CONVEXITY THEORY FOR THE TERM STRUCTURE EQUATION

ERIK EKSTRÖM AND JOHAN TYSK

ABSTRACT. We study convexity and monotonicity properties for prices of bonds and bond options when the short rate is modeled by a diffusion process. We provide conditions under which convexity of the price in the short rate is guaranteed. Under these conditions the price is decreasing in the drift and increasing in the volatility of the short rate. We also study convexity properties of the logarithm of the price.

1. Introduction

Already in the seminal paper [24], convexity of the option price in the underlying asset is discussed. In the last decade this issue has attracted renewed interest in the literature, compare [1], [4], [5], [6], [12], [13], [14], [15], [17], [18], [19] and [20]. Given a convex pay-off function one asks for what models this convexity is preserved in the sense that the price also is a convex function of the underlying asset at any fixed time prior to maturity. This question is studied for diffusion models, for models with jumps and for various option types including options written on several underlying assets. The interest in convexity has at least three reasons. Firstly, convexity is a fundamental qualitative property of option prices (an even more fundamental qualitative property would be monotonicity in the underlying asset provided the pay-off is monotone, but such properties can usually be derived immediately). Secondly, if the price is convex then it is also typically increasing in the volatility, and in the case of jump-diffusion models, also in the jump parameters. Thirdly, if a delta-hedger uses a model that overestimates the true volatility, he or she will obtain a superhedge for the claim provided the price is convex.

Our aim with the present paper is to continue this study to bonds and bond options, for which we regard the short rate as the underlying process. Thus we study preservation of convexity for the term structure equation instead of variants of the Black-Scholes equation as is the case in the references above. Surprisingly little has been done in this direction, with [2] as a notable exception. A motivation for this is perhaps that the third reason for studying convexity mentioned above is not directly applicable since the short rate is not a traded asset. However, the first two stated motives remain valid also in interest rate theory.

We assume that the short rate is modeled under some given risk neutral probability measure as a stochastic process $X = (X_t)_{t\geq 0}$ with dynamics

$$dX_t = \beta(X_t, t)\, dt + \sigma(X_t, t)\, dB_t,$$

where B is a standard Brownian motion and β and σ are given functions of time and the current short rate. We first investigate the convexity properties of the T-bond price

$$u(x,t) = E_{x,t}\left[\exp\left\{-\int_t^T X_s\, ds\right\}\right], \tag{1}$$

2000 *Mathematics Subject Classification.* Primary 91B28; Secondary 35K99.

Key words and phrases. Interest rate theory, bond options, convexity, parameter monotonicity, log-convexity, log-concavity, affine term structure.

where the indices indicate that $X_t = x$. Using the Feynman-Kac theorem it follows that the bond price u satisfies the term structure equation

$$\begin{cases} u_t + \frac{\sigma^2}{2}u_{xx} + \beta u_x - xu = 0 & \text{for } t < T \\ u = 1 & \text{for } t = T. \end{cases}$$

This enables us to use a mixture of stochastic techniques and methods from the theory of parabolic partial differential equations.

One should note that our pay-off function in the case of bonds is identically equal to one, so it is both convex and concave. However, convexity in x is the natural property to consider: since the bond price declines with x, convexity means that the absolute value of this decline decreases with x. We certainly expect bond prices to suffer a smaller decline if short rates move from 5% to 6% than if the short rates move from 1% to 2%! In Section 5 we find conditions on the model for X that guarantee that convexity is preserved. More precisely, if $\beta_{xx} \leq 2$ (in the sense of distributions if β is not twice continuously differentiable), then the model is convexity preserving. To our knowledge, this condition is indeed fulfilled for all models of the short rate that are used in practice, compare Table 1 below. Moreover, the condition is sharp in the sense that it is also a necessary condition for preservation of convexity provided the coefficients of the model are regular enough, see Theorem 5.2. Using the pathwise non-crossing property of diffusions, it is easily seen that the bond price u is decreasing in the drift β. In Section 6 we show that the bond price in a convexity preserving model is also increasing in the volatility σ. Thus the general relationship between convexity and monotonicity in the volatility known from option pricing theory extends to interest rate theory.

Our study of convexity and monotonicity properties is formulated for prices of options written with the short rate as the underlying asset. Thus we study the function

$$U(x,t) = E_{x,t}\left[\exp\left\{-\int_t^T X_s\,ds\right\} g(X_T)\right], \tag{2}$$

where g is a given convex pay-off function (note that the case $g \equiv 1$ corresponds to a bond). The main reason for extending the study from bonds to options on the short rate is to be able to study bond option prices, i.e. prices of options written on a bond price as the underlying asset. In fact, our general convexity and monotonicity results allow us in Section 7 to deduce properties of certain bond options. For example, the price of a bond call option is convex in x and therefore also increasing in σ for convexity preserving models.

It is also natural to consider convexity properties of the *logarithm* of the bond price. This is connected to the notion of *duration*, i.e. the negative of the derivative of the logarithm. The analogous concept for stock options is elasticity, compare [9], [10], [16] and [22]. We say that the price is log-convex if the logarithm of the price is convex in x and analogously for log-concavity. Again, since the pay-off is constant for the bond it is both log-convex and log-concave. Unlike the case of convexity, however, both of these cases deserve consideration. According to the discussion above, convexity of a bond price means that the *absolute* value of the decline is decreasing in x. In contrast to this, log-convexity means that the *relative* decline diminishes when x grows (a declining duration), and log-concavity means that the relative decline of the price increases (an increasing duration). In Section 8 we show that if the drift β is spatially concave and the diffusion coefficient σ^2 is spatially convex, then log-convexity is preserved. Similarly, in Section 9 we show that log-concavity is preserved provided β is convex and σ^2 is concave. If we insist that the model should preserve both log-convexity and log-concavity, we arrive at models where the logarithm of the bond price is both convex and concave, i.e. linear. These are, of course, the models that admit an affine term structure (in our context these bond prices would be referred to as being log-affine). Apart from admitting explicit bond prices, affine models play an important rôle in interest rate theory, compare for example [8]. Thus we recover the well-known sufficient condition that β and σ^2 are affine for the existence of an affine term structure.

In the next section we present the assumptions on the model parameters under which our results are presented. If additional regularity of the coefficients is assumed, bounds on the spatial derivatives of bond and option prices can be obtained, see Section 3. In Section 4 a continuity result is provided, thus allowing us to assume that the coefficients are regular, and the general results follow by approximating the coefficients. As mentioned above, our main results are presented in Sections 5-9, and a summary of these results is presented in Section 10.

One important feature of our approach is that it works just as well for models where X never reaches zero, for models where the rate can reach zero, as for models that allow negative interest rates. We thus avoid a case by case study. However, we will throughout the paper point out to which of the commonly studied models presented in Table 1 our results are applicable. For a more detailed discussion of these models, see for example [11].

Model	Dynamics	$X > 0$	AB	AO
V	$dX = k(\theta - X)\,dt + \sigma\,dB$	No	Yes	Yes
CIR	$dX = k(\theta - X)\,dt + \sigma\sqrt{X}\,dB$	Yes	Yes	Yes
D	$dX = bX\,dt + \sigma X\,dB$	Yes	Yes	No
EV	$dX = X(\eta - a\ln X)\,dt + \sigma X\,dB$	Yes	No	No
HW	$dX = k(\theta_t - X)\,dt + \sigma\,dB$	No	Yes	Yes
BK	$dX = X(\eta_t - a\ln X)\,dt + \sigma X\,dB$	Yes	No	No
MM	$dX = X(\eta_t - (\lambda - \frac{\gamma}{1+\gamma t})\ln X)\,dt + \sigma X\,dB$	Yes	No	No

TABLE 1. Some examples of short rate models. The table is copied from [11]. AB stands for analytic bond prices, and AO stands for analytic bond option prices. When referring to these models below we will assume that k, θ, a, η and γ are positive and that $\lambda \geq \gamma$.

2. Assumptions

As explained in the introduction, we assume that the short rate X is modeled as

$$dX_t = \beta(X_t, t)\,dt + \sigma(X_t, t)\,dB_t \tag{3}$$

for some Brownian motion B. Note that some of the short rate models in Table 1 are specified so that the interest rate cannot fall below zero, whereas some models allow for negative interest rates with positive probability. To unify the analysis, we view the short rate process X in (3) as specified on the whole real line $\mathbb{R}$ with σ and β suitably extended (for example to be 0) for negative values in case X is specified initially only on the positive real axis.

Throughout the article we make the following regularity and growth assumptions.

Assumption 2.1. *The functions $\beta, \sigma : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ are continuous functions, β is locally Lipschitz continuous in the x-variable, and σ is locally Hölder(1/2) in the x-variable. Moreover, there exists a constant D such that*

$$|\sigma(x, t)| \leq D(1 + x^+) \tag{4}$$

and

$$|\beta(x,t)| \leq D(1+|x|) \tag{5}$$

for all $(x,t) \in \mathbb{R} \times [0,\infty)$.

The conditions on σ and β guarantee a non-exploding unique strong solution to (3) for any initial point $(x,t) \in \mathbb{R} \times [0,\infty)$. The condition that σ is bounded for negative x implies that, for any T, the price u in (1) of a T-bond is finite at all points $(x,t) \in \mathbb{R} \times [0,T]$, see Corollary 3.3 below. Without this condition, the bond price u may be infinite. Indeed, models in which σ grows faster than $\sqrt{|x|}$ for negative x have typically infinite bond prices, compare Theorem 4.1 in [26]. One should note that all models in Table 1 satisfy the conditions of Assumption 2.1 in the following sense: those models in which the short rate may fall below zero (V and HW) clearly satisfy Assumption 2.1; the remaining models give rise to non-negative rates, so the parameters can be chosen arbitrarily for negative values so that Assumption 2.1 holds without affecting the bond value for positive rates.

3. Auxiliary estimates

When dealing with bonds and options on the short rate it is natural to study the corresponding term structure equation, i.e. the terminal value problem

$$\begin{cases} U_t + \alpha U_{xx} + \beta U_x - xU = 0 & \text{for } t < T \\ U = g & \text{for } t = T, \end{cases} \tag{6}$$

where $\alpha := \sigma^2/2$. Due to technical reasons we instead study the function

$$V(x,t) := V^f(x,t) := E_{x,t}\left[\exp\left\{-\int_t^T f(X_s)\,ds\right\} g(X_T)\right] \tag{7}$$

for some appropriate function f and then let f approach x. Note that the corresponding terminal value problem is

$$\begin{cases} V_t + \alpha V_{xx} + \beta V_x - fV = 0 & \text{for } t < T \\ V = g & \text{for } t = T. \end{cases} \tag{8}$$

Our choice of f is indicated in the following hypothesis.

