Multi-step Richardson-Romberg Extrapolation: Remarks on Variance Control and Complexity

Gilles Pagès [*]

January 17, 2007

Abstract

We propose a multi-step Richardson-Romberg extrapolation method for the computation of expectations $\mathbb{E}f(X_T)$ of a diffusion $(X_t)_{t\in[0,T]}$ when the weak time discretization error induced by the Euler scheme admits an expansion at an order $R \geq 2$. The complexity of the estimator grows as R^2 (instead of 2^R in the classical method) and its variance is asymptotically controlled by considering some consistent Brownian increments in the underlying Euler schemes. Some Monte Carlo simulations were carried with path-dependent options (lookback, barrier) which support the conjecture that their weak time discretization error also admits an expansion (in a different scale). Then an appropriate Richardson-Romberg extrapolation seems to outperform the Euler scheme with Brownian bridge.

Key words: SDE, Euler-Maruyama scheme, Romberg extrapolation, Vandermonde determinant, lookback option, barrier option.

MSC Classification (2000): 65C05, 60H35, 65B99, 65C30.

1 Introduction and preliminaries

One considers a d-dimensional Brownian diffusion process $(X_t)_{t\in[0,T]}$ solution of the following S.D.E.

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \qquad X_0 = x. \tag{1.1}$$

where $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0,T] \times \mathbb{R}^d \to \mathcal{M}(d \times q)$ are continuous functions and $(W_t)_{t\in[0,T]}$ denotes a q-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. We assume that b and σ are Lipschitz continuous in x uniformly with respect to $t \in [0,T]$ and that $b(.,0)$ and $\sigma(.,0)$ are bounded over $[0,T]$, see [12]. In fact what we will simply need is that, for every starting value $x \in \mathbb{R}^d$, the Brownian Euler scheme of (1.1) – $i.e.$ the Euler scheme based on the increments of W – converges to X in every $L^p(\mathbb{P})$, $p \in (0, +\infty)$, where X is the unique strong solution of the SDE starting at x.

[*]Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris 6, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail:`gpa@ccr.jussieu.fr`

Except for some very specific equations, it is impossible to process an exact simulation of the process X even at a fixed time T (by exact simulation, we mean writing $X_T = \chi(U)$, $U \sim U([0,1])$) (nevertheless, when $d = 1$ and $\sigma \equiv 1$, see [4], [5]). Consequently, to approximate $\mathbb{E}(f(X_T))$ by a Monte Carlo method, one needs to approximate X by a process that can be simulated (at least at a fixed number of instants). To this end one introduces the stepwise constant Brownian Euler scheme $\bar{X} = (\bar{X}_{\frac{kT}{n}})_{0 \le k \le n}$ with step $\frac{T}{n}$ associated to the SDE. It is defined by

$$\bar{X}_{t_{k+1}^n} = \bar{X}_{t_k^n} + b(t_k^n, \bar{X}_{t_k^n})\frac{T}{n} + \sigma(t_k^n, \bar{X}_{t_k^n})\sqrt{\frac{T}{n}} U_{k+1}, \qquad \bar{X}_0 = x, \quad k = 0, \ldots, n-1,$$

where $t_k^n = \frac{kT}{n}$, $k = 0, \ldots, n-1$ and $(U_k)_{1 \le k \le n}$ denotes a sequence of i.i.d. $\mathcal{N}(0;1)$-distributed random vectors given by

$$U_k := \sqrt{\frac{n}{T}} (W_{t_k^n} - W_{t_{k-1}^n}), \quad k = 1, \ldots, n.$$

Moreover, set for convenience $\bar{X}_t := \bar{X}_{\underline{t}}$ where $\underline{t} = t_k^n$ if $t \in [t_k^n, t_{k+1}^n)$.

Then, it is classical background that under the regularity and growth assumptions on the coefficients b and σ mentioned above, $\sup_{t \in [0,T]} |X_t - \bar{X}_t|$ goes to zero in every $L^p(\mathbb{P})$, $0 < p < \infty$, at a $O(\frac{1}{\sqrt{n}})$-rate.

However many authors in a long series of papers going back to the seminal papers by Talay-Tubaro ([16]) and Bally-Talay ([2], [3]), showed under various assumptions on the diffusion coefficients the existence of a vector space V of Borel functions $f : \mathbb{R}^d \to \mathbb{R}$ (bounded or with polynomial growth) for which one can expand the "weak" time discretization error induced by the Euler scheme into a power series of $\frac{1}{n}$. To be precise

$$(\mathcal{E}_R^V) \qquad \equiv \qquad \forall f \in V, \qquad \mathbb{E}(f(X_T)) = \mathbb{E}(f(\bar{X}_T)) + \sum_{k=1}^{R-1} \frac{c_k}{n^k} + O(n^{-R}) \tag{1.2}$$

where the real constants c_k, $k = 1, \ldots, R-1$, do not depend on the discretization parameter n (see [16] for smooth functions f, see [2] for bounded Borel functions and uniformly hypo-elliptic diffusions ([1])). In [3] is established the convergence of the p.d.f. of the Euler scheme at time T toward that of X_T. See also the recent work [11] for an extension to tempered distributions. Usually, this expansion is established in full details for $R = 2$ and is known as the (standard) Richardson-Romberg extrapolation. However, up to additional technicalities (and smoothness assumptions on the coefficients) the expansion holds true for larger values of R or even for every integer $R \ge 2$.

Furthermore, note that, the resulting vector space V of "admissible" functions is always characterizing for the $\mathbb{P}$-a.s. equality in the following sense: for every pair of $\mathbb{R}^d$-valued random vectors X, Y defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$,

$$(\forall f \in V, \ f(X) = f(Y) \quad \mathbb{P}\text{-a.s.}) \Longrightarrow (X = Y \quad \mathbb{P}\text{-a.s.})$$

---

[1]In [2], the diffusion coefficients are assumed to be $\mathcal{C}_b^\infty$ i.e. infinitely differentiable, bounded, with bounded derivatives; then $(\mathcal{E}_R^V)$ holds for any R if σ is uniformly elliptic. As concerns $R = 2$, the original method of proof (see [16]) based on an approximation by the parabolic PDE involving the infinitesimal generator works with finitely differentiable coefficients b, σ and function f. Although not proved in full details, an extension to Borel functions *with polynomial growth* is mentioned in [2]. See also [10], [13] for different approaches.

From now on, we always assume that V has this property.

Sharp rates of weak convergence (*i.e.* an expansion with $R = 2$) have been established for more general functionals $F((X_t)_{t \in [0,T]})$ depending on the whole trajectory of the diffusion. One uses the standard (stepwise constant) or the continuous Euler scheme to approximate $\mathbb{E}F(X)$. These contributions are usually motivated by the pricing of some classes of path-dependent European options like Asian options ([14]), lookback options (see [15]) or barrier options (see [9]). These sharp rates are c/n or $c/\sqrt{n}$ and one can reasonably conjecture that these rates can be extended into some expansions in an appropriate polynomial scale. The R-R extrapolation can be in such situations a way to support some conjectures concerning these expansions (and an heuristic way to speed up the Monte Carlo simulations).

In Section 2, we first briefly recall what the original Richardson-Romberg (R-R) extrapolation is (when $R = 2$). Then, we prove that the choice of consistent Brownian increments in the two involved Euler schemes is optimal. In Section 3, we propose a multi-step R-R extrapolation at order R and we show that choosing consistent Brownian increments still preserves the variance of the estimator for continuous functions or even Borel functions (under some uniform ellipticity assumption). We point out that, however, consistent increments is however not an optimal choice in general for a given function f. Then, in Section 4.4 we analyze the complexity of our approach and compare it to that of iterated R-R extrapolations. In Section 5, we provide for the first values of interest ($R = 2, 3, 4, 5$) some tables to simulate the needed consistent Brownian increments at this order. In Section 5.2, we illustrate the efficiency of this consistent R-R extrapolation with $R = 3, 4$ by pricing vanilla options in a high volatility B-S model. In Section 5.3, we make some further tests on partial lookback and Up & Out barrier options which support the existence of expansions of the error.

NOTATIONS: A^* is for transpose of the matrix A. $|u|$ denotes the canonical Euclidean norm on $\mathbb{R}^q$, $q \geq 1$.

2 The standard Richardson-Romberg extrapolation

Assume that $(\mathcal{E}_2^V)$ holds. Let $f \in V$ where V denotes a vector space of continuous functions with linear growth. The case of non continuous functions is investigated in the next section. For notational convenience we set $W^{(1)} = W$ and $X^{(1)} := X$. A regular Monte Carlo simulation based on M independent copies $(\bar{X}_T^{(1)})^m$, $m = 1, \ldots, M$, of the Euler scheme $\bar{X}_T^{(1)}$ with step T/n induces the following global (squared) quadratic error

$$
\begin{aligned}
\|\mathbb{E}(f(X_T)) - \frac{1}{M}\sum_{m=1}^{M} f((\bar{X}_T^{(1)})^m)\|_2^2 &= |\mathbb{E}(f(X_T)) - \mathbb{E}(f(\bar{X}_T^{(1)}))|^2 \\
&\quad + \|\mathbb{E}(f(\bar{X}_T^{(1)})) - \frac{1}{M}\sum_{m=1}^{M} f((\bar{X}_T^{(1)})^m)\|_2^2 \\
&= \frac{c_1^2}{n^2} + \frac{\text{Var}(f(\bar{X}_T^{(1)}))}{M} + O(n^{-3}).
\end{aligned} \tag{2.1}
$$

This quadratic error bound (2.1) does not take fully advantage of the above expansion $(\mathcal{E}_2)$. To take advantage of the expansion, one needs to make an R-R extrapolation. In that framework (originally introduced in [16]) one considers a second Brownian Euler scheme, this time of the solution $X^{(2)}$ of a "copy" of Equation (1.1) written with respect to a second Brownian motion $W^{(2)}$ defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. In fact, one may chose this Brownian motion by enlarging Ω if necessary. This second Euler scheme has *a twice smaller* step $\frac{T}{2n}$ and is denoted $\bar{X}^{(2)}$. Then, assuming $(\mathcal{E}_2^V)$to be more precise,

$$\mathbb{E}(f(X_T)) = \mathbb{E}(2f(\bar{X}_T^{(2)}) - f(\bar{X}_T^{(1)})) - \frac{1}{2}\frac{c_2}{n^2} + O(n^{-3})$$

Then, the new global (squared) quadratic error becomes

$$\|\mathbb{E}(f(X_T)) - \frac{1}{M}\sum_{m=1}^{M} 2f((\bar{X}_T^{(2)})^m) - f((\bar{X}_T)^m)\|_2^2 = \frac{c_2^2}{4n^4} + \frac{\mathrm{Var}(2f(\bar{X}_T^{(2)}) - f(\bar{X}_T^{(1)}))}{M} + O(n^{-5}). \tag{2.2}$$

The structure of this quadratic error suggests the following question: is it possible to reduce the (asymptotic) time discretization error *without increasing the Monte Carlo error?* *To what extend is it possible to control the variance term* $\mathrm{Var}(2f(\bar{X}_T^{(2)}) - f(\bar{X}_T^{(1)}))$?

It is clear that $\sup_{t\in[0,T]} |\bar{X}_t^{(i)} - X_t^{(i)}|$ converges to 0 in every $L^p(\mathbb{P})$, $0 < p < \infty$, $i = 1, 2$. Consequently

$$\sup_{t\in[0,T]} |(\bar{X}_t^{(1)}, \bar{X}_t^{(2)}) - (X_t^{(1)}, X_t^{(2)})| \overset{L^p(\mathbb{P}) \, \& \, a.s.}{\longrightarrow} 0 \quad \text{as} \quad n \to \infty.$$

In particular (keep in mind f has at most polynomial growth),

$$\mathrm{Var}(2f(\bar{X}_T^{(2)}) - f(\bar{X}_T^{(1)})) \longrightarrow \mathrm{Var}(2f(X_T^{(2)}) - f(X_T^{(1)})) \quad \text{as} \quad n \to \infty.$$

Then, straightforward computations show that

$$\mathrm{Var}\left(2f(X_T^{(2)}) - f(X_T^{(1)})\right) = 5\,\mathbb{E}(f(X_T^{(1)}))^2 - 4\,\mathbb{E}f(X_T^{(1)})f(X_T^{(2)}) - \left(\mathbb{E}(f(X_T^{(1)}))\right)^2 \tag{2.3}$$

where

$$\mathbb{E}\left(f(X_T^{(1)})f(X_T^{(2)})\right) \leq \|f(X_T^{(1)})\|_2\|f(X_T^{(2)})\|_2 = \|f(X_T^{(1)})\|_2^2 \tag{2.4}$$

by Schwarz's Inequality since $X_T^{(1)} \overset{d}{=} X_T^{(2)}$.

