

An application of the KMT construction to the pathwise weak error in the Euler approximation of one-dimensional diffusion process with linear diffusion coefficient

Emmanuelle Clément, Arnaud Gloter

▶ **To cite this version:**

Emmanuelle Clément, Arnaud Gloter. An application of the KMT construction to the pathwise weak error in the Euler approximation of one-dimensional diffusion process with linear diffusion coefficient. 2016. hal-01167276v2

HAL Id: hal-01167276
https://hal.science/hal-01167276v2

Preprint submitted on 24 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An application of the KMT construction to the pathwise weak error in the Euler approximation of one-dimensional diffusion process with linear diffusion coefficient

Emmanuelle Clément* Arnaud Gloter†

June 16, 2016

Abstract

It is well known that the strong error approximation, in the space of continuous paths equipped with the supremum norm, between a diffusion process, with smooth coefficients, and its Euler approximation with step $1/n$ is $O(n^{-1/2})$ and that the weak error estimation between the marginal laws, at the terminal time T, is $O(n^{-1})$. An analysis of the weak trajectorial error has been developed by Alfonsi, Jourdain and Kohatsu-Higa [1], through the study of the $p-$Wasserstein distance between the two processes. For a one-dimensional diffusion, they obtained an intermediate rate for the pathwise Wasserstein distance of order $n^{-2/3+\varepsilon}$. Using the Komlós, Major and Tusnády construction, we improve this bound, assuming that the diffusion coefficient is linear, and we obtain a rate of order $\log n/n$.

MSC 2010. 65C30, 60H35.

Key words: Diffusion process, Euler scheme, Wasserstein couplings, Komlós-Major-Tusnády construction.

1 Introduction

A classical problem in numerical probabilities is the computation of $Ef(X)$, where $X = (X_t)_{t\in[0,1]}$ is a stochastic process defined on the time interval $[0,1]$ and f a functional which may depend on the

*Université Paris-Est, LAMA, UMR 8050, UPEMLV, UPEC, CNRS, F-77454, Marne-la-Vallée, France.

†Université d'Évry Val d'Essonne, Laboratoire de Mathématiques et Modélisation d'Evry, UMR 8071, 91025 Évry Cedex, France. This author's research was supported by the chair Markets in transition (Fédération Bancaire Française) and the project ANR 11-LABX-0019

whole path of the process X. This problem appears for instance in finance where X represents the dynamic of a stock price and f the payoff of an option. The usual way to solve this problem is to approximate X by a numerical scheme and then to compute the expectation by using a Monte Carlo method.

Due to its implementation easiness, the most popular discretization scheme, when X is a diffusion process, is the Euler scheme. Denoting by $\overline{X}^n$, the Euler approximation of X with step $1/n$, it is well known that the pathwise strong order of convergence between X and $\overline{X}^n$ is $n^{-1/2}$, under regularity assumptions on the coefficients of the diffusion X (see for example [8]). Moreover the weak order of convergence at a fixed time t, evaluated by the difference $|Ef(X_t) - Ef(\overline{X}_t^n)|$, is n^{-1} (see [16]). However, for the pathwise weak approximation of X (when f depends on the whole trajectory of X), the order of convergence is still unknown, excepted for specific functionals f such as $f(X) = \int_0^1 X_s ds$ or $f(X) = \max_s X_s$. Recently Alfonsi, Jourdain and Kohatsu-Higa [1], [2] have proposed a general approach to control the pathwise weak approximation of a diffusion by its Euler scheme by considering the Wasserstein distance between the law of X and the law of $\overline{X}^n$.

For X and $\overline{X}$, two random variables with values in a normed vector space $(\mathcal{X}, \|.\|)$ and with finite p-moment for $1 \leq p < \infty$, the Wasserstein distance $\mathcal{W}_p$ between the law of X and the law of $\overline{X}$ is defined by :

$$\mathcal{W}_p(\mathcal{L}(X), \mathcal{L}(\overline{X})) = \inf_{(Y,\overline{Y}) \in \Pi(X,\overline{X})} E^{1/p} \left\| Y - \overline{Y} \right\|^p. \tag{1}$$

$\Pi(X, \overline{X})$ is the set of random variables $(Y, \overline{Y})$ with values in $\mathcal{X} \times \mathcal{X}$ with marginal laws respectively $\mathcal{L}(X)$ and $\mathcal{L}(\overline{X})$.

In our context, $\mathcal{X} = \mathcal{C}([0,1])$, equipped with the supremum norm $\|x\| = \sup_{t \in [0,1]} |x_t|$ or $\mathcal{X} = \mathbb{R}^n$, equipped with the norm of the maximum of the coordinates $\|x\| = \max_{i \in \{1,\dots,n\}} |x_i|$.

From the representation of $\mathcal{W}_1$ in the Kantorovitch duality (see for example [14]) :

$$\mathcal{W}_1(\mathcal{L}(X), \mathcal{L}(\overline{X})) = \sup_{f \in \mathcal{L}(1)} |Ef(X) - Ef(\overline{X})|$$

where $\mathcal{L}(1)$ is the set of Lipschitz functions $f : \mathcal{C}([0,1]) \mapsto \mathbb{R}$ with Lipschitz constant less than 1, and using the strong and weak orders of convergence of the Euler scheme, one can easily deduce the following upper and lower bounds :

$$\frac{c}{n} \leq \mathcal{W}_1(\mathcal{L}(X), \mathcal{L}(\overline{X}^n)) \leq \frac{C}{\sqrt{n}},$$

for some positive constants c and C.

For a one-dimensional uniformly elliptic diffusion, the main result of [1], is to construct a coupling between X and $\overline{X}^n$ which improves the preceding upper bound and leads to :

$$\mathcal{W}_p(\mathcal{L}(X), \mathcal{L}(\overline{X}^n)) \leq \frac{C}{n^{\frac{2}{3}-\varepsilon}},$$

for all $\varepsilon > 0$. This result gives an intermediate rate, for the pathwise weak approximation, between the strong order rate and the weak marginal rate and raises a natural question : is it possible to construct a coupling between a diffusion and its Euler scheme in such a way that the Wasserstein distance is of order $1/n$?

The aim of this paper is to give an answer to this question assuming that the diffusion coefficient is linear. In that case, we prove (see Theorem 2) :

$$\mathcal{W}_p(\mathcal{L}(X), \mathcal{L}(\overline{X}^n)) \leq C\frac{\log n}{n}.$$

This result is obtained by the construction of a sharp discrete time coupling between $(X_{k/n})_{1\leq k\leq n}$ and $(\overline{X}^n_{k/n})_{1\leq k\leq n}$, following the dyadic construction due to Komlós, Major and Tusnády ([10] [11]). We mention that recently, the KMT construction has been used in a series of papers (Davie [3], [4], Flint and Lyons [7]), to propose an approximation scheme, close to the Milstein scheme and with weak pathwise order of convergence $1/n$.

The KMT construction permits essentially to obtain an optimal coupling between a sequence of i.i.d. standard gaussian variables $(Y_i)_{1\leq i\leq n}$ and some other i.i.d. variables $(X_i)_{1\leq i\leq n}$ with finite Laplace transform in a neighbourhood of zero and such that $EX_1 = 0$, $VX_1 = 1$, in such a way that almost surely :

$$\max_{1\leq k\leq n} |\sum_{i=1}^{k} Y_i - \sum_{i=1}^{k} X_i| \leq C \log n.$$

In Section 2, we improve the KMT result when the variables X_i are equal in law to $Y_i - \frac{1}{2\sqrt{n}}(Y_i^2 - 1)$. In this particular case, we obtain, as a consequence of Theorem 1 below, that almost surely :

$$\max_{1\leq k\leq n} |\sum_{i=1}^{k} Y_i - \sum_{i=1}^{k} X_i| \leq C \log n/\sqrt{n}.$$

This is done through refined quantile coupling inequalities, which are established at the end of the paper, in Section 4. These results are applied in Section 3 to construct a coupling between a diffusion process with linear diffusion coefficient and its Euler approximation which achieves the pathwise weak order $\log n/n$.

In all the paper, C denotes a constant which value does not depend on n and may change from one line to the other.

2 A KMT type result

Let $(Y_i)_{i\geq 1}$ be a sequence of i.i.d. standard gaussian variables and let us consider the triangular array:

$$\overline{Y}_i^n = Y_i - \frac{1}{2\sqrt{n}}(Y_i^2 - 1), \quad 1 \leq i \leq n. \tag{2}$$

We set : $S_k = \sum_{i=1}^k Y_i$ and $\overline{S}_k^n = \sum_{i=1}^k \overline{Y}_i^n$.

In this framework, we can improve the classical KMT result.

Theorem 1 *One can construct on the same probability space a sequence of i.i.d. standard gaussian variables $(Y_i)_{1\leq i\leq n}$ and a sequence of i.i.d. variables $(X_i^n)_{1\leq i\leq n}$, with X_i^n equal in law to $\overline{Y}_i^n$, such that for positive constants C , K and λ, we have, for n large enough and for all $x > 0$:*

$$P(\sqrt{n} \sup_{1\leq k\leq n} |S_k - \overline{T}_k^n| \geq K \log n + x) \leq Ce^{-\lambda x}, \tag{3}$$

where $\overline{T}_k^n = \sum_{i=1}^k X_i^n$, and $S_k = \sum_{i=1}^k Y_i$.

A straightforward consequence of this result is that almost surely :

$$\max_{1\leq k\leq n} |S_k - \overline{T}_k^n| \leq C \log n/\sqrt{n}.$$

The coupling given in Theorem 1 improves the classical KMT result with a factor $1/\sqrt{n}$ and permits to control the Wasserstein distance between the law of $(S_k)_k$ and the law of $(\overline{S}_k^n)_k$ with the rate $\log n/n$ (see Corollary 1 below).

Remark 1 *1) If we consider directly the coupling between the random walks S and $\overline{S}^n$ (based on the same gaussian variables) , we have*

$$\max_{1\leq k\leq n} \left|S_k - \overline{S}_k^n\right| = \max_{1\leq k\leq n} \left|\frac{1}{2\sqrt{n}} \sum_{i=1}^k (Y_i^2 - 1)\right|$$

and consequently, from Donsker's theorem we deduce that $\max_{1\leq k\leq n} \left|S_k - \overline{S}_k^n\right|$ converges in law to $\sup_{t\in[0,1]} \frac{1}{\sqrt{2}}|B_t|$, where B is a standard Brownian motion. Observing moreover that (3) can be rewritten as

$$P(\sup_{1\leq k\leq n} |S_k - \overline{T}_k^n| \geq x) \leq Ce^{-\lambda\sqrt{n}x + K\log n}$$

we see that the result of Theorem 1 can not be obtain from the basic coupling between S and $\overline{S}^n$ and that the KMT coupling leads to a better result, and in turn to a sharper bound for the Wasserstein distance.

2) It is known that the classical KMT coupling result is optimal for random walks based on i.i.d sequences (see Theorem 2 in [10]). In Theorem 1, we improve the rate of the KMT result in the situation where $(\overline{Y}_i^n)$ is a triangular array of random variables, whose law depends on n. It seems crucial, here, that the law of $\overline{Y}_i^n$ becomes close to a Gaussian law as n growths.

The proof of Theorem 1 is postponed in Section 4. It is obtained by using the KMT method developed in [10], [11]. The main tool for this construction is a gaussian coupling to the partial sums $\overline{S}_k^n$, which is based essentially on a large deviation expansion of $\overline{p}_k^n(x)/\phi(x)$ where $\overline{p}_k^n$ is the density function of $\frac{1}{\sqrt{k}}\overline{S}_k^n$ and ϕ the density function of the standard gaussian law. We state and prove this large deviation expansion and the associated coupling inequalities at the end of Section 4.

As a consequence of this theorem, we deduce an upper bound for the Wasserstein distance $\mathcal{W}_p(\mathcal{L}((S_k)_{1\leq k\leq n}), \mathcal{L}((\overline{S}_k^n)_{1\leq k\leq n}))$.

Corollary 1 *For all $p \geq 1$, there exists a positive constant C such that :*

$$\mathcal{W}_p(\mathcal{L}((S_k)_{1\leq k\leq n}), \mathcal{L}((\overline{S}_k^n)_{1\leq k\leq n})) \leq C\frac{\log n}{\sqrt{n}}.$$

Proof Let $(S_k)_{1\leq k\leq n}$ and $(\overline{T}_k^n)_{1\leq k\leq n}$ be constructed as in Theorem 1. From the definition of the Wasserstein distance, one has :

$$\mathcal{W}_p(\mathcal{L}((S_k)_{1\leq k\leq n}), \mathcal{L}((\overline{S}_k^n)_{1\leq k\leq n})) \leq E^{1/p} \sup_{1\leq k\leq n} |S_k - \overline{T}_k^n|^p,$$

and so we just have to prove that :

$$E(\sqrt{n} \sup_{1\leq k\leq n} |S_k - \overline{T}_k^n|)^p \leq C(\log n)^p. \tag{4}$$

Recalling that for any positive random variable Z, and any $p \geq 1$:

$$EZ^p = \int_0^\infty pz^{p-1}P(Z > z)dz, \tag{5}$$

we deduce that

$$\begin{aligned} E(\sqrt{n} \sup_{1\leq k\leq n} |S_k - \overline{T}_k^n|)^p &\leq \int_0^\infty pz^{p-1}P(\sqrt{n} \sup_{1\leq k\leq n} |S_k - \overline{T}_k^n|) > z)dz, \\ &\leq \int_0^{K\log n} pz^{p-1}dz + \int_{K\log n}^\infty pz^{p-1}P(\sqrt{n} \sup_{1\leq k\leq n} |S_k - \overline{T}_k^n|) > z)dz, \end{aligned} \tag{6}$$

where K is the constant given in Theorem 1. The first integral in the righthandside of (6) is clearly bounded by $C(\log n)^p$. For the second one, we have, using successively the change of variables $z = x + K\log n$ and Theorem 1

$$\int_{K\log n}^{\infty} pz^{p-1}P(\sqrt{n}\sup_{1\le k\le n}|S_k - \overline{T}_k^n|) > z)dz \le C\int_0^{\infty} p(x+K\log n)^{p-1}e^{-\lambda x}dx \le C(\log n)^p.$$

This gives (4) and Corollary 1 is proved.

$\diamond$

3 Application to the Euler approximation of a diffusion

In this section, we apply the preceding results to bound the pathwise Wasserstein distance between a diffusion with linear diffusion coefficient and its Euler approximation. Let $X = (X_t)_{t\in[0,1]}$ be the solution of the stochastic differential equation :

$$X_0 = x_0, \quad dX_t = b(X_t)dt + X_t dB_t, \tag{7}$$

where (B_t) is a standard Brownian motion, $x_0 \in \mathbb{R}$. We assume that b admits a derivative denoted by $b^{(1)}$, and that b and $b^{(1)}$ are Lipschitz functions.

We consider the continuous time Euler approximation of X, with step $1/n$, defined by :

$$\overline{X}_0^n = x_0, \quad d\overline{X}_t^n = b(\overline{X}_{\varphi_n(t)}^n)dt + \overline{X}_{\varphi_n(t)}^n dB_t, \tag{8}$$

where $\varphi_n(t) = \frac{[nt]}{n}$, $t \in [0,1]$.

