Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

The principle of not feeling the boundary for the SABR model

Nan Chen & Nian Yang

To cite this article: Nan Chen & Nian Yang (2019) The principle of not feeling the boundary for the SABR model, Quantitative Finance, 19:3, 427-436, DOI: 10.1080/14697688.2018.1486037

To link to this article: https://doi.org/10.1080/14697688.2018.1486037

Published online: 03 Sep 2018.

Quantitative Finance, 2019
Vol. 19, No. 3, 427–436, https://doi.org/10.1080/14697688.2018.1486037

The principle of not feeling the boundary for the SABR model

NAN CHEN† and NIAN YANG‡*

†Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, Sha Tin, Hong Kong
‡Department of Finance and Insurance, Nanjing University, 16 Jinyin Street, Gulou, Nanjing 210093, People's Republic of China

(*Received 7 August 2017; accepted 1 June 2018; published online 3 September 2018*)

The stochastic alpha–beta–rho (SABR) model is widely used in fixed income and foreign exchange markets as a benchmark. The underlying process may hit zero with a positive probability and therefore an absorbing boundary at zero should be specified to avoid arbitrage opportunities. However, a variety of numerical methods choose to ignore the boundary condition to maintain the tractability. This paper develops a new principle of not feeling the boundary to quantify the impact of this boundary condition on the distribution of underlying prices. It shows that the probability of the SABR hitting zero decays to 0 exponentially as the time horizon shrinks. Applying this principle, we further show that conditional on the volatility process, the distribution of the underlying process can be approximated by that of a time-changed Bessel process with an exponentially negligible error. This discovery provides a theoretical justification for many almost exact simulation algorithms for the SABR model in the literature. Numerical experiments are also presented to support our results.

Keywords: SABR model; Probability of hitting zero; Principle of not feeling the boundary; Time-changed Bessel process

JEL Classification: C63, G13

1. Introduction

The stochastic alpha–beta–rho (SABR) model introduced in Hagan *et al.* (2002) has become very popular with practitioners in interest rate and foreign exchange markets for valuing European-style options. It can produce analytical asymptotic expressions for implied volatility, fitting the observed smile reasonably well and capturing the correct co-movement between the smile dynamics and the underlying asset price.

The model is a special class of stochastic volatility models. In particular, its underlying process is given by a constant elasticity of variance (CEV) type diffusion and its volatility process follows a geometric Brownian motion. Exactly due to this structural feature, one can show that the underlying process may hit zero with positive probability.† Therefore, we have to specify an absorbing boundary condition at zero to avoid arbitrage opportunities (see, e.g. Delbaen and Shirakawa 2002, Rebonato and McKay 2009). However, some intensively used approximate formulas for the Black model implied volatility given by SABR, such as those in Hagan *et al.* (2002), Obłój (2008), and Paulot (2015), were derived by simply ignoring the boundary condition for the pricing partial differential equation (PDE) system to maintain mathematical tractability. This therefore poses one interesting research problem upon us: how should we quantify the impact of the boundary condition?

To address this issue, we manage to develop a *principle of not feeling the boundary* in this paper. It shows that the probability of the SABR model hitting zero decays exponentially as the time horizon shrinks. Furthermore, the convergence rate of this hitting probability to zero largely depends on the modelling parameters: it becomes faster for a model with larger

*Corresponding author. Email: yangnian@nju.edu.cn
† For example, if the correlation is zero and the parameter 'beta' is less than 1/2 (see the SDE (1) for the SABR model), then the mass at zero is positive.

initial underlying price or beta (i.e. the index of the CEV component of the SABR model) or smaller initial volatility or the volatility of volatility. To the best of our knowledge, we characterize the exponentially decaying order of hitting probability for the first time.

Early works on the principle of not feeling the boundary can be traced back to Kac (1951) and Varadhan (1967), where the authors investigated the case of diffusions on Euclidean space generated by the second order, uniformly elliptic operators with Hölder continuous coefficients. Hsu (1995) extended the discussion to diffusions on a more general manifold. In terms of its financial applications, Gatheral *et al.* (2012) use the principle to explain why the boundary behaviour of local volatility models will not affect the asymptotic expansions of the transition density and the European call price written on it. This paper contributes to this literature because we are the first to present a rigorous characterization about the exponentially decaying order of the hitting probability in the case of the SABR, one important class of stochastic volatility models. We also note there are several works relating to this topic. For instance, Doust (2012) computed the probability of hitting zero via Monte–Carlo simulations. Bayer *et al.* (2013) cited this principle without any rigorous establishment to argue for the validity of their computational method in the SABR model. Gulisashvili *et al.* (2016) derived the formula of the hitting probability for the normal SABR model when the time horizon tends to infinity. For the uncorrelated SABR model, Gulisashvili *et al.* (2018) derived the behaviour of the atom at the origin for short and large times; they also referred to the small hitting probability for large initial values as the principle in a numerical example, but without indicating the decaying order. Using the PDE-based method, Yang and Wan (2018) obtained asymptotic formulas with a polynomial error bound for the survival probability (i.e. the probability of not hitting a nonnegative lower boundary) by solving a hierarchy of PDEs. It is worth noting that Hagan *et al.* (2014) proposed a numerical scheme to solve a simplified one-dimensional pricing PDE for the SABR model by considering the boundary conditions at zero.

Intuitively, our result implies that the specification of boundary conditions will have a limited influence on the distributional law of the SABR model for a small time. That explains why a variety of numerical methods for the SABR model perform well for short-run option pricing, even though they do not incorporate the boundary condition into consideration. As the second layer of contributions to the literature, the paper also develops some theoretical bounds on the bias of some almost exact SABR simulation algorithms that recently emerged in the literature (see, e.g. Chen and Liu 2011, Chen *et al.* 2012, Cai *et al.* 2017, Leitao *et al.* 2017). This research line of simulation stems from Islah (2009), in which the author found that the marginal distribution of the underlying price in SABR can be approximated by a noncentral chi-square distribution conditional on the volatility process. This approximation turns out to be quite accurate if we use it to compute short-term option prices. Meanwhile, the aforementioned papers also report significant approximation error in the simulated outcomes as the time horizon gets longer. Applying the principle of not feeling the boundary, we manage to identify the cause of such error—the impact of the absorbing boundary condition starts to kick in when we consider a long time horizon. Along this line, we fill the gap in the existing literature by presenting an analysis that the approximation error will be exponentially negligible as the time horizon shrinks.

The rest of this paper is organized as follows. Section 2 introduces the SABR model, followed by the main results about the principle of not feeling the boundary. Some numerical evidences are also presented to support our discovery. All the proofs are deferred to Section 3. We conclude the paper in Section 4.

2. The SABR model and the main results

2.1. The SABR model

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space, where $\mathbb{P}$ is the T-forward martingale measure. Two independent standard Brownian motions $\{B_t; 0 \le t \le T\}$ and $\{W_t; 0 \le t \le T\}$ are defined on $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ with their natural filtrations $\{\mathcal{F}_t^1\}$ and $\{\mathcal{F}_t^2\}$, respectively. Let the sample space Ω, the σ-algebra $\mathcal{F}$, and the filtration $\mathcal{F}_t$ be $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}^1 \otimes \mathcal{F}^2$, and $\mathcal{F}_t = \mathcal{F}_t^1 \otimes \mathcal{F}_t^2$. Denote F_t and A_t to be the forward price and volatility at time $t \in [0, T]$, respectively. The SABR model is then defined as a solution to the following system of stochastic differential equations (SDEs):

$$\begin{aligned} dF_t &= A_t F_t^{\beta}[\sqrt{1-\rho^2}dB_t + \rho dW_t], \\ dA_t &= \nu A_t dW_t, \end{aligned} \tag{1}$$

where the parameter beta β and the correlation correlation ρ satisfy $\beta \in (0, 1)$ and $\rho \in (-1, 1)$, respectively; the forward price F_0, the initial volatility A_0, and the volatility of volatility ν are positive. It is a local stochastic volatility model, in which the forward price process $\{F_t; 0 \le t \le T\}$ follows a CEV-type diffusion process and the dynamic of the volatility process $\{A_t; 0 \le t \le T\}$ is given by a geometric Brownian motion.

We need to specify the boundary condition at $F=0$ for SDE (1) in order to determine the existence and uniqueness of the model, because the CEV-type dynamic specification in F allows it to hit zero with positive probability. A reflecting boundary will obviously lead to an arbitrage opportunity: one can buy the forward at zero cost when it hits 0 and sell it for profit when it reflects back to the positive region; refer to Section 3.10 of Rebonato and McKay (2009) or Delbaen and Shirakawa (2002) for a detailed discussion on the issue. To rule out the arbitrage opportunity, we thus impose the following assumption on the model from now on.