Hypothesis 3.1. *f is smooth, concave and there exists a constant $K' > 0$ such that*

$$f(x) = \begin{cases} x & \textit{if } x \leq K' \\ \textit{constant} & \textit{if } x \geq K'+1. \end{cases}$$

The main goal of this section is to provide some estimates on option prices and the derivatives of V, compare Corollary 3.3 and Proposition 3.5. We first claim that V grows at most exponentially as $x \to -\infty$. To be more precise, assume that the pay-off function g satisfies

$$0 \leq g(x) \leq M \max\{e^{-Kx}, 1\} \tag{9}$$

for some non-negative constants M and K (if $g = 1$ then (9) holds with $M = 1$, $K = 0$). Moreover, let D be the constant from Hypothesis 3.4. From (5) we have

$$\beta(x,t) \geq -D(1-x) \text{ for all negative } x. \tag{10}$$

Define

$$W(x,t) := e^{f(x)h(t)}V(x,t), \tag{11}$$

where

$$h(t) = \frac{e^{D(T-t)}-1}{D} + Ke^{D(T-t)}.$$

The choice of the function h is motivated by the following observation. If $\sigma(x,t)=0$ and $\beta(x,t)=Dx$ for negative x, and if $X_0=x\leq 0$, then X satisfies $X_t=xe^{Dt}$. Consequently,

$$\begin{aligned} V(x,t) &= E_{x,t}\left[\exp\left\{-\int_t^T f(X_s)\,ds\right\}g(X_T)\right] = \exp\left\{-\int_t^T xe^{D(s-t)}\,ds\right\}g(xe^{D(T-t)}) \\ &\leq \exp\left\{-x\frac{e^{D(T-t)}-1}{D}\right\}M\exp\left\{-Kxe^{D(T-t)}\right\} = Me^{-xh(t)}, \end{aligned}$$

where we have used the bound (9). This indicates that the function W is bounded (at least for negative values of x). As is shown in Lemma 3.2 below, this intuition is indeed correct.

Lemma 3.2. *Assume Hypothesis 3.1 and the bound* (9)*. Then the function W defined in* (11) *is bounded on $\mathbb{R}\times[0,T]$, i.e. there exists a constant C such that $0\leq W\leq C$.*

Proof. By the bound (9) and the Cauchy-Schwartz inequality we have

$$\begin{aligned} \frac{W^2(x,t)}{M^2} &= \left(\frac{e^{f(x)h(t)}}{M}E_{x,t}\left[\exp\left\{-\int_t^T f(X_s)\,ds\right\}g(X_T)\right]\right)^2 \\ &\leq \left(E_{x,t}\left[\exp\left\{\int_t^T e^{D(s-t)}f(x)-f(X_s)\,ds\right\}\max\{e^{-KX_T},1\}e^{Kf(x)e^{D(T-t)}}\right]\right)^2 \qquad (12) \\ &\leq E_{x,t}\left[\exp\left\{2\int_t^T e^{D(s-t)}f(x)-f(X_s)\,ds\right\}\right] \\ &\quad\times E_{x,t}\left[\max\{e^{-2KX_T},1\}e^{2Kf(x)e^{D(T-t)}}\right]. \end{aligned}$$

By Jensen's inequality (applied to the exponential function and the time integral) we have

$$\begin{aligned} & E_{x,t}\left[\exp\left\{2\int_t^T e^{D(s-t)}f(x)-f(X_s)\,ds\right\}\right] \\ &\qquad \leq E_{x,t}\left[\frac{1}{T-t}\int_t^T \exp\left\{2(e^{D(s-t)}f(x)-f(X_s))(T-t)\right\}\,ds\right] \\ &\qquad = \frac{1}{T-t}\int_t^T E\left[\exp\left\{-2e^{D(s-t)}(T-t)Y_s\right\}\right]\,ds \end{aligned}$$

where $Y_s=e^{-D(s-t)}f(X_s)-f(x)$ with $Y_t=0$. By Ito's lemma we have

$$dY_s=\tilde\beta_s\,ds+\tilde\sigma_s\,dB_s \quad \text{for } s\geq t,$$

where

$$\tilde\beta_s=e^{-D(s-t)}\left(f'(X_s)\beta(X_s,s)-Df(X_s)+\frac{1}{2}f''(X_s,s)\sigma^2(X_s,s)\right)$$

and

$$\tilde\sigma_s=e^{-D(s-t)}f'(X_s)\sigma(X_s,s).$$

It follows from the assumption (10) and Hypothesis 3.1 that $\tilde\beta$ is bounded from below, i.e. there exists a constant C such that $\tilde\beta_s\geq -C$ for all s almost surely. Similarly, the bound (4) and Hypothesis 3.1 yield that $|\tilde\sigma_s|\leq C$. Since $y\mapsto\exp\{-2e^{D(s-t)}(T-t)y\}$ is decreasing and convex, we have

$$E\left[\exp\{-2e^{D(s-t)}(T-t)Y_s\}\right]\leq E\left[\exp\{-2e^{D(s-t)}(T-t)\tilde Y_s\}\right] \qquad (13)$$

for every $s \in [t,T]$, where

$$\tilde{Y}_s = -C(s-t) + CB_{s-t}$$

is a Brownian motion with drift starting from 0 at time t. Indeed, monotonicity in the drift is immediate, and monotonicity in the volatility for processes with constant drift and convex pay-off functions is well-known, compare Theorem 6.2 in [17] and Theorem 7 in [20]. Consequently,

$$E_{x,t}\left[\exp\left\{2\int_t^T f(x)e^{D(s-t)} - f(X_s)\,ds\right\}\right] \leq \frac{1}{T-t}\int_t^T E\left[\exp\left\{-2e^{D(s-t)}(T-t)\tilde{Y}_s\right\}\right]ds \leq C'$$

for some constant C' (in fact, $\tilde{Y}_s$ is $N(-C(s-t), C^2(s-t))$-distributed, so explicit bounds of the above integral are readily derived). Consequently, the first factor in (12) is bounded. For the second factor in (12), note that

$$E_{x,t}\left[\max\{e^{-2KX_T},1\}e^{2Kf(x)e^{D(T-t)}}\right] \leq E_{x,t}\left[e^{2K(f(x)e^{D(T-t)}-f(X_T))}\right] + e^{2Kf(x)e^{D(T-t)}} \tag{14}$$

The first term in (14) is of the form shown above to be bounded, and the second one is bounded since f is bounded from above. This shows that W is bounded, which finishes the proof. □

Corollary 3.3. *Assume that g satisfies* (9). *Then the option price U defined in* (2) *above is finite. In fact, there exists a constant $M > 0$ such that*

$$U(x,t) \leq M\max\{1, e^{-Mx}\}$$

for all $(x,t) \in \mathbb{R} \times [0,T]$.

Proof. Since $f(x) \leq x$, we have that $U(x,t) \leq V(x,t)$ for all x and t. Consequently, it follows from Lemma 3.2 that for every choice of a function f satisfying Hypothesis 3.1 there exists a constant C such that $U(x,t) \leq Ce^{-f(x)(e^{D(T-t)}-1)/D - Ke^{D(T-t)}f(x)}$ for all $(x,t) \in \mathbb{R} \times [0,T]$. Thus

$$U(x,t) \leq M\max\{1, e^{-Mx}\}$$

for some large constant M. □

We next provide conditions under which the kth spatial derivative of the function W decays like $|x|^{-k}$, $k = 1,2,3$. We need the following assumptions.

Hypothesis 3.4. *The coefficients α and β are smooth and $\alpha > 0$ at all points. Moreover,*

(15) $|\beta(x,t) - Dx|$ *does not depend on x for $(x,t) \in \mathbb{R} \times [0,T]$ with x large negative,*

and the derivatives satisfy the growth conditions

$$|\partial_x^k \beta(x,t)| \leq C(1+|x|)^{1-k} \qquad (k = 0,1,2,3)$$

for $(x,t) \in [0,\infty) \times [0,T]$,

$$|\partial_x^k \alpha(x,t)| \leq C(1+|x|)^{2-k} \qquad (k = 0,1,2,3)$$

for $(x,t) \in [0,\infty) \times [0,T]$, and

$$|\partial_x^k \alpha(x,t)| \leq C(1+|x|)^{-k} \qquad (k = 0,1,2,3)$$

for $(x,t) \in (-\infty,0] \times [0,T]$.

Proposition 3.5. *Assume Hypotheses 3.1 and 3.4, that the pay-off function g is smooth, satisfies (9) and that $e^{f(x)K}g(x)$ is constant for large $|x|$. Then there exists a constant C such that the function W defined in (11) satisfies*

$$|\partial_x^k W| \leq \frac{C}{1+|x|^k} \tag{16}$$

for $k=0,1,2,3$.

Proof. The case $k=0$ follows from Lemma 3.2. Thus we know that W is a bounded solution to

$$\begin{cases} W_t + \alpha W_{xx} + \hat{\beta} W_x + \gamma W = 0 \\ W(x,T) = \hat{g}(x) \end{cases} \tag{17}$$

where

$$\hat{g}(x) = e^{f(x)K} g(x),$$

$$\hat{\beta} = \beta - 2 f_x \alpha h$$

and

$$\gamma = (Df - f_x\beta)h + f_x^2 \alpha h^2 - f_{xx}\alpha h.$$

Let $\hat{W}(x,t) = W(x,t) - \hat{g}(x)$. Then

$$\begin{cases} \hat{W}_t + \alpha \hat{W}_{xx} + \hat{\beta}\hat{W}_x + \gamma \hat{W} + \hat{h} = 0 \\ \hat{W}(x,T) = 0, \end{cases} \tag{18}$$

where

$$\hat{h} = \alpha \hat{g}_{xx} + \hat{\beta}\hat{g}_x + \gamma \hat{g}.$$

Since $\hat{g}$ is constant for large values of $|x|$, W satisfies (16) if and only if $\hat{W}$ does. Using Hypotheses 3.1 and 3.4 we note that

$$\begin{array}{ll} |\partial_x^k \alpha| \leq C(1+|x|)^{2-k} & |\partial_x^k \hat{\beta}| \leq C(1+|x|)^{1-k} \\ |\partial_x^k \gamma| \leq C(1+|x|)^{-k} & |\partial_x^k \hat{h}| \leq C(1+|x|)^{-k} \end{array}$$

for some constant C and for all $k=0,1,2,3$. By the Feynman-Kac representation theorem,

$$\hat{W}(x,t) = E\left[\int_t^T \exp\left\{\int_t^s \gamma(Y_r^{x,t},r)\,dr\right\} \hat{h}(Y_s^{x,t})\,ds\right],$$

where $Y^{x,t}$ is the diffusion

$$dY_s^{x,t} = \hat{\beta}(Y_s^{x,t},s)\,ds + \sigma(Y_s^{x,t},s)\,dB_s$$

with $Y_t^{x,t} = x$. Now, let $x<y$. Then

$$\begin{aligned} |\hat{W}(x,t) - \hat{W}(y,t)| &\leq E\left[\int_t^T \exp\left\{\int_t^s \gamma(Y_r^{x,t},r)\,dr\right\} \left|\hat{h}(Y_s^{x,t}) - \hat{h}(Y_s^{y,t})\right| ds\right] \\ &\quad + E\left[\int_t^T \left|\exp\left\{\int_t^s \gamma(Y_r^{x,t},r)\,dr\right\} - \exp\left\{\int_t^s \gamma(Y_r^{y,t},r)\,dr\right\}\right| \hat{h}(Y_s^{y,t})\,ds\right] \\ &= I_1 + I_2. \end{aligned}$$

