Efficient one and multiple time-step simulation of the SABR model

Álvaro Leitao, Lech A. Grzelak and Cornelis W. Oosterlee

Delft University of Technology - Centrum Wiskunde & Informatica

CWI - February 15, 2016

“Our” definition of simulation

- Generate samples from (sampling) stochastic processes.
- The standard approach to sample from a given distribution, Z:

$$F_Z(Z) \stackrel{\mathrm{d}}{=} U \quad \text{thus} \quad z_n = F_Z^{-1}(u_n),$$

- F_Z is the cumulative distribution function (CDF).
- $\stackrel{\mathrm{d}}{=}$ means equality in the distribution sense.
- $U \sim \mathcal{U}([0,1])$ and u_n is a sample from $\mathcal{U}([0,1])$.
- The computational cost depends on inversion F_Z^{-1}.

Outline

SABR model

- The formal definition of the SABR model [5] reads

$$\begin{aligned} \mathrm{d}f(t) &= \sigma(t) f^{\beta}(t) \mathrm{d}W_f(t), \quad f(0) = S_0, \\ \mathrm{d}\sigma(t) &= \alpha\sigma(t)\mathrm{d}W_\sigma(t), \qquad \sigma(0) = \sigma_0, \end{aligned}$$

- $f(t) = S(t)e^{rt}$ is forward price of the underlying asset $S(t)$.
- $\sigma(t)$ is the stochastic volatility.
- $W_f(t)$ and $W_\sigma(t)$ are two correlated Brownian motions
- SABR parameters:
  - The volatility of the volatility, $\alpha > 0$.
  - The CEV elasticity, $0 \leq \beta \leq 1$.
  - The correlation coefficient, ρ ($W_f W_\sigma = \rho t$)

"Exact" simulation of SABR model

- Based on Islah [6], the conditional cumulative distribution function (CDF) of $f(t)$ in a generic interval $[s,t]$, $0 \leq s \leq t \leq T$:

$$Pr\left(f(t) \leq K | f(s) > 0, \sigma(s), \sigma(t), \int_s^t \sigma^2(z)\mathrm{d}z\right) = 1 - \chi^2(a; b, c),$$

where

$$a = \frac{1}{\nu(t)}\left(\frac{f(s)^{1-\beta}}{(1-\beta)} + \frac{\rho}{\alpha}\left(\sigma(t) - \sigma(s)\right)\right)^2,$$

$$c = \frac{K^{2(1-\beta)}}{(1-\beta)^2\nu(t)},$$

$$b = 2 - \frac{1 - 2\beta - \rho^2(1-\beta)}{(1-\beta)(1-\rho^2)},$$

$$\nu(t) = (1-\rho^2)\int_s^t \sigma^2(z)\mathrm{d}z,$$

and $\chi^2(x;\delta,\lambda)$ is the non-central chi-square CDF.

- Exact in the case of $\rho = 0$, an *approximation* otherwise.

Simulation of SABR model

- Simulation of the volatility process, $\sigma(t)|\sigma(s)$:

$$\sigma(t) \sim \sigma(s)\exp(\alpha \hat{W}_\sigma(t) - \frac{1}{2}\alpha^2 t),$$

  where $\hat{W}_\sigma(t)$ is a independent Brownian motion.
- Simulation of the integrated variance process, $\int_s^t \sigma^2(z)\mathrm{d}z|\sigma(t),\sigma(s)$.
- Simulation of the forward process, $f(t)|f(s), \int_s^t \sigma^2(z)\mathrm{d}z, \sigma(t), \sigma(s)$ by inverting the CDF.
- The conditional integrated variance is a challenging part. We propose:
  - Approximate the conditional distribution by using Fourier techniques and copulas.
  - Marginal distribution based on COS method [3].
  - Conditional distribution based on copulas.
  - Improvements for a fast computations.

