Efficient Techniques for Simulation of Interest Rate Models Involving Non-Linear Stochastic Differential Equations

by

Leif B. G. Andersen

Vice President, Research
General Re Financial Products Corp.
630 Fifth Avenue, Suite 450
New York, NY 10111
(212)-307-2339
<landerse@genre.com>

Preliminary Version

Efficient Techniques for Simulation of Interest Rate Models involving Non-Linear Stochastic Differential Equations

Abstract

This paper examines methods to reduce systematic and random errors in simulations of interest rate models based on non-solvable, non-linear stochastic differential equations (SDEs). The paper illustrates how application of high-order Ito-Taylor discretization schemes in combination with appropriate variance reduction techniques can yield very significant improvements in speed and accuracy. Besides discussing and testing several traditional approaches to variance reduction, we consider the more recent techniques of Girsanov measure transformation and quasi-random sequences. Using the Cox, Ingersoll, and Ross square-root diffusion model as a specific example, we find that applying a second- or third-order discretization scheme in combination with a probability measure transformation and the antithetic variate method yields the best results overall. Although quite effective in problems of low dimension, quasi-random sequences suffer from certain fundamental problems that limit their usefulness in the applications considered here.

Efficient Techniques for Simulation of Interest Rate Models involving Non-Linear Stochastic Differential Equations

1. Introduction

Driven by the increasing sophistication and competitiveness of the financial markets, much recent finance research emphasizes that theoretical models be "realistic", i.e. closely reflect observable market behavior. Whereas the structure of most classical models has been specified primarily with analytical tractability in mind, many new models deliberately sacrifice such tractability for a closer fit to the market. The lack of tractability, however, makes the practical usage of such models reliant upon the development of efficient numerical schemes. Indeed, the finance literature has recently seen an abundance of ingenious schemes designed either to support complicated, non-linear models or to provide automatic calibration to observable prices. Most of these schemes are based on binomial or trinomial discretizations of either the underlying stochastic processes or the fundamental no-arbitrage PDE. Examples from the interest rate markets include Jamshidian (1991), Hull and White (1990a, 1994a-b), Black, Derman and Toy (1990), Ho and Lee (1986), Jensen & Nielsen (1991) just to name a few. In equity markets Wilmott, Dewynne and Howison (1993), extending the work of Brennan and Schwartz (1978), have illustrated how many types of exotic options can be priced efficiently in either finite difference or finite element lattices. Recently, Dupire (1993, 1994), Derman and Kani (1994), Rubinstein (1994), and Andersen and Brotherton-Ratcliffe (1995) have shown how to use forward induction principles to extend lattice and finite difference models to incorporate the effects of the so-called "volatility smile" in options markets.

A widespread alternative to lattice schemes is the method of Monte Carlo simulation. Introduced to finance by Boyle (1977), the Monte Carlo method is typically applied to problems too complex for the lattice techniques. Such problems include the pricing of instruments that contain strongly path dependent options and/or depend on multiple

IV.3

stochastic factors[1]. As complexity of traded instruments has grown, Monte Carlo techniques have become increasingly important, a trend that is likely to continue in the future. This is particularly true in the fixed income markets, where multiple factors and path-dependency are embedded in a wide variety of new structured or mortgage-based contracts. Despite its increasing importance and wide applicability, the Monte Carlo method has received much less attention in the literature than lattice models. One reason for this is, of course, the low order of convergence ($O(N^{-1/2})$) which tends to make practical usage painfully slow. Recent research, however, suggests that application of results from the field of high-dimensional integration can, under certain circumstances, improve this convergence rate significantly. As shown in Brotherton-Ratcliffe (1994a-b), Joy (1994), and Paskov (1994), replacing the traditional pseudo-random number generators by low-discrepancy quasi-random sequences can lead to significant improvements in the computational effort required to value a wide array of path-dependent options and CMOs. Further, by applying a martingale control variate technique to option pricing, Clewlow and Carverhill (1994) report significant improvements over traditional variance reduction methods[2]. Another line of research has focused on extending the Monte Carlo technique to pricing problems that were previously considered beyond the capabilities of the method. Tilley (1993) and, more recently, Broadie and Glasserman (1994) have shown how, in principle, Monte Carlo simulation can be used to price American-style equity options. Andersen (1996), and Andersen and Brotherton-Ratcliffe (1996) illustrate how the Monte Carlo method can successfully be applied to the pricing of lookback and barrier options with continuous or high-frequency sampling of the underlying asset process.

Almost all recent research on Monte Carlo methods share the common assumption that the underlying asset process is simple deterministic geometric Brownian motion -- a process class characterized by a linear stochastic differential equation (SDE) that permits a closed-form solution. Given the price at time t, S_t, of an asset that follows geometric Brownian motion, recall that the asset price at time $s \geq t$ can be written explicitly in the familiar expression

IV.4

$$S_s = S_t \exp\left(\int_t^s [\mu(u) - \frac{1}{2}\sigma^2(u)]du + \tilde{z}\sqrt{\int_t^s \sigma^2(u)du} \right) \qquad (1)$$

where $\tilde{z}$ is a standard Gaussian variable, and the deterministic functions of time $\mu(t)$ and $\sigma(t)$ are the drift and volatility of S, respectively. In the case of geometric Brownian motion, simulation of asset-paths is thus essentially a matter of repeated application of (1). In interest rate modeling, which will be the main focus of the present paper, it is, however, highly questionable whether we can assume that market dynamics are governed by simple linear SDEs. Many popular interest rate models thus involve highly non-linear SDEs, including those of Pearson and Sun (1990), Cox (1975), Black, Derman and Toy (1990), and Cox, Ingersoll and Ross (1985)[3]. Empirical studies by, among others, Das (1993), Chan *et al* (1992), and Apabhai *et al* (1994) confirm the non-linearity of interest rate SDEs[4].

Without an explicitly solvable SDE, design of an efficient Monte Carlo simulation scheme involves two challenges: i) how to ensure rapid convergence; and ii) how to discretize the process such that the convergence level is close to the continuous-time solution of the SDE. Although Monte Carlo techniques, as discussed above, offer some help towards resolving the former problem, they do not provide much insight into the latter. To handle this problem, we instead need methods from the field of stochastic numerical methods. Pioneered by Milstein (1974, 1978), numerical schemes for solution of SDEs have been developed throughout the 1980s to levels of sophistication approaching those of numerical schemes for PDEs and ODEs. A comprehensive reference to many of these developments is Kloeden and Platen (1992). Despite their success in many physics and engineering problems, application of numerical methods for SDEs to finance has been quite limited and mostly confined to an equity setting, see Platen and Schweizer (1994), and Hofmann, Platen and Schweizer (1992).

In this paper, we will investigate the application of techniques from both Monte Carlo theory and the field of stochastic numerical methods to an interest rate setting. In particular, we will illustrate how careful application of high-order simulation schemes combined with SDE- and Monte Carlo-based variance reduction techniques can lead to significant improvements in convergence of interest rate models. For purposes of exposition, we will limit our discussion to one-factor short rate models; most of the results and general techniques can easily be extended to the case of multiple factors (as in Brennan and Schwartz (1982)) and should also prove useful for forward rate models (as in Heath, Jarrow and Morton (1992)). Further, we will focus most of our discussion to the fundamental problem of determining bond prices from given values of the short rate. As in practice all interest rate contingent claims depend on some, possibly path dependent, function of one or more bond prices (rather than the instantaneous short rate), the capability of quickly extracting accurate bond prices is a prerequisite for simulation of more complicated contingent claims. Throughout the paper we will use the Cox, Ingersoll and Ross (1985) square-root diffusion model to illustrate key results. This model is sufficiently complicated to illustrate the fundamental problems of simulating non-linear SDEs, yet permits a closed-form bond-pricing solution and hence allows for monitoring and testing of simulation results.

The paper is organized as follows. In Section 2, we provide notation and present the theoretical framework for the subsequent sections. Section 3 discusses general numerical schemes for discretization of the short rate SDE and presents initial simulation results for the Cox, Ingersoll, and Ross model. In Section 4, we present and test various techniques for reducing the variance of the bond price estimator. Section 5 contains selected bond simulation results to illustrate the effectiveness of the techniques discussed in previous sections. In Section 6, we briefly discuss the issues involved in application of the developed simulation techniques to more complex interest rate derivatives. Finally, Section 7 concludes the paper with a brief summary of the results and a discussion of potential directions of future research.

IV.6

2. Setup and Model Assumptions

We consider a frictionless intertemporal economy with a bounded trading horizon $[t_0, \tau]$, t_0 being the present and τ some arbitrary, fixed future time. As usual, we represent uncertainty by a probability space (Ω, F, P) where Ω is the set of possible states, F is a σ-field of subsets of Ω, and P is a (real world) probability measure. We will limit our discussions to models where the stochastic dynamics are driven by a single Brownian motion $\{\hat{W}_t, t \in [t_0, \tau]\}$. Information flow in the economy is represented by the (augmented) filtration $\{\Im_t, t \in [t_0, \tau]\}$ generated by $\hat{W}$ and satisfying $\Im_\tau = F$. As always, we assume that $\Im_t$ meets the usual conditions, i.e. is right-continuous and contains the null-sets of P.

To characterize the term structure of risk-free interest rates, we let[5] the strictly positive, adapted process $P(t, T)$, $t_0 \le t \le T \le \tau$ denote the time t price of a bond that matures at time T with certain payout $P(T, T) = \$1$. As discussed in Harrison and Kreps (1979), under technical conditions[6] there are no arbitrage opportunities if and only if there exists a probability measure Q, equivalent to P, under which the expected instantaneous rate of return on any discount bond equals the risk-free rate. Q is frequently referred to as the "risk-neutral" probability measure. Under Q, we will assume that the dynamics of the yield curve can be described by a one-factor diffusion process in the instantaneous short rate r taking values in a subset D of $\Re$:

$$dr_t = \mu_r(r_t, t)dt + \sigma_r(r_t, t)dW_t \qquad (2)$$

where $\{W_t, t \in [t_0, \tau]\}$ is a Brownian motion under Q and where $\mu_r, \sigma_r : D \times [t_0, \tau] \to \Re$ have sufficient regularity to ensure the existence of a unique solution to (2) (see for example Arnold (1992), chapter 6)[7]. In most cases D will be the set of all non-negative real numbers $\Re^+$, although some (Gaussian) models have $D = \Re$.

As discussed in Duffie and Kan (1994) most models of the form (2) appearing in the literature can be written as

$$dr_t = [c_1(t) + c_2(t)r_t + c_3(t)r_t \ln r_t]dt + [k_1(t) + k_2(t)r_t]^\gamma dW_t \tag{3}$$

for constant γ and deterministic coefficients c_1, c_2, c_3, k_1, k_2: $[t_0, \tau] \rightarrow \Re$. In this paper, two special cases of (3) will be of special importance, namely the *Cox, Ingersoll, and Ross (CIR) (1985) model* ($c_3 = 0$, $k_1 = 0$, $\gamma = 0.5$):

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}\,dW_t \tag{4}$$

and the *Vasicek (1977) model* ($c_3 = 0$, $k_2 = 0$, $\gamma = 1.0$):

$$dr_t = \alpha(\beta - r_t)dt + \kappa dW_t \tag{5}$$

for positive constants[8] a, b, σ, α, β, and κ. In the CIR model, the distribution of the short rate can be shown to be non-central chi-square with mean and variance

CIR:
$$E^Q(r_s \mid r_t) = r_t e^{-a(s-t)} + b(1 - e^{-a(s-t)}), \ \ s \geq t \tag{6a}$$

$$V^Q[r_s \mid r_t] = \frac{r_t \sigma^2}{a}(e^{-a(s-t)} - e^{-2a(s-t)}) + \frac{b\sigma^2}{2a}(1 - e^{-a(s-t)})^2, \ \ s \geq t \tag{6b}$$

In the Vasicek model, the short rate is Gaussian with moments

Vasicek:
$$E^Q(r_s \mid r_t) = r_t e^{-\alpha(s-t)} + \beta(1 - e^{-\alpha(s-t)}), \ \ s \geq t \tag{7a}$$

$$V^Q[r_s \mid r_t] = \frac{\kappa^2}{2\alpha}(1 - e^{-2\alpha(s-t)}), \ \ s \geq t \tag{7b}$$

The short rate under the CIR and Vasicek models takes values[9] on $D = \Re^+$ and $D = \Re$, respectively. The positive probability of the short rate being negative under the Vasicek

IV.8

model is clearly a problem; due to its analytical tractability, the Vasicek model is nevertheless quite popular in the literature.

In most applications, we are interested in determining the value of discount bonds given the process dynamics in (2). As shown in Harrison and Kreps (1979), discount bond prices are given by the risk-neutral conditional expectation

$$P(t,T) = E^Q \left[\exp(-\int_t^T r_u du) \Big| \mathfrak{I}_t \right] = E_t^Q \left[\exp(-\int_t^T r_u du) \right] \tag{8}$$

where we have introduced the notation $E_t^Q[\cdot] = E^Q[\cdot \mid \mathfrak{I}_t]$. For later uses, we note that (8) alternatively can be written

$$P(t,T) = E_t^Q[I_T] \tag{9a}$$

where I is given by the SDE

$$dI_s = -r_s I_s ds , \ \ I_t = 1, \ t_0 \leq t \leq s \leq \tau \tag{9b}$$

In both the CIR and Vasicek models, (8) can be evaluated analytically. The results are listed below for future reference.

CIR:

$$P(t,T) = A(t,T)e^{-r_t B(t,T)} \tag{10}$$

$$A(t,T) = \left[\frac{2\eta e^{(a+\eta)(T-t)/2}}{(\eta+a)(e^{\eta(T-t)}-1)+2\eta} \right]^{2ab/\sigma^2}$$

$$B(t,T) = \frac{2(e^{\eta(T-t)}-1)}{(\eta+a)(e^{\eta(T-t)}-1)+2\eta}$$

$$\eta = \sqrt{a^2 + 2\sigma^2}$$

Vasicek:

$$P(t,T) = C(t,T)e^{-r_t D(t,T)} \tag{11}$$

$$D(t,T) = \frac{1-e^{-\alpha(T-t)}}{\alpha}$$

$$C(t,T) = \exp\left[\frac{(D(t,T)-T+t)(\alpha^2\beta-\kappa^2/2)}{\alpha^2} - \frac{\kappa^2 D(t,T)^2}{4\alpha} \right]$$

3. Discretization Schemes for Simulation of Discount Bond Prices

3.1. The General Case

To develop schemes for the evaluation of the bond pricing equation (8) or, equivalently, (9a-b), we consider the vector SDE (under Q)

$$d\begin{pmatrix} r_t \\ I_t \end{pmatrix} = \begin{pmatrix} \mu_r(r_t,t) \\ -r_t I_t \end{pmatrix} dt + \begin{pmatrix} \sigma_r(r_t,t) \\ 0 \end{pmatrix} dW_t \,, \ t_0 \leq t \leq \tau \,, \qquad r_{t_0} = const. \in D, \ I_{t_0} = 1$$

$$\tag{12}$$

Assuming that σ_r and μ_r are sufficiently smooth, we can expand (12) in a first-order Ito-Taylor series[10]

$$\begin{pmatrix} r_s \\ I_s \end{pmatrix} = \begin{pmatrix} r_t \\ I_t \end{pmatrix} + \begin{pmatrix} \mu_r(r_t,t) \\ -r_t I_t \end{pmatrix}(s-t) + \begin{pmatrix} \sigma_r(r_t,t) \\ 0 \end{pmatrix}(W_s - W_t) + \vec{R}_1, \quad t_0 \le t \le s \le \tau$$

(13)

where the remainder vector $\vec{R}_1$ consists of multiple stochastic integrals of second order. The second-order Ito-Taylor expansion of (12) takes the form

$$\begin{pmatrix} r_s \\ I_s \end{pmatrix} = \begin{pmatrix} r_t \\ I_t \end{pmatrix} + \begin{pmatrix} \mu_r(r_t,t) \\ -r_t I_t \end{pmatrix}(s-t) + \begin{pmatrix} \sigma_r(r_t,t) \\ 0 \end{pmatrix}(W_s - W_t) +$$
$$\begin{pmatrix} \dfrac{1}{2} L^0 \mu_r(r_t,t) \\ \dfrac{1}{2} I_t \left(r_t^2 - \mu_r(r_t,t) \right) \end{pmatrix}(s-t)^2 + \begin{pmatrix} \dfrac{1}{2} L^1 \sigma_r(r_t,t) \\ 0 \end{pmatrix}\left((W_s - W_t)^2 - (s-t) \right) +$$
$$\begin{pmatrix} L^1 \mu_r(r_t,t) \\ -I_t \sigma_r(r_t,t) \end{pmatrix}(V_s - V_t) + \begin{pmatrix} L^0 \sigma_r(r_t,t) \\ 0 \end{pmatrix}\left((W_s - W_t)(s-t) - (V_s - V_t) \right) + \vec{R}_2$$

(14)

where the remainder $\vec{R}_2$ involves third-order stochastic integrals. In (14), we have introduced the operators

$$L^0 = \frac{\partial}{\partial t} + \mu_r(\cdot)\frac{\partial}{\partial r} + \frac{1}{2}\sigma_r^2(\cdot)\frac{\partial^2}{\partial r^2} ; \quad L^1 = \sigma_r(\cdot)\frac{\partial}{\partial r}$$

(15)

and a new Gaussian process[11] V_t given by $E_t^Q[V_s - V_t] = 0$, $E_t^Q[(V_s - V_t)^2] = (s-t)^3/3$, and $E_t^Q[(V_s - V_t)(W_s - W_t)] = (s-t)^2/2$.

To simulate a single path of r and I over the interval $[t_0, T]$, we discretize the time interval into K equidistant time steps of length $\delta = (T - t_0)/K$. Let $\hat{I}$ and $\hat{r}$ be the discrete-time approximations of the continuous-time variables r and I. By truncating off the remainder in the above Taylor expansions, we arrive at the following iterative schemes:

First-order Taylor (Euler):

$$\hat{I}_{i+1} = \hat{I}_i(1 - \hat{r}_i\delta) , \tag{16a}$$

$$\hat{r}_{i+1} = \hat{r}_i + \mu_r(\hat{r}_i, t_0 + i\delta)\delta + \sigma_r(\hat{r}_i, t_0 + i\delta)\tilde{z}_{i+1}\sqrt{\delta} ,$$
$$i = 0,1,...,K-1 , \; \hat{r}_0 = r_{t_0} , \; \hat{I}_0 = 1 \tag{16b}$$

Second-order Taylor (Milstein):

$$\hat{I}_{i+1} = \hat{I}_i\left(1 - \hat{r}_i\delta + \frac{1}{2}\left(\hat{r}_i^2 - \mu_r(\hat{r}_i, t_0 + i\delta)\right)\delta^2 - \frac{1}{2}\sigma_r(\hat{r}_i, t_0 + i\delta)\tilde{z}_{i+1}\delta\sqrt{\delta}\right), \tag{17a}$$

$$\hat{r}_{i+1} = \hat{r}_i + \left(\mu_r(\hat{r}_i, t_0 + i\delta) - \frac{1}{2}L^1\sigma_r(\hat{r}_i, t_0 + i\delta)\right)\delta + \sigma_r(\hat{r}_i, t_0 + i\delta)\tilde{z}_{i+1}\sqrt{\delta} +$$

$$\frac{1}{2}\left(L^1\mu_r(\hat{r}_i, t_0 + i\delta) + L^0\sigma_r(\hat{r}_i, t_0 + i\delta)\right)\tilde{z}_{i+1}\delta\sqrt{\delta} + \frac{1}{2}L^0\mu_r(\hat{r}_i, t_0 + i\delta)\delta^2 +$$

$$\frac{1}{2}L^1\sigma_r(\hat{r}_i, t_0 + i\delta)\tilde{z}_{i+1}^2\delta , \; i = 0,1,...,K-1 , \; \hat{r}_0 = r_{t_0} , \; \hat{I}_0 = 1$$

$$\tag{17b}$$

In (16a-b) and (17a-b) we have set $W_{i+1} - W_i = \tilde{z}_{i+1}\sqrt{\delta}$, where $\tilde{z}_1, \tilde{z}_2, ..., \tilde{z}_K$ is an independent sequence of standard Gaussian variables, $\tilde{z}_i \in N(0,1)$. In (17a-b) we have further set $V_{i+1} - V_i = \tilde{z}_{i+1}\delta\sqrt{\delta}/2$; this simplification can be shown to be valid for the purpose of evaluation *expectations* of the joint vector processes (12) (which is what we are interested in here); see Kloeden and Platen (1992), chapter 14 for a proof. In the literature, the first-order Taylor approximation (16a-b) is known as the *Euler scheme*, whereas (17a-b) is an extended version of the *Milstein scheme*.

