



Fisher information bounds and applications to SDEs with small noise

Nguyen Tien Dung^{a,*}, Nguyen Thu Hang^b^a Department of Mathematics, VNU University of Science, Vietnam National University, Hanoi, 334 Nguyen Trai, Thanh Xuan, Hanoi, 084, Viet Nam^b Department of Mathematics, Hanoi University of Mining and Geology, 18 Pho Vien, Bac Tu Liem, Hanoi, 084, Viet Nam

ARTICLE INFO

MSC:

60H07

94A17

60H10

Keywords:

Malliavin calculus

Fisher information

Small noise SDEs

ABSTRACT

In this paper, we first establish general bounds on the Fisher information distance to the class of normal distributions of Malliavin differentiable random variables. We then study the rate of Fisher information convergence in the central limit theorem for the solution of small noise stochastic differential equations and its additive functionals. We also show that the convergence rate is of optimal order.

1. Introduction

Given a random variable F with an absolutely continuous density p_F , the Fisher information of F (or its distribution) is defined by

$$I(F) = \int_{-\infty}^{+\infty} \frac{p'_F(x)^2}{p_F(x)} dx = E[\rho_F^2(F)],$$

where p'_F denotes a Radon–Nikodym derivative of p_F and $\rho_F := p'_F/p_F$ is the score function. Furthermore, the Fisher information distance of F to the normal distribution $N \sim \mathcal{N}(\mu, \sigma^2)$ is defined by

$$I(F \parallel N) := E \left[\left(\rho_F(F) + \frac{F - \mu}{\sigma^2} \right)^2 \right].$$

If the derivative p'_F does not exist then the Fisher information distance is defined to be infinite. We refer the reader to the monograph [5] for various properties of $I(F)$ and $I(F \parallel N)$. Here we note that $I(F \parallel N)$ is not a distance. However, it dominates many important distances in statistical applications such as the relative entropy (or Kullback–Leibler distance) and the supremum distance between densities. It also dominates the traditional distances such as Kolmogorov distance, total variation distance and Wasserstein distance, etc. Thus $I(F \parallel N)$ provides a very strong measure of convergence to normality when studying the central limit theorem.

The study of Fisher information convergence has a long history beginning in 1959 with the results of Linnik [9]. However, most of the existing results are devoted to the sums of independent random variables. For such sums, the quantitative estimates for the rate of convergence have been well studied, see [1,5–7]. Recently, Nourdin & Nualart [10] used the techniques of Malliavin calculus to obtain quantitative Fisher information bounds for the multiple Wiener–Itô integrals. This is a remarkable contribution

* Corresponding author.

E-mail address: dung@hus.edu.vn (N.T. Dung).

to the literature because, for the first time, the quantitative estimates for $I(F \parallel N)$ were obtained for the random variables F not to be a sum of independent random variables.

In the present paper, our first purpose is to extend the method developed in [10] to a general class of Malliavin differentiable random variables. We provide two explicit estimates for the Fisher information distance in [Theorems 3.1](#) and [3.2](#). Our second purpose is to investigate the rate of Fisher information convergence in the central limit theorem for the solution of stochastic differential equations (SDEs) with small noise:

$$X_{\varepsilon,t} = X_0 + \int_0^t b(s, X_{\varepsilon,s})ds + \varepsilon \int_0^t \sigma(s, X_{\varepsilon,s})dB_s, \quad t \in [0, T], \quad (1.1)$$

where the initial condition X_0 is a real number, $b, \sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ are deterministic functions, $(B_t)_{t \in [0, T]}$ is a standard Brownian motion and $\varepsilon \in (0, 1)$ is a small parameter. Let us consider the ordinary differential equation

$$x_t = X_0 + \int_0^t b(s, x_s)ds, \quad t \in [0, T]. \quad (1.2)$$

It is well known that, in theory of stochastic differential equations with small noise, one of the fundamental problems is to study the convergence of $X_{\varepsilon,t}$ to x_t as ε tends zero. The convergence can be described via large deviation principle, central limit theorem and moderate deviation principle, etc. (see two monographs [4,8] for more details). In particular, the central limit theorem results have been discussed by various authors. Under suitable assumptions, for example, it follows from [14,16] that, as $\varepsilon \rightarrow 0$,

$$\tilde{X}_{\varepsilon,t} := \frac{X_{\varepsilon,t} - x_t}{\varepsilon} \text{ converges in distribution to } N_t$$

for every $t \in [0, T]$, where N_t is a centered normal random variable with appropriate variance. Naturally, one may wonder whether the convergence also holds in the Fisher information distance (i.e. convergence holds in a stronger sense). Our [Theorem 4.1](#) below not only gives an affirmative answer, but also provides a convergence rate of order $O(\varepsilon^2)$. In addition, in [Theorem 4.2](#), we show that the rate $O(\varepsilon^2)$ is optimal as $\varepsilon \rightarrow 0$. In [Theorem 4.3](#), as a further illustration, we also obtain the optimal rate $O(\varepsilon^2)$ of Fisher information convergence for the additive functional of the form

$$Y_{\varepsilon,t} = \int_0^t f(s, X_{\varepsilon,s})ds, \quad t \in [0, T]. \quad (1.3)$$

The rest of the paper is organized as follows. In [Section 2](#), we recall some fundamental concepts of Malliavin calculus. [Section 3](#) contains the abstract results of this paper, two upper bounds on the Fisher information distance are given in [Theorems 3.1](#) and [3.2](#). [Section 4](#) is devoted to the study of Fisher information convergence for the solution of [\(1.1\)](#) and its additive functional [\(1.3\)](#). The main results of this section is formulated and proved in [Theorems 4.1](#), [4.2](#) and [4.3](#). An useful estimate for the negative moment of Volterra functionals is given in [Appendix](#).

2. Preliminaries

As we have said in the introduction, this paper is concerned with the Fisher information distance via the techniques of Malliavin calculus. For the reader's convenience, let us recall some elements of Malliavin calculus (for more details see [12]). We suppose that $(B_t)_{t \in [0, T]}$ is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P}, P)$, where $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ is a natural filtration generated by the Brownian motion B . For $h \in L^2[0, T]$, we denote by $B(h)$ the Wiener integral

$$B(h) = \int_0^T h(t)dB_t.$$

Let \mathcal{S} denote a dense subset of $L^2(\Omega, \mathcal{F}, P)$ that consists of smooth random variables of the form

$$F = f(B(h_1), B(h_2), \dots, B(h_n)), \quad (2.1)$$

where $n \in \mathbb{N}$, $f \in C_b^\infty(\mathbb{R}^n)$ and $h_1, h_2, \dots, h_n \in L^2[0, T]$. If F has the form [\(2.1\)](#), we define its Malliavin derivative as the process $DF := D_t F, t \in [0, T]$ given by

$$D_t F = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(B(h_1), B(h_2), \dots, B(h_n))h_k(t).$$

More generally, for each $k \geq 1$, we can define the iterated derivative operator on a cylindrical random variable by setting

$$D_{t_1, \dots, t_k}^k F = D_{t_1} \dots D_{t_k} F.$$

For any $1 \leq p, k < \infty$, we denote by $\mathbb{D}^{k,p}$ the closure of \mathcal{S} with respect to the norm

$$\|F\|_{k,p}^p := E|F|^p + E \left[\left(\int_0^T |D_u F|^2 du \right)^{\frac{p}{2}} \right] + \dots + E \left[\left(\int_0^T \dots \int_0^T |D_{t_1, \dots, t_k}^k F|^2 dt_1 \dots dt_k \right)^{\frac{p}{2}} \right].$$

A random variable F is said to be Malliavin differentiable if it belongs to $\mathbb{D}^{1,2}$. For any $F \in \mathbb{D}^{1,2}$, the Clark–Ocone formula says that

$$F - E[F] = \int_0^T E[D_s F | \mathcal{F}_s] dB_s.$$

In particular, we have

$$\text{Var}(F) = \int_0^T E(E[D_s F | \mathcal{F}_s])^2 ds = \int_0^T E[D_s F E[D_s F | \mathcal{F}_s]] ds.$$

An important operator in the Malliavin's calculus theory is the divergence operator δ . It is the adjoint of derivative operator D . The domain of δ is the set of all functions $u \in L^2(\Omega \times [0, T])$ such that

$$E|\langle DF, u \rangle_{L^2[0, T]}| \leq C(u)\|F\|_{L^2(\Omega)},$$

where $C(u)$ is some positive constant depending on u . In particular, if $u \in \text{Dom}\delta$, then $\delta(u)$ is characterized by following duality relationships

$$\delta(uF) = F\delta(u) - \langle DF, u \rangle_{L^2[0, T]} \quad (2.2)$$

$$E[\langle DF, u \rangle_{L^2[0, T]}] = E[F\delta(u)] \text{ for any } F \in \mathbb{D}^{1,2}. \quad (2.3)$$

It is known that any random variable F in $L^2(\Omega, \mathcal{F}, P)$ can be expanded into an orthogonal sum of its Wiener chaos:

$$F = \sum_{n=0}^{\infty} J_n F,$$

where $J_0 = E[F]$ and J_n denotes the projection onto the n th Wiener chaos. From this chaos expansion one may define the Ornstein–Uhlenbeck operator L by

$$LF = \sum_{n=0}^{\infty} -nJ_n F.$$

The domain of L is

$$\text{Dom}L = \{F \in L^2(\Omega) : \sum_{n=1}^{\infty} n^2 E|J_n F|^2 < \infty\} = \mathbb{D}^{2,2}.$$

Moreover, a random variable F belongs to $\text{Dom}L$ if and only if $F \in \text{Dom}\delta D$ (i.e. $F \in \mathbb{D}^{1,2}$ and $DF \in \text{Dom}\delta$), and in this case: $\delta DF = -LF$. We also define the operator L^{-1} as follows: for every $F \in L^2(\Omega)$ with zero mean, we set

$$L^{-1}F = \sum_{n=1}^{\infty} -\frac{1}{n} J_n F.$$

Note that, for any $F \in L^2(\Omega)$ with zero mean, we have that $L^{-1}F \in \text{Dom}L$, and

$$LL^{-1}F = F.$$

3. General Fisher information bounds

We first construct the representation formula for the score function.

Lemma 3.1. *Let $F \in \mathbb{D}^{1,2}$ and $u : \Omega \rightarrow L^2[0, T]$, and suppose that $\langle DF, u \rangle_{L^2[0, T]} \neq 0$ a.s. and $\frac{u}{\langle DF, u \rangle_{L^2[0, T]}}$ belongs to the domain of δ . Then the law of F has an absolutely continuous density and its score function ρ_F is given by*

$$\rho_F(x) := p'_F(x)/p_F(x) = -E\left[\delta\left(\frac{u}{\langle DF, u \rangle_{L^2[0, T]}}\right) \middle| F = x\right], \quad x \in \text{supp } p_F. \quad (3.1)$$

Proof. This lemma is not new. For the sake of completeness, we shall give a proof here. According to Exercise 2.1.3 in [12], the law of F has a continuous density given by

$$p_F(x) = E\left[\mathbf{1}_{\{F > x\}} \delta\left(\frac{u}{\langle DF, u \rangle_{L^2[0, T]}}\right)\right], \quad x \in \text{supp } p_F. \quad (3.2)$$

Note that the proof of (3.2) is similar to that of Proposition 2.1.1 in [12]. Since $F \in \mathbb{D}^{1,2}$, this implies that $\text{supp } p_F$ is a closed interval of \mathbb{R} (see Proposition 2.1.7 in [12]): $\text{supp } p_F = [\alpha, \beta]$ with $-\infty \leq \alpha < \beta \leq \infty$. It follows from (3.2) that

$$\begin{aligned} p_F(x) &= E\left[\mathbf{1}_{\{F > x\}} E\left[\delta\left(\frac{u}{\langle DF, u \rangle_{L^2[0, T]}}\right) \middle| F\right]\right] \\ &= \int_x^\beta E\left[\delta\left(\frac{u}{\langle DF, u \rangle_{L^2[0, T]}}\right) \middle| F = y\right] p_F(y) dy. \end{aligned}$$

So p_F is absolutely continuous and the representation (3.1) is verified. The proof of the lemma is complete. \square

We now are ready to establish the Fisher information bounds by using suitable choices of the function u .

Theorem 3.1. Let $F \in \mathbb{D}^{2,4}$ and N be a normal random variable with mean μ and variance σ^2 . Define

$$\Theta := \langle DF, u \rangle_{L^2[0,T]},$$

where $u_t := E[D_t F | \mathcal{F}_t]$, $t \in [0, T]$. Assume that $\Theta \neq 0$ a.s. Then, we have

$$I(F \parallel N) \leq c \left(\frac{1}{\sigma^4} (E[F] - \mu)^2 + A_F |\text{Var}(F) - \sigma^2|^2 + C_F \left(E \|D\Theta\|_{L^2[0,T]}^4 \right)^{1/2} \right), \quad (3.3)$$

where c is an absolute constant and A_F, C_F are positive constants given by

$$A_F := \frac{1}{\sigma^4} \left(E \|u\|_{L^2[0,T]}^8 E |\Theta|^{-8} \right)^{1/4}, \quad C_F := A_F + \left(E \|u\|_{L^2[0,T]}^8 E |\Theta|^{-16} \right)^{1/4}.$$

Proof. For simplicity, we write $\langle \cdot, \cdot \rangle$ instead of $\langle \cdot, \cdot \rangle_{L^2[0,T]}$ and $\|\cdot\|$ instead of $\|\cdot\|_{L^2[0,T]}$. In Lemma 3.1, we use $u_t := E[D_t F | \mathcal{F}_t]$, $t \in [0, T]$. Note that, by the Clark–Ocone formula, we have

$$F - E[F] = \int_0^T E[D_t F | \mathcal{F}_t] dB_t = \delta(u).$$

Hence, by the integration-by-part formula (2.2), we get the following representation for the score function

$$\rho_F(F) = -E \left[\frac{F - E[F]}{\Theta} + \frac{\langle D\Theta, u \rangle}{\Theta^2} \middle| F \right].$$

As a consequence, we obtain

$$\begin{aligned} I(F \parallel N) &= E \left(\rho_F(F) + \frac{F - \mu}{\sigma^2} \right)^2 \\ &= E \left(E \left[-\frac{F - E[F]}{\Theta} - \frac{\langle D\Theta, u \rangle}{\Theta^2} + \frac{F - \mu}{\sigma^2} \middle| F \right] \right)^2 \\ &\leq E \left(-\frac{(F - E[F])(\sigma^2 - \Theta)}{\sigma^2 \Theta} - \frac{\langle D\Theta, u \rangle}{\Theta^2} + \frac{E[F] - \mu}{\sigma^2} \right)^2. \end{aligned}$$