Assumption 2.1 0 is an absorbing boundary of $\{F_t; 0 \le t \le T\}$.

We show that the solution to equation (1) uniquely exists under this assumption in Lemma 3.1.

2.2. The main results

Let

$$\tau_0^F = \inf\{t \in [0, T] : F_t = 0\},$$

the first time the forward price process hits zero. The first main result of the paper establishes a probability bound on $\mathbb{P}(\tau_0^F \le T)$. More specifically, we have

THEOREM 2.1 (Principle of Not Feeling the Boundary) *Under Assumption* 2.1, *there exists a positive constant C (depending on ν, β, A_0, F_0) such that,*

$$\limsup_{T\downarrow 0} T \ln \mathbb{P}(\tau_0^F \le T) \le -C. \tag{2}$$

In words, the theorem states that the probability of the event that the forward price F_t hits 0 before T will vanish exponentially as T tends to zero. Since the SABR model (1) is a diffusion process changing its value continuously over time, intuitively Theorem 2.1 implies that the existence of the boundary at zero will not affect the probability law of F_t for small time. In this sense, we refer to it as the *principle of not feeling the boundary*.

It is worthwhile pointing out that using such principle to quantify the impact of the boundary to the probability law of a diffusion process is familiar to probabilists. Kac (1951) pioneered the study for Brownian motion. Varadhan (1967) considered the case of diffusion processes in a Euclidean space generated by second order, uniformly elliptic operators. Hsu (1995) extended the principle to diffusions on a general manifold. One interesting application of the principle in option pricing appeared in Gatheral *et al.* (2012), in which the authors used it to obtain asymptotic expansions about the transitional probability function of a local volatility model and the associated call option price.

Turn to the implications of Theorem 2.1 on the SABR model. Define a function $g(\cdot)$ such that for $F \ge 0$,

$$g(F) = \frac{F^{1-\beta}}{1-\beta}.$$

Let $X_t = g(F_t)$. Applying the local Itô formula (Kallenberg 1997, Corollary 15.20) up to the stopping time τ_0^F, we have

$$X_{T\wedge\tau_0^F} = X_0 + \frac{\rho}{\nu}(A_{T\wedge\tau_0^F} - A_0) + \sqrt{1-\rho^2}\int_0^{T\wedge\tau_0^F} A_s\,\mathrm{d}B_s + \int_0^{T\wedge\tau_0^F} \frac{(1-2\theta)(1-\rho^2)A_s^2}{2X_s}\,\mathrm{d}s, \tag{3}$$

with $X_0 = g(F_0)$ and

$$\theta = \frac{1}{2} + \frac{\beta}{2(1-\beta)(1-\rho^2)}. \tag{4}$$

Note that the function g defined above is invertible. We have $F_T = F_{T\wedge\tau_0^F} = g^{-1}(X_{T\wedge\tau_0^F})$. To derive the probability law of F_T, it suffices to determine the probability law of $X_{T\wedge\tau_0^F}$.

Along the sample path that satisfies $\tau_0^F > T$, the representation of X in equation (3) reduces to

$$X_T = X_0 + \frac{\rho}{\nu}(A_T - A_0) + \sqrt{1-\rho^2}\int_0^T A_s\,\mathrm{d}B_s + \int_0^T \frac{(1-2\theta)(1-\rho^2)A_s^2}{2X_s}\,\mathrm{d}s. \tag{5}$$

Conditioning on the volatility process $\{A_t; 0 \le t \le T\}$, the distribution law of X_T given by equation (5) should be the same as the marginal distribution of a *time-changed Bessel process* of parameter $(1-2\theta)/2$ at T, starting from $X_0 + \rho/\nu(A_T - A_0)$ and with the changed time clock $(1-\rho^2)\int_0^T A_s^2\,\mathrm{d}s$. Since the probability of $\{\tau_0^F \le T\}$ is exponentially negligible for small time T according to Theorem 2.1, we expect that the probability distribution of X_T should provide a good approximation to the distribution of $X_{T\wedge\tau_0^F}$.

Lemma 3.4 in Section 3 presents explicitly the cumulative distribution function of a time-changed Bessel process in terms of noncentral chi-square distributions. Inspired by all the above observations, we introduce a new random variable (r.v.) $\tilde{F}_T$ whose conditional distribution, given both the values of A_T and $\int_0^T A_s^2\,\mathrm{d}s$, satisfies

$$\mathbb{P}\left(\tilde{F}_T \le g(U) \,\middle|\, \int_0^T A_s^2\,\mathrm{d}s, A_0, A_T\right) = \begin{cases} 1 - Q\left(\dfrac{\tilde{g}^2(F_0)}{\Delta}; 2\theta, \dfrac{g^2(U)}{\Delta}\right), & U > 0; \\ 1 - Q\left(\dfrac{\tilde{g}^2(F_0)}{\Delta}; 2\theta\right), & U = 0, \end{cases} \tag{6}$$

for any $U \ge 0$, where

$$\Delta = (1-\rho^2)\int_0^T A_s^2\,\mathrm{d}s \tag{7}$$

and

$$\tilde{g}(F_0) := \left(X_0 + \frac{\rho}{\nu}(A_T - A_0)\right)^+ = \left(g(F_0) + \frac{\rho}{\nu}(A_T - A_0)\right)^+. \tag{8}$$

Here, $Q(x;\mu,\lambda)$ is the cumulative distribution function of a noncentral chi-square random variable with degree of freedom μ and noncentrality λ. $Q(x;\mu)$ is its degenerate special case when $\lambda = 0$.

The following theorem characterizes the error bound if we use the aforementioned r.v. $\tilde{F}_T$ to build up an approximation to the original model F_T. More precisely, we have

THEOREM 2.2 (Approximate Conditional Marginal Distribution) *Suppose Assumption* 2.1 *holds. For any Lipschitz function $h(\cdot)$, there exists a positive constant C (depending on h, ν, β, ρ, A_0, F_0) such that*

$$\limsup_{T\downarrow 0} T \ln \left|\mathbb{E}[h(F_T)\,|\,A_0, F_0] - \mathbb{E}\left[\mathbb{E}\left[h(\tilde{F}_T) \,\middle|\, \int_0^T A_s^2\,\mathrm{d}s, A_T\right] \middle|\, A_0, F_0\right]\right| \le -C. \tag{9}$$

In equation (9), the first expectation $\mathbb{E}[h(F_T)\,|\,A_0, F_0]$ is taken with respect to the original SABR model; the inner one in the second iterated expectations is computed from the probability distribution given in equation (6) and the outer one is taken with respect to the joint distribution of $\int_0^T A_s^2\,\mathrm{d}s$ and A_T.

Roughly, we know from this theorem

$$\left|\mathbb{E}[h(F_T)\,|\,A_0, F_0] - \mathbb{E}\left[\mathbb{E}\left[h(\tilde{F}_T)\,\middle|\,\int_0^T A_s^2\,\mathrm{d}s, A_T\right]\middle|\,A_0, F_0\right]\right| \le \exp(-C/T),$$

i.e. the difference of these two terms will vanish exponentially as $T \to 0$. Specifically, if the correlation is zero, i.e. $\rho = 0$, then the difference of the above two expectations is zero. Because the marginal distribution of the forward price is exactly given by equation (6), which is an immediate consequence of the results of Islah (2009), Cai *et al.* (2017), and Leitao *et al.* (2017).

The iterated expectation in equation (9) can be evaluated efficiently through Monte–Carlo simulation. To this end, we may use the following three-step procedure:

Step 1 Given A_0, simulate A_T.
Step 2 Draw a sample of $\int_0^T A_s^2\,\mathrm{d}s$, given A_0 and A_T.
Step 3 Given A_0, A_T, and $\int_0^T A_s^2\,\mathrm{d}s$, simulate $\tilde{F}_T$ from the distribution (6).

Several papers in the literature, including Chen and Liu (2011), Cai *et al.* (2017), and Leitao *et al.* (2017), developed different simulation schemes to materialize these steps. We include the detail of the algorithm presented in Cai *et al.* (2017) in Appendix 1. All of these papers document the accurate performance of the above approximation in numerical experiments when it is used to compute option prices written on the SABR model, especially for short-term options. But none produces any theoretical guarantees. Theorem 2.2 fills the gap by showing that the approximation error in using these Monte–Carlo simulation schemes to price short-term options under the SABR model is exponentially negligible.