If we apply (16a-b) or (17a-b) to generate N sample values of the process $\hat{I}$ at time T, $\hat{I}_K^1, \hat{I}_K^2, ..., \hat{I}_K^N$, we can estimate the expectation (9) as

IV.12

$$\hat{P}(t_0, T) = \frac{1}{N} \sum_{j=1}^{N} \hat{I}_K^j \tag{18}$$

To analyze the quality of the estimate $\hat{P}(t_0, T)$, it is convenient to decompose the total mean error

$$\hat{e} = \hat{P}(t_0, T) - P(t_0, T) \tag{19}$$

into two components, a systematic error e_{sys} and a statistical error $\hat{e}_{stat}$:

$$\hat{e} = e_{sys} + \hat{e}_{stat} \tag{20}$$

where e_{sys} is a constant

$$e_{sys} = E_{t_0}^Q[\hat{e}] = E_{t_0}^Q[\hat{P}(t_0, T)] - P(t_0, T) = E_{t_0}^Q[\hat{I}_K] - P(t_0, T) \tag{21}$$

and $\hat{e}_{stat}$ is a random variable with mean $E_{t_0}^Q[\hat{e}_{stat}] = 0$ and variance

$$V_{t_0}^Q[\hat{e}_{stat}] = V_{t_0}^Q[\hat{e}] = V_{t_0}^Q[\hat{P}(t_0, T)] = \frac{1}{N} V_{t_0}^Q[\hat{I}_K] \tag{22}$$

Under certain smoothness and regularity conditions on μ_r and σ_r, it can be shown that both the Euler and Milstein schemes have the property that $e_{sys} \to 0$ for $\delta \to 0$; however, the speed at which e_{sys} approaches zero depends on which scheme is used. More precisely, one can show (Kloeden and Platen (1992), p. 473-74) that the following inequalities hold

Euler Scheme:	$\left\|e_{sys}\right\| \leq C_1 \delta$	(23a)
Milstein Scheme:	$\left\|e_{sys}\right\| \leq C_2 \delta^2$	(23b)

where C_1 and C_2 are positive real constants independent of the step-size δ. Based on the exponent of δ in the above inequalities we say that the *Euler and Milstein scheme converge weakly with order 1 and 2, respectively*.

Schemes with weak convergence orders greater than 2 can, in principle, be constructed from high-order Ito-Taylor expansions of (12). Unfortunately, such expansions generally lead to non-trivial multiple stochastic integrals which frequently require the usage of quite complicated approximation techniques. An attractive alternative to the truncated Taylor series is *Richardson extrapolation* which essentially allows us to construct high-order schemes by combining results from low-order schemes. Suppose for example that we use the Milstein scheme (17a-b) to construct two bond price estimates, $\hat{P}_\delta(t_0, T)$ and $\hat{P}_{2\delta}(t_0, T)$, for step-sizes δ and 2δ, respectively. Under technical conditions on μ_r and σ_r, an error expansion result by Talay and Tubaro (1990) can be used to show that the combination

$$\hat{P}_{R,\delta}(t_0, T) = \frac{1}{3}\left(4\hat{P}_\delta(t_0, T) - \hat{P}_{2\delta}(t_0, T)\right) \tag{24}$$

converges with weak order 3. As in the deterministic case, Richardson extrapolation works through cancellation of leading order error terms.

As a final remark on SDE discretizations, we note that Kloeden and Platen (1992) have developed several derivative-free implicit and explicit discretization schemes (including versions of the well-known Runge-Kutta scheme frequently applied in the numerical solution of ODEs). Although we will not discuss such methods further here, they should prove helpful in problems where drift and volatility functions are not explicitly given but only known in a finite number of observed points.

3.2 Discretization and Test of the CIR Model

Applying the general formulas (16a-b) and (17a-b) to the CIR process[12] (4) yields the following discretization schemes:

Euler Scheme:

$$\hat{I}_{i+1} = \hat{I}_i(1-\hat{r}_i\delta) , \tag{25a}$$

$$\hat{r}_{i+1} = \hat{r}_i + a(b-\hat{r}_i)\delta + \sigma\sqrt{|\hat{r}_i|}\tilde{z}_{i+1}\sqrt{\delta} , \quad i = 0,1,...,K-1 , \; \hat{r}_0 = r_{t_0} , \; \hat{I}_0 = 1 \tag{25b}$$

Milstein Scheme:

$$\hat{I}_{i+1} = \hat{I}_i\left(1-\hat{r}_i\delta+\frac{1}{2}\left(\hat{r}_i^2 - a(b-\hat{r}_i)\right)\delta^2 - \frac{1}{2}\sigma\sqrt{|\hat{r}_i|}\tilde{z}_{i+1}\delta\sqrt{\delta}\right) , \tag{26a}$$

$$\hat{r}_{i+1} = \hat{r}_i + \left(a(b-\hat{r}_i)-\frac{\sigma^2}{4}\right)\delta + \sigma\sqrt{|\hat{r}_i|}\tilde{z}_{i+1}\sqrt{\delta} +$$

$$\frac{\sigma}{4}\left(\frac{1}{\sqrt{|\hat{r}_i|}}\left(a(b-\hat{r}_i)-\frac{\sigma^2}{4}\right)-2\sqrt{|\hat{r}_i|}a\right)\tilde{z}_{i+1}\delta\sqrt{\delta} - \frac{1}{2}a^2(b-\hat{r}_i)\delta^2 + \tag{26b}$$

$$\frac{\sigma^2}{4}\tilde{z}_{i+1}^2\delta , \quad i = 0,1,...,K-1 , \; \hat{r}_0 = r_{t_0} , \; \hat{I}_0 = 1$$

As the discretized CIR process -- unlike its continuous-time equivalent -- has a non-zero probability of generating negative short rates, notice that we have taken the precaution of using absolute values of $\hat{r}_i$ in all square-root operations in (25a-b) and (26a-b). Since the CIR model permits a closed-form bond pricing equation (see (10)), we can, in principle, test the convergence properties of (25a-b) and (26a-b). One complicating factor in convergence tests using a finite number of simulation paths is the presence of the statistical error term $\hat{e}_{stat}$ (see (22)) which necessitates construction of confidence intervals around the error estimate $\hat{e}$. One way of generating such confidence intervals involves arranging

the total number N of path simulations into B batches of S simulations each, $N = SB$. Provided that S is sufficiently large (say, larger than 20), the Central Limit Theorem implies that the batch errors will be approximately Gaussian. Using the standard Student T-test we can then use the sample variance

$$\hat{V}_B = \frac{1}{B(B-1)} \left(B \sum_{k=1}^{B} \hat{P}_k^2 - \left(\sum_{k=1}^{B} \hat{P}_k \right)^2 \right)$$

of the B batch averages $\hat{P}_1, \hat{P}_2, ..., \hat{P}_B$ to construct the necessary confidence intervals around the total mean error $\hat{e}$. For example, the 90% confidence interval for $\hat{e}$ has the form $\hat{e} \pm c_{0.9}$, $c_{0.9} = t_{0.9,B-1} \sqrt{\hat{V}_B / B}$ where the multiplier $t_{0.9,B-1}$ is the two-tailed 90th percentile of a Student T-distribution with B-1 degrees of freedom. Some examples of $t_{0.9,B-1}$ include: $t_{0.9,9} = 1.83$, $t_{0.9,99} = 1.66$, and $t_{0.9,999} = 1.65$.

In Table 1 below, we have listed simulation results for a 2-year discount bond assuming model parameters[13] of $a = 0.4$, $b = 0.1$, $\sigma = 0.1$, and an initial short rate of $r_0 = 0.06$. The table covers both the Euler scheme (25a-b), the Milstein scheme (26a-b), and the extrapolated third-order scheme (24). In an attempt to minimize the effects of the statistical error $\hat{e}_{stat}$ in (21), we have set $N = 1,000,000$ ($S = B = 1,000$). The table contains the 90% confidence intervals of the absolute error, $\hat{e} \pm c_{0.9}$, as well as a "worst-case" absolute error, defined by

$$\left| \hat{e}_{wc} \right| = MAX \left(\left| \hat{e} + c_{0.9} \right|, \left| \hat{e} - c_{0.9} \right| \right) \tag{27}$$

For comparisons of computational efficiency, the table lists calculation times T_{comp} (in seconds) of all simulations. The computations were run on a DEC Alpha 7610 mainframe using the standard pseudo-random generator `ran1()` (see Press *et al* (1992)) in combination with the Box-Mueller transformation.

Crude Monte Carlo Simulation Results for CIR Model
2-Year Discount Bond, Theoretical Value = 0.8655244

$N = 1,000,000; \ S = 1,000; \ B = 1,000$

$a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$

	K	1	2	4	8	16	32	64	128	256
	δ	2	1	0.5	0.25	0.125	0.0625	0.03125	0.01563	0.00781
	$\log_2(\delta)$	1.00	0.00	-1.00	-2.00	-3.00	-4.00	-5.00	-6.00	-7.00
Euler	$\hat{e}$	1.45E-2	3.06E-3	9.93E-4	3.62E-4	1.79E-4	7.60E-5	4.15E-5	4.53E-5	5.15E-5
	$\log_2|\hat{e}|$	-6.11	-8.35	-9.98	-11.43	-12.45	-13.68	-14.56	-14.43	-14.25
	$\sqrt{\hat{V}_B}$	0	7.49E-4	7.97E-4	8.39E-4	8.63E-4	8.50E-4	8.43E-4	8.53E-4	8.46E-4
	$c_{0.9}$	0	3.91E-5	4.16E-5	4.38E-5	4.50E-5	4.44E-5	4.40E-5	4.45E-5	4.42E-5
	$|\hat{e}_{wc}|$	1.45E-2	3.10E-3	1.03E-3	4.05E-4	2.24E-4	1.20E-4	8.54E-5	8.98E-5	9.56E-5
	$\log_2|\hat{e}_{wc}|$	-6.11	-8.33	-9.92	-11.27	-12.12	-13.02	-13.51	-13.44	-13.35
	T_{comp}	10.90	19.33	35.42	68.87	137.11	274.12	537.30	1065.74	2130.95
	$\log_2(T_{comp})$	3.45	4.27	5.15	6.11	7.10	8.10	9.07	10.06	11.06
Mil-stein	$\hat{e}$	-1.03E-2	-1.88E-3	-3.81E-4	-1.07E-4	-8.36E-6	-5.54E-6	2.88E-6	2.73E-5	4.28E-5
	$\log_2|\hat{e}|$	-6.60	-9.05	-11.36	-13.19	-16.87	-17.46	-18.40	-15.16	-14.51
	$\sqrt{\hat{V}_B}$	1.08E-3	9.08E-4	8.65E-4	8.69E-4	8.78E-4	8.56E-4	8.46E-4	8.55E-4	8.47E-4
	$c_{0.9}$	5.62E-5	4.74E-5	4.51E-5	4.54E-5	4.58E-5	4.47E-5	4.41E-5	4.46E-5	4.42E-5
	$|\hat{e}_{wc}|$	1.03E-2	1.93E-3	4.26E-4	1.53E-4	5.42E-5	5.02E-5	4.70E-5	7.19E-5	8.70E-5
	$\log_2|\hat{e}_{wc}|$	-6.60	-9.02	-11.20	-12.68	-14.17	-14.28	-14.38	-13.76	-13.49
	T_{comp}	14.15	27.62	56.10	103.90	206.21	404.36	822.91	1626.24	3260.38
	$\log_2(T_{comp})$	3.82	4.79	5.81	6.70	7.69	8.66	9.68	10.67	11.67
3rd Order	$\hat{e}$	NA	9.28E-4	1.32E-4	-2.42E-6	2.15E-5	-1.71E-6	1.75E-6	2.72E-5	4.80E-5
	$\log_2|\hat{e}|$	NA	-10.07	-12.89	-18.66	-15.51	-19.16	-19.13	-15.17	-14.35
	$\sqrt{\hat{V}_B}$	NA	1.12E-3	1.00E-3	9.88E-4	9.85E-4	9.65E-4	9.45E-4	9.49E-4	9.54E-4
	$c_{0.9}$	NA	5.83E-5	5.23E-5	5.16E-5	5.14E-5	5.04E-5	4.93E-5	4.95E-5	4.98E-5
	$|\hat{e}_{wc}|$	NA	9.86E-4	1.84E-4	5.40E-5	7.29E-5	5.21E-5	5.10E-5	7.67E-5	9.78E-5
	$\log_2|\hat{e}_{wc}|$	NA	-9.99	-12.41	-14.18	-13.74	-14.23	-14.26	-13.67	-13.32
	T_{comp}	NA	34.11	64.13	123.22	243.48	478.95	951.60	1887.33	3842.11
	$\log_2(T_{comp})$	NA	5.09	6.00	6.94	7.93	8.90	9.89	10.88	11.91

Table 1

IV.17

Similar simulations have been run for 1-, 5-, and 10-year discount bonds; the results have been placed in Appendix A. The behavior of the 1-, 5-, and 10-year simulated bond prices closely resemble that of the 2-year bond prices.

For the 2-year discount bond, Figures 1 and 2 show logarithmic (base 2) graphs of the absolute mean error and, more importantly, the absolute worst-case error against the size of the time-step. Although the different convergence orders of the used schemes are apparent for large values of the time-step $(\delta \geq 2^{-2})$, the Milstein and third-order schemes quickly reach an error level of around $|\hat{e}_{wc}| \approx 2^{-14} = 0.61$ basis points (bp) beyond which increasing the number of steps in the path simulation yields no additional improvements. The existence of this bound on accuracy is caused by the variance (22) of the bond price estimator. Despite the fact that we have simulated as many as 1,000,000 sample paths, the convergence behavior of the higher-order schemes thus essentially "drowns" in random noise. Indeed, as is apparent from Figure 3, when $\delta < 2^{-2}$ the third-order scheme has converged to within the statistical error on $\hat{e}$ (as given by the 90% confidence interval). The same effect occurs in the Milstein scheme when $\delta < 2^{-3}$.

Figure 1

Figure 2

IV.19

3rd Order Scheme: Confidence Interval of ê vs. Time-Step (δ)

$T - t_0 = 2, \ a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$

$S = 1,000; \ B = 1,000; \ N = 1,000,000$
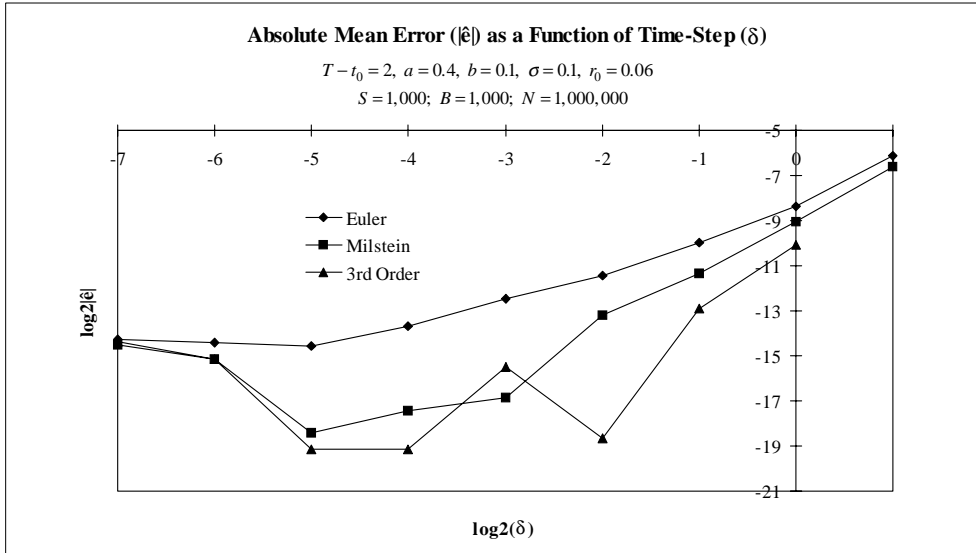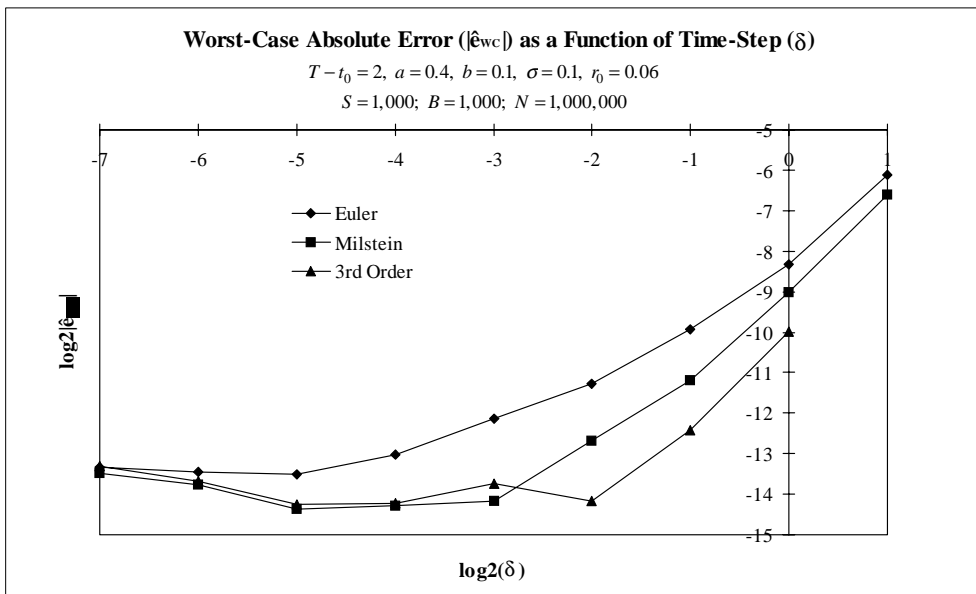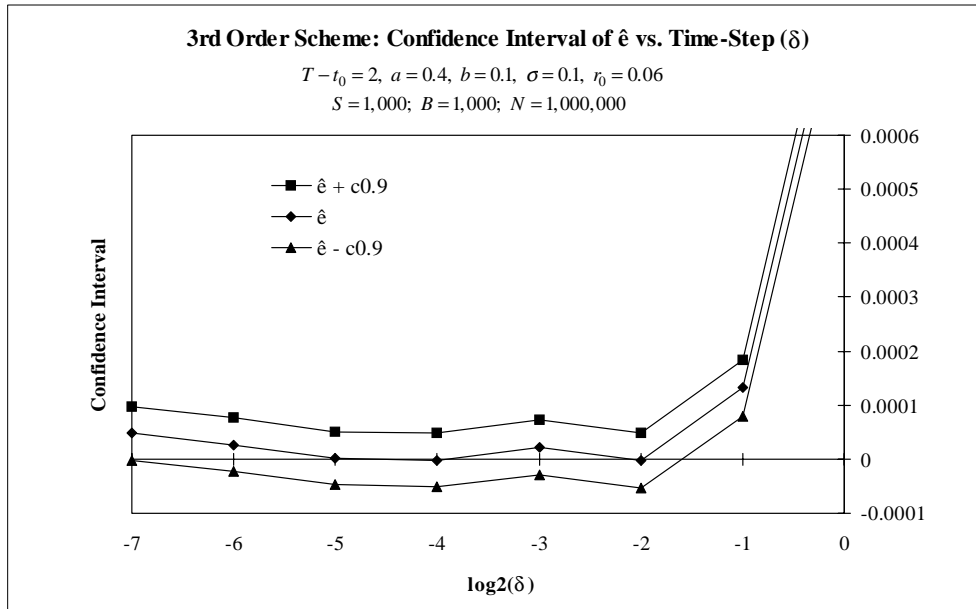
Figure 3