Since γ is bounded and $\hat{h}$ is Lipschitz continuous, we find that

$$I_1 \leq C\int_t^T E\left[|Y_s^{x,t} - Y_s^{y,t}|\right] ds = C\int_t^T E\left[Y_s^{y,t} - Y_s^{x,t}\right] ds,$$

where the equality holds since $x < y$ implies $Y_s^{x,t} \leq Y_s^{y,t}$ for $s \geq t$, compare Theorem IX.3.7 in [25]. Since the drift $\hat{\beta}$ is Lipschitz continuous, we have

$$E\left[Y_s^{y,t} - Y_s^{x,t}\right] = y - x + \int_t^s E\left[\beta(Y_s^{y,t}, r) - \beta(Y_r^{x,t}, r)\right]\, dr \leq y - x + C\int_t^s E\left[Y_r^{y,t} - Y_r^{x,t}\right]\, dr$$

for any $s \geq t$. It follows from Gronwall's lemma that

$$I_1 \leq C|y - x|(T - t).$$

Similarly, since $\hat{h}$ is bounded and γ is bounded and Lipschitz it can also be shown that

$$I_2 \leq C|y - x|(T - t)$$

for some constant C. It follows that $\hat{W}_x$ is bounded on $\mathbb{R} \times [0, T]$ and that

$$\lim_{(x,t)\to(x_0,T)} \hat{W}_x(x, t) = 0.$$

Therefore, by differentiating the equation (18) we know that the first spatial derivative $p := \hat{W}_x$ is a bounded classical solution to

$$\left\{ \begin{array}{l} p_t + \alpha p_{xx} + (\alpha_x + \hat{\beta})p_x + (\gamma + \hat{\beta}_x)p + \gamma_x \hat{W} - \hat{h}_x = 0 \\ p(x, T) = 0 \end{array} \right. \tag{19}$$

It is straightforward to check that

$$\tilde{p}(x, t) := e^{C(T-t)} l(x),$$

where l is a smooth positive function with $l(x) = 1/(1 + |x|)$ for large $|x|$, is a supersolution to (19) for some large constant C, and $-\tilde{p}$ is a subsolution. By the maximum principle we have $|\hat{W}_x| = |p| \leq \tilde{p} \leq e^{CT}/(1 + |x|)$. This proves (16) for $k = 1$.

Next, the solution to equation (19) also has a stochastic representation. Using this representation, the same analysis as above can be applied to show that p_x is bounded, and using the maximum principle again it is straightforward to check that $p_x = \hat{W}_{xx}$ decays like $|x|^{-2}$. Finally, the same reasoning can be applied to prove (16) for $k = 3$. □

4. Continuity of bond prices in the model parameters

As mentioned in the introduction, it is central to our approach to be able to approximate a bond price (or an option price) by a sequence of bond prices (option prices) in models with regular coefficients. Thus we need to know that bond prices are continuous in the model parameters. To formulate such a result, let σ_n and β_n satisfy the Assumption 2.1 uniformly in n (i.e. the bounds (4) and (5) hold with the same constant D). Also assume that $\alpha_n \to \alpha$ and $\beta_n \to \beta$ uniformly on compacts as $n \to \infty$. Let X^n be the solution of the stochastic differential equation

$$dX_s^n = \beta_n(X_s^n, s)\, ds + \sigma_n(X_s^n, s)\, dB_s \qquad X_t^n = x.$$

Under these conditions it is known that

$$E_{x,t}\left[\sup_{t\leq s\leq T} |X_s^n - X_s|^2\right] \to 0 \tag{20}$$

as $n \to \infty$, see Theorem 2.5 in [3]. Now let

$$U_n(x, t) = E_{x,t}\left[\exp\left\{-\int_t^T X_s^n\, ds\right\} g(X_T^n)\right]. \tag{21}$$

Here g is assumed to have the following property: for each positive constant k there exists a constant C_k such that

$$|g(x) \wedge k - g(y) \wedge k| \leq C_k|x - y| \tag{22}$$

for all $x, y \in \mathbb{R}$. Note that this condition is satisfied for example if g is convex and decreasing.

Proposition 4.1. *Assume that σ_n, β_n and X^n are as described above. Also assume that g satisfies* (9) *and* (22)*. Then*

$$U(x,t) = \lim_{n\to\infty} U_n(x,t)$$

for all $(x,t) \in \mathbb{R}\times[0,T]$.

Proof. Fix a point $(x,t) \in \mathbb{R}\times[0,T]$. For positive constants k, define $U^k(x,t)$ and $U_n^k(x,t)$ by

$$U^k(x,t) = E_{x,t}\left[\exp\left\{\int_t^T k\wedge -X_s\,ds\right\}(g(X_T)\wedge k)\right]$$

and

$$U_n^k(x,t) = E_{x,t}\left[\exp\left\{\int_t^T k\wedge -X_s^n\,ds\right\}(g(X_T^n)\wedge k)\right].$$

Then

$$\left|U_n^k(x,t) - U^k(x,t)\right| \le I_1 + I_2,$$

where

$$I_1 = E_{x,t}\left[\exp\left\{\int_t^T k\wedge -X_s^n\,ds\right\}|g(X_T^n)\wedge k - g(X_T)\wedge k|\right]$$

and

$$I_2 = E_{x,t}\left[\left|\exp\left\{\int_t^T k\wedge -X_s^n\,dt\right\} - \exp\left\{\int_t^T k\wedge -X_s\,ds\right\}\right|(g(X_T)\wedge k)\right].$$

Note that

$$I_1 \le e^{k(T-t)}E_{x,t}\left[|g(X_T^n)\wedge k - g(X_T)\wedge k|\right] \le e^{k(T-t)}C_k E_{x,t}\left[|X_T^n - X_T|\right]$$

since g satisfies (22). It thus follows from (20) that $I_1 \to 0$ as $n\to\infty$. Similarly, using the inequality $|e^x - e^y| \le (e^x + e^y)|x-y|$, we find

$$\begin{aligned} I_2 &\le kE_{x,t}\left[\left|\exp\left\{\int_t^T k\wedge -X_s^n\,ds\right\} - \exp\left\{\int_t^T k\wedge -X_s\,ds\right\}\right|\right] \\ &\le 2ke^{k(T-t)}E_{x,t}\left[\int_t^T |k\wedge -X_s^n - k\wedge -X_s|\,ds\right] \\ &\le 2ke^{k(T-t)}E_{x,t}\left[(T-t)\sup_{t\le s\le T}|X_s^n - X_s|\right] \to 0 \end{aligned}$$

as $n\to\infty$ by (20). Consequently,

$$\lim_{n\to\infty} U_n^k(x,t) = U^k(x,t) \tag{23}$$

for each fixed k. It then follows from the monotone convergence theorem that

$$\begin{aligned} \lim_{k\to\infty}\lim_{n\to\infty} U_n^k(x,t) &= \lim_{k\to\infty} E_{x,t}\left[\exp\left\{\int_t^T k\wedge -f(X_s)\,ds\right\}(g(X_T)\wedge k)\right] \\ &= E_{x,t}\left[\exp\left\{-\int_t^T f(X_s)\,ds\right\}g(X_T)\right] = U(x,t). \end{aligned} \tag{24}$$

Since

$$U_n(x,t) = \lim_{k\to\infty} U_n^k(x,t)$$

by the monotone convergence theorem, it suffices to show that the order of the limits in (24) can be interchanged. To see this, first note that

$$
\begin{aligned}
(25)\qquad 0 &\leq U_n(x,t) - U_n^k(x,t) \\
&\leq E_{x,t}\left[\exp\left\{-\int_t^T X_s^n\,ds\right\} g(X_T^n) - \exp\left\{\int_t^T k\wedge -X_s^n\,ds\right\} g(X_T^n)\right] \\
&\quad + E_{x,t}\left[\exp\left\{-\int_t^T X_s^n\,ds\right\}\left(g(X_T^n) - k\wedge g(X_T^n)\right)\right] = I_3 + I_4.
\end{aligned}
$$

We first claim that

$$
\lim_{k\to\infty}\sup_n I_i = 0 \qquad \text{for } i = 3,4. \tag{26}
$$

To see this, note that

$$
\begin{aligned}
I_3 &= E_{x,t}\left[\exp\left\{-\int_t^T X_s^n\,ds\right\} g(X_T^n)\left(1 - \exp\left\{\int_t^T \min(k + X_s^n, 0)\,ds\right\}\right)\right] \\
&\leq E_{x,t}\left[\exp\left\{-\int_t^T X_s^n\,ds\right\} g(X_T^n)\int_t^T (-k - X_s^n)^+\,ds\right] \\
&\leq \left(E_{x,t}\left[\exp\left\{-2\int_t^T X_s^n\,ds\right\}(g(X_T^n))^2\right] E_{x,t}\left[\left(\int_t^T (-k - X_s^n)^+\,ds\right)^2\right]\right)^{1/2}
\end{aligned}
$$

where we have used the inequality $1 - e^x \leq -x$ and the Cauchy-Schwartz inequality. The first factor is bounded uniformly in n for each fixed x (this can be proved analogously to Lemma 3.2). As for the second factor, an application of Jensen's inequality gives

$$
E_{x,t}\left[\left(\int_t^T (-k - X_s^n)^+\,ds\right)^2\right] \leq (T-t)\int_t^T E_{x,t}\left[\left((-k - X_s^n)^+\right)^2\right]\,ds.
$$

Proceeding as in the proof of Lemma 3.2 above, it is straightforward to check that

$$
\begin{aligned}
E_{x,t}\left[((-k - X_s^n)^+)^2\right] &\leq e^{2DT} E_{x,t}\left[\left((-ke^{-DT} - e^{-D(s-t)}X_s^n)^+\right)^2\right] \\
&\leq e^{2DT} E_{x,t}\left[\left((-ke^{-DT} - \tilde{Y}_s)^+\right)^2\right]
\end{aligned}
$$

for a Brownian motion $\tilde{Y}$ with negative drift. Moreover, since $((-ke^{-DT} - \tilde{Y}_s)^+)^2$ is a submartingale we have

$$
E_{x,t}\left[\left((-ke^{-DT} - \tilde{Y}_s)^+\right)^2\right] \leq E_{x,t}\left[\left((-ke^{-DT} - \tilde{Y}_T)^+\right)^2\right] \to 0
$$

as $k \to \infty$ by dominated convergence. Consequently, (26) holds for $i = 3$. Similar methods can be employed to establish (26) also for $i = 4$.

