Central Limit Theorem for the Multilevel Monte Carlo Euler Method

Mohamed Ben Alaya[1] & Ahmed Kebaier*[1]

[1] Université Paris 13
Sorbonne Paris Cité, LAGA, CNRS (UMR 7539)
99, av. J.B. Clément 93430 Villetaneuse, France
mba@math.univ-paris13.fr kebaier@math.univ-paris13.fr

September 16, 2013

Abstract

This paper focuses on studying the multilevel Monte Carlo method recently introduced by Giles [8] which is significantly more efficient than the classical Monte Carlo one. Our aim is to prove a central limit theorem of Lindeberg Feller type for the multilevel Monte Carlo method associated with the Euler discretization scheme. To do so, we prove first a stable law convergence theorem, in the spirit of Jacod and Protter [16], for the Euler scheme error on two consecutive levels of the algorithm. This leads to an accurate description of the optimal choice of parameters and to an explicit characterization of the limiting variance in the central limit theorem of the algorithm. A complexity of the multilevel Monte Carlo algorithm is carried out.

AMS 2000 Mathematics Subject Classification. 60F05, 62F12, 65C05, 60H35.
Key Words and Phrases. Central limit theorem, Multilevel Monte Carlo methods, Euler scheme, finance.

1 Introduction

In many applications, in particular in the pricing of financial securities, we are interested in the effective computation by Monte Carlo methods of the quantity $\mathbb{E}f(X_T)$, where $X := (X_t)_{0\leq t\leq T}$ is a diffusion process and f a given function. The Monte Carlo Euler method consists of two steps. First, approximate the diffusion process $(X_t)_{0\leq t\leq T}$ by the Euler scheme $(X_t^n)_{0\leq t\leq T}$ with time step T/n. Then, approximate $\mathbb{E}\, f\,(X_T^n)$ by $\frac{1}{N}\sum_{i=1}^{N} f(X_{T,i}^n)$, where $f(X_{T,i}^n)_{1\leq i\leq N}$ is

*This research benefited from the support of the chair "Risques Financiers", Fondation du Risque.

a sample of N independent copies of $f(X_T^n)$. This approximation is affected respectively by a discretization error and a statistical error

$$\varepsilon_n := \mathbb{E}\left(f(X_T^n) - f(X_T)\right) \text{ and } \frac{1}{N}\sum_{i=1}^{N} f(X_{T,i}^n) - \mathbb{E}f(X_T^n).$$

On one hand, Talay and Tubaro [21] prove that if f is sufficiently smooth, then $\varepsilon_n \sim c/n$ with c a given constant and in a more general context, Kebaier [17] proves that the rate of convergence of the discretization error ε_n can be $1/n^\alpha$ for all values of $\alpha \in [1/2, 1]$ (see e.g. Kloeden and Platen [18] for more details on discretization schemes). On the other hand, the statistical error is controlled by the central limit theorem with order $1/\sqrt{N}$. Further, the optimal choice of the sample size N in the classical Monte Carlo method mainly depends on the order of the discretization error. More precisely, it turns out that for $\varepsilon_n = 1/n^\alpha$ the optimal choice of N is $n^{2\alpha}$. This leads to a total complexity in the Monte Carlo method of order $C_{MC} = n^{2\alpha+1}$ (see Duffie and Glynn [5] for related results). Let us recall that the complexity of an algorithm is proportional to the maximum number of basic computations performed by this one. Hence, expressing this complexity in terms of the discretization error ε_n we get $C_{MC} = \varepsilon_n^{-2-1/\alpha}$.

In order to improve the performance of this method, Kebaier introduced a two-level Monte Carlo method [17] (called the statistical Romberg method) reducing the complexity C_{MC} while maintaining the convergence of the algorithm. This method uses two Euler schemes with time steps T/n and T/n^β, $\beta \in (0, 1)$ and approximates $\mathbb{E}\, f(X_T)$ by

$$\frac{1}{N_1}\sum_{i=1}^{N_1} f(\hat{X}_{T,i}^{n^\beta}) + \frac{1}{N_2}\sum_{i=1}^{N_2} f(X_{T,i}^n) - f(X_{T,i}^{n^\beta}),$$

where $\hat{X}_T^{n^\beta}$ is a second Euler scheme with time step T/n^β and such that the Brownian paths used for X_T^n and $X_T^{n^\beta}$ has to be independent of the Brownian paths used to simulate $\hat{X}_T^{n^\beta}$. It turns out that for a given discretization error $\varepsilon_n = 1/n^\alpha$ ($\alpha \in [1/2, 1]$), the optimal choice is obtained for $\beta = 1/2$, $N_1 = n^{2\alpha}$ and $N_2 = n^{2\alpha-(1/2)}$. With this choice, the complexity of the statistical Romberg method is of order $C_{SR} = n^{2\alpha+(1/2)} = \varepsilon_n^{-2-1/(2\alpha)}$ which is lower than the classical complexity in the Monte Carlo method.

More recently, Giles [8] generalized the statistical Romberg method of Kebaier [17] and proposed the multilevel Monte Carlo algorithm, in a similar approach to Heinrich's multilevel method for parametric integration [12] (see also Creutzig, Dereich, Müller-Gronbach and Ritter [3], Dereich [4], Giles [7], Giles, Higham and Mao [9], Giles and Szpruch [10], Heinrich [11], Heinrich and Sindambiwe [13] and Hutzenthaler, Jentzen and Kloeden [14] for related results). The multilevel Monte Carlo method uses information from a sequence of computations with decreasing step sizes and approximates the quantity $\mathbb{E}f(X_T)$ by

$$Q_n = \frac{1}{N_0}\sum_{k=1}^{N_0} f(X_{T,k}^1) + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{k=1}^{N_\ell}\left(f(X_{T,k}^{\ell,m^\ell}) - f(X_{T,k}^{\ell,m^{\ell-1}})\right),\ m \in \mathbb{N}\setminus\{0,1\},$$

where the fine discretization step is equal to T/n thereby $L = \frac{\log n}{\log m}$. For $\ell \in \{1, \cdots, L\}$, processes $(X_{t,k}^{\ell,m^\ell}, X_{t,k}^{\ell,m^{\ell-1}})_{0\le t\le T}$, $k \in \{1, \cdots, N_\ell\}$, are independent copies of $(X_t^{\ell,m^\ell}, X_t^{\ell,m^{\ell-1}})_{0\le t\le T}$

whose components denote the Euler schemes with time steps $m^{-\ell}T$ and $m^{-(\ell-1)}T$. However, for fixed ℓ, the simulation of $(X_t^{\ell,m^\ell})_{0\leq t\leq T}$ and $(X_t^{\ell,m^{\ell-1}})_{0\leq t\leq T}$ has to be based on the same Brownian path. Concerning the first empirical mean, processes $(X_{t,k}^1)_{0\leq t\leq T}$, $k\in\{1,\cdots,N_0\}$, are independent copies of $(X_t^1)_{0\leq t\leq T}$ which denotes the Euler scheme with time step T. Here, it is important to point out that all these $L+1$ Monte Carlo estimators have to be based on different independent samples. Due to the above independence assumption for the paths, the variance of the multilevel estimator is given by

$$\sigma^2 := Var(Q_n) = N_0^{-1}Var(f(X_T^1)) + \sum_{\ell=1}^{L} N_\ell^{-1}\sigma_\ell^2,$$

where $\sigma_\ell^2 = Var\left(f(X_T^{\ell,m^\ell}) - f(X_T^{\ell,m^{\ell-1}})\right)$. Assuming that the diffusion coefficients of X and the function f are Lipschitz continuous then it is easy to check, using properties of the Euler scheme, that

$$\sigma^2 \leq c_2 \sum_{\ell=0}^{L} N_\ell^{-1} m^{-\ell}$$

for some positive constant c_2 (see Proposition 1 for more details). Giles [8] uses this computation in order to find the optimal choice of the multilevel Monte Carlo parameters. More precisely, to obtain a desired root mean squared error (RMSE), say of order $1/n^\alpha$, for the multilevel estimator, Giles [8] uses the above computation on σ^2 to minimize the total complexity of the algorithm. It turns out that the optimal choice is obtained for (see Theorem 3.1 of [8])

$$N_\ell = 2c_2 n^{2\alpha}\left(\frac{\log n}{\log m}+1\right)\frac{T}{m^\ell}, \quad \text{for } \ell\in\{0,\cdots,L\} \quad \text{and} \quad L=\frac{\log n}{\log m}. \tag{1}$$

Hence, for an error $\varepsilon_n = 1/n^\alpha$, this optimal choice leads to a complexity for the multilevel Monte Carlo Euler method proportional to $n^{2\alpha}(\log n)^2 = \varepsilon_n^{-2}(\log \varepsilon_n)^2$. Interesting numerical tests, comparing three methods (crude Monte Carlo, statistical Romberg and the multilevel Monte Carlo), were processed in Korn, Korn and Kroisandt [19].

In the present paper, we focus on central limit theorems for the inferred error; a question which has not been addressed in previous research. To do so, we use techniques adapted to this setting, based on a central limit theorem for triangular array (see Theorem 2) together with Toeplitz lemma. It is worth to note that our approach improves techniques developed by Kebaier [17] in his study of the statistical Romberg method (see Remark 2 for more details). Hence, our main result is a Lindeberg Feller central limit theorem for the multilevel Monte Carlo Euler algorithm (see Theorem 4). Further, this allows us to prove a Berry-Essen type bound on our central limit theorem.

