An Antithetic Approach of Multilevel Richardson-Romberg Extrapolation Estimator for Multidimensional SDES

Cheikh Mbaye[1(✉)], Gilles Pagès[2], and Frédéric Vrins[1]

[1] Université Catholique de Louvain, Louvain-la-Neuve, Belgium
{cheikh.mbaye,frederic.vrins}@uclouvain.be
[2] Université Pierre et Marie Curie, Paris, France
gilles.pages@upmc.fr

Abstract. The Multilevel Richardson-Romberg (*ML2R*) estimator was introduced by Pagès & Lemaire in [1] in order to remove the bias of the standard Multilevel Monte Carlo (*MLMC*) estimator in the 1D Euler scheme. Milstein scheme is however preferable to Euler scheme as it allows to reach the optimal complexity $O(\varepsilon^{-2})$ for each of these estimators. Unfortunately, Milstein scheme requires the simulation of Lévy areas when the SDE is driven by a multidimensional Brownian motion, and no efficient method is currently available to this purpose so far (except in dimension 2). Giles and Szpruch [2] recently introduced an antithetic multilevel correction estimator avoiding the simulation of these areas without affecting the second order complexity. In this work, we revisit the *ML2R* and *MLMC* estimators in the framework of the antithetic approach, thereby allowing us to remove the bias whilst preserving the optimal complexity when using Milstein scheme.

Keywords: Multilevel Monte Carlo · Antithetic Multilevel Monte Carlo · Richardson-Romberg extrapolation · Milstein scheme · Option pricing

1 Introduction

Monte Carlo methods provide a flexible and elegant tool for pricing derivatives securities that do not require the calculation of optimal exercise decision, like European options, exotic options (Asian, barrier, lookback, etc.). Unfortunately, these methods suffer from chronic deficit in performance. Various techniques have been implemented to accelerate their convergence: antithetic prints, control variate, importance sampling, just to name a few. Except the first, all of these techniques rely on the specificities of the product we want to price and hence fail by lack of genericness. Multilevel Monte Carlo (*MLMC*) aims at providing a universal solution applicable to this convergence problem. In the case of Euler scheme, Lemaire and Pagès introduced a more efficient method called Multilevel Richardson-Romberg (*ML2R*). This leads to a reduction of the computational

I. Dimov et al. (Eds.): NAA 2016, LNCS 10187, pp. 482–491, 2017.
DOI: 10.1007/978-3-319-57099-0_54

cost from $O(\varepsilon^{-2}(\log(1/\varepsilon))^2)$ to $O(\varepsilon^{-2}(\log(1/\varepsilon)))$ (see [1]), for a given accuracy $\varepsilon > 0$. The Milstein scheme allows to reach the optimal complexity $O(\varepsilon^{-2})$ for each of these estimators. In the multidimensional case, an antithetic approach avoiding the simulation of Lévy area allows to achieve this optimal complexity for MLMC estimator [2].

In this paper, we introduce a suitable antithetic correction for *ML2R* estimator without affecting the optimal complexity. We show that this complexity is achieved by both *ML2R* and *MLMC* but our new *antithetic ML2R* is better in term of performance than the standard *antithetic MLMC* (see Sect. 4).

This paper is organized as follows: in Sect. 2, we introduce a general framework to formalize the optimization of a biased Monte Carlo simulation as described in [1]. In Sect. 3, we introduce our new antithetic *ML2R* estimtor (see Eq. (8)) with its optimal allocation given by Theorem 3 and the standard antithetic *MLMC* after a brief background on the Multilevel estimators (*MLMC* & *ML2R*). In the last section, we present the numerical experiments.

2 General Framework

In this section, we will revisit in detail the preliminaries we need before introducing our Multilevel estimators.

