

Efficient unbiased simulation scheme for the SABR stochastic volatility model

Bin Chen*, Cornelis W. Oosterlee

MAC-2, Center for Mathematics and Computer Science,
Amsterdam, the Netherlands

Hans van der Weide

Faculty of Electrical Engineering, Mathematics and Computer Science,
TU Delft, the Netherlands

Ver. 4.02 January 19, 2011

Abstract

The stochastic Alpha Beta Rho stochastic volatility (SABR-SV) model is widely used in the financial industry for the pricing of fixed income instruments. In this paper we develop an unbiased simulation scheme for the SABR-SV model, which deals efficiently with (undesired) possible negative values, the martingale property of the discrete scheme and the discretization bias of commonly used Euler discretization schemes. The proposed algorithm is based on the analytic properties of the governing distribution. Experiments with realistic model parameters show that this scheme is robust for interest rate valuation.

1 Introduction

The Stochastic Alpha Beta Rho (SABR) model by Hagan [12] is a popular model in the financial industry because of the availability of an analytic asymptotic implied volatility formula. Practical applications of the SABR model include interpolation of volatility surfaces and the hedging of volatility risk. In the context of pricing interest rate derivatives, the combination of the SABR model and the market standard Libor Market Models (LMM) [27] is of particular interest. Other references on this topic include Morini & Mercurio [26], Hagan & Lesniewski [13] or Labordere [21].

The constant elasticity of variance (CEV) process, introduced by Cox [9], is an important ingredient of the SABR model. The CEV process has appeared in several other models in finance, including the CEV Libor Market Model (LMM) by Andersen & Andreasen [2].

Pricing and hedging of path-dependent securities are applications that typically require the use of Monte Carlo methods. Despite the fact that the CEV model has been introduced more than 30 years ago and that various researchers have shown evidence of significant bias in the basic Euler scheme for the CEV model, only a few references devising efficient unbiased Monte Carlo schemes were found in the literature.

It was shown by Schroder [31] that the CEV process, representing the asset price dynamics of the SABR model, equals a space transformed *squared Bessel* process. As the volatility process in the SABR model is driven by a geometric Brownian motion, a close relation between the SABR model and the Heston model [14] exists: In the Heston model the asset price dynamics follow geometric Brownian motion, whereas the volatility is governed by a squared Bessel process. Due to this, it seems natural to generalize the unbiased simulation schemes for the Heston stochastic volatility model to the SABR case. Broadie & Kaya [6]'s so-called *exact simulation scheme* (the BK scheme) is based on the insight in Willard [33] that the conditional distribution, given the terminal volatility and the integrated variance in a time interval, is log-normal. In their scheme, an acceptance-rejection technique is employed to sample the variance process, and a Fourier inversion technique is applied to recover the variance process. Although the BK scheme is free of bias by construction, its practical

*Corresponding author. Postal address: P.O. Box 94079, 1090 GB, Amsterdam, the Netherlands. Email: B.Chen@cwi.nl

application is hampered by its computational speed. Andersen [1] developed two efficient variants of the BK scheme, the truncated Gaussian (TG) and the quadratic exponential (QE) schemes, that are both based on the moment matching technique. Essentially, the noncentral chi-square distribution is approximated by a distribution whose moments are matched with those of the exact distribution. Since the QE scheme is based on transformations to uniform and normal random numbers, it can be implemented efficiently [11].

A direct application of the QE scheme to the SABR model does not work well, because the QE scheme is based on a squared Bessel process with a *reflecting* boundary at zero volatility, which gives rise to a sub-martingale process. It is therefore not suited to model SABR's asset price dynamics. Instead, a squared Bessel process with an *absorbing* boundary is the specification which is in agreement with the arbitrage-free constraints (and thus produces a true martingale process). Accurate handling of the absorbing boundary behavior is nontrivial, as the transition density of the absorbed process does not integrate to unity and the moments are not known in closed form.

Some simulation algorithms for the squared Bessel processes exist in the literature. Andersen & Andreasen [2] investigated basic Euler as well as log-Euler schemes for the CEV model in a Monte Carlo setting and mentioned that 'the simulated prices of caps, floors and swaptions exhibit a bias relative to the continuous-time prices ... even for an infinite number of Monte Carlo trials'. Kahl & Jäckel [19] proposed a higher-order Monte Carlo scheme based on an implicit Milstein discretization for the CEV process. This scheme does not perform well when zero is attainable, due to the discontinuity of the first derivative of the diffusion coefficient. Campolieti & Makarov [7] developed an exact scheme based on an acceptance-rejection sampling of the Bessel bridge process. The Bessel bridge scheme is however quite complex and its computational time is relatively high. Lord et al. [24] consider an Euler scheme in combination with certain rules to deal with negative paths produced by the Euler scheme. The authors conclude that the computational efficiency of the Euler scheme with these fixes is superior to the more complicated schemes. For certain relevant parameter configurations the scheme produces a significant bias for practical sizes of the time steps.

In this paper we propose an unbiased path simulation scheme for the continuous-time CEV and SABR models, also based on Willard's [33] idea of mixing the conditional distributions [23] of a stochastic volatility model (given the terminal volatility and integrated variance). Our contribution is threefold. First of all, we derive the conditional distribution of the SABR model over a discrete time step and show that, conditioned on the terminal volatility and the integrated variance, it is a space transformed squared Bessel process with a shifted initial condition. Secondly, we propose an efficient easy-to-implement algorithm to simulate the squared Bessel process with an absorbing boundary at zero. Thirdly, we provide a simple approximation formula for the conditional moments of the integrated variance by means of the small disturbance expansion method (see Kunitomo & Takahashi [20] and Chen [8]), which facilitates effective sampling from the joint distribution of the terminal volatility and the integrated variance.

The paper is organized as follows. In Section 2 we describe the basic SABR model and summarize some analytic properties that are relevant for simulation. In Section 3, we review some existing discretization schemes, for later comparison. In Section 3.4 we present the unbiased discretization scheme to simulate the asset and the variance processes. Section 4 discusses the performance of the whole algorithm. Section 5 concludes.

2 Some analytic features of the SABR model

Given a time interval Δ and an arbitrary set \mathcal{T} of discrete times $s < s + \Delta \dots < s + N\Delta$ and a stochastic process $X = \{X_t; t > 0\}$, a discretized simulation scheme generates a skeleton $X_s, X_{s+\Delta}, \dots$ of a sample path of the stochastic process X . To device such a scheme, we start sampling from the marginal distribution of $X_{s+\Delta}$. A repetition of such a one-period scheme may produce the full time-discrete paths for X . Since here we consider the discrete scheme generating paths for a stochastic volatility model, $X_t = (S_t, \sigma_t)$, for all $t \in \mathcal{T}$, the asset price process S_t itself is not a Markov process. The fundamental question, as argued by Andersen [1], is how to generate a random sample of $S_{s+\Delta}$ from the conditional distribution of $S_{s+\Delta}$, given $(S_s, \sigma_s, \sigma_{s+\Delta}, \int_s^{s+\Delta} \sigma_u^2 du)$.

In this section, we focus on the analysis of the conditional distribution function; the sampling technique is discussed in the next section. For notational convenience, we do not use the notation s and $s + \Delta$ for the initial and terminal time points, respectively, but use simply 0 for the initial time

point, t, s represent the running time and Δ is the terminal time for each discrete time interval.

The Stochastic Alpha Beta Rho Stochastic Volatility (SABR SV) model [12] is given by the following system of stochastic differential equations (SDEs) with constant parameters α and β :

$$\begin{aligned} dS_t &= \sigma_t S_t^\beta dW_{1,t}, \\ d\sigma_t &= \alpha \sigma_t dW_{2,t}, \end{aligned} \quad (2.1)$$

where $E[dW_{1,t}dW_{2,t}] = \rho dt$. Since the asset price S_t itself follows a CEV process, one can expect that the conditional SABR process, given σ_Δ and $\int_0^\Delta \sigma_s^2 ds$, is a CEV process as well. The next step will be to ‘mix’ the conditional CEV process with the joint distribution of σ_Δ and $\int_0^\Delta \sigma_s^2 ds$.

We will show, in the subsequent section, that, conditional on σ_Δ and $\int_0^\Delta \sigma_s^2 ds$, the coordinate transformed asset price process defined by the invertible transformation $X(S) = S^{1-\beta}/(1-\beta)$ is a time-changed Bessel process of dimension $(1-2\beta-\rho^2(1-\beta))/((1-\beta)(1-\rho^2))$, starting at $S_0^{1-\beta}/(1-\beta)$. Based on this, the analytic distribution function for S_Δ will be derived. After that we show that we can sample this conditional distribution efficiently by a direct inversion scheme.

2.1 The distribution of the CEV process

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space generated by $\{W_t\}$, a one-dimensional Brownian motion. For all $0 \leq t \leq T$, the CEV process can be described by the following stochastic differential equation:

$$dS_t = \sigma_t S_t^\beta dW_t \quad (2.2)$$

with initial condition $S_0 = \zeta_0$ which we assume to be \mathcal{F}_0 -measurable.

Here we simply choose σ_t to be a constant, i.e. $\sigma_t \equiv \sigma$. Following Schroder [31], we consider an invertible transformation $X_t = S_t^{1-\beta}/(1-\beta)$ for $\beta \neq 1$. Application of Itô’s lemma gives us the following SDE for X_t which we recognize as a time-changed Bessel process:

$$\begin{aligned} dX_t &= (1-\beta) \frac{S_t^{-\beta}}{1-\beta} \sigma S_t^\beta dW_{1,t} - \frac{1}{2} \beta(1-\beta) \frac{S_t^{-1-\beta}}{1-\beta} \sigma^2 S_t^{2\beta} dt \\ &= \sigma dW_t - \frac{\beta\sigma^2}{(2-2\beta)X_t} dt. \end{aligned} \quad (2.3)$$

Then we define a second transformation, $Y_t = X_t^2$, which results in a time-changed squared Bessel process of dimension $\delta := (1-2\beta)/(1-\beta)$, that thus satisfies the following SDE:

$$dY_t = 2\sqrt{|Y_t|} \sigma dW_t + \delta \sigma^2 dt. \quad (2.4)$$

Let $\nu(t)$ be a time-change function, so that $\nu(t) = \sigma^2 t$. Then, $Y_t = Z_{\nu(t)}$, where $\{Z_t\}$ is a δ -dimensional squared Bessel process, i.e., the strong solution of the SDE:

$$dZ_t = 2\sqrt{|Z_t|} dW_t + \delta dt \quad (2.5)$$

with degree of freedom δ . The squared Bessel process is a Markov process and its transition densities are known explicitly.

The next step is to sample random numbers from the analytic transition density in Z -space and to apply an inverse variable transformation to obtain the random numbers in the original coordinates.

First, a few technical details need to be discussed. They are presented in the form of results and propositions below.

RESULT 2.1. [2]: For a standard squared Bessel process, as defined by SDE (2.5), the following statements hold true:

1. All solutions to SDE (2.5) are non-explosive.
2. For $\delta < 2$, $Z = 0$ is an attainable boundary for Process (2.5).

3. For $\delta \geq 2$ SDE (2.5) has a unique solution and zero is not attainable.
4. For $0 < \delta < 2$ the SDE (2.5) does not have a unique solution, unless the boundary condition is specified for the solution to Eq. (2.5) at $Z = 0$.
5. For $\delta \leq 0$, there is a unique strong solution to the SDE (2.2), and boundary condition zero is absorbing.

Proof. The results have been proved in Appendix A of [2] based on the theory presented in [5]. \square

For the latter two cases, the transition densities are known:

RESULT 2.2 (Transition density for squared Bessel process). *The transition density, $q^\delta(t, x, y)$, for the squared Bessel process reads:*

1. For $\delta \leq 0$ and for $0 < \delta < 2$ in Eq. (2.5) but only when the boundary is absorbing at $y = 0$:

$$q^\delta(t, x, y) = \frac{1}{2t} \left(\frac{y}{x}\right)^{\frac{\delta-2}{4}} \exp\left(-\frac{x+y}{2t}\right) I_{|\frac{\delta-2}{2}|}\left(\frac{\sqrt{xy}}{t}\right), \quad y \geq 0, t > 0. \quad (2.6)$$

2. For $0 < \delta < 2$ when $y = 0$ is a reflecting boundary:

$$q^\delta(t, x, y) = \frac{1}{2t} \left(\frac{y}{x}\right)^{\frac{\delta-2}{4}} \exp\left(-\frac{x+y}{2t}\right) I_{\frac{\delta-2}{2}}\left(\frac{\sqrt{xy}}{t}\right), \quad y \geq 0, t > 0. \quad (2.7)$$

Here we denote by $I_a(x)$ the Bessel function, defined by

$$I_a(x) := \sum_{j=0}^{\infty} \frac{(x/2)^{2j+a}}{j! \Gamma(a+j+1)},$$

and by $\Gamma(x)$ the Gamma function, $\Gamma(x) := \int_0^\infty u^{x-1} e^{-u} du$.

Proof. See Borodin [5] p. 136 for squared Bessel process transition densities. \square

Table 1: The mapping of three parameter ranges.

