- Confidence interval: there is only one parameter to reduce the simulation size M: the variance $\sigma_X^2 = \mathrm{Var}(X)$;
- Idea: Variance reduction !
- If X and X' satisfy

$$\mathbb{E}\,X' = \mathbb{E}\,X \quad \text{and} \quad \mathrm{Var}(X') < \mathrm{Var}(X),$$

then switch

$$X \rightsquigarrow X'$$

since

$$M' \simeq \frac{\sigma^2_{X'}}{\varepsilon^2} q_\alpha^2 < M \simeq \frac{\sigma^2_{X}}{\varepsilon^2} q_\alpha^2.$$

- To be continued... Let us come back for a whole to simulation...

d-dimensional standard normal vectors

Theorem (Box-Muller method)

Let R^2 and $\Theta : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ be two independent r.v. with distributions $\mathcal{L}(R^2) = \mathcal{E}(\frac{1}{2})$ and $\mathcal{L}(\Theta) = U([0, 2\pi])$ respectively. Then

$$X := (R\cos\Theta, R\sin\Theta) \stackrel{d}{=} \mathcal{N}(0, I_2)$$

where $R := \sqrt{R^2}$.

Corollary (Box-Muller method on a computer)

One can simulate a distribution $\mathcal{N}(0; I_2)$ from a couple (U_1, U_2) of independent random variable with distribution $\mathcal{U}([0, 1])$ by setting

$$X := \left(\sqrt{-2\log(U_1)}\cos(2\pi U_2), \sqrt{-2\log(U_1)}\sin(2\pi U_2)\right).$$

Simulation of the multivariate distribution $\mathcal{N}(0; I_d)$

- Let $(U_1, \ldots, U_d) \stackrel{d}{=} \mathcal{U}([0,1]^d)$ be random vector (so that $U_1, \ldots, U_d$ are i.i.d. with distribution $\mathcal{U}([0,1])$).

- Set for $i = 1, \ldots, d/2$,

$$(X^{2i-1}, X^{2i}) = \sqrt{-2\log(U_{2i-1})}\,(\cos(2\pi U_{2i}), \sin(2\pi U_{2i}))\,.$$

- The second popular method to simulate bi-variate normal distributions is the Polar method due to Marsaglia.

Theorem (Marsaglia's Polar method)

Let $(V_1, V_2) \stackrel{d}{=} U(B(0;1))$ where $B(0,1)$ denotes the canonical Euclidean unit ball in $\mathbb{R}^2$. Set $R^2 = V_1^2 + V_2^2$. Show that

$$X := \sqrt{-2\log(R^2)/R^2}\Big(V_1, V_2\Big) \stackrel{d}{=} \mathcal{N}(0; I_2).$$

Correlated d-dimensional Gaussian vectors, Gaussian processes

- Let $X = (X^1, \ldots, X^d)$ be a centered $\mathbb{R}^d$-valued Gaussian vector with covariance matrix $\Sigma = \big[\mathrm{Cov}(X^i, X^j)\big]_{1\leqslant i,j\leqslant d} \in \mathcal{S}(d, \mathbb{R})$.

$$\Phi_X(u) := \mathbb{E}\left(e^{i(u|X)}\right) = e^{-\frac{1}{2}u^*\Sigma u}, \quad u \in \mathbb{R}^d.$$

- As a well-known consequence, the covariance matrix Σ characterizes the distribution of X, hence the notation $\mathcal{N}(0; \Sigma)$.

Lemma (Key lemma)

Let $Y : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^q$ be an $L^2(\mathbb{P})$-random vector with covariance matrix $C_Y = \big[\mathrm{Cov}(Y^i, Y^j)\big]_{1\leqslant i,j\leqslant q}$ and let $A = [a_{ij}]_{1\leqslant i\leqslant d, 1\leqslant j\leqslant q} \in \mathcal{M}(d, q)$ be a $d \times q$ matrix. Then the $\mathbb{R}^d$-valued random vector AY has a covariance matrix C_{AY} given by

$$C_{AY} = AC_Y A^*$$

where $A^* = [a_{ji}]$ stands for the transpose of A.

This result can be used in two different ways.

