# E-Companion to "Exact Simulation of the SABR Model"

## EC.1. Noncentral Chi-Squared Distribution and Squared Bessel Processes

A *squared Bessel process* is defined as the strong solution to the following SDE:

$$X_t = x + 2\int_0^t \sqrt{X_s}dW_s + \delta t, \tag{EC.1}$$

where $x \geq 0$ and $\delta$ are two constants. The state space of this diffusion is $[0, +\infty)$. Whether the boundary 0 is reachable or not depends on the value of $\delta$. More specifically, when $\delta \geq 2$, the boundary 0 is *entrance-not-exit*; that is, the process cannot reach 0 if it starts from an interior point in $[0, +\infty)$. When $0 < \delta < 2$, the boundary 0 is *non-singular*; that is, the process can reach 0 in a finite time horizon with positive probability and it is possible to start the diffusion from 0. When $\delta \leq 0$, the boundary 0 is *exit-not-entrance*; that is, 0 is reachable from an interior point of $[0, +\infty)$ with positive probability but it is impossible to start the diffusion from 0.

Following Borodin and Salminen (2002), when $0 < \delta < 2$, one can specify the behavior of $X_t$ at the boundary 0 as either absorbing or reflecting once it reaches 0. When $\delta > 2$, or $0 < \delta < 2$ and 0 is specified as a reflecting boundary, the transition density of $X_t$ is known as

$$p(t;x,dy) = \mathbb{P}(X_t \in dy | X_0 = x) = \frac{1}{2t}\left(\frac{y}{x}\right)^{(\delta-2)/4} \exp\left(-\frac{x+y}{2t}\right) I_{\frac{\delta}{2}-1}\left(\frac{\sqrt{xy}}{t}\right) dy \tag{EC.2}$$

for $x > 0, y > 0$. When $\delta \leq 0$, or $0 < \delta < 2$ and 0 is specified as an absorbing boundary,

$$p(t;x,dy) = \mathbb{P}(X_t \in dy | X_0 = x) = \frac{1}{2t}\left(\frac{y}{x}\right)^{(\delta-2)/4} \exp\left(-\frac{x+y}{2t}\right) I_{1-\frac{\delta}{2}}\left(\frac{\sqrt{xy}}{t}\right) dy, \tag{EC.3}$$

for $x > 0, y > 0$. The densities (EC.2) and (EC.3) point to a linkage between the probability law of $X_t$ and the distribution of noncentral chi-squared random variables. Comparing (EC.2) with the density of noncentral chi-squared random variables presented at the end of the introduction, we can see easily that the transitional law of $X_t$ in (EC.2) can be expressed as

$$X_t = t \cdot \chi'^2\left(\delta; \frac{x}{t}\right).$$

The derivation in the case of (EC.3) is not as straightforward as the above case. Schroder (1989) was the first to point out its relationship to the complementary noncentral chi-squared distribution. We summarize it in the following lemma:

**Lemma EC.1.1** *When $\delta \le 0$, or $0 < \delta < 2$ and 0 is an absorbing boundary,*

$$\mathbb{P}[X_t \ge y | X_0 = x] = Q_{\chi'^2}\left(\frac{x}{t}; 2-\delta, \frac{y}{t}\right)$$

*for $y > 0$. The probability that $X_t = 0$ is not zero. In particular,*

$$\mathbb{P}[X_t = 0 | X_0 = x] = 1 - Q_{\chi^2}\left(\frac{x}{t}; 2-\delta\right).$$

Taking the square root of $X_t$ leads us to another well-known process called the *Bessel process.* Itô's lemma implies that $Z_t := \sqrt{X_t}$ follows an SDE such that

$$Z_t = \sqrt{x} + \int_0^t \frac{\delta-1}{2Z_s} ds + W_t.$$

The state space of the process $Z$ is $[0, +\infty)$. It inherits the boundary classification from $X$. When $\delta \ge 2$, the boundary 0 is entrance-not-exit; when $0 < \delta < 2$, 0 is nonsingular; and when $\delta \le 0$, the boundary 0 is exit-not-entrance. The transitional density of $Z_t$ is also obtainable through that of $X_t$. For more discussions on Bessel processes, one may refer to Borodin and Salminen (2002).