Let $\varepsilon > 0$ be given. From (25) and (26) it follows that there exists a k_0 such that

$$
0 \leq U_n(x,t) - U_n^k(x,t) \leq \varepsilon
$$

for all $k \geq k_0$ and all n. Thus, for such a k we have

$$
\lim_{n\to\infty} U_n^k(x,t) \leq \liminf_{n\to\infty} U_n(x,t) \leq \limsup_{n\to\infty} U_n(x,t) \leq \lim_{n\to\infty} U_n^k(x,t) + \varepsilon
$$

where we recall from (23) that the outer limits exist. Letting $k \to \infty$ and remembering (24) above we find that

$$
U(x,t) \leq \liminf_{n\to\infty} U_n(x,t) \leq \limsup_{n\to\infty} U_n(x,t) \leq U(x,t) + \varepsilon.
$$

Since ε is arbitrary,

$$\lim_{n\to\infty} U_n(x,t) = U(x,t)$$

as required. □

5. Convexity of bond prices

In this section we show that interest rate models have convex bond prices provided the drift β is not "too convex". More precisely, we need to require that $\beta_{xx} \le 2$ (in the sense of distributions). To see why this condition comes into play, consider the corresponding term structure equation, i.e. the terminal value problem

$$\begin{cases} U_t + \alpha U_{xx} + \beta U_x - xU = 0 \\ U(x,T) = g(x) \end{cases}$$

for some convex and decreasing pay-off function g. When using the PDE-approach we will in the sequel simplify the presentation by performing a standard change $\tau = T - t$ of the direction of time. By a slight abuse of notation, we use the same symbols β, σ, U, V etc. to denote the new functions depending on time τ to maturity rather than on the actual time t. With this new convention, the term structure equation becomes an *initial* value problem

$$\begin{cases} U_\tau = \alpha U_{xx} + \beta U_x - xU \\ U(x,0) = g(x). \end{cases}$$

Assume for the moment that all coefficients are regular enough. Also assume that there is a first point (x_0, τ_0) at which convexity is almost lost, i.e. that $U(x,\tau)$ is convex for $0 \le \tau \le \tau_0$ and $U_{xx}(x_0,\tau_0) = 0$. Then

$$\begin{aligned} \partial_\tau U_{xx} &= \partial_x^2 U_\tau = \partial_x^2(\alpha U_{xx} + \beta U_x - xU) \\ &= \alpha U_{xxxx} + (2\alpha_x + \beta)U_{xxx} + (2\beta_x - x)U_{xx} + (\beta_{xx} - 2)U_x. \end{aligned}$$

Since $x \mapsto U_{xx}(x,\tau_0)$ has a minimum at $x = x_0$, we have $U_{xxxx} \ge U_{xxx} = U_{xx} = 0$ at (x_0,τ_0). Thus we find that

$$\partial_\tau U_{xx} \ge (\beta_{xx} - 2)U_x \ge 0$$

at this point provided $\beta_{xx} \le 2$, i.e. the infinitesimal change of U at the critical point (x_0,τ_0) is convex. This suggests that convexity will be preserved if $\beta_{xx} \le 2$. Below we make this argument rigorous.

Theorem 5.1. *Assume that $\beta_{xx}(x,t) \le 2$ (in the sense of distributions) at all points $(x,t) \in \mathbb{R} \times [0,T]$. Also assume that the pay-off function g is convex, decreasing and satisfies (9). Then the corresponding option price $U(x,t)$ is convex in x at all times $t \in [0,T]$.*

Remark Note that all models in Table 1 satisfy the condition $\beta_{xx} \le 2$. Consequently, all those models give rise to convex bond prices.

Proof. We first assume that Hypothesis 3.4 holds, that $\beta_{xx} = 0$ for $x \ge C$ for some constant C, and that the pay-off function g is smooth with $e^{Kf(x)}g(x)$ being constant for large $|x|$, where K is the constant from (9).

Instead of studying the option price U directly, we first study the function $V = V^f$ defined in (7) for some function f satisfying Hypothesis 3.1 such that

$$\beta_{xx}(x,\tau) - 2f_x(x) \le 0 \qquad (27)$$

for all $(x,\tau) \in \mathbb{R} \times [0,T]$. Under these assumptions we claim that the function V is spatially convex.

To see this, note that it follows from Proposition 3.5 that there exists a constant C such that

$$|V_{xx}(x,\tau)| \leq Cp(x,\tau) \tag{28}$$

at all points, where

$$p(x,\tau) = e^{-f(x)K_0 e^{(D+1)\tau}}$$

(here K_0 is a large constant so that $K_0 \geq (e^{DT}-1)/D + Ke^{DT}$). For $\varepsilon > 0$, consider the function

$$V^\varepsilon(x,\tau) := V(x,\tau) + \varepsilon e^{M\tau}(x^2 + x^N p(x,\tau)).$$

Here the even number $N > 2$ is chosen large so that $x^2 + x^N p(x,\tau)$ has a strictly positive second spatial derivative at all points $(x,\tau) \in \mathbb{R} \times [0,T]$ (the constant M will be chosen below). Since V satisfies $\partial_\tau V = \mathcal{L}^f V$ where

$$\mathcal{L}^f = \alpha\partial_x^2 + \beta\partial_x - f,$$

we find that

$$\partial_x^2(\partial_\tau V^\varepsilon - \mathcal{L}^f V^\varepsilon) = \varepsilon\partial_x^2(\partial_\tau - \mathcal{L}^f)\left(e^{M\tau}(x^2 + x^N p)\right) = \varepsilon e^{M\tau}(MI_1 - I_2), \tag{29}$$

where

$$I_1 = \partial_x^2(x^2 + x^N p) > 0$$

and

$$I_2 = \partial_x^2\left(\alpha\partial_x^2(x^2 + x^N p) + \beta\partial_x(x^2 + x^N p) - f(x^2 + x^N p) + (D+1)K_0 e^{(D+1)\tau} f x^N\right).$$

For large positive x we have that f is constant, so p is bounded. Thus the term I_1 grows like x^{N-2}, whereas the term I_2 grows at most like x^{N-2} for large positive x. Similarly, for large negative x we have $f(x) = x$, so

$$I_1 \sim K_0^2 e^{2(D+1)\tau} x^N p,$$

whereas

$$I_2 \sim \alpha x^N p_{xxxx} + (-\beta K_0 e^{(D+1)\tau} - x + (D+1)K_0 e^{(D+1)\tau} x)x^N p_{xx}.$$

Using (5) we find that the highest order terms of I_2 behaves at most like

$$\alpha x^N p_{xxxx} + (-DK_0 e^{(D+1)\tau} - 1 + (D+1)K_0 e^{(D+1)\tau})x^{N+1} p_{xx},$$

and since $-DK_0 e^{(D+1)\tau} - 1 + (D+1)K_0 e^{(D+1)\tau} \geq 0$ we find that there exists a positive constant C such that

$$I_2 \leq Cx^N p$$

for large negative x. Consequently, I_1 grows at least as fast as I_2 for large values of $|x|$, so M can be chosen large so that MI_1 dominates I_2 everywhere. Actually, we will assume that M is chosen so large that

$$MI_1 > I_2 - I_3 \tag{30}$$

at all points, where

$$I_3 := (\beta_{xx} - 2f_x)\partial_x(x^2 + x^N p)$$

(compare (32) below). This can be done since $\beta_{xx} = 0$ and thus $I_3 = 0$ for large x, and $I_3 \geq 0$ for large negative x, compare (27).

Now, let

$$\Gamma := \{(x,\tau) : V^\varepsilon_{xx}(x,\tau) < 0\}.$$

We claim that the set Γ is empty. To see this, note that the second spatial derivative of $x^2 + x^N p$ grows at least like $Cx^N p$ as $|x| \to \infty$. Consequently, it follows from (28) that

$\Gamma \subseteq [-R,R]\times[0,T]$ for some $R>0$. Thus Γ is bounded and $\overline{\Gamma}$ is compact. Suppose that $\Gamma \neq \emptyset$, and define

$$\tau_0 := \inf\{\tau \geq 0 : (x,\tau) \in \overline{\Gamma} \text{ for some } x \in \mathbb{R}\}.$$

Due to compactness, the infimum is attained at (x_0,τ_0) for some x_0, and by continuity of V_{xx} we have $V^\varepsilon_{xx}(x_0,\tau_0)=0$, so $\tau_0>0$. Since $V^\varepsilon_{xx}(x_0,\tau)\geq 0$ for $\tau \leq \tau_0$, we must have

$$\partial_x^2\partial_\tau V^\varepsilon(x_0,\tau_0) = \partial_\tau\partial_x^2 V^\varepsilon(x_0,\tau_0) \leq 0. \tag{31}$$

Moreover, straightforward calculations give

$$\begin{aligned}\partial_x^2(\mathcal{L}^f V^\varepsilon) &= \partial_x^2(\alpha V^\varepsilon_{xx} + \beta V^\varepsilon_x - fV^\varepsilon)\\ &= \alpha V^\varepsilon_{xxxx} + (2\alpha_x+\beta)V^\varepsilon_{xxx} + (\alpha_{xx}+2\beta_x - f)V^\varepsilon_{xx}\\ &\quad +(\beta_{xx}-2f_x)V^\varepsilon_x - f_{xx}V^\varepsilon.\end{aligned}$$

Since $V^\varepsilon_{xx}(x_0,t_0)=0$ and V^ε is convex, the function $x\mapsto V^\varepsilon_{xx}(x,\tau_0)$ has a minimum point at x_0. Consequently, $V^\varepsilon_{xxx}(x_0,\tau_0)=0$ and $V^\varepsilon_{xxxx}(x_0,\tau_0)\geq 0$. Therefore, at the point (x_0,τ_0) we have

$$\begin{aligned}\partial_x^2(\mathcal{L}^f V^\varepsilon) &= \alpha V^\varepsilon_{xxxx} + (\beta_{xx}-2f_x)V^\varepsilon_x - f_{xx}V^\varepsilon\\ &\geq \varepsilon e^{M\tau}(\beta_{xx}-2f_x)\partial_x(x^2+x^N p) = \varepsilon e^{M\tau} I_3\end{aligned} \tag{32}$$

where we have used $V_x \leq 0$, (27) and that f is concave. Combining (29), (30), (31) and (32) yields

$$\partial_x^2(\partial_\tau V^\varepsilon - \mathcal{L}^f V^\varepsilon) = \varepsilon e^{M\tau}(MI_1 - I_2) > -\varepsilon e^{M\tau} I_3 \geq \partial_x^2(\partial_\tau V^\varepsilon - \mathcal{L}^f V^\varepsilon)$$

at (x_0,τ_0). This contradiction shows that the set E is empty, i.e. V^ε is convex. Letting $\varepsilon \downarrow 0$ it follows that also V is spatially convex at all times $\tau \in [0,T]$.