In order to show this central limit theorem, we first prove a stable law convergence theorem, for the Euler scheme error on two consecutive levels $m^{\ell-1}$ and m^ℓ, of the type obtained in Jacod and Protter [16]. Indeed, we prove the following functional result (see Theorem 3)

$$\sqrt{\frac{m^\ell}{(m-1)T}}(X^{\ell,m^\ell} - X^{\ell,m^{\ell-1}}) \Rightarrow^{stably} U, \quad \text{as} \quad \ell\to\infty,$$

where U is the same limit process given in Theorem 3.2 of Jacod and Protter [16]. Our result uses standard tools developed in their paper but it can not be deduced without a specific and laborious study. Further, their result, namely

$$\sqrt{\frac{m^\ell}{T}}(X^{\ell,m^\ell} - X) \Rightarrow^{stably} U, \quad \text{as} \quad \ell \to \infty,$$

is neither sufficient nor appropriate to prove our Theorem 4, since the multilevel Monte Carlo Euler method involves the error process $X^{\ell,m^\ell} - X^{\ell,m^{\ell-1}}$ rather than $X^{\ell,m^\ell} - X$.

Thanks to Theorem 4 we obtain a precise description for the choice of the parameters to run the multilevel Monte Carlo Euler method. Afterward, by a complexity analysis we obtain the optimal choice for the multilevel Monte Carlo Euler method. It turns out that for a total error of order $\varepsilon_n = 1/n^\alpha$ the optimal parameters are given by

$$N_\ell = \frac{(m-1)T}{m^\ell \log m} n^{2\alpha} \log n, \quad \text{for } \ell \in \{0, \cdots, L\} \quad \text{and} \quad L = \frac{\log n}{\log m}. \tag{2}$$

This leads us to a complexity proportional to $n^{2\alpha}(\log n)^2 = \varepsilon_n^{-2}(\log \varepsilon_n)^2$ which is the same order obtained by Giles [8]. By comparing relations (1) and (2), we note that our optimal sequence of sample sizes $(N_\ell)_{0\le\ell\le L}$ does not depend on any given constant, since our approach is based on proving a central limit theorem and not on obtaining an upper bound for the variance of the algorithm. However, some numerical tests comparing the runtime with respect to the root mean square error, show that we are in line with the original work of Giles [8]. Nevertheless, the major advantage of our central limit theorem is that it fills the gap in the literature for the multilevel Monte Carlo Euler method and allows to construct a more accurate confidence interval compared to the one obtained using Chebyshev's inequality. All these results are stated and proved in section 3. The next section is devoted to recall some useful stochastic limit theorems and to introduce our notations.

2 General framework

2.1 Preliminaries

Let (X_n) be a sequence of random variables with values in a Polish space E defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ be an extension of $(\Omega, \mathcal{F}, \mathbb{P})$, and let X be an E-valued random variable on the extension. We say that (X_n) converges in law to X stably and write $X_n \Rightarrow^{stably} X$, if

$$\mathbb{E}(Uh(X_n)) \to \tilde{\mathbb{E}}(Uh(X))$$

for all $h : E \to \mathbb{R}$ bounded continuous and all bounded random variable U on $(\Omega, \mathcal{F})$. This convergence is obviously stronger than convergence in law that we will denote here by "$\Rightarrow$". According to section 2 of Jacod [15] and Lemma 2.1 of Jacod and Protter [16], we have the following result.

Lemma 1 *let V_n and V be defined on $(\Omega, \mathcal{F})$ with values in another metric space E'.*

$$\textit{if} \quad V_n \xrightarrow{\mathbb{P}} V, \quad X_n \Rightarrow^{stably} X \quad \textit{then} \quad (V_n, X_n) \Rightarrow^{stably} (V, X).$$

Conversely, if $(V, X_n) \Rightarrow (V, X)$ *and* V *generates the* σ*-field* $\mathcal{F}$*, we can realize this limit as* (V, X) *with* X *defined on an extension of* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $X_n \Rightarrow^{stably} X$*.*

Now, we recall a result on the convergence of stochastic integrals formulated from Theorem 2.3 in Jacod and Protter [16]. This is a simplified version but it is sufficient for our study. Let $X^n = (X^{n,i})_{1\leq i\leq d}$ be a sequence of $\mathbb{R}^d$-valued continuous semimartingales with the decomposition

$$X_t^{n,i} = X_0^{n,i} + A_t^{n,i} + M_t^{n,i}, \quad 0 \leq t \leq T$$

where, for each $n \in \mathbb{N}$ and $1 \leq i \leq d$, $A^{n,i}$ is a predictable process with finite variation, null at 0 and $M^{n,i}$ is a martingale null at 0.

Theorem 1 *Assume that the sequence* (X^n) *is such that*

$$\langle M^{n,i}\rangle_T + \int_0^T \left|dA_s^{n,i}\right|$$

is tight. Let H^n *and* H *be a sequence of adapted, right-continuous and left-hand limited processes all defined on the same filtered probability space. If* $(H^n, X^n) \Rightarrow (H, X)$ *then* X *is a semimartingale with respect to the filtration generated by the limit process* (H, X)*, and we have* $(H^n, X^n, \int H^n dX^n) \Rightarrow (H, X, \int H dX)$*.*

We recall also the following Lindeberg Feller central limit theorem that will be used in the sequel (see for instance Theorem 7.2 and 7.3 in [1]).

Theorem 2 (central limit theorem for triangular array) *Let* $(k_n)_{n\in\mathbb{N}}$ *be a sequence such that* $k_n \longrightarrow \infty$ *as* $n \longrightarrow \infty$*. For each* n*, let* $X_{n,1}, \cdots, X_{n,k_n}$ *be* k_n *independent random variables with finite variance such that* $\mathbb{E}(X_{n,k}) = 0$ *for all* $k \in \{1, \cdots, k_n\}$*. Suppose that the following conditions hold.*

A1. $\lim_{n\to\infty} \sum_{k=1}^{k_n} \mathbb{E}|X_{n,k}|^2 = \sigma^2, \sigma > 0.$

A2. Lindeberg's condition: for all $\varepsilon > 0$*,* $\lim_{n\to\infty} \sum_{k=1}^{k_n} \mathbb{E}\left(|X_{n,k}|^2 1_{\{|X_{n,k}|>\varepsilon\}}\right) = 0$*. Then*

$$\sum_{k=1}^{k_n} X_{n,k} \Rightarrow \mathcal{N}(0, \sigma^2) \quad \text{as} \quad n \to \infty.$$

Moreover, if the $X_{n,k}$ *have moments of order* $p > 2$*, then the Lindeberg's condition can be obtained by the following one*

A3 Lyapunov's condition: $\lim_{n\to\infty} \sum_{k=1}^{k_n} \mathbb{E}|X_{n,k}|^p = 0.$

2.2 The Euler scheme

Let $X := (X_t)_{0\leq t\leq T}$ be the process with values in $\mathbb{R}^d$, solution to

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d \tag{3}$$

where $W = (W^1, \ldots, W^q)$ is a q-dimensional Brownian motion on some given filtered probability space $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ with $(\mathcal{F}_t)_{t\geq 0}$ is the standard Brownian filtration, b and σ are respectively $\mathbb{R}^d$ and $\mathbb{R}^{d\times q}$ valued functions. We consider the continuous Euler approximation X^n with step $\delta = T/n$ given by:

$$dX_t^n = b(X_{\eta_n(t)})dt + \sigma(X_{\eta_n(t)})dW_t, \quad \eta_n(t) = [t/\delta]\delta.$$

It is well known that under the global Lipschitz condition

$$(\mathcal{H}_{b,\sigma}) \quad \exists\ C_T > 0, \text{ such that, } |b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq C_T|y - x|, \quad x, y \in \mathbb{R}^d,$$

the Euler scheme satisfies the following property (see e.g. Bouleau and Lépingle [2])

$$\mathcal{P})\ \ \forall p \geq 1, \quad X, X^n \in L^p \quad \text{and} \quad \mathbb{E}\left[\sup_{0\leq t\leq T}|X_t - X_t^n|^p\right] \leq \frac{K_p(T)}{n^{p/2}}, \quad \text{with } K_p(T) > 0.$$

Note that according to Theorem 3.1 of Jacod and Protter [16], under the weaker condition

$$(\tilde{\mathcal{H}}_{b,\sigma}) \qquad b \text{ and } \sigma \text{ are locally Lipschitz with linear growth,}$$

we have only the uniform convergence in probability, namely the property

$$\tilde{\mathcal{P}}) \qquad \sup_{0\leq t\leq T}|X_t - X_t^n| \xrightarrow{\mathbb{P}} 0.$$

Following the notation of Jacod and Protter [16], we rewrite diffusion (3) as follows

$$dX_t = \varphi(X_t)dY_t = \sum_{j=0}^{q}\varphi_j(X_t)dY_t^j$$

where φ_j is the j-th column of the matrix σ, for $1 \leq j \leq q$, $\varphi_0 = b$ and $Y_t := (t, W_t^1, \cdots, W_t^q)'$. Then, the continuous Euler approximation X^n with time step $\delta = T/n$ becomes

$$dX_t^n = \varphi(X_{\eta_n(t)}^n)dY_t = \sum_{j=0}^{q}\varphi_j(X_{\eta_n(t)}^n)dY_t^j, \quad \eta_n(t) = [t/\delta]\delta. \tag{4}$$