Consequently, minimizing the variance term amounts to maximizing $\mathbb{E}(f(X_T^{(1)})f(X_T^{(2)}))$. It follows from the equality case in Schwarz's Inequality that the equality in the left hand side of (2.4) holds iff $f(X_T^{(2)}) = \lambda f(X_T^{(1)})$ $\mathbb{P}$-*a.s.* for some $\lambda \in \mathbb{R}_+$ or $f(X_T^{(1)}) = 0$ $\mathbb{P}$-*a.s..* But since $X_T^{(1)}$ and $X_T^{(2)}$ have the same distribution, this implies $f(X_T^{(2)}) = f(X_T^{(1)})$ $\mathbb{P}$-*a.s..*

Finally minimizing the asymptotic variance term for every $f \in V$ is possible if and only if

$$X_T^{(2)} = X_T^{(1)} \quad \mathbb{P}\text{-}a.s. \tag{2.5}$$

4

A sufficient (and sometimes necessary, see Annex) condition to ensure (2.5) clearly is that

$$W^{(2)} = W^{(1)}.$$

This means that the white noise $(U_k^{(2)})_{1 \leq k \leq n}$ of the Euler scheme $\bar{X}^{(2)}$ satisfies

$$U_k^{(2)} = \sqrt{\frac{2n}{T}}(W_{\frac{kT}{2n}}^{(1)} - W_{\frac{(k-1)T}{2n}}^{(1)}), \quad k = 1, \ldots, 2n,$$

so that, from a simulation point of view, one needs to simulate $2n$ i.i.d. copies $U_k^{(2)}$ of the normal distribution and then sets

$$U_k^{(1)} = \frac{U_{2k}^{(2)} + U_{2k-1}^{(2)}}{\sqrt{2}}, \ k = 1, \ldots, n.$$

Note that all what precedes (as well as all what follows in fact) can be extended to continuous functionals $F(X)$ for which an expansion similar to (1.2) holds (see *e.g.* the pricing of Asian options in [14] or Section 5) or for different dynamics like SDE with delay studied in [6].

This result provides an (at least asymptotic) positive answer to a question raised in [7] (see p. 361): Self-consistent Brownian increments do minimize the (asymptotic) variance in the *R-R* extrapolation. In the next section we show that it is possible to design some multi-step *R-R* extrapolations with a reasonable complexity for which the control of the variance is still preserved.

As concerns variance control, note that if one proceeds using two *independent* sequences of Gaussian white noises $U^{(1)}$ and $U^{(2)}$, the expansion (2.3) yields

$$\mathrm{Var}\left(2f(X_T^{(2)}) - f(X_T^{(1)})\right) = 5\,\mathrm{Var}\left(f(X_T^{(1)})\right)$$

Such a choice is then the worst possible one. It induces an increase of the Monte Carlo simulation by a factor 5.

3   Multi-step Richardson-Romberg extrapolation

Iterated *R-R* extrapolations are usually not implemented, essentially because their numerical efficiency (when the white noises are independent) becomes less and less obvious. The first reason is the increase of the complexity, the second one is the "explosion" of the variance term (especially when implemented with independent Brownian motions) and the third one is the absence of control of the coefficients c_k as k increases. In what follows we propose a solution to the first two problems: we propose a multi-step *R-R* extrapolation with consistent Brownian increments. Doing so, we will control the variance in the Monte Carlo error term and limit the increase of the complexity of the procedure. One proceeds as follows: let $R \geq 2$ be an integer. We wish to obtain a time discretization error behaving like n^{-R} as $n \to \infty$. To this end, we introduce R Gaussian Euler schemes $\bar{X}^{(r)} = (\bar{X}_{\frac{kT}{rn}}^{(r)})_{0 \leq k \leq r\,n}$

with step $\frac{T}{rn}$, $r = 1, \ldots, R$. Each Euler scheme is designed using a Gaussian white noise $(U_k^{(r)})_{1 \le k \le rn}$ obtained from the increments of a standard Brownian motion $W^{(r)}$ by setting

$$U_k^{(r)} := \sqrt{\frac{rn}{T}} \left(W_{\frac{kT}{rn}}^{(r)} - W_{\frac{(k-1)T}{rn}}^{(r)} \right), \quad k = 1, \ldots, rn.$$

The Brownian motions $W^{(1)}, \ldots, W^{(R)}$ are all defined on the same probability space. We will come back further on about the practical simulation of these increments. Our "meta-assumption" $(\mathcal{E}_{R+1}^V)$ implies that for a function $f \in V$, one has

$$\mathbb{E}(f(X_T)) = \mathbb{E}(f(\bar{X}_T^{(r)})) + \sum_{\ell=1}^{R-1} \frac{c_\ell}{r^\ell} \frac{1}{n^\ell} + \frac{c_R}{r^R} \frac{1}{n^R} (1 + O(1/(rn))), \qquad r = 1, \ldots, R.$$

One defines the $R \times (R-1)$ matrix

$$A = \left[\frac{1}{r^\ell} \right]_{1 \le r \le R, 1 \le \ell \le R-1}$$

and $\mathbf{1}$ as the unit (column) vector of $\mathbb{R}^R$. Then, the above system of equations reads

$$\mathbb{E}(f(X_T)) \, \mathbf{1} = \left[\begin{array}{c} \vdots \\ \mathbb{E}(f(\bar{X}_T^{(r)})) \\ \vdots \end{array} \right]_{1 \le r \le R} + A \left[\begin{array}{c} \vdots \\ \frac{c_r}{n^r} \\ \vdots \end{array} \right]_{1 \le r \le R-1} + c_R \left[\begin{array}{c} \vdots \\ \frac{1}{n^R r^R}(1 + O(\frac{1}{rn})) \\ \vdots \end{array} \right]_{1 \le r \le R}. \quad (3.1)$$

Now, let $\alpha \in \mathbb{R}^R$ satisfying

$$\alpha^* \mathbf{1} = 1 \qquad \text{and} \qquad \alpha^* A = 0.$$

This reads

$$\alpha^* \widetilde{A} = e_1 \qquad (3.2)$$

where $e_1 = (1, 0, \ldots, 0)^*$ is the first (column) vector of the canonical basis of $\mathbb{R}^R$ and

$$\widetilde{A} = \left[\frac{1}{r^{\ell-1}} \right]_{1 \le r \le R, 1 \le \ell \le R} = \text{Vandermonde}\left(\frac{1}{r}, r = 1, \ldots, R \right).$$

Multiplying (3.1) on the left by α^* yields

$$\mathbb{E}(f(X_T)) = \mathbb{E}\left(\sum_{r=1}^R \alpha_r f(\bar{X}_T^{(r)}) \right) + \frac{\widetilde{c}_R}{n^R} (1 + O(1/n)) \qquad (3.3)$$

where

$$\widetilde{c}_R = c_R \sum_{r=1}^R \frac{\alpha_r}{r^R}.$$

Solving the Cramer linear system (3.2) yields after some elementary computations based on the Vandermonde determinant the following lemma.

Lemma 3.1 *For every integer $R \ge 1$,*

$$\alpha_r = \alpha(r, R) := (-1)^{R-r} \frac{r^R}{r!(R-r)!}, \quad 1 \le r \le R. \qquad (3.4)$$

6

Remarks. • As a consequence of (3.4), one shows that for every integer $R \geq 1$,

$$\sum_{r=1}^{R} \frac{\alpha_r}{r^R} = \sum_{r=0}^{R-1} \frac{(-1)^r}{r!} \frac{1}{(R-r)!} = \frac{(-1)^{R-1}}{R!}$$

(this follows from the comparison of the coefficient of x^R in the expansion $e^x e^{-x} = 1$) so that

$$|\widetilde{c}_R| = \frac{|c_R|}{R!}.$$

This bound emphasizes that this multi-step extrapolation *does not* "hide" an "exploding" behaviour in the first non vanishing order of the expansion. In fact it has a damping effect on this coefficient c_R. However, the fact that we have in practice no control on the coefficient c_R (and its successors) at a reasonable cost remains and is not overcome by this approach.

• If $\mathcal{E}_R^V$ holds, then the above computations show that (3.3) becomes

$$\mathbb{E}(f(X_T)) = \mathbb{E}\left(\sum_{r=1}^{R} \alpha_r f(\bar{X}_T^{(r)}) \right) + O(n^{-R}).$$

Now we pass to the choice of the Brownian motions $W^{(r)}$ in order to preserve the variance of such a combination.

3.1 The case of continuous functions with polynomial growth

In this section, we assume that the vector space V is made of continuous functions with polynomial growth and is stable by product. First note that for every $p > 0$

$$\sup_{t \in [0,T]} |(\bar{X}_t^{(1)}, \ldots, \bar{X}_t^{(r)}, \ldots, \bar{X}_t^{(R)}) - (X_t^{(1)}, \ldots, X_t^{(r)}, \ldots, X_t^{(R)})| \overset{L^p(\mathbb{P}) \,\&\, a.s.}{\longrightarrow} 0. \qquad (3.5)$$

Let $f \in V$. Then

$$\mathrm{Var}\left(\sum_{r=1}^{R} \alpha_r f(\bar{X}_T^{(r)}) \right) \longrightarrow \mathrm{Var}\left(\sum_{r=1}^{R} \alpha_r f(X_T^{(r)}) \right) \qquad \text{as} \quad n \to \infty \qquad (3.6)$$

since f^2 is continuous with polynomial growth. When

$$W^{(r)} = W, \quad r = 1, \ldots, R,$$

then, $f(X_T^{(r)}) = f(X_T)$ for every $r \in \{1, \ldots, R\}$ so that

$$\mathrm{Var}\left(\sum_{r=1}^{R} \alpha_r f(X_T^{(r)}) \right) = \left(\sum_{r=1}^{R} \alpha_r \right)^2 \mathrm{Var}(f(X_T)) = \mathrm{Var}(f(X_T)).$$

This obvious remark shows that this choice for the Brownian motions $W^{(r)}$ lead to a control of the variance of the multi-step R-R estimator by that of $f(X_T^{(1)})$. However, the optimality of this choice turns out to be a less straightforward question.

7

On the opposite, if one considers some mutually independent Brownian motions $W^{(r)}$ (which is in some way the "laziest" choice from the programming viewpoint), then

$$\mathrm{Var}\left(\sum_{r=1}^{R}\alpha_r f(X_T^{(r)})\right) = \left(\sum_{r=1}^{R}\alpha_r^2\right)\mathrm{Var}(f(X_T))$$

having in mind that

$$\sum_{r=1}^{R}\alpha_r^2 \geq \alpha_R^2 = \left(\frac{R^R}{R!}\right)^2 \sim \frac{e^{2R}}{2\pi R} \quad \text{as} \quad R \to \infty.$$

The practical aspects of the above analysis can be summed up as follows.

