

Multilevel Richardson-Romberg extrapolation

Séminaire Bachelier

V. Lemaire & G. Pagès

LPMA-UPMC

June 27, 2014

1 Introduction

- Abstract framework and assumptions
- Monte Carlo type estimators
- Crude Monte Carlo estimator in a biased framework
- Multistep Richardson-Romberg estimator

2 Multilevel estimators

- Definition
- Majoration of the effort
- specification(s) of the allocation matrix \mathbf{T}
- Convergence for a fixed $R \geq 2$
- Main result

3 Applications and numerical experiments

- Parameters and methodology
- Discretization of Brownian diffusions
- Nested Monte Carlo

1 Introduction

- Abstract framework and assumptions
- Monte Carlo type estimators
- Crude Monte Carlo estimator in a biased framework
- Multistep Richardson-Romberg estimator

2 Multilevel estimators

- Definition
- Majoration of the effort
- specification(s) of the allocation matrix \mathbf{T}
- Convergence for a fixed $R \geq 2$
- Main result

3 Applications and numerical experiments

- Parameters and methodology
- Discretization of Brownian diffusions
- Nested Monte Carlo

Abstract linear simulation framework

- $Y_0 \in \mathbf{L}^2(\mathbf{P})$ **not simulatable** real-valued random variable.
- Aim: compute $I_0 = \mathbf{E}[Y_0]$ with a given accuracy $\varepsilon > 0$.
- Suppose we have family $(Y_h)_{h \in \mathcal{D}}$ of real-valued random variables in $\mathbf{L}^2(\mathbf{P})$ satisfying:
 - ▶ Bias error expansion (weak error expansion):

$$\exists \alpha > 0, \mathbf{E}[Y_h] = \mathbf{E}[Y_0] + c_1 h^\alpha + c_2 h^{2\alpha} + \dots + c_R h^{\alpha R} (1 + \eta_R(h)) \quad (WE_{\alpha,R})$$

- ▶ Strong approximation error assumption:

$$\exists \beta > 0, \quad \|Y_h - Y_0\|_2^2 = \mathbf{E}[|Y_h - Y_0|^2] \leq V_1 h^\beta \quad (SE_\beta)$$

- h is the *bias parameter* taking value in $\mathcal{D} \subset \{\frac{T}{n}, n \geq T/\mathbf{h}\}$ (so that $\mathbf{h} = \max \mathcal{D}$).

Monte Carlo type estimators

We consider **Monte Carlo type estimators** I_π^N depending on some parameters π to compute

$$I_0 = \mathbf{E}[Y_0]$$

i.e. satisfying

- **Constant bias:** $\mathbf{E}[I_\pi^N] = \mathbf{E}[I_\pi^1]$ for all $N \geq 1$
- **Inverse linear variance:** $\text{var}(I_\pi^N) = \frac{\mathfrak{v}(\pi)}{N}$ where $\mathfrak{v}(\pi) = \text{var}(I_\pi^1)$
- **Linear cost:** $\text{Cost}(I_\pi^N) = N \kappa(\pi)$ where $\kappa(\pi) = \text{Cost}(I_\pi^1)$

Typically I_π^N is the N -empirical mean of i.d.d. r.v.

Aim: minimize the global cost of the estimator for a given error $\varepsilon > 0$.

$$(\pi(\varepsilon), N(\varepsilon)) = \underset{\|I_\pi^N - I_0\|_2 \leq \varepsilon}{\operatorname{argmin}} \text{Cost}(I_\pi^N).$$

Definition

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

$$\phi(\pi) = \nu(\pi) \kappa(\pi).$$

- If the estimator I_π^N is *unbiased* then $\|I_\pi^1 - I_0\|_2^2 = \nu(\pi)$ so that

$$\pi(\varepsilon) = \pi_* = \operatorname{argmin}_{\pi \in \Pi} \phi(\pi), \quad N(\varepsilon) = \frac{\nu(\pi_*)}{\varepsilon^2} = \frac{\phi(\pi_*)}{\kappa(\pi_*)\varepsilon^2}.$$

- If the estimator I_π^N is *biased* then $\|I_\pi^1 - I_0\|_2^2 = \nu(\pi) + \mu^2(\pi)$ where

$$\mu^2(\pi) = (\mathbf{E}[I_\pi^N] - I_0)^2 = (\mathbf{E}[I_\pi^1] - I_0)^2$$

so that

$$\pi(\varepsilon) = \operatorname{argmin}_{\pi \in \Pi, |\mu(\pi)| < \varepsilon} \left(\frac{\phi(\pi)}{\varepsilon^2 - \mu^2(\pi)} \right), \quad N(\varepsilon) = \frac{\phi(\pi(\varepsilon))}{\kappa(\pi(\varepsilon))(\varepsilon^2 - \mu^2(\pi(\varepsilon)))}.$$

Theorem (Crude Monte Carlo, $\pi = h$)

The Monte Carlo estimator of $\mathbf{E}[Y_0]$ defined by $\bar{Y}_h^N = \frac{1}{N} \sum_{k=1}^N Y_h^k$ satisfies

$$\mu(h) \sim c_1 h^\alpha, \quad \kappa(h) = \frac{1}{h}, \quad \phi(h) = \frac{\text{var}(Y_h)}{h}.$$

Then, the optimal parameters $h^*(\varepsilon)$ and $N^*(\varepsilon)$ are

$$h^*(\varepsilon) = \frac{(1 + 2\alpha)^{-\frac{1}{2\alpha}}}{|c_1|^{1/\alpha}} \varepsilon^{\frac{1}{\alpha}}, \quad N^*(\varepsilon) = \left(1 + \frac{1}{2\alpha}\right) \frac{\text{var}(Y_0)(1 + \theta h^*(\varepsilon)^{\frac{\beta}{2}})^2}{\varepsilon^2}.$$

where $\theta = \sqrt{\frac{V_1}{\text{var}(Y_0)}}$. Furthermore, we have

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^{2 + \frac{1}{\alpha}} \min_{\substack{h \in \mathcal{D}, \\ |\mu(h)| < \varepsilon}} \text{Cost}(\bar{Y}_h^N) \leq |c_1|^{\frac{1}{\alpha}} \left(1 + \frac{1}{2\alpha}\right) (1 + 2\alpha)^{\frac{1}{2\alpha}} \text{var}(Y_0).$$