At the end of this subsection, we need to stress that the discussion ahead of Theorem 2.2 is not rigorous. More strict proofs of the two theorems in this subsection can be found in Section 3.

2.3. Numerical evidences

In this subsection, we shall provide more numerical evidences about the principle of not feeling the boundary and its implications in option pricing. Note that equation (2) in Theorem 2.1 implies that $\ln(\mathbb{P}(\tau_0^F \le T))$ is in proportion to $1/T$. To numerically corroborate this discovery, figure 1 displays the relationship between the logarithm of the hitting probability and the reciprocal of the maturity under different values of initial forward price F_0, beta, initial volatility A_0, and volatility of volatilities ν. The values of parameters we use as the benchmark in this experiment are $F_0 = 0.1$, $A_0 = 0.2$, $\beta = 0.1$, $\nu = 0.1$, and $\rho = -0.5$, respectively. We change the value of one parameter in each subfigure while fixing the others. All the data points fall onto straight lines,

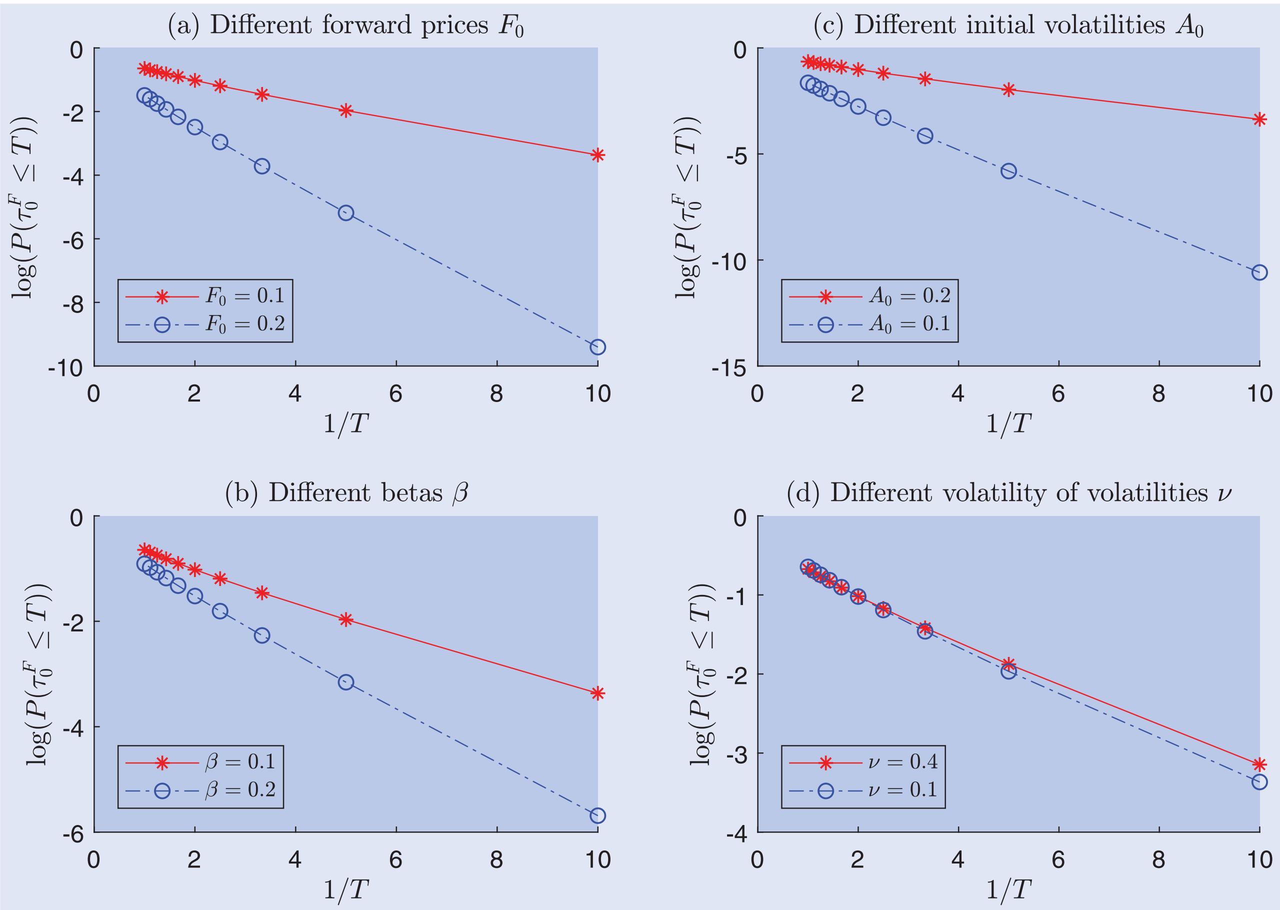

Figure 1. The linear relation between $\log(\mathbb{P}(\tau_0^F \le T))$ and $1/T$ under different values of F_0, A_0, β, and ν. The hitting probability of the SABR model is computed from the finite-difference method, see Appendix 1 for more details.

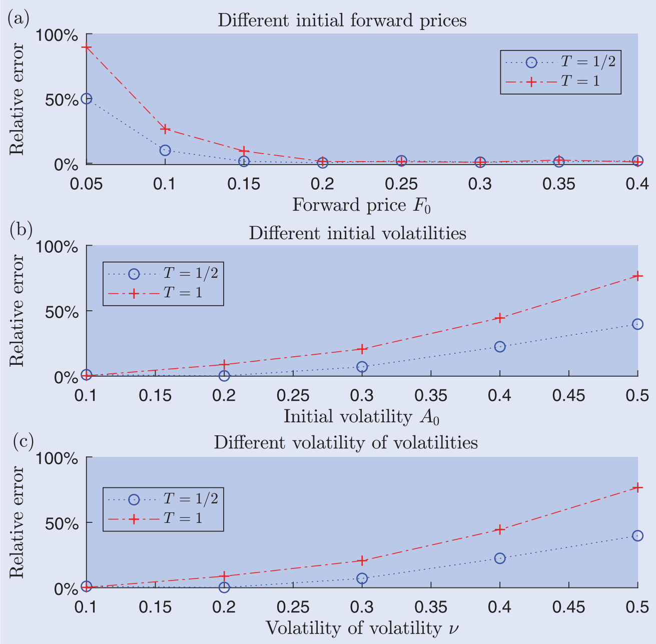

Figure 2. The relative errors of the simulation under different values of F_0, A_0, and ν. Consider at-the-money call option. We fix $\beta = 0.1$ and $\rho = -0.5$ in all three subfigures. The other parameter values used in this example are $A_0 = 0.3$ and $\nu = 0.1$ in (a), $F_0 = 0.1$ and $\nu = 0.1$ in (b), $F_0 = 0.1$ and $A_0 = 0.1$ in (c), respectively. We compute the benchmark price D using the finite-difference method in Appendix 1. The simulation algorithm of Cai *et al.* (2017) is used to generate Monte Carlo price S. The relative error is defined as $|S - D|/D$. The number of simulation trials is 10,000.

strongly suggesting that the principle of not feeling the boundary holds for the SABR model under a variety of parameter values. Moreover, the slope of the lines, i.e., the constant on the right-hand side of equation (2), characterizes how fast the probability decays as $T \to 0$. A clear pattern arises in figure 1 that $\mathbb{P}(\tau_0^F \leq T)$ converges to 0 faster for the SABR model with a larger initial forward price, larger beta, smaller initial volatility, or smaller volatility of volatility.

Figure 2 shows the impact of the boundary on the option pricing. Using the values computed from the finite-difference method as the 'true' option price, we plot the relative error in the computation of the simulation scheme presented in the last subsection. For a given T, the error tends to be significant as $F_0 \to 0$, $A_0 \to +\infty$, or $\nu \to +\infty$. This is consistent with the conclusions from the two theorems in the last subsection. As noted before, $\mathbb{P}(\tau_0^F \leq T)$ becomes not negligible for a model with small F_0, large A_0 or ν, and will lead to the failure of the distribution approximation presented in Theorem 2.2. In addition, through the comparison between $T = 1/2$ and $T = 1$, we can see that the boundary condition should have a smaller impact to the error of the simulation-based pricing schemes for a shorter time horizon.

3. Proofs

This section provides proofs for Theorems 2.1 and 2.2. In Section, 3.1 we present some technical lemmas for subsequent analysis. The proofs for Theorems 2.1 and 2.2 are presented in Sections 3.2, and 3.3, respectively. For convenience, we will use the following notations throughout this section.