Using the Cauchy–Schwarz and Hölder inequalities we deduce

$$\begin{aligned} I(F \parallel N) &\leq 3E \left[\frac{(F - E[F])^2 (\Theta - \sigma^2)^2}{\sigma^4 \Theta^2} \right] + 3E \left[\frac{\langle D\Theta, u \rangle^2}{\Theta^4} \right] + \frac{3}{\sigma^4} (E[F] - \mu)^2 \\ &\leq \frac{3}{\sigma^4} (E|F - E[F]|^8 E|\Theta|^{-8})^{1/4} (E|\Theta - \sigma^2|^4)^{1/2} + 3E \left[\frac{\|D\Theta\|^2 \|u\|^2}{|\Theta|^4} \right] + \frac{3}{\sigma^4} (E[F] - \mu)^2 \\ &\leq \frac{3}{\sigma^4} (E|F - E[F]|^8 E|\Theta|^{-8})^{1/4} (E|\Theta - \sigma^2|^4)^{1/2} + 3(E\|D\Theta\|^4)^{1/2} (E\|u\|^8 E|\Theta|^{-16})^{1/4} \\ &\quad + \frac{3}{\sigma^4} (E[F] - \mu)^2. \end{aligned} \quad (3.4)$$

We note that $E|\Theta| = \text{Var}(F)$. Then, by the inequality (3.19) in [11] we have

$$\begin{aligned} E|\Theta - \sigma^2|^4 &\leq 8|\text{Var}(F) - \sigma^2|^4 + 8E|\Theta - \text{Var}(F)|^4 \\ &\leq 8|\text{Var}(F) - \sigma^2|^4 + 72E\|D\Theta\|^4. \end{aligned} \quad (3.5)$$

On the other hand, by the Burkholder–Davis–Gundy inequality, there exists $C > 0$ such that

$$\begin{aligned} E|F - E[F]|^8 &= E \left[\left(\int_0^T E[D_t F | \mathcal{F}_t] dB_t \right)^8 \right] \\ &\leq CE \left[\left(\int_0^T |E[D_t F | \mathcal{F}_t]|^2 dt \right)^4 \right] \\ &= CE\|u\|^8. \end{aligned} \quad (3.6)$$

So we can get the desired estimate (3.3) by inserting (3.5) and (3.6) into (3.4). This completes the proof of the theorem. \square

Theorem 3.2. Let $F \in \mathbb{D}^{2,4}$ and N be a normal random variable with mean μ and variance σ^2 . Define

$$\Gamma := \langle DF, -DL^{-1}F \rangle_{L^2[0,T]}.$$

Assume that $\Gamma \neq 0$ a.s. Then, we have

$$I(F \parallel N) \leq c \left(\frac{1}{\sigma^4} (E[F] - \mu)^2 + A_F |\text{Var}(F) - \sigma^2|^2 + C_F \left(E\|D\Gamma\|_{L^2[0,T]}^4 \right)^{1/2} \right), \quad (3.7)$$

where c is an absolute constant and A_F, C_F are positive constants given by

$$A_F := \frac{1}{\sigma^4} \left(E \|DF\|_{L^2[0,T]}^8 E |\Gamma|^{-8} \right)^{1/4}, \quad C_F := A_F + \left(E \|DF\|_{L^2[0,T]}^8 E |\Gamma|^{-16} \right)^{1/4}.$$

Proof. Consider the stochastic process $u_t := -D_t L^{-1} F$, $t \in [0, T]$. Note that

$$F - E[F] = LL^{-1}F = -\delta DL^{-1}F = \delta(u). \quad (3.8)$$

Hence, with the exact proof of (3.4), we obtain the following.

$$\begin{aligned} I(F \| N) &\leq \frac{3}{\sigma^4} \left(E |F - E[F]|^8 E |\Gamma|^{-8} \right)^{1/4} \left(E |\Gamma - \sigma^2|^4 \right)^{1/2} \\ &\quad + 3 \left(E \|D\Gamma\|_{L^2[0,T]}^4 \right)^{1/2} \left(E \|u\|_{L^2[0,T]}^8 E |\Gamma|^{-16} \right)^{1/4} + \frac{3}{\sigma^4} (E[F] - \mu)^2. \end{aligned} \quad (3.9)$$

By (2.3) and (3.8) we have $\text{Var}(F) = E[F\delta(u)] = E[\Gamma]$. Hence, by the inequality (3.19) in [11], we obtain

$$\begin{aligned} E |\Gamma - \sigma^2|^4 &\leq 8 |\text{Var}(F) - \sigma^2|^4 + 8 E |\Gamma - \text{Var}(F)|^4 \\ &\leq 8 |\text{Var}(F) - \sigma^2|^4 + 72 E \left[\|D\Gamma\|_{L^2[0,T]}^4 \right]. \end{aligned} \quad (3.10)$$

We also have

$$E |F - E[F]|^8 \leq 7^4 E \|DF\|_{L^2[0,T]}^8. \quad (3.11)$$

On the other hand, by the inequality (3.17) in [11], we have

$$E \|u\|_{L^2[0,T]}^8 \leq E \|DF\|_{L^2[0,T]}^8. \quad (3.12)$$

Inserting (3.10), (3.11) and (3.12) into (3.9) yields the desired bound (3.7). So the proof of the theorem is complete. \square

Remark 3.1. (i) We have implicitly assumed that the bounds (3.3) and (3.7) both involve finite quantities, as otherwise there is nothing to prove.

(ii) In general, the random variables Θ and Γ are different from each other. However, they satisfy the relationship: $E[\Theta|F] = E[\Gamma|F]$ a.s. (see Proposition 2.3 in [2]).

Remark 3.2. (i) Theorem 3.1 is of interest for the readers who are not used to working with the Ornstein–Uhlenbeck operator. On the other hand, comparing with Theorem 3.1, the advantage of Theorem 3.2 lies in the fact that it can be extended to a more general setting: Suppose that \mathfrak{H} is a real separable Hilbert space with scalar product denoted by $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$. We denote by $W = \{W(h) : h \in \mathfrak{H}\}$ an isonormal Gaussian process defined in a complete probability space (Ω, \mathcal{F}, P) , \mathcal{F} is the σ -field generated by W . Now Malliavin derivative operator is with respect to W . Then, we have

$$I(F \| N) \leq c \left(\frac{1}{\sigma^4} (E[F] - \mu)^2 + A_F |\text{Var}(F) - \sigma^2|^2 + C_F (E \|D\Gamma\|_{\mathfrak{H}}^4)^{1/2} \right), \quad (3.13)$$

where $\Gamma := \langle DF, -DL^{-1}F \rangle_{\mathfrak{H}}$ and

$$A_F := \frac{1}{\sigma^4} \left(E \|DF\|_{\mathfrak{H}}^8 E |\Gamma|^{-8} \right)^{1/4}, \quad C_F := A_F + \left(E \|DF\|_{\mathfrak{H}}^8 E |\Gamma|^{-16} \right)^{1/4}.$$

The bound (3.13) thus provides us a potential tool to study the Fisher information distance for stochastic differential equations driven by fractional Brownian motion, or stochastic partial differential equations.

(ii) Let $F = I_q(f)$ be a multiple Wiener–Itô integral of order $q \geq 2$. We have $-DL^{-1}F = \frac{1}{q}DF$ and hence, $\Gamma = \frac{1}{q} \|DF\|_{\mathfrak{H}}^2$. We now use the moment estimates for $D^2F \otimes_1 DF$ provided in [10] and we obtain

$$(E \|D\Gamma\|_{\mathfrak{H}}^4)^{1/2} \leq c(E|\Gamma|^4 - 3),$$

where c is a positive constant. So our bound (3.13) recovers the fourth moment bound established in [10] for the multiple Wiener–Itô integrals.

4. Optimal Fisher information bounds for small noise SDEs

In this Section, we apply Theorem 3.1 to investigate the rate of Fisher information convergence for the solution to Eq. (1.1) and its additive functional (1.3). Although theory of stochastic differential equations with small noise is very rich, to the best of our knowledge, the results of this section are new. Given a function $h(t, x)$, we use the notations

$$h'(t, x) = \frac{\partial h(t, x)}{\partial x} \quad \text{and} \quad h''(t, x) = \frac{\partial^2 h(t, x)}{\partial x^2}.$$

We make the use of the following assumptions:

(A₁) The coefficients $b, \sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ are measurable functions having linear growth, that is, there exists $L > 0$ such that

$$|b(t, x)| + |\sigma(t, x)| \leq L(1 + |x|) \quad \forall x \in \mathbb{R}, t \in [0, T].$$

(A₂) $\sigma(t, x)$ and $b(t, x)$ are twice continuously differentiable in x with the derivatives bounded by L .

(A₃) $f(t, x)$ is twice differentiable in x , $f(t, x)$ together with its derivatives have polynomial growth and $\|f'\|_0 := \inf_{(t,x) \in [0,T] \times \mathbb{R}} f'(t, x) > 0$.

The main results of this section are stated in the following theorems.

Theorem 4.1. Let $(X_{\varepsilon,t})_{t \in [0,T]}$ and $(x_t)_{t \in [0,T]}$ be the solutions to the Eqs. (1.1) and (1.2), respectively. Define

$$\tilde{X}_{\varepsilon,t} := \frac{X_{\varepsilon,t} - x_t}{\varepsilon}, \quad \beta_t^2 := \int_0^t \sigma^2(r, x_r) \exp \left(2 \int_r^t b'(u, x_u) du \right) dr, \quad t \in [0, T].$$

Suppose the assumptions (A₁)-(A₂) and that, for some $p_0 > 16$,

$$E \left[\left(\int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right)^{-p_0} \right] < \infty \quad \forall \varepsilon \in (0, 1), t \in (0, T]. \quad (4.1)$$

Then, for all $\varepsilon \in (0, 1)$ and $t \in (0, T]$, we have

$$\begin{aligned} I(\tilde{X}_{\varepsilon,t} \parallel N_t) &\leq C \left(\frac{t^4}{\beta_t^4} + \frac{t^4}{\beta_t^4} \left(E \left| \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{2}{p_0}} \right. \\ &\quad \left. + t^4 \left(E \left| \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{4}{p_0}} \right) \varepsilon^2, \end{aligned} \quad (4.2)$$

where N_t denotes a normal random variable with mean zero and variance β_t^2 and C is a positive constant not depending on t and ε .

We can write down C in an explicit form ($\log C$ is a polynomial in T and L), but it is not our goal here. The non-degeneracy condition (4.1) make Theorem 4.1 not easy to use in practical applications. Hence, it is necessary to obtain the sufficient conditions which are easy to check. We have the following.

Corollary 4.1. Suppose the assumptions (A₁)-(A₂). We assume, in addition, that $\sigma(0, X_0) \neq 0$ and $|\sigma(t, x) - \sigma(s, x)| \leq L|t - s|^{\delta_1}$ for all $x \in \mathbb{R}$ and $s, t \in [0, T]$, where L, δ_1 are positive real numbers. Then, we have

$$I(\tilde{X}_{\varepsilon,t} \parallel N_t) \leq C \left(\frac{t^4}{\beta_t^4} + \frac{t^2}{\beta_t^4} + 1 \right) \varepsilon^2 \quad \forall \varepsilon \in (0, 1), t \in (0, T], \quad (4.3)$$

where C is a positive constant not depending on t and ε .

Proof. It is easy to see that $E|X_{\varepsilon,t} - X_0|^p \leq Ct^{\frac{p}{2}}$ for all $\varepsilon \in (0, 1), t \in [0, T]$, where C is a positive constant not depending on t and ε . On the other hand, the function σ satisfies

$$|\sigma(t, x) - \sigma(s, y)| \leq L(|t - s|^{\delta_1} + |x - y|) \quad \forall x, y \in \mathbb{R}, s, t \in [0, T].$$

Hence, we can apply Lemma A (given in Appendix) to $Y_t = X_{\varepsilon,t}$, $h(t, x) = \sigma(t, x)$ and $k(t, s) = 1$ and we obtain

$$E \left[\left(\int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right)^{-p_0} \right] \leq Ct^{-p_0}, \quad t \in (0, T]$$

for all $p_0 > 0$. So the bound (4.3) follows directly from (4.2). \square

Remark 4.1. Generally, the Berry-Esseen bounds for the rate of convergence are more informative in practice. As an application of Corollary 4.1, we obtain the following

$$\sup_{x \in \mathbb{R}} |P(\tilde{X}_{\varepsilon,t} \leq x) - P(N_t \leq x)| \leq \sqrt{I(\tilde{X}_{\varepsilon,t} \parallel N_t)} \leq C \left(\frac{t^2}{\beta_t^2} + \frac{t}{\beta_t^2} + 1 \right) \varepsilon \quad \forall \varepsilon \in (0, 1), t \in (0, T].$$

The bound (4.1) provides us the convergence rate of order $O(\varepsilon^2)$ as $\varepsilon \rightarrow 0$. One may wonder whether this rate is optimal. Interestingly, the answer is affirmative as in the next theorem.

Theorem 4.2. Suppose the assumptions (A₁)-(A₂). Then, for each $t \in (0, T]$, we have

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} I(\tilde{X}_{\varepsilon,t} \parallel N_t) \geq \frac{1}{4\beta_t^4} \left(E \left[E \left[\delta(V_t DU_t) \mid U_t \right] \right)^2 \right), \quad (4.4)$$

where β_t^2 is as in Theorem 4.1, $(U_t)_{t \in [0,T]}$ and $(V_t)_{t \in [0,T]}$ are stochastic processes defined by

$$U_t = \int_0^t b'(s, x_s) U_s ds + \int_0^t \sigma(s, x_s) dB_s, \quad 0 \leq t \leq T, \quad (4.5)$$

$$V_t = \int_0^t \left(\frac{1}{2} b''(s, x_s) U_s^2 + b'(s, x_s) V_s \right) ds + \int_0^t \sigma'(s, x_s) U_s dB_s, \quad 0 \leq t \leq T. \quad (4.6)$$

Remark 4.2. The stochastic differential equations (4.5) and (4.6) can be solved explicitly. We have

$$U_t = \int_0^t \sigma(s, x_s) e^{\int_s^t b'(u, x_u) du} dB_s, \quad 0 \leq t \leq T$$

and

$$V_t = \int_0^t \frac{1}{2} b''(s, x_s) U_s^2 e^{\int_s^t b'(u, x_u) du} ds + \int_0^t \sigma'(s, x_s) U_s e^{\int_s^t b'(u, x_u) du} dB_s, \quad 0 \leq t \leq T.$$

Furthermore, the random variables U_t and V_t are Malliavin differentiable and their derivatives are given by

$$D_r U_t = \sigma(r, x_r) e^{\int_r^t b'(u, x_u) du}, \quad 0 \leq r \leq t \leq T,$$

$$D_r V_t = \sigma'(r, x_r) U_r e^{\int_r^t b'(u, x_u) du} + \int_r^t b''(s, x_s) U_s D_r U_s e^{\int_s^t b'(u, x_u) du} ds + \int_r^t \sigma'(s, x_s) D_r U_s e^{\int_s^t b'(u, x_u) du} dB_s, \quad 0 \leq r \leq t \leq T.$$

It is also easy to see that $U_t, V_t \in \mathbb{D}^{k,p}$ for all $k \geq 1, p \geq 2$.