2.1 Milstein Scheme

Let $T > 0$ and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ satisfying the usual conditions. We consider the multidimensional $\mathbb{R}^d$-valued SDEs of the form

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 \in \mathbb{R}^d,\ t\in[0,T] \tag{1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{M}(d,q)$ are smooth functions and W is a q-dimensional Brownian motion. Define the tensor $g_{ijk}(x)$ as

$$g_{ijk}(x) = \frac{1}{2}\sum_{m=1}^{d} \frac{\partial \sigma_{ij}(x)}{\partial x_m}\sigma_{mk}(x), \quad i = 1,\ldots,d,\ j = 1,\ldots,q.$$

Then X can be approximated by the Milstein scheme with uniform step $h = T/n$

$$\Delta\tilde{X}_{i,\ell} = b_i(\tilde{X}_\ell)h + \sum_{j=1}^{q}\sigma_{ij}(\tilde{X}_\ell)\Delta W_{j,\ell} + \sum_{j,k=1}^{q} g_{ijk}(\tilde{X}_\ell)(\Delta W_{j,\ell}\Delta W_{k,\ell} - hC_{i,k} - A_{jk,\ell})$$

where C is the correlation matrix for the driving Brownian paths $\Delta W_{j,\ell} = W^j_{t_{\ell+1}} - W^j_{t_\ell}$, $\Delta\tilde{X}_{i,\ell} = \tilde{X}_{i,\ell+1} - \tilde{X}_{i,\ell}$ with $t_\ell = \ell h$, and $A_{jk,\ell}$ is the Lévy area of (W^j, W^k) between t_ℓ and $t_{\ell+1}$ defined as (see [2])

$$A_{jk,\ell} = \int_{t_\ell}^{t_{\ell+1}} (W^j_t - W^j_{t_\ell})dW^k_t - \int_{t_\ell}^{t_{\ell+1}} \left(W^k_t - W^k_{t_\ell}\right) dW^j_t.$$

The Milstein scheme without Lévy area, or *approximate Milstein scheme*, is obtained by setting $A \equiv 0$, namely:

$$\Delta \bar{X}_{i,\ell} = b_i(\bar{X}_\ell)h + \sum_{j=1}^{q} \sigma_{ij}(\bar{X}_\ell)\Delta W_{j,\ell} + \sum_{j,k=1}^{q} g_{ijk}(\bar{X}_\ell)(\Delta W_{j,\ell}\Delta W_{k,\ell} - hC_{i,k}). \quad (2)$$

2.2 Variance and Complexity (Effort)

We consider a family of linear statistical estimators $(I_\pi^N)_{N\geq 1}$ of $I_0 \in \mathbb{R}$ where π lies in a parameter set $\Pi \subset \mathbb{R}^d$. By linear, we mean, on the one hand, that

$$\mathbb{E}[I_\pi^N] = \mathbb{E}[I_\pi^1], \quad N \geq 1,$$

and, on the other hand, that the numerical cost, $\mathrm{Cost}(I_\pi^N)$ induced by the simulation of I_π^N is given by

$$\mathrm{Cost}(I_\pi^N) = N\kappa(\pi)$$

where $\kappa(\pi) = \mathrm{Cost}(I_\pi^1)$ is the cost of a single simulation or *unitary complexity*. We also assume that our estimator is of *Monte Carlo type* in the sense that its variance is *inverse linear* in the size N of the simulation:

$$\mathrm{var}(I_\pi^N) = \frac{\nu(\pi)}{N}$$

where $\nu(\pi) = \mathrm{var}(I_\pi^1)$ denotes the variance of one simulation. We are looking for the "best" estimator in the family $\{(I_\pi^N), \pi \in \Pi\}$ *i.e.* the estimator minimizing the computational cost for a given $\varepsilon > 0$. This generic problem reads:

$$(\pi(\varepsilon), N(\varepsilon)) = \underset{\|I_\pi^N - I_0\|_2 \leq \varepsilon}{\mathrm{argmin}} \; \mathrm{Cost}(\mathrm{I}_\pi^\mathrm{N}). \quad (3)$$

In order to solve this minimization problem, we introduce the notion of *effort* $\phi(\pi)$ of a linear Monte Carlo type estimator I_π^N.

Definition 1. *The effort of the estimator I_π^N is defined for every $\pi \in \Pi$ by*

$$\phi(\pi) = \nu(\pi)\kappa(\pi) \quad \text{so that} \quad \mathrm{Cost}(I_\pi^N) = N\frac{\phi(\pi)}{\nu(\pi)}. \quad (4)$$

When the estimators $(I_\pi^N)_{N\geq 1}$ are *biased*, the $\mathbb{L}^2$-error becomes

$$\mathbb{E}[(I_\pi^N - I_0)^2] = \mu^2(\pi) + \frac{\nu(\pi)}{N} \quad \text{where} \quad \mu(\pi) = \mathbb{E}[I_\pi^N] - I_0 = \mathbb{E}[I_\pi^1] - I_0$$

denotes the bias (which does not depend on N).