Multistep Richardson-Romberg estimator [P. 2007, MCMA]

- Aim: Kill the bias using the weak error expansion ($WE_{\alpha,R}$)
- *Refiners*: R -tuple of integers $\underline{n} := (n_1, n_2, \dots, n_R) \in \mathbf{N}^R$ satisfying

$$n_1 = 1 < n_2 < \dots < n_R.$$

- We consider $Y_{h,\underline{n}}$ the \mathbf{R}^R -valued random vector

$$Y_{h,\underline{n}} = \left(Y_h, Y_{\frac{h}{n_2}}, \dots, Y_{\frac{h}{n_R}} \right)$$

- We define the vector $\mathbf{w} = (\mathbf{w}_1, \dots, \mathbf{w}_R)$ as the unique solution to the *Vandermonde system* $V \mathbf{w} = e_1$ where

$$V = V(1, n_2^{-\alpha}, \dots, n_R^{-\alpha}) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & n_2^{-\alpha} & \dots & n_R^{-\alpha} \\ \vdots & \vdots & \dots & \vdots \\ 1 & n_2^{-\alpha(R-1)} & \dots & n_R^{-\alpha(R-1)} \end{pmatrix}.$$

- Key point: The solution \mathbf{w} has a **closed form** given by Cramer's rule:

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

- We also derive the following identity of interest in what follows

$$\tilde{\mathbf{w}}_{R+1} := \sum_{i=1}^R \frac{\mathbf{w}_i}{n_i^{\alpha R}} = \frac{(-1)^{R-1}}{\underline{n}!^\alpha}.$$

Multistep Richardson-Romberg estimator

Definition (Multistep Richardson-Romberg estimator)

$$\begin{aligned}\bar{Y}_{h,\underline{n}}^N &= \frac{1}{N} \sum_{k=1}^N \langle \mathbf{w}, Y_{h,\underline{n}}^k \rangle, \quad (Y_{h,\underline{n}}^k)_{k \geq 1} \text{ i.i.d.} \\ &= \frac{1}{N} \sum_{k=1}^N \sum_{i=1}^R \mathbf{w}_i Y_{\frac{h}{n_i}}^k\end{aligned}$$

Note that $\pi = (h, \underline{n}, R)$ and

$$\begin{aligned}\mu(\pi) &= \mathbf{E} \left[\langle \mathbf{w}, Y_{h,\underline{n}}^1 \rangle \right] = \langle \mathbf{w}, \mathbf{E} \left[Y_{h,\underline{n}}^1 \right] \rangle \\ &= \mathbf{E} [Y_0] + \tilde{\mathbf{w}}_{R+1} c_R h^{\alpha_R} (1 + \eta_{R,n}(h))\end{aligned}$$

Let $|\underline{n}| = n_1 + \cdots + n_R$ and $\underline{n}! = n_1 \cdots n_R$.

Theorem (Multistep Richardson-Romberg)

The Multistep Richardson-Romberg estimator of $\mathbf{E}[Y_0]$ satisfies

$$\mu(h) \simeq (-1)^{R-1} c_R \left(\frac{h^R}{\underline{n}!} \right)^\alpha, \quad \kappa(h) = \frac{|\underline{n}|}{h}, \quad \phi(h) \simeq \frac{|\underline{n}| \operatorname{var}(Y_0)}{h}$$

and for a fixed $R \geq 2$ the optimal parameters $h^*(\varepsilon)$ and $N^*(\varepsilon)$ are

$$h^*(\varepsilon) = \frac{(1 + 2\alpha R)^{-\frac{1}{2\alpha R}}}{|c_R|^{1/(\alpha R)}} \varepsilon^{\frac{1}{\alpha R}}, \quad N^*(\varepsilon) = \left(1 + \frac{1}{2\alpha R}\right) \frac{\operatorname{var}(Y_0) (1 + \theta h^*(\varepsilon)^{\frac{\beta}{2}})^2}{\varepsilon^2}.$$

Furthermore, we have

$$\inf_{\substack{h \in \mathcal{D} \\ |\mu(h)| < \varepsilon}} \operatorname{Cost}(\bar{Y}_h^N) \sim \left(\frac{(1 + 2\alpha R)^{1 + \frac{1}{2\alpha R}}}{2\alpha R} \right) \frac{|c_R|^{\frac{1}{\alpha R}} \operatorname{var}(Y_0)}{\varepsilon^{2 + \frac{1}{\alpha R}}} \frac{|\underline{n}|}{\underline{n}!^{\frac{1}{R}}} \quad \text{as } \varepsilon \rightarrow 0.$$

Remarks on the Multistep estimator

- If $n_i = i$ then $|\underline{n}| = \frac{R(R+1)}{2}$ and $(n!)^{\frac{1}{R}} \sim \frac{R}{e}$ so that $\frac{|\underline{n}|}{n!^{\frac{1}{R}}} \sim \frac{e(R+1)}{2}$.
- If $n_i = M^{i-1}$ ($M \geq 2$) then $\frac{|\underline{n}|}{n!^{\frac{1}{R}}} \sim M^{\frac{R-1}{2}}$

How to avoid this drawback ?

