Non-asymptotic error bounds for the multilevel Monte Carlo Euler method applied to SDEs with constant diffusion coefficient*

Benjamin Jourdain†   Ahmed Kebaier‡

Abstract

In this paper, we are interested in deriving non-asymptotic error bounds for the multilevel Monte Carlo method. As a first step, we deal with the explicit Euler discretization of stochastic differential equations with a constant diffusion coefficient. We prove that, as long as the deviation is below an explicit threshold, a Gaussian-type concentration inequality optimal in terms of the variance holds for the multilevel estimator. To do so, we use the Clark-Ocone representation formula and derive bounds for the moment generating functions of the squared difference between a crude Euler scheme and a finer one and of the squared difference of their Malliavin derivatives.

Keywords: non asymptotic bounds; Euler scheme; multilevel Monte Carlo methods; Malliavin calculus.
AMS MSC 2010: 60H35; 65C30; 65C05; 60H07.
Submitted to EJP on September 4, 2017, final version accepted on January 23, 2019.
Supersedes `arXiv:1708.07064v2`.

1 Introduction

We are interested in deriving non asymptotic error estimations for the multilevel Monte Carlo estimators introduced by Giles [4]. In this paper, as a first step, we deal with estimators of $\mathbb{E}\left[f(X_T)\right]$ where $f : \mathbb{R}^d \to \mathbb{R}$ is Lipschitz continuous with constant $[\dot{f}]_\infty$, $T \in (0, +\infty)$ is a deterministic time horizon and $X := (X_t)_{0 \le t \le T}$ is the $\mathbb{R}^d$-valued solution to the stochastic differential equation with additive noise

$$dX_t = b(X_t)dt + dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \tag{1.1}$$

driven by the d-dimensional Brownian motion $W = (W^1, \ldots, W^d)$ and with Lipschitz drift function $b : \mathbb{R}^d \to \mathbb{R}^d$:

$$(\mathcal{H}_{GL}) \qquad\qquad \exists C_b < +\infty, \ \forall x, y \in \mathbb{R}^d \quad |b(x) - b(y)| \leq C_b |x - y|.$$

When $d = 1$, this additive noise setting is not restrictive. Indeed any stochastic differential equation $dY_t = \sigma(Y_t)dW_t + \eta(Y_t)dt$ with multiplicative noise given by some function $\sigma : \mathbb{R} \to \mathbb{R}_+^*$ such that $\frac{1}{\sigma}$ is locally integrable can be reduced to (1.1) by the Lamperti transformation : for $\varphi(y) = \int_{y_0}^y \frac{dz}{\sigma(z)}$, $X_t = \varphi(Y_t)$ solves (1.1) with $b(x) = \left(\frac{\eta}{\sigma} - \frac{\sigma'}{2} \right)(\varphi^{-1}(x))$.

For $n \in \mathbb{N}^*$, we consider the simple Euler-Maruyama approximation X^n with time step T/n and we introduce its continuous version given by

$$dX_t^n = b(X_{\eta_n(t)}^n)dt + dW_t, \quad \eta_n(t) = \left\lfloor \frac{nt}{T} \right\rfloor \frac{T}{n}, \ X_0^n = x_0. \tag{1.2}$$

When b is smooth, both the strong and the weak errors of this scheme converge to 0 with order 1 as $n \to \infty$. According to [4], the complexity for the multilevel Monte Carlo estimator of $\mathbb{E}[f(X_T)]$ based on this scheme to achieve a root mean square error ε is $\mathcal{O}(\varepsilon^{-2})$ in the limit $\varepsilon \to 0$, the same as in a standard Monte Carlo method with i.i.d. unbiased samples. For positive integers m and L and $(N_\ell)_{0 \leq \ell \leq L}$, the Multilevel Monte Carlo method approximates the expectation of interest $\mathbb{E}[f(X_T)]$ by

$$\hat{Q} = \frac{1}{N_0} \sum_{k=0}^{N_0} f(X_{T,k}^1) + \sum_{\ell=0}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) \right). \tag{1.3}$$

The processes $((X_{t,k}^{m^\ell})_{0 \leq t \leq T})_k$ denote independent copies of the Euler scheme with time step $m^{-\ell} T$ for $\ell \in \{0, \cdots, L\}$. Here, it is important to point out that all these $L + 1$ Monte Carlo estimators have to be based on different, independent samples. However, for fixed k and ℓ, the simulations $f(X_{T,k}^{m^\ell})$ and $f(X_{T,k}^{m^{\ell-1}})$ have to be based on the same Brownian path but with different times steps $m^{-\ell} T$ and $m^{-(\ell-1)} T$.

Our main motivation is the derivation of Gaussian type concentration inequalities for $\hat{Q} - \mathbb{E}[f(X_T)]$, a natural question, which, to our knowledge has not been addressed in the literature. Frikha and Menozzi [3] obtained concentration inequalities for $f(X_T^n) - \mathbb{E}[f(X_T^n)]$. Deriving estimations of the moment generating function of the differences $f(X_T^{mn}) - f(X_T^n) - \mathbb{E}[f(X_T^{mn}) - f(X_T^n)]$ which are optimal in terms of their variances is a much more delicate task and adapting their approach seems to be problematic. However, the boundedness of the Malliavin derivatives $\mathcal{D}X_T^n$ and $\mathcal{D}X_T^{mn}$ in the additive noise setting permits to follow the approach of Houdré and Privault [6] based on the Clark-Ocone formula and this is one reason why we focus on this setting. Another reason is that for stochastic differential equations with multiplicative noise, more sophisticated schemes, like the Milstein scheme in the commutative case or the Giles and Szpruch [5] scheme in the general case, are necessary to improve to two the order one of convergence of $\text{Var}[f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})]$ and recover the unbiased Monte Carlo complexity.

In Section 2, when b is $\mathscr{C}^2$, Lipschitz continuous and the Laplacians of its coordinates have an affine growth, we first derive non-asymptotic estimates of the squared error $\mathbb{E}[(\hat{Q} - \mathbb{E}[f(X_T)])^2]$ of the multilevel Monte Carlo estimator (MLMC) (1.3) for a Lipschitz continuous test function f by computing explicit bounds for the bias $\mathbb{E}[f(X_T^{m^L}) - f(X_T)]$ and variance $\text{Var}[f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})]$. Then we optimize the parameters $(L, (N_\ell)_{0 \leq \ell \leq L})$ in order to minimize the computation cost needed to achieve a root mean square error smaller than a given precision ε. It turns out that, as $\varepsilon \to 0$, the optimal bias is of order $\mathcal{O}(\varepsilon^{4/3})$, which, to our knowledge, has not been pointed out in the MLMC literature so far. Notice that, for stochastic differential equations with a non constant diffusion

coefficient (multiplicative noise), this property remains true for the multilevel Monte Carlo estimator based on the Giles and Szpruch scheme [5], since it exhibits the same orders of convergence of the bias and the variance within a given level as (1.3).

In Section 3, we state and derive our main result : as long as the deviation is below an explicit threshold, a Gaussian-type concentration inequality optimal in terms of the variance holds for the multilevel estimator (1.3). Denoting by $\hat{Q}_\varepsilon$ the multilevel Monte Carlo estimator corresponding to the optimal choice of parameters discussed in Section 2.1, we obtain the existence of explicit positive constants c_1, c_2 and c_3 such that

$$\forall \varepsilon \in (0, c_1), \ \forall \alpha \in (0, c_2 \varepsilon^{2/3}), \ \mathbb{P}\left(|\hat{Q}_\varepsilon - \mathbb{E}f(X_T)| \geq \alpha\right) \leq 2e^{\frac{2}{c_3}} e^{-\frac{\alpha^2}{c_3 \varepsilon^2}}. \tag{1.4}$$

In view of the last factor, this bound is optimal in terms of the precision ε (up to the value of the multiplicative constant c_3). For deviations $\alpha(\varepsilon)$ depending on ε and such that $\lim_{\varepsilon \to 0} \frac{\alpha(\varepsilon)}{\varepsilon} = \infty$, the right-hand side of (1.4) converges to 0 far quicker than the one of the bound $\mathbb{P}\left(|\hat{Q}_\varepsilon - \mathbb{E}f(X_T)| \geq \alpha\right) \leq \frac{\varepsilon^2}{\alpha^2}$ consequence of the Markov inequality. We show in Corollary 3.6 that the same inequality holds for deviations α up to the order $\ln(1/\varepsilon)^{-1/\beta}$ with $\beta > 1$ for a multilevel estimator with increased numbers of simulations in the high levels but with computation cost still of order $\mathcal{O}(\varepsilon^{-2})$ as $\varepsilon \to 0$. Moreover, we derive a comparison between the root mean square error (RMSE) and Orlicz norm for both standard and multilevel Monte Carlo. It turns out that compared to standard Monte Carlo, the MLMC estimator achieves the same complexity reduction for Orlicz norm as for the RMSE (see Section 3.3). The limitation, mentioned above, on the range of deviations α for which the Gaussian-type concentration inequality holds is related to a corresponding limitation on the range of parameters for which we are able to estimate (optimally in terms of the variance) the moment generating function of $\hat{Q} - \mathbb{E}f(X_T)$. This comes from the quadratic contributions of the Brownian increments that one obtains when applying Itô's formula twice to exhibit the order of the difference $f(X_T^{mn}) - f(X_T^n)$ for $n \in \{1, m, \ldots, m^{L-1}\}$. Maybe these restrictions could be relaxed when replacing the Brownian increments in the Euler schemes by Rademacher random variables like in the weak MLMC method introduced by Belomestny and Nagapetyan [2]. Nonetheless the derivation of concentration bounds for the weak MLMC estimators would require a different approach. Indeed, we use the Clark-Ocone formula as suggested in Houdré and Privault [6], to relate the estimation of the moment generating function of $f(X_T^{mn}) - f(X_T^n) - \mathbb{E}[f(X_T^{mn}) - f(X_T^n)]$ for $n \in \{1, m, \ldots, m^{L-1}\}$, to the ones of the squared difference between the crude Euler scheme with n steps and the finer one with mn steps and of the squared difference of their Malliavin derivatives. Such estimations are respectively proved in sections 4 and 5 by using a clever decomposition of the difference between the two schemes. They are combined in Section 6 to estimate the moment generating function of $\hat{Q} - \mathbb{E}f(X_T)$.

Notations Throughout this paper, we shall use the following notations.

- We denote by $\mathscr{C}_p^\infty(\mathbb{R}^d, \mathbb{R}^q)$ the set of all infinitely differentiable functions $g : \mathbb{R}^d \to \mathbb{R}^q$ such that g and all of its partial derivatives have at most polynomial growth.

- For $n \in \mathbb{N}^*$, we denote by $\mathscr{C}^n(\mathbb{R}^d, \mathbb{R}^q)$ the set of all n times continuously differentiable functions $g : \mathbb{R}^d \to \mathbb{R}^q$.

- For $g : \mathbb{R}^d \to \mathbb{R}^d$, we denote by ∇g the Jacobian matrix defined for all $i, j \in \{1, \ldots, d\}$ and $x \in \mathbb{R}^d$ by $(\nabla g)_{ij}(x) = \partial_{x_j} g_i(x)$.

- For $g : \mathbb{R}^d \to \mathbb{R}^d$, $\Delta g : \mathbb{R}^d \to \mathbb{R}^d$ denotes the function obtained by applying the Laplacian to each coordinate of g.

- The ceiling function and floor function are denoted respectively by $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ (i.e. for $x \in \mathbb{R}$, $\lceil x \rceil$ represents the smallest integer no less than x; $\lfloor x \rfloor$ represents the largest integer no greater than x).

- For $d \in \mathbb{N}^*$, $\mathcal{M}_d$ denotes the set of real d-square matrices with identity matrix I_d.

- For any matrix sequence $(A_k)_{k \in \mathbb{N}} \in \mathcal{M}_d$, we use the following convention

$$\prod_{k=n_2}^{n_1} A_k = A_{n_2} \cdots A_{n_1}, \ \ \forall n_1, n_2 \in \mathbb{N} \text{ s.t. } n_1 \leq n_2.$$

- The Euclidean inner product and the associated norm are respectively denoted by $\cdot$ and $|\cdot|$.

- For $M \in \mathcal{M}_d$, the matrix norm induced by the Euclidean norm $|\cdot|$ is denoted by

$$\|M\| = \sup_{x \in \mathbb{R}^d : |x|=1} |Mx|.$$

- For any adapted $\mathbb{R}^d$-valued process $(H_t)_{0 \leq t \leq T}$ and $\mathcal{M}_d$-valued process $(M(t))_{0 \leq t \leq T}$, we denote

$$|H| := \left\| \left(\int_0^T |H_t|^2 dt \right)^{1/2} \right\|_\infty \quad \text{and} \quad |M| := \left\| \left(\int_0^T \text{Tr}\left[M(t) M(t)^\top \right] dt \right)^{1/2} \right\|_\infty,$$

where for $A \in \mathcal{M}_d$, $A^\top$ and $\text{Tr}[A]$ denote respectively the transpose and the trace of matrix A.

2 Non-asymptotic mean square error of the MLMC estimator

2.1 Assumptions and strong error analysis

It is well known that under assumption $(\mathcal{H}_{GL})$ we have
$(\mathcal{P})$
$$\forall p \geq 1, \ \sup_{0 \leq t \leq T} |X_t|, \ \sup_{0 \leq t \leq T} |X_t^n| \in L^p \text{ and } \mathbb{E}\left[\sup_{0 \leq t \leq T} |X_t - X_t^n|^p \right] \leq \frac{K_p(T)}{n^{p/2}}, \text{ with } K_p(T) < \infty.$$

Moreover, since the diffusion coefficient is constant, the Euler scheme coincides with the Milstein scheme and if b belongs to $\mathscr{C}^2(\mathbb{R}^d, \mathbb{R}^d)$ with bounded derivatives, then the strong error estimation improves to $\mathbb{E}\left[\sup_{0 \leq t \leq T} |X_t - X_t^n|^p \right] \leq \frac{K_p(T)}{n^p}$, with $K_p(T) < \infty$ (see for instance [7]). In order to get a non-asymptotic control of the bias and the variance of the multilevel Monte Carlo estimator, we are now going to state an explicit bound for the terminal quadratic strong error $\mathbb{E}\left[|X_T^{mn} - X_T^n|^2 \right]$ for $(n, m) \in \mathbb{N}^* \times \bar{\mathbb{N}}$ (with the convention $X^{mn} = X$ for $m = \infty$) under the following assumption. The constancy of the diffusion coefficient ensures that the bias can be estimated with the right order of convergence using this strong error analysis instead of the more complicated weak error analysis.

Assumption (R1) The function $b \in \mathscr{C}^2(\mathbb{R}^d, \mathbb{R}^d)$ and there exist finite constants $[\dot{b}]_\infty \in (0, +\infty)$ and $a_{\Delta b} \in [0, +\infty)$ such that

$$\forall x \in \mathbb{R}^d, \quad \|\nabla b(x)\| \leq [\dot{b}]_\infty,$$
$$\forall x \in \mathbb{R}^d, \ |\Delta b(x)| \leq 2a_{\Delta b}(1 + |x - x_0|). \tag{2.1}$$

Proposition 2.1. *Assume* $(\mathbf{R1})$. *Then, for all* $(n, m) \in \mathbb{N}^* \times \bar{\mathbb{N}}$,

$$\mathbb{E}\left[|X_T^{mn} - X_T^n|^2\right] \leq K_{1,m} \frac{(m-1)T^2}{mn^2} \text{ and } \mathbb{E}\left[\max_{0 \leq k \leq n} |X_{\frac{kT}{n}}^{mn} - X_{\frac{kT}{n}}^n|^2\right] \leq K_{2,m} \frac{(m-1)T^2}{mn^2},$$

where $\sqrt{K_{1,m}} = C_{(4.9)}[\dot{b}]_\infty \sqrt{\frac{d(2m-1)}{6m}} + e^{[\dot{b}]_\infty T} \sqrt{\frac{m-1}{m}}$

$$\times \left(\frac{|b(x_0)|}{2} \left(a_{\Delta b} \frac{[\dot{b}]_\infty T - 1 + e^{-[\dot{b}]_\infty T}}{2[\dot{b}]_\infty^2} + [\dot{b}]_\infty T \right) + a_{\Delta b} \frac{1 - e^{-[\dot{b}]_\infty T}}{[\dot{b}]_\infty} + \frac{2}{3}\sqrt{d}([\dot{b}]_\infty^2 + a_{\Delta b})T^{3/2} \right)$$

and $K_{2,m}$ *is defined like* $K_{1,m}$ *but with* $C_{(4.9)} + \sqrt{T}$ *replacing the constant* $C_{(4.9)} := \sqrt{T} + [\dot{b}]_\infty \int_0^T e^{[\dot{b}]_\infty (T-t)} \sqrt{t} dt$.

In the estimations (and in the remaining of the paper), m only appears through ratios which have a limit as $m \to \infty$ and when $m = \infty$, we consider that they are equal to this limit. The proof is postponed to Section 4.

2.2 MLMC parameters optimization revisited

In what follows let us assume that $f \in \mathscr{C}^1(\mathbb{R}^d, \mathbb{R})$ is a Lipschitz continuous function with constant $[\dot{f}]_\infty$. For the Multilevel Monte Carlo estimator

$$\hat{Q} = \frac{1}{N_0} \sum_{k=0}^{N_0} f(X_{T,k}^1) + \sum_{\ell=1}^{L} \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) \right),$$

defined in (1.3), the expectation leads to a telescoping summation so that

$$\left| \mathbb{E}\left[f(X_T) - \hat{Q} \right] \right| = \left| \mathbb{E}\left[f(X_T) - f(X_T^{m^L}) \right] \right| \leq [\dot{f}]_\infty \mathbb{E}^{1/2}\left[\left| X_T - X_T^{m^L} \right|^2 \right] \leq [\dot{f}]_\infty \frac{T\sqrt{K_{1,\infty}}}{m^L}, \tag{2.2}$$

where we used Proposition 2.1 for the last inequality. On the other hand, again by Proposition 2.1,

$$\text{Var}\left[f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) \right] \leq [\dot{f}]_\infty^2 \mathbb{E}\left[\left| X_T^{m^\ell} - X_T^{m^{\ell-1}} \right|^2 \right] \leq [\dot{f}]_\infty^2 \frac{K_{1,m}(m-1)T^2}{m^{2\ell-1}}.$$

Last, as $X_T^1 \sim \mathcal{N}(x_0 + b(x_0)T, TI_d)$, we use the logarithmic Sobolev inequality for the Gaussian measure and the Herbst's argument (see e.g. propositions 5.5.1 and 5.4.1 in [1]) to get for all $\lambda \in \mathbb{R}$

$$\mathbb{E}\left[\exp\left(\lambda(f(X_T^1) - \mathbb{E}[f(X_T^1)]) \right) \right] \leq \exp\left(\frac{\lambda^2 [\dot{f}]_\infty^2 T}{2} \right). \tag{2.3}$$

By performing Taylor expansions as $\lambda \to 0$, we easily deduce that $\text{Var}\left[f(X_T^1) \right] \leq [\dot{f}]_\infty^2 T$. As a consequence, the following non-asymptotic estimation of the mean square error of $\hat{Q}$ holds.