To deduce the convexity of U, consider an increasing sequence $\{f_i\}_{i=1}^\infty$ of functions f_i satisfying Hypothesis 3.1 such that $f_i(x)\to x$ as $i\to\infty$ for all x. Then

$$\int_t^T f_i(X_s)\,ds \to \int_t^T X_s\,ds$$

almost surely as $i\to\infty$. Therefore

$$\begin{aligned}V^{f_i}(x,t) &= E_{x,t}\left[\exp\left\{-\int_t^T f_i(X_s)\,ds\right\} g(X_T)\right]\\ &\to E_{x,t}\left[\exp\left\{-\int_t^T X_s\,ds\right\} g(X_T)\right] = U(x,t)\end{aligned}$$

as $i\to\infty$ by monotone convergence. Since each function V^{f_i} is spatially convex it follows that the option price $U(x,t)$ is spatially convex.

To remove the Hypothesis 3.4 and the assumption that $\beta_{xx}=0$ for large x, assume only Assumption 2.1 and that $\beta_{xx}\leq 2$ (in the sense of distributions). Let $\{\sigma_n\}_{n=1}^\infty$ and $\{\beta_n\}_{n=1}^\infty$ be sequences of smooth continuous coefficients satisfying Hypothesis 3.4 such that $(\beta_n)_{xx}=0$ for large x and $(\beta_n)_{xx}\leq 2$. Moreover, assume that σ_n and β_n satisfy the growth conditions (4) and (5) uniformly in n and that $\sigma_n\to\sigma$ and $\beta_n\to\beta$ uniformly on compacts as $n\to\infty$. It follows from Proposition 4.1 that

$$U(x,t) = \lim_{n\to\infty} U_n(x,t),$$

where the function U_n is defined as in (21). Since the pointwise limit of a sequence of convex functions is convex, it follows that U is convex.

Finally, to remove the assumptions about the smoothness of g and that $e^{Kf(x)}g(x)$ is constant outside a compact we approximate g from above by a sequence $\{g_n\}_{n=1}^\infty$

of smooth pay-offs behaving like $e^{-Kf(x)}$ outside compacts, such that $g_n(x) \downarrow g(x)$ as $n \to \infty$. By monotone convergence it follows that U is convex also without the smoothness requirements on g. □

The heuristic calculations in the beginning of this section indicate that the condition $\beta_{xx} \leq 2$ is not only a sufficient condition, but also a *necessary* condition for preservation of convexity for the term structure equation. Our next result shows that this is indeed true provided the coefficients are regular enough.

Theorem 5.2. *Assume that α and β are smooth. Also assume that $\alpha > 0$ and $\beta_{xx} > 2$ at some point (x_0, T). Then there exists an option with maturity T and with a decreasing and convex pay-off g such that the corresponding price $U(x,t)$ is non-convex at some time $t < T$.*

Proof. Let g be a smooth convex pay-off function which is linear and strictly decreasing in a neighborhood of x_0. Since $U(x,\tau)$ is a solution of a parabolic differential equation with regular coefficients, its derivatives exist and are continuous up to the boundary $\tau = 0$, see [23] (here we again let $\tau = T - t$). Straightforward calculations yield

$$\begin{aligned}\partial_\tau U_{xx} &= \partial_x^2 U_\tau = \partial_x^2(\alpha U_{xx} + \beta U_x - xU) \\ &= \alpha U_{xxxx} + (2\alpha_x + \beta)U_{xxx} + (\alpha_{xx} + 2\beta_x)U_{xx} + (\beta_{xx} - 2)U_x.\end{aligned}$$

Since g is linear in a neighborhood of x_0, we find that

$$\partial_\tau U_{xx}(x_0, 0) = (\beta_{xx} - 2)g_x > 0.$$

Since $U_{xx}(x_0, 0) = 0$, this means that U is not convex at some time $\tau > 0$. This finishes the proof. □

6. Parameter monotonicity

In this section we utilize the well-known connection between convexity and parameter monotonicity, see for instance [17] or [14]. We thus show how preservation of convexity implies that bond prices and prices of convex options are monotonically increasing in the volatility. To formulate the result, let X and $\tilde{X}$ be two diffusion processes satisfying

$$dX_t = \beta(X_t, t)\, dt + \sigma(X_t, t)\, dB_t$$

and

$$d\tilde{X}_t = \tilde{\beta}(\tilde{X}_t, t)\, dt + \tilde{\sigma}(\tilde{X}_t, t)\, dB_t,$$

respectively. Let

$$U(x,t) = E_{x,t}\left[\exp\left\{-\int_t^T X_s\, ds\right\} g(X_T)\right]$$

and

$$\tilde{U}(x,t) = E_{x,t}\left[\exp\left\{-\int_t^T \tilde{X}_s\, ds\right\} g(\tilde{X}_T)\right]$$

be the corresponding option prices.

Theorem 6.1. *Assume that $\beta(x,t) \leq \tilde{\beta}(x,t)$ and $|\sigma(x,t)| \geq |\tilde{\sigma}(x,t)|$ for all $(x,t) \in \mathbb{R} \times [0,T]$, and that either $\beta_{xx} \leq 2$ at all points or $\tilde{\beta}_{xx} \leq 2$ at all points (both in the distributional sense). Also assume that the pay-off function g is convex, decreasing and satisfies* (9)*. Then $\tilde{U}(x,t) \leq U(x,t)$ for all $(x,t) \in \mathbb{R} \times [0,T]$.*

Remark Monotonicity in the drift is immediate since g is decreasing and it is known that $\beta(x,t) \leq \tilde{\beta}(x,t)$ for all x and t implies that $X_T \leq \tilde{X}_T$ (if $\sigma(x,t) = \tilde{\sigma}(x,t)$ for all x and t). Theorem 6.1 tells us that bond prices are also increasing in the volatility, provided the condition $\beta_{xx} \leq 2$ is fulfilled. In particular, all models in Table 1 are monotonically increasing in the volatility and decreasing in the drift.

Proof. Let f be a given function satisfying Hypothesis 3.1, and define the functions

$$V(x,t) = E_{x,t}\left[\exp\left\{-\int_t^T f(X_s)\,ds\right\} g(X_T)\right]$$

and

$$\tilde{V}(x,t) = E_{x,t}\left[\exp\left\{-\int_t^T f(\tilde{X}_s)\,ds\right\} g(\tilde{X}_T)\right].$$

As in the proof of Theorem 5.1, we will use the parametrization in remaining time $\tau = T - t$ to maturity. Let

$$p(x,\tau) = e^{-f(x)K_0 e^{(D+1)\tau}},$$

where the positive constant D satisfies (4) and (5) for both models, and the constant K_0 is large so that $K_0 \geq (e^{DT} - 1)/D + Ke^{DT}$. For $\varepsilon > 0$, define

$$V^\varepsilon(x,\tau) = V(x,\tau) + \varepsilon e^{M\tau}(1 + x^2 p(x,\tau)).$$

It is straightforward to check that the constant M can be chosen so large that

$$M(1+x^2p) > -x^2\partial_\tau p + \alpha\partial_x^2(1+x^2p) + \beta\partial_x(1+x^2p) - f(1+x^2p) \tag{33}$$

at all points $(x,\tau) \in \mathbb{R}\times[0,T]$. Now, let

$$\Gamma := \{(x,\tau) \in \mathbb{R}\times[0,T] : \tilde{V}(x,\tau) < V^\varepsilon(x,\tau)\},$$

and suppose that $\Gamma \neq \emptyset$. Since $\tilde{V}p^{-1}$ is bounded by Lemma 3.2, and since $V^\varepsilon p^{-1}$ grows like x^2 for large $|x|$, there exists $R > 0$ such that $\Gamma \subseteq (-R,R)\times[0,T]$. Therefore, Γ is bounded and $\overline{\Gamma}$ is compact. Define

$$\tau_0 := \inf\{\tau : (x,\tau) \in \overline{\Gamma} \text{ for some } x\}.$$

By compactness, there exists $x_0 \in \mathbb{R}$ such that $(x_0,\tau_0) \in \overline{\Gamma}$. Since $V^\varepsilon - \tilde{V}$ is continuous, we have $V^\varepsilon(x_0,\tau_0) = \tilde{V}(x_0,\tau_0)$. Therefore, $V^\varepsilon(x,0) > \tilde{V}(x,0)$ implies that $\tau_0 > 0$. By the definition of τ_0 we have $V^\varepsilon(x_0,\tau) - \tilde{V}(x_0,\tau) \geq 0$ for $0 < \tau < \tau_0$, so

$$\partial_\tau(V^\varepsilon - \tilde{V}) \leq 0 \tag{34}$$

at (x_0,τ_0). On the other hand, V and $\tilde{V}$ satisfy the parabolic equations $V_\tau = \mathcal{L}^f V$ and $\tilde{V}_\tau = \tilde{\mathcal{L}}^f\tilde{V}$, respectively, where

$$\mathcal{L}^f = \alpha\partial_x^2 + \beta\partial_x - f \qquad \text{and} \qquad \tilde{\mathcal{L}} = \tilde{\alpha}\partial_x^2 + \tilde{\beta}\partial_x - f.$$

Consequently, at the point (x_0,τ_0) we have

$$\begin{aligned}
\partial_\tau(V^\varepsilon - \tilde{V}) &= \mathcal{L}^f V^\varepsilon - \tilde{\mathcal{L}}^f\tilde{V} + \varepsilon e^{M\tau}M(1+x^2p) \\
&\quad + \varepsilon e^{M\tau}\left(x^2\partial_\tau p - \alpha\partial_x^2(1+x^2p) - \beta\partial_x(1+x^2p) + f(1+x^2p)\right) \\
&> \alpha V^\varepsilon_{xx} + \beta V^\varepsilon_x - fV^\varepsilon - (\tilde{\alpha}\tilde{V}_{xx} + \tilde{\beta}\tilde{V}_x - f\tilde{V})
\end{aligned}$$

by (33). The function $x \mapsto V^\varepsilon(x,\tau_0) - \tilde{V}(x,\tau_0)$ attains its minimum 0 at $x = x_0$. Thus we have $V^\varepsilon = \tilde{V}$, $V^\varepsilon_x = \tilde{V}_x$ and $V^\varepsilon_{xx} \geq \tilde{V}_{xx}$ at the point (x_0,τ_0). Since at least one of the two models is convexity preserving, $V^\varepsilon_{xx} \geq 0$ at this point, so $\alpha V^\varepsilon_{xx} \geq \tilde{\alpha}\tilde{V}_{xx}$. Therefore, at (x_0,τ_0),

$$\partial_\tau(V^\varepsilon - \tilde{V}) > \alpha V^\varepsilon_{xx} - \tilde{\alpha}\tilde{V}_{xx} + (\beta - \tilde{\beta})\tilde{V}_x \geq 0,$$

where we also have used $\beta \leq \tilde{\beta}$ and $\tilde{V}_x \leq 0$. This contradicts (34). Consequently, Γ is empty, so $V^\varepsilon \geq \tilde{V}$ everywhere. Letting $\varepsilon \searrow 0$ it follows that $V \geq \tilde{V}$.