- C is a generic positive constant.
- $C(\varpi)$ is a generic positive constant depending on the parameter vector ϖ, which can be one or a group of β, ν, F_0, and so on. The explicit dependence will be indicated in the following analysis.

3.1. Technical lemmas

LEMMA 3.1 (Strong Solution up to Explosion) *Under Assumption* 2.1, *the SABR model* (1) *has a unique strong solution up to the explosion time S, where $S = \inf\{t > 0 : F_t = 0\}$.*†

Proof Under Assumption 2.1, Lions and Musiela (2007) and Hobson (2010) have proved that the equation system (1) admits a unique weak solution up to the explosion. Moreover, the solution to the SDE (1) exists in a strong sense. Note that the diffusion coefficients are locally Lipschitz. We then have strong uniqueness for the solution by simply following the proof of Theorem 5.2.5 in Karatzas and Shreve (1991). Therefore, the SDE (1) exists a strong solution up to explosion, which is implied by the weak existence and strong uniqueness (Karatzas and Shreve 1991, Corollary 5.3.23). ∎

LEMMA 3.2 *Consider the volatility process $\{A_t; 0 \leq t \leq T\}$ in the SABR model* (1). *Define*

$$\Omega_T^{X_0} = \left\{ \inf_{s\in[0,T]} \left(X_0 + \frac{\rho}{\nu}(A_s - A_0) \right) \leq 0 \right\}. \tag{10}$$

Let $C_a = \nu X_0/\rho A_0$ if $\rho \neq 0$. Then, for (i) $\rho = 0$, (ii) $\rho > 0$ *and $C_a \geq 1$, we have $\Omega_T^{X_0} = \varnothing$; otherwise,*

$$\mathbb{P}(\Omega_T^{X_0}) \leq \frac{1}{\sqrt{1 - C_a}} \frac{\nu\sqrt{T}}{|\ln(1 - C_a)|} \exp\left(-\frac{\ln^2(1 - C_a)}{2\nu^2 T} \right). \tag{11}$$

Proof We first consider the case $\rho > 0$. Note that $A_s = A_0 \exp(-\nu^2 s/2 + W_s)$ for all $s \in [0, T]$. Then,

$$\begin{aligned} &\inf_{s\in[0,T]} (X_0 + \rho/\nu(A_s - A_0)) \leq 0 \\ &\iff \inf_{s\in[0,T]} \exp(-\nu^2 s/2 + \nu W_s) \\ &\leq 1 - C_a. \end{aligned}$$

If $C_a \geq 1$, then $\Omega_T^{X_0} = \varnothing$. Thus, the lemma holds obviously. If $0 < C_a < 1$, then

$$\Omega_T^{X_0} = \left\{ \inf_{s\in[0,T]} (-\nu s/2 + W_s) \leq \frac{\ln(1 - C_a)}{\nu} \right\}.$$

This set corresponds to the event that a drifted Brownian motion $\{-\nu s/2 + W_s : s \geq 0\}$ hits the level $b := \ln(1 -$

† $\{F_t; 0 \leq t \leq T\}$ has a finite moment (Andersen and Piterbarg 2007, Lions and Musiela 2007) for $\beta \in [0, 1)$, thus $S = \lim_{n\to\infty}\{t > 0 : F_t \leq 1/n, \text{or } F_t \geq n\}$.

$C_a)/\nu < 0$ before T. By the distribution of first passage times of the drifted Brownian motion (Karatzas and Shreve 1991, formula (3.5.12)), we have

$$\mathbb{P}(\Omega_T^{X_0}) = \int_0^T \frac{-b}{\sqrt{2\pi s^3}} \exp\left(-\frac{b^2 + \nu bs + \nu^2 s^2/4}{2s}\right) \mathrm{d}s \\ \leq \int_0^T \frac{-b}{\sqrt{2\pi s^3}} \exp\left(-\frac{b^2}{2s}\right) \cdot \exp\left(-\frac{\nu b}{2}\right) \mathrm{d}s,$$

where the inequality is due to the fact that $\exp(-\nu^2 s/8) \leq 1$. Letting $z = -b/\sqrt{s}$, the above inequality yields that

$$\mathbb{P}(\Omega_T^{X_0}) \leq \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{1-C_a}} \int_{-b/\sqrt{T}}^{+\infty} \mathrm{e}^{-(1/2)z^2} \,\mathrm{d}z.$$

By the inequality in Problem 2.9.22 of Karatzas and Shreve (1991), we have equation (11). A similar argument applies to the case when $\rho \leq 0$. ■

Lemma 3.3 *Consider the volatility process $\{A_t; 0 \leq t \leq T\}$ in the SABR model* (1). *Recall that $\Delta = (1-\rho^2)\int_0^T A_s^2 \,\mathrm{d}s$ in* (7). *For arbitrary $C_0 > \Delta_0 \equiv (1-\rho^2)A_0^2 T$, then we have*

$$\mathbb{P}(\Delta \geq C_0) \leq C_1(\nu\sqrt{T}) \exp\left(-\frac{C_2}{\nu^2 T}\right),$$

where $C_1 = 2\sqrt{2/\pi}/\ln(C_0/\Delta_0)$ and $C_2 = (\ln(C_0/\Delta_0))^2/8$.

Proof Note that $A_t = A_0 \exp(-\nu^2 t + 2\nu W_t)$ for $t \in [0,T]$. Then, we have

$$\mathbb{P}(\Delta \geq C_0) = \mathbb{P}\left(\int_0^T \exp(-\nu^2 t)\exp(2\nu W_t)\,\mathrm{d}t \geq \frac{C_0 T}{\Delta_0}\right) \\ < \mathbb{P}\left(\int_0^T \exp(2\nu \sup_{t\in[0,T]} W_t)\,\mathrm{d}t \geq \frac{C_0 T}{\Delta_0}\right).$$

Therefore, using the density of the running maximum of Brownian motion (Karatzas and Shreve 1991, formulas (2.8.3)), we have

$$\mathbb{P}(\Delta \geq C_0) \leq \mathbb{P}\left(\sup_{0\leq t\leq T} W_t \geq \frac{1}{2\nu}\ln\left(\frac{c}{\Delta_0}\right)\right) \\ = \sqrt{\frac{2}{\pi}} \int_{\ln(c/\Delta_0)/2\nu\sqrt{T}}^{\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x \\ \leq C_1(\nu\sqrt{T}) \exp\left(-\frac{C_2}{\nu^2 T}\right).$$

The last inequality holds due to the formula (2.9.20) in Karatzas and Shreve (1991). ■

Lemma 3.4 (Time-changed Bessel process) *For any given deterministic positive continuous function $\varphi : [0,T] \to (0,+\infty)$, let*

$$Y_t = Y_0 + \int_0^t \varphi(s)\,\mathrm{d}B_s + \int_0^t \frac{(1-2\theta)\varphi^2(s)}{2Y_s}\,\mathrm{d}s, \tag{12}$$

where B_t is a standard Brownian motion. For any fixed $t > 0$, the transition density of Y_t in equation (12), *starting from $Y_0 > 0$, is given as follows*:

$$p_\varphi(t; Y_0, y) \\ = \begin{cases} \dfrac{Y_0^\theta y^{1-\theta}}{\int_0^t \varphi^2(s)\,\mathrm{d}s} \exp\left(-\dfrac{Y_0^2 + y^2}{2\int_0^t \varphi^2(s)\,\mathrm{d}s}\right) I_\theta\left(\dfrac{Y_0 y}{\int_0^t \varphi^2(s)\,\mathrm{d}s}\right), & y > 0; \\ \dfrac{1}{\Gamma(1+\theta)}\Gamma\left(\theta, \dfrac{Y_0^2}{2\int_0^t \varphi^2(s)\,\mathrm{d}s}\right), & y = 0; \end{cases} \tag{13}$$

where $\Gamma(\theta) = \int_0^\infty x^{\theta-1}\mathrm{e}^{-x}\,\mathrm{d}x$, $\Gamma(\theta,z) = \int_z^\infty x^{\theta-1}\mathrm{e}^{-x}\,\mathrm{d}x$, and $I_\theta(z) = \sum_{m=0}^\infty ((z/2)^{2m+\theta}/m!\Gamma(m+\theta+1))$. Moreover, Y_t admits the following distribution function:

$$\mathbb{P}(Y_t \leq y | Y_0) \\ = \begin{cases} 1 - Q\left(\dfrac{Y_0^2}{\int_0^t \varphi^2(s)}; 2\theta, \dfrac{y^2}{\int_0^t \varphi^2(s)}\right), & y > 0; \\ 1 - Q\left(\dfrac{Y_0^2}{\int_0^t \varphi^2(s)}; 2\theta\right), & y = 0, \end{cases} \tag{14}$$

where $Q(x;\mu,\lambda)$ is the cumulative distribution function of a noncentral chi-square random variable with degree of freedom μ and noncentrality λ. $Q(x;\mu)$ is its degenerate special case when $\lambda = 0$.