In the next theorem, for the additive functional of solutions, we also obtain the convergence rate of optimal order $O(\varepsilon^2)$.

Theorem 4.3. Consider the stochastic process $(Y_{\varepsilon,t})_{t \in [0,T]}$ defined by (1.3). Define $y_t := \int_0^t f(s, x_s) ds$ and

$$\tilde{Y}_{\varepsilon,t} := \frac{Y_{\varepsilon,t} - y_t}{\varepsilon}, \quad \gamma_t^2 := \int_0^t \left(\int_r^t f'(s, x_s) \sigma(r, x_r) e^{\int_r^s b'(u, x_u) du} ds \right)^2 dr, \quad 0 \leq t \leq T.$$

Suppose the assumptions (A_1) – (A_3) and that, for some $p_0 > 16$,

$$E \left[\left(\int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right)^{-p_0} \right] < \infty \quad \forall \varepsilon \in (0, 1), t \in (0, T]. \quad (4.7)$$

Then, for all $\varepsilon \in (0, 1)$ and $t \in (0, T]$, we have

$$I(\tilde{Y}_{\varepsilon,t} \parallel Z_t) \leq C \left(\frac{t^4}{\gamma_t^4} + \frac{t^{10}}{\gamma_t^4} \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{2}{p_0}} + t^{10} \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{4}{p_0}} \right) \varepsilon^2, \quad (4.8)$$

where Z_t denotes a normal random variable with mean zero and variance γ_t^2 and C is a positive constant not depending on t and ε .

Corollary 4.2. Suppose the assumptions (A_1) – (A_3) . We assume, in addition, that $\sigma(0, X_0) \neq 0$ and $|\sigma(t, x) - \sigma(s, x)| \leq L|t - s|^{\delta_1}$ for all $x \in \mathbb{R}$ and $s, t \in [0, T]$, where L, δ_1 are positive real numbers. Then, we have

$$I(\tilde{Y}_{\varepsilon,t} \parallel Z_t) \leq C \left(\frac{t^4}{\gamma_t^4} + \frac{1}{t^2} \right) \varepsilon^2 \quad \forall \varepsilon \in (0, 1), t \in (0, T], \quad (4.9)$$

where C is a positive constant not depending on t and ε .

Proof. For every $t_0 \in (0, T]$, we have

$$\int_0^{t_0} (t_0 - r)^2 \sigma^2(r, X_{\varepsilon,r}) dr = t_0^3 \int_0^1 (1 - r)^2 \sigma^2(t_0 r, X_{\varepsilon, t_0 r}) dr.$$

We consider the stochastic process $Y_t := X_{\varepsilon, t_0 t}, 0 \leq t \leq 1$. Then, $Y_0 = X_0$ is deterministic and $E|Y_t - X_0|^p = E|X_{\varepsilon, t_0 t} - X_0|^p \leq C t^{\frac{p}{2}}$ for all $\varepsilon \in (0, 1), t \in [0, 1]$, where C is a positive constant not depending on t_0, t and ε . In addition, the functions $h(t, x) = \sigma(t_0 t, x)$ and $k(t, s) = (t - s)^2$ satisfy

$$|h(t, x) - h(s, y)| \leq L(|t - s|^{\delta_1} + |x - y|) \quad \forall x, y \in \mathbb{R}, s, t \in [0, 1],$$

$$|k(t, s) - k(t, 0)| \leq L|t - s| \quad \forall s, t \in [0, 1]$$

for some positive constant L not depending on t_0 . Thus, for all $p_0 > 0$, we can use Lemma A to get

$$E \left[\left(\int_0^1 (1 - r)^2 \sigma^2(t_0 r, X_{\varepsilon, t_0 r}) dr \right)^{-p_0} \right] \leq C < \infty,$$

where C is a positive constant not depending on t_0 and ε . Consequently,

$$E \left[\left(\int_0^{t_0} (t_0 - r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right)^{-p_0} \right] \leq C t_0^{-3p_0} \quad \forall t_0 \in (0, T],$$

and hence, the bound (4.9) follows directly from (4.8). \square

Remark 4.3. (i) In the assumption (A_3) , the condition $\|f'\|_0 := \inf_{(t,x)} f'(t, x) > 0$ can be replaced by $\|f'\|_0 := \sup_{(t,x) \in [0,T] \times \mathbb{R}} f'(t, x) < 0$.

(ii) Consider the stochastic processes

$$\bar{U}_t = \int_0^t f'(s, x_s) U_s ds, \quad 0 \leq t \leq T,$$

$$\bar{V}_t = \frac{1}{2} \int_0^t (f''(s, x_s) U_s^2 + 2f'(s, x_s) V_s) ds, \quad 0 \leq t \leq T,$$

where $(U_t)_{t \in [0,T]}$ and $(V_t)_{t \in [0,T]}$ are as in Theorem 4.2. The reader can verify that

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} I(\tilde{X}_{\varepsilon,t} \parallel Z_t) \geq \frac{1}{4\gamma_t^4} \left(E \left[E \left[\delta \left(\bar{V}_t D \bar{U}_t \right) \mid \bar{U}_t \right] \right] \right)^2, \quad t \in (0, T].$$

The proof is similar to that of (4.4). So we omit it.

4.1. Estimates for Malliavin derivatives

Hereafter, we denote by C a generic constant which may vary at each appearance. Let us collect some fundamental results about the Malliavin differentiability of solutions to Eq. (1.1).

Proposition 4.1. Suppose the assumptions (A_1) and (A_2) . Then, Eq. (1.1) has a unique solution $(X_{\varepsilon,t})_{t \in [0,T]}$ satisfying, for each $p \geq 2$,

$$\sup_{t \in [0,T]} E |X_{\varepsilon,t}|^p \leq C, \quad \forall \varepsilon \in (0, 1), \quad (4.10)$$

where C is a positive constant not depending on ε . Moreover, for each $t \in [0, T]$, the random variable $X_{\varepsilon,t}$ is twice Malliavin differentiable and the derivatives satisfy the following linear equations, for all $0 \leq r, \theta \leq t \leq T$,

$$D_r X_{\varepsilon,t} = \varepsilon \sigma(r, X_{\varepsilon,r}) + \int_r^t b'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} ds + \varepsilon \int_r^t \sigma'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} dB_s \quad (4.11)$$

and

$$\begin{aligned} D_\theta D_r X_{\varepsilon,t} &= \varepsilon \sigma'(r, X_{\varepsilon,r}) D_\theta X_{\varepsilon,r} + \int_{r \vee \theta}^t [b''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} + b'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}] ds \\ &+ \varepsilon \sigma'(\theta, X_{\varepsilon,\theta}) D_r X_{\varepsilon,\theta} + \varepsilon \int_{r \vee \theta}^t [\sigma''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} + \sigma'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}] dB_s. \end{aligned} \quad (4.12)$$

Proof. See Theorems 2.2.1 and 2.2.2 in [12]. \square

Proposition 4.2. Suppose the assumptions (A_1) and (A_2) . Then, for each $p \geq 2$, we have

$$\sup_{0 \leq r \leq t \leq T} E |D_r X_{\varepsilon,t}|^p \leq C \varepsilon^p \quad \forall \varepsilon \in (0, 1), \quad (4.13)$$

and

$$\sup_{0 \leq r, \theta \leq t \leq T} E |D_\theta D_r X_{\varepsilon,t}|^p \leq C \varepsilon^{2p} \quad \forall \varepsilon \in (0, 1), \quad (4.14)$$

where C is a positive constant not depending on ε .

Proof. The proof is similar to that of Theorems 2.2.1 and 2.2.2 in [12]. For each $p \geq 2$, by the fundamental inequality $(|a_1| + |a_2| + |a_3|)^p \leq 3^{p-1}(|a_1|^p + |a_2|^p + |a_3|^p)$, we obtain from (4.11) that

$$\begin{aligned} |D_r X_{\varepsilon,t}|^p &\leq 3^{p-1} |\varepsilon \sigma(r, X_{\varepsilon,r})|^p + 3^{p-1} \left| \int_r^t b'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} ds \right|^p \\ &+ 3^{p-1} \left| \varepsilon \int_r^t \sigma'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} dB_s \right|^p \end{aligned}$$

for all $0 \leq r \leq t \leq T$. From the linear growth property of σ and the estimate (4.10), we have

$$\sup_{0 \leq r \leq T} E |\sigma(r, X_{\varepsilon,r})|^p \leq C \quad \forall \varepsilon \in (0, 1), \quad (4.15)$$

where C is a positive constant not depending on ε . Furthermore, by using the Hölder and Burkholder–Davis–Gundy inequalities, we deduce

$$\begin{aligned} E \left| \int_r^t b'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} ds \right|^p &\leq L^p (t-r)^{p-1} \int_r^t E |D_r X_{\varepsilon,s}|^p ds \\ &\leq L^p T^{p-1} \int_r^t E |D_r X_{\varepsilon,s}|^p ds, \quad 0 \leq r \leq t \leq T \end{aligned}$$

and, for some $C_p > 0$,

$$\begin{aligned} E \left| \varepsilon \int_r^t \sigma'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} dB_s \right|^p &\leq C_p \varepsilon^p E \left(\int_r^t |\sigma'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s}|^2 ds \right)^{\frac{p}{2}} \\ &\leq C_p (t-r)^{\frac{p}{2}-1} \varepsilon^p L^p \int_r^t E |D_r X_{\varepsilon,s}|^p ds \\ &\leq C_p T^{\frac{p}{2}-1} L^p \varepsilon^p \int_r^t E |D_r X_{\varepsilon,s}|^p ds, \quad 0 \leq r \leq t \leq T. \end{aligned}$$

We therefore obtain, for all $\varepsilon \in (0, 1)$,

$$E |D_r X_{\varepsilon,t}|^p \leq C \varepsilon^p + C \int_r^t E |D_r X_{\varepsilon,s}|^p ds, \quad 0 \leq r \leq t \leq T,$$

where C is a positive constant not depending on r, t and ε . Using Gronwall's lemma, we get

$$E |D_r X_{\varepsilon,t}|^p \leq C \varepsilon^p e^{C(t-r)} \leq C \varepsilon^p \quad \forall 0 \leq r \leq t \leq T.$$

This finishes the proof of (4.13). The proof of (4.14) can be done similarly. Indeed, we obtain from Eq. (4.12) that

$$\begin{aligned} E |D_\theta D_r X_{\varepsilon,t}|^p &\leq 4^{p-1} E |\varepsilon \sigma'(r, X_{\varepsilon,r}) D_\theta X_{\varepsilon,r}|^p + 4^{p-1} E |\varepsilon \sigma'(\theta, X_\theta) D_r X_{\varepsilon,\theta}|^p \\ &\quad + 4^{p-1} E \left| \int_{r \vee \theta}^t [b''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} + b'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}] ds \right|^p \\ &\quad + 4^{p-1} E \left| \varepsilon \int_{r \vee \theta}^t [\sigma''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} + \sigma'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}] dB_s \right|^p \end{aligned}$$

for all $0 \leq \theta, r \leq t \leq T$. By using the estimate (4.13) and the boundedness of $b', b'', \sigma', \sigma''$ we obtain

$$E |\varepsilon \sigma'(r, X_{\varepsilon,r}) D_\theta X_{\varepsilon,r}|^p + E |\varepsilon \sigma'(\theta, X_\theta) D_r X_{\varepsilon,\theta}|^p \leq C \varepsilon^{2p},$$

$$\begin{aligned} E \left| \int_{r \vee \theta}^t [b''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} + b'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}] ds \right|^p \\ \leq C \varepsilon^{2p} + C \int_{r \vee \theta}^t E |D_\theta D_r X_{\varepsilon,s}|^p ds, \end{aligned}$$

and

$$\begin{aligned} E \left| \varepsilon \int_{r \vee \theta}^t [\sigma''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} + \sigma'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}] dB_s \right|^p \\ = \varepsilon^p E \left| \int_{r \vee \theta}^t [\sigma''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} + \sigma'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}]^2 ds \right|^{\frac{p}{2}} \\ \leq C \varepsilon^{3p} + C \varepsilon^p \int_{r \vee \theta}^t E |D_\theta D_r X_{\varepsilon,s}|^p ds. \end{aligned}$$

As a consequence, for all $\varepsilon \in (0, 1)$,

$$E |D_\theta D_r X_{\varepsilon,t}|^p \leq C \varepsilon^{2p} + C \int_{r \vee \theta}^t E |D_\theta D_r X_{\varepsilon,s}|^p ds, \quad 0 \leq \theta, r \leq t \leq T$$

and we obtain (4.14) by using Gronwall's lemma again.

The proof the proposition is complete. \square

4.2. Proof of Theorem 4.1

The proof of Theorem 4.1 will be given at the end of this subsection. In order to be able to apply Theorem 3.1, we need the following technical results.