Proposition 3.1 *Let $T > 0$. One considers the SDE (1.1). Assume that on every filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ on which exists a standard $\mathcal{F}_t$-Brownian motion W, (1.1) has a strong solution on $[0,T]$ and that its Euler scheme with step T/n (with Brownian increments) converges in every $L^p(\mathbb{P})$ for the sup-norm over $[0,T]$ for every $p \in [1,\infty)$. Furthermore assume that (1.1) admits an expansion $(\mathcal{E}_R^V)$ at an order $R \geq 1$ for a characterizing vector space V of continuous functions f with polynomial growth, stable by product.*
Let $\alpha \in \mathbb{R}^R$ be defined by (3.4). For every $r \in \{1,\ldots,R\}$ set

$$U_k^{(r)} := \sqrt{\frac{rn}{T}}\left(W_{\frac{kT}{rn}} - W_{\frac{(k-1)T}{rn}}\right), \ k \geq 1. \tag{3.7}$$

If $\bar{X}^{(r)}$ denotes an Euler scheme of (1.1) with step $\frac{T}{rn}$ associated to the Gaussian white noise $(U_k^{(r)})_{k\geq1}$, then

$$\mathbb{E}\left(\sum_{r=1}^{R}\alpha_r f(\bar{X}_T^{(r)})\right) = \mathbb{E}(f(X_T)) + O(n^{-R}) \ \text{ and } \ \lim_{n}\mathrm{Var}\left(\sum_{r=1}^{R}\alpha_r f(\bar{X}_T^{(r)})\right) = \mathrm{Var}(f(X_T)).$$

Remark. A first extension of this result to non continuous Borel functions can be made easily as follows: Convergence (3.5) implies that $(\bar{X}_T^{(1)}, \ldots, \bar{X}_T^{(r)}, \ldots, \bar{X}_T^{(R)})$ weakly converges in $(\mathbb{R}^d)^R$ to $(X_T^{(1)}, \ldots, X_T^{(1)})$. Then, if f is $\mathbb{P}_{X_T}(dx)$-a.s. continuous with polynomial growth, so is $(\sum_r \alpha_r f(x_r))^2$ with respect to $\mathbb{P}_{(X_T^{(1)}, \ldots, X_T^{(1)})}$. This implies that the convergence of the variances (3.6) still holds true.

3.2 The case of bounded Borel functions (elliptic diffusions)

We no longer assume that V is a subspace of continuous functions. As a counterpart, we make some more stringent assumptions on the coefficients b and σ of the diffusion (1.1), namely that they satisfy the following assumption:

$$(UE) \equiv \begin{cases} (i) & b_i, \sigma_{ij} \in \mathcal{C}_b^\infty(\mathbb{R}^d, \mathbb{R}), \ i \in \{1,\ldots,d\}, \ j \in \{1,\ldots,q\} \\ (ii) & \exists \varepsilon_0 > 0 \text{ such that } \forall x \in \mathbb{R}^d, \ \sigma\sigma^*(x) \geq \varepsilon_0 I_d. \end{cases}$$

(In particular this implies $q \geq d$). Then we know from [2] (see also [11]) that $(\mathcal{E}_R^V)$ holds at any order $R \geq 1$ with $V = \mathcal{B}_b(\mathbb{R}^d, \mathbb{R})$ where $\mathcal{B}_b(\mathbb{R}^d, \mathbb{R})$ denotes the set of bounded Borel functions $f : \mathbb{R}^d \to \mathbb{R}$. In that setting, the following result holds

Proposition 3.2 *If Assumption* (UE) *holds, then Proposition 3.1 still holds at any order* $R \geq 1$ *with* $V = \mathcal{B}_b(\mathbb{R}^d, \mathbb{R})$.

Proof. The conclusion of Proposition 3.1 in Section 3.1 remains valid for a subspace V of possibly non continuous functions provided the convergence of the variance term in Equation (3.6) still holds for every $f \in V$. Developing the variance term as follows

$$\mathrm{Var}\left(\sum_{r=1}^{R} \alpha_r f(\bar{X}_T^{(r)})\right) = \sum_{r,r'=1}^{R} \alpha_r \alpha_{r'} \mathbb{E}\left(f(\bar{X}_T^{(r)}) f(\bar{X}_T^{(r')})\right) - \left(\sum_{r=1}^{R} \alpha_r \mathbb{E} f(\bar{X}_T^{(r)})\right)^2$$

shows that it amounts to proving that

$$\forall\, r,\, r' \in \{1, \ldots, R\}, \quad \mathbb{E}\left(f(\bar{X}_T^{(r)}) f(\bar{X}_T^{(r')})\right) \longrightarrow \mathbb{E}\left(f(X_T^{(r)}) f(X_T^{(r')})\right) \quad \text{as} \quad n \to \infty. \quad (3.8)$$

In order to prove this convergence, we rely on the following lemma established in [3].

Lemma 3.2 *If* (UE) *holds then the distributions of* X_T *and its Brownian Euler schemes with step* T/n *, starting at* $x \in \mathbb{R}^d$ *are absolutely continuous with distributions* $p_T(x, y) \lambda_d(dy)$ *and* $\bar{p}_T^n(x, y) \lambda_d(dy)$ *respectively. Furthermore, they satisfy for every* $x, y \in \mathbb{R}^d$ *and every* $n \geq 1$,

$$\bar{p}_T^n(x, y) = p_T(x, y) + \frac{\pi_n(x, y)}{n} \quad \text{with} \quad |\pi_n(x, y)| \leq C_1 e^{-C_2 |x-y|^2}$$

for a real constants $C_1, C_2 > 0$.

Let μ denote the finite positive measure defined by $\mu(dy) = (p_T(x, y) + C_1 e^{-C_2 |x-y|^2}) \lambda_d(dy)$. The set $\mathrm{Lip}_b(\mathbb{R}^d, \mathbb{R})$ of bounded Lipschitz functions is everywhere dense in $L^1_{\mathbb{R}}(\mathbb{R}^d, \mu)$. Consequently, there exists a sequence $(f_k)_{k \geq 1}$ of functions of $\mathrm{Lip}_b(\mathbb{R}^d, \mathbb{R})$ such that $\lim_k \|f - f_k\|_1 = 0$. Furthermore, since one can replace each f_k by $((-\|f\|_\infty) \vee f_k) \wedge \|f\|_\infty$, one can assume without loss of generality that $\|f_k\|_\infty \leq \|f\|_\infty$. Then one gets

$$\begin{aligned}
\left|\mathbb{E}\left(f(\bar{X}_T^{(r)}) f(\bar{X}_T^{(r')}) - f(X_T^{(r)}) f(X_T^{(r')})\right)\right| \leq\ & \left|\mathbb{E}\left(f(\bar{X}_T^{(r)}) f(\bar{X}_T^{(r')}) - f_k(\bar{X}_T^{(r)}) f_k(\bar{X}_T^{(r')})\right)\right| \\
& + \left|\mathbb{E}\left(f_k(\bar{X}_T^{(r)}) f_k(\bar{X}_T^{(r')}) - f_k(X_T^{(r)}) f_k(X_T^{(r')})\right)\right| \\
& + \left|\mathbb{E}\left(f_k(X_T^{(r)}) f_k(X_T^{(r')}) - f(X_T^{(r)}) f(X_T^{(r')})\right)\right| \\
\leq\ & 2\|f\|_\infty \max_{1 \leq r \leq R} \mathbb{E}\left(|f - f_k|(\bar{X}_T^{(r)})\right) \\
& + 2\|f\|_\infty [f_k]_{\mathrm{Lip}} \max_{1 \leq r \leq R} \mathbb{E}|X_T^{(r)} - \bar{X}_T^{(r)}| \\
& + 2\|f\|_\infty \max_r \mathbb{E}\left(|f - f_k|(X_T^{(r)})\right).
\end{aligned}$$

Now, for every $k \geq 1$ and every $r \in \{1, \ldots, R\}$,

$$\mathbb{E}\left(|f - f_k|(\bar{X}_T^{(r)})\right) + \mathbb{E}\left(|f - f_k|(X_T^{(r)})\right) \leq 2\|f - f_k\|_{L^1(\mu)}$$

so that, for every $r,\, r' \in \{1, \ldots, R\}$ and every $k \geq 1$,

$$\limsup_n \left|\mathbb{E}\left(f(\bar{X}_T^{(r)}) f(\bar{X}_T^{(r')}) - f(X_T^{(r)}) f(X_T^{(r')})\right)\right| \leq 2\|f\|_\infty \|f - f_k\|_{L^1(\mu)}$$

which in turn implies (3.8) by letting k go to infinity. $\quad \diamond$

9

3.3 About the optimality of the Brownian specification when $R \geq 3$

We proved in Section 2 that for the standard R-R extrapolation ($R = 2$), the choice of the same underlying Brownian motion W for both schemes is optimal *i.e.* leads to the lowest possible asymptotic variance for any admissible function f. In fact this is the only case of optimality: the choice of consistent Brownian increments in the higher order R-R extrapolation is never optimal in terms of variance reduction when $R \geq 3$.

Let $f \in V$ denote a (fixed) function in V such that $\mathrm{Var}(f(X_T^{(1)})) > 0$. One may assume without loss of generality that

$$\mathrm{Var}(f(X_T^{(1)})) = 1.$$

Let $S^f := [\mathrm{Cov}(f(X_T^{(r)}), f(X_T^{(r')}))]_{1 \leq r, r' \leq R}$ denote the covariance matrix of the variables $f(X_T^{(r)})$. Since $\mathrm{Var}(f(X_T^{(r)})) = 1$, $r = 1, \ldots, R$, the minimization of the variance of the estimator $\alpha^* [f(X_T^{(r)})]_{1 \leq r \leq R}$ is "lower-bounded" in a natural way by the following problem

$$\min \left\{ \alpha^* S \alpha, \ S \in \mathcal{S}^+(d, \mathbb{R}), \ S_{ii} = 1 \right\}. \tag{3.9}$$

Choosing like in Propositions 3.1 and 3.2, $f(X_T^{(r)}) = f(X_T) \ \mathbb{P}\text{-}a.s.$, $r = 1, \ldots, R$ corresponds to $S^f = S = \mathbf{1}\,\mathbf{1}^*$ (*i.e.* $S_{ij}^f = 1$, $1 \leq i, j \leq R$).

We use the term "lower-bounding" to emphasize that for a given $f \in V$, there are possibly admissible matrices in (3.9) which cannot be covariance matrices for $(f(X_T^{(r)}))_{1 \leq r \leq R}$.

Proposition 3.3 (*a*) *When $R \geq 3$,*

$$\min \left\{ \alpha^* S \alpha, \ S \in \mathcal{S}^+(R, \mathbb{R}), \ S_{ii} = 1 \right\} = 0,$$

hence, the unit matrix $S_1 := \mathbf{1}\,\mathbf{1}^*$ is never a solution to the minimization problem (3.9).
(*b*) *When $R = 3$, if $f(X_T) = \mathbb{E}f(X_T) + \mathrm{StD}(f(X_T))\varepsilon$, $\mathbb{P}(\varepsilon = \pm 1) = 1/2$, then the extrapolation formula stands as an exact quadrature formula.*

Remark. Although it has no practical interest for applications, the situation described in (*b*) may happen: assume the drift b is odd, the diffusion coefficient σ is even and f is the sign function $f(x) = \mathrm{sign}(x)$. When $R \geq 4$, for a given function f, the solution(s) of the abstract minimization problem (3.9) do not correspond in general to an admissible covariance matrix for $(f(X_T^{(r)}))_{1 \leq r \leq R}$ (see below the proof for a short discussion on simple examples). Furthermore, even when it happens to be the case, the variance control provided by $\mathbf{1}\,\mathbf{1}^*$ is obviously more straightforward than a numerical search of the appropriate underlying covariance structure of the Brownian motions $(W^{(r)})_{1 \leq r \leq R}$ driving the diffusions. However, one must keep in mind that multi-step R-R extrapolation leaves some degrees of freedom to some variance reduction method as soon as $R \geq 3$ and the opportunity in terms of computational cost to design an online optimization procedure (depending on the function f) may be interesting in some cases. The appropriate approach is then an online recursive stochastic approximation method somewhat similar to that introduced in [1] for online variance reduction. This is beyond the scope of the present paper.