Using $\overline{X}_t^n - \overline{X}_{\varphi_n(t)}^n = b(\overline{X}_{\varphi_n(t)}^n)(t-\varphi_n(t)) + \overline{X}_{\varphi_n(t)}^n(B_t - B_{\varphi_n(t)})$, we can write heuristically the dynamic of $\overline{X}^n$ as :

$$\overline{X}_t^n = x_0 + \int_0^t b(\overline{X}_s^n)ds + \int_0^t \overline{X}_s^n(1-(B_s - B_{\varphi_n(s)}))dB_s + O(\frac{1}{n}).$$

We can observe that this dynamic is mainly driven by the process (L_t^n) defined by :

$$L_t^n = B_t - \int_0^t (B_s - B_{\varphi_n(s)})dB_s, \quad t \in [0,1]. \tag{9}$$

From this observation, to study the Wasserstein distance between $\mathcal{L}(X)$ and $\mathcal{L}(\overline{X}^n)$, a natural way is to introduce the process $Y_t = x_0 + \int_0^t b(Y_s)ds + \int_0^t Y_s dL_s^n$.

Following this heuristic idea, we consider the auxiliary process $\tilde{X}^n$, which approximates $\overline{X}^n$ with pathwise strong order $1/n$ (see Lemma 1) :

$$\tilde{X}_0^n = x_0, \quad d\tilde{X}_t^n = b(\tilde{X}_t^n)dt - \frac{1}{2}\tilde{X}_t^n(1-(1-(B_t - B_{\varphi_n(t)}))^2)dt + \tilde{X}_t^n dL_t^n. \tag{10}$$

The addition of the drift term $\frac{1}{2}\tilde{X}^n_t(1-(1-(B_t-B_{\varphi_n(t)}))^2)dt$ in the dynamic of the process is not essential but permits to obtain the representation formula (12) below.

Following Doss [5] and using Itô's formula, we first remark that the processes X and $\tilde{X}^n$ admit the representations :

$$X_t = e^{B_t}\left(x_0 + \int_0^t e^{-B_s}(b(X_s) - \frac{1}{2}X_s)ds\right), \quad t \in [0,1], \tag{11}$$

$$\tilde{X}^n_t = e^{L^n_t}\left(x_0 + \int_0^t e^{-L^n_s}(b(\tilde{X}^n_s) - \frac{1}{2}\tilde{X}^n_s)ds\right), \quad t \in [0,1]. \tag{12}$$

Based on these representations, a first step to control the Wasserstein distance between $\mathcal{L}(X)$ and $\mathcal{L}(\tilde{X}^n)$ is to bound $\mathcal{W}_p(\mathcal{L}(B), \mathcal{L}(L^n))$. This can be done by using the results of Section 2.

More precisely, observing that from Itô's formula $\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t)$, we deduce that the discrete process $(L^n_{\frac{k}{n}})$ satisfies :

$$L^n_{\frac{k}{n}} = B_{\frac{k}{n}} - \frac{1}{2}\sum_{i=1}^k((B_{\frac{i}{n}} - B_{\frac{i-1}{n}})^2 - \frac{1}{n}) = \sum_{i=1}^k\left(B_{\frac{i}{n}} - B_{\frac{i-1}{n}} - \frac{1}{2}((B_{\frac{i}{n}} - B_{\frac{i-1}{n}})^2 - \frac{1}{n})\right), \quad 1 \le k \le n,$$

and consequently, $(L^n_{\frac{k}{n}})_{1\le k\le n}$ is equal in law to $\frac{1}{\sqrt{n}}(\overline{S}^n_k)_{1\le k\le n}$, where using the notations of Section 2, $\overline{S}^n_k = \sum_{i=1}^k \overline{Y}^n_i$, with $\overline{Y}^n_i$ defined by (2). Similarly, we observe that $(B_{\frac{k}{n}})_{1\le k\le n}$ is equal in law to $\frac{1}{\sqrt{n}}(S_k)_{1\le k\le n}$. This permits to derive immediately from Corollary 1 the following result.

Corollary 2 *For $p \ge 1$, there exists a positive constant C, such that for n large enough :*

$$\mathcal{W}_p(\mathcal{L}((B_{\frac{k}{n}})_{1\le k\le n}), \mathcal{L}((L^n_{\frac{k}{n}})_{1\le k\le n})) \le C\frac{\log n}{n}.$$

Next, we can extend this result to the continuous processes $B = (B_t)_{t\in[0,1]}$ and $L^n = (L^n_t)_{t\in[0,1]}$ using the strong approximation error on each interval with length $1/n$.

Proposition 1 *a) For $p \ge 1$, there exists a positive constant C, such that for n large enough :*

$$\mathcal{W}_p(\mathcal{L}(B), \mathcal{L}(L^n)) \le C\frac{\log n}{n}.$$

b) Let $F : \mathcal{C}([0,1]) \mapsto \mathcal{C}([0,1])$ satisfying :

$$\forall f, g \in \mathcal{C}([0,1]), \quad \|F(f) - F(g)\| \le C(f,g)\|f - g\|.$$

Assuming that $\forall p \ge 1$, $\sup_n EC(B, L^n)^p < \infty$, then for $p \ge 1$, there exists a positive constant C, such that for n large enough :

$$\mathcal{W}_p(\mathcal{L}(F(B)), \mathcal{L}(F(L^n))) \le C\frac{\log n}{n}.$$

Proof a) We consider the process (L_t^n) defined by (9), driven by the Brownian motion $(\tilde{B}_t)_t$, and we introduce the process B_t^n :

$$B_t^n = L^n_{\frac{k-1}{n}} + \tilde{B}_t - \tilde{B}_{\frac{k-1}{n}}, \quad \text{for} \quad \frac{k-1}{n} \leq t < \frac{k}{n}. \tag{13}$$

The process (B_t^n) is discontinuous and coincide with (L_t^n) at the discretization times $(k/n)_{0\leq k\leq n-1}$.

First, we prove the following strong approximation result :

$$E \max_{1\leq k\leq n} \sup_{t\in[\frac{k-1}{n},\frac{k}{n}]} |B_t^n - L_t^n|^p \leq C\frac{(\log n)^p}{n^p}. \tag{14}$$

To prove (14), we will use (5) with $Z = n \max_{1\leq k\leq n} \sup_{t\in[\frac{k-1}{n},\frac{k}{n}]} |B_t^n - L_t^n|$ and so we have to control $P(Z > z)$, for $z > 0$. We have :

$$P(Z > z) \leq n \max_k P(n \sup_{t\in[\frac{k-1}{n},\frac{k}{n}]} |B_t^n - L_t^n| > z),$$

with $n(B_t^n - L_t^n) = n\int_{\frac{k-1}{n}}^t (\tilde{B}_s - \tilde{B}_{\varphi_n(s)})d\tilde{B}_s$, for $\frac{k-1}{n} \leq t < \frac{k}{n}$. Observing that the processes $(\int_{\frac{k-1}{n}}^t (\tilde{B}_s - \tilde{B}_{\varphi_n(s)})d\tilde{B}_s)_{t\in[\frac{k-1}{n},\frac{k}{n}]}$ and $(\int_0^t \tilde{B}_s d\tilde{B}_s)_{t\in[0,\frac{1}{n}]}$ have the same law, we deduce :

$$P(Z > z) \leq nP(n \sup_{t\in[0,\frac{1}{n}]} |\int_0^t \tilde{B}_s d\tilde{B}_s| > z).$$

But since $\int_0^t \tilde{B}_s d\tilde{B}_s = \frac{1}{2}(\tilde{B}_t^2 - t)$, we have :

$$P(Z > z) \leq nP\left(n \sup_{t\in[0,\frac{1}{n}]} |\tilde{B}_t^2| > 2(z - \frac{1}{2})\right),$$

and by time rescaling :

$$P(Z > z) \leq nP\left(\sup_{t\in[0,1]} |\tilde{B}_t^2| > 2(z - \frac{1}{2})\right).$$

Using the exponential inequality for the Brownian motion (see Proposition 1.8 in [15]), we have $P(\sup_{t\in[0,1]} |\tilde{B}_t| > a) \leq 2e^{-a^2/2}$, and this finally leads to :

$$P(Z > z) \leq Cne^{-(z-\frac{1}{2})}. \tag{15}$$

Turning back to (5), we have :

$$EZ^p \leq \int_0^{\log n} pz^{p-1}dz + \int_{z>\log n} pz^{p-1}P(Z > z)dz. \tag{16}$$

Reporting (15) in the second integral of (16) and using the change of variables $x = z - \log n$, we deduce :

$$EZ^p \leq C(\log n)^p.$$

This proves the strong approximation result (14).

We end the proof as in [1] (Proof of Theorem 3.2.). The Wasserstein distance in the left hand side of Corollary 2 is attained for a probability measure π on $\mathbb{R}^n \times \mathbb{R}^n$ with marginal laws respectively the law of a Brownian motion at times $(k/n)_{0\leq k\leq n}$ and the law of $(L_{\frac{k}{n}})_k$. We fix $(L^n_{\frac{k}{n}})_k$ to be the discretization of the solution of (9) for a Brownian motion $(\tilde{B}_t)_t$ and let $(\overline{B}_{\frac{k}{n}})_k$ be distributed according to the first marginal of π given the second one equal to $(L^n_{\frac{k}{n}})_k$. The random variable $((\overline{B}_{\frac{k}{n}})_k, L^n_{\frac{k}{n}})_k)$ in $\mathbb{R}^n \times \mathbb{R}^n$ is distributed according to π and attains the Wasserstein distance between $\mathcal{L}((B_{\frac{k}{n}})_k)$ and $\mathcal{L}((L^n_{\frac{k}{n}})_k)$.

By the triangle inequality, we have:

$$\mathcal{W}_p(\mathcal{L}(B), \mathcal{L}(L^n)) \leq \mathcal{W}_p(\mathcal{L}(B), \mathcal{L}(B^n)) + \mathcal{W}_p(\mathcal{L}(B^n), \mathcal{L}(L^n)), \tag{17}$$

where $B^n = (B^n_t)_{t\in[0,1]}$ is defined by (13). Let us note that the process B^n is not continuous and so the associated Wasserstein distance is defined in $D([0,1])$, the space of c-à-d-l-à-g functions, equipped with the supremum norm.

From the strong error approximation (14), the second right handside term in (17) is bound by $C \log n/n$ and to end the proof it remains to estimate

$$\mathcal{W}_p(\mathcal{L}(B), \mathcal{L}(B^n)).$$

For this we consider a Brownian motion $(W_t)_{t\in[0,1]}$ independent on $((\overline{B}_{\frac{k}{n}})_k, (\tilde{B}_t)_t)$ and we construct the two Brownian Bridges driven by (W_t) : $(W_t^{\overline{B}_{\frac{k-1}{n}}, \overline{B}_{\frac{k}{n}}})_{t\in[\frac{k-1}{n}, \frac{k}{n}]}$ (starting from $\overline{B}_{\frac{k-1}{n}}$ and ending at $\overline{B}_{\frac{k}{n}}$), and $(W_t^{B^n_{\frac{k-1}{n}}, B^n_{\frac{k}{n}-}})_{t\in[\frac{k-1}{n}, \frac{k}{n}]}$ (starting from $B^n_{\frac{k-1}{n}}$ and ending at $B^n_{\frac{k}{n}-}$, where $B^n_{\frac{k}{n}-}$ is the left hand limit at time $\frac{k}{n}$ of (B^n_t)). We set for $t \in [\frac{k-1}{n}, \frac{k}{n})$ and $1 \leq k \leq n$:

$$W^1_t = W_t^{\overline{B}_{\frac{k-1}{n}}, \overline{B}_{\frac{k}{n}}} = \overline{B}_{\frac{k-1}{n}} n(\frac{k}{n} - t) + \overline{B}_{\frac{k}{n}} n(t - \frac{k-1}{n}) + W_t - W_{\frac{k-1}{n}} - n(t - \frac{k-1}{n})(W_{\frac{k}{n}} - W_{\frac{k-1}{n}}),$$

$$W^2_t = W_t^{B^n_{\frac{k-1}{n}}, B^n_{\frac{k}{n}-}} = B^n_{\frac{k-1}{n}} n(\frac{k}{n} - t) + B^n_{\frac{k}{n}-} n(t - \frac{k-1}{n}) + W_t - W_{\frac{k-1}{n}} - n(t - \frac{k-1}{n})(W_{\frac{k}{n}} - W_{\frac{k-1}{n}}).$$

One can check that $\mathcal{L}((W_t^1)_t) = \mathcal{L}((B_t)_t)$ and $\mathcal{L}((W_t^2)_t) = \mathcal{L}((B_t^n)_t)$, consequently,

$$\begin{aligned}\mathcal{W}_p(\mathcal{L}(B), \mathcal{L}(B^n)) &\leq E^{1/p} \sup_{t\in[0,1]} |W_t^1 - W_t^2|^p \\ &\leq E^{1/p} \max_k (|\overline{B}_{\frac{k-1}{n}} - B^n_{\frac{k-1}{n}}|^p \vee |\overline{B}_{\frac{k}{n}} - B^n_{\frac{k}{n}-}|^p) \\ &\leq E^{1/p} \max_k |\overline{B}_{\frac{k}{n}} - L^n_{\frac{k}{n}}|^p + E^{1/p} \max_k |L^n_{\frac{k}{n}} - B^n_{\frac{k}{n}-}|^p.\end{aligned}$$

We have

$$E^{1/p} \max_k |L^n_{\frac{k}{n}} - B^n_{\frac{k}{n}-}|^p \leq E^{1/p} \max_k \sup_{t\in[\frac{k-1}{n}, \frac{k}{n}]} |B_t^n - L_t^n|^p,$$

and by construction of the process $(\overline{B}_{\frac{k}{n}})_k$,

$$E^{1/p} \max_k |\overline{B}_{\frac{k}{n}} - L^n_{\frac{k}{n}}|^p = \mathcal{W}_p(\mathcal{L}((B_{\frac{k}{n}})_k), \mathcal{L}((L^n_{\frac{k}{n}})_k)),$$

Consequently, using Corollary 2 and (14), we finally obtain :

$$\mathcal{W}_p(\mathcal{L}(B), \mathcal{L}(B^n)) \leq C \log n/n,$$

and a) is proved.

b) Let $((\overline{B}_t)_{t\in[0,1]}, (\overline{L}_t^n)_{t\in[0,1]})$ be a random variable in $\mathcal{C}([0,1]) \times \mathcal{C}([0,1])$ which attains the Wasserstein distance $\mathcal{W}_{2p}$ in Proposition 1 :

$$\mathcal{W}_{2p}(\mathcal{L}(\overline{B}), \mathcal{L}(\overline{L}^n)) = E^{\frac{1}{2p}} \sup_{t\in[0,1]} |\overline{B}_t - \overline{L}_t^n|^{2p} \leq C \frac{\log n}{n}.$$

Then we have :

$$\mathcal{W}_p(\mathcal{L}(F(B)), \mathcal{L}((F(L^n))) \leq E^{1/p} \left(C(\overline{B}, \overline{L}^n) \left\|\overline{B} - \overline{L}^n\right\|\right)^p,$$

and b) follows from the Cauchy-Schwarz inequality.

This achieves the proof of Proposition 1.

◇

From this proposition, we deduce a bound for the Wasserstein distance between $\mathcal{L}(X)$ and $\mathcal{L}(\tilde{X}^n)$.