Using control variance technique such as Multilevel:

- We consider R *independent copies* $(Y_{h,\underline{n}}^{(j)})$, $j = 1, \dots, R$ of the random vector $Y_{h,\underline{n}}$ and we *split* $\langle \mathbf{w}, Y_{h,\underline{n}} \rangle$ into

$$\langle \mathbf{w}, Y_{h,\underline{n}} \rangle \rightsquigarrow \sum_{j=1}^R \langle \mathbf{T}^j, Y_{h,\underline{n}}^{(j)} \rangle = \sum_{i,j=1}^R \mathbf{T}_i^j Y_{\frac{h}{n_i}}^{(j)}$$

where $\mathbf{T} = [\mathbf{T}^1 \dots \mathbf{T}^R]$ is an $R \times R$ matrix such that $\sum_j \mathbf{T}^j = \mathbf{w}$.

1 Introduction

- Abstract framework and assumptions
- Monte Carlo type estimators
- Crude Monte Carlo estimator in a biased framework
- Multistep Richardson-Romberg estimator

2 Multilevel estimators

- Definition
- Majoration of the effort
- specification(s) of the allocation matrix \mathbf{T}
- Convergence for a fixed $R \geq 2$
- Main result

3 Applications and numerical experiments

- Parameters and methodology
- Discretization of Brownian diffusions
- Nested Monte Carlo

Definitions

- **Allocation matrix:** An $R \times R$ -matrix $\mathbf{T} = [\mathbf{T}^1 \dots \mathbf{T}^R]$ satisfying

$$\mathbf{T}^1 = e_1 \quad \text{and} \quad \forall j \in \{2, \dots, R\}, \quad \sum_{i=1}^R \mathbf{T}_i^j = 0. \quad (1)$$

so that

$$\sum_{i,j=1}^d \mathbf{T}_i^j = 1.$$

- **Stratification/Allocation strategy:** Let $q = (q_1, \dots, q_R)$, $q_j > 0$, $j = 1, \dots, R$ and $\sum_j q_j = 1$. Let $N_j = \lceil q_j N \rceil$.
- **Multilevel estimator of order R**

$$\bar{Y}_{h,\underline{n}}^{N,q} = \sum_{j=1}^R \frac{1}{N_j} \sum_{k=1}^{N_j} \langle \mathbf{T}^j, Y_{h,\underline{n}}^{(j),k} \rangle \approx \frac{1}{N} \sum_{j=1}^R \frac{1}{q_j} \sum_{k=1}^{N_j} \langle \mathbf{T}^j, Y_{h,\underline{n}}^{(j),k} \rangle \quad (2)$$

- ▶ **Multilevel Monte Carlo** $\equiv \sum_{j=1}^R \mathbf{T}^j = e_R$
- ▶ **Multilevel Richardson-Romberg** $\equiv \sum_{j=1}^R \mathbf{T}^j = \mathbf{w}$

Effort of a Multilevel estimator

- Parameters: $\pi = (\pi_0, q)$ with $\pi_0 = (h, \underline{n}, R, \mathbf{T})$.
- (Unitary) Cost of a Multilevel estimator

$$\text{Cost}(\bar{Y}_{h,\underline{n}}^{N,q}) = \sum_{j=1}^R N_j \sum_{i=1}^R \frac{1}{h} n_i \mathbf{1}_{\{\mathbf{T}_i^j \neq 0\}} = N \kappa(\pi)$$

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

$$\kappa(\pi) = \frac{1}{h} \sum_{j=1}^R q_j \sum_{i=1}^R n_i \mathbf{1}_{\{\mathbf{T}_i^j \neq 0\}}.$$

- Effort of a Multilevel estimator:

$$\phi(\pi) = \mathfrak{v}(\pi) \kappa(\pi) = \left(\sum_{j=1}^R \frac{1}{q_j} \text{var} \left(\langle \mathbf{T}^j, Y_{h,\underline{n}}^{(j)} \rangle \right) \right) \kappa(\pi)$$

Theorem (Bias)

Assume $(WE_{\alpha,R})$ holds.

(i) *Multilevel Richardson-Romberg estimator*

$$\mu(\pi_0, q) \simeq (-1)^{R-1} c_R \left(\frac{h^R}{\underline{n}!} \right)^\alpha$$

(ii) *Multilevel Monte Carlo estimator*

$$\mu(\pi_0, q) \simeq c_1 \left(\frac{h}{n_R} \right)^\alpha$$

In both cases, the bias **does not depend on the stratification strategy q selected in the simplex $\mathcal{S}_+(R)$.**

Steps to minimize the global cost

Step 1: Minimize the effort ϕ over admissible stratification strategies

$$q^* = q^*(\pi_0) = \operatorname{argmin}_{q \in \mathcal{S}_+(R)} \bar{\phi}(\pi_0, q) \quad \text{where} \quad \bar{\phi}(\pi) \geq \phi(\pi),$$

Let $\phi^*(\pi_0) = \phi(\pi_0, q^*)$ be the **optimally stratified effort**.

Step 2: Minimizing the resulting cost as a function of the remaining parameters π_0 for a given target error ε

$$\pi_0(\varepsilon) = \operatorname{argmin}_{\substack{\pi_0 \in \Pi_0 \\ |\mu(\pi_0, q^*)| < \varepsilon}} \left(\frac{\phi^*(\pi_0)}{\varepsilon^2 - \mu^2(\pi_0, q^*)} \right),$$

$$N(\pi_0(\varepsilon)) = \frac{\phi^*(\pi_0(\varepsilon))}{\kappa(\pi_0(\varepsilon), q^*)(\varepsilon^2 - \mu^2(\pi_0(\varepsilon), q^*))}.$$

Theorem (Optimal stratification (no weak error expansion needed))