## EC.2. On the Proof of Proposition 2.1 and Proposition 6.1

In this section, we shall provide an outline for the proof of Propositions 2.1 and 6.1. We focus on the case of $\beta \in [0, 1)$. Define a transform

$$G(x) = \frac{x^{1-\beta}}{1-\beta}, \qquad \text{for } x \ge 0.$$

Applying Itô's lemma to the transformed process $\{Y_t := G(F_t)\}$ yields

$$Y_T = Y_0 + \sqrt{1-\rho^2}\int_0^T \alpha_u dW_u^{(1)} + \rho\int_0^T \alpha_u dW_u^{(2)} - \int_0^T \frac{\beta\alpha_u^2}{2(1-\beta)Y_u} du. \tag{EC.4}$$

Represent $\int_0^T \alpha_u dW_u^{(2)}$ in terms of $\alpha_T$ based on Eq. (4) and substitute it into Eq. (EC.4). Then

$$Y_T = Y_0 + \frac{\rho}{\nu}(\alpha_T - \alpha_0) + \sqrt{1-\rho^2}\int_0^T \alpha_u dW_u^{(1)} - \int_0^T \frac{\beta\alpha_u^2}{2(1-\beta)Y_u} du. \tag{EC.5}$$

Note that Eq. (EC.4) is not well defined when $Y=0$ and we need additional specifications about its behavior at 0 to complement the definition. Both the absorbing and reflecting boundary specifications of $\{F_t\}$ are translated to the boundary conditions on $\{Y_t\}$ through the transform $G(\cdot)$.

Define

$$V_t = (1-\rho^2)\int_0^t \alpha_u^2 du, \quad \text{for any } t \geq 0,$$

which is strictly increasing in $t$. Denote its inversion by $A_s = \inf\{t \geq 0 : V_t \geq s\}$ for all $s \geq 0$. Then by the time-change for martingales (see, e.g., Karatzas and Shreve 1992, Theorem 3.4.6), the time changed process

$$B_s := \sqrt{1-\rho^2}\int_0^{A_s} \alpha_u dW_u^{(1)}, \quad s \geq 0$$

is a standard Brownian motion, and

$$B_{V_t} = \sqrt{1-\rho^2}\int_0^t \alpha_u dW_u^{(1)}, \quad a.s. \tag{EC.6}$$

Consider the following Bessel process:

$$Z_t = Z_0 + B_t - \int_0^t \frac{\beta}{2(1-\beta)(1-\rho^2)Z_s} ds.$$

If we set $t$ to be $V_T$ in the above SDE,

$$Z_{V_T} = Z_0 + B_{V_T} - \int_0^{V_T} \frac{\beta}{2(1-\beta)(1-\rho^2)Z_s} ds. \tag{EC.7}$$

Applying a change-of-variable $u = A_s$ to the last integral on the RHS of (EC.7) yields

$$\int_0^{V_T} \frac{\beta}{2(1-\beta)(1-\rho^2)Z_s} ds = \int_0^T \frac{\beta\alpha_u^2}{2(1-\beta)Z_{V_u}} du. \tag{EC.8}$$

Combining (EC.6), (EC.7) with (EC.8) results in

$$Z_{V_T} = Z_0 + \sqrt{1-\rho^2}\int_0^T \alpha_u dW_u^{(1)} - \int_0^T \frac{\beta\alpha_u^2}{2(1-\beta)Z_{V_u}} du. \tag{EC.9}$$

Conditional on the whole sample path of $\{\alpha_u, 0 \leq u \leq T\}$ and then letting $Z_0 = Y_0 + \rho(\alpha_T - \alpha_0)/\nu$ in (EC.9), Chen et al. (2012) conclude that $Y_T$ and $Z_{V_T}$ have the same conditional distribution. However, in general this is not correct because $Y_u$ and $Z_{V_u}$ $(0 < u < T)$, which are involved on the

RHSs of (EC.5) and (EC.9), have different conditional distributions for any $0<u<T$, unless $\rho=0$. Specifically, for any $0<u<T$,

$$\begin{aligned} Y_u &= Y_0+\rho\frac{\alpha_u-\alpha_0}{\nu}+\sqrt{1-\rho^2}\int_0^u \alpha_s dW_s^{(1)}-\int_0^u \frac{\beta\alpha_s^2}{2(1-\beta)Y_s}ds, \\ Z_{V_u} &= Y_0+\rho\frac{\alpha_T-\alpha_0}{\nu}+\sqrt{1-\rho^2}\int_0^u \alpha_s dW_s^{(1)}-\int_0^u \frac{\beta\alpha_s^2}{2(1-\beta)Z_{V_s}}ds. \end{aligned} \tag{EC.10}$$

Clearly, only if $\rho=0$, do $Y_u$ and $Z_{V_u}$ (with $Z_0=Y_0+\rho(\alpha_T-\alpha_0)/\nu$) have the same conditional distribution for any $0\le u\le T$. If $\rho\neq 0$, the conditional distribution of $Z_{V_u}$ might be a good approximation of that of $Y_u$ if $T$ is small such that $\alpha_u\approx\alpha_T$.