3 The Multilevel Monte Carlo Euler method

Let $(X_t^{m^\ell})_{0\leq t\leq T}$ denotes the Euler scheme with time step $m^{-\ell}T$ for $\ell \in \{0, \cdots, L\}$, where $L = \log n/\log m$. Noting that

$$\mathbb{E}f(X_T^n) = \mathbb{E}f(X_T^1) + \sum_{\ell=1}^{L}\mathbb{E}\left(f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})\right), \tag{5}$$

the multilevel method is to estimate independently by the Monte Carlo method each of the expectations on the right-hand side of the above relation. Hence, we approximate $\mathbb{E}f(X_T^n)$ by

$$Q_n = \frac{1}{N_0}\sum_{k=1}^{N_0} f(X_{T,k}^1) + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{k=1}^{N_\ell}\Big(f(X_{T,k}^{\ell,m^\ell}) - f(X_{T,k}^{\ell,m^{\ell-1}})\Big). \tag{6}$$

Here, it is important to point out that all these $L+1$ Monte Carlo estimators have to be based on different, independent samples. More precisely, for each $\ell \in \{1,\cdots,L\}$ the samples $(X_{T,k}^{\ell,m^\ell}, X_{T,k}^{\ell,m^{\ell-1}})_{1\le k\le N_\ell}$ are independent copies of $(X_T^{\ell,m^\ell}, X_T^{\ell,m^{\ell-1}})$ whose components are the Euler schemes with time steps $m^{-\ell}T$ respectively $m^{-(\ell-1)}T$ and simulated with the same Brownian path. Concerning the first empirical mean, the samples $(X_{T,k}^1)_{1\le k\le N_0}$ are independent copies of X_T^1. The following result gives us a first description of the asymptotic behavior of the variance in the multilevel Monte Carlo Euler method.

Proposition 1 *Assume that b and σ are functions satisfying condition $(\mathcal{H}_{b,\sigma})$. For a Lipschitz continuous function $f:\mathbb{R}^d \longrightarrow \mathbb{R}$ we have*

$$Var(Q_n) = O\big(\sum_{\ell=0}^{L} N_\ell^{-1} m^{-\ell}\big). \tag{7}$$

Proof : We have

$$\begin{aligned}
Var(Q_n) &= N_0^{-1}Var\left(f(X_T^1)\right) + \sum_{\ell=1}^{L} N_\ell^{-1} Var\left(f(X_T^{\ell,m^\ell}) - f(X_T^{\ell,m^{\ell-1}})\right)\\
&\le N_0^{-1}Var\left(f(X_T^1)\right) + 2\sum_{\ell=1}^{L} N_\ell^{-1}\left(Var(f(X_T^{m^\ell}) - f(X_T)) + Var(f(X_T^{m^{\ell-1}}) - f(X_T))\right)\\
&\le N_0^{-1}Var\left(f(X_T^1)\right) + 2[f]_{lip}\sum_{\ell=1}^{L} N_\ell^{-1}\mathbb{E}\left[\sup_{0\le t\le T}\left|X_t^{m^\ell} - X_t\right|^2 + \sup_{0\le t\le T}\left|X_t^{m^{\ell-1}} - X_t\right|^2\right],
\end{aligned}$$

where $[f]_{lip} := \sup_{u\neq v}\frac{|f(u)-f(v)|}{|u-v|}$. We complete the proof by using $\mathcal{P}$) on the strong convergence of the Euler scheme.

□

Inequality (7) indicates the dependence of the variance of Q_n on the choice of the parameters $N_0,\ldots,N_L$. This variance can be smaller than the variance of $f(X_T^n)$, so that Q_n appears as a good candidate for the variance reduction.

The main result of this section is a Lindeberg Feller central limit theorem (see Theorem 4 below). In order to prove this result, we need to prove first a new stable law convergence theorem for the Euler scheme error adapted to the setting of multilevel Monte Carlo algorithm. This is crucial and is the aim of the following subsection.

3.1 Stable convergence

In what follows, we prove a stable law convergence theorem, for the Euler scheme error on two consecutive levels $m^{\ell-1}$ and m^{ℓ}, of the type obtained in Jacod and Protter [16]. Our result in Theorem 3 below is an innovative contribution on the Euler scheme error that is different and more tricky than the original work by Jacod and Protter [16] since it involves the error process $X^{\ell,m^\ell} - X^{\ell,m^{\ell-1}}$ rather than $X^{\ell,m^\ell} - X$. Note that the study of the error $X^{\ell,m^\ell} - X^{\ell,m^{\ell-1}}$ as $\ell \to \infty$ can be reduced to the study of the error $X^{mn} - X^n$ as $n \to \infty$ where X^{mn} and X^n stand for two Euler schemes with time steps $T/(mn)$ and T/n constructed on the same Brownian path.

Theorem 3 *Assume that b and σ are $\mathcal{C}^1$ with linear growth then the following result holds.*

$$\textit{For all } m \in \mathbb{N} \setminus \{0,1\}, \quad \sqrt{\frac{mn}{(m-1)T}}(X^{mn} - X^n) \Rightarrow^{\textit{stably}} U, \qquad \textit{as } n \to \infty,$$

with $(U_t)_{0\le t\le T}$ the d-dimensional process satisfying

$$U_t = \frac{1}{\sqrt{2}} \sum_{i,j=1}^{q} Z_t \int_0^t H_s^{i,j} dB_s^{ij}, \quad t \in [0,T], \tag{8}$$

where

$$H_s^{i,j} = (Z_s)^{-1} \dot{\varphi}_{s,j} \bar{\varphi}_{s,i}, \ \textit{with } \dot{\varphi}_{s,j} := \nabla \varphi_j(X_s) \ \textit{and } \bar{\varphi}_{s,i} := \varphi_i(X_s), \tag{9}$$

and $(Z_t)_{0\le t\le T}$ is the $\mathbb{R}^{d\times d}$ valued process solution of the linear equation

$$Z_t = I_d + \sum_{j=0}^{q} \int_0^t \dot{\varphi}_{s,j} dY_s^j Z_s, \quad t \in [0,T].$$

Here, $\nabla\varphi_j$ is a $d \times d$ matrix with $(\nabla\varphi_j)_{ik}$ is the partial derivative of φ_{ij} with respect to the k-th coordinate, and $(B^{ij})_{1\le i,j\le q}$ is a standard q^2-dimensional Brownian motion independent of W. This process is defined on an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t\ge 0}, \tilde{\mathbb{P}})$ of the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$.

Note that by letting formally m tend to infinity we recover the Jacod and Protter's result [16].

Proof: Consider the error process $U^{mn,n} = (U_t^{mn,n})_{0\le t\le T}$, defined by

$$U_t^{mn,n} := X_t^{mn} - X_t^n, \quad t \in [0,T].$$

Combining relation (4), for both processes X^{mn} and X^n, together with a Taylor expansion

$$dU_t^{mn,n} = \sum_{j=0}^{q} \dot{\varphi}_{t,j}^n (X^{mn}_{\eta_{mn}(t)} - X^n_{\eta_n(t)})\, dY_t^j,$$

where $\dot{\varphi}_{t,j}^n$ is the $d \times d$ matrix whose i-th row is the gradient of the real-valued function φ_{ij} at a point between $X^n_{\eta_n(t)}$ and $X^{mn}_{\eta_{mn}(t)}$. Therefore, the equation satisfied by U^n can be written as

$$U_t^{mn,n} = \int_0^t \sum_{j=0}^{q} \dot{\varphi}_{s,j}^n\, U_s^{mn,n} dY_s^j + G_t^{mn,n},$$

with

$$G_t^{mn,n} = \int_0^t \sum_{j=0}^{q} \dot{\varphi}_{s,j}^n (X_s^n - X_{\eta_n(s)}^n)\, dY_s^j - \int_0^t \sum_{j=0}^{q} \dot{\varphi}_{s,j}^n (X_s^{mn} - X_{\eta_{mn}(s)}^{mn})\, dY_s^j.$$

In the following, let $(Z_t^{mn,n})_{0\le t\le T}$ be the $\mathbb{R}^{d\times d}$ valued solution of

$$Z_t^{mn,n} = I_d + \int_0^t \left(\sum_{j=0}^{q} \dot{\varphi}_{s,j}^n \, dY_s^j \right) Z_s^{mn,n}.$$

Theorem 48 p.326 in [20], ensures existence of the process $((Z_t^{mn,n})^{-1})_{0\le t\le T}$ defined as the solution of

$$(Z_t^{mn,n})^{-1} = I_d + \int_0^t (Z_s^{mn,n})^{-1} \sum_{j=1}^{q} (\dot{\varphi}_{s,j}^n)^2 ds - \int_0^t (Z_s^{mn,n})^{-1} \sum_{j=0}^{q} \dot{\varphi}_{s,j}^n dY_s^j.$$

Thanks to Theorem 56 p.333 in the same reference [20], we get

$$U_t^{mn,n} = Z_t^{mn,n} \Big\{ \int_0^t (Z_s^{mn,n})^{-1} dG_s^{mn,n} - \int_0^t (Z_s^{mn,n})^{-1} \sum_{j=1}^{q} (\dot{\varphi}_{s,j}^n)^2 (X_s^n - X_{\eta_n(s)}^n)\, ds + \int_0^t (Z_s^{mn,n})^{-1} \sum_{j=1}^{q} (\dot{\varphi}_{s,j}^n)^2 (X_s^{mn} - X_{\eta_{mn}(s)}^{mn})\, ds \Big\}.$$