Assume (SE_β) holds, and let $\theta = \sqrt{\frac{V_1}{\text{var}(Y_0)}}$. Then the *optimally stratified effort* ϕ^* satisfies

$$\phi^*(\pi_0) \leq \frac{\text{var}(Y_0)}{h} \left(1 + \theta h^{\frac{\beta}{2}} \sum_{j=1}^R \left(\sum_{i=1}^R |\mathbf{T}_i^j| n_i^{-\frac{\beta}{2}} \right) \left(\sum_{i=1}^R n_i \mathbf{1}_{\{\mathbf{T}_i^j \neq 0\}} \right)^{\frac{1}{2}} \right)^2$$

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

$$\begin{cases} q_1^*(\pi_0) = \mu^*(1 + \theta h^{\frac{\beta}{2}}) \\ q_j^*(\pi_0) = \mu^* \theta h^{\frac{\beta}{2}} \left(\sum_{i=1}^R |\mathbf{T}_i^j| n_i^{-\frac{\beta}{2}} \right) \left(\sum_{i=1}^R n_i \mathbf{1}_{\{\mathbf{T}_i^j \neq 0\}} \right)^{-\frac{1}{2}}, \quad j = 2, \dots, R, \end{cases} \quad (3)$$

with μ^* is the normalizing constant such that $\sum_{j=1}^R q_j^* = 1$.

Key Lemma for the stratification

Lemma (Hölder Inequality!)

For all $j \in \{1, \dots, R\}$ let $a_j > 0$, $b_j > 0$ and $q > 0$ such that $\sum q_j = 1$.
Then

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

and the equality case occurs if and only if $q_j = \mu \sqrt{a_j b_j^{-1}}$ with

$$\mu = \left(\sum_{k=1}^R \sqrt{a_k b_k^{-1}} \right)^{-1}.$$

Upper bound of the effort

- Using strong error assumption we prove that $\phi(\pi) \leq \bar{\phi}(\pi)$

$$\bar{\phi}(\pi) = \frac{\text{var}(Y_0)}{h} \left(\frac{(1 + \theta h^{\frac{\beta}{2}})^2}{q_1} + \theta^2 h^\beta \sum_{j=2}^R \frac{1}{q_j} \left(\sum_{i=1}^R |\mathbf{T}_i^j| n_i^{-\frac{\beta}{2}} \right)^2 \right) \times \left(\sum_{i,j=1}^R q_j n_i \mathbf{1}_{\{\mathbf{T}_i^j \neq 0\}} \right).$$

- Applying Lemma with $a_1 = (1 + \theta h^{\frac{\beta}{2}})^2$, $b_1 = 1$ and

$$a_j = \theta^2 h^\beta \left(\sum_{i=1}^R |\mathbf{T}_i^j| n_i^{-\frac{\beta}{2}} \right)^2, \quad b_j = \sum_{i=1}^R n_i \mathbf{1}_{\{\mathbf{T}_i^j \neq 0\}}, \quad j \in \{2, \dots, R\}$$

completes the proof.

MLRR

Among several possibilities we specify the allocation matrix \mathbf{T} as follows:

$$\mathbf{T} = \begin{pmatrix} \mathbf{W}_1 & -\mathbf{W}_2 & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{W}_2 & -\mathbf{W}_3 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \mathbf{W}_{R-1} & -\mathbf{W}_R \\ 0 & \cdots & \cdots & \cdots & 0 & \mathbf{W}_R \end{pmatrix}.$$

where

$$\mathbf{W}_j = \sum_{k=j}^R \mathbf{w}_k, \quad j = 1, \dots, R$$

so that $\mathbf{W}_1 = 1$.

MLMC

$$\mathbf{T} = \begin{pmatrix} 1 & -1 & 0 & \dots & \dots & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 & -1 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Let $R \geq 2$ and the n_i , $i = 1, \dots, R$ be fixed parameters.

Theorem (Bias parameter optimization (MLRR))

Assume $(WE_{\alpha,R})$ and (SE_{β}) hold.

A *Multilevel Richardson-Romberg estimator* satisfies

$$\inf_{\substack{h \in \mathcal{D} \\ |\mu(h, q^*)| < \varepsilon}} \text{Cost} \left(\bar{Y}_{h, \underline{n}}^{N, q^*} \right) \sim \left(\frac{(1 + 2\alpha R)^{1 + \frac{1}{2\alpha R}}}{2\alpha R} \right) \frac{|c_R^{\frac{1}{\alpha R}}| \text{var}(Y_0)}{\varepsilon^{2 + \frac{1}{\alpha R}}} \frac{1}{\underline{n}!^{\frac{1}{R}}} \quad \text{as } \varepsilon \rightarrow 0,$$

with q^* defined by (3). *Note that $|\underline{n}|$ disappeared...*

This asymptotically optimal bound is achieved with the bias parameter:

$$h^*(\varepsilon, R) = (1 + 2\alpha R)^{-\frac{1}{2\alpha R}} \left(\frac{\varepsilon}{|c_R|} \right)^{\frac{1}{\alpha R}} \underline{n}!^{\frac{1}{R}}. \quad (4)$$

Let $R \geq 2$ and let n_i , $i = 1, \dots, R$ be fixed parameters.

Theorem (Bias parameter optimization (MLMC))

Assume $(WE_{\alpha,R})$ and (SE_{β}) hold.

A *Multilevel Monte Carlo estimator* satisfies

$$\inf_{\substack{h \in \mathcal{D} \\ |\mu(h, q^*)| < \varepsilon}} \text{Cost} \left(\bar{Y}_{h, \underline{n}}^{N, q^*} \right) \sim \left(\frac{(1 + 2\alpha)^{1 + \frac{1}{2\alpha}}}{2\alpha} \right) \frac{|c_1|^{\frac{1}{\alpha}} \text{var}(Y_0)}{n_R \varepsilon^{2 + \frac{1}{\alpha}}} \quad \text{as } \varepsilon \rightarrow 0,$$

with q^* defined by (3).