2.3 Assumptions on Weak and Strong Approximation Errors

We consider a family of real-valued random variables $(Y_h)_{h\in\mathcal{H}}$ associated to a random variable $Y_0 \in \mathbb{L}^2$. The index set $\mathcal{H}$ is a *consistent* set of *step* parameters in the sense that $\mathcal{H} = \{\mathbf{h}/n, n \geq 1\}$ for a fixed $\mathbf{h} \in (0, +\infty)$. The family satisfies two assumptions which formalize the strong and weak rates of approximation of Y_0 by Y_h when the bias parameter $h \to 0$ in $\mathcal{H}$.

Bias Error Expansion (Weak Error Rate):

$$\exists\, \alpha > 0,\ \bar{R} \in \mathbb{N},\ \mathbb{E}[Y_h] = \mathbb{E}[Y_0] + \sum_{k=1}^{\bar{R}} c_k h^{\alpha k} + h^{\alpha \bar{R}} \eta_{\bar{R}(h)},\ \lim_{h\to 0} \eta_{\bar{R}(h)} = 0,\ (WE_{\alpha,\bar{R}})$$

where c_k, $k = 1, \ldots, \bar{R}$ are real coefficients and $\eta_{\bar{R}}$ is a real function defin ed on $\mathcal{H}$.

Strong Approximation Error Assumption:

$$\exists\, \beta > 0,\ V_1 \in \mathbb{R}_+, \quad \|Y_h - Y_0\|_2^2 = \mathbb{E}\left[|Y_h - Y_0|^2\right] \leq V_1 h^{\beta}. \qquad (SE_\beta)$$

The parameters α, β and $\bar{R}$ are structural parameters which depend on the family $(Y_h)_{h\in\mathcal{H}}$. In the sequel, we will consider a free parameter $R \in \{2, \ldots, \bar{R}\}$ for which $(WE_{\alpha,R})$ is always satisfied.

We associate to the family $(Y_h)_{h\in\mathcal{H}}$, the $\mathbb{R}^R$-valued random vector

$$Y_{h,\underline{n}} = (Y_h, Y_{\frac{h}{n_2}}, \ldots, Y_{\frac{h}{n_R}}) \qquad (5)$$

where the R-tuple of integers $\underline{n} := (n_1, n_2, \ldots, n_R) \in \mathbb{N}^R$, called *refiners*, satisfy

$$n_1 = 1 < n_2 < \cdots < n_R. \qquad (6)$$

3 Antithetic Multilevel Monte Carlo

Let $R \geq 2$ and let $(Y_{h,\underline{n}}^{(j),k})_{k\geq 1}$ be an i.i.d sequence of copies of $Y_{h,\underline{n}}$.

3.1 Brief Recall of Multilevel Monte Carlo Estimators

A *Multilevel Richardson-Romberg (ML2R) estimator* of depth R attached to an *allocation strategy* $q = (q_1, \ldots, q_R)$ with $q_j > 0$, $j = 1, \ldots, R$, $\sum_j q_j = 1$ and a weight vector $\mathbf{W}$, is defined for every integer $N \geq 1$ and $h \in \mathcal{H}$ by (see [1])

$$\bar{Y}_{h,\underline{n}}^{N,q} = \frac{1}{N_1} \sum_{k=1}^{N_1} Y_h^{(1),k} + \sum_{j=2}^{R} \frac{\mathbf{W}_j}{N_j} \sum_{k=1}^{N_j} \left(Y_{\frac{h}{n_j}}^{(j),k} - Y_{\frac{h}{n_{j-1}}}^{(j),k} \right)$$

where for all $j \in 1, \ldots, R$, $N_j = \lceil q_j N \rceil$, $\mathbf{W}_j = \mathbf{w}_j + \ldots + \mathbf{w}_R$, and

$$\mathbf{w}_i = \frac{(-1)^{R-i} n_i^{-\alpha(R-1)}}{\prod_{1\leq j<i} (n_i^\alpha - n_j^\alpha) \prod_{i<j\leq R} (n_j^\alpha - n_i^\alpha)}.$$

The vector $\mathbf{w} = (\mathbf{w}_1, \ldots, \mathbf{w}_R)$ is defined as the unique solution of a Vandermonde system (see [1] for more details).