- *Square root of Σ.* There exists a unique non-negative symmetric matrix $\sqrt{\Sigma}$ s.t. $\Sigma = \sqrt{\Sigma}^2 = \sqrt{\Sigma}\sqrt{\Sigma}^*$. Morever, $\sqrt{\Sigma}$ commutes with Σ. Consequently

$$\text{If} \qquad Z \stackrel{d}{=} \mathcal{N}(0; I_d) \qquad \text{then} \qquad \sqrt{\Sigma}Z \stackrel{d}{=} \mathcal{N}(0; \Sigma).$$

  Computing $\sqrt{\Sigma}$ needs diagonalization: it is costly.

- *Cholesky decomposition of Σ.* When the covariance matrix Σ is invertible (*i.e.* definite), one may decompose Σ as

$$\Sigma = T\,T^* \text{ with } T \text{ lower triangular matrix}$$

  (*i.e.* such that $T_{ij} = 0$ if $i < j$). Then,

$$T\,Z \stackrel{d}{=} \mathcal{N}(0; \Sigma).$$

  Cholesky's decomposition is the matrix version of the Hilbert-Schmidt orthonormalization procedure: there exists *a unique such lower triangular matrix with $T_{ii} > 0$ $i = 1, \ldots, d$ (diagonal terms)* called the Cholesky matrix of Σ.

- The Cholesky based procedure is more performing: it approximately divides the complexity of the simulation by a 2 factor.

Simulation of the standard Brownian motion at fixed times

- Let $W = (W_t)_{t \geqslant 0}$ be a standard Brownian motion defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ i.e. a centered Gaussian process with covariance given by

$$C^W(s,t) = \mathbb{E}\,(W_s W_t) = s \wedge t.$$

- Let $0 \leqslant t_1 < t_2 < \ldots, t_{n-1} < t_n$ be an increasing sequence of n instants. A first method relies on the Cholesky decomposition of the covariance structure of the Gaussian vector $(W_{t_1}, \ldots, W_{t_n})$ given by

$$\Sigma^W_{(t_1,\ldots,t_n)} = [t_i \wedge t_j]_{1 \leqslant i,j \leqslant n}.$$

- However, it may seem more natural to use the independence and the stationarity of the increments *i.e.* which imply

$$\Big(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}\Big) \stackrel{d}{=} \mathcal{N}\Big(0; \mathrm{Diag}(t_1, t_2 - t_1, \ldots, t_n - t_{n-1})\Big) \ldots$$

- ...so that

$$\begin{bmatrix} W_{t_1} \\ W_{t_2} - W_{t_1} \\ \vdots \\ W_{t_n} - W_{t_{n-1}} \end{bmatrix} \stackrel{d}{=} \mathrm{Diag}\big(\sqrt{t_1}, \sqrt{t_2 - t_1}, \ldots, \sqrt{t_n - t_{n-1}}\big) \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$$

  where $(Z_1, \ldots, Z_n) \stackrel{d}{=} \mathcal{N}\Big(0; I_n\Big)$.

- The simulation of $(W_{t_1}, \ldots, W_{t_n})$ follows by summing up the increments.

- So what? Noting that

$$\begin{bmatrix} W_{t_1} \\ W_{t_2} \\ \vdots \\ W_{t_n} \end{bmatrix} = L \begin{bmatrix} W_{t_1} \\ W_{t_2} - W_{t_1} \\ \vdots \\ W_{t_n} - W_{t_{n-1}} \end{bmatrix} \qquad \text{where} \qquad L = \big[\mathbf{1}_{\{i \geqslant j\}}\big]_{1 \leqslant i,j \leqslant n}$$

  is lower triangular ...

…and if we set

$$T = L\,\mathrm{Diag}\big(\sqrt{t_1}, \sqrt{t_2 - t_1}, \ldots, \sqrt{t_n - t_{n-1}}\big) = \left[\sqrt{t_j - t_{j-1}}\mathbf{1}_{\{i \geqslant j\}}\right]_{1 \leqslant i,j \leqslant n}$$

one checks that

$$\begin{bmatrix} W_{t_1} \\ W_{t_2} \\ \vdots \\ W_{t_n} \end{bmatrix} = T \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}.$$

We derive, owing to the key lemma

$$TT^* = TI_nT^* = \Sigma^W_{t_1,\ldots,t_n}.$$

that is the Cholesky decomposition of the covariance matrix $\Sigma^W_{t_1,\ldots,t_n}$!