This asymptotically optimal bound is achieved with a bias parameter given by

$$h^*(\varepsilon, R) = (1 + 2\alpha)^{-\frac{1}{2\alpha}} \left(\frac{\varepsilon}{|c_1|} \right)^{\frac{1}{\alpha}} n_R. \quad (5)$$

Theorem (Main (MLRR) result when $n_i = M^{i-1}$, $i = 1, \dots, R$.)

Assume $\lim_R |c_R|^{1/R} = \tilde{c} \in (0, +\infty)$. The *Multilevel Richardson-Romberg estimator* satisfies

$$\overline{\lim}_{\varepsilon \rightarrow 0} v(\beta, \varepsilon) \times \inf_{\substack{h \in \mathcal{D}, R \geq 2 \\ |\mu(h, R, q^*)| < \varepsilon}} \text{Cost} \left(Y_{h, \underline{n}}^{N, q} \right) \leq K(\alpha, \beta, M)$$

with

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

These bounds are achieved with an order

$$R^*(\varepsilon) = \left\lfloor \frac{1}{2} + \frac{\log(\mathbf{h} \tilde{c}_\alpha^1)}{\log(M)} + \sqrt{\left(\frac{1}{2} + \frac{\log(\mathbf{h} \tilde{c}_\alpha^1)}{\log(M)} \right)^2 + 2 \frac{\log(1/\varepsilon)}{\alpha \log(M)}} \right\rfloor$$

and the bias parameter $h^* = h^*(\varepsilon, R(\varepsilon)) \in (0, \mathbf{h}]$ is given by (4).

Theorem (Main result (MLMC))

The *Multilevel Monte Carlo estimator* satisfies

$$\overline{\lim}_{\varepsilon \rightarrow 0} v(\beta, \varepsilon) \times \inf_{\substack{h \in \mathcal{D}, R \geq 2 \\ |\mu(h, R, q^*)| < \varepsilon}} \text{Cost} \left(Y_{h, \underline{n}}^{N, q} \right) \leq K(\alpha, \beta, M)$$

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

These bounds are achieved with an order

$$R^*(\varepsilon) = \left\lceil 1 + \frac{\log((1 + 2\alpha)^{\frac{1}{2\alpha}} |c_1|^{\frac{1}{\alpha}} \mathbf{h})}{\log(M)} + \frac{\log(1/\varepsilon)}{\alpha \log(M)} \right\rceil,$$

and a bias parameter $h^* = h^*(\varepsilon, R(\varepsilon)) \in (0, \mathbf{h}]$ given by (5).

1 Introduction

- Abstract framework and assumptions
- Monte Carlo type estimators
- Crude Monte Carlo estimator in a biased framework
- Multistep Richardson-Romberg estimator

2 Multilevel estimators

- Definition
- Majoration of the effort
- specification(s) of the allocation matrix \mathbf{T}
- Convergence for a fixed $R \geq 2$
- Main result

3 Applications and numerical experiments

- Parameters and methodology
- Discretization of Brownian diffusions
- Nested Monte Carlo

Optimal parameters for MLRR (with $n_i = M^{i-1}$)

R	$\left\lfloor \frac{1}{2} + \frac{\log(\tilde{c}^{\frac{1}{\alpha}} \mathbf{h})}{\log(M)} + \sqrt{\left(\frac{1}{2} + \frac{\log(\tilde{c}^{\frac{1}{\alpha}} \mathbf{h})}{\log(M)} \right)^2 + 2 \frac{\log(1/\varepsilon)}{\alpha \log(M)}} \right\rfloor$
h^{-1}	$\left\lceil (1 + 2\alpha R)^{\frac{1}{2\alpha R}} \varepsilon^{-\frac{1}{\alpha R}} M^{-\frac{R-1}{2}} \right\rceil$
q	$q_1 = \mu^*(1 + \theta h^{\frac{\beta}{2}})$ $q_j = \mu^* \theta h^{\frac{\beta}{2}} \left(\mathbf{W}_j(R, M) \frac{n_{j-1}^{-\frac{\beta}{2}} + n_j^{-\frac{\beta}{2}}}{\sqrt{n_{j-1} + n_j}} \right),$
N	$\frac{\text{var}(Y_0) \left(1 + \theta h^{\frac{\beta}{2}} \sum_{j=1}^R \mathbf{W}_j(R, M) \left(n_{j-1}^{-\frac{\beta}{2}} + n_j^{-\frac{\beta}{2}} \right) \sqrt{n_{j-1} + n_j} \right)^2}{\varepsilon^2 \left(1 + \frac{1}{2\alpha R} \right)^{-1} \sum_{j=1}^R q_j (n_{j-1} + n_j)}$

Optimal parameters for MLMC (with $n_i = M^{i-1}$)

R	$\left\lceil 1 + \frac{\log(c_1/\varepsilon)}{\alpha \log M} \right\rceil$
h^{-1}	$\left\lceil (1 + 2\alpha)^{\frac{1}{2\alpha}} \left(\frac{\varepsilon}{c_1} \right)^{-\frac{1}{\alpha}} M^{-(R-1)} \right\rceil$
q	$q_1 = \mu^*(1 + \theta h^{\frac{\beta}{2}})$ $q_j = \mu^* \theta h^{\frac{\beta}{2}} \left(\frac{n_{j-1}^{-\frac{\beta}{2}} + n_j^{-\frac{\beta}{2}}}{\sqrt{n_{j-1} + n_j}} \right),$
N	$\frac{\text{var}(Y_0) \left(1 + \theta h^{\frac{\beta}{2}} \sum_{j=1}^R \left(n_{j-1}^{-\frac{\beta}{2}} + n_j^{-\frac{\beta}{2}} \right) \sqrt{n_{j-1} + n_j} \right)^2}{\varepsilon^2 \left(1 + \frac{1}{2\alpha} \right)^{-1} \sum_{j=1}^R q_j (n_{j-1} + n_j)}$