Lemma 4.1. *Suppose the assumptions (A_1) – (A_2) . Then, for each $p \geq 2$, we have*

$$E |X_{\varepsilon,t} - x_t|^p \leq C t^{\frac{p}{2}} \varepsilon^p \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.16)$$

where C is a positive constant not depending on t and ε . Moreover, if $h(t, x)$ is a differentiable function in x and its derivative has polynomial growth, we also have

$$E|h(t, X_{\varepsilon,t}) - h(t, x_t)|^p \leq Ct^{\frac{p}{2}} \varepsilon^p \quad \forall \varepsilon \in (0, 1), t \in [0, T]. \quad (4.17)$$

Proof. We have

$$X_{\varepsilon,t} - x_t = \int_0^t (b(s, X_{\varepsilon,s}) - b(s, x_s))ds + \varepsilon \int_0^t \sigma(s, X_{\varepsilon,s})dB_s, \quad t \in [0, T].$$

Hence, by Hölder and Burkholder–Davis–Gundy inequalities, we deduce

$$\begin{aligned} E|X_{\varepsilon,t} - x_t|^p &= E \left| \int_0^t (b(s, X_{\varepsilon,s}) - b(s, x_s))ds + \varepsilon \int_0^t \sigma(s, X_{\varepsilon,s})dB_s \right|^p \\ &\leq 2^{p-1} E \left| \int_0^t (b(s, X_{\varepsilon,s}) - b(s, x_s))ds \right|^p + 2^{p-1} \varepsilon^p E \left| \int_0^t \sigma(s, X_{\varepsilon,s})dB_s \right|^p \\ &\leq Ct^{p-1} \int_0^t E|b(s, X_{\varepsilon,s}) - b(s, x_s)|^p ds + Ct^{\frac{p}{2}-1} \varepsilon^p \int_0^t E|\sigma(s, X_{\varepsilon,s})|^p ds, \quad t \in [0, T], \end{aligned}$$

where C is a positive constant not depending on t and ε . So by the boundedness of b' and the estimate (4.15) we obtain

$$E|X_{\varepsilon,t} - x_t|^p \leq C \int_0^t E|X_{\varepsilon,s} - x_s|^p ds + Ct^{\frac{p}{2}} \varepsilon^p, \quad t \in [0, T],$$

which, together with Gronwall's lemma, yields

$$E|X_{\varepsilon,t} - x_t|^p \leq Ct^{\frac{p}{2}} \varepsilon^p e^{Ct} \leq Ct^{\frac{p}{2}} \varepsilon^p, \quad t \in [0, T].$$

It remains to prove (4.17). For each $t \in [0, T]$, using the Taylor's expansion, we have

$$h(t, X_{\varepsilon,t}) - h(t, x_t) = h'(t, x_t + \eta_t(X_{\varepsilon,t} - x_t))(X_{\varepsilon,t} - x_t),$$

where η_t is a random variable lying between 0 and 1. By the polynomial growth property of h' and the estimate (4.10), we have $\sup_{t \in [0, T]} E|h'(t, x_t + \eta_t(X_{\varepsilon,t} - x_t))|^p \leq C$ for all $\varepsilon \in (0, 1)$, where C is a positive constant not depending on ε . We now use the Cauchy–Schwarz inequality and the estimate (4.16) to get

$$E|h(t, X_{\varepsilon,t}) - h(t, x_t)|^p \leq \sqrt{E|h'(t, x_t + \eta_t(X_{\varepsilon,t} - x_t))|^{2p} E|X_{\varepsilon,t} - x_t|^{2p}} \leq Ct^{\frac{p}{2}} \varepsilon^p.$$

This finishes the proof of the proposition. \square

Proposition 4.3. Suppose the assumptions (A_1) – (A_2) . Let $(\tilde{X}_{\varepsilon,t})_{t \in [0, T]}$ be as in Theorem 4.1. Then, we have

$$|E[\tilde{X}_{\varepsilon,t}]| \leq Ct^2 \varepsilon \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.18)$$

$$|\text{Var}(\tilde{X}_{\varepsilon,t}) - \beta_t^2| \leq Ct^{3/2} \varepsilon \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.19)$$

where C is a positive constant not depending on t and ε .

Proof. We first verify the estimate (4.18). By using the Taylor's expansion, we obtain

$$\begin{aligned} \tilde{X}_{\varepsilon,t} &= \frac{1}{\varepsilon} \int_0^t (b(s, X_{\varepsilon,s}) - b(s, x_s))ds + \int_0^t \sigma(s, X_{\varepsilon,s})dB_s \\ &= \int_0^t b'(s, x_s) \tilde{X}_{\varepsilon,s} ds + \frac{1}{2\varepsilon} \int_0^t b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))(X_{\varepsilon,s} - x_s)^2 ds + \int_0^t \sigma(s, X_{\varepsilon,s})dB_s, \end{aligned} \quad (4.20)$$

where, for each $0 \leq s \leq t$, θ_s is a random variable lying between 0 and 1. Taking the expectation of $\tilde{X}_{\varepsilon,t}$ gives us

$$E[\tilde{X}_{\varepsilon,t}] = \int_0^t b'(s, x_s) E[\tilde{X}_{\varepsilon,s}] ds + \frac{1}{2\varepsilon} \int_0^t E[b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))(X_{\varepsilon,s} - x_s)^2] ds$$

This is a linear differential equation and its solution is given by

$$E[\tilde{X}_{\varepsilon,t}] = \frac{1}{2\varepsilon} \int_0^t e^{\int_s^t b'(u, x_u) du} E[b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))(X_{\varepsilon,s} - x_s)^2] ds.$$

Consequently, we can use the boundedness of the derivatives b', b'' and the estimate (4.16) to get

$$\begin{aligned} |E[\tilde{X}_{\varepsilon,t}]| &\leq \frac{Le^{LT}}{2\varepsilon} \int_0^t E|X_{\varepsilon,s} - x_s|^2 ds \\ &\leq Ct^2 \varepsilon, \quad t \in [0, T]. \end{aligned}$$

So the estimate (4.18) holds. It remains to prove the estimate (4.19). Solving Eq. (4.11) we obtain

$$\begin{aligned} D_r X_{\varepsilon,t} &= \varepsilon \sigma(r, X_{\varepsilon,r}) \exp \left(\int_r^t \left(b'(u, X_{\varepsilon,u}) - \frac{1}{2} \varepsilon^2 \sigma'^2(u, X_{\varepsilon,u}) \right) du + \varepsilon \int_r^t \sigma'(u, X_{\varepsilon,u}) dB_u \right) \\ &= \varepsilon \sigma(r, X_{\varepsilon,r}) \exp \left(\int_r^t b'(u, X_{\varepsilon,u}) du \right) Z_{r,t}, \quad 0 \leq r \leq t \leq T. \end{aligned} \quad (4.21)$$

where $Z_{r,t}$ is given by

$$Z_{r,t} := \exp \left(\varepsilon \int_r^t \sigma'(u, X_{\varepsilon,u}) dB_u - \frac{1}{2} \varepsilon^2 \int_r^t \sigma'^2(u, X_{\varepsilon,u}) du \right), \quad 0 \leq r \leq t \leq T. \quad (4.22)$$

Note that, by the Itô differential formula, $Z_{r,t}$ satisfies

$$Z_{r,t} = 1 + \int_r^t \varepsilon \sigma'(s, X_{\varepsilon,s}) Z_{r,s} dB_s, \quad 0 \leq r \leq t \leq T. \quad (4.23)$$

So $E[Z_{r,t} | \mathcal{F}_r] = 1$. Furthermore, for each $p \geq 2$, it is easy to see that

$$\sup_{0 \leq r \leq t \leq T} E|Z_{r,t}|^p \leq C \quad (4.24)$$

for some $C > 0$ not depending on $\varepsilon \in (0, 1)$. By the Clark–Ocone formula, the Itô isometry and (4.21), we have

$$\begin{aligned} \text{Var}(\tilde{X}_{\varepsilon,t}) &= E \left[\int_0^t (E[D_r \tilde{X}_{\varepsilon,t} | \mathcal{F}_r])^2 dr \right] = \frac{1}{\varepsilon^2} E \left[\int_0^t (E[D_r X_{\varepsilon,t} | \mathcal{F}_r])^2 dr \right] \\ &= E \left[\int_0^t \sigma^2(r, X_{\varepsilon,r}) \left(E \left[e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{r,t} \middle| \mathcal{F}_r \right] \right)^2 dr \right], \end{aligned}$$

and hence,

$$\begin{aligned} \text{Var}(\tilde{X}_{\varepsilon,t}) - \beta_t^2 &= E \left[\int_0^t \sigma^2(r, X_{\varepsilon,r}) \left(E \left[e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{r,t} \middle| \mathcal{F}_r \right] \right)^2 dr \right] - \int_0^t \sigma^2(r, x_r) e^{2 \int_r^t b'(u, x_u) du} dr \\ &= E \left[\int_0^t (\sigma^2(r, X_{\varepsilon,r}) - \sigma^2(r, x_r)) \left(E \left[e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{r,t} \middle| \mathcal{F}_r \right] \right)^2 dr \right] \\ &\quad + E \left[\int_0^t \sigma^2(r, x_r) \left(\left(E \left[e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{r,t} \middle| \mathcal{F}_r \right] \right)^2 - e^{2 \int_r^t b'(u, x_u) du} \right) dr \right]. \end{aligned} \quad (4.25)$$

Consequently, since b' is bounded by L , $E[Z_{r,t} | \mathcal{F}_r] = 1$ and $\sup_{0 \leq r \leq T} \sigma^2(r, x_r) < \infty$, we can infer from (4.25) that

$$\begin{aligned} |\text{Var}(\tilde{X}_{\varepsilon,t}) - \beta_t^2| &\leq C \int_0^t E|\sigma^2(r, X_{\varepsilon,r}) - \sigma^2(r, x_r)| dr + C \int_0^t E \left| E \left[e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{r,t} \middle| \mathcal{F}_r \right] - e^{\int_r^t b'(u, x_u) du} \right| dr \\ &\leq C \int_0^t E|\sigma^2(r, X_{\varepsilon,r}) - \sigma^2(r, x_r)| dr + C \int_0^t E \left| \left(e^{\int_r^t b'(u, X_{\varepsilon,u}) du} - e^{\int_r^t b'(u, x_u) du} \right) Z_{r,t} \right| dr \\ &\leq C \int_0^t E|\sigma^2(r, X_{\varepsilon,r}) - \sigma^2(r, x_r)| dr + C \int_0^t \int_r^t E|b'(u, X_{\varepsilon,u}) - b'(u, x_u)| Z_{r,t} du dr \\ &\leq C \int_0^t E|\sigma^2(r, X_{\varepsilon,r}) - \sigma^2(r, x_r)| dr + C \int_0^t \int_r^t \sqrt{E|b'(u, X_{\varepsilon,u}) - b'(u, x_u)|^2 E|Z_{r,t}|^2} du dr, \end{aligned} \quad (4.26)$$

where C is a positive constant not depending on t and ε . Hence, in view of the estimates (4.17) and (4.24), we get

$$|\text{Var}(\tilde{X}_{\varepsilon,t}) - \beta_t^2| \leq C t^{\frac{3}{2}} \varepsilon + C t^{\frac{5}{2}} \varepsilon \leq C t^{\frac{3}{2}} \varepsilon. \quad (4.27)$$

So the estimate (4.19) is proved. This completes the proof of the proposition. \square

Proposition 4.4. Let $(\tilde{X}_{\varepsilon,t})_{t \in [0,T]}$ be as in Theorem 4.1. Define

$$\Theta_{\tilde{X}_{\varepsilon,t}} := \int_0^t D_r \tilde{X}_{\varepsilon,t} E[D_r \tilde{X}_{\varepsilon,t} | \mathcal{F}_r] dr, \quad t \in [0, T].$$

Then, under the assumption of Theorem 4.1, we have

$$E|\Theta_{\tilde{X}_{\varepsilon,t}}|^{-p} \leq C \left(E \left| \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{p}{p_0}} \quad \forall \varepsilon \in (0, 1), t \in (0, T], \quad (4.28)$$

where $0 < p < p_0$ and $C > 0$ is a positive constant not depending on t and ε .

Proof. Since $D_r \tilde{X}_{\varepsilon,t} = \frac{D_r X_{\varepsilon,t}}{\varepsilon}$, it follows from (4.21) that

$$\begin{aligned}\Theta_{\tilde{X}_{\varepsilon,t}} &= \frac{1}{\varepsilon^2} \int_0^t D_r X_{\varepsilon,t} E[D_r X_{\varepsilon,t} | \mathcal{F}_r] dr \\ &= \int_0^t \sigma^2(r, X_{\varepsilon,r}) e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{t,r} E \left[e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{t,r} \middle| \mathcal{F}_r \right] dr,\end{aligned}$$

where $Z_{t,r}$ is defined by (4.22). Since b', σ' are bounded by L , we deduce $e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{t,r} \geq e^{-\frac{3LT}{2}} e^{\varepsilon \int_r^t \sigma'(u, X_{\varepsilon,u}) dB_u}$ and $E \left[e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{t,r} \middle| \mathcal{F}_r \right] \geq e^{-LT} E \left[Z_{t,r} \middle| \mathcal{F}_r \right] = e^{-LT}$. Then, we obtain

$$\begin{aligned}\Theta_{\tilde{X}_{\varepsilon,t}} &\geq e^{-\frac{5LT}{2}} \int_0^t \sigma^2(r, X_{\varepsilon,r}) e^{\varepsilon \int_r^t \sigma'(u, X_{\varepsilon,u}) dB_u} dr \\ &\geq e^{-\frac{5LT}{2}} e^{M_t - \max_{0 \leq t \leq T} M_t} \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \\ &\geq e^{-\frac{5LT}{2}} e^{-2 \max_{0 \leq t \leq T} M_t} \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr,\end{aligned}$$

where $M_t := \varepsilon \int_0^t \sigma'(u, X_{\varepsilon,u}) dB_u, 0 \leq t \leq T$. We observe that M_t is a martingale with the bounded quadratic variation. Indeed, $\langle M \rangle_t = \varepsilon^2 \int_0^t \sigma'^2(s, X_{\varepsilon,s}) ds \leq L^2 T$ for all $\varepsilon \in (0, 1)$ and $t \in [0, T]$. By Dubin and Schwarz's theorem, there exists a one dimensional Brownian motion $(m_t)_{t \geq 0}$ such that $M_t = m_{\langle M \rangle_t}$. Then we arrive at the following

$$|\Theta_{\tilde{X}_{\varepsilon,t}}| \geq e^{-\frac{5LT}{2}} e^{-2 \max_{0 \leq t \leq T} M_t} \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \quad \forall \varepsilon \in (0, 1), t \in (0, T].$$

Note that, by Fernique's theorem, we always have $E \left[e^{4q \max_{0 \leq t \leq T} M_t} \right] < \infty$ for all $q > 0$. Hence, for $0 < p < p_0$, we use Hölder's inequality to deduce the following

$$\begin{aligned}E|\Theta_{\tilde{X}_{\varepsilon,t}}|^{-p} &\leq e^{\frac{5pLT}{2}} E \left[e^{2p \max_{0 \leq t \leq T} M_t} \left(\int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right)^{-p} \right] \\ &\leq C \left(E \left[\int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right]^{-p_0} \right)^{\frac{p}{p_0}} \quad \forall \varepsilon \in (0, 1), t \in (0, T],\end{aligned}$$

where $C > 0$ is a positive constant not depending on t and ε . The proof of the proposition is complete. \square

Proposition 4.5. Let $(\Theta_{\tilde{X}_{\varepsilon,t}})_{t \in [0, T]}$ be as in Proposition 4.4. Suppose the assumptions (A_1) and (A_2) . Then, we have

$$E \| D\Theta_{\tilde{X}_{\varepsilon,t}} \|_{L^2[0, T]}^4 \leq C \varepsilon^4 t^6 \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.29)$$

where C is a positive constant not depending on t and ε .