CEV exponent	Squared Bessel δ
$-\infty < \beta < \frac{1}{2}$	$0 < \delta < 2$
$\frac{1}{2} \leq \beta < 1$	$-\infty < \delta \leq 0$
$\beta > 1$	$2 < \delta < \infty$

By solving a series of inequalities (see the results in Table 1), we find essentially three different parameter ranges which determine the behavior of the CEV process at the boundary and the form of the transition densities:

1. For $\beta > 1$, SDE (2.2) has a unique solution and boundary condition zero is not attainable. The density function integrates to unity over $S \in (0, \infty)$ for all $t \geq 0$ and the process S_t is a strict local martingale.
2. For $\beta < \frac{1}{2}$, SDE (2.2) does not have a unique solution, unless a separate boundary condition is specified for the boundary behavior at $S = 0$.
 - The density integrates to unity, if the boundary is *reflecting* and process S_t is a strict sub-martingale.
 - The density will not integrate to unity if the boundary at $S = 0$ is *absorbing*¹ and process S_t is a true martingale.

¹There is a degenerate part with an atom in the boundary and an absolutely continuous part over $(0, \infty)$.

3. For $\frac{1}{2} \leq \beta < 1$, a unique strong solution to SDE (2.2) exists, and boundary value zero is absorbing. The density function does not integrate to unity for $t > 0$ and process S_t is a *true martingale*.

For most financial applications, parameter β ranges between 0 to 1, which is included in Cases 2 and 3 in the list above. We therefore focus on these two cases, and, correspondingly, on the Items 4 and 5 in Result 2.1.

Based on the transition density of the squared Bessel diffusions in X -space given in Result 2.2, one can easily obtain the transition density for the CEV process (2.2) in S -space. Note first of all that

$$S_\Delta = \left((1 - \beta) \sqrt{|Z_{\nu(\Delta)}|} \right)^{\frac{1}{1-\beta}}.$$

Let us define a map

$$h : s \rightarrow \left((1 - \beta) \sqrt{s} \right)^{\frac{1}{1-\beta}}, \quad s \geq 0,$$

with inverse

$$h^{-1} : y \rightarrow \frac{y^{2(1-\beta)}}{(1-\beta)^2}, \quad y \geq 0.$$

So, $S_\Delta = h(Z_{\nu(\Delta)})$ and $Z_0 = h^{-1}(S_0) = S_0^{2(1-\beta)} / (1-\beta)^2$. Then $Z_{\nu(\Delta)}$ has density $q^\delta(\nu(\Delta), Z_0, y)$ and it follows that the density for S_Δ is given by

$$p(S|S_0) = q^\delta(\nu(\Delta), Z_0, h^{-1}(s)) \frac{dh^{-1}(s)}{dS},$$

where we use $p(S|S_0)$ to denote the conditional transition density for the CEV process. By combining the two cases considered in Result 2.2, the related transition densities for the CEV process, S_t , in Eq. (2.2) are of the following form:

1. For $0 < \beta < \frac{1}{2}$ with absorption at zero and for $\frac{1}{2} \leq \beta < 1$:

$$p(S_\Delta|S_0) = \frac{1}{\nu(\Delta)} \left(\frac{S_\Delta}{S_0} \right)^{-\frac{1}{2}} \exp \left(- \frac{S_\Delta^{2(1-\beta)} + S_0^{2(1-\beta)}}{2(1-\beta)^2 \nu(\Delta)} \right) I_{|\frac{\delta-2}{2}|} \left(\frac{(S_0 S_\Delta)^{1-\beta}}{\nu(\Delta)(1-\beta)^2} \right) \frac{S_\Delta^{1-2\beta}}{1-\beta},$$

where $\nu(\Delta) = \sigma^2 \Delta$ and $\delta = \frac{1-2\beta}{1-\beta}$.

2. For $0 < \beta < \frac{1}{2}$ with a reflecting boundary at $S = 0$:

$$p(S_\Delta|S_0) = \frac{1}{\nu(\Delta)} \left(\frac{S_\Delta}{S_0} \right)^{-\frac{1}{2}} \exp \left(- \frac{S_\Delta^{2(1-\beta)} + S_0^{2(1-\beta)}}{2(1-\beta)^2 \nu(\Delta)} \right) I_{\frac{\delta-2}{2}} \left(\frac{(S_0 S_\Delta)^{1-\beta}}{\nu(\Delta)(1-\beta)^2} \right) \frac{S_\Delta^{1-2\beta}}{1-\beta}.$$

By integrating these identities, we find the cumulative distribution functions:

RESULT 2.3. *The cumulative distribution function of the CEV price process as in Eq. (2.2) is given by the following formula:*

1. For $0 < \beta < \frac{1}{2}$ with absorption at zero and for $\frac{1}{2} \leq \beta < 1$:

$$\Pr(S_\Delta \leq x|S_0) = 1 - \chi^2(a; b, c). \quad (2.8)$$

2. For $0 < \beta < \frac{1}{2}$ with a reflecting boundary at $S = 0$:

$$\Pr(S_\Delta \leq x|S_0) = \chi^2(c; 2 - b, a), \quad (2.9)$$

with the following parameters:

$$a = \frac{S_0^{2(1-\beta)}}{(1-\beta)^2\nu(\Delta)}, \quad b = \frac{1}{1-\beta}, \quad c = \frac{x^{2(1-\beta)}}{(1-\beta)^2\nu(\Delta)}, \quad \nu(\Delta) = \sigma^2\Delta,$$

and $\chi^2(x; \delta, \lambda)$ is the noncentral chi-square cumulative distribution function for random variable x with non-centrality parameter λ and degree of freedom δ .

Proof. The proofs of these results can be found in Schroder [31] using classic results for Bessel processes. An alternative proof based on Green function theory can be found in Lesniewski [22]. \square

As stated in Result 2.1, the density will not integrate to unity when the boundary is absorbing. The shortage in the total probability mass is the probability absorbed at $S = 0$. Following the result of Result 2.3, a formula for the absorption probability can be obtained:

COROLLARY 2.1. For $0 < \beta < 1$, the probability of S_Δ given by SDE (2.2) and initial conditional S_0 reads

$$\Pr(S_\Delta = 0|S_0) = 1 - \gamma\left(\frac{1}{2(1-\beta)}, \frac{S_0^{2(1-\beta)}}{2(1-\beta)^2\nu(\Delta)}\right) / \Gamma\left(\frac{1}{2(1-\beta)}\right), \quad (2.10)$$

where $\gamma(\alpha, \beta)$ is the lower incomplete Gamma function and $\Gamma(\alpha)$ is the Gamma function.

Proof. Choosing x in Eq. (2.8) to be zero, we find that

$$\begin{aligned} \Pr(S_\Delta = 0|S_0) &= 1 - \chi^2\left(\frac{S_0^{2(1-\beta)}}{(1-\beta)^2\nu(\Delta)}; \frac{1}{1-\beta}, 0\right) \\ &= 1 - \text{Chi}^2\left(\frac{S_0^{2(1-\beta)}}{(1-\beta)^2\nu(\Delta)}; \frac{1}{1-\beta}\right), \end{aligned} \quad (2.11)$$

where the last equality sign is because the noncentral chi-square distribution with a zero non-centrality parameter reduces to a chi-square distribution. A chi-square distribution has an explicit cumulative function in terms of gamma functions:

$$\text{Chi}^2(x; \delta) = \frac{\gamma(\delta/2, x/2)}{\Gamma(\delta/2)}, \quad (2.12)$$

where $\gamma(\alpha, \beta)$ is again the lower incomplete Gamma function. We substitute Eq. (2.12) into Eq. (2.11) and prove the claim. \square

In interest rate derivative pricing, initial value, S_0 , is often very small, hence it is likely for S_t to reach zero. The specification of the boundary condition for process S_t at zero may have a significant impact on the distribution, as is evident from the distributions in Figure 1. The two plots show a comparison of the exact cumulative distribution of the CEV process versus the log-normal and normal distributions at $T = 0.25$ for S_Δ , given two different levels of S_0 , i.e. $S_0 = 6\%$ in left-hand plot and $S_0 = 2\%$ in the right-hand plot. The model parameters chosen were $\sigma = 0.3$ and $\beta = 0.4$. We have included the normal and log-normal distributions by matching the first two moments of S_Δ . The labels ‘absorbing’ and ‘reflecting’ distinguish the distributions of the CEV process with an absorption and a reflection boundary condition, respectively. The two matched distributions do not represent accurate approximations for the true distributions of S_Δ .

Andersen & Andreasen [2] and Rebonato [28], pp.48, argue that if the asset price follows a CEV process under a certain measure, there is only one acceptable boundary condition at zero to ensure the arbitrage-free conditions, which is the absorption condition. If there were a reflecting boundary at zero, then for an initially worthless portfolio, we would wait until zero is reached (which would happen with a strictly positive probability). Once the price is zero, we take a long position in the asset and sell it immediately when the 0 barrier has reflected the price process, realizing a risk-free profit.

We therefore also assume that the boundary is absorbing at $S_t = 0$ and develop a method to sample from distribution function (2.8). The strategy of sampling will be discussed in full detail in the sections to follow.

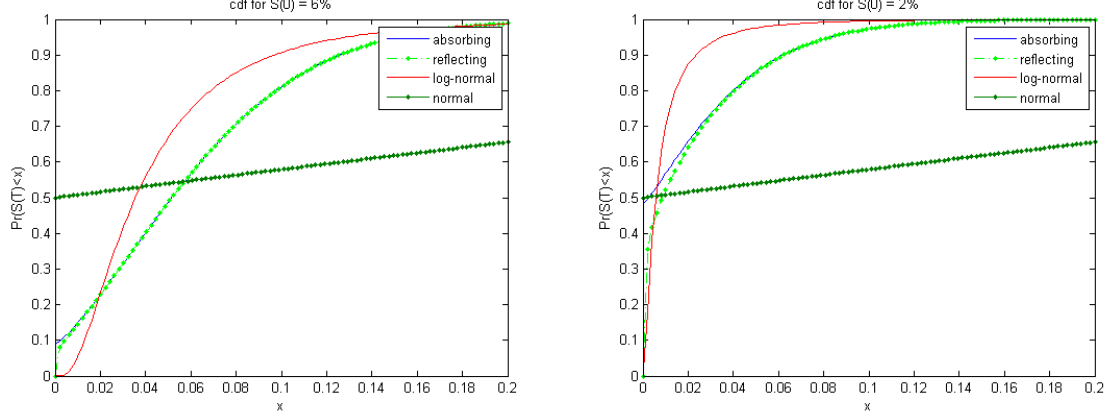


Figure 1: Comparison of the exact cumulative distribution of the CEV process versus the log-normal and normal distributions at $T = 0.25$ for S_Δ given different levels of S_0 .

2.2 SABR conditional distribution

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \Pr)$ be a filtered probability space generated by two Brownian motions $\{U_t, W_{2,t}\}$. We denote the probability space as the product of two filtered probability spaces generated by two independent Brownian motions, i.e. $\Omega = \Omega_1 \times \Omega_2, \mathcal{F} = \mathcal{F}^1 \times \mathcal{F}^2, \{\mathcal{F}_t\} = \{\mathcal{F}_t^1 \times \mathcal{F}_t^2\}, \Pr = \Pr_1 \times \Pr_2$.

Whereas a conventional numerical scheme requires the realizations of S_Δ to be $\mathcal{F}_0 = \mathcal{F}_0^1 \times \mathcal{F}_0^2$ measurable, here, S_Δ depends on the realization of σ_Δ (i.e. \mathcal{F}_Δ^2 measurable).

Based on the closed-form distribution function of the CEV process, we show that, conditional on the levels of σ_Δ and $\int_0^\Delta \sigma_s^2 ds$, the transformed asset price process $S_t^{2-2\beta}/(1-\beta)^2$ is a shifted squared Bessel process with initial condition $(S_0^{1-\beta}/(1-\beta) + \frac{\rho}{\alpha}(\sigma_\Delta - \sigma_0))^2$. Following the arguments of Islah [15], we present the following results:

RESULT 2.4 (SABR Conditional Distribution [15]). *In the context of SABR model (2.1), and conditional on the level of terminal volatility σ_Δ and integrated variance $\int_0^\Delta \sigma_s^2 ds$, let $\omega_2 \in \Omega_2, t \rightarrow \sigma_t(\omega_2)$ be a volatility path. The followings statements hold:*

1. For an invertible variable transformation $X(S) = S^{1-\beta}/(1-\beta)$, application of Itô's lemma gives us

$$X_\Delta = X_0 + \frac{\rho}{\alpha}(\sigma_\Delta - \sigma_0) + \sqrt{1-\rho^2} \int_0^\Delta \sigma_s dU_s - \int_0^\Delta \frac{\beta \sigma_s^2}{(2-2\beta)X_s} ds, \quad (2.13)$$

where U_s is a standard Brownian motion, independent of $W_{2,s}$ in System (2.1).

2. With $Y(S) = S^{2-2\beta}/(1-\beta)^2$ and application of a time-change $\nu(t) = (1-\rho^2) \int_0^t \sigma_s^2 ds$, Y_t is a squared Bessel process of dimension $\frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)}$ solving the SDE:

$$dY_{\nu(t)} = 2\sqrt{Y_{\nu(t)}} dU_{\nu(t)} + \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)} d\nu(t), \quad (2.14)$$

with initial condition $Y_0 = (\frac{S_0^{1-\beta}}{1-\beta} + \frac{\rho}{\alpha}(\sigma_\Delta - \sigma_0))^2$.