Choice of the **root** $M = M(\varepsilon)$

- Refiners: $n_i = M^{i-1}$, $i = 1, \dots, R$, M integer, $M \geq 2$.
- Optimal parameters are all functions of $M \Rightarrow$ Numerical optimization of the cost function

$$\mathbf{C}_\varepsilon(M) = \text{Cost}(Y_{h(M), \underline{n}}^{N(M), q(M)}) := N(M) \kappa(h(M), R(M), q(M)),$$

for $M \in \{2, \dots, 10\}$, where

$$\kappa(h, R, q) = \frac{1}{h} \sum_{j=1}^R q_j (n_{j-1} + n_j)$$

(with $n_0 = 0$) or (for **nested simulations**)

$$\kappa(h, R, q) = \frac{1}{h} \sum_{j=1}^R q_j n_j.$$

Black-Scholes *Call* by Euler scheme ($\alpha = \beta = 1$)

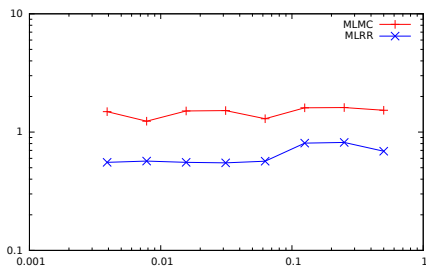
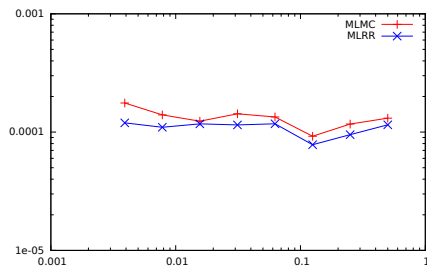
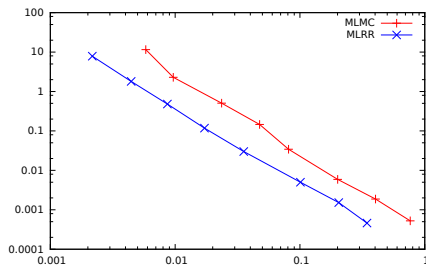
- Model for X : (geometric B.M.) $x_0 = 100$, $r = 0.05$, $\sigma = 0.2$.
- Payoff $\phi = (x(T) - K)_+$ (Call), strike $K = 100$, maturity $T = 5$ years.
- $Y_h = \phi(\bar{X}_T^h)$ with \bar{X}^h **stepwise constant** Euler scheme with step $\frac{T}{n}$.
- Structural parameters
 - ▶ $\text{var} = 704$
 - ▶ $\theta \approx 0.43$ (see estimation $\hat{V}_1(h)$ of V_1 below)

▷ L^2 -error using 256 runs to estimate V_1 by

$$\hat{V}_1(h) = (1 + M_{\max}^{-\frac{\beta}{2}})^{-2} h^{-\beta} \|Y_h - Y_{\frac{h}{M_{\max}}}\|_2^2.$$

with $h \in (0, \mathbf{h}]$, $M_{\max} \geq 10$.

▷ Bias computed using the true price 29.1386.

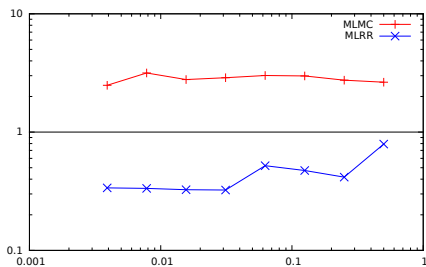
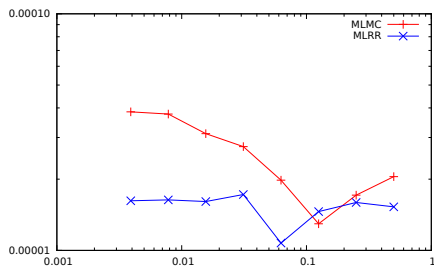
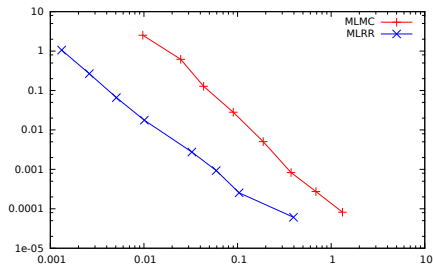

 $\tilde{\varepsilon}_L/\varepsilon$ as a function of ε

CPU-time $\times \varepsilon^2$ as a function of ε

CPU-time as a function of $\tilde{\varepsilon}_L$

Parameters and complexity for a given ε

k	$\varepsilon = 2^{-k}$	<i>MLRR</i>				<i>MLMC</i>			
		R	M	h^{-1}	Cost	R	M	h^{-1}	Cost
1	$5.00 \cdot 10^{-01}$	2	5	1	$4.00 \cdot 10^{+04}$	2	4	1	$3.56 \cdot 10^{+04}$
2	$2.50 \cdot 10^{-01}$	2	9	1	$1.73 \cdot 10^{+05}$	2	7	1	$1.67 \cdot 10^{+05}$
3	$1.25 \cdot 10^{-01}$	2	8	2	$8.47 \cdot 10^{+05}$	2	8	2	$8.88 \cdot 10^{+05}$
4	$6.25 \cdot 10^{-02}$	3	4	1	$5.16 \cdot 10^{+06}$	3	6	1	$5.95 \cdot 10^{+06}$
5	$3.12 \cdot 10^{-02}$	3	5	1	$2.09 \cdot 10^{+07}$	3	8	1	$2.69 \cdot 10^{+07}$
6	$1.56 \cdot 10^{-02}$	3	6	1	$8.63 \cdot 10^{+07}$	3	8	2	$1.26 \cdot 10^{+08}$
7	$7.81 \cdot 10^{-03}$	3	7	1	$3.59 \cdot 10^{+08}$	4	7	1	$7.41 \cdot 10^{+08}$
8	$3.91 \cdot 10^{-03}$	3	9	1	$1.55 \cdot 10^{+09}$	4	8	1	$3.16 \cdot 10^{+09}$