Proposition 2.2. *Under* $(\mathbf{R1})$,

$$\mathbb{E}\left[\left(\hat{Q} - \mathbb{E}[f(X_T)] \right)^2 \right] \leq [\dot{f}]_\infty^2 \left(\frac{K_{1,\infty}T^2}{m^{2L}} + \frac{T}{N_0} + \sum_{\ell=1}^{L} \frac{K_{1,m}(m-1)T^2}{N_\ell m^{2\ell-1}} \right).$$

According to the above proposition, to achieve a root mean square error $\varepsilon > 0$, one should choose

$$\frac{[\dot{f}]_\infty T\sqrt{K_{1,\infty}}}{m^L} < \varepsilon \text{ i.e. } L \geq \lfloor \ln([\dot{f}]_\infty T\sqrt{K_{1,\infty}}/\varepsilon)/\ln(m) \rfloor + 1. \tag{2.4}$$

For such a choice, one should then choose $(N_\ell)_{0 \leq \ell \leq L}$ such that

$$\sum_{\ell=1}^{L} \frac{m+1}{N_\ell m^{2\ell+1}} + \frac{C_{(2.5)}^2}{N_0} \leq \frac{m+1}{K_{1,m} m^2(m-1)} \left(\frac{\varepsilon^2}{[\dot{f}]_\infty^2 T^2} - \frac{K_{1,\infty}}{m^{2L}} \right), \qquad (2.5)$$

where $C_{(2.5)} = \frac{1}{m}\sqrt{\frac{m+1}{K_{1,m}(m-1)T}}$ minimizing the computation cost which is equal to $N_0 + \sum_{\ell=1}^{L} N_\ell(m+1)m^{\ell-1}$. Note that for $\ell \in \{1, \ldots, L\}$, $(m+1)m^{\ell-1} = m^\ell + m^{\ell-1}$ is the number of grid values of the Euler schemes which are computed for each Brownian path at the level ℓ. This constrained minimization problem leads to $N_0 = N\frac{C_{(2.5)}}{C_{(2.5)}+\sum_{\ell=1}^{L} m^{-3\ell/2}}$ and $N_\ell = N\frac{m^{-3\ell/2}}{C_{(2.5)}+\sum_{\ell=1}^{L} m^{-3\ell/2}}$ where the total number N of simulations is chosen in order to achieve equality in (2.5) :

$$N = \left(C_{(2.5)} + \sum_{\ell=1}^{L} m^{-3\ell/2} \right) \left(C_{(2.5)} + \frac{m+1}{m} \times \frac{1 - m^{-L/2}}{\sqrt{m} - 1} \right)$$

$$\times \frac{K_{1,m} m^2(m-1)[\dot{f}]_\infty^2 T^2}{(m+1)(\varepsilon^2 - [\dot{f}]_\infty^2 T^2 K_{1,\infty} m^{-2L})}. \qquad (2.6)$$

Then the computation cost is given by $\text{Cost}(m, m^{-L})$ where

$$\text{Cost}(m, x) = \left(C_{(2.5)} + \frac{m+1}{m} \times \frac{1 - \sqrt{x}}{\sqrt{m} - 1} \right)^2 \frac{K_{1,m} m^2(m-1)[\dot{f}]_\infty^2 T^2}{(m+1)(\varepsilon^2 - [\dot{f}]_\infty^2 T^2 K_{1,\infty} x^2)}.$$

Notice that for fixed m, $\text{Cost}(m, x)$ is up to some positive multiplicative factor not depending on x equal to $g(x) = \frac{(\sqrt{\alpha} - \sqrt{x})^2}{\beta_\varepsilon^2 - x^2}$ with $\sqrt{\alpha} = 1 + \frac{m(\sqrt{m}-1)}{m+1} C_{(2.5)} > 1$ and $\beta_\varepsilon = \frac{\varepsilon}{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}$ not depending on x. We thus want to find $L \in \mathbb{N}$ minimizing $g(m^{-L})$ under the constraint (2.4) which writes $m^{-L} < \beta_\varepsilon$. We have $g'(x) = \frac{\sqrt{\alpha} - \sqrt{x}}{(\beta_\varepsilon^2 - x^2)^2} \left(2\sqrt{\alpha} x - x^{3/2} - \frac{\beta_\varepsilon^2}{\sqrt{x}} \right)$. Since, as $\alpha > 1$, $x \mapsto h(x) = 2\sqrt{\alpha} x - x^{3/2} - \frac{\beta_\varepsilon^2}{\sqrt{x}}$ is increasing on $(0, 1]$,

- either $2\sqrt{\alpha} - 1 - \beta_\varepsilon^2 \leq 0$, which implies $\beta_\varepsilon > 1$ and $\inf_{x \in [0,1]} g(x) = g(1)$ so that $L = 0$ solves the constrained minimization problem,
- or $2\sqrt{\alpha} - 1 - \beta_\varepsilon^2 > 0$ so that, since $\lim_{x \to 0^+} 2\sqrt{\alpha} x - x^{3/2} - \frac{\beta_\varepsilon^2}{\sqrt{x}} = -\infty$ and $2\sqrt{\alpha}\beta_\varepsilon - \beta_\varepsilon^{3/2} - \frac{\beta_\varepsilon^2}{\sqrt{\beta_\varepsilon}} = 2\beta_\varepsilon(\sqrt{\alpha} - \sqrt{\beta_\varepsilon}) > 2\beta_\varepsilon(1 - \sqrt{\beta_\varepsilon})$, there exists $x_\varepsilon^\star \in (0, 1 \wedge \beta_\varepsilon)$ such that g is decreasing on $(0, x_\varepsilon^\star)$ and increasing on $(x_\varepsilon^\star, 1 \wedge \beta_\varepsilon)$ and the value of L solving the constrained minimization problem belongs to $\{\lfloor -\frac{\ln x_\varepsilon^\star}{\ln m} \rfloor, \lceil -\frac{\ln x_\varepsilon^\star}{\ln m} \rceil\}$.

We denote by L^ε this optimal value of L, by N^ε (resp. N_ℓ^ε) the corresponding total number of samples (resp. number of samples in the level ℓ) and by $\hat{Q}_\varepsilon$ the multi-level Monte Carlo estimator (1.3) with those optimal parameters. Note that when $\varepsilon^2 < [\dot{f}]_\infty^2 T^2 K_{1,\infty} \left(1 + \frac{2m(\sqrt{m}-1)C_{(2.5)}}{m+1} \right)$ which is equivalent to $2\sqrt{\alpha} - 1 - \beta_\varepsilon^2 > 0$, since $h\left(\frac{\beta_\varepsilon^{4/3}}{2^{2/3}\alpha^{1/3}} \right) = -\frac{\beta_\varepsilon^2}{2\sqrt{\alpha}} < 0$ and $\inf_{x \in (0,1]} h'(x) \geq 2\sqrt{\alpha} - \frac{3}{2} > 0$, we have $\frac{\beta_\varepsilon^{4/3}}{2^{2/3}\alpha^{1/3}} < x_\varepsilon^\star < \frac{\beta_\varepsilon^{4/3}}{2^{2/3}\alpha^{1/3}} + \frac{1}{2\sqrt{\alpha} - 3/2} \times \frac{\beta_\varepsilon^2}{2\sqrt{\alpha}}$. Hence, as $\varepsilon \to 0$, $x_\varepsilon^\star \sim \frac{\beta_\varepsilon^{4/3}}{2^{2/3}\alpha^{1/3}}$, $L^\varepsilon \sim \frac{4\ln(1/\varepsilon)}{3\ln m}$ and the bias term $[\dot{f}]_\infty T\sqrt{K_{1,\infty}} m^{-L^\varepsilon}$ behaves as $\mathcal{O}(\varepsilon^{4/3})$. More precisely, when $\varepsilon < c_1 := [\dot{f}]_\infty T\sqrt{K_{1,\infty}} \left(1 + \frac{2m(\sqrt{m}-1)C_{(2.5)}}{m+1} \right)^{1/2}$,

$$m^{-L^\varepsilon} \geq e^{-\lceil -\frac{\ln x_\varepsilon^\star}{\ln m} \rceil \ln m} > \frac{x_\varepsilon^\star}{m} > \frac{\varepsilon^{4/3}}{2^{2/3} m([\dot{f}]_\infty T\sqrt{K_{1,\infty}})^{4/3}(1 + \frac{m(\sqrt{m}-1)}{m+1}C_{(2.5)})^{2/3}}, \qquad (2.7)$$

$$m^{-L^\varepsilon} \le e^{-\lfloor -\frac{\ln x^\star_\varepsilon}{\ln m} \rfloor \ln m} < m x^\star_\varepsilon < \frac{m\varepsilon^{4/3}}{(1 + \frac{m(\sqrt{m}-1)}{m+1} C_{(2.5)})([\dot{f}]_\infty T \sqrt{K_{1,\infty}})^{4/3}}$$

$$\times \left(\frac{(1 + \frac{m(\sqrt{m}-1)}{m+1} C_{(2.5)})^{1/3}}{2^{2/3}} + \frac{(1 + \frac{2m(\sqrt{m}-1)}{m+1} C_{(2.5)})^{1/3}}{1 + \frac{4m(\sqrt{m}-1)}{m+1} C_{(2.5)}} \right).$$

(2.8)

We easily deduce that, as expected from [4], $\mathrm{Cost}(m, m^{-L^\varepsilon}) = \mathcal{O}(\varepsilon^{-2})$ as $\varepsilon \to 0$ for fixed m and $N^\varepsilon = \mathcal{O}(\varepsilon^{-2})$. One could also consider minimizing $m \mapsto \mathrm{Cost}(m, L^\varepsilon(m))$ numerically, where we used the notation $L^\varepsilon(m)$ to make the dependence on m explicit.

Remark 2.3. For this optimal choice which clearly differs from the one in [4], the bias of the Multilevel Monte Carlo method is not of the same order of magnitude as the precision ε but much smaller. To the best of our knowledge, such a $4/3$ order of convergence of the bias does not appear in the existing multilevel Monte Carlo methods literature. Notice that, for stochastic differential equations with a non constant diffusion coefficient (multiplicative noise), this property remains true for the multilevel Monte Carlo estimator based on the Giles and Szpruch scheme [5], since it exhibits the same orders of convergence of the bias and the variance within a given level as (1.3).

3 Concentration bounds for the MLMC Euler method

The main result of this paper is a concentration inequality for the Multilevel Monte Carlo estimator $\hat{Q}$ defined in (1.3). To prove this result, we are going to estimate the moment generating function of $\hat{Q} - \mathbb{E}[f(X_T^{m^L})] = \sum_{\ell=0}^{L} \hat{Q}_\ell$, where $\hat{Q}_0 := \frac{1}{N_0} \sum_{k=1}^{N_0} f(X_{T,k}^1) - \mathbb{E}f(X_T^1)$ and, for $\ell \ge 1$,

$$\hat{Q}_\ell := \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) - \mathbb{E}[f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})] \right).$$

3.1 Estimation of the moment generating function

We are first going to derive an exponential type upper bound with the optimal rate of convergence for the moment generating function of the square of the error $U_T = X_T^n - X_T^{mn}$ between the Euler schemes with n and mn steps. The proof of the following result is postponed to Section 4.

Theorem 3.1. Let $(n, m) \in \mathbb{N}^* \times \bar{\mathbb{N}}$, $t_k = \frac{kT}{n}$ for $k \in \{0, \dots, n\}$ and ρ be a constant satisfying

$$0 \le \rho \le \frac{9mn^2}{4T^2(m-1)\left(C_{(4.9)}[\dot{b}]_\infty \sqrt{3d(2m-1)/m} + C_{(4.11)}T^{3/2}\sqrt{2(m-1)/m} \right)^2} := \rho_{(3.1)} n^2,$$

(3.1)

where $C_{(4.9)} = \sqrt{T} + [\dot{b}]_\infty \int_0^T e^{[\dot{b}]_\infty (T-t)} \sqrt{t} dt$ with $C_{(4.11)} = e^{T[\dot{b}]_\infty}([\dot{b}]_\infty^2 + a_{\Delta b})$ and by convention $\frac{m}{m-1} = \frac{m-1}{m} = 1$ and $\frac{2m-1}{m} = 2$ when $m = \infty$. Under assumption $(\mathbf{R1})$, we have for all $x \ge 0$

$$\mathbb{E}\left[\exp\left\{ \rho\left(\frac{(m-1)T^2 x}{2mn} + \max_{1 \le k \le n} |X_{t_k}^{mn} - X_{t_k}^n| \right)^2 \right\} \right] \le \exp\left\{ \rho C_{(3.2)}(x) \frac{(m-1)T^2}{mn^2} \right\}, \quad \text{with}$$

(3.2)

$$C_{(3.2)}(x) = \left(C_{(4.9)}[\dot{b}]_\infty \sqrt{\frac{3d(2m-1)}{m}} + C_{(4.11)}T^{3/2}\sqrt{\frac{2(m-1)}{m}} \right)$$

$$\times \left(C_{(4.9)}[\dot{b}]_\infty \frac{2\ln 2}{3\sqrt{3}} \sqrt{\frac{d(2m-1)}{m}} + C_{(4.11)} \sqrt{\frac{(m-1)T}{2m}} \right.$$
$$\left. \times \left(\frac{(3d+1)(|b(x_0)| + [\dot{b}]_\infty + \tilde{x})^2}{4[\dot{b}]_\infty^2} + \frac{4d\sqrt{T}(|b(x_0)| + [\dot{b}]_\infty + \tilde{x})}{3\sqrt{\pi}[\dot{b}]_\infty} + \frac{4dT\ln 2}{9} \right) \right)$$

and $\tilde{x} = \frac{[\dot{b}]_\infty e^{-[b]_\infty T}}{([\dot{b}]_\infty^2 + a_{\Delta b})} x.$

To derive our main result, we need to reinforce our assumption (**R1**) since our approach relies on Malliavin calculus that requires additional smoothness on the coefficient b.

Assumption (R2) The function $b \in \mathscr{C}^3(\mathbb{R}^d, \mathbb{R}^d)$ and satisfies assumption (**R1**). Moreover, there exist finite constants $[\ddot{b}]_\infty \in (0, +\infty)$ and $a_{\nabla \Delta b} \in [0, +\infty)$ such that

$$\forall j \in \{1, \dots, d\}, \ \forall x \in \mathbb{R}^d, \ \left\| \frac{\partial \nabla b}{\partial x_j}(x) \right\| \leq [\ddot{b}]_\infty$$
$$\forall x \in \mathbb{R}^d, \ \|\nabla \Delta b(x)\| \leq 2a_{\nabla \Delta b}(1 + |x - x_0|). \tag{3.3}$$

To state our next results we introduce the following finite quantities.

Constants Notations (CN)

$$\rho_{(5.14)} := \frac{m^2}{2C_{(5.13)}^2 T^2 (m-1)^2}, \quad \rho_{(5.16)} := \frac{3n^2}{4T^2 d[\ddot{b}]_\infty^2 (2m-1)(m-1)},$$

$$C_{(5.13)} := (\sqrt{d}[\dot{b}]_\infty [\ddot{b}]_\infty \vee [\dot{b}]_\infty^2 + a_{\nabla \Delta b}),$$

$$\Phi_1(r) := \frac{\sqrt{d}[\ddot{b}]_\infty}{[\dot{b}]_\infty}(e^{[\dot{b}]_\infty (T-r)} - 1), \quad \Phi_2(r) := \int_r^T e^{[\dot{b}]_\infty s} \sqrt{s}\, ds,$$

$$\Phi_3(r) := \sqrt{\frac{1 - e^{-2[\dot{b}]_\infty (T-r)}}{2[\dot{b}]_\infty}} + \sqrt{\frac{[\dot{b}]_\infty}{2}} \int_r^T \sqrt{1 - e^{-2[\dot{b}]_\infty (t-r)}}\, dt,$$

$$\hat{\rho}(r) := \frac{\frac{\rho_{(3.1)}}{\Phi_1^2(r)} \frac{\rho_{(5.14)}}{\Phi_2^2(r)} \frac{\rho_{(5.16)}}{\Phi_3^2(r)}}{\left(\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)} \frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} + \frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)} \frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)} + \frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} \frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)} \right)^2},$$

$$\Phi(r, x) := \left(\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)} \frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} + \frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)} \frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)} + \frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} \frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)} \right)$$
$$\times \left(\frac{\Phi_3^2(r) C_{(5.16)}}{\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)} \frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)}} + \frac{\phi_2(r, x)}{\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)} \frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}} + \frac{\Phi_1^2(r) C_{(3.2)}(0)}{\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} \frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}} \right), \quad x \geq 0,$$

$$\phi_2(r, x) := \frac{(m-1)}{m} \left((3d+1) \left(C_{(5.13)} \frac{|b(x_0)| + [\dot{b}]_\infty}{2[\dot{b}]_\infty^2}(e^{[\dot{b}]_\infty T} - e^{[\dot{b}]_\infty r}) + x \right)^2 \right.$$
$$+ \frac{4dC_{(5.13)}\Phi_2(r)}{\sqrt{\pi}} \left(C_{(5.13)} \frac{|b(x_0)| + [\dot{b}]_\infty}{2[\dot{b}]_\infty^2}(e^{[\dot{b}]_\infty T} - e^{[\dot{b}]_\infty r}) + x \right) + d\ln 2 C_{(5.13)}^2 \Phi_2^2(r) \right), x \geq 0,$$

$$C_{(6.7)} := dT \int_0^T \frac{e^{2[\dot{b}]_\infty (T-t)}}{\hat{\rho}(t)} \left(\sqrt{\rho_{(3.1)}} + \frac{[\dot{f}]_{\mathrm{lip}}}{[\dot{f}]_\infty} \sqrt{\hat{\rho}(t)} \right)^2 dt$$
$$\times \sup_{r \in [0,T)} \frac{[\dot{f}]_{\mathrm{lip}} \sqrt{\hat{\rho}(r)} C_{(3.2)}(2[\dot{b}]_\infty [\dot{f}]_\infty / T[\dot{f}]_{\mathrm{lip}}) + [\dot{f}]_\infty \hat{\rho}(r) \Phi(r, 0)/\sqrt{\rho_{(3.1)}}}{[\dot{f}]_\infty \sqrt{\rho_{(3.1)}} + [\dot{f}]_{\mathrm{lip}} \sqrt{\hat{\rho}(r)}}.$$

The proof of the following theorem is postponed to Section 6.

Theorem 3.2. *Let assumption* $(\mathbf{R2})$ *hold and* $f \in \mathscr{C}^1(\mathbb{R}^d, \mathbb{R})$ *be a Lipschitz continuous function with constant* $[\dot{f}]_\infty$ *and such that* ∇f *is also Lipchitz with constant* $[\dot{f}]_{\mathrm{lip}}$. *For all* $\lambda \leq \mathcal{C} \min_{1 \leq \ell \leq L} N_\ell m^\ell$, *where*

$$\mathcal{C} = \left(\frac{\rho_{(3.1)}}{2dm^2 T \int_0^T \frac{e^{2[b]_\infty (T-t)}}{\hat{\rho}(t)} \left([\dot{f}]_\infty \sqrt{\rho_{(3.1)}} + [\dot{f}]_{\mathrm{lip}} \sqrt{\hat{\rho}(t)} \right)^2 dt} \right)^{1/2},$$

we have

$$\mathbb{E}\left[\exp\left(\lambda[\hat{Q} - \mathbb{E}f(X_T^{m^L})] \right) \right] \leq \exp\left\{ \lambda^2 [\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right) \right\}.$$

According to Section 2, the bias satisfies

$$\left| \mathbb{E}f(X_T^{m^L}) - \mathbb{E}f(X_T) \right| \leq \mathbb{E}^{1/2} |f(X_T^{m^L}) - f(X_T)|^2 \leq \frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L},$$

so we easily deduce the following corollary.

Corollary 3.3. *Under assumptions of Theorem 3.2, we have* $\forall |\lambda| \leq \mathcal{C} \min_{1 \leq \ell \leq L} N_\ell m^\ell$,

$$\mathbb{E}\left[\exp\left(\lambda[\hat{Q} - \mathbb{E}f(X_T)] \right) \right] \leq$$
$$\exp\left\{ \lambda^2 [\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right) + |\lambda| \frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} \right\}.$$

3.2 Concentration bounds

Using the above corollary, for all $\lambda \in [0, \mathcal{C} \min_{1 \leq \ell \leq L} m^\ell N_\ell]$ and $\alpha \geq 0$, we get

$$\mathbb{P}\left(\hat{Q} - \mathbb{E}f(X_T) \geq \alpha \right) \leq \exp\{\psi_\alpha(\lambda)\}, \tag{3.4}$$

with

$$\psi_\alpha(\lambda) := \lambda^2 [\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right) + \lambda \left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} - \alpha \right).$$

Now, when $0 \leq \alpha \leq 2[\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right) \mathcal{C} \min_{1 \leq \ell \leq L} m^\ell N_\ell + \frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L}$,

$$\min_{\lambda \in [0, \mathcal{C} \min_{1 \leq \ell \leq L} m^\ell N_\ell]} \psi_\alpha(\lambda) = - \frac{\left(\alpha - \frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} \right)_+^2}{4[\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right)},$$

where $(x)_+ = \max(x, 0)$ and otherwise

$$\min_{\lambda \in [0, \mathcal{C} \min_{1 \leq \ell \leq L} m^\ell N_\ell]} \psi_\alpha(\lambda) = \mathcal{C} \min_{1 \leq \ell \leq L} m^\ell N_\ell$$
$$\times \left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} - \alpha + [\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right) \mathcal{C} \min_{1 \leq \ell \leq L} m^\ell N_\ell \right)$$
$$< - \frac{\mathcal{C} \min_{1 \leq \ell \leq L} m^\ell N_\ell}{2} \alpha.$$

Dealing with $\mathbb{P}\left(\hat{Q} - \mathbb{E}f(X_T) \leq -\alpha\right)$ in a symmetric way we end up with the concentration inequality,

$$\forall 0 \leq \alpha \leq 2[\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}}\right) \mathcal{C} \min_{1\leq\ell\leq L} m^\ell N_\ell + \frac{[\dot{f}]_\infty T\sqrt{K_{1,\infty}}}{m^L},$$

$$\mathbb{P}\left(|\hat{Q} - \mathbb{E}f(X_T)| \geq \alpha\right) \leq 2\exp\left(-\frac{\left(\alpha - \frac{[\dot{f}]_\infty T\sqrt{K_{1,\infty}}}{m^L}\right)^2}{4[\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}}\right)}\right)$$

$$\forall \alpha \geq 2[\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}}\right) \mathcal{C} \min_{1\leq\ell\leq L} m^\ell N_\ell + \frac{[\dot{f}]_\infty T\sqrt{K_{1,\infty}}}{m^L},$$

$$\mathbb{P}\left(|\hat{Q} - \mathbb{E}f(X_T)| \geq \alpha\right) \leq 2\exp\left(-\frac{\mathcal{C}\min_{1\leq\ell\leq L} m^\ell N_\ell}{2}\alpha\right).$$

Hence, we proved our main result, which we now state.

