FDIC Center for Financial Research
Working Paper

No. 2008-02

Valuing Convertible Bonds with Stock Price, Volatility, Interest Rate, and Default Risk*

January 2008

Federal Dposit Insurance Corporation • Center for Financial Research

Valuing Convertible Bonds with Stock Price, Volatility, Interest Rate, and Default Risk*

Pavlo Kovalov† **Vadim Linetsky**‡

January 4, 2008

Abstract

This paper develops a computational framework to value convertible bonds in general multi-factor Markovian models with credit risk. We show that the convertible bond value function satisfies a variational inequality formulation of the stochastic game between the bondholder and the issuer. We approximate the variational inequality by a penalized nonlinear partial differential equation (PDE). We solve the penalized PDE formulation numerically by applying a finite element spatial discretization and an adaptive time integrator. To provide specific examples, we value and study convertible bonds in affine, as well as non-affine, models with four risk factors, including stochastic interest rate, stock price, volatility, and default intensity.

Keywords: Convertible bonds, credit risk, volatility skew, credit spreads, stochastic games, variational inequalities, penalty approximation, finite element method-of-lines

*This research was supported by the grants from the FDIC Center for Financial Research, Moody's Credit Markets Research Fund, and the National Science Foundation under grant DMI-0422937.

†Quantitative Risk Management, Inc., Chicago, IL.

‡Department of Industrial Engineering and Management Sciences, McCormick School of Engineering and Applied Sciences, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, Phone: (847) 491-2084, E-mail: linetsky@iems.northwestern.edu, Web: http://users.iems.northwestern.edu/∼linetsky.

1 Introduction

Convertible bonds are (generally callable) corporate bonds with an additional contractual feature: the bondholder is allowed to convert the bond into a pre-specified number of shares of common stock at any time prior to maturity. With well-developed convertible bond markets in the U.S., Europe, and Japan, the problem of consistent modeling, valuation, and risk management of convertible bonds is of significant practical importance. As any corporate debt, convertible bonds are subject to interest rate risk and credit risk. The additional conversion option explicitly introduces dependence on the stock price. It is a well-established empirical fact that the stock price volatility is stochastic and is negatively correlated with the stock price (the so-called *leverage effect*). This negative dependence manifests itself in the *implied volatility smile/skew* patterns across strikes observed in stock options prices in the equity options markets. Furthermore, it is also a well-established empirical fact that corporate bond and credit default swap (CDS) spreads are positively correlated with the stock volatility of the underlying reference firm. Thus, in order to accurately model, value, and risk manage convertible bonds one needs a model that incorporates stochastic dynamics of interest rates, stock price, stochastic volatility, and credit risk, and captures the key empirical relationships among the stock price process, the volatility process, and corporate credit spreads. Consistent modeling of these risk factors and their dependences is especially important to firms engaged in *convertible bond arbitrage*, a significant segment of the global hedge fund universe.

Starting from Ingersoll (1977) and Brennan and Schwartz (1977), the early literature on convertible bonds followed a one-factor structural firm-value approach along the lines pioneered by Merton (1974). In this approach, one starts with a stochastic process for the firm value (typically geometric Brownian motion), and all the corporate securities issued by the firm are treated as contingent claims on the firm value. Brennan and Schwartz (1980) introduced a second stochastic factor into this structural framework, by modeling stochastic default-free interest rates (see also Nyborg (1996) for a good summary of this early literature; for recent advances in the structural approach see Sirbu et al. (2004), Sirbu and Shreve (2005), and Bermudez and Webber (2003)). While appealing from the corporate finance theory standpoint, the structural approach has not been widely adopted by the convertible bond practitioner community for at least two reasons. First, the structural approach is necessarily highly stylized. A detailed model would require modeling all of the corporate securities issued by the firm simultaneously. This might involved modeling dozens of different straight and convertible bond issues, taking into account seniority. Secondly, the firm value process is not directly observable. The observable data in the market include the stock price, credit spreads (corporate bond yields and, more recently, CDS spreads in the credit derivatives market), and prices (and implied volatilities) of stock options. In practice, a practitioner would have a difficult time ascertaining any precise numerical value for the value of the firm, not to mention the difficulties in estimating its process parameters, such as volatility. In contrast, the stock price, corporate credit spreads, and implied volatilities of equity options are continuously observed in the market. From the practical standpoint, it makes sense to develop convertible bond models that can be calibrated to the prices of liquid benchmark securities, such as stock options and CDS, and then used to value convertible bonds. Moreover, since the firm's common stock, plain vanilla stock options, CDS, and straight corporate debt are all available for trading, convertible bond arbitrageurs need models that price convertible bonds consistently with these more liquid instruments.

This dissatisfaction with the structural models lead practitioners to propose a number of low-factor (typically one- or two-factor) models where the issues of credit risk were handled in a somewhat ad-hoc manner by either splitting the convertible bond into its fixed income and equity components and discounting the components at different rates (e.g., Tsiveriotis and Fernandes (1998)) or adjusting discount rates according to somewhat ad-hoc rules depending on the stock price level (e.g., Bardhan et al. (1994)). However, these models lack consistent theoretical underpinnings.

On the other hand, the intensity-based reduced-form credit risk modeling literature has enjoyed remarkable development over the past decade (see the recent monographs by Bielecki and Rutkowski (2002), Duffie and Singleton (2003), Lando (2004), and Schonbucher (2003) for state-of-the-art surveys). Modeling of convertible bonds in this modern intensity based framework was initiated by Davis and Lischka (2002), who proposed a convertible bond model that incorporated a Black-Scholes stock price (equity risk), a stochastic short rate (interest rate risk), and a default intensity (hazard rate of default) dependent on the stock price (credit risk linked with the stock price level). Such models with stochastic stock price, stochastic interest rate, and default intensity taken to be a deterministic function of the stock price became known in the industry as "two-and-a-half-factor models" (see Andersen and Buffum (2004), Takahashi et al. (2001), and Ayche et al. (2003) for detailed studies of "one-and-a-half-factor models" with the default intensity taken to be a deterministic function of the stock price, and deterministic interest rates). Several typical specifications of the default intensity as a function of the underlying stock price were later solved in closed form in the case of European-style securities by Linetsky (2006) and Carr and Linetsky (2006) (unfortunately these solutions do not extend to convertible bonds, which are American-style securities). However, all of these models still fail to account for the empirical fact that while the default intensity and credit spreads are strongly influenced by the stock price, they are not perfectly correlated. Moreover, these models fail to take into account stochastic volatility of the stock price, an essential empirical feature in the realm of stock options modeling.

Recently, Carr and Wu (2005) introduced an interesting three-factor reduced-form affine model of default. In this model, the stock price drops to zero at default. Prior to default, the stock price follows a continuous process with stochastic volatility. The default intensity and the instantaneous stock variance follow a bi-variate diffusion process intricately specified to capture empirical evidence on stock option prices and corporate debt and CDS spreads. Carr and Wu (2005) show in an extensive empirical study that their model is well suited for joint modeling and valuation of stock options and credit default swaps. This model is also well suited to serve as a basis for a convertible bond model. However, while Carr and Wu's main interest is in relatively shorter maturity stock options and CDS spreads, they take the short rate as deterministic. In contrast, convertible bonds have maturities up to thirty years. For example, a recent Eastman Kodak convertible issue has thirty year maturity (on October 15, 2003 Eastman Kodak & Co. issued $575 million of 3.375% convertible senior notes due October 15, 2033). While it may be a reasonable approximation to assume constant interest rates for the valuation of shorter maturity securities, valuing long-term bonds assuming constant interest rates is highly unrealistic and prone to serious errors.

The contributions of the present paper to the convertible bond literature are four-fold.

- (i) First, we develop a mathematical framework to value convertible bonds in general multi-factor Markovian models with credit risk. We show that, when the underlying uncertainty is modeled as a Markov process with killing (to account for default), the convertible bond value function satisfies a variational inequality formulation of the stochastic game between the bondholder and the issuer (fundamental references on the theory of variational inequalitites and stochastic games with stopping times include Friedman (1976) and Bensoussan and Lions (1982, 1984)). We approximate the variational inequality by a penalized nonlinear partial differential equation (PDE) with two penalty terms corresponding to call and conversion constraints. In particular, we introduce a new class of penalty terms with continuous Jacobians that have computational advantages over the standard penalty term with discontinuous Jacobian that has been previously used in the literature on American option valuation.

- (ii) Second, we develop a computational framework to solve the multi-dimensional nonlinear PDE with two penalty terms in the finite element method-of-lines framework. We localize the PDE to a bounded computational domain. We then re-formulate it in the variational (weak) form and discretize it spatially using multi-dimensional finite elements. The result is a non-linear ODE system. We integrate it in time using an adaptive time integrator SUNDIALS based on the backward differentiation formulae (BDF). We provide an error analysis, and show that adaptivity of the time integrator is essential for efficient numerical solution of the problem.

- (iii) Third, to provide specific examples, we develop convertible bond valuation in affine, as well as non-affine, four-factor models with stochastic interest rate, stock price, volatility, and default intensity. Our affine model formulation is a four-factor extension of the three-factor Carr and Wu (2005) reduced-form affine model. At default, the stock price drops to zero. Prior to default, the stock price follows a continuous process with stochastic volatility. The default intensity, the instantaneous stock variance, and the default-free short rate follow a tri-variate diffusion process specified to capture empirical evidence on stock option prices and CDS spreads. We also introduce two non-affine extensions of this model that allow the default intensity to explicitly depend on the stock price to account for the so-called *collapse of the bond floor* well known to practitioners.

- (iv) Fourth, we study the convertible bond value function and its Greeks (sensitivities with respect to the underlying risk factors) necessary for hedging convertible bonds as functions of the underlying factors, and determine and analyze optimal conversion and call strategies in these models. To the best of our knowledge, this paper is the first attempt at convertible bond modeling that includes four stochastic risk factors influencing convertible bond values. The existing literature has so far been limited to one- and two-factor models.

We note that the finite element method has been recently applied to convertible bond valuation in an interesting work by Barone-Adesi et al. (2003), who applied it to solve the stochastic game between the bond issuer and the bondholder. However, Barone-Adesi et al. (2003) do not consider credit risk at all. They work with two-factor models with stochastic stock price and stochastic default-free interest rate, and assume no credit risk. This assumption is introduced to simplify the model and its numerical solution, but is unrealistic, especially in light of the fact that it is typically the riskier firms with lower credit ratings that issue convertible bonds (e.g., the previously cited recent Eastman Kodak convertible bond issue is currently rated B by Standard & Poor — a "junk" bond rating).[1] In contrast, the focus of the present paper is on convertible bond valuation with credit risk.

[1] According to a recent study by Moody's Investor Services (see Hamilton et al.(2001)), between 1970 and

To conclude this brief survey of the convertible bonds literature, we mention an interesting recent work by Bielecki et al. (2005) who study a general formulation of the convertible bond problem as a game option in an abstract semimartingale market model set-up (see also related work by Kallsen and Kuhn (2005)). This paper develops rigorous theoretical underpinnings of the convertible bond valuation in the general semimartingale framework. In the present paper we limit ourselves to the Markovian framework leading to PDEs, as our primary focus here is computational.

The rest of this paper is organized as follows. In Section 2 we describe the structure of a typical convertible bond contract and fix our notation. In Section 3 we introduce a general Markovian modeling framework with credit risk, as well as our example four-factor affine model for the underlying economic uncertainty, its non-affine extensions, and three-factor special cases (reductions). In Section 4, following the theoretical framework of Bensoussan and Lions (1982, 1984), we formulate the convertible bond valuation problem in the Markovian setting as a stochastic game between the issuer and the bondholder and present its variational inequality formulation. In Section 5 we present a penalty approximation, which approximates the value function of the variational inequality with the solution of a non-linear PDE with the appropriately chosen penalty terms approximating the action of conversion and call constraints. We note that, in contrast to the rest of the literature on American-style options that use penalty formulations with penalty terms with discontinuous Jacobians (e.g., Forsyth and Vetzal (2002), Sapariuc et al. (2004)), we use penalty terms with continuous Jacobians to improve computational performance of the PDE solver. In Section 6 we present a variational (weak) formulation of the penalized non-linear PDE localized to a bounded domain in the state space. In Section 7 we describe our approach to the numerical solution. We work in the finite element method-of-lines framework. We discretize the problem spatially with finite element basis functions. The resulting system of non-linear ODEs is integrated in time using an adaptive solver SUNDIALS developed by the Lawrence Livermore National Laboratory. This solver features adaptively variable integration order and step size selectors and is suitable for non-linear and non-smooth problems. Equipped with this computational framework, in Section 8, after presenting extensive numerical experiments and convergence studies to validate our computational approach, we value some representative convertible bonds, compute their Greeks, study the behavior of convertible bond prices and Greeks as functions of the underlying risk factors, and analyze optimal conversion and call strategies. Section 9 concludes the paper. Analytical solutions for European-style securities in the four-factor affine model are presented in the Appendix. They are used as benchmarks for convergence studies in Section 8.

2 Convertible Bonds

Consider a convertible bond issued by a certain firm at time $t = 0$ with maturity date $T > 0$, face value $F > 0$ to be repaid at maturity, and coupon amounts $C_i > 0$ paid at coupon payment dates $\{t_i = i\delta, i = 1, 2, \ldots, N\}$, where $t_N = T$ and $\delta > 0$ is the time interval between two coupons (typically $\delta = 1/2$ in the U.S. corporate bond market, corresponding to semiannual coupon

2000, 280 convertible bond issuers have defaulted on US$86.7 billion of long-term convertible debt. Default rates for convertible debt issuers are higher than for those issuers without convertible bonds in their capital structure. Furthemore, recovery rates for defaulted convertible bonds are significantly lower than those for non-convertible bonds, recovering $29 on average compared with $43 per $100 par for straight bonds. Thus, the issue of credit risk is especially important in convertible bond valuation and analysis.

payments). According to the market practice, we assume that the coupon payments are equal and $C_i = c\delta F$, where c is the annualized percentage rate to be applied to the face value F. We assume that the coupon payments are equally spaced with the interval δ (assumptions of equal coupons and equally spaced coupon payment dates are to simplify notation and are without loss of generality). The *accrued interest* at time $t \in (i\delta, (i+1)\delta]$ between the two coupons is given by: $A(t) = c(t - i\delta)F$. The bondholder has an option to convert the bond into $k > 0$ shares of common stock of the issuer firm at any time prior to and including maturity. A *conversion factor* $k > 0$ specifies the number of shares of stock to be received in exchange for surrendering the bond. In the U.S. convertible market, no accrued interest is paid at the time of conversion.

The issuer has an option to call the bond after an initial *call protection period* $[0, t_c)$, $t_c \in [0, T]$ (t_c is called the *first call* date). At the time of call $t \in [t_c, T)$, the issuer re-purchases the bond from the bondholder for a (dirty) *call price* $C_d(t)$. The bond prospectus typically specifies a *call schedule* that stipulates the *clean call price* $C_c(t)$ in effect during each year of the bond's life when the bond is callable. If the bond is called at some time t during the year, the actual (or *dirty*) call price $C_d(t)$ to be paid to the bondholder is the clean call price plus the interest $A(t)$ accrued since the last coupon payment, $C_d(t) = C_c(t) + A(t)$. The bondholder's conversion option has a precedence over the issuer's call option. That is, if the issuer calls a bond at time t, the bondholder has the right to either convert the bond at t and receive k shares of the underlying stock worth kS_t or receive the dirty call price $C_d(t)$ equal to the clean call price plus the accrued interest.

Sometimes the bondholder also has an option to put the bond back to the issuer at some specified times during the life of the bond for some pre-specified amount (typically the bond face value plus accrued interest). In the U.S. convertible market, the put option is typically *Bermudan-style* with only several exercise opportunities during the life of the bond (typically several years apart). This is in contrast to the call and conversion options, which are American-style, granting continuous exercise opportunities. In this paper to simplify notation we do not consider the put option (but it can be easily incorporated in our valuation framework, by checking optimality of exercising the put at the eligible exercise dates). In this paper we also do not consider the so-called *soft call protection* features in some convertible bonds that restrict the issuer's right to call the bond based on the price history of the stock, as this introduces path-dependencies that significantly complicate the analysis. For more detailed discussions of various features of convertible bonds and market practices we refer the reader to the literature (e.g., Grau et al. (2003), Greemwood and Hodges (2002), Howard and O'Connor (2001), Lau and Kwok (2004)).

As any corporate debt, the convertible bond is subject to default by the issuer firm. In the event of default, we assume that the value of the common stock drops to zero, and the convertible bond pays a fixed recovery payment equal to a fraction $R \in [0, 1]$ of the face value F (the *recovery of face value* assumption). Other recovery assumptions can also be considered in our framework, but for simplicity we limit ourselves to the recovery of face value. This is also consistent with the actual market practice of what typically happens in bankruptcy proceedings (see Andersen and Buffum (2003)). We also make a simplifying assumption that at default the stock price drops to zero (no recovery to the stockholders). This simplification is quite reasonable for practical purposes. At the time of default the stock typically trades for relatively low values (typically for pennies in the U.S.). However, if desired, some non-zero recovery on the common stock can be incorporated in the model as shown by Ayche et al. (2003).