As noted in Section EC.1 in the e-companion, the transition probability density of $\{Z_t^2\}$ is related to the noncentral chi-squared distribution. According to the discussion therein, there are several possible scenarios. When $\delta:=1-\frac{\beta}{(1-\beta)(1-\rho^2)}\ge 0$ and 0 is specified as a reflecting boundary for the process $\{Z_t^2\}$, the conditional distribution of $Z_{V_T}^2$ with $Z_0=Y_0+\rho(\alpha_T-\alpha_0)/\nu$, given the whole sample path of $\{\alpha_u, 0\le u\le T\}$, is

$$Z_{V_T}^2\sim V_T\cdot\chi'^2\left(\delta;\frac{(Y_0+\rho(\alpha_T-\alpha_0)/\nu)^2}{V_T}\right).$$

or explicitly,

$$\mathbb{P}\left(Z_{V_T}^2\le u\middle|\alpha_u, 0\le u\le T\right)=Q_{\chi'^2}(C^*;\delta,A),$$

where $A=\frac{(Y_0+\rho(\alpha_T-\alpha_0)/\nu)^2}{V_T}$ and $C^*=\frac{u}{V_T}$. Note that the conditional distribution above depends only on $\alpha_0$, $\alpha_T$, $\int_0^T\alpha_s^2 ds$ and $Y_0$. Therefore, it is also the conditional distribution of $Z_{V_T}^2$ with $Z_0=Y_0+\rho(\alpha_T-\alpha_0)/\nu$, given $\alpha_0$, $\alpha_T$, $\int_0^T\alpha_s^2 ds$ and $Y_0$.

Similarly, when $\delta\ge 0$ and 0 is an absorbing boundary for the process $\{Z_t^2\}$, or $\delta<0$, the conditional distribution of $Z_{V_T}^2$ is given by

$$\begin{aligned} \mathbb{P}\left(Z_{V_T}^2\ge u\middle|\alpha_0,\alpha_T,\int_0^T\alpha_u^2du, Y_0\right) &= Q_{\chi'^2}(A;2-\delta,C^*), \quad \textit{for any } u>0; \\ \mathbb{P}\left(Z_{V_T}^2 = 0\middle|\alpha_0,\alpha_T,\int_0^T\alpha_u^2du, Y_0\right) &= 1-Q_{\chi'^2}(A;2-\delta,0)=1-Q_{\chi^2}(A;2-\delta). \end{aligned}$$

Then the conditional distribution of $Y_T$ follows (exactly when $\rho=0$ and approximately otherwise). Utilizing the invertibility of $G$, we obtain the conditional distribution for $F_T$ immediately.

## EC.3. Proof of Proposition 3.1

*Proof.* Conditioning on $\alpha_0$ and $\alpha_T$, we can perform the following transformation for $\int_0^T \alpha_s^2 ds$:

$$\begin{aligned}\int_0^T \alpha_s^2 ds &= \alpha_0^2 \int_0^T \exp\left(-\nu^2 s + 2\nu W_s^{(2)}\right) ds = \alpha_0^2 \int_0^{\nu^2 T} \exp\left(-s + 2\nu W_{s/\nu^2}^{(2)}\right) \frac{ds}{\nu^2} \\ &= \frac{\alpha_0^2}{\nu^2} \int_0^{\nu^2 T} \exp\left(-s + 2B_s\right) ds \equiv \frac{\alpha_0^2}{\nu^2} A_{\nu^2 T}^{(-1/2)},\end{aligned}$$

where $\{B_s \triangleq \nu W_{s/\nu^2}^{(2)} : s \geq 0\}$ is still a standard Brownian motion, $A_t^{(\mu)} = \int_0^t \exp(2B_s^{(\mu)}) ds$, and $B_t^{(\mu)} = B_t + \mu t$. It follows that

$$\alpha_T = \alpha_0 \exp\left(-\frac{\nu^2 T}{2} + \nu W_T^{(2)}\right) = \alpha_0 \exp\left(-\frac{\nu^2 T}{2} + B_{\nu^2 T}\right) = \alpha_0 \exp\left(B_{\nu^2 T}^{(-1/2)}\right).$$

Consequently, the conditional distribution of $\int_0^T \alpha_s^2 ds$, given $\alpha_0$ and $\alpha_T$, is identical to that of $\alpha_0^2 A_{\nu^2 T}^{(-1/2)}/\nu^2$, given $\alpha_0$ and $B_{\nu^2 T}^{(-1/2)} = \ln(\alpha_T/\alpha_0)$. Therefore, by the following result (see (1) on p. 1051 in Barrieu et al. 2004),

$$\mathbb{P}(A_t^{(\mu)} \in du \big| B_t^{(\mu)} = x) = \frac{\sqrt{2\pi t}}{u} \exp\left(\frac{x^2}{2t} - \frac{1 + e^{2x}}{2u}\right) I_0\left(\frac{e^x}{u}\right) f_{e^x/u}(t) du, \tag{EC.11}$$