Remark 1. If the R-level weight vector $\mathbf{W}$ satisfies $\mathbf{W} = (1, \ldots, 1)$, the estimator is called a Multilevel Monte Carlo *MLMC* estimator.

3.2 Bias Parameter and Depth *R* Optimization for Geometric Refiners

In this approach, we consider geometric refiners with *root* $M \geq 2$ of the form

$$n_i = M^{i-1},\ i = 1, \ldots, R. \tag{7}$$

Theorem 1 (see[1]). *The* ML2R *estimator satisfies*

$$\limsup_{\varepsilon \to 0} v(\beta, \varepsilon) \times \inf_{\substack{h \in \mathcal{H}, R \geq 2 \\ |\mu(h,R,q^*)| < \varepsilon}} \mathrm{Cost}\left(\bar{Y}^{N,q^*}_{h,\underline{n}}\right) \leq K(\alpha, \beta, M) < +\infty$$

$$\text{with} \quad v(\beta, \varepsilon) = \begin{cases} \varepsilon^2 (\log(1/\varepsilon))^{-1} & \text{if } \beta = 1, \\ \varepsilon^2 & \text{if } \beta > 1, \\ \varepsilon^2 e^{-\frac{1-\beta}{\sqrt{\alpha}}\sqrt{2\log(1/\varepsilon)\log(M)}} & \text{if } \beta < 1. \end{cases}$$

The MLMC *estimator satisfies (see also Giles in [3])*

$$\limsup_{\varepsilon \to 0} v(\beta, \varepsilon) \times \inf_{\substack{h \in \mathcal{H}, R \geq 2 \\ |\mu(h,R,q^*)| < \varepsilon}} \mathrm{Cost}\left(\bar{Y}^{N,q^*}_{h,\underline{n}}\right) \leq K(\alpha, \beta, M) < +\infty$$

$$\text{with} \quad v(\beta, \varepsilon) = \begin{cases} \varepsilon^2 (\log(1/\varepsilon))^{-2} & \text{if } \beta = 1, \\ \varepsilon^2 & \text{if } \beta > 1, \\ \varepsilon^{2+\frac{1-\beta}{\alpha}} & \text{if } \beta < 1. \end{cases}$$

We refer to [1] for explicit formulas of $K(\alpha, \beta, M)$.

3.3 Antithetic Multilevel Estimators

In this section we set $M = 2$. Let $Y = f(\bar{X})$ where f is a smooth payoff function and $\bar{X}$ the approximation of the Milstein scheme without Lévy area.
First we define $\bar{X}^f$ the approximate Milstein scheme with the fine time steps and its antithetic "twin" $\bar{X}^a$. This second approximate Milstein scheme is defined (or simulated) by simply inverting two by two the successive Brownian increments used to simulate $\bar{X}^f$. Finally, the approximate Milstein scheme on the coarse time step $\bar{X}^c$, associated to the payoff Y^c, is still simulated using the above Brownian increments, but grouped two by two. Hence $\bar{X}^f$ and $\bar{X}^a$ have the same conditional distribution, given $\bar{X}^c$. In particular $\mathbb{E}[Y^f] = \mathbb{E}[Y^a]$ where Y^f

and Y^a are the payoffs "written" on $\bar{X}^f$ and $\bar{X}^a$ respectively. At the same time, one has

$$(\bar{X}^f - \bar{X}^a) \approx -(\bar{X}^a - \bar{X}^c) \quad \text{and} \quad (Y^f - Y^c) \approx -(Y^a - Y^c)$$

so that $\frac{1}{2}(Y^f + Y^a) \approx Y^c$. This suggests that $\frac{1}{2}(Y^f + Y^a) - Y^c$ has a much smaller variance than the standard estimator $Y^f - Y^c$ (see [2]).

In this case, an *antithetic ML2R estimator* with weight vector $\mathbf{W}$ is given by

$$\bar{Y}^{N,q,a}_{h,\underline{n}} = \frac{1}{N_1}\sum_{k=1}^{N_1} Y_h^{(1),k,c} + \sum_{j=2}^{R} \frac{\mathbf{W}_j}{N_j}\sum_{k=1}^{N_j}\left(\frac{1}{2}\left(Y^{(j),k,f}_{\frac{h}{n_j}} + Y^{(j),k,a}_{\frac{h}{n_j}}\right) - Y^{(j),k,c}_{\frac{h}{n_{j-1}}}\right) \tag{8}$$

Remark 2. If the weight vector $\mathbf{W}$ satisfies $\mathbf{W} = (1,\ldots,1)$, we obtain the original antithetic *MLMC* estimator devised in [2].