Proposition 2 *For $p \geq 1$, there exists a positive constant C, such that for n large enough :*

$$\mathcal{W}_p(\mathcal{L}(X), \mathcal{L}((\tilde{X}^n)) \leq C \frac{\log n}{n}.$$

Proof

The proof is based on the representation formulas (11) and (12), and Proposition 1 b).

We introduce the notation :

$$D_t = x_0 + \int_0^t e^{-B_s}(b(X_s) - \frac{1}{2}X_s)ds, \tag{18}$$

and

$$D_t^n = x_0 + \int_0^t e^{-L_s^n}(b(\tilde{X}_s^n) - \frac{1}{2}\tilde{X}_s^n)ds, \tag{19}$$

so that $X_t = e^{B_t}D_t$ and $\tilde{X}_t^n = e^{L_t^n}D_t^n$. From the triangle inequality :

$$|X_t - \tilde{X}_t^n| \leq |e^{B_t} - e^{L_t^n}||D_t| + e^{L_t^n}|D_t - D_t^n|.$$

Since b is Lipschitz :

$$\begin{aligned} |D_t - D_t^n| &\leq C\left(\int_0^t e^{-B_s}|X_s - \tilde{X}_s^n|ds + \int_0^t |e^{-B_s} - e^{-L_s^n}|(|\tilde{X}_s^n| + 1)ds\right), \\ &\leq C\left(\int_0^t |D_s - D_s^n|ds + \int_0^t e^{-B_s}|e^{B_s} - e^{L_s^n}||D_s^n|ds + \int_0^t |e^{-B_s} - e^{-L_s^n}|(|\tilde{X}_s^n| + 1)ds\right), \end{aligned}$$

and from Gronwall's Lemma,

$$\|D - D^n\| \leq C(B, L^n, D^n, \tilde{X}^n)\|B - L^n\|,$$

with

$$C(B, L^n, D^n, \tilde{X}^n) = C[1 + \left\|e^{-B_\cdot}\right\|\|D^n\|(\left\|e^{B_\cdot}\right\| + \left\|e^{L_\cdot^n}\right\|) + (\left\|\tilde{X}^n\right\| + 1)(\left\|e^{-B_\cdot}\right\| + \left\|e^{-L_\cdot^n}\right\|],$$

where we have used : $\left\|e^{B_\cdot} - e^{L_\cdot^n}\right\| \leq \left[\left\|e^{B_\cdot}\right\| + \left\|e^{L_\cdot^n}\right\|\right] \|B - L^n\|$. This yields :

$$\left\|X - \tilde{X}^n\right\| \leq \left[(\left\|e^{B_\cdot}\right\| + \left\|e^{L_\cdot^n}\right\|)\|D\| + \left\|e^{L_\cdot^n}\right\|C(B, L^n, D^n, \tilde{X}^n)\right] \|B - L^n\|.$$

But, for all $p \geq 1$, we have $E\left\|e^{|B_\cdot|}\right\|^p < \infty$, $E\|X\|^p < \infty$ (see [9] p306) and consequently $E\|D\|^p < \infty$. So from Proposition 1 b), to finish the proof of Proposition 2, we just have to verify

$$\forall p \geq 1, \quad \sup_n E\left\|\tilde{X}^n\right\|^p < \infty, \tag{20}$$

$$\forall p \geq 1, \quad \sup_n E\left\|e^{|L_\cdot^n|}\right\|^p < \infty. \tag{21}$$

To prove (21), we just prove that $\forall p \geq 1$, $\sup_n E\left\|e^{L_\cdot^n}\right\|^p < \infty$, since we obtain with similar arguments that $\sup_n E\left\|e^{-L_\cdot^n}\right\|^p < \infty$.

From (9), the martingale $(L_t^n)_{t\in[0,1]}$ can be written as :

$$L_t^n = \int_0^t (1 - (B_s - B_{\varphi_n(s)}))dB_s.$$

We first bound $Ee^{a\langle L^n,L^n\rangle_1} = E(e^{a\int_0^1(1-(B_s-B_{\varphi_n(s)}))^2ds})$, for $a>0$. We have

$$E(e^{a\int_0^1(1-(B_s-B_{\varphi_n(s)}))^2ds}) \leq e^{2a}Ee^{2a\sum_{k=0}^{n-1}\int_{\frac{k}{n}}^{\frac{k+1}{n}}(B_s-B_{\frac{k}{n}})^2ds}.$$

Since the random variables $(\int_{\frac{k}{n}}^{\frac{k+1}{n}}(B_s-B_{\frac{k}{n}})^2ds)_{0\leq k\leq n-1}$ are independent and identically distributed, we deduce

$$E(e^{a\int_0^1(1-(B_s-B_{\varphi_n(s)}))^2ds}) \leq e^{2a}(Ee^{2a\int_0^{\frac{1}{n}}B_s^2ds})^n \leq e^{2a}(Ee^{\frac{2a}{n}\sup_{s\leq 1/n}B_s^2})^n.$$

By time rescaling, $\sup_{0\leq s\leq 1/n}B_s^2$ is equal in law to $\frac{1}{n}\sup_{0\leq s\leq 1}B_s^2$ and consequently

$$E(e^{a\int_0^1(1-(B_s-B_{\varphi_n(s)}))^2ds}) \leq e^{2a}(Ee^{\frac{2a}{n^2}\sup_{s\leq 1}B_s^2})^n.$$

From Hölder's inequality and Doob's maximal inequality applied to the positive submartingale $(e^{\frac{2a}{n^2}B_t^2})_t$, we have for $q>1$,

$$Ee^{\frac{2a}{n^2}\sup_{s\leq 1}B_s^2} \leq (E\sup_{s\in[0,1]}e^{\frac{2aq}{n^2}B_s^2})^{1/q} \leq \frac{q}{q-1}(Ee^{\frac{2aq}{n^2}B_1^2})^{1/q}.$$

Remarking that for $\alpha<1/2$, $Ee^{\alpha B_1^2}=1/\sqrt{1-2\alpha}$, this gives, for n large enough and choosing $q=n$:

$$Ee^{\frac{2a}{n^2}\sup_{s\leq 1}B_s^2} \leq \frac{n}{n-1}(1-\frac{4a}{n})^{-\frac{1}{2n}}.$$

This permits to obtain

$$Ee^{a\langle L^n,L^n\rangle_1} = E(e^{a\int_0^1(1-(B_s-B_{\varphi_n(s)}))^2ds}) \leq e^{2a}\left(\frac{n}{n-1}\right)^n(1-\frac{4a}{n})^{-\frac{1}{2}} \leq C_a, \tag{22}$$

where C_a is a constant depending on a but not on n. From Novikov's criterion, we deduce that for any $a>0$, $(\mathcal{E}(aL_t^n))_{t\in[0,1]} = (e^{aL_t^n-\frac{a^2}{2}\langle L^n,L^n\rangle_t})_{t\in[0,1]}$ is a martingale. Observing that

$$e^{pL_t^n} = \mathcal{E}(2pL_t^n)^{1/2}e^{p^2\langle L^n,L^n\rangle_t},$$

is a positive submartingale, and applying Doob's maximal inequality, we have :

$$E(\sup_{t\in[0,1]}e^{pL_t^n}) \leq C(Ee^{2pL_1^n})^{1/2} \leq C(Ee^{8p^2\langle L^n,L^n\rangle_1})^{1/4},$$

and so from (22), this gives :

$$E(\sup_{t\in[0,1]}e^{pL_t^n}) \leq C_p.$$

This achieves the proof of (21).

It remains to prove (20). We recall that $\tilde{X}_t^n = e^{L_t^n} D_t^n$, with D_t^n given by (19). Since b is Lipschitz, we have :

$$|D_t^n| \leq |x_0| + C\left\|e^{-L_\cdot^n}\right\| + C\int_0^t |D_s^n| ds,$$

so from Gronwall's Lemma : $\|D^n\| \leq C\left\|e^{-L_\cdot^n}\right\|$ and then (20) is a straightforward consequence of (21).

$\diamond$

From these intermediate results, we deduce a bound for the Wasserstein distance between the law of the diffusion and the law of its Euler approximation. Our main result is the following.

Theorem 2 *For $p \geq 1$, there exists a positive constant C, such that for n large enough :*

$$\mathcal{W}_p(\mathcal{L}(X), \mathcal{L}(\overline{X}^n)) \leq C\frac{\log n}{n}.$$

Proof

The result of Theorem 2 follows from Proposition 2 and Lemma 1 below , applying the triangle inequality and observing that $\mathcal{W}_p(\mathcal{L}((\overline{X}_t^n)_{t\in[0,1]}), \mathcal{L}((\tilde{X}_t^n)_{t\in[0,1]})) \leq E^{1/p} \sup_{t\in[0,1]} |\overline{X}_t^n - \tilde{X}_t^n|^p$.

$\diamond$

For the statement of Lemma 1, we recall that $(\overline{X}_t^n)$ and $(\tilde{X}_t^n)$ are defined on the same probability space by (8) and (10).

Lemma 1 *For $p \geq 1$, there exists a positive constant C, such that $\forall n \geq 1$:*

$$\left(E \sup_{t\in[0,1]} |\overline{X}_t^n - \tilde{X}_t^n|^p\right)^{1/p} \leq C/n.$$

Proof To simplify the notation, we write $\Delta B_t = B_t - B_{\varphi_n(t)}$ and $\Delta_t = t - \varphi_n(t)$ and we denote by $(U_t^n)_{t\in[0,1]} = (\tilde{X}_t^n - \overline{X}_t^n)_{t\in[0,1]}$ the error process.

We first remark that for all $p \geq 1$, $E\sup_{t\in[0,1]} |\overline{X}_t^n|^p \leq C$ (see [9] p306) and $E\sup_{t\in[0,1]} |\tilde{X}_t^n|^p \leq C$ (see (20)) , so $E\sup_{t\in[0,1]} |U_t^n|^p \leq C$, for some positive constant C. Moreover from these bounds, it is easy to see that for $p \geq 1$, there exists $C > 0$ such that

$$\forall t \in [0,1], \quad E|\overline{X}_t^n - \overline{X}_{\varphi_n(t)}^n|^p \leq C/n^{p/2}, \tag{23}$$

$$\forall t \in [0,1], \quad E|\tilde{X}_t^n - \tilde{X}_{\varphi_n(t)}^n|^p \leq C/n^{p/2}. \tag{24}$$

From (8) and using the preceding notations, we have : $\overline{X}^n_t - \overline{X}^n_{\varphi_n(t)} = b(\overline{X}^n_{\varphi_n(t)})\Delta_t + \overline{X}^n_{\varphi_n(t)}\Delta B_t$, and we can write the dynamic of the Euler scheme as :

$$d\overline{X}^n_t = b(\overline{X}^n_t)dt + \overline{X}^n_t dL^n_t - (b(\overline{X}^n_t) - b(\overline{X}^n_{\varphi_n(t)}))dt + (\overline{X}^n_t - \overline{X}^n_{\varphi_n(t)})\Delta B_t dB_t - b(\overline{X}^n_{\varphi_n(t)})\Delta_t dB_t.$$

Now, it is easy to verify from the expressions of $\tilde{X}^n$ and L^n (equations (10) and (9)) and the preceding equation that (U^n_t) satisfies the equation :

$$U^n_t = \int_0^t (b(\tilde{X}^n_s) - b(\overline{X}^n_s))ds + \int_0^t U^n_s(1 - \Delta B_s)dB_s + R^n_t, \tag{25}$$

with

$$dR^n_t = -\frac{1}{2}\tilde{X}^n_t(2\Delta B_t - (\Delta B_t)^2)dt + (b(\overline{X}^n_t) - b(\overline{X}^n_{\varphi_n(t)}))dt - (\overline{X}^n_t - \overline{X}^n_{\varphi_n(t)})\Delta B_t dB_t + b(\overline{X}^n_{\varphi_n(t)})\Delta_t dB_t. \tag{26}$$

From Burkholder-Davis-Gundy inequality and using the Lipschitz assumption on b, we have

$$E \sup_{t\in[0,1]} |\int_0^t b(\overline{X}^n_{\varphi_n(s)})\Delta_s dB_s|^p \leq C/n^p$$

and since $\int_0^t \tilde{X}^n_{\varphi_n(s)}\Delta B_s ds = \int_0^t \tilde{X}^n_{\varphi_n(s)}[(t \wedge (\varphi_n(s) + \frac{1}{n})) - s]dB_s$, we obtain

$$E \sup_{t\in[0,1]} |\int_0^t \tilde{X}^n_{\varphi_n(s)}\Delta B_s ds|^p \leq C/n^p,$$

and similarly

$$E \sup_{t\in[0,1]} |\int_0^t \overline{X}^n_{\varphi_n(s)}\Delta B_s ds|^p \leq C/n^p.$$

Moreover we have the expansion, for $\eta_t \in [\overline{X}^n_{\varphi_n(t)}, \overline{X}^n_t]$:

$$b(\overline{X}^n_t) - b(\overline{X}^n_{\varphi_n(t)}) = b^{(1)}(\eta_t)(\overline{X}^n_t - \overline{X}^n_{\varphi_n(t)}) = [b^{(1)}(\overline{X}^n_{\varphi_n(t)}) + (b^{(1)}(\eta_t) - b^{(1)}(\overline{X}^n_{\varphi_n(t)}))](\overline{X}^n_t - \overline{X}^n_{\varphi_n(t)}).$$

This permits to conclude, after a few computation involving the estimations (23)–(24) and the Lipschitz assumption on b and $b^{(1)}$, that :

$$E \sup_{t\in[0,1]} |R^n_t|^p \leq C_p/n^p. \tag{27}$$

Turning back to (25), we deduce, using once again the Lipschitz assumption on b together with convexity and Burkholder-Davis-Gundy inequalities and the bound (27), for $p \geq 2$:

$$E \sup_{v\leq t} |U^n_v|^p \leq C\left(\int_0^t E\left((1 + |\Delta B_s|)^p \sup_{v\leq s} |U^n_v|^p\right) ds + \frac{1}{n^p}\right). \tag{28}$$

Moreover, introducing the truncation $1_{|\Delta B_s| \leq C_1}$, we have for any $q \geq 1$:

$$P(|\Delta B_s| > C_1) \leq C/n^q,$$

and so from Cauchy Schwarz inequality

$$E\left((1+|\Delta B_s|)^p \sup_{v \leq s} |U_v^n|^p 1_{|\Delta B_s| > C_1}\right) \leq C/n^p,$$

this gives :

$$E\left((1+|\Delta B_s|)^p \sup_{v \leq s} |U_v^n|^p\right) \leq C(E \sup_{v \leq s} |U_v^n|^p + \frac{1}{n^p}),$$

Reporting this in (28), we deduce :

$$E \sup_{v \leq t} |U_v^n|^p \leq C\left(\int_0^t E \sup_{v \leq s} |U_v^n|^p ds + \frac{1}{n^p}\right),$$

and the result of Lemma 1 follows from Gronwall's Lemma.

◇

4 Quantile coupling inequalities and proof of Theorem 1

This section is devoted to the proof of Theorem 1 and is organized as follows. In Section 4.1 we construct a coupling between a sequence of i.i.d. standard Gaussian variables $(Y_k)_{1 \leq k \leq n}$ and a sequence of i.i.d. variables $(X_k^n)_{1 \leq k \leq n}$ such that X_k^n has the same distribution as $\overline{Y}_k^n$ given in (2). Then, in Section 4.2 we prove that (3) of Theorem 1 holds true for this specific coupling. In Sections 4.3–4.4. we provide the proof of technical lemmas which are used in Sections 4.1–4.2.