we obtain

$$\begin{aligned}&\mathbb{P}\left(\int_0^T \alpha_s^2 ds \in dw \middle| \alpha_0, \alpha_T\right) \\ &= \mathbb{P}\left(\frac{\alpha_0^2}{\nu^2} A_{\nu^2 T}^{(-1/2)} \in dw \middle| \alpha_0, B_{\nu^2 T}^{(-1/2)} = \ln\left(\frac{\alpha_T}{\alpha_0}\right)\right) \\ &= \frac{\sqrt{2\pi y}}{u} \exp\left(\frac{x^2}{2y} - \frac{1 + e^{2x}}{2u}\right) I_0\left(\frac{e^x}{u}\right) f_{e^x/u}(y) du \Bigg|_{x = \ln(\alpha_T/\alpha_0), y = \nu^2 T, u = \nu^2 w/\alpha_0^2} \\ &= \frac{\sqrt{2\pi T}\nu}{w} \exp\left\{\frac{1}{2\nu^2}\left(\frac{1}{T}\left[\ln\left(\frac{\alpha_T}{\alpha_0}\right)\right]^2 - \frac{\alpha_0^2 + \alpha_T^2}{w}\right)\right\} I_0\left(\frac{\alpha_0 \alpha_T}{\nu^2 w}\right) f_{\frac{\alpha_0 \alpha_T}{\nu^2 w}}(\nu^2 T) dw,\end{aligned}$$

which is exactly (10). □

## EC.4. Proof of Proposition 3.2

*Proof.* By the same transformation as in the proof for Proposition 3.1 and with the aid of the identity (12), we can obtain

$$\mathbb{E}\left[\exp\left\{-\theta\left(\int_0^T \alpha_s^2 ds\right)^{-1}\right\} \middle| \alpha_0, \alpha_T\right] = \mathbb{E}\left[\exp\left(-\frac{\theta \nu^2}{\alpha_0^2} \frac{1}{A_{\nu^2 T}^{(-1/2)}}\right) \middle| \alpha_0, B_{\nu^2 T}^{(-1/2)} = \ln\left(\frac{\alpha_T}{\alpha_0}\right)\right]$$

$$= \exp\left(-\frac{\phi_x(\lambda)^2 - x^2}{2y}\right)\Bigg|_{x=\ln(\alpha_T/\alpha_0), y=\nu^2 T, \lambda=\theta\nu^2/\alpha_0^2}$$

$$= \exp\left(-\frac{[\phi_{\ln(\alpha_T/\alpha_0)}(\theta\nu^2/\alpha_0^2)]^2 - \ln(\alpha_T/\alpha_0)^2}{2\nu^2 T}\right).$$

Then the Laplace transform of the conditional cdf of $\left(\int_0^T \alpha_s^2 ds\right)^{-1}$ given $\alpha_0$ and $\alpha_T$ is

$$\widehat{L}_h(\theta) = \int_0^{+\infty} e^{-\theta s} L_h(s) ds = \frac{1}{\theta}\int_0^{+\infty} e^{-\theta s} dL_h(s) = \frac{1}{\theta}\exp\left(-\frac{[\phi_{\ln(\alpha_T/\alpha_0)}(\theta\nu^2/\alpha_0^2)]^2 - \ln(\alpha_T/\alpha_0)^2}{2\nu^2 T}\right),$$

where the second equality holds due to the integration by parts. □

## EC.5. The Euler Inversion for (15) When $T$ is Small

When $T$ is small, the exponent on the RHS of (15) could be large, either positive or negative. Here we plan to study this effect on the accuracy of the Euler inversion algorithm as well as the selection of involved algorithm parameters. Note that three algorithm parameters are involved, $M$, $m$ and $n$. $M$ is used to control the discretization error, while $m$ and $n$ are used in the Euler transformation to control the truncation error. As suggested in Abate and Whitt (1992), we fix $M$ to be 20. By (17), we can see that this selection makes the discretization error no greater than $\frac{e^{-M}}{1-e^{-M}} = 2\times 10^{-9}$ regardless of the value of $T$.

The remainder is devoted to studying the selection of $m$ and $n$ (in particular when $T$ is small). Note that in our simulation method, we apply the Euler inversion to evaluate the function $L_h(\cdot)$ that is required for solving the equation $L_h(V) = U$ for $V$. Given a sample $U \sim$Unif $(0,1)$, denote the solution by $V_{m,n}(U)$ to indicate the parameters used in the Euler inversion. We increase $m$ and $n$ simultaneously by $d$ and define the relative change in the solution

$$R(m,n,d;U) := \left|\frac{V_{m+d,n+d}(U) - V_{m,n}(U)}{V_{m,n}(U)}\right|.$$

Then the maximum relative change $\max_{U\in(0,1)} R(m,n,d;U)$ can reflect the accuracy of the Euler inversion with parameters $m$ and $n$. Here we approximate $\max_{U\in(0,1)} R(m,n,d;U)$ by $\max_{i=1,\cdots,999} R(m,n,d;U_i)$, where $U_i = i/1000$.

