Ninomiya-Victoir scheme: strong convergence, antithetic version and application to multilevel estimators

Anis AL GERBI

CERMICS École des Ponts ParisTech
project team ENPC-INRIA-UPEM Mathrisk

July 7, 2016

Joint work with Benjamin Jourdain and Emmanuelle Clément

Outline

Goal

- We are interested in the computation, by Monte Carlo methods, of the expectation $Y = \mathbb{E}\left[f\left(X_T\right)\right]$, where $X = (X_t)_{0 \le t \le T}$ is the solution to a multidimensional stochastic differential equation (SDE) and $f : \mathbb{R}^n \mapsto \mathbb{R}$ a given function such that $\mathbb{E}\left[f\left(X_T\right)^2\right] < +\infty$.
- We will focus on minimizing the computational complexity subject to a given target error $\epsilon \in \mathbb{R}_+^*$.
- To measure the accuracy of an estimator $\hat{Y}$, we will consider the root mean squared error:

$$RMSE\left(\hat{Y}; Y\right) = \mathbb{E}^{\frac{1}{2}}\left[\left|Y - \hat{Y}\right|^2\right].$$

Itô-type SDE

We consider a general Itô-type SDE of the form

$$\begin{cases} dX_t = b(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t)dW_t^j \\ X_0 = x \end{cases}$$

where:

- $x \in \mathbb{R}^n$,
- $(X_t)_{0 \le t \le T}$ is a $n-$dimensional stochastic process,
- $W = \left(W^1, \ldots, W^d\right)$ is a $d-$dimensional standard Brownian motion,
- $b, \sigma^1, \ldots, \sigma^d : \mathbb{R}^n \to \mathbb{R}^n$ are Lipschitz continuous.

Stratonovich form

Assuming $\mathcal{C}^1$ regularity for diffusion coefficients $\sigma^1, \ldots, \sigma^d$, the Itô-type SDE can be written in Stratonovich form:

$$\begin{cases} dX_t = \sigma^0(X_t)dt + \sum_{j=1}^{d} \sigma^j(X_t) \circ dW_t^j \\ X_0 = x \end{cases}$$

where $\sigma^0 = b - \frac{1}{2} \sum_{j=1}^{d} \partial\sigma^j \sigma^j$ and $\partial\sigma^j$ is the Jacobian matrix of σ^j defined as follows

$$\partial\sigma^j = \left(\partial_{x_k} \sigma^{ij}\right)_{1 \le i,k \le n}.$$

The Ninomiya-Victoir scheme

Notations

- $\left(t_k = k\frac{T}{N}\right)_{0\leq k\leq N}$ is the subdivision of $[0, T]$,
- $\eta = (\eta_1, \ldots, \eta_N)$ is a sequence of independent, identically distributed Rademacher random variables independent of W,
- for all $j \in \{1, \ldots, d\}$, $\Delta W^j_{t_{k+1}} = W^j_{t_{k+1}} - W^j_{t_k}$,
- for $j \in \{0, \ldots, d\}$ and $x_0 \in \mathbb{R}^d$, let $(\exp(t\sigma^j)x_0)_{t\in\mathbb{R}}$ solve the ODE

$$\begin{cases} \frac{dx(t)}{dt} = \sigma^j\left(x\left(t\right)\right) \\ x\left(0\right) = x_0. \end{cases}$$

Scheme

If $\eta_{k+1} = 1$

$$X^{NV,N,\eta}_{t_{k+1}} = \exp\left(\frac{T}{2N}\sigma^0\right)\exp\left(\Delta W^d_{t_{k+1}}\sigma^d\right)\ldots\exp\left(\Delta W^1_{t_{k+1}}\sigma^1\right)\exp\left(\frac{T}{2N}\sigma^0\right)X^{NV,N,\eta}_{t_k},$$

and if $\eta_{k+1} = -1$

$$X^{NV,N,\eta}_{t_{k+1}} = \exp\left(\frac{T}{2N}\sigma^0\right)\exp\left(\Delta W^1_{t_{k+1}}\sigma^1\right)\ldots\exp\left(\Delta W^d_{t_{k+1}}\sigma^d\right)\exp\left(\frac{T}{2N}\sigma^0\right)X^{NV,N,\eta}_{t_k}.$$

Link between ODEs and SDEs

Link between ODEs and SDEs

for $j \in \{1, \dots, d\}$ and $y \in \mathbb{R}^n$, the dynamics of $Y_t = \exp\left(W_t^j \sigma^j\right) y$ is given by

$$dY_t = \sigma^j \left(Y_t\right) \circ dW_t^j = \frac{1}{2} \partial\sigma^j \sigma^j \left(Y_t\right) + \sigma^j \left(Y_t\right) dW_t^j.$$

Splitting of the infinitesimal generator

The infinitesimal generator

$$\mathcal{L} = b.\nabla_x + \frac{1}{2} \mathrm{Tr}\left[(\sigma^1, \dots, \sigma^d)(\sigma^1, \dots, \sigma^d)^* \nabla_x^2\right]$$

is then split into

$$\mathcal{L} = \mathcal{L}_0 + \frac{1}{2} \sum_{j=1}^{d} \mathcal{L}_j^2,$$

where

$$\mathcal{L}_0 u\left(t, x\right) = \sigma^0 \left(x\right) . \ \nabla_x u\left(t, x\right),$$

and

$$\mathcal{L}_j u\left(t, x\right) = \sigma^j \left(x\right) . \ \nabla_x u\left(t, x\right).$$