Proof. The Malliavin derivative of $\Theta_{\tilde{X}_{\varepsilon,t}}$ can be computed as follows

$$D_\theta \Theta_{\tilde{X}_{\varepsilon,t}} = \frac{1}{\varepsilon^2} \int_0^t (D_\theta D_r X_{\varepsilon,t} E[D_r X_{\varepsilon,t} | \mathcal{F}_r] + D_r X_{\varepsilon,t} E[D_\theta D_r X_{\varepsilon,t} | \mathcal{F}_r]) dr. \quad (4.30)$$

Hence, we get

$$\| D\Theta_{\tilde{X}_{\varepsilon,t}} \|_{L^2[0, T]}^4 = \frac{1}{\varepsilon^8} \left(\int_0^t \left(\int_0^t (D_\theta D_r X_{\varepsilon,t} E[D_r X_{\varepsilon,t} | \mathcal{F}_r] + D_r X_{\varepsilon,t} E[D_\theta D_r X_{\varepsilon,t} | \mathcal{F}_r]) dr \right)^2 d\theta \right)^2.$$

Using the Hölder inequality gives us

$$\begin{aligned}E \| D\Theta_{\tilde{X}_{\varepsilon,t}} \|_{L^2[0, T]}^4 &\leq \frac{t}{\varepsilon^8} \int_0^t E \left| \int_0^t (D_\theta D_r X_{\varepsilon,t} E[D_r X_{\varepsilon,t} | \mathcal{F}_r] + D_r X_{\varepsilon,t} E[D_\theta D_r X_{\varepsilon,t} | \mathcal{F}_r]) dr \right|^4 d\theta \\ &\leq \frac{t^4}{\varepsilon^8} \int_0^t \int_0^t E \left| D_\theta D_r X_{\varepsilon,t} E[D_r X_{\varepsilon,t} | \mathcal{F}_r] + D_r X_{\varepsilon,t} E[D_\theta D_r X_{\varepsilon,t} | \mathcal{F}_r] \right|^4 dr d\theta \\ &\leq \frac{8t^4}{\varepsilon^8} \int_0^t \int_0^t (E |D_\theta D_r X_{\varepsilon,t} E[D_r X_{\varepsilon,t} | \mathcal{F}_r]|^4 + E |D_r X_{\varepsilon,t} E[D_\theta D_r X_{\varepsilon,t} | \mathcal{F}_r]|^4) dr d\theta \\ &\leq \frac{16t^4}{\varepsilon^8} \int_0^t \int_0^t \sqrt{E |D_\theta D_r X_{\varepsilon,t}|^8 E |D_r X_{\varepsilon,t}|^8} dr d\theta.\end{aligned}$$

Hence, recalling the estimates (4.13) and (4.14), we get

$$E \| D\Theta_{\tilde{X}_{\varepsilon,t}} \|_{L^2[0, T]}^4 \leq C \varepsilon^4 t^6 \quad \forall \varepsilon \in (0, 1), t \in [0, T],$$

where $C > 0$ is a finite constant not depending on t and ε . The proof of the proposition is complete. \square

Proof of Theorem 4.1. Put $\tilde{u}_r = E[D_r \tilde{X}_{\varepsilon,t} | \mathcal{F}_r]$ for $0 \leq r \leq t \leq T$. Then, we have

$$\|\tilde{u}\|_{L^2[0,T]}^8 = \left(\int_0^t |E[D_r \tilde{X}_{\varepsilon,t} | \mathcal{F}_r]|^2 dr \right)^4 = \frac{1}{\varepsilon^8} \left(\int_0^t |E[D_r X_{\varepsilon,t} | \mathcal{F}_r]|^2 dr \right)^4.$$

Using Hölder's inequality and the estimate (4.13) we have

$$\begin{aligned} E\|\tilde{u}\|_{L^2[0,T]}^8 &\leq \frac{t^3}{\varepsilon^8} \int_0^t E|E[D_r X_{\varepsilon,t} | \mathcal{F}_r]|^8 dr \\ &\leq \frac{t^3}{\varepsilon^8} \int_0^t E|D_r X_{\varepsilon,t}|^8 dr \\ &\leq Ct^4 \quad \forall \varepsilon \in (0, 1), t \in [0, T]. \end{aligned} \quad (4.31)$$

By applying Theorem 3.1 to $F = \tilde{X}_{\varepsilon,t}$ and $N = N_t$ we get

$$I(\tilde{X}_{\varepsilon,t} \parallel N_t) \leq c \left(\frac{1}{\beta_t^4} (E[\tilde{X}_{\varepsilon,t}])^2 + A_{\tilde{X}_{\varepsilon,t}} |\text{Var}(\tilde{X}_{\varepsilon,t}) - \beta_t^2|^2 + C_{\tilde{X}_{\varepsilon,t}} (E\|D\Theta_{\tilde{X}_{\varepsilon,t}}\|^4)^{1/2} \right),$$

where c is an absolute constant and

$$A_{\tilde{X}_{\varepsilon,t}} := \frac{1}{\beta_t^4} \left(E\|\tilde{u}\|_{L^2[0,T]}^8 E|\Theta_{\tilde{X}_{\varepsilon,t}}|^{-8} \right)^{1/4}, \quad C_{\tilde{X}_{\varepsilon,t}} := A_{\tilde{X}_{\varepsilon,t}} + \left(E\|\tilde{u}\|_{L^2[0,T]}^8 E|\Theta_{\tilde{X}_{\varepsilon,t}}|^{-16} \right)^{1/4}.$$

Using the estimates (4.28) and (4.31) we obtain

$$\begin{aligned} A_{\tilde{X}_{\varepsilon,t}} &\leq \frac{Ct}{\beta_t^4} \left(E \left| \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{2}{p_0}}, \\ C_{\tilde{X}_{\varepsilon,t}} &\leq \frac{Ct}{\beta_t^4} \left(E \left| \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{2}{p_0}} + Ct \left(E \left| \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{4}{p_0}}. \end{aligned}$$

Furthermore, thanks to Propositions 4.3 and 4.5, we have

$$\begin{aligned} |E[\tilde{X}_{\varepsilon,t}]|^2 &\leq Ct^4 \varepsilon^2, \\ |\text{Var}(\tilde{X}_{\varepsilon,t}) - \beta_t^2|^2 &\leq Ct^3 \varepsilon^2, \\ (E\|D\Theta_{\tilde{X}_{\varepsilon,t}}\|^4)^{1/2} &\leq Ct^3 \varepsilon^2. \end{aligned}$$

Combining the above computations gives us

$$\begin{aligned} I(\tilde{X}_{\varepsilon,t} \parallel N_t) &\leq C \left(\frac{t^4}{\beta_t^4} + \frac{t^4}{\beta_t^4} \left(E \left| \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{2}{p_0}} \right. \\ &\quad \left. + t^4 \left(E \left| \int_0^t \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{4}{p_0}} \right) \varepsilon^2. \end{aligned}$$

So the proof of Theorem 4.1 is complete. \square

4.3. Proof of Theorem 4.2

Our idea is to use the following relation between the Fisher information and total variation distances, see [13,15]:

$$\sqrt{I(\tilde{X}_{\varepsilon,t} \parallel N_t)} \geq d_{TV}(\tilde{X}_{\varepsilon,t} \parallel N_t) := \frac{1}{2} \sup_g |E[g(\tilde{X}_{\varepsilon,t})] - E[g(N_t)]|, \quad (4.32)$$

where the supremum is running over all measurable functions g bounded by 1. Thus our main task is to find a lower bound for $\lim_{\varepsilon \rightarrow 0} d_{TV}(\tilde{X}_{\varepsilon,t} \parallel N_t)$.

Proposition 4.6. Suppose the assumptions (A_1) – (A_2) . Then, for each $p \geq 2$, we have

$$\sup_{t \in [0,T]} E|\tilde{X}_{\varepsilon,t} - U_t|^p \leq C\varepsilon^p \quad \forall \varepsilon \in (0, 1), \quad (4.33)$$

$$\sup_{t \in [0,T]} \int_0^t E|D_r \tilde{X}_{\varepsilon,t} - D_r U_t|^2 dr \leq C\varepsilon^2 \quad \forall \varepsilon \in (0, 1), \quad (4.34)$$

where C is a positive constant not depending on ε .

Proof. Recalling (4.5) and (4.20) we have

$$\tilde{X}_{\varepsilon,t} - U_t = \int_0^t b'(s, x_s)(\tilde{X}_{\varepsilon,s} - U_s) ds + \frac{1}{2\varepsilon} \int_0^t b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))(X_{\varepsilon,s} - x_s)^2 ds$$

$$+ \int_0^t (\sigma(s, X_{\varepsilon,s}) - \sigma(s, x_s)) dB_s, \quad t \in [0, T]. \quad (4.35)$$

Hence, under the assumption (A_2) , we can use Hölder and Burkholder–Davis–Gundy inequalities to get

$$\begin{aligned} E|\tilde{X}_{\varepsilon,t} - U_t|^p &\leq C \int_0^t E|\tilde{X}_{\varepsilon,s} - U_s|^p ds + \frac{C}{\varepsilon^p} \int_0^t E|X_{\varepsilon,s} - x_s|^{2p} ds \\ &\quad + C \int_0^t E|X_{\varepsilon,s} - x_s|^p ds, \quad t \in [0, T] \end{aligned}$$

for all $\varepsilon \in (0, 1)$ and for some $C > 0$. Consequently, it follows from the estimate (4.16) that

$$E|\tilde{X}_{\varepsilon,t} - U_t|^p \leq C\varepsilon^p + C \int_0^t E|\tilde{X}_{\varepsilon,s} - U_s|^p ds, \quad t \in [0, T],$$

where C is a positive constant not depending on t and ε . So we obtain the estimate (4.33) by using Gronwall's lemma. Indeed, we have

$$E|\tilde{X}_{\varepsilon,t} - U_t|^p \leq C\varepsilon^p e^{Ct} \leq C\varepsilon^p \quad \forall \varepsilon \in (0, 1), t \in [0, T].$$

Let us now prove (4.34). We have, for $0 \leq r \leq T$,

$$\begin{aligned} D_r \tilde{X}_{\varepsilon,t} - D_r U_t &= \frac{1}{\varepsilon} D_r X_{\varepsilon,t} - D_r U_t \\ &= \sigma(r, X_{\varepsilon,r}) e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{r,t} - \sigma(r, x_r) e^{\int_r^t b'(u, x_u) du}, \end{aligned}$$

where $Z_{r,t}$ is defined by (4.22). We deduce

$$\begin{aligned} |D_r \tilde{X}_{\varepsilon,t} - D_r U_t| &\leq |\sigma(r, X_{\varepsilon,r}) - \sigma(r, x_r)| e^{\int_r^t b'(u, X_{\varepsilon,u}) du} Z_{r,t} \\ &\quad + \sigma(r, x_r) \left| e^{\int_r^t b'(u, X_{\varepsilon,u}) du} - e^{\int_r^t b'(u, x_u) du} \right| Z_{r,t} + \sigma(r, x_r) e^{\int_r^t b'(u, x_u) du} |Z_{r,t} - 1| \end{aligned}$$

and

$$\begin{aligned} \int_0^t E|D_r \tilde{X}_{\varepsilon,t} - D_r U_t|^2 dr &\leq 3 \int_0^t E \left[|\sigma(r, X_{\varepsilon,r}) - \sigma(r, x_r)|^2 e^{2 \int_r^t b'(u, X_{\varepsilon,u}) du} Z_{r,t}^2 \right] dr \\ &\quad + 3 \int_0^t \sigma^2(r, x_r) E \left[\left| e^{\int_r^t b'(u, X_{\varepsilon,u}) du} - e^{\int_r^t b'(u, x_u) du} \right|^2 Z_{r,t}^2 \right] dr \\ &\quad + 3 \int_0^t \sigma^2(r, x_r) e^{2 \int_r^t b'(u, x_u) du} E|Z_{r,t} - 1|^2 dr. \end{aligned}$$

Hence, using the same arguments as in the proof (4.26) and (4.27), we get

$$\int_0^t E|D_r \tilde{X}_{\varepsilon,t} - D_r U_t|^2 dr \leq C\varepsilon^2 + 3 \int_0^t \sigma^2(r, x_r) e^{2 \int_r^t b'(u, x_u) du} E|Z_{r,t} - 1|^2 dr.$$

On the other hand, it follows from Eq. (4.23) that

$$\sup_{0 \leq r \leq t \leq T} E|Z_{r,t} - 1|^2 \leq C\varepsilon^2 \quad \forall \varepsilon \in (0, 1).$$

So we obtain

$$\int_0^t E|D_r \tilde{X}_{\varepsilon,t} - D_r U_t|^2 dr \leq C\varepsilon^2 \quad \forall \varepsilon \in (0, 1), t \in [0, T],$$

where C is a positive constant not depending on t and ε . This finishes the proof of the proposition. \square

Proposition 4.7. Suppose the assumptions (A_1) – (A_2) . Then, for each $p \geq 2$, we have

$$\lim_{\varepsilon \rightarrow 0} E \left| \frac{\tilde{X}_{\varepsilon,t} - U_t}{\varepsilon} - V_t \right|^p = 0, \quad t \in [0, T]. \quad (4.36)$$

Proof. We rewrite (4.35) as follows

$$\begin{aligned} \tilde{X}_{\varepsilon,t} - U_t &= \int_0^t b'(s, x_s)(\tilde{X}_{\varepsilon,s} - U_s) ds + \frac{1}{2\varepsilon} \int_0^t b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))(X_{\varepsilon,s} - x_s)^2 ds \\ &\quad + \int_0^t \sigma'(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))(X_{\varepsilon,s} - x_s) dB_s, \quad t \in [0, T]. \end{aligned}$$

where, for each $0 \leq s \leq t$, θ_s, η_s are random variables lying between 0 and 1. Hence, by the definition (4.6) of V_t , we obtain

$$\frac{\tilde{X}_{\varepsilon,t} - U_t}{\varepsilon} - V_t = \int_0^t b'(s, x_s) \left(\frac{\tilde{X}_{\varepsilon,s} - U_s}{\varepsilon} - V_s \right) ds$$

$$\begin{aligned}
& + \frac{1}{2} \int_0^t (b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))\tilde{X}_{\varepsilon,s}^2 - b''(s, x_s)U_s^2) ds \\
& + \int_0^t (\sigma'(s, x_s + \eta_s(X_{\varepsilon,s} - x_s))\tilde{X}_{\varepsilon,s} - \sigma'(s, x_s)U_s) dB_s, \quad t \in [0, T].
\end{aligned}$$