3. Let τ be the stopping time for which process Y hits zero, i.e. $\tau = \inf\{\nu(s) | Y_{\nu(s)} = 0\}$, the 'stopped' process Y reads

$$Y_{\nu(t) \wedge \tau} = Y_0 + 2 \int_0^{\nu(t) \wedge \tau} \sqrt{Y_{\nu(s)}} dU_{\nu(s)} + \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)} (\nu(t) \wedge \tau). \quad (2.15)$$

Proof. In order to derive the conditional dynamics of the SABR model in Eq. (2.1), we first integrate the SDE for the volatility σ_t :

$$\int_0^\Delta d\sigma_s = \alpha \int_0^\Delta \sigma_s dW_{2,s} \Rightarrow \int_0^\Delta \sigma_s dW_{2,s} = \frac{1}{\alpha}(\sigma_\Delta - \sigma_0). \quad (2.16)$$

When conditioning on the volatility level σ_Δ , the above identity becomes a constant. It plays an important role in the following derivation.

Based on the arguments in Section 2.1, the application of Itô's lemma to $X = S^{1-\beta}/(1-\beta)$ results in

$$X_\Delta = X_0 + \int_0^\Delta \sigma_s dW_{1,s} - \int_0^\Delta \frac{\beta\sigma_s^2}{(2-2\beta)X_s} ds. \quad (2.17)$$

We now employ the Cholesky decomposition of the two correlated Brownian motions,

$$dW_{1,t} = \rho dW_{2,t} + \sqrt{1-\rho^2} dU_t, \quad dW_{2,t} = dW_{2,t}.$$

After substitution of the Cholesky decomposition into Eq. (2.17), we arrive at

$$X_\Delta = X_0 + \frac{\rho}{\alpha}(\sigma_\Delta - \sigma_0) + \sqrt{1-\rho^2} \int_0^\Delta \sigma_s dU_s - \int_0^\Delta \frac{\beta\sigma_s^2}{(2-2\beta)X_s} ds, \quad (2.18)$$

where we used Identity (2.16).

In order to prove the second claim in the Result, we write Eq. (2.18) as,

$$\begin{aligned} d\tilde{X}_t &= \sqrt{1-\rho^2}\sigma_t dU_t - \frac{\beta\sigma_t^2}{(2-2\beta)\tilde{X}_t} dt, \quad \text{with:} \\ \tilde{X}_0 &= X_0 + \frac{\rho}{\alpha}(\sigma_\Delta - \sigma_0), \end{aligned}$$

where the notation \tilde{X}_t denotes process X_t with a shifted initial condition. Despite the difference in notation, \tilde{X}_t and X_t are the same processes.

We define the variable transformation $Y = \tilde{X}^2$, which, after applying Itô's lemma, gives

$$\begin{aligned} dY_t &= 2\tilde{X}_t d\tilde{X}_t + d\tilde{X}_t^2 \\ &= 2\sqrt{Y_t}\sqrt{1-\rho^2}\sigma_t dU_t + \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)}(1-\rho^2)\sigma_t^2 dt. \end{aligned} \quad (2.19)$$

Due to the independence of the Brownian motions U_t and $W_{2,t}$, the integral $\int_0^t \sigma_s dU_s$ is a Gaussian distribution with mean zero and variance $\int_0^t \sigma_s^2 ds$.

We now consider the time-change $\nu(t) = (1-\rho^2) \int_0^t \sigma_s^2 ds$. A Brownian motion under this 'clock' will have the same distribution as $\sqrt{1-\rho^2} \int_0^t \sigma_s dU_s$, i.e.

$$U_{\nu(t)} = \int_0^{\nu(t)} dU_s = \sqrt{1-\rho^2} \int_0^t \sigma_s dU_s.$$

We substitute the time-changed Brownian motion into Eq. (2.19), which gives us

$$dY_{\nu(t)} = 2\sqrt{Y_{\nu(t)}} dU_{\nu(t)} + \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)} d\nu(t),$$

i.e. a time-changed squared Bessel process of dimension $\delta = \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)}$, starting at

$$Y_0 = \tilde{X}_0^2 = \left(X_0 + \frac{\rho}{\alpha}(\sigma_\Delta - \sigma_0)\right)^2.$$

Inclusion of a stopping time, $\tau = \inf\{\nu(s) | Y_{\nu(s)} = 0\}$, to the second result will prove the third claim. □

PROPOSITION 2.2 (Cumulative Distribution for Conditional SABR Process). *For some S_0 , strictly larger than 0, the conditional cumulative distribution of S_Δ with an absorbing boundary at $S_t = 0$ given σ_Δ and $\int_0^\Delta \sigma_s^2 ds$ reads*

$$\Pr(S_\Delta \leq K | S_0 > 0, \sigma_\Delta, \int_0^\Delta \sigma_s^2 ds) = 1 - \chi^2(a; b, c), \quad (2.20)$$

where

$$\begin{aligned} a &= \frac{1}{v(\Delta)} \left(\frac{S_0^{1-\beta}}{(1-\beta)} + \frac{\rho}{\alpha} (\sigma_\Delta - \sigma_0) \right)^2, & b &= 2 - \frac{1 - 2\beta - \rho^2(1-\beta)}{(1-\beta)(1-\rho^2)}, \\ c &= \frac{K^{2(1-\beta)}}{(1-\beta)^2 v(\Delta)}, & v(\Delta) &= (1-\rho^2) \int_0^\Delta \sigma_s^2 ds. \end{aligned} \quad (2.21)$$

$\chi^2(x; \delta, \lambda)$ is again the noncentral chi-square cumulative distribution function.

Proof. Given that $Y_{v(t)}$ is a time-changed Bessel process, we substitute parameter δ and the non-centrality parameter Y_0 into the distribution function presented in Result 2.3. \square

Note that the condition $S_0 > 0$ is crucial, because the paths that reach zero should stay in zero, due to the stopping time τ defined in Result 2.4.

REMARK. *It was argued by Andersen [4] that the continuous time process, S_t , will be a martingale with*

$$E[S_{t+\Delta} | S_t] = S_t < \infty.$$

The equivalent discrete-time process, \hat{S}_t , generated by the unbiased simulation scheme may not satisfy the martingale condition,

$$E[\hat{S}_{t+\Delta} | \hat{S}_t] \neq \hat{S}_t.$$

The net drift, away from the martingale, is visible for parameter sets with small β and close-to-zero rates. However, this combination of parameters does not appear in practical applications as it gives rise impractical implied volatility levels. For practical SABR parameters, the martingale bias is very small and can be controlled by reducing the size of the time step.

3 The discretization scheme for the SABR model

In this section we will present the unbiased simulation scheme for the SABR model. Before that, we review some existing path simulation schemes for the stochastic volatility model. We will denote time discrete approximations to S_t and σ_t by \hat{S}_t and $\hat{\sigma}_t$, respectively.

3.1 Taylor based time discrete approximation schemes

The basic first-order Taylor approximation scheme for (2.1) takes the following form:

$$\begin{aligned} \hat{S}_\Delta &= \hat{S}_0 + \hat{\sigma}_0 \hat{S}_0^\beta Z_1 \sqrt{\Delta}, \\ \hat{\sigma}_\Delta &= \hat{\sigma}_0 \exp\left(-\frac{1}{2}\alpha^2 \Delta + \alpha Z_2 \sqrt{\Delta}\right), \end{aligned} \quad (3.1)$$

with Z_1 and Z_2 two correlated standard normal random variables, i.e. $E[Z_1 Z_2] = \rho dt$.

This Euler scheme represents an $O(\sqrt{\Delta})$ -accurate Taylor approximation for the asset price process, S_t . To reduce the bias introduced by the first-order approximation, Elerian [10] suggests using a transition density derived from the scheme due to Milstein [25]. Including higher-order expansion terms gives us the Milstein scheme:

$$\begin{aligned} \hat{S}_\Delta &= \hat{S}_0 + \hat{\sigma}_0 \hat{S}_0^\beta Z_1 \sqrt{\Delta} + \frac{1}{2} \beta \hat{\sigma}_0^2 \hat{S}_0^{2\beta-1} (Z_1^2 \Delta - \Delta), \\ \hat{\sigma}_\Delta &= \hat{\sigma}_0 \exp\left(-\frac{1}{2}\alpha^2 \Delta + \alpha Z_2 \sqrt{\Delta}\right). \end{aligned} \quad (3.2)$$

A common problem shared by the Taylor-based approximation schemes is the possibility of generating negative asset prices, \hat{S}_t , that give rise to financially meaningless solutions. A remedy for the negative prices is to transform the SDE to logarithmic coordinates using Itô's lemma [3]. For the CEV asset price process the log-Euler scheme is defined as:

$$\begin{aligned}\hat{S}_\Delta &= \hat{S}_0 \exp\left(-\frac{1}{2}\hat{\sigma}_0^2 S_0^{2\beta-2}\Delta + \hat{\sigma}_0 \hat{S}_0^{\beta-1} Z_1 \sqrt{\Delta}\right), \\ \hat{\sigma}_\Delta &= \hat{\sigma}_0 \exp\left(-\frac{1}{2}\alpha^2 \Delta + \alpha Z_2 \sqrt{\Delta}\right).\end{aligned}\quad (3.3)$$

Whereas the log-Euler scheme preserves positivity of the asset price process, numerical experiments show that the scheme may become unstable for specific time-step sizes [3]. The instabilities occur because the diffusion terms in Eq. (3.3) approach infinity quickly when \hat{S}_t reaches zero (see some results in Table 2).

Table 2: Percentages of Taylor-based simulation experiments with failure for the CEV model, with different step-sizes over a 5 years interval. The parameters are $\beta = 0.3$, $\sigma = 0.3$ and $S_0 = 4\%$.

Issue	Stepsize	Euler	Milstein	Log-Euler
Negativity	$\Delta = 0.5$	85%	49%	0%
	$\Delta = 0.25$	88%	56%	0%
	$\Delta = 0.125$	94%	59%	0%
Infinity	$\Delta = 0.5$	0%	34%	96%
	$\Delta = 0.25$	0%	31%	96%
	$\Delta = 0.125$	0%	37%	95%

3.2 Local linearization

A second approximation is proposed in Shoji & Ozaki [32] and referred to as *local linearization approach*. The original approach is based a model with constant volatility, which is true for the target SDE, after the variable transformation. Shoji & Ozaki suggest to apply Itô's lemma to the drift of the model SDE:

$$dS_t = \bar{\mu}(S_t)dt + \bar{\sigma}dW_t,$$

with $\bar{\sigma}$ constant and fixed initial value, x_0 ,

$$d\bar{\mu}(S_t) = \frac{1}{2}\bar{\sigma}^2 \bar{\mu}''(S_t)dt + \bar{\mu}'(S_t)dS_t.$$

Using the first-order Taylor expansion, the approximate drift $\bar{\mu}$ is defined by:

$$\bar{\mu}(x_\Delta) = \bar{\mu}(x_0) + \bar{\mu}'(x_0)(x_\Delta - x_0) + \frac{1}{2}\bar{\sigma}^2 \bar{\mu}''(x_0)(t - s).$$

The approximate density would then be obtained from $d\tilde{S}_t = \bar{\mu}(\tilde{S}_t)dt + \bar{\sigma}dW_t$, which is in the form of an Ornstein-Uhlenbeck process. From this approximation, one can easily obtain its transition density from the analytic results for Ornstein-Uhlenbeck processes:

$$p_{\text{Sh\&Oz}}(x_\Delta, t; x_0, s) = \phi(x_\Delta; \bar{\mu}, \bar{\sigma}^2) \quad (3.4)$$

where ϕ refers to the probability density function of a normal distribution, and

$$\begin{aligned}\bar{\mu} &= x_0 + \frac{\bar{\mu}(x_0)}{\bar{\mu}'(x_0)}L + \frac{\bar{\sigma}^2 \bar{\mu}''(x_0)}{2[\bar{\mu}'(x_0)]^2} (L - \bar{\mu}'(x_0)\Delta), \\ \bar{\sigma}^2 &= \frac{\bar{\sigma}^2}{2\bar{\mu}'(x_0)} \left(e^{2\bar{\mu}'(x_0)\Delta} - 1 \right), \quad L = e^{\bar{\mu}'(x_0)\Delta} - 1.\end{aligned}$$

This approximation seems attractive due to the closed-form transition density formula. However, it suffers from several numerical issues; the first one being the negativity problem. Since density (3.4)

is normal, it has a non-zero probability to become negative-valued. Secondly, the drift term, $\bar{\mu}$, and its derivatives, i.e.

$$\bar{\mu}(x_0) = -\frac{\beta\bar{\sigma}^2}{(2-2\beta)x_0}, \quad \bar{\mu}'(x_0) = \frac{\beta\bar{\sigma}^2}{(2-2\beta)x_0^2}, \quad \bar{\mu}''(x_0) = -\frac{2\beta\bar{\sigma}^2}{(2-2\beta)x_0^3},$$

diverge for x_0 close to zero. Due to these drawbacks, we do not consider this method here.

3.3 Exact scheme of Broadie and Kaya

Broadie and Kaya [6] proposed the BK scheme, an exact simulation for the Heston model. Although the Heston dynamics are different from the SABR dynamics, the exact simulation concept serves as the basis to construct the discrete approximation schemes for the SABR model here.

The BK scheme is based on sampling σ_Δ from its distribution function. Given σ_Δ (and σ_0), a sample from $\int_0^\Delta \sigma_s^2 ds$ is drawn. Conditional on σ_Δ and $\int_0^\Delta \sigma_s^2 ds$, process $\ln S_\Delta$ is from a Gaussian distribution in the Heston model.