3 The Markovian Modeling Framework

3.1 General Framework

Let λ_t denote the (generally stochastic) arrival rate of the default event. Let S_t denote the time-t stock price. We assume that *prior to default* the risk-neutral stock price dynamics is:

$$dS_t = (r_t - q + \lambda_t)S_t dt + \sqrt{V_t} S_t dW_t^S, \tag{3.1}$$

where r_t is the (generally stochastic) instantaneous default-free interest rate in the economy (the short rate), q is the dividend yield on the stock (assumed constant), V_t is the (generally stochastic) instantaneous variance rate, and W_t^S is a standard Brownian motion. When default occurs, we assume that the stock price drops to zero, where it remains forever (zero is a *cemetery* state for the stock price process). The default intensity λ_t in the drift in (3.1) compensates for the possibility of a jump to zero, so that the discounted total return process (including stock price changes, dividends, and a possible jump to default) remains a martingale in the risk-neutral economy (in this paper we take a risk-neutral probability measure as given and consider all processes under the given risk-neutral measure).

In our general Markovian framework, we assume that the default intensity and the instantaneous stock return variance are functions of the stock price and an n-dimensional state vector Z following an n-dimensional Markov diffusion process $\{Z_t, t \geq 0\}$, and the short rate is a function of Z:

$$\lambda_t = \lambda(Z_t, S_t), \;\; V_t = V(Z_t, S_t), \;\; r_t = r(Z_t),$$

$$dZ_t^i = \mu_i(Z_t)dt + \sigma_i(Z_t)dW_t^i, \;\; i = 1, ..., n,$$

where W_t^i is an n-dimensional Brownian motion with the correlation structure $dW_t^i dW_t^j = \rho_{i,j}dt$ and $dW^S dW^i = \rho_{S,i}dt$, and μ and σ are the (generally state-dependent) drift and volatility of Z. They are assumed to satisfy the technical conditions ensuring that the SDE for the state vector Z has a unique strong solution.

3.2 A Four-Factor Affine Specification

To simplify the model calibration process, it is convenient to specify the model dynamics so that it is in the affine class of models (see Duffie et al. (2000)). For affine models the characteristic function can be obtained analytically, and European-style securities, such as options and straight corporate bonds and CDS, can be valued by inverting the Fourier transform. In this paper we study the following affine model specification. We assume that the short rate r_t, the instantaneous variance V_t, and the default intensity λ_t follow the joint dynamics:

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r \sqrt{r_t}\, dW_t^r, \tag{3.2a}$$

$$dV_t = \kappa_V(\theta_V - V_t)dt + \sigma_V \sqrt{V_t}\, dW_t^V, \tag{3.2b}$$

$$dz_t = \kappa_z(\theta_z + \gamma V_t - z_t)dt + \sigma_z \sqrt{z_t}\, dW_t^z, \tag{3.2c}$$

$$\lambda_t = z_t + \alpha V_t + \beta r_t, \tag{3.2d}$$

$$dW_t^S dW_t^V = \rho_{SV} dt, \tag{3.2e}$$

$$dW_t^S dW_t^r = dW_t^S dW_t^z = dW_t^r dW_t^V = dW_t^r dW_t^z = dW_t^V dW_t^z = 0. \tag{3.2f}$$

The short rate is assumed to follow a Cox-Ingersoll-Ross (CIR) (1985) process with the long-run level $\theta_r > 0$, rate of mean reversion $\kappa_r > 0$, and short rate volatility $\sigma_r > 0$. To calibrate to the initial yield curve we can also take $\theta_r = \theta_r(t)$ to be a deterministic function of time, but for simplicity here we assume it is constant. The instantaneous variance is also assumed to follow a CIR process with the long-run variance level $\theta_V > 0$, rate of mean reversion $\kappa_V > 0$, and volatility of variance $\sigma_V > 0$. The default intensity λ_t is specified to be a linear combination of three factors: a process z_t with dynamics (3.2c), a stock price variance dependent contribution αV_t with $\alpha \geq 0$ and an interest rate dependent contribution βr_t with $\beta \geq 0$. The correlation structure among the four driving standard Brownian motions is given by (3.2e-f). The only non-zero correlation is ρ_{SV} between W^S and W^V. We assume that $\rho_{SV} < 0$ to capture the leverage effect, i.e., the negative correlation between the stock price and volatility.

This model is a four-factor extension of the three-factor model recently proposed by Carr and Wu (2005). Carr and Wu's model obtains if we assume constant short rate and set $\beta = 0$. Since in this paper we are dealing with long-term convertible bonds, it is important to allow for stochastic default-free interest rates in addition to the stochastic stock price, volatility, and default intensity. We refer the reader to Carr and Wu (2005) for a thorough discussion and empirical and economic justification of their three-factor model and, in particular, of their specification of default intensity $\lambda_t = z_t + \alpha V_t$ (linear dependence of the default intensity on the instantaneous stock price variance rate is also employed by Carr and Linetsky (2006) in the context of their jump-to-default extended CEV model). All the arguments of Carr and Wu (2005) remain valid in our four-factor setting. The rather intricate specification of default intensity is in agreement with vast empirical evidence on the joint dynamics of the stock price volatility and credit default swap spreads on the same reference firm presented in Carr and Wu (2005). In particular, the stochastic variance V_t influences the default intensity in two ways. In addition to the obvious linear functional dependence of the default intensity on the variance in (3.2d), the process (3.2e) for the default intensity factor z_t features a mean-reverting drift with the stochastic long-run level $\theta_z + \gamma V_t$ with $\gamma \geq 0$ and $\theta_z > 0$. As the stock volatility increases, the long-run level of the default intensity factor z_t increases, and the process is pulled towards this increasing long-run level with the mean-reversion rate $\kappa_z > 0$. The state space of the four-dimensional process $\{(S_t, V_t, r_t, z_t), t \geq 0\}$ is $\Omega = (\mathbb{R}_+)^4$. The affine model (3.2) is a particular instance of the general Markovian framework of Section 3.1 with the three-dimensional process Z: $Z_t^1 = V_t$, $Z_t^2 = r_t$, $Z_t^3 = z_t$, and $\lambda = Z_t^3 + \alpha Z_t^1 + \beta Z_t^2$.

Carr and Wu's model can be thought of as a jump-to-default extended Heston's stochastic volatility model with stochastic default intensity that is affine in the instantaneous stock return variance. We further extend it by introducing stochastic interest rates. We also include the interest rate contribution to the default intensity βr_t with $\beta \geq 0$. As interest rates increase, the cost of issuing new debt increases, which puts upward pressure on default intensity.

Remark 3.1 (On Nested Lower-Dimensional Affine Models). The four-factor model (3.1)-(3.2) nests three three-dimensional affine models as special cases.

(1) If we set $\kappa_r = 0$ and $\sigma_r = 0$ we obtain Carr and Wu's three-factor model with constant interest rates. This model can be used to value shorter-term securities, where the impact of stochastic interest rates is relatively less important than the impact of stochastic volatility and stochastic default intensity.

(2) Another possibility is to set $\kappa_z = 0$ and $\sigma_z = 0$ so that z is constant. In this model the default intensity is $\lambda_t = z + \alpha V_t + \beta r_t$ with constant z, and all the dynamics of the intensity comes from changes in the stock price volatility and the short rate. This three-factor specification is

a simpler extension of Heston's stochastic volatility model than the four-factor specification. However, empirical results in Carr and Wu suggest that the CDS spread dynamics includes an additional independent factor in addition to the stock volatility.

(3) Another three-factor affine model specification is obtained by setting $\kappa_V = 0$ and $\sigma_V = 0$ so that V is constant. In this specification the volatility is constant, and the default intensity is independent of the stock price, as the Brownian motions W^S and W^z are independent. Note that if we make W^S and W^z correlated, the model will cease to be in the affine class and will have no analytical solutions for European securities (however, the numerical method developed in this paper will equally apply). These three-factor reductions provide lower-dimensional alternatives to the four-factor model, if constraints on the computing burden prevent one from using the full four-factor model.

Remark 3.2 (On the Calibration of Affine Models). A major advantage of the affine model specification is the availability of analytical solutions for European-style securities that greatly simplify the calibration process. The user can calibrate the model to a cross section of stock options with different strike prices and maturities and to credit spreads (e.g. CDS spreads) of different maturities (as well as the risk-free yield curve) by using the analytical solutions. The calibration process fixes model parameter values in Eqs.(3.1)-(3.2). Then, after the model is calibrated, the user prices a convertible bond by a numerical PDE method. Note that the calibration is done using only the analytical solutions, and does *not* require solving the PDE numerically. The PDE is solved numerically *only once* at the stage of pricing the convertible bond after the model has been calibrated to interest rate, CDS and stock options data using the affine solution. Moreover, the availability of analytical solutions for European-style contracts allows one to benchmark numerical methods needed to value American-style convertible bonds. This is a second important advantage of affine models. The computational performance of numerical implementation can be optimized using the analytical benchmark for the European case, and the optimized implementation can then be used to value the American-style CB more efficiently.

3.3 Some Non-Affine Model Specifications

One limitation of the affine specification is the inability to make the default intensity λ explicitly depend on the stock price S. Indeed, including in the default intensity an additional term with the negative power of the stock price S_t^{-p} that explodes as the stock falls towards zero as in, e.g., Carr and Linetsky (2006) and Linetsky (2006), destroys the affine structure. While in the affine model (3.1)-(3.2) the stock price does influence the default intensity dynamics through the negative correlation between the stock price S and the stock variance V, as well as through the contribution of V into the drift of z, the default intensity remains finite as the stock price falls to zero. In the context of convertible bond modeling, this limits our ability to model the so-called *collapse of the bond floor* well-known to practitioners. For low stock prices, the conversion option is deep out of the money and the bond behaves essentially as a straight corporate bond (this is known as the *bond floor* for the convertible bond). It is well known from empirical data that when the stock price falls to very low levels, market prices of corporate bonds issued by the firm rapidly fall (and credit spreads increase sharply) as market participants anticipate that a default event is much more likely when the firm's stock price falls to very low levels. To capture this feature, we consider several non-affine modifications of the affine model presented in the previous section. One possibility is to add a stock-dependent term into the default intensity

(3.2d):

$$\lambda_t = z_t + \alpha V_t + \beta r_t + a S_t^{-p}$$

with some $p > 0$ and $a \geq 0$. When z, V, and r are all constant, this model reduces to the jump-to-default extended Black-Scholes model recently solved by Linetsky (2006). However, this analytical solution of the one-factor model does not extend to multi-factor models.

An alternative specification is to make the instantaneous stock return variance depend both on the stock price itself in the constant-elasticity-of-variance (CEV) fashion and on an independent stochastic volatility factor:

$$V_t = v_t S_t^{2\beta},$$

where $\beta < 0$ is the CEV elasticity parameter, and v_t follows a CIR process:

$$dv_t = \kappa_v(\theta_v - v_t)dt + \sigma_v \sqrt{v_t}\, dW_t^v,$$

where the Brownian motion can be taken to be independent of W^S due to the fact that we already have a built-in leverage effect between the variance V and the stock price S through the negative CEV power $\beta < 0$. If we substitute this hybrid stochastic volatility-CEV (SVCEV) specification for the instantaneous stock variance into the default intensity specification (3.2d), we obtain:

$$\lambda_t = z_t + \alpha v_t S_t^{2\beta} + \beta r_t.$$

If z, v, and r are all constant, this model reduces to the one-factor jump-to-default extended CEV model (or JDCEV for short) recently introduced and solved in closed form by Carr and Linetsky (2006). However, this analytical solution does not extend to multi-factor models.

From the stand-point of our numerical method, the non-affine models discussed in this section do not present any additional computational difficulties in pricing convertible bonds. The advantage of such non-affine specifications is the ability to capture the explosion of the default intensity and the collapse of the bond floor in the convertible bond valuation. However, their disadvantage is that the lack of analytical solutions for stock options, corporate bonds, and CDS significantly complicates model calibration, as one now needs to solve for option prices and credit spreads by numerically solving the PDE within the calibration process, which will be too slow for practical purposes. Alternatively, one could use Monte Carlo simulation to estimate option prices and credit spreads by simulation nested within the calibration procedure. This would still be significantly slower than the calibration of affine models using the analytical solutions and the Fourier transform inversion.

3.4 Pricing European-style Securities

Consider a European-style contingent claim that delivers some payoff $\Psi(S_T)$ at expiration time $T > 0$ if the company does not default prior to and including time T, and delivers a fixed recovery payment $R \geq 0$ at the time of default if default occurs prior to and including expiration. According to the standard intensity-based credit risk theory (see the monographs cited in the Introduction), assuming no default by time $t \geq 0$, the time-t price of such a contingent claim can be written as

$$U(t, S, V, r, z) = \mathbb{E}_t \left[e^{-\int_t^T (r_u + \lambda_u) du} \Psi(S_T) \right] \tag{3.3a}$$

$$+\mathbb{E}_t \left[\int_t^T e^{-\int_t^s (r_u + \lambda_u) du} \lambda_s R ds \right], \tag{3.3b}$$

where $\mathbb{E}_t$ is the expectation operator under the risk-neutral probability measure and conditional on the information at time t. Since the model is Markovian, the only relevant information at time t are the values of the four state variables $S_t = S$, $V_t = V$, $r_t = r$, and $z_t = z$.

In particular, for a defaultable zero-coupon bond with unit face value and with recovery $R \in [0, 1]$ at the time of default we have:

$$B_R(t, S, V, r, z) = \mathbb{E}_t \left[e^{-\int_t^T (r_u + \lambda_u) du} \right] \tag{3.4a}$$

$$+R \int_t^T \mathbb{E}_t \left[e^{-\int_t^s (r_u + \lambda_u) du} \lambda_s \right] ds. \tag{3.4b}$$

For a European-style call option with payoff $(S_T - K)^+$ if the firm has not defaulted prior to and including T and zero if default occured by T, we have:

$$C(t, S, V, r, z) = \mathbb{E}_t \left[e^{-\int_t^T (r_u + \lambda_u) du} (S_T - K)^+ \right]. \tag{3.5}$$

The European-style put price can then be determined from the put-call parity.

A key feature of the three-factor Carr-Wu model and its four-factor version with stochastic interest rates considered here is that they are in the affine class (Duffie et al. (2000)). Accordingly, the expectation (3.4a) appearing in the defaultable bond pricing formula is exponential affine in the values of the factors r_t, V_t, and z_t at time t with time-dependent coefficients satisfying a system of ODEs of the Riccati type. The expectation appearing in the call pricing formula (3.5) can be calculated by inverting the following generalized discounted characteristic function of the log-return $\ln(S_T/S_t)$ (Fourier transform):

$$\varphi(u) := \mathbb{E}_t \left[e^{-\int_t^T (r_u + \lambda_u) du} e^{u \ln(S_T/S_t)} \right], \tag{3.6}$$

where u is a complex variable taking values in the domain $\mathcal{D} \subset \mathbb{C}$ of the complex plane where the expectation in (3.6) is well-defined (note that it is always well defined at least for purely imaginary u). This characteristic function has the exponential affine form and is calculated in Carr and Wu (2005) in the three-factor model. The four-factor extension with stochastic short rate is given in the Appendix. To calculate the recovery part (3.4b) of the defaultable zero-coupon bond, we need to compute the expectation $\mathbb{E}_t \left[e^{-\int_t^s (r_u + \lambda_u) du} \lambda_s \right]$. Since λ_t is affine in the underlying factors, this expectation also has the exponential affine structure. See the Appendix for details.

A *European-style convertible bond* with conversion allowed only at maturity and no call feature can be expressed as a portfolio of defaultable zero-coupon bonds with no recovery corresponding to the coupon payments prior to maturity, a contingent claim that pays

$$\Psi(S_T) = \max\{(1 + c\delta)F, kS_T\} = (1 + c\delta)F + k \max\{S_T - (1 + c\delta)F/k, 0\}$$

at bond maturity T if there is no default prior to and including maturity, and a recovery payment RF at the time of default if default occurs prior to and including maturity. The time t-value of the European-style convertible bond is then (here $[x]$ denotes the integer part of $x \in \mathbb{R}$ and the sum is over the remaining coupons):

$$ECB(t, S, V, r, z) = \sum_{i=[t/\delta]+1}^{N-1} c\delta F \mathbb{E}_t \left[e^{-\int_t^{i\delta} (r_u + \lambda_u) du} \right] + (1 + c\delta) F \mathbb{E}_t \left[e^{-\int_t^T (r_u + \lambda_u) du} \right]$$

$$+k\mathbb{E}_t\left[e^{-\int_t^T(r_u+\lambda_u)du}(S_T-K)^+\right]+RF\int_t^T\mathbb{E}_t\left[e^{-\int_t^s(r_u+\lambda_u)du}\lambda_s\right]ds, \tag{3.7}$$

where the strike of the conversion option component is

$$K:=(1+c\delta)F/k. \tag{3.8}$$

All the terms in this expression for the European-style convertible bond can be computed in the affine framework by Fourier inversion (see the Appendix).

While our main interest in the present paper is in the valuation of American-style convertible bonds that are analytically intractable due to American-style call and conversion options and require numerical PDE methods, the analytical solutions for European-style contingent claims such as defaultable zero-coupon bonds and call and put options provide useful benchmarks for numerical methods. Furthermore, the availability of analytical solutions for corporate bonds, CDS and options greatly facilitates calibration of the model to the credit and equity options data as demonstrated in Carr and Wu (2005) in the three-factor case. The calibrated model can then be used to value convertible bonds via the numerical method studied in this paper.