Theorem 3.4. *Under assumptions of Theorem 3.2, the multilevel Monte Carlo estimator* (1.3) *satisfies,* $\forall\ 0 \leq \alpha \leq 2[\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}}\right) \mathcal{C} \min_{1\leq\ell\leq L} m^\ell N_\ell + \frac{[\dot{f}]_\infty T\sqrt{K_{1,\infty}}}{m^L}$,

$$\mathbb{P}\left(|\hat{Q} - \mathbb{E}f(X_T)| \geq \alpha\right) \leq 2\exp\left(-\frac{\left(\alpha - \frac{[\dot{f}]_\infty T\sqrt{K_{1,\infty}}}{m^L}\right)_+^2}{2[\dot{f}]_\infty^2 \left(\frac{T}{N_0} + \sum_{\ell=1}^L \frac{2C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}}\right)}\right),$$

with $(x)_+ = \max(x, 0)$.

Notice that the factor $[\dot{f}]_\infty^2 \left(\frac{T}{N_0} + \sum_{\ell=1}^L \frac{2C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}}\right)$ in the denominator is closely related to the non-asymptotic upper-bound $[\dot{f}]_\infty^2 \left(\frac{T}{N_0} + \sum_{\ell=1}^L \frac{K_{1,m}(m-1)T^2}{N_\ell m^{2\ell-1}}\right)$ of the variance of $\hat{Q}$ derived in Section 2.2. The only difference is the replacement of $K_{1,m}$ by $2C_{(6.7)}$. Let us now discuss the constraint on α under which we proved Gaussian type concentration and see that in the limit $\varepsilon \to 0$, for the optimal parameters discussed in Section 2.2, we can choose $\alpha = \mathcal{O}(\varepsilon^{2/3})$ i.e. much larger than the root mean square error ε.

Following the discussion and notations of Section 2.2, for $\varepsilon > 0$, we consider $\hat{Q}_\varepsilon$ the MLMC estimator (1.3) with the optimal parameters L^ε, N^ε,

$$N_0^\varepsilon = N^\varepsilon \frac{C_{(2.5)}}{C_{(2.5)} + \sum_{\ell=1}^{L^\varepsilon} m^{-3\ell/2}} \text{ and } N_\ell^\varepsilon = N^\varepsilon \frac{m^{-3\ell/2}}{C_{(2.5)} + \sum_{\ell=1}^{L^\varepsilon} m^{-3\ell/2}}, \ell \geq 1. \tag{3.5}$$

One has $\min_{1\leq\ell\leq L^\varepsilon} m^\ell N_\ell^\varepsilon = \frac{N^\varepsilon m^{-L^\varepsilon/2}}{C_{(2.5)} + \sum_{\ell=1}^{L^\varepsilon} m^{-3\ell/2}}$ and therefore $\frac{1}{N_0^\varepsilon} \min_{1\leq\ell\leq L^\varepsilon} m^\ell N_\ell^\varepsilon = \frac{m^{-L^\varepsilon/2}}{C_{(2.5)}}$. As a consequence,

$$\frac{[\dot{f}]_\infty T\sqrt{K_{1,\infty}}}{m^{L^\varepsilon}} + 2[\dot{f}]_\infty^2 \left(\frac{T}{2N_0^\varepsilon} + \sum_{\ell=1}^{L^\varepsilon} \frac{C_{(6.7)}(m-1)T^2}{N_\ell^\varepsilon m^{2\ell-1}}\right) \mathcal{C} \min_{1\leq\ell\leq L^\varepsilon} m^\ell N_\ell^\varepsilon \geq \frac{\mathcal{C}[\dot{f}]_\infty^2 T}{C_{(2.5)}} m^{-L^\varepsilon/2}$$

where, according to (2.7), when $\varepsilon < c_1 := [\dot{f}]_\infty T\sqrt{K_{1,\infty}}\left(1 + \frac{2m(\sqrt{m}-1)C_{(2.5)}}{m+1}\right)^{1/2}$, the right-hand side is larger than

$$c_2\varepsilon^{2/3} \text{ with } c_2 := \frac{\mathcal{C}[\dot{f}]_\infty^2 T}{2^{1/3}C_{(2.5)}\sqrt{m}([\dot{f}]_\infty T\sqrt{K_{1,\infty}})^{2/3}(1 + \frac{m(\sqrt{m}-1)}{m+1}C_{(2.5)})^{1/3}}.$$

Under the same condition on ε, according to (2.8),

$$[\dot{f}]_\infty T\sqrt{K_{1,\infty}}m^{-L^\varepsilon} < \frac{m\varepsilon^{4/3}}{(1+\frac{m(\sqrt{m}-1)}{m+1}C_{(2.5)})([\dot{f}]_\infty T\sqrt{K_{1,\infty}})^{1/3}}$$
$$\times\left(\frac{(1+\frac{m(\sqrt{m}-1)}{m+1}C_{(2.5)})^{1/3}}{2^{2/3}} + \frac{(1+\frac{2m(\sqrt{m}-1)}{m+1}C_{(2.5)})^{1/3}}{1+\frac{4m(\sqrt{m}-1)}{m+1}C_{(2.5)}}\right).$$

On the other hand, one has

$$2[\dot{f}]_\infty^2\left(\frac{T}{N_0^\varepsilon} + \sum_{\ell=1}^{L^\varepsilon}\frac{2C_{(6.7)}(m-1)T^2}{N_\ell^\varepsilon m^{2\ell-1}}\right) \tag{3.6}$$

$$\leq \left(2\vee\frac{4C_{(6.7)}}{K_{1,m}}\right)[\dot{f}]_\infty^2\left(\frac{K_{1,\infty}T^2}{m^{2L^\varepsilon}} + \frac{T}{N_0^\varepsilon} + \sum_{\ell=1}^{L^\varepsilon}\frac{K_{1,m}(m-1)T^2}{N_\ell^\varepsilon m^{2\ell-1}}\right) \leq \left(2\vee\frac{4C_{(6.7)}}{K_{1,m}}\right)\varepsilon^2, \tag{3.7}$$

according to the optimization of parameters which follows Proposition 2.2. Combining the two last inequalities and the fact that for positive α and x, $(\alpha-x)_+^2 \geq \frac{\alpha^2}{2}-x^2$, we obtain that for $\varepsilon < c_1$,

$$\frac{\left(\alpha-\frac{[\dot{f}]_\infty T\sqrt{K_{1,\infty}}}{m^{L^\varepsilon}}\right)_+^2}{2[\dot{f}]_\infty^2\left(\frac{T}{N_0^\varepsilon}+\sum_{\ell=1}^{L^\varepsilon}\frac{2C_{(6.7)}(m-1)T^2}{N_\ell^\varepsilon m^{2\ell-1}}\right)} \geq \frac{\alpha^2}{c_3\varepsilon^2} - c_4\varepsilon^{2/3},$$

with

$$c_3 = \left(4\vee\frac{8C_{(6.7)}}{K_{1,m}}\right)$$

$$c_4 = \frac{2}{c_3}\left(\frac{m}{(1+\frac{m(\sqrt{m}-1)}{m+1}C_{(2.5)})([\dot{f}]_\infty T\sqrt{K_{1,\infty}})^{1/3}}\right.$$
$$\left.\times\left(\frac{(1+\frac{m(\sqrt{m}-1)}{m+1}C_{(2.5)})^{1/3}}{2^{2/3}} + \frac{(1+\frac{2m(\sqrt{m}-1)}{m+1}C_{(2.5)})^{1/3}}{1+\frac{4m(\sqrt{m}-1)}{m+1}C_{(2.5)}}\right)\right)^2.$$

Then

$$\forall\varepsilon\in(0,c_1),\ \forall\alpha\in(0,c_2\varepsilon^{2/3}),\ \mathbb{P}\left(|\hat{Q}_\varepsilon - \mathbb{E}f(X_T)|\geq\alpha\right) \leq 2e^{-\frac{\alpha^2}{c_3\varepsilon^2}+c_4\varepsilon^{2/3}}.$$

Notice that the fact $[\dot{f}]_\infty T\sqrt{K_{1,\infty}}m^{-L^\varepsilon} < \varepsilon$ that the bias is smaller than the precision ε leads, by a similar reasoning, to the following result.

Corollary 3.5. *Under assumptions of Theorem 3.2, the multilevel Monte Carlo estimator $\hat{Q}_\varepsilon$ (1.3) equipped with the optimal parameters (3.5) satisfies*

$$\forall\varepsilon\in(0,c_1),\ \forall\alpha\in(0,c_2\varepsilon^{2/3}),\ \mathbb{P}\left(|\hat{Q}_\varepsilon - \mathbb{E}f(X_T)|\geq\alpha\right) \leq 2e^{-\frac{\alpha^2}{c_3\varepsilon^2}}e^{\frac{2}{c_3}\wedge c_4\varepsilon^{2/3}}.$$

At this stage, a natural question arises: is there an alternative choice of the parameters that does not increase neither the root mean square error ε nor the order in ε of the computational cost of the multilevel Monte Carlo estimator and for which the upper bound on the deviation parameter α is larger than $c_1\varepsilon^{2/3}$?

For $\beta > 1$, we set $N_\ell^{\varepsilon,\beta} = N_\ell^\varepsilon \times \frac{m^{\frac{\ell-1}{2}}}{m^{\frac{\ell-1}{2}}\wedge\ell^\beta}$ for $\ell\in\{1,\dots,L^\varepsilon\}$ and

$$\hat{Q}_{\varepsilon,\beta} = \frac{1}{N_0^\varepsilon}\sum_{k=0}^{N_0^\varepsilon}f(X_{T,k}^1) + \sum_{\ell=1}^{L^\varepsilon}\frac{1}{N_\ell^{\varepsilon,\beta}}\sum_{k=1}^{N_\ell^{\varepsilon,\beta}}\left(f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}})\right). \tag{3.8}$$

Since for each ℓ, $N_\ell^{\varepsilon,\beta} \geq N_\ell^\varepsilon$, the root mean square error and the statistical error of $\hat{Q}_{\varepsilon,\beta}$ are not greater than the ones of $\hat{Q}_\varepsilon$ and therefore than ε. The two estimators share the same bias. Moreover, $\min_{1 \leq \ell \leq L^\varepsilon} m^\ell N_\ell^{\varepsilon,\beta} \geq \frac{N^\varepsilon m^{-1/2}}{(L^\varepsilon)^\beta (C_{(2.5)} + \sum_{\ell=1}^{L^\varepsilon} m^{-3\ell/2})}$ so that $\frac{1}{N_0^\varepsilon} \min_{1 \leq \ell \leq L^\varepsilon} m^\ell N_\ell^{\varepsilon,\beta} \geq \frac{1}{C_{(2.5)}\sqrt{m}(L^\varepsilon)^\beta}$. As a consequence,

$$\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^{L^\varepsilon}} + 2[\dot{f}]_\infty^2 \left(\frac{T}{2N_0^\varepsilon} + \sum_{\ell=1}^{L^\varepsilon} \frac{C_{(6.7)}(m-1)T^2}{N_\ell^{\varepsilon,\beta} m^{2\ell-1}} \right) \mathcal{C} \min_{1 \leq \ell \leq L^\varepsilon} m^\ell N_\ell^{\varepsilon,\beta} \geq \frac{\mathcal{C}[\dot{f}]_\infty^2 T}{C_{(2.5)}\sqrt{m}(L^\varepsilon)^\beta}$$

where, when $\varepsilon < c_1$, the right-hand side is larger than

$$c_5(\varepsilon) := \frac{\mathcal{C}[\dot{f}]_\infty^2 T}{C_{(2.5)}\sqrt{m}} \left(1 + \frac{4}{3\ln m} \ln \left(\frac{[\dot{f}]_\infty T \sqrt{2K_{1,\infty}(1 + \frac{m(\sqrt{m}-1)}{m+1}C_{(2.5)})}}{\varepsilon} \right) \right)^{-\beta}.$$

Reasoning like in the above derivation of concentration inequalities for $\hat{Q}_\varepsilon$, we easily get the following result.

Corollary 3.6. *Under assumptions of Theorem 3.2, the multilevel Monte Carlo estimator $\hat{Q}_{\varepsilon,\beta}$ (3.8) satisfies*

$$\forall \varepsilon \in (0, c_1), \ \forall \alpha \in (0, c_5(\varepsilon)), \ \mathbb{P}\left(|\hat{Q}_{\varepsilon,\beta} - \mathbb{E}f(X_T)| \geq \alpha \right) \leq 2e^{-\frac{\alpha^2}{c_3\varepsilon^2} e^{\frac{2}{c_3} \wedge c_4 \varepsilon^{2/3}}}.$$

Moreover, the computational cost of $\hat{Q}_{\varepsilon,\beta}$ is proportional to

$$N_0^\varepsilon + \frac{m+1}{m} \sum_{\ell=1}^{L^\varepsilon} m^\ell N_\ell^{\varepsilon,\beta} = \frac{N^\varepsilon}{C_{(2.5)} + \sum_{\ell=1}^{L^\varepsilon} m^{-3\ell/2}} \left(C_{(2.5)} + \frac{m+1}{m\sqrt{m}} \sum_{\ell=1}^{L^\varepsilon} (m^{\frac{\ell-1}{2}} \wedge \ell^\beta)^{-1} \right).$$

Since $\beta > 1$, it is of the same order $\mathcal{O}(\varepsilon^{-2})$ as N^ε and therefore as the computational cost of $\hat{Q}_\varepsilon$ in the limit $\varepsilon \to 0$.

3.3 Error control in Orlicz norm

In view of our previous results, it is natural and fruitful to generalize the RMS error analysis developed in Section 2.2 by means of Young functions that are increasing convex functions $\Psi : \mathbb{R}^+ \to \mathbb{R}^+$ satisfying $\Psi(0) = 0$ and $\lim_{x \to +\infty} \Psi(x) = +\infty$. For a given Young function Ψ the associated Orlicz norm $\|X\|_\Psi$ of a random variable X is defined by

$$\|X\|_\Psi := \inf\{c > 0, \ : \mathbb{E}[\Psi(X/c)] \leq 1\} \text{ with } \inf \emptyset = \infty.$$

We set $\Psi_e(x) := (e^x - 1)/(e - 1)$ as a fixed Young function. At first, let us deal with a standard Monte Carlo algorithm that approximates $\mathbb{E}[f(X_T)]$ by $\bar{Q} := \frac{1}{N_0} \sum_{k=1}^{N_0} f(X_{T,k}^{m^L})$. Then, one can use Section 4 of [3] to bound $\mathbb{E}\left[\exp(\lambda(\bar{Q} - \mathbb{E}[f(X_T^{m^L})])) \right]$ from above. More precisely, taking advantage of the constant diffusion coefficient to improve the bound in Proposition 4.1 of [3] to $[f_i^\Delta]_1 \leq [\dot{f}]_\infty e^{[\dot{b}]_\infty(T-t_i)}$, we get

$$\mathbb{E}\left[\exp(\lambda(\bar{Q} - \mathbb{E}[f(X_T^{m^L})])) \right] \leq \exp\left\{ \frac{\lambda^2 C_{(3.9)}}{N_0} \right\} \text{ with } C_{(3.9)} = \frac{T[\dot{f}]_\infty^2}{2m^L} \sum_{k=1}^{m^L} e^{2[\dot{b}]_\infty(T - \frac{Tk}{m^L})}. \tag{3.9}$$

With (2.2), we easily deduce that for all $c > 0$

$$\mathbb{E}\left[\Psi_e \left(\frac{\bar{Q} - \mathbb{E}[f(X_T)]}{c} \right) \right] \leq \frac{1}{e-1} \left[\exp\left(\frac{C_{(3.9)}}{c^2 N_0} + \frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{cm^L} \right) - 1 \right]$$

and then

$$\left\| \bar{Q} - \mathbb{E}[f(X_T)] \right\|_{\Psi_e} \leq \frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{2m^L} + \sqrt{\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{2m^L} \right)^2 + \frac{C_{(3.9)}}{N_0}}.$$

Recall that the RMS error in this case is given by

$$\mathbb{E}^{1/2}\left[(\bar{Q} - \mathbb{E}[f(X_T)])^2 \right] \leq \sqrt{\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} \right)^2 + \mathrm{Var}(\bar{Q}_0)}.$$

By performing Taylor expansions as $\lambda \to 0$ in (3.9), we easily get $\mathrm{Var}\left[\bar{Q} \right] \leq \frac{2C_{(3.9)}}{N_0}$. As a consequence, we get

$$\mathbb{E}^{1/2}\left[(\bar{Q} - \mathbb{E}[f(X_T)])^2 \right] \leq \sqrt{\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} \right)^2 + \frac{2C_{(3.9)}}{N_0}}. \qquad (3.10)$$

Note that using

$$\sqrt{\frac{a^2 + b}{2}} \leq \frac{a}{2} + \sqrt{\left(\frac{a}{2} \right)^2 + \frac{b}{2}} \leq \sqrt{\frac{(3 + \sqrt{3})}{4} (a^2 + b)}, \text{ for } a, b > 0, \qquad (3.11)$$

we easily get

$$\sqrt{\frac{\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} \right)^2 + \frac{2C_{(3.9)}}{N_0}}{2}} \leq \left\| \bar{Q} - \mathbb{E}[f(X_T)] \right\|_{\Psi_e} \leq$$
$$\sqrt{\frac{(3 + \sqrt{3})}{4} \left[\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} \right)^2 + \frac{2C_{(3.9)}}{N_0} \right]}. \quad (3.12)$$

Then the RMS error and the Orlicz norm of the standard Monte Carlo share, up to a constant factor, the same order.

Now, we proceed similarly for the multilevel Monte Carlo algorithm and thanks to Corollary 3.3 we write, for all $c \geq 1/(\mathcal{C} \min_{1 \leq \ell \leq L} N_\ell m^\ell)$

$$\mathbb{E}\left[\Psi_e \left(\frac{\hat{Q} - \mathbb{E}[f(X_T)]}{c} \right) \right] \leq \frac{1}{e - 1}$$
$$\times \left[\exp\left(\frac{[\dot{f}]_\infty^2}{c^2} \left(\frac{T}{2N_0} + \sum_{\ell=1}^{L} \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right) + \frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{cm^L} \right) - 1 \right].$$

Hence, if

$$\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{2m^L} + \sqrt{\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{2m^L} \right)^2 + [\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^{L} \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right)} \right) \geq$$
$$\frac{1}{\mathcal{C} \min_{1 \leq \ell \leq L} N_\ell m^\ell}, \quad (3.13)$$

then

$$\left\| \hat{Q} - \mathbb{E}[f(X_T)] \right\|_{\Psi_e} \leq$$

$$\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{2m^L} + \sqrt{\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{2m^L} \right)^2 + [\dot{f}]_\infty^2 \left(\frac{T}{2N_0} + \sum_{\ell=1}^L \frac{C_{(6.7)}(m-1)T^2}{N_\ell m^{2\ell-1}} \right)} \right)$$

and thus by (3.11), we get

$$\left\| \hat{Q} - \mathbb{E}[f(X_T)] \right\|_{\Psi_e} \leq c_6 \sqrt{\left(\frac{[\dot{f}]_\infty T \sqrt{K_{1,\infty}}}{m^L} \right)^2 + [\dot{f}]_\infty^2 \left(\frac{T}{N_0} + \sum_{\ell=1}^L \frac{K_{1,m}(m-1)T^2}{N_\ell m^{2\ell-1}} \right)},$$

(3.14)

where $c_6 := \left[\left(\frac{3+\sqrt{3}}{4} \right) \left(1 \vee \frac{2C_{(6.7)}}{K_{1,m}} \right) \right]^{1/2}$.

Note that the above upper bound is equal, up to a constant factor, to the upper bound given by the RMS error estimate of Proposition 2.2. Combining this result with (3.10), (3.12) and (3.14), we conclude that compared to the standard Monte Carlo method, under constraint (3.13), the multilevel Monte Carlo estimator achieves the same complexity reduction for the Orlicz norm $\| \cdot \|_{\Psi_e}$ as for the RMS error.

