Bernoulli **22**(4), 2016, 2442–2485
DOI: 10.3150/15-BEJ734

Asymptotic development for the CLT in total variation distance

VLAD BALLY[1] and LUCIA CARAMELLINO[2]

[1]*Université Paris-Est, LAMA (UMR CNRS, UPEMLV, UPEC), MathRisk INRIA, F-77454 Marne-la-Vallée, France. E-mail: bally@univ-mlv.fr*
[2]*Dipartimento di Matematica, Università di Roma – Tor Vergata, Via della Ricerca Scientifica 1, I-00133 Roma, Italy. E-mail: caramell@mat.uniroma2.it*

The aim of this paper is to study the asymptotic expansion in total variation in the central limit theorem when the law of the basic random variable is locally lower-bounded by the Lebesgue measure (or equivalently, has an absolutely continuous component): we develop the error in powers of $n^{-1/2}$ and give an explicit formula for the approximating measure.

Keywords: abstract Malliavin calculus; integration by parts; regularizing functions; total variation distance

1. Introduction

The aim of this paper is to study the convergence in total variation in the Central Limit Theorem (CLT) under a certain regularity condition for the random variable at hand. Given two measures μ, ν in $\mathbb{R}^N$, we recall that the distance in total variation is defined as

$$d_{\mathrm{TV}}(\mu,\nu) = \sup\left\{\left|\int f\,d\mu - \int f\,d\nu\right| : \|f\|_\infty \le 1\right\}.$$

Let F be a centred r.v. in $\mathbb{R}^N$ with identity covariance matrix and let F_k, $k \in \mathbb{N}$, denote independent copies of F. We set

$$S_n = \frac{1}{\sqrt{n}}\sum_{k=1}^n F_k.$$

We also define μ_n the law of S_n and Γ the standard Gaussian law in $\mathbb{R}^N$.

The problem of the convergence in total variation for the CLT, that is, $d_{\mathrm{TV}}(\mu_n,\Gamma) \to 0$ as $n \to \infty$, is very old. Prohorov [14] in 1952 proved that, in dimension 1, a necessary and sufficient condition in order to get the result is that there exists n_0 such that the law of $\sum_{k=1}^{n_0} F_k$ has an absolutely continuous component (see next Definition 2.1). Then many related problems have been considered in the literature, such as the generalization to the multidimensional case, the study of the speed of convergence, the convergence and the development of the density of S_n, if it exists, or the case of a r.v. F whose law has not necessarily an absolutely continuous component, the latter implying the use of a different distance, which is similar to the total variation one but defined on a special class of test functions, typically indicator functions of special sets.

A first class of results has been obtained by Rao [15] and then improved by Battacharaya [6]: in [15], one proves that the convergence in the CLT holds when the test function is the indicator function of a convex set D. This result is improved in [6] where D is no more a convex set but a set with a boundary which is small in some sense. An overview on this topic is given in [5]. But it turns out that one is not generally able to extend the above mentioned results to a general set D (and so to general measurable and bounded test functions), because thanks to the Prohorov's result, one needs to assume a little bit of regularity on the law of the basic random variable F which comes on in the CLT. In such a case, Sirazhdinov and Mamatov [16] prove that if $F \in L^3(\Omega)$ then the density of the absolutely continuous component of the law μ_n converges in $L^1(\mathbb{R}^N)$ to the standard Gaussian density and, therefore, the convergence of the CLT holding in total variation distance, at speed $1/\sqrt{n}$. This is done in the one-dimensional case, but it works as well in the multidimensional case. The second part of the book [5] gives a complete view on the recent research on this topic, mainly on the development of the density of S_n around the standard Gaussian density. Results concerning the convergence in the entropy distance (under the same type of hypothesis) has been recently obtained in [7].

This paper contributes in this direction by giving the precise expansion of the CLT in total variation distance. More precisely, we assume that the law of F is locally lower bounded by the Lebesgue measure Leb_N on $\mathbb{R}^N$ in the following sense: there exists an open set D_0 and $\varepsilon_0 > 0$ such that for every Borel set A one has

$$\mathbb{P}(F \in A) \geq \varepsilon_0 \times \mathrm{Leb}_N(A \cap D_0). \tag{1.1}$$

We will show that this is equivalent to the request that the law of F has an absolutely continuous component (and moreover, we can construct such absolutely continuous measure in order that the associated density is a non-negative lower semicontinuous function, see Appendix A). So it is clear that our hypotheses overlaps the assumption of the existence of the density but one cannot reduce one to another (if the law of F gives positive probability to the rational points then it is not absolutely continuous; and doing convolutions does not help). Let us give a non-trivial example. Consider a functional F on the Wiener space and assume that F is twice differentiable in Malliavin sense: $F \in \mathbb{D}^{2,p}$ with $p > N$ where N is the dimension of F. Let σ_F be the Malliavin covariance matrix of F. If $\mathbb{P}(\det \sigma_F > 0) = 1$ then the celebrated criterion of Bouleau and Hirsh ensures that the law of F is absolutely continuous, so we are in the classical case (in fact it suffices that $F \in \mathbb{D}^{1,2}$). But if $\mathbb{P}(\det \sigma_F > 0) < 1$ this criterion does no more work (and one may easily produce examples when the law of F is not absolutely continuous). In [3], we proved that if $\mathbb{P}(\det \sigma_F > 0) > 0$ then the law of F has the property (1.1). Notice also that in the one-dimensional case ($N = 1$) the fact that F is not constant immediately implies that $\mathbb{P}(\sigma_F > 0) > 0$. Indeed, in this case $\sigma_F = |DF|^2$ and if this is almost surely null, then F is constant.