Order 2 of weak convergence

Denoting by $(X_t^x)_{t\geq 0}$ the solution to the SDE starting from $X_0^x = x \in \mathbb{R}^n$, for $f : \mathbb{R}^n \to \mathbb{R}^n$ smooth, $u(t,x) = \mathbb{E}\left[f(X_t^x)\right]$ solves the Feynman-Kac PDE

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) = Lu(t,x), \ (t,x) \in [0,\infty) \times \mathbb{R}^n \\ u(0,x) = f(x), \ x \in \mathbb{R}^n \end{cases}$$

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial t}\mathcal{L}u = \mathcal{L}\frac{\partial}{\partial t}u = \mathcal{L}^2 u$$

and $u(t_1,x) = f(x) + t_1\mathcal{L}f(x) + \frac{t_1^2}{2}\mathcal{L}^2 f(x) + \mathcal{O}(t_1^3).$

Ninomiya and Victoir have designed their scheme so that

$$\mathbb{E}[f(X_{t_1}^{NV,\eta})] = f(x) + t_1\mathcal{L}f(x) + \frac{t_1^2}{2}\mathcal{L}^2 f(x) + \mathcal{O}(t_1^3).$$

One step error $\mathcal{O}(\frac{1}{N^3}) \overset{N\text{steps}}{\longrightarrow} \mathcal{O}(\frac{1}{N^2})$ global error.

Order 1/2 of strong convergence

Theorem (Strong convergence)

Assume that the vector fields, σ^0, σ^j and $\partial\sigma^j\sigma^j$, for all $j \in \{1, \ldots, d\}$, are Lipschitz continuous. Then, for all $p \geq 1$, there exists a constant $C_{NV} \in \mathbb{R}_+^*$ such that for all $N \in \mathbb{N}^*$,

$$\mathbb{E}\left[\max_{0\leq k\leq N}\left\|X_{t_k} - X_{t_k}^{NV,\eta}\right\|^{2p}\middle|\eta\right] \leq \frac{C_{NV}}{N^p}.$$

Stable convergence of the normalized error

Theorem (Stable convergence)

Assume that

- $\sigma^0 \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ *and is a Lipschitz continuous function with polynomially growing second order derivatives,*
- *for all* $j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ *and is Lipschitz continuous and its first order derivative is locally Lipschitz with polynomially growing Lipschitz constant,*
- *for all* $j, m \in \{1, \ldots, d\}, \partial\sigma^j\sigma^m$ *is Lipschitz continuous,*
- *for all* $j \in \{1, \ldots, d\}, \partial\sigma^j\sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ *with polynomially growing second order derivatives.*

Then

$$V^N := \sqrt{N}\left(X - X^{NV,\eta}\right) \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} V,$$

where V *is the unique solution of the following affine equation*

$$V_t = \sqrt{\frac{T}{2}} \sum_{j=1}^{d} \sum_{m=1}^{j-1} \int_0^t \left[\sigma^j, \sigma^m\right](X_s)\, dB_s^{j,m} + \int_0^t \partial b(X_s)\, V_s ds + \sum_{j=1}^{d} \int_0^t \partial\sigma^j(X_s)\, V_s dW_s^j,$$

with $\left[\sigma^j, \sigma^m\right] = \partial\sigma^m\sigma^j - \partial\sigma^j\sigma^m$, *and* $(B_t)_{0\le t\le T}$ *is a standard* $d(d-1)/2$*-dimensional Brownian motion independent of* W.

Stable convergence

Remark

- The limit does not depend on η,
- the strong convergence rate is actually $1/2$ when at least two Brownian vector fields do not commute,
- if the Brownian vector fields commute, i.e.

$$\text{for all } j, m \in \{1, \dots, d\}, \partial\sigma_j\sigma_m = \partial\sigma_m\sigma_j,$$

then the limit is 0.

Commutation of the Brownian vector fields

Theorem (Strong convergence)

Assume that

$$\text{for all } j, m \in \{1, \ldots, d\}, \partial\sigma_j\sigma_m = \partial\sigma_m\sigma_j,$$

(the order of integration of these Brownian vector fields no longer matters and η is useless) and that

- *for all* $j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ *with bounded first order derivatives,*
- $\sigma^0 \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ *with bounded first order derivatives and polynomially growing second order derivatives,*
- $\sum_{j=1}^{d} \partial\sigma^j\sigma^j$ *is a Lipschitz continuous function,*

Then, there exists a constant $C'_{NV} \in \mathbb{R}_+^*$ *such that for all* $N \in \mathbb{N}^*$*,*

$$\mathbb{E}\left[\sup_{t\leq T}\left\|X_t - X_t^{NV}\right\|^{2p}\right] \leq \frac{C'_{NV}}{N^{2p}}.$$

Under the commutativity of the Brownian vector fields, it is possible to implement the Milstein scheme which also exhibits order one of strong convergence.