Distribution of the integrated variance

- Not available.
- For notational convenience, we will use $Y(s,t) := \int_s^t \sigma^2(z)\mathrm{d}z$.
- Discrete equivalent, M monitoring dates:

$$Y(s,t) := \int_s^t \sigma^2(z)\mathrm{d}z \approx \sum_{j=1}^{M} \Delta t \sigma^2(t_j) =: \hat{Y}(s,t)$$

  where $t_j = s + j\Delta t$, $j = 1, \ldots, M$ and $\Delta t = \frac{t-s}{M}$.
- In the logarithmic domain, where we aim to find an approximation of $F_{\log \hat{Y} | \log \sigma(s)}$:

$$F_{\log \hat{Y} | \log \sigma(s)}(x) = \int_{-\infty}^{x} f_{\log \hat{Y} | \log \sigma(s)}(y)\mathrm{d}y,$$

  where $f_{\log \hat{Y} | \log \sigma(s)}$ is the *probability density function* (PDF) of $\log \hat{Y}(s,t) | \log \sigma(s)$.

PDF of the integrated variance

- Equivalent: Characteristic function and inversion (Fourier pair).
- Recursive procedure to derive an approximated $\phi_{\log \hat{Y} | \log \sigma(s)}$.
- We start by defining the logarithmic increment of $\sigma^2(t)$:

$$R_j = \log\left(\frac{\sigma^2(t_j)}{\sigma^2(t_{j-1})}\right), j = 1, \ldots, M$$

- $\sigma^2(t_j)$ can be written:

$$\sigma^2(t_j) = \sigma^2(t_0)\exp(R_1 + R_2 + \cdots + R_j).$$

- We introduce the iterative process

$$\begin{aligned} Y_1 &= R_M, \\ Y_j &= R_{M+1-j} + Z_{j-1}, \quad j = 2, \ldots, M. \end{aligned}$$

with $Z_j = \log(1 + \exp(Y_j))$.

PDF of the integrated variance (cont.)

- $\hat{Y}(s,t)$ can be expressed:

$$\hat{Y}(s,t) = \sum_{i=1}^{M} \sigma^2(t_i)\Delta t = \Delta t \sigma^2(s) \exp(Y_M).$$

- And, we compute $\phi_{\log \hat{Y} | \log \sigma(s)}(u)$, as follows:

$$\phi_{\log \hat{Y} | \log \sigma(s)}(u) = \exp\left(iu \log(\Delta t \sigma^2(s))\right) \phi_{Y_M}(u).$$

- By applying COS method in the support $[\hat{a}, \hat{b}]$:

$$f_{\log \hat{Y} | \log \sigma(s)}(x) \approx \frac{2}{\hat{b}-\hat{a}} \sum_{k=0}^{N-1}{}' C_k \cos\left((x-\hat{a}) \frac{k\pi}{\hat{b}-\hat{a}} \right),$$

with

$$C_k = \Re\left(\phi_{\log \hat{Y} | \log \sigma(s)}\left(\frac{k\pi}{\hat{b}-\hat{a}}\right) \exp\left(-i \frac{\hat{a}k\pi}{\hat{b}-\hat{a}}\right)\right).$$

CDF of the integrated variance

- The CDF of $\log \hat{Y}(s,t) | \log \sigma(s)$:

$$F_{\log \hat{Y} | \log \sigma(s)}(x) = \int_{-\infty}^{x} f_{\log \hat{Y} | \log \sigma(s)}(y) \mathrm{d}y$$

$$\approx \int_{\hat{a}}^{x} \frac{2}{\hat{b} - \hat{a}} \sum_{k=0}^{N-1\,'} C_k \cos\left((y - \hat{a}) \frac{k\pi}{\hat{b} - \hat{a}} \right) \mathrm{d}y.$$

- The efficient computation of $\phi_{Y_M}(u)$ is crucial for the performance of the whole procedure (specially, one-step case).
- The inversion of $F_{\log \hat{Y} | \log \sigma(s)}$ is relatively expensive (unafforable in the multi-step case).