Cost, Complexity and Effort of an Antithetic Multilevel Estimator. For a simulation size N, the numerical cost induced by the antithetic estimators $Y^{N,q,a}_{h,\underline{n}}$, $N \geq 1$, with the convention $n_0 = (n_0)^{-1} = 0$, is

$$\text{Cost}(\bar{Y}^{N,q,a}_{n,\underline{n}}) = \sum_{j=1}^{R} \frac{N_j}{h}(n_j + n_{j-1}) = N\kappa(\pi) \tag{9}$$

where $\kappa(\pi)$ is the unitary complexity given by

$$\kappa(\pi) = \frac{1}{h}\sum_{j=1}^{R} q_j(n_j + n_{j-1}). \tag{10}$$

The effort (product of variance by complexity) of such an estimator reads

$$\phi(\pi) = \left(\sum_{j=1}^{R}\frac{1}{q_j}\text{var}\left[\mathbf{W}_j\left(\frac{1}{2}\left(Y^{(j),f}_{\frac{h}{n_j}} + Y^{(j),a}_{\frac{h}{n_j}}\right) - Y^{(j),c}_{\frac{h}{n_{j-1}}}\right)\right]\right)\kappa(\pi) \tag{11}$$

with the convention $Y^{(1),c}_{\frac{h}{n_0}} \equiv 0$. The proof of Theorem 3 below relies on the following lemma which is a straightforward consequence of Schwartz's Inequality.

Lemma 1. *For all* $j = 1,\ldots,R$, *let* $a_j > 0$, $b_j > 0$ *and* $q_j > 0$ *such that* $\sum\limits_{j=1}^{R} q_j = 1$. *Then*

$$\left(\sum_{j=1}^{R}\frac{a_j}{q_j}\right)\left(\sum_{j=1}^{R} b_j q_j\right) \geq \left(\sum_{j=1}^{R}\sqrt{a_j b_j}\right)^2 \tag{12}$$

and equality holds iff $q_j = \sqrt{a_j b_j^{-1}}\left(\sum_{k=1}^{R}\sqrt{a_k b_k^{-1}}\right)^{-1}$, $j = 1,\ldots,R$.

3.4 Optimization of the Effort for Antithetic Multilevel Estimators

In this section, we first present a lemma and theorem which give an upper-bound of the antithetic Multilevel estimator and then allow us to prove the last theorem of this section which gives the optimal allocation of simulated paths across the levels of our new antithetic *ML2R* estimator. This allows to achieve the optimal complexity without simulating the Lévy areas.

Theorem 2 (see [2]).

(a) For $p \geq 2$, there exists a constant K_p, independent of step $h = \frac{T}{n}$, such that the approximate Milstein scheme $\bar{X}$ satisfies

$$\forall \ell \geq 1 \qquad \mathbb{E}\left[\max_{0\leq \ell \leq n} \|\bar{X}_\ell\|^p\right] \leq K_p. \tag{13}$$

(b) For all $p \geq 2$, there exists a constant K'_p such that

$$\forall \ell \geq 1 \qquad \mathbb{E}\left[\max_{0\leq \ell \leq n} \left\|\frac{1}{2}(\bar{X}^f_\ell + \bar{X}^a_\ell) - \bar{X}^c_\ell\right\|^p\right] \leq K'_p h^p \tag{14}$$

with $h = \frac{T}{n}$ the coarse time step. Hence one has $\beta = 2$, for $p = 2$.

The next theorem is the main result for the antithetic estimators.

Theorem 3 (Allocation optimization). *The minimal effort ϕ^* of an antithetic multilevel estimator satisfies*

$$\phi^*(\pi_0) \leq \bar{\phi}(\pi_0, q^*) = \frac{1}{h}\left(\sqrt{K_p} + h\sqrt{K'_p}\sum_{j=2}^{R} |\mathbf{W}_j| \frac{\sqrt{n_j + n_{j-1}}}{n_{j-1}}\right)^2$$

where $q^* = q^*(\pi_0)$ is the allocation strategy given by

$$\begin{cases} q^*_1(\pi_0) = \mu^* \sqrt{K_p} \\ q^*_j(\pi_0) = \mu^* h\sqrt{K'_p}\left(|\mathbf{W}_j| \frac{n_{j-1}^{-1}}{\sqrt{n_j + n_{j-1}}}\right), \quad j = 2, \ldots, R \end{cases}$$

and μ^* is the normalizing constant such that $\sum_{j=1}^{R} q^*_j = 1$ (since $\beta = 2$, see [1]). The constants K_p and K'_p are respectively defined in (13) and (14).