For technical reasons, essentially the non integrability of the characteristic function of the random variables $\overline{Y}_k^n$, we regularize them by adding independent normally distributed random variables. For that, we consider a sequence of independent identically distributed standard Gaussian variables $(\xi_k)_{k \geq 1}$, independent of the sequence $(Y_k)_{k \geq 1}$, and we set :

$$\overline{Y}_k^{\star,n} = \overline{Y}_k^n + \frac{1}{n}\xi_k, \tag{29}$$

$$\overline{S}_k^{\star,n} = \sum_{i=1}^k \overline{Y}_i^{\star,n}. \tag{30}$$

We need to introduce some notations for the law of the variables we will consider in the construction of the coupling. We denote by ϕ the density of a standard Gaussian law, and by Φ its cumulative distribution function.

We let $\overline{p}_k^{\star,n}$ be the density function of $\frac{1}{\sqrt{k}}\overline{S}_k^{\star,n}$ and denote by F_k its cumulative distribution function. To simplify the notations, we have omitted the dependence upon n for F_k.

For $k = 2p$ an even integer in $\{2, \dots, n\}$, we define

$$\tilde{S}_k^{\star,n} = 2\overline{S}_{k/2}^{\star,n} - \overline{S}_k^{\star,n} = \sum_{i=1}^{k/2} \overline{Y}_i^{\star,n} - \sum_{i=k/2+1}^{k} \overline{Y}_i^{\star,n}.$$

We denote by $\tilde{p}_k^{\star,n}(\cdot \mid y)$ the conditional density of $\frac{1}{\sqrt{k}}\widetilde{S}_k^{\star,n}$ given $\frac{1}{\sqrt{k}}\overline{S}_k^{\star,n} = y$. The associated conditional cumulative distribution function is denoted by $\tilde{F}_k(\cdot \mid y) = \int_{-\infty}^{\cdot} \tilde{p}_k^{\star,n}(x \mid y)\mathrm{d}x$, where again we have suppressed the dependence upon n in the notation.

In the sequel we denote by F_k^{-1} the generalized inverse of the function F_k, and $\tilde{F}_k^{-1}(\cdot \mid y)$ the generalized inverse of the function $x \mapsto \tilde{F}_k(x \mid y)$.

4.1 The dyadic KMT construction

The construction of the coupling follows the dyadic construction scheme introduced by [10] p.116–118 (see also [6] p. 51–53) and we give it for the sake of completeness. We adopt the notation of [10]. In the sequel, it will be convenient to assume that n is a dyadic number, $n = 2^N$. Remark that if n is not a dyadic number, the construction below gives a coupling between the random variables (Y_k) and (X_k^n) for $k \in \{1, \dots, 2^N\}$ where $2^{N-1} < n \leq 2^N$, and the deviation bound (3) in Theorem 1 holds true for $\sqrt{n}\sup_{1\leq k\leq 2^N} \left|S_k - \overline{S}_k^n\right|$. Consequently we can assume that $n = 2^N$ without loss of generality.

Assume we are given a sequence of independent standard Gaussian variables $(Y_k)_{k\geq 1}$ on some probability space.

For k equal 1 to 2^N, we set :

$$S_k := \sum_{i=1}^{k} Y_i, \quad S_0 = 0,$$

and

$$V_{m,k} := S_{(k+1)2^m} - S_{k2^m}, \quad , 0 \leq k < 2^{N-m}, \quad 0 \leq m \leq N,$$

$$\tilde{V}_{m,k} := V_{m-1,2k} - V_{m-1,2k+1}, \quad , 0 \leq k < 2^{N-m}, \quad 1 \leq m \leq N.$$

Remark that we have $V_{m,0} = S_{2^m}$. Moreover, for $1 \leq m \leq N$, $0 \leq k < 2^{N-m}$, the Gaussian variables $V_{m,k}$ and $\tilde{V}_{m,k}$ are independent and $(\tilde{V}_{m,k})_{0\leq k<2^{N-m}}$ is an i.i.d. sequence for all m.

We now construct a sequence of independent identically distributed variables $(X_k^n)_{1\leq k\leq 2^N}$ with distribution defined by (2). We first construct some independent variables $(X_k^{\star,n})_{1\leq k\leq 2^N}$ equal in law to $(\overline{Y}_k^{\star,n})_{1\leq k\leq 2^N}$, defined by (29). The procedure consists in constructing by induction the sums of size

2^m, $(\sum_{i=k2^m+1}^{(k+1)2^m} X_i^{\star,n})_{k\in\{0,\dots,2^{N-m}-1\}}$, starting the construction with the level $m=N$ and eventually obtaining the variable $X_i^{\star,n}$ at level $m=0$.

First, for $m=N$, we set :

$$U_{N,0} := 2^{N/2} F_{2^N}^{-1}\left(\Phi\left(\frac{V_{N,0}}{2^{N/2}}\right)\right), \quad \tilde{U}_{N,0} := 2^{N/2}\tilde{F}_{2^N}^{-1}\left(\Phi\left(\frac{\tilde{V}_{N,0}}{2^{N/2}}\right)|\frac{1}{2^{N/2}}U_{N,0}\right). \tag{31}$$

We define then $U_{N-1,0}$ and $U_{N-1,1}$, by the relations :

$$U_{N-1,0} := \frac{1}{2}(U_{N,0}+\tilde{U}_{N,0}), \quad U_{N-1,1} := \frac{1}{2}(U_{N,0}-\tilde{U}_{N,0}).$$

By construction, $U_{N,0}$ is distributed as $\overline{S}_{2^N}^{\star,n}$ and using that $\tilde{V}_{N,0}$ is independent of $U_{N,0}$, it is easy to verify that $(U_{N,0},\tilde{U}_{N,0})$ is distributed as $(\overline{S}_{2^N}^{\star,n},\tilde{S}_{2^N}^{\star,n})$. Hence, the random variables $U_{N-1,0}$ and $U_{N-1,1}$ are independent and both distributed according to $\overline{S}_{2^{N-1}}^{\star,n}$.

We next continue the construction by induction on m. Assuming that $U_{m,k}$ is constructed, for $0\le k<2^{N-m}$, we set :

$$\tilde{U}_{m,k} := 2^{m/2}\tilde{F}_{2^m}^{-1}\left(\Phi\left(\frac{\tilde{V}_{m,k}}{2^{m/2}}\right)|\frac{1}{2^{m/2}}U_{m,k}\right), \tag{32}$$

and we define

$$U_{m-1,2k} := \frac{1}{2}(U_{m,k}+\tilde{U}_{m,k}), \quad U_{m-1,2k+1} := \frac{1}{2}(U_{m,k}-\tilde{U}_{m,k}). \tag{33}$$

We can observe that the joint distribution of the pair $(U_{m,k},\tilde{U}_{m,k})$ is the one of $(\overline{S}_{2^m}^{\star,n},\tilde{S}_{2^m}^{\star,n})$. In turn, $U_{m-1,2k}$ and $U_{m-1,2k+1}$ are independent and distributed according to $\overline{S}_{2^{m-1}}^{\star,n}$.

Moreover, it can be verified that, for any fixed $0\le m\le N$, the constructed random variables $U_{m,k}$, for $0\le k<2^{N-m}$, are independent.

At the final step, $m=0$, this permits to construct a sequence of independent random variables $(X_k^{\star,n})_{1\le k\le 2^N}$, equal in law to $(\overline{Y}_k^{\star,n})_{1\le k\le 2^N}$ by setting

$$X_k^{\star,n} := U_{0,k-1}, \quad 1\le k\le 2^N. \tag{34}$$

We end the construction with the variables $(X_k^n)_{1\le k\le 2^N}$, distributed according to the law of $(\overline{Y}_k^n)_{1\le k\le 2^N}$. Let $F^\star(.|y)$ be the cumulative density function of $\overline{Y}_k^n$ given $\overline{Y}_k^{\star,n}=y$. It is clear by (2) and (29) that $F^\star$ does not depend on k. We set :

$$X_k^n := (F^\star)^{-1}(\eta_k|X_k^{\star,n}), \quad 1\le k\le 2^N,$$

where $(\eta_k)_{k\ge1}$ is a sequence of independent random variables, uniformly distributed on $[0,1]$, and independent of $(Y_k)_{k\ge1}$. The pair $(X_k^n,X_k^{\star,n})$ has the distribution of $(\overline{Y}_k^n,\overline{Y}_k^{\star,n})$ and the difference $X_k^n-X_k^{\star,n}$ has the distribution of a centered Gaussian variable with variance $1/n^2$.

In that follows, we set

$$\overline{T}_k^{\star,n} := \sum_{i=1}^{k} X_i^{\star,n}, \quad \overline{T}_k^n := \sum_{i=1}^{k} X_i^n, \quad 1 \le k \le 2^N, \quad \overline{T}_0^{\star,n} := 0, \quad \overline{T}_0^n := 0,$$

where $X_i^{\star,n}$ and X_i^n are constructed above.

From (33) and (34), we easily deduce that,

$$U_{m,k} = \sum_{i=k2^m}^{(k+1)2^m-1} U_{0,i} = \sum_{i=k2^m+1}^{(k+1)2^m} X_i^{\star,n} = \overline{T}_{(k+1)2^m}^{\star,n} - \overline{T}_{k2^m}^{\star,n}, \quad , 0 \le k < 2^{N-m}, \quad 0 \le m \le N, \tag{35}$$

$$\tilde{U}_{m,k} = U_{m-1,2k} - U_{m-1,2k+1}, \quad , 0 \le k < 2^{N-m}, \quad 1 \le m \le N. \tag{36}$$

Moreover, from the dyadic construction, we have the following representation (see Lemma 5 in [6]), for $1 \le k \le 2^N$:

$$\overline{T}_k^{\star,n} = \frac{k}{2^N} U_{N,0} + \sum_{m=1}^{N} c_m \tilde{U}_{m,l(m,k)}, \quad S_k = \frac{k}{2^N} V_{N,0} + \sum_{m=1}^{N} c_m \tilde{V}_{m,l(m,k)}, \tag{37}$$

where $c_m \in [0,1]$ and $l(m,k)$ is defined by

$$l(m,k)2^m < k \le (l(m,k)+1)2^m.$$

Remark that (37) can be obtained as a consequence of the following decomposition on the Haar basis of $\ell^2(\{1,\dots,2^N\})$,

$$1_{\{1,\dots,k\}}(u) = \frac{k}{2^N} 1_{\{1,\dots,2^N\}}(u) + \sum_{m=1}^{N} c_m \psi_{m,l(m,k)}(u), \quad \forall u \in \{1\dots,2^N\},$$

where $\psi_{m,l} = 1_{\{l2^m+1,\dots,l2^m+2^{m-1}\}} - 1_{\{l2^m+2^{m-1}+1,\dots,(l+1)2^m\}}$.

In the next section, we will assess the probability of deviation between the random walks $(S_k)_k$ and $(\overline{T}_k^n)_k$. This crucially relies on the two following lemmas, which assess the difference between the random variables $U_{N,0}$, $V_{N,0}$ and $\tilde{U}_{m,k}$, $\tilde{V}_{m,k}$ related by (31)-(32).

Lemma 2 *There exist $\varepsilon > 0$ and $C > 0$ such that for n large enough, we have :*

$$|U_{N,0} - V_{N,0}| \le \frac{C}{\sqrt{n}} \left(\frac{|U_{N,0}|^2}{2^N} + 1 \right), \text{ if } |U_{N,0}| \le \varepsilon 2^N \sqrt{n}.$$

Lemma 3 *There exist $\varepsilon > 0$ and $C > 0$ such that, for all $m \in \{1,\dots,N\}$, $k \in \{0,\dots,2^{N-m}-1\}$ and n large enough :*

$$\left| \tilde{U}_{m,k} - \tilde{V}_{m,k} \right| \le \frac{C}{\sqrt{n}} \left(\frac{|U_{m,k}|^2 + |\tilde{U}_{m,k}|^2}{2^m} + 1 \right), \quad \text{if} \quad \max(|U_{m,k}|, |\tilde{U}_{m,k}|) \le \varepsilon 2^m \sqrt{n}.$$

The proof of these two lemmas are postponed to Sections 4.3–4.4.

4.2 Proof of Theorem 1

In this section, we prove that the control (3) holds true for the variables constructed in Section 4.1.

We first prove that for any positive constant K and λ, we have for n large enough and $x > 0$:

$$P(\sqrt{n} \sup_{1\leq k\leq n} |\overline{T}_k^{\star,n} - \overline{T}_k^n| \geq K \log n + x) \leq Ce^{-\lambda x}. \tag{38}$$

By construction $\sqrt{n}\sup_{1\leq k\leq n} |\overline{T}_k^{\star,n} - \overline{T}_k^n|$ is equal in law to $\frac{1}{\sqrt{n}} \sup_{1\leq k\leq n} |\sum_{i=1}^k \xi_i|$, where $(\xi_i)_i$ are i.i.d. standard gaussian variables. So we deduce, using the exponential inequality for the Brownian motion (see Proposition 1.8 in [15]),

$$P(\sqrt{n} \sup_{1\leq k\leq n} |\overline{T}_k^{\star,n} - \overline{T}_k^n| \geq K \log n + x) \leq 2e^{-\frac{1}{2}(K\log n+x)^2},$$

and (38) is proved. Consequently to prove Theorem 1, it is sufficient to prove that for some positive constant C, K and λ (independent of n and x), we have, for n large enough and for all $x > 0$:

$$P(\sqrt{n} \sup_{1\leq k\leq n} |\overline{T}_k^{\star,n} - S_k| \geq K \log n + x) \leq Ce^{-\lambda x},$$

or equivalently, there exist positive constants C, α and λ, such that for all x and N :

$$P(\sup_{1\leq k\leq n} |\overline{T}_k^{\star,n} - S_k| \geq x) \leq Ce^{\alpha N - \lambda 2^{N/2} x}. \tag{39}$$

To prove (39), we distinguish between the cases $x < 8\varepsilon 2^{N/2}$ and $x \geq 8\varepsilon 2^{N/2}$ where ε is deduced from Lemmas 2–3.

In that follows, we note

$$\Delta = \sup_{1\leq k\leq n} |\overline{T}_k^{\star,n} - S_k|. \tag{40}$$

First case : $x < 8\varepsilon 2^{N/2}$

We define the event

$$A = \bigcap_{k=0}^{2^N-1} \{|U_{0,k}| \leq \varepsilon 2^{N/2}\}, \tag{41}$$

which will be useful for the application of Lemmas 2 and 3.