In some more detail, with $V_t = \sigma_t^2$ the Heston stochastic volatility [14] process can be described as:

$$\begin{aligned} dS_t &= \sqrt{V_t} S_t dW_{1,t}, \\ dV_t &= \kappa(\theta - V_t)dt + \eta\sqrt{V_t}dW_{2,t} \end{aligned} \quad (3.5)$$

with $E[dW_{1,t}dW_{2,t}] = \rho dt$. To obtain a bias-free scheme, first the SDE for the volatility is integrated, i.e.,

$$\int_0^\Delta \sqrt{V_s}dW_{2,s} = \frac{1}{\eta}(V_\Delta - V_0 - \kappa\theta\Delta + \kappa \int_0^\Delta V_s ds). \quad (3.6)$$

Application of the Cholesky decomposition, i.e.,

$$dW_{1,t} = \rho dW_{2,t} + \sqrt{1-\rho^2}dU_t, \quad dW_{2,t} = dW_{2,t},$$

gives us for $\ln S_t$:

$$d \ln S_t = -\frac{1}{2}V_t dt + \rho\sqrt{V_t}dW_{2,t} + \sqrt{1-\rho^2}\sqrt{V_t}dU_t,$$

with Brownian motion U_t , independent of $W_{2,t}$. In integral form, we then obtain

$$\begin{aligned} \ln S_\Delta &= \ln S_0 - \frac{1}{2} \int_0^\Delta V_s ds + \rho \int_0^\Delta \sqrt{V_s}dW_{2,s} + \sqrt{1-\rho^2} \int_0^\Delta \sqrt{V_s}dU_s \\ &= \ln S_0 + \frac{\rho}{\eta}(V_\Delta - V_0 - \kappa\theta\Delta) + \left(\frac{\rho\kappa}{\eta} - \frac{1}{2}\right) \int_0^\Delta V_s ds + \sqrt{1-\rho^2} \int_0^\Delta \sqrt{V_s}dU_s. \end{aligned} \quad (3.7)$$

Due to the independence of U_t and $W_{2,t}$, the Itô integral $\int_0^\Delta V_s dU_s$ is Gaussian with mean zero and variance $\int_0^\Delta V_s ds$. It is easy to see from Eq. (3.7) that $\ln S_\Delta$ is normally distributed, conditional on V_Δ and $\int_0^\Delta V_s ds$. By aggregation of all conditional Gaussian distributed samples, we obtain the desired distribution for the stochastic volatility model.

For the Heston model, the distribution of V_Δ is known in closed form, but the conditional distribution of $\int_0^\Delta V_s ds$ is not known explicitly. Broadie and Kaya [6] derive the characteristic function, which is based on two modified Bessel functions that contain infinite series expansions. A numerical Fourier-inversion step is then necessary to generate the desired conditional cumulative distribution function. The evaluation of the characteristic function for $\int_0^\Delta V_s ds$ as well as the required Fourier inversions require significant computational effort. The implementation of these steps has to be done with great care to avoid bias from the numerical inversion.

3.4 An unbiased scheme for SABR simulation

By utilizing the analytic results for the conditional SABR process presented in Proposition 2.2, and the above mentioned technique of ‘mixing conditional distributions’, we present an unbiased discretization scheme for the SABR model. We start to simulate the conditional SABR process, which is a space-transformed squared Bessel process with an absorbing boundary at zero, and then mix the conditional process by the joint dynamics of the terminal volatility and integrated variance.

Despite the fact that the volatility σ_Δ can easily be sampled from a log-normal distribution, it is not straightforward to sample the integrated variance. It is also challenging to simulate a CEV process exactly. Hence, we devote the following two subsections to these issues.

3.4.1 Sampling the conditional CEV process

As discussed in Section 2.2, the CEV process is a space-transformed squared Bessel process, whose distribution function is known in closed-form as the noncentral chi-square distribution. For large initial asset prices, S_0 , the probability of hitting zero is almost zero, i.e. $\Pr(\inf\{t|S_t = 0\} < \Delta) \rightarrow 0$, and the distribution function approaches an ordinary noncentral chi-square distribution:

$$\Pr(S_\Delta \leq K) = 1 - \chi^2(a; b, c) = \chi^2(c; 2 - b, a) + \Pr(\inf\{t|S_t = 0\} < \Delta) \approx \chi^2(c; 2 - b, a). \quad (3.8)$$

where a , b and c are as defined in Proposition 2.2.

It is well-known (see Johnson et al. [17], pp. 450) that the noncentral chi-square distribution approaches a Gaussian distribution as the non-centrality parameter goes to infinity. In [1] it was stated that the noncentral chi-square distribution with a sufficiently large non-centrality parameter can be accurately approximated by a quadratic function of Gaussian variables.

Andersen’s Quadratic Exponential (QE) scheme does not perform satisfactory for small values of S_0 , as the moment-matching method then breaks down. This is due to the fact that for small values of S_0 , the probability of reaching zero is high, so that the approximation in Eq. (3.8) does not hold and the analytic moments for the distribution $1 - \chi^2(a; b, c)$ are not known. For these small values, we therefore propose to use a Newton-type root finding method to invert the distribution function (2.20) directly.

More specifically, we determine a value, c^* , which solves the equation $1 - \chi^2(a; b, c^*) - U = 0$ with a high accuracy. To compute the noncentral chi-square distribution, we use Schroder’s [31] recurrence formula, in which the evaluation of an infinite double sum of gamma functions is required. This computation takes only a small fraction of the costs for the computation of the original series [31].

The partial derivative of the cumulative distribution is the probability density function (PDF) which is known analytically as the transition density for a squared Bessel process, given in Eq. (2.6). Using this derivative information gives a substantial enhancement of the computational performance².

Moment-matched quadratic Gaussian approximation

The mean and the variance of the noncentral chi-square distribution, $\chi^2(x; k, \lambda)$, are $k + \lambda$ and $2(k + 2\lambda)$, respectively. Here we determine, as in Proposition 5 in [1], the values of the relevant parameters by moment matching.

RESULT 3.1 (Moment-Matched Quadratic Gaussian Approximation [1]). *We denote the mean and variance of a noncentral chi-square distribution, $\chi^2(x; k, \lambda)$, by $m := k + \lambda$ and $s^2 := 2(k + 2\lambda)$, and we define $\psi := s^2 / m^2$. With*

$$e^2 = 2\psi^{-1} - 1 + \sqrt{2\psi^{-1} - 1} \sqrt{2\psi^{-1} - 1} \geq 0 \quad \text{and} \quad d = \frac{m}{1 + e^2}.$$

A non-central chi-square distributed random variable Y is computed by

$$Y = d(e + Z)^2 \quad Z \sim N(0, 1),$$

where $E[Y] = m$ and $\text{Var}[Y] = s^2$.

²When the gradient information is not supplied, the algorithm requires many extra function evaluations of the noncentral chi-square distribution function, which is generally expensive to compute.

In our case, we have

$$k = \frac{1 - 2\beta - \rho^2(1 - \beta)}{(1 - \beta)(1 - \rho^2)}, \quad \lambda = \frac{1}{\nu(\Delta)} \left(\frac{S_0^{1-\beta}}{(1 - \beta)} + \frac{\rho}{\alpha}(\sigma_\Delta - \sigma_0) \right)^2,$$

with $\nu(\Delta) = (1 - \rho^2) \int_0^\Delta \sigma_s^2 ds$. We then compute $S_\Delta^{2(1-\beta)}/(1 - \beta)^2 \nu(\Delta)$ by the quadratic normal approximation:

$$\frac{S_\Delta^{2(1-\beta)}}{(1 - \beta)^2 \nu(\Delta)} = d(e + Z)^2 \quad \Rightarrow \quad S_\Delta = \left((1 - \beta)^2 \nu(\Delta) \cdot d(e + Z)^2 \right)^{\frac{1}{2(1-\beta)}}, \quad (3.9)$$

with $Z \sim N(0, 1)$.

The constants d and e are as in Result 3.1.

Direct inversion scheme

As mentioned earlier, the quadratic Gaussian approximation is accurate only if S_0 is sufficiently large, or, equivalently, when the probability of absorption is small. For small values of S_0 , we invert the distribution function (2.20) directly by a Newton-type method. As in Andersen [1], variable $\psi = s^2/m^2$, defined in Result 3.1, is used as the threshold level to determine which algorithm (either the moment-matched Quadratic Gaussian or the direct inversion) is to be employed.

We first assume that the integrated variance, $\nu(\Delta) = (1 - \rho^2) \int_0^\Delta \sigma_s^2 ds$, has been determined. An algorithm to sample the conditional CEV process in terms of the transformed squared Bessel process with an absorbing boundary at zero, starting at time 0 until Δ , then reads:

1. Compute $a = \frac{1}{\nu(\Delta)} \left(\frac{S_0^{1-\beta}}{(1-\beta)} + \frac{\rho}{\alpha}(\sigma_\Delta - \sigma_0) \right)^2$ and $b = 2 - \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)}$ by Result 2.2;
2. Draw a (vector of) uniform random numbers, U ;
3. Compute the absorption probability $\Pr(S_\Delta = 0|S_0)$ by Eq. (2.10);
 - (a) **If** $S_0 = 0$: $S_\Delta = 0$ and return;
 - (b) **Else if** $U < \Pr(S_\Delta = 0|S_0)$: $S_\Delta = 0$ and return;
 - (c) **Otherwise**: Go to the next step;
4. Compute $\psi := s^2/m^2$ with $m := k + \lambda$ and $s^2 := 2(k + 2\lambda)$;
5. Select a threshold value, $\psi^{\text{thres}} \in [1, 2]$, as in [1]. Here we set $\psi^{\text{thres}} = 2$ for numerical efficiency;
6. **If** $\{0 < \psi \leq \psi^{\text{thres}}\} \cap \{m \geq 0\}$: We sample S_Δ by Eq. (3.9);
7. **Otherwise** if $\psi > \psi^{\text{thres}}$ or $\{m < 0\} \cap \{0 < \psi \leq \psi^{\text{thres}}\}$: We determine the root c^* of the equation $H(a, b, c) := 1 - \chi^2(a; b, c) - U = 0$ with initial guess $c_0 = a$, and repeat the Newton method until the prescribed tolerance ϵ is reached:

$$c_{n+1} = c_n - \frac{H(a, b, c_n)}{\bar{q}(a, b, c_n)},$$

where

$$\bar{q}(a, b, c) = \frac{1}{2} \left(\frac{c}{a} \right)^{\frac{b-2}{4}} \exp \left(-\frac{a+c}{2} \right) I_{|\frac{b-2}{2}|}(\sqrt{ac}).$$

The desired random number, c^* , is from a squared Bessel distribution with an absorbing boundary at zero.

8. We apply the inverse coordinate transform to recover the random numbers in asset price (or physical) space:

$$S_\Delta = \left(c^* (1 - \beta)^2 \nu(\Delta) \right)^{\frac{1}{2-2\beta}}.$$

This root finding method consists of only basic operations, so that the whole procedure can be vectorized and a vector of uniform random numbers can be processed simultaneously.

Enhanced direct inversion procedure

Although the above method is accurate, it appears to be rather slow, due to a significant number of evaluations of the (expensive) noncentral chi-square CDF. In order to speed up the procedure, we determine an *accurate initial solution* by a cheap numerical procedure. With an accurate initial solution, this root finding procedure will converge in only a few iterations.

Yuan [34] gives numerical evidence for the fact that the noncentral chi-square distribution ‘converges’ to a normal distribution when non-centrality parameter c or random number a increase in value. A number of normal approximations to the noncentral chi-square distribution has been developed, see Johnson and Kotz [16] for a review. A particular accurate approximation is derived by Sankaran [29]:

$$\chi^2(a; b, c) \sim N \left(\frac{1 - hp(1 - h + \frac{1}{2}(2 - h)mp) - (\frac{a}{b+c})^h}{h\sqrt{2p(1 + mp)}} \right) \quad (3.10)$$

where N denotes the normal cumulative density function and

$$\begin{aligned} h &= 1 - \frac{2(b+c)(b+3c)}{3(b+2c)^2}, \\ p &= \frac{b+2c}{(b+c)^2}, \\ m &= (h-1)(1-3h). \end{aligned} \quad (3.11)$$

This approximation consists of basic functions only, e.g. square roots, powers and normal distribution functions, that can be executed quickly on modern hardware. The approximation is sufficiently accurate for a wide range of parameters, but it varies across different sets of parameters. Especially for small values of parameter a in Eq. (2.20), this approximation is less accurate.

To illustrate the performance of Sankaran’s approximation, we consider two different test cases with different values for parameter a (keeping the other parameters the same). Figure 2 presents the results. In both plots, the difference between the exact curve and Sankaran’s approximation is visible, but for the larger value of a the discrepancy is substantially smaller. For the purpose of generating an initial solution for the second stage of the root finding method, however, the approximation quality is fully satisfactory.

After the inversion of a vector of uniform variables, U , into normal variables, $X = N^{-1}(U)$, we find a vector of roots, c , which solve the following equation

$$G(a, b, c) := \frac{1 - hp(1 - h + \frac{1}{2}(2 - h)mp) - (\frac{a}{b+c})^h}{h\sqrt{2p(1 + mp)}} - X = 0, \quad (3.12)$$

with h , p and m as defined in Eq. (3.11). This can be performed very efficiently by Newton’s method with analytic derivative information.

In the next stage, we choose the c_0 in Step 7 of the *direct inversion scheme* to be the solution of Eq. (3.12) and execute that step. The result is a significant improvement in the number of function evaluations required for Newton’s root finding algorithm to converge.