We set $d=5$ and start with $m=5$ and $n=20$. Table EC.1 shows how the maximum relative change decreases as $m$ and $n$ rise in three cases, $T=1$ (1 year), $T=1/2$ (0.5 year) and $T=1/12$ (1 month). It can be seen that the smaller the value of $T$ is, the more slowly the maximum relative change decreases. In other words, we need larger $m$ and $n$ to achieve the same accuracy as $T$ becomes smaller. The numerical results also indicate that when $T$ is no less than 1 month, which is typical for common financial applications, setting $m=20$ and $n=35$ has been quite accurate to control the maximum relative change under $10^{-7}$. If an even smaller value of $T$ is considered, a similar test as above can be used to make a proper choice of $m$ and $n$.

**Table EC.1** How the maximum relative change decreases as $m$ and $n$ rise in three cases, $T=1$ (1 year), $T=1/2$ (0.5 year) and $T=1/12$ (1 month).

| Maximum Relative Changes in $V_{m,n}(U)$ when $T=1$ | | | | | |
|---|---|---|---|---|---|
| | $(m,n)$ | | | | |
| $\alpha_T$ | (5,20) | (10,25) | (15,30) | (20,35) | (25,40) |
| 0.1 | 2.22E-05 | 2.64E-10 | 5.77E-13 | 6.18E-13 | 6.85E-13 |
| 0.5 | 1.93E-05 | 2.59E-10 | 1.17E-12 | 1.18E-12 | 6.18E-13 |
| 0.9 | 2.22E-05 | 2.65E-10 | 7.15E-13 | 1.32E-12 | 1.08E-12 |
| Maximum Relative Changes in $V_{m,n}(U)$ when $T=1/2$ | | | | | |
| | $(m,n)$ | | | | |
| $\alpha_T$ | (5,20) | (10,25) | (15,30) | (20,35) | (25,40) |
| 0.1 | 2.51E-04 | 3.47E-08 | 1.86E-12 | 2.68E-13 | 1.23E-13 |
| 0.5 | 2.03E-04 | 2.73E-08 | 4.74E-13 | 3.83E-13 | 4.09E-13 |
| 0.9 | 2.51E-04 | 3.47E-08 | 1.80E-12 | 3.92E-13 | 3.12E-13 |
| Maximum Relative Changes in $V_{m,n}(U)$ when $T=1/12$ | | | | | |
| | $(m,n)$ | | | | |
| $\alpha_T$ | (5,20) | (10,25) | (15,30) | (20,35) | (25,40) |
| 0.1 | 8.91E-03 | 7.02E-04 | 4.47E-06 | 1.12E-08 | 1.35E-11 |
| 0.5 | 3.05E-01 | 6.21E-04 | 1.90E-06 | 9.57E-09 | 4.74E-12 |
| 0.9 | 3.10E-01 | 7.02E-04 | 4.47E-06 | 1.59E-08 | 1.36E-11 |

*Notes.* The parameters are $\alpha_0=0.3$, $\nu=0.4$ and $d=5$.

## EC.6. Proof of Theorem 5.1

*Proof.* When $\beta=1$, $F_T$ has a conditional normal distribution. The result (20) then follows immediately from the Black-Scholes formula. When $\beta<1$, given $F_0>0$ and $\alpha_0>0$, the conditional distribution of $F_T$ is given by (6) or (8). Based on the two identities below for the noncentral chi-squared density

$$\frac{\partial}{\partial x}q_{\chi'^2}(x;\mu+2,\lambda) = -\frac{1}{2}q_{\chi'^2}(x;\mu+2,\lambda)+\frac{1}{2}q_{\chi'^2}(x;\mu,\lambda) \quad \text{and}$$

$$\frac{\partial}{\partial \lambda} q_{\chi'^2}(x;\mu,\lambda) = -\frac{1}{2} q_{\chi'^2}(x;\mu,\lambda) + \frac{1}{2} q_{\chi'^2}(x;\mu+2,\lambda),$$

we deduce that for $\mu > 0$,

$$\begin{aligned}\frac{\partial}{\partial \lambda} Q_{\chi'^2}(x;\mu,\lambda) &= \frac{\partial}{\partial \lambda}\int_0^x q_{\chi'^2}(y;\mu,\lambda)dy = -\frac{1}{2}\int_0^x [q_{\chi'^2}(y;\mu,\lambda) - q_{\chi'^2}(y;\mu+2,\lambda)]dy \\ &= -\int_0^x \frac{\partial}{\partial y} q_{\chi'^2}(y;\mu+2,\lambda)dy = -q_{\chi'^2}(x;\mu+2,\lambda).\end{aligned}$$