Proof Given any deterministic positive continuous function $\varphi : [0,T] \to (0,+\infty)$, note that the coefficients of the process Y (cf. (12)), $\varphi(s)$ and $(1-2\theta)\varphi^2(s)/(2Y_s)$, are locally Lipschitz in the space variable Y_s when $Y_s > 0$. Mimicking the proof of Theorem 5.2.5 of Karatzas and Shreve (1991), then we can easily show the strong uniqueness of the solution to equation (12) up to the explosion time $\tau^Y = \{t \geq 0 : Y_t = 0\}$. Let

$$\phi(t) := \int_0^t \varphi^2(\gamma)\,\mathrm{d}\gamma \quad \text{and} \quad \psi(t) := \inf\{s > 0 : \phi(s) > t\}.$$

Since $\varphi(\cdot)$ is strictly positive and continuous, we know that $\phi(t)$ and $\psi(t)$ are both continuous and monotonously increasing. Define $M_t = \int_0^{\psi(t)} \varphi(\gamma)\,\mathrm{d}B_\gamma$. Note that $\langle M, M\rangle_t = \int_0^{\psi(t)} \varphi^2(\gamma)\,\mathrm{d}\gamma = t$. From Theorem 3.3.16 of Karatzas and Shreve (1991), $\{M_t\}$ is a Brownian motion with respect to the filtration $\{\mathcal{F}_{\psi(t)}^B : t \in [0,T]\}$. Given $Y_0 > 0$, we know that

$$Z_t = Y_0 + M_t + \int_0^t \frac{1-2\theta}{2Z_s}\,\mathrm{d}s \tag{15}$$

is a Bessel process with dimension $2-2\theta$. Therefore, it is well known that the weak solution to equation (15) exists up to $\tau^Z = \inf\{t \geq 0 : Z_t = 0\}$, and under the assumption that $Z = 0$ is an absorbing boundary for the process, its transition density should be given by

$$p^Z(t, Z_0, Z_t) = \begin{cases} \dfrac{Y_0^\theta Z_t^{1-\theta}}{t} \exp\left(-\dfrac{Y_0^2 + Z_t^2}{2t}\right) I_\theta\left(\dfrac{Y_0 Z_t}{t}\right), & Z_t > 0; \\ \dfrac{1}{\Gamma(1+\theta)} \Gamma\left(\theta, \dfrac{Y_0^2}{2t}\right), & Z_t = 0. \end{cases} \tag{16}$$

See Borodin and Salminen (2002) for detailed discussions on the Bessel process.

Let $Y_t = Z_{\phi(t)}$ for any $t \geq 0$. It is easy to see that

$$Y_t = Y_0 + \int_0^t \varphi(\gamma)\, \mathrm{d}B_\gamma + \int_0^t (1-2\theta) \frac{\varphi^2(\gamma)}{2Y_\gamma}\, \mathrm{d}\gamma.$$

So far, we have shown the SDE (12) admits a weak solution. Combining with the strong uniqueness, we know that the existence of strong solution to equation (12). Furthermore, by equation (16), we can also see that the transition density of Y is given by equation (13). Then, using (13), similar to the arguments in Appendix 2 of Yang *et al.* (2017), we can derive the cumulative distribution function in equation (14). ■

In the remark below, we compare Lemma 3.4 with Section 4.3 of Yang and Wan (2018) as well as Results 2.2 and 2.4 of Chen *et al.* (2012).

Remark 3.1 In Section 4.3 of Yang and Wan (2018), they achieve a PDE with a small perturbation parameter, in which the infinitesimal generator of a Bessel process is the leading order operator. Using the transition density of a Bessel process, Yang and Wan (2018) solve a hierarchy of PDEs to obtain the asymptotic formulas for the probability that the forward price hits zero. Result 2.2 of Chen *et al.* (2012) reviews the transition density for a squared Bessel process (Borodin and Salminen 2002). Then, using Result 2.2, Chen *et al.* (2012) arrive at Result 2.4, which is originated from Islah (2009). However, the argument of Result 2.4, especially equation (2.17), is not correct because they have overlooked the stopping time τ_0^F when applying Itô's formula. Lemma 3.4 presents the probability density function and cumulative distribution function of a time-changed Bessel process. With the help of Lemma 3.4, we then show in Theorem 2.2 that Result 2.4 of Chen *et al.* (2012) holds with an exponential negligible error.

3.2. *Proof of Theorem 2.1*

Recall $\{X_t; 0 \leq t \leq T\}$ defined in equation (3). Let $\rho^\perp = \sqrt{1-\rho^2}$, and let $\tau_0 = \inf\{t \in [0,T] : X_t = 0\}$ be the first time that the process $\{X_t; 0 \leq t \leq T\}$ hits zero. Rewriting equation (3), then we have

$$X_{T\wedge\tau_0} = X_0 + \frac{\rho}{\nu}(A_{T\wedge\tau_0} - A_0) + \rho^\perp \int_0^{T\wedge\tau_0} A_s\, \mathrm{d}B_s + \int_0^{T\wedge\tau_0} \frac{(1-2\theta)(\rho^\perp A_s)^2}{2X_s}\, \mathrm{d}s. \tag{17}$$

Given a sample path of $\{A_t(\omega) : 0 \leq t \leq T\}$ from

$$\Omega := \left\{\omega : \inf_{s\in[0,T]} \left(X_0 + \frac{\rho}{\nu}(A_s(\omega) - A_0)\right) > \bar{X}_0\right\}, \tag{18}$$

consider a new process $\{\bar{X}_t\}$ satisfying the following SDE

$$\bar{X}_t = \bar{X}_0 + \rho^\perp \int_0^t A_s\, \mathrm{d}B_s + \int_0^t \frac{(1-2\theta)(\rho^\perp A_s)^2}{2\bar{X}_t}\, \mathrm{d}s, \tag{19}$$

where $\bar{X}_0 \in (0, X_0)$. The strong uniqueness and existence up to τ_0 for $\{X_t; 0 \leq t \leq T\}$ is presented in Lemma 3.1. If we specify an absorbing boundary at zero for the SDE (19), Lemma 3.4 indicates that $\{\bar{X}_t; 0 \leq t \leq T\}$ must exist uniquely in a strong sense up to $\bar{\tau}_0$, where

$$\bar{\tau}_0 = \inf\{t \in [0,T] : \bar{X}_t = 0\}.$$

Conditional on Ω defined in equation (18), the initial point $X_0 + \rho/\nu(A_{T\wedge\tau_0} - A_0)$ is larger than $\bar{X}_0$. Similar to the comparison principle (Karatzas and Shreve 1991, Proposition 5.2.18), we have that

$$X_{T\wedge\tau_0\wedge\bar{\tau}_0} \geq \bar{X}_{T\wedge\tau_0\wedge\bar{\tau}_0}. \tag{20}$$

Therefore, conditional on Ω, we have that $\{\tau_0 \leq T\} \cap \Omega \subseteq \{\bar{\tau}_0 \leq T\} \cap \Omega$. More precisely, on the event $\{\tau_0 \leq T\} \cap \Omega$, if $\tau_0 < \bar{\tau}_0$ for some sample paths, then by equation (20), we have $0 = X_{\tau_0} = X_{T\wedge\tau_0\wedge\bar{\tau}_0} \geq \bar{X}_{T\wedge\tau_0\wedge\bar{\tau}_0} = \bar{X}_{\tau_0} > 0$. Contradiction! This implies that

$$\{\tau_0 \leq T\} \cap \Omega = \{\tau_0 \leq T, \tau_0 \geq \bar{\tau}_0\} \cap \Omega \subseteq \{\bar{\tau}_0 \leq T\} \cap \Omega.$$

Combining the above formula and the law of total probability, we have

$$\begin{aligned} \mathbb{P}(\tau_0 \leq T) &= \mathbb{P}(\{\tau_0 \leq T\} \cap \Omega) + \mathbb{P}(\tau_0 \leq T | \Omega^c)\mathbb{P}(\Omega^c) \\ &\leq \mathbb{P}(\bar{\tau}_0 \leq T) + \mathbb{P}(\Omega^c), \end{aligned} \tag{21}$$

where Ω^c is the complementary set of Ω defined in equation (18). From Lemma 3.2, we have

$$\mathbb{P}(\Omega^c) \leq \frac{1_{\{\rho\neq 0\}} 1_{\{C_a<1\}}}{\sqrt{1-C_a}} \frac{\nu\sqrt{T}}{|\ln(1-C_a)|} \exp\left(-\frac{\ln^2(1-C_a)}{2\nu^2 T}\right), \tag{22}$$

where $C_a = \nu(X_0 - \bar{X}_0)/\rho A_0$ if $\rho \neq 0$.