For practical pricing problems where we cannot rely on as many as 1,000,000 simulations, the problems associated with the estimator variance are clearly going to be even more severe. For a more realistic number of simulated paths, say $N = 10,000$, the worst-case error plateau can be expected to increase by around a factor 10 ($= \sqrt{1,000,000/10,000}$), as confirmed by Figure 4 below. Notice in the figure that the third-order scheme is completely overshadowed by the statistical error and exhibits no convergence behavior whatsoever.
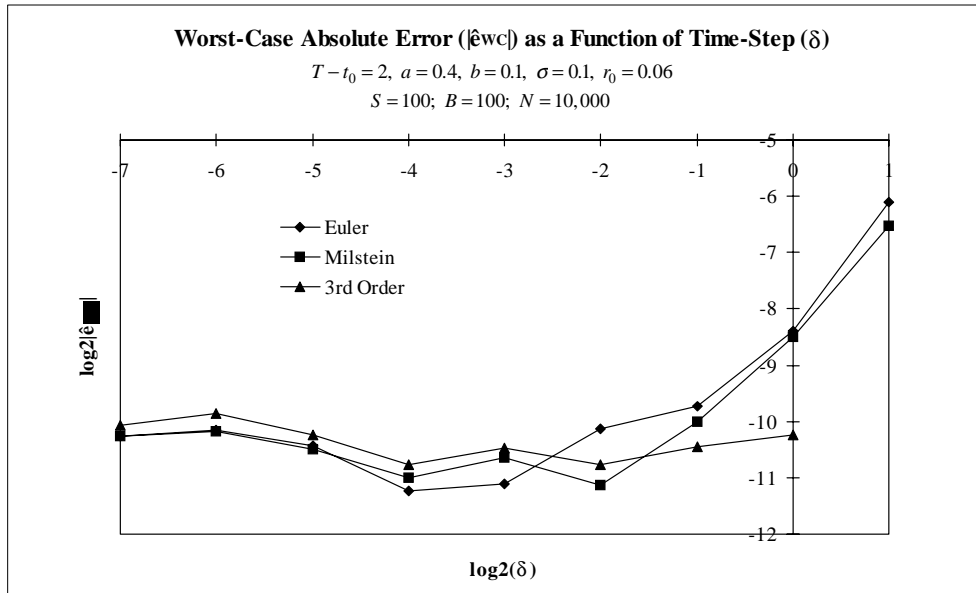
Figure 4

Despite the problems surrounding the high variance of the bond price estimator (which we will discuss further in the next section), it is obvious from Figures 1 and 2 that the higher-order schemes nevertheless exhibit promising convergence behavior. For example, to reach a (worst-case) accuracy below 2^{-12}, the Euler scheme needs 32 time-steps per path, whereas the Milstein and third-order schemes only require 8 and 4 steps, respectively. To properly account for the additional computational burden of the higher-order schemes, Figure 5 graphs the worst-case error of the different schemes against the required computation time. It is clear from the figure that even after adjusting for the increased computational overhead, the higher-order schemes are more efficient than the Euler scheme. For our particular example, it also appears that the third-order scheme is slightly more efficient than the Milstein scheme.

IV.21
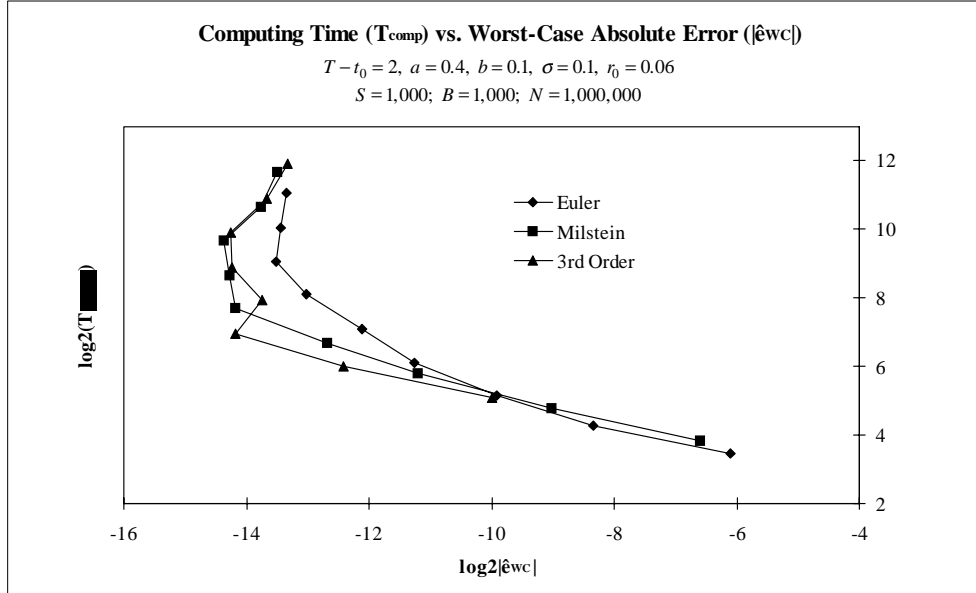
Figure 5

4. Variance Reduction Methods

4.1 Traditional Variance Reduction Methods

It is evident from the results so far, that for high-order discretization schemes to be of any practical use, we need methods to reduce the variance of the bond price estimator (18). Initially, we will focus on two widely used techniques from the classical Monte Carlo theory, namely *antithetic variates* and *control variates*. To briefly introduce these methods, let us restate the bond pricing problem in a slightly modified form. From (18), we know that discount bond prices can be approximated from the random variables

$$\hat{I}_K^j = f(\tilde{z}_1^j, \tilde{z}_2^j, ..., \tilde{z}_K^j) \ , \ \ j = 1, 2, ..., N \tag{28}$$

IV.22

where $\tilde{z}_1^{\,j}, \tilde{z}_2^{\,j},..., \tilde{z}_K^{\,j}$ is an independent sequence of standard Gaussian variables and $f : \Re^K \to \Re$ is an (implicit) well-behaved function given by the discretization scheme used. Instead of using (28) directly to evaluate (18), the *antithetic variate method* uses the N variables

$$\hat{I}_{K,A}^{\,j} = \frac{1}{2}\Big(f(\tilde{z}_1^{\,j}, \tilde{z}_2^{\,j},..., \tilde{z}_K^{\,j}) + f(-\tilde{z}_1^{\,j}, -\tilde{z}_2^{\,j},..., -\tilde{z}_K^{\,j})\Big) = \frac{1}{2}\Big(\hat{I}_K^{\,j} + \hat{I}_{K-}^{\,j}\Big), \quad j = 1,2,...,N$$

(29)

where the "mirror" process $\hat{I}_{K-}$ is obtained by changing the sign on all draws from the standard Gaussian distribution. The mean of $\hat{I}_{K,A}$ is clearly identical to that of $\hat{I}_K$, whereas the variance is

$$V_{t_0}^Q[\hat{I}_{K,A}] = \frac{1}{2} V_{t_0}^Q[\hat{I}_K]\Big(\rho(\hat{I}_K, \hat{I}_{K-}) + 1\Big)$$

(30)

where $\rho(\hat{I}_K, \hat{I}_{K-})$ denotes the correlation coefficient between $\hat{I}_K$ and $\hat{I}_{K-}$. If this correlation coefficient is close to -1, the variance of $\hat{I}_{K,A}$ will be significantly smaller[14] than the variance of $\hat{I}_K$.

Whereas the antithetic variate method relies upon the existence of a process negatively correlated to $\hat{I}_K$, the *control variate method* is based upon sampling a process positively correlated to $\hat{I}_K$. One way[15] to introduce a control variate is to consider an alternative short rate model, say

$$d\begin{pmatrix} r_t^* \\ I_t^* \end{pmatrix} = \begin{pmatrix} \mu_r^*(r_t^*, t) \\ -r_t^* I_t^* \end{pmatrix} dt + \begin{pmatrix} \sigma_r^*(r_t^*, t) \\ 0 \end{pmatrix} dW_t, \quad t_0 \le t \le \tau, \quad r_{t_0}^* = r_{t_0} = const. \in D^*, \quad I_{t_0}^* = 1$$

(31)

IV.23

for which an analytical solution to discount bond prices is known (like the CIR model (4) or the Vasicek model (5)). Now instead of (18), we write

$$\hat{P}_C(t_0,T) = \frac{1}{N}\sum_{j=1}^{N} \hat{I}_K^j - \frac{1}{N}\sum_{j=1}^{N} \hat{I}_K^{*j} + P^*(t_0,T) = \frac{1}{N}\sum_{j=1}^{N}(\hat{I}_K^j - \hat{I}_K^{*j}) + P^*(t_0,T)$$

(32)

where $P^*(t_0,T)$ is the known theoretical value of the control variate bond price. The mean and variance of $\hat{P}_C(t_0,T)$ are given by

$$E_{t_0}^Q[\hat{P}_C(t_0,T)] = P(t_0,T) + e_{sys} - e_{sys}^*$$

(33a)

$$\begin{aligned} V_{t_0}^Q[\hat{P}_C(t_0,T)] &= \frac{1}{N}\left(V_{t_0}^Q[\hat{I}_K] + V_{t_0}^Q[\hat{I}_K^*] - 2COV_{t_0}^Q(\hat{I}_K,\hat{I}_K^*)\right) \\ &= V_{t_0}^Q[\hat{e}_{stat}] + V_{t_0}^Q[\hat{e}_{stat}^*] - \frac{2}{N}COV_{t_0}^Q(\hat{I}_K,\hat{I}_K^*) \end{aligned}$$

(33b)

It follows easily from the triangle inequality that

$$|e_{sys}| \leq C\delta^n \ , \ |e_{sys}^*| \leq C^*\delta^n \ \Rightarrow \ |e_{sys} - e_{sys}^*| \leq |e_{sys}| + |e_{sys}^*| \leq (C + C^*)\delta^n$$

(34)

From (33a) we can therefore conclude that if both (12) and (31) are discretized by nth-order weak schemes, the estimate $\hat{P}_C(t_0,T)$ will converge with (at least) weak order n as well. Further, from (33b) we see that the variance of $\hat{P}_C(t_0,T)$ will be less than the variance of estimate $\hat{P}(t_0,T)$ (see (22)) if

$$COV_{t_0}^Q(\hat{I}_K,\hat{I}_K^*) > \frac{1}{2}V_{t_0}^Q(\hat{I}_K^*) \Rightarrow \rho(\hat{I}_K,\hat{I}_K^*) > \frac{1}{2}\sqrt{\frac{V_{t_0}^Q(\hat{I}_K^*)}{V_{t_0}^Q(\hat{I}_K)}}$$

(35)

To ensure that (35) holds, it is important that the parameters of the control variate process (31) are chosen to match the dynamics of the original process (12) as closely as possible.

4.2 Application of Traditional Variance Reduction to the CIR Model

Whereas implementation of the antithetic variate method on the CIR model is straightforward, implementation of the control variate technique involves selecting an appropriate alternative process (31). Here, we will use the Vasicek process (5) which is structurally quite similar to the CIR process (4); Appendix B contains the Euler and Milstein discretization schemes for the Vasicek model. The third-order Vasicek scheme can be generated from (24). There are several ways to pick the parameters α, β, and κ in (5); we will choose the parameters to match the first and second moments of the CIR model[16]. From (6) and (7) we thus get

$$\alpha = a \ , \ \ \beta = b \ , \ \ \kappa^2 = \frac{\sigma^2 \left(2r_{t_0} \left(e^{-a(T-t_0)} - e^{-2a(T-t_0)} \right) + b \left(1 - e^{-a(T-t_0)} \right)^2 \right)}{\left(1 - e^{-2a(T-t_0)} \right)}$$

(36)

In Figure 6 we have shown some comparisons of CIR and Vasicek bond prices (see (10) and (11)) using the parameter choice (36).

PCIR - PVasicek for Different Bond Maturities

b = 0.1, r₀ = 0.06
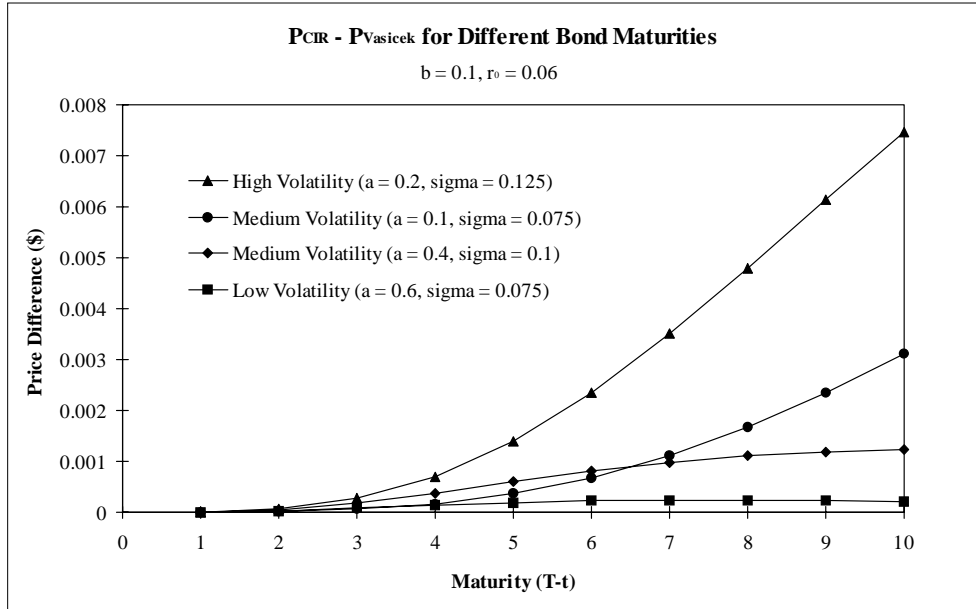
Figure 6

Although, as expected, the CIR and Vasicek bond prices diverge somewhat as the bond volatility increases (that is, when the mean reversion rate a decreases, the short rate volatility σ increases, and/or the maturity of the bond increases), the quality of the Vasicek control variate nevertheless appears to be quite satisfactory. Using the same CIR parameters as in Section 3.2, Table 2 on the next page shows batch variance and confidence intervals for 10,000 simulations of a 5-year discount bond. The data have been generated using i) crude Monte Carlo; ii) antithetic variate method; iii) control variate method; and iv) control variate method + antithetic variate method.

Simulation Results for CIR Model
5-Year Discount Bond, Theoretical Value = 0.6642841

$N = 10,000;\ S = 100;\ B = 100$
$a = 0.4,\ b = 0.1,\ \sigma = 0.1,\ r_0 = 0.06$

	K	4	8	16	32	64	128	256
	δ	1.25	0.625	0.3125	0.1563	0.07813	0.03906	0.01953
	Crude Monte Carlo							
Euler	$\hat{V}_B^{1/2}$	0.00579	0.00541	0.00557	0.00510	0.00576	0.00610	0.00686
	$c_{0.9}$	9.61E-4	8.98E-4	9.25E-4	8.46E-4	9.57E-4	1.01E-3	1.14E-3
	T_{comp}	0.34	0.68	1.35	2.75	5.34	10.78	21.3
Milstein	$\hat{V}_B^{1/2}$	0.00543	0.00519	0.00540	0.00506	0.00577	0.00611	0.00687
	$c_{0.9}$	9.01E-4	8.62E-4	8.97E-4	8.40E-4	9.57E-4	1.01E-3	1.14E-3
	T_{comp}	0.55	1.03	2.05	4.02	8.23	16.28	32.65
3rd Order	$\hat{V}_B^{1/2}$	0.00632	0.00590	0.00633	0.00577	0.00665	0.00718	0.00789
	$c_{0.9}$	1.05E-3	9.80E-4	1.05E-3	9.57E-4	1.10E-3	1.19E-3	1.31E-3
	T_{comp}	0.64	1.23	2.43	4.79	9.53	18.99	38.44
	Antithetic Variate							
Euler	$\hat{V}_B^{1/2}$	7.26E-4	7.30E-4	6.98E-4	6.33E-4	6.57E-4	6.10E-4	5.79E-4
	$c_{0.9}$	1.21E-4	1.21E-4	1.16E-4	1.05E-4	1.09E-4	1.01E-4	9.61E-5
	T_{comp}	0.52	0.97	1.89	3.71	7.49	15.21	30.14
Milstein	$\hat{V}_B^{1/2}$	5.52E-4	5.90E-4	5.48E-4	6.02E-4	6.57E-4	6.11E-4	5.83E-4
	$c_{0.9}$	9.17E-5	9.79E-5	9.09E-5	1.00E-4	1.09E-4	1.01E-4	9.68E-5
	T_{comp}	0.85	1.68	3.34	6.60	12.90	25.91	51.87
3rd Order	$\hat{V}_B^{1/2}$	7.08E-4	6.74E-4	6.41E-4	7.50E-4	7.96E-4	7.31E-4	6.96E-4
	$c_{0.9}$	1.18E-4	1.12E-4	1.06E-4	1.24E-4	1.32E-4	1.21E-4	1.16E-4
	T_{comp}	1.07	2.05	4.00	8.18	15.68	31.82	63.64

Table 2 (Part 1)

	K	4	8	16	32	64	128	256
	δ	1.25	0.625	0.3125	0.1563	0.07813	0.03906	0.01953
	Control Variate							
Euler	$\hat{V}_B^{1/2}$	0.00160	0.00136	0.00130	0.00124	0.00141	0.00153	0.00138
	$c_{0.9}$	2.65E-4	2.25E-4	2.16E-4	2.06E-4	2.35E-4	2.54E-4	2.30E-4
	T_{comp}	0.42	0.78	1.54	3.10	5.83	11.76	23.80
Milstein	$\hat{V}_B^{1/2}$	0.00130	0.00126	0.00120	0.00123	0.00141	0.00151	0.00137
	$c_{0.9}$	2.15E-4	2.08E-4	2.00E-4	2.05E-4	2.34E-4	2.51E-4	2.28E-4
	T_{comp}	0.65	1.24	2.37	4.64	9.11	18.03	36.21
3rd Order	$\hat{V}_B^{1/2}$	0.00165	0.00144	0.00142	0.00152	0.00164	0.00175	0.00160
	$c_{0.9}$	2.73E-4	2.40E-4	2.36E-4	2.53E-4	2.72E-4	2.90E-4	2.65E-4
	T_{comp}	0.71	1.39	2.86	5.43	10.92	21.8	42.30
	Antithetic + Control Variate							
Euler	$\hat{V}_B^{1/2}$	0.00111	0.00112	0.00110	0.00105	0.00113	0.00114	0.00097
	$c_{0.9}$	1.84E-4	1.86E-4	1.83E-4	1.74E-4	1.87E-4	1.90E-4	1.62E-4
	T_{comp}	0.59	1.22	2.31	4.49	7.87	16.99	32.91
Milstein	$\hat{V}_B^{1/2}$	0.00104	0.00103	0.00099	0.00104	0.00114	0.00114	0.00098
	$c_{0.9}$	1.73E-4	1.71E-4	1.64E-4	1.73E-4	1.89E-4	1.90E-4	1.63E-4
	T_{comp}	0.99	1.93	3.59	7.01	13.98	28.61	54.88
3rd Order	$\hat{V}_B^{1/2}$	0.00131	0.00119	0.00118	0.00131	0.00135	0.00136	0.00119
	$c_{0.9}$	2.17E-4	1.97E-4	1.96E-4	2.18E-4	2.24E-4	2.25E-4	1.97E-4
	T_{comp}	1.29	2.49	4.63	8.99	17.81	35.72	71.03

Table 2 (Part 2)

To aggregate the results in Table 2 and to properly account for the differences in computation time, we define an average efficiency ratio

$$E = \frac{const.}{\left(\hat{V}_B^{Euler} + \hat{V}_B^{Milstein} + \hat{V}_B^{3rd}\right)\left(T_{comp}^{Euler} + T_{comp}^{Milstein} + T_{comp}^{3rd}\right)} \quad (37)$$

where we pick the constant in the numerator to normalize E to 1 for crude Monte Carlo simulation. The results of this calculation are shown in Table 3 and graphed in Figure 7.