Stable convergence of the normalized error

Theorem (Stable convergence)

Assume that

- *for all* $j \in \{0, \ldots, d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ *with bounded first order derivatives,* $\partial\sigma^j$ *and* $\partial^2\sigma^j$ *are locally Lipschitz with polynomially growing Lipschitz constant,*
- *for all* $j \in \{1, \ldots, d\}, \partial\sigma^j\sigma^j$ *is a Lipschitz continuous function,*

and that the commutativity condition holds. Then:

$$U^N = N\left(X - X^{NV}\right) \underset{N\to+\infty}{\overset{stably}{\Longrightarrow}} U,$$

where U *is the unique solution of the following affine equation:*

$$U_t = \frac{T}{2\sqrt{3}} \sum_{j=1}^{d} \int_0^t \left[\sigma^0, \sigma^j\right](X_s)\, d\tilde{B}^j_s + \int_0^t \partial b(X_s)\, U_s ds + \sum_{j=1}^{d} \int_0^t \partial\sigma^j(X_s)\, U_s dW^j_s,$$

and $\left(\tilde{B}_t\right)_{0\le t\le T}$ *is a standard* d*-dimensional Brownian motion independent of* W*.*

Outline

The Multilevel Monte Carlo

The main idea of this technique is to use the following telescopic summation to control the bias:

$$\mathbb{E}\left[f\left(X_T^{2^L}\right)\right] = \mathbb{E}\left[f\left(X_T^1\right)\right] + \sum_{l=1}^{L} \mathbb{E}\left[f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right)\right].$$

Then, a generalized multilevel Monte Carlo estimator is built as follows:

$$\hat{Y}_{MLMC} = \sum_{l=0}^{L} \frac{1}{M_l} \sum_{k=1}^{M_l} Z_k^l$$

where $\left(Z_k^l\right)_{0\leq l\leq L, 1\leq k\leq M_l}$ are independent random variables such that:

$$\mathbb{E}\left[Z^0\right] = \mathbb{E}\left[f\left(X_T^1\right)\right]$$

and for all $l \in \{1, \ldots, L\}$,

$$\mathbb{E}\left[Z^l\right] = \mathbb{E}\left[f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right)\right].$$

Cost and canonical exemple

Cost

For a given discretization level $l \in \{0, \ldots, L\}$, the computational cost of simulating one sample Z^l is $C\lambda_l 2^l$, where:

- $C \in \mathbb{R}_+$ is a constant, depending only on the discretization scheme,
- for all $l \in \{0, \ldots, L\}, \lambda_l \in \mathbb{Q}_+^*$ is a weight, depending only on l,

$$\mathcal{C}_{MLMC} = C \sum_{l=0}^{L} M_l \lambda_l 2^l.$$

Natural choice for $Z^l, l \in \{0, \ldots, L\}$

$$Z^0 = f\left(X_T^1\right)$$

$$\text{for all } l \in \{1, \ldots, L\}, Z^l = f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right).$$

For this canonical choice, it is natural to take $\lambda_0 = 1$ and $\lambda_l = \frac{3}{2}$.

Complexity analysis

Bias

$$B\left(\hat{Y}_{MLMC};Y\right)=\mathbb{E}\left[\hat{Y}_{MLMC}\right]-Y=\mathbb{E}\left[f\left(X_T^{2^L}\right)\right]-\mathbb{E}\left[f\left(X_T\right)\right].$$

The bias is related to the weak error of the scheme:

$$\mathbb{E}\left[f\left(X_T^{2^L}\right)-f\left(X_T\right)\right]=\frac{c_1}{2^{\alpha L}}+o\left(\frac{1}{2^{\alpha L}}\right).$$

Variance

$$\mathbb{V}\left[\hat{Y}_{MLMC}\right]=\sum_{l=0}^{L}\frac{1}{M_l}\mathbb{V}\left[Z^l\right].$$

If the simulation of X^{2^l} and $X^{2^{l-1}}$ comes from the same Brownian path, then $\mathbb{V}\left[Z^l\right]$ converges to 0 as l goes to infinity. The rate β of convergence to zero of $\mathbb{V}\left[Z^l\right]$ is related to the strong convergence order γ of the scheme ($\beta \geq 2\gamma$).

Optimal complexity

Theorem (Complexity theorem (Giles))

Assume that there exist $(\alpha, c_1) \in \mathbb{R}_+^* \times \mathbb{R}^*$ *and* $(\beta, c_2) \in \left(\mathbb{R}_+^*\right)^2$ *such that for all* $l \in \mathbb{N}$*:*

$$\mathbb{E}\left[f\left(X_T^{2^l}\right)\right] - Y = \frac{c_1}{2^{\alpha l}} + o\left(\frac{1}{2^{\alpha l}}\right)$$

and

$$\mathbb{V}\left[Z^l\right] = \frac{c_2}{2^{\beta l}} + o\left(\frac{1}{2^{\beta l}}\right).$$

Then, the optimal complexity is given by:

$$\begin{cases} \mathcal{C}_{MLMC}^* = O\left(\epsilon^{-2}\right) & \text{if } \beta > 1, \\ \mathcal{C}_{MLMC}^* = O\left(\epsilon^{-2}\left(\log\left(\frac{1}{\epsilon}\right)\right)^2\right) & \text{if } \beta = 1, \\ \mathcal{C}_{MLMC}^* = O\left(\epsilon^{-2+\frac{\beta-1}{\alpha}}\right) & \text{if } \beta < 1. \end{cases}$$