By Lemma 3.4, $\{\bar{X}_t; 0 \leq t \leq T\}$ defined in equation (19) is also a time-changed Bessel process. Moreover,

$$\mathbb{P}(\bar{\tau}_0 \leq T) = \frac{\mathbb{E}[\Gamma(\theta, \bar{X}_0^2/(2\Delta))]}{\Gamma(1+\theta)}.$$

Note that $\max_{x>0} x^\alpha\, \mathrm{e}^{-x}$ is bounded for $\alpha > 0$, then there exists a positive constant C such that

$$\Gamma(\theta, x) = \int_x^\infty z^{\beta/(1-\beta)(1-\rho^2)-1/2}\, \mathrm{e}^{-z}\, \mathrm{d}z < Cx^{-1/2}\, \mathrm{e}^{-Cx}.$$

Combining the above inequality with the Cauchy-Schwartz inequality, then we have

$$\begin{aligned}\mathbb{P}(\bar{\tau}_0 \le T) &< \frac{\mathbb{E}[C\Delta^{1/2}/\bar{X}_0 \cdot \exp(-C\bar{X}_0^2/\Delta)]}{\Gamma(1+\theta)} \\ &\le C(\beta)/\bar{X}_0\sqrt{\mathbb{E}[\Delta]}\sqrt{\mathbb{E}[\exp(-C\bar{X}_0^2/\Delta)]} \\ &< C(\beta)/\bar{X}_0(A_0\sqrt{T}\,\mathrm{e}^{\nu^2T/2})(\mathbb{E}[\exp(-C\bar{X}_0^2/\Delta) \\ &\quad \times (1_{\{\Delta>2\Delta_0\}} + 1_{\{\Delta\le 2\Delta_0\}})])^{1/2}.\end{aligned}$$

Furthermore, taking $\bar{X}_0 = X_0/2$ and applying Lemma 3.3, then we have

$$\begin{aligned}\mathbb{P}(\bar{\tau}_0 \le T) &< C(\beta)/X_0(A_0\sqrt{T}\,\mathrm{e}^{\nu^2T/2})(\mathbb{E}[1_{\{\Delta>2\Delta_0\}}] \\ &\quad + \exp(-CX_0^2(A_0^2T)^{-1}))^{1/2} \\ &< C(\beta)/X_0(A_0\sqrt{T}\,\mathrm{e}^{\nu^2T/2})(\nu\sqrt{T}\exp(-C(\nu\sqrt{T})^{-2}) \\ &\quad + \exp(-CX_0^2(A_0^2T)^{-1}))^{1/2}. \qquad (23)\end{aligned}$$

Therefore, combining equations (21), (22), and (23), we have

$$\mathbb{P}(\tau_0 \le T) \le C_1\sqrt{T}(1 + \mathrm{e}^{\nu^2T/2}(1+\nu\sqrt{T})^{1/2})\exp(-C_2/T), \qquad (24)$$

where

$$C_1 = \max\left\{\frac{1_{\{\rho\neq 0\}}1_{\{C_a<1\}}\nu}{\sqrt{1-C_a}|\ln(1-C_a)|}, \frac{C(\beta)A_0}{X_0}\right\},$$

$$C_2 = \min\left\{\frac{\ln^2(1-C_a)}{4\nu^2}, \frac{C}{2\nu^2}, \frac{CX_0}{2A_0^2}\right\},$$

and $C_a = (\nu X_0/2\rho A_0)1_{\{\rho\neq 0\}}$.

Finally, taking logarithm on both sides of equation (24), we have

$$\limsup_{T\downarrow 0} T\ln\mathbb{P}(\tau_0 \le T) \le -C_1,$$

where $C_1 = C(\nu, \beta, A_0, F_0)$ is a positive constant. The proof completes.

Remark 3.2 If $\beta = 0$, going through the above proof, we find the derivations are still valid. Thus, the conclusion of Theorem 2.1 still holds for the case $\beta = 0$. It is worth to note that the last term on the right-hand side (equation (17)) disappears, and Lemma 3.4 still holds without the drift term in equation (12) in this case.

3.3. *Proof of Theoremp 2.2*

Recall the process $X_t = g(F_t)$ defined in eqaution (17). Given a path of $\{A_t(\omega) : t \in [0,T]\}$, consider a new process $\{\tilde{X}_t; 0 \le t \le T\}$ on the probability space $(\Omega, \mathcal{F})$ to approximate $X_{\cdot\wedge\tau_0}$.

$$\tilde{X}_T = \left(X_0 + \frac{\rho}{\nu}(A_T - A_0)\right)^+ + (1-\rho^2)\int_0^T A_s\,\mathrm{d}B_s + \int_0^T \frac{(1-2\theta)(1-\rho^2)A_s^2}{2\tilde{X}_s}\,\mathrm{d}s. \qquad (25)$$

Note that our construction is feasible because the driving Brownian motion W_t of the volatility process is independent of B_t. Define $\tilde{F}_T := g^{-1}(\tilde{X}_T)$. By Lemma 3.4, we know that the distribution function of $\tilde{F}_T$ is given by equation (6).

We now use the distribution of $\tilde{F}_T = g^{-1}(\tilde{X}_T)$ to approximate the distribution of $F_T = g^{-1}(X_{\tau_0\wedge T})$ determined by equation (17). Note that the distributions of $\tilde{F}_T$ and F_T are exactly the same if the correlation is zero (see, e.g. Islah 2009, Cai *et al.* 2017, Leitao *et al.* 2017). Therefore, we only need to prove the approximation error (9) holds when $\rho \neq 0$.

Let $S_n = \inf\{t \in [0,T] : X_t \le 1/n \text{ or } X_t \ge n\}$, $\tilde{S}_n = \inf\{t \in [0,T] : \tilde{X}_t \le 1/n \text{ or } \tilde{X}_t \ge n\}$, and $\sigma_n = S_n \wedge \tilde{S}_n$. Moreover, $\lim_{n\to\infty}\sigma_n = \tau_0 \wedge \tilde{\tau}_0$ where $\tilde{\tau}_0 = \inf\{t \in [0,T] : \tilde{X}_t = 0\}$. Given a Lipschitz function $h(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ and recalling $g(\cdot)$ in equation (8), the composition $h \circ g^{-1}(\cdot)$ is a locally Lipschitz function. Thus, we have

$$\begin{aligned}\mathbb{E}[|h(F_T) - h(\tilde{F}_T)|] &\le C(\beta,h)\mathbb{E}[|X_T - \tilde{X}_T|\mathbf{1}_{\{\sigma_n>T\}}] \\ &\quad + \mathbb{E}[|h\circ g^{-1}(X_T) - h\circ g^{-1}(\tilde{X}_T)|\mathbf{1}_{\{\sigma_n\le T\}}]. \qquad (26)\end{aligned}$$

Note that

$$\begin{aligned}\mathbb{E}[|X_T - \tilde{X}_T|\mathbf{1}_{\{\sigma_n>T\}}] &= \int_0^T \frac{(1-\rho^2)A_t^2(1-2\theta)}{2} \times \mathbb{E}\left[\left|\frac{1}{X_t} - \frac{1}{\tilde{X}_t}\right|\mathbf{1}_{\{\sigma_n>T\}}\right]\mathrm{d}t \\ &\le \frac{(1-2\theta)(1-\rho^2)n^2}{2} \times \int_0^T A_t^2\mathbb{E}[|X_t - \tilde{X}_t|\mathbf{1}_{\{\sigma_n>T\}}]\,\mathrm{d}t.\end{aligned}$$

By the Gronwall's inequality, we have

$$\mathbb{E}[|X_T - \tilde{X}_T \mid \mathbf{1}_{\{\sigma_n>T\}}] = 0. \qquad (27)$$