Efficiency Ratios for Simulation of 5-Year Discount Bond in CIR Model

$$N = 10,000; \; S = 100; \; B = 100$$
$$a = 0.4, \; b = 0.1, \; \sigma = 0.1, \; r_0 = 0.06$$

K	4	8	16	32	64	128	256
δ	1.25	0.625	0.3125	0.1563	0.07813	0.03906	0.01953
Crude Monte Carlo	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Antithetic Variate	5.54	5.18	5.79	5.01	5.52	6.27	7.38
Control Variate	3.32	3.51	3.80	3.49	3.64	3.61	4.48
Antithetic + Control Variate	2.71	2.58	2.93	2.64	2.93	3.01	4.00

Table 3

IV.29

Efficiency Ratios for Traditional Variance Reduction Methods

$T - t_0 = 5, \ a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$

$S = 100; \ B = 100; \ N = 10,000$
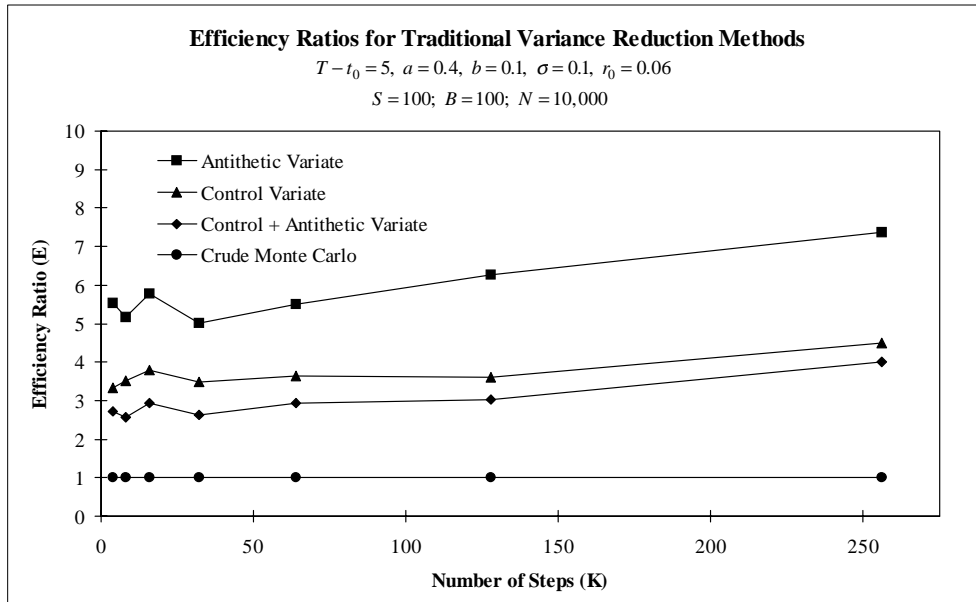
Figure 7

It is obvious from Table 2-3 and Figure 7 that all three tested variance reduction methods yield significant improvements over crude Monte Carlo. Somewhat surprisingly, however, the basic antithetic method appears to outperform the methods that involve control variates, even the combined control/antithetic variate approach. To explain this last result, we used raw simulation data to calculate the following sample correlation coefficients (Euler scheme, $K = 32$ steps):

IV.30

Sample Correlation Coefficients for CIR Model
Euler Simulation of 5-Year Discount Bond

$$N = 10{,}000; \ S = 100; \ B = 100; \ K = 32$$
$$a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$$

Correlation Coefficient	Value
$\hat{\rho}(\hat{I}_K, \hat{I}_{K-})$	-0.978
$\hat{\rho}(\hat{I}_K, \hat{I}_K^*)$	0.983
$\hat{\rho}((\hat{I}_K + \hat{I}_{K-}), (\hat{I}_K^* + \hat{I}_{K-}^*))$	-0.758

Table 4

The correlation coefficients for both the antithetic and the control variate methods are very close to their optimal values of -1 and +1, respectively. Interestingly, however, the correlation between the combined variates $(\hat{I}_K + \hat{I}_{K-})/2$ and $(\hat{I}_K^* + \hat{I}_{K-}^*)/2$ is *negative*. According to (35), adding the control variate method to the antithetic method will thus cause an increase in variance, consistent with the experimental results in Table 2. Although there are other applications (see for example Clewlow and Carverhill (1994)) where the combination of the antithetic and control variate methods will outperform either method alone, it is obvious from the above findings that an uncritical combination of variance reduction techniques can lead to suboptimal results.

For completeness, Appendix C lists efficiency ratios for 1-, 2-, and 10-year discount bonds. Except for the 1-year bond where the control variate method slightly outperforms the antithetic variate method, the efficiency results are very similar to those of the 5-year bond. Notice, that the efficiency of the control variate method falls with increasing bond maturities, an effect that can easily be understood from Figure 6: the higher the bond maturity, the poorer the quality of the Vasicek control variate.

4.3 The Measure Transform Method

In the previous sections, we illustrated how knowledge about a simpler control variate process can, in principle, be used to improve simulation results. For the specific example of using the Vasicek model as a control variate process for the CIR model the results, however, were somewhat disappointing. In this section we will discuss an alternative, SDE-based method to incorporate results obtained from simpler processes into the simulation procedure. The method was proposed by Milstein (1988) and is based on a (reverse) application of the Girsanov Theorem for shift of probability measure (Karatzas and Shreve (1991), p. 190-201).

To introduce this technique, we consider a process θ_t and an equivalent probability measure Z under which

$$d\widetilde{W}_t = dW_t + \theta_t dt \tag{38}$$

is a standard Brownian motion on (Ω, F, Z). Under technical conditions on θ_t (see footnote 6), the Girsanov Theorem asserts that the measure Z is related to Q through the Radon-Nikodym derivative

$$\frac{dQ}{dZ} = \exp\left(\int_{t_0}^{\tau} \theta_s d\widetilde{W}_s - \frac{1}{2} \int_{t_0}^{\tau} \theta_s^2 ds \right) \tag{39}$$

where as before $[t_0, \tau]$ represents our bounded trading horizon. Corresponding to (39), we introduce the *likelihood ratio process*

$$\xi_t = E_t^Z\left(\frac{dQ}{dZ} \right) = \exp\left(\int_{t_0}^{t} \theta_s d\widetilde{W}_s - \frac{1}{2} \int_{t_0}^{t} \theta_s^2 ds \right)$$

or

$$d\xi_t = \theta_t \xi_t d\widetilde{W}_t , \quad \xi_{t_0} = 1 \tag{40}$$

which obviously is a martingale under Z. Expectations under Z and Q can now be shown to be related through

$$E_{t_0}^Q[X_t] = E_{t_0}^Z[X_t \xi_t] \tag{41}$$

for any absolutely integrable $\mathfrak{I}_t$-measurable random variable X_t, $t \in [t_0, \tau]$. Notice that (41) is independent of the actual choice of ξ_t; if we can chose ξ_t such that the variance of the product $X_t \xi_t$ (under Z) is smaller than the variance of X_t (under Q), (41) can be used as a variance reduction scheme.

We now return to the joint process (12) which we amend to include the likelihood ratio process (40). Under Z, we have

$$d \begin{pmatrix} r_t \\ I_t \\ \xi_t \end{pmatrix} = \begin{pmatrix} \mu_r(r_t,t) - \theta_t \sigma_r(r_t,t) \\ -r_t I_t \\ 0 \end{pmatrix} dt + \begin{pmatrix} \sigma_r(r_t,t) \\ 0 \\ \theta_t \xi_t \end{pmatrix} d\widetilde{W}_t, \quad r_{t_0} = const. \in D, \ I_{t_0} = \xi_{t_0} = 1 \tag{42}$$

and the corresponding bond pricing equation

$$P(t_0,T) = E_t^Z[I_T \xi_T] \tag{43}$$

The question now arises: how do we pick θ_t to minimize the variance of $I_T \xi_T$? If we introduce the function[17] $h_T : D \times [t_0,T] \to \mathfrak{R}^+$ such that $P(t,T) = h_T(r_t,t)$, Appendix D (equation (D.8)) shows that choosing

IV.33

$$\theta_t = \frac{-\partial h_T(r_t,t)}{\partial r} \frac{\sigma_r(r_t,t)}{h_T(r_t,t)} = g_T(r_t,t) , \quad g_T : D \times [t_0,T] \to \Re \qquad (44)$$

will reduce the variance of $I_T \xi_T$ to zero (under Z). Unfortunately, we generally do not know the explicit relation $P(t,T) = h_T(r_t,t)$ -- if we did, there would be no need to evaluate (43) through simulation. However, in many circumstances we can come up with a good guess for (44), for example by using known bond formulas for simpler interest rate models. Suppose, say, we believe that our interest rate model yields bond prices which are reasonably close to those of the Vasicek (5) model (with appropriately chosen parameters). Applying the Vasicek bond formula (11) to (44) yields the simple result

$$g_T^{Vasicek}(r_t,t) = D(t,T)\sigma_r(r_t,t), \ \ D(t,T) = \frac{1 - e^{-\alpha(T-t)}}{\alpha} \qquad (45)$$

Discretization schemes for the SDE (42) can be derived along the same lines as in Section 3.1. The Euler and Milstein schemes for r_t and I_t will be identical to (16a-b) and (17a-b) provided $\mu_r(r_t,t)$ is replaced by $\mu_{r,MT}(r_t,t) = \mu_r(r_t,t) - g_T(r_t,t)\sigma_r(r_t,t)$. The schemes for the likelihood ratio process ξ_t are as follows

Euler:

$$\hat{\xi}_{i+1} = \hat{\xi}_i(1 + g_T(\hat{r}_i, t_0 + i\delta)\tilde{z}_{i+1}\sqrt{\delta}) , \quad i = 0,1,...,K-1 , \ \hat{r}_0 = r_{t_0} , \ \hat{\xi}_0 = 1 \ \ (46)$$

Milstein:

$$\frac{\hat{\xi}_{i+1}}{\hat{\xi}_i} = 1 - \frac{1}{2}\left(L^1 g_T(\hat{r}_i, t_0 + i\delta) + g_T(\hat{r}_i, t_0 + i\delta)^2\right)\delta + g_T(\hat{r}_i, t_0 + i\delta)\tilde{z}_{i+1}\sqrt{\delta} +$$

$$\frac{1}{2}\left(L^0_{MT} g_T(\hat{r}_i, t_0 + i\delta) + g_T(\hat{r}_i, t_0 + i\delta)L^1 g_T(\hat{r}_i, t_0 + i\delta)\right)\tilde{z}_{i+1}\delta\sqrt{\delta} +$$

$$\frac{1}{2}\left(L^1 g_T(\hat{r}_i, t_0 + i\delta) + g_T(\hat{r}_i, t_0 + i\delta)^2\right)\tilde{z}_{i+1}^2\delta, \quad i = 0,1,...,K-1, \quad \hat{r}_0 = r_{t_0}, \quad \hat{\xi}_0 = 1$$

$$(47)$$

where the differential operator L^0_{MT} is identical to L^0 in (15), except that $\mu_r(r_t, t)$ must be replaced by $\mu_{r,MT}(r_t, t) = \mu_r(r_t, t) - g_T(r_t, t)\sigma_r(r_t, t)$.

4.4 Application of the Measure Transform Method to the CIR Model

We now turn to applying the measure transform method to the CIR process (4). We will use the Vasicek model (with $\alpha = a$, see (36)) to generate the function $h_T(r_t, t)$; consequently, the volatility of the likelihood ratio process is given by (45). From (10), we note that the (continuous-time) perfect choice of $g_T(r_t, t)$ is

$$g_T^{CIR}(r_t, t) = B(t,T)\sigma\sqrt{r_t}, \quad B(t,T) = \frac{2(e^{\sqrt{a^2 + 2\sigma^2}(T-t)} - 1)}{(\sqrt{a^2 + 2\sigma^2} + a)(e^{\sqrt{a^2 + 2\sigma^2}(T-t)} - 1) + 2\sqrt{a^2 + 2\sigma^2}}$$

$$(48)$$

The graph below compares $B(t,T)$ to $D(t,T)$ for selected values of σ, a, and *T-t*; as for the control variate technique (see Figure 6), the difference between $B(t,T)$ and $D(t,T)$ grows with increasing bond volatility. Nevertheless, the fit appears to be quite good, particularly for short- to medium-term discount bonds.

B(t,T) and D(t,T) for Different Bond Maturities

b = 0.1, r₀ = 0.06
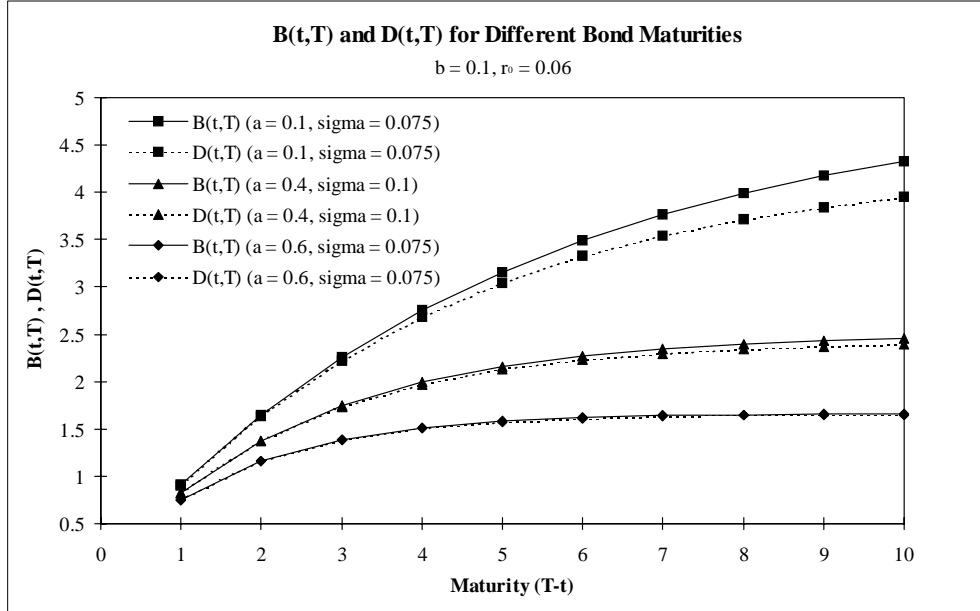
Figure 8

Applying (16a-b), (17a-b), (46), and (47) to the CIR process now yields the following schemes

Euler Scheme:

$$\hat{I}_{i+1} = \hat{I}_i(1 - \hat{r}_i\delta) \,, \tag{48a}$$

$$\hat{\xi}_{i+1} = \hat{\xi}_i(1 + D(t_0 + i\delta, T)\sigma\sqrt{|\hat{r}_i|}\tilde{z}_{i+1}\sqrt{\delta}) \,, \tag{48b}$$

$$\hat{r}_{i+1} = \hat{r}_i + \left(ab - \hat{r}_i(a + \sigma^2 D(t_0 + i\delta, T))\right)\delta +$$

$$\sigma\sqrt{|\hat{r}_i|}\tilde{z}_{i+1}\sqrt{\delta} \,, \quad i = 0,1,...,K-1 \,, \quad \hat{r}_0 = r_{t_0} \,, \quad \hat{I}_0 = \hat{\xi}_0 = 1 \tag{48c}$$

Milstein Scheme:

$$\hat{I}_{i+1} = \hat{I}_i\left(1 - \hat{r}_i\delta + \frac{1}{2}(\hat{r}_i(\psi_i + \hat{r}_i) - ab)\delta^2 - \frac{1}{2}\sigma\sqrt{|\hat{r}_i|}\tilde{z}_{i+1}\delta\sqrt{\delta}\right) \,, \tag{49a}$$

IV.36

$$\frac{\hat{\xi}_{i+1}}{\hat{\xi}_i} = 1 - \frac{(\psi_i - a)}{2}\left(D(t_0 + i\delta, T)\hat{r}_i + \frac{1}{2}\right)\delta + D(t_0 + i\delta, T)\sigma\sqrt{|\hat{r}_i|}\tilde{z}_{i+1}\sqrt{\delta} +$$

$$\frac{D(t_0 + i\delta, T)\sigma}{4}\left(\sqrt{|\hat{r}_i|}\left(a - \frac{2}{D(t_0 + i\delta, T)}\right) + \frac{1}{\sqrt{|\hat{r}_i|}}\left(ab - \frac{\sigma^2}{4}\right)\right)\tilde{z}_{i+1}\delta\sqrt{\delta} +$$

$$\frac{(\psi_i - a)}{2}\left(D(t_0 + i\delta, T)\hat{r}_i + \frac{1}{2}\right)\tilde{z}_{i+1}^2\delta,$$

<div align="right">(49b)</div>

$$\hat{r}_{i+1} = \hat{r}_i + \left(ab - \hat{r}_i\psi_i - \frac{\sigma^2}{4}\right)\delta + \sigma\sqrt{|\hat{r}_i|}\tilde{z}_{i+1}\sqrt{\delta} + \frac{\sigma}{4}\left(\frac{1}{\sqrt{|\hat{r}_i|}}\left(ab - \frac{\sigma^2}{4}\right) - 3\sqrt{|\hat{r}_i|}\psi_i\right)\tilde{z}_{i+1}\delta\sqrt{\delta} +$$

$$\frac{1}{2}\left(\hat{r}_i\left(a\psi_i + \sigma^2 + D(t_0 + i\delta, T)^2\sigma^4\right) - ba\psi_i\right)\delta^2 +$$

$$\frac{\sigma^2}{4}\tilde{z}_{i+1}^2\delta, \quad i = 0,1,...,K-1, \quad \hat{r}_0 = r_{t_0}, \quad \hat{I}_0 = 1$$

<div align="right">(49c)</div>

where

$$\psi_i \equiv a + D(t_0 + i\delta, T)\sigma^2$$

We notice that, similar to the control variate technique, the measure transformation method can be combined with the antithetic method. As the measure transformation method is not based upon a correlation argument, we are less likely to experience the difficulties we encountered in Section 4.2 when we attempted to combine antithetic and control variates.

