sum_{j=1}^n \int_0^{t\wedge\tau} \alpha_{ij}\left(s, X_s^{\tau_k}\right) dW_s^{(j)^{\tau_k}} + \int_0^{t\wedge\tau_k\wedge\tau} \beta_i\left(s, X_s^{\tau_k}\right) ds \\ &= X_0^{(i)} + \sum_{j=1}^n \int_0^{t\wedge\tau\wedge\tau_k} \alpha_{ij}\left(s, X_s^{\tau}\right) dW_s^{(j)^{\tau}} + \int_0^{t\wedge\tau\wedge\tau_k} \beta_i\left(s, X_s^{\tau}\right) ds \quad P\text{–a.s.} \end{aligned}$$

We can therefore find a null set N in $(\Omega, \mathcal{F}, P)$, so that for $\omega \in N^c$ the following holds:

$$\forall k \in \mathbb{N}: \quad X_t^{(i)^{\tau_k \wedge \tau}}(\omega) = X_0^{(i)}(\omega) \;+\; \left(\sum_{j=1}^{n} \int_0^{t \wedge \tau \wedge \tau_k} \alpha_{ij}\left(s, X_s^{\tau}\right) dW_s^{(j)^{\tau}}\right)(\omega) \;+\; \int_0^{t \wedge \tau(\omega) \wedge \tau_k(\omega)} \beta_i\left(s, X_s^{\tau}(\omega)\right) ds.$$

Now fix an arbitrary $\omega \in N^c$. Since $t \wedge \tau(\omega) < e_X(\omega)$, there is a $k_0 \in \mathbb{N}$ fulfilling $t \wedge \tau(\omega) \leq \tau_{k_0}(\omega)$. We have

$$\begin{aligned} X^{(i)}{}_t^{\tau}(\omega) &= X_t^{(i)^{\tau_{k_0} \wedge \tau}}(\omega) \\ &= X_0^{(i)}(\omega) + \left(\sum_{j=1}^{n} \int_0^{t \wedge \tau \wedge \tau_{k_0}} \alpha_{ij}\left(s, X_s^{\tau}\right) dW_s^{(j)^{\tau}}\right)(\omega) + \int_0^{t \wedge \tau(\omega) \wedge \tau_{k_0}(\omega)} \beta_i\left(s, X_s^{\tau}(\omega)\right) ds \\ &= X_0^{(i)}(\omega) + \left(\sum_{j=1}^{n} \int_0^{t} \alpha_{ij}\left(s, X_s^{\tau}\right) dW_s^{(j)^{\tau}}\right)(\omega) + \int_0^{t \wedge \tau(\omega)} \beta_i\left(s, X_s^{\tau}(\omega)\right) ds. \qquad \square \end{aligned}$$

The result we use in the main body of the paper is the following

A.3. Theorem. *Suppose that any solution of the SDE determined by α and β satisfying a deterministic initial condition does not explode. Then non–explosion also holds for solutions with random initial conditions.*

For the convenience of the reader, we will provide a complete proof of this intuitive result. Our approach follows that of Yeh (1995), § 18. In particular, we make use of regular conditional probabilities. To be assured of their existence, we must first transport our solution onto a suitably nice probability space. This is done in the next section.

A.1. **Function Space Representation of Solutions.** We first introduce a suitable analogy of Wiener space. As already mentioned, $\widehat{\mathcal{U}}$ is a Polish space, therefore the space $C(\mathbb{R}_+, \widehat{\mathcal{U}})$ of all continuous functions from $\mathbb{R}_+$ to $\widehat{\mathcal{U}}$ endowed with the topology of uniform convergence on compacts is Polish (cf. Bauer (1990), Theorem 31.6). We set

$$\tilde{C}(\mathbb{R}_+, \widehat{\mathcal{U}}) := \left\{ w \in C(\mathbb{R}_+, \widehat{\mathcal{U}}) \middle| \, w(0) \in \mathcal{U} \right\}.$$

Now $\tilde{C}(\mathbb{R}_+, \widehat{\mathcal{U}})$ is an open subset of $C(\mathbb{R}_+, \widehat{\mathcal{U}})$ and therefore also Polish. Finally, we define

$$\widehat{\mathbf{W}} = \tilde{C}(\mathbb{R}_+, \widehat{\mathcal{U}}) \times C(\mathbb{R}_+, \mathbb{R}^n)$$

and endow $\widehat{\mathbf{W}}$ with the product topology, making $\widehat{\mathbf{W}}$ into a Polish space also.
For every $t \in \mathbb{R}_+$ we have the canonical projection mappings

$$\begin{aligned} p_t: &\quad \widehat{\mathbf{W}} \to \widehat{\mathcal{U}}, \quad (\mathtt{w}, \mathtt{w}') \mapsto \mathtt{w}(t) \\ q_t: &\quad \widehat{\mathbf{W}} \to \mathbb{R}^n, \quad (\mathtt{w}, \mathtt{w}') \mapsto \mathtt{w}'(t). \end{aligned}$$

If we denote the Borel–σ–Algebra of $\widehat{\mathbf{W}}$ by $\mathfrak{W}$, we have

$$\mathfrak{W} = \sigma\left(p_s, q_s; s \in \mathbb{R}_+\right).$$

The canonical filtration $\mathbb{W} = \{\mathfrak{W}_t\}_{t \in \mathbb{R}_+}$ on $(\widehat{\mathbf{W}}, \mathfrak{W})$ is given by $\mathfrak{W}_t = \sigma(p_s, q_s; s \in [0, t])$ for every $t \in \mathbb{R}_+$. We also have two canonical stochastic processes $Y = \{Y_t\}_{t \in \mathbb{R}_+}$ and

$W = \{W_t\}_{t\in\mathbb{R}_+}$ on $(\widehat{\mathbf{W}}, \mathbb{W})$ given by

$$\begin{aligned} Y_t(w,w') &:= p_t(\mathtt{w},\mathtt{w}') = \mathtt{w}(t) \\ W_t(w,w') &:= q_t(\mathtt{w},\mathtt{w}') = \mathtt{w}'(t) \end{aligned}$$

for every $t \in \mathbb{R}_+$. The processes Y and W are obviously continuous and $\mathbb{W}$–adapted, the explosion time e_Y of Y is a $\mathbb{W}$–stopping time with $e_Y > 0$.
Now let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a stochastic basis on which we have a normal n–dimensional Brownian motion B and let X be a solution of the SDE

$$dZ_t = \alpha(t,Z_t)dB_t + \beta(t,Z_t)dt. \tag{A.2}$$

This solution induces a canonical map $(X,B) : \Omega \to \widehat{\mathbf{W}}$ defined by

$$(X,B)(\omega) := (X_\bullet(\omega), B_\bullet(\omega)).$$

The mapping (X,B) is $\mathcal{F}$–$\mathfrak{W}$–measurable and also $\mathcal{F}_t$–$\mathfrak{W}_t$–measurable for every $t \in \mathbb{R}_+$. Let $P_{(X,B)}$ denote the image of P under (X,B). We denote by $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}, \mathbb{W}^* := \{\mathfrak{W}^*_t\}_{t\in\mathbb{R}_+}, P_{(X,B)})$ the usual augmentation of the stochastic basis $(\widehat{\mathbf{W}}, \mathfrak{W}, \mathbb{W}, P_{(X,B)})$. Observe that, as $(\Omega, \mathcal{F}, \mathbb{F}, P)$ fulfills the usual hypotheses by assumption, the mapping (X,B) is in fact $\mathcal{F}$–$\overline{\mathfrak{W}}$–measurable and $\mathcal{F}_t$–$\mathfrak{W}^*_t$–measurable for every $t \in \mathbb{R}_+$. It is trivial but useful to note that for every $t \in \mathbb{R}_+$ we have

$$Y_t \circ (X,B) = X_t, \quad W_t \circ (X,B) = B_t.$$

In particular, $e_Y \circ (X,B) = e_X$, and therefore:

$$P_{(X,B)}\left[e_Y = +\infty\right] = P\left[(X,B)^{-1}\left(\{e_Y = +\infty\}\right)\right] = P\left[e_X = +\infty\right].$$

A.1. Theorem. *The process W is a standard, n–dimensional $(P_{(X,B)}, \mathbb{W}^*)$–Brownian motion and the process Y is a solution of*

$$dZ_t = \alpha(t,Z_t)dW_t + \beta(t,Z_t)dt. \tag{A.3}$$

Proof: 1. The paths of W are obviously continuous, W is $\mathbb{W}^*$–adapted. Let $s,t \in \mathbb{R}_+$ with $s < t$. Then we have

$$(W_t - W_s) \circ (X,B) = B_t - B_s.$$

Therefore the distribution of $W_t - W_s$ under $P_{(X,B)}$ is just the distribution of $B_t - B_s$ under P. For the same reason we have

$$P_{(X,B)}\left[W_0 = 0\right] = P\left[B_0 = 0\right] = 1.$$

To prove that W is a Brownian motion, it only remains to show that $W_t - W_s$ is independent of $\mathfrak{W}^*_s$. We denote the Borel σ–algebra of $\mathbb{R}^n$ by $\mathcal{B}^n$. Suppose that $C \in \mathcal{B}^n$ and $A \in \mathfrak{W}^*_s$. Since $(X,B)^{-1}(A) \in \mathcal{F}_s$, we have

$$\begin{aligned} & P_{(X,B)}\left[A \cap (W_t - W_s)^{-1}(C)\right] \\ = \;& P\left[(X,B)^{-1}(A) \cap (B_t - B_s)^{-1}(C)\right] \\ = \;& P\left[(X,B)^{-1}(A)\right] P\left[(B_t - B_s)^{-1}(C)\right] \\ = \;& P_{(X,B)}[A] \cdot P_{(X,B)}\left[(W_t - W_s)^{-1}(C)\right]. \end{aligned}$$