3.5 The integrated variance

The sampling of the conditionally integrated variance process may be an expensive procedure. There are several methods in the literature to *approximate* the integrated variance process, $\int_0^\Delta V_s ds | V_0, V_\Delta$, or, equivalently, $\int_0^\Delta \sigma_s^2 ds | \sigma_0, \sigma_\Delta$. Notable examples include the Fourier inversion technique in [6] and the drift interpolation in [1]. In the present paper, we propose to use an approximation based on a *moment-matched drift interpolation technique*. Instead of dealing with the distribution of the quantity $\int_0^\Delta \sigma_s^2 ds | \sigma_0, \sigma_\Delta$, we match the first two conditional moments of $\int_0^\Delta \sigma_s^2 ds$ given σ_0 (and σ_Δ) in a log-normal distribution, and sample the distribution of $\int_0^\Delta \sigma_s^2 ds$ from the moment-matched log-normal distribution.

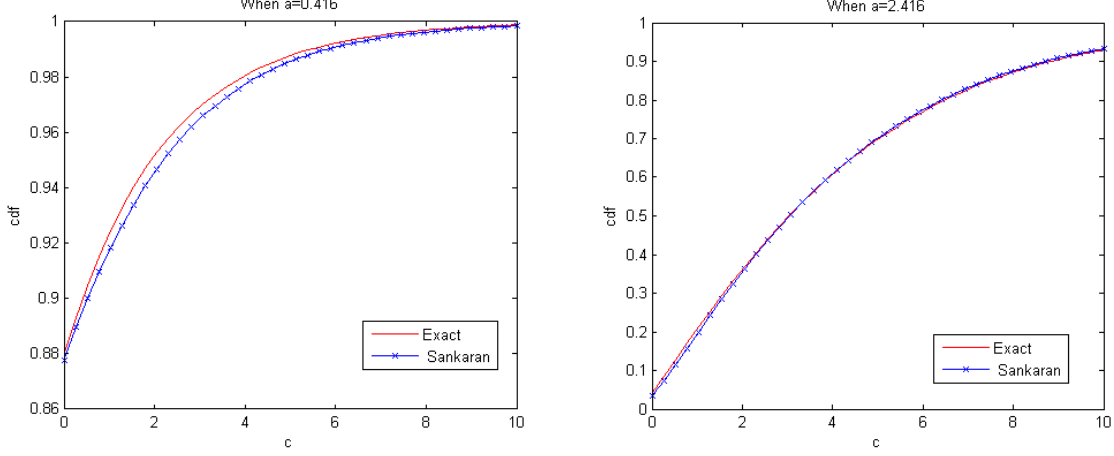


Figure 2: Comparison of the quality of Sankaran's approximation for two sets of parameters. Degree of freedom parameter b is set to 0.423 for both cases, but parameter a is lower in the left-side plot, $a = 0.416$ than in the right-side plot, $a = 2.416$.

3.5.1 Small disturbance expansion

In a first step, we derive the conditional moments of $\int_0^\Delta \sigma_s^2 ds$, given σ_Δ (and σ_0), by the *small disturbance expansion*, proposed by Kunitomo [20]. The small disturbance expansion method is closely related to the stochastic Taylor expansion, which is especially accurate when the quantity $\alpha^2 \Delta$ is small. In order to apply the small disturbance expansion technique, we first reformulate the log-normal volatility process by introducing a small parameter $0 < \epsilon \ll 1$ in the diffusion coefficient, so that $\alpha = \epsilon \tilde{\alpha}$:

$$\sigma_t^{(\epsilon)} = \sigma_0 + \epsilon \tilde{\alpha} \int_0^t \sigma_s^{(\epsilon)} dW_{2,s}. \quad (3.13)$$

We construct an expansion of $\sigma_t^{(\epsilon)}$ around σ_0 by $\epsilon \rightarrow 0$.

PROPOSITION 3.1. *For stochastic integral (3.13), we have the formal small disturbance expansion*

$$\sigma_t^{(\epsilon)} = \sigma_0 + \epsilon \left. \frac{\partial \sigma_t^{(\epsilon)}}{\partial \epsilon} \right|_{\epsilon=0} + \frac{1}{2} \epsilon^2 \left. \frac{\partial^2 \sigma_t^{(\epsilon)}}{\partial \epsilon^2} \right|_{\epsilon=0} + \frac{1}{6} \epsilon^3 \left. \frac{\partial^3 \sigma_t^{(\epsilon)}}{\partial \epsilon^3} \right|_{\epsilon=0} + \frac{1}{24} \epsilon^4 \left. \frac{\partial^4 \sigma_t^{(\epsilon)}}{\partial \epsilon^4} \right|_{\epsilon=0} + O(\epsilon^5),$$

where

$$\begin{aligned} \left. \frac{\partial \sigma_t^{(\epsilon)}}{\partial \epsilon} \right|_{\epsilon=0} &= \sigma_0 \tilde{\alpha} \int_0^t dW_{2,s} = \sigma_0 \tilde{\alpha} W_{2,t}, \\ \left. \frac{\partial^2 \sigma_t^{(\epsilon)}}{\partial \epsilon^2} \right|_{\epsilon=0} &= 2\sigma_0 \tilde{\alpha}^2 \int_0^t \int_0^{s_1} dW_{2,s_2} dW_{2,s_1} = \sigma_0 \tilde{\alpha}^2 (W_{2,t}^2 - t), \\ \left. \frac{\partial^3 \sigma_t^{(\epsilon)}}{\partial \epsilon^3} \right|_{\epsilon=0} &= 6\sigma_0 \tilde{\alpha}^3 \int_0^t \int_0^{s_1} \int_0^{s_2} dW_{2,s_3} dW_{2,s_2} dW_{2,s_1} = \sigma_0 \tilde{\alpha}^3 (W_{2,t}^3 - 3W_{2,t}t), \\ \left. \frac{\partial^4 \sigma_t^{(\epsilon)}}{\partial \epsilon^4} \right|_{\epsilon=0} &= 24\sigma_0 \tilde{\alpha}^3 \int_0^t \int_0^{s_1} \int_0^{s_2} \int_0^{s_3} dW_{2,s_4} dW_{2,s_3} dW_{2,s_2} dW_{2,s_1} = \sigma_0 \tilde{\alpha}^3 (W_{2,t}^4 - 6W_{2,t}^2 t + 3t^2). \end{aligned} \quad (3.14)$$

Proof. It is easy to see from Eq. (3.13) that $\sigma_s^{(0)} = \sigma_0$. Straightforward application of the deterministic calculus rules gives us

$$\frac{\partial \sigma_t^{(\epsilon)}}{\partial \epsilon} = \tilde{\alpha} \int_0^t \sigma_s^{(\epsilon)} dW_{2,s} + \epsilon \tilde{\alpha} \int_0^t \frac{\partial \sigma_t^{(\epsilon)}}{\partial \epsilon} dW_{2,s}.$$

We take the limit $\epsilon \rightarrow 0$, and get the expression for the first-order expansion term:

$$\left. \frac{\partial \sigma_t^{(\epsilon)}}{\partial \epsilon} \right|_{\epsilon=0} = \tilde{\alpha} \int_0^t \sigma_s^{(0)} dW_{2,s} = \sigma_0 \tilde{\alpha} \int_0^t dW_{2,s}.$$

The higher-order expansion terms follow by repeating these rules. \square

COROLLARY 3.2. *The formal small disturbance expansion of the integrated volatility $A_\Delta^{(\epsilon)} = \int_0^\Delta (\sigma_t^{(\epsilon)})^2 dt$ is now given by*

$$A_\Delta^{(\epsilon)} = A_\Delta^{(0)} + \epsilon A_\Delta^{(1)} + \epsilon^2 A_\Delta^{(2)} + \epsilon^3 A_\Delta^{(3)} + \epsilon^4 A_\Delta^{(4)} + O(\epsilon^5) \quad (3.15)$$

where

$$\begin{aligned} A_\Delta^{(0)} &= \sigma_0^2 \Delta, \\ A_\Delta^{(1)} &= \left. \frac{\partial A_\Delta^{(\epsilon)}}{\partial \epsilon} \right|_{\epsilon=0} = 2\sigma_0 \int_0^\Delta \frac{\partial \sigma_t^{(0)}}{\partial \epsilon} dt, \\ A_\Delta^{(2)} &= \left. \frac{\partial^2 A_\Delta^{(\epsilon)}}{\partial \epsilon^2} \right|_{\epsilon=0} = \int_0^\Delta \left(\left(\frac{\partial \sigma_t^{(0)}}{\partial \epsilon} \right)^2 + \sigma_0 \frac{\partial^2 \sigma_t^{(0)}}{\partial \epsilon^2} \right) dt, \\ A_\Delta^{(3)} &= \left. \frac{\partial^3 A_\Delta^{(\epsilon)}}{\partial \epsilon^3} \right|_{\epsilon=0} = \int_0^\Delta \left(\frac{\partial \sigma_t^{(0)}}{\partial \epsilon} \frac{\partial^2 \sigma_t^{(0)}}{\partial \epsilon^2} + \frac{1}{3} \sigma_0 \frac{\partial^3 \sigma_t^{(0)}}{\partial \epsilon^3} \right) dt, \\ A_\Delta^{(4)} &= \left. \frac{\partial^4 A_\Delta^{(\epsilon)}}{\partial \epsilon^4} \right|_{\epsilon=0} = \int_0^\Delta \left(\frac{1}{4} \left(\frac{\partial^2 \sigma_t^{(0)}}{\partial \epsilon^2} \right)^2 + \frac{1}{3} \frac{\partial \sigma_t^{(0)}}{\partial \epsilon} \frac{\partial^3 \sigma_t^{(0)}}{\partial \epsilon^3} + \frac{1}{12} \sigma_0 \frac{\partial^4 \sigma_t^{(0)}}{\partial \epsilon^4} \right) dt. \end{aligned}$$

Proof. The above result is a natural extension of Proposition 3.1. One can easily check the validity of the above expression by a derivation from a different point-of-departure, i.e.,

$$\begin{aligned} A_\Delta^{(\epsilon)} &= \int_0^\Delta (\sigma_t^{(\epsilon)})^2 dt \\ &= \int_0^\Delta \left(\sigma_0 + \epsilon \left. \frac{\partial \sigma_t^{(\epsilon)}}{\partial \epsilon} \right|_{\epsilon=0} + \frac{1}{2} \epsilon^2 \left. \frac{\partial^2 \sigma_t^{(\epsilon)}}{\partial \epsilon^2} \right|_{\epsilon=0} + \frac{1}{6} \epsilon^3 \left. \frac{\partial^3 \sigma_t^{(\epsilon)}}{\partial \epsilon^3} \right|_{\epsilon=0} + \frac{1}{24} \epsilon^4 \left. \frac{\partial^4 \sigma_t^{(\epsilon)}}{\partial \epsilon^4} \right|_{\epsilon=0} + O(\epsilon^5) \right)^2 dt. \end{aligned}$$

Expand the inner quadratic expression and collect the terms up to order $O(\epsilon^4)$ gives us the expansion above. \square

Based on the expressions above, the first conditional moment of $A_\Delta^{(\epsilon)}$, given $W_2(\cdot)$ at terminal time Δ , can be computed by substituting the expansion terms (3.14) in Identity (3.15):

$$\begin{aligned} E[A_\Delta^{(\epsilon)} | W_{2,\Delta}] &= E[A_\Delta^{(0)} + \epsilon A_\Delta^{(1)} + \epsilon^2 A_\Delta^{(2)} + \epsilon^3 A_\Delta^{(3)} + \epsilon^4 A_\Delta^{(4)} + O(\epsilon^5) | W_{2,\Delta}] \\ &= E \left[\sigma_0^2 \Delta + 2\epsilon \sigma_0^2 \tilde{\alpha} \int_0^\Delta W_{2,t} dt + \epsilon^2 \sigma_0^2 \tilde{\alpha}^2 \int_0^\Delta (2W_{2,t}^2 - t) dt + \epsilon^3 \sigma_0^2 \tilde{\alpha}^3 \int_0^\Delta \left(\frac{4}{3} W_{2,t}^3 - 2W_{2,t} t \right) dt \right. \\ &\quad \left. + \epsilon^4 \sigma_0^2 \tilde{\alpha}^4 \int_0^\Delta \left(\frac{2}{3} W_{2,t}^4 - 2W_{2,t}^2 t + \frac{t^2}{2} \right) dt + O(\epsilon^5) | W_{2,\Delta} \right] \\ &= \sigma_0^2 \Delta + 2\epsilon \sigma_0^2 \tilde{\alpha} \frac{1}{2} W_{2,\Delta} \Delta + \frac{1}{3} \epsilon^2 \sigma_0^2 \tilde{\alpha}^2 (2W_{2,\Delta}^2 \Delta - \frac{\Delta^2}{2}) + \frac{1}{3} \epsilon^3 \sigma_0^2 \tilde{\alpha}^3 (W_{2,\Delta}^3 \Delta - W_{2,\Delta} \Delta^2) \\ &\quad + \epsilon^4 \sigma_0^2 \tilde{\alpha}^4 \left(\frac{2}{15} W_{2,\Delta}^4 \Delta - \frac{3}{10} W_{2,\Delta}^2 \Delta^2 + \frac{2}{5} \Delta^3 \right) + O(\epsilon^5) \\ &= \sigma_0^2 \Delta \left\{ 1 + \epsilon \tilde{\alpha} W_{2,\Delta} + \frac{1}{3} \epsilon^2 \tilde{\alpha}^2 (2W_{2,\Delta}^2 - \frac{\Delta}{2}) + \frac{1}{3} \epsilon^3 \tilde{\alpha}^3 (W_{2,\Delta}^3 - W_{2,\Delta} \Delta) \right. \\ &\quad \left. + \frac{1}{5} \epsilon^4 \tilde{\alpha}^4 \left(\frac{2}{3} W_{2,\Delta}^4 - \frac{3}{2} W_{2,\Delta}^2 \Delta + 2\Delta^2 \right) \right\} + O(\epsilon^5), \quad (3.16) \end{aligned}$$

where the derivations involving the computation of the time integral of the Wiener processes can be found in Kahl [18] (Table 4.1). The expressions for $E[\int_0^\Delta W_{2,t}^4 dt | W_{2,\Delta}]$ and $E[\int_0^\Delta W_{2,t}^2 t dt | W_{2,\Delta}]$ have not been provided in [18], but have been derived by ourselves as a straightforward (but tedious) extension of the derivations in [18].