Applying the above schemes to the example in Section 4.2 (5-year bond, $N = 10,000$) yields the following results:

Simulation Results for CIR Model
5-Year Discount Bond, Theoretical Value = 0.6642841

$N = 10{,}000;\ S = 100;\ B = 100$

$a = 0.4,\ b = 0.1,\ \sigma = 0.1,\ r_0 = 0.06$

	K	4	8	16	32	64	128	256
	δ	1.25	0.625	0.3125	0.1563	0.07813	0.03906	0.01953
	Measure Transformation							
Euler	$\hat{V}_B^{1/2}$	2.21E-3	1.08E-3	5.05E-4	2.67E-4	1.51E-4	8.80E-5	7.53E-5
	$c_{0.9}$	3.68E-4	1.79E-4	8.38E-5	4.43E-5	2.51E-5	1.46E-5	1.25E-5
	T_{comp}	0.44	0.87	1.70	3.44	6.83	13.58	26.80
Milstein	$\hat{V}_B^{1/2}$	5.30E-4	1.84E-4	9.00E-5	6.04E-5	5.82E-5	6.75E-5	7.25E-5
	$c_{0.9}$	8.80E-5	3.06E-5	1.49E-5	1.00E-5	9.66E-6	1.12E-5	1.20E-5
	T_{comp}	0.69	1.36	2.66	5.43	10.76	21.48	42.75
3rd Order	$\hat{V}_B^{1/2}$	1.08E-3	4.96E-4	2.50E-4	1.49E-4	9.20E-5	8.46E-5	8.68E-5
	$c_{0.9}$	1.79E-4	8.23E-5	4.15E-5	2.48E-5	1.53E-5	1.40E-5	1.44E-5
	T_{comp}	0.88	1.70	3.37	6.83	13.25	26.40	53.03
	Measure Transformation + Antithetic Variate							
Euler	$\hat{V}_B^{1/2}$	7.11E-4	3.02E-4	1.64E-4	8.71E-5	6.74E-5	4.97E-5	3.24E-5
	$c_{0.9}$	1.18E-4	5.02E-5	2.72E-5	1.45E-5	1.12E-5	8.25E-6	5.37E-6
	T_{comp}	0.69	1.34	2.62	5.11	10.22	20.45	40.93
Milstein	$\hat{V}_B^{1/2}$	4.69E-4	1.67E-4	5.72E-5	2.14E-5	1.18E-5	9.51E-6	8.46E-6
	$c_{0.9}$	7.79E-5	2.77E-5	9.49E-6	3.56E-6	1.96E-6	1.58E-6	1.41E-6
	T_{comp}	1.20	2.25	4.45	8.91	17.79	35.56	70.81
3rd Order	$\hat{V}_B^{1/2}$	4.79E-4	1.99E-4	7.06E-5	2.98E-5	1.70E-5	1.29E-5	1.02E-5
	$c_{0.9}$	7.95E-5	3.31E-5	1.17E-5	4.95E-6	2.82E-6	2.15E-6	1.70E-6
	T_{comp}	1.48	2.92	5.78	11.45	22.80	45.49	90.69

Table 5

Efficiency Ratios for Simulation of 5-Year Discount Bond in CIR Model

$$N = 10,000; \; S = 100; \; B = 100$$
$$a = 0.4, \; b = 0.1, \; \sigma = 0.1, \; r_0 = 0.06$$

K	4	8	16	32	64	128	256
δ	1.25	0.625	0.3125	0.1563	0.07813	0.03906	0.01953
Crude Monte Carlo	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Antithetic Variate	5.54	5.18	5.79	5.01	5.52	6.27	7.38
Measure Transformation	3.49	7.02	15.45	24.58	45.17	60.53	69.48
Measure Transformation + Antithetic Variate	4.80	11.15	26.92	52.22	85.96	122.01	193.22

Table 6

The data in Table 6 is graphed in Figure 9. Except for very large time-steps (>1 year in our example), the combined method of measure transformation and antithetic variates is far superior to any of the traditional variance reduction methods tested in Section 4.2. Notice that the quality of the measure transformation method improves significantly as the number of time steps in each simulation path is increased. This behavior is not surprising given that the method has been designed around the continuous-time limit of the discretized processes. The tendency of the measure transform method to improve with increasing number of discretization steps is attractive since it complements the behavior of the systematic error in the SDE discretization scheme; increasing the number of discretization steps will improve both e_{sys} and $\hat{e}_{stat}$.

For completeness, efficiency results for 1-, 2-, and 10-year bonds are listed in Appendix C. For all of these bonds, the measure transformation technique yields very significant improvements over the results obtained by traditional methods. As was the case for the control variate method, the efficiency of the measure transformation technique decreases with increasing bond maturities.
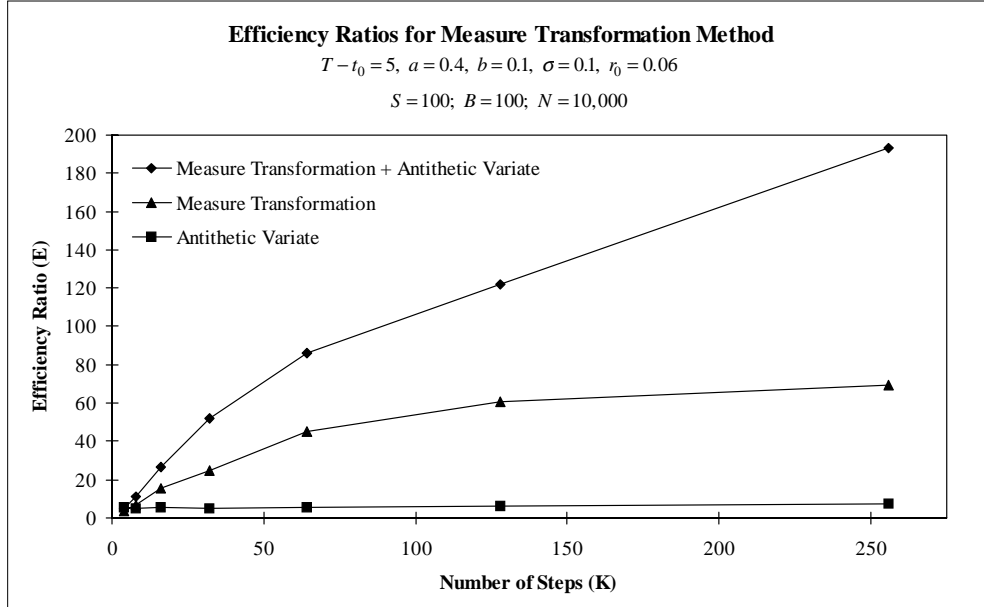
Figure 9

4.5 Quasi-Random Sequences

As mentioned in the introduction, so-called quasi-random sequences have recently been applied quite successfully to certain classes of finance problems involving explicitly solvable SDEs. In this and the following section, we will investigate whether this promising technique is equally useful for the simulation of non-solvable SDEs.

To introduce the method of quasi-random sequences, consider writing equation (28) as

$$\hat{I}_K^j = f(\tilde{z}_1^j, \tilde{z}_2^j, ..., \tilde{z}_K^j) = k(\tilde{u}_1^j, \tilde{u}_2^j, ..., \tilde{u}_K^j), \quad j = 1, 2, ..., N \tag{50}$$

where $\tilde{u}_1^j, \tilde{u}_2^j, ..., \tilde{u}_K^j$ is an independent sequence of standard *uniform* U(0,1)-variables and $k: \Re^K \to \Re$ is an (implicit) well-behaved function. The transformation from Gaussian to uniform variates can be accomplished through inversion of the cumulative Gaussian

distribution function or through the (inverse) Box-Mueller transformation. With (50), the expectation of $\hat{I}_K$ can now be written formally as an integral over the K-dimensional hypercube

$$E_{t_0}^Q[\hat{I}_K] = \int_{[0;1]^K} k(x_1, x_2, ..., x_K) dx_1 dx_2 \cdots dx_K = \int_{[0;1]^K} k(\vec{x}) d\vec{x} \qquad (51)$$

Given some deterministic or random scheme to sample N K-tuples $\vec{x}_1, \vec{x}_2, ..., \vec{x}_N$, we consider estimating (51) through

$$E_{t_0}^Q[\hat{I}_K] \approx \frac{1}{N} \sum_{j=1}^{N} k(\vec{x}_j) \qquad (52)$$

which is identical to (18), except that we have not in (52) specified which sampling scheme is used. If the sampling scheme is Monte-Carlo simulation (i.e. based on pseudo-random number generators), we know that the expected error on (52) is independent of the dimension K and proportional to $N^{-1/2}$. To improve the convergence properties of (52), several deterministic sampling algorithms have been suggested instead of Monte Carlo simulation. One class of such algorithms is based on the generation of *quasi-random* or *low discrepancy* sequences, i.e. sequences which fill out the hypercube in a cluster-free, self-avoiding way[18]. Specific algorithms for generating quasi-random numbers have been suggested by Halton (1960), Sobol (1967), and Faure (1982), among others. For a very readable introduction to Halton and Sobol sequences, see Press *et al* (1992), chapter 7. Faure sequences are discussed in Bouleau and Lepingle (1994), chapter 2C, and Joy (1994).

From the Koksma-Hlawka Inequality (see Niederreiter (1992), p. 20) one can show that under technical conditions on the function k, the Halton, Sobol, and Faure sequences generate errors which decrease with N at least as $(\ln N)^K / N$:

$$\left| \int_{[0;1]^K} k(\vec{x})d\vec{x} - \frac{1}{N} \sum_{j=1}^{N} k(\vec{x}_j) \right| \leq Var(k)O\left(\frac{(\ln N)^K}{N} \right) \qquad (53)$$

where $Var(k)$ is the (bounded) variation of k on $[0;1]^K$ in the sense of Hardy and Krause (Niederreiter (1992), p. 19). Although the $O\left((\ln N)^K / N\right)$ error bound is smaller[19] than $O(N^{-1/2})$ as $N \to \infty$, it is not small for realistic N for problems of large dimension K:

Convergence Orders for Quasi-Random Sequences and Crude Monte Carlo

K	N=10,000		N=1,000,000	
	Quasi-Random	Monte Carlo	Quasi-Random	Monte Carlo
1	9.21E-4	1.00E-2	1.38E-5	1.00E-3
5	6.63E+0	1.00E-2	5.03E-1	1.00E-3
10	4.39E+5	1.00E-2	2.53E+5	1.00E-3
50	1.64E+44	1.00E-2	1.04E+51	1.00E-3

Table 7

In practice, however, the upper bound provided by the Koksma-Hlawka inequality often turns out to significantly understate the true convergence speed of quasi-random sequences; in Brotherton-Ratcliffe (1994b), for example, Sobol sequences outperform crude Monte Carlo for option pricing applications with $K = 48$ dimensions. For problems involving numerical solution of SDEs, it is nevertheless worrying that the performance of quasi-random sequences decreases when the dimension goes up. An attempt to improve the accuracy of the systematic error e_{sys} through an increase in the number of discretization steps might thus be countered by decreased accuracy on the random error $\hat{e}_{stat}$.

As before, we can combine quasi-random sequences with other variance reduction techniques, including the control variate method and the measure transform method. It is,

however, not recommended to combine Sobol sequences with antithetic variates, as the inclusion of "mirror paths" is likely to affect the discrepancy of the sequence adversely.

4.6 Application of Sobol Sequences to the CIR Model

As experimental results by Paskov (1994) and Brotherton-Ratcliffe (1994b) indicate that Sobol sequences frequently outperform both Halton and Faure sequences, this paper will only discuss the application of Sobol sequences. For the practical generation of Sobol points, we have relied upon the highly efficient algorithm by Antonov and Saleev (1979) which is described in detail in Press *et al* (1992). The generation of the primitive polynomials needed in the Anotonov and Saleev's algorithm has been based on Knuth (1981), chapter 3.

Again using the example of a 5-year bond with $N = 10,000$ (results for other bonds are listed in Appendix C), we get the results shown in Tables 8 and 9. As a reflection of the deterministic nature of Sobol sequences, in Table 8 we have replaced standard deviation $\sqrt{\hat{V}_B}$ with the root-mean-square error (relative to the sample mean) RMS_B; both quantities are, however, calculated identically. Notice that we cannot generate any confidence intervals when Sobol sequences are used.

Simulation Results for CIR Model
5-Year Discount Bond, Theoretical Value = 0.6642841

$N = 10,000; \ S = 100; \ B = 100$

$a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$

	K	4	8	16	32	64	128	256
	δ	1.25	0.625	0.3125	0.1563	0.07813	0.03906	0.01953
		Sobol Sequence						
Euler	RMS_B	1.42E-3	1.78E-3	1.85E-3	2.44E-3	3.55E-3	3.86E-3	3.87E-3
	T_{comp}	0.34	0.59	1.14	2.2	4.32	8.32	16.76
Milstein	RMS_B	1.28E-3	1.61E-3	1.81E-3	2.41E-3	3.53E-3	3.82E-3	3.87E-3
	T_{comp}	0.50	0.96	1.80	3.54	7.01	13.94	27.88
3rd Order	RMS_B	1.41E-3	1.78E-3	2.08E-3	2.88E-3	4.17E-3	4.39E-3	4.17E-3
	T_{comp}	0.63	1.13	2.19	4.27	8.47	16.81	33.09
		Sobol Sequence + **Control Variate**						
Euler	RMS_B	6.76E-4	1.07E-3	1.20E-3	1.41E-3	1.31E-3	1.43E-3	1.67E-3
	T_{comp}	0.43	0.71	1.21	2.43	4.78	9.49	18.66
Milstein	RMS_B	5.78E-4	8.98E-4	1.08E-3	1.31E-3	1.25E-3	1.40E-3	1.65E-3
	T_{comp}	0.62	1.12	2.04	3.75	7.45	14.49	28.91
3rd Order	RMS_B	7.55E-4	9.65E-4	1.21E-3	1.52E-3	1.50E-3	1.70E-3	2.03E-3
	T_{comp}	0.69	1.24	2.54	4.63	9.36	18.47	37.08
		Sobol Sequence + **Measure Transformation**						
Euler	RMS_B	6.87E-4	3.03E-4	2.83E-4	1.92E-4	1.10E-4	5.03E-5	4.31E-5
	T_{comp}	0.46	0.82	1.56	2.98	5.80	11.62	22.36
Milstein	RMS_B	2.46E-4	8.55E-5	2.84E-5	2.19E-5	2.75E-5	3.62E-5	3.83E-5
	T_{comp}	0.66	1.31	2.44	4.68	9.30	18.88	36.44
3rd Order	RMS_B	3.27E-4	1.81E-4	9.73E-5	8.67E-5	6.42E-5	4.78E-5	4.74E-5
	T_{comp}	0.85	1.64	3.32	5.99	11.87	24.52	46.63

Table 8

IV.44

Efficiency Ratios for Simulation of 5-Year Discount Bond in CIR Model

$$N = 10,000;\ S = 100;\ B = 100$$
$$a = 0.4,\ b = 0.1,\ \sigma = 0.1,\ r_0 = 0.06$$

K	4	8	16	32	64	128	256
δ	1.25	0.625	0.3125	0.1563	0.07813	0.03906	0.01953
Crude Monte Carlo	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Antithetic Variate	5.54	5.18	5.79	5.01	5.52	6.27	7.38
Sobol Sequence	4.44	3.50	3.42	2.38	1.89	1.89	2.16
Sobol Sequence + Control Variate	7.68	5.39	4.99	4.02	4.78	4.64	4.42
Sobol Sequence + Measure Transformation	*10.81*	*22.57*	*33.68*	44.93	77.23	120.81	147.07
Antitethetic Variate + Measure Transformation	4.80	11.15	26.92	*52.22*	*85.96*	*122.01*	*193.22*

Note: The highest efficiency ratio for each step-size has been highlighted

Table 9

Although in our example the deterioration of the Sobol sequence technique certainly is not as rapid as might be expected from the Koksma-Hlawka inequality, Table 9 show that performance of the Sobol sequence still drops by roughly a factor 2 when the dimension of the bond pricing problem is increased from 4 to 256. Further, despite being about twice faster than the simple antithetic variate technique, the efficiency of the Sobol sequence is generally lower than that of antithetic variates, even after combining the Sobol sequence with a Vasicek control variate. Sobol sequences, however, appear to interact quite well with the measure transformation method: for $K \leq 16$, the combined method of Sobol sequence and measure transformation edges out the combined method of antithetic variate and measure transformation. Unfortunately, the good properties of Sobol sequences for problems with low values of K do not benefit us much in the applications considered in this paper: when K is small, the systematic error in the discretization scheme will frequently overshadow the random error associated with the variance of the bond price estimator (see for example Figure 3). In general, we prefer to use the combined method of antithetic

variates and measure transformation as this method produces the best results in problems where low variance matters the most, namely for problems involving high values of K.

5. Combining High-Order Simulation Schemes with Variance Reduction Methods: Experimental Results

In this section we will briefly show some further simulation results obtained by combining the discretization schemes in Section 3 with the variance reduction techniques from Section 4. In particular, we wish to measure the effect of the joint variance reduction method of antithetic variates and measure transformation on the convergence results for the 2-year discount bond in Section 3.2. For the sake of brevity, we only show graphs of the simulation results; the raw simulation results are tabulated in Appendix E.

In Figures 10 and 11 below, we show the convergence profiles for $N = 1,000,000$ simulations of the Euler, Milstein, and third order schemes in combination with antithetic variates and measure transformation. Compared to Figures 1 and 2 in Section 3.2, it is obvious that adding the variance reduction technique has increased the maximum attainable accuracy considerably: whereas the lowest possible worst-case error in Section 3.2 was around $|\hat{e}_{wc}| \approx 2^{-14} = 0.61$ bp, the third-order scheme now reaches an accuracy level of $|\hat{e}_{wc}| \approx 2^{-25} = 3.0 \cdot 10^{-4}$ bp. The effectiveness of the variance reduction method also allows us to fully capitalize on the advantages of the high-order discretization schemes: to reach the accuracy of the 256-step (variance reduced) Euler scheme, the third-order scheme and Milstein schemes need only 8 and 16 steps, respectively. As Figure 12 shows, the advantages of the higher-order methods hold even after accounting for their higher computational overhead.
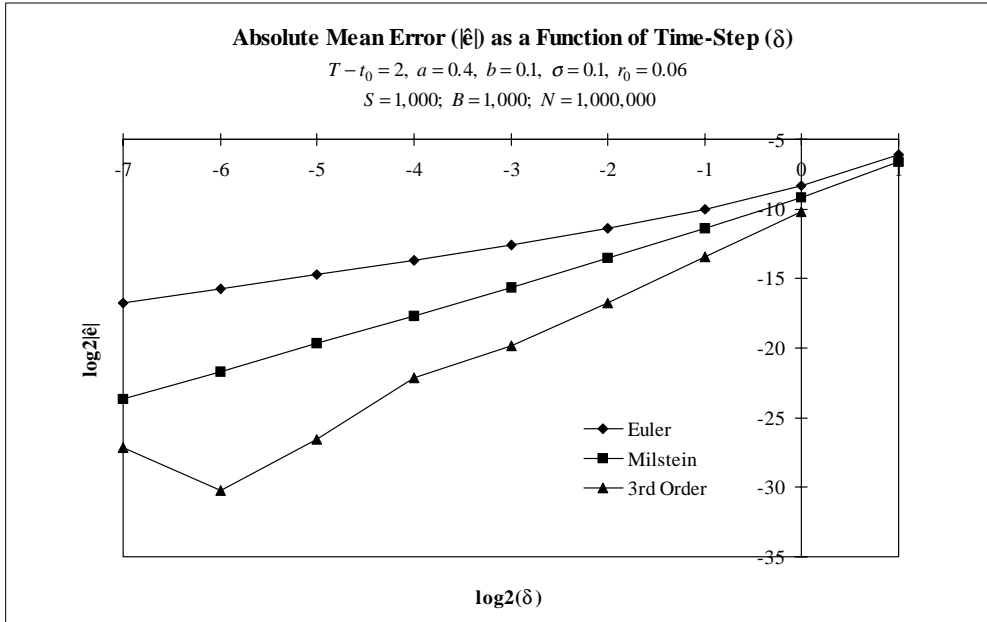
Figure 10