B-S Lookback by Euler scheme ($\alpha = 0.5, \beta = 1$)

- Model for X : (geometric B.M.) $x_0 = 100$, $r = 0.15$, $\sigma = 0.1$.
- Lookback Payoff
$$\phi = e^{-rT} (x(T) - \lambda \min_{t \in [0, T]} x(t))_+, \quad x \in \mathcal{C}([0, T], \mathbf{R}),$$
- $\lambda = 1.1$, maturity $T = 1$ year.
- $Y_h = \phi(\bar{X}_T^h)$ with \bar{X}^h **stepwise constant** Euler scheme with step $\frac{T}{n}$.
- True price $I_0 = 8.89343$ (used to compute the bias).

 $\tilde{\varepsilon}_L / \varepsilon$ as a function of ε  $\text{CPU-time} \times \varepsilon^2$ as a function of ε CPU-time in second of one estimator as a function of $\tilde{\varepsilon}_L$

Parameters and complexity for a given ε

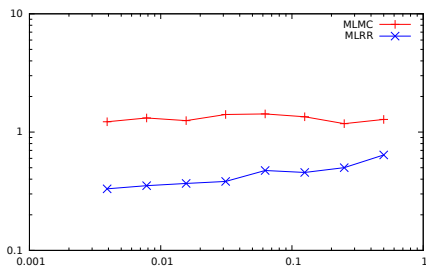
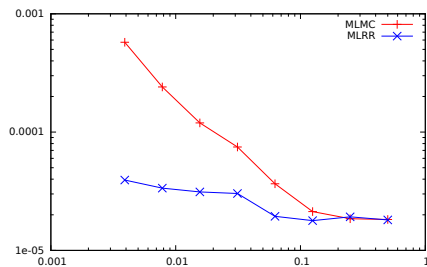
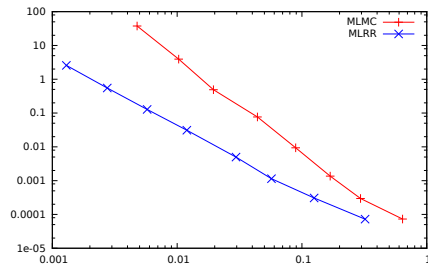
k	$\varepsilon = 2^{-k}$	<i>MLRR</i>				<i>MLMC</i>			
		R	M	h^{-1}	Cost	R	M	h^{-1}	Cost
1	$5.00 \cdot 10^{-01}$	2	7	2	$3.13 \cdot 10^{+03}$	2	8	1	$2.05 \cdot 10^{+03}$
2	$2.50 \cdot 10^{-01}$	3	6	1	$1.76 \cdot 10^{+04}$	3	6	1	$1.61 \cdot 10^{+04}$
3	$1.25 \cdot 10^{-01}$	3	7	1	$7.07 \cdot 10^{+04}$	3	8	2	$8.88 \cdot 10^{+04}$
4	$6.25 \cdot 10^{-02}$	3	7	2	$3.37 \cdot 10^{+05}$	4	8	1	$5.04 \cdot 10^{+05}$
5	$3.12 \cdot 10^{-02}$	4	5	1	$1.68 \cdot 10^{+06}$	5	7	1	$2.93 \cdot 10^{+06}$
6	$1.56 \cdot 10^{-02}$	4	6	1	$6.74 \cdot 10^{+06}$	5	10	1	$1.40 \cdot 10^{+07}$
7	$7.81 \cdot 10^{-03}$	4	7	1	$2.74 \cdot 10^{+07}$	6	8	1	$7.17 \cdot 10^{+07}$
8	$3.91 \cdot 10^{-03}$	4	9	1	$1.16 \cdot 10^{+08}$	6	10	2	$3.59 \cdot 10^{+08}$

B-S barrier option by Euler scheme ($\alpha = \beta = 0.5$)

- Model for X : (geometric B.M.) $x_0 = 100$, $r = 0.0$, $\sigma = 0.15$.
- Up & Out Barrier Call Payoff

$$\phi = e^{-rT}(x(T) - K)_+ \mathbf{1}_{\left\{\max_{t \in [0, T]} x(t) \leq L\right\}}, \quad x \in \mathcal{C}([0, T], \mathbf{R}),$$

- strike $K = 100$, barrier $L = 130$ and maturity $T = 1$ year.
- $Y_h = \phi(\bar{X}_T^h)$ with \bar{X}^h **stepwise constant** Euler scheme with step $\frac{T}{n}$.
- True price $I_0 = 3.869693$ (used to compute the bias).

 $\tilde{\varepsilon}_L / \varepsilon$ as a function of ε CPU-time $\times \varepsilon^2$ as a function of ε CPU-time in second of one estimator as a function of $\tilde{\varepsilon}_L$

Parameters and complexity for a given ε

k	$\varepsilon = 2^{-k}$	<i>MLRR</i>				<i>MLMC</i>			
		R	M	h^{-1}	Cost	R	M	h^{-1}	Cost
1	$5.00 \cdot 10^{-01}$	2	6	2	$7.74 \cdot 10^{+03}$	2	8	1	$3.97 \cdot 10^{+03}$
2	$2.50 \cdot 10^{-01}$	3	4	1	$6.06 \cdot 10^{+04}$	3	6	1	$4.73 \cdot 10^{+04}$
3	$1.25 \cdot 10^{-01}$	3	7	1	$2.69 \cdot 10^{+05}$	3	8	2	$3.54 \cdot 10^{+05}$
4	$6.25 \cdot 10^{-02}$	3	7	2	$1.60 \cdot 10^{+06}$	4	8	1	$3.06 \cdot 10^{+06}$
5	$3.12 \cdot 10^{-02}$	4	5	1	$1.01 \cdot 10^{+07}$	5	7	1	$3.04 \cdot 10^{+07}$
6	$1.56 \cdot 10^{-02}$	4	6	1	$4.50 \cdot 10^{+07}$	5	8	2	$2.17 \cdot 10^{+08}$
7	$7.81 \cdot 10^{-03}$	4	7	1	$2.01 \cdot 10^{+08}$	6	8	1	$1.78 \cdot 10^{+09}$
8	$3.91 \cdot 10^{-03}$	4	9	1	$9.88 \cdot 10^{+08}$	6	10	2	$1.63 \cdot 10^{+10}$