Copula-based simulation of $\int_s^t \sigma^2(z)\mathrm{d}z | \sigma(t), \sigma(s)$

- In order to apply copulas, we need (logarithmic domain):
  - $F_{\log \hat{Y} | \log \sigma(s)}$.
  - $F_{\log \sigma(t) | \log \sigma(s)}$.
  - Correlation between $\log Y(s,t)$ and $\log \sigma(t)$.
- The distribution of $\log \sigma(t) | \log \sigma(s) = z$ is

$$\mathcal{N}\left(\mu_{\log \sigma(t)} + \mathcal{P}_{\log \sigma(t), \log \sigma(s)} \frac{s_{\log \sigma(t)}}{s_{\log \sigma(s)}} (z - \mu_{\log \sigma(t)}), s_{\log \sigma(t)} \sqrt{1 - \mathcal{P}^2_{\log \sigma(t), \log \sigma(s)}}\right),$$

  where all the quantities are known.
- Approximated Pearson's correlation coefficient:

$$\mathcal{P}_{\log Y, \log \sigma(t)} \approx \frac{t^2 - s^2}{2\sqrt{\left(\frac{1}{3}t^4 + \frac{2}{3}ts^3 - t^2 s^2\right)}}.$$

- For some copulas, like Archimedean, Kendall's τ is required:

$$\mathcal{P} = \sin\left(\frac{\pi}{2}\tau\right).$$

Sampling $\int_s^t \sigma^2(z)\mathrm{d}z|\sigma(t), \sigma(s)$: Steps

1. Determine $F_{\log\sigma(t)|\log\sigma(s)}$ and $F_{\log\hat{Y}|\log\sigma(s)}$.
2. Determine the correlation between $\log Y(s,t)$ and $\log\sigma(t)$.
3. Generate correlated uniform samples, $U_{\log\sigma(t)|\log\sigma(s)}$ and $U_{\log\hat{Y}|\log\sigma(s)}$ by means of copula.
4. From $U_{\log\sigma(t)|\log\sigma(s)}$ and $U_{\log\hat{Y}|\log\sigma(s)}$ invert original marginal distributions.
5. The samples of $\sigma(t)|\sigma(s)$ and $Y(s,t) = \int_s^t \sigma^2(z)\mathrm{d}z|\sigma(t), \sigma(s)$ are obtained by taking exponentials.

One time-step simulation of the SABR model

- $s = 0$ and $t = T$, with T the maturity time.
- The use is restricted to price European options up to $T = 2$.
- $\log \sigma(s)$ becomes constant.
- $F_{\log \sigma(t) | \log \sigma(s)}$ and $F_{\log \hat{Y} | \log \sigma(s)}$ turn into $F_{\log \sigma(T)}$ and $F_{\log \hat{Y}(T)}$.
- The computation of $\phi_{\log \hat{Y}(T)}$ is much simpler and very fast.
- The approximated Pearson's coefficient results in a constant value:

$$\mathcal{P}_{\log Y(T), \log \sigma(T)} \approx \frac{T^2}{2\sqrt{\frac{1}{3}T^4}} = \frac{\sqrt{3}}{2}.$$

Approximated correlation

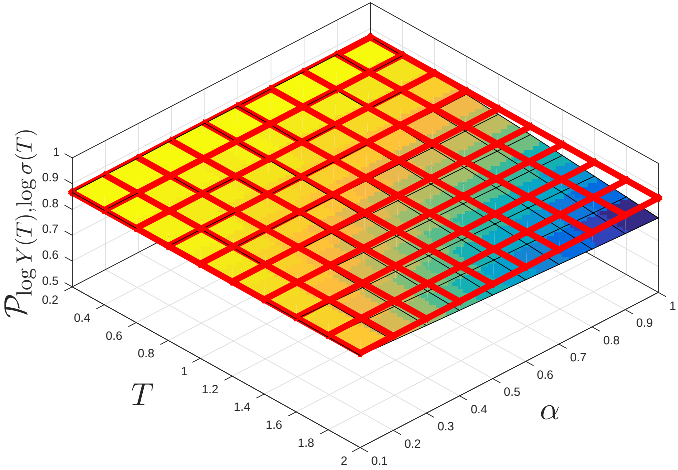


Figure: Pearson's coefficient: Empirical (surface) vs. approximation (red grid).