Figure 11

IV.47
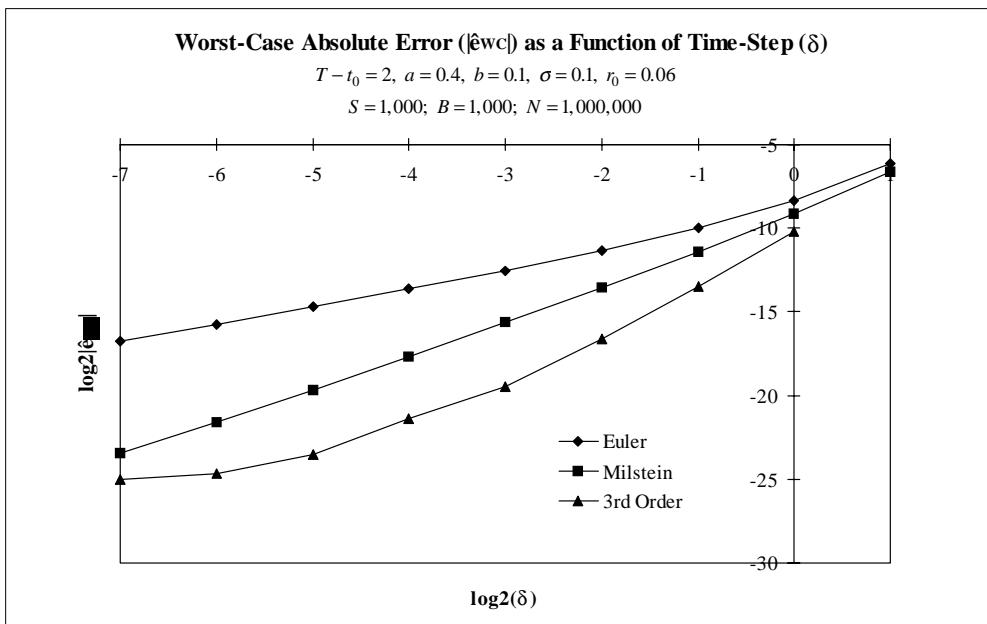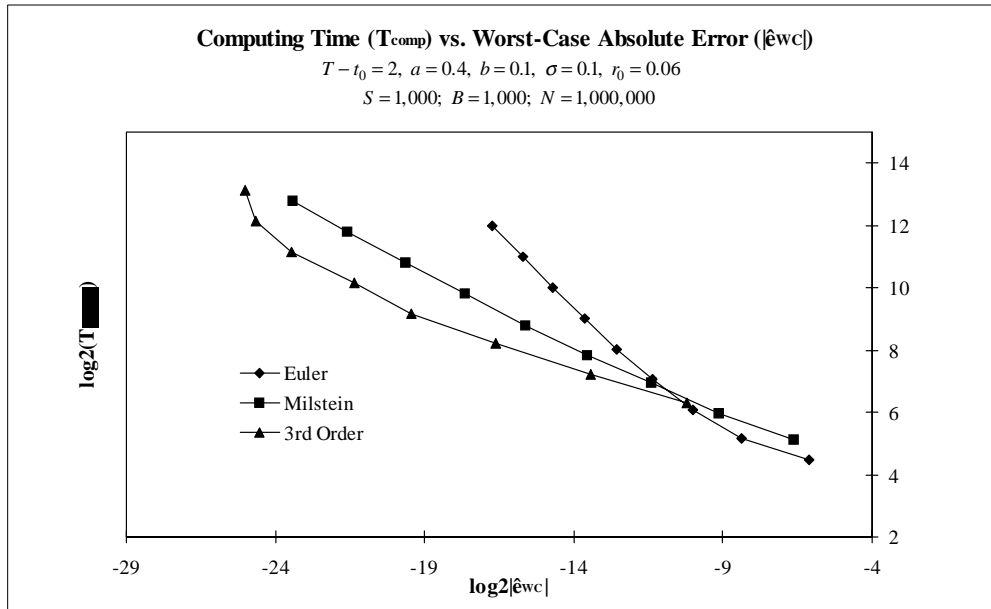
Figure 12

In the examples above, the large number of simulations ($N = 1,000,000$) allowed us to reach very high levels of accuracy on the bond price estimates. In many practical applications, however, speed is more important than accuracy which dictates the usage of significantly fewer simulation paths. In Figure 13 below, we have used results from simulations of $N = 1,000,000$, $N = 10,000$, and $N = 100$ paths to draw pieces of the (worst case) efficiency profiles of the third-order discretization method with antithetic variate and measure transform variance reduction (the "advanced approach"). In the graph, we have also included the efficiency profile ($N = 1,000,000$) of the Euler scheme simulated with crude Monte Carlo (the "naive approach").

Computing Time (T$_{comp}$) vs. Worst-Case Absolute Error (|ê$_{wc}$|)

$T - t_0 = 2, \ a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$
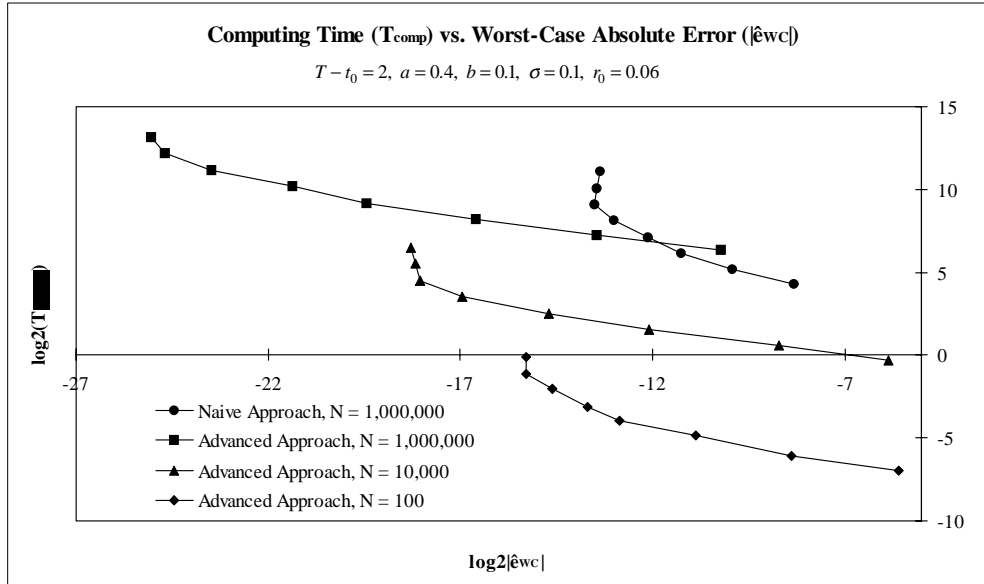
Figure 13

From the figure, we can for example conclude that the accuracy of 100 simulations of a 64-step third-order scheme combined with antithetic variates and measure transformation exceeds the highest attainable accuracy of 1,000,000 crude Monte Carlo simulations of the Euler scheme. 1 bp (worst-case) accuracy is reached by the variance reduced third-order scheme in 32 steps and 100 simulations, requiring as little computation time as 0.116 seconds. The naive Euler scheme, on the other hand, requires 64 steps and 1,000,000 simulations, for a computation time of 537.30 seconds -- more than 4,500 times slower than the variance reduced third-order scheme.

6. Simulation of Options.

So far, the focus of this paper has been the simplest and most fundamental of all interest rate contingent claims, namely the zero-coupon bond with a certain $1 payout at maturity. The simulation techniques discussed, however, are sufficiently general to allow for the

pricing of instruments with significantly more complicated dependence of the interest rate path $\{r_t, t \in [t_0, \tau]\}$. We will leave a systematic investigation of the simulation of general interest rate contingent claims to future research and here merely consider a few selected issues.

To focus on a simple example, consider a T_c-maturity European call c, struck at K, on a zero-coupon bond maturing at T ($t_0 \le T_c \le T$). Following Harrison and Kreps (1979) the price of the call is

$$
\begin{aligned}
c(t_0, T_c, T) &= E_{t_0}^Q \left[\exp(-\int_{t_0}^{T_c} r_s ds) MAX\{P(T_c, T) - K, 0\} \right] \\
&= E_{t_0}^Q \left[I_{T_c} MAX \left\{ E_{T_c}^Q \left[\exp(-\int_{T_c}^{T} r_s ds) \right] - K, 0 \right\} \right] \\
&= E_{t_0}^Q \left[I_{T_c} MAX \left\{ h_T(r_{T_c}, T_c) - K, 0 \right\} \right] \\
&= E_{t_0}^Q [F(r_{T_c}, I_{T_c}; T_c, T)]
\end{aligned}
\qquad (54)
$$

where the payout function $F(\cdot)$ is known explicitly only in cases where the bond pricing function $h_T(r_t, t) = P(t, T)$ can be written in closed form (e.g. the Vasicek model). In most cases, however, we need a *local simulation* from T_c to T to estimate the bond $P(T_c, T)$ -- a problem which has been extensively discussed in the previous 5 sections of this paper.

Although the expectation in (54) involves a function of r and I (and not just I itself), it can be shown that the weak convergence orders for the simulation schemes in section 3 hold unchanged, provided that the payout function F is sufficiently smooth and satisfies polynomial growth conditions[20]. Similarly, the variance reduction techniques of section 4 can all be applied virtually unchanged, although some care must be taken in the choice of probability measure in the measure transform technique. To expand on this latter point, consider the call pricing equation under a transformed measure Z

IV.50

$$c(t_0, T_c, T) = E_{t_0}^Z\left[\xi_{T_c} F(r_{T_c}, I_{T_c}; T_c, T)\right] = E_{t_0}^Z\left[\xi_{T_c} I_{T_c} MAX\{P(T_c, T) - K, 0\}\right]$$

(55)

In choosing a likelihood ratio process, we have several alternatives, including i) minimizing the variance of $\xi_{T_c} F(r_{T_c}, I_{T_c}; T_c, T)$; ii) minimizing the variance of $\xi_{T_c} I_{T_c}$; iii) minimizing the variance of $\xi_{T_c} P(T_c, T)$; and iv) minimizing the variance of $\xi_{T_c} I_{T_c} P(T_c, T)$. Notice, that one should *not* attempt to consider a term of the form $\xi_{T_c}[P(T_c, T) - K]$ due to potential singularities in (44). Since i) aims at reducing the variance of the total option price, it is obviously theoretically preferable to ii), iii) and iv) which target the variance of various combinations of the discount bond $P(t, T_c)$ and the payout bond $P(T_c, T)$. Similar to the case of the bond pricing problem, the necessary approximation (see (44)) of $E_t^Q[F(r_{T_c}, I_{T_c}; T_c, T)]$ can be based on the known closed-form pricing formula for European discount bond options in the Vasicek model (Vasicek (1977)). As the derivatives-based high-order simulation schemes become quite complex, it is in general recommended to keep the guess for $E_t^Q[F(r_{T_c}, I_{T_c}; T_c, T)]$ as simple as possible and use Richardson extrapolation wherever feasible.

7. Conclusions

In this paper we have considered general methods to improve speed and accuracy of simulation models based on one-factor SDEs in the instantaneous risk free interest rate. As discussed, two types of simulation errors must be dealt with: i) systematic deviations from the continuous-time limit introduced by the discretization of the SDE, and ii) random errors introduced by the stochastic nature of the bond price estimator. We have shown how second-order Ito-Taylor expansions, either directly or through extrapolation methods, can form the basis of high-order schemes which significantly improve the convergence properties of the systematic error. In particular, we have shown that the second-order extended Milstein scheme and the Richardson extrapolated third-order scheme both

outperform the "naive" Euler scheme, even after adjustments for increased computation time. As we have seen in Section 3.2, the improvements obtained by using high-order discretization schemes can, however, easily be overshadowed the random errors of the bond price estimator. In practice, high-order discretization methods must thus generally be supplemented by variance reduction techniques. In this paper, we have outlined and tested four methods: i) antithetic variates, ii) control variates, iii) measure transformation, and iv) Sobol quasi-random sequences. For the tested example of the Cox, Ingersoll, and Ross (CIR) (1985) model, the combined method of antithetic variates and measure transformation (based on closed-form bond prices in the Vasicek (1978) model) generates the best results. We have shown that using a third-order scheme combined with antithetic variates and the measure transformation method can, in some cases, improve the accuracy and speed of bond price estimates by factors in excess of 10^3. We should, of course, point out that these results are unique to the CIR model and might not be as significant for other models. In particular, when applying the measure transformation method, we have benefited from the structural similarity of the Vasicek and the CIR models. Nevertheless, the results obtained in this paper are very encouraging and hopefully will stimulate further empirical research using alternative short rate models.

Besides investigating the application of the techniques discussed in this paper to alternative models, several other interesting areas of research remain open. An obvious topic is the extension of the results in this paper to more complicated interest rate derivatives, particularly those involving path-dependency. Such problems might involve several "local" bond price simulations along each interest rate path and raises some interesting questions of how to combine "global" and "local" simulation schemes to achieve maximum speed and accuracy. Another line of research involves testing the quality of the many traditional Monte Carlo variance reduction methods we did not cover in this paper, including importance sampling, stratified sampling, and adaptive Monte Carlo (see for example Press *et al* (1992) for a good discussion of these methods). A final open topic is the application of

implicit and explicit discretization schemes to problems where the interest rate process parameters are not given directly, but must be extracted from market data.

References

Andersen, L., "Monte Carlo Simulation of Barrier and Lookback Options with Continuous or High-Frequency Monitoring of the Underlying Asset," Working Paper, General Re Financial Products, February 1996

Andersen, L. and R. Brotherton-Ratcliffe, "Using Monte Carlo Simulation to Value Barrier and Lookback Options", Working Paper, General Re Financial Products, March 1996

Andersen, L. and R. Brotherton-Ratcliffe, "The Equity Option Volatility Smile: An Implicit Finite Difference Approach," Working Paper, General Re Financial Products, December 1995

Antonov, I. and V. Saleev, *USSR Computational Mathematics and Mathematical Physics*, 1, 1979, 252-256

Apabhai, M., K. Choe, F. Khennach and P. Wilmott, "One-Factor Interest Rate Modeling -- Empirical Evidence," Proceedings of *RISK Magazine* Conference: Yield Curve Analysis, November 1994

Arnold, L., *Stochastic Differential Equations: Theory and Practice*, Krieger Publishing Company, 1992

Bally, V. and D. Talay, "The Law of the Euler Scheme for Stochastic Differential Equations: I. Convergence Rate of the Distribution Function," Working Paper, University of Paris, 1995

Black, F., E. Derman and W. Toy, "A One-Factor Model of Interest Rates and Its Application," *Financial Analysts Journal*, January-February 1990, 33-39

Bouleau, N. and D. Lepingle, *Numerical Methods for Stochastic Processes*, John Wiley & Sons, 1994

Boyle, P., "Options: A Monte Carlo Approach," *Journal of Financial Economics*, 4 (1977), 323-38

Boyle, P., "A Lattice Framework for Option Pricing with two State Variables," *Journal of Financial and Quantitative Analysis*, 23 (March 1988), 1-12

Brennan, M. and E. Schwartz, "Finite Difference Methods and Jump Processes Arising in the Pricing of Contingent Claims: A Synthesis," *Journal of Financial and Quantitative Analysis*, 13 (September 1978), 462-74

Brennan, M. and E. Schwartz, "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency," *Journal of Financial and Quantitative Analysis*, 17, 3 (September 1982), 301-29

Broadie, M. and P. Glasserman, "Pricing American-Style Securities Using Simulation," Working Paper, Graduate School of Business, Columbia University, 1994

Brotherton-Ratcliffe, R., "Monte Carlo Motoring," *RISK Magazine*, December 1994a, 53-57

Brotherton-Ratcliffe, R. "Effective Variance Reduction Techniques for Monte-Carlo Simulations," Proceedings of *RISK Magazine* Seminar: Advanced Mathematics for Derivatives, January 1995

Chan, K., G. Karolyi, F. Longstaff and A. Sanders, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47 (July 1992), 1209-1227

Clewlow, L. and A. Carverhill, "On the Simulation of Contingent Claims," *Journal of Derivatives*, 2, 2 (Winter 1994), 66-75

Cox, J., "Notes on Option Pricing I: Constant Elasticity of Variance Diffusions," Working Paper, Stanford University, 1975

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

Das, S., "Mean Rate Shifts and Alternative Models of the Interest Rate: Theory and Evidence," Working Paper, Department of Finance, New York University, 1993

Derman, E. and I. Kani, "Riding on a Smile," *RISK Magazine*, February 1994, 32-39

Devroye, L., *Non-Uniform Random Variate Generation*, Springer Verlag, 1986

Duffie, D., *Dynamic Asset Pricing Theory*, Princeton University Press, 1992

Duffie, D. and R. Kan, "Multi-Factor Term Structure Models," *Philosophical Transactions of the Royal Society*, 347, 1684 (June 1994), 577-586

Dupire, B., "Model Art," *RISK Magazine*, September 1993, 118-24

Dupire, B., "Pricing with a Smile," *RISK Magazine*, January 1994, 18-20

Faure, H., "Discrépance de Suites associeés à un Système de Numération (en Dimension s)," *Acta Arithmetica*, XLI (1982), 337-351

Halton, J., "On the Efficiency of certain Quasi-Random Sequences of Points in evaluating Multi-Dimensional Integrals," *Numerische Mathematik*, 2 (1960), 84-90

Hammersley, J. and D. Handscomb, *Monte Carlo Methods*, John Wiley and Sons, 1964

Harrison, M. and D. Kreps, "Martingales and Arbitrage in Multi-Period Security Markets," *Journal of Economic Theory*, 20 (1979), 381-408

Harrison, M. and S. Pliska, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications*, 11 (1981), 215-260

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

Ho, T. and S. Lee, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance*, 41 (December 1986), 1011-29

Hofmann, N., E. Platen and M. Schweizer, "Option Pricing under Incompleteness and Stochastic Volatility," *Mathematical Finance*, 2, 3 (1992), 153-187

Hull, J. and A. White, "Valuing Derivative Securities using the Explicit Finite Difference Method," *Journal of Financial and Quantitative Analysis*, 25 (March 1990a), 87-100

Hull, J. and A. White, "Pricing Interest Rate Derivative Securities," *Review of Financial Studies*, 3, 4 (1990b), 87-100

Hull, J. and A. White, "Efficient Procedures for Valuing European and American Path-Dependent Options", *Journal of Derivatives*, 1 (1993), 21-31

Hull, J and A. White, "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models," *Journal of Derivatives*, 2, 1 (Fall 1994a), 7-16

Hull, J. and A. White, "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models," *Journal of Derivatives*, 2, 2 (Winter 1994b), 37-48

Jamshidian, F., "Forward Induction and Construction of Yield Curve Diffusion Models," *Journal of Fixed Income*, 1, 1 (June 1991), 62-74

Jensen, B. and J. Nielsen, "The Structure of Binomial Lattice Models for Bonds," Working Paper, Department of Finance, Copenhagen School of Economics and Business Administration, 1991

Joergensen, P., "American Option Pricing," Ph.D. Thesis, Department of Finance, Aarhus School of Business, November 1994

Joy, C., "Quasi Monte Carlo Techniques in Option Pricing," Working Paper, Enron Corporation, 1994

Karatzas, I. and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer Verlag, 1991

Kloeden, P. and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer Verlag, 1992

Knuth, D., *The Art of Computer Programming Vol. 2: Semi-Numerical Algorithms*, Addison Wesley, 1981

Milstein, G., "Approximate Integration of Stochastic Differential Equations," *Theory of Probability and Its Applications*, 19 (1974), 557-62

Milstein, G., "A Method of Second-Order Integration of Stochastic Differential Equations," *Theory of Probability and Its Applications*, 23 (1978), 396-401

Milstein, G., *The Numerical Integration of Stochastic Differential Equations*, Urals University Press, 1988

Newton, N., "Variance Reduction for Simulated Diffusions," *SIAM Journal of Applied Mathematics*, 54 (December 1994), p. 1780-1805

Niederreiter, H., *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, 1992

Paskov, S., "Computing High Dimensional Integrals with Applications to Finance," Working Paper, Department of Computer Science, Columbia University, October 1994

Pearson, N. and T. Sun, "An Empirical Examination of the Cox-Ingersoll-Ross Model of the Term Structure of Interest Rates using the Method of Maximum Likelihood," Working Paper, Graduate School of Business, Columbia University, 1990

Platen, E. and M. Schweizer, "On Smile and Skewness," Working Paper, Centre for Mathematics and Its Applications, School of Mathematical Sciences, Australian National University, August 1994

Pliska, S., "A Mathematical Analysis of Interest Rate Derivatives Pricing Models," Proceedings of *RISK Magazine* Seminar: Advanced Mathematics for Derivatives, September 1994

Press, W., S. Teukolsky, W. Vetterling and B. Flannery, *Numerical Recipes in C*, Cambridge University Press, 1992

Rubinstein, M., "Implied Binomial Trees," *Journal of Finance*, 49 (July 1994), 771-818

Sobol, I.. "On the Distribution of Points in a Cube and the approximate Evaluation of Integrals," *USSR Computational Mathematics and Mathematical Physics*, 7 (1967), 86-112

Talay, D. and Tubaro, L., "Expansions of the Global Error for Numerical Schemes solving Stochastic Differential Equations," *Stochastic Analysis and Applications*, 8, 4 (1990), 483-509