2. We must now show that the process Y is indeed a solution of (A.3). By definition of $\widehat{\mathbf{W}}$, Y_0 assumes only values in $\mathcal{U}$. The processes Y and Y^{e_Y} are continuous and $\mathbb{W}$–adapted, so that $\{Y = Y^{e_Y}\} \in \mathfrak{W}$. Furthermore, the following holds:

$$\forall t \in \mathbb{R}_+ : \quad Y^{e_Y}_t \circ (X,B) = X^{e_X}_t.$$

Therefore

$$P_{(X,B)}\left[Y = Y^{e_Y}\right] = P\left[(X,B)^{-1}\left(\{Y = Y^{e_Y}\}\right)\right] = P\left[X = X^{e_X}\right] = 1.$$

Let τ be a $\mathbb{W}^*$–stopping time with $[\![0,\tau]\!] \subset [\![0,e_Y[\![$, fix $i \in \{1,\dots,d\}$ and $t \in \mathbb{R}_+$. We must show:

$$Y_t^{(i)^\tau} = Y_0^{(i)} + \sum_{j=1}^{n} \int_0^t \alpha_{ij}\left(s, Y_s^\tau\right) dW_s^{(j)^\tau} + \int_0^{t\wedge\tau} \beta_i\left(s, Y_s^\tau\right) ds \quad P_{(X,B)}\text{–a.s.}$$

We define $\tilde{\tau} : \Omega \to \overline{\mathbb{R}}_+$ by $\tilde{\tau} := \tau \circ (X,B)$. One immediately sees that $\tilde{\tau}$ is an $\mathbb{F}$–stopping time with $[\![0,\tilde{\tau}]\!] \subset [\![0,e_X[\![$. Since X is a solution of (A.2), we know that

$$\text{(A.4)} \qquad X_t^{(i)^{\tilde{\tau}}} = X_0^{(i)} + \sum_{j=1}^{n} \int_0^t \alpha_{ij}\left(s, X_s^{\tilde{\tau}}\right) dB_s^{(j)^{\tilde{\tau}}} + \int_0^{t\wedge\tilde{\tau}} \beta_i\left(s, X_s^{\tilde{\tau}}\right) ds \quad P\text{–a.s.}$$

For fixed i and $t \in \mathbb{R}_+$ we define two real random variables Ψ and Φ as follows:

$$\begin{aligned}
\Psi &:= \sum_{j=1}^{n} \int_0^t \alpha_{ij}\left(s, Y_s^\tau\right) dW_s^{(j)^\tau} + \int_0^{t\wedge\tau} \beta_i\left(s, Y_s^\tau\right) ds + Y_0^{(i)} - Y_t^{(i)^\tau} \\
\Phi &:= \sum_{j=1}^{n} \int_0^t \alpha_{ij}\left(s, X_s^{\tilde{\tau}}\right) dB_s^{(j)^{\tilde{\tau}}} + \int_0^{t\wedge\tilde{\tau}} \beta_i\left(s, X_s^{\tilde{\tau}}\right) ds + X_0^{(i)} - X_t^{(i)^{\tilde{\tau}}}
\end{aligned}$$

Now (A.4) is equivalent to the fact that the distribution of Φ is δ_0, the Dirac measure at the origin. It is clearly sufficient to prove that Φ and Ψ are identically distributed.
To this end, we choose a sequence $(\mathcal{Z}_m)_{m\in\mathbb{N}}$ of partitions $\mathcal{Z}_m : 0 = t_0^m < \ldots < t_{k_m}^m = t$ of $[0,t]$, so that $|\mathcal{Z}_m| \to 0$. Fix $\mathtt{w} \in \widehat{\mathbf{W}}$. For every $m \in \mathbb{N}$ we define $\beta_i^{(m)} : [0,t] \to \mathbb{R}$ by

$$\beta_i^{(m)} := \beta_i\left(0, Y_0(\mathtt{w})\right)\chi_{\{0\}} + \sum_{\nu=1}^{k_m} \beta_i\left(t_{\nu-1}^m, Y_{t_{\nu-1}^m}^\tau(\mathtt{w})\right)\chi_{]t_{\nu-1}^m, t_\nu^m]}.$$

The continuous function mapping $[0,t]$ to $\mathbb{R}$ by $s \mapsto \beta_i(s, Y_s^\tau(\mathtt{w}))$ is the pointwise limit of the sequence $(\beta_i^{(m)})_{m\in\mathbb{N}}$. The sequence $(\beta_i^{(m)})_{m\in\mathbb{N}}$ is uniformly bounded by $\sup_{s\in[0,t]} |\beta_i(s, Y_s^\tau(\mathtt{w}))|$ $< +\infty$ on $[0,t]$. Therefore, by the dominated convergence theorem we have

$$\begin{aligned}
&\int_0^{t\wedge\tau(\mathtt{w})} \beta_i\left(s, Y_s^\tau(\mathtt{w})\right) ds \\
&= \lim_{m\to\infty} \int_0^{t\wedge\tau(\mathtt{w})} \beta_i^{(m)}(s) ds \\
&= \lim_{m\to\infty} \sum_{\nu=1}^{k_m} \beta_i\left(t_{\nu-1}^m, Y_{t_{\nu-1}^m}^\tau(\mathtt{w})\right)\left(t_\nu^m \wedge \tau(\mathtt{w}) - t_{\nu-1}^m \wedge \tau(\mathtt{w})\right).
\end{aligned}$$

As $\mathtt{w} \in \widehat{\mathbf{W}}$ was fixed arbitrarily, the following equation holds in the sense of pointwise convergence on $\widehat{\mathbf{W}}$:

$$\int_0^{t\wedge\tau} \beta_i\left(s, Y_s^\tau\right) ds = \lim_{m\to\infty} \sum_{\nu=1}^{k_m} \beta_i\left(t_{\nu-1}^m, Y_{t_{\nu-1}^m}^\tau\right)\left(t_\nu^m \wedge \tau - t_{\nu-1}^m \wedge \tau\right).$$

By exactly the same reasoning, we obtain

$$\int_0^{t\wedge\tilde\tau} \beta_i\left(s, X_s^{\tilde\tau}\right) ds = \lim_{m\to\infty} \sum_{\nu=1}^{k_m} \beta_i \left(t_{\nu-1}^m, X_{t_{\nu-1}^m}^{\tilde\tau}\right)\left(t_\nu^m \wedge \tilde\tau - t_{\nu-1}^m \wedge \tilde\tau\right).$$

In this case convergence is pointwise on Ω.

We now also fix $j \in \{1, \dots, n\}$. Since the process $(s, \mathtt{w}) \mapsto \alpha_{ij}(s, Y_s^\tau(\mathtt{w}))$ is in particular left continuous, we can approximate the stochastic integral by Stieltjes sums. Denoting stochastic convergence on $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}, P_{(X,B)})$ by $P_{(X,B)}$–lim, we have:

$$\int_0^t \alpha_{ij}(s, Y_s^\tau) dW_s^{(j)^\tau} = P_{(X,B)}\text{–}\lim_{m\to\infty} \sum_{\nu=1}^{k_m} \alpha_{ij}\left(t_{\nu-1}^m, Y_{t_{\nu-1}^m}^\tau\right)\left(W_{t_\nu^m}^{(j)^\tau} - W_{t_{\nu-1}^m}^{(j)^\tau}\right).$$

Analogously, we obtain:

$$\int_0^t \alpha_{ij}(s, X_s^{\tilde\tau}) dB_s^{(j)^{\tilde\tau}} = P\text{–}\lim_{m\to\infty} \sum_{\nu=1}^{k_m} \alpha_{ij}\left(t_{\nu-1}^m, X_{t_{\nu-1}^m}^{\tilde\tau}\right)\left(B_{t_\nu^m}^{(j)^{\tilde\tau}} - B_{t_{\nu-1}^m}^{(j)^{\tilde\tau}}\right).$$