4 Valuing Convertible Bonds: A Stochastic Game with Stopping Times

Let

$$\psi_1(t,S):=kS$$

be the payoff to the bondholder in the case of conversion at time t with the stock price equal to S. Let

$$\psi_2(t,S):=\max\{C(t),kS\}$$

be the payoff to the bondholder in the case of call at time t by the issuer with the stock price equal to S at t, where $C(t)$ is the dirty call price (including interest accrued since the last coupon). The payoff of the convertible bond at maturity T is

$$\Psi(S_T)=\max\{(1+c\delta)F,kS_T\},$$

Let $\mathbb{F}=\{\mathcal{F},t\geq 0\}$ be the filtration generated by the Brownian motion W driving our model. For $t\in[0,T]$, let $\Theta_{t,T}$ be the set of all $\mathbb{F}$-stopping times taking values in $[t,T]$. For $t\in[0,T]$, let $\theta_1\in\Theta_{t,T}$ and $\theta_2\in\Theta_{t\vee t_c,T}$ be two stopping times. The θ_1 will model the time of conversion by the bondholder and the θ_2 — the time of call by the issuer (recall that conversion is allowed at any time during the life of the bond, while the bond is callable only during the period $[t_c,T]$ for some $t_c\geq 0$). Let $J(\theta_1,\theta_2;t,S,V,r,z)$ be the time-t value of the bond, assuming the bondholder's conversion strategy is given by the stopping time θ_1 and the issuer's call strategy is given by the stopping time θ_2. Then, assuming no default by time $t\in[0,T]$, we have (here and in what follows $x\wedge y:=\min\{x,y\}$ and $x\vee y:=\max\{x,y\}$):

$$J(\theta_1,\theta_2;t,S,V,r,z)=\mathbb{E}_t\left[e^{-\int\limits_t^T(r_s+\lambda_s)ds}\Psi(S_T)\mathbf{1}_{\{T\leq\theta_1\wedge\theta_2\}}\right.$$

$$
+\sum_{i=[t/\delta]+1}^{N-1} \mathbb{E}_t \left[e^{-\int\limits_t^{i\delta} (r_s+\lambda_s)ds} c\delta F \mathbf{1}_{\{i\delta<\theta_1\wedge\theta_2\}} \right] + \mathbb{E}_t \left[e^{-\int\limits_t^{\theta_1} (r_s+\lambda_s)ds} \psi_1(\theta_1, S_{\theta_1}) \mathbf{1}_{\{\theta_1<T\wedge\theta_2\}} \right]
$$

$$
+\mathbb{E}_t \left[e^{-\int\limits_t^{\theta_2} (r_s+\lambda_s)ds} \psi_2(\theta_2, S_{\theta_2}) \mathbf{1}_{\{\theta_2\leq\theta_1,\theta_2<T\}} \right] + \mathbb{E}_t \left[\int_t^{\theta_1\wedge\theta_2\wedge T} e^{-\int\limits_t^{s} (r_u+\lambda_u)du} \lambda_s R ds \right]. \tag{4.1}
$$

The first term is the time-t present value of the terminal payoff to be received if there is no default, no call, and no conversion prior to maturity. The second term is the present value of the coupon stream, where each coupon is to be received if there is no default, no call, and no conversion prior to the coupon date $t_i = i\delta$, $i = [t/\delta]+1, \ldots, N-1$ (the final coupon at maturity $t_n = T$ is included in the terminal payoff and, thus, is not included in the coupon stream). The third term is the present value of the payoff at the time of conversion, assuming the bondholder converts prior to the issuer's call, prior to default, and prior to maturity. The fourth term is the present value of the payoff at the time of call, assuming the issuer calls prior to conversion, prior to default, and prior to maturity. The fifth and final term is the present value of the recovery payment in the case of default prior to call, conversion, and maturity. For the future development it will also be convenient to write the value of the coupon stream in the continuous form:

$$
\sum_{i=[t/\delta]+1}^{N-1} \mathbb{E}_t \left[e^{-\int\limits_t^{i\delta} (r_s+\lambda_s)ds} c\delta F \mathbf{1}_{\{i\delta<\theta_1\wedge\theta_2\}} \right] = \mathbb{E}_t \left[\int_t^{\theta_1\wedge\theta_2\wedge T} e^{-\int_t^s (r_u+\lambda_u)du} \mathcal{C}(s) ds \right], \tag{4.2}
$$

where we introduced a *coupon rate*: $\mathcal{C}(t) := \sum_{i=[t/\delta]+1}^{N-1} C_i \delta(t-t_i)$, where $\delta(x)$ is the Dirac delta function, and $C_i = \delta c F$ and $t_i = i\delta$ are the coupon amounts and payment dates, respectively. This framework accommodates both discrete and continuous coupons. If one wishes to consider a continuous coupon stream paid at the continuous rate c, then one simply sets $\mathcal{C}(t) = cF$, so that the interest earned in the infinitesimal time interval dt is $cFdt$. Continuous coupons are often considered in the literature to simplify the analysis, although in practice coupons are always paid at discrete time intervals.

We assume that the issuer and the bondholder behave rationally, i.e., the bondholder is value-maximizing and the issuer is value-minimizing by optimally choosing the stopping times θ_1 and θ_2, respectively. That is, the bondholder chooses the conversion time to maximize the bond value, while the issuer chooses the call time to minimize the bond value. Then, the convertible bond value is a saddle point of $J(\theta_1, \theta_2; t, S, V, r, z)$:

$$
\begin{aligned}
U(t,S,V,r,z) &= \inf_{\theta_2\in\Theta_{t\vee t_c,T}} \sup_{\theta_1\in\Theta_{t,T}} J(\theta_1,\theta_2;t,S,V,r,z) \\
&= \sup_{\theta_1\in\Theta_{t,T}} \inf_{\theta_2\in\Theta_{t\vee t_c,T}} J(\theta_1,\theta_2;t,S,V,r,z).
\end{aligned} \tag{4.3}
$$

This is an example of a *stochastic game with stopping times.* For the mathematical theory of stochastic games with stopping times and, in particular, for general existence and uniqueness results for saddle points and for the variational inequality formulation we refer to the fundamental references Friedman (1975, Chapter 16, Sections 16.9-16.12) and Bensoussan and Lions (1982, Chapter 5, Section 5.2) and Bensoussan and Lions (1984, Chapter 2, Section 2.9).

The time-t price of the convertible bond $U = U(t, S, V, r, z)$ (the value of the stochastic game (4.3)) can be determined as a solution of the following system of parabolic variational inequalities. Find a function U, $\psi_1 \leq U \leq \psi_2$, such that (see Bensoussan and Lions (1982), Eqs. (5.87-5.88) and Theorem 5.5, pp. 489-490 and Bensoussan and Lions (1984), Eqs. (9.52) and (9.54) and Theorem 9.5, pp. 166–167):

$$\begin{aligned} &\tfrac{\partial U}{\partial t} - \mathcal{A}U - (r+\lambda)U + \lambda R + \mathcal{C} = 0 \ \text{ if } \ \psi_1 < U < \psi_2, \\ &\tfrac{\partial U}{\partial t} - \mathcal{A}U - (r+\lambda)U + \lambda R + \mathcal{C} \leq 0 \ \text{ if } \ U = \psi_1, \\ &\tfrac{\partial U}{\partial t} - \mathcal{A}U - (r+\lambda)U + \lambda R + \mathcal{C} \geq 0 \ \text{ if } \ U = \psi_2, \end{aligned} \tag{4.4}$$

with the terminal condition at maturity T:

$$U(T, S, V, r, z) = \Psi(S). \tag{4.5}$$

Here $\mathcal{C} = \mathcal{C}(t)$ is the previously defined coupon rate, $\lambda = z + \alpha V + \beta r$, and $\mathcal{A}$ is the second-order differential operator $\mathcal{A} = -\mathcal{G}$, a negative of the infinitesimal generator $\mathcal{G}$ of the underlying diffusion process. In the particular case of the four-dimensional process (3.1)–(3.4) the infinitesimal generator has the form:

$$\mathcal{G}u := \frac{1}{2}\left(VS^2\frac{\partial^2 u}{\partial S^2} + \sigma_V^2 V\frac{\partial^2 u}{\partial V^2} + \sigma_r^2 r\frac{\partial^2 u}{\partial r^2} + \sigma_z^2 z\frac{\partial^2 u}{\partial z^2} + 2\rho_{SV}\sigma_V VS\frac{\partial^2 u}{\partial S\partial V}\right)$$
$$+(r - q + \lambda)S\frac{\partial u}{\partial S} + \kappa_V(\theta_V - V)\frac{\partial u}{\partial V} + \kappa_z(\theta_z + \gamma V - z)\frac{\partial u}{\partial z} + \kappa_r(\theta_r - r)\frac{\partial u}{\partial r}.$$

The problem is considered on the time interval $[0, T]$. In the absence of call and conversion options prior to maturity, the problem reduces to pricing a European-style convertible bond that can only be converted at maturity. Its value function is given by Eq. (3.14), where the expectations can be computed in the affine framework. The value function of the bond also satisfies the PDE $U_t - \mathcal{A}U - (r + \lambda)U + \lambda R + \mathcal{C} = 0$ in $[0, T] \times \Omega$ subject to the terminal payoff condition (4.5). American-style call and conversion options constraint the value function of the convertible bond to stay between the conversion and call values, $\psi_1 \leq U \leq \psi_2$. Intuitively, the space $\mathcal{Q} = [0, T] \times \Omega$ separates into three regions. The conversion region (it is optimal for the bondholder to convert the bond), where the conversion constraint is binding, $U = \psi_1$, and the inequality $U_t - \mathcal{A}U - (r + \lambda)U + \lambda R + \mathcal{C} \leq 0$ holds, the call region (it is optimal for the issuer to call the bond), where the call constraint is binding, $U = \psi_2$, and the inequality $U_t - \mathcal{A}U - (r+\lambda)U + \lambda R + \mathcal{C} \geq 0$ holds, and the continuation region (neither call nor conversion are optimal, and neither call nor conversion constraints are binding) where U satisfies the PDE $U_t - \mathcal{A}U - (r + \lambda)U + \lambda R + \mathcal{C} = 0$. The three regions are separated by the two boundaries: the optimal conversion boundary and the optimal call boundary. The stopping times θ_1^* and θ_2^* that realize the optimal strategies for the bondholder and the issuer in (4.3) are given by (Bensoussan and Lions (1982, Theorem 5.5, p. 490), and Bensoussan and Lions (1984, Theorem 9.6, Eq. (9.60), p. 167):

$$\begin{aligned} \hat{\theta}_1 &= \inf\{s \in [t, T] : U(s, S_s, V_s, r_s, z_s) = \psi_1(s, S_s)\}, \\ \hat{\theta}_2 &= \inf\{s \in [t \vee t_c, T] : U(s, S_s, V_s, r_s, z_s) = \psi_2(s, S_s)\}, \end{aligned} \tag{4.6}$$

where $U = U(t, S, V, r, z)$ is the solution of the parabolic variational inequality (4.4)–(4.5).

If the convertible bond is callable, it is always optimal to convert the bond at time t if $kS_t \geq C(t)$. Thus, the value function satisfies

$$U(t, S, V, r, z) = kS, \tag{4.7}$$

for those (t, S, V, r, z) for which $S \geq S^*(t)$, where $S^*(t) = C(t)/k$. Therefore, it is sufficient to consider the problem (4.4)–(4.5) only for those (t, S, V, r, z) for which $S \leq S^*(t)$, and then extend the solution to the whole space $[0, T] \times \Omega$ by using (4.7).

5 Penalty Approximation and Localization to a Bounded Domain

In order to numerically solve the variational inequality formulation (4.4)–(4.5), we construct a penalty approximation as follows. Fix some small $\varepsilon > 0$ and consider a non-linear PDE of the form (here $x^+ := \max\{x, 0\}$ and $x^- := \max\{-x, 0\}$):

$$\frac{\partial U_\varepsilon}{\partial t} - \mathcal{A}U_\varepsilon - (r + \lambda)U_\varepsilon + \lambda R + \mathcal{C} + \frac{1}{\varepsilon}(U_\varepsilon - \psi_1)^- - \frac{1}{\varepsilon}(U_\varepsilon - \psi_2)^+ = 0 \tag{5.1}$$

subject to the terminal condition

$$U_\varepsilon(T, S, V, r, z) = \Psi(S). \tag{5.2}$$

The two non-linear penalty terms $\frac{1}{\varepsilon}(U_\varepsilon - \psi_1)^-$ and $-\frac{1}{\varepsilon}(U_\varepsilon - \psi_2)^+$ approximate the action of the conversion and call constraints, respectively. According to Friedman (1976, Chapter 16) and Bensoussan and Lions (1982, Theorem 5.6, p. 490) and Bensoussan and Lions (1984, Theorem 9.5, p. 167), the solution U_ϵ of the non-linear penalized PDE problem (5.1)–(5.2) converges to the solution of the variational inequality (4.4)–(4.5) as $\varepsilon \to 0$. In particular, the following penalization error estimate holds for the penalty approximation (Brezzi and Norrie (1987), Section 9.4b, Boman (2001), Sapariuc et al. (2004)):

$$\max_{t \in [0,T]} \|U_\varepsilon(t, \cdot) - U(t, \cdot)\|_{L^\infty(\Omega)} \leq C\varepsilon \tag{5.3}$$

for some constant[2] $C > 0$ independent of ε, and

$$\psi_1 - C\varepsilon \leq U_\varepsilon \leq \psi_2 + C\varepsilon. \tag{5.4}$$

The intuition is as follows. While the solution U of the variational inequality is constrained to be not less than the conversion value and not greater than the call value, $\psi_1 \leq U \leq \psi_2$, the solution of the non-linear penalized PDE U_ε can fall below the conversion value ψ_1 or raise above the call value ψ_2. However, while when $\psi_1 \leq U \leq \psi_2$ both penalty terms vanish, when $U_\varepsilon < \psi_1$, the first penalty term $\frac{1}{\varepsilon}(U_\varepsilon - \psi_1)^-$ has a positive value that rapidly increases as the solution falls below ψ_1, forcing the solution above ψ_1. Similarly, when $U_\varepsilon > \psi_2$, the second penalty term $-\frac{1}{\varepsilon}(U_\varepsilon - \psi_2)^+$ has a negative value with the rapidly increasing absolute value as the solution

[2] In sections 5-7 C denotes positive constants in various error estimates. It should not lead to any confusion with the previously used notation for the call price and for the coupon rate.

raises above ψ_2, forcing the solution below ψ_2. The smaller the value of ε, the larger the value of the coefficient $1/\varepsilon$ in the penalty terms and the more closely the penalty terms approximate the action of the conversion and call constraints. The penalized solution U_ε can fall below ψ_1 and raise above ψ_2, but $\psi_1 - U_\varepsilon \leq C\varepsilon$ and $U_\varepsilon - \psi_2 \leq C\varepsilon$ for some constant $C > 0$ independent of ε.

While the specific functional form of the penalty terms in (5.1) is commonly used in the literature on the penalty approximation of variational inequalities (e.g., Bensoussan and Lions (1982, 1984), Forsyth and Vetzal (2002), Friedman (1976), Glowinski (1981, 1984), Marcozzi (2001), Sapariuc et al. (2004), Zvan et al. (1998)), more general forms of the penalty term can be considered (see Glowinski (1984) for general results about the penalty method). In fact, the penalty terms $\frac{1}{\varepsilon}(U_\varepsilon - \psi_1)^-$ and $-\frac{1}{\varepsilon}(U_\varepsilon - \psi_2)^+$ have discontinuous first derivatives with respect to U_ε. In the numerical solution, one needs to use Newton-type iterations to solve a non-linear system of algebraic equations resulting from the discretization of the PDE. The discontinuity in the Jacobian of this system stemming from the discontinuity in the derivative of the penalty term with respect to U_ε presents some computational challenges, as non-smooth Newton-type iterative schemes for non-linear systems with discontinuous Jacobians need to be used (see, e.g., Forsyth and Vetzal (2002)). In this paper we consider more general penalty terms of the form $(\frac{1}{\varepsilon}(U - \psi_1)^-)^p$ and $-(\frac{1}{\varepsilon}(U - \psi_2)^+)^p$ for some $p \geq 1$, which we call power-p penalty terms. Taking $p > 1$ restores the continuity of the derivative of the penalty term with respect to U_ε, and standard Newton iterations with continuous Jacobian can be used. In the numerical experiments in this paper we take $p = 2$ and verify that the estimate (5.3) holds in this case as well. Thus, for some small $\varepsilon > 0$, we solve the non-linear PDE

$$\frac{\partial U_\varepsilon}{\partial t} - \mathcal{A}U_\varepsilon - (r + \lambda)U_\varepsilon + \lambda R + \mathcal{C} + \left(\frac{1}{\varepsilon}(U_\varepsilon - \psi_1)^-\right)^p - \left(\frac{1}{\varepsilon}(U_\varepsilon - \psi_2)^+\right)^p = 0 \tag{5.5}$$

subject to the terminal condition (5.2).