Now, it remains to check the validity of constraint (3.13) when choosing the multilevel Monte Carlo algorithm optimal parameters derived in Section 2.2. Let us recall that in this setting the RMS error upper bound is equal to ε. On the one hand, using (3.11), we deduce that the term in the left hand side of (3.13) is larger than $c_6'\varepsilon$, where $c_6' := \left[\left(\frac{1}{2} \wedge \frac{C_{(6.7)}}{K_{1,m}} \right) \right]^{1/2}$. On the other hand, combining $\min_{1 \leq \ell \leq L^\varepsilon} m^\ell N_\ell^\varepsilon = \frac{N^\varepsilon m^{-L^\varepsilon/2}}{C_{(2.5)} + \sum_{\ell=1}^{L^\varepsilon} m^{-3\ell/2}}$ together with (2.6) and (2.7), we get for all $0 \leq \varepsilon \leq c_1$

$$\mathcal{C} \min_{1 \leq \ell \leq L} N_\ell m^\ell \geq \mathcal{C} C_{(2.5)} \frac{K_{1,m} m^2 (m-1)[\dot{f}]_\infty^2 T^2 m^{-L^\varepsilon/2}}{(m+1)\varepsilon^2} \geq c_7 \varepsilon^{-4/3},$$

where $c_7 := \mathcal{C} C_{(2.5)} \frac{K_{1,m} m^{3/2}(m-1)([\dot{f}]_\infty T)^{4/3}}{2^{1/3}(m+1)(K_{1,\infty})^{1/3}(1+\frac{m(\sqrt{m}-1)}{m+1}C_{(2.5)})^{1/3}}$. Hence, for all $0 \leq \varepsilon \leq (c_6' c_7)^3 \wedge c_1$ condition (3.13) is satisfied and

$$\left\| \hat{Q}_\varepsilon - \mathbb{E}[f(X_T)] \right\|_{\Psi_e} \leq c_6 \varepsilon.$$

4 Error expansion and moment generating function of $\max_{1 \leq k \leq n} |X_{t_k}^{mn} - X_{t_k}^n|^2$

For $(n, m) \in \mathbb{N}^* \times \bar{\mathbb{N}}$, we consider the Euler scheme X^n on the grid $(t_k = \frac{kT}{n})_{0 \leq k \leq n}$ and the process X^{mn} which is the Euler scheme on the finer grid $(\frac{jT}{mn})_{0 \leq j \leq mn}$ when m is finite and the solution to (1.1) when $m = \infty$. We introduce the difference $U_{t_k} = X_{t_k}^{mn} - X_{t_k}^n$ between the two processes and define $U_T^\star = \max_{0 \leq k \leq n} |U_{t_k}|$. For $s \in [0, T]$, we set

$$\eta_l(s) = \left\lfloor \frac{ls}{T} \right\rfloor \frac{T}{l}, \; \hat{\eta}_l(s) = \left\lceil \frac{ls}{T} \right\rceil \frac{T}{l} \text{ when } l \in \mathbb{N}^* \text{ and } \eta_\infty(s) = \hat{\eta}_\infty(s) = s.$$

We have $U_0 = 0$ and for all $k \in \{0, \dots, n-1\}$,

$$U_{t_{k+1}} = U_{t_k} + \int_{t_k}^{t_{k+1}} b(X_{\eta_{mn}(s)}^{mn})ds - \frac{T}{n}b(X_{t_k}^n)$$

$$= U_{t_k} + \frac{T}{n}\left(b(X_{t_k}^{mn}) - b(X_{t_k}^n) \right) + \int_{t_k}^{t_{k+1}} \left(b(X_{\eta_{mn}(s)}^{mn}) - b(X_{t_k}^{mn}) \right) ds.$$

To deal with this induction equation, it is convenient to introduce the matrices

$$A_k = I_d + \mathbf{1}_{\{U_{t_k} \neq 0\}} \frac{T}{n} \times \frac{\left(b(X_{t_k}^{mn}) - b(X_{t_k}^n)\right) U_{t_k}^{\top}}{|U_{t_k}|^2},$$

$$\forall l \leq k, \ \mathcal{A}_k^l := A_k A_{k-1} \ldots A_l \ \text{and} \ \mathcal{A}_k^{k+1} = I_d.$$

This way, we have

$$U_{t_{k+1}} = A_k U_{t_k} + \int_{t_k}^{t_{k+1}} \left(b(X_{\eta_{mn}(s)}^{mn}) - b(X_{t_k}^{mn})\right) ds. \tag{4.1}$$

Let us introduce $V_{t_k} = \int_0^{t_k} \left(b(X_{\eta_{mn}(s)}^{mn}) - b(X_{\eta_n(s)}^{mn})\right) ds$. One can check by induction that

$$U_{t_k} = \sum_{l=1}^{k-1} \mathcal{A}_{k-1}^{l+1} (A_l - I_d) V_{t_l} + V_{t_k}. \tag{4.2}$$

Indeed, since

$$\sum_{l=1}^{k} \mathcal{A}_k^{l+1} (A_l - I_d) V_{t_l} + V_{t_{k+1}} = (A_k - I_d) V_{t_k} + A_k \left(\sum_{l=1}^{k-1} \mathcal{A}_{k-1}^{l+1} (A_l - I_d) V_{t_l} + V_{t_k} \right)$$

$$- A_k V_{t_k} + V_{t_{k+1}}$$

$$= A_k \left(\sum_{l=1}^{k-1} \mathcal{A}_{k-1}^{l+1} (A_l - I_d) V_{t_l} + V_{t_k} \right) + V_{t_{k+1}} - V_{t_k}$$

and $V_{t_{k+1}} - V_{t_k} = \int_{t_k}^{t_{k+1}} \left(b(X_{\eta_{mn}(s)}^{mn}) - b(X_{\eta_n(s)}^{mn})\right) ds$ both sides of (4.2) statisfy the induction equality (4.1). By Itô's formula and the integration by parts formula, for $k \in \{1, \ldots, n\}$,

$$V_{t_k} = \int_0^{t_k} \gamma(s) \left((\nabla b(X_s^{mn}) b(X_{\eta_{mn}(s)}^{mn}) + \frac{\Delta b}{2}(X_s^{mn})) ds + \nabla b(X_s^{mn}) dW_s \right),$$

where $\gamma(s) = (\hat{\eta}_n(s) - \hat{\eta}_{mn}(s))$. Therefore

$U_{t_k} = (U_{t_k}^{(1)} + U_{t_k}^{(2)})$ with

$$U_{t_k}^{(1)} = \sum_{l=1}^{k-1} \mathcal{A}_{k-1}^{l+1} (A_l - I_d) \int_0^{t_l} \gamma(t) \nabla b(X_t^{mn}) dW_t + \int_0^{t_k} \gamma(t) \nabla b(X_t^{mn}) dW_t \tag{4.3}$$

and $\displaystyle U_{t_k}^{(2)} = \sum_{l=1}^{k-1} \mathcal{A}_{k-1}^{l+1} (A_l - I_d) \int_0^{t_l} \gamma(t) (\nabla b(X_t^{mn}) b(X_{\eta_{mn}(t)}^{mn}) + \frac{1}{2} \Delta b(X_t^{mn})) dt$

$$+ \int_0^{t_k} \gamma(t) (\nabla b(X_t^{mn}) b(X_{\eta_{mn}(t)}^{mn}) + \frac{1}{2} \Delta b(X_t^{mn})) dt$$

respectively giving the contributions of the stochastic and the deterministic integrals. One can take advantage of the simpler expression

$$U_{t_k} = \sum_{l=1}^{k-1} \mathcal{A}_{k-1}^{l+1} \int_{t_l}^{t_{l+1}} \left(b(X_{\eta_{mn}(s)}^{mn}) - b(X_{\eta_n(s)}^{mn})\right) ds$$

(also proved by induction) and the linearity of the decomposition into stochastic (w.r.t. dW_t) and standard (w.r.t. dt) integrals to rewrite

$$U_{t_k}^{(2)} = \int_0^{t_k} \mathcal{A}_{k-1}^{\lceil \frac{nt}{T} \rceil} \gamma(t) (\nabla b(X_t^{mn}) b(X_{\eta_{mn}(t)}^{mn}) + \frac{1}{2} \Delta b(X_t^{mn})) dt$$

but the analogous expression of U_T^1 does not make sense because of the non-adapted factor $\mathcal{A}_{k-1}^{\lceil \frac{nt}{T} \rceil}$. This is the reason why we introduced the more complicated decomposition (4.2). Notice however than $n > T[\dot{b}]_\infty$, then the matrices A_k and therefore $\mathcal{A}_k^1$ are invertible for $k \in \{0, \ldots, n\}$ so that $U_{t_k}^{(1)} = \mathcal{A}_{k-1}^1 \int_0^{t_k} (\mathcal{A}_{\lfloor \frac{nt}{T} \rfloor}^1)^{-1} \gamma(t) \nabla b(X_t^{mn}) dW_t$. With Lemma 4.1 just below, we conclude that

$$U_T^\star \leq U_{(1)}^\star + U_{(2)}^\star \text{ where}$$

$$U_{(1)}^\star = \frac{[\dot{b}]_\infty T}{n} \sum_{k=1}^{n-1} e^{[\dot{b}]_\infty (T-t_{k+1})} \left| \int_0^{t_l} \gamma(t) \nabla b(X_t^{mn}) dW_t \right| + \max_{1 \leq k \leq n} \left| \int_0^{t_k} \gamma(t) \nabla b(X_t^{mn}) dW_t \right|$$

(4.4)

$$U_{(2)}^\star = \int_0^T e^{[\dot{b}]_\infty (T-t)} \gamma(t) \left| \nabla b(X_t^{mn}) b(X_{\eta_{mn}(t)}^{mn}) + \frac{1}{2} \Delta b(X_t^{mn}) \right| dt.$$

(4.5)

Let us state two lemmas that will be used to deal with $U_{(1)}^\star$ and $U_{(2)}^\star$ in the proof of Theorem 3.1. The first one follows from usual linear algebra arguments and the submultiplicative property of the matrix norm.

Lemma 4.1. *One has* $\forall k \in \{1, \ldots, n-1\}$,

$$\|A_k - I_d\| \leq \frac{T[\dot{b}]_\infty}{n}, \quad \|A_k\| \leq 1 + \frac{T[\dot{b}]_\infty}{n} \quad and \quad \forall l \in \{0, \ldots, k\}, \|\mathcal{A}_k^{l+1}\| \leq e^{[\dot{b}]_\infty (t_k - t_l)}. \quad (4.6)$$

Lemma 4.2. *When b is Lipschitz continuous with constant $[\dot{b}]_\infty \in (0, +\infty)$, one has*

1. $\forall t \in [0,T]$, $\sup_{n \in \bar{\mathbb{N}}} \sup_{s \leq t} |X_s^n - x_0| \leq e^{[\dot{b}]_\infty t} \sup_{s \leq t} |W_s| + \frac{e^{[\dot{b}]_\infty t} - 1}{[\dot{b}]_\infty} |b(x_0)|$,

2. $\sup_{n \in \bar{\mathbb{N}}} \sup_{s \leq t} |b(X_s^n)| \leq e^{[\dot{b}]_\infty t} \left([\dot{b}]_\infty \sup_{s \leq t} |W_s| + |b(x_0)| \right)$.

3. *Moreover, under* (**R1**), *we have*

$$\sup_{n \in \bar{\mathbb{N}}} \sup_{s \leq t} |\Delta b(X_s^n)| \leq 2a_{\Delta b} \left(e^{[\dot{b}]_\infty t} \sup_{s \leq t} |W_s| + \frac{e^{[\dot{b}]_\infty t} - 1}{[\dot{b}]_\infty} |b(x_0)| + 1 \right)$$

4. *and under* (**R2**), *we have*

$$\sup_{n \in \bar{\mathbb{N}}} \sup_{s \leq t} |\nabla \Delta b(X_s^n)| \leq 2a_{\nabla \Delta b} \left(e^{[\dot{b}]_\infty t} \sup_{s \leq t} |W_s| + \frac{e^{[\dot{b}]_\infty t} - 1}{[\dot{b}]_\infty} |b(x_0)| + 1 \right).$$

Notice that the function b is Lipschitz continuous with constant $[\dot{b}]_\infty$ under (**R1**).

Proof. For $n \in \bar{\mathbb{N}}$, one has

$$\forall t \in [0,T], \ X_t^n - x_0 = W_t + \int_0^t \left(b(X_{\eta_n(r)}^n) - b(x_0) \right) + b(x_0) dr.$$

Since b is Lipschitz continuous with constant $[\dot{b}]_\infty$,

$$\forall x \in \mathbb{R}^d, \ |b(x)| \leq |b(x) - b(x_0)| + |b(x_0)| \leq [\dot{b}]_\infty |x - x_0| + |b(x_0)|. \quad (4.7)$$

One deduces that

$$\sup_{s \leq t} |X_s^n - x_0| \leq \sup_{s \leq t} |W_s| + [\dot{b}]_\infty \int_0^t \sup_{s \leq r} |X_s^n - x_0| + \frac{|b(x_0)|}{[\dot{b}]_\infty} dr.$$

Applying Gronwall's lemma to the function $t \mapsto \sup_{s \le t} |X^n_s - x_0| + \frac{|b(x_0)|}{[\dot{b}]_\infty}$, one obtains that

$$\forall t \in [0, T], \; \sup_{s \le t} |X^n_s - x_0| \le e^{[\dot{b}]_\infty t} \sup_{s \le t} |W_s| + \frac{e^{[\dot{b}]_\infty t} - 1}{[\dot{b}]_\infty} |b(x_0)|.$$

The second (resp. third and fourth) inequality follows by using (4.7) (resp. (2.1) and (3.3)). $\qquad\square$

Proof of Theorem 3.1. Let $x \ge 0$. Since $\max_{1 \le k \le n} |U_{t_k}| = U^\star_T \le U^\star_{(1)} + U^\star_{(2)}$, Jensen's inequality implies that

$$\left(\frac{(m-1)T^2 x}{2mn} + \max_{1 \le k \le n} |U_{t_k}| \right)^2 \le \frac{(U^\star_{(1)})^2}{q} + \frac{(\frac{(m-1)T^2 x}{2mn} + U^\star_{(2)})^2}{1-q},$$

where $q \in (0, 1)$ is a parameter to be optimized later. With Hölder's inequality, we deduce that

$$\mathbb{E}\left[\exp\left\{ \rho \left(\frac{(m-1)T^2 x}{2mn} + U^\star_T \right)^2 \right\} \right] \le \mathbb{E}^q \left[\exp\left\{ \frac{\rho (U^\star_{(1)})^2}{q^2} \right\} \right]$$

$$\times \mathbb{E}^{1-q} \left[\exp\left\{ \frac{\rho(\frac{(m-1)T^2 x}{2mn} + U^\star_{(2)})^2}{(1-q)^2} \right\} \right]. \quad (4.8)$$

• **First term**. Let us first deal with the contribution of $U^\star_{(1)}$. Let us introduce the quantities

$$C = \sqrt{T} + \sum_{k=1}^{n-1} \frac{T[\dot{b}]_\infty e^{[\dot{b}]_\infty (T - t_{k+1})} \sqrt{t_k}}{n} \le \sqrt{T} + [\dot{b}]_\infty \int_0^T e^{[\dot{b}]_\infty (T-t)} \sqrt{t} dt := C_{(4.9)}, \quad (4.9)$$

$p_n = \frac{\sqrt{T}}{C}$ and $p_k = \frac{[\dot{b}]_\infty T e^{[\dot{b}]_\infty (T - t_{k+1})} \sqrt{t_k}}{Cn}$ for $1 \le k \le n-1$. Notice that $\sum_{k=1}^n p_k = 1$ so that we have defined a probability measure. By (4.4),

$$U^\star_{(1)} = C \sum_{k=1}^{n-1} \frac{p_k}{\sqrt{t_k}} \left| \int_0^{t_k} \gamma(t) \nabla b(X^{mn}_t) dW_t \right| + \frac{C p_n}{\sqrt{T}} \max_{1 \le k \le n} \left| \int_0^{t_k} \gamma(t) \nabla b(X^{mn}_t) dW_t \right|$$

$$\le C_{(4.9)} \sum_{k=1}^n \frac{p_k}{\sqrt{t_k}} \max_{1 \le l \le k} \left| \int_0^{t_l} \gamma(t) \nabla b(X^{mn}_t) dW_t \right|.$$

Applying Jensen's inequality to the convex function $\mathbb{R} \ni x \mapsto \exp\left\{ \frac{\rho x^2}{q^2} \right\}$, we deduce that

$$\mathbb{E}\left[\exp\left\{ \frac{\rho}{q^2} (U^\star_{(1)})^2 \right\} \right] \le \mathbb{E}\left[\sum_{k=1}^n p_k \exp\left\{ \frac{C^2_{(4.9)} \rho}{q^2 t_k} \max_{1 \le l \le k} \left| \int_0^{t_l} \gamma(t) \nabla b(X^{mn}_t) dW_t \right|^2 \right\} \right].$$

Now, using the periodicity of the function γ with period $t_1 = T/n$, we get

$$\int_0^{t_k} \gamma(t)^2 \, \mathrm{Tr} \left[\nabla b(X^{mn}_t) \nabla b(X^{mn}_t)^\top \right] dt \le d[\dot{b}]^2_\infty \int_0^{t_k} \gamma^2(t) dt = d[\dot{b}]^2_\infty t_k T^2 \frac{(2m-1)(m-1)}{6(mn)^2}.$$

Then, by the second assertion in Lemma A.1,

$$\forall \rho \in \left[0, \frac{3(qmn)^2}{4C^2_{(4.9)} T^2 d[\dot{b}]^2_\infty (2m-1)(m-1)} \right],$$

$$\mathbb{E}\left[\exp\left\{\frac{\rho(U_{(1)}^\star)^2}{q^2}\right\}\right] \leq \exp\left\{\frac{2\ln(2)C_{(4.9)}^2\rho T^2 d[\dot{b}]_\infty^2(2m-1)(m-1)}{3(qmn)^2}\right\}. \quad (4.10)$$

• **Second term**. On the other hand, by (4.5) and Lemma 4.2, one has

$$U_{(2)}^\star \leq \int_0^T \gamma(t)e^{[\dot{b}]_\infty(T-t)}\left(e^{[\dot{b}]_\infty t}([\dot{b}]_\infty^2 + a_{\Delta b})\left(\sup_{s\leq t}|W_s| + \frac{|b(x_0)|}{[\dot{b}]_\infty}\right) + a_{\Delta b}\right)dt.$$

Using $\int_0^T \gamma(t)dt = \frac{(m-1)T^2}{2mn}$ and $\tilde{x} = \frac{[\dot{b}]_\infty e^{-[\dot{b}]_\infty T}}{([\dot{b}]_\infty^2 + a_{\Delta b})}x$, one deduces that

$$\frac{(m-1)T^2 x}{2mn} + U_{(2)}^\star \leq e^{T[\dot{b}]_\infty}([\dot{b}]_\infty^2 + a_{\Delta b})\int_0^T \gamma(t)\left(\sup_{s\leq t}|W_s| + \frac{|b(x_0)| + [\dot{b}]_\infty + \tilde{x}}{[\dot{b}]_\infty}\right)dt$$

$$= e^{T[\dot{b}]_\infty}([\dot{b}]_\infty^2 + a_{\Delta b})\left(\int_0^T \gamma(t)\sup_{s\leq t}|W_s|dt + \frac{(m-1)T^2}{2mn}\times\frac{|b(x_0)| + [\dot{b}]_\infty + \tilde{x}}{[\dot{b}]_\infty}\right)$$

$$= e^{T[\dot{b}]_\infty}([\dot{b}]_\infty^2 + a_{\Delta b})\int_0^T \gamma(t)\left(\sup_{s\leq t}|W_s| + \frac{(m-1)T^2\sqrt{t}}{2mn\int_0^T \sqrt{r}\gamma(r)dr}\times\frac{|b(x_0)| + [\dot{b}]_\infty + \tilde{x}}{[\dot{b}]_\infty}\right)dt,$$

where we used the periodicity of $\gamma(t)$ for the first equality. Setting $C_{(4.11)} = e^{T[\dot{b}]_\infty}([\dot{b}]_\infty^2 + a_{\Delta b})$ and using Jensen's inequality for the probability density $p(t) = \frac{\sqrt{t}\gamma(t)}{\int_0^T \sqrt{s}\gamma(s)ds}$ on $[0, T]$, we obtain that

$$\mathbb{E}\left[\exp\left\{\frac{\rho(\frac{(m-1)T^2 x}{2mn} + U_{(2)}^\star)^2}{(1-q)^2}\right\}\right] \leq \int_0^T \mathbb{E}\left[\exp\left\{\frac{\rho C_{(4.11)}^2(\int_0^T \sqrt{r}\gamma(r)dr)^2}{(1-q)^2}\right.\right.$$

$$\left.\left.\times\left(\frac{1}{\sqrt{t}}\sup_{s\leq t}|W_s| + \frac{(m-1)T^2}{2mn\int_0^T \sqrt{r}\gamma(r)dr}\times\frac{|b(x_0)| + [\dot{b}]_\infty + \tilde{x}}{[\dot{b}]_\infty}\right)^2\right\}\right]p(t)dt. \quad (4.11)$$

Since $\sup_{s\leq 1}|W_s| \leq \sqrt{\sum_{i=1}^d \sup_{s\leq 1}|W_s^i|^2} \leq \sum_{i=1}^d \sup_{s\leq 1}|W_s^i|$, for $\delta \geq 0$,

$$\left(\sup_{s\leq 1}|W_s| + \delta\right)^2 \leq \sum_{i=1}^d \sup_{s\leq 1}|W_s^i|^2 + 2\delta\sum_{i=1}^d \sup_{s\leq 1}|W_s^i| + \delta^2 \leq \sum_{i=1}^d\left(\sup_{s\leq 1}|W_s^i| + \delta\right)^2 - (d-1)\delta^2,$$

where the random variables in the sum in the right-hand side are independent. Setting $\delta = \frac{(m-1)T^2}{2mn\int_0^T \sqrt{r}\gamma(r)dr}\times\frac{|b(x_0)| + [\dot{b}]_\infty + \tilde{x}}{[\dot{b}]_\infty}$, plugging this inequality in (4.11) after using the scaling property of the Brownian motion W, we deduce that

$$\mathbb{E}\left[\exp\left\{\frac{\rho(\frac{(m-1)T^2 x}{2mn} + U_{(2)}^\star)^2}{(1-q)^2}\right\}\right] \leq$$

$$\exp\left\{-\frac{(d-1)(m-1)^2\rho C_{(4.11)}^2 T^4(|b(x_0)| + [\dot{b}]_\infty + \tilde{x})^2}{4((1-q)mn)^2[\dot{b}]_\infty^2}\right\}$$

$$\times \mathbb{E}^d\left[\exp\left\{\frac{\rho C_{(4.11)}^2(\int_0^T \sqrt{r}\gamma(r)dr)^2}{(1-q)^2}\left(\sup_{s\leq 1}|W_s^1| + \delta\right)^2\right\}\right].$$