For every $m \in \mathbb{N}$, we now define random variables $\Psi^{(m)}$ and $\Phi^{(m)}$ as follows:

$$\begin{aligned}
\Psi^{(m)} &:= \sum_{j=1}^n \sum_{\nu=1}^{k_m} \alpha_{ij}\left(t_{\nu-1}^m, Y_{t_{\nu-1}^m}^\tau\right)\left(W_{t_\nu^m}^{(j)^\tau} - W_{t_{\nu-1}^m}^{(j)^\tau}\right) + \sum_{\nu=1}^{k_m} \beta_i\left(t_{\nu-1}^m, Y_{t_{\nu-1}^m}^\tau\right)\left(t_\nu^m \wedge \tau - t_{\nu-1}^m \wedge \tau\right) \\
&\quad + Y_0^{(i)} - Y_t^{(i)^\tau} \\
\Phi^{(m)} &:= \sum_{j=1}^n \sum_{\nu=1}^{k_m} \alpha_{ij}\left(t_{\nu-1}^m, X_{t_{\nu-1}^m}^{\tilde\tau}\right)\left(B_{t_\nu^m}^{(j)^{\tilde\tau}} - B_{t_{\nu-1}^m}^{(j)^{\tilde\tau}}\right) + \sum_{\nu=1}^{k_m} \beta_i\left(t_{\nu-1}^m, X_{t_{\nu-1}^m}^{\tilde\tau}\right)\left(t_\nu^m \wedge \tilde\tau - t_{\nu-1}^m \wedge \tilde\tau\right) \\
&\quad + X_0^{(i)} - X_t^{(i)^{\tilde\tau}}.
\end{aligned}$$

From the arguments offered above, it follows that

$$\begin{aligned}
\Psi &= P_{(X,B)}\text{–}\lim_{m\to\infty} \Psi^{(m)} \\
\Phi &= P\text{–}\lim_{m\to\infty} \Phi^{(m)}.
\end{aligned}$$

For $m \in \mathbb{N}$, we let μ_m denote the distribution of $\Psi^{(m)}$ under $P_{(X,B)}$ and denote the distribution of Ψ under $P_{(X,B)}$ by μ. The stochastic convergence of $(\Psi^{(m)})_{m\in\mathbb{N}}$ to Ψ implies the weak convergence of $(\mu_m)_{m\in\mathbb{N}}$ to μ. From the definition of Y, W and $\tilde\tau$ it follows that

$$\forall m \in \mathbb{N}: \quad \Psi^{(m)} \circ (X, B) = \Phi^{(m)}.$$

Therefore, for every $m \in \mathbb{N}$ the distribution of $\Phi^{(m)}$ under P is just μ_m. This shows that the sequence $(\mu_m)_{m\in\mathbb{N}}$ converges to δ_0 , the distribution of Φ under P. As the limit of a weakly convergent sequence of probability measures is uniquely determined, we conclude that $\mu = \delta_0$. □

A.2. **Constructing Solutions with Deterministic Initial Conditions from (Y, W).**

A.1. Definition. **(Regular factorized conditional probability)**: Let $(\Omega, \mathcal{F}, P)$ be a probability space and Z be a random variable with values in a measure space (S, Σ). A regular factorized conditional probability for P given Z is a Markov kernel K from (S, Σ) to $(\Omega, \mathcal{F})$ so that for every $A \in \mathcal{F}$ we have:

$$P[A|Z] = K(\cdot, A) \circ Z \quad P\text{–a.s.}$$

The following theorem ensures the existence of regular factorized conditional probabilities (cf. Bauer (1991)).

A.2. THEOREM. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, so that Ω is a Polish space and $\mathcal{F}$ is the Borel σ–algebra of Ω. Let Z be a random variable with values in a measure space (S, Σ). Then there exists a regular factorized conditional probability for P given Z.*

A.3. REMARK. 1. Let $(\Omega, \mathcal{F}, P)$ be a probability space and Z a random variable with values in a measure space (S, Σ). Suppose a regular factorized conditional probability for P given Z exists, then we denote it by $\{P^x\}_{x\in S}$, so that $\{P^x\}_{x\in S}$ is a family of probability measures on $(\Omega, \mathcal{F})$ and the Markov kernel is in fact given by the mapping $(x, A) \mapsto P^x[A]$. Let μ denote the distribution of Z under P. For every $A \in \mathcal{F}$ we have

$$P[A] = \int_\Omega P[A|Z]dP = \int_\Omega P^{Z(\omega)}[A]\ P(d\omega) = \int_S P^x[A]\mu(dx).$$

In particular, if $P[A] = 0$, there is a null set Λ in (S, Σ, μ), so that the following is true:

$$\forall x \in \Lambda^c : \quad P^x[A] = 0.$$

2. If the σ–algebra Σ is countably generated, one can show the existence of a null set Λ in (S, Σ, μ), so that for every $x \in \Lambda^c$ the following holds:

$$\forall C \in \Sigma : \quad P^x[Z \in C] = \chi_C(x).$$

In particular, if $x \in \Lambda^c$ and $\{x\} \in \Sigma$ we have

$$Z = x \quad P^x\text{–a.s.}$$

We now return to the space $\widehat{\mathbf{W}}$ and our solution (Y, W) of the SDE determined by α and β. Since $(\widehat{\mathbf{W}}, \mathfrak{W})$ is a Polish space, we can find a regular factorized conditional probability for $P_{(X,B)}$ on $(\widehat{\mathbf{W}}, \mathfrak{W})$ given Y_0, which we denote by $\{P^x_{(X,B)}\}_{x\in\mathcal{U}}$. For each $x \in \mathcal{U}$, we denote by $(\widehat{\mathbf{W}}, \mathfrak{W}^x, \mathbb{W}^{*,x} := \{\mathfrak{W}^{*,x}_t\}_{t\in\mathbb{R}_+}, P^x_{(X,B)})$ the usual augmentation of $(\widehat{\mathbf{W}}, \mathfrak{W}, \mathbb{W}, P^x_{(X,B)})$. We let μ denote the distribution of Y_0 on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ under $P_{(X,B)}$.
The remainder of this section is devoted to proving the following result:

A.4. THEOREM. *There is a null set Λ in $(\mathcal{U}, \mathcal{B}(\mathcal{U}), \mu)$ so that for every $x \in \Lambda^c$ the process W is a normal, n–dimensional $(P^x_{(X,B)}, \mathbb{W}^{*,x})$–Brownian motion and furthermore the process Y is a solution of*

$$dZ_t = \alpha(t, Z_t)dW_t + \beta(t, Z_t)dt$$

on the stochastic basis $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}^x, \mathbb{W}^{*,x}, P^x_{(X,B)})$ fulfilling the initial condition

$$Y_0 = x \quad P^x_{(X,B)}\text{–a.s.}$$

We first obtain several partial results.

A.5. LEMMA. *There is a null set Λ_1 in $(\mathcal{U}, \mathcal{B}(\mathcal{U}), \mu)$, so that for every $x \in \Lambda_1^c$ the process W is a normal, n–dimensional $(P^x_{(X,B)}, \mathbb{W}^{*,x})$–Brownian motion.*

PROOF: It suffices to find a null set Λ_1 in $(\mathcal{U}, \mathcal{B}(\mathcal{U}), \mu)$, so that for every $x \in \Lambda_1^c$ W is a $(P^x_{(X,B)}, \mathbb{W})$–Brownian motion. Since W is a $(P_{(X,B)}, \mathbb{W})$–Brownian motion, we have for every $y \in \mathbb{R}^n$ and all $s, t \in \mathbb{R}_+$ with $s < t$:

$$E\left[e^{i\langle y, W_t - W_s\rangle}\middle|\, \mathfrak{W}_s\right] = e^{-\frac{1}{2}\|y\|^2(t-s)}\ P_{(X,B)}\text{–a.s.}$$

This implies that for every $A \in \mathfrak{W}_s$ and every $C \in \mathcal{B}(\mathcal{U})$ we have:

$$\int_{A\cap\{Y_0\in C\}} e^{i\langle y,W_t-W_s\rangle} dP_{(X,B)} = e^{-\frac{1}{2}\|y\|^2(t-s)} P_{(X,B)}\left[A\cap Y_0^{-1}(C)\right].$$

Using the regular factorized conditional probability $\{P^x_{(X,B)}\}_{x\in\mathcal{U}}$ we can rewrite this equation as follows:

$$\int_C\int_{\widehat{\mathbf{W}}} \chi_A\, e^{i\langle y,W_t-W_s\rangle} dP^x_{(X,B)}\mu(dx) = e^{-\frac{1}{2}\|y\|^2(t-s)}\int_C P^x_{(X,B)}(A)\mu(dx).$$

Since C was an arbitrary set in $\mathcal{B}(\mathcal{U})$, there is a null set $\Lambda_{y,s,t,A}$ in $(\mathcal{U},\mathcal{B}(\mathcal{U}),\mu)$, so that for $x\in\Lambda^c_{y,s,t,A}$ we have:

$$\int_{\widehat{\mathbf{W}}} \chi_A e^{i\langle y,W_t-W_s\rangle} dP^x_{(X,B)} = e^{-\frac{1}{2}\|y\|^2(t-s)} P^x_{(X,B)}(A).$$

The σ–algebra $\mathfrak{W}_s$ possesses a countable generator $\mathcal{E}_s$, by replacing $\mathcal{E}_s$ with the algebra it generates, we can assume that $\widehat{\mathbf{W}}\in\mathcal{E}_s$ and that $\mathcal{E}_s$ is stable under intersections. We define

$$\Lambda_{y,s,t} := \bigcup_{E\in\mathcal{E}_s}\Lambda_{y,s,t,E}.$$

Obviously $\mu(\Lambda_{y,s,t})=0$. Now fix $x\in\Lambda^c_{y,s,t}$. The system

$$\mathcal{D} := \left\{A\in\mathfrak{W}_s \;\middle|\; \int_{\widehat{\mathbf{W}}}\chi_A e^{i\langle y,W_t-W_s\rangle} dP^x_{(X,B)} = e^{-\frac{1}{2}\|y\|^2(t-s)}P^x_{(X,B)}(A)\right\}$$

is a Dynkin–system containing $\mathcal{E}_s$, therefore $\mathcal{D}=\mathfrak{W}_s$. To recapitulate, we have shown:

$$\forall x\in\Lambda^c_{y,s,t}\quad \forall A\in\mathfrak{W}_s:\quad \int_{\widehat{\mathbf{W}}}\chi_A e^{i\langle y,W_t-W_s\rangle} dP^x_{(X,B)} = e^{-\frac{1}{2}\|y\|^2(t-s)}P^x_{(X,B)}(A).$$