Since the state space Ω is unbounded in our problem, in order to be able to solve the problem numerically we need to localize it to a bounded computational domain. To this end, we consider a sequence of problems formulated on increasing bounded domains that exhaust the state space Ω. Let $\{\Omega_k\}_{k=1}^\infty$ be a sequence of increasing bounded open domains such that $\cup_{k=1}^\infty \Omega_k = \Omega$. We consider a sequence of PDE problems (5.5) subject to the terminal condition (5.2) posed on bounded domains Ω_k. For the problem on the bounded domain Ω_k to be well posed, we introduce artificial boundary conditions on the boundary $\partial\Omega_k$. For the localized problem, we have

$$\max_{t \in [0,T]} ||U_{\varepsilon,k}(t, \cdot) - U_\varepsilon(t, \cdot)||_{L^\infty(G)} \to 0 \text{ as } k \to \infty \tag{5.6}$$

for any compact set $G \subset \Omega$ such that $G \subset \Omega_k$ for all k and for essentially arbitrary choice of the artificial boundary conditions, as long as the problem is well-posed (see Friedman (1976, Chapter 16), Bensoussan and Lions (1982, 1984), Marcozzi (2001), Sapariuc et al. (2004)). The result (5.6) says that the behavior of the solution near the distant boundary $\partial\Omega_k$ does not affect the solution on any fixed bounded region $G \subset \Omega_k$ in the limit $k \to \infty$. Therefore, any well-posed problem on Ω_k is suitable as an approximation of the original problem on the unbounded state space Ω, provided that k is taken sufficiently large and the quality of the approximation is considered only on some fixed approximation domain $G \subset \Omega_k$. This justifies the use of essentially arbitrary artificial boundary conditions. The domain $G \subset \Omega_k$ where we are interested in the solution is referred to as the *approximation domain*. The domain $\Omega_k \subset \Omega$ where the computation is performed is referred to as the *computational domain* (see Marcozzi (2001) for details). *A priori* determination of the size of the computational domain required for a given error tolerance is possible as shown in Kangro and Nicolaides (2000) for the case of the multi-dimensional Black-Scholes-Merton PDE. Namely, after changing the variables in the Black-Scholes-Merton PDE that reduces it to the heat equation,

$$\max_{t\in[0,T]} ||U_k(t,\cdot) - U(t,\cdot)||_{L^\infty(G)} \le Ce^{-cR_k}, \tag{5.7}$$

where R_k is the radius of the computational domain Ω_k and $C > 0$ and $c > 0$ are independent of R_k.

While from the theoretical standpoint any choice of artificial boundary conditions will work as long as the problem on the bounded domain is well-posed, from the practical standpoint of the numerical computation, one would like the artificial boundary conditions to approximate the solution on the boundary as closely as possible. If the chosen artificial boundary conditions closely approximate the solution, then the bounded computational domain can be chosen to be of relatively moderate size, reducing the size of the numerical computation and gaining in computational efficiency. If the chosen artificial boundary conditions do not approximate the solution well enough, one needs to choose a large enough computational domain to mitigate the pollution of the solution from the artificial boundary conditions, thus increasing the size of the numerical computation. In the case of our convertible bond problem, we fix an approximation domain $G = (\underline{S}, \overline{S}) \times (\underline{V}, \overline{V}) \times (\underline{r}, \overline{r}) \times (\underline{z}, \overline{z})$, where we are interested in the solution, and consider a sequence of computational domains $\Omega_k = (\underline{S}_k, \overline{S}_k) \times (\underline{V}_k, \overline{V}_k) \times (\underline{r}_k, \overline{r}_k) \times (\underline{z}_k, \overline{z}_k)$, $0 < \underline{S}_k < \underline{S} < \overline{S} < \overline{S}_k < \infty$, $0 < \underline{V}_k < \underline{V} < \overline{V} < \overline{V}_k < \infty$, $0 < \underline{r}_k < \underline{r} < \overline{r} < \overline{r}_k < \infty$, $0 < \underline{z}_k < \underline{z} < \overline{z} < \overline{z}_k < \infty$. In this paper we use artificial Dirichlet boundary conditions for all $t \in [0,T]$:

$$U_k|_{\partial\Omega_k} = g|_{\partial\Omega_k}, \tag{5.8}$$

where $g = g(t,S,V,r,z)$ is an appropriately chosen function to approximate the solution U on the boundary. In particular, the following choice of g approximates the solution on the boundary reasonably well:

$$g(t,S,V,r,z) = \max\{\min\{C(t), ECB(t,S,V,z,r)\}, kS\}, \tag{5.9}$$

where $ECB(t,S,V,z,r)$ is the price of the European-style convertible bond (3.14). Note that

$$g(T,S,V,r,z) = \Psi(S),$$

because $ECB(T,S,V,z,r) = \Psi(S)$, and $\psi_1 \le g \le \psi_2$ for all $t \in [0,T]$, i.e., the call and conversion constraints are enforced on the boundary. The European-style convertible bond price can be computed analytically in the affine framework, or can be approximated by using some suitable simplifying assumptions, e.g., letting the volatilities of V, r, and z be all zero, and using the Black-Scholes-Merton formula in the European call option portion of the expression (3.14). In the latter case, we have a simple explicit expression for the European-style convertible bond to be used in the artificial boundary condition (5.9):

$$ECB(t,S,V,r,z) = \sum_{i=[t/\delta]+1}^{N-1} c\delta F e^{-(r+\lambda)(i\delta-t)} + (1+c\delta)Fe^{-(r+\lambda)(T-t)}$$

$$+kEC_{BSM}(K,T;t,S,\sqrt{V},r+\lambda,q)+\frac{\lambda RF}{r+\lambda}(1-e^{-(T-t)(r+\lambda)}), \tag{5.10}$$

where $\lambda = z + \alpha V + \beta r$ and $EC_{BSM}(K,T;t,S,\sqrt{V},r+\lambda,q)$ is the time-t Black-Scholes-Merton price of the European call with strike K (where $K = (1+c\delta)F/k$) and expiration T when the current stock price at time t is S, the volatility of the stock price is $\sigma = \sqrt{V}$, the interest rate is $r+\lambda$, and dividend yield is q. We use this expression in our choice of the artificial boundary conditions in our numerical experiments in this paper.

Remark 5.1. The probabilistic interpretation of the localization to a bounded domain is as follows. Fix a bounded open domain Ω_k in the state space Ω and add the following artificial knock-out clause to the convertible bond contract. If the underlying state variable process hits the boundary $\partial\Omega_k$ at some time t prior to maturity T, conversion θ_1, call θ_2, and default, then the bond is retired (knocked out) and the bondholder receives a rebate $g(t, S_t, V_t, r_t, z_t)$. Clearly, the value function of this artificial knock-out convertible bond with rebate g will approximate the value function of the convertible bond with no knock-out feature better and better as we enlarge the domain Ω_k, shifting the boundary further and further away from the initial state of the process. This probabilistic interpretation is discussed in Friedman (1976) and Bensoussan and Lions (1982, 1984) in the general context of stochastic games with stopping times. These references also provide a probabilistic interpretation of the penalty approximation (see Bensoussan and Lions (1982, Section 5 in Chapter 3 in particular). In fact, the theorems in Friedman (1976) and Bensoussan and Lions (1982, 1984) we cite throughout this paper are formulated for stochastic games with stopping times in a bounded domain with a boundary. In order to apply these results in our setting, one needs to localize the stochastic game with stopping times to a bounded domain by adjoining an artificial knock-out condition with rebate, and then consider the variational inequality formulation of the game on the bounded domain, as well as its penalty approximation.

6 The Weak (Variational) Formulation

In order to solve the non-linear PDE (5.5) on a bounded computational domain Ω_k with the Dirichlet boundary conditions (5.8)–(5.9) and the terminal condition (5.2) numerically in the finite element method-of-lines framework, in this section we develop a weak (variational) formulation of the PDE. In what follows we consider a fixed bounded computational domain Ω_k and drop the index k to simplify notation (i.e., Ω is now a bounded domain as described in section 5). For future convenience, we change the stock price variable to a dimensionless variable $x := \ln(S/S_0)$, where S_0 is some reference price level, and transform the terminal value problem with non-homogeneous Dirichlet boundary conditions into an initial value problem with homogeneous Dirichlet boundary conditions by defining

$$u_\varepsilon(t,x,V,r,z) := U_\varepsilon(T-t, S_0e^x, V, r, z) - g(T-t, S_0e^x, V, r, z), \tag{6.1}$$

$$\phi_i(t,x,V,r,z) := \psi_i(T-t, S_0e^x) - g(T-t, S_0e^x, V, r, z), \;\; i = 1,2. \tag{6.2}$$

The function u_ε solves the PDE

$$\frac{\partial u_\varepsilon}{\partial t} + \mathcal{A}u_\varepsilon + (r+\lambda)u_\varepsilon + \pi_\varepsilon(u_\varepsilon) = f \tag{6.3}$$

with the non-linear penalty term

$$\pi_\varepsilon(u_\varepsilon) = -\left(\frac{1}{\varepsilon}(u_\varepsilon - \phi_1)^-\right)^p + \left(\frac{1}{\varepsilon}(u_\varepsilon - \phi_2)^+\right)^p \tag{6.4}$$

and the non-homogeneous term (we denote $\mathbf{x} = (x, V, r, z)$, the four-dimensional state vector)

$$f(t, \mathbf{x}) = \lambda R + \mathcal{C}(T-t) + \frac{\partial g}{\partial t}(T-t, \mathbf{x}) - \mathcal{A}g(T-t, \mathbf{x}) - (r+\lambda)g(T-t, \mathbf{x}) \tag{6.5}$$

with the homogeneous boundary and initial conditions:

$$u_\varepsilon(t, \mathbf{x})|_{\partial\Omega} = 0, \ \ \forall t \in [0, T], \ \text{ and } \ u_\varepsilon(0, \mathbf{x}) = 0, \ \ \forall \mathbf{x} \in \Omega. \tag{6.6}$$

Here $\mathcal{A}$ is the second-order elliptic differential operator on the bounded domain Ω (a restriction of the negative of the infinitesimal generator $\mathcal{G}$ of the ddiffusion process[3] to the bounded domain):

$$\mathcal{A}u := -\frac{1}{2}\left(V\frac{\partial^2 u}{\partial x^2} + \sigma_V^2 V \frac{\partial^2 u}{\partial V^2} + \sigma_r^2 r \frac{\partial^2 u}{\partial r^2} + \sigma_z^2 z \frac{\partial^2 u}{\partial z^2} + 2\rho_{SV}\sigma_V V \frac{\partial^2 u}{\partial x \partial V}\right)$$
$$-(r - q + \lambda - \frac{1}{2}V)\frac{\partial u}{\partial x} - \kappa_V(\theta_V - V)\frac{\partial u}{\partial V} - \kappa_z(\theta_z + \gamma V - z)\frac{\partial u}{\partial z} - \kappa_r(\theta_r - r)\frac{\partial u}{\partial r} \tag{6.7}$$

with the homogeneous Dirichlet boundary conditions on $\partial\Omega$.

A weak (variational) formulation of the PDE problem (6.3)–(6.7) is obtained by considering a space of test functions square-integrable on Ω, with their (weak) first derivatives square-integrable on Ω, and vanishing on the boundary $\partial\Omega$ (the Sobolev space $H_0^1(\Omega) := \{u \in L^2(\Omega) : \nabla u \in L^2(\Omega), u|_{\partial\Omega} = 0\}$). Multiplying the PDE (6.3) with a test function $v = v(\mathbf{x})$, integrating over Ω and applying Green's formula, we arrive at the weak (variational) formulation of the PDE:

$$(\frac{\partial u_\varepsilon}{\partial t}, v) + a(u_\varepsilon, v) + (\pi_\varepsilon(u_\varepsilon), v) = (f, v), \tag{6.8}$$

where

$$(u, v) = \int_\Omega u(\mathbf{x})v(\mathbf{x})d\mathbf{x}$$

is the inner product in $L^2(\Omega)$, and the bilinear form $a(\cdot, \cdot)$ is defined by:

$$a(u,v) = \frac{1}{2}\int_\Omega \left(V\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \sigma_V^2 V \frac{\partial u}{\partial V}\frac{\partial v}{\partial V} + \sigma_r^2 r \frac{\partial u}{\partial r}\frac{\partial v}{\partial r} + \sigma_z^2 z \frac{\partial u}{\partial z}\frac{\partial v}{\partial z} + \rho_{SV}\sigma_V V\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial V} + \frac{\partial v}{\partial x}\frac{\partial u}{\partial V}\right)\right) d\mathbf{x}$$
$$-\int_\Omega \left([r - q + \lambda - \frac{1}{2}V - \frac{1}{2}\rho_{SV}\sigma_V]\frac{\partial u}{\partial x}v + [\kappa_V(\theta_V - V) - \frac{1}{2}\sigma_V^2]\frac{\partial u}{\partial V}v + [\kappa_z(\theta_z + \gamma V - z) - \frac{1}{2}\sigma_z^2]\frac{\partial u}{\partial z}v \right.$$
$$\left. + [\kappa_r(\theta_r - r) - \frac{1}{2}\sigma_r^2]\frac{\partial u}{\partial r}v \right) d\mathbf{x} + \int_\Omega (r+\lambda)uv d\mathbf{x}.$$

To solve the variational formulation, we seek a function $u_\varepsilon = u_\varepsilon(t, \mathbf{x})$ that vanishes on the boundary $\partial\Omega$, satisfies a vanishing initial condition (6.6), and such that Eq. (6.8) holds for any test function $v \in H_0^1(\Omega)$. More details on the variational formulation of parabolic PDEs associated with diffusion processes can be found in Quarteroni and Valli (1997) and Thomee (1997), where the relevant functional analytic background can be found.

[3]To be specific, in this section all expressions are written for the infinitesimal generator of the four-factor affine model in section 3.2. However, the framework is valid for any other Markovian model of the general form in section 3.1, including non-affine models in section 3.3.

7 Numerical Solution by the Finite Element Method-of-Lines

7.1 Finite Element Approximation

We now develop a spatial discretization of the variational formulation (6.8), (6.6) on the bounded rectangular domain $\overline{\Omega} = [\underline{x}, \overline{x}] \times [\underline{V}, \overline{V}] \times [\underline{r}, \overline{r}] \times [\underline{z}, \overline{z}]$ by the Galerkin finite element method (see Ciarlet (1978), Larson and Thomee (2003), Quarteroni and Valli (1997) and Thomee (1997) for textbook treatments of the finite element method). Each of the four variables is discretized as follows: $x_i = \underline{x} + ih_x$, $i = 0, \ldots, m_x + 1$, with $h_x = (\overline{x} - \underline{x})/(m_x + 1)$, $V_i = \underline{V} + ih_V$, $i = 0, \ldots, m_V + 1$, with $h_V = (\overline{V} - \underline{V})/(m_V + 1)$, $z_i = \underline{z} + ih_z$, $i = 0, \ldots, m_z + 1$, with $h_z = (\overline{z} - \underline{z})/(m_z + 1)$, and $r_i = \underline{r} + ih_r$, $i = 0, \ldots, m_r + 1$, with $h_r = (\overline{z} - \underline{z})/(m_r + 1)$, where h_x, h_V, h_r, and h_z are discretization step sizes in each of the variables. Then the finite element grid nodes in the four-dimensional rectangular domain $\overline{\Omega}$ are $(x_{i_x}, V_{i_V}, z_{i_z}, r_{i_r})$, $0 \leq i_x \leq m_x + 1$, $0 \leq i_V \leq m_V + 1$, $0 \leq i_z \leq m_z + 1$, $0 \leq i_r \leq m_r + 1$. Since our problem is homogeneous with the vanishing Dirichlet boundary conditions, so that the solution vanishes on the boundary, we will associate finite element basis functions only with the inner nodes: $(x_{j_x}, x_{j_V}, x_{j_z}, x_{j_r}), 1 \leq j_x \leq m_x$, $1 \leq j_V \leq m_V$, $1 \leq j_z \leq m_z$, $1 \leq j_r \leq m_r$. We introduce a natural ordering for the inner nodes as follows. For $\mathbf{x}_J = (x_{j_x}, x_{j_V}, x_{j_z}, x_{j_r})$, define $J : \mathbb{N}^4 \to \mathbb{N}$ as follows:

$$J(j_x, j_V, j_z, j_r) = j_x + (j_V - 1)m_x + (j_z - 1)m_x m_V + (j_r - 1)m_x m_V m_z.$$

The total number of inner nodes is $M = m_x m_V m_z m_r$. We define the finite-element basis functions $\{\varphi_J(\mathbf{x})\}_{J=1}^M$ associated with the inner nodes in the four-dimensional rectangular domain Ω as the product of one-dimensional hat functions:

$$\varphi_J(\mathbf{x}) = \varphi_{i_x}(x)\varphi_{i_V}(V)\varphi_{i_z}(z)\varphi_{i_r}(r), \quad J = J(i_x, i_V, i_z, i_r),$$

$$1 \leq i_x \leq m_x, \ 1 \leq i_V \leq m_V, \ 1 \leq i_z \leq m_z, \ 1 \leq i_r \leq m_r, \ 1 \leq J \leq M,$$

where each one-dimensional hat function $\varphi_i(x)$ associated with the node x_i is defined by:

$$\varphi_i(x) = \begin{cases} (x - x_{i-1})/h, & x_{i-1} \leq x \leq x_i \\ (x_{i+1} - x)/h, & x_i \leq x \leq x_{i+1} \\ 0\ , & x \notin [x_{i-1}, x_{i+1}] \end{cases} .$$

The Jth basis function $\varphi_J(\mathbf{x})$ is a hat function equal to one at the inner node $\mathbf{x}_J = (x_{j_x}, x_{j_V}, x_{j_z}, x_{j_r})$ and zero outside of the four-dimensional cube $[x_{j_x-1}, x_{j_x+1}] \times [V_{j_v-1}, V_{j_V+1}] \times [z_{j_z-1}, z_{j_z+1}] \times [r_{j_r-1}, r_{j_r+1}]$.