To prove (39), we use the decomposition :

$$P(\Delta \geq x) \leq P(\{\Delta \geq x\} \cap A) + P(A^c).$$

We first control $P(A^c)$. By (41), we have for $t_0 > 0$:

$$\begin{aligned} P(A^c) \leq 2^N P(|\overline{X}_1^{\star,n}| > \varepsilon 2^{N/2}) &\leq 2^N (P(\overline{X}_1^{\star,n} > \varepsilon 2^{N/2}) + P(-\overline{X}_1^{\star,n} > \varepsilon 2^{N/2})) \\ &\leq 2^N ((E^{t_0 \overline{X}_1^{\star,n}} + E^{-t_0 \overline{X}_1^{\star,n}}) e^{-t_0 \varepsilon 2^{N/2}}). \end{aligned}$$

Since $\overline{X}_1^{\star,n}$ is equal in law to $\overline{Y}_1^{\star,n}$, we have $E[e^{t\overline{X}_1^{\star,n}}] = e^{\Psi(t)}$ with Ψ given by (54), and we deduce for $t > -2^{N/2}/2$:

$$E^{t\overline{X}_1^{\star,n}} \leq e^{\frac{3}{2}t^2 + \frac{t}{2^{2^{N/2}}}} \frac{1}{\sqrt{1 + t/2^{N/2}}},$$

so by choosing $t_0 = \varepsilon 2^{N/2}/2$, we obtain :

$$P(A^c) \leq C 2^N e^{-\varepsilon^2 2^N/8},$$

where C is a positive constant depending on ε. Since $x < 8\varepsilon 2^{N/2}$, we deduce the bound :

$$P(A^c) \leq C 2^N e^{-\varepsilon x 2^{N/2}/64}. \tag{42}$$

Turning to the control of $P(\{\Delta \geq x\} \cap A)$, we first have :

$$P(\{\Delta \geq x\} \cap A) \leq 2^N \sup_{1 \leq k \leq 2^N} P(\{|\overline{T}_k^{\star,n} - S_k| > x\} \cap A).$$

Now from (37), we deduce :

$$|\overline{T}_k^{\star,n} - S_k| \leq |U_{N,0} - V_{N,0}| + \sum_{m=1}^{N} |\tilde{U}_{m,l(m,k)} - \tilde{V}_{m,l(m,k)}|, \tag{43}$$

where $l(m,k)$ satisfies $l(m,k)2^m < k \leq (l(m,k)+1)2^m$.

But on A, we have for all m and for all k (this is immediate from the definition of A and (35)–(36)):

$$|U_{m,k}| \leq \varepsilon 2^m 2^{N/2}, \quad |\tilde{U}_{m,k}| \leq \varepsilon 2^m 2^{N/2},$$

consequently, using the results of Lemmas 2 and 3, we obtain on A :

$$|U_{N,0} - V_{N,0}| \leq \frac{C}{2^{N/2}}\left(\frac{|U_{N,0}|^2}{2^N} + 1\right),$$

$$|\tilde{U}_{m,l(m,k)} - \tilde{V}_{m,l(m,k)}| \leq \frac{C}{2^{N/2}}\left(\frac{|U_{m,l(m,k)}|^2 + |\tilde{U}_{m,l(m,k)}|^2}{2^m} + 1\right).$$

Plugging these bounds in (43) , this gives on A :

$$|\overline{T}_k^{\star,n} - S_k| \leq \frac{C}{2^{N/2}}\left(\frac{|U_{N,0}|^2}{2^N} + N + 1 + \delta_k\right). \tag{44}$$

where

$$\delta_k = \sum_{m=1}^{N} \frac{|U_{m,l(m,k)}|^2 + |\tilde{U}_{m,l(m,k)}|^2}{2^m}.$$

From Lemma 4 below, we observe that on A we have the equality in law :

$$\delta_k =_{\mathcal{L}} \delta_1 = \sum_{m=1}^{N} \frac{|U_{m,0}|^2 + |\tilde{U}_{m,0}|^2}{2^m},$$

and we can write :

$$P(\{\Delta \geq x\} \cap A) \leq 2^N P\left(\{\frac{C}{2^{N/2}}(\frac{|U_{N,0}|^2}{2^N} + N + 1 + \delta_1) \geq x\} \cap A\right).$$

Since $U_{m,0} = U_{m-1,0} + U_{m-1,1}$ and $\tilde{U}_{m,0} = U_{m-1,0} - U_{m-1,1}$, we remark that :

$$\delta_1 = 2\sum_{m=1}^{N} \frac{|U_{m-1,0}|^2 + |U_{m-1,1}|^2}{2^m} = \sum_{m=0}^{N-1} \frac{|U_{m,0}|^2 + |U_{m,1}|^2}{2^m}$$

and we obtain

$$P(\{\Delta \geq x\} \cap A) \leq 2^N P\left(\{\frac{C}{2^{N/2}}(\frac{|U_{N,0}|^2}{2^N} + N + 1 + \sum_{m=0}^{N-1} \frac{|U_{m,0}|^2 + |U_{m,1}|^2}{2^m}) \geq x\} \cap A\right).$$

Now, we get to bound $\sum_{m=0}^{N} \frac{|U_{m,0}|^2}{2^m}$ by $\sum_{m=0}^{N-1} \frac{|U_{m,1}|^2}{2^m}$. First we have, using (35),

$$|U_{m,0}|^2 = \left|U_{0,0} + \sum_{j=0}^{m-1} \sum_{i=2^j}^{2^{j+1}-1} U_{0,i}\right|^2 = (U_{0,0} + \sum_{j=0}^{m-1} U_{j,1})^2,$$

so by Cauchy-Schwarz inequality with $q = 1/\sqrt{2}$, we obtain :

$$|U_{m,0}|^2 \leq (\sum_{j\geq 0} q^j)(\frac{1}{q^m}U_{0,0}^2 + \sum_{j=0}^{m-1} \frac{1}{q^{m-j-1}}U_{j,1}^2) \leq C(\frac{1}{q^m}U_{0,0}^2 + \sum_{j=0}^{m-1} \frac{1}{q^{m-j-1}}U_{j,1}^2).$$

It yields after some calculus :

$$\sum_{m=0}^{N} \frac{|U_{m,0}|^2}{2^m} \leq C(U_{0,0}^2 + \sum_{m=0}^{N-1} \frac{U_{m,1}^2}{2^m}),$$

and finally :

$$P(\{\Delta \geq x\} \cap A) \leq 2^N P\left(\{\frac{C}{2^{N/2}}(U_{0,0}^2 + N + 1 + \sum_{m=0}^{N-1} \frac{U_{m,1}^2}{2^m}) \geq x\} \cap A\right). \tag{45}$$

To end the proof, we introduce the notation, for $0 \leq m \leq N-1$,

$$\tau_m = \frac{U_{m,1}^2}{2^m}\mathbf{1}_{\{|U_{m,1}| \leq \varepsilon 2^m 2^{N/2}\}}, \tag{46}$$

and

$$\overline{\tau}_0 = U_{0,0}^2\mathbf{1}_{\{|U_{0,0}| \leq \varepsilon 2^{N/2}\}}. \tag{47}$$

The random variables $(\overline{\tau}_0, \tau_0, \tau_1, \ldots, \tau_{N-1})$ are independent and this permits to deduce

$$\begin{aligned} P(\{\Delta \geq x\} \cap A) &\leq 2^N P\left(\tfrac{C}{2^{N/2}}(\overline{\tau}_0 + N + 1 + \textstyle\sum_{m=0}^{N-1} \tau_m) \geq x\right) \\ &\leq 2^N e^{Ct(N+1)} E e^{Ct\overline{\tau}_0} \textstyle\prod_{m=0}^{N-1} E e^{Ct\tau_m} e^{-t2^{N/2}x}, \end{aligned}$$

where the last inequality hods for any $t > 0$. Consequently to obtain the bound

$$P(\{\Delta \geq x\} \cap A) \leq C e^{\alpha N - \lambda 2^{N/2} x},$$

for some positive constants C, α and λ independent of N and x, it is sufficient to prove that

$$\exists t_0 > 0, \quad \exists C > 0, \quad \text{such that} \quad \forall 0 \leq m \leq N-1 \quad E e^{t_0 \tau_m} \leq C. \tag{48}$$

Integrating by parts, we first remark that :

$$E e^{t_0 \tau_m} = 1 + \int_0^{\varepsilon^2 2^m 2^N} t_0 e^{t_0 y} P(\tau_m > y) dy,$$

so to prove (48) we just have to prove that for $0 \leq y \leq \varepsilon^2 2^m 2^N$:

$$P(\tau_m > y) \leq C e^{-\eta y}, \tag{49}$$

for positive constant C and η (independent of m, y, N). We have

$$\begin{aligned} P(\tau_m > y) &\leq P(U_{m,1}^2 > 2^m y) \\ &\leq P(U_{m,1} > 2^{m/2}\sqrt{y}) + P(-U_{m,1} > 2^{m/2}\sqrt{y}), \\ &\leq E e^{tU_{m,1}} e^{-t2^{m/2}\sqrt{y}} + E e^{-tU_{m,1}} e^{-t2^{m/2}\sqrt{y}}, \end{aligned}$$

where the last inequality holds for any $t > 0$. We recall that $U_{m,1}$ is equal in law to $\sum_{i=1}^{2^m} \overline{Y}_i^{\star,n}$, this gives $E e^{tU_{m,1}} = e^{2^m \Psi(t)}$, where Ψ is defined by (54). Moreover, using the notation

$$\Lambda^*(u) = \inf_t (\Psi(t) - tu), \tag{50}$$

one has (since $\Psi^{(1)}(0) = 0$) : for $u \geq 0$, $\Lambda^*(u) = \inf_{t \geq 0}(\Psi(t) - tu)$ and $\Lambda^*(-u) = \inf_{t \leq 0}(\Psi(t) + tu)$. This gives :

$$P(\tau_m > y) \leq e^{2^m \Lambda^*(2^{-m/2}\sqrt{y})} + e^{2^m \Lambda^*(-2^{-m/2}\sqrt{y})}.$$

From the estimation (61) (where t is defined by (59) with $k = 2^m$ and $x = \sqrt{y}$), we deduce, since $\sqrt{y} \leq \varepsilon 2^{m/2} 2^{N/2}$, that (choosing ε small enough and N large enough),

$$2^m \Lambda^*(2^{-m/2}\sqrt{y}) \leq -\eta y \quad \text{and} \quad 2^m \Lambda^*(2^{-m/2}\sqrt{y}) \leq -\eta y,$$

for some $\eta > 0$ and (49) is proved. This achieves the proof of (39) in the first case.

Second case : $x \geq 8\varepsilon 2^{N/2}$

Following [10] and [6], we first choose an integer M such that

$$\frac{x}{8\varepsilon} < 2^M 2^{N/2} \leq \frac{x}{4\varepsilon}.$$

Such an integer exists since $x \geq 8\varepsilon 2^{N/2}$. Without loss of generality, we may assume that $M \leq N$ (if it is not the case the proof reduces to consider Δ_1 and Δ_2 below with $M = N$). We set

$$B = \bigcap_{k=0}^{2^{N-M}-1} \{|U_{M,k}| \leq \varepsilon 2^M 2^{N/2}\}. \tag{51}$$

We remark that on B, we have for all $m \geq M+1$ and for all k :

$$|U_{m,k}| \leq \varepsilon 2^m 2^{N/2}, \quad |\tilde{U}_{m,k}| \leq \varepsilon 2^m 2^{N/2}.$$

Moreover we define :

$$\Delta_1 = \max_{0 \leq k \leq 2^{N-M}} \max_{1 \leq l \leq 2^M} |\overline{T}^{\star,n}_{k2^M+l} - \overline{T}^{\star,n}_{k2^M}|,$$

$$\Delta_2 = \max_{0 \leq k \leq 2^{N-M}} \max_{1 \leq l \leq 2^M} |S_{k2^M+l} - S_{k2^M}|,$$

$$\Delta_3 = \max_{0 \leq k \leq 2^{N-M}} |\overline{T}^{\star,n}_{k2^M} - S_{k2^M}|.$$

We immediately see that

$$\Delta \leq \Delta_1 + \Delta_2 + \Delta_3,$$

where Δ is defined by (40). Moreover, observing that $B^c \subset \{\Delta_1 \geq \varepsilon 2^M 2^{N/2}\}$, we have :

$$\{\Delta \geq x\} \subset \{\Delta_1 \geq \varepsilon 2^M 2^{N/2}\} \cup \{\Delta_2 \geq \varepsilon 2^M 2^{N/2}\} \cup (\{\Delta_3 \geq x/2\} \cap B),$$

and so

$$P(\Delta \geq x) \leq P(\Delta_1 \geq \varepsilon 2^M 2^{N/2}) + P(\Delta_2 \geq \varepsilon 2^M 2^{N/2}) + P(\Delta_3 \geq x/2\} \cap B). \tag{52}$$

We first bound $P(\Delta_3 \geq x/2\} \cap B)$. Starting with the decomposition (similar to (37))

$$\overline{T}^{\star,n}_{k2^M} = \frac{k}{2^N} U_{N,0} + \sum_{m=M+1}^{N} c_m \tilde{U}_{m,l(m,k2^M)},$$

and proceeding as in the proof of the first case, one can show that

$$P(\Delta_3 \geq x/2\} \cap B) \leq 2^N P(\frac{C}{2^N}(N+1+\sum_{m=M}^{N-1} \tau_m) \geq x),$$

where the variables τ_m are defined by (46). This permits to conclude, using the same arguments as previously, that

$$P(\Delta_3 \geq x/2\} \cap B) \leq Ce^{\alpha N - \lambda 2^{N/2} x}.$$

Turning to $P(\Delta_1 \geq \varepsilon 2^M 2^{N/2})$, we have :

$$P(\Delta_1 \geq \varepsilon 2^M 2^{N/2}) \leq 2^N P(\max_{1 \leq k \leq 2^M} |\overline{T}_k^{\star,n}| \geq \varepsilon 2^M 2^{N/2})$$

Applying Doob's maximal inequality to the positive submartingales $(e^{t\overline{T}_k^{\star,n}})$ and $(e^{-t\overline{T}_k^{\star,n}})$, where $t > 0$, and optimizing on t, we deduce using the notation (50):

$$P(\Delta_1 \geq \varepsilon 2^M 2^{N/2}) \leq 2^N (e^{2^M \Lambda^*(\varepsilon 2^{N/2})} + e^{2^M \Lambda^*(-\varepsilon 2^{N/2})})$$

and from (61)

$$P(\Delta_1 \geq \varepsilon 2^M 2^{N/2}) \leq 2^{N+1} e^{-\eta \varepsilon^2 2^N 2^M} \leq 2^{N+1} e^{-\eta \varepsilon x 2^{N/2}/8},$$

for some $\eta > 0$. Similarly for the standard gaussian variables, we have :

$$P(\max_{1 \leq k \leq 2^M} |S_k| \geq \varepsilon 2^M 2^{N/2}) \leq 2e^{-\varepsilon^2 2^M 2^N/2},$$

and so

$$P(\Delta_2 \geq \varepsilon 2^M 2^{N/2}) \leq 2^{N+1} e^{-\varepsilon x 2^{N/2}/16}.$$

This achieves the proof of (3) and hence of Theorem 1. ◇

Lemma 4 *We have the equality in law, for all $k \in \{1, \ldots, 2^N\}$,*

$$\left(1_A, (|U_{m,l(m,k)}|)_{m=0,\ldots,N}, (|\tilde{U}_{m,l(m,k)}|)_{m=1,\ldots,N}\right) =_{\mathcal{L}} \left(1_A, (|U_{m,0}|)_{m=0,\ldots,N}, (|\tilde{U}_{m,0}|)_{m=1,\ldots,N}\right). \quad (53)$$

Proof The proof is based on the fact that the law of the vector $(U_{0,i})_i$ is invariant by permutation, and that one can find permutations that transform the left hand side of (53) into the right hand side. Let us sketch how to construct these permutations.