We set

$$\tilde{\Lambda}_1 := \bigcup_{\substack{u,v\in\mathbb{Q}_+,u<v\\ p\in\mathbb{Q}^n}}\Lambda_{p,u,v}.$$

Suppose that $s,t\in\mathbb{R}_+$, $s<t$, $y\in\mathbb{R}^n$ are given. We choose sequences (u_k), (v_k) in $\mathbb{Q}_+$ and (p_k) in $\mathbb{Q}^n$, so that $(u_k)\downarrow s$, $(v_k)\uparrow t$, $(p_k)\to y$ and for every $k\in\mathbb{N}$ we have $u_k<v_k$. Let $x\in\tilde{\Lambda}^c_1$ and $A\in\mathfrak{W}_s$ be arbitrary. By the construction of $\tilde{\Lambda}_1$, the following equation holds for every $k\in\mathbb{N}$:

$$\int_{\widehat{\mathbf{W}}}\chi_A e^{i\langle p_k,W_{v_k}-W_{u_k}\rangle} dP^x_{(X,B)} = e^{-\frac{1}{2}\|p_k\|^2(v_k-u_k)}P^x_{(X,B)}(A).$$

By the dominated convergence theorem it follows that

$$\int_{\widehat{\mathbf{W}}}\chi_A e^{i\langle y,W_t-W_s\rangle} dP^x_{(X,B)} = \lim_{k\to\infty} e^{-\frac{1}{2}\|p_k\|^2(v_k-u_k)}P^x_{(X,B)}(A) = e^{-\frac{1}{2}\|y\|^2(t-s)}P^x_{(X,B)}(A).$$

Since A was arbitrary we have shown:

$$\forall x\in\tilde{\Lambda}^c_1\quad \forall s,t\in\mathbb{R}_+, s<t, y\in\mathbb{R}^n:\quad E\left[e^{i\langle y,W_t-W_s\rangle}\middle|\,\mathfrak{W}_s\right] = e^{-\frac{1}{2}\|y\|^2(t-s)}\; P^x_{(X,B)}\text{–a.s.}$$

This proves that for every $x\in\tilde{\Lambda}^c_1$, the process W is a $(P^x_{(X,B)},\mathbb{W})$–Brownian motion. Using remark A.3, it is now trivial to enlarge $\tilde{\Lambda}_1$ slightly so as to ensure that $W_0=0$ $P^x_{(X,B)}$–a.s. $\square$

To prove theorem A.4, we will use lemma A.2. To this end we need to fix an announcing sequence for the explosion time e_Y of Y. A natural choice is the following one. We can find a sequence $(\mathcal{U}_k)$ of open subsets of $\mathcal{U}$ with $(\mathcal{U}_k) \uparrow \mathcal{U}$, and so that for every $k \in \mathbb{N}$, $\mathcal{U}_k$ is relatively compact with $\overline{\mathcal{U}}_k \subset \mathcal{U}_{k+1}$. We define for every $k \in \mathbb{N}$:

$$\sigma_k := \inf \left\{ t \in \mathbb{R}_+ \,\middle|\, Y_t \in \widehat{\mathcal{U}} \backslash \mathcal{U}_k \right\}.$$

Since Y is continuous and $\mathbb{W}$–adapted, and $\widehat{\mathcal{U}} \backslash \mathcal{U}_k$ is a closed subset of the metric space $\widehat{\mathcal{U}}$, σ_k is a $\mathbb{W}$–stopping time for each $k \in \mathbb{N}$. One easily checks that $(\sigma_k)_{k \in \mathbb{N}}$ is indeed an announcing sequence for e_Y. Since σ_k is a $\mathbb{W}$–stopping time, it is automatically a $\mathbb{W}^{*,x}$–stopping time for every $x \in \mathcal{U}$.

A.6. Lemma. *Fix $i \in \{1, \ldots, d\}$, $j \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$. Let $I = (I_t)_{t \in \mathbb{R}_+}$ be a fixed version of the stochastic integral $(\int_0^t \alpha_{ij}(s, Y_s^{\sigma_k}) dW_s^{(j)^{\sigma_k}})_{t \in \mathbb{R}_+}$, where the stochastic integral refers to the stochastic basis $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}, \mathbb{W}^*, P_{(X,B)})$. For each $x \in \Lambda_1^c$ let $I^x = (I_t^x)_{t \in \mathbb{R}_+}$ be a fixed version of the same integral, now taken with respect to the stochastic basis $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}^x, \mathbb{W}^{*,x}, P_{(X,B)}^x)$. Then there exists a null set Λ_2 in $(\mathcal{U}, \mathcal{B}(\mathcal{U}), \mu)$, so that for every $x \in (\Lambda_1 \cup \Lambda_2)^c$ the two processes I and I^x are $P_{(X,B)}^x$–indistinguishable.*

Proof: Since we are dealing with continuous processes, it suffices to show that for every $t \in \mathbb{R}_+$ there exists a null set $\Lambda_{2,t}$ in $(\mathcal{U}, \mathcal{B}(\mathcal{U}), \mu)$, so that the following holds:

$$\forall x \in (\Lambda_1 \cup \Lambda_{2,t})^c : \quad I_t = I_t^x \quad P_{(X,B)}^x\text{–a.s.}$$

We fix $t \in \mathbb{R}_+$ and choose a sequence $(\mathcal{Z}_m)_{m \in \mathbb{N}}$ of partitions $\mathcal{Z}_m : 0 = t_0^m < \ldots < t_{l_m}^m = t$ of $[0, t]$ with $|\mathcal{Z}_m| \to 0$. For each $m \in \mathbb{N}$ we define

$$\Phi_m := \sum_{\nu=1}^{l_m} \alpha_{ij} \left(t_{\nu-1}^m, Y_{t_{\nu-1}^m}^{\sigma_k} \right) \left(W_{t_\nu^m}^{(j)^{\sigma_k}} - W_{t_{\nu-1}^m}^{(j)^{\sigma_k}} \right).$$

The sequence $(\Phi_m)_{m \in \mathbb{N}}$ converges $P_{(X,B)}$–stochastically to I_t, this is just the approximation of the stochastic integral by Stieltjes sums. Choosing a subsequence if necessary, we can assume that the sequence $(\Phi_m)_{m \in \mathbb{N}}$ converges $P_{(X,B)}$–a.s. to I_t . In particular, there is a set $N \in \mathfrak{W}$ with $P_{(X,B)}[N] = 0$, so that

$$\forall \omega \in N^c : \quad I_t(\omega) = \lim_{m \to \infty} \Phi_m(\omega).$$

Again by remark A.3, there is a null set $\Lambda_{2,t}$ in $(\mathcal{U}, \mathcal{B}(\mathcal{U}), \mu)$, so that

$$\forall x \in \Lambda_{2,t}^c : \quad P_{(X,B)}^x[N] = 0.$$

This shows that for $x \in \Lambda_{2,t}^c$ the sequence (Φ_m) converges $P_{(X,B)}^x$–almost surely to I_t, but for $x \in \Lambda_1^c$ we already know that it converges $P_{(X,B)}^x$–stochastically to I_t^x. Therefore we have

$$\forall x \in (\Lambda_1 \cup \Lambda_{2,t})^c : \quad I_t = I_t^x \quad P_{(X,B)}^x\text{–a.s.}$$

□

Proof of Theorem A.4: By remark A.3, there is a μ–null set Λ_0, so that for $x \in \Lambda_0^c$ the two processes Y and Y^{e_Y} are $P_{(X,B)}^x$–indistinguishable and $Y_0 = x$ $P_{(X,B)}^x$–a.s. For $i \in \{1, \ldots, d\}$, $j \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$, let $I^{(i,j,k)} = (I_t^{(i,j,k)})_{t \in \mathbb{R}_+}$ be a fixed version of the stochastic integral $(\int_0^t \alpha_{ij}(s, Y^{\sigma_k}) dW_s^{(j)^{\sigma_K}})_{t \in \mathbb{R}_+}$ computed with respect to

$(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}, \mathbb{W}^*, P_{(X,B)})$. For each $x \in \Lambda_1^c$, let $I^{(i,j,k,x)}$ be a fixed version of the same integral computed with respect to $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}^x, \mathbb{W}^{*,x}, P^x_{(X,B)})$. As Y is a solution of

$$dZ_t = \alpha(t, Y_t) dW_t$$

on $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}, \mathbb{W}^*, P_{(X,B)})$, there is a $P_{(X,B)}$–null set $N \in \mathfrak{W}$, so that for every $i \in \{1, \dots, d\}$, $k \in \mathbb{N}$ and $t \in \mathbb{R}_+$ we have:

$$\forall \omega \in N^c : \quad Y_t^{(i)^{\sigma_k}}(\omega) = Y_0^{(i)}(\omega) + \sum_{j=1}^n I_t^{(i,j,k)}(\omega) + \int_0^{t\wedge\sigma_k(\omega)} \beta_i\left(s, Y_s^{\sigma_k}(\omega)\right) ds.$$