Tilley, J., "Valuing American Options in a Path Simulation Model," *Transactions of the Society of Actuaries*, XLV (1993), 83-104

Vasicek, O., "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5 (1977), 177-88

Wilmott, P., J. Dewynne and S. Howison, *Option Pricing*, Oxford Financial Press, 1993

Appendix A

Additional Crude Monte Carlo Simulation Results for CIR Model

1-, 5-, and 10-year Discount Bonds

Crude Monte Carlo Simulation Results for CIR Model
1-Year Discount Bond, Theoretical Value = 0.9352398

$$N = 1,000,000; \ S = 1,000; \ B = 1,000$$
$$a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$$

	K	1	2	4	8	16	32	64	128	256
	δ	2	1	0.5	0.25	0.125	0.0625	0.03125	0.01563	0.00781
	$\log_2(\delta)$	1.00	0.00	-1.00	-2.00	-3.00	-4.00	-5.00	-6.00	-7.00
Euler	$\hat{e}$	4.76E-3	1.79E-3	7.87E-4	3.57E-4	1.82E-4	8.68E-5	4.54E-5	3.37E-5	2.93E-5
	$\log_2|\hat{e}|$	-7.71	-9.13	-10.31	-11.45	-12.42	-13.49	-14.43	-14.86	-15.06
	$\sqrt{\hat{V}_B}$	0	2.73E-4	3.18E-4	3.46E-4	3.63E-4	3.60E-4	3.58E-4	3.64E-4	3.61E-4
	$c_{0.9}$	0	1.43E-5	1.66E-5	1.81E-5	1.89E-5	1.88E-5	1.87E-5	1.90E-5	1.88E-5
	$|\hat{e}_{wc}|$	4.76E-3	1.80E-3	8.04E-4	3.75E-4	2.01E-4	1.06E-4	6.41E-5	5.27E-5	4.81E-5
	$\log_2|\hat{e}_{wc}|$	-7.71	-9.11	-10.28	-11.38	-12.28	-13.21	-13.93	-14.21	-14.34
	T_{comp}	10.90	19.33	35.42	68.87	137.11	274.12	537.30	1065.74	2130.95
	$\log_2(T_{comp})$	3.45	4.27	5.15	6.11	7.10	8.10	9.07	10.06	11.06
Mil-stein	$\hat{e}$	-1.42E-3	-2.95E-4	-5.97E-5	-2.27E-5	2.70E-6	-1.47E-7	2.43E-6	1.25E-5	1.88E-5
	$\log_2|\hat{e}|$	-9.46	-11.73	-14.03	-15.43	-18.50	-22.70	-18.65	-16.28	-15.70
	$\sqrt{\hat{V}_B}$	-1.42E-3	3.78E-4	3.67E-4	3.70E-4	3.75E-4	3.66E-4	3.61E-4	3.65E-4	3.62E-4
	$c_{0.9}$	-7.41E-5	1.97E-5	1.91E-5	1.93E-5	1.95E-5	1.91E-5	1.88E-5	1.91E-5	1.89E-5
	$|\hat{e}_{wc}|$	1.49E-3	3.15E-4	7.89E-5	4.20E-5	2.22E-5	1.92E-5	2.13E-5	3.16E-5	3.76E-5
	$\log_2|\hat{e}_{wc}|$	-9.39	-11.63	-13.63	-14.54	-15.46	-15.67	-15.52	-14.95	-14.70
	T_{comp}	14.15	27.62	56.10	103.90	206.21	404.36	822.91	1626.24	3260.38
	$\log_2(T_{comp})$	3.82	4.79	5.81	6.70	7.69	8.66	9.68	10.67	11.67
3rd Order	$\hat{e}$	NA	8.44E-5	2.37E-5	-4.07E-6	9.79E-6	2.98E-7	1.72E-6	1.25E-5	2.08E-5
	$\log_2|\hat{e}|$	NA	-13.53	-15.37	-17.91	-16.64	-21.68	-19.15	-16.29	-15.55
	$\sqrt{\hat{V}_B}$	NA	4.77E-4	4.29E-4	4.23E-4	4.21E-4	4.13E-4	4.02E-4	4.06E-4	4.07E-4
	$c_{0.9}$	NA	2.49E-5	2.24E-5	2.20E-5	2.20E-5	2.15E-5	2.10E-5	2.12E-5	2.13E-5
	$|\hat{e}_{wc}|$	NA	1.09E-4	4.60E-5	2.61E-5	3.18E-5	2.18E-5	2.27E-5	3.36E-5	4.21E-5
	$\log_2|\hat{e}_{wc}|$	NA	-13.16	-14.41	-15.22	-14.94	-15.48	-15.43	-14.86	-14.54
	T_{comp}	NA	34.11	64.13	123.22	243.48	478.95	951.60	1887.33	3842.11
	$\log_2(T_{comp})$	NA	5.09	6.00	6.94	7.93	8.90	9.89	10.88	11.91

IV.61

Crude Monte Carlo Simulation Results for CIR Model
5-Year Discount Bond, Theoretical Value = 0.6642841

$$N = 1{,}000{,}000; \ S = 1{,}000; \ B = 1{,}000$$
$$a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$$

	K	1	2	4	8	16	32	64	128	256
	δ	2	1	0.5	0.25	0.125	0.0625	0.03125	0.01563	0.00781
	$\log_2(\delta)$	1.00	0.00	-1.00	-2.00	-3.00	-4.00	-5.00	-6.00	-7.00
Euler	$\hat{e}$	3.57E-2	-2.67E-2	-1.17E-2	-5.68E-3	-2.74E-3	-1.38E-3	-6.73E-4	-2.78E-4	-6.93E-5
	$\log_2|\hat{e}|$	-4.81	-5.23	-6.42	-7.46	-8.51	-9.50	-10.54	-11.81	-13.82
	$\sqrt{\hat{V_B}}$	0	2.68E-3	2.12E-3	2.03E-3	2.01E-3	1.95E-3	1.92E-3	1.92E-3	1.90E-3
	$c_{0.9}$	0	1.40E-4	1.11E-4	1.06E-4	1.05E-4	1.01E-4	9.99E-5	1.00E-4	9.94E-5
	$|\hat{e}_{wc}|$	3.57E-2	2.68E-2	1.18E-2	5.78E-3	2.85E-3	1.48E-3	7.73E-4	3.78E-4	1.69E-4
	$\log_2|\hat{e}_{wc}|$	-4.81	-5.22	-6.41	-7.43	-8.46	-9.40	-10.34	-11.37	-12.53
	T_{comp}	10.90	19.33	35.42	68.87	137.11	274.12	537.30	1065.74	2130.95
	$\log_2(T_{comp})$	3.45	4.27	5.15	6.11	7.10	8.10	9.07	10.06	11.06
Mil-stein	$\hat{e}$	-1.19E-1	-1.67E-2	-2.34E-3	-4.58E-4	-5.78E-5	-2.08E-5	8.09E-6	6.51E-5	1.03E-4
	$\log_2|\hat{e}|$	-3.07	-5.90	-8.74	-11.09	-14.08	-15.56	-16.92	-13.91	-13.25
	$\sqrt{\hat{V_B}}$	-1.19E-1	2.15E-3	1.96E-3	1.95E-3	1.97E-3	1.92E-3	1.91E-3	1.92E-3	1.90E-3
	$c_{0.9}$	-6.21E-3	1.12E-4	1.02E-4	1.02E-4	1.03E-4	1.00E-4	9.95E-5	1.00E-4	9.93E-5
	$|\hat{e}_{wc}|$	1.25E-1	1.68E-2	2.44E-3	5.60E-4	1.61E-4	1.21E-4	1.08E-4	1.65E-4	2.02E-4
	$\log_2|\hat{e}_{wc}|$	-3.00	-5.89	-8.68	-10.80	-12.60	-13.01	-13.18	-12.56	-12.27
	T_{comp}	14.15	27.62	56.10	103.90	206.21	404.36	822.91	1626.24	3260.38
	$\log_2(T_{comp})$	3.82	4.79	5.81	6.70	7.69	8.66	9.68	10.67	11.67
3rd Order	$\hat{e}$	NA	1.75E-2	2.48E-3	1.91E-4	6.83E-5	-3.62E-6	7.15E-6	6.62E-5	1.17E-4
	$\log_2|\hat{e}|$	NA	-5.84	-8.65	-12.35	-13.84	-18.08	-17.09	-13.88	-13.06
	$\sqrt{\hat{V_B}}$	NA	2.28E-3	2.24E-3	2.21E-3	2.21E-3	2.16E-3	2.13E-3	2.13E-3	2.14E-3
	$c_{0.9}$	NA	1.19E-4	1.17E-4	1.15E-4	1.15E-4	1.13E-4	1.11E-4	1.11E-4	1.12E-4
	$|\hat{e}_{wc}|$	NA	1.76E-2	2.60E-3	3.07E-4	1.84E-4	1.16E-4	1.18E-4	1.77E-4	2.28E-4
	$\log_2|\hat{e}_{wc}|$	NA	-5.83	-8.59	-11.67	-12.41	-13.07	-13.04	-12.46	-12.10
	T_{comp}	NA	34.11	64.13	123.22	243.48	478.95	951.60	1887.33	3842.11
	$\log_2(T_{comp})$	NA	5.09	6.00	6.94	7.93	8.90	9.89	10.88	11.91

Crude Monte Carlo Simulation Results for CIR Model
10-Year Discount Bond, Theoretical Value = 0.4125989

$$N = 1,000,000; \ S = 1,000; \ B = 1,000$$
$$a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$$

	K	1	2	4	8	16	32	64	128	256
	δ	2	1	0.5	0.25	0.125	0.0625	0.03125	0.01563	0.00781
	$\log_2(\delta)$	1.00	0.00	-1.00	-2.00	-3.00	-4.00	-5.00	-6.00	-7.00
Euler	$\hat{e}$	-1.26E-2	-2.02E-1	-5.40E-2	-2.35E-2	-1.10E-2	-5.40E-3	-2.64E-3	-1.24E-3	-5.34E-4
	$\log_2|\hat{e}|$	-6.31	-2.30	-4.21	-5.41	-6.50	-7.53	-8.56	-9.65	-10.87
	$\sqrt{\hat{V}_B}$	0	6.24E-3	2.91E-3	2.53E-3	2.43E-3	2.33E-3	2.28E-3	2.28E-3	2.25E-3
	$c_{0.9}$	0	3.25E-4	1.52E-4	1.32E-4	1.27E-4	1.22E-4	1.19E-4	1.19E-4	1.18E-4
	$|\hat{e}_{wc}|$	1.26E-2	2.03E-1	5.41E-2	2.36E-2	1.12E-2	5.52E-3	2.76E-3	1.36E-3	6.52E-4
	$\log_2|\hat{e}_{wc}|$	-6.31	-2.30	-4.21	-5.41	-6.49	-7.50	-8.50	-9.52	-10.58
	T_{comp}	10.90	19.33	35.42	68.87	137.11	274.12	537.30	1065.74	2130.95
	$\log_2(T_{comp})$	3.45	4.27	5.15	6.11	7.10	8.10	9.07	10.06	11.06
Mil-stein	$\hat{e}$	-6.32E-1	-1.16E-1	-7.72E-3	-4.39E-4	5.62E-5	1.50E-5	2.73E-5	8.62E-5	1.29E-4
	$\log_2|\hat{e}|$	-0.66	-3.11	-7.02	-11.15	-14.12	-16.02	-15.16	-13.50	-12.92
	$\sqrt{\hat{V}_B}$	-6.32E-1	3.10E-3	2.31E-3	2.29E-3	2.33E-3	2.28E-3	2.26E-3	2.27E-3	2.25E-3
	$c_{0.9}$	-3.30E-2	1.62E-4	1.20E-4	1.19E-4	1.22E-4	1.19E-4	1.18E-4	1.18E-4	1.17E-4
	$|\hat{e}_{wc}|$	6.65E-1	1.16E-1	7.84E-3	5.59E-4	1.78E-4	1.34E-4	1.45E-4	2.05E-4	2.47E-4
	$\log_2|\hat{e}_{wc}|$	-0.59	-3.10	-6.99	-10.81	-12.46	-12.87	-12.75	-12.26	-11.99
	T_{comp}	14.15	27.62	56.10	103.90	206.21	404.36	822.91	1626.24	3260.38
	$\log_2(T_{comp})$	3.82	4.79	5.81	6.70	7.69	8.66	9.68	10.67	11.67
3rd Order	$\hat{e}$	NA	5.59E-2	2.85E-2	2.00E-3	2.13E-4	4.30E-6	1.77E-5	8.72E-5	1.47E-4
	$\log_2|\hat{e}|$	NA	-4.16	-5.13	-8.96	-12.20	-17.83	-15.79	-13.49	-12.73
	$\sqrt{\hat{V}_B}$	NA	3.22E-3	2.50E-3	2.59E-3	2.61E-3	2.55E-3	2.53E-3	2.52E-3	2.52E-3
	$c_{0.9}$	NA	1.68E-4	1.31E-4	1.35E-4	1.36E-4	1.33E-4	1.32E-4	1.31E-4	1.32E-4
	$|\hat{e}_{wc}|$	NA	5.60E-2	2.86E-2	2.14E-3	3.49E-4	1.37E-4	1.50E-4	2.19E-4	2.79E-4
	$\log_2|\hat{e}_{wc}|$	NA	-4.16	-5.13	-8.87	-11.48	-12.83	-12.71	-12.16	-11.81
	T_{comp}	NA	34.11	64.13	123.22	243.48	478.95	951.60	1887.33	3842.11
	$\log_2(T_{comp})$	NA	5.09	6.00	6.94	7.93	8.90	9.89	10.88	11.91

Appendix B

Discretization Schemes for the Vasicek Model

Euler Scheme:

$$\hat{I}_{i+1} = \hat{I}_i(1 - \hat{r}_i\delta) \,, \tag{B.1a}$$

$$\hat{r}_{i+1} = \hat{r}_i + \alpha(\beta - \hat{r}_i)\delta + \kappa\tilde{z}_{i+1}\sqrt{\delta} \,, \quad i = 0,1,\ldots,K-1 \,, \; \hat{r}_0 = r_{t_0} \,, \; \hat{I}_0 = 1 \tag{B.1b}$$

Milstein Scheme:

$$\hat{I}_{i+1} = \hat{I}_i\left(1 - \hat{r}_i\delta + \frac{1}{2}\left(\hat{r}_i^2 - \alpha(\beta - \hat{r}_i)\right)\delta^2 - \frac{1}{2}\kappa\tilde{z}_{i+1}\delta\sqrt{\delta}\right) \,, \tag{B.2a}$$

$$\hat{r}_{i+1} = \hat{r}_i + \alpha(\beta - \hat{r}_i)\delta + \kappa\tilde{z}_{i+1}\sqrt{\delta} - \frac{\kappa\alpha}{2}\tilde{z}_{i+1}\delta\sqrt{\delta} -$$
$$\frac{1}{2}\alpha^2(\beta - \hat{r}_i)\delta^2 \,, \quad i = 0,1,\ldots,K-1 \,, \; \hat{r}_0 = r_{t_0} \,, \; \hat{I}_0 = 1 \tag{B.2b}$$

Appendix C

Additional Efficiency Ratio Results for CIR Model

1-, 2-, and 10-year Discount Bonds

Efficiency Ratios for Simulation of 1-Year Discount Bond in CIR Model

$N = 10,000;\ S = 100;\ B = 100$

$a = 0.4,\ b = 0.1,\ \sigma = 0.1,\ r_0 = 0.06$

K	4	8	16	32	64	128	256
δ	0.25	0.125	0.0625	0.03125	0.01563	0.007813	0.003906
Crude Monte Carlo	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Antithetic Variate	6.76	6.12	6.76	6.01	6.52	6.68	8.92
Control Variate	7.30	7.05	7.98	7.34	7.66	7.71	9.85
Antithetic Variate + Control Variate	5.20	4.63	5.36	4.89	5.37	5.39	7.38
Measure Transformation	3.77	7.06	16.89	30.02	64.55	149.05	263.86
Measure Transformation + Antithetic Variate	10.17	22.65	*53.18*	*123.94*	*263.20*	*532.54*	*1238.48*
Sobol Sequence	4.57	3.62	3.70	2.61	2.08	2.02	2.21
Sobol Sequence + Control Variate	*14.85*	9.57	9.27	7.37	8.52	9.17	8.81
Sobol Sequence + Measure Transformation	14.58	*23.88*	45.82	63.47	103.10	253.87	470.90

Note: The highest efficiency ratio for each step-size has been highlighted
.

IV.66

Efficiency Ratios for Simulation of 2-Year Discount Bond in CIR Model

$N = 10,000;\ S = 100;\ B = 100$

$a = 0.4,\ b = 0.1,\ \sigma = 0.1,\ r_0 = 0.06$

K	4	8	16	32	64	128	256
δ	0.25	0.125	0.0625	0.03125	0.01563	0.007813	0.003906
Crude Monte Carlo	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Antithetic Variate	6.01	5.17	5.73	5.08	5.48	5.74	7.51
Control Variate	5.38	5.04	5.68	5.24	5.41	5.42	6.77
Antithetic Variate + Control Variate	4.13	3.51	4.05	3.68	4.03	4.09	5.55
Measure Transformation	4.11	7.19	16.92	29.48	61.75	120.90	175.80
Measure Transformation + Antithetic Variate	8.08	17.29	41.29	*94.17*	*184.38*	*320.54*	*613.20*
Sobol Sequence	4.79	3.54	3.58	2.51	1.99	1.97	2.18
Sobol Sequence + Control Variate	11.78	7.25	6.93	5.54	6.44	6.75	6.52
Sobol Sequence + Measure Transformation	*14.80*	*23.67*	*42.94*	58.93	97.62	222.34	345.35

Note: The highest efficiency ratio for each step-size has been highlighted

Efficiency Ratios for Simulation of 10-Year Discount Bond in CIR Model

$N = 10,000$; $S = 100$; $B = 100$

$a = 0.4$, $b = 0.1$, $\sigma = 0.1$, $r_0 = 0.06$

K	4	8	16	32	64	128	256
δ	0.25	0.125	0.0625	0.03125	0.01563	0.007813	0.003906
Crude Monte Carlo	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Antithetic Variate	5.29	4.98	5.23	4.87	6.11	6.29	7.26
Control Variate	2.62	3.12	3.29	2.98	3.25	3.22	4.05
Antithetic Variate + Control Variate	2.20	2.23	2.52	2.26	2.57	2.65	3.40
Measure Transformation	2.21	5.49	11.67	18.26	31.49	36.08	39.28
Measure Transformation + Antithetic Variate	2.52	6.47	16.21	28.70	47.46	63.11	**95.55**
Sobol Sequence	4.34	3.41	3.26	2.27	1.83	1.84	2.13
Sobol Sequence + Control Variate	5.47	4.95	4.41	3.56	4.35	3.84	3.30
Sobol Sequence + Measure Transformation	*6.58*	*17.40*	*27.19*	*32.80*	*57.93*	*74.58*	83.77

Note: The highest efficiency ratio for each step-size has been highlighted

IV.68

Appendix D

Proof of (44) -- Optimal Choice of Likelihood Ratio Process

Let us introduce the adapted process

$$u(t,T) = E_t^Q \left[\exp(-\int_{t_0}^T r_u du) \right] = E_t^Q[I_T] , \ t_0 \leq t \leq T \leq \tau \tag{D.1}$$

where I_t is given by (12). We note that $u(t,T)$ can alternatively be written

$$u(t,T) = E_t^Q \left[\exp(-\int_t^T r_u du - \int_{t_0}^t r_u du) \right] = P(t,T)I_t \ , \ t_0 \leq t \leq T \leq \tau \tag{D.2}$$

By the law of conditional iterated expectations

$$E_t^Q[u(s,T)] = E^Q \left[E^Q[I_T | \Im_s] | \Im_t \right] = E^Q[I_T | \Im_t] = u(t,T) \ , \quad t \leq s \leq T \tag{D.3}$$

which implies that $u(t,T)$ is a martingale under Q. Since we know that $P(t,T) = h_T(r_t,t)$ and thus $u(t,T) = h_T(r_t,t)I_t$, Ito's lemma yields

$$du(t,T) = \sigma_r(r_t,t)I_t \frac{\partial h_T(r_t,t)}{\partial r} dW_t \tag{D.4}$$

where we have used the martingale property (D.3) of $u(t,T)$ to set the dt-term in (D.4) equal to zero. Under the equivalent measure Z, (D.4) becomes (from (38))

$$du(t,T) = -\sigma_r(r_t,t)I_t \frac{\partial h_T(r_t,t)}{\partial r} \theta_t dt + \sigma_r(r_t,t)I_t \frac{\partial h_T(r_t,t)}{\partial r} d\tilde{W}_t \tag{D.5}$$

Given the likelihood ratio process (40), Ito's lemma asserts that the product $\xi_t u(t,T)$ follows (under Z)

$$d(\xi_t u(t,T)) = \xi_t \left(\sigma_r(r_t,t)I_t \frac{\partial h_T(r_t,t)}{\partial r} + h_T(r_t,t)I_t\theta_t \right) d\tilde{W}_t \qquad (D.6)$$

Since $u(t,T)$ is an adapted process, we know from (D.1) (or (D.2)) that $u(T,T) = I_T$ and thus

$$P(t_0,T) = E_{t_0}^Z[I_T\xi_T] = E_{t_0}^Z[u(T,T)\xi_T] \qquad (D.7)$$

From (D.6), if we set

$$\theta_t = \frac{-\sigma_r(r_t,t)}{h_T(r_t,t)} \frac{\partial h_T(r_t,t)}{\partial r} \qquad (D.8)$$

the product $\xi_t u(t,T)$ *becomes non-random under Z.* (D.7) then implies that

$$I_T\xi_T = u(T,T)\xi_T = u(t_0,T)\xi_{t_0} = u(t_0,T) = P(t_0,T) \qquad (D.9)$$

which is obviously non-random as well. Q.E.D.