Finally, an application of the monotone convergence theorem (as f approaches x) yields that $U \geq \tilde{U}$. □

7. Bond options

In this section we apply the results of Sections 5 and 6 to study convexity and monotonicity properties of bond option prices. We consider a European call option with time of maturity T_1 and strike price $K > 0$ on a bond maturing at T_2, where naturally $T_2 > T_1$. We denote the price of this option at time $t \leq T_1$ by $C(x, t; T_1, T_2)$ with x denoting the short rate. Thus

$$C(x, t; T_1, T_2) = E_{x,t}\left[\exp\left\{-\int_t^{T_1} X_s\,ds\right\}(u(X_{T_1}, T_1) - K)^+\right],$$

where $U(x, t)$ is the value function of a T_2-bond. Assuming that the short rate dynamics is given by equation (3), we then have the following result.

Theorem 7.1. *Assume that $\beta_{xx}(x, t) \leq 2$ (in the sense of distributions). Then the bond call option price $C(x, t; T_1, T_2)$ defined above is convex in x at all times $t \in [0, T_1]$.*

Proof. We know from Theorem 5.1 (applied with $g = 1$) that the price u of the T_2-bond is convex in x. Note that an increasing function of the bond price is decreasing as a function of the short rate. Thus, since the pay-off function of the call is convex and increasing, it follows that $C(x, T_1; T_1, T_2) = (u(x, T_1) - K)^+$ is convex. Moreover, since u is decreasing in x so is $C(x, T_1; T_1, T_2)$. Since there exists a constant $M > 0$ such that

$$(u(x, T_1) - K)^+ \leq u(x, T_1) \leq M \max\{e^{Mx}, 1\},$$

see Corollary 3.3, the result follows from another application of Theorem 5.1. □

We also have a related monotonicity result. To formulate this, let X and $\tilde{X}$ be two diffusion processes satisfying

$$dX_t = \beta(X_t, t)\,dt + \sigma(X_t, t)\,dB_t$$

and

$$d\tilde{X}_t = \tilde{\beta}(\tilde{X}_t, t)\,dt + \tilde{\sigma}(X_t, t)\,dB_t.$$

Let C and $\tilde{C}$ be the corresponding call option prices written on the T_2-bond prices u and $\tilde{u}$, respectively. Then we have the following result.

Theorem 7.2. *Assume that $\beta(x, t) \leq \tilde{\beta}(x, t)$ and $|\sigma(x, t)| \geq |\tilde{\sigma}(x, t)|$ for all x and t. Also assume that either $\beta_{xx} \leq 2$ (in the distributional sense) at all points or $\tilde{\beta}_{xx} \leq 2$ at all points. Then*

$$\tilde{C}(x, t; T_1, T_2) \leq C(x, t; T_1, T_2)$$

for $t \leq T_1$.

Proof. First note that

$$g(x) := (u(x, T_1) - K)^+ \geq (\tilde{u}(x, T_1) - K)^+ =: \tilde{g}(x) \tag{35}$$

since $u \geq \tilde{u}$ by Theorem 6.1. If $\beta_{xx} \leq 2$ at all points, then g is decreasing and convex. It thus follows from Theorem 6.1 and (35) that

$$\begin{aligned} C(x,t;T_1,T_2) &= E_{x,t}\left[\exp\left\{-\int_t^{T_1} X_s\,ds\right\} g(X_{T_1})\right] \\ &\geq E_{x,t}\left[\exp\left\{-\int_t^{T_1} \tilde{X}_s\,ds\right\} g(\tilde{X}_{T_1})\right] \\ &\geq E_{x,t}\left[\exp\left\{-\int_t^{T_1} \tilde{X}_s\,ds\right\} \tilde{g}(\tilde{X}_{T_1})\right] = \tilde{C}(x,t;T_1,T_2). \end{aligned}$$

A similar argument can be applied if instead $\tilde{\beta}_{xx} \leq 2$ at all points. This finishes the proof. □

Remark As noted in Section 5, all models in Table 1 satisfy $\beta_{xx} \leq 2$. Thus all those models have bond call option prices which are convex, decreasing in the drift and increasing in the volatility.

8. Log-convexity of bond prices

In this section we study convexity properties of the logarithms of bond prices. Recall that a non-negative function u is said to be log-convex if $u(\lambda x_1 + (1-\lambda)x_2) \leq u(x_1)^\lambda u(x_2)^{1-\lambda}$ for all $\lambda \in (0,1)$ and x_1, x_2. If u is strictly positive, then log-convexity is equivalent to the function $x \mapsto \ln u(x)$ being convex.

For simplicity, we only deal with log-convexity of bond prices, so we assume that the pay-off function $g \equiv 1$. Thus

$$W(x,t) = e^{f(x)(e^{D(T-t)}-1)/D} E_{x,t}\left[\exp\left\{-\int_t^T f(X_s)\,ds\right\}\right]. \tag{36}$$

Recall from Lemma 3.2 that the function W is bounded on $\mathbb{R} \times [0,T]$. We first show that if the drift β is such that (15) is satisfied, then the function W is also bounded away from zero.

Lemma 8.1. *Assume that $g \equiv 1$, and that Hypothesis 3.1 and* (15) *hold. Then W defined in* (36) *is also bounded away from 0, i.e. there exists a constant $C > 0$ such that*

$$\frac{1}{C} \leq W(x,t) \leq C$$

on $\mathbb{R} \times [0,T]$.

Proof. Using Jensen's inequality, applied to the exponential function and the expectation, yields

$$\begin{aligned} W(x,t) &= E_{x,t}\left[\exp\left\{\int_t^T e^{D(s-t)} f(x) - f(X_s)\,ds\right\}\right] \\ &\geq \exp\left\{E_{x,t}\left[\int_t^T e^{D(s-t)} f(x) - f(X_s)\,ds\right]\right\} \\ &= \exp\left\{-\int_t^T e^{D(s-t)} E\left[Y_s\right]\,ds\right\}, \end{aligned}$$

where again $Y(s) = e^{-D(s-t)} f(X_s) - f(x)$ with initial condition $Y_t = 0$. Note that the condition (15) yields that the drift $\tilde{\beta}$ of Y is bounded also from above by some constant

C, so $E[Y_s] \leq C(s-t)$ for $s \in [t,T]$. Consequently,

$$W(x,t) \geq \exp\left\{\int_t^T -e^{D(s-t)}C(s-t)\,ds\right\} \geq \exp\{-Ce^{DT}T^2\} > 0$$

for all $t \in [0,T]$, which finishes the proof of the lemma. □

Theorem 8.2. *Assume that α is spatially convex and β is spatially concave. Then the bond price $u(x,t)$ is log-convex in the spot rate x.*

Remark It follows that all the models in Table 1 have log-convex bond prices.

Proof. We need to show that $\ln u(x,t)$ is spatially convex. Similar to the proof of Theorem 5.1, we approximate the function $\ln u$ by the function $\ln V = \ln V^f$ for some function f satisfying Hypothesis 3.1. We first assume that Hypothesis 3.4 holds. In addition to the bounds on the derivatives in Hypothesis 3.4, we also assume that

$$|\partial_x^k \alpha(x,t)| \leq C(1+x)^{1-k} \qquad k = 0,1,2 \tag{37}$$

for positive x.

It is straightforward to check that the function $P := \ln V^f$ satisfies the non-linear equation $P_\tau = \hat{\mathcal{L}}P$, where

$$\hat{\mathcal{L}}P = \alpha P_{xx} + \alpha P_x^2 + \beta P_x - f,$$

with initial condition $P(x,0) = 0$ (here we have again used the parametrization in terms of time τ to maturity rather than the physical time t). Moreover, there exists a constant C such that

$$|P_{xx}(x,\tau)| \leq \frac{C}{1+|x|^2} \tag{38}$$

for all $(x,\tau) \in \mathbb{R} \times [0,T]$. Indeed, for large $|x|$ we have $f_{xx}(x) = 0$, so

$$P_{xx} = (\ln V)_{xx} = (\ln W - fh)_{xx} = (\ln W)_{xx} = \frac{WW_{xx} - W_x^2}{W^2},$$

which decays like $|x|^{-2}$ according to Proposition 3.5 and Lemma 8.1.

Let $q : \mathbb{R} \to \mathbb{R}$ be a decreasing and convex function with strictly positive second derivative such that

$$q(x) = \begin{cases} -x - (\ln|x|)^2 & \text{if } x \leq -K \\ -(\ln x)^2 & \text{if } x \geq K \end{cases}$$

for some positive constant K. We now claim that there exists a positive constant M such that

$$Mq_{xx} > \partial_x^2(\alpha q_{xx} + \beta q_x + 2\alpha P_x q_x) + |\partial_x^2(\alpha q_x^2)| \tag{39}$$

at all points $(x,\tau) \in \mathbb{R} \times [0,T]$. To see this, note that q_{xx} behaves like $|x|^{-2}\ln|x|$ for large $|x|$. The estimates

$$|P_x + f_x(e^{D\tau} - 1)/D| \leq \frac{C}{1+|x|} \qquad \text{and} \qquad |P_{xxx}| \leq \frac{C}{1+|x|^3}$$

can be obtained in the same way as (38) was derived. Using these estimates, Hypothesis 3.4 and (37), it is straightforward to check that all terms in

$$\partial_x^2(\alpha q_{xx} + \beta q_x + 2\alpha P_x q_x) + |\partial_x^2(\alpha q_x^2)|$$

decay at least like $|x|^{-2}\ln|x|$. Consequently we can choose M large so that (39) holds.

Next, for $\varepsilon \in (0, e^{-MT})$, define

$$P^\varepsilon(x,\tau) := P(x,\tau) + \varepsilon e^{M\tau} q(x).$$

Since $P_\tau = \hat{\mathcal{L}}P$, straightforward calculations yield that

$$\begin{aligned} (40)\, \partial_x^2(P^\varepsilon_\tau - \hat{\mathcal{L}}P^\varepsilon) &= \varepsilon e^{M\tau} M q_{xx} - \varepsilon e^{M\tau}\partial_x^2(\alpha q_{xx} + \varepsilon e^{M\tau}\alpha q_x^2 + \beta q_x + 2\alpha P_x q_x) \\ &\geq \varepsilon e^{M\tau}\left(M q_{xx} - \partial_x^2(\alpha q_{xx} + \beta q_x + 2\alpha P_x q_x) - |\partial_x^2(\alpha q_x^2)|\right) > 0, \end{aligned}$$

where we used $\varepsilon e^{M\tau} \leq 1$ and (39).