Therefore the conditional density of $F_T$ is given by

$$\mathbb{P}\left(F_T \in du \middle| F_0, \alpha_0, \alpha_T, \int_0^T \alpha_s^2 ds\right) = -\frac{\partial}{\partial \lambda} Q_{\chi'^2}\left(A; 1+\gamma, C(u)\right) dC(u) = q_{\chi'^2}\left(A; 3+\gamma, C(u)\right) dC(u),$$

where $\gamma = \frac{\beta}{(1-\beta)(1-\rho^2)} > 0$. Note that the European option price can be expressed as

$$\begin{aligned}\mathbb{E}[(F_T - K)^+] &= \mathbb{E}\left[F_T 1_{\{F_T > K\}}\right] - \mathbb{E}\left[K 1_{\{F_T > K\}}\right] \\ &= \mathbb{E}\left[\int_K^{+\infty} u q_{\chi'^2}(A; 3+\gamma, C(u)) dC(u)\right] - K\mathbb{E}\left[Q_{\chi'^2}\left(A; 1+\gamma, C(K)\right)\right] \quad \text{(EC.12)}\end{aligned}$$

where the second equality holds only when $\rho = 0$ and becomes an approximation otherwise. Now we focus on computing the first term on the RHS of (EC.12). By the change of variable $C(u) = x$,

$$\int_K^{+\infty} u q_{\chi'^2}(A; 3+\gamma, C(u)) dC(u) = \int_{C(K)}^{+\infty} u(x) q_{\chi'^2}(A; 3+\gamma, x) dx,$$

where $u(x) \equiv C^{-1}(x) = \left[(1-\rho^2)(1-\beta)^2 \int_0^T \alpha_s^2 ds \cdot x\right]^{\frac{1}{2(1-\beta)}} = [x/C(1)]^{\frac{1}{2(1-\beta)}}$. Noticing that $q_{\chi'^2}(x;\mu,\lambda) = q_{\chi'^2}(\lambda;\mu,x)(\lambda/x)^{\frac{\mu-2}{2}}$, then we obtain

$$\begin{aligned}\int_K^{+\infty} u q_{\chi'^2}(A; 3+\gamma, C(u)) dC(u) &= \int_{C(K)}^{+\infty} \left[\frac{x}{C(1)}\right]^{\frac{1}{2(1-\beta)}} q_{\chi'^2}(x; 3+\gamma, A) \left(\frac{A}{x}\right)^{\frac{\gamma+1}{2}} dx \\ &= C(1)^{-\frac{1}{2(1-\beta)}} A^{\frac{1+\gamma}{2}} \int_{C(K)}^{+\infty} q_{\chi'^2}(x; 3+\gamma, A) x^{-\frac{\rho^2\gamma}{2}} dx \\ &= C(1)^{-\frac{1}{2(1-\beta)}} A^{\frac{1+\gamma}{2}} \Phi^+\left(-\frac{\rho^2\gamma}{2}, C(K); 3+\gamma, A\right),\end{aligned}$$

where $\Phi^+(p,k;\delta,\alpha) \equiv \mathbb{E}[[\chi'^2(\delta,\alpha)]^p 1_{\{\chi'^2(\delta,\alpha)>k\}}]$ is the truncated $p$-th moments of $\chi'^2(\delta,\alpha)$. Substituting the above into (EC.12) yields

$$\mathbb{E}[(F_T - K)^+] = \mathbb{E}\left[C(1)^{-\frac{1}{2(1-\beta)}} A^{\frac{1+\gamma}{2}} \Phi^+\left(-\frac{\rho^2\gamma}{2}, C(K); 3+\gamma, A\right)\right] - K\mathbb{E}\left[Q_{\chi'^2}\left(A; 1+\gamma, C(K)\right)\right].$$

When $\rho = 0$, the above formula is exact and reduces to (21) thanks to $\Phi^+(0,k;\delta,\alpha) = Q_{\chi'^2}(k;\delta,\alpha)$. When $\rho \neq 0$, $\Phi^+(0,k;\delta,\alpha)$ has a series representation (22) (see Carr and Linetsky 2006). □

## EC.7. Proof of Proposition 6.2

*Proof.* When $\rho = 0$, $\delta_0 \in (0,1)$ and 0 is specified as a reflecting boundary, denote by $p(u,\alpha,w)$ the joint density of $F_T$, $\alpha_T$, and $\int_0^T \alpha_s^2 ds$. Then

$$p(u,\alpha,w) = p_{F_T}(u|\alpha,w) \cdot p_{\alpha_T,I_T}(\alpha,w),$$

where $p_{F_T}(u|\alpha,w)$ denotes the pdf of $F_T$ conditional on $\alpha_T = \alpha$ and $\int_0^T \alpha_s^2 ds = w$, and $p_{\alpha_T,I_T}(\alpha,w)$ the joint pdf of $\alpha_T$ and $\int_0^T \alpha_s^2 ds$. It follows that the marginal density of $F_T$ is given by

$$p_{F_T}(u) = \int_0^{+\infty}\int_0^{+\infty} p(u,\alpha,w)d\alpha dw = \int_0^{+\infty}\int_0^{+\infty} p_{F_T}(u|\alpha,w)\cdot p_{\alpha_T,I_T}(\alpha,w)d\alpha dw.$$