Finally, we substitute $\alpha = \epsilon \tilde{\alpha}$ back in the expressions, collect the first three terms of the expansion to approximate the solution of the original model.

REMARK. *The expansion for the conditional mean (3.16) has to be performed up to fourth order, because the third-order solution gives rise to negative values (for negative values of $W_{2,\Delta}$), see Figure 3. However the variance should be non-negative, and the same holds for the integrated variance.*

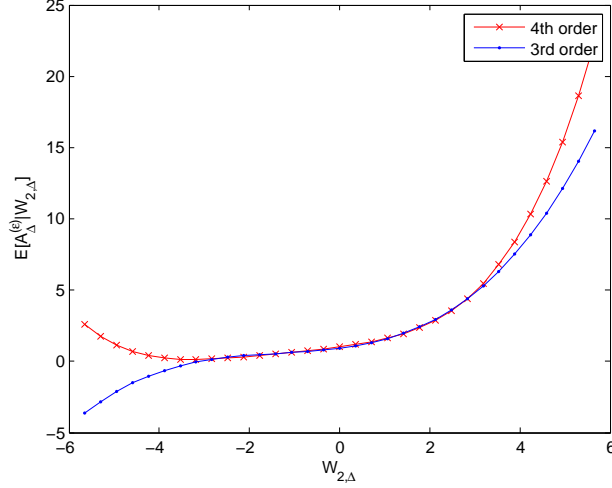


Figure 3: Illustration of formula (3.16) for the conditional mean $E[A_\Delta | W_{2,\Delta}]$; including the first three order terms versus the first four terms.

The computation of the conditional variance is however involved. Nevertheless, we identify the leading term as $E[\epsilon^2(A_\Delta^{(1)} - E[A_\Delta^{(1)}])^2 | W_{2,\Delta}]$:

$$\begin{aligned}
\text{Var}[A_\Delta^{(\epsilon)} | W_{2,\Delta}] &= E[(A_\Delta^{(\epsilon)} - E[A_\Delta^{(\epsilon)}])^2 | W_{2,\Delta}] \\
&= E[(A_\Delta^{(0)} + \epsilon A_\Delta^{(1)} + \frac{1}{2}\epsilon^2 A_\Delta^{(2)} + O(\epsilon^3) - E[A_\Delta^{(0)} + \epsilon A_\Delta^{(1)} + \frac{1}{2}\epsilon^2 A_\Delta^{(2)} + O(\epsilon^3)])^2 | W_{2,\Delta}] \\
&= E[(\epsilon(A_\Delta^{(1)} - E[A_\Delta^{(1)}]) + \frac{1}{2}\epsilon^2(A_\Delta^{(2)} - E[A_\Delta^{(2)}]) + O(\epsilon^3))^2 | W_{2,\Delta}] \\
&= E[\epsilon^2(A_\Delta^{(1)} - E[A_\Delta^{(1)}])^2 | W_{2,\Delta}] + O(\epsilon^3).
\end{aligned}$$

So, we find:

$$\begin{aligned}
E[\epsilon^2(A_\Delta^{(1)} - E[A_\Delta^{(1)}])^2 | W_{2,\Delta}] &= 4\sigma_0^4 \alpha^2 E[(\int_0^\Delta W_{2,t} dt)^2 | W_{2,\Delta}] - 4\sigma_0^4 \alpha^2 E[\int_0^\Delta W_{2,t} dt | W_{2,\Delta}]^2 \\
&= 4\sigma_0^4 \alpha^2 (E[(W_{2,\Delta}\Delta - \int_0^\Delta t dW_{2,t})^2 | W_{2,\Delta}] - E[(W_{2,\Delta}\Delta - \int_0^\Delta t dW_{2,t}) | W_{2,\Delta}]^2) \\
&= 4\sigma_0^4 \alpha^2 (W_{2,\Delta}^2 \Delta^2 - 2W_{2,\Delta}\Delta \frac{1}{2}W_{2,\Delta}\Delta + E[(\int_0^\Delta t dW_{2,t})^2 | W_{2,\Delta}] - \frac{1}{4}W_{2,\Delta}^2 \Delta^2) \\
&= 4\sigma_0^4 \alpha^2 (\frac{1}{4}W_{2,\Delta}^2 \Delta^2 + \frac{1}{12}\Delta^3 - \frac{1}{4}W_{2,\Delta}^2 \Delta^2) \\
&= \frac{1}{3}\sigma_0^4 \alpha^2 \Delta^3, \tag{3.17}
\end{aligned}$$

In the derivation above we have used the relation $d(W_t \cdot t) = W_t dt + t dW_t$.

The conditional variance does not depend on $W_{2,\Delta}$, which suggests that the conditional distribution of the integrated variance is 'shifted by the time Δ realization of the Brownian motion W_2 ', but its variance is not affected by $W_{2,\Delta}$.

3.5.2 Conditional moment-matched log-normal sampling scheme

Under common market conditions, i.e. $\sigma_0 < 1$ and $\alpha < 1$, the conditional variance is a very small value, for small Δ . In other words, the randomness of the conditional distribution of $A_\Delta^{(\epsilon)}$ is low in this situation. This suggests that one can accurately reproduce the conditional distribution of the integrated variance, $A_\Delta^{(\epsilon)}$, by an approximate distribution having the same mean and variance. One could choose a Gaussian distribution for this purpose, however, a disadvantage is that in that case

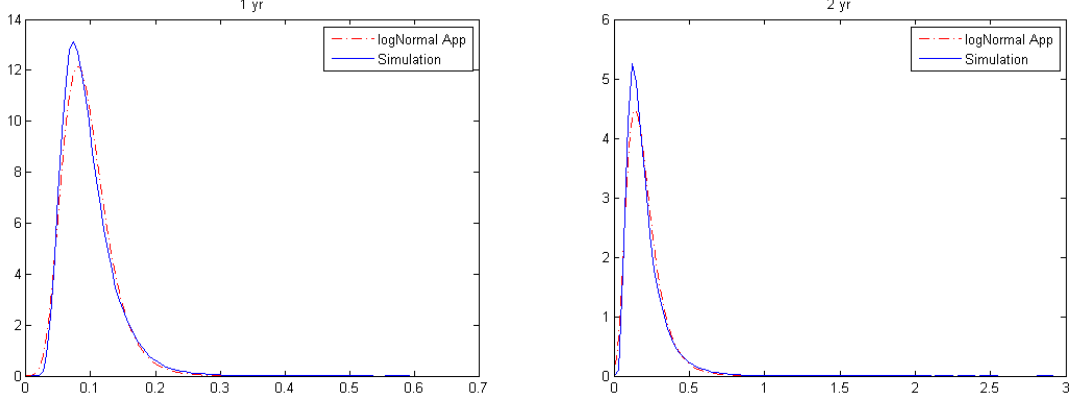


Figure 4: Quality of the approximation of the conditional moment-matched log-normal density compared to the true density of integrated variance, A_Δ . The parameters chosen are $\sigma_0 = 0.4$, $\alpha = 0.5$. In the LHS plot, $T = 1$; in the RHS plot $T = 2$.

large weights are assigned to the negative part of the real axis (whereas A_Δ cannot be negative). Therefore, we choose a (conditional) moment-matched *log-normally distributed* random variable to approximate the conditional distribution of $A_\Delta^{(\epsilon)}$.

More precisely, we denote the conditional mean and variance of $A_\Delta^{(\epsilon)}$ obtained from formula (3.16) and (3.17) by:

$$\begin{aligned} m &= \sigma_0^2 \Delta \left(1 + \alpha W_{2,\Delta} + \frac{1}{3} \alpha^2 (2W_{2,\Delta}^2 - \frac{\Delta}{2}) + \frac{1}{3} \alpha^3 (W_{2,\Delta}^3 - W_{2,\Delta} \Delta) + \frac{1}{5} \alpha^4 \left(\frac{2}{3} W_{2,\Delta}^4 - \frac{3}{2} W_{2,\Delta}^2 \Delta + 2\Delta^2 \right) \right) \\ v &= \frac{1}{3} \sigma_0^4 \alpha^2 \Delta^3, \end{aligned} \quad (3.18)$$

respectively. Then, we define a log-normal random variable, $\log(X) \sim N(\mu, \sigma)$, with mean and variance m and v , respectively, i.e., $E[X] = m$ and $\text{Var}[X] = v$. Parameters μ and σ can be easily obtained if the values of mean and variance are known:

$$\mu = \ln(m) - \frac{1}{2} \ln \left(1 + \frac{v}{m^2} \right), \quad \sigma^2 = \ln \left(1 + \frac{v}{m^2} \right).$$

The quality of this approximation is compared with simulation results for two time intervals, $\Delta = 1$ and $\Delta = 2$, in Figure 4. The shape of the density function for X is ‘flatter’ for larger values of Δ which reflects a higher uncertainty in the realizations of integrated variance A_Δ .

As $W_{2,\Delta}$ is normally distributed with variance Δ , it is straightforward to compute the joint density of $W_{2,\Delta}$ and $A_\Delta^{(\epsilon)}$ by recalling:

$$\Pr(A_\Delta, W_{2,\Delta}) \approx \Pr(A_\Delta^{(\epsilon)}, W_{2,\Delta}) = \Pr(A_\Delta^{(\epsilon)} | W_{2,\Delta}) \Pr(W_{2,\Delta}).$$

The above formula suggests that if we first draw normal random numbers, $W_{2,\Delta}$, and then sample the integrated variance from the conditional distribution, A_Δ given one realization of $W_{2,\Delta}$, the joint realization of A_Δ and $W_{2,\Delta}$ reconstructs the desired joint density.

3.6 Discretization scheme for a full SABR model

We combine the two components described above and arrive at the *unbiased scheme for the SABR model* with a correlation structure. We start from the SABR model as in Eq. (2.1) with calibrated parameters α , β and ρ . With an initial asset price and volatility at time 0, i.e. S_0 and σ_0 , we simulate the discrete paths with an absorbing boundary at zero for the next time point, Δ , as follows:

1. Draw samples from a normal distribution, $W_{2,\Delta} \sim N(0, \sqrt{\Delta})$, The volatility at time step Δ reads:

$$\sigma_\Delta = \sigma_0 \exp \left(\alpha W_{2,\Delta} - \frac{1}{2} \alpha^2 \Delta \right).$$

2. Compute the asymptotic conditional mean, m , and variance, v , for the integrated variance, $A_\Delta^{(\epsilon)}$, by

$$\begin{aligned} m &= \sigma_0^2 \Delta \left(1 + \alpha W_{2,\Delta} + \frac{1}{3} \alpha^2 (2W_{2,\Delta}^2 - \frac{\Delta}{2}) + \frac{1}{3} \alpha^3 (W_{2,\Delta}^3 - W_{2,\Delta} \Delta) + \frac{1}{5} \alpha^4 \left(\frac{2}{3} W_{2,\Delta}^4 - \frac{3}{2} W_{2,\Delta}^2 \Delta + 2\Delta^2 \right) \right), \\ v &= \frac{1}{3} \sigma_0^4 \alpha^2 \Delta^3 \end{aligned}$$

3. Compute the parameters of the moment-matched log-normal distribution by

$$\mu = \ln(m) - \frac{1}{2} \ln \left(1 + \frac{v}{m^2} \right), \quad \sigma^2 = \ln \left(1 + \frac{v}{m^2} \right).$$

4. Draw (a vector of) uniform random numbers, U_1 , and determine their inverse according to the log-normal distribution (defined by μ and σ):

$$A_\Delta = \exp(\sigma \cdot N^{-1}(U_1) + \mu).$$

5. Insert A_Δ and σ_Δ in the algorithm described in Sec. 3.4.1 to sample the conditional CEV process.

4 Numerical experiments

To analyze the validity and efficiency of our discretization schemes numerically, we price some European options based on the parameter sets in Table 3.

The first two test cases represent two limiting cases for the SABR model, i.e., with $\beta = 1$ and $\alpha = 0$, respectively. Our aim is here to check the efficiency of two components of the scheme proposed, i.e. the moment-matched log-normal integrated variance sampling scheme and the direct inversion scheme for the conditional CEV process. In Case I, $\beta = 1$, the asset price follows a basic geometric Brownian motion and there is no complication with the absorbing boundary at zero. The main pricing bias will then be from the moment-matching algorithm to sample the integrated variance. In Case II, we set $\alpha = 0$, so that we isolate the part which is related to the simulation of the CEV process. In this second test, we give special attention to the martingale property of the simulated path.

For Test Case I, we use the option prices generated by a small time step Euler Monte Carlo simulation as reference prices, whereas for Test Case II we use an analytic option pricing formula, derived by Schroder [31], as the reference.