Since the increments of the Euler scheme satisfy

$$X_s^n - X_{\eta_n(s)}^n = \sum_{i=0}^{q} \bar{\varphi}_{s,i}^n (Y_s^i - Y_{\eta_n(s)}^i) \quad \text{and} \quad X_s^{mn} - X_{\eta_{mn}(s)}^{mn} = \sum_{i=0}^{q} \bar{\varphi}_{s,i}^{mn} (Y_s^i - Y_{\eta_{mn}(s)}^i),$$

with $\bar{\varphi}_{s,i}^n = \varphi_i(X_{\eta_n(s)}^n)$ and $\bar{\varphi}_{s,i}^{mn} = \varphi_i(X_{\eta_{mn}(s)}^{mn})$, it is easy to check that

$$U_t^{mn,n} = \sum_{i,j=1}^{q} Z_t^{mn,n} \int_0^t H_s^{i,j,mn,n} (Y_s^i - Y_{\eta_n(s)}^i)\, dY_s^j + R_{t,1}^{mn,n} + R_{t,2}^{mn,n} - \sum_{i,j=1}^{q} Z_t^{mn,n} \int_0^t \tilde{H}_s^{i,j,mn,n} (Y_s^i - Y_{\eta_{mn}(s)}^i)\, dY_s^j - \tilde{R}_{t,1}^{mn,n} - \tilde{R}_{t,2}^{mn,n} \quad (10)$$

with

$$R_{t,1}^{mn,n} = \sum_{i=0}^{q} Z_t^{mn,n} \int_0^t K_s^{i,mn,n} (Y_s^i - Y_{\eta_n(s)}^i)\, ds, \quad R_{t,2}^{mn,n} = \sum_{j=1}^{q} Z_t^{mn,n} \int_0^t H_s^{0,j,mn,n} (s - \eta_n(s))\, dY_s^j,$$

and

$$\tilde{R}_{t,1}^{mn,n} = \sum_{i=0}^{q} Z_t^{mn,n} \int_0^t \tilde{K}_s^{i,mn,n} (Y_s^i - Y_{\eta_{mn}(s)}^i)\, ds, \quad \tilde{R}_{t,2}^{mn,n} = \sum_{j=1}^{q} Z_t^{mn,n} \int_0^t \tilde{H}_s^{0,j,mn,n} (s - \eta_{mn}(s))\, dY_s^j.$$

where, for $(i,j) \in \{0,\cdots,q\} \times \{1,\cdots,q\}$,

$$K_s^{i,mn,n} = (Z_s^{mn,n})^{-1} \left(\dot{\varphi}_{s,0}^n \bar{\varphi}_{s,i}^n - \sum_{j=1}^q (\dot{\varphi}_{s,j}^n)^2 \bar{\varphi}_{s,i}^n \right), \quad H_s^{i,j,mn,n} = (Z_s^{mn,n})^{-1} \dot{\varphi}_{s,j}^n \bar{\varphi}_{s,i}^n,$$

and

$$\tilde{K}_s^{i,mn,n} = (Z_s^{mn,n})^{-1} \left(\dot{\varphi}_{s,0}^n \bar{\varphi}_{s,i}^{mn} - \sum_{j=1}^q (\dot{\varphi}_{s,j}^n)^2 \bar{\varphi}_{s,i}^{mn} \right), \quad \tilde{H}_s^{i,j,mn,n} = (Z_s^{mn,n})^{-1} \dot{\varphi}_{s,j}^n \bar{\varphi}_{s,i}^{mn}.$$

Now, let us introduce

$$Z_t = I_d + \int_0^t \sum_{j=0}^q \left(\dot{\varphi}_{s,j} \, dY_s^j \right) Z_s, \qquad \text{with } \dot{\varphi}_{t,j} = \nabla \varphi_j(X_t).$$

Moreover, $((Z_t)^{-1})_{0 \le t \le T}$ exists and satisfies the following explicit linear stochastic differential equation

$$(Z_t)^{-1} = I_d + \int_0^t (Z_s)^{-1} \sum_{j=1}^q (\dot{\varphi}_{s,j})^2 ds - \int_0^t (Z_s)^{-1} \sum_{j=0}^q \dot{\varphi}_{s,j} dY_s^j.$$

Thanks to the uniform convergence in probability of the Euler scheme and according to Theorem 2.5 in Jacod and Protter [16], we have

$$\sup_{0 \le t \le T} |Z_t^{mn,n} - Z_t| \xrightarrow{\mathbb{P}} 0 \text{ and } \sup_{0 \le t \le T} \left| (Z_t^{mn,n})^{-1} - (Z_t)^{-1} \right| \xrightarrow{\mathbb{P}} 0. \tag{11}$$

Furthermore, in relation (10), one can replace respectively $H_s^{i,j,mn,n}$ and $\tilde{H}_s^{i,j,mn,n}$ by their common limit $H_s^{i,j}$ given by relation (9). So that relation (10) becomes

$$U_t^{mn,n} = \sum_{i,j=1}^q Z_t^{mn,n} \int_0^t H_s^{i,j} (Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)}) \, dY_s^j + R_t^{mn,n}, \tag{12}$$

with

$$R_t^{mn,n} = R_{t,1}^{mn,n} + R_{t,2}^{mn,n} + R_{t,3}^{mn,n} - \tilde{R}_{t,1}^{mn,n} - \tilde{R}_{t,2}^{mn,n} - \tilde{R}_{t,3}^{mn,n}$$

where $R_{t,i}^{mn,n}$ and $\tilde{R}_{t,i}^{mn,n}$, $i \in \{1,2\}$, are introduced by relation (10) and

$$\begin{aligned} R_{t,3}^{mn,n} &= \sum_{i,j=1}^q Z_t^{mn,n} \int_0^t (H_s^{i,j,mn,n} - H_s^{i,j})(Y_s^i - Y^i_{\eta_n(s)}) \, dY_s^j \\ \tilde{R}_{t,3}^{mn,n} &= \sum_{i,j=1}^q Z_t^{mn,n} \int_0^t (\tilde{H}_s^{i,j,mn,n} - H_s^{i,j})(Y_s^i - Y^i_{\eta_{mn}(s)}) \, dY_s^j, \end{aligned}$$

The remainder term process $R^{mn,n}$ vanishes with rate $\sqrt{n}$ in probability. More precisely, we have the following convergence result.

Lemma 2 *The rest term introduced in relation (12) is such that* $\sup_{0\leq t\leq T}\left|\sqrt{n}R_t^{mn,n}\right|$ *converges to zero in probability as* n *tends to infinity.*

For the reader's convenience, the proof of this lemma is postponed to the end of the current subsection.

The task is now to study the asymptotic behavior of the process given by relation (12)

$$\sum_{i,j=1}^{q}\sqrt{n}Z_t^{mn,n}\int_0^t H_s^{i,j}(Y^i_{\eta_{mn}(s)}-Y^i_{\eta_n(s)})\,dY_s^j.$$

In order to study this process, we introduce the martingale process,

$$M_t^{n,i,j}=\int_0^t(Y^i_{\eta_{mn}(s)}-Y^i_{\eta_n(s)})\,dY_s^j,\quad (i,j)\in\{1,\cdots,q\}^2,$$

and we proceed to a preliminary calculus of the expectation of its bracket. Let (i,j) and $(i',j')\in\{1,\cdots,q\}^2$, we have

- for $j\neq j'$, the bracket $\langle M^{n,i,j},M^{n,i',j'}\rangle=0$
- for $j=j'$ and $i\neq i'$, $\mathbb{E}\langle M^{n,i,j},M^{n,i',j}\rangle=0$
- for $j=j'$ and $i=i'$, $\mathbb{E}\langle M^{n,i,j}\rangle_t=\int_0^t(\eta_{mn}(s)-\eta_n(s))\,ds$, $t\in[0,T]$ and we have

$$\begin{aligned}
\mathbb{E}(\langle M^{n,i,j}\rangle_t) &= \int_0^{\eta_n(t)}(\eta_{mn}(s)-\eta_n(s))ds+O(\frac{1}{n^2})\\
&= \sum_{\ell=0}^{m-1}\sum_{k=0}^{[t/\delta]-1}\int_{(mk+\ell)\delta/m}^{(mk+\ell+1)\delta/m}(\eta_{mn}(s)-\eta_n(s))\,ds+O(\frac{1}{n^2})\\
&= \sum_{\ell=0}^{m-1}\sum_{k=0}^{[t/\delta]-1}\frac{\delta^2}{m}\left(\frac{mk+\ell}{m}-k\right)+O(\frac{1}{n^2})=\frac{(m-1)\delta^2}{2m}[t/\delta]+O(\frac{1}{n^2})\\
&= \frac{(m-1)T}{2mn}t+O(\frac{1}{n^2}).
\end{aligned}\tag{13}$$

Having disposed of this preliminary evaluations, we can now study the stable convergence of $\left(\sqrt{\frac{2mn}{(m-1)T}}M^{n,i,j}\right)_{1\leq i,j\leq q}$. By virtue of Theorem 2-1 in [15], we need to study the asymptotic behavior of both brackets $n\langle M^{n,i,j},M^{n,i',j'}\rangle_t$ and $\sqrt{n}\langle M^{n,i,j},Y^{j'}\rangle_t$, for all $t\in[0,T]$ and all $(i,j,i',j')\in\{1,\cdots,q\}^4$. The case $j\neq j'$ is obvious and we only proceed to prove that

- for $j=j'$, $\sqrt{n}\langle M^{n,i,j},Y^j\rangle_t\xrightarrow[n\to\infty]{\mathbb{P}}0$, for all $t\in[0,T]$.
- for $j=j'$ and $i\neq i'$, $n\langle M^{n,i,j},M^{n,i',j}\rangle_t\xrightarrow[n\to\infty]{\mathbb{P}}0$, for all $t\in[0,T]$.
- for $j=j'$ and $i=i'$, $n\langle M^{n,i,j}\rangle_t\xrightarrow[n\to\infty]{\mathbb{P}}\frac{(m-1)T}{2m}t$, for all $t\in[0,T]$.