Proof. (*a*) Assume $R \geq 4$. We will show that the minimum in (3.9) is in fact 0. Let $I^{\pm} := \{i \in \{1, \ldots, R\} \,|\, \alpha_i \gtrless 0\}$. Let $\rho \in (0, 1)$ and let $A(\rho) := \sqrt{(1 - \rho)I_R + \rho \mathbf{1}\mathbf{1}^*}$ denote

the symmetric square root of the positive definite matrix $(1 - \rho)I_R + \rho \mathbf{1}\mathbf{1}^*$. Then set for every $i \in I^+$, $C_i = A(\rho)e_i$ where $(e_1, \ldots, e_R)$ still denotes the canonical basis of $\mathbb{R}^R$ and

$$\theta_i := \frac{\alpha_i}{\sum_{j \in I^-} |\alpha_j|}, \qquad i \in I^+.$$

First note that $|C_i|^2 = \sum_j (A(\rho)_{ji})^2 = (A(\rho)^2)_{ii} = 1$ and $(C_i | C_j) = \rho$, $i \neq j$. Then,

$$\sum_{i \in I^+} \theta_i = \frac{\sum_{i \in I^+} |\alpha_i|}{\sum_{j \in I^-} |\alpha_j|} = \frac{1 + \sum_{j \in I^-} |\alpha_j|}{\sum_{j \in I^-} |\alpha_j|} > 1.$$

On the other hand

$$\sum_{i \in I^+} \theta_i^2 = \frac{\sum_{i \in I^+} \alpha_i^2}{(\sum_{j \in I^-} |\alpha_j|)^2} = \frac{\sum_{i \in I^+} \alpha_i^2}{(\sum_{i \in I^+} \alpha_i - 1)^2} = \frac{\sum_{i \in I^+} \alpha_i^2}{\sum_{i \in I^+} \alpha_i^2 + 1 + \sum_{i \in I^+} \alpha_i (\sum_{j \neq i, j \in I^+} \alpha_j - 2)} < 1$$

since for every $i \in I^+$, $\sum_{j \neq i, j \in I^+} \alpha_j \geq \min(\alpha_{R-2}, \alpha_R) > 2$ provided $R \geq 4$. Consequently the function $g : [0, 1] \to \mathbb{R}$ defined by $g(\rho) := |\sum_{i \in I^+} \theta_i C_i|^2$ satisfies $g(0) = \sum_{i \in I^+} \theta_i^2 < 1$ and $g(1) = (\sum_{i \in I^+} \theta_i)^2 > 1$ so that there exists a real number $\rho_0 \in (0, 1)$ such that $g(\rho_0) = 1$. From now on assume that $\rho = \rho_0$.

Set $C_i = C := \sum_{j \in I^+} \theta_j C_j$, for every $i \in I^-$ so that $|C| = 1$ and

$$\sum_{r=1}^R \alpha_i C_i = (\sum_{i \in I^-} \alpha_i)C + \sum_{i \in I^+} \alpha_i C_i = 0.$$

Consequently the symmetric nonnegative matrix

$$S := [C_1 \cdots C_R]^* [C_1 \cdots C_R] \tag{3.10}$$

satisfies $S_{ii} = 1$ for every $i \in \{1, \ldots, R\}$ and $S\alpha = 0$.

When $R = 3$, one checks that $S^f = u\, u^*$, $u^* = [1 \ -1 \ -1]$, satisfies $S^f \alpha = 0$ since $\alpha_2 + \alpha_3 = 1/2 = \alpha_1$.

(b) The situation for $R = 3$ corresponds to $f(X_T^{(2)}) = f(X_T^{(3)}) = 2\,\mathbb{E}(f(X_T^{(1)})) - f(X_T^{(1)})$ so that $f(X_T^{(1)}) + f(X_T^{(2)}) = 2\,\mathbb{E}f(X_T^{(1)})$ $\mathbb{P}$-a.s.. The same identity holds for $f^2 \in V$, so that $f(X_T^{(1)}) \times f(X_T^{(2)}) = \mathbb{E}f(X_T^{(1)})f(X_T^{(2)})$ $\mathbb{P}$-a.s. (use $ab = \frac{1}{2}((a+b)^2 - a^2 - b^2)$). Consequently $f(X_T^{(1)})$ and $f(X_T^{(2)})$ are the zeros of $z^2 - uz + v$ for some deterministic real constants u and v. The conclusion follows once noticed that $f(X_T^{(1)}) \overset{d}{=} f(X_T^{(2)})$. $\quad \diamond$

TWO TOY SITUATIONS: (a) Assume – which is of no numerical interest – that $X_T^{(r)} = W_T^{(r)}$, $r = 1, \ldots, R$, where the covariance of the Brownian motions $W^{(r)}$ is given by (3.10). Then $S\alpha = 0$ so that the multi-step R-R extrapolation is an exact quadrature formula (to compute $0 \ldots$).

(b) If one considers some correlated geometrical Brownian motions $X_t^{(r)} = e^{-\frac{\sigma^2 t}{2} + \sigma W_t^{(r)}}$ $r = 1, \ldots, R$, where $(W^{(1)}, \ldots, W^{(R)})$ has $C_W = [c_{rr'}]_{1 \leq r, r' \leq R}$ as a covariance matrix (at time $t = 1$), then the covariance matrix of $(X^{(1)}, \ldots, X^{(R)})$ is given by

$$C_X = [e^{c_{rr'} \sigma} - 1]_{1 \leq r, r' \leq R}.$$

One easily checks that the mapping $\Psi : [c_{rr'}] \mapsto [e^{c_{rr'}\sigma} - 1]$ is not bijective on the set of non negative symmetric matrices with constant diagonal coefficients: let $c = \log(1 + \kappa)$, $\rho = \log(1 + \vartheta)$ where $\vartheta \geq -\frac{\kappa}{R-1}$ so that the matrix $[(\kappa - \vartheta)\delta_{rr'} + \vartheta]$ is nonnegative. Assume $\vartheta = -\frac{\lambda}{R-1}$ with $\lambda \in (\log(1 + \kappa), \kappa)$. Then this covariance structure cannot be obtained from some correlated Brownian motions since $c + (R-1)\rho = \log(1 + \kappa) + (R-1)\log(1 + \vartheta) < 0$.

4 Complexity of the multi-step R-R procedure (with consistent Brownian increments)

Throughout this section, we assume that the R Gaussian white noises $(U_k^{(r)})_{1 \leq k \leq nr}$ are consistent Brownian increments given by (3.7).

4.1 Complexity

The simulation of the correlated sequences $U_k^{(r)}$ can often be neglected in terms of complexity with respect to the computation of one step of the R Euler schemes (the main concern being in fact to spare the random number generator, see below). Then, one can easily derive the complexity of one Monte Carlo path of this R-order R-R extrapolation procedure: for every $r \in \{1, \ldots, R\}$, one has to compute rn values of an Euler scheme. So the complexity $\kappa(R)$ of the procedure is given by

$$\kappa(R) = C_{b,\sigma} \times n \times \sum_{r=1}^{R} r = C_{b,\sigma} \times n \times \frac{R(R+1)}{2}$$

where $C_{b,\sigma}$ denotes the complexity of a one time step computation in (1.1). Note that using a recursively iterated R-R procedure would have lead to consider some Euler schemes with steps $\frac{T}{2^{r-1}n}$, $1 \leq r \leq R$, (see e.g. [7], p.362 when $R = 3$) which in turn would have induced a complexity of order $C_{b,\sigma} \times n \times (2^R - 1)$.

Thus the global complexity of a Monte Carlo simulation of size M is then

$$N := C_{b,\sigma} \times \frac{R(R+1)}{2} \times n \times M.$$

Up to a scaling of the global complexity by $1/C_{b,\sigma}$ one may assume without loss of generality that $C_{b,\sigma} = 1$. So evaluating the performances of the R-R extrapolation as a function of its complexity amounts to solving the following minimization problem:

$$\mathcal{E}(R, N) := \min_{\frac{R(R+1)}{2} \times n \times M = N} \sqrt{\frac{\mathrm{Var}(f(X_T))}{M}(1 + \varepsilon(n)) + \frac{\widetilde{c}_R^2}{n^{2R}}(1 + O(1/n))}$$

where $\lim_n \varepsilon(n) = 0$, since

$$\left\| \mathbb{E}(f(X_T)) - \frac{1}{M} \sum_{m=1}^{M} \sum_{r=1}^{R} \alpha_r f((\bar{X}_T^{(r)})^m) \right\|_2^2 = \frac{\mathrm{Var}(f(X_T))}{M}(1 + \varepsilon(n)) + \frac{\widetilde{c}_R^2}{n^{2R}}(1 + O(1/n)).$$

Standard computations lead to the following Proposition.

Proposition 4.4 *Let $R \geq 2$ be the order of expansion of the time discretization error.*

(a) Then,

$$\lim_{N \to \infty} N^{-\frac{1}{2}(1-\frac{1}{2R+1})} \mathcal{E}(R,N) = \frac{1}{\sqrt{2}} \times \Theta(R) \times R\, |\widetilde{c}_R|^{\frac{1}{2R+1}} \times (\mathrm{Var}(f(X_T)))^{\frac{R}{2R+1}}$$

where $\Theta(R) = 2^{\frac{1}{2(2R+1)}} R^{-\frac{1}{2R+1}} (1 + \frac{1}{R})^{\frac{R}{2R+1}} \left((2R)^{-\frac{2R}{2R+1}} + (2R)^{\frac{1}{2R+1}} \right)^{\frac{1}{2}} \to 1$ *as* $R \to \infty$.

(b) Furthermore, for a fixed complexity level N, the solution $(n(N), M(N))$ of the minimization problem satisfies

$$M(N) \sim \frac{2}{R(R+1)} \left(\frac{R+1}{4} \right)^{\frac{1}{2R+1}} \left(\frac{\mathrm{Var}(f(X_T))}{\widetilde{c}_R^2} \right)^{\frac{1}{2R+1}} N^{\frac{2R}{2R+1}}$$

and

$$n(N) = \left(\frac{R+1}{4} \right)^{-\frac{1}{2R+1}} \left(\frac{\widetilde{c}_R^2}{\mathrm{Var}(f(X_T))} \right)^{\frac{1}{2R+1}} N^{\frac{1}{2R+1}}$$

as $N \to \infty$ so that

$$M(N) \sim \frac{1}{2R} \frac{\mathrm{Var}(f(X_T))}{\widetilde{c}_R^2} n(N)^{2R}.$$

COMMENTS:

– The rate of convergence increases with R and tends toward $N^{\frac{1}{2}}$. This means that, as expected, the higher one expands the time discretization error, the less this error slows down the asymptotic global rate of convergence.

– However, an expansion of order R being fixed, the range at which the theoretical rate of convergence $N^{\frac{1}{2}(1-\frac{1}{2R+1})}$ does occur depends on the value of the term $R|\widetilde{c}_R|^{\frac{1}{2R+1}} = \frac{R|c_R|^{\frac{1}{2R+1}}}{(R!)^{\frac{1}{2R+1}}}$ ($\sim \sqrt{eR}|c_R|^{\frac{1}{2R+1}}$ as $R \to \infty$) for which no bound or estimate is available in practice. Although some theoretical explicit expression do exist for the coefficients c_R (at least for small values of R, see [16], [11], [6]), it is usually not possible to have some numerical estimates.

– The time discretization parameter $n = n(N)$ and the size $M = M(N)$ of the Monte Carlo simulation are explicit functions of N which involve the unknown coefficient $|c_R|$. So the above *sharp* L^2-rate of convergence is essentially a theoretical bound that cannot be achieved at a reasonable cost in practice. However one could imagine to produce a rough estimate of c_R by normalizing the R-R extrapolation (by n^R) using a small preliminary MC simulation so as to design $n(N)$ and $M(N)$ using the above rules. However the recursive feature of the resulting MC would be lost which seems not very realistic in practice.

4.2 Efficient simulation of consistent Brownian increments

The question of interest in this section is the simulation of the Brownian increments. Several methods can be implemented, the main concern being to spare the random number generator (its contribution to the global complexity of the simulation can be neglected in practice as soon as b and σ have "not too simple" expressions).

The problem amounts to simulating the increments of the R Euler schemes, say between absolute time $t = 0$ and $t = T/n$. This means to simulate the upper triangular matrix

$$\left[W_{\frac{\ell T}{rn}} - W_{\frac{(\ell-1)T}{rn}}\right]_{1 \leq \ell \leq r, 1 \leq r \leq R}.$$

4.2.1 A lazy approach

Let $R \geq 2$ be a fixed integer. For every $r \in \{1, \ldots, R\}$, set

$$M(R) := \mathrm{lcm}(1, 2, \ldots, R) \qquad \text{and} \qquad m(r) := M(R)/r, \ r = 1, \ldots, R.$$

Then, one simulates $nM(R)$ independent copies $\xi_1, \ldots, \xi_{nM(R)}$ of $\mathcal{N}(0; 1)$ and one sets

$$U_k^{(r)} = \frac{1}{\sqrt{m(r)}} \sum_{i=(k-1)m(r)+1}^{km(r)} \xi_i, \qquad 1 \leq k \leq r\,n.$$

Such a simulation strategy consumes $n \times M(R)$ random numbers to complete the simulation of the R Euler schemes until maturity T.