It suffices to show that for any $\alpha > 0$ and $w > 0$,

$$\lim_{u\to 0+} p_{F_T}(u|\alpha,w) = +\infty. \tag{EC.13}$$

Then the result of the proposition follows immediately due to Fatou's Lemma. By Proposition 6.1,

$$p_{F_T}(u|\alpha,w) = \frac{d}{du} Q_{\chi'^2}(C_0(u);\delta_0,A_0) = q_{\chi'^2}(C_0(u);\delta_0,A_0)\cdot C_0'(u), \tag{EC.14}$$

where $C_0(u) = \frac{u^{2(1-\beta)}}{(1-\beta)^2 w}$. Since $\beta < 1$, it follows that $C_0(u)$ goes to 0 as $u$ goes to 0. On the other hand, $\delta_0 > 0$ implies that $\frac{\delta_0}{2} - 1 \neq -1, -2, \cdots$. Hence we have (see 9.6.7 on p. 375 of Abramowitz and Stegun 1972)

$$I_{\frac{\delta_0}{2}-1}(x) \sim \frac{1}{\Gamma(\frac{\delta_0}{2})}\left(\frac{x}{2}\right)^{\frac{\delta_0}{2}-1}, \quad \text{as } x \to 0.$$

Accordingly,

$$q_{\chi'^2}(C_0(u);\delta_0,A_0) \sim \frac{1}{2\Gamma(\frac{\delta_0}{2})} e^{-\frac{C_0(u)+A_0}{2}}\left(\frac{C_0(u)}{2}\right)^{\frac{\delta_0}{2}-1}, \quad \text{as } u \to 0. \tag{EC.15}$$

Combining (EC.14) with (EC.15) yields

$$p_{F_T}(u|\alpha,w) \sim \frac{2}{\Gamma(\frac{\delta_0}{2})}\cdot e^{-\frac{C_0(u)+A_0}{2}}\left(\frac{1}{2w}\right)^{\frac{\delta_0}{2}}\cdot\frac{u^{\delta_0(1-\beta)-1}}{(1-\beta)^{\delta_0-1}}.$$

Since $\delta_0 \neq 1$, we obtain $\beta \neq 0$. Thus $\delta_0(1-\beta) - 1 = -2\beta < 0$, which further implies that $\lim_{u\to 0+} p_{F_T}(u|\alpha,w) = +\infty$. Thus (EC.13) is proved. $\square$

## EC.8. Introductions to Some Existing Numerical Methods for Pricing European Options under the SABR Model

In this section, we provide some details of existing numerical methods for pricing European options under the SABR model which have been used in our paper.

### EC.8.1. The Expansion Formula of Hagan et al. (2002)

Consider a European call option on the forward price with the strike price $K$. Applying the singular perturbation techniques in PDE, Hagan et al. (2002) prove that the celebrated Black's formula can still be applied to price this European call option under the SABR model. Specifically, the option price at time 0 is given by

$$\mathbb{E}[(F_T-K)^+|F_0,\alpha_0]=F_0N(d_+)-KN(d_-),$$

where $T$ is the time to maturity of the option, $N(\cdot)$ is the cdf of a standard normal random variable, $d_\pm=\frac{\log(F_0/K)\pm\frac{1}{2}\sigma_{im}^2T}{\sigma_{im}\sqrt{T}}$, and the implied volatility $\sigma_{im}$ is given by the following asymptotic expansion formula (see Hagan et al. 2002, (2.17a)–(2.17c), p. 89)

$$\begin{aligned}\sigma_{im}=&\frac{\alpha_0}{(F_0K)^{(1-\beta)/2}\{1+\frac{(1-\beta)^2}{24}\log^2\frac{F_0}{K}+\frac{(1-\beta)^4}{1920}\log^4\frac{F_0}{K}+\cdots\}}\cdot\frac{z}{x(z)}\\&\cdot\left\{1+\left[\frac{(1-\beta)^2}{24}\frac{\alpha_0}{(F_0K)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha_0}{(F_0K)^{(1-\beta)/2}}+\frac{2-3\rho^2}{24}\nu^2\right]T+\cdots\right\},\end{aligned}$$

where

$$\begin{aligned}z=&\frac{\nu}{\alpha_0}(F_0K)^{(1-\beta)/2}\log\frac{F_0}{K},\\x(z)=&\log\left\{\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right\}.\end{aligned}$$