Next to these two tests, we also consider two practically relevant, yet challenging, parameter sets, i.e. Cases III and IV, with parameters often observed in fixed income instrument pricing. Parameter set III is representative for a low interest rate market, as observed for example in Japan. This parameter set is often embedded in the popular power-reverse dual contract. In this set, the impact of the behavior at the zero boundary on the price should be clearly visible. Parameter set IV then describes volatile market conditions as a high volatility-of-volatility parameter gives rise to a heavy tailed distribution of the asset prices.

Table 3: Parameters in Test Cases I to IV for the numerical experiments.

	Set I	Set II	Set III	Set IV
S_0	4%	4%	0.5%	7%
σ_0	20%	20%	20%	40%
α	0.3	0	0.3	0.8
β	1	0.4	0.7	0.4
ρ	-0.5	0	-0.3	-0.6

The benchmark Monte Carlo schemes, for the unbiased scheme proposed here, are the Full Truncation Euler scheme from [24]. All Monte Carlo simulations have been performed simulating 10^5 paths.

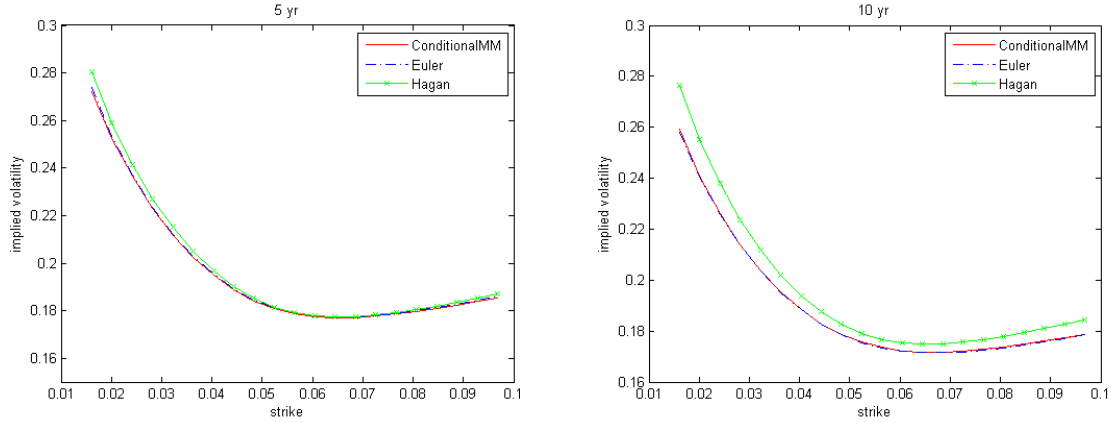


Figure 5: Result of the conditional moment-matching log-normal sampling scheme for the integrated variance, simulating 5 and 10 year call option prices under a double log-normal model. A comparison is made with the truncated Euler Monte Carlo scheme, and Hagan’s asymptotic SABR formula. Parameters used are $\alpha = 0.3$, $\rho = -0.5$, $S_0 = 0.04$, $\sigma_0 = 0.2$.

Table 4: Results of the unbiased SABR scheme with $\beta = 1$, Test Case I.

K	Δ	40%	80%	100%	120%	160%	200%
T= 5							
Euler	1/200	0.02077	0.00889	0.00512	0.00279	0.00083	0.00029
Unbiased	1/2	0.02079	0.00887	0.00510	0.00277	0.00082	0.00029
	1/4	0.02076	0.00890	0.00512	0.00279	0.00082	0.00029
Hagan		0.02083	0.00894	0.00514	0.00279	0.00082	0.00030
T= 10							
Euler	1/200	0.02198	0.01124	0.00758	0.00503	0.00234	0.00125
Unbiased	1/2	0.02196	0.01122	0.00756	0.00502	0.00233	0.00124
	1/4	0.02198	0.01124	0.00758	0.00504	0.00235	0.00126
Hagan		0.02230	0.01154	0.00781	0.00521	0.00248	0.00139

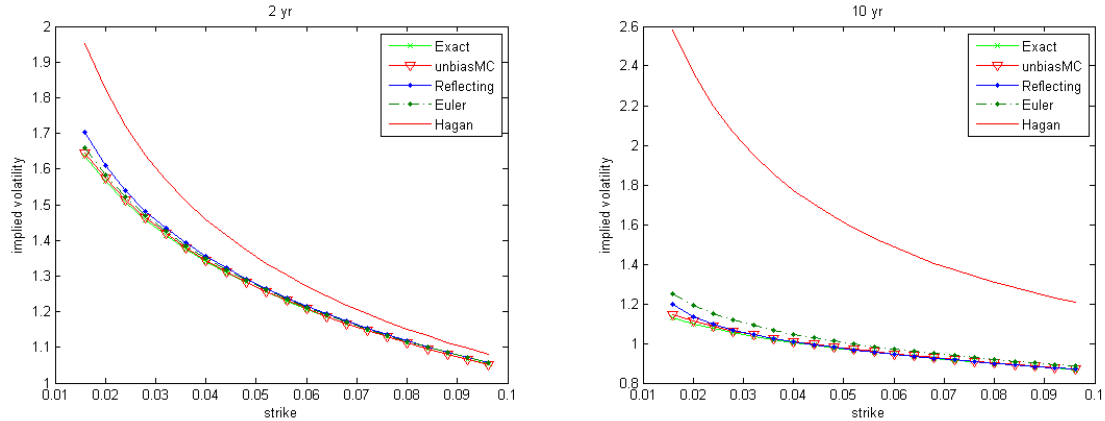


Figure 6: Comparison for the CEV process of the implied volatilities generated by different methods; Maturities are 2 and 10 years, and parameters $\alpha = 0.0$, $S_0 = 0.04$, $\sigma_0 = 0.2$ and $\beta = 0.4$.

4.1 Results for test case I

As in a double log-normal model (i.e. $\beta = 1$) $S_t = 0$ cannot be reached, the Euler scheme performs well. It is also reasonable to expect that the Euler Monte Carlo scheme with a sufficiently large number of time steps is stable and converges to the correct solution. Here, we perform two tests, with $T = 5$ and $T = 10$, respectively.

Table 4 shows that the conditional moment-matched integrated variance sampling scheme produces a significantly smaller bias for a practical number of time steps, like 2 or 4 time steps per year. The accuracy of the unbiased scheme is comparable to that of an Euler scheme with 50 times more time steps. To illustrate the accuracy of the conditional moment-matching scheme, we present the implied option volatilities from the moment-matching scheme with 4 time steps a year together with the 200 time steps Euler scheme and Hagan’s asymptotic formula in Figure 5. In the two figures, we observe that the difference between the implied volatilities from the 200 time steps Euler scheme and the unbiased SABR scheme with 4 time steps is negligible, whereas Hagan’s asymptotic formula produces a visible pricing error for the 5 year maturity and is inexact for the 10 year maturity.

4.2 Results for test case II

With $\alpha = 0$, the stochastic volatility part vanishes and the system reduces to a plain CEV model. An option pricing formula for the CEV model is known in closed form [31]. More recently, Lesniewski [22] provided a classification (and an explicit option pricing formula) for the CEV process with absorbing and with reflecting boundaries. Hence, we have an analytic reference value so that we can determine the accuracy of our discretization scheme and the price impact of the assumptions related to the boundary condition at zero. In this subsection, we also examine the martingale property of the discrete process in the algorithm proposed here.

In detail, we have implemented an Euler scheme with full truncation at zero, i.e. $\hat{S}_\Delta = \max(\hat{S}_\Delta, 0)$, the direct inversion scheme for the CEV process but with a reflecting boundary as well as our proposed unbiased scheme. We present the implied volatilities obtained by these numerical schemes and compare them to the exact CEV option pricing formula with absorbing boundaries in [22], in Figure 6. We also include Hagan’s formula in the comparison by choosing a very small volatility-of-volatility parameter and a small correlation coefficient³, i.e. $\alpha = 0.0001$ and $\rho = 0.0001$. In all experiments the Euler scheme consists of 50 times more time steps than the unbiased SABR scheme.

From Figure 6, we see that:

- The unbiased scheme has a low bias for all strikes and maturities. In most of the cases the implied volatilities obtained by the unbiased scheme are highly accurate when compared to the exact solution. The unbiased scheme is essentially free of bias with merely four time steps a year, whereas the Full Truncation Euler scheme requires more than 200 time steps to converge.

³This makes sense because the CEV model is a special case of the SABR model with a zero volatility-of-volatility parameter.

Table 5: Martingale properties of several discretization schemes for a pure CEV process; Test Case II. The Full Truncation Euler schemes considered are implemented with 50 times more time steps than the unbiased scheme.

Δ	Unbiased	Truncated Euler	Reflecting
		$T = 5$	
1/2	4%	4.05%	4.39%
1/4	4%	4.06%	4.22%
1/8	4%	4.07%	4.13%
1/16	4%	4.04%	4.07%
		$T = 10$	
1/2	4.01%	4.09%	4.47%
1/4	3.99%	4.05%	4.25%
1/8	4%	4.02%	4.16%
		$T = 15$	
1/2	3.99%	4.06%	4.47%
1/4	4.01%	4.04%	4.31%
1/8	4%	3.99%	4.15%

- The unbiased scheme with the reflecting boundary agrees very well with the exact solution and with the proposed unbiased scheme for strikes that are far away from zero. For small strike values there is a substantial miss-pricing by the direct inversion scheme with reflecting boundary and its pricing bias increases with maturity.
- Hagan’s formula is not an accurate approximation for the CEV model in the parameter range of small β (i.e. $\beta \leq 0.5$).
- The small time step Full Truncation Euler scheme performs reasonably well for all maturities, in particular for short maturities. However, we observe an upward shift in the implied volatility curve (the shift is larger for long maturities). This upward ‘bias’ is the result of the truncation and can not be removed completely, not even by smaller time steps Δ .
- The pricing biases from the Full Truncation Euler scheme as well as from the direct inversion scheme with reflecting boundary are most significant for small values of the strike.

We focus on the martingale property of the discretized processes generated by the different simulation schemes. In Table 5 we see that the direct inversion scheme with reflecting boundary gives rise to a positive drift which decreases with smaller time steps. In contrast, the proposed unbiased SABR scheme does not generate any statistically significant drift, and the martingale property is preserved. The Euler scheme with Full Truncation does not preserve the martingale property, although the drift decreases (but does not disappear) with smaller time step Δ .

To show the order of convergence of the unbiased SABR scheme, we compute the root-mean-squared (RMS) errors of at-the-money (ATM) option prices obtained by the Euler and the unbiased SABR scheme for different numbers of time steps, see Figure 7. The convergence behavior of the unbiased scheme appears to be superior (which is also the case for other K -values, not shown).

4.3 Results for cases III and IV

For Test Cases III and IV, we consider the full correlation SABR model and some practical yet challenging parameter settings, like rates almost zero or high volatility.

Parameter set III is particularly challenging for the Euler discretization scheme and for Hagan’s SABR formula, because the initial rates are close to zero. It is known that when initial asset prices are close to zero, many paths may reach negative values. The Full Truncation Euler scheme will project the negative values to zero. The drawback is that the truncated Gaussian process is not a martingale anymore, and an increasing number of time steps has to be employed to reduce the resulting bias. On the other hand, the asymptotic SABR formula by Hagan is not valid for strikes $K \rightarrow 0$. The

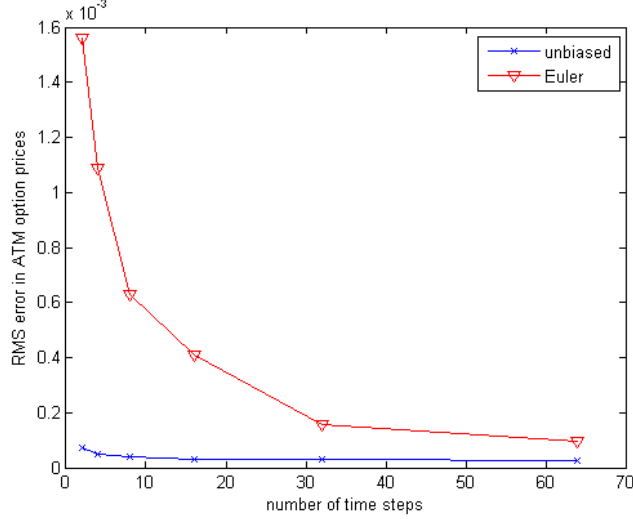


Figure 7: Test Case II; Convergence of the estimated RMS error for call options with decreasing time step Δ .

Table 6: Martingale bias of discrete processes for Test Case III.

Δ	$T = 2$		$T = 5$		$T = 10$	
	Unbiased	Euler	Unbiased	Euler	Unbiased	Euler
1	0.005	0.0057	0.005	0.0064	0.005	0.007
1/2	0.005	0.0054	0.005	0.0059	0.005	0.0061
1/4	0.005	0.0052	0.005	0.0054	0.005	0.0054
1/8	0.005	0.0051	0.005	0.0052	0.005	0.0052
1/16	0.005	0.0051	0.005	0.0051	0.005	0.0051

formula is a result of keeping $\log(f/K)$ constant and taking $T \rightarrow 0$. However, when $K \rightarrow 0$, it follows that $\log(f/K) \rightarrow \infty$, which is an incorrect way of approaching the asymptotic limit. As a result, Hagan's formula is not accurate for very low strike values.