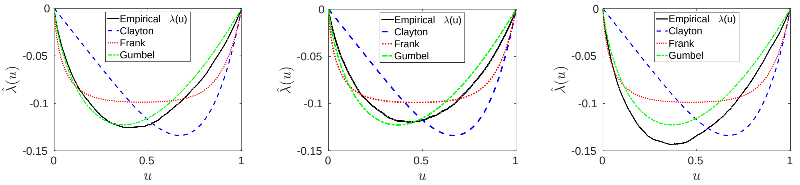
Copula analysis

- Based on the one-step simulation, a copula analysis is carried out.
- Gaussian, Student **t** and Archimedean (Clayton, Frank and Gumbel).
- A *goodness-of-fit (GOF)* for copulas needs to be evaluated.
- Archimedean: graphic GOF based on Kendall's processes.
- Generic GOF based on the so-called *Deheuvels or empirical copula*.

	S_0	σ_0	α	β	ρ	T
Set I	1.0	0.5	0.4	0.7	0.0	2
Set II	0.05	0.1	0.4	0.0	-0.8	0.5
Set III	0.04	0.4	0.8	1.0	-0.5	2

Table: Data sets.

GOF - Archimedean



(a) Set I. (b) Set II. (c) Set III.

Figure: Archimedean GOF test: $\hat{\lambda}(u)$ vs. empirical $\lambda(u)$.

	Clayton	Frank	Gumbel
Set I	1.3469×10^{-3}	2.9909×10^{-4}	5.1723×10^{-5}
Set II	1.0885×10^{-3}	2.1249×10^{-4}	8.4834×10^{-5}
Set III	2.1151×10^{-3}	7.5271×10^{-4}	2.6664×10^{-4}

Table: MSE of $\hat{\lambda}(u) - \lambda(u)$.

Generic GOF

	Gaussian	Student $\mathbf{t}$ $(\nu = 5)$	Gumbel
Set I	5.0323×10^{-3}	5.0242×10^{-3}	3.8063×10^{-3}
Set II	3.1049×10^{-3}	3.0659×10^{-3}	4.5703×10^{-3}
Set III	5.9439×10^{-3}	6.0041×10^{-3}	4.3210×10^{-3}

Table: Generic GOF: D_2.

- The three copulas perform very similarly.
- For longer maturities: Gumbel performs better.
- The Student **t** copula is discarded: very similar to the Gaussian copula and the calibration of the ν parameter adds extra complexity.
- As a general strategy, the Gumbel copula is the most robust choice.
- With short maturities, the Gaussian copula may be a satisfactory alternative.

Pricing European options

- The strike values X_i are chosen following the expression:

$$X_i(T) = f(0)\exp(0.1 \times T \times \delta_i),$$
$$\delta_i = -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5.$$

- Forward asset, $f(T)$: Bin Chen's enhanced inversion [2].
- Convergence and execution time in term of number of samples, n:

	$n = 1000$	$n = 10000$	$n = 100000$	$n = 1000000$
	Gaussian (Set I, X_1)			
Error	519.58	132.39	37.42	16.23
Time	0.3386	0.3440	0.3857	0.5733
	Gumbel (Set I, X_1)			
Error	151.44	−123.76	34.14	11.59
Time	0.3492	0.3561	0.3874	0.6663

Table: Convergence in n: error (basis points) and time (*sec.*).