We look for a finite element approximation $u_{\varepsilon,h}$ to the solution u_ε of the variational formulation (6.8), (6.6) as a linear combination of the finite element basis functions with time-dependent coefficients:

$$u_{\varepsilon,h}(t, \mathbf{x}) = \sum_{J=1}^{M} u_{\varepsilon,J}(t)\varphi_J(\mathbf{x}), \ \ t \in [0, T]. \tag{7.1}$$

Note that, by construction, $u_{\varepsilon,h}$ vanishes on the boundary $\partial\Omega$ (since the basis functions associated with the inner nodes vanish on the boundary). Thus, we look for an approximation $u_{\varepsilon,h}$ to the true solution u_ε in the finite element space V_h spanned by the finite element basis functions

$\{\varphi_J(\mathbf{x})\}_{J=1}^M$. We also approximate the conversion and call constraints and the penalty term in the finite element basis:

$$\phi_{i,h}(t,\mathbf{x}) = \sum_{J=1}^M \phi_{i,J}(t)\varphi_J(\mathbf{x}), \quad i=1,2, \ \ t \in [0,T], \tag{7.2}$$

$$\pi_{\varepsilon,h}(u_{\varepsilon,h})(t,\mathbf{x}) = \sum_{J=1}^M \pi_\varepsilon(u_{\varepsilon,h})(t,\mathbf{x}_J)\varphi_J(\mathbf{x})$$

$$= \sum_{J=1}^M \left\{ -\left(\frac{1}{\varepsilon}(u_{\varepsilon,J}(t)-\phi_{1,J}(t))^-\right)^p + \left(\frac{1}{\varepsilon}(u_{\varepsilon,J}(t)-\phi_{2,J}(t))^+\right)^p \right\} \varphi_J(\mathbf{x}). \tag{7.3}$$

Denote by $\mathbf{u}_\varepsilon(t) = (u_{\varepsilon,1}(t),\dots,u_{\varepsilon,M}(t))^\top$ the M-dimensional vector of time-dependent coefficients in (7.1) to be determined. Substituting (7.1) into (6.8) and letting the test functions v in (6.8) run through the set of all basis functions $\{\varphi_J\}_{j=1}^M$ (i.e., we approximate the test function by $v_h(\mathbf{x}) = \sum_{j=1}^M v_J\varphi_J(\mathbf{x})$), we obtain the following M-dimensional system of ODEs:

$$\mathbb{M}\dot{\mathbf{u}}_\varepsilon(t) + \mathbb{A}\mathbf{u}_\varepsilon(t) + \mathbb{M}\pi_\varepsilon(\mathbf{u}_\varepsilon)(t) - \mathbf{F}(t) = 0, \quad t \in (0,T] \tag{7.4}$$

subject to the homogeneous initial condition $\mathbf{u}_\varepsilon(0)=0$. Here $\dot{\mathbf{u}}_\varepsilon(t) = (\dot{u}_{\varepsilon,1}(t),\dots,\dot{u}_{\varepsilon,M}(t))^\top$, $\dot{u}_{\varepsilon,J}(t) \equiv du_{\varepsilon,J}(t)/dt$, $\mathbb{M} = (m_{IJ})_{I,J=1}^M$, $m_{IJ} = (\varphi_J,\varphi_I)$, $\mathbb{A} = (a_{IJ})_{I,J=1}^M$, $a_{IJ} = a(\varphi_J,\varphi_I)$, $\pi_\varepsilon(\mathbf{u}_\varepsilon)(t) = (\pi_\varepsilon(u_{\varepsilon,1})(t),\dots,\pi_\varepsilon(u_{\varepsilon,M})(t))^\top$,

$$\pi_\varepsilon(u_{\varepsilon,J})(t) = -\left(\frac{1}{\varepsilon}(u_{\varepsilon,J}(t)-\phi_{1,J}(t))^-\right)^p + \left(\frac{1}{\varepsilon}(u_{\varepsilon,J}(t)-\phi_{2,J})^+\right)^p,$$

and $\mathbf{F}(t) = (f_1(t),\dots,f_M(t))^\top$,

$$f_J(t) = (\lambda R + \mathcal{C}(T-t) + \dot{g}(T-t),\varphi_J) - a(g(T-t),\varphi_J).$$

The ODE system (7.4) is referred to as the *semi-discretization* of the variational problem (spatially discrete and continuous in time). The problem is now reduced to the numerical integration of this ODE system. This is referred to as the *finite element method-of-lines* (MOL). Due to the origins of the finite element method in structural engineering, $\mathbb{M}$ is referred to as the *mass matrix*, $\mathbb{A}$ is referred to as the *stiffness matrix*, and $\mathbf{F}$ is referred to as the *load vector*. The ODE (7.4) also has a non-linear penalty term $\pi_\varepsilon(\mathbf{u}_\varepsilon)$. For each $t \in [0,T]$, on the bounded computational domain Ω the semi-discrete finite element approximation is known to be second order accurate in the spatial discretization step size h:

$$||u_{\varepsilon,h}(t,\cdot) - u_\varepsilon(t,\cdot)|| \le Ch^2 \tag{7.5}$$

both in the $L^2(\Omega)$ and $L^\infty(\Omega)$ norms.

7.2 Integrating the ODE System

We have reduced the convertible bond pricing problem to the numerical solution of the non-linear ODE system (7.4). We observe that $\mathbb{M} \sim O(h_x h_V h_z h_r)$ and $\mathbb{A} \sim O((h_x h_V h_z h_r)^{-1})$. Hence, the system (7.4) is *stiff* for small spatial discretization steps h. In this paper we use a variable integration step-size and variable integration order backward differentiation formula (BDF) based ODE package SUNDIALS (SUite of Nonlinear and DIfferential/ALgebraic equation Solvers) available from the Lawrence Livermore National Laboratory (http://www.llnl.gov/CASC/sundials/). In particular, we use the library IDA included in the SUNDIALS suite. IDA is a package for the solution of differential-algebraic equation (DAE) systems in the form

$$F(t, \mathbf{u}, \dot{\mathbf{u}}) = 0, \tag{7.6}$$

where t is an independent variable, $\mathbf{u} \in \mathbb{R}^n$ is a vector of dependent variables, $\dot{\mathbf{u}} = d\mathbf{u}/dt$, and $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is a non-linear function, $F(t, \mathbf{u}, \dot{\mathbf{u}}) = \mathbb{M}\dot{\mathbf{u}}_\varepsilon + \mathbb{A}\mathbf{u}_\varepsilon + \mathbb{M}\pi_\varepsilon(\mathbf{u}_\varepsilon) - \mathbf{F}$. It is written in C, but is derived from the package DASPK which is written in Fortran. The integration method in IDA is variable-order, variable-coefficient BDF, in fixed-leading-coefficient form, making it suitable for the stiff systems integration (see Brenan et al. (1996)). The method order varies between 1 and 5. The backward differentiation formula (BDF) of order q is obtained by approximating the value of the derivative $\dot{\mathbf{u}}(t_n)$ at time t_n through the derivative of a polynomial of degree q interpolating $\mathbf{u}(t)$ at $q + 1$ time steps t_n, t_{n-1}, $\dots$, t_{n-q}:

$$\dot{\mathbf{u}}(t_n) \sim \Delta t_n^{-1} \sum_{i=0}^{q} \alpha_{n,i} \mathbf{u}_{n-i}.$$

The coefficients $\alpha_{n,i}$ are uniquely determined by the order q and the history of time steps Δt_n, Δt_{n-1},...,Δt_{n-q} and are independent of the function $\mathbf{u}(t)$. In the case of equal time steps $\Delta t_1 = \Delta t_2 = \cdots = \Delta t_n$, the coefficients $\alpha_{n,i} = \alpha_i$ of the BDF scheme of order up to five are:

q	α_0	α_1	α_2	α_3	α_4	α_5
1	1	-1				
2	$\frac{3}{2}$	-2	$\frac{1}{2}$			
3	$\frac{11}{6}$	-3	$\frac{3}{2}$	-$\frac{1}{3}$		
4	$\frac{25}{12}$	-4	3	-$\frac{4}{3}$	$\frac{1}{4}$	
5	$\frac{137}{60}$	-5	5	-$\frac{10}{3}$	$\frac{5}{4}$	-$\frac{1}{5}$

For the variable step case, the coefficients $\alpha_{n,i}$ depend on the time step history.

The application of the BDF formula reduces (7.6) to a non-linear algebraic system with the unknown vector $\mathbf{u}_n = \mathbf{u}(t_n)$ that has to be solved at each step from the knowledge of the vectors $\mathbf{u}_{n-i} = \mathbf{u}(t_{n-i})$, $i = 1, .., q_n$ (in IDA the order of the BDF scheme q_n is variable from time step to time step):

$$G(\mathbf{u}_n) \equiv F(t, \mathbf{u}_n, \Delta t_n^{-1} \sum_{i=0}^{q_n} \alpha_{n,i} \mathbf{u}_{n-i}) = 0.$$

The resulting non-linear algebraic system is solved by Newton iteration. Starting with the solution obtained at the previous step, $\mathbf{u}_{n(0)} = \mathbf{u}_{n-1}$, the Newton correction is determined iteratively by solving the linear system:

$$\mathbb{J}(\mathbf{u}_{n(m+1)} - \mathbf{u}_{n(m)}) = -G(\mathbf{u}_{n(m)}), \tag{7.7}$$

where $\mathbf{u}_{n(m)}$ is the Newton approximation of the solution $\mathbf{u}_n$ on the m-th iteration and $\mathbb{J}$ is the Jacobian matrix of G,

$$\mathbb{J} = \frac{\partial G}{\partial \mathbf{u}} = \frac{\partial F}{\partial \mathbf{u}} + \frac{\alpha_{n,0}}{\Delta t_n} \frac{\partial F}{\partial \dot{\mathbf{u}}}$$

evaluated at the previous value $\mathbf{u}_{n(m)}$. For the ODE system (7.4), the Jacobian matrix is:

$$\mathbb{J} = \mathbb{A} + \mathbb{M}\mathbf{\Pi}_\varepsilon(\mathbf{u}_\varepsilon) + \frac{\alpha_{n,0}}{\Delta t_n}\mathbb{M},$$

where $\mathbf{\Pi}_\varepsilon(\mathbf{u}_\varepsilon) = \left(\frac{\partial \pi_\varepsilon(u_{\varepsilon,J})}{\partial u_{\varepsilon,I}}\right)_{I,J=1}^{M}$ is a diagonal matrix:

$$\frac{\partial \pi_\varepsilon(u_{\varepsilon,J})}{\partial u_{\varepsilon,I}} = \delta_{I,J}\left(\frac{p}{\varepsilon}\right)\left\{\left(\frac{1}{\varepsilon}(u_{\varepsilon,J}(t) - \phi_{1,J}(t))^-\right)^{p-1}\mathbf{1}_{\{u_{\varepsilon,J}(t)<\phi_{1,J}(t)\}}\right.$$

$$\left.+\left(\frac{1}{\varepsilon}(u_{\varepsilon,J}(t) - \phi_{2,J})^+\right)^{p-1}\mathbf{1}_{\{u_{\varepsilon,J}(t)>\phi_{2,J}(t)\}}\right\},$$

where $\delta_{I,J}$ is the identity matrix and $\mathbf{1}_{\{x>y\}}$ is the indicator function. For $p > 1$ this Jacobian is continuous in $\mathbf{u}_\varepsilon$. We use the Scaled Preconditioned Generalized Method of Residuals (SPGMRES) to solve the linear system (7.7) at each step of the Newton iteration. The preconditioned system is $(\mathbb{P}^{-1}\mathbb{J})(\mathbf{u}_{n(m+1)} - \mathbf{u}_{n(m)}) = -\mathbb{P}^{-1}G$ with the user-defined preconditioner $\mathbb{P}$. We use a diagonal preconditioner $\mathbb{P} = diag(\mathbb{J})$, which is the simplest to invert preconditioner suitable to reduce stiffness resulting from the large penalty term for small ε.

During the integration of the system, IDA adaptively selects the time step Δt_n and the order of the BDF scheme q_n so that the error remains within the desired error tolerance. IDA starts the integration with the order $q = 1$ and the initial step size which is adjusted to satisfy the initial error test. During the integration, the solver tries to increase both the order of integration and the step size. If the error test fails with the given integration step and order, the step size and the order of integration are reduced. The quantities controlling the change of the step size and integration order are estimated from the asymptotic properties of the BDF scheme. The details of the adaptive order and step selection can be found in the SUNDIALS documentation Hindmarsh et al. (2005 and 2006).

8 Analysis of Convertible Bonds

8.1 The Affine Model

We will assume the following parameters for the four-factor affine model of section 3.2 that will be used in numerical examples in this section:

- Default-free interest rate process: $\theta_r = 0.06$, $\kappa_r = 0.25$, $\sigma_r = 0.1$, initial short rate $r_0 = 0.05$.

- Stock process: dividend yield $q = 0.02$.

- Instantaneous stock price variance process: $\kappa_V = 1.0$, $\theta_V = 0.0625$, $\sigma_V = 0.8$, initial variance level $V_0 = 0.0625$ (corresponding to the initial stock price volatility $\sigma_0 = 0.25$ or 25% per annum), correlation between the stock return and the variance processes $\rho_{SV} = -0.5$.

- Default intensity factor z: $\theta_z = 0.005$, $\kappa_z = 0.5$, $\gamma = 0.3$, $\sigma_z = 0.8$, $z_0 = 0.023$.

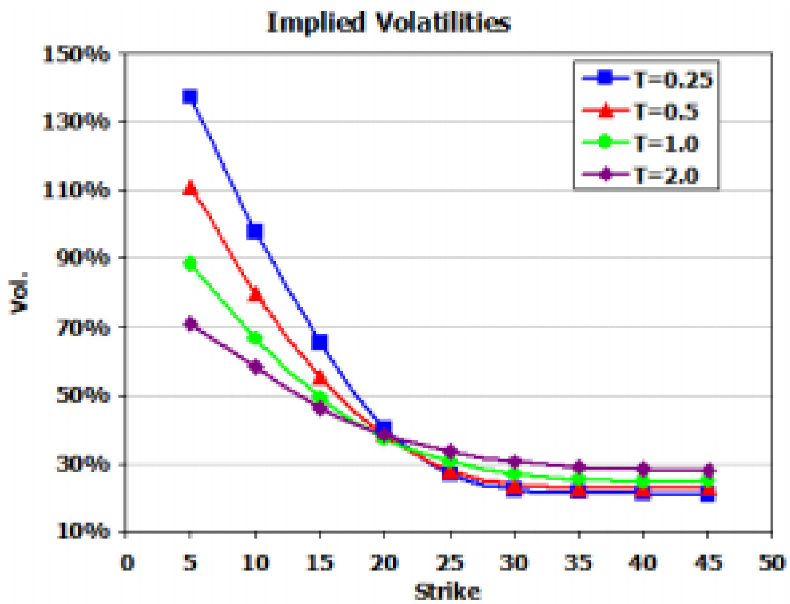


Figure 1: **Implied volatility skews for times to expiration $T = 0.25, 0.5, 1,$ and 2 years.**

- Default intensity process: $\alpha = 0.3$, $\beta = 0.1$. Initial default intensity with these parameters: $\lambda_0 = 0.04675$ or 4.675% per annum.

Figure 1 plots the implied volatility skews for European-style stock options with times to expiration $T = 0.25, 0.5, 1,$ and 2 years in this model. For each time to expiration, it graphs the Black-Scholes implied volatility as a function of the strike price, assuming the current stock price is $S_0 = 25$. The call option prices are computed using the affine solution presented in the Appendix, and the Black-Scholes volatilities are implied from these model option prices. We observe that the implied volatility skews have empirically realistic shapes. For each time to expiration, the volatility is a decreasing convex function of strike. This so-called *volatility skew* is induced by the negative correlation between the stock price and the instantaneous variance, as well as by the possibility of default (jump to zero). As time to expiration increases, the volatility skew gradually flattens out, in accordance with the pattern empirically observed in equity options markets.

Figure 2 plots the term structure of zero-coupon credit spreads $s(T)$ in this model, assuming zero recovery in default. Defaultable zero-coupon bonds with different times to maturity T are priced, and credit spreads are then computed according to:

$$s(T) = -\frac{1}{T}\ln(B(T)) - R(T),$$

where $B(T)$ is the T-maturity defaultable zero-coupon bond price and $R(T)$ is the default-free T-maturity zero-coupon interest rate (yield-to-maturity on the T-maturity bond). We observe that the instantaneous credit spread is equal to λ_0, and, for these parameter values, the term structure of credit spreads is hump-shaped, first upward-slopping, and then downward-slopping after about two years. The term structure of default-free zero-coupon interest rates $R(T)$ is also plotted in Figure 2. It is upward-slopping for these parameter values.