Optimal parameters

Optimal parameters

$$L^* = \left\lceil \frac{\log_2\left(\frac{\sqrt{2}|c_1|}{\epsilon}\right)}{\alpha} \right\rceil$$

$$M_0^* = \left\lceil \frac{2}{\epsilon^2}\sqrt{\frac{\mathbb{V}\left[Z^0\right]}{\lambda_0}}\left(\sqrt{\lambda_0 \mathbb{V}\left[Z^0\right]} + \sum_{l=1}^{L^*}\sqrt{c_2\lambda_l 2^{l(1-\beta)}}\right)\right\rceil$$

and for all $l \in \{1, \ldots, L^*\}$

$$M_l^* = \left\lceil \frac{2}{\epsilon^2}\sqrt{\frac{c_2}{\lambda_l 2^{l(\beta+1)}}}\left(\sqrt{\lambda_0 \mathbb{V}\left[Z^0\right]} + \sum_{l=1}^{L^*}\sqrt{c_2\lambda_l 2^{l(1-\beta)}}\right)\right\rceil.$$

Regression

One can estimate $(\alpha, \beta, c_1, c_2)$ by using a regression:

$$\mathbb{V}\left[Z^l\right] \sim \frac{c_2}{2^{\beta l}}$$

$$\mathbb{E}\left[Z^l\right] \sim \frac{c_1\left(1 - 2^{\alpha}\right)}{2^{\alpha l}}.$$

Theoretical computing time

Denoting by τ^l the theoretical computing time of level $l \in \{0, \ldots, L^*\}$, one ha

$$\tau^l \propto M_l^* 2^l.$$

Replacing M_l^*, one can write

$$\tau^l \propto 2^{-l\left(\frac{\beta+1}{2}\right)} 2^l = 2^{-l\left(\frac{\beta-1}{2}\right)}.$$

When $\beta = 1$, for the Euler scheme for example, τ^l is constant.

Acceleration of the multilevel Monte Carlo

Debrabant and Rössler consider

$$\hat{Y}^{DR}_{MLMC} = \frac{1}{M_0}\sum_{k=1}^{M_0} f\left(X_T^{1,0,k}\right) + \sum_{l=1}^{L-1}\frac{1}{M_l}\sum_{k=1}^{M_l}\left(f\left(X_T^{2^l,l,k}\right) - f\left(X_T^{2^{l-1},l,k}\right)\right)$$
$$+ \frac{1}{M_L}\sum_{k=1}^{M_L}\left(f\left(\hat{X}_T^{2^L,L,k}\right) - f\left(X_T^{2^{L-1},L,k}\right)\right),$$

where $\hat{X}$ is a scheme with high order of weak convergence and such that

$$\mathbb{V}\left[f\left(\hat{X}_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right)\right] = \frac{c_2'}{2^{\beta l}} + o\left(\frac{1}{2^{\beta l}}\right).$$

Outline

The Multilevel Richardson-Romberg Extrapolation

Adapting the notation of Pagès and Lemaire, the multilevel Richardson-Romberg extrapolation estimator is built as follows:

$$\hat{Y}_{ML2R} = \sum_{l=0}^{L} \frac{W_l}{M_l} \sum_{k=0}^{M_l} Z_k^l,$$

where $\left(Z_{kl}^l\right)_{0\le l\le L, 1\le k\le M_l}$ are independent random variables satisfying

$$\mathbb{E}\left[Z^0\right] = \mathbb{E}\left[f\left(X_T^1\right)\right],$$

and for all $l \in \{1, \ldots, L\}$,

$$\mathbb{E}\left[Z^l\right] = \mathbb{E}\left[f\left(X_T^{2^l}\right) - f\left(X_T^{2^{l-1}}\right)\right].$$

Actually, the multilevel Richardson-Romberg extrapolation can be seen as a weighted version of the Multilevel Monte Carlo estimator.

Optimal complexity

Theorem (Complexity theorem (Pagès, Lemaire))

Assume that there exist $R \in \mathbb{N}^*, \alpha \in \mathbb{R}_+^*, c'_1, \dots, c'_R \in \mathbb{R}^*$ *and* $(\beta, c_2) \in \left(\mathbb{R}_+^*\right)^2$ *such that for all* $l \in \mathbb{N}$*:*

$$\mathbb{E}\left[f\left(X_T^{2^l}\right)\right] - Y = \sum_{j=1}^{R} \frac{c'_j}{2^{\alpha l j}} + o\left(\frac{1}{2^{\alpha l R}}\right),$$

and

$$\mathbb{V}\left[Z^l\right] = \frac{c_2}{2^{\beta l}} + o\left(\frac{1}{2^{\beta l}}\right).$$

Then, the optimal complexity is given by:

- $\mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\right)$ *if* $\beta > 1$,
- $\mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\log\left(\frac{1}{\epsilon}\right)\right)$ *if* $\beta = 1$,
- $\mathcal{C}^*_{ML2R} = O\left(\epsilon^{-2}\exp\left(-\frac{\beta-1}{\sqrt{\alpha}}\sqrt{2\log(2)\log\left(\frac{1}{\epsilon}\right)}\right)\right)$ *if* $\beta < 1$.