As a consequence, we deduce

$$E \left| \frac{\tilde{X}_{\varepsilon,t} - U_t}{\varepsilon} - V_t \right|^p \leq C \int_0^t E \left| \frac{\tilde{X}_{\varepsilon,s} - U_s}{\varepsilon} - V_s \right|^p ds + CK_\varepsilon, \quad t \in [0, T]$$

for some $C > 0$ and for all $\varepsilon \in (0, 1)$, where K_ε is given by

$$\begin{aligned}
K_\varepsilon := & \int_0^T E \left| b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))\tilde{X}_{\varepsilon,s}^2 - b''(s, x_s)U_s^2 \right|^p ds \\
& + \int_0^T E \left| \sigma'(s, x_s + \eta_s(X_{\varepsilon,s} - x_s))\tilde{X}_{\varepsilon,s} - \sigma'(s, x_s)U_s \right|^p ds.
\end{aligned}$$

An application of Gronwall's lemma gives us

$$E \left| \frac{\tilde{X}_{\varepsilon,t} - U_t}{\varepsilon} - V_t \right|^p \leq CK_\varepsilon e^{Ct} \leq CK_\varepsilon \quad \forall \varepsilon \in (0, 1), t \in [0, T]. \quad (4.37)$$

We observe that

$$\begin{aligned}
& \int_0^T E \left| b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))\tilde{X}_{\varepsilon,s}^2 - b''(s, x_s)U_s^2 \right|^p ds \\
& \leq 2^{p-1} \int_0^T E \left| b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))(\tilde{X}_{\varepsilon,s}^2 - U_s^2) \right|^p ds \\
& + 2^{p-1} \int_0^T E \left| (b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s)) - b''(s, x_s))U_s^2 \right|^p ds \\
& \leq 2^{p-1} L^p \int_0^T \sqrt{E|\tilde{X}_{\varepsilon,s} - U_s|^{2p} E|\tilde{X}_{\varepsilon,s} + U_s|^{2p}} ds \\
& + 2^{p-1} \int_0^T \sqrt{E|b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s)) - b''(s, x_s)|^{2p} E|U_s|^{4p}} ds.
\end{aligned}$$

Furthermore, $E|U_s|^q < \infty$ for all $q > 1$. Hence, by the estimate (4.33) and the dominated convergence theorem, we get

$$\int_0^T E \left| b''(s, x_s + \theta_s(X_{\varepsilon,s} - x_s))\tilde{X}_{\varepsilon,s}^2 - b''(s, x_s)U_s^2 \right|^p ds \rightarrow 0 \text{ as } \varepsilon \rightarrow 0.$$

Similarly, we also have

$$\int_0^T E \left| \sigma'(s, x_s + \eta_s(X_{\varepsilon,s} - x_s))\tilde{X}_{\varepsilon,s} - \sigma'(s, x_s)U_s \right|^p ds \rightarrow 0 \text{ as } \varepsilon \rightarrow 0.$$

Those imply that $K_\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$. So the desired relation (4.36) follows from (4.37). The proof of the proposition is complete. \square

Proposition 4.8. Suppose the assumptions (A_1) – (A_2) . Then, for any continuous and bounded function g and for each $t \in (0, T]$, we have

$$\lim_{\varepsilon \rightarrow 0} \frac{E[g(\tilde{X}_{\varepsilon,t})] - E[g(U_t)]}{\varepsilon} = \frac{1}{\beta_t^2} E[g(U_t)\delta(V_t DU_t)]. \quad (4.38)$$

Proof. For simplicity, we write $\langle \cdot, \cdot \rangle$ instead of $\langle \cdot, \cdot \rangle_{L^2[0,T]}$ and $\|\cdot\|$ instead of $\|\cdot\|_{L^2[0,T]}$. Fix $t \in (0, T]$, by using the formula (3.2) in [3], we get

$$E[g(\tilde{X}_{\varepsilon,t})] - E[g(U_t)] = E \left[\int_{U_t}^{\tilde{X}_{\varepsilon,t}} g(z) dz \delta \left(\frac{DU_t}{\|DU_t\|^2} \right) \right] - E \left[\frac{g(\tilde{X}_{\varepsilon,t}) \langle D\tilde{X}_{\varepsilon,t} - DU_t, DU_t \rangle}{\|DU_t\|^2} \right].$$

We observe that $\|DU_t\|^2 = \beta_t^2$ and $\delta \left(\frac{DU_t}{\|DU_t\|^2} \right) = U_t/\beta_t^2$. So we obtain

$$E[g(\tilde{X}_{\varepsilon,t})] - E[g(U_t)] = \frac{1}{\beta_t^2} E \left[U_t \int_{U_t}^{\tilde{X}_{\varepsilon,t}} g(z) dz \right] - \frac{1}{\beta_t^2} E \left[g(\tilde{X}_{\varepsilon,t}) \langle D\tilde{X}_{\varepsilon,t} - DU_t, DU_t \rangle \right].$$

Then, for $\varepsilon \neq 0$,

$$\begin{aligned}
& \frac{E[g(\tilde{X}_{\varepsilon,t})] - E[g(U_t)]}{\varepsilon} - \frac{1}{\beta_t^2} E[g(U_t)V_t U_t] + \frac{1}{\beta_t^2} E[g(U_t)\langle DV_t, DU_t \rangle] \\
& = \frac{1}{\beta_t^2} E \left[\left(\frac{1}{\varepsilon} \int_{U_t}^{\tilde{X}_{\varepsilon,t}} g(z) dz - g(U_t)V_t \right) U_t \right] - \frac{1}{\beta_t^2} E \left[(g(\tilde{X}_{\varepsilon,t}) - g(U_t)) \left\langle \frac{D\tilde{X}_{\varepsilon,t} - DU_t}{\varepsilon}, DU_t \right\rangle \right]
\end{aligned}$$

$$- \frac{1}{\beta_t^2} E \left[g(U_t) \left\langle \frac{D\tilde{X}_{\varepsilon,t} - DU_t}{\varepsilon} - DV_t, DU_t \right\rangle \right], \quad 0 < t \leq T. \quad (4.39)$$

We observe that

$$\begin{aligned} \frac{1}{\varepsilon} \int_{U_t}^{\tilde{X}_{\varepsilon,t}} g(z) dz - g(U_t) V_t &= \frac{\tilde{X}_{\varepsilon,t} - U_t}{\varepsilon} \int_0^1 g(U_t + z(\tilde{X}_{\varepsilon,t} - U_t)) dz - g(U_t) V_t \\ &= \left(\frac{\tilde{X}_{\varepsilon,t} - U_t}{\varepsilon} - V_t \right) \int_0^1 g(U_t + z(\tilde{X}_{\varepsilon,t} - U_t)) dz + V_t \int_0^1 (g(U_t + z(\tilde{X}_{\varepsilon,t} - U_t)) - g(U_t)) dz, \end{aligned}$$

and hence,

$$\begin{aligned} E \left| \left(\frac{1}{\varepsilon} \int_{U_t}^{\tilde{X}_{\varepsilon,t}} g(z) dz - g(U_t) V_t \right) U_t \right| &\leq \|g\|_{\infty} E \left| \left(\frac{\tilde{X}_{\varepsilon,t} - U_t}{\varepsilon} - V_t \right) U_t \right| \\ &\quad + E \left| V_t U_t \int_0^1 (g(U_t + z(\tilde{X}_{\varepsilon,t} - U_t)) - g(U_t)) dz \right|. \end{aligned}$$

Because the random variables U_t and V_t belong to $L^2(\Omega, \cdot)$ we have

$$\lim_{\varepsilon \rightarrow 0} E \left| \left(\frac{\tilde{X}_{\varepsilon,t} - U_t}{\varepsilon} - V_t \right) U_t \right| = 0 \quad \text{by the estimate (4.36),}$$

and

$$\lim_{\varepsilon \rightarrow 0} E \left| V_t U_t \int_0^1 (g(U_t + z(\tilde{X}_{\varepsilon,t} - U_t)) - g(U_t)) dz \right| = 0$$

by the dominated convergence theorem. So it holds that

$$\lim_{\varepsilon \rightarrow 0} E \left[\left(\frac{1}{\varepsilon} \int_{U_t}^{\tilde{X}_{\varepsilon,t}} g(z) dz - g(U_t) V_t \right) U_t \right] = 0. \quad (4.40)$$

On the other hand, we have

$$\begin{aligned} E \left[(g(\tilde{X}_{\varepsilon,t}) - g(U_t)) \left\langle \frac{D\tilde{X}_{\varepsilon,t} - DU_t}{\varepsilon}, DU_t \right\rangle \right] &\leq \frac{1}{\beta_t} E \left[\frac{|g(\tilde{X}_{\varepsilon,t}) - g(U_t)| \|D\tilde{X}_{\varepsilon,t} - DU_t\|}{\varepsilon} \right] \\ &\leq \frac{1}{\beta_t} (E |g(\tilde{X}_{\varepsilon,t}) - g(U_t)|^2)^{\frac{1}{2}} \left(\frac{E \|D\tilde{X}_{\varepsilon,t} - DU_t\|^2}{\varepsilon^2} \right)^{\frac{1}{2}}. \end{aligned}$$

Once again, by (4.34) and the dominated convergence theorem, we derive

$$\lim_{\varepsilon \rightarrow 0} E \left[(g(\tilde{X}_{\varepsilon,t}) - g(U_t)) \left\langle \frac{D\tilde{X}_{\varepsilon,t} - DU_t}{\varepsilon}, DU_t \right\rangle \right] = 0. \quad (4.41)$$

In view of Lemma 1.2.3 in [12], it follows from (4.34) and (4.36) that

$$\lim_{\varepsilon \rightarrow 0} E \left[g(U_t) \left\langle \frac{D\tilde{X}_{\varepsilon,t} - DU_t}{\varepsilon} - DV_t, DU_t \right\rangle \right] = 0. \quad (4.42)$$

Combining (4.39)–(4.42) yields

$$\lim_{\varepsilon \rightarrow 0} \frac{E[g(\tilde{X}_{\varepsilon,t})] - E[g(U_t)]}{\varepsilon} = \frac{1}{\beta_t^2} E[g(U_t) V_t U_t] - \frac{1}{\beta_t^2} E[g(U_t) \langle DV_t, DU_t \rangle].$$

Then we obtain (4.38) by using the duality relationships (2.2). This finishes the proof of the proposition. \square

Proof of Theorem 4.2. We first note that, for each $t \in (0, T]$, U_t is also a normal random variable with mean zero and variance β_t^2 . This means that $I(\tilde{X}_{\varepsilon,t} \| N_t) = I(\tilde{X}_{\varepsilon,t} \| U_t)$. Hence, it follows from the relations (4.32) and (4.38) that

$$\lim_{\varepsilon \rightarrow 0} \sqrt{I(\tilde{X}_{\varepsilon,t} \| N_t)} \geq \frac{1}{2\beta_t^2} \left| E[g(U_t) \delta(V_t DU_t)] \right| = \frac{1}{2\beta_t^2} \left| E[g(U_t) E[\delta(V_t DU_t) | U_t]] \right|$$

for any continuous function g bounded by 1. Then, by the routine approximation argument, we can choose to use $g(x) = \text{sign} \left(E[\delta(V_t DU_t) | U_t = x] \right)$ and we obtain (4.4).

The proof of Theorem 4.2 is complete. \square

4.4. Proof of Theorem 4.3

We recall that

$$Y_{\varepsilon,t} = \int_0^t f(s, X_{\varepsilon,s}) ds, \quad t \in [0, T].$$

By the chain rule for Malliavin derivatives, we have

$$D_r Y_{\varepsilon,t} = \int_r^t f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} ds, \quad (4.43)$$

$$D_\theta D_r Y_{\varepsilon,t} = \int_{\theta \vee r}^t [f''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} + f'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}] ds \quad (4.44)$$

for all $0 \leq r, \theta \leq t \leq T$, where $f'(t, x) = \frac{\partial f(t, x)}{\partial x}$, $f''(t, x) = \frac{\partial^2 f(t, x)}{\partial x^2}$.

For the proof of [Theorem 4.3](#) we will need the following [Propositions 4.9–4.12](#).

Proposition 4.9. Suppose the assumptions (A_1) – (A_3) . Then, we have

$$\sup_{0 \leq r \leq t} E |D_r Y_{\varepsilon,t}|^p \leq C t^p \varepsilon^p \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.45)$$

$$\sup_{0 \leq r, \theta \leq t} E |D_\theta D_r Y_{\varepsilon,t}|^p \leq C t^p \varepsilon^{2p} \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.46)$$

where C is a positive constant not depending on t and ε .

Proof. We first note that, by the polynomial growth property of f', f'' and the estimate [\(4.10\)](#), we have

$$\sup_{0 \leq s \leq T} E |f'(s, X_{\varepsilon,s})|^{2p} + \sup_{0 \leq s \leq T} E |f''(s, X_{\varepsilon,s})|^{2p} \leq C \quad (4.47)$$

for all $\varepsilon \in (0, 1)$, where C is a positive constant not depending on ε .

By using Hölder's inequality, it follows from [\(4.43\)](#) that

$$\begin{aligned} E |D_r Y_{\varepsilon,t}|^p &= E \left| \int_r^t f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} ds \right|^p \\ &\leq (t-r)^{p-1} \int_r^t E |f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s}|^p ds \\ &\leq t^{p-1} \int_r^t \sqrt{E |f'(s, X_{\varepsilon,s})|^{2p} E |D_r X_{\varepsilon,s}|^{2p}} ds \\ &\leq C t^{p-1} \int_r^t \sqrt{E |D_r X_{\varepsilon,s}|^{2p}} ds, \end{aligned}$$

which, together with the estimate [\(4.13\)](#), implies [\(4.45\)](#). The estimate [\(4.46\)](#) can be proved similarly. Indeed, we obtain from [\(4.44\)](#) that

$$\begin{aligned} E |D_\theta D_r Y_{\varepsilon,t}|^p &\leq 2^{p-1} \left(E \left| \int_{\theta \vee r}^t f''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s} ds \right|^p + E \left| \int_{\theta \vee r}^t f'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s} ds \right|^p \right) \\ &\leq 2^{p-1} t^{p-1} \left(\int_{\theta \vee r}^t E |f''(s, X_{\varepsilon,s}) D_\theta X_{\varepsilon,s} D_r X_{\varepsilon,s}|^p ds + \int_{\theta \vee r}^t E |f'(s, X_{\varepsilon,s}) D_\theta D_r X_{\varepsilon,s}|^p ds \right) \\ &\leq 2^{p-1} t^{p-1} \left(\int_{\theta \vee r}^t \sqrt[3]{E |f''(s, X_{\varepsilon,s})|^{3p} E |D_\theta X_{\varepsilon,s}|^{3p} E |D_r X_{\varepsilon,s}|^{3p}} ds \right. \\ &\quad \left. + \int_{\theta \vee r}^t \sqrt{E |f'(s, X_{\varepsilon,s})|^{2p} E |D_\theta D_r X_{\varepsilon,s}|^{2p}} ds \right) \\ &\leq C t^{p-1} \left(\int_{\theta \vee r}^t \sqrt[3]{E |D_\theta X_{\varepsilon,s}|^{3p} E |D_r X_{\varepsilon,s}|^{3p}} ds + \int_{\theta \vee r}^t \sqrt{E |D_\theta D_r X_{\varepsilon,s}|^{2p}} ds \right). \end{aligned}$$

So we obtain [\(4.46\)](#) by using the estimates [\(4.13\)](#) and [\(4.14\)](#). The proof of the proposition is complete. \square

Proposition 4.10. Suppose the assumptions (A_1) – (A_3) . Let $(\tilde{Y}_{\varepsilon,t})_{t \in [0, T]}$ be as in [Theorem 4.3](#). Then, we have

$$|E[\tilde{Y}_{\varepsilon,t}]| \leq C t^2 \varepsilon \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.48)$$

$$|\text{Var}(\tilde{Y}_{\varepsilon,t}) - \gamma_t^2| \leq C t^{7/2} \varepsilon \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.49)$$

where C is a positive constant not depending on t and ε .