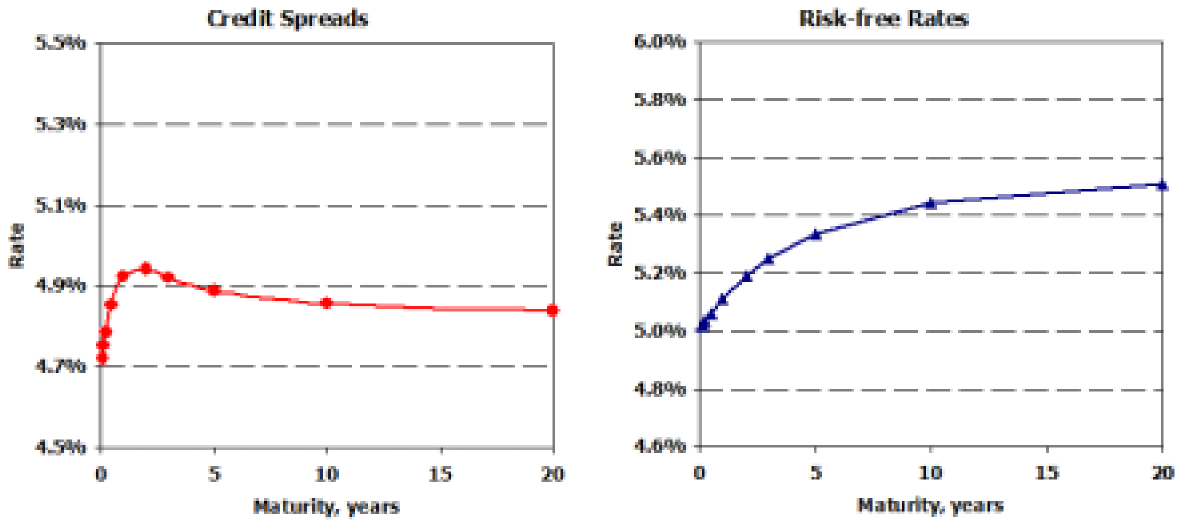


Figure 2: **The term structures of default-free interest rates and credit spreads.**

Figure 3 plots the value function of a European-style 5-year convertible bond with the following parameters: coupon rate $c = 0.03$ (3%) per annum, coupon period $\delta = 0.5$ year (semiannual coupon), face value $F = 100$ dollars, conversion factor $k = 4$ (convertible into four shares of stock per \$100 face value), conversion is allowed only at maturity, no call, and zero recovery in default. The bond price is computed according to Eq. (3.7), where the expectations are computed as described in the Appendix. The bond price is plotted as a function of the stock price and time remaining to maturity with all other parameters kept fixed (in particular, $V_0 = 0.0625$, $z_0 = 0.023$, $r_0 = 0.02$).

8.2 Convergence of the Numerical Solution

In this section we experimentally study convergence of our numerical approximation and, in particular, experimentally validate the error estimates (5.3) and (7.5) for the penalty approximation and the spatial discretization, as well as study computational performance of the SUNDIALS temporal integration solver. We consider two 5-year convertible bonds: a European-style convertible bond with the same parameters as in section 8.1 and an American-style convertible bond with the clean call price of \$140, conversion allowed at any time during the five-year period, call protection during the first two years (call allowed at any time during the subsequent three years) and all other parameters the same as in section 8.1. For the European-style convertible bond we are able to use the affine solution for the analytical expression (3.7) as our benchmark (see the Appendix). The computed value function of the European-style bond is obtained by numerically solving the PDE (5.1)–(5.2) without the penalty terms (corresponding to $\varepsilon \to \infty$). The numerical solution is then compared to the analytical affine solution. The discrete coupon payments are handled as follows. The time interval $[0, T]$ with $T = 5$ years is split into subintervals (t_{i-1}, t_i) between the coupon payment dates $t_i = i\delta$, $i = 1, 2, \ldots, N$ ($N = 10$ and $\delta = 0.5$ in this case). The PDE (5.1) (correspondingly the ODE (7.4) after spatial discretization) is solved between the coupon payment dates with the so-called *jump condition* enforced at the coupon

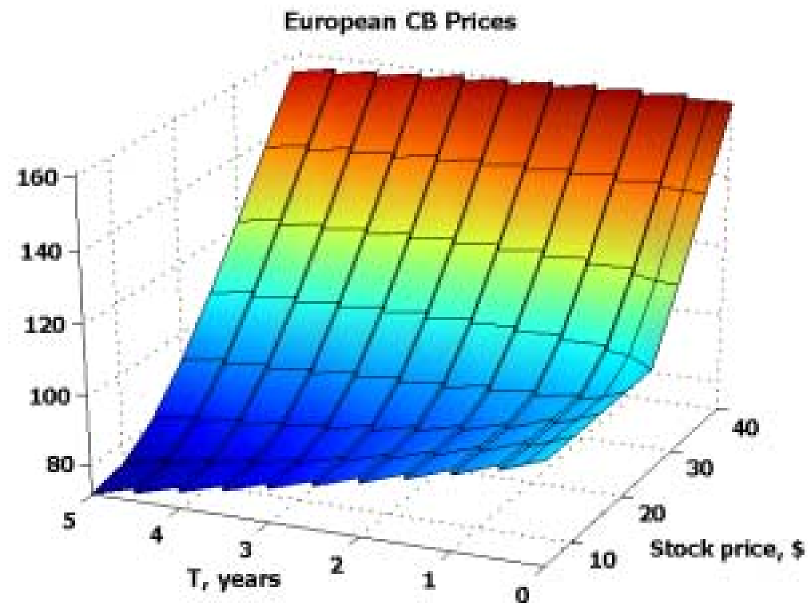


Figure 3: **European-style convertible bond value function.**

payment dates:

$$U(t_i-, \mathbf{x}) = U(t_i+, \mathbf{x}) + C_i$$

with $C_i = \delta c F$, the coupon paid at time t_i. This jump condition enforces the absence of arbitrage across the coupon payment date. Indeed, to prevent arbitrage, the value of the bond just before the coupon is paid should be equal to the value of the bond just after the coupon is paid plus the coupon amount. In other words, the bond value drops by the amount of the coupon at each of the coupon payment dates. Between the coupon payment dates, the PDE (5.1) is solved with the coupon rate $\mathcal{C}$ in the PDE set equal to zero, as there are no payments between the coupon dates.

The value function of the American-style convertible bond with call and conversion is computed by solving the PDE (5.1)–(5.2) with the penalty term, where we select a penalization parameter $\varepsilon << 1$. In this case we do not have an analytical benchmark. To study convergence, we compute a benchmark value function by selecting a small ε, taking a large computational domain, selecting a small spatial discretization step h, and solving the ODE (7.4) with a small error tolerance parameter. The computed numerical solution serves as the benchmark for solutions on coarser spatial discretization grids, with larger values of the penalization parameter, and less stringent error tolerances in the solver.

Figure 4 illustrates the results for the European-style and American-style convertible bond. The graphs (a) and (b) plot the L^2 norms of the errors over the approximation domain as functions of the spatial discretization step size h in the log-log scale for the European and American cases, respectively (the temporal integration error tolerance and the penalization parameter in the American case are selected small enough to isolate the spatial discretization error). The affine solution is used as a benchmark in the European case, and a pre-computed numerical solution is used as a benchmark in the American case. The plots evidence the h^2 convergence as $h \to 0$, validating the estimate (7.5). Sample solution values are given in Table 1.

In this example, the convertible bond prices are computed for $S = S_0 = 20$, $V = 0.0625$ (25% volatility of the stock), $z = 0.023$, and $r = 0.05$. In this example the American-style bond is worth less than the European-style bond due to the call option retained by the bond issuer that significantly limits the possible gains from the equity price appreciation and subsequent conversion by forcing the bond holder to convert early.

	European			American		
Benchmark Solution	119.402			93.211		
Mesh size, h	0.2	0.1	0.05	0.2	0.1	0.05
FEM solution	120.067	119.518	119.375	92.058	93.008	93.188
Error	−0.665	−0.116	0.027	1.153	0.204	0.023
Error as % of par	−0.67%	−0.12%	0.03%	1.15%	0.20%	−0.02%

Table 1: The FEM solution for the European and American convertible bond problems.

The graph (c) in the second row plots the maximum norm error as a function of the penalization parameter ε for the penalty terms with $p = 1, 2, 3$ (the spatial discretization step size and the temporal discretization error tolerance are selected small enough to isolate the penalization error). The plot evidences the linear error decay as $\varepsilon \to 0$, validating the estimate (5.3).

The graph (d) illustrates the performance of the SUNDIALS IDA temporal integration solver in integrating the non-linear ODE system resulting from the spatial discretization of the penalized PDE in the American case. It plots the time step history of the SUNDIALS solver. We observe that after each coupon payment date (that introduces non-smoothness into the solution, as the solution jumps by the amount of the coupon), the solver restarts the computation with a small time step, and then rapidly increases the time step size as the solution smooths out, until it reaches the next coupon payment date, where the process is repeated. The solver uses the first order integration right after each coupon payment date, and increases to between two and four as the temporal integration progresses and the solution smooths out. The average integration order of the solver in this example is 2.07. The step size during the action of the call constraint is slightly smaller than during the call-protected period.

Table 2 presents CPU times needed to achieve practically relevant accuracy of $0.1 - 0.2\%$ of the convertible bond value. According to Bloomberg, typical bid/ask spreads on convertible bonds quoted by dealers are in the range between 0.5% and 2%, depending on the liquidity of a particular issue. The accuracy of $0.1 - 0.2\%$ ensures that computational errors are significantly smaller than the bid/ask spread. The table presents CPU times for the 5-year convertible bond studied in previous examples in this section. The bond is valued using the full four-factor model, as well as the three three-factor models (with constant interest rate, constant volatility, or constant z).

8.3 The Convertible Bond Value Function and Optimal Call and Conversion Boundaries

The value function of the American convertible bond $CB(t, S, V, z, r)$ as a function of the stock price S and time to maturity $\tau = T - t$ is shown in Figure 5 for fixed $V = 0.0625$, $z = 0.023$, and $r = 0.05$. The convertible bond value function exhibits downward jumps after each coupon payment, and becomes equal to the payoff at the expiration time. Four sections of the value function at $\tau = 5$ years to maturity are shown in Figure 6. The value function is an increasing and convex function of S, and a decreasing and convex function of the variables z and r (Figures 6 (a)–(d)). The value of the convertible bond is increasing in V for relatively high values of the stock price S due to the high positive effect of volatility on the embedded conversion option (equity component of the convertible bond). In contrast, the value of the convertible bond decreases with the increase in V for relatively low values of the stock price S (see Figure 6 (a)). This is due to the increase in the default intensity λ associated with higher equity volatility and resulting decline in the bond (debt) component due to increased default risk that out-weights the increase of the value of the embedded conversion option.

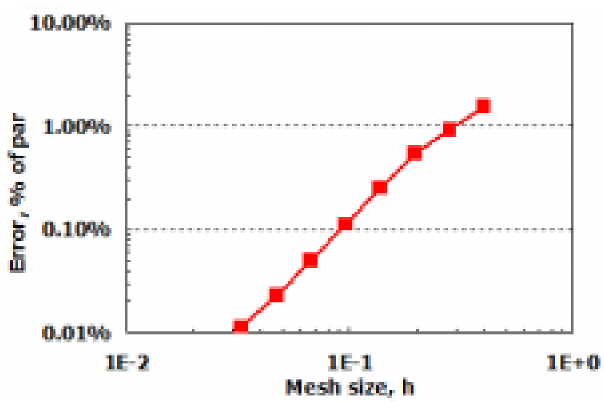
In Figure 5 we also show the call (solid line) and conversion (dashed line) boundaries on the solution surface and their projections onto the (S, t)-plane. It is optimal for the bond issuer to call the bond at time t if $S_t \geq S^{call}(t)$, where $S^{call}(t)$ is the call boundary, and it is optimal for the bond holder to convert the bond into shares at time t if $S_t \geq S^{conv}(t)$, where $S^{conv}(t)$ is the conversion boundary. Note that during the call protection period, the bond holder will not convert the bond into shares just before the coupon payment time, as it is optimal to postpone the conversion until after the coupon payment date to receive the coupon and then convert the bond after the coupon is received. As a result, the conversion boundary asymptotically goes to infinity with the approach of the coupon payment date. When the bond is callable, both the call and conversion boundaries should be equal to or lower than the point of intersection of the call and conversion payoffs: $S^{conv}(t), S^{call}(t) \leq C(t)/k$. In our solution, the call boundary is equal to this quantity during the time when the bond is callable, $S^{call}(t) = C(t)/k$, up until the final coupon period before maturity (see also Figure 7 (a)–(d)). With the approach of maturity (when the value of the bond approaches its face value F), the conversion boundary declines to F/k (which is less than $C(t)/k$). The F/k is the stock price that makes conversion value equal to the notional amount ($kS = F$).

Since the bond conversion generally happens at relatively high stock prices, where the variance V has a positive effect on the bond price through its effect on the conversion option value, the conversion boundary $S^{conv}(t)$ is increasing in V (Figure 7 (a)). Similarly, because of the negative effect of z and r on the solution, the boundaries are decreasing in these variables (Figure 7 (b)–(c)). The correlation between the stock price and the variance processes, ρ_{SV}, has a negative effect on the boundaries, similar to the effects of α, β and γ (see Table 2). The volatility of variance σ_V has a positive effect on the boundaries, while σ_z and σ_r have negative effects.

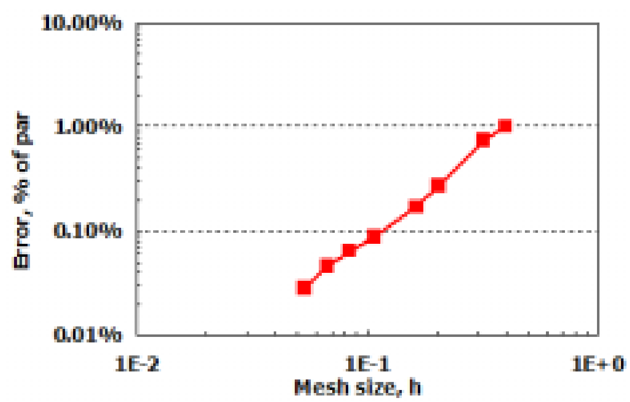
An interesting question is how the call and conversion boundaries are influenced by the number of stochastic factors used in the model. To assess this, we consider cases when one of the three processes V_t, z_t, and r_t is deterministic, i.e., in each of the three cases we set one of the three volatilities σ_V, σ_z, σ_r equal to zero. Figure 7 (d) plots the resulting boundaries in these three three-factor models, as well as the boundaries for the complete four-factor model. We see that the conversion boundary in the call-protected period, as well as the conversion boundary in the last coupon period before maturity are sensitive to the effects of volatilities of all of the state variables. Making one of the four factors deterministic and reducing the four-factor model to the three-factor one has a significant effect on the optimal conversion boundaries. This is further illustrated in Table 2 that gives convertible bond values and the critical conversion levels for $T = 5$ years to maturity for the four-factor model and the three cases where we set one of the volatilities σ_V, σ_z, σ_r equal to zero.

	Four-factor model					Three-factor models		
	Baseline	$\rho_{SV}=0$	$\alpha=0$	$\beta=0$	$\gamma=0$	$\sigma_V=0$	$\sigma_z=0$	$\sigma_r=0$
Price	93.21	93.26	97.13	94.15	93.27	94.25	91.55	93.19
S^{conv}	41.47	40.67	50.47	43.07	41.97	39.57	43.07	41.37

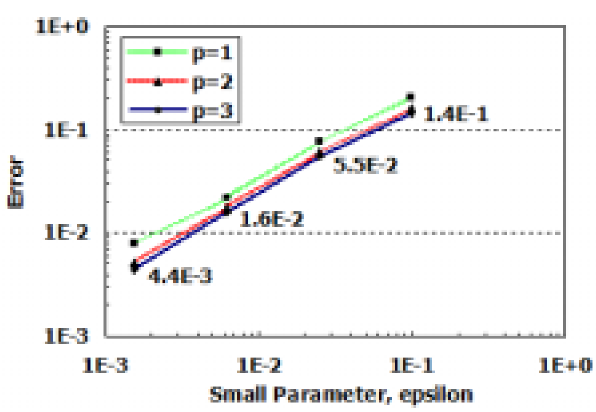
Table 2: Convertible bond price and critical value S^{conv} for the four-factor model and three three-factor models computed at $t=5$, $V=0.0625$, $z=0.023$, and $r=0.05$. The bond price is computed at $S=20$. The four-factor model parameters are: $\rho_{SV}=-0.5$, $\alpha=0.3$, $\beta=0.1$, $\gamma=0.1$, $\sigma_V=0.8$, $\sigma_z=0.8$, $\sigma_z=0.1$. In each of the three-factor models one of the volatilities σ_V, σ_z, or σ_z is set equal to zero.



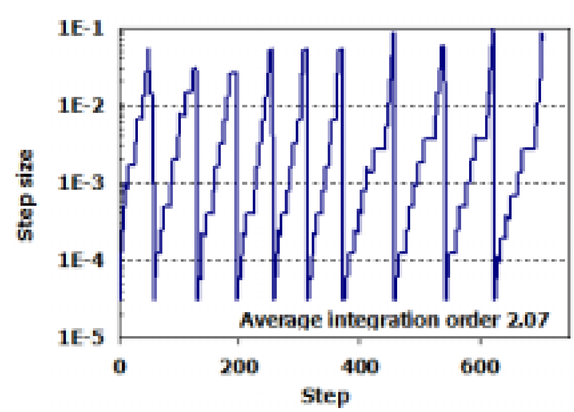
(a) Convergence of European CB solution



(b) Convergence of American CB solution



(c) Convergence of Penalty Approximation



(d) SUNDIALS IDA integration step size

Figure 4: Convergence of solutions for European and American convertible bonds.