4.2.2 Saving the random number generator

One only simulates what is needed to get the above matrix. This means to simulate the Brownian increments between the points of the set

$$S_R := \left\{\frac{\ell}{r}\frac{T}{n}, \ 1 \leq \ell \leq r, \ 1 \leq r \leq R\right\}.$$

Such a simulation strategy consumes $n \times \mathrm{card}\, S_R$ random numbers to complete the simulation of the R Euler schemes. Obviously

$$S_R = \left\{\frac{\ell}{r}\frac{T}{n}, \ 1 \leq \ell \leq r, \ 1 \leq r \leq R, \ \gcd(\ell, r) = 1\right\}$$

so that

$$\mathrm{card}\, S_R = \sum_{r=1}^{R} \varphi(r)$$

where φ denotes the Euler function with the convention $\varphi(1) = 1$. Classical Number Theory results say that

$$\sum_{r=1}^{R} \varphi(r) \sim \frac{3}{\pi^2} R^2 \qquad \text{as} \qquad R \to +\infty.$$

whereas $M(R) = e^{(1+o(1))R}$ as $R \to \infty$. In practice, only the first values of $M(R)$ and $\mathrm{card}(S_R)$ are of interest. They are are reported in the table below.

R	1	2	3	4	5	6	7
$M(R)$	1	2	6	12	60	60	420
$\mathrm{card}(S_R)$	1	2	4	6	10	12	18

For a given value of R, it is necessary to sort the values $\frac{\ell}{r} \in S_R$ and to design the array of resulting lengths in order to simulate the above Brownian matrix at each "global" time step T/n. But this can be done once since it is universal.

5 Numerical experiments

5.1 Some tables

We specify below how to simulate the Gaussian white noises $(U_k^{(r)})_{1 \le k \le r}$ of the R coupled Euler schemes. Note that these tables are equivalent to a "multistep" Brownian bridge based recursive simulation of the underlying Wiener process.

– $\boxed{R = 2:}$ Let U_1, U_2 be two i.i.d. copies of $\mathcal{N}(0; I_q)$. Set

$$U_i^{(2)} = U_i, \; i = 1, 2, \qquad U_1^{(1)} = \frac{U_1 + U_2}{\sqrt{2}}$$

and
$$\alpha_1 = -1, \qquad \alpha_2 = 2.$$

(Note that $\sum_i \alpha_i^2 = 5$)).

– $\boxed{R = 3:}$ Let U_1, U_2, U_3, U_4 be four i.i.d. copies of $\mathcal{N}(0; I_q)$. Set

$$U_1^{(3)} = U_1, \quad U_2^{(3)} = \frac{U_2 + U_3}{\sqrt{2}}, \quad U_3^{(3)} = U_4,$$

$$U_1^{(2)} = \frac{\sqrt{2}\,U_1 + U_2}{\sqrt{3}}, \quad U_2^{(2)} = \frac{U_3 + \sqrt{2}\,U_4}{\sqrt{3}},$$

$$U_1^{(1)} = \frac{U_1^{(2)} + U_2^{(2)}}{\sqrt{2}}.$$

and
$$\alpha_1 = \frac{1}{2}, \qquad \alpha_2 = -4, \qquad \alpha_3 = \frac{9}{2}.$$

(Note that $\sum_i \alpha_i^2 = \frac{73}{2}$).

– $\boxed{R = 4:}$ Let $U_1, \ldots, U_6$ be six i.i.d. copies of $\mathcal{N}(0; I_q)$. Set

$$U_1^{(4)} = U_1, \quad U_2^{(4)} = \frac{U_2 + \sqrt{2}\,U_3}{\sqrt{3}}, \quad U_3^{(4)} = \frac{\sqrt{2}\,U_4 + U_5}{\sqrt{3}}, \quad U_4^{(4)} = U_6.$$

$$U_1^{(3)} = \frac{\sqrt{3}\,U_1 + U_2}{2}, \quad U_2^{(3)} = \frac{U_3 + U_4}{\sqrt{2}}, \quad U_3^{(3)} = \frac{U_5 + \sqrt{3}\,U_6}{2},$$

$$U_1^{(2)} = \frac{U_1^{(4)} + U_2^{(4)}}{\sqrt{2}}, \quad U_2^{(2)} = \frac{U_3^{(4)} + U_4^{(4)}}{\sqrt{2}},$$

$$U_1^{(1)} = \frac{U_1^{(2)} + U_2^{(2)}}{\sqrt{2}}.$$

and
$$\alpha_1 = -\frac{1}{6}, \qquad \alpha_2 = 4, \qquad \alpha_3 = -\frac{27}{2}, \qquad \alpha_4 = \frac{32}{3}.$$

(Note that $\sum_i \alpha_i^2 = \frac{5617}{18} \approx 312.06$).

– $\boxed{R = 5:}$

$$\alpha_1 = \frac{1}{24}, \qquad \alpha_2 = -\frac{8}{3}, \qquad \alpha_3 = \frac{81}{4}, \qquad \alpha_4 = -\frac{128}{3}, \qquad \alpha_5 = \frac{625}{24}.$$

5.2   Numerical test with a high volatility Black-Scholes model

To illustrate the efficiency of the use of consistent white noises in R-R extrapolation, we consider the simplest option pricing model, the (risk-neutral) Black-Scholes dynamics, but with unusually high volatility. (We are aware this model used to price Call options does not fulfill the theoretical assumptions made above). To be precise

$$dX_t = X_t\,(rdt + \sigma dW_t),$$

with the following values for the parameters

$$X_0 = 100,\ K = 100,\ r = 0.15,\ \sigma = 1.0,\ T = 1.$$

Note that a volatility $\sigma = 100\%$ per year is equivalent to a 4 year maturity with volatility 50% (or 16 years with volatility 25%). The reference Black-Scholes premium is $C_0^{BS} = 42.96$. We consider the Euler scheme with step T/n of this equation.

$$\bar{X}_{t_{k+1}} = \bar{X}_{t_k}\left(1 + r\frac{T}{n} + \sigma\sqrt{\frac{T}{n}}U_{k+1}\right), \qquad \bar{X}_0 = X_0,$$

where $t_k = \frac{kT}{n}$, $k = 0, \ldots, n$. We want to price a vanilla Call option *i.e.* to compute

$$C_0 = e^{-rT}\mathbb{E}((X_T - K)_+)$$

using a Monte Carlo simulation with M sample paths.

• THREE STEP R-R EXTRAPOLATION ($R = 3$): We processed three Monte Carlo simulations of common size $M = 10^6$ to evaluate the efficiency of the R-R extrapolation with $R = 3$, having in mind that, for a given complexity N, teh size of teh Monte Carlo simulation and the time discritization parameter satisfy $M(N) \propto n(N)^{2R}$, see Proposition 4.4(b).

– An R-R extrapolation of order $R = 3$ as defined by (3.3) and (3.4) with $n = 2, 4, 6, 8, 10$ with *consistent increments* (the maximal number of steps is $Rn = 3n$). Note these specifications for n are quite low in comparison with the high volatility of the B-S model.

– An R-R extrapolation of order $R = 3$ with the same architecture but implemented with *independent increments*.

– A *regular Euler scheme with the same complexity i.e.* with $R(R+1)/2 \times n = 6\,n$ steps.

The results are depicted in the Figures 1 and 2 below. In Figure 1, the abscissas represent the size $(3n)$ of the Euler scheme with the highest discretization frequency used in the procedure. In Figure 2, the abscissas represent the size $(6n)$ of the standard Euler scheme with the same complexity as the R-R extrapolation ($R = 3$). The main conclusions are the following:

– The standard deviation of the R-R extrapolation with independent noises is 5 times greater than the one observed with consistent increments. This makes this higher order R-R expansions ($R = 3$) useless at least within the usual range of our Monte Carlo simulations
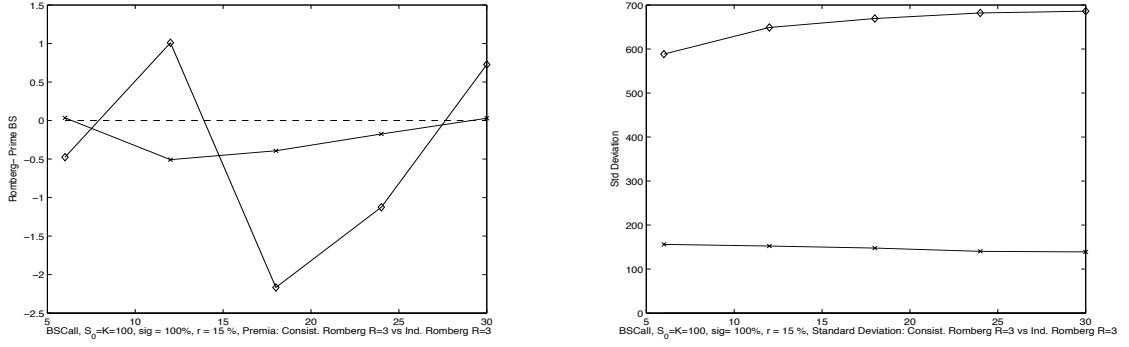
Figure 1: B-S EURO CALL OPTION. $M = 10^6$. *R-R extrapolations* $R = 3$. *Consistent Brownian increments* (×—×—×) *vs independent Brownian increments* (◇—◇—◇). $X_0 = K = 100$, $\sigma = 100\%$, $r = 15\%$. *Abscissas:* $3n$, $n = 2, 4, 6, 8, 10$. *Left:* Premia. *Right:* Standard Deviations.

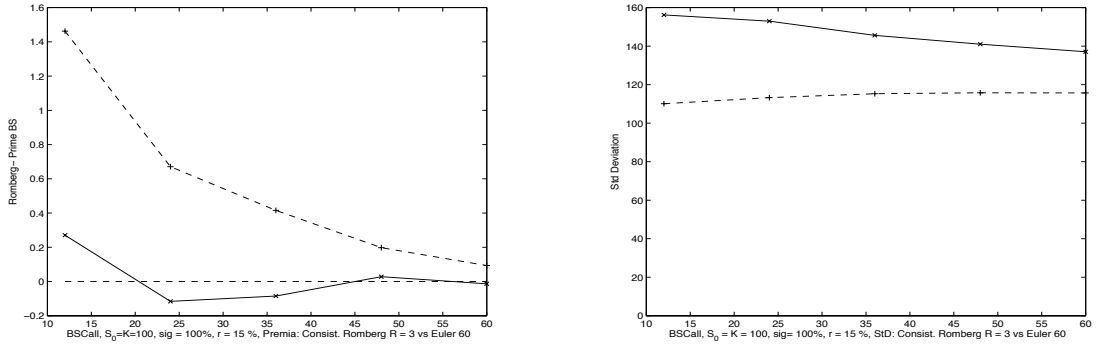Figure 2: B-S EURO CALL OPTION. $M = 10^6$. *R-R extrapolation* $R = 3$. *Consistent Brownian increments* (×—×—×) *vs Euler scheme with equivalent complexity* (+ - - + - - +). $X_0 = K = 100$, $\sigma = 100\%$, $r = 15\%$. *Abscissas:* $3n$, $n = 2, 4, 6, 8, 10$. *Abscissas:* $6n$, $n = 2, 4, 6, 8, 10$. *Left:* Premia. *Right:* Standard Deviations.

(see Fig. 1): in fact it is less efficient in our high volatility setting than the Euler scheme with equivalent complexity and less efficient than a standard R-R extrapolation ($R = 2$).

– Considering consistent increments in the R-R extrapolation gives the method its full efficiency as emphasized by Figure 2 : R-R extrapolation is clearly much more efficient than the Euler scheme of equal complexity in a Monte Carlo simulation of size $M = 10^6$.