Using the fact that for each $k \in \{1, \ldots, n\}$, $r \mapsto \gamma(r)$ is non-increasing on $[t_{k-1}, t_k]$ while $r \mapsto \sqrt{r}$ is increasing, we obtain that

$$\int_0^T \sqrt{r}\gamma(r)dr = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \sqrt{r}\gamma(r)dr \leq \sum_{k=1}^n \frac{n}{T}\int_{t_{k-1}}^{t_k}\gamma(r)dr\int_{t_{k-1}}^{t_k}\sqrt{r}dr$$

$$= \frac{(m-1)T}{2mn}\int_0^T \sqrt{r}dr = \frac{(m-1)T^{5/2}}{3mn}.$$

Applying the first assertion in Lemma A.1 with $|H| = 1$ and $\mu = \frac{\rho C_{(4.11)}^2 (\int_0^T \sqrt{r}\gamma(r)dr)^2}{(1-q)^2}$, we deduce that if $0 \leq \rho \leq \frac{9((1-q)mn)^2}{8C_{(4.11)}^2 T^5 (m-1)^2}$, then

$$\mathbb{E}\left[\exp\left\{\frac{\rho(\frac{(m-1)T^2 x}{2mn} + U_{(2)}^\star)^2}{(1-q)^2}\right\}\right] \leq \exp\left\{\frac{\rho C_{(4.11)}^2 T^4 (m-1)^2}{((1-q)mn)^2}\right.$$

$$\times \left. \left(\frac{(3d+1)(|b(x_0)| + [\dot{b}]_\infty + \tilde{x})^2}{4[\dot{b}]_\infty^2} + \frac{4d\sqrt{T}(|b(x_0)| + [\dot{b}]_\infty + \tilde{x})}{3\sqrt{\pi}[\dot{b}]_\infty} + \frac{4dT\ln 2}{9}\right)\right\}. \quad (4.12)$$

• **Conclusion**. We now choose $q = \frac{C_{(4.9)}[\dot{b}]_\infty \sqrt{3d(2m-1)/m}}{C_{(4.9)}[\dot{b}]_\infty \sqrt{3d(2m-1)/m} + C_{(4.11)}T^{3/2}\sqrt{2(m-1)/m}}$ to obtain the same constraint on ρ for the two terms $U_T^{(1)}$ and $U_T^{(2)}$ and conclude by combining (4.8), (4.10) and (4.12) that if $0 \leq \rho \leq \frac{9mn^2}{4T^2(m-1)\left(C_{(4.9)}[\dot{b}]_\infty \sqrt{3d(2m-1)/m} + C_{(4.11)}T^{3/2}\sqrt{2(m-1)/m}\right)^2}$,

$$\mathbb{E}\left[\exp\left\{\rho(U_T^\star)^2\right\}\right] \leq \exp\left\{\rho C_{(3.2)}(x)\frac{(m-1)T^2}{mn^2}\right\}. \qquad \square$$

Remark 4.3. • When $n > [\dot{b}]_\infty T$, one could consider using the alternative expression $U_{t_k}^{(1)} = \mathcal{A}_{k-1}^1 \int_0^{t_k} (\mathcal{A}_{\lfloor \frac{nt}{T}\rfloor}^1)^{-1}\gamma(t)\nabla b(X_t^{mn})dW_t$ to replace in the first step $U_{(1)}^\star$ by

$$e^{[\dot{b}]_\infty T} \max_{1\leq k \leq n}\left|\int_0^{t_k} (\mathcal{A}_{\lfloor \frac{nt}{T}\rfloor}^1)^{-1}\gamma(t)\nabla b(X_t^{mn})dW_t\right|.$$

Lemma 4.1, implies that for $k \in \{0,\ldots,n\}$, $\|(\mathcal{A}_k^1)^{-1}\| \leq (1 - \frac{[\dot{b}]_\infty T}{n})^{-k}$. This leads to replace $C_{(4.9)}$ by some constant not smaller than

$$e^{[\dot{b}]_\infty T}\left(\int_0^T e^{2[\dot{b}]_\infty t}dt\right)^{1/2} \geq e^{[\dot{b}]_\infty T}\sqrt{T} = \sqrt{T}\left(1 + [\dot{b}]_\infty \int_0^T e^{[\dot{b}]_\infty(T-t)}dt\right) \geq C_{(4.9)}.$$

• In the last step of the derivation of (4.8), one could choose a constant $\tilde{q} \in (0,1)$ different from q to obtain

$$\mathbb{E}\left[\exp\left\{\rho(U_T^\star)^2\right\}\right] \leq \mathbb{E}^{\tilde{q}}\left[\exp\left\{\frac{\rho(U_{(1)}^\star)^2}{q\tilde{q}}\right\}\right]\mathbb{E}^{1-\tilde{q}}\left[\exp\left\{\frac{\rho(U_{(2)}^\star)^2}{(1-q)(1-\tilde{q})}\right\}\right],$$

but, following the reasoning in the above proof, this leads to the same upper-bound but under a stronger constraint on ρ. Indeed, for a fixed value of $q\tilde{q}$, the maximal value of $(1-q)(1-\tilde{q}) = 1 - (q+\tilde{q}) + q\tilde{q}$ is attained for $q = \tilde{q}$.

• On the other hand, if $q(t)$ is some probability density on $[0,T]$ and $(Y_t)_{t\in[0,T]}$ an $\mathbb{R}^d$-valued process, applying Jensen's inequality to the convex function $\mathbb{R}^d \ni x \mapsto \exp\left\{|x|^2\right\}$ leads to

$$\mathbb{E}\left[\exp\left\{\left|\int_0^T Y_t q(t)dt\right|^2\right\}\right] \leq \int_0^T \mathbb{E}\left[\exp\{|Y_t|^2\}\right]q(t)dt.$$

whereas applying Jensen's inequality to $\mathbb{R}^d \ni x \mapsto |x|^2$ then Hölder's inequality leads to the upper-bound $\exp\left\{\int_0^T \ln \mathbb{E}\left[\exp\{|Y_t|^2\}\right] q(t)dt\right\}$ which is smaller when $\mathbb{E}\left[\exp\{|Y_t|^2\}\right]$ is not constant dt a.e.. Nevertheless, in the above proof, the repeated use of Jensen's inequality for $x \mapsto \exp\left\{(x)^2\right\}$ did not worsen the final estimation because this estimation relies on uniform in $t \in [0, T]$ bounds for $\mathbb{E}\left[\exp\{|Y_t|^2\}\right]$.

Proof of Proposition 2.1. By (4.3) and Lemma 4.1,

$$|U_T^{(1)}| \leq \frac{[\dot{b}]_\infty T}{n} \sum_{k=1}^{n-1} e^{[\dot{b}]_\infty (T-t_{k+1})} \left| \int_0^{t_l} \gamma(t) \nabla b(X_t^{mn}) dW_t \right| + \left| \int_0^T \gamma(t) \nabla b(X_t^{mn}) dW_t \right|.$$

The difference between the right-hand side and the definition (4.4) of $U_{(1)}^\star$ is that in the latter $\left| \int_0^T \gamma(t) \nabla b(X_t^{mn}) dW_t \right|$ is replaced by $\max_{1 \leq k \leq n} \left| \int_0^{t_k} \gamma(t) \nabla b(X_t^{mn}) dW_t \right|$. Reasoning like in estimation of the first term in the proof of Theorem 3.1, we obtain that for the probability measure $(p_k)_{1 \leq k \leq n}$ introduced in this estimation,

$$
\begin{aligned}
\mathbb{E}\left[|U_T^1|^2\right] &\leq C_{(4.9)}^2 \sum_{k=1}^n \frac{p_k}{t_k} \mathbb{E}\left[\left| \int_0^{t_k} \gamma(t) \nabla b(X_t^{mn}) dW_t \right|^2\right] \\
&= C_{(4.9)}^2 \sum_{k=1}^n \frac{p_k}{t_k} \int_0^{t_k} \gamma(t)^2 \operatorname{Tr}\left[\nabla b(X_t^{mn}) \nabla b(X_t^{mn})^\top\right] dt \\
&\leq C_{(4.9)}^2 [\dot{b}]_\infty^2 dT^2 \frac{(2m-1)(m-1)}{6(mn)^2}.
\end{aligned}
\tag{4.13}
$$

When estimating $\mathbb{E}\left[(U_{(1)}^\star)^2\right]$ one needs to apply Doob's inequality to deal with the last and different term. Modifying the probability by giving weights proportional to $2\sqrt{T}$ for $k = n$ and to $\frac{[\dot{b}]_\infty T e^{[\dot{b}]_\infty (T-t_{k+1})} \sqrt{t_k}}{n}$ for $1 \leq k \leq n-1$ leads to

$$\mathbb{E}\left[(U_{(1)}^\star)^2\right] \leq (C_{(4.9)} + \sqrt{T})^2 [\dot{b}]_\infty^2 dT^2 \frac{(2m-1)(m-1)}{6(mn)^2}. \tag{4.14}$$

On the other hand, by the Minkowski inequality and (4.5),

$$\mathbb{E}^{1/2}\left[(U_{(2)}^\star)^2\right] \leq e^{[\dot{b}]_\infty T} \left(\int_0^T e^{-[\dot{b}]_\infty s} \gamma(s) \mathbb{E}^{1/2}\left[\left| \nabla b(X_s^{mn}) b(X_{\eta_{mn}(s)}^{mn}) + \frac{\Delta b}{2}(X_s^{mn}) \right|^2\right] ds \right). \tag{4.15}$$

Using Lemma 4.2 and that

$$\mathbb{E}^2\left[\sup_{r \leq s} |W_r|\right] \leq \mathbb{E}\left[\sup_{r \leq s} W_r^2\right] \leq \mathbb{E}\left[\sum_{i=1}^d \sup_{r \leq s}(W_r^i)^2\right] = 4d \times s,$$

we have

$$
\begin{aligned}
e^{-[\dot{b}]_\infty s} &\mathbb{E}^{1/2}\left[\left| \nabla b(X_s^{mn}) b(X_{\eta_{mn}(s)}^{mn}) + \frac{\Delta b}{2}(X_s^{mn})) \right|^2\right] \\
&\leq [\dot{b}]_\infty \left(2[\dot{b}]_\infty \sqrt{d \times s} + |b(x_0)| \right) + a_{\Delta b} \left(2\sqrt{d \times s} + \frac{1 - e^{-[\dot{b}]_\infty s}}{[\dot{b}]_\infty} |b(x_0)| + e^{-[\dot{b}]_\infty s} \right) \\
&\leq |b(x_0)| \left(a_{\Delta b} \frac{1 - e^{-[\dot{b}]_\infty s}}{[\dot{b}]_\infty} + [\dot{b}]_\infty \right) + 2([\dot{b}]_\infty^2 + a_{\Delta b})\sqrt{d \times s} + a_{\Delta b} e^{-[\dot{b}]_\infty s} \\
&:= g(s) + a_{\Delta b} e^{-[\dot{b}]_\infty s}.
\end{aligned}
$$

Using the fact that for each $k \in \{1, \ldots, n\}$, $s \mapsto g(s)$ is non-decreasing on $[t_{k-1}, t_k)$ while $s \mapsto \gamma(s)$ is non-increasing and bounded by $\frac{(m-1)T}{mn}$, we get that

$$
\begin{aligned}
e^{-[\dot{b}]_\infty T} \mathbb{E}^{1/2} \left[(U_{(2)}^\star)^2 \right] &\leq \sum_{k=1}^n \frac{n}{T} \int_{t_{k-1}}^{t_k} \gamma(s) ds \int_{t_{k-1}}^{t_k} g(s) ds + a_{\Delta b} \frac{(m-1)T}{mn} \int_0^T e^{-[\dot{b}]_\infty s} ds \\
&= \frac{(m-1)T}{2mn} \int_0^T g(s) ds + a_{\Delta b} \frac{(m-1)T}{mn} \times \frac{1 - e^{-[\dot{b}]_\infty T}}{[\dot{b}]_\infty} \\
&\leq \frac{(m-1)T}{mn} \Bigg(\frac{|b(x_0)|}{2} \bigg(a_{\Delta b} \frac{[\dot{b}]_\infty T - 1 + e^{-[\dot{b}]_\infty T}}{[\dot{b}]_\infty^2} + [\dot{b}]_\infty T \bigg) \\
&\qquad\qquad + a_{\Delta b} \frac{1 - e^{-[\dot{b}]_\infty T}}{[\dot{b}]_\infty} + \frac{2}{3}\sqrt{d}([\dot{b}]_\infty^2 + a_{\Delta b})T^{3/2} \Bigg).
\end{aligned}
$$

The estimation of $\mathbb{E}\left[|U_T|^2\right]$ (resp. $\mathbb{E}\left[|U_T^\star|^2\right]$) is obtained by plugging this inequality together with (4.13) (resp. (4.14)) into the inequality

$$
\mathbb{E}\left[|U_T|^2\right] \leq \frac{\mathbb{E}\left[|U_T^{(1)}|^2\right]}{q} + \frac{\mathbb{E}\left[(U_{(2)}^\star)^2\right]}{1-q} \text{ resp. } \mathbb{E}\left[|U_T^\star|^2\right] \leq \frac{\mathbb{E}\left[|U_{(1)}^\star|^2\right]}{q} + \frac{\mathbb{E}\left[(U_{(2)}^\star)^2\right]}{1-q}
$$

for $q \in (0,1)$ and optimizing over q : for $a, b > 0$, $\min_{q \in (0,1)} \frac{a}{q} + \frac{b}{1-q} = (\sqrt{a} + \sqrt{b})^2$ attained for $q = \frac{\sqrt{a}}{\sqrt{a} + \sqrt{b}}$. $\qquad \square$

5 Error expansion and moment generating function of its Malliavin derivative

5.1 Basic facts on Malliavin calculus

In this work, we follow the notations, definitions and results of [8]. Let $(W_t)_{0 \leq t \leq T}$ be a d-dimensional Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Let $\mathcal{D}$ denote the Malliavin derivative operator taking values in the real separable Hilbert space $H := L^2([0,T], \mathbb{R}^d)$ whose norm is denoted by $|\cdot|_H$. More precisely, for $h \in H$, we denote $W(h)$ the Wiener integral $W(h) = \int_0^T h(t) \cdot dW_t$. Let $\mathcal{S}$ denote the class of random variables of the form $F = f(W(h^1), \cdots, W(h^n))$, for $n \geq 1$, with $(h^1, \cdots, h^n) \in H^{\otimes n}$ and $f \in \mathscr{C}_p^\infty(\mathbb{R}^n, \mathbb{R})$. Then, for $F \in \mathcal{S}$, the Malliavin derivative of F denoted $\mathcal{D}F = (\mathcal{D}_t^i F, 0 \leq t \leq T, 1 \leq i \leq d)$ is defined by

$$
\mathcal{D}_t^i F = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(W(h^1), \cdots, W(h^n)) h_i^k(t),
$$

where h_i^k denotes the i-th coordinate of h^k. The operator $\mathcal{D}$ is closable as an operator from $L_p(\Omega)$ to $L_p(\Omega, H)$, for any $p \geq 1$. Its domain with respect to the norm $|F|_{1,p} := [\mathbb{E}(|F|^p) + \mathbb{E}(|\mathcal{D}F|_H^p)]^p$ is denoted by $\mathbb{D}^{1,p}$.

We now state some essential properties, which are going to be useful in the sequel.

Proposition 5.1 (Chain's rule)**.** *Let $\phi \in \mathscr{C}^1(\mathbb{R}^q, \mathbb{R})$ with bounded first order derivatives and $F = (F^1, \cdots, F^q)$ be an $\mathbb{R}^q$-valued random vector with $F^k \in \mathbb{D}^{1,p}$ for $k = 1, \ldots, q$. Then, $\phi(F) \in \mathbb{D}^{1,p}$ and for each $i = 1, \ldots, d$*

$$
\mathcal{D}_t^i \phi(F) = \sum_{j=1}^q \frac{\partial \phi}{\partial x^j}(F) \mathcal{D}_t^i F^j.
$$

Proposition 5.2 (Clark-Ocone formula)**.** *Let F be a $\mathcal{F}_T$-measurable random variable that belongs to $\mathbb{D}^{1,p}$ for some $p \geq 1$. Then,*

$$F = \mathbb{E}(F) + \sum_{i=1}^{d} \int_0^T \mathbb{E}(\mathcal{D}_s^i F | \mathcal{F}_s) dW_s^i, \quad \text{a.s.}$$

A preliminary essential result is on the boundedness of Malliavin's derivative of the diffusion and its Euler scheme given by (1.2)

Lemma 5.3. *Let $X = (X_t)_{t \in [0,T]}$ be the solution to Eq. (1.1), where the coefficient b satisfies our global Lipschitz condition $(\mathcal{H}_{GL})$. Then, $X_t, X_t^n \in \mathbb{D}^{1,\infty} := \cap_{p \geq 1} \mathbb{D}^{1,p}$ for any $t \in [0,T]$. Moreover, assume that $b \in \mathscr{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ with $\|\nabla b\| \leq [\dot{b}]_\infty$ for some finite constant $[\dot{b}]_\infty$. Then, for all $1 \leq j \leq d$ we have*

$$\forall r, t \in [0,T], \ |\mathcal{D}_r^j X_t| \vee \sup_{n \in \mathbb{N}^*} |\mathcal{D}_r^j X_t^n|) \leq \mathbf{1}_{\{r \leq t\}} e^{[\dot{b}]_\infty (t-r)}.$$

Proof. Under our assumptions, for any $t \in [0,T]$, the random variables X_t, X_t^n belong to $\mathbb{D}^{1,\infty}$ (see [8], Section 2.2). The estimation of the Malliavin derivative of X_t is straightforward. We only give a proof for the estimation of the Malliavin derivative of the Euler scheme. For $r, t \in [0,T]$,

$$\mathcal{D}_r^j X_t^n = \begin{cases} 0 \text{ if } t < r \\ \mathbf{e}_j \text{ if } r \leq t \leq \lceil \frac{rn}{T} \rceil \frac{T}{n} \\ \mathbf{e}_j + \int_{\lceil \frac{rn}{T} \rceil \frac{T}{n}}^{t} \nabla b(X_{\eta_n(u)}^n) \mathcal{D}_r^j X_{\eta_n(u)}^n du \text{ if } t \geq \lceil \frac{rn}{T} \rceil \frac{T}{n}, \end{cases} \quad (5.1)$$

where $(\mathbf{e}_j)_{j=1,\dots,d}$ denotes the canonical basis of $\mathbb{R}^d$. Hence $\mathcal{D}_r^j X_{\lceil \frac{rn}{T} \rceil \frac{T}{n}}^n = \mathbf{e}_j$ and for $t \geq \lceil \frac{rn}{T} \rceil \frac{T}{n}, \forall k \in \{\lceil \frac{rn}{T} \rceil, \dots, \lfloor \frac{tn}{T} \rfloor\}$,

$$\mathcal{D}_r^j X_{t \wedge \frac{(k+1)T}{n}}^n = \mathcal{D}_r^j X_{\frac{kT}{n}}^n + \left(t \wedge \frac{(k+1)T}{n} - \frac{kT}{n}\right) \nabla b(X_{\frac{kT}{n}}^n) \mathcal{D}_r^j X_{\frac{kT}{n}}^n,$$

so that

$$\mathcal{D}_r^j X_t^n = \left(I_d + (t - \eta_n(t)) \nabla b(X_{\eta_n(t)}^n)\right) \prod_{k=\lfloor \frac{tn}{T} \rfloor - 1}^{\lceil \frac{rn}{T} \rceil} \left(I_d + \frac{T}{n} \nabla b(X_{\frac{kT}{n}}^n)\right) \mathbf{e}_j.$$

Using the boundedness of the first order derivatives of b, we deduce that for $t \geq \lceil \frac{rn}{T} \rceil \frac{T}{n}$,

$$|\mathcal{D}_r^j X_t^n| \leq \left(1 + (t - \eta_n(t))[\dot{b}]_\infty\right) \left(1 + \frac{T}{n}[\dot{b}]_\infty\right)^{\lfloor \frac{tn}{T} \rfloor - \lceil \frac{rn}{T} \rceil} \leq e^{[\dot{b}]_\infty (t - \lceil \frac{rn}{T} \rceil \frac{T}{n})}. \qquad \square$$

5.2 Moment generating function of $\max_{1 \leq k \leq n} |\mathcal{D}_r^j X_{t_k}^{mn} - \mathcal{D}_r^j X_{t_k}^n|^2$

The next theorem states an exponential type upper bound for the moment generating function of $\max_{1 \leq k \leq n} |\mathcal{D}_r^j X_{t_k}^{mn} - \mathcal{D}_r^j X_{t_k}^n|^2$. In what follows, we refer to constants notation (**CN**) introduced in Section 3.1.