There is a μ–null set Λ_3, so that

$$\forall x \in \Lambda_3^c : \quad P^x_{(X,B)}[N] = 0.$$

We therefore have

$$\forall x \in \Lambda_3^c \quad \forall i \in \{1, \dots, d\}, k \in \mathbb{N}, t \in \mathbb{R}_+ :$$
$$Y_t^{(i)^{\sigma_k}} = Y_0^{(i)} + \sum_{j=1}^n I_t^{(i,j,k)} + \int_0^{t\wedge\sigma_k} \beta_i\left(s, Y_s^{\sigma_k}\right) ds \quad P^x_{(X,B)}\text{–a.s.}$$

By Lemma A.6, there is a μ–null set $\widehat{\Lambda}_2$, so that for $x \in (\Lambda_1 \cup \widehat{\Lambda}_2)^c$ and any choice of $i \in \{1, \dots, d\}$, $j \in \{1, \dots, n\}$, $k \in \mathbb{N}$ the two processes $I^{(i,j,k)}$ and $I^{(i,j,k,x)}$ are indistinguishable. This implies

$$\forall x \in \left(\Lambda_1 \cup \widehat{\Lambda}_2 \cup \Lambda_3\right)^c \quad \forall i \in \{1, \dots, d\}, k \in \mathbb{N}, t \in \mathbb{R}_+ :$$
$$Y_t^{(i)^{\sigma_k}} = Y_0^{(i)} + \sum_{j=1}^n I_t^{(i,j,k,x)} + \int_0^{t\wedge\sigma_k} \beta_i\left(s, Y_s^{\sigma_k}\right) ds \quad P^x_{(X,B)}\text{–a.s}$$

Setting $\Lambda := \Lambda_0 \cup \Lambda_1 \cup \widehat{\Lambda}_2 \cup \Lambda_3$ and using Lemma A.2, we see that for $x \in \Lambda^c$ the process Y is a solution of $dZ_t = \alpha(t, Z_t)dW_t + \beta(t, Z_t)dt$ on the stochastic basis $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}^x, \mathbb{W}^{*,x}, P^x_{(X,B)})$ with the deterministic initial condition $Y_0 = x$ $P^x_{(X,B)}$–a.s.

□

A.3. **Concluding Proof of Theorem A.3:** We return to our original solution (X, B) on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P)$. As we have already shown

$$P\left[e_X < +\infty\right] = P_{(X,B)}\left[e_Y < +\infty\right].$$

Therefore we need only show that $P_{(X,B)}[e_Y < +\infty] = 0$. By theorem A.4 there exists a μ–null set Λ, so that for $x \in \Lambda^c$ the process W is a standard n–dimensional Brownian motion on $(\widehat{\mathbf{W}}, \overline{\mathfrak{W}}^x, \mathbb{W}^{*,x}, P^x_{(X,B)})$ and Y is a solution of our SDE with $Y_0 = x$ $P_{(X,B)}$–a.s. By assumption, solutions with a deterministic initial condition do not explode, therefore

$$\forall x \in \Lambda^c : \quad P^x_{(X,B)}\left[e_Y < +\infty\right] = 0.$$

Now we have

$$P_{(X,B)}\left[e_Y < +\infty\right] = \int_{\mathcal{U}} P^x_{(X,B)}\left[e_Y < +\infty\right] \mu(dx) = 0.$$

With this, the desired result is finally proved. □

Appendix B. Proof of theorem 2.1.1

We proceed inductively. We define $r^{(0)} : \{0\} \times \Omega \to]0; +\infty[$ by setting $r^{(0)}(0, \omega) := f(\omega)$. Now let $i \in \{1, \ldots, n\}$ be given. Suppose that we have already constructed a continuous, $\mathbb{F}$–adapted process $r^{(i-1)} : [0; T_{i-1}] \times \Omega \to]0; +\infty[$, so that for every $t \in [0; T_{i-1}]$ the following holds:

$$r^{(i-1)}(t) = f + \int_0^t \alpha\left(s, r^{(i-1)}(s)\right) dW(s) + \int_0^t \beta\left(s, r^{(i-1)}(s)\right) ds \quad Q\text{–a.s.}$$

In the case of $i = 1$ this equation is trivial. We now introduce a deterministic time change $\langle\tau\rangle = \{\tau_t\}_{t \in \mathbb{R}_+}$ by setting $\tau_t := t + T_{i-1}$ for every $t \in \mathbb{R}_+$. We define a new filtration $\mathbb{G} = \{\mathcal{G}_t\}_{t \in \mathbb{R}_+}$ by setting $\mathcal{G}_t = \mathcal{F}_{\tau_t} = \mathcal{F}_{t+T_{i-1}}$ and define a process $\tilde{W} = \left(\tilde{W}(t)\right)_{t \in \mathbb{R}_+}$ via $\tilde{W}(t) := W(t + T_{i-1}) - W(T_{i-1})$. The stochastic basis $(\Omega, \mathcal{F}, \mathbb{G}, Q)$ also fulfills the usual hypotheses and it is easy to see that $\tilde{W}$ is a standard $(Q, \mathbb{G})$–Brownian motion. The random variable $r^{(i-1)}(T_{i-1})$ is $\mathcal{G}_0$–measurable. From our above remarks it follows that there is a continuous, $\mathbb{G}$–adapted process y with values in $]0; +\infty[$ so that $r(0) = r^{(i-1)}(T_{i-1})$ and

$$dr(t) = (\theta_i - a_i r(t))dt + \sigma_i \sqrt{r(t)} d\tilde{W}(t).$$

We define the process $r^{(i)} : [0; T_i] \times \Omega \to]0; +\infty[$ as follows:

$$r^{(i)}(t, \omega) := \begin{cases} r^{(i-1)}(t, \omega) & \text{if } t \in [0; T_{i-1}], \\ r(t - T_{i-1}, \omega) & \text{if } t \in]T_{i-1}; T_i]. \end{cases} \tag{B.1}$$

The process $r^{(i)}$ is $\mathbb{F}$–adapted and possesses continuous paths. Obviously for $t \in [0; T_{i-1}]$ we have

$$r^{(i)}(t) = f + \int_0^t \alpha\left(s, r^{(i)}(s)\right) dW(s) + \int_0^t \beta\left(s, r^{(i)}(s)\right) ds \quad Q\text{–a.s.} \tag{B.2}$$

Now assume $t \in]T_{i-1}; T_i]$. Then:

$$\begin{aligned} & f + \int_0^t \beta\left(s, r^{(i)}(s)\right) ds + \int_0^t \alpha\left(s, r^{(i)}(s)\right) dW(s) \\ = \; & r^{(i)}(T_{i-1}) + \int_{T_{i-1}}^t \beta\left(s, r^{(i)}(s)\right) ds + \int_{T_{i-1}}^t \alpha\left(s, r^{(i)}(s)\right) dW(s) \quad Q\text{–a.s.} \end{aligned}$$

We will now write the right hand side of this equation somewhat differently. First of all:

$$\begin{aligned} \int_{T_{i-1}}^t \beta\left(s, r^{(i)}(s)\right) ds &= \int_{T_{i-1}}^t \left(\theta_i - a_i r^{(i)}(s)\right) ds \\ &= \int_0^{t-T_{i-1}} \left(\theta_i - a_i r^{(i)}(s + T_{i-1})\right) ds \\ &= \int_0^{t-T_{i-1}} (\theta_i - a_i r(s))\, ds. \end{aligned}$$

Furthermore:

$$\int_{T_{i-1}}^t \alpha\left(s, r^{(i)}(s)\right) dW(s) = \int_{T_{i-1}}^t \sigma_i \sqrt{r^{(i)}(s)} dW(s) \quad Q\text{–a.s.}$$

Denoting stochastic integration by a $\bullet$, the transformation property of the stochastic integral under time change implies the following identity up to indistinguishability

$$\left(\sigma_i\sqrt{r^{(i)}} \bullet W\right)_{\langle\tau\rangle} - \left(\sigma_i\sqrt{r^{(i)}} \bullet W\right)_{\tau_0} = \sigma_i\sqrt{r^{(i)}_{\langle\tau\rangle}} \bullet W_{\langle\tau\rangle}.$$

The stochastic integral on the right hand side of this equation is taken with respect to the stochastic basis $(\Omega, \mathcal{F}, \mathbb{G}, Q)$. Explicitly, we have for $t \in]T_{i-1}; T_i]$:

$$\begin{aligned} & \int_{T_{i-1}}^{t} \sigma_i\sqrt{r^{(i)}(s)}dW(s) \\ = & \int_0^{t-T_{i-1}} \sigma_i\sqrt{r^{(i)}(s+T_{i-1})}d\tilde{W}(s) \\ = & \int_0^{t-T_{i-1}} \sigma_i\sqrt{r(s)}d\tilde{W}(s) \quad Q\text{–a.s.} \end{aligned}$$