ℵ Practitioner's corner (Warning!)

- When modeling the dynamics of several risky assets, say d, by a d-dimensional Black-Scholes model one often faces this kind of formula

$$dX_t^i = X_t^i\big(rdt + \sum_{\ell=1}^{q} \sigma_{i\ell} dW_t^\ell\big), \quad i = 1\ldots, d$$

  where $W = (W^1, \ldots, W^q)$ is a standard q-dimensional Brownian motion.

- This can be rewritten

$$dX_t^i = X_t^i\big(rdt + |\sigma_{i\cdot}| dB_t^i\big), \quad i = 1\ldots, d$$

  where, $\forall\, i \in \{1, \ldots, d\}$, $\sigma_{i.}$ is the column vector $[\sigma_{ij}]_{1\leqslant j\leqslant q}$, $|\sigma_{i\cdot}|$ its Euclidean norm and

$$B_t^i = \sum_{\ell=1}^{q} \frac{\sigma_{i\ell}}{|\sigma_{i\cdot}|} W_t^\ell.$$

  Then $B = (B^1, \ldots, B^d)$ are the correlated normalized stochastic sources attached to the log-return $\log(X_t^i)$.

- The $d-$dim process B is in fact a correlated Brownian motion as a continuous centered Gaussian process with independent increments such that

$$\mathrm{Var}(B_t^i) = \sum_{\ell,\ell'=1}^{q} \frac{\sigma_{i\ell}\sigma_{i\ell'}}{|\sigma_{i.}|^2} \underbrace{\mathrm{Cov}(W_t^\ell, W_t^{\ell'})}_{=0 \text{ if } \ell\neq\ell',\, =t \text{ if } \ell=\ell'} = \frac{t}{|\sigma_{i.}|^2}\sum_{\ell=1}^{q}\sigma_{i\ell}^2 = t$$

- Likewise, the normalized covariance matrix of B (at time 1) is then given by

$$\mathrm{Cov}\big(B_t^i, B_t^j,\big) = t\sum_{\ell=1}^{q}\frac{\sigma_{i\ell}\sigma_{j\ell}}{|\sigma_{i.}||\sigma_{j.}|} = t\frac{\big(\sigma_{i.}|\sigma_{j.}\big)}{|\sigma_{i.}||\sigma_{j.}|} = \underbrace{\frac{(\sigma\sigma^*)_{ij}}{|\sigma_{i.}||\sigma_{j.}|}}_{=:\rho_{ij}\in[-1,1]}\, t.$$

where $(.|.)$ denotes here the canonical inner product on $\mathbb{R}^q$.

- $R := \Sigma_{B_1} = [\rho_{ij}]_{1\leqslant i,j\leqslant d}$ is called the correlation matrix of the Brownian motion B.

- So, provided $q \geqslant d$ and prior to the simulation, one can (should!) perform a Cholesky decomposition on the symmetric non-negative matrix
$$R = \left[\frac{(\sigma_{i.}|\sigma_{j.})}{|\sigma_{i.}||\sigma_{j.}|}\right]_{1\leqslant i,j\leqslant d}$$
to speed up the simulation (if $\sigma\sigma^*$ is definite!).
- If $R = TT^*$, T lower triangular, $T_{ii} > 0$, then $\widetilde{B} := T\widetilde{W}$, $\widetilde{W}$ standard d-dim B.M., is a correlated d-dim Brownian motion with the same covariance matrix as B. Hence B and $\widetilde{B}$ have the same distribution but is faster to simulate.
- However, have in mind that, if $q < d$, then Σ_{B_1} has rank at most q and cannot be invertible so that Cholesky decomposition is not possible (at least in the present form).

Simulation of the Fractional Brownian motion at fixed times.