IV.70

Appendix E

Simulation Results for CIR Model using Antithetic Variates and Measure Transform

2-year Discount Bond

Simulation Results using Antithetic Variates + Measure Transform in CIR Model
2-Year Discount Bond, Theoretical Value = 0.8655244

$$N = 1{,}000{,}000; \; S = 1{,}000; \; B = 1{,}000$$
$$a = 0.4, \; b = 0.1, \; \sigma = 0.1, \; r_0 = 0.06$$

	K	1	2	4	8	16	32	64	128	256
	δ	2	1	0.5	0.25	0.125	0.0625	0.03125	0.01563	0.00781
	$\log_2(\delta)$	1.00	0.00	-1.00	-2.00	-3.00	-4.00	-5.00	-6.00	-7.00
Euler	$\hat{e}$	1.45E-2	3.04E-3	9.69E-4	3.78E-4	1.65E-4	7.76E-5	3.72E-5	1.83E-5	9.12E-6
	$\log_2|\hat{e}|$	-6.11	-8.36	-10.01	-11.37	-12.57	-13.65	-14.71	-15.73	-16.74
	$\sqrt{\hat{V_B}}$	0	1.16E-4	5.56E-5	2.60E-5	1.35E-5	7.54E-6	4.42E-6	2.70E-6	1.71E-6
	$c_{0.9}$	0	6.04E-6	2.90E-6	1.35E-6	7.04E-7	3.93E-7	2.30E-7	1.41E-7	8.92E-8
	$|\hat{e}_{wc}|$	1.45E-2	3.04E-3	9.72E-4	3.79E-4	1.66E-4	7.80E-5	3.74E-5	1.85E-5	9.21E-6
	$\log_2|\hat{e}_{wc}|$	-6.11	-8.36	-10.01	-11.37	-12.56	-13.65	-14.71	-15.72	-16.73
	T_{comp}	22.12	35.43	67.81	136.15	263.31	513.01	1022.29	2044.11	4096.38
	$\log_2(T_{comp})$	4.47	5.15	6.08	7.09	8.04	9.00	10.00	11.00	12.00
Mil-stein	$\hat{e}$	-9.98E-3	-1.78E-3	-3.68E-4	-8.32E-5	-1.96E-5	-4.76E-6	-1.19E-6	-2.99E-7	-7.62E-8
	$\log_2|\hat{e}|$	-6.65	-9.13	-11.41	-13.55	-15.64	-17.68	-19.68	-21.67	-23.65
	$\sqrt{\hat{V_B}}$	2.55E-4	1.10E-4	3.78E-5	1.35E-5	4.67E-6	1.72E-6	6.60E-7	3.21E-7	2.51E-7
	$c_{0.9}$	1.33E-5	5.74E-6	1.97E-6	7.05E-7	2.44E-7	8.99E-8	3.44E-8	1.67E-8	1.31E-8
	$|\hat{e}_{wc}|$	9.99E-3	1.79E-3	3.70E-4	8.39E-5	1.98E-5	4.84E-6	1.22E-6	3.16E-7	8.93E-8
	$\log_2|\hat{e}_{wc}|$	-6.64	-9.13	-11.40	-13.54	-15.62	-17.66	-19.64	-21.59	-23.42
	T_{comp}	34.51	62.08	123.37	227.11	443.34	892.41	1777.44	3558.20	7082.72
	$\log_2(T_{comp})$	5.11	5.96	6.95	7.83	8.79	9.80	10.80	11.80	12.79
3rd Order	$\hat{e}$	NA	8.39E-4	8.77E-5	9.20E-6	1.06E-6	2.15E-7	-9.90E-9	8.00E-10	-6.60E-9
	$\log_2|\hat{e}|$	NA	-10.22	-13.48	-16.73	-19.85	-22.15	-26.59	-30.22	-27.17
	$\sqrt{\hat{V_B}}$	NA	1.10E-4	4.07E-5	1.65E-5	6.62E-6	2.92E-6	1.43E-6	6.96E-7	4.27E-7
	$c_{0.9}$	NA	5.75E-6	2.12E-6	8.62E-7	3.45E-7	1.52E-7	7.48E-8	3.63E-8	2.23E-8
	$|\hat{e}_{wc}|$	NA	8.44E-4	8.98E-5	1.01E-5	1.40E-6	3.67E-7	8.47E-8	3.71E-8	2.89E-8
	$\log_2|\hat{e}_{wc}|$	NA	-10.21	-13.44	-16.60	-19.44	-21.38	-23.49	-24.68	-25.04
	T_{comp}	NA	79.66	150.12	296.20	580.15	1145.34	2280.14	4547.46	9073.21
	$\log_2(T_{comp})$	NA	6.32	7.23	8.21	9.18	10.16	11.15	12.15	13.15

Simulation Results using Antithetic Variates + Measure Transform in CIR Model
2-Year Discount Bond, Theoretical Value = 0.8655244

$$N = 10{,}000; \ S = 100; \ B = 100$$
$$a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$$

	K	1	2	4	8	16	32	64	128	256
	δ	2	1	0.5	0.25	0.125	0.0625	0.03125	0.01563	0.00781
	$\log_2(\delta)$	1.00	0.00	-1.00	-2.00	-3.00	-4.00	-5.00	-6.00	-7.00
Euler	$\hat{e}$	1.45E-2	3.06E-3	9.72E-4	3.75E-4	1.70E-4	7.88E-5	3.81E-5	1.95E-5	9.80E-6
	$\log_2\lvert\hat{e}\rvert$	-6.11	-8.35	-10.01	-11.38	-12.52	-13.63	-14.68	-15.65	-16.64
	$\sqrt{\hat{V_B}}$	0	1.24E-3	6.34E-4	2.61E-4	1.51E-4	6.50E-5	4.20E-5	2.84E-5	1.63E-5
	$c_{0.9}$	0	6.52E-5	3.33E-5	1.37E-5	7.91E-6	3.41E-6	2.21E-6	1.49E-6	8.54E-7
	$\lvert\hat{e}_{wc}\rvert$	1.45E-2	3.12E-3	1.01E-3	3.89E-4	1.78E-4	8.22E-5	4.03E-5	2.10E-5	1.07E-5
	$\log_2\lvert\hat{e}_{wc}\rvert$	-6.11	-8.32	-9.96	-11.33	-12.46	-13.57	-14.60	-15.54	-16.52
	T_{comp}	0.22	0.36	0.69	1.34	2.62	5.11	10.22	20.45	40.93
	$\log_2(T_{comp})$	-2.18	-1.47	-0.54	0.42	1.39	2.35	3.35	4.35	5.36
Mil-stein	$\hat{e}$	-9.91E-3	-1.77E-3	-3.65E-4	-8.14E-5	-2.00E-5	-4.79E-6	-1.10E-6	-3.63E-7	-3.20E-8
	$\log_2\lvert\hat{e}\rvert$	-6.66	-9.14	-11.42	-13.58	-15.61	-17.67	-19.80	-21.39	-24.90
	$\sqrt{\hat{V_B}}$	2.48E-3	1.03E-3	3.70E-4	1.41E-4	4.74E-5	1.62E-5	6.71E-6	3.62E-6	2.48E-6
	$c_{0.9}$	1.30E-4	5.39E-5	1.94E-5	7.40E-6	2.49E-6	8.50E-7	3.52E-7	1.90E-7	1.30E-7
	$\lvert\hat{e}_{wc}\rvert$	1.00E-2	1.82E-3	3.84E-4	8.88E-5	2.24E-5	5.64E-6	1.45E-6	5.53E-7	1.62E-7
	$\log_2\lvert\hat{e}_{wc}\rvert$	-6.64	-9.10	-11.35	-13.46	-15.44	-17.44	-19.40	-20.79	-22.55
	T_{comp}	0.33	0.60	1.20	2.25	4.45	8.91	17.79	35.56	70.81
	$\log_2(T_{comp})$	-1.60	-0.74	0.26	1.17	2.15	3.16	4.15	5.15	6.15
3rd Order	$\hat{e}$	NA	8.45E-4	9.69E-5	9.96E-6	2.56E-6	7.63E-8	-2.05E-7	-2.85E-7	8.04E-8
	$\log_2\lvert\hat{e}\rvert$	NA	-10.21	-13.33	-16.62	-18.58	-23.64	-22.22	-21.74	-23.57
	$\sqrt{\hat{V_B}}$	NA	1.08E-3	3.99E-4	1.77E-4	6.90E-5	2.80E-5	1.35E-5	7.18E-6	4.01E-6
	$c_{0.9}$	NA	5.68E-5	2.09E-5	9.28E-6	3.62E-6	1.47E-6	7.10E-7	3.77E-7	2.10E-7
	$\lvert\hat{e}_{wc}\rvert$	NA	9.01E-4	1.18E-4	1.92E-5	6.18E-6	1.54E-6	9.15E-7	6.62E-7	2.91E-7
	$\log_2\lvert\hat{e}_{wc}\rvert$	NA	-10.12	-13.05	-15.67	-17.30	-19.31	-20.06	-20.53	-21.71
	T_{comp}	NA	0.79	1.48	2.92	5.78	11.45	22.80	45.49	90.69
	$\log_2(T_{comp})$	NA	-0.34	0.57	1.55	2.53	3.52	4.51	5.51	6.50

Simulation Results using Antithetic Variates + Measure Transform in CIR Model
2-Year Discount Bond, Theoretical Value = 0.8655244

$$N = 100; \ S = 10; \ B = 10$$
$$a = 0.4, \ b = 0.1, \ \sigma = 0.1, \ r_0 = 0.06$$

	K	1	2	4	8	16	32	64	128	256
	δ	2	1	0.5	0.25	0.125	0.0625	0.03125	0.01563	0.00781
	$\log_2(\delta)$	1.00	0.00	-1.00	-2.00	-3.00	-4.00	-5.00	-6.00	-7.00
Euler	$\hat{e}$	1.45E-2	3.05E-3	1.05E-3	2.65E-4	1.88E-4	8.08E-5	4.65E-5	2.14E-5	6.80E-6
	$\log_2|\hat{e}|$	-6.11	-8.36	-9.89	-11.88	-12.38	-13.59	-14.39	-15.51	-17.17
	$\sqrt{\hat{V_B}}$	0	1.50E-2	6.74E-3	1.56E-3	1.12E-3	6.54E-4	2.75E-4	2.99E-4	1.29E-4
	$c_{0.9}$	0	8.68E-4	3.90E-4	9.03E-5	6.46E-5	3.78E-5	1.59E-5	1.73E-5	7.44E-6
	$|\hat{e}_{wc}|$	1.45E-2	3.92E-3	1.44E-3	3.55E-4	2.52E-4	1.19E-4	6.24E-5	3.87E-5	1.42E-5
	$\log_2|\hat{e}_{wc}|$	-6.11	-7.99	-9.44	-11.46	-11.95	-13.04	-13.97	-14.66	-16.10
	T_{comp}	0.002	0.004	0.007	0.015	0.031	0.052	0.102	0.221	0.428
	$\log_2(T_{comp})$	-8.97	-7.97	-7.16	-6.06	-5.01	-4.27	-3.29	-2.18	-1.22
Mil-stein	$\hat{e}$	-9.34E-3	-1.62E-3	-2.66E-4	-4.44E-5	-9.07E-6	3.29E-6	-5.37E-7	5.09E-7	-2.57E-7
	$\log_2|\hat{e}|$	-6.74	-9.27	-11.88	-14.46	-16.75	-18.21	-20.83	-20.91	-21.89
	$\sqrt{\hat{V_B}}$	3.72E-2	1.25E-2	3.74E-3	1.13E-3	2.23E-4	1.03E-4	7.55E-5	3.49E-5	2.79E-5
	$c_{0.9}$	2.15E-3	7.24E-4	2.16E-4	6.52E-5	1.29E-5	5.98E-6	4.37E-6	2.02E-6	1.61E-6
	$|\hat{e}_{wc}|$	1.15E-2	2.35E-3	4.82E-4	1.10E-4	2.20E-5	9.28E-6	4.90E-6	2.53E-6	1.87E-6
	$\log_2|\hat{e}_{wc}|$	-6.44	-8.73	-11.02	-13.16	-15.47	-16.72	-17.64	-18.59	-19.03
	T_{comp}	0.003	0.006	0.012	0.022	0.041	0.087	0.189	0.338	0.712
	$\log_2(T_{comp})$	-8.38	-7.38	-6.38	-5.51	-4.61	-3.52	-2.40	-1.56	-0.49
3rd Order	$\hat{e}$	NA	8.78E-4	1.23E-4	1.80E-5	9.20E-6	2.94E-6	1.87E-6	4.20E-7	-1.40E-6
	$\log_2|\hat{e}|$	NA	-10.15	-12.99	-15.76	-16.73	-18.37	-19.03	-21.18	-19.45
	$\sqrt{\hat{V_B}}$	NA	1.45E-2	4.15E-3	1.98E-3	6.24E-4	2.50E-4	1.35E-4	7.14E-5	4.56E-5
	$c_{0.9}$	NA	8.38E-4	2.40E-4	1.15E-4	3.61E-5	1.45E-5	7.83E-6	4.13E-6	2.64E-6
	$|\hat{e}_{wc}|$	NA	1.72E-3	3.63E-4	1.33E-4	4.53E-5	1.74E-5	9.70E-6	4.55E-6	4.04E-6
	$\log_2|\hat{e}_{wc}|$	NA	-9.19	-11.43	-12.88	-14.43	-15.81	-16.65	-17.75	-17.92
	T_{comp}	NA	0.008	0.015	0.034	0.065	0.116	0.241	0.456	0.942
	$\log_2(T_{comp})$	NA	-6.97	-6.06	-4.88	-3.94	-3.11	-2.05	-1.13	-0.09

IV.74

Endnotes

[1] It is well known that the computational expense of Monte Carlo methods grows linearly with the number of stochastic factors whereas lattice methods are characterized by an exponential relationship between the number of factors and computational expense. In general, lattice methods become impractical when the number of factors exceed 2 or 3; for a discussion of lattices with two state variables, see for example Boyle (1988). For an application of lattice approaches to mildly path-dependent structures, see Hull and White (1993).

[2] For an extensive list of traditional approaches to variance reduction, see the standard reference Hammersley and Handscomb (1964).

[3] For a more complete description of short rate processes proposed in the literature, see Duffie and Kan (1994) or Joergensen (1994), p. 95-99.

[4] All these studies assume that interest rate evolution can be described by a (risk-neutral) short rate process of the form $dr_t = \mu(r_t, t)dt + cr_t^{\gamma} dW_t$. For the US market, the three papers estimate the exponent γ to be 1.3438, 1.4999, and 1.13, respectively.

[5] It is probably slightly confusing that we use P to denote both zero-coupon bond prices and the real-life probability measure. Fortunately, we will not need the real-life probability measure at any further point in this paper. Indeed, as mentioned in Harrison and Pliska (1981), the sole role of the real-life probability measure is to define the null sets.

[6] These conditions are given by Girsanov's Theorem (see Karatzas and Shreve (1991), p. 190-201) and roughly require the market price of risk of discount bonds to exist and be sufficiently well-behaved (Novikov's condition).

[7] Being an applied study, this paper will not delve too deeply on technical regularity conditions. We do point out, however, that the existence of a unique solution to (2) is ensured if μ_r and σ_r satisfy Lipschitz and linear growth conditions. The delicate issue of

existence and uniqueness for the Cox-Ingersoll-Ross square-root model (which violates these conditions) is discussed by Duffie (1992), Appendix E.

[8]We note that both (4) and (5) can be extended to the case of deterministic parameters and still be (somewhat) analytically tractable (see Hull and White (1990b) and Pliska (1994)).

[9]As shown in Cox, Ingersoll, and Ross (1980), p. 391, if the drift parameters of the CIR process are sufficiently high, the origin becomes inaccessible. In this case, $D = \Re^+ \setminus \{0\}$.

[10]For an introduction to Ito-Taylor expansions, see Kloeden and Platen (1992), chapter 5.

[11]The process $V_s - V_t$ is the representation of the double stochastic integral $\int_t^s \int_t^u dW_v du$.

[12]We note that the CIR process, similar to geometric Brownian motion (1), can be simulated bias-free using the known solution to the SDE. As mentioned in section 2, this would involve generating random draws from the non-central chi-square distribution; a simple algorithm for generating such draws can be found in Devroye (1986), p. 468-471.

[13]These parameters are identical to the ones chosen (and justified) by Hull and White (1990a).

[14]Notice, that the variance of $\hat{I}_{K,A}$ will be smaller than the variance of $\hat{I}_K$ if just $\rho(\hat{I}_K, \hat{I}_{K-}) < 1$ which, in practice, will always be the case. As the computation time of the antithetic method is roughly twice that of crude Monte Carlo, the method, however, does not necessarily produce any gains in *efficiency* (see equation (37) for a definition of efficiency).

[15]The approach taken in this paper is based on traditional Monte Carlo methods. Newton (1994) has introduced an SDE-based control variate technique quite different from the one applied in this paper.

[16]In the general case where the moments of the short rate might not be explicitly known, some other matching criteria must be used. For example, in the CIR model an alternative

"naive" approach could be to simply set $\kappa = \sigma \sqrt{r_{t_0}}$ (which actually generates quite reasonable results).

[17]Under technical conditions on the short rate process parameters, the existence of such a function follows from the Feynman-Kac theorem, see for example Duffie (1992), p. 129.

[18]For an exact definition of discrepancy, see for example Bouleau and Lepingle (1994).

[19]Using L'Hospital's rule, one can show that $(\ln N)^K / N$ behaves almost as $O(1/N)$ for large N.

[20]Many contingent claims -- including the simple European puts and calls on zero-coupon bonds -- have kinks that violate the classical smoothness conditions. In practice, this hardly represents a problem as we can easily ensure that the conditions are satisfied by introducing a slight local smoothing of the payout function (as in Duffie (1992), p. 203). More rigorously, Bally and Talay (1995) show that in many cases it is possible to relax the classical smoothness conditions.