Theorem 5.4. *Let assumption (**R2**) hold, $(n, m) \in \mathbb{N}^* \times \bar{\mathbb{N}}$, $t_k = \frac{kT}{n}$ for $k \in \{0, \dots, n\}$ and ρ be a constant satisfying $0 \leq \rho \leq e^{-2[\dot{b}]_\infty (T-r)} \hat{\rho}(r) n^2$. Then,*

$$\mathbb{E}\left[\exp\left\{\rho \max_{0 \leq k \leq n} |\mathcal{D}_r^j X_{t_k}^{mn} - \mathcal{D}_r^j X_{t_k}^n|^2\right\}\right] \leq \exp\left\{\rho e^{2[\dot{b}]_\infty (T-r)} \Phi(r, [\dot{b}]_\infty) \frac{(m-1)T^2}{mn^2}\right\}. \quad (5.2)$$

Proof. Let $j \in \{1, \ldots, d\}$ and $r \in [0, T]$. By (5.1), for $k \in \{0, \ldots, \lfloor \frac{rn}{T} \rfloor\}$, $\mathcal{D}_r^j U_{t_k} = 0$,

$$\mathcal{D}_r^j U_{t_{\lceil \frac{rn}{T} \rceil}} = \int_{\hat{\eta}_{mn}(r)}^{\hat{\eta}_n(r)} \nabla b(X_{\eta_{mn}(s)}^{mn}) \mathcal{D}_r^j X_{\eta_{mn}(s)}^{mn} ds \tag{5.3}$$

and for $k \in \{\lceil \frac{rn}{T} \rceil, \ldots, n-1\}$,

$$\begin{aligned}
\mathcal{D}_r^j U_{t_{k+1}} =& \mathcal{D}_r^j U_{t_k} + \frac{T}{n} \nabla b(X_{t_k}^{mn}) \mathcal{D}_r^j U_{t_k} + \frac{T}{n}(\nabla b(X_{t_k}^{mn}) - \nabla b(X_{t_k}^n)) \mathcal{D}_r^j X_{t_k}^n \\
&+ \int_{t_k}^{t_{k+1}} \nabla b(X_{\eta_{mn}(s)}^{mn}) \mathcal{D}_r^j X_{\eta_{mn}(s)}^{mn} - \nabla b(X_{t_k}^{mn}) \mathcal{D}_r^j X_{t_k}^{mn} ds.
\end{aligned}$$

Setting

$$B_k := I_d + \frac{T}{n} \nabla b(X_{t_k}^{mn}),$$

and defining $(V_{t_k}^{(r,j)})_{\lceil \frac{rn}{T} \rceil \le k \le n}$ inductively by $V_{t_{\lceil \frac{rn}{T} \rceil}}^{(r,j)} = 0$ and

$$\begin{aligned}
V_{t_{k+1}}^{(r,j)} - V_{t_k}^{(r,j)} =& \frac{T}{n} \left(\nabla b(X_{t_k}^{mn}) - \nabla b(X_{t_k}^n) \right) \mathcal{D}_r^j X_{t_k}^n \\
&+ \int_{t_k}^{t_{k+1}} \nabla b(X_{\eta_{mn}(s)}^{mn}) \mathcal{D}_r^j X_{\eta_{mn}(s)}^{mn} - \nabla b(X_{t_k}^{mn}) \mathcal{D}_r^j X_{t_k}^{mn} ds, \tag{5.4}
\end{aligned}$$

we deduce that for $k \in \{\lceil \frac{rn}{T} \rceil, \ldots, n-1\}$,

$$\mathcal{D}_r^j U_{t_{k+1}} - V_{t_{k+1}}^{(r,j)} = B_k \mathcal{D}_r^j U_{t_k} - V_{t_k}^{(r,j)} = B_k(\mathcal{D}_r^j U_{t_k} - V_{t_k}^{(r,j)}) + (B_k - I_d)V_{t_k}^{(r,j)},$$

equality similar to (4.1). Let us introduce

$$\forall l \le k, \ \mathcal{B}_k^l := B_k B_{k-1} \ldots B_l \text{ and } \mathcal{B}_k^{k+1} = I_d.$$

One can check by induction on k that for all $k \in \{\lceil \frac{rn}{T} \rceil, \ldots, n\}$ and any sequence $(\tilde{V}_{t_k}^{(r,j)})_{\lceil \frac{rn}{T} \rceil \le k \le n}$ such that $\tilde{V}_{t_{\lceil \frac{rn}{T} \rceil}}^{(r,j)} = 0$, we have

$$\begin{aligned}
D_r^j U_{t_k} =& \mathcal{B}_{k-1}^{\lceil \frac{rn}{T} \rceil} D_r^j U_{t_{\lceil \frac{rn}{T} \rceil}} + \sum_{l=\lceil \frac{rn}{T} \rceil}^{k-1} \mathcal{B}_{k-1}^{l+1}(\tilde{V}_{t_{l+1}}^{(r,j)} - \tilde{V}_{t_l}^{(r,j)}) \\
&+ \sum_{l=\lceil \frac{rn}{T} \rceil+1}^{k-1} \mathcal{B}_{k-1}^{l+1}(B_l - I_d)(V_{t_l}^{(r,j)} - \tilde{V}_{t_l}^{(r,j)}) + V_{t_k}^{(r,j)} - \tilde{V}_{t_k}^{(r,j)}. \tag{5.5}
\end{aligned}$$

Let us explicit the right-hand side of (5.4). We have

$$\frac{T}{n} \left(\nabla b(X_{t_k}^{mn}) - \nabla b(X_{t_k}^n) \right) \mathcal{D}_r^j X_{t_k}^n = \frac{T}{n} H_{t_k}^b U_{t_k} \mathcal{D}_r^j X_{t_k}^n,$$

where for $u \in \mathbb{R}^d$, the lth column of the $d \times d$ matrix $(H_{t_{k_1}}^b u)$ is given by

$$\left(H_{t_k}^b u \right)_{\cdot l} := \left(\int_0^1 \frac{\partial \nabla b}{\partial x_l}(\theta X_{t_k}^n + (1-\theta)X_{t_k}^{mn}) d\theta \right) u.$$

Then, we write

$$\nabla b(X^{mn}_{\eta_{mn}(s)})\mathcal{D}^j_r X^{mn}_{\eta_{mn}(s)} - \nabla b(X^{mn}_{\eta_n(s)})\mathcal{D}^j_r X^{mn}_{\eta_n(s)} = (\nabla b(X^{mn}_{\eta_{mn}(s)}) - \nabla b(X^{mn}_{\eta_n(s)}))\mathcal{D}^j_r X^{mn}_{\eta_n(s)}$$
$$+ \nabla b(X^{mn}_{\eta_{mn}(s)})(\mathcal{D}^j_r X^{mn}_{\eta_{mn}(s)} - \mathcal{D}^j_r X^{mn}_{\eta_n(s)}).$$

For the first term, by Itô's formula and the integration by parts formula,

$$\int_{t_k}^{t_{k+1}} (\nabla b(X^{mn}_{\eta_{mn}(s)}) - \nabla b(X^{mn}_{\eta_n(s)}))\mathcal{D}^j_r X^{mn}_{\eta_n(s)} ds$$
$$= \int_{t_k}^{t_{k+1}} \gamma(s)\left(\left(\nabla^2 b(X^{mn}_s)b(X^{mn}_{\eta_{mn}(s)}) + \frac{1}{2}\nabla[\Delta b(X^{mn}_s)] \right)\mathcal{D}^j_r X^{mn}_{\eta_n(s)} ds \right.$$
$$\left. + \nabla^2 b(X^{mn}_s)\mathcal{D}^j_r X^{mn}_{\eta_n(s)} dW_s \right),$$

where $\gamma(s) = (\hat{\eta}_n(s) - \hat{\eta}_{mn}(s))$ and for $u \in \mathbb{R}^d$, the lth columns of the $d \times d$ matrices $(\nabla^2 b(X^n_t)u)$ and $\left(\nabla[\Delta b(X^n_t)] \right)$ are given by

$$\left(\nabla^2 b(X^n_t)u \right)_{\cdot l} := \frac{\partial \nabla b(X^n_t)}{\partial x_l}u \text{ and } \left(\nabla[\Delta b(X^n_t)] \right)_{\cdot l} := \frac{\partial \Delta b(X^n_t)}{\partial x_l}, \quad 1 \le l \le d.$$

Concerning the second term we use (5.1) to write

$$\mathcal{D}^j_r X^{mn}_{\eta_{mn}(s)} - \mathcal{D}^j_r X^{mn}_{\eta_n(s)} = \int_{\eta_n(s)}^{\eta_{mn}(s)} \nabla b(X^{mn}_{\eta_{mn}(u)})\mathcal{D}^j_r X^{mn}_{\eta_{mn}(u)} du + 1_{\{\eta_n(s)<r\le\eta_{mn}(s)\}}\mathbf{e}_j.$$

For $k \in \{\lceil \frac{rn}{T}\rceil, \ldots, n-1\}$, remarking that $\int_{t_k}^{t_{k+1}} \nabla b(X^{mn}_{\eta_{mn}(s)})1_{\{\eta_n(s)<r\le\eta_{mn}(s)\}}\mathbf{e}_j ds = 0$ and using Fubini's theorem, we get

$$\int_{t_k}^{t_{k+1}} \nabla b(X^{mn}_{\eta_{mn}(s)})(\mathcal{D}^j_r X^{mn}_{\eta_{mn}(s)} - \mathcal{D}^j_r X^{mn}_{\eta_n(s)})ds$$
$$= \int_{t_k}^{t_{k+1}} \left(\int_{\hat{\eta}_{mn}(s)}^{\hat{\eta}_n(s)} \nabla b(X^{mn}_{\eta_{mn}(u)})du \right)\nabla b(X^{mn}_{\eta_{mn}(s)})\mathcal{D}^j_r X^{mn}_{\eta_{mn}(s)} ds.$$

Hence,

$$V^{(r,j)}_{t_{k+1}} - V^{(r,j)}_{t_k} = \int_{t_k}^{t_{k+1}} \gamma(s)\nabla^2 b(X^{mn}_s)\mathcal{D}^j_r X^{mn}_{\eta_n(s)} dW_s + \frac{T}{n}H^b_{t_k}U_{t_k}\mathcal{D}^j_r X^n_{t_k}$$
$$+ \int_{t_k}^{t_{k+1}} \gamma(s)G^{b,mn}_s ds$$

$$\text{with } G^{b,mn}_s := \left(\nabla^2 b(X^{mn}_s)b(X^{mn}_{\eta_{mn}(s)}) + \frac{1}{2}\nabla[\Delta b(X^{mn}_s)] \right)\mathcal{D}^j_r X^{mn}_{\eta_n(s)}$$
$$+ \left(\frac{1_{\{\gamma(s)>0\}}}{\gamma(s)}\int_{\hat{\eta}_{mn}(s)}^{\hat{\eta}_n(s)} \nabla b(X^{mn}_{\eta_{mn}(u)})du \right)\nabla b(X^{mn}_{\eta_{mn}(s)})\mathcal{D}^j_r X^{mn}_{\eta_{mn}(s)}.$$

Choosing $\tilde{V}^{(r,j)}_{t_{k+1}} - \tilde{V}^{(r,j)}_{t_k}$ equal to the sum of the second and third terms in the above expression of $V^{(r,j)}_{t_{k+1}} - V^{(r,j)}_{t_k}$ and applying (5.5), we conclude that for $k \in \{\lceil \frac{rn}{T}\rceil, \ldots, n\}$,

$$\mathcal{D}^j_r U_{t_k} = \mathcal{D}^j_r U^{(0)}_{t_k} + \mathcal{D}^j_r U^{(1)}_{t_k} + \mathcal{D}^j_r U^{(2)}_{t_k} + \mathcal{D}^j_r U^{(3)}_{t_k} \text{ with} \tag{5.6}$$

$$\mathcal{D}^j_r U^{(0)}_{t_k} := \mathcal{B}^{\lceil \frac{rn}{T}\rceil}_{k-1}\mathcal{D}^j_r U_{t_{\lceil \frac{rn}{T}\rceil}}$$

$$\mathcal{D}^j_r U^{(1)}_{t_k} := \frac{T}{n}\sum_{l=\lceil \frac{rn}{T}\rceil}^{k-1} \mathcal{B}^{l+1}_{k-1}H^b_{t_l}U_{t_l}\mathcal{D}^j_r X^n_{t_l}$$

$$\mathcal{D}_r^j U_{t_k}^{(2)} := \sum_{l=\lceil \frac{rn}{T} \rceil}^{k-1} \mathcal{B}_{k-1}^{l+1} \int_{t_l}^{t_{l+1}} \gamma(s) G_s^{b,mn} ds$$

$$\mathcal{D}_r^j U_{t_k}^{(3)} := \sum_{l=\lceil \frac{rn}{T} \rceil+1}^{k-1} \mathcal{B}_{k-1}^{l+1}(B_l - I_d) \int_{\hat{\eta}_n(r)}^{t_l} \gamma(s) \nabla^2 b(X_s^{mn}) \mathcal{D}_r^j X_{\eta_n(s)}^{mn} dW_s$$

$$+ \int_{\hat{\eta}_n(r)}^{t_k} \gamma(s) \nabla^2 b(X_s^{mn}) \mathcal{D}_r^j X_{\eta_n(s)}^{mn} dW_s.$$

Combining assumption (**R2**), standard linear algebra arguments and Lemma 5.3, we easily prove the following result.

Lemma 5.5. *One has* $\forall k \in \{1, \dots, n-1\}$,

$$\|B_k - I_d\| \le \frac{T[\dot{b}]_\infty}{n}, \quad \|B_k\| \le 1 + \frac{T[\dot{b}]_\infty}{n}, \quad \forall l \in \{0, \dots, k\}, \|\mathcal{B}_k^{l+1}\| \le e^{[\dot{b}]_\infty(t_k-t_l)}$$

and $\forall s \in [\hat{\eta}_n(r), T]$

$$|G_s^{b,mn}| \le e^{[\dot{b}]_\infty(\eta_{mn}(s)-r)} \left(\sqrt{d}[\ddot{b}]_\infty |b(X_{\eta_{mn}(s)}^{mn})| + \frac{1}{2}\|\nabla[\Delta b(X_s^{mn})]\| + [\dot{b}]_\infty^2 \right).$$

Combining this result, (5.3) and Lemma 5.3, we easily get an upper bound for $\mathcal{D}_r^j U_{t_k}^{(0)}$

$$|\mathcal{D}_r^j U_{t_k}^{(0)}| \le \|\mathcal{B}_{k-1}^{\lceil \frac{rn}{T} \rceil}\| |\mathcal{D}_r^j U_{t_{\lceil \frac{rn}{T} \rceil}}| \le e^{[\dot{b}]_\infty(t_k-\hat{\eta}_n(r))}[\dot{b}]_\infty e^{[\dot{b}]_\infty(\hat{\eta}_n(r)-r)}\gamma(r)$$

$$= [\dot{b}]_\infty e^{[\dot{b}]_\infty(t_k-r)}\gamma(r). \tag{5.7}$$

In the same way as in the previous section where we introduced $U_T^\star = \max_{0 \le k \le n} |U_{t_k}|$, we also define for all $r \in [0, T]$, the process $\mathcal{D}_r^j U_T^\star = \max_{0 \le k \le n} |\mathcal{D}_r^j U_{t_k}^{(0)} + \mathcal{D}_r^j U_{t_k}^{(1)} + \mathcal{D}_r^j U_{t_k}^{(2)} + \mathcal{D}_r^j U_{t_k}^{(3)}|$. According to our assumption (**R2**) and by Lemmas 5.3 and 5.5, we have

$$\mathcal{D}_r^j U_T^\star \le \mathcal{D}_r^j U_{(0)}^\star + \mathcal{D}_r^j U_{(1)}^\star + \mathcal{D}_r^j U_{(2)}^\star + \mathcal{D}_r^j U_{(3)}^\star, \text{ with}$$

$$\mathcal{D}_r^j U_{(0)}^\star \le [\dot{b}]_\infty e^{[\dot{b}]_\infty(T-r)}\frac{(m-1)T}{mn} \tag{5.8}$$

$$\mathcal{D}_r^j U_{(1)}^\star := \frac{T}{n}e^{[\dot{b}]_\infty(T-r)} \sum_{k=\lceil \frac{rn}{T} \rceil}^{n-1} e^{[\dot{b}]_\infty(T-t_{k+1})}\|H_{t_k}^b U_{t_k}\|$$

$$\mathcal{D}_r^j U_{(2)}^\star = \int_{\hat{\eta}_n(r)}^T e^{[\dot{b}]_\infty(T-\hat{\eta}_n(s))}|\gamma(s)G_s^{b,mn}|ds$$

$$\mathcal{D}_r^j U_{(3)}^\star := \frac{T}{n} \sum_{k=\lceil \frac{rn}{T} \rceil+1}^{n-1} [\dot{b}]_\infty e^{[\dot{b}]_\infty(T-t_{k+1})}| \int_{\hat{\eta}_n(r)}^{t_k} \gamma(s)\nabla^2 b(X_s^{mn})\mathcal{D}_r^j X_{\eta_n(s)}^{mn} dW_s|$$

$$+ \max_{\lceil \frac{rn}{T} \rceil+1 \le k \le n} | \int_{\hat{\eta}_n(r)}^{t_k} \gamma(s)\nabla^2 b(X_s^{mn})\mathcal{D}_r^j X_{\eta_n(s)}^{mn} dW_s|. \tag{5.9}$$

By Jensen's inequality we get

$$|\mathcal{D}_r^j U_T^\star|^2 \le \frac{|\mathcal{D}_r^j U_{(1)}^\star|^2}{(1-q)} + \frac{|\mathcal{D}_r^j U_{(0)}^\star + \mathcal{D}_r^j U_{(2)}^\star|^2}{q(1-\bar{q})} + \frac{|\mathcal{D}_r^j U_{(3)}^\star|^2}{q\bar{q}},$$

where $q, \bar{q} \in (0, 1)$ are two parameters to be optimized later. Then, by Hölder and Jensen inequalities, we deduce that for all $r \in [0, T]$, we have

$$\mathbb{E}\left[\exp\left\{\rho|\mathcal{D}_r^j U_T^\star|^2\right\}\right] \le \mathbb{E}^{1-q}\left[\exp\left\{\frac{\rho}{(1-q)^2}|\mathcal{D}_r^j U_{(1)}^\star|^2\right\}\right]$$

$$\times \, \mathbb{E}^{q(1-\bar{q})} \left[\exp \left\{ \frac{\rho}{q^2(1-\bar{q})^2} |\mathcal{D}_r^j U_{(0)}^\star + \mathcal{D}_r^j U_{(2)}^\star|^2 \right\} \right]$$

$$\times \, \mathbb{E}^{q\bar{q}} \left[\exp \left\{ \frac{\rho}{q^2\bar{q}^2} |\mathcal{D}_r^j U_{(3)}^\star|^2 \right\} \right]. \tag{5.10}$$

• **First term**. In this part, we focus on the contribution of the term $\mathcal{D}_r^j U_{(1)}^\star$. Note that under assumption (**R2**) we have, for all $k \in \{\lceil \frac{rn}{T} \rceil, \cdots, n-1\}$

$$\|H_{t_k}^b U_{t_k}\| \leq \left(\sum_{j=1}^d \int_0^1 \left\| \frac{\partial \nabla b}{\partial x_j} (\theta X_{t_k}^n + (1-\theta) X_{t_k}^{mn}) \right\|^2 d\theta \, |U_{t_k}|^2 \right)^{\frac{1}{2}} \leq \sqrt{d}[\ddot{b}]_\infty |U_{t_k}|.$$

Therefore,

$$\mathcal{D}_r^j U_{(1)}^\star \leq \frac{T\sqrt{d}[\ddot{b}]_\infty}{n} U_T^\star e^{[\dot{b}]_\infty (T-r)} \sum_{k=\lceil \frac{rn}{T} \rceil}^{n-1} e^{[\dot{b}]_\infty (T-t_{k+1})}$$

$$\leq \sqrt{d}[\ddot{b}]_\infty U_T^\star e^{[\dot{b}]_\infty (T-r)} \int_{\hat{\eta}_{n(r)}}^T e^{[\dot{b}]_\infty (T-t)} dt$$

$$\leq e^{[\dot{b}]_\infty (T-r)} \frac{\sqrt{d}[\ddot{b}]_\infty}{[\dot{b}]_\infty} (e^{[\dot{b}]_\infty (T-r)} - 1) U_T^\star := e^{[\dot{b}]_\infty (T-r)} \Phi_1(r) U_T^\star. \tag{5.11}$$

Hence, it follows that for all $r \in [0, T]$, we have

$$\mathbb{E} \left[\exp \left\{ \frac{\rho}{(1-q)^2} |\mathcal{D}_r^j U_{(1)}^\star|^2 \right\} \right] \leq \mathbb{E} \left[\exp \left\{ \frac{\rho e^{2[\dot{b}]_\infty (T-r)} \Phi_1^2(r)}{(1-q)^2} |U_T^\star|^2 \right\} \right].$$