Proof. Consider Theorem 2 with ($p = 2$). Using the independence of the random variables, one has, $\forall j \geq 2$,

$$\begin{aligned} \mathrm{var}(\bar{Y}^{N,q,a}_{h,\underline{n}}) &= \frac{\mathrm{var}(Y^{(1),f}_h)}{N_1} + \sum_{j=2}^{R} \frac{(\mathbf{W}_j)^2}{N_j} \mathrm{var}\left[\frac{1}{2}\left(Y^{(j),f}_{\frac{h}{n_j}} + Y^{(j),a}_{\frac{h}{n_j}}\right) - Y^{(j),c}_{\frac{h}{n_{j-1}}}\right] \\ &\leq \frac{K_p}{N_1} + \sum_{j=2}^{R} \frac{(\mathbf{W}_j)^2}{N_j} K'_p (h/n_{j-1})^2 = \frac{K_p}{N_1} + \sum_{j=2}^{R} \frac{(\mathbf{W}_j)^2}{N_j} K'_p h^2 n_{j-1}^{-2}. \end{aligned}$$

Using (10), one derives an upper-bound for the effort $\phi(\pi) \leq \bar{\phi}(\pi)$, with

$$\bar{\phi}(\pi) = \frac{1}{h}\left(\frac{K_p}{q_1} + \sum_{j=2}^{R} \frac{(\mathbf{W}_j)^2}{q_j} K'_p h^2 n_{j-1}^{-2}\right)\left(\sum_{j=1}^{R} q_j(n_j + n_{j-1})\right). \qquad (15)$$

Using Lemma 1 with $a_1 = K_p$, $b_1 = 1$, $a_j = K'_p h^2 (\mathbf{W}_j)^2\, n_{j-1}^{-2}$ et $b_j = (n_j + n_{j-1})$, $j = 2, \ldots, R$, the proof is complete.

Remark 3. If $\mathbf{W}_j = 1$, $j = 1, \ldots, R$ then, one recovers Giles and Szpruch' antithetic *MLMC* estimator.

4 Numerical Experiments

Practitioner's Corner. To implement the antithetic Multilevel estimators (8), we need to know both the weak and strong rates of convergence of the biased estimator Y_h toward Y_0. The Milstein scheme is a second order scheme which satisfies (SE_β) with $\beta = 2$ and $(WE_{\alpha,\bar{R}})$ for $\alpha = 1$, at least when dealing with "vanilla" Lipschitz poyoffs.
Note that in the antithetic case for multidimensional SDEs, the root M of Eq. (7) is set to $M = 2$ to allow the permutation of the two successive pair of Brownian increments coming from the fine time step to simulate the antithetic paths. A natural approximation of K_p and K'_p respectively from (13) and (14) is $K_p \sim \operatorname{var}(Y_h^{(1)})$ and $K'_p \sim \operatorname{var}\left[\frac{1}{2}(Y_{\frac{h}{2}}^{(1),f} + Y_{\frac{h}{2}}^{(1),a}) - Y_h^{(1),c}\right] h^2$.

Best-of-Call option by 2D Milstein scheme, $\alpha = 1$, $\beta = 2$. We consider the diffusion Eq. (1) with $d = 2$, $b(X_t) = rX_t$, $\sigma(X_t) = \sigma X_t$ (2D Black-Scholes model), $X_0^1 = 100$, $X_0^2 = 100$, $r = 0.15$, $\sigma_1 = 0.3$, $\sigma_1 = 0.4$ and the Milstein scheme without Lévy area defined by (2) with $q = 2$ and $C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The payoff of a Best-of-Call option is given by

$$f(x_1, x_2) = e^{-rT}(\max(x_1(T), x_2(T)) - K)_+ \quad (x_1, x_2) \in \mathcal{C}([0, T], \mathbb{R}^2)$$

with maturity $T = 1$ and strike $K = 80$ and the biased estimator $Y_h = f(\bar{X}_T^1, \bar{X}_T^2)$. For the numerical simulation, we apply the antithetic procedure clearly detailed in [2]. The constants K_p and K'_p have been estimated by $K_p = 7.55$ and $K'_p = 965.04$ and the bias is computed using the exact price $I_0 = 51.0867$. The results are summarized in Table 1 for *ML2R* estimator and in Table 2 for *MLMC* estimator (see the Appendix). Note that the depth R required to achieve the optimal complexity is more important for the *MLMC* estimator (Tables 1 and 2).