• FOUR STEP R-R EXTRAPOLATION ($R = 4$): With a size of $M = 10^6$, the R-R-extrapolation with $R = 4$ is not completely convincing: at this range, the variance of the estimator is not yet controlled by $\mathrm{Var}(f(X_T))$ for the selected (small) values of the discretization parameter n.

However, for a larger simulation, say $M = 10^8$, the multistep extrapolation of order $R = 4$ clearly becomes the most efficient one as illustrated by Figure 3 (right) and Table 1.
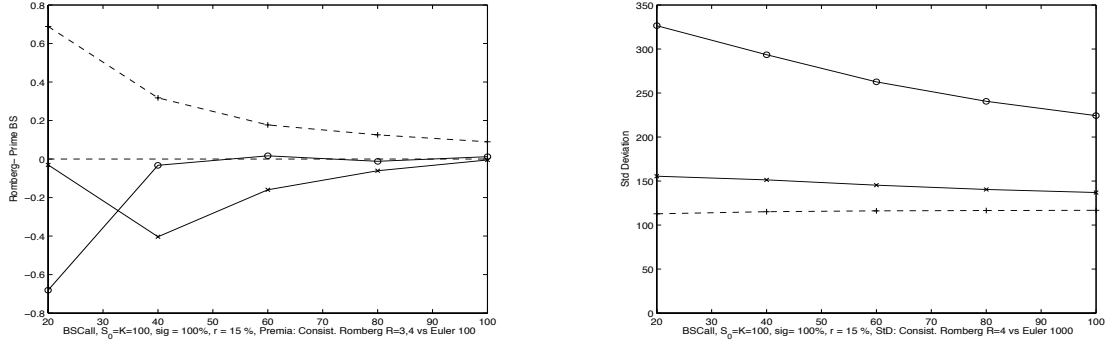
To carry out a comparision, we implemented this time:

17

Figure 3: B-S EURO CALL option. $M = 10^8$. *R-R extrapolation $R = 3, 4$. Consistent Brownian increments $R = 3$: —×—×—×—; Consistent Brownian increments $R = 4$: —o—o—o—; Euler scheme with equivalent complexity (+ - -+ - -+). $X_0 = K = 100$, $\sigma = 100\%$, $r = 15\%$. Abscissas: $10n$, $n = 2, 4, 6, 8, 10$. Left: Premia. Right: Standard Deviations.*

– Two R-R extrapolations with orders $R = 3$ and $R = 4$ as defined by (3.3) and (3.4) with $n = 2, 4, 6, 8, 10$ with *consistent increments*.

– A *regular Euler scheme with the same complexity i.e.* with $4 \times 5/2 \times n = 10\,n$ steps.

In Figure 3, the abscissas represent the size $(10n)$ of the Euler scheme with same complexity as the R-R extrapolation with $R = 4$.

n	2	4	6	8	10
BS *Ref.*			42.96		
$R = 3$	42.93	42.55	42.80	42.90	42.95
	(-0.07%)	(-0.94%)	(-0.37%)	(-0.14%)	(-0.01%)
$R = 4$	42.28	42.92	42.97	42.94	42.97
	(-1.59%)	(-0.08%)	(0.04%)	(-0.03%)	(0.03%)

Table 1. *R-R extrapolation with $R = 3, 4$ vs the Euler scheme with step $1/(10n)$.*

This emphasizes the natural field of application of multistep R-R extrapolation for numerical applications $(R \geq 4)$: this is the most efficient method to obtain accurate results in a high variance framework: it allows a smaller size of discretization step.

5.3 Further numerical experiments: path dependent options

In this section we will consider some path-dependent (European) options *i.e.* related to some payoffs $F((X_t)_{t \in [0,T]})$ where F is a functional defined on the set $\mathbb{D}([0,T], \mathbb{R}^d)$ of right continuous left-limited functions $x : [0,T] \to \mathbb{R}$. It is clear that all the asymptotic control of the variance obtained in Section 3.1 for the estimator $\sum_{r=1}^{R} \alpha_r f(\bar{X}_T^{(r)})$ of $\mathbb{E}(f(X_T))$ when f is continuous can be extended to functionals $F : \mathbb{D}([0,T], \mathbb{R}^d) \to \mathbb{R}$ which are $\mathbb{P}_X$-*a.s.* continuous with respect to the sup-norm defined by $\| x \|_{\sup} := \sup_{t \in [0,T]} |x(t)|$ with polynomial growth (*i.e.* $|F(x)| = O(\|x\|_{\sup}^{\ell})$ for somme natural integer ℓ as $\|x\|_{\sup} \to \infty$).

18

This simply follows from the fact that the (piecewise constant) Euler scheme $\bar{X}$ (with step T/n) converges for the sup-norm toward X in $L^2(\mathbb{P})$.

The same result holds with the continuous Euler scheme $\bar{X}^c$ defined by

$$\forall\, t \in [0,T], \qquad \bar{X}_t^c = x_0 + \int_0^t b(\bar{X}_{\underline{s}}^c)ds + \int_0^t \sigma(\bar{X}_{\underline{s}}^c)dW_s, \ \underline{t} = \lfloor nt/T \rfloor$$

(the diffusion coefficients have been frozen between time discretization instants). With this scheme, one can simply consider the path set $\mathcal{C}([0,T], \mathbb{R}^d)$ instead of $\mathbb{D}([0,T], \mathbb{R}^d)$.

Furthermore, this asymptotic control of the variance holds true with any R-tuple $\alpha = (\alpha_r)_{1 \le r \le R}$ of weights coefficents satisfying $\sum_{1 \le r \le R} \alpha_r = 1$, so these coefficients can be adapted to the structure of the weak error expansion.

On the other hand, in the recent past years, several papers provided some weak rates of convergence for some families of functionals F. These works were essentially motivated by the pricing of path-dependent (European) options, like Asian, lookback or barrier options. This corresponds to functionals

$$F(x) := \Phi\Big(\int_0^T x(s)ds\Big), \quad F(x) := \Phi\big(x(T), \sup_{t \in [0,T]} x(t), \inf_{t \in [0,T]} x(t)\big), \quad F(x) = \Phi(x(T))\mathbf{1}_{\{\tau_D(x) \le T\}}$$

where Φ is usually at least Lipschitz and $\tau_D := \inf\{s \in [0,T], x(s\pm) \in {}^c\!D\}$ is the hitting time of ${}^c\!D$ by x ([2]). Let us briefly mention two well-known examples:

– In [9], it is established that if the domain D has a smooth enough boundary, $b, \sigma \in \mathcal{C}^3(\mathbb{R}^d)$, σ uniformly elliptic on D, then for every Borel bounded function f vanishing in a neighbourhood of ∂D,

$$\mathbb{E}(f(\bar{X}_T)\mathbf{1}_{\{\tau(\bar{X}) > T\}}) - \mathbb{E}(f(X_T)\mathbf{1}_{\{\tau(X) > T\}}) = O\left(\frac{1}{\sqrt{n}}\right) \quad \text{as} \quad n \to \infty. \qquad (5.11)$$

If furthermore, b and σ are $\mathcal{C}^5$, then

$$\mathbb{E}(f(\bar{X}_T^c)\mathbf{1}_{\{\tau(\bar{X}^c) > T\}}) - \mathbb{E}(f(X_T)\mathbf{1}_{\{\tau(X) > T\}}) = O\left(\frac{1}{n}\right) \quad \text{as} \quad n \to \infty. \qquad (5.12)$$

Note however that these assumptions are not satisfied by usual barrier options (see below).

– it is suggested in [15] (including a rigorous proof when $X = W$) that if $b, \sigma \in \mathcal{C}_b^4(\mathbb{R})$, σ is uniformly elliptic and $\Phi \in \mathcal{C}^{4,2}(\mathbb{R}^2)$ (with some partial derivatives with polynomial growth), then

$$\mathbb{E}(\Phi(\bar{X}_T, \min_{0 \le k \le n} \bar{X}_{t_k})) - \mathbb{E}(\Phi(X_T, \min_{t \in [0,T]} X_t)) = O\left(\frac{1}{\sqrt{n}}\right) \quad \text{as} \quad n \to \infty. \qquad (5.13)$$

A similar improvement – $O(\frac{1}{n})$ rate – as above can be expected when replacing $\bar{X}$ by the continuous Euler scheme $\bar{X}^c$.

For both classes of functionals (with D as a half-line in 1-dimension in the first setting), the practical implementation of the continuous Euler scheme is known as the *Brownian bridge method*. It relies on the simulations of the distribution of $\min_{t \in [0,T]} \bar{X}_t^c$ and

---

[2] when x is stepwise constant and càdlàg, one can write "s" instead of "$s\pm$".

$\max_{t \in [0,T]} \bar{X}_t^c$ given the event $\{\bar{X}_{t_k}^c = x_k, \ k = 0, \dots, n\}$ ($= \{\bar{X}_{t_k} = x_k, \ k = 0, \dots, n\}$). This distribution is known since

$$\mathcal{L}(\max_{t \in [0,T]} \bar{X}_t^c \mid \{\bar{X}_{t_k} = x_k, \ k = 0, \dots, n\}) = \mathcal{L}(\max_{0 \le k \le n-1} G_{x_k, x_{k+1}}^{-1}(U_k))$$

where

$$G_{x,y}^{-1}(u) = \frac{1}{2}\left(x + y + \sqrt{(y-x)^2 - 2T\sigma^2(x)\log(u)/n}\right)$$

and $(U_k)_{0 \le k \le n-1}$ are $i.i.d.$ uniformly distributed random variables over the unit interval. A similar formula holds for the minimum using now the inverse distribution function

$$F_{x,y}^{-1}(u) = \frac{1}{2}\left(x + y - \sqrt{(y-x)^2 - 2T\sigma^2(x)\log(u)/n}\right).$$

At this stage there are two ways to implement the (multistep) R-R extrapolation with consistent Brownian increments in order to improve the performances of the original (stepwise constant or continuous) Euler schemes. Both rely on natural conjectures about the existence of a higher order expansion of the time discretization error suggested by the above rates of convergence (5.11), (5.12) and (5.13).

• *Standard Euler scheme:* As concerns the standard Euler scheme, this means the existence of a vector space V (stable by product) of admissible functionals satisfying

$$(\mathcal{E}_R^{\frac{1}{2}, V}) \qquad \equiv \qquad \forall F \in V, \qquad \mathbb{E}(F(X)) = \mathbb{E}(F(\bar{X})) + \sum_{k=1}^{R-1} \frac{c_k}{n^{\frac{k}{2}}} + O(n^{-\frac{R}{2}}). \qquad (5.14)$$

The main point for practical application is to compute the weights $\alpha^{(\frac{1}{2})} = (\alpha_r^{(\frac{1}{2})})_{1 \le r \le R}$ of the extrapolation are modified. Namely

$$\alpha_r^{(\frac{1}{2})} = \alpha^{(\frac{1}{2})}(r, R) := \frac{(-1)^{R-r}}{2} \frac{r^R}{r!(R-r)!} \prod_{k=1}^{R}\left(1 + \sqrt{\frac{k}{r}}\right), \qquad 1 \le r \le R.$$

For small values of R, we have

$\boxed{R = 2}$ $\quad \alpha_1^{(\frac{1}{2})} = -(1 + \sqrt{2}), \qquad \alpha_2^{(\frac{1}{2})} = \sqrt{2}(1 + \sqrt{2}).$

$\boxed{R = 3}$ $\quad \alpha_1^{(\frac{1}{2})} = \dfrac{\sqrt{3} - \sqrt{2}}{2\sqrt{2} - \sqrt{3} - 1}, \qquad \alpha_2^{(\frac{1}{2})} = -2\dfrac{\sqrt{3} - 1}{2\sqrt{2} - \sqrt{3} - 1}, \qquad \alpha_3^{(\frac{1}{2})} = 3\dfrac{\sqrt{2} - 1}{2\sqrt{2} - \sqrt{3} - 1}.$

$\boxed{R = 4}$ $\quad \alpha_1^{(\frac{1}{2})} = -\dfrac{(1 + \sqrt{2})(1 + \sqrt{3})}{2}, \qquad \alpha_2^{(\frac{1}{2})} = 4(\dfrac{3}{2} + \sqrt{2})(\sqrt{3} + \sqrt{2}),$

$\qquad \alpha_3^{(\frac{1}{2})} = -\dfrac{3}{2}(\sqrt{3} + \sqrt{2})(2 + \sqrt{3})(3 + \sqrt{3}), \qquad \alpha_4^{(\frac{1}{2})} = 4(2 + \sqrt{2})(2 + \sqrt{3}).$

Note that these coefficients have greater absolute values than in the standard case. Thus if $R = 4$, $\sum_{1 \le r \le 4}(\alpha_r^{(\frac{1}{2})})^2 \approx 10\,900!$ which induces an increase of the variance term for too small values of the time discretization parameters n even when increments are consistently generated. The complexity computations carried out in Section 4.1 need to be updated but

grosso modo the optimal choice for the time discretization parameter n as a function of the MC size M is

$$M \propto n^R.$$

• *The continuous Euler scheme:* The conjecture is simply to assume that the expansion $(\mathcal{E}_R^V)$ now holds for a vector space V of functionals F (with linear growth with respect to the sup-norm). The increase of the complexity induced by the Brownian bridge method is difficult to quantize: it amounts to computing $\log(U_k)$ and the inverse distribution functions $F_{x,y}^{-1}$ and $G_{x,y}^{-1}$.