- The fractional Brownian motion (fBm) $W^H = (W_t^H)_{t\geqslant 0}$ with Hurst coefficient $H \in (0,1)$ is defined as a centered Gaussian process with a correlation function given for every $s,\ t \in \mathbb{R}_+$ by

$$C^{W^H}(s,t) = \frac{1}{2}\Big(t^{2H} + s^{2H} - |t-s|^{2H}\Big).$$

- When $H = \frac{1}{2}$, W^H is simply a standard B.M.
- When $H \neq \frac{1}{2}$, W^H has none of the usual properties of the Brownian motion (except the stationarity of its increments and some self-similarity properties):
  - it has dependent increments,
  - it is not a martingale (or even a semi-martingale),
  - it is not either a Markov process.

- To simulate the fBm W^H at times $t_1, \ldots, t_n$ use the Cholesky decomposition of its covariance matrix

$$[C^{W^H}(t_i, t_j)]_{1 \leqslant i,j \leqslant n}.$$

But be careful, this matrix may become ill-conditioned if n is large and $t_k = t_k^n = \frac{kT}{n}$, $t_k = \frac{kT}{n}$, $k = 1, \ldots, n$, which may induce instability in the Cholesky transform.

- This should be out in balance with the fact that such a computation can be made only once and offline to be stored (at least for usual values of the t_i like the above $t_k = \frac{kT}{n}$, $k = 1, \ldots, n$).
- Other methods have been introduced in which the covariance matrix is embed in a circuit matrix. It relies on a (fast) Fourier transform procedure.

Pricing a *Best-of Call* by a Monte Carlo simulation

- $2D$-Black-Scholes model: interest rate r, volatility $\sigma > 0$.

$$X_t^i = x_0^i \exp\left((r - \frac{\sigma_i^2}{2})t + \sigma_i W_t^i\right),\ i = 1, 2, \quad \langle W^1, W^2\rangle_t = \rho t,\ \rho \in [-1, 1]$$

- Best-of Call: $h_T = \big(\max(X_T^1, X_T^2) - K\big)^+$.

$$(W_T^1, W_T^2) \stackrel{d}{=} \sqrt{T}(Z^1, \rho Z^1 + \sqrt{1-\rho^2}Z^2),\ (Z^1, Z^2) \stackrel{d}{=} \mathcal{N}(0; I_2).$$

so that

$$\begin{aligned} e^{-rT} h_T &= \varphi(Z^1, Z^2) \\ &:= \left(\max\left(x_0^1 \exp\left(-\frac{\sigma_1^2}{2}T + \sigma_1\sqrt{T}Z^1\right), x_0^2 \exp\left(-\frac{\sigma_2^2}{2}T + \sigma_2\sqrt{T}(\rho Z^1 + \sqrt{1-\rho^2}Z^2)\right)\right) - Ke^{-rT}\right)_+ . \end{aligned}$$

- Box-Muller method:

$$(Z^1, Z^2) = \left(\sqrt{-2\log(U_1)}\cos(2\pi U_2), \sqrt{-2\log(U_1)}\sin(2\pi U_2)\right).$$

$$\textit{Best-of Call}_0 = e^{-rT}\mathbb{E}\left(\big(\max(X_T^1, X_T^2) - K\big)^+\right) = \mathbb{E}\,\widetilde{\varphi}(U_1, U_2)$$

Numerical Application

$$r = 0.1,\ \sigma_i = 20\,\%,\ i = 1, 2,\ \rho = 0,\ X_0^i = 100,\ T = 1,\ K = 100$$

▷ Confidence level is set at $\alpha = 0.95$. $M = 2^m$, $m = 10, \ldots, 30$

$$\overline{\varphi}_M := \frac{1}{M}\sum_{m=1}^{M} \widetilde{\varphi}(U_{2m-1}, U_{2m}).$$

M	$\overline{\varphi}_M$	$I_{\alpha,M}$
$2^{10} = 1\,024$	18.5133	[17.4093; 19.6173]
$2^{11} = 2\,048$	18.7398	[17.9425; 19.5371]
$2^{12} = 4\,096$	18.8370	[18.2798; 19.3942]
$2^{13} = 8\,192$	19.3155	[18.9196; 19.7114]
$2^{14} = 16\,384$	19.1575	[18.8789; 19.4361]
$2^{15} = 32\,768$	19.0911	[18.8936; 19.2886]
$2^{16} = 65\,536$	19.0475	[18.9079; 19.1872]
$2^{17} = 131\,072$	19.0566	[18.9576; 19.1556]
$2^{18} = 262\,144$	19.0373	[18.9673; 19.1073]
$2^{19} = 524\,288$	19.0719	[19.0223; 19.1214]
$2^{20} = 1\,048\,576$	19.0542	[19.0191; 19.0892]