This leads to a reduction of the computational cost with our new *ML2R* estimator which is more efficient than the *MLMC*. Figure 2 shows that *MLMC* takes 50 to 60% more time than *ML2R*. Other experiments carried out on Margrabe model lead to the same conclusions.

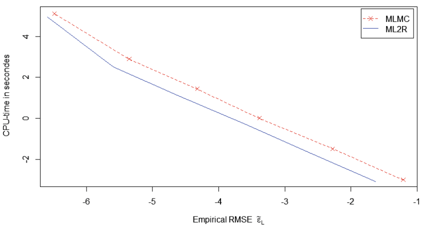

Fig. 1. CPU-time (y-axis log scale) as a function of $\tilde{\varepsilon}_L$ (x-axis $\log_2$ scale)

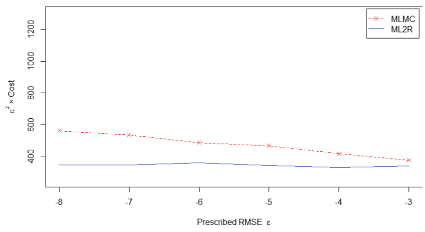

Fig. 2. $\varepsilon^2 \times$Cost as a function of ε ($\log_2$ scale for the x-axis)

Acknowledgements. The work of Cheikh Mbaye is supported by the *Fédération Wallonie-Bruxelles* and the *National Bank of Belgium* via an *FSR* grant.

The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the *National Bank of Belgium.*

Appendix: Main Results for Antithetic Estimators

Table 1. Best-of-Call option: parameters and results of *ML2R* estimator.

k	$\varepsilon = 2^{-k}$	$\mathbb{L}^2$-error	Time (s)	Bias	R	M	$\mathbf{h}^{-1}$	N	Cost
3	$1.25 \cdot 10^{-01}$	$3.23 \cdot 10^{-01}$	$4.69 \cdot 10^{-02}$	$-1.85 \cdot 10^{-02}$	4	2	1	$1.64 \cdot 10^{+04}$	$2.18 \cdot 10^{+04}$
4	$6.25 \cdot 10^{-02}$	$1.64 \cdot 10^{-02}$	$1.84 \cdot 10^{-01}$	$-2.07 \cdot 10^{-03}$	4	2	1	$6.38 \cdot 10^{+04}$	$8.46 \cdot 10^{+04}$
5	$3.12 \cdot 10^{-02}$	$8.33 \cdot 10^{-02}$	$7.50 \cdot 10^{-01}$	$8.36 \cdot 10^{-03}$	5	2	1	$2.62 \cdot 10^{+05}$	$3.53 \cdot 10^{+05}$
6	$1.56 \cdot 10^{-02}$	$4.05 \cdot 10^{-02}$	$3.15 \cdot 10^{+00}$	$1.99 \cdot 10^{-03}$	5	2	1	$1.09 \cdot 10^{+06}$	$1.48 \cdot 10^{+06}$
7	$7.81 \cdot 10^{-03}$	$2.07 \cdot 10^{-02}$	$1.24 \cdot 10^{+01}$	$-1.03 \cdot 10^{-03}$	5	2	1	$4.20 \cdot 10^{+06}$	$5.67 \cdot 10^{+06}$
8	$3.91 \cdot 10^{-03}$	$1.04 \cdot 10^{-02}$	$1.40 \cdot 10^{+02}$	$9.00 \cdot 10^{-04}$	5	2	1	$1.67 \cdot 10^{+07}$	$2.25 \cdot 10^{+07}$

Table 2. Best-of-Call option: parameters and results of *MLMC* estimator.