The second difficulty is that simulating (the extrema of) some of continuous Euler schemes using the Brownian bridge in a consistent way is not straightforward at all. However, one can reasonably expect that using independent Brownian bridges "relying" on stepwise constant Euler schemes with consistent Brownian increments will have a small impact on the global variance (although slightly increasing it).

To illustrate and compare these approaches we carried some numerical tests on partial lookback and barrier options in the Black-Scholes model presented in the previous section.

▷ PARTIAL LOOKBACK OPTIONS: The partial lookback Call option is defined by its payoff functional

$$F(x) = e^{-rT} \left(x(T) - \lambda \min_{s \in [0,T]} x(s) \right)_+, \quad x \in \mathcal{C}([0,T], \mathbb{R}),$$

where $\lambda > 0$ (if $\lambda \leq 1$, the $(\,.\,)_+$ can be dropped). The premium

$$\text{Call}_0^{Lkb} = e^{-rT} \mathbb{E}((X_T - \lambda \min_{t \in [0,T]} X_t)_+)$$

is given by

$$\text{Call}_0^{Lkb} = X_0 \text{Call}^{BS}(1, \lambda, \sigma, r, T) + \lambda \frac{\sigma^2}{2r} X_0 \text{Put}^{BS}\left(\lambda^{\frac{2r}{\sigma^2}}, 1, \frac{2r}{\sigma}, r, T\right).$$

We took the same values for the B-S parameters as in the former section and set the coefficient λ at $\lambda = 1.1$. For this set of parameters $\text{Call}_0^{Lkb} = 57.475$.

As concerns the MC simulation size, we still set $M = 10^6$. We compared the following three methods for every choice of n:

– A 3-step R-R extrapolation ($R = 3$) of the stepwise constant Euler scheme (for which a $O(n^{-\frac{3}{2}})$-rate can be expected from the conjecture).

– A 3-step R-R extrapolation ($R = 3$) based on the continuous Euler scheme(Brownian bridge method) for which a $O(\frac{1}{n^3})$-rate can be conjectured (see [9]).

– A continuous Euler scheme (Brownian bridge method) of equivalent complexity *i.e.* with discretization parameter $6n$ for which a $O(\frac{1}{n})$-rate can be expected (see [9]).

The three procedures have the same complexity if one neglects the cost of the bridge simulation with respect to that of the diffusion coefficients (note this is very conservative in favour of "bridged schemes").

We do not reproduce the results obtained for the standard stepwise constant Euler scheme which are clearly out of the game (as already emphasized in [9]). In Figure 4, the abscissas represent the size of Euler scheme with equivalent complexity (*i.e.* $6n$, $n = 2, 4, 6, 8, 10$). Figure 4(*a*) (left) shows that both 3-step R-R extrapolation methods converge significantly faster than the "bridged" Euler scheme with equivalent complexity in this high volatility framework. The standard deviations depicted in Figure 4(*a*) (right) show that the 3-step R-R extrapolation of the Brownian bridge is controlled even for small values of n. This is not the case with the 3-step R-R extrapolation method of the stepwise constant Euler scheme. Other simulations – not reproduced here – show this is already true for the standard R-R extrapolation and the bridged Euler scheme. In any case the multistep R-R extrapolation with $R = 3$ significantly outperforms the bridged Euler scheme.

When $M = 10^8$, one verifies (see Figure 4(*b*)) that the time discretization error of the 3-step R-R extrapolation vanishes like for the partial lookback option. In fact for $n = 10$ the 3-step bridged Euler scheme yields a premium equal to 57.480 which corresponds to less than half a cent error, *i.e.* 0.05 % accuracy! This result being obtained without any control variate variable.

The R-R extrapolation of the standard Euler scheme also provides excellent results. In fact it seems difficult to discriminate them with those obtained with the bridged schemes, which is slightly unexpected if one think about the natural conjecture about the time discretization error expansion.

As a theoretical conclusion, these results strongly support both conjectures about the existence of expansion for the weak error in the $(n^{-p/2})_{p \geq 1}$ and $(n^{-p})_{p \geq 1}$ scales respectively.

$\triangleright$ Up & out Call option: Let $0 \leq K \leq L$. The Up-and-Out Call option with strike K and barrier L is defined by its payoff functional

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

It is again classical background, that in a B-S model

$$\mathrm{Call}^{U \& O}(X_0, r, \sigma, T) = \mathrm{Call}^{BS}(X_0, K, r, \sigma, T) - \mathrm{Call}^{BS}(X_0, L, r, \sigma, T) - e^{-rT}(L-K)\Phi(d^-(L))$$
$$- \left(\frac{L}{X_0} \right)^{1+\mu} \left(\mathrm{Call}^{BS}(X_0, K', r, \sigma, T) - \mathrm{Call}^{BS}(X_0, L', r, \sigma, T) - e^{-rT}(L'-K')\Phi(d^-(L')) \right)$$

with

$$K' = K \left(\frac{X_0}{L} \right)^2, \ L' = L \left(\frac{X_0}{L} \right)^2, \ d^-(L) = \frac{\log(X_0/L) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \ \text{and} \ \Phi(x) := \int_{-\infty}^x e^{-\frac{\xi^2}{2}} \frac{d\xi}{\sqrt{2\pi}}$$

and $\mu = \dfrac{2r}{\sigma^2}$.

We took again the same values for the B-S parameters as for the vanilla call. We set the barrier value at $L = 300$. For this set of parameters $C_0^{UO} = 8.54$. We tested the same three schemes. The numerical results are depicted in Figure 5.

The conclusion (see Figure 5(*a*) (left)) is that, at this very high level of volatility, when $M = 10^6$ (which is a standard size given the high volatility setting) the (quasi-)consistent
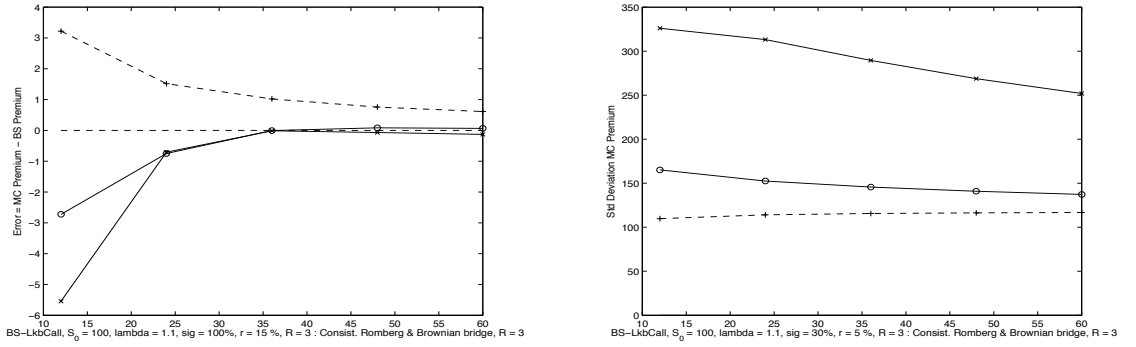
(a)
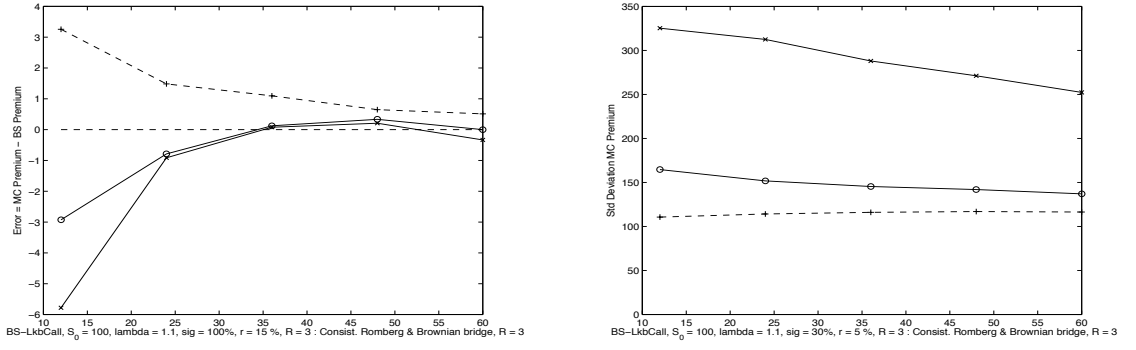
(b)

Figure 4: *B-S* Euro Partial Lookback Call option. (*a*) $M = 10^6$. *R-R extrapolation (R=3) of the Euler scheme with Brownian bridge: o—o—o. Consistent R-R extrapolation (R=3): ×—×—×. Euler scheme with Brownian bridge with equivalent complexity: +−−+−−+. $X_0 = 100$, $\sigma = 100\%$, $r = 15\%$, $\lambda = 1.1$. Abscissas: $6n$, $n = 2, 4, 6, 8, 10$.* Left: *Premia*. Right: *Standard Deviations*. (*b*) *Idem with $M = 10^8$.*

3-step R-R extrapolation with Brownian bridge clearly outperforms the continuous Euler scheme (Brownian bridge) of equivalent complexity while the 3-step R-R extrapolation based on the stepwise constant Euler schemes with consistent Brownian increments is not competitive at all: it suffers from both a too high variance (see Figure 5(a) (right)) for the considered sizes of the Monte Carlo simulation and from its too slow rate of convergence in time.

When $M = 10^8$ (see Figure 5(b) (left)), one verifies again that the time discretization error of the 3-step R-R extrapolation almost vanishes like for the partial lookback option. This no longer the case with the 3-step R-R extrapolation of stepwise constant Euler schemes. It seems clear that the discretization time error is more prominent for the barrier option: thus with $n = 10$, the relative error is $\frac{9.09-8.54}{8.54} \approx 6.5\%$ by this first R-R extrapolation whereas, the 3-step R-R method based on the quasi-consistent "bridged" method yields a an approximate premium of 8.58 corresponding to a relative error of $\frac{8.58-8.54}{8.54} \approx 0.4\%$. These specific results (obtained without any control variate) are representative of the global behaviour of the methods as emphasized by Figure 5(b)(left).

6   Conclusion

The multi-step R-R extrapolation method with consistent Brownian increments provides an efficient method to evaluate expectations of functionals of diffusions having a very high diffusion coefficient using a simple Monte Carlo simulation based on stepwise constant or continuous Euler schemes of reasonable size (in terms of discretization). This is made possible by the control of both variance and complexity.

However the asymptotic variance control may have not produced its effect when the time discretization parameter n is too small. Then it could be useful to explore some on-line variance reduction method: the idea would be to use some stochastic approximation methods as introduced in [1] to specify directly the optimal variance structure of the Euler schemes involved in the extrapolation for a given value of n.

References

[1] AROUNA, B (2004) Adaptative Monte Carlo method, a variance reduction technique. *Monte Carlo Methods Appl.* , **10**(1), 1–24.

[2] V. BALLY, D. TALAY (1996): The distribution of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function, *Probab. Theory Related Fields*, **104**(1), 43-60 (1996).