For the first point, we consider the L^2 convergence

$$\begin{aligned}\mathbb{E}\langle M^{n,i,j}, Y^j\rangle_t^2 &= \mathbb{E}\left(\int_0^t (Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)})ds\right)^2 \\ &= \int_0^t\int_0^t \mathbb{E}\left((Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)})(Y^i_{\eta_{mn}(u)} - Y^i_{\eta_n(u)})\right) dsdu \\ &= 2\int_{0<s<u<t} g(s,u)dsdu\end{aligned}$$

with

$$g(s,u) = \eta_{mn}(s)\wedge\eta_{mn}(u) - \eta_{mn}(s)\wedge\eta_n(u) - \eta_n(s)\wedge\eta_{mn}(u) + \eta_n(s)\wedge\eta_n(u). \tag{14}$$

It is worthy to note that

$$\eta_n(s) \le \eta_{mn}(s) \le s \le \eta_n(u) \le \eta_{mn}(u) \le u, \quad \forall\ s \le \eta_n(u). \tag{15}$$

Hence $g(s,u) = 0$, for $s \le \eta_n(u)$, $g(s,u) = \eta_{mn}(s) - \eta_n(s)$, for $\eta_n(u) < s < u$, and

$$\mathbb{E}\ \langle M^{n,i,j}, Y^j\rangle_t^2 = 2\int_{0<\eta_n(u)<s<u<t} (\eta_{mn}(s) - \eta_n(s))\, dsdu \le 2\frac{T}{n}\int_0^t (u - \eta_n(u))du \le 2\frac{T^2}{n^2}t.$$

This yields the desired result. Concerning the second point, the L^2 norm is given by

$$\begin{aligned}\mathbb{E}\langle M^{n,i,j}, M^{n,i',j}\rangle_t^2 &= \mathbb{E}\left(\int_0^t (Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)})(Y^{i'}_{\eta_{mn}(s)} - Y^{i'}_{\eta_n(s)})ds\right)^2 \\ &= \int_0^t\int_0^t \left(\mathbb{E}\left((Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)})(Y^i_{\eta_{mn}(u)} - Y^i_{\eta_n(u)})\right)\right)^2 dsdu \\ &= 2\int_{0<s<u<t} g(s,u)^2 dsdu,\end{aligned}$$

with the same function g given in relation (14). By properties of g developed above, we have in the same manner

$$\mathbb{E}\langle M^{n,i,j}, M^{n,i',j}\rangle_t^2 = 2\int_{0<\eta_n(u)<s<u<t} (\eta_{mn}(s) - \eta_n(s))^2 dsdu \le 2\frac{T^3}{n^3}t,$$

which proves our claim. For the last point, that is the essential one, taking into account the development of $\mathbb{E}\langle M^{n,i,j}\rangle_t$ given by relation (13) we obtain

$$\mathbb{E}\left(n\langle M^{n,i,j}\rangle_t - \frac{(m-1)T}{2m}t\right)^2 = n^2\mathbb{E}\langle M^{n,i,j}\rangle_t^2 - \frac{(m-1)^2T^2}{4m^2}t^2 + O(\frac{1}{n}). \tag{16}$$

Otherwise, we have

$$\begin{aligned}\mathbb{E}\langle M^{n,i,j}\rangle_t^2 &= \mathbb{E}\left(\int_0^t (Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)})^2 ds\right)^2 \\ &= \int_0^t\int_0^t \mathbb{E}\left((Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)})^2 (Y^i_{\eta_{mn}(u)} - Y^i_{\eta_n(u)})^2\right) dsdu \\ &= 2\int_{0<s<u<t} h(s,u)dsdu \qquad (17)\end{aligned}$$

with

$$h(s,u) = \mathbb{E}\left((Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)})^2 (Y^i_{\eta_{mn}(u)} - Y^i_{\eta_n(u)})^2\right). \tag{18}$$

On one hand, for $s \leq \eta_n(u)$, by property (15) and since the increments $Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)}$ and $Y^i_{\eta_{mn}(u)} - Y^i_{\eta_n(u)}$ are independent, it follows immediately that

$$h(s,u) = (\eta_{mn}(s) - \eta_n(s))(\eta_{mn}(u) - \eta_n(u)).$$

On the other hand, in relation (18) we use the Cauchy-Schwartz inequality to get $h(s,u) = O(\frac{1}{n^2})$ and this yields

$$\int_{0<\eta_n(u)<s<u<t} h(s,u)dsdu = O(\frac{1}{n^3}).$$

Now, noting that $(\eta_{mn}(s) - \eta_n(s))(\eta_{mn}(u) - \eta_n(u)) = O(\frac{1}{n^2})$, relation (17) becomes

$$\begin{aligned}\mathbb{E}\left(\langle M^{n,i,j}\rangle_t^2\right) &= 2\int_{0<s<u<t} (\eta_{mn}(s) - \eta_n(s))(\eta_{mn}(u) - \eta_n(u))dsdu + O(\frac{1}{n^3}) \\ &= \left(\int_0^t (\eta_{mn}(s) - \eta_n(s))ds\right)^2 + O(\frac{1}{n^3}).\end{aligned}$$

Once again thanks to the development of $\mathbb{E}(\langle M^{n,i,j}\rangle_t)$ given by relation (13), we deduce that

$$\mathbb{E}\langle M^{n,i,j}\rangle_t^2 = \frac{(m-1)^2T^2}{4m^2n^2}t^2 + O(\frac{1}{n^3}). \tag{19}$$

Combining relations (16) and (19), we deduce the convergence in L^2 of $n\langle M^{n,i,j}\rangle_t$ towards $\frac{(m-1)T}{2m}t$. By virtue of Theorem 2-1 in Jacod [15], $\left(\sqrt{\frac{2mn}{(m-1)T}}M^{n,i,j}\right)_{1\leq i,j\leq q}$ converges in law stably to a standard q^2-dimensional Brownian motion $(B^{ij})_{1\leq i,j\leq q}$ independent of W. Consequently, by Lemma 1 and Theorem 1, we obtain

$$\left(\sqrt{\frac{mn}{(m-1)T}}\int_0^t H^{i,j}_s (Y^i_{\eta_{mn}(s)} - Y^i_{\eta_n(s)})\, dY^j_s,\ t\geq 0\right)_{1\leq i,j\leq q} \Rightarrow^{stably} \left(\int_0^t H^{i,j}_s \frac{dB^{ij}_s}{\sqrt{2}},\ t\geq 0\right)_{1\leq i,j\leq q}$$

Finally, we complete the proof using relations (11), (12), Lemma 2 and once again Lemma 1 to obtain

$$\sqrt{\frac{mn}{(m-1)T}}\, U^{mn,n} \Rightarrow^{stably} U, \quad \text{where} \quad U_t = \frac{1}{\sqrt{2}}\sum_{i,j=1}^q Z_t \int_0^t H^{i,j}_s dB^{ij}_s.$$

□

Proof of Lemma 2 : At first, we prove the uniform convergence in probability toward zero of the normalized rest terms $\sqrt{n}R^{mn,n}_{t,i}$ for $i \in \{1,2\}$. The convergence of $\sqrt{n}\tilde{R}^{mn,n}_{t,i}$ $i \in \{1,2\}$ is a straightforward consequence of the previous one. The main part of these rest terms can be represented as integrals with respect to three types of supermartingales that can be classified through the following three cases

$$D^{n,0,0}_t = \sqrt{n}\int_0^t (s - \eta_n(s))\, ds, \quad D^{n,i,0}_t = \sqrt{n}\int_0^t (Y^i_s - Y^i_{\eta_n(s)})\, ds, \quad M^{n,0,j}_t = \sqrt{n}\int_0^t (s - \eta_n(s))\, dY^j_s,$$

where $(i,j) \in \{1,\cdots,q\}^2$ and $t \in [0,T]$. In the first case the supermartingale is deterministic of finite variation and its total variation on the interval $[0,T]$ has the following expression

$$\int_0^T \left|dD_t^{n,0,0}\right| = \sqrt{n}\int_0^T (s-\eta_n(s))\,ds \leq \frac{T^2}{\sqrt{n}}.$$

So, the process $D^{n,0,0}$ converges to 0 and is tight. In the second case, for $i \in \{1,\cdots,q\}$, the supermartingale is also of finite variation and its total variation on the interval $[0,T]$ has the following expression

$$\int_0^T \left|dD_t^{n,i,0}\right| = \sqrt{n}\int_0^T |Y_s^i - Y_{\eta_n(s)}^i|\,ds.$$

It is clear that $\sup_n \mathbb{E}\left(\int_0^T |dD_s^{n,i,0}|\right) < \infty$ which ensures the tightness of the process $D^{n,i,0}$. Therefore, we only need to establish the convergence of $D_t^{n,i,0}$ towards 0 in $L^2(\Omega)$, for $t \in [0,T]$. In fact, we have

$$\mathbb{E}\left((D_t^{n,i,0})^2\right) = 2n\int_{0<s<u<t} \mathbb{E}\left((Y_s^i - Y_{\eta_n(s)}^i)(Y_u^i - Y_{\eta_n(u)}^i)\right)\,ds\,du.$$

When $s \leq \eta_n(u)$, we have $\eta_n(s) \leq s \leq \eta_n(u) \leq u$ and by independence of the Brownian motion increments, we deduce that the integrand term is equal to 0. Otherwise, when $s \geq \eta_n(u)$, we apply the Cauchy Schwartz inequality to get

$$\mathbb{E}\left((D_t^{n,i,0})^2\right) \leq 2T\int_0^t (u-\eta_n(u))du \leq 2\frac{T^2}{n}t.$$

It follows from all these that $D^{n,i,0} \Rightarrow 0$. In the last case, for $j \in \{1,\cdots,q\}$, the process $M_t^{n,0,j}$ is a square integrable martingale and its bracket has the following expression

$$\langle M^{n,0,j}\rangle_T = n\int_0^T (s-\eta_n(s))^2\,ds \leq \frac{T^3}{n}.$$

It is clear that $\sup_n \mathbb{E}\langle M^{n,0,j}\rangle_T < \infty$, so we deduce the tightness of the process $\langle M^{n,0,j}\rangle$ and the convergence $M^{n,0,j} \Rightarrow 0$.