Therefore

$$\begin{aligned} & f + \int_0^t \beta\left(s, r^{(i)}(s)\right) ds + \int_0^t \alpha\left(s, r^{(i)}(s)\right) dW(s) \\ = \;& r^{(i)}(T_{i-1}) + \int_{T_{i-1}}^t \beta\left(s, r^{(i)}(s)\right) ds + \int_{T_{i-1}}^t \alpha\left(s, r^{(i)}(s)\right) dW(s) \\ = \;& r^{(i)}(T_{i-1}) + \int_0^{t-T_{i-1}} (\theta_i - a_i r(s))\, ds + \int_0^{t-T_{i-1}} \sigma_i\sqrt{r(s)}d\tilde{W}(s) \\ = \;& r(t-T_{i-1}) = r^{(i)}(t) \quad Q\text{–a.s.} \end{aligned}$$

We have now shown that the process $r^{(i)}$ satisfies (B.2) for all $t \in [0; T_i]$. After N such induction steps we obtain a solution of the equation (2) on $[0; T_N]$. One more induction step (with a trivial change of notation) allows us to extend the solution to all of $\mathbb{R}_+$. □

Appendix C. Lemma for Calculations Involving the Noncentral χ^2 Distribution

C.1. Lemma. *Denote the density function of a noncentral chi–square distribution with ν degrees of freedom and noncentrality parameter λ by $q_{\chi^2}(\cdot, \nu, \lambda)$. Then for $b, r > 0$ and an arbitrary constant L the following holds:*

$$e^{-rL} q_{\chi^2}(b\,r, \nu, \lambda) = \exp\left\{-\frac{L}{b+2L}\lambda\right\} \left(\frac{b}{b+2L}\right)^{\frac{1}{2}\nu-1} q_{\chi^2}\left((b+2L)r, \nu, \frac{\lambda b}{b+2L}\right)$$

PROOF: Substituting for q_{χ^2} its infinite sum expression (see Johnson and Kotz (1970b), Chapter 28, eq. 3), we get

$$\begin{aligned}
& e^{-rL}q_{\chi^2}(b\ r,\nu,\lambda) \\
= \ & e^{-rL}2^{-\frac{1}{2}\nu}\exp\left\{-\frac{1}{2}(b\ r+\lambda)\right\}\sum_{j=0}^{\infty}\frac{(b\ r)^{\frac{1}{2}\nu+j-1}\lambda^j}{\Gamma(\frac{1}{2}\nu+j)2^{2j}j!} \\
= \ & 2^{-\frac{1}{2}\nu}\exp\left\{-\frac{1}{2}\left(r(b+2L)+\frac{\lambda b}{b+2L}\right)\right\}\exp\left\{\frac{1}{2}\left(\frac{\lambda b}{b+2L}-\lambda\right)\right\}\left(\frac{b}{b+2L}\right)^{\frac{1}{2}\nu-1} \\
& \cdot\sum_{j=0}^{\infty}\frac{((b+2L)r)^{\frac{1}{2}\nu+j-1}\left(\frac{\lambda b}{b+2L}\right)^j}{\Gamma(\frac{1}{2}\nu+j)2^{2j}j!} \\
= \ & \exp\left\{\frac{1}{2}\left(\frac{\lambda b}{b+2L}-\lambda\right)\right\}\left(\frac{b}{b+2L}\right)^{\frac{1}{2}\nu-1}q_{\chi^2}\left((b+2L)r,\nu,\frac{\lambda b}{b+2L}\right)
\end{aligned}$$

□

C.2. REMARK. If one interprets q_{χ^2} not as a density, but as a function with complex arguments, then it is easy to see that lemma C.1 is also valid for complex L.

APPENDIX D. MULTIPLE COMPOUND χ^2 DISTRIBUTION

Let b_r, b_x, η_r, and λ be strictly positive constants and $\nu_r \geq 1$, $\nu_x \geq 1$. Let r and x be strictly positive random variables such that $b_x x$ has a χ^2–distribution with ν_x degrees of freedom and noncentrality parameter λ. Furthermore let $b_r r$ — given that $b_x x = z_x$ — have a χ^2–distribution with ν_r degrees of freedom and noncentrality parameter $\eta_r x = \eta_r b_x^{-1} z_x$. The joint distribution of $b_r r$ and $b_x x$ is given by the probability density function

$$p(b_r r = z_r,\ b_x x = z_x) = p(b_r r = z_r \mid b_x x = z_x)\cdot p(b_x x = z_x).$$

We are interested in the marginal distribution of $b_r r$, which is given by integrating over z_x:

$$p(b_r r = z_r) = \int_0^\infty p(b_r r = z_r|b_x x = z_x)p(b_x x = z_x)dz_x.$$

Both $p(b_r r = z_r|b_x x = z_x)$ and $p(b_x x = z_x)$ are noncentral χ^2 density functions. Writing these as a mixture of central χ^2 probability density functions, we have[8]

$$\begin{aligned}
p(b_r r = z_r) \ &= \ \int_0^\infty\left(\sum_{j=0}^{\infty}\frac{\left(\frac{1}{2}\eta_r b_x^{-1}z_x\right)^j}{j!}\exp\left\{-\frac{1}{2}\eta_r b_x^{-1}z_x\right\}p_{\chi^2_{\nu_r+2j}}(z_r)\right) \\
& \quad \cdot\left(\sum_{j=0}^{\infty}\frac{\left(\frac{1}{2}\lambda\right)^j}{j!}\exp\left\{-\frac{1}{2}\lambda\right\}p_{\chi^2_{\nu_x+2j}}(z_x)\right)dz_x
\end{aligned}$$

$$\begin{aligned}
\text{(D.1)} \qquad &= \ \int_0^\infty\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}\eta_r b_x^{-1}\right)^j}{j!}\frac{\left(\frac{1}{2}\lambda\right)^k}{k!}\exp\left\{-\frac{1}{2}\lambda\right\}p_{\chi^2_{\nu_r+2j}}(z_r) \\
& \quad z_x^j\exp\left\{-\frac{1}{2}\eta_r b_x^{-1}z_x\right\}p_{\chi^2_{\nu_x+2k}}(z_x)dz_x\ .
\end{aligned}$$

[8]See Johnson and Kotz (1970b), chapter 28, eq. (3).

where $p_{\chi^2_m}$ denotes the probability density function of a central χ^2–distribution with m degrees of freedom.

Applying lemma C.1 to (D.1) and interchanging integration and addition, we get

$$\text{(D.2)}\; p(b_r r = z_r) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\eta_r b_x^{-1}\right)^j}{j!}\frac{\left(\frac{1}{2}\lambda\right)^k}{k!} \exp\left\{-\frac{1}{2}\lambda\right\} p_{\chi^2_{\nu_r+2j}}(z_r)$$
$$\int_0^\infty z_x^j \left(\frac{b_x}{b_x+\eta_r}\right)^{\frac{1}{2}(\nu_x+2k)} p_{\chi^2_{\nu_x+2k}}\left(\frac{b_x+\eta_r}{b_x}z_x\right) d\left(\frac{b_x+\eta_r}{b_x}z_x\right).$$

The integral (D.2) can be expressed in terms of the j-th moment about zero of a central χ^2 distribution with $\nu_x + 2k$ degrees of freedom[9]:

$$\left(\frac{b_x}{b_x+\eta_r}\right)^{\frac{1}{2}\nu_x+k+j} 2^j \frac{\Gamma(\frac{1}{2}(\nu_x+2k)+j)}{\Gamma(\frac{1}{2}(\nu_x+2k))}.$$

D.1. Definition. Let strictly positive random variables $r_1, \dots, r_n$ as well as strictly positive constants r_0, $b_1, \dots, b_n$, $\eta_1, \dots, \eta_n$, $\nu_1, \dots, \nu_n$ be given. Suppose that for each $j \in \{1; \dots; n\}$ the random variable $(b_j r_j)$ conditioned on r_{j-1} is noncentral χ^2 distributed with ν_j degrees of freedom and noncentrality parameter $\eta_j r_{j-1}$. Then we call r_n *n times multiple compound noncentral χ^2 distributed with transformation coefficients b_j.*

D.2. Lemma. *The probability density function $p(b_n \cdot r_n = z_n)$ of the multiple compound noncentral χ^2 distributed random variable $(b_n r_n)$ is*

$$p(b_n\cdot r_n = z_n) = \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty} \frac{\left(\frac{1}{2}\eta_1 r_0\right)^{j_1}}{j_1!}\exp\left\{-\frac{1}{2}\eta_1 r_0\right\} p_{\chi^2_{\nu_n+2j_n}}(z_n)$$
$$\text{(D.3)}\quad \prod_{k=2}^{n} \frac{\left(\frac{1}{2}\eta_k b_{k-1}^{-1}\right)^{j_k}}{j_k!}\left(\frac{b_{k-1}}{b_{k-1}+\eta_k}\right)^{\frac{1}{2}(\nu_{k-1}+2j_{k-1})+j_k} 2^{j_k}\frac{\Gamma(\frac{1}{2}(\nu_{k-1}+2j_{k-1})+j_k)}{\Gamma(\frac{1}{2}(\nu_{k-1}+2j_{k-1}))}.$$