Optimal parameters

$$L^* = \left\lfloor \sqrt{\left(\frac{1}{2} + \log_2(T)\right)^2 + \frac{2}{\alpha}\log_2\left(\frac{\sqrt{1+4\alpha}}{\epsilon}\right)} + \log_2(T) - \frac{1}{2} \right\rfloor$$

$$W_l = \sum_{j=l}^{L^*} w_j$$

$$M_l^* = \lceil q_l^* N^* \rceil$$

where:

$$w_j = (-1)^{L^*-j} \frac{2^{-\frac{\alpha}{2}(L^*-j)(L^*-j+1)}}{\prod_{k=1}^{j}\left(1-2^{-k\alpha}\right) \prod_{k=1}^{L^*-j}\left(1-2^{-k\alpha}\right)}$$

$$\begin{cases} q_0^* \propto (1+\theta) \\ q_l^* \propto \theta \left|W_l\right| \frac{2^{-\frac{\beta}{2}l} + 2^{-\frac{\beta}{2}(l-1)}}{\sqrt{2^l + 2^{l-1}}} \quad \forall l \in \{1, \ldots, L^*\} \\ \sum_{l=0}^{L^*} q_l^* = 1 \end{cases}$$

$$N^* = \left(1 + \frac{1}{2\alpha(L^*+1)}\right) \frac{\mathbb{V}\left[f(X_T)\right]\left(1 + \theta\left(1 + \sum_{l=1}^{L^*}\left|W_l\right|\left(2^{-\frac{\beta}{2}l} + 2^{-\frac{\beta}{2}(l-1)}\right)\sqrt{2^l + 2^{l-1}}\right)\right)^2}{\epsilon^2\left(q_0^* + \sum_{l=1}^{L^*} q_l^*\left(2^l + 2^{l-1}\right)\right)}$$

and

$$\theta = T^{-\frac{\beta}{2}} \sqrt{\frac{c_2}{\mathbb{V}\left[f(X_T)\right]}}.$$

Outline

The Giles-Szpruch scheme

The Giles-Szpruch scheme is a modified Milstein scheme. The terms involving the Lévy areas $\int_{t_k}^{t_{k+1}} \Delta W_s^j dW_s^m - \int_{t_k}^{t_{k+1}} \Delta W_s^m dW_s^j$ have been removed:

$$
\begin{cases}
X_{t_{k+1}}^{GS} = X_{t_k}^{GS} + b\left(X_{t_k}^{GS}\right)(t_{k+1} - t_k) + \sum_{j=1}^{d} \sigma^j \left(X_{t_k}^{GS}\right) \Delta W_{t_{k+1}}^j \\
\qquad + \frac{1}{2} \sum_{j,m=1}^{d} \partial \sigma^j \sigma^m \left(X_{t_k}^{GS}\right) \left(\Delta W_{t_{k+1}}^j \Delta W_{t_{k+1}}^m - \mathbb{1}_{\{j=m\}} (t_{k+1} - t_k)\right) \\
X_{t_0}^{GS} = x.
\end{cases}
$$

The Giles-Szpruch scheme: antithetic version

We consider two grids: a coarse grid with time step $h_{l-1} = \frac{T}{2^{l-1}}$, a fine grid with time step $h_l = \frac{T}{2^l}$ and we introduce some notations:

- $\forall k \in \left\{0, \dots, 2^{l-1}\right\}, t_k = kh_{l-1}$,
- $\forall k \in \left\{0, \dots, 2^{l-1} - 1\right\}, t_{k+\frac{1}{2}} = \left(k + \frac{1}{2}\right) h_{l-1}$,
- $\Delta W^c_{t_{k+1}} = W_{t_{k+1}} - W_{t_k}$, $\Delta W^f_{t_{k+\frac{1}{2}}} = W_{t_{k+\frac{1}{2}}} - W_{t_k}$ and $\Delta W^f_{t_{k+1}} = W_{t_{k+1}} - W_{t_{k+\frac{1}{2}}}$.

On the coarsest grid, $X^{GS,2^{l-1}}$ is defined inductively by:

$$X^{GS,2^{l-1}}_{t_{k+1}} = X^{GS,2^{l-1}}_{t_k} + b\left(X^{GS,2^{l-1}}_{t_k}\right) h_{l-1} + \sum_{j=1}^{d} \sigma^j \left(X^{GS,2^{l-1}}_{t_k}\right) \Delta W^{j,c}_{t_{k+1}}$$

$$+ \frac{1}{2} \sum_{j,m=1}^{d} \partial\sigma^j \sigma^m \left(X^{GS,2^{l-1}}_{t_k}\right) \left(\Delta W^{j,c}_{t_{k+1}} \Delta W^{m,c}_{t_{k+1}} - \mathbb{1}_{\{m=j\}} h_{l-1}\right).$$