If $k = 1$, then $l(m,k) = 0$ for all m, and the result (53) is immediate. Otherwise, we write $k - 1 = \sum_{m=0}^{N-1} d_m 2^m$, where $d_m \in \{0,1\}$. Let us remark that we have $l(m,k) = \sum_{j=m}^{N-1} d_j 2^{j-m}$ for all $m \in \{0, \ldots, N-1\}$. Let us denote $m_0 = \sup\{m \mid d_m = 1, 0 \leq m \leq N-1\}$ which is well defined since

$k-1\neq 0$. We define $(U'_{0,i})_i$ a permutation of the random vector $(U_{0,i})_i$ as follows :

$$U'_{0,i} := \begin{cases} U_{0,i+2^{m_0}} & \text{if } 0 \le i < 2^{m_0}, \\ U_{0,i-2^{m_0}} & \text{if } 2^{m_0} \le i < 2^{m_0+1}, \\ U_{0,i} & \text{if } 2^{m_0+1} \le i < 2^N - 1, \end{cases}$$

and set $U'_{m,l} = \sum_{i=l2^m}^{l2^m+2^m-1} U'_{0,i}$, $\tilde{U}'_{m,l} = U'_{m-1,2l} - U'_{m-1,2l+1}$ for $1 \le m \le N$ and $0 \le l \le 2^{N-m}-1$. We set $k' = k - 2^{m_0}$. Then, one can check that, $l(m,k') = l(m,k) = 0$ for $m > m_0$ and $l(m,k') = l(m,k) - 2^{m_0-m}$ for $0 \le m \le m_0$. Moreover, we easily get $U'_{m,l} = U_{m,l}$ and $\tilde{U}'_{m,l} = \tilde{U}_{m,l}$ for all $m > m_0+1$, $0 \le l \le 2^{N-m}-1$. By construction, we have $U'_{m_0+1,l(m_0+1,k')} = U'_{m_0+1,0} = U_{m_0+1,0} = U_{m_0+1,l(m_0+1,k)}$ and $\tilde{U}'_{m_0+1,l(m+1,k')} = \tilde{U}'_{m_0+1,0} = -\tilde{U}_{m_0+1,0} = -\tilde{U}_{m_0+1,l(m_0+1,k)}$. For $1 \le m \le m_0$, we can write

$$\begin{aligned} U'_{m,l(m,k')} &= \sum_{i=l(m,k')2^m}^{l(m,k')2^m+2^m-1} U'_{0,i} = \sum_{i=l(m,k)2^m-2^{m_0}}^{l(m,k)2^m-2^{m_0}+2^m-1} U'_{0,i} = \sum_{i=l(m,k)2^m}^{l(m,k)2^m+2^m-1} U'_{0,i-2^{m_0}} \\ &= \sum_{i=l(m,k)2^m}^{l(m,k)2^m+2^m-1} U_{0,i} = U_{m,l(m,k)}, \text{ since } l(m,k)2^m \in \{2^{m_0}, \dots, 2^{m_0+1}-1\} . \end{aligned}$$

Using that the set A is invariant by permutation of the $U_{0,i}$ (see (41)), we deduce from the discussion above that

$$\left(1_A, (|U_{m,l(m,k)}|)_{m=0,\dots,N}, (|\tilde{U}_{m,l(m,k)}|)_{m=1,\dots,N}\right) =_{\mathcal{L}} \left(1_A, (|U_{m,l(m,k')}|)_{m=0,\dots,N}, (|\tilde{U}_{m,l(m,k')}|)_{m=1,\dots,N}\right),$$

where $k' < k$. Hence, for $k \neq 1$, this shows that we can replace k by $k' < k$ in the left hand side of (53) without changing its law. By a finite number of iterations of the procedure we deduce (53). $\diamond$

4.3 Quantile coupling inequalities

In this section, we prove Lemma 2. We first establish a sharp expansion for the law of $\frac{1}{\sqrt{k}}\overline{S}_k^{\star,n}$.

Lemma 5 *Let $\overline{p}_k^{\star,n}$ be the density function of $\frac{1}{\sqrt{k}}\overline{S}_k^{\star,n}$, ϕ and Φ be respectively the density and the cumulative distribution function of the standard gaussian law. There exist some constants $\varepsilon > 0$ and $C > 0$ such that for all $k \ge 1$ and n large enough, we have :*

i) for $|x| \le \varepsilon\sqrt{kn}$:

$$\overline{p}_k^{\star,n}(x) = \phi(x) e^{\frac{1}{\sqrt{n}} T_k^1(x)},$$

ii) for $0 \le x \le \varepsilon\sqrt{kn}$:

$$P(\frac{1}{\sqrt{k}}\overline{S}_k^{\star,n} > x) = (1-\Phi(x)) e^{\frac{1}{\sqrt{n}} T_k^2(x)},$$

$$P(\frac{1}{\sqrt{k}}\overline{S}_k^{\star,n} \leq -x) = \Phi(-x)e^{\frac{1}{\sqrt{n}}T_k^3(x)},$$

where $|T_k^j(x)| \leq C\frac{(1+|x|^3)}{\sqrt{k}}$, *for* $1 \leq j \leq 3$.

Remark 2 *For* $k = n$, *the approximation of* $\overline{p}_k^{\star,n}(x)$ *by* $\phi(x)$ *is of order* $1/n$, *for* $|x| \leq \varepsilon n$. *In the classical KMT result, the order of approximation is* $1/\sqrt{n}$ *for* $|x| \leq \varepsilon\sqrt{n}$. *It is important to have a better approximation result which holds for larger values of* x *to improve the final bound in the KMT construction (compare for example to the refined quantile inequalities given in [13]).*

An inspection of the proof below shows that without the regularization technique (i.e. without adding the small Gaussian variables ξ_k in (29)), the result of Lemma 5 still holds but with $k \geq 3$ only.

Proof of Lemma 5

We can prove i) and ii) by the technique of conjugated random variables (see [6],[11], [12]) . We only give the proof of i) in details, the proof of ii) being very similar (see for example [12]). We first compute the Laplace transform of the variables $\overline{Y}_k^{\star,n}$. For $t > -\sqrt{n}$, let $R(t) = Ee^{t\overline{Y}_k^{\star,n}}$ and $\Psi(t) = \log R(t)$. A simple computation gives :

$$\Psi(t) = \frac{t}{2\sqrt{n}} + \frac{t^2}{2(1+t/\sqrt{n})} - \frac{1}{2}\log(1+t/\sqrt{n}) + \frac{t^2}{2n^2}. \tag{54}$$

In particular, we have $\Psi(0) = 0$, $\Psi^{(1)}(0) = 0$, $\Psi^{(2)}(0) = 1 + \frac{1}{2n} + \frac{1}{n^2}$ and it is easy to verify from the computation of $\Psi^{(2)}$ and $\Psi^{(3)}$, that, for $0 \leq |t| \leq c\sqrt{n}$, $0 < 1/C \leq \Psi^{(2)}(t) \leq C$ and $|\Psi^{(3)}(t)| \leq C/\sqrt{n}$, for some positive constants c and C. Fixing t such that $|t| \leq c\sqrt{n}$, we consider the sequence of independent random variables $(Z_k^n)_{k\geq 1}$ such that, Z_k^n admits the density function $\frac{e^{tx}}{R(t)}f_{\overline{Y}_k^{\star,n}}(x)$, where $f_{\overline{Y}_k^{\star,n}}$ denotes the density function of $\overline{Y}_k^{\star,n}$. One can easily verify that $E(Z_k^n) = \Psi^{(1)}(t)$ and $V(Z_k^n) = \Psi^{(2)}(t)$. We denote by q_k^n the density function of the normalized sum $\frac{1}{\sqrt{k\Psi^{(2)}(t)}}\sum_{i=1}^k (Z_i^n - \Psi^{(1)}(t))$. The following relation holds between $\overline{p}_k^{\star,n}$ and q_k^n:

$$\overline{p}_k^{\star,n}(x) = \frac{e^{k\Psi(t)-tx\sqrt{k}}}{\sqrt{\Psi^{(2)}(t)}} q_k^n\left(\frac{x\sqrt{k} - k\Psi^{(1)}(t)}{\sqrt{k\Psi^{(2)}(t)}}\right). \tag{55}$$

The next step to obtain the result of Lemma 5 is to prove that that for $|t| \leq c\sqrt{n}$

$$\sup_x |q_k^n(x) - \phi(x)| \leq C\frac{1}{\sqrt{kn}}. \tag{56}$$

Let $\hat{q}_k^n(u)$ be the Fourier transform of q_k^n, we have $\hat{q}_k^n(u) = \left(Ee^{\frac{iu}{\sqrt{k\Psi^{(2)}(t)}}(Z_1^n - \Psi^{(1)}(t))} \right)^k$ and consequently

$$\hat{q}_k^n(u) = \left(\frac{R\left(t + \frac{iu}{\sqrt{k\Psi^{(2)}(t)}}\right)}{R(t)} \right)^k e^{-\frac{iu\Psi^{(1)}(t)}{\sqrt{\Psi^{(2)}(t)}}\sqrt{k}}. \tag{57}$$

Now, by the Fourier inversion formula,

$$\sup_x |q_k^n(x) - \phi(x)| \leq \frac{1}{2\pi} \int |\hat{q}_k^n(u) - e^{-\frac{u^2}{2}}| du \leq I_k^{n,0} + I_k^{n,1} + I_k^{n,2},$$

where

$$I_k^{n,0} = C \int_{|u| \leq \alpha\sqrt{kn}} |\hat{q}_k^n(u) - e^{-\frac{u^2}{2}}| du,$$

$$I_k^{n,1} = C \int_{|u| > \alpha\sqrt{kn}} |\hat{q}_k^n(u)| du,$$

$$I_k^{n,2} = C \int_{|u| > \alpha\sqrt{kn}} e^{-\frac{u^2}{2}} du,$$

and α is a positive constant which will be precised below.

Since for $x > 0$, $\int_x^\infty e^{-\frac{u^2}{2}} du \leq \frac{1}{x} e^{-\frac{x^2}{2}}$, one can easily see that

$$I_k^{n,2} \leq \frac{C}{\sqrt{kn}}.$$

Turning back to $I_k^{n,1}$, a tedious computation using (54) and (57) gives

$$|\hat{q}_k^n(u)| = \frac{e^{-\frac{u^2}{2n^2\Psi^{(2)}(t)}}}{\left(1 + \frac{u^2}{kn\Psi^{(2)}(t)(1+t/\sqrt{n})^2}\right)^{k/4}} \times \frac{e^{\frac{kt^2(1+t/\sqrt{n})}{2((1+t/\sqrt{n})^2 + \frac{u^2}{kn\Psi^{(2)}(t)})}}}{e^{\frac{kt^2}{2(1+t/\sqrt{n})}}} g(u),$$

where

$$\log g(u) = \frac{-u^2(1 - t/\sqrt{n})}{2\Psi^{(2)}(t)\left((1 + t/\sqrt{n})^2 + \frac{u^2}{kn\Psi^{(2)}(t)}\right)}.$$

We deduce then the bound

$$|\hat{q}_k^n(u)| \leq \frac{e^{-\frac{u^2}{2n^2\Psi^{(2)}(t)}}}{\left(1 + \frac{u^2}{kn\Psi^{(2)}(t)(1+t/\sqrt{n})^2}\right)^{k/4}} g(u). \tag{58}$$

We remark that the contribution of the regularization variables (ξ_k) in the numerator of equation (58) (the term $e^{-\frac{u^2}{2n^2\Psi^{(2)}(t)}}$) ensures the integrability of $|\hat{q}_k^n(u)|$ for the small values of k ($k = 1$ and $k = 2$).

Recalling that if $|t| \leq c\sqrt{n}$, we have $0 < 1/C \leq \Psi^{(2)}(t) \leq C$, we deduce that for $|u| > \alpha\sqrt{kn}$, $0 \leq g(u) \leq e^{-Ckn}$. This finally yields :

$$I_k^{n,1} \leq Ce^{-Ckn} \int e^{-C\frac{u^2}{2n^2}} du \leq Cne^{-Ckn} \leq \frac{C}{\sqrt{kn}}.$$

It remains to bound the main term $I_k^{n,0}$. With the previous notations, we rewrite (57) as :

$$\hat{q}_k^n(u) = e^{k\Psi\left(t+\frac{iu}{\sqrt{k\Psi^{(2)}(t)}}\right)-k\Psi(t)-\frac{iu\Psi^{(1)}(t)}{\sqrt{\Psi^{(2)}(t)}}\sqrt{k}}.$$

A Taylor expansion up to order three of $v \mapsto \Psi\left(t + \frac{iv}{\sqrt{k\Psi^{(2)}(t)}}\right)$ on $[0, u]$ gives :

$$\hat{q}_k^n(u) = e^{-\frac{u^2}{2}-\frac{iu^3}{6\Psi^{(2)}(t)\sqrt{k\Psi^{(2)}(t)}}\eta_u},$$

where $|\eta_u| \leq \sup_{|y|\leq|u|} |\Psi^{(3)}(t + \frac{iy}{\sqrt{k\Psi^{(2)}(t)}})|$. We deduce then that :

$$|\hat{q}_k^n(u) - e^{-\frac{u^2}{2}}| = e^{-\frac{u^2}{2}}|1 - e^{-\frac{iu^3}{6\Psi^{(2)}(t)\sqrt{k\Psi^{(2)}(t)}}\eta_u}|.$$

Using the inequality $|1 - e^z| \leq |z|e^{|z|}$ for any complex number z, we obtain by choosing α such that $|\frac{u}{\Psi^{(2)}(t)\sqrt{k\Psi^{(2)}(t)}}\eta_u| \leq 1$, for $|u| \leq \alpha\sqrt{kn}$,

$$|\hat{q}_k^n(u) - e^{-\frac{u^2}{2}}| \leq C\frac{|u|^3}{\sqrt{kn}}e^{-\frac{u^2}{2}}e^{\frac{u^2}{6}} = C\frac{|u|^3}{\sqrt{kn}}e^{-\frac{u^2}{3}}.$$

This gives

$$I_k^{n,0} \leq \frac{C}{\sqrt{kn}}.$$

This achieves the proof of (56).

We turn back to (55). We first recall that for $|t| \leq c\sqrt{n}$, $0 \leq 1/C \leq \Psi^{(2)}(t) \leq C$. As a consequence, $\Psi^{(1)}$ is increasing, for $|t| \leq c\sqrt{n}$, with values in $[-C\sqrt{n}, C\sqrt{n}]$, for some constant C. It follows that for $|x| \leq \varepsilon\sqrt{kn}$, the equation $x = \Psi^{(1)}(t)\sqrt{k}$ admits a unique solution. In the sequel, we fix t to be the unique solution of :

$$x = \Psi^{(1)}(t)\sqrt{k}. \tag{59}$$

We have $|t| \leq c\sqrt{n}$ and so combining (55) with (56), we obtain :

$$\overline{p}_k^{\star,n}(x) = \frac{e^{k\Psi(t)-tx\sqrt{k}}}{\sqrt{\Psi^{(2)}(t)}}\frac{1}{\sqrt{2\pi}}\left(1 + \frac{O(1)}{\sqrt{kn}}\right). \tag{60}$$

where $O(1)$ is a bounded function. Next considering (59), by a Taylor expansion of $[\Psi^{(1)}]^{-1}$ on $[0, x/\sqrt{k}]$ up to order two, we obtain :

$$t = \frac{x}{\sqrt{k}\Psi^{(2)}(0)} + \frac{x^2}{k\sqrt{n}}O(1),$$

where we have used $[\Psi^{(1)}]^{-1}(0) = 0$, $([\Psi^{(1)}]^{-1})^{(1)}(0) = 1/\Psi^{(2)}(0)$, $([\Psi^{(1)}]^{-1})^{(2)}(u) = \frac{\psi^{(3)}([\Psi^{(1)}]^{-1}(u))}{[\Psi^{(2)}([\Psi^{(1)}]^{-1}(u))]^2}$ and recalling that for $|u| \leq c\sqrt{n}$, we have $0 \leq 1/C \leq \Psi^{(2)}(u) \leq C$ and $|\Psi^{(3)}(u)| \leq C/\sqrt{n}$.