Pricing European options - Implied volatilities

Strikes	X_1	X_2	X_3	X_4	X_5	X_6	X_7
	Set I (Reference: Antonov [1])						
Hagan	55.07	52.34	50.08	N/A	47.04	46.26	45.97
MC	23.50	21.41	19.38	N/A	16.59	15.58	14.63
Gaussian	16.23	20.79	24.95	N/A	33.40	37.03	40.72
Gumbel	11.59	15.57	19.12	N/A	25.41	28.66	31.79
	Set II (Reference: Korn [7])						
Hagan	−558.82	−492.37	−432.11	−377.47	−327.92	−282.98	−242.22
MC	5.30	6.50	7.85	9.32	10.82	12.25	13.66
Gaussian	9.93	9.98	10.02	10.20	10.57	10.73	11.04
Gumbel	−9.93	−9.38	−8.94	−8.35	−7.69	−6.83	−5.79
	Set III (Reference: MC Milstein)						
Hagan	287.05	252.91	220.39	190.36	163.87	141.88	126.39
Gaussian	16.10	16.76	16.62	15.22	13.85	12.29	10.67
Gumbel	6.99	3.79	0.67	−2.27	−5.57	−9.79	−14.06

Table: Implied volatility: errors in basis points.

- One-step SABR simulation is a fast alternative to Hagan formula.
- Overcomes the known issues, like low strikes and high volatilities.
- For long maturities: multiple time-step version.

Multiple time-step simulation of the SABR model

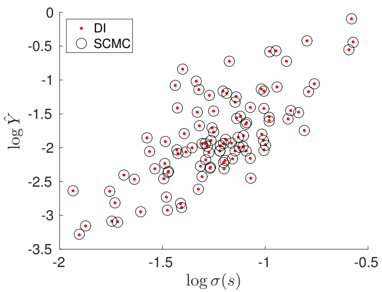
- In intermediate steps, $\phi_{\log \hat{Y}|\log \sigma(s)}$ becomes "stochastic".
- $f_{\log \hat{Y}|\log \sigma(s)}$ needs to be computed for each sample of $\log \sigma(s)$.
- Consequently, the inversion of $F_{\log \hat{Y}|\log \sigma(s)}$ is unafforable ($n \uparrow\uparrow$).
- Solution: *Stochastic Collocation Monte Carlo* (SCMC) sampler [4].

$$y_n|v_n \approx g_{L_{\hat{Y}},L_\sigma}(x_n) = \sum_{i=1}^{L_{\hat{Y}}} \sum_{j=1}^{L_\sigma} F^{-1}_{\log \hat{Y}|\log \sigma(s)=v_j}(F_X(x_i))\ell_i(x_n)\ell_j(v_n),$$

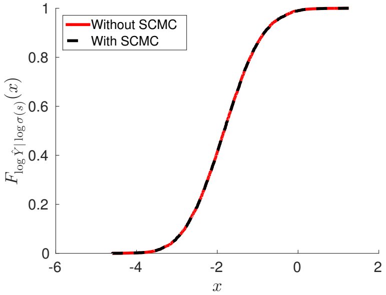
where x_n are the samples from the *cheap variable*, X, and v_n the given samples of $\log \sigma(s)$. x_i and v_j are the *collocation points* of X and $\log \sigma(s)$, respectively. ℓ_i and ℓ_j are the Lagrange polynomials defined by

$$\ell_i(x_n) = \prod_{k=1,k\neq i}^{L_{\hat{Y}}} \frac{x_n - x_k}{x_i - x_k}, \quad \ell_j(v_n) = \prod_{k=1,k\neq j}^{L_\sigma} \frac{v_n - v_k}{v_i - v_k}.$$

Application of 2D SCMC to $F_{\log \hat{Y} | \log \sigma(s)}$



(a) $\log \hat{Y} | \log \sigma(s)$ - Direct inversion vs. SCMC.



(b) $F_{\log \hat{Y} | \log \sigma(s)}(x)$.

Samples	Without SCMC	With SCMC		
		$L_{\hat{Y}} = L_\sigma = 3$	$L_{\hat{Y}} = L_\sigma = 7$	$L_{\hat{Y}} = L_\sigma = 11$
100	1.0695	0.0449	0.0466	0.0660
10000	16.3483	0.0518	0.0588	0.0798
1000000	1624.3019	0.2648	0.5882	1.0940

Table: SCMC time.

Multi-step SABR simulation - Pricing

- Data sets.