[3] V. BALLY, D. TALAY (1996): The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density, *Monte Carlo Methods Appl.*, **2**(2), 93-128.

[4] A. BESKOS, O. PAPASPILIOPOULOS, G.O. ROBERTS (2004): Retrospective exact simulation of diffusion sample paths with applications, *Working paper of Lancaster University*, http://www.maths.lancs.ac.uk/~papspil/research.html.
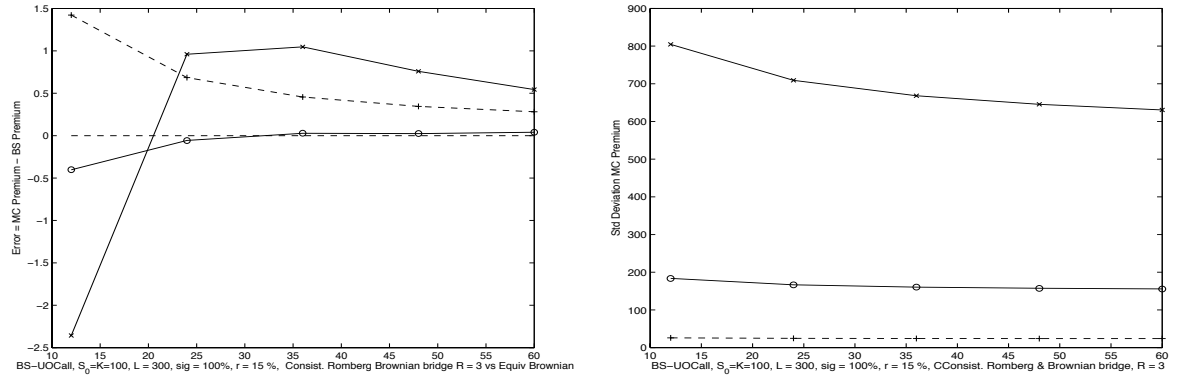
(a)
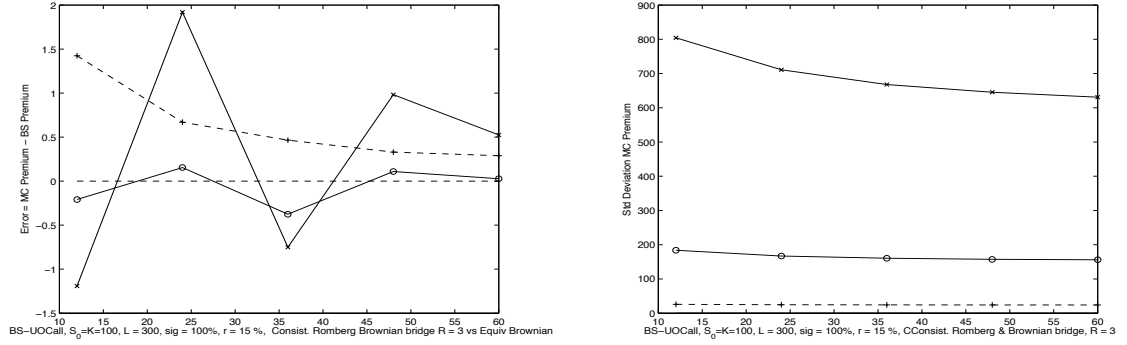
(b)

Figure 5: B-S Euro up-&-out Call option. (a) $M = 10^6$. *R-R extrapolation ($R=3$) of the Euler scheme with Brownian bridge:* o——o——o. *Consistent R-R extrapolation ($R=3$):* —×—×—×—. *Euler scheme with Brownian bridge and equivalent complexity:* + − − + − − +. $X_0 = K = 100$, $L = 300$, $\sigma = 100\%$, $r = 15\%$. *Abscissas:* $6n$, $n = 2, 4, 6, 8, 10$. *Left: Premia. Right: Standard Deviations.* (b) *Idem with* $M = 10^8$.

[5] A. BESKOS, G.O. ROBERTS (2005): Exact simulation of diffusions, *Ann. appl. Prob.*, **15**(4), 2422-2444.

[6] E. CLÉMENT, A. KOHATSU-HIGA, D. LAMBERTON (2006): A duality approach for the weak approximation of Stochastic Differential Equations, *Annals of Applied Probability*, **16**(3), 449-471.

[7] P. GLASSERMAN (2003): Monte Carlo Methods in Financial Engineering, Springer-Verlag, New York, 596p.

[8] E. GOBET, S. MENOZZI (2004): Exact approximation rate of killed hypo-elliptic diffusions using the discrete Euler scheme, *Stochastic Process. Appl.*, **112**(2), 201-223.

[9] E. GOBET (2000): Weak approximation of killed diffusion using Euler schemes, *Stoch. Proc. and their Appl.*, **87**, 167-197.

[10] E. GOBET, R. MUNOS (2005): Sensitivity analysis using Itô-Malliavin calculus and martingales. Application to optimal control, *SIAM J. Control Optim.*, **43**(5), 1676-1713.

[11] J. GUYON (2006): Euler scheme and tempered distributions, *Stochastic Processes and their Applications*, **116**(6), 877-904.

[12] I. KARATZAS, S.E. SCHREVE (1998): *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Springer-Verlag, New York, 470p.

[13] A. KOHATSU-HIGA, R. PETTERSSON (2002): Variance reduction methods for simulation of densities on Wiener space, *SIAM J. Numer. Anal.*, **40**(2), 431-450.

[14] B. LAPEYRE, E. TEMAM (2001): Competitive Monte Carlo methods for the pricing of Asian Options, *J. of Computational Finance*, **5**(1), 39-59.

[15] P. SEUMEN-TONOU (1997): *Méthodes numériques probabilistes pour la résolution d'équations du transport et pour l'évaluation d'options exotiques*, Thèse de l'Université de Provence (Marseille, France), 116p.

[16] D. TALAY, L. TUBARO (1990): Expansion of the global error for numerical schemes solving stochastic differential equations, *Stoch. Anal. Appl.*, **8**, 94-120.

Annex

We will show that in some situations the equality $X_T^{(1)} = X_T^{(2)}$ *a.s.* may imply that $W^{(1)} = W^{(2)}$ $\mathbb{P}$-*a.s.* as processes defined on $[0, T]$. Assume that $(W^{(1)}, W^{(2)})$ is a $\mathbb{R}^{2q}$-dimensional Brownian motion with respect to its (augmented) natural filtration denoted $(\mathcal{F}_t^{1,2})_{t \in [0,T]}$, both marginals $W^{(1)}$ and $W^{(2)}$ being standard d-dimensional Brownian motions. Let $R_W := \left[\mathbb{E}(W_1^{(1)} W_1^{(2)*}) \right] \in \mathcal{M}(q \times q)$ denote the correlation matrix of $W_1^{(1)}$ and $W_1^{(2)}$.

Proposition 6.5 *Let $F_{b,\sigma} : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ denote the function defined by*

$$F_{b,\sigma}(t, x_1, x_2) = (x_1 - x_2)^* (b(t, x_1) - b(t, x_2)) + \frac{1}{2} \sum_{i=1}^d \left(|\sigma_{i\cdot}(t, x_1)| - |\sigma_{i\cdot}(t, x_2)| \right)^2.$$

Assume that, for every $t \in [0, T]$,

$$\forall x_1, x_2 \in \mathbb{R}^d, \qquad x_1 \neq x_2 \implies F_{b,\sigma}(t, x_1, x_2) > 0. \tag{6.15}$$

Furthermore, assume

$$\exists\, i \in \{1, \ldots, d\}, \quad \forall\, j \in \{1, \ldots, q\}, \qquad \int_0^T \mathbb{E}\left(\sigma_{ij}^2(t, X_t)\right) dt > 0. \qquad (6.16)$$

If $X_T^{(1)} = X_T^{(2)}$ $\mathbb{P}$-a.s., then $W^{(1)} = W^{(2)}$ (and $X^{(1)} = X^{(2)}$) $\mathbb{P}$-a.s..

Remark. Assumption (6.15) is always satisfied if b is increasing *i.e.*

$$\forall\, x_1,\, x_2 \in \mathbb{R}^d, \qquad x_1 \neq x_2 \Longrightarrow (x_1 - x_2)^*(b(t, x_1) - b(t, x_2)) > 0.$$

Proof. (a) It follows from Itô's formula that

$$|X_T^{(1)} - X_T^{(2)}|^2 \;\; = \;\; \int_0^T \Phi(t, X_t^{(1)}, X_t^{(2)}) dt + M_T$$

where

$$\Phi(t, x_1, x_2) \;\; = \;\; 2(x_1 - x_2)^* \left(b(t, x_1) - b(t, x_2)\right) + \mathrm{Tr}(\sigma\sigma^*(t, x_1)) + \mathrm{Tr}(\sigma\sigma^*(t, x_2))$$
$$-2\sum_{i=1}^d \sigma_{i.}(t, x_1)^* R_W \sigma_{i.}(t, x_2)$$

and $M_t = 2\int_0^t (X_s^{(1)} - X_s^{(2)})(\sigma(t, X_s^{(1)}) dW_s^{(1)} - \sigma(t, X_s^{(2)}) dW_s^{(2)})$ is an $\mathcal{F}_t$-local martingale null at zero. In fact it is a true martingale since all the coefficients have linear growth and $\sup_{t \in [0,T]} |X_t^{(1)}|$ lies in every $L^p(\mathbb{P})$, $p > 0$. One checks from the definition of R_W that, for every $u, v \in \mathbb{R}^q$, $u^* R_W v \leq |u|\,|v|$. Consequently, for every $t \in [0, T]$, every $x_1, x_2 \in \mathbb{R}^d$,

$$\Phi(t, x_1, x_2) \geq 2F_{b,\sigma}(t, x_1, x_2) \geq 0.$$

If $X_T^{(1)} = X_T^{(2)}$ $\mathbb{P}$-a.s., then

$$M_T = -\int_0^T \Phi(t, X_t^{(1)}, X_t^{(2)}) dt \leq 0 \qquad \mathbb{P}\text{-a.s..}$$

Hence $M_T = 0$ $\mathbb{P}$-a.s. since $\mathbb{E}M_T = 0$. In turn this implies that $\int_0^T \Phi(t, X_t^{(1)}, X_t^{(2)}) dt = 0$ $\mathbb{P}$-a.s.. The continuous function Φ being non negative and $(X_t^{(1)}, X_t^{(2)})$ being pathwise continuous,

$$\mathbb{P}\text{-a.s.} \qquad \forall\, t \in [0, T], \qquad \Phi(t, X_t^{(1)}, X_t^{(2)}) = F_{b,\sigma}(t, X_t^{(1)}, X_t^{(2)}) = 0.$$

Consequently, (6.15)
$$\mathbb{P}\text{-a.s.} \qquad \forall\, t \in [0, T], \qquad X_t^{(1)} = X_t^{(2)}$$

Elementary computations show that

$$\Phi(t, \xi, \xi) = 2\sum_{i=1}^d \sigma_{i.}(t, \xi)^* (I_q - R_W)\sigma_{i.}(t, \xi).$$

27

The symmetric matrix $I_q - R_W$ being nonnegative,

$$\forall\, i \in \{1, \ldots, d\}, \quad \sigma_{i.}(t, X_t^{(1)})^*(I_q - R_W)\sigma_{i.}(t, X_t^{(1)}) = 0 \quad i.e. \quad \sigma_{i.}(t, X_t^{(1)}) \in \mathrm{Ker}(I_q - R_W).$$

If $I_q \neq R_W$, this (nonnegative symmetric) matrix has at least one positive eigenvalue $\lambda > 0$. Let $u \in \mathbb{R}^q \setminus \{0\}$ be an eigenvector associated to λ. Then $u \perp \mathrm{Ker}(I_q - R_W)$ so that $u^* \sigma_{i.}(t, X_t^{(1)}) = 0$ $\mathbb{P}$-a.s.. This cannot be satisfied by an index i satisfying (6.16). Hence $R_W = I_q$ which in turn implies that $W^{(1)} = W^{(2)}$ since $(W^{(1)}, W^{(2)})$ is a Gaussian centered process. $\quad\diamond$