Now, since $\Psi(t) = \frac{t^2}{2}\Psi^{(2)}(0) + \frac{t^3}{6}\eta_t$, with $|\eta_t| \leq \sup_{|u|\leq|t|} |\Psi^{(3)}(u)|$, we deduce the expansion :

$$k\Psi(t) = \frac{x^2}{2\Psi^{(2)}(0)} + \frac{x^3}{\sqrt{kn}}O(1),$$

where $O(1)$ is a function which is bounded uniformly in k and n, for $|x| \leq \varepsilon\sqrt{kn}$. Using $\Psi^{(2)}(0) = 1 + \frac{1}{2n} + \frac{1}{n^2}$, this finally leads to :

$$k\Psi(t) - t\sqrt{k}x = -\frac{x^2}{2} + \frac{x^2}{n}O(1) + \frac{x^3}{\sqrt{kn}}O(1). \tag{61}$$

Reporting (61) in (60), it yields :

$$\overline{p}_k^{\star,n}(x) = \phi(x)\frac{e^{\frac{x^2}{n}O(1)+\frac{x^3}{\sqrt{kn}}O(1)}}{\sqrt{\Psi^{(2)}(t)}}\left(1 + \frac{O(1)}{\sqrt{kn}}\right).$$

We conclude, observing that $\Psi^{(2)}(t) = \Psi^{(2)}(0) + \frac{x}{\sqrt{kn}}O(1)$,

$$\overline{p}_k^{\star,n}(x) = \phi(x)e^{O(1)\frac{(1+x^3)}{\sqrt{kn}}},$$

and i) is proved. ◇

Based on the refined quantile inequalities for the law of $\frac{1}{\sqrt{k}}\overline{S}^{\star,n}$ stated in Lemma 5 ii), we deduce the result of Lemma 2.

Proof of Lemma 2

The result of Lemma 2 is a consequence of the more general following result, applied to the particular case $m = N$.

There exist $\varepsilon > 0$ and $C > 0$ such that, for all $m \in \{0, \ldots, N\}$ and $n = 2^N$ large enough, we have:

$$\left|2^{m/2}F_{2^m}^{-1}\left(\Phi\left(\frac{V_{m,0}}{2^{m/2}}\right)\right) - V_{m,0}\right| \leq \frac{C}{\sqrt{n}}\left(\left|F_{2^m}^{-1}\left(\Phi\left(\frac{V_{m,0}}{2^{m/2}}\right)\right)\right|^2 + 1\right),$$

if $\left|2^{m/2}F_{2^m}^{-1}\left(\Phi\left(\frac{V_{m,0}}{2^{m/2}}\right)\right)\right| \leq \varepsilon 2^m\sqrt{n}$.

To get this result, we have to prove that for $0 \leq |x| \leq \varepsilon\sqrt{kn}$,

$$\Phi(x - u(x)) \leq F_k(x) \leq \Phi(x + u(x)), \tag{62}$$

where $u(x) = \frac{C}{\sqrt{n}}\left(\frac{x^2}{\sqrt{k}} + \frac{1}{\sqrt{k}}\right)$. Indeed, observing that (62) is equivalent to $x - u(x) \leq \Phi^{-1}(F_k(x)) \leq x + u(x)$, we obtain the above result for $x = F_k^{-1}(\Phi(\frac{1}{\sqrt{k}}V_{m,0}))$ and $k = 2^m$.

From Lemma 5, part ii), we have for $0 \leq x \leq \varepsilon\sqrt{kn}$:

$$(1 - \Phi(x))e^{-C\frac{(1+x^3)}{\sqrt{kn}}} \leq 1 - F_k(x) \leq (1 - \Phi(x))e^{C\frac{(1+x^3)}{\sqrt{kn}}}, \tag{63}$$

$$\Phi(-x)e^{-C\frac{(1+x^3)}{\sqrt{kn}}} \leq F_k(-x) \leq \Phi(-x)e^{C\frac{(1+x^3)}{\sqrt{kn}}}. \tag{64}$$

Now from Mason and Zhou [13], Lemma 3, we have for all $A > 0$, $k \geq 64A^2$ and $0 \leq x \leq \sqrt{k}/(8A)$:

$$\log\left(\frac{\Phi(-x+u)}{\Phi(-x)}\right) = \log\left(\frac{1 - \Phi(x-u)}{1 - \Phi(x)}\right) \geq A(\frac{1+x^3}{\sqrt{k}}), \tag{65}$$

$$\log\left(\frac{\Phi(-x-u)}{\Phi(-x)}\right) = \log\left(\frac{1 - \Phi(x+u)}{1 - \Phi(x)}\right) \leq -A(\frac{1+x^3}{\sqrt{k}}), \tag{66}$$

where $u = 2A\frac{1+x^2}{\sqrt{k}}$. Combining (65) and (66) with $A = C/\sqrt{n}$ and n large enough, with (63) and (64), we deduce that $\forall k \geq 1$ and $0 \leq |x| \leq \varepsilon\sqrt{kn}$, (62) holds. ⋄

4.4 Conditional quantile inequalities

In this section we prove Lemma 3. Recall that, for k even, $\tilde{p}_k^{\star,n}(\cdot \mid y)$ is the conditional density of $\frac{1}{\sqrt{k}}\tilde{S}_k^{\star,n}$ given $\frac{1}{\sqrt{k}}\overline{S}_k^{\star,n} = y$, where the joint law of $(\tilde{S}_k^{\star,n}, \overline{S}_k^{\star,n}) = (2\overline{S}_{k/2}^{\star,n} - \overline{S}_k^{\star,n}, \overline{S}_k^{\star,n})$ is defined via (29)–(30). In the following three lemmas, we establish some expansions for the conditional density $\tilde{p}_k^{\star,n}(\cdot \mid y)$ and the associated conditional quantile inequalities. Then, we will deduce Lemma 3.

Lemma 6 *There exist some constants $\varepsilon > 0$ and $C > 0$ such that for all $k \geq 1$ and n large enough, we have, for $|x| \leq \varepsilon\sqrt{kn}, |y| \leq \varepsilon\sqrt{kn}$:*

$$\tilde{p}_k^{\star,n}(x \mid y) = \phi(x)e^{\frac{1}{\sqrt{n}}\tilde{T}_k^1(x,y)},$$

where $\left|\tilde{T}_k^1(x,y)\right| \leq C\frac{1+|x|^3+|y|x^2+|y|}{\sqrt{k}}$.

Proof First we show the following expansion for the density of $\frac{1}{\sqrt{k}}\overline{S}_k^{\star,n}$,

$$\exists\varepsilon > 0, \forall\, |x| \leq \varepsilon\sqrt{nk}, \quad \overline{p}_k^{\star,n}(x) = \phi(x)\exp\left(\frac{x^3}{\sqrt{nk}}r_n(\frac{x}{\sqrt{nk}}) + c_n x^2 + \frac{1+|x|}{\sqrt{nk}}B_{n,k}(x)\right), \tag{67}$$

where r_n is a sequence of smooth functions defined on some neighbourhood $[-\varepsilon,\varepsilon]$ of 0 and whose derivatives up to order two are bounded independently of n; c_n is a sequence such that $c_n = O(1/n)$; and $B_{n,k}(\cdot)$ is some measurable function bounded independently of k and n.

Recalling the representation (60), where t is the unique solution of (59), $\Psi^{(2)}(t) = \Psi^{(2)}(0) + \frac{|x|}{\sqrt{kn}}O(1)$, and $\Psi^{(2)}(0) = 1 + \frac{1}{2n} + \frac{1}{n^2}$ yields to the representation

$$\overline{p}_k^{\star,n}(x) = \frac{1}{\sqrt{2\pi}} \exp\left(k\Psi(t) - tx\sqrt{k} + \frac{1+|x|}{\sqrt{nk}}O(1)\right)$$

Let us denote $\Phi_n(s) = \frac{1}{n}\Psi(\sqrt{n}s)$ and by h_n the inverse of the function $s \mapsto \frac{1}{\sqrt{n}}\Psi^{(1)}(\sqrt{n}s)$. Due to the expression (54), it is simple to check that both functions are well defined on some neighborhoods of 0 independent of n, and we can assume that h_n is well defined on the interval $[-\varepsilon,\varepsilon]$, up to reducing the value of ε. Using these notations, we get $t = \sqrt{n}h_n(\frac{x}{\sqrt{nk}})$, and in turn, $\overline{p}_k^{\star,n}(x) = \frac{1}{\sqrt{2\pi}} \exp\Big(nk\Phi_n(h_n(\frac{x}{\sqrt{nk}})) - \sqrt{nk}h_n(\frac{x}{\sqrt{nk}})x + \frac{1+|x|}{\sqrt{nk}}O(1)\Big)$. Now, since $\Phi_n(0) = \Phi_n^{(1)}(0) = 0$, $\Phi_n^{(2)}(0) = 1 + \frac{1}{2n} + \frac{1}{n^2}$, we can write $\Phi_n(s) = \Phi_n^{(2)}(0)\frac{s^2}{2} + s^3\gamma_n(s) = (1/2 + O(1/n))s^2 + s^3\gamma_n(s)$, where γ_n is some function. Using that Φ_n, and all its derivatives, are bounded independently of n on $[-\varepsilon,\varepsilon]$, we deduce that the same property holds true for γ_n. Analogously, we can show that

$$h_n(s) = h_n^{(1)}(0)s + s^2\beta_n(s) = (1 + O(1/n))s + s^2\beta_n(s) \tag{68}$$

where the function β_n, and its derivatives of any order, are bounded independently of n on $[-\varepsilon,\varepsilon]$. With simple computation, we deduce that

$$\begin{aligned} nk\Phi_n(h_n(\frac{x}{\sqrt{nk}})) - \sqrt{nk}h_n(\frac{x}{\sqrt{nk}})x &= nk\Phi_n\left((1+O(1/n))\frac{x}{\sqrt{nk}} + \frac{x^2}{nk}\beta_n(\frac{x}{\sqrt{nk}})\right) \\ &\quad - \sqrt{nk}x\left((1+O(1/n))\frac{x}{\sqrt{nk}} + \frac{x^2}{nk}\beta_n(\frac{x}{\sqrt{nk}})\right) \\ &= -\frac{1}{2}x^2 + x^2O(1/n) + \frac{x^3}{\sqrt{nk}}r_n(\frac{x}{\sqrt{nk}}) \end{aligned}$$

where r_n is some bounded function, with bounded derivatives. This gives (67).

Now, using the independence of the random variables $\overline{S}_k^{\star,n} + \tilde{S}_k^{\star,n}$ and $\overline{S}_k^{\star,n} - \tilde{S}_k^{\star,n}$, we easily deduce

$$\forall x,\ y, \quad \tilde{p}_k^{\star,n}(x \mid y) = \frac{\overline{p}_{k/2}^{\star,n}(\frac{x+y}{\sqrt{2}})\overline{p}_{k/2}^{\star,n}(\frac{-x+y}{\sqrt{2}})}{\overline{p}_k^{\star,n}(y)}. \tag{69}$$

From (67), we readily get

$$\begin{aligned} \tilde{p}_k^{\star,n}(x \mid y) &= \frac{\phi(\frac{x+y}{\sqrt{2}})\phi(\frac{-x+y}{\sqrt{2}})}{\phi(y)} \exp\left(\delta_n(x,y) + \frac{1+|x|+|y|}{\sqrt{nk}}O(1) + c_nx^2\right), \\ &= \phi(x)\exp\left(\delta_n(x,y) + \frac{1+|x|+|y|}{\sqrt{nk}}O(1) + c_nx^2\right), \end{aligned} \tag{70}$$

where $\delta_n(x,y) = \frac{nk}{2}\left[(\frac{x+y}{\sqrt{nk}})^3 r_n(\frac{x+y}{\sqrt{nk}}) + (\frac{-x+y}{\sqrt{nk}})^3 r_n(\frac{-x+y}{\sqrt{nk}}) - 2(\frac{y}{\sqrt{nk}})^3 r_n(\frac{y}{\sqrt{nk}})\right]$. From a second order Taylor expansion of $z \mapsto z^3 r_n(z)$ around $\frac{y}{\sqrt{nk}}$ it can be shown that $|\delta_n(x,y)| = x^2 O(\sum_{i=1}^3 (\left|\frac{x}{\sqrt{nk}}\right|^i + \left|\frac{y}{\sqrt{nk}}\right|^i))$. Using $|x| \leq \varepsilon\sqrt{nk}$ and $|y| \leq \varepsilon\sqrt{nk}$, this yields to $|\delta_n(x,y)| = \frac{1}{\sqrt{nk}} O(|x|^3 + x^2 |y|)$. Using the expansion (70) we deduce the lemma. ◇

Lemma 7 *Let $\varepsilon > 0$, then there exist $0 < \varepsilon'' < \varepsilon' < \varepsilon$, and $C > 0$, such that for n large enough, and all k,*

$$\forall\, |y| \leq \varepsilon''\sqrt{nk}, \forall\, |x| \geq \varepsilon'\sqrt{nk}, \text{ we have, } \tilde{p}_k^{\star,n}(x \mid y) \leq C \exp\left(-\frac{\sqrt{nk}}{C}[|x| - \frac{\varepsilon'\sqrt{nk}}{2}]\right). \tag{71}$$

Proof We just consider the case $x > 0$, since the proof for $x < 0$ is similar.

We first need to prove the following upper bound on the density of $\overline{S}_k^{\star,n}$,

$$\exists \varepsilon' > 0, \forall 0 \leq y \leq \varepsilon'\sqrt{nk}, \forall z \geq 0, \quad \overline{p}_k^{\star,n}(y+z) \leq \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2} - \frac{yz}{2} + O(\frac{1+|y|^3}{\sqrt{nk}})}. \tag{72}$$

From (55) with $x = y + z$ and for t given by $t = \sqrt{n} h_n(\frac{y}{\sqrt{kn}})$ where h_n is defined in the proof of Lemma 6 we have

$$\overline{p}_k^{\star,n}(y+z) = \frac{e^{k\Psi(t) - ty\sqrt{k}}}{\sqrt{\Psi^{(2)}(t)}} q_k^n \left(\frac{(y+z)\sqrt{k} - k\Psi^{(1)}(t)}{\sqrt{k\Psi^{(2)}(t)}}\right) e^{-tz\sqrt{k}}.$$

Remark that t is well defined for $\frac{y}{\sqrt{nk}}$ in a neighbourhood $[-\varepsilon', \varepsilon']$ of 0 and is solution to $\Psi^{(1)}(t) = y/\sqrt{k}$. Then, proceeding exactly as in the proof of Lemma 5 we can deduce that

$$\overline{p}_k^{\star,n}(y+z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2} + O(\frac{1+|y|^3}{\sqrt{nk}})} e^{-tz\sqrt{k}}. \tag{73}$$

Now, from (68), we deduce $-tz\sqrt{k} = -yz\left[1 + O(\frac{1}{n})\right] + O(\frac{|z|y^2}{\sqrt{kn}})$. Using that $\frac{|y|}{\sqrt{kn}} \leq \varepsilon' < 1/4$, as soon as we choose the value of ε' small enough, we get $-tz\sqrt{k} \leq -yz/2$, for n large. Finally, (72) follows from (73).