Proof. We recall that

$$\tilde{Y}_{\varepsilon,t} = \frac{1}{\varepsilon} \int_0^t (f(s, X_{\varepsilon,s}) - f(s, x_s)) ds, \quad t \in [0, T].$$

Hence, in view of the estimate [\(4.17\)](#), we obtain

$$|E[\tilde{Y}_{\varepsilon,t}]| \leq \frac{1}{\varepsilon} \int_0^t E |f(s, X_{\varepsilon,s}) - f(s, x_s)| ds$$

$$\begin{aligned} &\leq \frac{1}{\varepsilon} \int_0^t \sqrt{E|f(s, X_{\varepsilon,s}) - f(s, x_s)|^2} ds \\ &\leq Ct^2\varepsilon. \end{aligned}$$

This completes the proof of (4.48). Let us prove the estimate (4.49). By the Clark–Ocone formula and the Itô isometry we have

$$\begin{aligned} \text{Var}(\tilde{Y}_{\varepsilon,t}) &= E \left[\int_0^T E[D_r \tilde{Y}_{\varepsilon,t} | \mathcal{F}_r] dB_r \right]^2 = \frac{1}{\varepsilon^2} E \left[\int_0^T (E[D_r Y_{\varepsilon,t} | \mathcal{F}_r])^2 dr \right] \\ &= \frac{1}{\varepsilon^2} E \left[\int_0^t \left(E \left[\int_r^t f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} ds \middle| \mathcal{F}_r \right] \right)^2 dr \right]. \end{aligned}$$

Then, we obtain

$$\begin{aligned} &\text{Var}(\tilde{Y}_{\varepsilon,t}) - \gamma_t^2 \\ &= \frac{1}{\varepsilon^2} E \left[\int_0^t \left(\int_r^t E \left[f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} \middle| \mathcal{F}_r \right] ds \right)^2 dr \right] - \int_0^t \left(\int_r^t f'(s, x_s) \sigma(r, x_r) e^{\int_r^s b'(u, x_u) du} ds \right)^2 dr \\ &= \int_0^t E[H_{t,r} G_{t,r}] dr, \end{aligned} \quad (4.50)$$

where, for $0 \leq r \leq t \leq T$, $H_{t,r}$ and $G_{t,r}$ are defined by

$$\begin{aligned} H_{t,r} &:= \frac{1}{\varepsilon} \int_r^t E \left[f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} \middle| \mathcal{F}_r \right] ds - \int_r^t f'(s, x_s) \sigma(r, x_r) e^{\int_r^s b'(u, x_u) du} ds, \\ G_{t,r} &:= \frac{1}{\varepsilon} \int_r^t E \left[f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} \middle| \mathcal{F}_r \right] ds + \int_r^t f'(s, x_s) \sigma(r, x_r) e^{\int_r^s b'(u, x_u) du} ds. \end{aligned}$$

Recalling the representation (4.21) we have

$$\begin{aligned} H_{t,r} &:= (\sigma(r, X_{\varepsilon,r}) - \sigma(r, x_r)) \int_r^t E \left[f'(s, X_{\varepsilon,s}) e^{\int_r^s b'(u, X_{\varepsilon,u}) du} Z_{r,s} \middle| \mathcal{F}_r \right] ds \\ &\quad + \sigma(r, x_r) \int_r^t E \left[(f'(s, X_{\varepsilon,s}) - f'(s, x_s)) e^{\int_r^s b'(u, x_u) du} Z_{r,s} \middle| \mathcal{F}_r \right] ds \\ &\quad + \sigma(r, x_r) \int_r^t f'(s, x_s) E \left[(e^{\int_r^s b'(u, X_{\varepsilon,u}) du} - e^{\int_r^s b'(u, x_u) du}) Z_{r,s} \middle| \mathcal{F}_r \right] ds \end{aligned}$$

and hence, note that b' is bounded by L and $\sup_{0 \leq r \leq T} |\sigma(r, x_r)| + \sup_{0 \leq r \leq T} |f'(r, x_r)| < \infty$, we deduce

$$\begin{aligned} |H_{t,r}| &\leq C |\sigma(r, X_{\varepsilon,r}) - \sigma(r, x_r)| \int_r^t E \left[|f'(s, X_{\varepsilon,s})| Z_{r,s} \middle| \mathcal{F}_r \right] ds \\ &\quad + C \int_r^t E \left[|f'(s, X_{\varepsilon,s}) - f'(s, x_s)| Z_{r,s} \middle| \mathcal{F}_r \right] ds + C \int_r^t E \left[\int_r^s |b'(u, X_{\varepsilon,u}) - b'(u, x_u)| du Z_{r,s} \middle| \mathcal{F}_r \right] ds. \end{aligned}$$

By using the Cauchy–Schwarz inequality we get

$$\begin{aligned} |H_{t,r}|^2 &\leq Ct |\sigma(r, X_{\varepsilon,r}) - \sigma(r, x_r)|^2 \int_r^t E \left[|f'(s, X_{\varepsilon,s})|^2 Z_{r,s}^2 \middle| \mathcal{F}_r \right] ds \\ &\quad + Ct \int_r^t E \left[|f'(s, X_{\varepsilon,s}) - f'(s, x_s)|^2 Z_{r,s}^2 \middle| \mathcal{F}_r \right] ds \\ &\quad + Ct^2 \int_r^t E \left[\int_r^s |b'(u, X_{\varepsilon,u}) - b'(u, x_u)|^2 du Z_{r,s}^2 \middle| \mathcal{F}_r \right] ds \end{aligned}$$

and

$$\begin{aligned} E|H_{t,r}|^2 &\leq Ct^{\frac{3}{2}} \sqrt{E|\sigma(r, X_{\varepsilon,r}) - \sigma(r, x_r)|^4} \int_r^t E \left[|f'(s, X_{\varepsilon,s})|^4 Z_{r,s}^4 \right] ds \\ &\quad + Ct \int_r^t E \left[|f'(s, X_{\varepsilon,s}) - f'(s, x_s)|^2 Z_{r,s}^2 \right] ds + Ct^2 \int_r^t \int_r^s E \left[|b'(u, X_{\varepsilon,u}) - b'(u, x_u)|^2 Z_{r,s}^2 \right] dud s \\ &\leq Ct^{\frac{3}{2}} \sqrt{E|\sigma(r, X_{\varepsilon,r}) - \sigma(r, x_r)|^4} \int_r^t \sqrt{E|f'(s, X_{\varepsilon,s})|^8 E|Z_{r,s}|^8} ds \\ &\quad + Ct \int_r^t \sqrt{E|f'(s, X_{\varepsilon,s}) - f'(s, x_s)|^4 E|Z_{r,s}|^4} ds \\ &\quad + Ct^2 \int_r^t \int_r^s \sqrt{E|b'(u, X_{\varepsilon,u}) - b'(u, x_u)|^4 E|Z_{r,s}|^4} dud s. \end{aligned}$$

In view of the estimates (4.17), (4.24) and (4.47), we therefore obtain

$$E|H_{t,r}|^2 \leq Ct^3 \varepsilon^2 \quad \forall \varepsilon \in (0, 1), 0 \leq r \leq t \leq T, \quad (4.51)$$

where C is a positive constant not depending on t and ε . On the other hand, we have

$$\begin{aligned} E|G_{t,r}|^2 &= E\left|H_{t,r} + 2 \int_r^t f'(s, x_s) \sigma(r, x_r) e^{\int_r^s b'(u, x_u) du} ds\right|^2 \\ &\leq 2E|H_{t,r}|^2 + 8\left|2 \int_r^t f'(s, x_s) \sigma(r, x_r) e^{\int_r^s b'(u, x_u) du} ds\right|^2 \\ &\leq Ct^2 \quad \forall \varepsilon \in (0, 1), 0 \leq r \leq t \leq T. \end{aligned} \quad (4.52)$$

Combining (4.50), (4.51) and (4.52) gives us

$$\begin{aligned} |\text{Var}(\tilde{Y}_{\varepsilon,t}) - \sigma_t^2| &= \int_0^t E|H_{t,r} G_{t,r}| dr \\ &\leq \int_0^t \sqrt{E|H_{t,r}|^2 E|G_{t,r}|^2} dr \\ &\leq Ct^{\frac{7}{2}} \varepsilon \quad \forall \varepsilon \in (0, 1), t \in [0, T]. \end{aligned}$$

This finishes the proof of (4.49). The proof of the proposition is complete. \square

Proposition 4.11. Let $(\tilde{Y}_{\varepsilon,t})_{t \in [0, T]}$ be as in Theorem 4.3. Define

$$\Theta_{\tilde{Y}_{\varepsilon,t}} := \int_0^t D_r \tilde{Y}_{\varepsilon,t} E[D_r \tilde{Y}_{\varepsilon,t} | \mathcal{F}_r] dr, \quad t \in [0, T].$$

Then, under the assumption of Theorem 4.3, we have

$$E|\Theta_{\tilde{Y}_{\varepsilon,t}}|^{-p} \leq C \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{p}{p_0}} \quad \forall \varepsilon \in (0, 1), t \in (0, T], \quad (4.53)$$

where $0 < p < p_0$ and $C > 0$ is a positive constant not depending on t and ε .

Proof. We have

$$\begin{aligned} \Theta_{\tilde{Y}_{\varepsilon,t}} &= \frac{1}{\varepsilon^2} \int_0^t D_r Y_{\varepsilon,t} E[D_r Y_{\varepsilon,t} | \mathcal{F}_r] dr \\ &= \frac{1}{\varepsilon^2} \int_0^t \left(\int_r^t f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} ds \right) \left(\int_r^t E[f'(s, X_{\varepsilon,s}) D_r X_{\varepsilon,s} | \mathcal{F}_r] ds \right) dr. \end{aligned}$$

Note that $\|f'\|_0 := \inf_{(t,x)} f'(t, x) > 0$. Hence, by using the same arguments as in the proof of Proposition 4.4, we obtain

$$\Theta_{\tilde{Y}_{\varepsilon,t}} \geq \|f'\|_0^2 e^{-\frac{5LT}{2}} e^{-2 \max_{0 \leq t \leq T} M_t} \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr$$

and, for $0 < p < p_0$, we get

$$E|\Theta_{\tilde{Y}_{\varepsilon,t}}|^{-p} \leq C \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{p}{p_0}} \quad \forall \varepsilon \in (0, 1), t \in (0, T],$$

where $C > 0$ is a positive constant not depending on t and ε . The proof of the proposition is complete. \square

Proposition 4.12. Let $(\Theta_{\tilde{Y}_{\varepsilon,t}})_{t \in [0, T]}$ be as in Proposition 4.11. Suppose the assumptions (A_1) – (A_3) . Then, we have

$$E\|D\Theta_{\tilde{Y}_{\varepsilon,t}}\|_{L^2[0, T]}^4 \leq Ct^{14} \varepsilon^4 \quad \forall \varepsilon \in (0, 1), t \in [0, T], \quad (4.54)$$

where C is a positive constant not depending on t and ε .

Proof. The Malliavin derivative of $\Theta_{\tilde{Y}_{\varepsilon,t}}$ can be computed by

$$D_\theta \Theta_{\tilde{Y}_{\varepsilon,t}} = \frac{1}{\varepsilon^2} \int_0^t (D_\theta D_r Y_{\varepsilon,t} E[D_r Y_{\varepsilon,t} | \mathcal{F}_r] + D_r Y_{\varepsilon,t} E[D_\theta D_r Y_{\varepsilon,t} | \mathcal{F}_r]) dr, \quad 0 \leq \theta \leq t.$$

Hence,

$$\begin{aligned} \|D\Theta_{\tilde{Y}_{\varepsilon,t}}\|_{L^2[0, T]}^2 &= \int_0^T |D_\theta \Theta_{\tilde{Y}_{\varepsilon,t}}|^2 d\theta \\ &= \int_0^t \frac{1}{\varepsilon^4} \left(\int_0^t (D_\theta D_r Y_{\varepsilon,t} E[D_r Y_{\varepsilon,t} | \mathcal{F}_r] + D_r Y_{\varepsilon,t} E[D_\theta D_r Y_{\varepsilon,t} | \mathcal{F}_r]) dr \right)^2 d\theta. \end{aligned}$$

Using the Hölder's inequality, we obtain

$$E\|D\Theta_{\tilde{Y}_{\varepsilon,t}}\|_{L^2[0, T]}^4 \leq \frac{t}{\varepsilon^8} \int_0^t E \left| \int_0^t (D_\theta D_r Y_{\varepsilon,t} E[D_r Y_{\varepsilon,t} | \mathcal{F}_r] + D_r Y_{\varepsilon,t} E[D_\theta D_r Y_{\varepsilon,t} | \mathcal{F}_r]) dr \right|^4 d\theta$$

$$\begin{aligned}
&\leq \frac{t^4}{\varepsilon^8} \int_0^t \int_0^t E \left| D_\theta D_r Y_{\varepsilon,t} E[D_r Y_{\varepsilon,t} | \mathcal{F}_r] + D_r Y_{\varepsilon,t} E[D_\theta D_r Y_{\varepsilon,t} | \mathcal{F}_r] \right|^4 dr d\theta \\
&\leq \frac{8t^4}{\varepsilon^8} \int_0^t \int_0^t \left(E \left| D_\theta D_r Y_{\varepsilon,t} E[D_r Y_{\varepsilon,t} | \mathcal{F}_r] \right|^4 + E \left| D_r Y_{\varepsilon,t} E[D_\theta D_r Y_{\varepsilon,t} | \mathcal{F}_r] \right|^4 \right) dr d\theta \\
&\leq \frac{16t^4}{\varepsilon^8} \int_0^t \int_0^t \sqrt{E[D_\theta D_r Y_{\varepsilon,t}]^8 E[D_r Y_{\varepsilon,t}]^8} dr d\theta.
\end{aligned}$$