The Giles-Szpruch scheme: antithetic version

Similarly, on the finest grid:

$$\begin{cases} X^{GS,2^l}_{t_{k+\frac{1}{2}}} & = X^{GS,2^l}_{t_k} + b\left(X^{GS,2^l}_{t_k}\right) h_l + \sum_{j=1}^{d} \sigma^j \left(X^{GS,2^l}_{t_k}\right) \Delta W^{j,f}_{t_{k+\frac{1}{2}}} \\ & + \frac{1}{2} \sum_{j,m=1}^{d} \partial\sigma^j \sigma^m \left(X^{GS,2^l}_{t_k}\right) \left(\Delta W^{j,f}_{t_{k+\frac{1}{2}}} \Delta W^{m,f}_{t_{k+\frac{1}{2}}} - \mathbb{1}_{\{m=j\}} h_l\right) \\ X^{GS,2^l}_{t_{k+1}} & = X^{GS,2^l}_{t_{k+\frac{1}{2}}} + b\left(X^{GS,2^l}_{t_{k+\frac{1}{2}}}\right) h_l + \sum_{j=1}^{d} \sigma^j \left(X^{GS,2^l}_{t_{k+\frac{1}{2}}}\right) \Delta W^{j,f}_{t_{k+1}} \\ & + \frac{1}{2} \sum_{j,m=1}^{d} \partial\sigma^j \sigma^m \left(X^{GS,2^l}_{t_{k+\frac{1}{2}}}\right) \left(\Delta W^{j,f}_{t_{k+1}} \Delta W^{m,f}_{t_{k+1}} - \mathbb{1}_{\{m=j\}} h_l\right). \end{cases}$$

The antithetic scheme $\tilde{X}^{GS,2^l}$ is defined by the same discretization, except that the Brownian increment $\Delta W^f_{t_{k+\frac{1}{2}}}$ and $\Delta W^f_{t_{k+1}}$ are swapped.

Strong coupling with order one between successive levels

Considering, for all $l \in \{1, \ldots, L\}$,

$$Z_{GS}^{l} = \frac{1}{2}\left(f\left(\tilde{X}_{T}^{GS,2^{l}}\right) + f\left(X_{T}^{GS,2^{l}}\right)\right) - f\left(X_{T}^{GS,2^{l-1}}\right),$$

Giles and Szpruch obtain a first order of convergence.

Theorem (Giles-Szpruch)

Assume that $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ and $b, \forall j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ with bounded first and second order derivatives. Then:

$$\forall p \geq 1, \exists c \in \mathbb{R}_+^*, \forall l \in \mathbb{N}^*,\ \mathbb{E}\left[\left|Z_{GS}^{l}\right|^{2p}\right] \leq \frac{c}{2^{2pl}}.$$

Outline

Coupling between the Ninomiya-Victoir scheme and the Giles-Szpruch scheme

Theorem (Strong convergence)

Assume that $b \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}^n)$ *with bounded first and second order derivatives, and,* $\forall j \in \{1, \dots, d\}, \sigma^j \in \mathcal{C}^3(\mathbb{R}^n; \mathbb{R}^n)$ *with bounded first and second order derivatives and with polynomially growing third order derivatives, and that,* $\forall j, m \in \{1, \dots, d\}$, $\partial\sigma^j\sigma^m$ *has bounded first order derivatives. Then:*

$$\exists C_{GS} \in \mathbb{R}_+^*, \forall N \in \mathbb{N}^*, \; \mathbb{E}\left[\max_{0 \leq k \leq N} \left\| \bar{X}_{t_k}^{NV,\eta} - X_{t_k}^{GS} \right\|^{2p} \middle| \eta\right] \leq \frac{C_{GS}}{N^{2p}},$$

where

$$\bar{X}^{NV,\eta} = \frac{1}{2}\left(X^{NV,\eta} + X^{NV,-\eta}\right).$$

Strong coupling with order one between successive levels

Considering:

$$Z_{GS-NV}^{l} = \frac{1}{4} f\left(\tilde{X}_T^{NV,2^l,\eta}\right) + \frac{1}{4} f\left(\tilde{X}_T^{NV,2^l,-\eta}\right) + \frac{1}{4} f\left(X_T^{NV,2^l,\eta}\right) + \frac{1}{4} f\left(X_T^{NV,2^l,-\eta}\right) - f\left(X_T^{GS,2^{l-1}}\right),$$

we have a first order of convergence.

Corollary

Assume that $f \in \mathcal{C}^2\left(\mathbb{R}^n, \mathbb{R}\right)$ *and* $b \in \mathcal{C}^2\left(\mathbb{R}^n, \mathbb{R}^n\right)$ *with bounded first and second order derivatives, and,* $\forall j \in \{1, \ldots, d\}, \sigma^j \in \mathcal{C}^3\left(\mathbb{R}^n, \mathbb{R}^n\right)$ *with bounded first and second order derivatives and with polynomially growing third order derivatives. Then:*

$$\forall p \geq 1, \exists c \in \mathbb{R}_+^*, \forall l \in \mathbb{N}^*, \; \mathbb{E}\left[\left|Z_{GS-NV}^{l}\right|^{2p}\right] \leq \frac{c}{2^{2pl}}.$$

Strong coupling with order one between successive levels

Considering:

$$Z_{NV}^{l} = \frac{1}{4}\left(f\left(\tilde{X}_T^{NV,2^l,\eta}\right) + f\left(\tilde{X}_T^{NV,2^l,-\eta}\right) + f\left(X_T^{NV,2^l,\eta}\right) + f\left(X_T^{NV,2^l,-\eta}\right)\right) - \frac{1}{2}\left(f\left(X_T^{NV,2^{l-1},\eta}\right) + f\left(X_T^{NV,2^{l-1},-\eta}\right)\right), \forall l \in \{1,\ldots,L\},$$

we have a first order of convergence.