k	$\varepsilon = 2^{-k}$	$\mathbb{L}^2$-error	Time (s)	Bias	R	M	$\mathbf{h}^{-1}$	N	Cost
3	$1.25 \cdot 10^{-01}$	$4.33 \cdot 10^{-01}$	$5.08 \cdot 10^{-02}$	$-3.18 \cdot 10^{-01}$	5	2	1	$1.86 \cdot 10^{+04}$	$2.41 \cdot 10^{+04}$
4	$6.25 \cdot 10^{-02}$	$2.07 \cdot 10^{-01}$	$2.30 \cdot 10^{-01}$	$-1.47 \cdot 10^{-01}$	6	2	1	$8.08 \cdot 10^{+04}$	$1.06 \cdot 10^{+05}$
5	$3.12 \cdot 10^{-02}$	$9.59 \cdot 10^{-02}$	$1.02 \cdot 10^{+00}$	$-6.49 \cdot 10^{-02}$	7	2	1	$3.57 \cdot 10^{+05}$	$4.80 \cdot 10^{+05}$
6	$1.56 \cdot 10^{-02}$	$4.99 \cdot 10^{-02}$	$4.23 \cdot 10^{+00}$	$-3.53 \cdot 10^{-02}$	8	2	1	$1.47 \cdot 10^{+06}$	$2.00 \cdot 10^{+06}$
7	$7.81 \cdot 10^{-03}$	$2.45 \cdot 10^{-02}$	$1.80 \cdot 10^{+01}$	$-1.75 \cdot 10^{-02}$	9	2	1	$6.37 \cdot 10^{+06}$	$8.78 \cdot 10^{+06}$
8	$3.91 \cdot 10^{-03}$	$1.12 \cdot 10^{-03}$	$1.65 \cdot 10^{+02}$	$-7.31 \cdot 10^{-03}$	10	2	1	$2.64 \cdot 10^{+07}$	$3.67 \cdot 10^{+07}$

Optimal Parameters for the Antithetic Estimators. In all the numerical experiments presented in the paper, we set $\mathbf{h} = T$ in the definition of the parameters of the simulation for both *MLMC* and *ML2R* estimators. In the table below, keep in mind that $n_0 = 0$ by convention (Table 3).

Table 3. Optimal parameters for antithetic multilevel estimators.

	ML2R	*MLMC*
$q(\varepsilon)$	$q_1 = \mu^* \sqrt{K_p}$ $q_j = \mu^* h \sqrt{K'_p} \left(\lvert \mathbf{W}_j \rvert \frac{n_{j-1}^{-1}}{\sqrt{n_j + n_{j-1}}} \right),\ j = 2 \ldots, R$	$q_1 = \mu^* \sqrt{K_p}$ $q_j = \mu^* h \sqrt{K'_p} \left(\frac{n_{j-1}^{-1}}{\sqrt{n_j + n_{j-1}}} \right)\ j = 2 \ldots, R$
$N(\varepsilon)$	$\left(1 + \frac{1}{2\alpha R}\right) \frac{\left(1 + h\sqrt{K'_p} \sum_{j=1}^{R} \lvert \mathbf{W}_j \rvert \left(n_{j-1}^{-1}\right) \sqrt{n_j + n_{j-1}}\right)^2}{\varepsilon^2 \sum_{j=1}^{R} q_j (n_j + n_{j-1})}$	$\left(1 + \frac{1}{2\alpha}\right) \frac{\left(1 + h\sqrt{K'_p} \sum_{j=1}^{R} \left(n_{j-1}^{-1}\right) \sqrt{n_j + n_{j-1}}\right)^2}{\varepsilon^2 \sum_{j=1}^{R} q_j (n_j + n_{j-1})}$

Note that in the table above, μ^* is defined such that $\sum_{j=1}^{R} q_j = 1$. For the optimal parameters R and h, their formulas remain the same as in [1].

References

1. Lemaire, V., Pagès, G.: Multilevel Richardson-Romberg extrapolation. pre-pub. LPMA 1603 arXiv:1401.1177 (2014). (forthcoming in Bernoulli)
2. Giles, M.B., Szpruch, L.: Antithetic multilevel Monte Carlo for multi-dimensional SDEs without Lévy area simulation. Ann. Appl. Prob. **24**(4), 1585–1620 (2014)
3. Giles, M.B.: Multilevel Monte Carlo path simulation. Oper. Res. **56**(3), 607–617 (2008)