Proof: We prove the lemma by induction: For $n = 2$ we get (D.2). Now let (D.3) be valid for some n. For $n+1$ we have

$$p(b_{n+1}\cdot r_{n+1} = z_{n+1})$$
$$= \int_0^\infty p(b_{n+1}\cdot r_{n+1} = z_{n+1} \mid b_n r_n = z_n) p(b_n r_n = z_n) dz_n$$
$$= \int_0^\infty \left(\sum_{j_{n+1}=0}^{\infty}\frac{\left(\frac{1}{2}\eta_{n+1}b_n^{-1}z_n\right)^{j_{n+1}}}{j_{n+1}!}\exp\left\{-\frac{1}{2}\eta_{n+1}b_n^{-1}z_n\right\} p_{\chi^2_{\nu_{n+1}+2j_{n+1}}}(z_{n+1})\right)$$
$$\cdot\left(\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\frac{\left(\frac{1}{2}\eta_1 r_0\right)^{j_1}}{j_1!}\exp\left\{-\frac{1}{2}\eta_1 r_0\right\} p_{\chi^2_{\nu_n+2j_n}}(z_n)\right.$$
$$\left.\prod_{k=2}^{n}\frac{\left(\frac{1}{2}\eta_k b_{k-1}^{-1}\right)^{j_k}}{j_k!}\left(\frac{b_{k-1}}{b_{k-1}+\eta_k}\right)^{\frac{1}{2}(\nu_{k-1}+2j_{k-1})+j_k} 2^{j_k}\frac{\Gamma(\frac{1}{2}(\nu_{k-1}+2j_{k-1})+j_k)}{\Gamma(\frac{1}{2}(\nu_{k-1}+2j_{k-1}))}\right) dz_n$$

and analogously to ((D.1) $\Leftrightarrow$ (D.2)) we get (D.3) for $n+1$.

□

[9]See Johnson and Kotz (1970a), ch. 17, p. 168.

D.3. Remark. Since only the values of the central χ^2 densities in (D.3) depend on z_n, we can interchange addition and integration and write the multiple compound noncentral χ^2 distribution function as

$$P(b_n \cdot r_n \leq z_n) = \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty} \frac{\left(\frac{1}{2}\eta_1 r_0\right)^{j_1}}{j_1!} \exp\left\{-\frac{1}{2}\eta_1 r_0\right\} P_{\chi^2_{\nu_n+2j_n}}(z_n)$$

$$\prod_{k=2}^{n} \frac{\left(\frac{1}{2}\eta_k b_{k-1}^{-1}\right)^{j_k}}{j_k!} \left(\frac{b_{k-1}}{b_{k-1}+\eta_k}\right)^{\frac{1}{2}(\nu_{k-1}+2j_{k-1})+j_k} 2^{j_k} \frac{\Gamma(\frac{1}{2}(\nu_{k-1}+2j_{k-1})+j_k)}{\Gamma(\frac{1}{2}(\nu_{k-1}+2j_{k-1}))}. \tag{D.4}$$

Due to the nested infinite sums, calculating (D.4) appears to be of exponential complexity in n. Fortunately, this is not the case. In fact, were it not for the term $\Gamma(\frac{1}{2}\nu_{k-1}+j_{k-1}+j_k)$, it would be possible to separate the terms and calculate (D.4) as a product of n one–dimensional sums, obviously a problem of linear complexity in n. As it is, the number of operations necessary to determine $P(b_n \cdot r_n \leq z_n)$ still only increases linearly in n.
We start by calculating the terms

$$\frac{\left(\frac{1}{2}\eta_1 r_0\right)^{j_1}}{j_1!} \exp\left\{-\frac{1}{2}\eta_1 r_0\right\} \tag{D.5}$$

for all $j_1 \in J_1 := \{\underline{j_1};\ldots;\overline{j_1}\}$. Note that (D.5) is unimodal in j_1, therefore $\underline{j_1}$ and $\overline{j_1}$ can be chosen in such a manner that (D.5) is smaller than $\epsilon > 0$ for any $j_1 \notin J_1$. For each $j_1 \in J_1$ we then calculate

$$\frac{\left(\frac{1}{2}\eta_2 b_1^{-1}\right)^{j_2}}{j_2!} \left(\frac{b_1}{b_1+\eta_2}\right)^{\frac{1}{2}(\nu_1+2j_1)+j_2} 2^{j_2} \frac{\Gamma(\frac{1}{2}(\nu_1+2j_1)+j_2)}{\Gamma(\frac{1}{2}(\nu_1+2j_1))} \tag{D.6}$$

for all $j_2 \in J_2 := \{\underline{j_2};\ldots;\overline{j_2}\}$, where (D.6) is again unimodal in j_2. We multiply (D.6) with (D.5) for each $(j_1;j_2) \in J_1 \times J_2$ and then sum over j_1 for each j_2, reducing the index dimension to 1 again. Substituting the result for (D.5) and defining (D.6) analogously for $(j_2;j_3)$, we iterate until we reach j_n, yielding a value for each $j_n \in J_n$, which we multiply with the respective value of $P_{\chi^2_{\nu_n+2j_n}}(z_n)$ and sum one last time to get $P(b_n \cdot r_n \leq z_n)$.

D.4. Lemma. *The derivative of (D.4) with respect to r_0 is*

$$\frac{\partial}{\partial r_0} P(b_n r_n \leq z_n) = -\eta_n \left(\prod_{j=1}^{n-1} \eta_j b_j^{-1}\right) \tilde{p}(b_n r_n = z_n) \tag{D.7}$$

where $\tilde{p}(b_n r_n = z_n)$ is defined as $p(b_n r_n = z_n)$, however with $\tilde{\nu}_j := \nu_j + 2\ \forall\ 1 \leq j \leq n$, all other coefficients identical.

Proof: For $n = 1$, we have

$$\frac{\partial}{\partial r_0} p(b_n r_n = z_n) = \frac{\partial}{\partial r_0} q_{\chi^2}(z, \nu_1, \eta_1 r_0) = \eta_1 \frac{\partial}{\partial(\eta_1 r_0)} q_{\chi^2}(z, \nu_1, \eta_1 r_0)$$

which is equal to[10]

$$= -\eta_1 \frac{\partial}{\partial z} q_{\chi^2}(z, \nu_1 + 2, \eta_1 r_0)$$

Now let

$$\text{(D.8)} \qquad \frac{\partial}{\partial r_0} p(b_n r_n = z_n) = -\eta_n \left(\prod_{j=1}^{n-1} \eta_j b_j^{-1} \right) \frac{\partial}{\partial z_n} \tilde{p}(b_n r_n = z_n)$$

for some n. Then for $n+1$

$$\frac{\partial}{\partial r_0} p(b_{n+1} r_{n+1} = z_{n+1}) = \frac{\partial}{\partial r_0} \int_0^\infty q_{\chi^2}(z_{n+1}, \nu_{n+1}, \eta_{n+1} r_n) p(b_n r_n = z_n) dz_n$$

$$= \int_0^\infty q_{\chi^2}(z_{n+1}, \nu_{n+1}, \eta_{n+1} b_n^{-1} z_n) \left(-\eta_n \left(\prod_{j=1}^{n-1} \eta_j b_j^{-1} \right) \frac{\partial}{\partial z_n} \tilde{p}(b_n r_n = z_n) \right) dz_n$$

and doing integration by parts:

$$= \eta_n \left(\prod_{j=1}^{n-1} \eta_j b_j^{-1} \right) \left(\underbrace{\left[-q_{\chi^2}(z_{n+1}, \nu_{n+1}, \eta_{n+1} b_n^{-1} z_n) \tilde{p}(b_n r_n = z_n) \right]_0^\infty}_{=0} \right.$$

$$\left. + \int_0^\infty \left(\frac{\partial}{\partial z_n} q_{\chi^2}(z_{n+1}, \nu_{n+1}, \eta_{n+1} b_n^{-1} z_n) \right) \tilde{p}(b_n r_n = z_n) dz_n \right)$$

$$= \eta_n \left(\prod_{j=1}^{n-1} \eta_j b_j^{-1} \right) \int_0^\infty -\eta_{n+1} b_n^{-1} \left(\frac{\partial}{\partial z_{n+1}} q_{\chi^2}(z_{n+1}, \nu_{n+1} + 2, \eta_{n+1} b_n^{-1} z_n) \right) \tilde{p}(b_n r_n = z_n) dz_n$$

$$= -\eta_{n+1} \left(\prod_{j=1}^{n} \eta_j b_j^{-1} \right) \frac{\partial}{\partial z_{n+1}} \tilde{p}(b_{n+1} r_{n+1} = z_{n+1})$$

We have thus shown (D.8) for all n by induction, and integrating with respect to z_n yields (D.7).