We now prove (71). Up to a modification of the value of ε', we can assume that the result of Lemma 5 i) holds true for $x \leq \varepsilon'\sqrt{nk}$. Thus, from (69), we get for $x \in \mathbb{R}$, $|y| \leq \varepsilon'\sqrt{nk}$

$$\tilde{p}_k^{\star,n}(x \mid y) \leq \sqrt{2\pi} e^{\frac{y^2}{2} + C\frac{1+|y|^3}{\sqrt{nk}}} \overline{p}_{k/2}^{\star,n}(\frac{x+y}{\sqrt{2}}) \overline{p}_{k/2}^{\star,n}(\frac{-x+y}{\sqrt{2}}),$$

where $C > 0$ is some constant. From (55)–(56), it is easily seen that $\overline{p}^{\star,n}$ is a bounded function. Hence, we deduce,

$$\tilde{p}_k^{\star,n}(x \mid y) \leq C e^{\frac{y^2}{2} + C\frac{1+|y|^3}{\sqrt{nk}}} \overline{p}_{k/2}^{\star,n}(\frac{x+y}{\sqrt{2}}), \tag{74}$$

where $C > 0$ is some constant.

We assume, for the sequel, that $x \geq \varepsilon' \sqrt{nk}$ and $|y| \leq \varepsilon'' \sqrt{nk}$ with $\varepsilon'' = \varepsilon'/8$. We write $\overline{p}^{\star,n}_{k/2}(\frac{x+y}{\sqrt{2}})$ as $\overline{p}^{\star,n}_{k/2}((\frac{\varepsilon'\sqrt{nk}}{2} + y)\frac{1}{\sqrt{2}} + (x - \frac{\varepsilon'\sqrt{nk}}{2})\frac{1}{\sqrt{2}})$ and use (72). Since $\frac{\varepsilon'\sqrt{nk}}{2} + y \in [\frac{3}{8}\varepsilon'\sqrt{nk}, \frac{5}{8}\varepsilon'\sqrt{nk}] \subset [\frac{1}{4}\varepsilon'\sqrt{nk}, \frac{3}{4}\varepsilon'\sqrt{nk}]$, we deduce

$$\overline{p}^{\star,n}_{k/2}(\frac{x+y}{\sqrt{2}}) \leq C \exp\left(-\frac{1}{4^3}\varepsilon'^2 nk + \frac{C}{\sqrt{nk}} + C\varepsilon'^3 nk - \frac{\varepsilon'\sqrt{nk}}{16}(x - \varepsilon' \frac{\sqrt{nk}}{2}) \right),$$

for some constant $C > 0$. From (74) and $|y| \leq \varepsilon'\sqrt{nk}/8$, we deduce,

$$\tilde{p}^{\star,n}_k(x \mid y) \leq C \exp\left(-\frac{1}{2.4^3}\varepsilon'^2 nk(1 - \varepsilon' C) + \frac{C}{\sqrt{nk}} - \frac{\varepsilon'\sqrt{nk}}{16}(x - \varepsilon' \frac{\sqrt{nk}}{2}) \right),$$

where C is some positive constant. Up to a modification of ε', we can assume that $\varepsilon' C \leq 1$ and the latter equation gives (71). $\diamond$

We recall that $\tilde{F}_k(x \mid y) = \int_{-\infty}^{x} \tilde{p}^{\star,n}_k(u \mid y)\mathrm{d}u$, is the conditional cumulative distribution function of $\frac{1}{\sqrt{k}}\tilde{S}^{\star,n}_k$ given $\frac{1}{\sqrt{k}}\overline{S}^{\star,n}_k = y$.

Lemma 8 *There exist $\varepsilon > 0$ and $C > 0$, such that for all $0 < x \leq \varepsilon\sqrt{nk}$ and $|y| \leq \varepsilon\sqrt{nk}$,*

$$1 - \tilde{F}_k(x \mid y) = (1 - \Phi(x)) e^{\frac{1}{\sqrt{n}}\tilde{T}^2_k(x,y)}, \tag{75}$$

$$\tilde{F}_k(-x \mid y) = \Phi(-x) e^{\frac{1}{\sqrt{n}}\tilde{T}^3_k(x,y)}, \tag{76}$$

where $\left|\tilde{T}^j_k(x,y)\right| \leq C\frac{1+|x|^3+|y|x^2+|y|}{\sqrt{k}}$, for $j = 2, 3$.

Proof We only prove (75), since the proof of (76) is similar.

Using Lemma 6, let us consider $\varepsilon_1 > 0$, such that for all $|x| \leq \varepsilon_1\sqrt{nk}$, $|y| \leq \varepsilon_1\sqrt{nk}$, we have

$$\tilde{p}^{\star,n}_k(x \mid y) \leq \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} e^{C_1 \frac{1+|x|^3+|y|x^2+|y|}{\sqrt{nk}}},$$

for some constant $C_1 > 0$. Hence, for $0 \leq x < A \leq \varepsilon_1\sqrt{nk}$ and $|y| \leq \varepsilon_1\sqrt{nk}$, we can write

$$\tilde{F}_k(A \mid y) - \tilde{F}_k(x \mid y) \leq e^{C_1 \frac{1+|y|}{\sqrt{nk}}} \int_x^A e^{-\frac{u^2}{2}(1 - 2C_1\frac{u+|y|}{\sqrt{nk}})} \frac{\mathrm{d}u}{\sqrt{2\pi}}. \tag{77}$$

Assume now on that $\varepsilon_1 < \frac{1}{8C_1}$, up to a modification of the value of ε_1. Then, the change of variable $v = u\left(1 - 2C_1\frac{u+|y|}{\sqrt{nk}}\right)^{1/2}$ is one to one as the variable u ranges in $[x, A]$ and it is easy to see that

$\left|\frac{\mathrm{d}v}{\mathrm{d}u}\right| \leq 1 + C_2 \frac{v+|y|}{\sqrt{nk}}$, with some constant $C_2 > 0$. As a result, after a change of variable, we get

$$\begin{aligned}\int_x^A e^{-\frac{u^2}{2}(1-2C_1\frac{u+|y|}{\sqrt{nk}})}\frac{\mathrm{d}u}{\sqrt{2\pi}} &\leq \int_{\hat{x}(y,n,k)}^{\infty} e^{-\frac{v^2}{2}}(1+C_2\frac{v+|y|}{\sqrt{nk}})\frac{\mathrm{d}v}{\sqrt{2\pi}} \\ &= [1-\Phi(\hat{x}(y,n,k))](1+\frac{C_2\,|y|}{\sqrt{nk}}) + \frac{C_2}{\sqrt{nk}}\phi(\hat{x}(y,n,k)) \end{aligned} \tag{78}$$

where we have noted $\hat{x}(y,n,k) = x\left(1-2C_1\frac{x+|y|}{\sqrt{nk}}\right)^{1/2}$.

From the mean value theorem,

$$\log\left(\frac{1-\Phi(\hat{x}(y,n,k))}{1-\Phi(x)}\right) = (x-\hat{x}(y,n,k))\frac{\phi(\xi)}{1-\Phi(\xi)} \leq C_3(x^2+x\,|y|)\frac{\phi(\xi)}{1-\Phi(\xi)},$$

where $\xi \in [\hat{x}(y,n,k), x]$ and $C_3 > 0$ is some constant. From Lemma 2 in [13] we know that $z \mapsto \frac{\phi(z)}{1-\Phi(z)}$ is increasing and Lemma 1 in [13] easily implies

$$\frac{\phi(z)}{1-\Phi(z)} \leq C_4(1+z) \tag{79}$$

for any $z \geq 0$ and $C_4 > 0$ some constant. We deduce

$$\log\left(\frac{1-\Phi(\hat{x}(y,n,k))}{1-\Phi(x)}\right) \leq C_5(x^2+x\,|y|)(1+x), \tag{80}$$

where C_5 is some constant. Putting together (77), (78) and (80) we deduce,

$$\tilde{F}_k(A \mid y) - \tilde{F}_k(x \mid y) \leq [1-\Phi(x)]e^{C_5\frac{1+|x|^3+|y|x^2+|y|}{\sqrt{nk}}}(1+C_2\frac{1+|y|}{\sqrt{nk}}) + \frac{C_2}{\sqrt{nk}}\varphi(\hat{x}(y,n,k)).$$

Using (79) and (80), we have $\phi(\hat{x}(y,n,k)) \leq C_4(1+|x|)(1-\Phi(x))e^{C_5(x^2+x|y|)(1+x)}$. As a consequence, we easily deduce

$$\tilde{F}_k(A \mid y) - \tilde{F}_k(x \mid y) \leq [1-\Phi(x)]e^{C_6\frac{1+|x|^3+|y|x^2+|y|}{\sqrt{nk}}}, \quad \forall 0 \leq x < A \leq \varepsilon_1\sqrt{nk}, |y| \leq \varepsilon_1\sqrt{nk}, \tag{81}$$

and where $C_6 > 0$ is some constant.

In order to prove (75), it remains to control $1 - \tilde{F}_k(A \mid y) = \int_A^\infty \tilde{p}_k^{\star,n}(u \mid y)\mathrm{d}u$. From Lemma 7, there exists $0 < \varepsilon_1'' < \varepsilon_1' < \varepsilon_1$ and $C_7 > 0$ such that for all $|y| \leq \varepsilon_1''\sqrt{nk}$, $u \geq \varepsilon_1'\sqrt{nk}$, $\tilde{p}_k^{\star,n}(u \mid y) \leq C_7 \exp\left(-\frac{\sqrt{nk}}{C_7}[|x| - \frac{\varepsilon_1'\sqrt{nk}}{2}]\right)$. We choose $A = \varepsilon_1'\sqrt{nk} \leq \varepsilon_1\sqrt{nk}$, and with easy computations deduce that $\int_A^\infty \tilde{p}_k^{\star,n}(u \mid y)\mathrm{d}u \leq \frac{C_7^2}{\sqrt{nk}}e^{-\frac{\varepsilon_1'}{4C_7}nk}$. If we let $\varepsilon_2 = \sqrt{\frac{\varepsilon_1'}{4C_7}} \wedge \varepsilon_1'$, and if $x \leq \varepsilon_2\sqrt{nk}$, we have, $1-\Phi(x) \geq \frac{\phi(x)}{C_4+C_4x} \geq \frac{e^{-\frac{\varepsilon_1' nk}{8C_7}}}{\sqrt{2\pi}(C_4+C_4\varepsilon_2\sqrt{nk})}$. This implies,

$$\forall 0 \leq x \leq \varepsilon_2\sqrt{nk}, \quad \int_A^\infty \tilde{p}_k^{\star,n}(u \mid y)\mathrm{d}u \leq C_8[1-\Phi(x)]\exp^{-\frac{nk}{C_8}}, \tag{82}$$

for some constant $C_8 > 0$. Joining (81) with (82) yields to the result (75). $\diamond$

Proof of Lemma 3.

Recalling (32) and repeating the same reasoning as in the proof of Lemma 2, it is sufficient to prove that for $0 \leq |x| \leq \varepsilon\sqrt{kn}$, and $0 \leq |y| \leq \varepsilon\sqrt{kn}$

$$1 - \Phi(x + u(x,y)) \leq 1 - \tilde{F}_k(x \mid y) \leq 1 - \Phi(x - u(x,y)), \tag{83}$$

where $u(x,y) = \frac{C(1+x^2+y^2)}{\sqrt{nk}}$. We focus on the case $x \geq 0$, as the proof is similar for $x \leq 0$. Using Lemma 8, there exist ε_1 and C_1 such that for $0 \leq x \leq \varepsilon_1\sqrt{nk}$, $|y| \leq \varepsilon_1\sqrt{nk}$,

$$(1 - \Phi(x))e^{-C_1 \frac{1+|x|^3+|y|x^2+|y|}{\sqrt{nk}}} \leq 1 - \tilde{F}_k(x \mid y) \leq (1 - \Phi(x))e^{C_1 \frac{1+|x|^3+|y|x^2+|y|}{\sqrt{nk}}}.$$

We set $A = \frac{C_1}{\sqrt{n}} \frac{1+|x|^3+x^2|y|+|y|}{1+|x|^3}$. Then it is simple to check that if ε_1 is small enough and n sufficiently large, we have $A^2 \leq \frac{2C_1^2}{n} + \frac{4C_1^2y^2}{n} \leq \frac{k}{64}$, for all $k \geq 1$. In the same way, we easily check that if ε_1 is small enough, we have $8Ax \leq \sqrt{k}$. As a consequence, we can apply (65)–(66) to get

$$1 - \Phi(x + u) \leq [1 - \Phi(x)]e^{-A\frac{1+x^3}{\sqrt{k}}} \leq [1 - \Phi(x)]e^{A\frac{1+x^3}{\sqrt{k}}} \leq 1 - \Phi(x - u), \tag{84}$$

where $u = 2A\frac{1+x^2}{\sqrt{k}} \leq C\frac{(1+x^2+y^2)}{\sqrt{nk}} = u(x,y)$, for some constant C. Since $A\frac{1+x^3}{\sqrt{k}} = C_1\frac{1+|x|^3+|y|x^2+|y|}{\sqrt{nk}}$, the equation (84) gives (83). $\diamond$

References

[1] A. Alfonsi, B. Jourdain, and A. Kohatsu-Higa. Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme. *Ann. Appl. Probab.*, 24(3):1049–1080, 2014.

[2] A. Alfonsi, B. Jourdain, and A. Kohatsu-Higa. Optimal transport bounds between the time-marginals of a multidimensional diffusion and its Euler scheme. *Electron. J. Probab.*, 20:no. 70, 31, 2015.

[3] A. Davie. Kmt theory applied to approximations of sde. *Stochastic Analysis and Applications*, pages 185–201, 2014.

[4] A. Davie. Pathwise approximation of stochastic differential equations using coupling. *Preprint*, 2014.

[5] Halim Doss. Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 13(2):99–125, 1977.

[6] Uwe Einmahl. Extensions of results of Komlós, Major, and Tusnády to the multivariate case. *J. Multivariate Anal.*, 28(1):20–68, 1989.

[7] G. Flint and T. Lyons. Pathwise approximation of sdes by coupling piecewise abelian rough paths. *arXiv:1505.01298v1*, 2015.

[8] Shūya Kanagawa. On the rate of convergence for Maruyama's approximate solutions of stochastic differential equations. *Yokohama Math. J.*, 36(1):79–86, 1988.

[9] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[10] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32:111–131, 1975.

[11] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV's, and the sample DF. II. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 34(1):33–58, 1976.

[12] Péter Major. The approximation of partial sums of independent RV's. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 35(3):213–220, 1976.

[13] David M. Mason and Harrison H. Zhou. Quantile coupling inequalities and their applications. *Probab. Surv.*, 9:439–479, 2012.

[14] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. I and II*. Probability and its Applications (New York). Springer-Verlag, New York, 1998. Theory and Applications.

[15] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

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