Corollary

Assume that $f \in \mathcal{C}^2(\mathbb{R}^n,\mathbb{R})$ and $b \in \mathcal{C}^2(\mathbb{R}^n,\mathbb{R}^n)$ with bounded first and second order derivatives, and, $\forall j \in \{1,\ldots,d\}, \sigma^j \in \mathcal{C}^3(\mathbb{R}^n,\mathbb{R}^n)$ with bounded first and second order derivatives and with polynomially growing third order derivatives. Then:

$$\forall p \geq 1, \exists c \in \mathbb{R}_+^*, \forall l \in \mathbb{N}^*,\ \mathbb{E}\left[\left|Z_{NV}^{l}\right|^{2p}\right] \leq \frac{c}{2^{2pl}}.$$

Derived MLMC estimators

- $\hat{Y}^{GS}_{MLMC}$ is the MLMC estimator with the Giles-Szpruch scheme:

$$\hat{Y}^{GS}_{MLMC} = \sum_{l=0}^{L^*} \frac{1}{M_l^*} \sum_{k=1}^{M_l^*} Z_{GS}^{l,k}$$

  where $Z^0_{GS} = f\left(X_T^{GS,1}\right)$.
- $\hat{Y}^{NV}_{MLMC}$ is the MLMC estimator with the Ninomiya-Victoir scheme:

$$\hat{Y}^{NV}_{MLMC} = \sum_{l=0}^{L^*} \frac{1}{M_l^*} \sum_{k=1}^{M_l^*} Z_{NV}^{l,k}$$

  where $Z^0_{NV} = f\left(X_T^{NV,1,\eta}\right)$ or $Z^0_{NV} = \frac{1}{2}\left(f\left(X_T^{NV,1,\eta}\right) + f\left(X_T^{NV,1,-\eta}\right)\right)$.
- $\hat{Y}^{GS-NV}_{MLMC}$ is the MLMC estimator with the Giles-Szpruch scheme from level 0 to level $L^* - 1$, and the coupling between the Ninomiya-Victoir and the Giles-Szpruch scheme at the last level L^*:

$$\hat{Y}^{GS-NV}_{MLMC} = \sum_{l=0}^{L^*-1} \frac{1}{M_l^*} \sum_{k=1}^{M_l^*} Z_{GS}^{l,k} + \frac{1}{M_{L^*}^*} \sum_{k=1}^{M_{L^*}^*} Z_{GS-NV}^{L^*,k}.$$

Derived ML2R estimators

- $\hat{Y}^{GS}_{ML2R}$ is the ML2R estimator with the Giles-Szpruch scheme:

$$\hat{Y}^{GS}_{ML2R} = \sum_{l=0}^{L^*} \frac{W_l}{M_l^*} \sum_{k=1}^{M_l^*} Z_{GS}^{l,k}.$$

- $\hat{Y}^{NV}_{ML2R}$ is the ML2R estimator with the Ninomiya-Victoir scheme:

$$\hat{Y}^{NV}_{ML2R} = \sum_{l=0}^{L^*} \frac{W_l}{M_l^*} \sum_{k=1}^{M_l^*} Z_{NV}^{l,k}.$$

Outline

The Clark-Cameron SDE

ClarkCameron SDE

$$\begin{cases} dX_t^1 = \mu dt + dW_t^1 \\ dX^2 = X_t^1 dW_t^2. \end{cases}$$

Parameters and Payoff

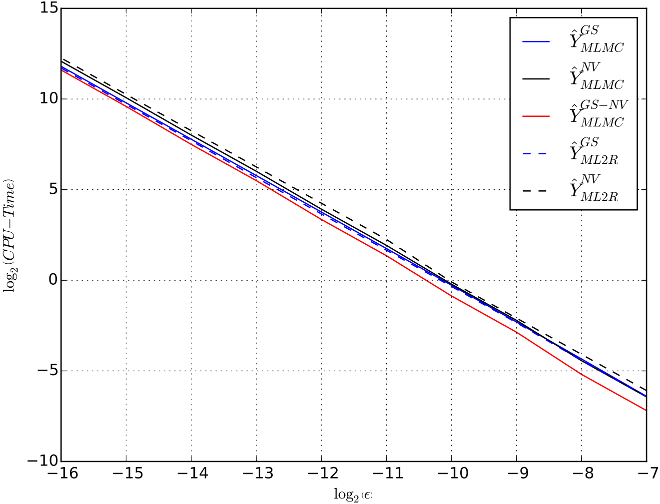
- $X_0^1 = X_0^2 = 0$ and $T = 1$.
- $\mu = 1$.
- $f(x_1, x_2) = \cos(x_2)$ and $f(x_1, x_2) = (x_2)_+$.