	S_0	σ_0	α	β	ρ	T
Set I	1	0.3	0.5	1.0	-0.8	5
Set II	0.5	0.5	0.4	0.5	0.0	2
Set III	0.035	0.01	0.5	0.0	0.0	30

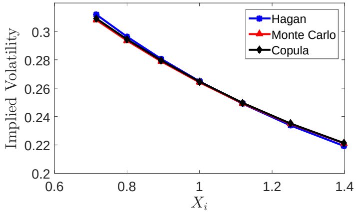
- Convergence in term of number of time-steps, m (Set II).

Strikes	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Antonov[1]	75.51%	74.18%	72.90%	N/A	70.47%	69.32%	68.22%
Copula ($m = T/2$)	72.43%	71.45%	70.49%	69.55%	68.63%	67.74%	66.88%
Diff.(bp)	-307.56	-272.86	-240.61	N/A	-183.50	-158.05	-134.38
Copula ($m = T$)	74.65%	73.43%	72.25%	71.11%	70.00%	68.93%	67.91%
Diff.(bp)	-85.84	-74.88	-64.74	N/A	-46.65	-38.66	-31.33
Copula ($m = 2T$)	75.43%	74.14%	72.89%	71.68%	70.51%	69.39%	68.31%
Diff.(bp)	-8.00	-4.70	-1.39	N/A	4.13	6.44	8.58
Copula ($m = 12T$)	75.55%	74.22%	72.93%	71.68%	70.48%	69.33%	68.23%
Diff.(bp)	4.45	3.70	2.82	N/A	1.58	0.92	0.36

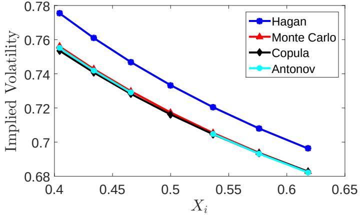
Table: Implied volatility: Antonov vs. Copula.

Pricing - Implied volatilities

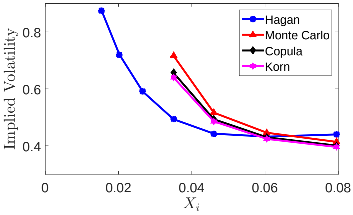
- $n = 1000000$ and $m = 2T$:



(c) Implied volatility set I.



(d) Implied volatility set II.



(e) Implied volatility set III.

Conclusions

- We propose an efficient SABR simulation based on Fourier and copula techniques.
- The one-step SABR is a fast alternative to Hagan formula for short maturities.
- Overcomes the known issues of Hagan's expression.
- When long maturities are considered, multi-step version.
- High accuracy with very few number of time-steps.

References

Alexandre Antonov, Michael Konikov, and Michael Spector.
SABR spreads its wings.
Risk Magazine, pages 58–63, August 2013.

Bin Chen, Cornelis W. Oosterlee, and Hans van der Weide.
A low-bias simulation scheme for the SABR stochastic volatility model.
International Journal of Theoretical and Applied Finance, 15(2), 2012.

Fang Fang and Cornelis W. Oosterlee.
A novel pricing method for European options based on Fourier-cosine series expansions.
SIAM Journal on Scientific Computing, 31:826 – 848, November 2008.

Lech A. Grzelak, Jeroen A. S. Witteveen, M. Suárez-Taboada, and Cornelis W. Oosterlee.
The Stochastic Collocation Monte Carlo Sampler: Highly efficient sampling from "expensive" distributions.
Preprint, 2014.

Patrick S. Hagan, Deep Kumar, Andrew S. Lesniewski, and Diana E. Woodward.
Managing smile risk.
Wilmott Magazine, pages 84–108, 2002.

Othmane Islah.
Solving SABR in exact form and unifying it with LIBOR market model, 2009.
Available at http://ssrn.com/abstract=1489428.

Ralf Korn and Songyin Tang.
Exact analytical solution for the normal SABR model.
Wilmott Magazine, 2013(66):64–69, 2013.

Álvaro Leitao, Lech A. Grzelak, and Cornelis W. Oosterlee.
On a one time-step SABR simulation approach: Application to European options.
Submitted to Applied Mathematics and Computation, 2016.

Questions

Thank you for your attention