It follows from (38) and the fact that q_{xx} behaves like $x^{-2}\ln|x|$ for large $|x|$ that the set

$$\Gamma := \{(x,\tau) \in \mathbb{R} \times [0,T] : P^\varepsilon_{xx}(x,\tau) < 0\}$$

is bounded. Thus, if Γ is non-empty, then there exists a point $(x_0, \tau_0) \in \overline{\Gamma}$ such that

$$\tau_0 = \inf\{\tau : (x,\tau) \in \overline{\Gamma} \text{ for some } x \in \mathbb{R}\}.$$

Since $P^\varepsilon_{xx}(x,0) > 0$ we have $\tau_0 > 0$. Consequently, at the point (x_0,τ_0) we have

$$\partial_x^2 P^\varepsilon_\tau = \partial_\tau P^\varepsilon_{xx} \leq 0.$$

Moreover,

$$\begin{aligned} \partial_x^2(\hat{\mathcal{L}}P^\varepsilon) &= \partial_x^2(\alpha P^\varepsilon_{xx} + \alpha(P^\varepsilon_x)^2 + \beta P^\varepsilon_x - f) \\ &= \alpha P^\varepsilon_{xxxx} + (2\alpha_x + \beta)P^\varepsilon_{xxx} + (\alpha_{xx} + 2\beta_x)P^\varepsilon_{xx} + \alpha_{xx}(P^\varepsilon_x)^2 \\ &\quad + 4\alpha_x P^\varepsilon_x P^\varepsilon_{xx} + 2\alpha P^\varepsilon_x V^\varepsilon_{xxx} + 2\alpha(P^\varepsilon_{xx})^2 + \beta_{xx}P^\varepsilon_x - f_{xx} \\ &\geq \beta_{xx}P^\varepsilon_x \geq 0, \end{aligned}$$

where the first inequality is due to $P^\varepsilon_{xxxx} \geq P^\varepsilon_{xxx} = P^\varepsilon_{xx} = 0$ at (x_0,τ_0), the assumption $\alpha_{xx} \geq 0$ and the concavity of f, and the second inequality follows from $P^\varepsilon_x \leq 0$ and the assumption $\beta_{xx} \leq 0$. Thus

$$\partial_x^2(P^\varepsilon_\tau - \hat{\mathcal{L}}P^\varepsilon) \leq 0 \tag{41}$$

at (x_0,τ_0). But this is a contradiction to (40), which shows that the set Γ is empty. Thus P^ε is convex. Letting ε tend to 0 we find that also P is convex at all times $\tau \in [0,T]$.

Now, if Hypothesis 3.4 and (37) are not satisfied, then we can approximate α and β with smooth coefficients as in the proof of Theorem 5.1. Using Proposition 4.1 and then letting $f \to x$ it is straightforward to check that also $\ln u$ is convex, which finishes the proof. □

9. Log-concavity of bond prices

In this section we discuss concavity properties of the logarithm $F = \ln u(x,t)$ of the bond price. Note that F satisfies the non-linear parabolic equation

$$F_\tau = \alpha F_{xx} + \alpha F_x^2 + \beta F_x - x \tag{42}$$

with initial condition $F(x,0) = 0$. In order to find the appropriate condition for preservation of log-concavity, we first present a heuristic argument similar to the one presented in the beginning of Section 5.

Assume that (x_0,τ_0) is a first point where concavity is almost lost, i.e. $x \mapsto F(x,\tau)$ is concave for all $\tau \leq \tau_0$, and $F_{xx}(x_0,\tau_0) = 0$. Differentiating (42) twice gives

$$\begin{aligned} \partial_\tau F_{xx} &= \partial_x^2 F_\tau = \partial_x^2(\alpha F_{xx} + \alpha F_x^2 + \beta F_x - x) \\ &= \alpha F_{xxxx} + (2\alpha_x + \beta)F_{xxx} + (\alpha_{xx} + 2\beta_x)F_{xx} + \beta_{xx}F_x \\ &\quad + \alpha_{xx}F_x^2 + 4\alpha_x F_x F_{xx} + 2\alpha(F_x F_{xxx} + F_{xx}^2). \end{aligned}$$

Since $F_{xxxx} \leq F_{xxx} = F_{xx} = 0$ at the point (x_0,τ_0) we get

$$\partial_\tau F_{xx} \leq \beta_{xx}F_x + \alpha_{xx}F_x^2$$

at that point. Thus, since $F_x \leq 0$, we see that a sufficient condition for preservation of log-concavity appears to be that β is convex and α concave. Since $\alpha = \sigma^2/2$ is non-negative, however, our convention that the model is specified on the whole real line is no longer convenient. Indeed, specifying σ to be 0 for negative short rates is not compatible with a concave infinitesimal variance α. The only possibility to have α concave on the whole real line is to require it to be constant in x. Such models can be shown to preserve log-concavity using the same methods as in Section 8.

Theorem 9.1. *Assume that α is only time-dependent, i.e. $\alpha(x,t) = \gamma(t)$ for some function γ. Also assume that β is convex in x for each fixed time t. Then the T-bond price $u(x,t)$ is log-concave in x at every fixed time $t \in [0,T]$.*

Proof. The proof follows along the same lines as Theorem 8.2 with some minor modifications. The function f of Hypothesis 3.1 needs to be replaced by a convex function which equals x for positive x and is constant for $x \leq -K$, where K is some positive constant. With this new f, Lemma 8.1 and Proposition 3.5 remain valid if we modify β to be linear for x large positive. □

The conditions in Theorem 9.1 are only satisfied by the models V and HW in Table 1. To investigate the remaining models, we are forced to leave the tractable setting of diffusions defined on the whole real line and instead consider models specified on a half-line. Such models, however, typically lead to partial differential equations with degenerate coefficients.

For simplicity, we assume that X is defined on the positive real axis $[0,\infty)$. We will also assume that $\alpha(0,t) = 0$ and $\beta(0,t) \geq 0$. Note that under these conditions, no boundary behavior of the process at $x = 0$ needs to be specified. Let

$$w(x,t) = e^{h(t)x}u(x,t), \tag{43}$$

where $h(t) = (e^{D(T-t)}-1)/D$. By arguing as in the proofs of Lemma 3.2 and Lemma 8.1, it can be shown that if $\beta - Dx$ and α are bounded for large x, then there exists a positive constant C such that

$$C^{-1} \leq w(x,t) \leq C \tag{44}$$

for all $(x,t) \in [0,\infty) \times [0,T]$. To apply the techniques used in previous sections, we also need estimates of the derivatives of the function w. We have the following result.

Lemma 9.2. *Assume that X is specified on the positive real axis with $\alpha(0,t) = 0$, $\beta(0,t) \geq 0$, $\alpha(x,t) > 0$ for $x \in (0,\infty)$, that α and β are smooth and that α and $\beta - Dx$ are constant in x for large x. Then there exists a constant $C > 0$ such that for $k = 0,1,2,...$ we have*

$$|\partial_x^k w(x,t)| \leq \frac{C}{1+x^k} \tag{45}$$

for $(x,t) \in (0,\infty) \times [0,T]$.

Proof. As noted above, the case $k = 0$ can be proven analogously to Lemma 3.2. It is straightforward to check that $\hat{w} := w - 1$ is the bounded classical solution to

$$\begin{cases} \hat{w}_t + \alpha\hat{w}_{xx} + \hat{\beta}\hat{w}_x + \gamma\hat{w} + \gamma = 0 & t < T \\ w = 0 & t = T, \end{cases} \tag{46}$$

where

$$\hat{\beta} = \beta - 2\alpha h \qquad \text{and} \qquad \gamma = \alpha h^2 + (Dx - \beta)h.$$

Note that

$$|\partial_x^k \hat{\beta}(x,t)| \leq C(1+x)^{1-k} \qquad \text{and} \qquad |\partial_x^k \gamma(x,t)| \leq C(1+x)^{-k} \tag{47}$$

for some constant C. By stochastic representation,

$$\hat{w}(x,t) = E\left[\int_t^T \exp\left\{\int_t^s \gamma(Y_r^{x,t}, r)\, dr\right\} \gamma(Y_s^{x,t})\, ds\right],$$

where $Y^{x,t}$ is a diffusion given by

$$dY_s^{y,t} = \hat{\beta}(Y_s^{x,t}, s) + \sigma(Y_s^{x,t}, s)\, dB_s$$

and the indices indicate that $X_t^{x,t} = x$. As in the proof of Lemma 3.5, the bounds in (47) and Gronwall's lemma can be applied to prove that $\hat{w}_x$ is bounded and satisfies $\hat{w}_x(x,T) = 0$. Thus $\hat{w}_x$ is a bounded solution to the equation obtained by differenting the equation (46), and using the maximum principle the estimate (45) can be established for $k = 1$. The rest of the proof follows inductively by treating the differentiated equation as above. □

Remark It follows from Lemma 9.2 that the derivatives $\partial_x^k \partial_t^l u$, $k + 2l \leq 4$, are continuous up to the spatial boundary $x = 0$.

Theorem 9.3. *Assume that X is specified on the positive real axis with $\alpha(0,t) = 0$ and $\beta(0,t) \geq 0$. Also assume that α is concave in x and β is convex in x at any fixed time $t \in [0,T]$. Then the bond price $u(x,t)$ is log-concave in x.*

Proof. We will assume that α and β satisfy the conditions in Lemma 9.2. We also assume that there exists a constant $\eta > 0$ such that $\alpha(x,t) = \gamma(t)x$ and $\beta_{xx}(x,t) = 0$ for $x \leq \eta$ and for some function $\gamma(t) \geq \eta$. The general case follows by approximation.

Define the function $F(x,\tau) = \ln u(x,\tau)$, where u is as in (1). As above, F satisfies the non-linear equation

$$F_\tau = \hat{\mathcal{L}}F$$

where

$$\hat{\mathcal{L}}F = \alpha F_{xx} + \alpha F_x^2 + \beta F_x - x.$$

It follows from (44) and Lemma 9.2 that

$$|F_x(x,t) + h(t)| \leq C(1+x)^{-1},$$
$$|F_{xx}(x,t)| \leq C(1+x)^{-2}$$

and

$$|F_{xxx}(x,t)| \leq C(1+x)^{-3}.$$

Let $q : (0,\infty) \to \mathbb{R}$ be a smooth, increasing and concave function with strictly negative second derivative such that

$$q(x) = \begin{cases} x - x^2 & \text{if } x < 1/C \\ (\ln x)^2 & \text{if } x > C \end{cases}$$

for some constant $C > 0$. We claim that there is a constant $M > 0$ so large that

$$Mq_{xx} < \partial_x^2(\alpha q_{xx} + \beta q_x + 2\alpha F_x q_x) - |\partial_x^2(\alpha q_x^2)| \tag{48}$$

at all points $(x,\tau) \in [0,\infty) \times [0,T]$. Indeed, the right hand side is bounded for small x, and for large values of x the right hand side decays at least as fast as $x^{-2}\ln x$. Consequently, M can be chosen so that (48) holds at all points.