Now thanks to property $\tilde{\mathcal{P}}$) and relation (11), it is easy to check that the integrand processes $K_s^{i,mn,n}$ and $H_s^{0,j,mn,n}$, introduced in relation (10), converge uniformly in probability to their respective limits $K_s^i = (Z_s)^{-1}\left(\dot{\varphi}_{s,0}\bar{\varphi}_{s,i} - \sum_{j=1}^q (\dot{\varphi}_{s,j})^2\bar{\varphi}_{s,i}\right)$ and $H_s^{0,j} = (Z_s)^{-1}\dot{\varphi}_{s,j}\bar{\varphi}_{s,i}$, where $\dot{\varphi}_{s,j} = \nabla\varphi_j(X_s)$ and $\bar{\varphi}_{s,i} = \varphi_i(X_s)$. Therefore, by Theorem 1 we deduce that the integral processes given by

$$\sqrt{n}\int_0^t K_s^{i,mn,n}(Y_s^i - Y_{\eta_n(s)}^i)\,ds \quad \text{and} \quad \sqrt{n}\int_0^t H_s^{0,j,mn,n}(s-\eta_n(s))\,dY_s^j$$

vanish. Consequently, we conclude using relation (11) that $\sqrt{n}R_i^{mn,n} \Rightarrow 0$ for $i \in \{1,2\}$.

We now proceed to prove that $R_3^{mn,n} \Rightarrow 0$. The convergence of the process $\tilde{R}_3^{mn,n}$ toward 0 is obviously obtained from the previous one. The main part of this rest term can be represented as a stochastic integral with respect to the martingale process given by

$$N_t^{n,i,j} = \sqrt{n} \int_0^t (Y_s^i - Y_{\eta_n(s)}^i)\, dY_s^j,$$

with $(i,j) \in \{1, \cdots, q\} \times \{1, \cdots, q\}$. It was proven in Jacod and Protter [16] that

$$\sqrt{\frac{n}{T}} N^{n,i,j} \Rightarrow^{stably} \frac{B^{ij}}{\sqrt{2}},$$

where $(B^{ij})_{1 \leq i,j \leq q}$ is a standard q^2-dimensional Brownian motion defined on an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$ of the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, which is independent of W. Thanks to property $\tilde{\mathcal{P}}$) and relation(11), the integrand process $H^{i,j,mn,n} - H^{i,j} \Rightarrow 0$ and once again by Theorem 1 we deduce that the integral processes given by

$$\sqrt{n} \int_0^t (H_s^{i,j,mn,n} - H_s^{i,j})(Y_s^i - Y_{\eta_n(s)}^i)\, dY_s^j$$

vanish. All this allows us to conclude using relation (11). □

3.2 Central limit theorem

Let us recall that the multilevel Monte Carlo method uses information from a sequence of computations with decreasing step sizes and approximates the quantity $\mathbb{E}f(X_T)$ by

$$Q_n = \frac{1}{N_0} \sum_{k=1}^{N_0} f(X_{T,k}^1) + \sum_{\ell=1}^{L} \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(f(X_{T,k}^{\ell,m^\ell}) - f(X_{T,k}^{\ell,m^{\ell-1}}) \right), \; m \in \mathbb{N} \setminus \{0,1\} \;\text{ and }\; L = \frac{\log n}{\log m}.$$

In the same way as in the case of a crude Monte Carlo estimation, let us assume that the discretization error

$$\varepsilon_n = \mathbb{E}f(X_T^n) - \mathbb{E}f(X_T)$$

is of order $1/n^\alpha$ for any $\alpha \in [1/2, 1]$. Taking advantage from the limit theorem proven in the above section, we are now able to establish a central limit theorem of Lindeberg Feller type on the multilevel Monte Carlo Euler method. To do so, we introduce a real sequence $(a_\ell)_{\ell \in \mathbb{N}}$ of positive terms such that

$$(\mathcal{W}) \qquad \lim_{L \to \infty} \sum_{\ell=1}^{L} a_\ell = \infty \quad \text{and} \quad \lim_{L \to \infty} \frac{1}{\left(\sum_{\ell=1}^{L} a_\ell \right)^{p/2}} \sum_{\ell=1}^{L} a_\ell^{p/2} = 0, \text{ for } p > 2.$$

and we assume that the sample size N_ℓ depends on the rest of parameters by the relation

$$N_\ell = \frac{n^{2\alpha}(m-1)T}{m^\ell a_\ell} \sum_{\ell=1}^{L} a_\ell, \quad \ell \in \{0, \cdots, L\} \quad \text{and} \quad L = \frac{\log n}{\log m}. \tag{20}$$

We choose this form for N_ℓ because it is a generic form allowing us a straightforward use of Toeplitz lemma that is a crucial tool used in the proof of our central limit theorem. Indeed, property $(\mathcal{W})$ implies that if $(x_\ell)_{\ell\geq 1}$ is a sequence converging to $x \in \mathbb{R}$ as ℓ tends to infinity then

$$\lim_{L\to+\infty} \frac{\sum_{\ell=1}^{L} a_\ell x_\ell}{\sum_{\ell=1}^{L} a_\ell} = x.$$

In the sequel, we will denote by $\tilde{\mathbb{E}}$ respectively $\tilde{Var}$ the expectation respectively the variance defined on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ introduced in Theorem 3. We can now state the central limit theorem under strengthened conditions on the diffusion coefficients.

Theorem 4 *Assume that b and σ are $\mathcal{C}^1$ functions satisfying the global Lipschitz condition $(\mathcal{H}_{b,\sigma})$. Let f be a real valued function satisfying*

$$(\mathcal{H}_{\mathbf{f}}) \quad |f(x) - f(y)| \leq C(1 + |x|^p + |y|^p)|x - y|, \quad \textit{for some } C, p > 0.$$

Assume $\mathbb{P}(X_T \notin \mathcal{D}_f) = 0$, where $\mathcal{D}_f := \{x \in \mathbb{R}^d; f \textit{ is differentiable at } x\}$, and that for some $\alpha \in [1/2, 1]$ we have

$$(\mathcal{H}_{\varepsilon_{\mathbf{n}}}) \quad \lim_{n\to\infty} n^\alpha \varepsilon_n = C_f(T, \alpha).$$

Then, for the choice of N_ℓ, $\ell \in \{0, 1, \cdots, L\}$ given by equation (20), we have

$$n^\alpha \big(Q_n - \mathbb{E}\left(f(X_T)\right)\big) \Rightarrow \mathcal{N}\left(C_f(T, \alpha), \sigma^2\right)$$

with $\sigma^2 = \tilde{V}ar\big(\nabla f(X_T).U_T\big)$ and $\mathcal{N}\left(C_f(T, \alpha), \sigma^2\right)$ denotes a normal distribution.

The global Lipschitz condition $(\mathcal{H}_{b,\sigma})$ seems to be essential to establish our result, since it ensures property $\mathcal{P}$). Otherwise, Hutzenthaler, Jentzen and Kloeden [14] prove that under weaker conditions on b and σ the multilevel Monte Carlo Euler method may diverges whereas the crude Monte Carlo method converges.

Proof: To simplify our notations we give the proof for $\alpha = 1$, the case $\alpha \in [1/2, 1)$ is a straightforward deduction. Combining relations (5) and (6) together we get

$$Q_n - \mathbb{E}\left(f(X_T)\right) = \hat{Q}_n^1 + \hat{Q}_n^2 + \varepsilon_n,$$

where

$$\begin{aligned}
\hat{Q}_n^1 &= \frac{1}{N_0} \sum_{k=1}^{N_0} \left(f(X_{T,k}^1) - \mathbb{E}\left(f(X_T^1)\right)\right) \\
\hat{Q}_n^2 &= \sum_{\ell=1}^{L} \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(f(X_{T,k}^{\ell,m^\ell}) - f(X_{T,k}^{\ell,m^{\ell-1}}) - \mathbb{E}\left(f(X_T^{\ell,m^\ell}) - f(X_T^{\ell,m^{\ell-1}})\right)\right).
\end{aligned}$$

Using assumption $(\mathcal{H}_{\varepsilon_{\mathbf{n}}})$ we obviously obtain the term $C_f(T,\alpha)$ in the limit. Taking $N_0 = \frac{n^2(m-1)T}{a_0}\sum_{\ell=1}^{L} a_\ell$, we can apply the classical central limit theorem to $\hat{Q}_n^1$. Then we have $n\hat{Q}_n^1 \xrightarrow{\mathbb{P}} 0$. Finally, we have only to study the convergence of $n\hat{Q}_n^2$ and we will conclude by establishing

$$n\hat{Q}_n^2 \Rightarrow \mathcal{N}\Big(0, \tilde{V}ar\big(\nabla f(X_T).U_T\big)\Big).$$

To do so, we plan to use Theorem 2 with the Lyapunov condition and we set

$$X_{n,\ell} := \frac{n}{N_\ell}\sum_{k=1}^{N_\ell} Z_{T,k}^{m^\ell,m^{\ell-1}} \text{ and } Z_{T,k}^{m^\ell,m^{\ell-1}} := f(X_{T,k}^{\ell,m^\ell}) - f(X_{T,k}^{\ell,m^{\ell-1}}) - \mathbb{E}\left(f(X_{T,K}^{\ell,m^\ell}) - f(X_{T,k}^{\ell,m^{\ell-1}})\right). \tag{21}$$

In other words, we will check the following conditions :

- $\lim_{n\to\infty}\sum_{\ell=1}^{L}\mathbb{E}(X_{n,\ell})^2 = \tilde{V}ar\big(\nabla f(X_T).U_T\big)$
- (Lyapunov condition) there exists $p>2$ such that $\lim_{n\to\infty}\sum_{\ell=1}^{L}\mathbb{E}\,|X_{n,\ell}|^p = 0$.