Nested Monte Carlo

- Aim: Computing

$$\mathbf{E} [f(\mathbf{E} [X | Y])]$$

where (X, Y) is couple of $\mathbf{R} \times \mathbf{R}^{q_Y}$ -valued random variable.

- Assume that X has the representation

$$X = F(Z, Y).$$

where : $F : \mathbf{R}^{q_Z} \times \mathbf{R}^{q_Y} \rightarrow \mathbf{R}$ is a Borel function

$Z : (\Omega, \mathcal{A}) \rightarrow \mathbf{R}^{q_Z}$ is independent of Y .

To comply with the multilevel framework, we set $\mathcal{D} = \{1/K, K \geq 1\}$

$$Y_0 = f(\mathbf{E} [X | Y]), \quad Y_{\frac{1}{K}} = f \left(\frac{1}{K} \sum_{k=1}^K F(Z^k, Y) \right).$$

Nested Monte Carlo

Theorem

If f is Lipschitz continuous and $2R$ times differentiable with $f^{(k)}$ bounded, $k = R, \dots, 2R$, the nested Monte Carlo satisfies

- (SE_β) with $\beta = 1$
- $(WE_{\alpha,R})$ with $\alpha = 1$.

Non smooth case:

- Gordy & Juneja, *Nested simulation in portfolio risk measurement*, 2010
- Broadie et al. *Efficient risk estimation via nested sequential simulation*, 2011
- Work in progress...

Put-on-call option

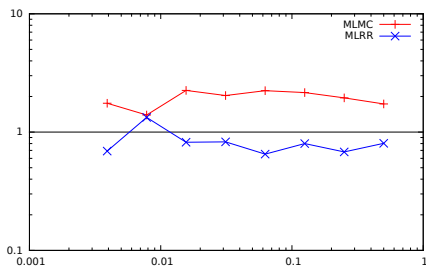
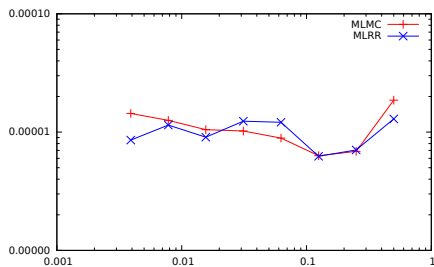
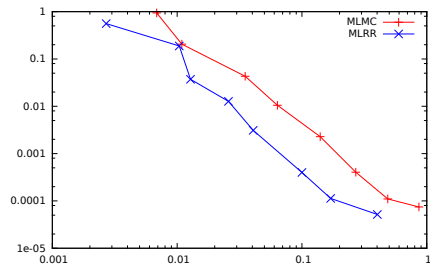
- Payoff of a **Put-on-Call** option (involves conditional expectation)

$$(K_1 - \mathbf{E}[(S_{T_2} - K_2)_+ | S_{T_1}])_+$$

with parameters $T_1 = 1/12$, $T_2 = 1/2$ and $K_1 = 6.5$, $K_2 = 100$.

- Black-Scholes model with $s_0 = 100$, $r = 0.03$ and $\sigma = 0.3$.
Note that

$$S_{T_2} = F(G, S_{T_1}) = S_{T_1} e^{(r - \frac{\sigma^2}{2})(T_2 - T_1) + \sigma \sqrt{T_2 - T_1} G}$$

 $\tilde{\varepsilon}_L / \varepsilon$ as a function of ε CPU-time $\times \varepsilon^2$ as a function of ε CPU-time in second of one estimator as a function of $\tilde{\varepsilon}_L$

Parameters and complexity for a given ε

k	$\varepsilon = 2^{-k}$	<i>MLRR</i>				<i>MLMC</i>			
		R	M	h^{-1}	Cost	R	M	h^{-1}	Cost
1	$5.00 \cdot 10^{-01}$	2	5	1	$1.37 \cdot 10^{+03}$	2	2	2	$1.08 \cdot 10^{+03}$
2	$2.50 \cdot 10^{-01}$	2	4	2	$6.24 \cdot 10^{+03}$	2	4	2	$5.39 \cdot 10^{+03}$
3	$1.25 \cdot 10^{-01}$	2	8	2	$2.79 \cdot 10^{+04}$	2	8	2	$2.84 \cdot 10^{+04}$
4	$6.25 \cdot 10^{-02}$	3	4	1	$1.87 \cdot 10^{+05}$	3	4	2	$2.04 \cdot 10^{+05}$
5	$3.12 \cdot 10^{-02}$	3	5	1	$7.81 \cdot 10^{+05}$	3	6	2	$9.94 \cdot 10^{+05}$
6	$1.56 \cdot 10^{-02}$	3	4	2	$3.29 \cdot 10^{+06}$	3	8	2	$4.64 \cdot 10^{+06}$
7	$7.81 \cdot 10^{-03}$	3	7	1	$1.40 \cdot 10^{+07}$	4	6	2	$2.96 \cdot 10^{+07}$
8	$3.91 \cdot 10^{-03}$	3	7	2	$6.13 \cdot 10^{+07}$	4	7	2	$1.29 \cdot 10^{+08}$