As assumption (**R1**) is satisfied under (**R2**), then Theorem 3.1 applies and for all $\rho \in [0, \frac{\rho_{(3.1)}(1-q)^2 n^2}{e^{2[\dot{b}]_\infty (T-r)} \Phi_1^2(r)}]$,

$$\mathbb{E}^{(1-q)} \left[\exp \left\{ \frac{\rho}{(1-q)^2} |\mathcal{D}_r^j U_{(1)}^\star|^2 \right\} \right] \leq \exp \left\{ \rho \frac{e^{2[\dot{b}]_\infty (T-r)} \Phi_1^2(r) C_{(3.2)}(0)}{(1-q)} \times \frac{(m-1)T^2}{mn^2} \right\}. \tag{5.12}$$

• **Second term**. By the second assertion of Lemma 5.5, we have

$$D_r^j U_{(2)}^\star \leq e^{[\dot{b}]_\infty (T-r)} \int_{\hat{\eta}_n(r)}^T \gamma(t) \left(\sqrt{d}[\ddot{b}]_\infty |b(X_{\eta_{mn}(t)}^{mn})| + \frac{1}{2} \|\nabla[\Delta b(X_t^{mn})]\| + [\dot{b}]_\infty^2 \right) dt.$$

Moreover, thanks to Lemma 4.2, we get

$$D_r^j U_{(2)}^\star \leq e^{[\dot{b}]_\infty (T-r)}$$

$$\times \int_{\hat{\eta}_n(r)}^T \gamma(t) \left(e^{[\dot{b}]_\infty t} (\sqrt{d}[\dot{b}]_\infty [\ddot{b}]_\infty + a_{\nabla \Delta b}) \left(\sup_{s \leq t} |W_s| + \frac{|b(x_0)|}{[\dot{b}]_\infty} \right) + a_{\nabla \Delta b} + [\dot{b}]_\infty^2 \right) dt$$

$$\leq C_{(5.13)} e^{[\dot{b}]_\infty (T-r)} \int_{\hat{\eta}_n(r)}^T \gamma(t) e^{[\dot{b}]_\infty t} \left(\sup_{s \leq t} |W_s| + \frac{|b(x_0)|}{[\dot{b}]_\infty} + 1 \right) dt, \tag{5.13}$$

where $C_{(5.13)} = (\sqrt{d}[\dot{b}]_\infty [\ddot{b}]_\infty \vee [\dot{b}]_\infty^2 + a_{\nabla \Delta b})$. With (5.8), we deduce that

$$\mathcal{D}_r^j U_{(0)}^\star + D_r^j U_{(2)}^\star \leq C_{(5.13)} e^{[\dot{b}]_\infty (T-r)} \times \int_{\hat{\eta}_n(r)}^T \gamma(t) e^{[\dot{b}]_\infty t} \left(\sup_{s \leq t} |W_s| \right.$$

$$+ \frac{\sqrt{t}}{\int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}\sqrt{s}ds} \left(\frac{|b(x_0)| + [\dot{b}]_\infty}{[\dot{b}]_\infty} \int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}ds + \frac{[\dot{b}]_\infty (m-1)T}{C_{(5.13)}mn} \right) \Bigg) dt.$$

Therefore, using Jensen's inequality for the probability density

$$p(t) = \frac{\sqrt{t}\gamma(t)e^{[\dot{b}]_\infty t}}{\int_{\hat{\eta}_n(r)}^T \sqrt{s}\gamma(s)e^{[\dot{b}]_\infty s}ds}, \quad t \in [\hat{\eta}_n(r), T],$$

we obtain that

$$\mathbb{E}\left[\exp\left\{\frac{\rho}{q^2(1-\bar{q})^2}|\mathcal{D}_r^j U_{(0)}^\star + \mathcal{D}_r^j U_{(2)}^\star|^2\right\}\right] \le$$

$$\int_{\hat{\eta}_n(r)}^T \mathbb{E}\left[\exp\left\{\frac{\rho C_{(5.13)}^2 e^{2[\dot{b}]_\infty(T-r)}(\int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}\sqrt{s}ds)^2}{q^2(1-\bar{q})^2}\left(\frac{1}{\sqrt{t}}\sup_{s\le t}|W_s| + \delta\right)^2\right\}\right]p(t)dt,$$

where $\delta = \frac{1}{\int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}\sqrt{s}ds}\left(\frac{|b(x_0)|+[\dot{b}]_\infty}{[\dot{b}]_\infty}\int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}ds + \frac{[\dot{b}]_\infty(m-1)T}{C_{(5.13)}mn}\right)$ and by the scaling property for the Brownian motion W, we may replace $\frac{1}{\sqrt{t}}\sup_{s\le t}|W_s|$ by $\sup_{s\le 1}|W_s|$. In the same way as we did for the second term of Section 4, we use that

$$\left(\sup_{s\le 1}|W_s| + \delta\right)^2 \le \sum_{i=1}^d \left(\sup_{s\le 1}|W_s^i| + \delta\right)^2 - (d-1)\delta^2,$$

to get

$$\mathbb{E}\left[\exp\left\{\frac{\rho}{q^2(1-\bar{q})^2}|\mathcal{D}_r^j U_{(0)}^\star + \mathcal{D}_r^j U_{(2)}^\star|^2\right\}\right] \le$$

$$\exp\left\{-\frac{\rho(d-1)C_{(5.13)}^2 e^{2[\dot{b}]_\infty(T-r)}}{q^2(1-\bar{q})^2}\left(\frac{|b(x_0)|+[\dot{b}]_\infty}{[\dot{b}]_\infty}\int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}ds + \frac{[\dot{b}]_\infty(m-1)T}{C_{(5.13)}mn}\right)^2\right\}$$

$$\times \mathbb{E}^d\left[\exp\left\{\frac{\rho C_{(5.13)}^2 e^{2[\dot{b}]_\infty(T-r)}(\int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}\sqrt{s}ds)^2}{q^2(1-\bar{q})^2}\left(\sup_{s\le 1}|W_s^1| + \delta\right)^2\right\}\right].$$

Applying the first assertion in Lemma A.1 with $|H| = 1$ and using that since $s \mapsto \sqrt{s}e^{[\dot{b}]_\infty s}$ and $s \mapsto e^{[\dot{b}]_\infty s}$ are non-decreasing, $\int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}\sqrt{s}ds \le \frac{T(m-1)}{2mn}\int_r^T e^{[\dot{b}]_\infty s}\sqrt{s}ds$ and $\int_{\hat{\eta}_n(r)}^T \gamma(s)e^{[\dot{b}]_\infty s}ds \le \frac{T(m-1)}{2mn} \times \frac{e^{[\dot{b}]_\infty T} - e^{[\dot{b}]_\infty r}}{[\dot{b}]_\infty}$, we deduce that if

$$\rho \in [0, \frac{\rho_{(5.14)}q^2(1-\bar{q})^2 n^2}{e^{2[\dot{b}]_\infty(T-r)}\Phi_2^2(r)}] \text{ with } \rho_{(5.14)} := \frac{m^2}{2C_{(5.13)}^2 T^2(m-1)^2} \text{ and } \Phi_2(r) := \int_r^T e^{[\dot{b}]_\infty s}\sqrt{s}ds,$$

then

$$\mathbb{E}^{q(1-\bar{q})}\left[\exp\left\{\frac{\rho}{q^2(1-\bar{q})^2}|\mathcal{D}_r^j U_{(0)}^\star + D_r^j U_{(2)}^\star|^2\right\}\right] \le$$

$$\exp\left\{\rho\frac{e^{2[\dot{b}]_\infty(T-r)}\phi_2(r, [\dot{b}]_\infty)}{q(1-\bar{q})} \times \frac{T^2(m-1)}{mn^2}\right\}, \quad (5.14)$$

where

$$\phi_2(r, x) = \frac{(m-1)}{m}\left((3d+1)\left(C_{(5.13)}\frac{|b(x_0)|+[\dot{b}]_\infty}{2[\dot{b}]_\infty^2}(e^{[\dot{b}]_\infty T} - e^{[\dot{b}]_\infty r}) + x\right)^2\right.$$

$$+ \frac{4dC_{(5.13)}\Phi_2(r)}{\sqrt{\pi}} \left(C_{(5.13)} \frac{|b(x_0)| + [\dot{b}]_\infty}{2[\dot{b}]_\infty^2} (e^{[\dot{b}]_\infty T} - e^{[\dot{b}]_\infty r}) + x \right) + d\ln 2C_{(5.13)}^2 \Phi_2^2(r) \right).$$

• **Third term**. Let us introduce the quantities

$$C = \sqrt{\frac{e^{2[\dot{b}]_\infty(T-\hat{\eta}_n(r))} - 1}{2[\dot{b}]_\infty}} + \sum_{k=\lceil \frac{rn}{T} \rceil + 1}^{n-1} \frac{T[\dot{b}]_\infty e^{[\dot{b}]_\infty(T-t_{k+1})} \sqrt{(e^{2[\dot{b}]_\infty(t_k-\hat{\eta}_n(r))} - 1)/2[\dot{b}]_\infty}}{n}$$

$$\leq e^{[\dot{b}]_\infty(T-r)} \left(\sqrt{\frac{1 - e^{-2[\dot{b}]_\infty(T-r)}}{2[\dot{b}]_\infty}} + \sqrt{\frac{[\dot{b}]_\infty}{2}} \int_r^T \sqrt{1 - e^{-2[\dot{b}]_\infty(t-r)}} dt \right) := e^{[\dot{b}]_\infty(T-r)} \Phi_3(r),$$

$$(5.15)$$

$p_n = \frac{1}{C} \sqrt{\frac{e^{2[\dot{b}]_\infty(T-\hat{\eta}_n(r))} - 1}{2[\dot{b}]_\infty}}$ and $p_k = \frac{T[\dot{b}]_\infty e^{[\dot{b}]_\infty(T-t_{k+1})} \sqrt{(e^{2[\dot{b}]_\infty(t_k-\hat{\eta}_n(r))} - 1)/2[\dot{b}]_\infty}}{Cn}$ for $\lceil \frac{rn}{T} \rceil + 1 \leq k \leq n - 1$. Notice that $\sum_{k=\lceil \frac{rn}{T} \rceil + 1}^{n} p_k = 1$ so that we have defined a probability measure. Therefore, by (5.9) we have

$$\mathcal{D}_r^j U_{(3)}^\star = \sum_{k=\lceil \frac{rn}{T} \rceil + 1}^{n-1} p_k \frac{C\sqrt{2[\dot{b}]_\infty}}{\sqrt{e^{2[\dot{b}]_\infty(t_k-\hat{\eta}_n(r))} - 1}} \left| \int_{\hat{\eta}_n(r)}^{t_k} \gamma(s) \nabla^2 b(X_s^{mn}) \mathcal{D}_r^j X_{\eta_n(s)}^{mn} dW_s \right|$$

$$+ p_n \frac{C\sqrt{2[\dot{b}]_\infty}}{\sqrt{e^{2[\dot{b}]_\infty(T-\hat{\eta}_n(r))} - 1}} \max_{\lceil \frac{rn}{T} \rceil + 1 \leq k \leq n} \left| \int_{\hat{\eta}_n(r)}^{t_k} \gamma(s) \nabla^2 b(X_s^{mn}) \mathcal{D}_r^j X_{\eta_n(s)}^{mn} dW_s \right|$$

and so

$$\mathcal{D}_r^j U_{(3)}^\star \leq \sum_{k=\lceil \frac{rn}{T} \rceil + 1}^{n} p_k \frac{C\sqrt{2[\dot{b}]_\infty}}{\sqrt{e^{2[\dot{b}]_\infty(t_k-\hat{\eta}_n(r))} - 1}} \max_{\lceil \frac{rn}{T} \rceil + 1 \leq l \leq k} \left| \int_{\hat{\eta}_n(r)}^{t_l} \gamma(t) \nabla^2 b(X_s^{mn}) \mathcal{D}_r^j X_{\eta_n(s)}^{mn} dW_s \right|.$$

Now, applying Jensen's inequality to the convex function $\mathbb{R} \ni x \mapsto \exp\left\{ \frac{\rho x^2}{q^2 \bar{q}^2} \right\}$, we deduce that for all $r \in [0, T]$, we have

$$\mathbb{E}\left[\exp\left\{ \frac{\rho}{q^2 \bar{q}^2} |\mathcal{D}_r^j U_{(3)}^\star|^2 \right\} \right] \leq \sum_{k=\lceil \frac{rn}{T} \rceil + 1}^{n} p_k$$

$$\times \mathbb{E}\left[\exp\left\{ \frac{\rho e^{2[\dot{b}]_\infty(T-r)} \Phi_3^2(r) 2[\dot{b}]_\infty}{q^2 \bar{q}^2 (e^{2[\dot{b}]_\infty(t_k-\hat{\eta}_n(r))} - 1)} \max_{\lceil \frac{rn}{T} \rceil + 1 \leq l \leq k} \left| \int_{\hat{\eta}_n(r)}^{t_l} \gamma(t) \nabla^2 b(X_s^{mn}) \mathcal{D}_r^j X_{\eta_n(s)}^{mn} dW_s \right|^2 \right\} \right].$$

Now, using assumption (**R2**), Lemma 5.3, than the periodicity of the function γ with period $t_1 = T/n$, we get

$$\int_{\hat{\eta}_n(r)}^{t_k} \gamma(t)^2 \operatorname{Tr}\left[\nabla^2 b(X_t^{mn}) \mathcal{D}_r^j X_{\eta_n(t)}^{mn} (\nabla^2 b(X_t^{mn}) \mathcal{D}_r^j X_{\eta_n(t)}^{mn})^\top \right] dt$$

$$\leq \int_{\hat{\eta}_n(r)}^{t_k} \gamma(t)^2 \sum_{j=1}^d \left\| \frac{\partial \nabla b(X_t^{mn})}{\partial x_j} \right\|^2 |\mathcal{D}_r^j X_{\eta_n(t)}^{mn}|^2 dt \leq d[\ddot{b}]_\infty^2 \int_{\hat{\eta}_n(r)}^{t_k} \gamma(t)^2 e^{2[\dot{b}]_\infty(\eta_n(t)-\hat{\eta}_n(r))} dt$$

$$= d[\ddot{b}]_\infty^2 T^2 \frac{(2m-1)(m-1)}{6(mn)^2} \frac{T}{n} \sum_{l=\lceil \frac{nr}{T} \rceil}^{k-1} e^{2[\dot{b}]_\infty(t_l-\hat{\eta}_n(r))}$$

$$\leq d[\ddot{b}]_\infty^2 T^2 \frac{(2m-1)(m-1)}{6(mn)^2} \times \frac{e^{2[\dot{b}]_\infty(t_k-\hat{\eta}_n(r))} - 1}{2[\dot{b}]_\infty}.$$

Then, by the second assertion in Lemma A.1,

$$\forall \rho \in [0, \frac{\rho_{(5.16)} q^2 \bar{q}^2 n^2}{e^{2[\dot{b}]_\infty (T-r)} \Phi_3^2(r)}], \text{ with } \rho_{(5.16)} := \frac{3m^2}{4T^2 d[\ddot{b}]_\infty^2 (2m-1)(m-1)}$$

$$\mathbb{E}^{q\bar{q}}\left[\exp\left\{\frac{\rho}{q^2 \bar{q}^2}|D_r^j U_{(3)}^\star|^2\right\}\right] \le \exp\left\{\rho \frac{e^{2[\dot{b}]_\infty (T-r)} \Phi_3^2(r) C_{(5.16)}}{q\bar{q}} \times \frac{(m-1)T^2}{mn^2}\right\}, \quad (5.16)$$

where $C_{(5.16)} := \frac{2}{3}\ln(2)d[\ddot{b}]_\infty^2 \frac{2m-1}{m}$.

• **Conclusion**. In order to have the same constraint on ρ for the three Laplace transforms, we choose

$$q = \frac{\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\left(\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} + \frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}\right)}{\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} + \frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)} + \frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)}\frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}} \text{ and } \bar{q} = \frac{\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)}}{\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} + \frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}}.$$

Then, by combining (5.10), (5.12), (5.14) and (5.16), we deduce that if

$$0 \le \rho \le \frac{\frac{\rho_{(3.1)}}{\Phi_1^2(r)}\frac{\rho_{(5.14)}}{\Phi_2^2(r)}\frac{\rho_{(5.16)}}{\Phi_3^2(r)}}{e^{2[\dot{b}]_\infty (T-r)}\left(\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} + \frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)} + \frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)}\frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}\right)^2} \times n^2$$

then as

$$\Phi(r, [\dot{b}]_\infty) = \left(\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)} + \frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)} + \frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)}\frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}\right)$$

$$\times \left(\frac{\Phi_3^2(r)C_{(5.16)}}{\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)}} + \frac{\phi_2(r, [\dot{b}]_\infty)}{\frac{\sqrt{\rho_{(3.1)}}}{\Phi_1(r)}\frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}} + \frac{\Phi_1^2(r)C_{(3.2)}(0)}{\frac{\sqrt{\rho_{(5.14)}}}{\Phi_2(r)}\frac{\sqrt{\rho_{(5.16)}}}{\Phi_3(r)}}\right),$$

we get

$$\mathbb{E}\left[\exp\left\{\rho|\mathcal{D}_r^j U_T^\star|^2\right\}\right] \le \exp\left\{\rho e^{2[\dot{b}]_\infty (T-r)}\Phi(r, [\dot{b}]_\infty)\frac{(m-1)T^2}{mn^2}\right\}. \qquad \square$$

6 Proof of Theorem 3.2

For $\lambda \in \mathbb{R}$, by independence,

$$\mathbb{E}\left[\exp\left(\lambda[\hat{Q} - \mathbb{E}f(X_T^{m^L})]\right)\right] = \prod_{\ell=0}^{L} \mathbb{E}\left[\exp(\lambda\hat{Q}_\ell)\right], \text{ where} \quad (6.1)$$

$$\mathbb{E}\left[\exp(\lambda\hat{Q}_\ell)\right] = \left(\mathbb{E}\left[\exp\left\{\frac{\lambda}{N_\ell}\left(f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) - \mathbb{E}[f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})]\right)\right\}\right]\right)^{N_\ell}$$

and $\mathbb{E}\left[\exp(\lambda\hat{Q}_0)\right] \le \exp\left\{\frac{\lambda^2 T[\dot{f}]_\infty^2}{2N_0}\right\}$,

where we used the Gaussian concentration bound (2.3) to get the last the inequality. For $\ell \in \{1, \dots, L\}$, we set $n \in \mathbb{N}^*$ and define

$$\Upsilon := f(X_T^{mn}) - f(X_T^n) - \mathbb{E}[f(X_T^{mn}) - f(X_T^n)].$$

For $\tilde{\lambda} \in \mathbb{R}$, we want to obtain an estimation of $\mathbb{E}\left[\exp(\tilde{\lambda}\Upsilon)\right]$ of the form $\exp\left\{C\tilde{\lambda}^2 \frac{(m-1)T^2}{mn^2}\right\}$ where C is an explicit constant and $\frac{(m-1)T^2}{mn^2}$ is the order of the variance of the centered random variable Υ according to Proposition 2.1. To do so, we assume that $f \in \mathscr{C}_b^1(\mathbb{R}^d, \mathbb{R})$

is Lipschitz continuous with constant $[\dot{f}]_\infty$ and such that ∇f is also Lipschitz with constant $[\dot{f}]_{\mathrm{lip}}$. By Clarck's Ocone formula we have

$$f(X_T^{mn}) - f(X_T^n) - \mathbb{E}[f(X_T^{mn}) - f(X_T^n)] = \int_0^T \mathbb{E}\left[K_r | \mathcal{F}_r\right] \cdot dW_r,$$

where, for $j \in \{1, \ldots, d\}$, the jth-component of the d-dimensional vector K_r is given by $K_{r,j} := \mathcal{D}_r^j f(X_T^{mn}) - \mathcal{D}_r^j f(X_T^n)$. For $p \in (0,1)$, we use Hölder's inequality to get

$$\mathbb{E}\left[\exp(\tilde{\lambda}\Upsilon)\right] \leq \mathbb{E}^p\left[\exp\left\{\frac{\tilde{\lambda}}{p}\int_0^T \mathbb{E}\left[K_r | \mathcal{F}_r\right] \cdot dW_r - \frac{\tilde{\lambda}^2}{2p^2}\int_0^T |\mathbb{E}\left[K_r | \mathcal{F}_r\right]|^2 dr\right\}\right]$$
$$\times \mathbb{E}^{1-p}\left[\exp\left\{\frac{\tilde{\lambda}^2}{2p(1-p)}\int_0^T |\mathbb{E}\left[K_r | \mathcal{F}_r\right]|^2 dr\right\}\right].$$

Now, by the Malliavin chain rule we have

$$K_r = \mathcal{D}_r X_T^{mn} \nabla f(X_T^{mn}) - \mathcal{D}_r X_T^n \nabla f(X_T^n), \tag{6.2}$$

where $\mathcal{D}_r X_T^n = (\mathcal{D}_r^i X_{T,j}^n)_{1 \leq i,j \leq d} \in \mathbb{R}^{d \times d}$.