For the first one, we have

$$\begin{aligned}\sum_{\ell=1}^{L}\mathbb{E}(X_{n,\ell})^2 = \sum_{\ell=1}^{L} Var(X_{n,\ell}) &= \sum_{\ell=1}^{L}\frac{n^2}{N_\ell}Var\left(Z_{T,1}^{m^\ell,m^{\ell-1}}\right)\\ &= \frac{1}{\sum_{\ell=1}^{L}a_\ell}\sum_{\ell=1}^{L} a_\ell\frac{m^\ell}{(m-1)T}Var\left(Z_{T,1}^{m^\ell,m^{\ell-1}}\right). \end{aligned}\tag{22}$$

Otherwise, since $\mathbb{P}(X_T\notin\mathcal{D}_f)=0$, applying the Taylor expansion theorem twice we get

$$\begin{aligned}f(X_T^{\ell,m^\ell}) - f(X_T^{\ell,m^{\ell-1}}) &= \nabla f(X_T).U_T^{m^\ell,m^{\ell-1}} +\\ &(X_T^{\ell,m^\ell}-X_T)\varepsilon(X_T,X_T^{\ell,m^\ell}-X_T) - (X_T^{\ell,m^{\ell-1}}-X_T)\varepsilon(X_T,X_T^{\ell,m^{\ell-1}}-X_T).\end{aligned}$$

The function ε is given by the Taylor-Young expansion, so it satisfies $\varepsilon(X_T,X_T^{\ell,m^\ell}-X_T)\xrightarrow[\ell\to\infty]{\mathbb{P}}0$ and $\varepsilon(X_T,X_T^{\ell,m^{\ell-1}}-X_T)\xrightarrow[\ell\to\infty]{\mathbb{P}}0$. By property $\mathcal{P}$) we get the tightness of $\sqrt{\frac{m^\ell}{(m-1)T}}(X_T^{\ell,m^\ell}-X_T)$ and $\sqrt{\frac{m^\ell}{(m-1)T}}(X_T^{\ell,m^{\ell-1}}-X_T)$ and we deduce

$$\sqrt{\frac{m^\ell}{(m-1)T}}\left((X_T^{\ell,m^\ell}-X_T)\varepsilon(X_T,X_T^{\ell,m^\ell}-X_T) - (X_T^{\ell,m^{\ell-1}}-X_T)\varepsilon(X_T,X_T^{\ell,m^{\ell-1}}-X_T)\right)\xrightarrow[\ell\to\infty]{\mathbb{P}}0.$$

So, according to Lemma 1 and Theorem 3 we conclude that

$$\sqrt{\frac{m^\ell}{(m-1)T}}\left(f(X_T^{\ell,m^\ell}) - f(X_T^{\ell,m^{\ell-1}})\right)\Rightarrow^{stably}\nabla f(X_T).U_T,\quad \text{as}\ \ \ell\to\infty. \tag{23}$$

Using $(\mathcal{H}_{\mathbf{f}})$ it follows from property $\mathcal{P}$) that

$$\forall\varepsilon>0,\quad \sup_\ell\mathbb{E}\left|\sqrt{\frac{m^\ell}{(m-1)T}}\left(f(X_T^{\ell,m^\ell}) - f(X_T^{\ell,m^{\ell-1}})\right)\right|^{2+\varepsilon} < \infty.$$

We deduce using relation (23) that

$$\mathbb{E}\left(\sqrt{\frac{m^\ell}{(m-1)T}}\left(f(X_T^{\ell,m^\ell})-f(X_T^{\ell,m^{\ell-1}})\right)\right)^k \to \tilde{\mathbb{E}}\Big(\nabla f(X_T).U_T\Big)^k < \infty \quad \text{for } k\in\{1,2\}.$$

Consequently,

$$\frac{m^\ell}{(m-1)T}Var(Z_{T,1}^{m^\ell,m^{\ell-1}}) \longrightarrow \tilde{V}ar\left(\nabla f(X_T).U_T\right) < \infty.$$

Hence combining this result with relation (22), we obtain the first condition using Toeplitz lemma. Concerning the second one, by Burkholder's inequality and elementary computations, we get for $p>2$

$$\mathbb{E}|X_{n,\ell}|^p = \frac{n^p}{N_\ell^p}\mathbb{E}\left|\sum_{\ell=1}^{N_\ell} Z_{T,1}^{m^\ell,m^{\ell-1}}\right|^p \le C_p\frac{n^p}{N_\ell^{p/2}}\mathbb{E}\left|Z_{T,1}^{m^\ell,m^{\ell-1}}\right|^p, \tag{24}$$

where C_p is a numerical constant depending only on p. Otherwise, property $\mathcal{P}$) ensures the existence of a constant $K_p>0$ such that

$$\mathbb{E}\big|Z_{T,1}^{m^\ell,m^{\ell-1}}\big|^p \le \frac{K_p}{m^{p\ell/2}}.$$

Therefore

$$\sum_{\ell=1}^{L}\mathbb{E}\left|X_{n,\ell}\right|^p \le \tilde{C}_p\sum_{\ell=1}^{L}\frac{n^p}{N_\ell^{p/2}m^{p\ell/2}} \le \frac{\tilde{C}_p}{\left(\sum_{\ell=1}^{L}a_\ell\right)^{p/2}}\sum_{\ell=1}^{L}a_\ell^{p/2} \underset{n\to\infty}{\longrightarrow} 0. \tag{25}$$

This completes the proof. □

Remark 1 *From Theorem 2 page 544 in [6], we prove a Berry-Essen type bound on our central limit theorem. This improves the relevance of the above result. Indeed, take $\alpha=1$ as in the proof, for $X_{n,0}=n\hat{Q}_n^1$ and $X_{n,\ell}$ given by relation (21), with $\ell\in\{1,\cdots,L\}$, put*

$$s_n^2=\sum_{\ell=0}^{L}\mathbb{E}|X_{n,\ell}|^2, \quad \rho_n=\sum_{\ell=0}^{L}\mathbb{E}|X_{n,\ell}|^3$$

and denote by F_n the distribution function of $n(Q_n-\mathbb{E}f(X_T^n))/s_n$. Then for all $x\in\mathbb{R}$ and $n\in\mathbb{N}^*$

$$|F_n(x)-G(x)| \le 6\frac{\rho_n}{s_n^3}, \tag{26}$$

where G is the distribution function of a standard Gaussian random variable. If we interpret the output of the above inequality as sum of independent individual path simulation, we get

$$s_n^2=\frac{1}{(m-1)T\sum_{\ell=1}^{L}a_\ell}\left(a_0Var\left(f(X_T^1)\right)+\sum_{\ell=1}^{L}a_\ell m^\ell Var\left(f(X_T^{\ell,m^\ell})-f(X_T^{\ell,m^{\ell-1}})\right)\right).$$

According to the above proof, it is clear that s_n behaves like a constant but getting lower bounds for s_n seems not to be a common result to our knowledge. Concerning ρ_n, taking $p = 3$ in both inequalities (24) and (25) gives us an upper bound. In fact, when f is Lipschitz, there exists a positive constant C depending on b, σ, T and f such that

$$\rho_n \leq \frac{C}{\left(\sum_{\ell=1}^{L} a_\ell\right)^{3/2}} \sum_{\ell=1}^{L} a_\ell^{3/2}.$$

For the optimal choice $a_\ell = 1$, given in the below subsection, the obtained Berry-Essen type bound is of order $1/\sqrt{\log n}$.