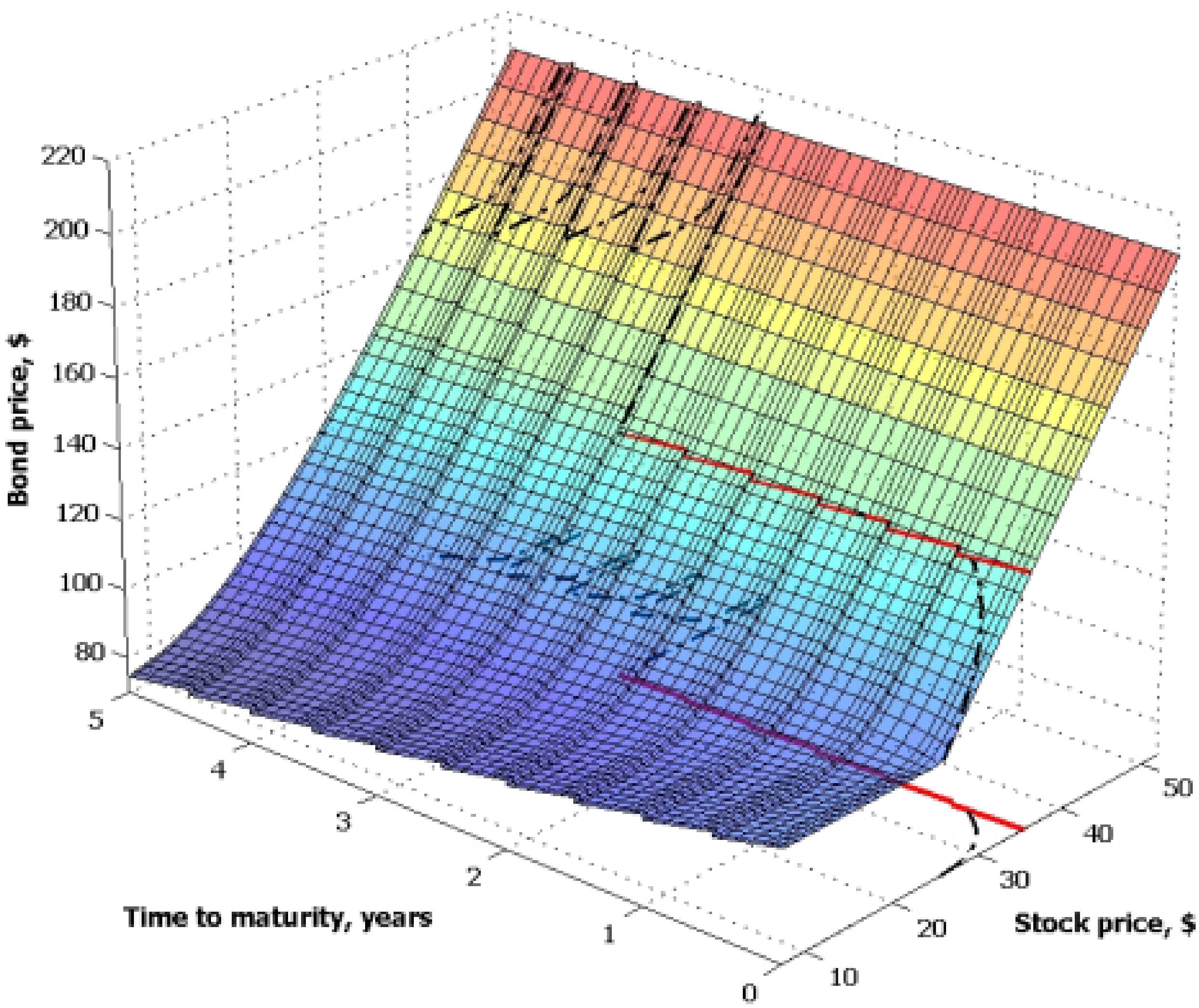
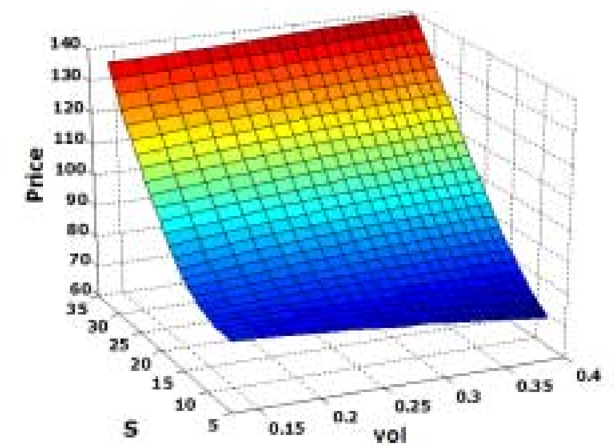
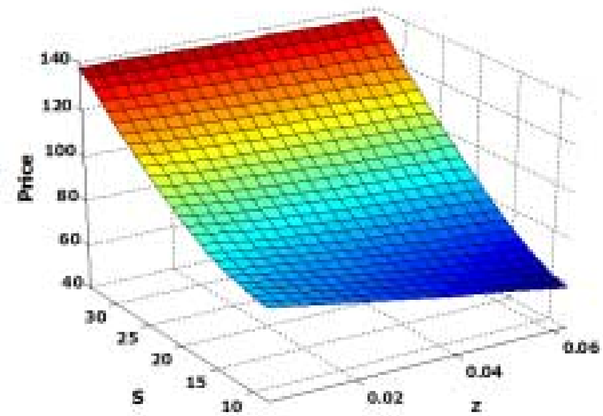


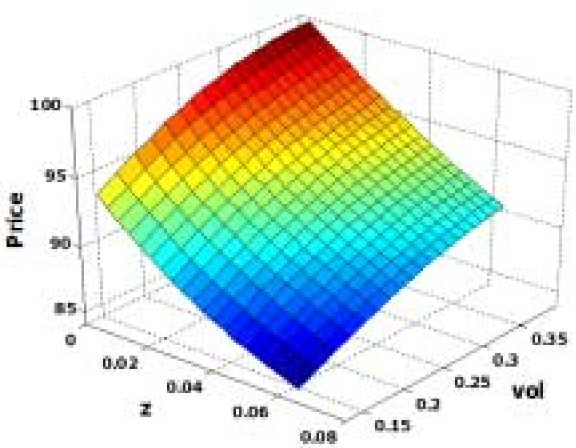
Figure 5: Solution for the 5-year Convertible Bond (semiannual 3% coupon, 2 years call protection period, and clean call price \$1400). Solid lines – optimal call boundary, dashed lines – optimal conversion boundary. The projection of the optimal call and conversion boundaries on the (S, T) plane is also shown.



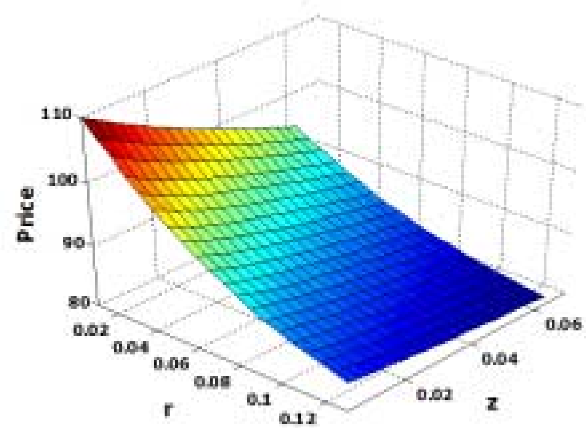
(a) Solution surface as a function of S and $\sqrt{V}$ for $z = 0.023$ and $r = 0.05$



(b) Solution surface as a function of S and z for $V = 0.0625$ and $r = 0.05$

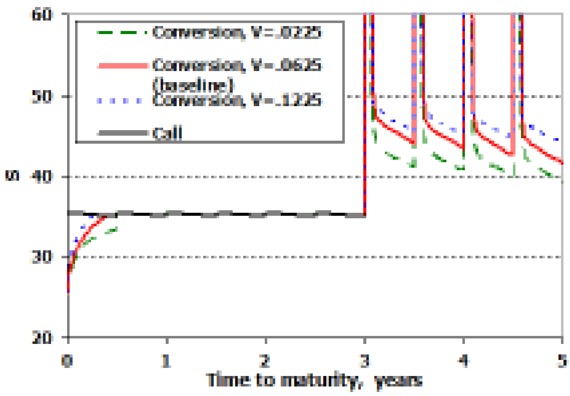


(c) Solution surface as a function of $\sqrt{V}$ and z for $S = 20.0$ and $r = 0.05$

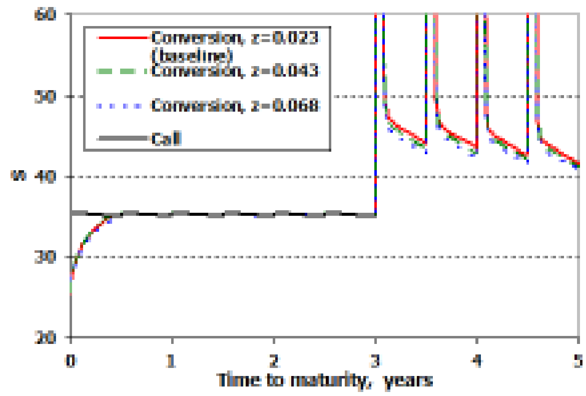


(d) Solution surface as a function of z and r for $S = 20.0$ and $V = 0.0625$

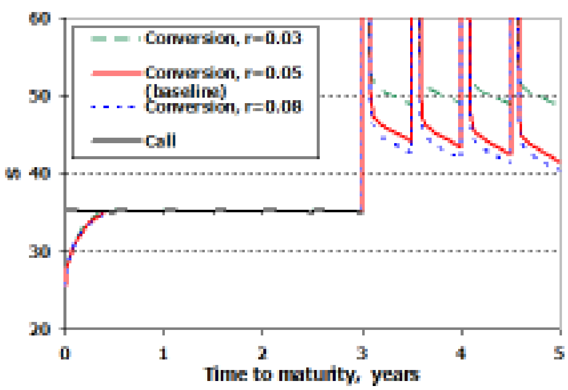
Figure 6: Two-dimensional sections of the convertible bond value function.



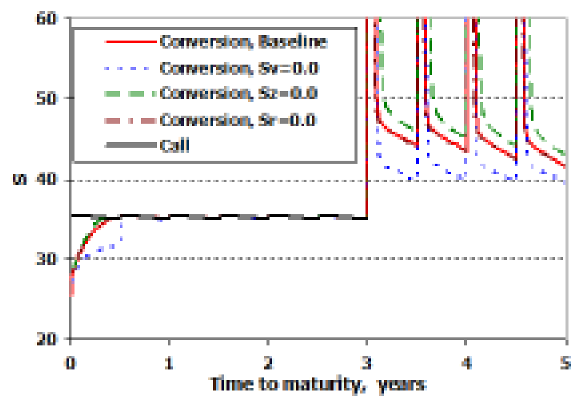
(a) Call and Conversion Boundaries for different V



(b) Call and Conversion Boundaries for different z



(c) Call and Conversion Boundaries for different r



(d) Dependance of call and conversion boundaries on volatilities of state variables.

Figure 7: Optimal call and conversion boundaries.

9 Conclusion

This paper develops and solves a four-factor convertible bond model with stochastic interest rate, stock price, volatility, and default intensity. Our model is a four-factor extension of the three-factor Carr and Wu (2005) reduced-form affine model of default applied to convertible bonds. At default, the stock price drops to zero. Prior to default, the stock price follows a continuous process with stochastic volatility. The default intensity, the instantaneous stock variance, and the default-free short rate follow a tri-variate diffusion process specified to capture empirical evidence on stock option prices and credit default swap spreads. We show that the value function of the convertible bond satisfies a variational inequality formulation of the stochastic game between the bondholder and the issuer. The variational inequality is approximated by a penalized non-linear partial differential equation. The penalized formulation is solved numerically by applying the finite element spatial discretization and an adaptive time integrator. This framework allows us to value and analyze convertible bonds in this empirically realistic set-up with four risk factors, in contrast with the existing convertible bond literature that has up to now been primarily limited to one- and two-factor models. Our framework naturally includes such empirical features of market data as the leverage effect and the dependence between credit spreads and equity volatility. The inclusion of credit risk, volatility dynamics, and interest rate dynamics are important in order to value and hedge convertible bonds consistently with more liquid securities such as straight corporate debt and CDS, equity, and equity options that can be used to hedge convertible bonds. Such consistency across all traded securities on the same reference company is particularly important for the firms engaged in convertible bond arbitrage. At the same time, accurate determination of the optimal conversion and call policies is important to bond holders and bond issuers in order to maximize the value of their assets and minimize the value of their liabilities, respectively.

A Analytical Solutions for European-style Securities in the Four-Factor Affine Model

Introduce a new state variable $x_t := \log(S_t/S_0)$, so that x_t solves the following SDE:

$$dx_t = (r_t - q + \lambda_t - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^S.$$

The dynamics of the four stochastic variables in our model can be expressed in the matrix form:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,$$

where $X_t = (x_t, V_t, z_t, r_t)^\top$ is the state vector and W_t is a vector of four independent Brownian motions. The model is such that both the drift vector $\mu(X)$ and the covariance matrix $\sigma(X)\sigma(X)^\top$ are affine functions of X, $\mu(X) = K_0 + K_1 X$, $K_0 \in \mathbb{R}^4$, $K_1 \in \mathbb{R}^{4\times4}$, and $\sigma(X)\sigma(X)^\top = H_0 + H_x x + H_V V + H_z z + H_r r$, $H_0, H_x, H_V, H_z, H_r \in \mathbb{R}^{4\times4}$, with the coefficients:

$$K_0 = \begin{pmatrix} -q_t \\ \kappa_V\theta_V \\ \kappa_z\theta_z \\ \kappa_r\theta_r \end{pmatrix}, \quad K_1 = \begin{pmatrix} 0 & \alpha - \frac{1}{2} & 1 & \beta + 1 \\ 0 & -\kappa_V & 0 & 0 \\ 0 & \kappa_z\gamma & -\kappa_z & 0 \\ 0 & 0 & 0 & -\kappa_r \end{pmatrix}, \quad H_0 = 0, \quad H_x = 0,$$

$$H_V = \begin{pmatrix} 1 & \rho_{SV}\sigma_V & 0 & 0 \\ \rho_{SV}\sigma_V & \sigma_V^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad H_z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_Z^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad H_r = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_r^2 \end{pmatrix}.$$

We also introduce a discount rate $R(X_t) := r_t + \lambda_t$ that is also an affine function of the state vector: $R(X) = R_0 + R_1^\top X$ with $R_0 = 0$ and the vector $R_1 = (0, \alpha, 1, \beta + 1)^\top$.

Since the model is in the affine class, its generalized characteristic function

$$\varphi(u, X_t, t, T) = \mathbb{E}_t \left[e^{-\int_t^T R(X_s)ds} e^{ux_T} \right], \quad u \in \mathbb{C}$$

is exponential affine in the state vector (see Duffie et al. (2000) and Carr and Wu (2005)):

$$\varphi(u, X, t, T) = e^{A(t)+B_x(t)x+B_V(t)V+B_z(t)z+B_r(t)r},$$

where the scalar $A(t)$ and the vector $B(t) = (B_x(t), B_V(t), B_z(t), B_r(t))^\top$ solve the system of complex-valued ODEs:

$$\dot{B}_x(t) = (R_1)_0 - (K_1^\top B(t))_0 - \frac{1}{2}B(t)^\top H_x B(t) = 0,$$

$$\dot{B}_V(t) = (R_1)_1 - (K_1^\top B(t))_1 - \frac{1}{2}B(t)^\top H_V B(t) = \alpha - (\alpha - \frac{1}{2})B_x(t)$$

$$+\kappa_V B_V(t) - \kappa_z \gamma B_z(t) - \frac{1}{2}(B_x(t)^2 + 2\rho_{SV}\sigma_V B_x(t) B_V(t) + \sigma_V^2 B_V(t)^2),$$

$$\dot{B}_z(t) = (R_1)_2 - (K_1^\top B(t))_2 - \frac{1}{2}B(t)^\top H_z B(t) = 1 - B_x(t) + \kappa_z B_z(t) - \frac{1}{2}\sigma_z^2 B_z(t)^2,$$

$$\dot{B}_r(t) = (R_1)_3 - (K_1^\top B(t))_3 - \frac{1}{2}B(t)^\top H_3 B(t) = \beta - (\beta + 1)B_x(t) + \kappa_r B_r(t) - \frac{1}{2}\sigma_r^2 B_r(t)^2,$$

$$\dot{A}(t) = R_0 - K_0^T B(t) - \frac{1}{2}B(t)^\top H_0 B(t) = qB_x(t) - \kappa_V \theta_V B_V(t) - \kappa_z \theta_z B_z(t) - \kappa_r \theta_r B_r(t)$$

with the terminal conditions $A(T) = 0$ and $B(T) = (u, 0, 0, 0)^\top$. Integrating the first ODE, we obtain $B_x(t) = u$, and the system of ODE reduces to

$$\dot{B}_V(t) = (\alpha - (\alpha - \frac{1}{2})u - \frac{1}{2}u^2) + (\kappa_V - \rho_{SV}\sigma_V u)B_V(t) - \kappa_z \gamma B_z(t) - \frac{1}{2}\sigma_V^2 B_V(t)^2, \quad \text{(A.1a)}$$

$$\dot{B}_z(t) = (1 - u) + \kappa_z B_z(t) - \frac{1}{2}\sigma_z^2 B_z(t)^2, \quad \text{(A.1b)}$$

$$\dot{B}_r(t) = \beta - (\beta + 1)u + \kappa_r B_r(t) - \frac{1}{2}\sigma_r^2 B_r(t)^2, \quad \text{(A.1c)}$$

$$\dot{A}(t) = qu - \kappa_V \theta_V B_V(t) - \kappa_z \theta_z B_z(t) - \kappa_r \theta_r B_r(t) \quad \text{(A.1d)}$$

with the terminal conditions $B_V(T) = B_z(T) = B_r(T) = A(T) = 0$. If $\gamma = 0$ (no effect of the stock volatility on the drift of the default intensity), then we have a system of independent Riccatiatti equations for $B_V(t)$, $B_z(t)$, and $B_r(t)$ that can be solved analytically. If $\gamma \neq 0$, the solution to this non-linear system of ODEs can be obtained numerically.