Combining equations (26) and (27), and letting $n \to +\infty$, the following inequality holds

$$\begin{aligned}&\mathbb{E}[|h(F_T) - h(\tilde{F}_T)|] \\ &\quad \le \mathbb{E}[|h\circ g^{-1}(X_T) - h\circ g^{-1}(\tilde{X}_T)|\mathbf{1}_{\{\tau_0\wedge\tilde{\tau}_0\le T\}}].\end{aligned}$$

By the Lipschtiz property of $h(\cdot)$ and the definition of $g(\cdot)$ in equation (8), we have

$$\begin{aligned}\mathbb{E}[|h(F_T) - h(\tilde{F}_T)|] &\le C(h)\mathbb{E}[F_T\mathbf{1}_{\{\tau_0\wedge\tilde{\tau}_0\le T\}}] \\ &\quad + C(\beta,h)\mathbb{E}[\tilde{X}_T^{1/(1-\beta)}\mathbf{1}_{\{\tau_0\wedge\tilde{\tau}_0\le T\}}]. \qquad (28)\end{aligned}$$

Recall the definition of θ in equation (4). Define $p := 2\theta(1-\beta) \equiv 1 + \beta\rho^2/(1-\rho^2) > 1$ $(\rho \neq 0)$, and let q satisfy $1/p + 1/q = 1$. By the Hölder inequality, we have

$$\begin{aligned}\mathbb{E}[F_T\mathbf{1}_{\{\tau_0\wedge\tilde{\tau}_0\le T\}}] &\le (\mathbb{E}[F_T^p])^{1/p}\mathbb{P}(\tau_0\wedge\tilde{\tau}_0 \le T)^{1/q} \\ &< C(\beta,\rho,F_0)\mathbb{P}(\tau_0\wedge\tilde{\tau}_0 \le T)^{1/q}, \qquad (29)\end{aligned}$$

$$\mathbb{E}[\tilde{X}_T^{1/(1-\beta)}\mathbf{1}_{\{\tau_0\wedge\tilde{\tau}_0\le T\}}] \le (\mathbb{E}[\tilde{X}_T^{2\theta}])^{1/p}\mathbb{P}(\tau_0\wedge\tilde{\tau}_0 \le T)^{1/q}. \qquad (30)$$

where the second inequality in equation (29) holds because $\mathbb{E}[F_T^p] < \infty$ (Andersen and Piterbarg 2007, Proposition 5.1).

Lemma 3.4 indicates that $\tilde{X}_T$ is a time-changed Bessel process. Letting $\tilde{X}_0 = (X_0 + (\rho/\nu)(A_T - A_0))^+$, then we have

$$\begin{aligned}\mathbb{E}[\tilde{X}_T^{2\theta}] &= \mathbb{E}\left[\int_0^\infty \tilde{X}_T^{2\theta} \frac{\tilde{X}_T}{\Delta}\left(\frac{\tilde{X}_0}{\tilde{X}_T}\right)^\theta \exp\left(-\frac{\tilde{X}_0^2 + \tilde{X}_T^2}{2\Delta}\right)\right. \\ &\quad \left. \times I_\theta\left(\frac{\tilde{X}_0\tilde{X}_T}{\Delta}\right) \mathrm{d}\tilde{X}_T\right] \\ &= \mathbb{E}\left[\tilde{X}_0^{2\theta}\int_0^\infty \frac{1}{2}\left(\frac{z}{\lambda}\right)^{\theta/2} \exp\left(-\frac{\lambda+z}{2}\right)\right. \\ &\quad \left. \times I_\theta\left(\sqrt{z\lambda}\right)\Big|_{\lambda=\tilde{X}_0^2/\Delta} \mathrm{d}z\right] \\ &= \mathbb{E}[\tilde{X}_0^{2\theta}],\end{aligned}$$

where the third equality holds due to the definition of a noncentral chi-square distribution's density function. Furthermore, by the Minkowski inequality, we have

$$\begin{aligned}\mathbb{E}[\tilde{X}_T^{2\theta}] &= \mathbb{E}\left[\left(\left(X_0 - \frac{\rho A_0}{\nu}\right) + \frac{\rho}{\nu}A_T\right)^{2\theta}\right] \\ &\leq \left(\left|X_0 - \frac{\rho A_0}{\nu}\right| + \frac{|\rho|}{\nu}(\mathbb{E}[A_T^{2\theta}])^{1/2\theta}\right)^{2\theta} \\ &< C(\nu,\beta,\rho,A_0,F_0). \qquad (31)\end{aligned}$$

Therefore, combining equations (28), (29), (30), and (31), we have

$$\begin{aligned}\mathbb{E}[|h(F_T) - h(\tilde{F}_T)|] &\leq C(\beta,\nu,A_0,F_0,h)(\mathbb{P}(\tau_0 \leq T) \\ &\quad + \mathbb{P}(\tilde{\tau}_0 \leq T))^{1/q}. \qquad (32)\end{aligned}$$

Noting the formulas (2) and (23) in Theorem 2.1 and its proof, we have the conclusion in equation (9).

Remark 3.3 The conclusion in Theorem 2.2 stills holds if we replace $\tilde{g}(F_0) = (g(F_0) + (\rho/\nu)(A_T - A_0))^+$ with $\tilde{g}(F_0) = |g(F_0) + (\rho/\nu)(A_T - A_0)|$. This is because the probability that $g(F_0) + (\rho/\nu)(A_T - A_0) \leq 0$ is exponentially negligible for small T (see Lemma 3.2).

4. Conclusions

This paper develops the principle of not feeling the boundary for the SABR model to quantify the impact of an absorbing boundary at zero on its probability distribution and the European option price. More precisely, we have the probability of the SABR hitting zero decays to zero exponentially as the time horizon tends to zero. With the help of the principle, we demonstrate that the distribution of the forward price conditional on the volatility can be approximated by that of a time-changed Bessel process with an exponentially negligible error, which provides a theoretical justification for a variety of almost exact simulation algorithms recently emerged in the literature.

Disclosure statement

No potential conflict of interest was reportedby the authors.

Funding

Nan Chen is supported by the Research Grant Council of Hong Kong SAR Government [grant number CUHK14201114]. Nian Yang gratefully acknowledges the financial support of the China National Social Science Fund [grant number 15BJL093].

References

Abate, J. and Whitt, W., The Fourier-series method for inverting transforms of probability distributions. *Queueing Syst.*, 1992, **10**, 5–87.

Andersen, L.B.G. and Piterbarg, V.V., Moment explosions in stochastic volatility models. *Finance Stoch.*, 2007, **11**, 29–50.

Bayer, C., Friz, P. and Loeffen, R., Semi-closed form cubature and applications to financial diffusion models. *Quant. Finance*, 2013, **13**, 769–782.

Borodin, N. and Salminen, P., *Handbook of Brownian Motion-facts and Formulae*, 2nd ed., 2002 (Birkhäuser-Verlag: Basel).

Cai, N., Song, Y. and Chen, N., Exact simulation of the SABR model. *Oper. Res.*, 2017, **65**, 931–951.

Chen, B., Oosterlee, W.C. and van der Weide, H., A low-bias simulation scheme for the SABR stochastic volatility model. *Int. J. Theor. Appl. Finan.*, 2012, **15**, 1250016-1–1250016-37.

Chen, N. and Liu, Y., Sensitivity estimation of SABR model via derivative of random variables. Proceedings of the 2011 Winter Simulation Conference, Vol. 114, issue 3, pp. 3871–3882, 2011.

Delbaen, F. and Shirakawa, H., A note on option pricing for the constant elasticity of variance model. *Asia-Pac. Financ. Markets*, 2002, **9**(2), 95–99.

Doust, P., No-arbitrage SABR. *J. Comput. Finance*, 2012, **15**, 3–31.

Gatheral, J., Hsu, E.P., Laurence, P., Ouyang, C. and Wang, T.-H., Asymptotics of implied volatility in local volatility models. *Math. Finance*, 2012, **22**, 591–762.

Gulisashvili, A., Horvath, B. and Jacquier, A., On the probability of hitting the boundary for Brownian motions on the SABR plane. *Electron. Commun. Probab.*, 2016, **21**, 1–13.

Gulisashvili, A., Horvath, B. and Jacquier, A., Mass at zero in the uncorrelated SABR modeland implied volatility asymptotics. *Quant. Finance*, 2018. doi:10.1080/14697688.2018.1432883.

Hagan, P.S., Kumar, D., Lesniewski, A. and Woodward, D., Managing smile risk. *Wilmott Mag.*, 2002, **1**, 84–108.