Remark 2 *Note that the above proof differs from the ones in Kebaier [17]. In fact, here our proof is based on the central limit theorem for triangular array which is adapted to the form of the multilevel estimator, whereas Kebaier used a crude approach based on studying the associated characteristic function. Further, this latter approach needs a control on the third moment, whereas we only need to control a moment strictly greater than two. Also, it is worth to note that the limit variance in Theorem 4 is smaller than the limit variance in Theorem 3.2 obtained by Kebaier in [17].*

3.3 Complexity analysis

From a complexity analysis point of view, we can interpret Theorem 4 as follows. For a total error of order $1/n^\alpha$ the computational effort necessary to run the multilevel Monte Carlo Euler method is given by the sequence of sample sizes specified by relation (20). The associated time complexity is given by:

$$\begin{aligned} C_{MMC} &= C \times \left(N_0 + \sum_{\ell=1}^{L} N_\ell (m^\ell + m^{\ell-1}) \right) \quad \text{with} \quad C > 0 \\ &= C \times \left(\frac{n^{2\alpha}(m-1)T}{a_0} \sum_{\ell=1}^{L} a_\ell + n^{2\alpha} \frac{(m^2-1)T}{m} \sum_{\ell=1}^{L} \frac{1}{a_\ell} \sum_{\ell=1}^{L} a_\ell \right). \end{aligned}$$

The minimum of the second term of this complexity is reached for the choice of weights $a_\ell^* = 1$, $\ell \in \{1, \cdots, L\}$, since the Cauchy-Schwartz inequality ensures that $L^2 \leq \sum_{\ell=1}^{L} \frac{1}{a_\ell} \sum_{\ell=1}^{L} a_\ell$, and the optimal complexity for the multilevel Monte Carlo Euler method is given by

$$C_{MMC} = C \times \left(\frac{(m-1)T}{a_0 \log m} n^{2\alpha} \log n + \frac{(m^2-1)T}{m(\log m)^2} n^{2\alpha} (\log n)^2 \right) = O\left(n^{2\alpha} (\log n)^2 \right).$$

It turns out that for a given discretization error $\varepsilon_n = 1/n^\alpha$ to be achieved the complexity is given by $C_{MMC} = O\left(\varepsilon_n^{-2} (\log \varepsilon_n)^2\right)$. Note that this optimal choice $a_\ell^* = 1$, $\ell \in \{1, \cdots, L\}$, with taking $a_0 = 1$ corresponds to the sample sizes given by

$$N_\ell = \frac{(m-1)T}{m^\ell \log m} n^{2\alpha} \log n, \quad \ell \in \{0, \cdots, L\}.$$

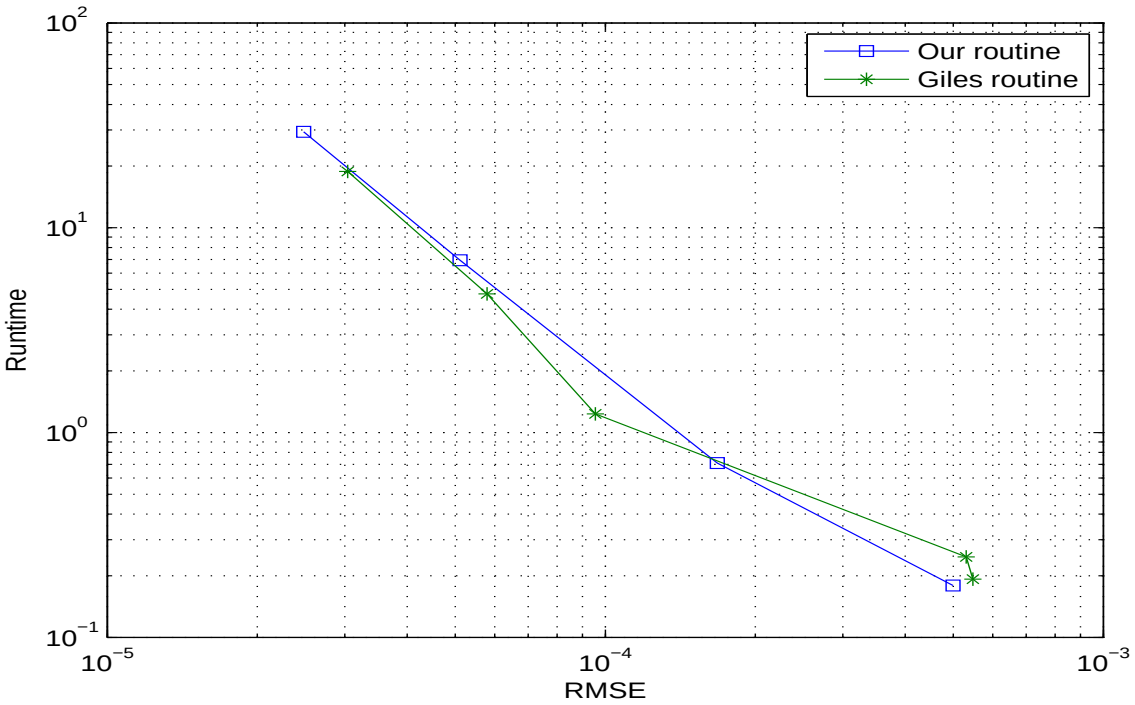

Figure 1: Comparison of both routines.

Hence, our optimal choice is consistent with that proposed by Giles [8]. Nevertheless, unlike the parameters obtained by Giles [8] for the same setting (see relation (1)), our optimal choice of the sample sizes N_ℓ, $\ell \in \{1, \cdots, L\}$ does not depend on any given constant, since our approach is based on proving a central limit theorem and not on getting upper bounds for the variance. Otherwise, for the same error of order $\varepsilon_n = 1/n^\alpha$ the optimal complexity of a Monte Carlo method is given by

$$C_{MC} = O\left(n^{2\alpha+1}\right) = O\left(\varepsilon_n^{-2-1/\alpha}\right)$$

which is clearly larger than C_{MMC}. So we deduce that the multilevel method is more efficient. Note also that the optimal choice of the parameter m is obtained for $m^* = 7$. Otherwise, any choice $N_0 = n^{2\alpha}(\log n)^\beta$, $0 < \beta < 2$, leads to the same result. Some numerical tests comparing original Giles work [8] with the one of us show that both error rates are in line. Here in Figure 1, we make a simple log-log scale plot of CPU time with respect to the root mean square error, for European call and with $N_0 = n^{2\alpha}(\log n)^{1.9}$.

It is worth to note that the advantage of the central limit theorem is to construct a more accurate confidence interval. In fact, for a given root mean square error RMSE, the radius of the 90%-confidence interval constructed by the central limit theorem is 1.64×RMSE. However, without this latter result one can only use Chebyshev's inequality which yields a radius equal to 3.16×RMSE. Finally note that, taking $\alpha = 1/2$ still gives the optimal rate and allows us to cancel the bias in the central limit theorem due to the Euler discretization.

4 Conclusion

The multilevel Monte Carlo algorithm is a method that can be used in a general framework: as soon as we use a discretization scheme in order to compute quantities such as $\mathbb{E}f\left(X_t, 0 \leq t \leq T\right)$, we can implement the statistical multilevel algorithm. And this is worth because it is an efficient method according to the original work by Giles [8]. The central limit theorems derived in this paper fill the gap in literature and confirm superiority of the multilevel method over the classical Monte Carlo approach.

Acknowledgments The authors are greatly indebted to the Associate Editor and the referees for their many suggestions and interesting comments which improve the paper.

References

[1] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

[2] N. Bouleau and D. Lépingle. *Numerical methods for stochastic processes*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Inc., New York, 1994. A Wiley-Interscience Publication.

[3] J. Creutzig, S. Dereich, T. Müller-Gronbach, and K. Ritter. Infinite-dimensional quadrature and approximation of distributions. *Found. Comput. Math.*, 9(4):391–429, 2009.

[4] S. Dereich. Multilevel Monte Carlo algorithms for Lévy-driven SDEs with Gaussian correction. *Ann. Appl. Probab.*, 21(1):283–311, 2011.

[5] D. Duffie and P. Glynn. Efficient Monte Carlo simulation of security prices. *Ann. Appl. Probab.*, 5(4):897–905, 1995.

[6] W. Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[7] M. B. Giles. Improved multilevel Monte Carlo convergence using the Milstein scheme. In *Monte Carlo and quasi-Monte Carlo methods 2006*, pages 343–358. Springer, Berlin, 2008.

[8] M. B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.

[9] M. B. Giles, D. J. Higham, and X. Mao. Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff. *Finance Stoch.*, 13(3):403–413, 2009.

[10] M. B. Giles and L. Szpruch. *Multilevel Monte Carlo methods for applications in finance*, chapter 1, pages 3–47.

[11] S. Heinrich. Monte Carlo complexity of global solution of integral equations. *J. Complexity*, 14(2):151–175, 1998.

[12] S. Heinrich. Multilevel monte carlo methods. *Lecture Notes in Computer Science, Springer-Verlag*, 2179(1):58–67, 2001.

[13] S. Heinrich and E. Sindambiwe. Monte carlo complexity of parametric integration. *J. Complexity*, 15(3):317–341, 1999. Dagstuhl Seminar on Algorithms and Complexity for Continuous Problems (1998).

[14] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Divergence of the multilevel monte carlo euler method for nonlinear stochastic differential equations. *Annals of Applied probability*, 23(5):1913–1966, 2013.

[15] J. Jacod. On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de Probabilités, XXXI*, volume 1655 of *Lecture Notes in Math.*, pages 232–246. Springer, Berlin, 1997.

[16] J. Jacod and P. Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26(1):267–307, 1998.

[17] A. Kebaier. Statistical romberg extrapolation: A new variance reduction method and applications to option pricing. *Annals of Applied probability*, 15(4):2681–2705, 2005.

[18] P. E. Kloeden and E. Platen. Numerical methods for stochastic differential equations. In *Nonlinear dynamics and stochastic mechanics*, CRC Math. Model. Ser., pages 437–461. CRC, Boca Raton, FL, 1995.

[19] R. Korn, E. Korn, and G. Kroisandt. *Monte Carlo methods and models in finance and insurance*. Chapman & Hall/CRC Financial Mathematics Series. CRC Press, Boca Raton, FL, 2010.

[20] P. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990.

[21] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.