According to Lemma 5.3 and under our assumption on the boundedness of ∇f, we easily check that $\sup_{r \in [0,T]} |K_r|^2 \leq 4d e^{2T[\dot{b}]_\infty}[\dot{f}]_\infty^2$. Therefore, the process

$$\left(\exp\left\{\frac{\tilde{\lambda}}{p}\int_0^t \mathbb{E}\left[K_r | \mathcal{F}_r\right] \cdot dW_r - \frac{\tilde{\lambda}^2}{2p^2}\int_0^t |\mathbb{E}\left[K_r | \mathcal{F}_r\right]|^2 dr\right\}\right)_{0 \leq t \leq T}$$

is a martingale, which together with the choice $p = 1/2$ which minimizes $\frac{1}{p(1-p)}$ leads us to

$$\mathbb{E}\left[\exp(\tilde{\lambda}\Upsilon)\right] \leq \mathbb{E}^{1/2}\left[\exp\left\{2\tilde{\lambda}^2\int_0^T |\mathbb{E}\left[K_r | \mathcal{F}_r\right]|^2 dr\right\}\right].$$

Applying Jensen's inequality twice, and now denoting by p a measurable positive function such that $\int_0^T p(r)dr = 1$, we obtain that

$$\mathbb{E}\left[\exp\left\{2\tilde{\lambda}^2\int_0^T |\mathbb{E}\left[K_r | \mathcal{F}_r\right]|^2 dr\right\}\right] \leq \int_0^T \mathbb{E}\left[\exp\left\{\frac{2\tilde{\lambda}^2 T}{p(r)}|K_r|^2\right\}\right] p(r)dr$$

and deduce that

$$\mathbb{E}\left[\exp(\tilde{\lambda}\Upsilon)\right] \leq \left(\int_0^T \mathbb{E}\left[\exp\left\{\frac{2\tilde{\lambda}^2 T}{p(r)}|K_r|^2\right\}\right] p(r)dr\right)^{1/2}. \tag{6.3}$$

We now want to estimate the moment generating function of $|K_r|^2$. Setting

$$U_T := X_T^{mn} - X_T^n,$$

remarking that $\|\mathcal{D}_r U_T\|^2 \leq \mathrm{Tr}\left[\mathcal{D}_r U_T (\mathcal{D}_r U_T)^\top\right] = \sum_{j=1}^d |\mathcal{D}_r^j U_T|^2$ and, by Lemma 5.3, $\|D_r X_T^{mn}\| \leq \sqrt{d} e^{[\dot{b}]_\infty(t-r)}$, we obtain that

$$|K_r| \leq \|D_r X_T^{mn}\|\|\nabla f(X_T^{mn}) - \nabla f(X_T^n)\| + \|\mathcal{D}_r X_T^{mn} - \mathcal{D}_r X_T^n\|\|\nabla f(X_T^n)\|$$
$$\leq \sqrt{d}[\dot{f}]_{\mathrm{lip}} e^{[\dot{b}]_\infty(t-r)}|U_T| + [\dot{f}]_\infty \left(\sum_{j=1}^d |\mathcal{D}_r^j U_T|^2\right)^{1/2}. \tag{6.4}$$

A careful look at the proof of this theorem shows that, in the decomposition (5.6) of $\mathcal{D}_r^j U_T$, the sum $\mathcal{D}_r^j U_T^{(1-3)} := \mathcal{D}_r^j U_T^{(1)} + \mathcal{D}_r^j U_T^{(2)} + \mathcal{D}_r^j U_T^{(3)}$ goes to 0 as $r \to T$ whereas $\mathcal{D}_r^j U_T^{(0)}$ does not. This indicates that it is not optimal to combine $\mathcal{D}_r^j U_T^{(0)}$ with $\mathcal{D}_r^j U_T^{(2)}$ as in this proof. We also notice that under the same constraint on ρ as in the theorem,

$$\mathbb{E}\left[\exp\left\{\rho|\mathcal{D}_r^j U_T^{(1-3)}|^2\right\}\right] \leq \exp\left\{\rho e^{2[\dot{b}]_\infty (T-r)}\Phi(r,0)\frac{(m-1)T^2}{mn^2}\right\}. \tag{6.5}$$

Since U_T does not depend on r, it should be better to combine $\mathcal{D}^j U_T^{(0)}$ with it by replacing (6.4) by the estimation

$$|K_r| \leq \sqrt{d}[\dot{f}]_{\mathrm{lip}} e^{[\dot{b}]_\infty (T-r)}|U_T| + [\dot{f}]_\infty \left(\sum_{j=1}^d |\mathcal{D}_r^j U_T^{(0)}|^2\right)^{1/2} + [\dot{f}]_\infty \left(\sum_{j=1}^d |\mathcal{D}_r^j U_T^{(1-3)}|^2\right)^{1/2}$$

$$\leq \sqrt{d}[\dot{f}]_{\mathrm{lip}} e^{[\dot{b}]_\infty (T-r)}\left(\frac{(m-1)T^2 x}{2mn} + |U_T|\right) + [\dot{f}]_\infty \left(\sum_{j=1}^d |\mathcal{D}_r^j U_T^{(1-3)}|^2\right)^{1/2},$$

where $x = \dfrac{2[\dot{b}]_\infty [\dot{f}]_\infty}{T[\dot{f}]_{\mathrm{lip}}}$,

which takes (5.7) into account. One deduces that for $\kappa(r) \in (0,1)$,

$$|K_r|^2 \leq \frac{1}{\kappa(r)}d[\dot{f}]_{\mathrm{lip}}^2 e^{2[\dot{b}]_\infty (T-r)}\left(\frac{(m-1)T^2 x}{2mn} + |U_T|\right)^2 + \frac{[\dot{f}]_\infty^2}{(1-\kappa(r))}\sum_{j=1}^d |\mathcal{D}_r^j U_T^{(1-3)}|^2.$$

Combining (6.3), Hölder's inequality and the convexity of the exponential function which ensures that the exponential of the mean of d terms is not greater than the mean of the exponentials, we deduce that

$$\mathbb{E}^2\left[\exp(\tilde{\lambda}\Upsilon)\right] \leq \int_0^T \mathbb{E}^{\kappa(r)}\left[\exp\left\{\frac{2d\tilde{\lambda}^2 T[\dot{f}]_{\mathrm{lip}}^2 e^{2[\dot{b}]_\infty (T-r)}}{p(r)\kappa^2(r)}\left(\frac{(m-1)T^2 x}{2mn} + |U_T|\right)^2\right\}\right]$$

$$\times \left(\frac{1}{d}\sum_{j=1}^d \mathbb{E}\left[\exp\left\{\frac{2d\tilde{\lambda}^2 T[\dot{f}]_\infty^2}{p(r)(1-\kappa(r))^2}|\mathcal{D}_r^j U_T^{(1-3)}|^2\right\}\right]\right)^{1-\kappa(r)} p(r)dr. \tag{6.6}$$

We now choose $\kappa(r) = \dfrac{[\dot{f}]_{\mathrm{lip}}\sqrt{\hat{\rho}(r)}}{[\dot{f}]_\infty \sqrt{\rho_{(3.1)}} + [\dot{f}]_{\mathrm{lip}}\sqrt{\hat{\rho}(r)}}$ to obtain the same constraint on $\tilde{\lambda}$ for the two expectations at time r and then $p(r) \propto \dfrac{e^{2[\dot{b}]_\infty (T-r)}}{\hat{\rho}(r)}\left([\dot{f}]_\infty \sqrt{\rho_{(3.1)}} + [\dot{f}]_{\mathrm{lip}}\sqrt{\hat{\rho}(r)}\right)^2$ to ensure that this common constraint does not depend on r. Notice that since the functions Φ_i are continuous on $[0,T]$, positive on $[0,T)$ and such that $\Phi_1(r) = \mathcal{O}(T-r)$, $\Phi_2(r) = \mathcal{O}(T-r)$ and $\Phi_3(r) = \mathcal{O}(\sqrt{T-r})$ so that $\hat{\rho}(r) = \mathcal{O}((T-r)^{-1})$ as $r \to T-$, the function p is bounded and therefore integrable on $[0,T]$. We conclude that if

$$\tilde{\lambda}^2 \leq \frac{\rho_{(3.1)}n^2}{2dT\int_0^T \frac{e^{2[\dot{b}]_\infty (T-t)}}{\hat{\rho}(t)}\left([\dot{f}]_\infty \sqrt{\rho_{(3.1)}} + [\dot{f}]_{\mathrm{lip}}\sqrt{\hat{\rho}(t)}\right)^2 dt},$$

$$\mathbb{E}^2\left[\exp(\tilde{\lambda}\Upsilon)\right] \leq$$
$$\int_0^T \exp\left\{\frac{2d\tilde{\lambda}^2 T([\dot{f}]_{\mathrm{lip}}\sqrt{\hat{\rho}(r)}C_{(3.2)}(2[\dot{b}]_\infty[\dot{f}]_\infty/T[\dot{f}]_{\mathrm{lip}}) + [\dot{f}]_\infty \hat{\rho}(r)\Phi(r,0)/\sqrt{\rho_{(3.1)}})}{[\dot{f}]_\infty \sqrt{\rho_{(3.1)}} + [\dot{f}]_{\mathrm{lip}}\sqrt{\hat{\rho}(r)}}\right.$$

$$\times \int_0^T \frac{e^{2[\dot{b}]_\infty(T-t)}}{\hat{\rho}(t)} \left([\dot{f}]_\infty\sqrt{\rho_{(3.1)}} + [\dot{f}]_{\text{lip}}\sqrt{\hat{\rho}(t)}\right)^2 dt\Big\}p(r)dr$$

$$\leq \exp\left\{\frac{2\tilde{\lambda}^2[\dot{f}]_\infty^2 C_{(6.7)}(m-1)T^2}{mn^2}\right\}, \tag{6.7}$$

where

$$C_{(6.7)} = dT \int_0^T \frac{e^{2[\dot{b}]_\infty(T-t)}}{\hat{\rho}(t)} \left(\sqrt{\rho_{(3.1)}} + \frac{[\dot{f}]_{\text{lip}}}{[\dot{f}]_\infty}\sqrt{\hat{\rho}(t)}\right)^2 dt$$

$$\times \sup_{r\in[0,T)} \frac{[\dot{f}]_{\text{lip}}\sqrt{\hat{\rho}(r)}C_{(3.2)}(2[\dot{b}]_\infty[\dot{f}]_\infty/T[\dot{f}]_{\text{lip}}) + [\dot{f}]_\infty\hat{\rho}(r)\Phi(r,0)/\sqrt{\rho_{(3.1)}}}{[\dot{f}]_\infty\sqrt{\rho_{(3.1)}} + [\dot{f}]_{\text{lip}}\sqrt{\hat{\rho}(r)}}$$

is finite since, as $r \to T-$, $\phi_2(r,0) = \mathcal{O}((T-r)^2)$ and $\Phi(r,0) = \mathcal{O}((T-r))$. We complete the proof using (6.1).

A Proofs of the technical lemmas

Lemma A.1. *Let $(H_t)_{t\leq T}$ be an adapted $\mathbb{R}^d$-valued process and $|H| := \left\|\left(\int_0^T |H_t|^2 dt\right)^{1/2}\right\|_\infty$.
Then $\forall \delta \geq 0$, $\forall \mu \in \left[0, \frac{1}{4|H|^2}\right)$,*

$$\mathbb{E}\left(e^{\mu(\delta+\sup_{t\in[0,T]}|\int_0^t H_s.dW_s|)^2}\right) \leq \exp\left\{\frac{2\mu\delta^2}{1-4\mu|H|^2}\right\}\left(\frac{8\mu\delta|H|}{\sqrt{2\pi}(1-4\mu|H|^2)} + \frac{1}{\sqrt{1-4\mu|H|^2}}\right),$$

where the right-hand side is smaller than $\exp\left\{4\mu\left(\delta^2 + \frac{2}{\sqrt{\pi}}\delta|H| + |H|^2\ln 2\right)\right\}$ when moreover $\mu \in \left[0, \frac{1}{8|H|^2}\right]$.

For $(M(t))_{t\leq T}$ an adapted $\mathbb{R}^{d\times d}$-valued process and $|M| := \left\|\left(\int_0^T \text{Tr}(M(t)M^*(t))dt\right)^{1/2}\right\|_\infty$,

$$\forall \mu \in \left[0, \frac{1}{4|M|^2}\right), \ \mathbb{E}\left(e^{\mu\sup_{t\in[0,T]}|\int_0^t M_s dW_s|^2}\right) \leq \frac{1}{\sqrt{1-4\mu|M|^2}}$$

$$\text{and } \forall \mu \in \left[0, \frac{1}{8|M|^2}\right], \ \mathbb{E}\left(e^{\mu\sup_{t\in[0,T]}|\int_0^t M_s dW_s|^2}\right) \leq e^{4\mu|M|^2\ln 2}.$$

Remark A.2. Let $h(t) = \|\|H_t\|\|_\infty$ for $t \in [0,T]$. Then

$$\forall \mu \in \left[0, \frac{1}{2\|h\|_2^2}\right], \ \sup_{t\in[0,T]} \mathbb{E}\left(e^{\mu(\int_0^t H_s.dW_s)^2}\right) \leq \frac{1}{\sqrt{1-2\mu\|h\|_2^2}},$$

where $\|h\|_2^2 = \int_0^T h^2(t)dt \geq |H|^2$.

When $\|h\|_2^2 < +\infty$, this is a consequence of the convexity of $x \mapsto e^{\mu x^2}$ and Jensen's inequality. Indeed introducing (on a possibly enlarged probability space) $(\beta_t)_{t\in[0,T]}$ a one-dimensional Brownian motion independent from $\mathcal{F}^{W,H} = \sigma((H_t, W_t), t \in [0,T])$, we obtain that for $t \in [0,T]$, $\int_0^t H_s.dW_s + \int_0^t \sqrt{h^2(s) - |H_s|^2}d\beta_s$ is a centered Gaussian random variable with variance equal to $\int_0^t h^2(s)ds$ such that

$$\mathbb{E}\left(\int_0^t H_s.dW_s + \int_0^t \sqrt{h^2(s) - |H_s|^2}d\beta_s\Big|\mathcal{F}^{W,H}\right) = \int_0^t H_s.dW_s.$$

Proof. The argument is based on the Dambins-Dubins-Schwarz (see e.g. [9]) theorem which ensures the existence of a one-dimensional standard Brownian motion $(\beta_t)_{t\geq 0}$ such that $\forall t \geq 0$, $\int_0^t H_s.dW_s = \beta_{\int_0^t |H_s|^2 ds}$. Hence

$$
\begin{aligned}
\mathbb{E}\left(e^{\mu(\delta + \sup_{t\in[0,T]}|\int_0^t H_s.dW_s|)^2}\right) &= \mathbb{E}\left(e^{\mu\left(\delta + \sup_{t\in[0,T]}|\beta_{\int_0^t |H_s|^2 ds}|\right)^2}\right) \\
&\leq \mathbb{E}\left(e^{\mu(\delta + \sup_{s\in[0,|H|^2]}|\beta_s|)^2}\right) \\
&\leq \mathbb{E}\left(e^{\mu(\delta + \sup_{s\in[0,|H|^2]}\beta_s)^2}e^{\mu(\delta - \inf_{s\in[0,|H|^2]}\beta_s)^2}\right) \\
&\leq \mathbb{E}^{1/2}\left(e^{2\mu(\delta + \sup_{s\in[0,|H|^2]}\beta_s)^2}\right)\mathbb{E}^{1/2}\left(e^{2\mu(\delta - \inf_{s\in[0,|H|^2]}\beta_s)^2}\right) \\
&= \mathbb{E}\left(e^{2\mu(\delta + |\beta_{|H|^2}|)^2}\right),
\end{aligned}
$$

where we have used that $\sup_{s\in[0,|H|^2]}\beta_s$ and $-\inf_{s\in[0,|H|^2]}\beta_s$ have the same law as $|\beta_{|H|^2}|$ for the last equality. Now, using the change of variables $y = x - 4\mu\delta|H|^2/(1-4\mu|H|)$ for the second equality, we obtain that

$$
\begin{aligned}
\mathbb{E}\left(e^{2\mu(\delta + |\beta_{|H|^2}|)^2}\right) &= 2\int_0^\infty \exp\left\{2\mu(\delta + x)^2 - \frac{x^2}{2|H|^2}\right\}\frac{dx}{|H|\sqrt{2\pi}} \\
&= 2\exp\left\{\frac{2\mu\delta^2}{1-4\mu|H|^2}\right\}\int_{-4\mu\delta|H|^2/(1-4\mu|H|^2)}^\infty \exp\left\{-\frac{(1-4\mu|H|^2)y^2}{2|H|^2}\right\}\frac{dy}{|H|\sqrt{2\pi}} \\
&\leq \exp\left\{\frac{2\mu\delta^2}{1-4\mu|H|^2}\right\}\left(\frac{8\mu\delta|H|}{\sqrt{2\pi}(1-4\mu|H|^2)} + \int_{-\infty}^\infty \exp\left\{-\frac{(1-4\mu|H|^2)y^2}{2|H|^2}\right\}\frac{dy}{|H|\sqrt{2\pi}}\right) \\
&= \exp\left\{\frac{2\mu\delta^2}{1-4\mu|H|^2}\right\}\left(\frac{8\mu\delta|H|}{\sqrt{2\pi}(1-4\mu|H|^2)} + \frac{1}{\sqrt{1-4\mu|H|^2}}\right).
\end{aligned}
$$

The concavity of the logarithm ensures that $\forall x \in [0, \frac{1}{2}]$, $\ln(1-x) \geq -2x\ln 2$ so that $\frac{1}{\sqrt{1-x}} = e^{-\frac{1}{2}\ln(1-x)} \leq e^{x\ln 2}$. Therefore when $\mu \in \left[0, \frac{1}{8|H|^2}\right]$,

$$
\begin{aligned}
\exp\left\{\frac{2\mu\delta^2}{1-4\mu|H|^2}\right\}&\left(\frac{8\mu\delta|H|}{\sqrt{2\pi}(1-4\mu|H|^2)} + \frac{1}{\sqrt{1-4\mu|H|^2}}\right) \\
&\leq e^{4\mu\delta^2}\left(\frac{8\mu\delta|H|}{\sqrt{\pi}} + 1\right)\frac{1}{\sqrt{1-4\mu|H|^2}} \leq e^{4\mu\delta^2}e^{\frac{8\mu\delta|H|}{\sqrt{\pi}}}e^{4\mu|H|^2\ln 2}.
\end{aligned}
$$

Let now for $i \in \{1, \ldots, d\}$, $M_i(t)$ denote the i-th line of the matrix $M(t)$ and $|M_i| := \left\|\left(\int_0^T \sum_{j=1}^d M_{ij}^2(t)dt\right)^{1/2}\right\|_\infty$. For $\mu < \frac{1}{4|M|^2}$, we have

$$
\begin{aligned}
\mathbb{E}\left(e^{\mu\sup_{t\in[0,T]}|\int_0^t M(s)dW_s|^2}\right) &= \mathbb{E}\left(e^{\mu\sup_{t\in[0,T]}\sum_{i=1}^d(\int_0^t M_i(s)dW_s)^2}\right) \\
&\leq \mathbb{E}\left(e^{\mu\sum_{i=1}^d \sup_{t\in[0,T]}(\int_0^t M_i(s)dW_s)^2}\right) \\
&\leq \mathbb{E}\left(\sum_{i=1}^d \frac{|M_i|^2}{|M|^2}e^{\frac{\mu|M|^2}{|M_i|^2}\sup_{t\in[0,T]}(\int_0^t M_i(s)dW_s)^2}\right) \\
&\leq \frac{1}{\sqrt{1-4\mu|M|^2}},
\end{aligned}
$$

where we used Jensen's inequality for the third inequality and the first statement of the Lemma for the fourth. $\qquad\square$

References

[1] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014. MR-3155209

[2] D. Belomestny and T. Nagapetyan, *Multilevel path simulation for weak approximation schemes with application to Lévy-driven SDEs*, Bernoulli **23** (2017), no. 2, 927–950. MR-3606755

[3] N. Frikha and S. Menozzi, *Concentration bounds for stochastic approximations*, Electron. Commun. Probab. **17** (2012), no. 47, 15. MR-2988393

[4] M. B. Giles, *Multilevel Monte Carlo path simulation*, Oper. Res. **56** (2008), no. 3, 607–617. MR-2436856

[5] M. B. Giles and L. Szpruch, *Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation*, Ann. Appl. Probab. **24** (2014), no. 4, 1585–1620. MR-3211005

[6] C. Houdré and N. Privault, *Concentration and deviation inequalities in infinite dimensions via covariance representations*, Bernoulli **8** (2002), no. 6, 697–720. MR-1962538

[7] G. N. Milstein, *Numerical integration of stochastic differential equations*, Mathematics and its Applications, vol. 313, Kluwer Academic Publishers Group, Dordrecht, 1995, Translated and revised from the 1988 Russian original. MR-1335454

[8] D. Nualart, *The Malliavin calculus and related topics*, Probability and its Applications (New York), Springer-Verlag, New York, 1995. MR-1344217

[9] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR-MR1725357