In the unbiased SABR scheme, option pricing at low strike values does not pose any problem. Since an accurate reference value is lacking for Test Case III, we use the martingale property, $E[\hat{S}_T] = \hat{S}_0$, of the discrete process to analyze the accuracy of the unbiased scheme. For comparison, we also simulate with the Full Truncation Euler scheme with the same number of time steps (see Table 6).

The results in Table 6 show that the unbiased scheme with the absorbing boundary does not generate any bias for this test. For the Euler scheme the drift decreases rapidly with the number of time steps.

For Test Case IV, due to the absence of reference values (by a semi-closed form solution or an accurate approximation), we make use of the following result to determine the convergence order of our algorithm:

RESULT 4.1 (Weak Convergence Order without an Exact Solution [30]). 1) If a discrete approximation \hat{X} to a continuous process X with time step Δ has a weak convergence order γ for some positive constant K_1 , i.e.:

$$|E[g(X(T))] - E[g(\hat{X}(T, \Delta))]| \leq K_1 \Delta^\gamma \quad (4.1)$$

Then, there exists a positive constant, K_2 , such that:

$$\left| E[g(\hat{X}(T, \Delta))] - E[g(\hat{X}(T, \frac{\Delta}{2}))] \right| \leq K_2 \Delta^\gamma. \quad (4.2)$$

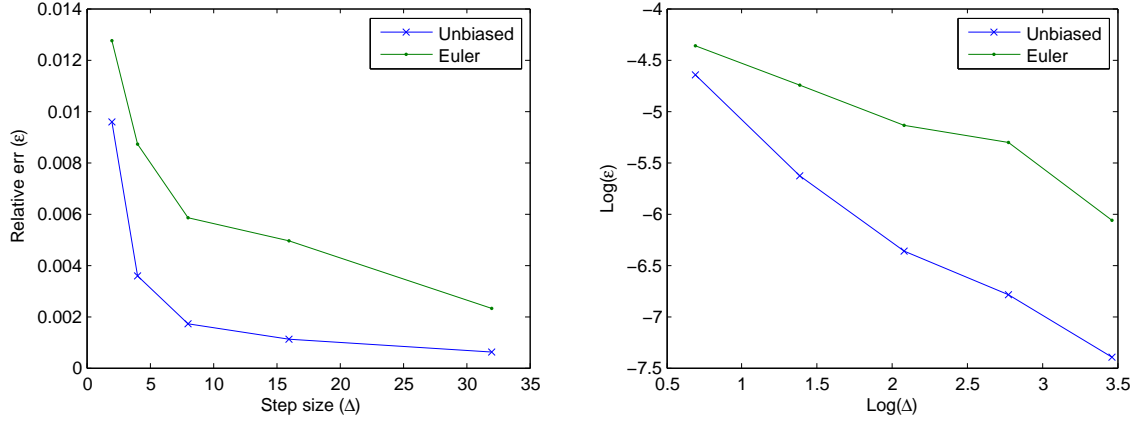


Figure 8: Convergence of relative errors; Test Case IV.

2) Conversely, if it is known that the discretization is weakly convergent and Eq. (4.2) holds for some positive constant K_2 , then the weak convergence order is γ . The proof can be found in Schmitz-Abe & Shaw [30].

For Test Case IV the resulting call option prices are presented in Table 7. We apply Formula (4.2) to the Monte Carlo prices of the ATM options with 5 years maturity from Table 7, and define the relative error to be

$$\epsilon = \left| \hat{C}(\hat{S}(T, \Delta)) - \hat{C}\left(\hat{S}\left(T, \frac{\Delta}{2}\right)\right) \right|.$$

Here, \hat{C} denotes the Monte Carlo estimate of the call option price for underlying discrete process \hat{S} . Clearly, the unbiased scheme produces convergent Monte Carlo prices (see Figure 8) and has a higher order of convergence than the truncated Euler scheme. Again, the relative error of the unbiased scheme with only 2 time steps per year is comparable to that of the Euler scheme with more than 32 time steps per year.

4.4 Computational time

From the numerical results presented, the accuracy of the unbiased scheme has been confirmed. In addition, the CPU time to compute one sample path is of great importance for practical application. The CPU time required for the unbiased scheme is largely dependent on the value of parameter β . Computations are faster for $\beta \approx 1$ and somewhat slower for $\beta \approx 0$. This can be seen from the space transform employed in the derivation of the unbiased SABR scheme, $Y = S^{2-2\beta}/(1-\beta)^2$. With $\beta \approx 1$, Y tends to infinity and the distribution approaches a Gaussian distribution. Most of the sample paths will then be drawn from the ‘cheap to evaluate’ quadratic Gaussian approximation. On the contrary, when $\beta \approx 0$, most of the draws would be performed by the direct inversion scheme, which takes more CPU time.

For reference, Table 8 presents CPU times for the Full Truncation Euler and the unbiased schemes used in Cases III and IV. In these two cases, the difference in CPU time for one time step in the unbiased scheme is roughly a factor of 20. Despite this, we argue that it is still preferable to choose the unbiased scheme in the latter case ($\beta \approx 0$, S_0 small) since for this parameter setting the Euler scheme is significantly biased⁴. This bias can not be reduced, not even by 80 – 100 time steps per year. The unbiased scheme is then more efficient to reach the same level of accuracy.

5 Conclusion

In this paper we have presented an unbiased SABR simulation scheme. We firstly reviewed some analytic properties of the CEV process, which is a space-transformed squared Bessel process, and

⁴The probability of hitting zero is high, so that a large number of truncations of the negative paths are expected.

Table 7: Estimated call option price for Test Case IV.

Δ	$K = 40\%$		$K = 100\%$		$K = 160\%$	
	Unbiased	Euler	Unbiased	Euler	Unbiased	Euler
$T = 2$						
1	0.0803	0.0922	0.0645	0.0714	0.0510	0.0538
1/2	0.0688	0.0838	0.0535	0.0649	0.0405	0.0489
1/4	0.0642	0.0771	0.0492	0.0596	0.0366	0.0449
1/8	0.0619	0.0708	0.0474	0.0546	0.0352	0.0409
1/16	0.0610	0.0673	0.0468	0.0518	0.0347	0.0388
1/32	0.0604	0.0643	0.0463	0.0494	0.0343	0.0368
$T = 5$						
1	0.0795	0.1074	0.0672	0.0895	0.0564	0.0735
1/2	0.0693	0.0928	0.0576	0.0767	0.0454	0.0624
1/4	0.0667	0.0827	0.0540	0.0680	0.0429	0.0550
1/8	0.0643	0.0759	0.0523	0.0621	0.0412	0.0499
1/16	0.0632	0.0702	0.0512	0.0572	0.0406	0.0457
1/32	0.0625	0.0674	0.0506	0.0548	0.0400	0.0437
$T = 10$						
1	0.0765	0.1133	0.0665	0.0966	0.0575	0.0815
1/2	0.0678	0.0975	0.0566	0.0823	0.0466	0.0687
1/4	0.0669	0.0850	0.0550	0.0713	0.0451	0.0591
1/8	0.0652	0.0774	0.0537	0.0645	0.0434	0.0530
1/16	0.0635	0.0705	0.0523	0.0584	0.0422	0.0476
1/32	0.0630	0.0678	0.0520	0.0561	0.0421	0.0456

Table 8: Computational time (in seconds) for a 10 year option with different step sizes, Δ .

Δ	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
Unbiased	8.95	16.20	30.8	59.7	118.4
Euler FT	0.49	0.90	1.73	3.36	6.7

discussed the classification of boundary conditions and the associated probability density and distribution functions. As the conditional SABR process, given the terminal volatility level and the integrated variance, is also a squared Bessel process, we can find an explicit distribution function for the conditional SABR process.

Based on the idea of mixing conditional distributions and a direct inversion of the noncentral chi-square distributions, we have proposed an unbiased SABR Monte Carlo scheme. The scheme proposed can deal with the – often problematic – behavior of the CEV process in the vicinity of the zero boundary. The unbiased scheme is stable and exhibits a highly satisfactory convergence behavior compared to the truncated Euler scheme. The scheme is an alternative when a truncated Euler scheme gives rise to significant bias, even with a very large number of time steps, which is the case, for example, when $S_0 \approx 0$ or when the skewness parameter, β , is less than $\frac{1}{2}$.

A multi-dimensional extension of the present scheme is an interesting topic of future research.

References

- [1] L. Andersen, Efficient simulation of the heston stochastic volatility model, *Journal of Computational Finance* 11:3 (2008) 1–22.
- [2] L. Andersen, J. Andreasen, Volatility skews and extensions of the libor market model, *Applied Mathematical Finance* 7:1 (2000) 1–32.

- [3] L. Andersen, R. Brotherton-Ratcliffe, Extended libor market models with stochastic volatility, *Journal of Computational Finance* 9:1 (2005) 1–27.
- [4] L. Andersen, V. Piterbarg, Moment explosions in stochastic volatility models, *Finance and Stochastics* 11:1 (2007) 29–50.
- [5] S. Borodin, *Handbook of Brownian Motion*, 2nd ed., Birkhäuser, 2002.
- [6] M. Broadie, O. Kaya, Exact simulation of stochastic volatility and other affine jump diffusion processes, *Operations Research* 54:2 (2006) 217 – 231.
- [7] G. Campolieti, R. Makarov, Pricing path-dependent options on state dependent volatility models with a Bessel bridge, *International Journal of Theoretical and Applied Finance* 10:1 (2007) 51 – 88.
- [8] B. Chen, C. Oosterlee, S. van Weeren, Analytic approximation to constant maturity swap convexity correction in a multi-factor SABR model, *International Journal of Theoretical and Applied Finance* 13:7 (2010) 1019–1046.
- [9] J. Cox, Note on option pricing I: Constant elasticity of variance diffusions, *Stanford University Working Paper* (1975) 229 – 263.
- [10] O. Elerian, A note on the existence of a closed form conditional transition density for the milstein scheme, *Nuffield College Working Paper*, Oxford University (1998).
- [11] A. V. Haastrecht, A. Pelsser, Efficient, Almost Exact Simulation of the Heston Stochastic Volatility Model, *International Journal of Theoretical and Applied Finance* 13:1 (2010) 1–43.
- [12] P. Hagan, D. Kumar, A. Lesniewski, D. Woodward, Managing smile risk, *Wilmott Magazine* (2002) 84–108.
- [13] P. Hagan, A. Lesniewski, Libor market model with SABR style stochastic volatility, working Paper (2008).
- [14] S. L. Heston, A closed form solution for options with stochastic volatility, *Review of financial studies* 6 (1993) 327–343.
- [15] O. Islah, Solving SABR in Exact Form and Unifying it with LIBOR Market Model, *SSRN eLibrary*.
- [16] N. Johnson, S. Kotz, *Distributions in Statistics: Continuous Univariate Distributions Vol. 2*, 1st ed., Boston: Houghton Mifflin Company, 1970.
- [17] N. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions Vol. 2*, 1st ed., Wiley Interscience, 1995.
- [18] C. Kahl, *Modelling and Simulation of Stochastic Volatility in Finance*, Dissertation.com, Boca Raton, Florida, 2007.
- [19] C. Kahl, P. Jäckel, Fast strong approximation Monte-Carlo schemes for stochastic volatility models, *Journal of Quantitative Finance* 6:6 (2006) 513–536.
- [20] N. Kunitomo, A. Takahashi, The asymptotic expansion approach to the valuation of interest rate contingent claims, *Mathematical Finance* 11 (2001) 117–151.
- [21] H. Labordere, Combining the SABR and IMM models, *Risk* October (2007) 68–74.
- [22] A. Lesniewski, Notes on the CEV model, NYU working paper (2009).
- [23] A. Lewis, The mixing approach to stochastic volatility and jump models, *Wilmott March* (2002) 24–45.
- [24] R. Lord, R. Koekkoek, D. van Dijk, A comparison of biased simulation schemes for stochastic volatility models, *Quantitative Finance* 10:2 (2010) 177–194.

- [25] G. Milstein, A method of second order accuracy integration of stochastic differential equations, *SIAM: Theory of Probability and Its Applications* 23 (1978) 396–401.
- [26] M. Morini, F. Mercurio, No-Arbitrage Dynamics for a Tractable SABR Term Structure Libor Model, SSRN eLibrary.
- [27] R. Rebonato, A time-homogeneous, sabr-consistent extension of the Imm, QUARC Working Paper, Royal Bank of Scotland (2007).
- [28] R. Rebonato, *The SABR/LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives*, 1st ed., John Wiley and Sons, 2009.
- [29] M. Sankaran, Approximations to the non-central chi-square distribution, *Biometrika* 50 (1963) 199 – 204.
- [30] K. Schmitz-Abe, W. Shaw, Measure order of convergence without an exact solution, euler vs. milstein scheme, *International Journal of Pure and Applied Mathematics* 24:3 (2005) 365 – 381.
- [31] M. Schroder, Computing the constant elasticity of variance option pricing formula, *Journal of Finance* 1:44 (1989) 211 – 218.
- [32] I. Shoji, T. Ozaki, Estimation for nonlinear stochastic differential equations by a local linearization method, *Stochastic Analysis and Applications* 16:4 (1998) 733 – 752.
- [33] G. Willard, Calculating prices and sensitivities for path-independent derivative securities in multifactor models, *Journal of Derivatives* 5 (1997) 45–61.
- [34] L. Yuan, J. Kalbfleisch, On the bessel distribution and related problems, *Annals of the Institute of Statistical Mathematics* 53:3 (2000) 438 – 477.