Now, for $\varepsilon \in (0, e^{-MT})$, define

$$F^\varepsilon(x,\tau) = F(x,\tau) + \varepsilon e^{M\tau} q(x).$$

Then

$$\begin{aligned} (49)\ \partial_x^2(F_\tau^\varepsilon - \hat{\mathcal{L}}F^\varepsilon) &= \varepsilon e^{M\tau}\left(Mq_{xx} - \partial_x^2(\alpha q_{xx} + \varepsilon e^{M\tau}\alpha q_x^2 + \beta q_x + 2\alpha F_x q_x)\right) \\ &\leq \varepsilon e^{M\tau}\left(Mq_{xx} - \partial_x^2(\alpha q_{xx} + \beta q_x + 2\alpha F_x q_x) + |\partial_x^2(\alpha q_x^2)|\right) < 0 \end{aligned}$$

according to (48). Next, define the set

$$\Gamma = \{(x,t) \in \mathbb{R} \times [0,T] : F^\varepsilon_{xx}(x,t) < 0\}.$$

Since F_{xx} decays at least like x^{-2} we have that $\Gamma \subseteq [0,R) \times [0,T]$ for some R, so $\overline{\Gamma}$ is compact. Let (x_0, τ_0) be a point such that

$$\tau_0 = \inf\{\tau > 0 : (x,\tau) \in \Gamma \text{ for some } x \in [0,\infty)\}.$$

If $x_0 > 0$, then arguing as before the inequality (41), it is straightforward to check that

$$\partial_x^2 (F^\varepsilon_\tau - \hat{\mathcal{L}} F^\varepsilon) \geq 0$$

at (x_0, τ_0), which contradicts (49). Therefore, assume that $x_0 = 0$. Then, at the point (x_0, τ_0) we have

$$\partial_x^2 F^\varepsilon_\tau = \partial_\tau F^\varepsilon_{xx} \geq 0 \tag{50}$$

and

$$\begin{aligned} \partial_x^2(\hat{\mathcal{L}} F^\varepsilon) &= \partial_x^2(\alpha F^\varepsilon_{xx} + \alpha (F^\varepsilon_x)^2 + \beta F^\varepsilon_x - x) \\ &= \alpha_{xx} F^\varepsilon_{xx} + 2\alpha_x F^\varepsilon_{xxx} + \alpha F^\varepsilon_{xxxx} + \alpha_{xx} F^\varepsilon_{xx} + 4\alpha_x F^\varepsilon_x F^\varepsilon_{xx} \\ &\quad + 2\alpha (F^\varepsilon_x F^\varepsilon_{xxx} + (F^\varepsilon_{xx})^2) + \beta_{xx} F^\varepsilon_x + 2\beta_x F^\varepsilon_{xx} + \beta F^\varepsilon_{xxx} \leq 0, \end{aligned} \tag{51}$$

where we have used $\alpha = 0$, $\beta \geq 0$, $F_{xx} = 0$ and $F_{xxx} \leq 0$ at the point (x_0, τ_0). The inequalities (49), (51) and (50) form a contradiction. This shows that the set Γ is empty, so F^ε is convex at all times. Letting $\varepsilon \to 0$, it follows that also F is convex in x. □

Remark From the results in Sections 8 and 9 we find that if α and β are both concave and convex, i.e. linear, then both log-convexity and log-concavity are preserved. In our terminology such models thus give rise to log-affine bond prices. Of course, these models are usually referred to as having an *affine term structure* and they play an important rôle in interest rate theory. Our sufficient conditions for the existence of an affine term structure are well-known, see for instance Chapter 17 of [7]. However, the results of Sections 8 and 9 offer a background to these seemingly ad hoc conditions.

10. Conclusions

In this paper we have conducted a study of convexity of solutions to the term structure equation. We show that if the drift β satisfies $\beta_{xx} \leq 2$, then the bond prices are convex in the current short rate, increasing in the volatility of the short rate and decreasing in the drift. Similar results hold for call options written on bond prices. For models with regular coefficients, the condition $\beta_{xx} \leq 2$ is also a necessary condition for preservation of convexity. We also have a general comparison theorem: if a model has smaller drift and larger volatility than another model, and at least one of them has a drift satisfying the condition above, then the first model has the larger bond prices. For bond call options the analogous result holds.

We also study convexity properties of the logarithm of a bond price corresponding to the relative sensitivity of the bond price to changes in the short rate. This relative sensitivity is often described by the duration, i.e. the negative of the derivative of the logarithm. We show that if the drift β is concave and the square σ^2 of the volatility is convex, then bond prices are log-convex (a decreasing duration). Similarly, if β is convex and σ^2 is concave, then bond prices are log-concave (an increasing duration). We also note that if we demand that the price is log-convex *and* log-concave, we recover the well-known sufficient conditions for a model to admit an affine term structure.

Our findings for some commonly used models are summarized in Table 2 below.

Model	Dynamics	AB	AO	C	LCV	LCC
V	$dX = k(\theta - X)\,dt + \sigma\,dB$	Yes	Yes	Yes	Yes	Yes
CIR	$dX = k(\theta - X)\,dt + \sigma\sqrt{X}\,dB$	Yes	Yes	Yes	Yes	Yes
D	$dX = bX\,dt + \sigma X\,dB$	Yes	No	Yes	Yes	No
EV	$dX = X(\eta - a\ln X)\,dt + \sigma X\,dB$	No	No	Yes	Yes	No
HW	$dX = k(\theta_t - X)\,dt + \sigma\,dB$	Yes	Yes	Yes	Yes	Yes
BK	$dX = X(\eta_t - a\ln X)\,dt + \sigma X\,dB$	No	No	Yes	Yes	No
MM	$dX = X(\eta_t - (\lambda - \frac{\gamma}{1+\gamma t})\ln X)\,dt + \sigma X\,dB$	No	No	Yes	Yes	No

Table 2. In the three last columns it is indicated which models preserve convexity for option prices and bond call options, log-convexity of bond prices and log-concavity of bond prices, respectively.

References

[1] Alvarez, L.H.R. On the properties of r-excessive mappings for a class of diffusions. Ann. Appl. Probab. 13 (2003) 1517-1533.

[2] Alvarez, L.H.R. On the form and risk-sensitivity of zero coupon bonds for a class of interest rate models. Insurance Math. Econom. 28 (2001) 83-90.

[3] Bahlali, K., Mezerdi, B. and Ouknine, Y. Pathwise uniqueness and approximation of solutions of stochastic differential equations. Séminaire de Probabilités, XXXII, 166-187, Lecture Notes in Math., 1686, Springer, Berlin, 1998.

[4] Bellamy, N. and Jeanblanc, M. Incompleteness of markets driven by a mixed diffusion. Finance Stoch. 4 (2000), 209-222.

[5] Bergenthum, J., and Ruschendorf, L. Comparison of option prices in semimartingale models. Finance Stoch. 10 (2006), 222-249.

[6] Bergman, Y., Grundy, B. and Wiener, Z. General properties of option prices. J. Finance 51 (1996) 1573-1610.

[7] Björk, T. Arbitrage Theory in Continuous Time, Oxford University Press, New York, (1998).

[8] Björk, T. and Svensson, L. On the existence of finite-dimensional realizations for nonlinear forward rate models. Math. Finance 11 (2001), 205-243.

[9] Borell, C. Geometric inequalitites in option pricing, in: Convex Geometric Analysis, in: Math. Sci. Res. Inst. Publ., vol. 34, Cambridge Univ. Press, Cambridge, 1999, 29-51.

[10] Borell, C. Isoperimetry, log-concavity, and elasticity of option prices, in: P. Wilmott, H. Rasmussen (Eds.), New Directions in Mathematical Finance, Wiley, 2002, 73-91.

[11] Brigo, D. and Mercurio, F. Interest rate models - theory and practice. Springer Finance. Springer-Verlag, Berlin, 2001.

[12] Ekström, E. Properties of American option prices. Stochastic Process. Appl. 114 (2004), 265-278.

[13] Ekström, E., Janson, S. and Tysk, J. Superreplication of options on several underlying assets. J. Appl. Probab. 42 (2005), 27-38.

[14] Ekström, E. and Tysk, J. Properties of option prices in models with jumps. (2005) To appear in Math. Finance.

[15] Ekström, E. and Tysk, J. Convexity preserving jump-diffusion models for option pricing. To appear in J. Math. Anal. Appl. (2006).

[16] Ekström, E. and Tysk, J. The American put is log-concave in the log-price. J. Math. Anal. Appl. 314 (2006), 710-723.

[17] El Karoui, N., Jeanblanc-Picqué, M. and Shreve, S. Robustness of the Black and Scholes formula. Math. Finance 8 (1998), no. 2, 93-126.

[18] Eriksson, J. Monotonicity in the volatility of single-barrier options, Int. J. Theor. Appl. Finance 9 (2006) 987-996.

[19] Hobson, D. Volatility misspecification, option pricing and superreplication via coupling. Ann. Appl. Probab. 8 (1998) 193-205.

[20] Janson, S. and Tysk, J. Volatility time and properties of option prices. Ann. Appl. Probab. 13 (2003) 890-913.

[21] Janson, S. and Tysk, J. Preservation of convexity of solutions to parabolic equations. J. Differential Equations 206 (2004) 182-226.
[22] Kolesnikov, A.V. On diffusion semigroups preserving the log-concavity, J. Funct. Anal. 186 (2001) 196-205.
[23] Lieberman, G. M. Initial regularity for solutions of degenerate equations. Nonlinear Anal. 14 (1990), 525-536.
[24] Merton, R. Theory of rational option pricing. Bell J. Econom. and Management Sci. 4 (1973), 141-183.
[25] Revuz, D. and Yor, M. Continuous martingales and Brownian motion. Third edition. Grundlehren der Mathematischen Wissenschaften, 293. Springer-Verlag, Berlin, 1999.
[26] Yong, J. Remarks on some short rate term structure models. J. Ind. Manag. Optim. 2 (2006) 119-134.

E-mail address: ekstrom@math.uu.se, johan.tysk@math.uu.se

Department of Mathematics, Uppsala University, Box 480, 75106 Uppsala, Sweden