### EC.8.2. Simulation by the Euler Scheme

Suppose $0=t_0<t_1<\cdots<t_M=T$. Define a time series $\{\tilde{F}_{t_i},i=0,1,\cdots,M-1\}$ by the recursion

$$\tilde{F}_{t_{i+1}}:=\tilde{F}_{t_i}+\alpha_{t_i}\tilde{F}_{t_i}^{\beta}\cdot(W_{t_{i+1}}-W_{t_i}),\quad i\geq 0;\quad \tilde{F}_0=F_0. \tag{EC.16}$$

It can be viewed as the discrete approximation of (1). Notice that the scheme (EC.16) may generate negative values even from a positive initial value. A conventional modification truncates the series at 0 whenever a negative value is generated:

$$\tilde{F}_{t_{i+1}} := \max\{\tilde{F}_{t_i} + \alpha_{t_i}\tilde{F}_{t_i}^{\beta} \cdot (W_{t_{i+1}} - W_{t_i}), 0\}, \quad i \geq 0; \quad \tilde{F}_0 = F_0. \tag{EC.17}$$

However, this introduces an extra distortion of the distribution for the model besides the one caused by the discretization.

### EC.8.3. The Finite Difference Method (FDM)

Define $V(T-t; f, \alpha) = \mathbb{E}[(F_T - K)^+ | F_t = f, \alpha_t = \alpha]$. Then $V(\tau; f, \alpha)$ solves the following PDE

$$\frac{\partial V}{\partial \tau} = \frac{1}{2}\alpha^2 f^{2\beta}\frac{\partial^2 V}{\partial f^2} + \rho\nu\alpha^2 f^{\beta}\frac{\partial^2 V}{\partial f \partial \alpha} + \frac{1}{2}\nu^2\alpha^2\frac{\partial^2 V}{\partial \alpha^2},$$

with boundary condition $V(0; f, \alpha) = (f - \alpha)^+$. The finite difference method aims to find an approximation to the solution, $V_{i,j}^k \approx V(s_k; x_i, y_j)$, on the three dimensional grid points $\{(s_k, x_i, y_j)\}$. In particular, we apply the Yanenko's scheme (see, e.g., Duffy 2006)

$$\begin{aligned}
\frac{\tilde{V}_{ij} - V_{ij}^k}{\Delta s} &= \frac{1}{2}y_j^2 x_i^{2\beta}\Delta_x^2 \tilde{V}_{ij} + \frac{1}{2}\rho\nu y_j^2 x_i^{\beta}\Delta_{xy}V_{ij}^k, \\
\frac{V_{ij}^{k+1} - \tilde{V}_{ij}}{\Delta s} &= \frac{1}{2}\nu^2 y_j^2 \Delta_y^2 V_{ij}^{k+1} + \frac{1}{2}\rho\nu y_j^2 x_i^{\beta}\Delta_{xy}\tilde{V}_{ij},
\end{aligned}$$

where for equally spaced grid points,

$$\begin{aligned}
\Delta_x^2 V_{ij}^k &= \frac{V_{i+1,j}^k - 2V_{ij}^k + V_{i-1,j}^k}{(\Delta x)^2}, \\
\Delta_y^2 V_{ij}^k &= \frac{V_{i,j+1}^k - 2V_{ij}^k + V_{i,j-1}^k}{(\Delta y)^2}, \\
\Delta_{xy} V_{ij}^k &= \frac{V_{i+1,j+1}^k - V_{i+1,j-1}^k - V_{i-1,j+1}^k + V_{i-1,j-1}^k}{4\Delta x \Delta y}.
\end{aligned}$$

There are also some techniques to accelerate the convergence of the finite difference scheme. For instance, we can construct non-uniformly spaced grids in the spatial dimensions and apply Richardson extrapolation in the temporal dimension.

## References


Abate, J., W. Whitt. 1992. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems* **10** 5–88.

Abramowitz, M., I. A. Stegun. 1972. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Courier Dover Publications.

Barrieu, P., A. Rouault, M. Yor. 2004. A study of the Hartman-Watson distribution motivated by numerical problems related to the pricing of Asian options. *J. Appl. Probab.* **41** 1049–1058.

Borodin, A. N., P. Salminen. 2002. *Handbook of Brownian motion–Facts and Formulae*. Springer.

Carr, P., V. Linetsky. 2006. A jump to default extended CEV model: an application of Bessel processes. *Finance and Stochastics* **10** 303–330.

Chen, B., C. W. Oosterlee, H. van der Weide. 2012. A low-bias simulation Scheme for the SABR stochastic volatility model. *Internat. J. Theor. Appl. Finance* **15** 125–161.

Duffy, D. J. 2006. *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach.* John Wiley and Sons.

Hagan, P. S., D. Kumar, A. S. Lesniewski, D. E. Woodward. 2002. Managing smile risk. *Wilmott Magazine* 84–108.

Karatzas, I., S. Shreve. 1992. *Brownian Motion and Stochastic Calculus*. Springer.

Schroder, M. 1989. Computing the constant elasticity of variance option pricing fromula. *J. Finance* **44**(1) 211–219.