Hagan, P.S., Kumar, D., Lesniewski, A. and Woodward, D., Arbitrage-free SABR. *Wilmott Mag.*, 2014, **69**, 60–75.

Hobson, D., Comparison results for stochastic volatility models via coupling. *Finance Stoch.*, 2010, **14**, 129–152.

Hsu, E.P., On the principle of not feeling the boundary for diffusion processes. *J. London Math. Soc.*, 1995, **51**(2), 373–382.

In 't Hout, K.J. and Foulon, S., ADI finite difference schemes for option pricing in the Heston model with correlation. *Int. J. Numer. Anal. Model.*, 2010, **7**(2), 303–320.

Islah, O., Solving SABR in exact form and unifying it with LIBOR market model, 2009. Available online at: https://ssrn.com/abstract=1489428 (accessed 5 December 2013).

Kac, M., On some connections between probability theory and differential and integral equations. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman, pp. 189–215, 1951 (University of California Press: Berkeley).

Kallenberg, O., *Foundations of Modern Probability*, 1997 (Springer-Verlag: New York).

Karatzas, I. and Shreve, S.E., *Brownian Motion and Stochastic Calculus*, 2nd ed., 1991 (Springer-Verlag: New York).
Leitao, A., Grzelak, L.A. and Oosterlee, C.W., On an efficient multiple time step Monte Carlo simulation of the SABR model. *Quant. Finance*, 2017, **17**, 1549–1565.
Lions, P.-L. and Musiela, M., Correlations and bounds for stochastic volatility models. *Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire*, 2007, **24**, 1–16.
Matsumoto, H. and Yor, M., Exponential functionals of Brownian motion, I: probability laws at fixed time. *Probability Surveys*, 2005, **2**, 312–347.
Obłój, J., Fine-tune your smile: correction to Hagan *et al.*. *Wilmott Magazine*, 2008, **35**, 102–104.
Paulot, L., Asymptotic implied volatility at the second order with application to the SABR model. In *Large Deviations and Asymptotic Methods in Finance*, edited by P. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier and J. Teichmann, pp. 37–69, 2015 (Springer-Verlag: Berlin).
Rebonato, R., McKay, K. and White, R., *The SABR/LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives*, 2009 (John Wiley & Sons: Chichester).
Varadhan, S.R.S., Diffusion processes in a small time interval. *Comm. Pure Appl. Math*, 1967, **20**, 659–685.
Yang, N., Chen, N., Liu, Y. and Wan, X., Approximate arbitrage-free option pricing under the SABR model. *J. Econ. Dyn. Control*, 2017, **83**, 198–214.
Yang, N. and Wan, X., The survival probability of the SABR model: asymptotics and application. *Quant. Finance*, 2018. doi:10.1080/14697688.2017.1422083

Appendix 1. Exact simulation of the approximate distribution

This subsection presents a method to simulate the sample path based on the conditional approximate distribution given in Theorem 2.2. Specifically, if $\tilde{F}_T$ is determined by the density in equation (6) conditional on A_0, A_T, and Δ, then the sample of $\tilde{F}_T$ can be generated exactly. The algorithm for the exact simulation of the approximate distribution is from Chen and Liu (2011) and Cai *et al.* (2017).

Step 1. Sampling from the distribution of A_T, given A_0. Recall equation (1), then, we have

$$A_T = A_0 \exp(-\tfrac{1}{2}\nu^2 T + \nu W_T) \overset{d}{=} A_0 \exp(-\tfrac{1}{2}\nu^2 T + \nu\sqrt{T}Z),$$

where Z follows from the standard normal distribution. Thus, we can generate a standard normal random variable $Z \sim \mathcal{N}(0,1)$ instead of A_T.

Step 2. Sampling from Δ, given A_0 and A_T. A Laplace transform inversion-based approach can be used to generate a sample from Δ conditional on A_0 and A_T (Chen and Liu 2011, Section 2.2.2; Cai *et al.* 2017, Section 3.2). More precisely, let $h(x) = (1-\rho^2)/x$ for $x > 0$. Recall Δ defined in equation (7). Denote

$$\begin{aligned} G_h(y) &:= \mathbb{P}(h(\Delta) \le y \,|\, A_0, A_T) \\ &= \mathbb{P}\left(\left(\left(\int_0^T A_t^2 dt \right)^{-1} \le y \,\middle|\, A_0, A_T \right) \right), \quad y \ge 0. \end{aligned}$$

For $\xi > 0$, the Laplace transform of $G_h(\cdot)$ is given by (Matsumoto and Yor 2005, Cai *et al.* 2017)

$$\begin{aligned} \hat{G}_h(\xi) &:= \int_{\mathbb{R}_+} \mathrm{e}^{-\xi y} G_h(y)\,\mathrm{d}y \\ &= \frac{1}{\xi} \exp\left(-\frac{\phi^2_{\ln(A_T/A_0)}(\xi\nu^2/A_0^2) - \ln^2(A_T/A_0)}{2\nu^2 T} \right), \end{aligned}$$

where $\phi_x(\lambda) = \arg\cosh(\lambda\,\mathrm{e}^{-x} + \cosh(x))$, and $\arg\cosh(z) = \ln(z + \sqrt{z^2-1})$, $\cosh(z) = (e^z + e^{-z})/2$. Therefore, we can obtain the function $G_h(\cdot)$ by numerically inverting the Laplace transform $\hat{G}_h(\cdot)$ via some algorithms such as Abate and Whitt (1992).

Generate a sample $U \sim \mathcal{U}(0,1)$ from the standard uniform distribution. Find the root of the equation $G_h(V) = U$. Then $h^{-1}(V) = (1-\rho^2)/V$ is a sample of Δ given A_0 and A_T.

Step 3. Sampling from the approximate distribution $\tilde{F}_T$, given F_0, A_0, A_T, and Δ. Recall the approximate distribution of F_T, conditional on F_0, A_0, A_T, and Δ, given by equation (6). Generate a sample $U \sim \mathcal{U}(0,1)$; if $U \le 1 - Q(\frac{\tilde{g}^2(F_0)}{\Delta}; 2\theta)$, then set $\tilde{F}_T = 0$; otherwise, find $\hat{U}$ which solves

$$\mathbb{P}(\tilde{F}_T \le \hat{U} \,|\, \Delta, A_0, A_T) = 1 - Q\left(\frac{\tilde{g}^2(F_0)}{\Delta}; 2\theta, \frac{g^2(\hat{U})}{\Delta} \right) = U,$$

and then set $\tilde{F}_T = \hat{U}$.

Appendix 2. PDE for the survival probability and call option price without arbitrage

Let $\tau_t^F := \min\{s \ge t : F_s = 0\}$. Consider the following conditional expectation with a payoff function $h(\cdot)$

$$\varphi_h(t,f,a) = \mathbb{E}[h(F_T)\mathbf{1}_{\{\tau_t^F > T\}} \,|\, F_t = f, A_t = a].$$

If $h(F) = (F-K)^+$, then $\varphi_h(t,f,a)$ corresponds to the price of a call without arbitrage (Yang *et al.* 2017, equation (4)). If $h(F) = 1$, then $\varphi_h(t,f,a)$ denotes the probability that the forward price does not hit 0 before T, i.e. the survival probability.

Moreover, the function $\varphi_h(t,f,a)$ is the solution to the following PDE (Yang *et al.* 2017, Theorem 1; Yang and Wan 2018, Theorem 2.1):

$$\frac{\partial \varphi_h}{\partial t} + \frac{1}{2}\left(a^2 f^{2\beta} \frac{\partial^2 \varphi_h}{\partial f^2} + 2\rho\nu a^2 f^{\beta} \frac{\partial^2 \varphi_h}{\partial f \partial a} + \nu^2 a^2 \frac{\partial^2 \varphi_h}{\partial a^2} \right) = 0, \tag{A1}$$

with boundary and terminal conditions

$$\varphi_h(t,0,a) = 0, \quad \varphi_h(T,f,a) = h(f). \tag{A2}$$

To obtain the benchmark for the call option price and the hitting (survival) probability, we numerically solve the PDE (equation (A1)) with boundary and terminal conditions (Equation (A2)). Specifically, we use the Alternative Direction Implicit (ADI) algorithm proposed by In 't Hout and Foulon (2010) to solve the related PDE; We truncate the region for (F,A) to $[0,2] \times [0,2]$ and discretize 2500 and 200 steps for the variable F and A, respectively. The number of steps for time is 500. All the numerical experiments are run in an environment of Matlab R2017b and a PC desktop with Intel(R) Core(TM)2 Quad CPU Q9400@2.66GHZ.