This, combined with (4.45) and (4.46), gives us the estimate (4.54). The proof of the proposition is complete. \square

Proof of Theorem 4.3. Set $\tilde{u}_r = E[D_r \tilde{Y}_{\varepsilon,t} | \mathcal{F}_r]$ for $0 \leq r \leq t \leq T$. We have

$$\|\tilde{u}\|_{L^2[0,T]}^8 = \left[\int_0^T \left| E[D_r \tilde{Y}_{\varepsilon,t} | \mathcal{F}_r] \right|^2 dr \right]^4 = \frac{1}{\varepsilon^8} \left[\int_0^t \left| E[D_r Y_{\varepsilon,t} | \mathcal{F}_r] \right|^2 dr \right]^4.$$

Using Hölder's inequality and the estimate (4.45) we deduce

$$\begin{aligned}
E\|\tilde{u}\|_{L^2[0,T]}^8 &\leq \frac{t^3}{\varepsilon^8} \int_0^t E \left| E[D_r Y_{\varepsilon,t} | \mathcal{F}_r] \right|^8 dr \\
&\leq \frac{t^3}{\varepsilon^8} \int_0^t E[|D_r Y_{\varepsilon,t}|^8] dr \leq Ct^{12}.
\end{aligned} \tag{4.55}$$

We now apply Theorem 3.1 to $F = \tilde{Y}_{\varepsilon,t}$ and $N = Z_t$ to get

$$I(\tilde{Y}_{\varepsilon,t} \parallel N_t) \leq c \left(\frac{1}{\gamma_t^4} (E[\tilde{Y}_{\varepsilon,t}]^2 + A_{\tilde{Y}_{\varepsilon,t}} |\text{Var}(\tilde{Y}_{\varepsilon,t}) - \gamma_t^2|^2 + C_{\tilde{Y}_{\varepsilon,t}} (E\|D\Theta_{\tilde{Y}_{\varepsilon,t}}\|_{L^2[0,T]}^4)^{1/2}) \right),$$

where c is an absolute constant and

$$A_{\tilde{Y}_{\varepsilon,t}} := \frac{1}{\gamma_t^4} \left(E\|\tilde{u}\|_{L^2[0,T]}^8 E|\Theta_{\tilde{Y}_{\varepsilon,t}}|^{-8} \right)^{1/4}, C_{\tilde{Y}_{\varepsilon,t}} := A_{\tilde{Y}_{\varepsilon,t}} + \left(E\|\tilde{u}\|_{L^2[0,T]}^8 E|\Theta_{\tilde{Y}_{\varepsilon,t}}|^{-16} \right)^{1/4}.$$

It follows from the estimates (4.53) and (4.55) that

$$\begin{aligned}
A_{\tilde{Y}_{\varepsilon,t}} &\leq \frac{Ct^3}{\gamma_t^4} \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{2}{p_0}}, \\
C_{\tilde{Y}_{\varepsilon,t}} &\leq \frac{Ct^3}{\gamma_t^4} \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{2}{p_0}} + Ct^3 \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{4}{p_0}}.
\end{aligned}$$

Furthermore, thanks to Propositions 4.10 and 4.12, we have

$$\begin{aligned}
|E[\tilde{Y}_{\varepsilon,t}]|^2 &\leq Ct^4 \varepsilon^2, \\
|\text{Var}(\tilde{Y}_{\varepsilon,t}) - \beta_t^2|^2 &\leq Ct^7 \varepsilon^2, \\
(E\|D\Theta_{\tilde{Y}_{\varepsilon,t}}\|^4)^{1/2} &\leq Ct^7 \varepsilon^2.
\end{aligned}$$

Combining the above computations gives us

$$\begin{aligned}
I(\tilde{Y}_{\varepsilon,t} \parallel Z_t) &\leq C \left(\frac{t^4}{\gamma_t^4} + \frac{t^{10}}{\gamma_t^4} \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{2}{p_0}} \right. \\
&\quad \left. + t^{10} \left(E \left| \int_0^t (t-r)^2 \sigma^2(r, X_{\varepsilon,r}) dr \right|^{-p_0} \right)^{\frac{4}{p_0}} \right) \varepsilon^2.
\end{aligned}$$

So the proof of Theorem 4.3 is complete. \square

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the anonymous referees for valuable comments which led to the improvement of this work.

Appendix. Negative moment of Volterra functionals

Lemma A. Let $(Y_t)_{t \in [0, T]}$ be a stochastic process such that Y_0 is deterministic and $E|Y_t - Y_0|^p \leq C t^{p\alpha}$ for some $C > 0, \alpha > 0$ and for all $t \in [0, T], p > 1$. Let $h : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ and $k : [0, T]^2 \rightarrow \mathbb{R}_+$ be continuous functions such that $h(0, X_0) \neq 0, k(t, 0) \neq 0$ for all $t \in (0, T]$ and

$$|h(t, x) - h(s, y)| \leq L(|t - s|^{\delta_1} + |x - y|^{\delta_2}) \quad \forall x, y \in \mathbb{R}, s, t \in [0, T] \quad (\text{A.1})$$

$$|k(t, s) - k(t, 0)| \leq L|t - s|^{\delta_3} \quad \forall s, t \in [0, T] \quad (\text{A.2})$$

for some positive constants L, δ_1, δ_2 and δ_3 . Then, for $\delta_0 := \min\{\delta_1, \alpha\delta_2, \delta_3\}$ and for all $p_0 > 0$, we have

$$E \left[\left(\int_0^t k(t, r) h^2(r, Y_r) dr \right)^{-p_0} \right] \leq C \left(\frac{1}{tk(t, 0)} \right)^{p_0} \left(1 + \left(\frac{t^{\delta_0}}{k(t, 0)} \right)^p \right) \quad (\text{A.3})$$

for all $t \in (0, T]$ and $p > \frac{p_0}{\delta_0}$, where C is a positive constant not depending on t .

Proof. Fixed $t \in (0, T]$. Put $y_0 := \frac{2}{th^2(0, Y_0)}$. Then, for any $y \geq y_0$, we have $\eta_y := \frac{2}{tyh^2(0, Y_0)} \in (0, 1)$. Hence,

$$\begin{aligned} \frac{1}{k(t, 0)} \int_0^t k(t, r) h^2(r, Y_r) dr &\geq \frac{1}{k(t, 0)} \int_0^{\eta_y t} k(t, r) h^2(r, Y_r) dr \\ &= h^2(0, Y_0) \eta_y t + \frac{1}{k(t, 0)} \int_0^{\eta_y t} (k(t, r) h^2(r, Y_r) - k(t, 0) h^2(0, Y_0)) dr \\ &\geq \frac{2}{y} - \frac{1}{k(t, 0)} \int_0^{\eta_y t} |k(t, r) h^2(r, Y_r) - k(t, 0) h^2(0, Y_0)| dr. \end{aligned}$$

Consequently, by Markov inequality, we obtain

$$\begin{aligned} P \left(\frac{1}{k(t, 0)} \int_0^t k(t, r) h^2(r, Y_r) dr \leq \frac{1}{y} \right) &\leq P \left(\frac{1}{k(t, 0)} \int_0^{\eta_y t} |k(t, r) h^2(r, Y_r) - k(t, 0) h^2(0, Y_0)| dr \geq \frac{1}{y} \right) \\ &\leq \frac{y^p}{k^p(t, 0)} E \left| \int_0^{\eta_y t} |k(t, r) h^2(r, Y_r) - k(t, 0) h^2(0, Y_0)| dr \right|^p \quad \forall p > 1. \end{aligned}$$

We now use Hölder's inequality to get

$$\begin{aligned} P \left(\frac{1}{k(t, 0)} \int_0^t k(t, r) h^2(r, Y_r) dr \leq \frac{1}{y} \right) &\leq \frac{y^p (\eta_y t)^{p-1}}{k^p(t, 0)} \int_0^{\eta_y t} E |k(t, r) h^2(r, Y_r) - k(t, 0) h^2(0, Y_0)|^p dr \\ &\leq \frac{2^{p-1} y^p (\eta_y t)^{p-1}}{k^p(t, 0)} \int_0^{\eta_y t} \left(k^p(t, r) E |h^2(r, Y_r) - h^2(0, Y_0)|^p + h^{2p}(0, Y_0) |k(t, r) - k(t, 0)|^p \right) dr \\ &\leq \frac{C y^p (\eta_y t)^{p-1}}{k^p(t, 0)} \int_0^{\eta_y t} \left(E |h^2(r, Y_r) - h^2(0, Y_0)|^p + r^{p\delta_3} \right) dr \quad \forall p > 1, \end{aligned}$$

where C is a positive constant not depending on t . We observe that

$$\begin{aligned} E |h(r, Y_r) - h(0, Y_0)|^{2p} &\leq 2^{2p-1} L^{2p} (r^{2p\delta_1} + E |Y_r - Y_0|^{2p\delta_2}) \\ &\leq C(r^{2p\delta_1} + r^{2p\alpha\delta_2}) \end{aligned}$$

for all $r \in [0, T]$. We also have

$$E |h(r, Y_r) + h(0, Y_0)|^{2p} \leq E |h(r, Y_r) - h(0, Y_0)|^{2p} + |2h(0, Y_0)|^{2p} \leq C$$

for all $r \in [0, T]$. So it holds that

$$\begin{aligned} E |h^2(r, Y_r) - h^2(0, Y_0)|^p &\leq \sqrt{E |h(r, Y_r) + h(0, Y_0)|^{2p} E |h(r, Y_r) - h(0, Y_0)|^{2p}} \\ &\leq C(r^{2p\delta_1} + r^{2p\alpha\delta_2}), \quad r \in [0, T]. \end{aligned}$$

Combining the above estimates yields

$$P \left(\frac{1}{k(t, 0)} \int_0^t k(t, r) h^2(r, Y_r) dr \leq \frac{1}{y} \right) \leq \frac{C y^p (\eta_y t)^{(1+\delta_0)p}}{k^p(t, 0)} \leq \frac{C y^{-p\delta_0}}{k^p(t, 0)} \quad \forall p > 1,$$

Then, for all $p > \frac{p_0}{\delta_0}$, we obtain

$$E \left[\left(\int_0^t k(t, r) h^2(r, Y_r) dr \right)^{-p_0} \right] = \frac{1}{k^{p_0}(t, 0)} E \left[\left(\frac{1}{k(t, 0)} \int_0^t k(t, r) h^2(r, Y_r) dr \right)^{-p_0} \right]$$

$$\begin{aligned}
&= \frac{1}{k^{p_0}(t, 0)} \int_0^\infty p_0 y^{p_0-1} P \left(\frac{1}{k(t, 0)} \int_0^t k(t, r) h^2(r, Y_r) dr \leq \frac{1}{y} \right) dy \\
&\leq \frac{1}{k^{p_0}(t, 0)} \left(\int_0^{y_0} p_0 y^{p_0-1} dy + \int_{y_0}^\infty p_0 y^{p_0-1} P \left(\int_0^t \frac{1}{k(t, 0)} h^2(r, Y_r) dr \leq \frac{1}{y} \right) dy \right) \\
&\leq \frac{1}{k^{p_0}(t, 0)} \left(y_0^{p_0} + \frac{C}{k^p(t, 0)} \int_{y_0}^\infty p_0 y^{p_0-1} y^{-p\delta_0} dy \right) \quad \forall p > 1.
\end{aligned}$$

We now choose to use $p = 1 + \frac{p_0}{\delta_0}$. This choice gives us the following

$$E \left[\left(\int_0^t k(t, r) h^2(r, Y_r) dr \right)^{-p_0} \right] \leq C \left(\frac{1}{tk(t, 0)} \right)^{p_0} \left(1 + \left(\frac{t^{\delta_0}}{k(t, 0)} \right)^p \right).$$

So we obtain the desired estimate (A.3) because $y_0 := \frac{2}{th^2(0, Y_0)}$. \square

References

- [1] S.G. Bobkov, G.P. Chistyakov, F. Götze, Berry–esseen bounds in the entropic central limit theorem, *Probab. Theory Related Fields* 159 (3–4) (2014) 435–478.
- [2] N.T. Dung, N. Privault, G.L. Torrisi, Gaussian estimates for the solutions of some one-dimensional stochastic equations, *Potential Anal.* 43 (2) (2015) 289–311.
- [3] N.T. Dung, T.C. Son, Lipschitz continuity in the Hurst index of the solutions of fractional stochastic volterra integro-differential equations, *Stoch. Anal. Appl.* 41 (4) (2023) 693–712.
- [4] M.I. Freidlin, A.D. Wentzell, *Random Perturbations of Dynamical Systems*, Vol. 260, Springer, Heidelberg, 2012, Translated from the 1979 Russian original by Joseph Szűcs. Third edition. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].
- [5] O. Johnson, *Information Theory and the Central Limit Theorem*, Imperial College Press, London, 2004.
- [6] O. Johnson, Maximal correlation and the rate of Fisher information convergence in the central limit theorem, *IEEE Trans. Inform. Theory* 66 (8) (2020) 4992–5002.
- [7] O. Johnson, A.R. Barron, Fisher information inequalities and the central limit theorem, *Probab. Theory Related Fields* 129 (3) (2004) 391–409.
- [8] Y. Kutoyants, Identification of Dynamical Systems with Small Noise, in: *Mathematics and its Applications*, vol. 300, Kluwer Academic Publishers Group, Dordrecht, 1994.
- [9] J.V. Linnik, An information-theoretic proof of the central limit theorem with Lindeberg conditions. (Russian), *Theor. Probab. Appl.* 4 (1959) 288–299.
- [10] I. Nourdin, D. Nualart, Fisher information and the fourth moment theorem, *Ann. Inst. Henri Poincaré Probab. Stat.* 52 (2) (2016) 849–867.
- [11] I. Nourdin, G. Peccati, G. Reinert, Second order Poincaré inequalities and CLTs on Wiener space, *J. Funct. Anal.* 257 (2) (2009) 593–609.
- [12] D. Nualart, *The Malliavin Calculus and Related Topics*, second ed., in: *Probability and its Applications*, Springer-Verlag, Berlin, 2006.
- [13] M.S. Pinsker, *Information and Information Stability of Random Variables and Processes*, San Francisco, Calif.-London-Amsterdam, 1964, Translated and edited by Amiel Feinstein Holden-Day, Inc..
- [14] K. Spiliopoulos, Fluctuation analysis and short time asymptotics for multiple scales diffusion processes, *Stoch. Dyn.* 14 (3) (2014) 1350026, 22pp.
- [15] A.J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Inf. Control* 2 (1959) 101–112.
- [16] Y. Suo, J. Tao, W. Zhang, Moderate deviation and central limit theorem for stochastic differential delay equations with polynomial growth, *Front. Math. China* 13 (4) (2018) 913–933.