Once the generalized characteristic function $\varphi(u, X_t, t, T)$ is determined, we can value defaultable bonds (3.10) without recovery ($R = 0$) directly by

$$B_0(t, X_t) = \varphi(0, X_t, t, T),$$

while European-style options can be valued by inverting the Fourier transform. In particular, due to the results in Duffie et al. (2000) and Carr and Wu (2005), the price of a call option with strike K is computed as follows:

$$C(t, X_t, K, T) = \mathbb{E}_t\left[e^{-\int_t^T R(X_s)ds}(S_0 e^{x_T} - K)^+\right]$$
$$= \frac{1}{2}\left(S_0\varphi(1, X_t, t, T) - K\varphi(0, X_t, t, T)\right)$$
$$-\frac{1}{\pi}\int_0^\infty \left\{\mathrm{Im}\left((S_0\varphi(1 - iv, X_t, t, T) - K\varphi(-iv, X_t, t, T))e^{ikv}\right)\right\}\frac{dv}{v},$$

where $k = \ln(K/S_0)$, $\mathrm{Im}(z)$ denotes the imaginary part of $z \in \mathbb{C}$, and the integral is computed by numerical quadrature.

To perform the computations of option prices more efficiently, Carr-Madan (1999), and Lee (2004) consider a *damped* option pricing function:

$$C_\delta(k) := S_0^{-1}e^{\delta k}C(t, X_t, S_0\exp(k), T),$$

with a damping constant $\delta > 0$, and $k = \ln(K/S_0)$. They show that the Fourier transform, $\hat{C}_\delta(u)$, of a *damped* option price $C_\delta(k)$, exists for the appropriately chosen damping constants, and can be computed as:

$$\frac{\varphi(u - (\delta + 1)i, X_t, t, T)}{\delta^2 + \delta - u^2 + i(2\delta + 1)u}.$$

As a result, the option price is given by the Fourier inversion formula involving only a single term:

$$C(t, X_t, K, T) = S_0\frac{e^{-\delta k}}{2\pi}\int_{-\infty}^\infty e^{-iuk}\frac{\varphi(u - (\delta + 1)i, X_t, t, T)}{\delta^2 + \delta - u^2 + i(2\delta + 1)u}du$$
$$= S_0\frac{e^{-\delta k}}{\pi}\int_0^\infty \mathrm{Re}\left[e^{-iuk}\frac{\varphi(u - (\delta + 1)i, X_t, t, T)}{\delta^2 + \delta - u^2 + i(2\delta + 1)u}\right]du, \tag{A.2}$$

where the integral is computed by numerical quadrature.

To value the recovery part (Eq. (3.11) and the last term in Eq. (3.14)), we also need to compute an extended generalized characteristic function $\phi(u, v, X_t, t, T)$ defined for $u \in \mathbb{C}$ and $v \in \mathbb{R}^n$ by

$$\phi(u, v, X_t, t, T) = \mathbb{E}_t\left[e^{-\int_t^T R(X_s)ds}(v \cdot X_T)e^{ux_T}\right].$$

This expectation can be computed as follows:

$$\phi(u, v, X_t, t, T) = \varphi(u, X_t, t, T)(C(t) + D(t)X),$$

where $\varphi(u, X_t, t, T)$ is the generalized characteristic function determined previously, and the scalar $C(t)$ and the vector $D(t) = (D_x(t), D_V(t), D_z(t), D_r(t))^\top$ solve the system of ODEs

$$\dot{D}_x(t) = -(K_1^\top D(t))_0 - \frac{1}{2}B(t)^\top H_x D(t), \tag{A.3a}$$

$$\dot{D}_V(t) = -(K_1^\top D(t))_1 - \frac{1}{2}B(t)^\top H_V D(t), \tag{A.3b}$$

$$\dot{D}_z(t) = -(K_1^\top D(t))_2 - \frac{1}{2}B(t)^\top H_z D(t), \tag{A.3c}$$

$$\dot{D}_r(t) = -(K_1^\top D(t))_3 - \frac{1}{2}B(t)^\top H_3 D(t), \tag{A.3d}$$

$$\dot{C}(t) = -K_0^T D(t) - \frac{1}{2}B(t)^\top H_0 D(t) \tag{A.3e}$$

with the terminal conditions $C(T) = 0$ and $D(T) = v^\top$, and the vector $B(t)$ determined previously when solving for $\varphi(u, X_t, t, T)$. Then the recovery part of the defaultable bond can be calculated as follows:

$$RF \int_t^T \mathbb{E}_t \left[e^{-\int_t^s (r_u + \lambda_u) du} \lambda_s \right] ds = RF \int_t^T \phi(u, v, X, t, s) ds$$

with $u = 0$ and $v = (0, \alpha, 1, \beta)^\top$.

In our implementation we usethe fifth-order Cash-Karp Runge-Kutta method with adaptive step-size control (see Gear (1971), Cash and Carp (1990), and the implementation in Press et al. (1995)) for solving the ODEs (A.1) and (A.3) required for computing the generalized characteristic function $\varphi(u, X_t, t, T)$ and the extended generalized characteristic function $\phi(u, v, X_t, t, T)$. The proposed integration method achieves sufficiently high accuracy with relatively small number of time steps and function evaluations. It allows for controlling both the absolute and relative error while solving the ODE. We use the absolute error tolerance of 10^{-5} and the initial step size of $dt = 10^{-4}$ to achieve the empirical error in the CB price smaller than 10^{-5}. It takes only 10 integration steps to solve each ODE to the desired accuracy. In contrast, the constant step 4-th order Runge-Kutta method requires more than 10000 time steps to achieve the comparable accuracy. The efficiency gains from using the Cash-Karp adaptive Runge-Kutta solver are dramatic in this application. The integration of the characteristic functions required for computing the option prices (A.2) is performed by using the trapezoidal rule, which is known to have exponentially decaying errors of discretization and truncation for the functions considered here (e.g., Lee (2004)). We use a constant discretization interval equal to 0.05, and stop the computations as soon as the integral on 10 consecutive intervals is smaller in absolute value than the desired accuracy 10^{-5}. The resulting empirical error of a CB price computation in this case is also of the order 10^{-5}.

References

[1] Andersen, L., and D. Buffum, 2004, "Calibration and Implementation of Convertible Bond Models," *Journal of Computational Finance*, **7** (2), 1–34.

[2] Ayache, E., P. Forsyth, and K. Vetzal, 2003, "The Valuation of Convertible Bonds with Credit Risk," *Journal of Derivatives*, **11**, Fall 2003, 9-29.

[3] Barone-Adesi G., A. Bermudez, and J. Hatgioannides, 2003, "Two-factor Convertible Bonds Valuation Using the Method of Characteristics/Finite Elements," *Journal of Economic Dynamics and Control*, **27**, 1801–1831.

[4] Bardhan, I., A. Bergier, E. Derman, C. Dosembet, I. Kani, 1994, "Valuing Convertible Bonds as Derivatives," Quantitative Strategies Research Note, November, Goldman Sachs, New York.

[5] Bielecki, T., S. Crepy, M. Jeanblanc, and M. Rutkowski, 2006, "Arbitrage Pricing of Convertible Securities with Credit Risk," Working Paper.

[6] Bielecki, T., and M. Rutkowski, 2002, *Credit Risk: Modeling, Valuation and Hedging*, Springer, Berlin.

[7] Bensoussan, A., and J. L. Lions, 1982, *Applications of Variational Inequalities in Stochastic Control*, Elsevier, Amsterdam.

[8] Bensoussan, A., and J. L. Lions, 1984, *Impulse Control and Quasi-Variational Inequalities*, Gauthier-Villars, Paris.

[9] Bermudez, A., and N. Webber, 2003, "An Asset Based Model of Defaultable Convertible Bonds with Endogenised Recovery," Working Paper, City University, Cass Business School, London.

[10] Boman, M., 2001, "A Posteriori Error Analysis in the Maximum Norm for a Penalty Finite Element method for the Time-Dependent Obstacle Problem," Preprint, Chalmers Finite Element Center, Goteborg, Sweden.

[11] Brenan, K. E., Campbell, S. L., and L. R. Petzold, 1996, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, SIAM, Philadelphia, PA.

[12] Brennan, M.J., and E.S. Schwartz, 1977, "Convertible Bonds: Valuation and Optimal Strategies for Call and Conversion," *Journal of Finance*, **XXXII** (5), 1699-1715.

[13] Brennan, M.J., and E.S. Schwartz, 1980, "Analyzing Convertible Bonds," *Journal of Financial and Quantitative Analysis*, **XV** (4), 907-929.

[14] Brezzi, F., D. H. Norrie, et al., 1987, *Finite Element Handbook*, McGraw-Hill, New York.

[15] Carr, P., and V. Linetsky, 2006, "A Jump to Default Extended CEV Model: An Application of Bessel Processes," *Finance and Stochastics*, 10(3), 303-330.

[16] Carr, P. and D. Madan, 1999, "Option Valuation using the Fast Fourier Transform," *Journal of Computational Finance*, **3**, 463–520.

[17] Carr P., and L. Wu, 2005, "Stock Options and Credit Default Swaps, A Joint Framework for Valuation and Estimation," Working Paper.

[18] Cash, J. R., and Karp, A. H., 1990, "A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides," *ACM Transactions on Mathematical Software*, 16, 3, 201-222.

[19] Ciarlet, P.G., 1978, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam (Revised Ed., SIAM, 2002).

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

[21] Davis, M., and F. Lischka, 2002, "Convertible Bonds with Market Risk and Credit Risk," in *Applied Probability*, eds. R. Chan et al., Studies in Advanced Mathematics, American Mathematical Society/International Press, 45-58.

[22] Duffie, D., and K. Singleton, 2003, *Credit Risk*, Princeton University Press, Princeton, NJ.

[23] Duffie, D., J. Pan, and K. Singleton, 2000, "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, **68**(6), 1343-1376.

[24] Friedman, A., 1988, *Variational Principles and Free-Boundary Problems*, R.E. Krieger, Malabar, FL.

[25] Friedman, A., 1976, *Stochastic Differential Equations and Applications*, Volume II, Academic Press, New York.

[26] Gear, C.W., 1971, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ.

[27] Grau, A.J., P.A. Forsyth, K.R. Vetzal, 2003, "Convertible Bonds with Call Notice Periods," Working Paper, University of Waterloo.

[28] Grimwood, R., and S. Hodges, 2002, "The Valuation of Convertible Bonds: A Study of Alternative Pricing Models," Working Paper, University of Warwick.

[29] Grossmann, C., 1991, "Penalties and Mixed Finite Element Analysis for Variational Inequalities," *International Studies of Numerical Mathematics*, **99**, 137-145.

[30] Forsyth, P.A. and K.R. Vetzal, 2002, "Quadratic Convergence for Valuing American Options Using a Penalty Method," *SIAM Journal on Scientific Computing*, 23(6), 2095-2122.

[31] Glowinski, R., J. L. Lions, and Tremolieris, 1981, *Numerical Analysis of Variational Inequalities*, North Holland, Amsterdam.

[32] Glowinski, R., 1984, *Numerical Methods for Nonlinear Variational Problems*, Springer, Berlin.

[33] Hamilton, D.T., P.M. Stumpp, and R. Cantor, 2001, "Default and Recovery Rates of Convertible Bond Issuers: 1970-2000," Working Paper, Moody's Investor Services, Global Credit Research.

[34] Hindmarsh, A. C., R. Serban, and A. Collier, 2006, "User Documentation for IDA, a Differential-Algebraic Equation Solver for Sequential and Parallel Computers," Technical Report UCRL-MA-136910, December, Lawrence Livermore National Laboratory.

[35] Hindmarsh, A. C., P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward, 2005, "SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers," *ACM Transactions on Mathematical Software*, **31**(3), 363-396. Also available as LLNL technical report UCRL-JP-200037.

[36] Howard, J., and M. O'Connor, 2001, "Convertible Securities: An Investor's Guide," Deutsche Bank Report.

[37] Hundsdorfer, W., and J.G. Verwer, 2003, *Numerical Solution of Time-dependent Advection-Diffusion-Reaction Equations*, Springer, Berlin.

[38] Ingersoll, J. E., 1977, "A Contingent-Claims Valuation of Convertible Securities," *Journal of Financial Economics*, **4**, 289-322.

[39] Johnson, C., 1976, "A Convergence Estimate for an Approximation of a Parabolic Variational Inequality," *SIAM Journal of Numerical Analysis*, **13**, 599-606.

[40] Johnson, C., 1987, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, Cambridge.

[41] Kallsen, J., and C. Kuhn, 2005, "Convertible Bonds: Financial Derivatives of Game Type," In: *Exotic Options and Advanced Lévy Models*, Kyprianou, A., Schoutens, W., and Wilmott, P., Eds., Wiley, pp.277-288.

[42] Kangro, R., and R. Nicolaides, 2000, "Far Field Boundary Conditions for Black-Scholes Equations," *SIAM Journal on Numerical Analysis*, **38** (4), 1357-1368.

[43] Lando, D., 2004, *Credit Risk Modeling*, Princeton University Press, Princeton, NJ.

[44] Larsson, S., and V. Thomee, 2003, *Partial Differential Equations with Numerical Methods*, Springer, Berlin.

[45] Lau, K. W., and Y.K. Kwok, "Anatomy of Option Features in Convertible Bonds," *Journal of Futures Markets*, **24**(6), 513-532.

[46] Lee, R., 2004, "Option Pricing by Transform Methods: Extensions, Unification, and Error Control," *Journal of Computational Finance*, **7**(3), 51-86.

[47] Linetsky, V., 2006, "Pricing Equity Derivatives subject to Bankruptcy," *Mathematical Finance*, **16** (2), 255-282.

[48] Marcozzi, M., 2001, "On the Approximation of Optimal Stopping Problems with Application to Financial Mathematics," *SIAM Journal on Scientific Computing*, **22** (5), 1865-1884.

[49] Merton, R., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance*, **29**: 449–70.

[50] Muromachi, Y., 1999, "The Growing Recognition of Credit Risk in Corporate and Financial Bond Markets," Working Paper, NLI Research Institute.

[51] Nyborg, K.G., 1996, "The Use and Pricing of Convertible Bonds," *Applied Mathematical Finance*, **3**, 167-190.

[52] Press, W. H., S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, 1995, "Numerical Recipes in C: the Art of Scientific Computing," 2nd. Ed., Cambridge U. Press.

[53] Quarteroni, A., and A. Valli, 1997, *Numerical Approximation of Partial Differential Equations*, Springer, Berlin.

[54] Sapariuc, I., M.D. Marcozzi, and J.E. Flaherty, 2004, "A Numerical Analysis of Variational Valuation Techniques for Derivative Securities," *Applied Mathematics and Computation*, **159** (1), 171-198.

[55] Schönbucher, P.J., 2003, *Credit Derivatives Pricing Models*, Wiley.

[56] Sirbu, M., I. Pikovsky, and S. Shreve, 2004, "Perpetual Convertible Bonds," *SIAM Journal on Control and Optimization*, **43**, 58-85.

[57] Sirbu, M., and S. Shreve, 2005, "A Two-Person Game for Pricing Convertible Bonds," Working Paper, Carnegie Mellon University.

[58] Takahashi, A, T. Kobayashi, and N. Nakagawa, 2001, "Pricing Convertible Bonds with Default Risk," *Journal of Fixed Income*, **11**, 20-29.

[59] Thomée, V., 1997, *Galerkin Finite Element Methods for Parabolic Problems*, Springer, Berlin.

[60] Tsiveriotis, K. and C. Fernandes, 1998, "Valuing Convertible Bonds with Credit Risk," *Journal of Fixed Income*, **8**, 95-102.

[61] Zvan, R., P.A. Forsyth, and K.R. Vetzal, 1998, "Penalty Methods for American Options with Stochastic Volatility," *Journal of Computational and Applied Mathematics*, **91** (2), 199-218.