□

In order to implement the multifactor version of the segmented square root model along the lines of Chen and Scott (1995), we need the following

D.5. PROPOSITION. *The characteristic function of the $(k-n)$ times multiple compound noncentral χ^2 distribution is given by*

$$\text{(D.9)} \qquad \Psi_{n+1,k}(x) = \left(\prod_{j=n+1}^{k} (1 + 2\mathcal{L}_{j,k}(x))^{-\frac{1}{2}\nu_j} \right) \exp\left\{ -\frac{\mathcal{L}_{n+1,k}(x)}{1 + 2\mathcal{L}_{n+1,k}(x)} \eta_{n+1} r_n \right\}$$

with $\mathcal{L}_{j,k}(x)$ recursively defined as

$$\text{(D.10)} \qquad \mathcal{L}_{j-1,k}(x) := \frac{\mathcal{L}_{j,k}(x) \eta_j b_{j-1}^{-1}}{1 + 2\mathcal{L}_{j,k}(x)} \qquad \textit{and} \qquad \mathcal{L}_{k,k}(x) := -ix$$

PROOF: For $k - n = 1$ we have

$$\Psi_{k,k}(x) = (1 - 2ix)^{-\frac{1}{2}\nu_k} \exp\left\{ ix \eta_k r_{k-1} (1 - 2ix)^{-1} \right\}$$

which is the characteristic function of the noncentral χ^2 distribution with ν_k degrees of freedom and noncentrality parameter $\eta_k r_{k-1}$[11]. Let (D.9) be valid for some $0 < n < k$.

[10] see Jamshidian (1995), p. 69

[11] See for example Johnson and Kotz (1970b).

Then we have for $n-1$:

$$\begin{aligned}\Psi_{n,k}(x) =& E\left[\exp\left\{ixb_k r_k\right\}\right]\\ =& \int_0^\infty e^{ixz_k} p_{n-1}(b_k r_k = z_k) dz_k\\ =& \int_0^\infty e^{ixz_k} \int_0^\infty p_n(b_k r_k = z_k \left|\eta_{n+1} b_n^{-1} z_n\right.) q_{\chi^2}(z_n, \nu_n, \eta_n r_{n-1}) dz_n dz_k\end{aligned}$$

interchanging the order of integration yields

$$= \int_0^\infty \Psi_{n+1,k}\left(x \left|\eta_{n+1} b_n^{-1} z_n\right.\right) q_{\chi^2}(z_n, \nu_n, \eta_n r_{n-1}) dz_n$$

and inserting (D.9)

$$\begin{aligned}=& \left(\prod_{j=n+1}^{k} \left(1+2\mathcal{L}_{j,k}(x)\right)^{-\frac{1}{2}\nu_j}\right)\\ &\cdot \int_0^\infty \exp\left\{-\frac{\mathcal{L}_{n+1,k}(x)}{1+2\mathcal{L}_{n+1,k}(x)} \eta_{n+1} b_n^{-1} z_n\right\} q_{\chi^2}(z_n, \nu_n, \eta_n r_{n-1}) dz_n\end{aligned}$$

and applying (D.10) and remark C.2

$$\begin{aligned}=& \left(\prod_{j=n}^{k} \left(1+2\mathcal{L}_{j,k}(x)\right)^{-\frac{1}{2}\nu_j}\right) \exp\left\{-\frac{\mathcal{L}_{n,k}(x)}{1+2\mathcal{L}_{n,k}(x)} \eta_n r_{n-1}\right\}\\ &\cdot \underbrace{\left(1+2\mathcal{L}_{n,k}(x)\right) \int_0^\infty q_{\chi^2}\left(\left(1+2\mathcal{L}_{n,k}(x)\right) z_n, \nu_n, \overbrace{\left(1+2\mathcal{L}_{n,k}(x)\right)^{-1} \eta_n r_{n-1}}^{=:\lambda}\right) dz_n}_{=:\xi(x)}.\end{aligned}$$

Now if $\xi(x) = 1$ for all x, then the proposition is proven. Inserting for q_{χ^2} its infinite sum expression[12]

$$\begin{aligned}\xi(x) =& \left(1+2\mathcal{L}_{n,k}(x)\right) \int_0^\infty \sum_{j=0}^{\infty} \frac{\left(\frac{1}{2}\lambda\right)^j}{j!} e^{-\frac{1}{2}\lambda} p_{\chi^2_{\nu_n+2j}}\left(\left(1+2\mathcal{L}_{n,k}(x)\right) z_n\right) dz_n\\ =& \sum_{j=0}^{\infty} \frac{\left(\frac{1}{2}\lambda\right)^j}{j!} e^{-\frac{1}{2}\lambda} \left(1+2\mathcal{L}_{n,k}(x)\right) \int_0^\infty p_{\chi^2_{\nu_n+2j}}\left(\left(1+2\mathcal{L}_{n,k}(x)\right) z_n\right) dz_n.\end{aligned}$$

[12]Note that q_{χ^2} and p_{χ^2} must now be interpreted as functions with complex arguments, not densities.

This equals 1 if $(1+2\mathcal{L}_{n,k}(x))\int_0^\infty p_{\chi^2_{\nu_n+2j}}\left((1+2\mathcal{L}_{n,k}(x))\,z_n\right)dz_n = 1\ \forall j\in\mathbb{N}$. We have

$$
\begin{aligned}
&(1+2\mathcal{L}_{n,k}(x))\int_0^\infty p_{\chi^2_{\nu_n+2j}}\left((1+2\mathcal{L}_{n,k}(x))\,z_n\right)dz_n\\
=&(1+2\mathcal{L}_{n,k}(x))\int_0^\infty 2^{-\frac{1}{2}\nu_n-j}\Gamma\left(\frac{1}{2}\nu_n+j\right)^{-1}\left((1+2\mathcal{L}_{n,k}(x))\,z_n\right)^{\frac{1}{2}\nu_n+j-1}\\
&\quad\cdot\exp\left\{-\frac{1}{2}(1+2\mathcal{L}_{n,k}(x))z_n\right\}dz_n\\
=&\Gamma\left(\frac{1}{2}\nu_n+j\right)^{-1}\left(\frac{1}{2}\left(1+2\mathcal{L}_{n,k}(x)\right)\right)^{\frac{1}{2}\nu_n+j}\int_0^\infty z_n^{\frac{1}{2}\nu_n+j-1}\exp\left\{-\frac{1}{2}(1+2\mathcal{L}_{n,k}(x))z_n\right\}dz_n
\end{aligned}
$$

which by formula 6.1.1 in Abramowitz and Stegun (1964) equals

$$
=\Gamma\left(\frac{1}{2}\nu_n+j\right)^{-1}\Gamma\left(\frac{1}{2}\nu_n+j\right)=1
$$

□

Acknowledgements

Financial support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, and Deutscher Akademischer Austauschdienst (DAAD) is gratefully acknowledged. Part of this work was completed while the first author was a guest of the Centre for Analytical Finance (CAF) at Aarhus University, Denmark. We would like to thank Jørgen Aase Nielsen, Klaus Schürger, Daniel Sommer, Dieter Sondermann and Asbjørn Trolle Hansen for helpful discussions. All remaining errors are ours.

References

Abramowitz, M. and Stegun, I. A. (eds) (1964), *Handbook of Mathematical Functions*, National Bureau of Standards.

Adams, K. J. and van Deventer, D. R. (1994), Fitting Yield Curves and Forward Rate Curves with Maximum Smoothness, *The Journal of Fixed Income* **4**(1), 52–62.

Bauer, H. (1990), *Maß– und Integrationstheorie*, Walter de Gruyter.

Bauer, H. (1991), *Wahrscheinlichkeitstheorie*, 4 edn, Walter de Gruyter.

Black, F. and Scholes, M. (1973), The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* pp. 637–654.

Chen, R.-R. and Scott, L. (1995), Interest Rate Options in Multifactor Cox-Ingersoll-Ross Models of the Term Structure, *The Journal of Derivatives* **3**(2), 52–72.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985), A Theory of the Term Structure of Interest Rates, *Econometrica* **53**(2), 385–407.

Duffie, J. D. and Kan, R. (1992), A Yield–Factor Model of Interest Rates, Stanford University, working paper.

Duffie, J. D. and Kan, R. (1996), A Yield Factor Model of Interest Rates, *Mathematical Finance* **6**(4), 379–406.

Heath, D., Jarrow, R. and Morton, A. (1992), Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation, *Econometrica* **60**(1), 77–105.

Hull, J. and White, A. (1990), Pricing Interest–Rate Derivative Securities, *The Review of Financial Studies* **3**(4), 573–592.

Jamshidian, F. (1987), Pricing of Contingent Claims in the One–Factor Term Structure Model, Financial Strategies Group, Merrill Lynch Capital Markets, working paper.

Jamshidian, F. (1989), An Exact Bond Option Formula, *Journal of Finance* **44**, 205–209.
Jamshidian, F. (1995), A Simple Class of Square–Root Interest–Rate Models, *Applied Mathematical Finance* **2**, 61–72.
Johnson, N. L. and Kotz, S. (1970a), *Continuous Univariate Distributions - 1*, The Houghton Mifflin Series in Statistics, John Wiley & Sons, Inc., New York, New York, USA.
Johnson, N. L. and Kotz, S. (1970b), *Continuous Univariate Distributions - 2*, The Houghton Mifflin Series in Statistics, John Wiley & Sons, Inc., New York, New York, USA.
Karatzas, I. and Shreve, S. E. (1988), *Brownian Motion and Stochastic Calculus*, 2 edn, Springer Verlag.
Scott, L. (1995), The Valuation of Interest Rate Derivatives in a Multi-Factor Term Structure Model with Deterministic Components, University of Georgia, working paper.
Yeh, J. (1995), *Martingales and Stochastic Analysis*, Vol. 1 of *Series on Multivariate Analysis*, World Scientific.