CPU-time ratios

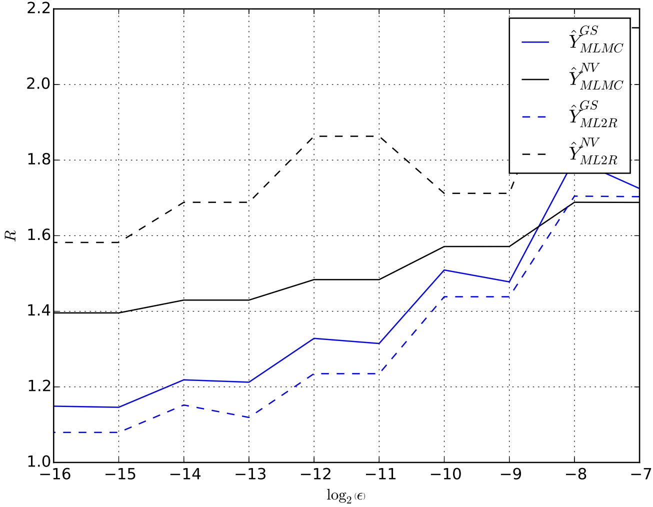
To measure the efficiency of $\hat{Y}_{MLMC}^{GS-NV}$ with respect to other estimators, we plot the following CPU-time ratios:

$$R = \frac{CPU - time\left(\hat{Y}\right)}{CPU - time\left(\hat{Y}_{MLMC}^{GS-NV}\right)}.$$

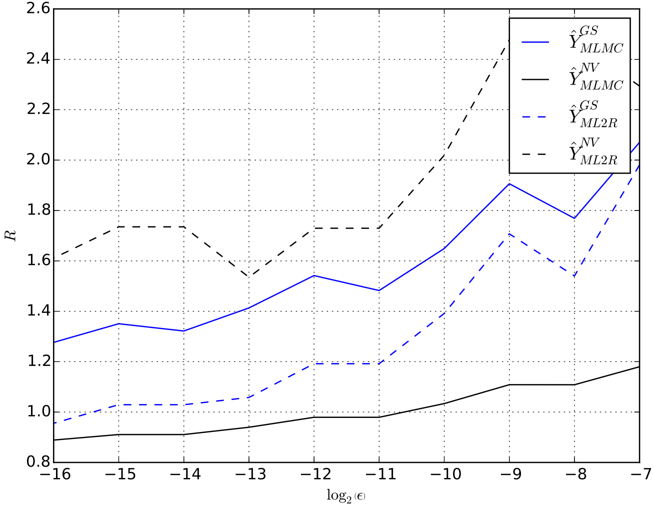
Numerical results: $f(x_1, x_2) = \cos(x_2)$



Numerical results $f(x_1, x_2) = \cos(x_2)$



Numerical results $f(x_1, x_2) = (x_2)_+$



Outline

The Heston model

The Heston model

$$\begin{cases} dU_t = (r - \delta - \frac{1}{2} V_t) dt + \sqrt{V_t} dW_t^1 \\ dV_t = \kappa(\theta - V_t) dt + \sigma \sqrt{V_t} \left(\rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2 \right), \end{cases}$$

where the asset price S is given by $S_t = \exp(U_t)$ and

- $\theta \in \mathbb{R}_+^*$ is the long implied variance, or long run average price variance; as t tends to infinity, the expected value of V_t tends to θ,
- $\kappa \in \mathbb{R}_+^*$ is the rate at which V_t reverts to θ,
- $\sigma \in \mathbb{R}_+^*$ is the volatility of the implied volatility and determines the variance of V_t,
- $r \in \mathbb{R}$ the annualized risk-free interest rate, continuously compounded,
- $\delta \in \mathbb{R}_+^*$ is the annualized continuous yield dividend,
- $\rho \in]-1, 1[$ is the correlation between the two Brownian motion (ie stock price and implied volatility).

The Heston model

We assume that $2\kappa\theta \geq \sigma^2$ to ensure that the zero boundary is not attainable for the volatility process.

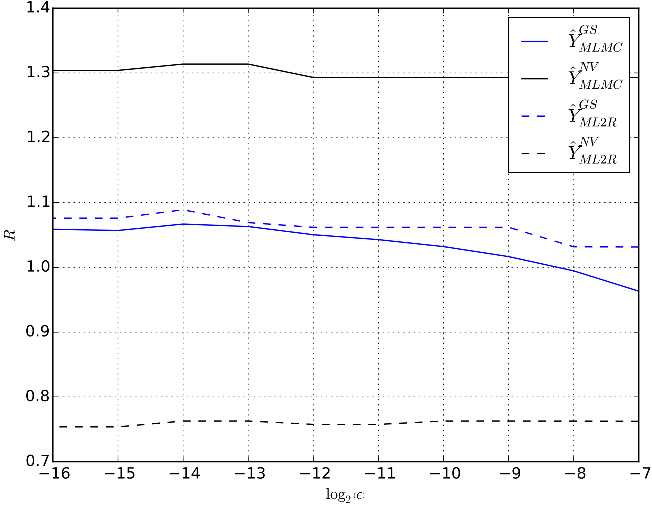
Parameters and Payoff

- $X_0 = 0,\ V_0 = 1$ and $T = 1$.
- $r = 0.05, \kappa = 0.5, \theta = 0.9,\ \sigma = 0.05$ and $\delta = \rho = 0$.
- $f(x, v) = \exp(-rT)(\exp(x) - 1)_+$.

Remark

The Ninomiya-Victoir scheme is well defined when $4\kappa\theta \geq \sigma^2$.

Numerical results



Thank you for your attention!