$\vdots$

$2^{30} = 1.0737.10^9$	19.0766	[19.0756; 19.0777]

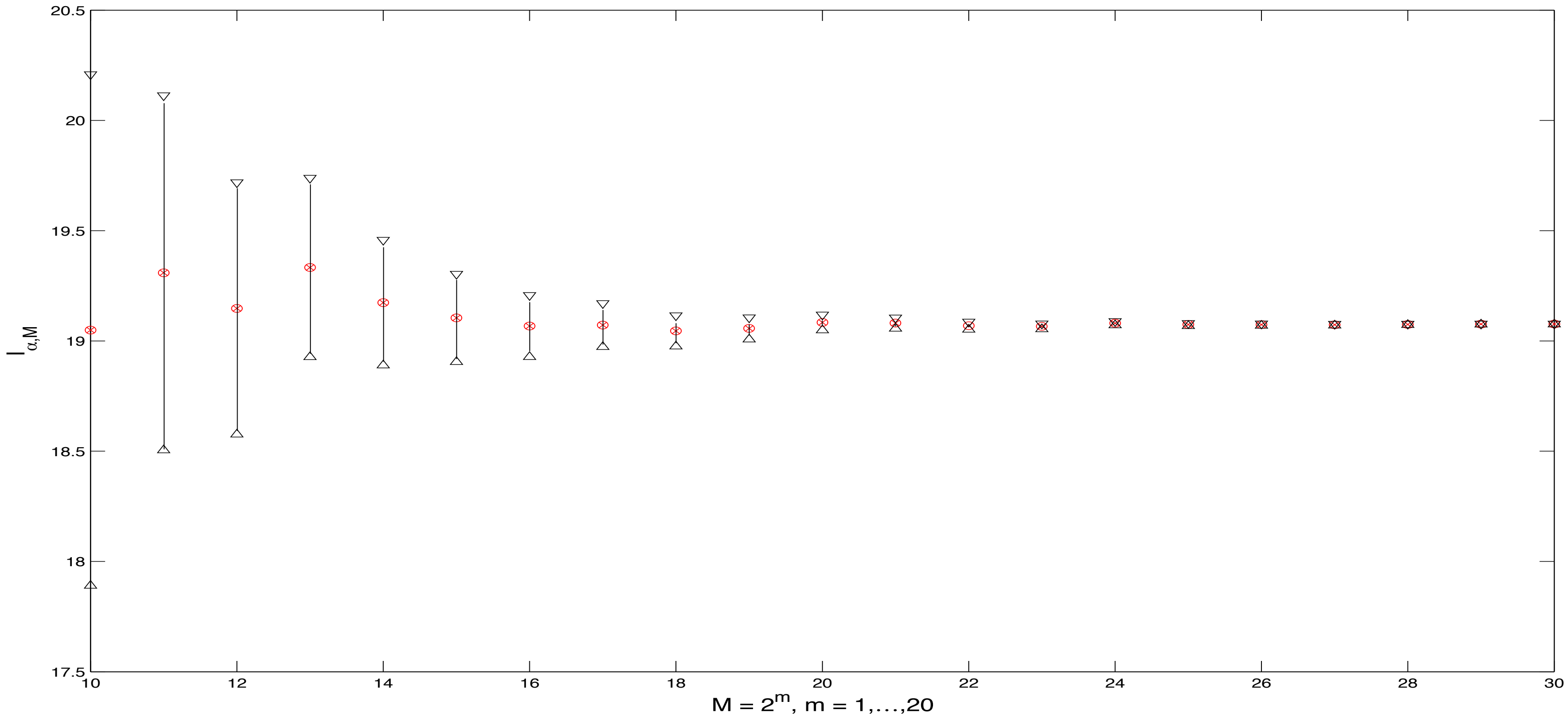

Figure: Black-Scholes Best of Call. *The Monte Carlo estimator* ($\otimes$) *and its confidence interval at level* $\alpha = 95\,\%$ *when* $M \in \{2^{10}, \ldots, 2^{30}\}$.