Let us introduce our results. We consider a random variable $F \in L^2(\mathbb{R}^N)$ which satisfies (1.1), such that $\mathbb{E}(F) = 0$ and the covariance matrix of F is the identity matrix. We take a sequence $F_k, k \in \mathbb{N}$ of independent copies of F and we denote by μ_n the law of $S_n = \frac{1}{n^{1/2}} \sum_{k=1}^n F_k$ and by Γ the standard Gaussian law on $\mathbb{R}^N$. Under these hypotheses, we first prove that $\lim_{n\to\infty} d_{\mathrm{TV}}(\mu_n, \Gamma) = 0$ where d_{TV} is the total variation distance. Then we give the asymptotic development, which we are able to find according to additional requests on the existence of the moments of F. More precisely, we get that, for $r \geq 2$, if $F \in L^{r+1}(\Omega)$ and if the moments of

F up to order r agree with the moments of the standard Gaussian law then under (1.1) one has

$$d_{\mathrm{TV}}(\mu_n, \Gamma) \le C\big(1 + \mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1} \times \frac{1}{n^{(r-1)/2}}.$$

In the general case, we obtain the following asymptotic expansion. For $r \ge 2$ and $n \ge 1$, we define a measure on $\mathbb{R}^N$ through

$$\Gamma_{n,r}(dx) = \gamma(x)\left(1 + \sum_{m=1}^{[r/3]} \frac{1}{n^{m/2}} \mathcal{K}_m(x)\right) dx, \tag{1.2}$$

where γ denotes the probability density function of a standard normal random variable in $\mathbb{R}^N$ and $\mathcal{K}_m(x)$ is a polynomial of order m ($[\cdot]$ standing for the integer part). Note that for $r = 2$ one gets $\Gamma_{n,r}(dx) = \gamma(x)\,dx = \Gamma(dx)$. So, we prove that if $F \in L^{r+1}(\Omega)$ with $r \ge 2$ then there exist polynomials $\mathcal{K}_m(x)$, $m = 1, \dots, [r/3]$ (no polynomials are needed for $r = 2$), such that, setting $\Gamma_{n,r}$ the measure in (1.2) and μ_n the law of S_n, under (1.1) one has

$$d_{\mathrm{TV}}(\mu_n, \Gamma_{n,r}) \le C\big(1 + \mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1+1} \times \frac{1}{n^{([r/3]+1)/2}}, \tag{1.3}$$

where $C > 0$ depends on r and N. So, in order to improve the development (and the rate of convergence) one needs to pass from the request $F \in L^{3k}$ to $F \in L^{3k+3}$, $k \ge 1$.

The development given in (1.3) is analogous to the one obtained in Theorem 19.5, page 199 in [5]. But our development is explicit: in [5], the result is obtained using the Fourier transform and consequently the coefficients in the development involve the inverse of the Fourier transform, whereas here we give an explicit expression for the polynomials $\mathcal{K}_m(x)$, as a linear combination of the Hermite polynomials (see next formula (4.38)).

The main instrument used in this paper is the Malliavin-type finite dimensional calculus defined in [4] and [3]. It turns out that for a random variable which satisfies (1.1) a very pleasant calculus may be settled. The idea is that (1.1) guarantees that the law of F contains some smooth noise. Then, using a splitting procedure (see Proposition 3.1 for details), we may isolate this noise and achieve integration by parts formulae based on it.

In the last years, a number of results concerning the weak convergence of functionals on the Wiener space using Malliavin calculus and Stein's method have been obtained by Nurdin, Peccati, Nualart and Poly; see, for example, [9–12]. In particular, in [10] and [9] the authors consider functionals living in a finite direct sum of chaoses and prove that under a very weak non-degeneracy condition (analogous to the one we consider here) the convergence in distribution of a sequence of such functionals imply the convergence in total variation. The results proved in these papers may be seen as variants of the CLT but for dependent random variables – so the framework and the arguments are rather different from the one considered here.

2. Main results

Let X be a random variable in $\mathbb{R}^N$ and let μ_X denote its law. The Lebesgue decomposition of μ_X says that there exist a measure $\mu(dx) = \mu(x)\,dx$, that is, μ is absolutely continuous w.r.t. the

Lebesgue measure, and a further measure ν which is singular, that is, concentrated on a set of null Lebesgue measure, such that

$$\mu_X(dx) = \mu(x)\,dx + \nu(dx). \tag{2.1}$$

Definition 2.1. *X is said to have an* absolutely continuous component *if the absolutely continuous measure μ in the decomposition* (2.1) *is not null*, *that is*, $\nu(\mathbb{R}^N) < 1$.

Definition 2.1 plays a crucial role when dealing with the convergence of the Central Limit Theorem (CLT) in the total variation distance d_{TV}. We recall the definition of d_{TV}: for any two measures μ and ν in $\mathbb{R}^N$ then

$$d_{\mathrm{TV}}(\mu,\nu) = \sup\left\{\left|\int f\,d\mu - \int f\,d\nu\right| : \|f\|_\infty \le 1\right\}.$$

We discuss here the CLT in total variation distance, so we consider a sequence $\{F_k\}_k$ of i.i.d. square integrable random variables, with null mean and covariance matrix $C(F)$. We set $A(F)$ the inverse of $C(F)^{1/2}$ and

$$S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n A(F) F_k.$$

We recall the following classical result, due to Prohorov [14].

Theorem 2.2 (Prohorov). *Let μ_n denote the law of S_n and Γ denote the standard Gaussian law in $\mathbb{R}^N$. The convergence in the CLT takes place w.r.t. the total variation distance*, *that is* $d_{\mathrm{TV}}(\mu_n,\Gamma) \to 0$ *as $n\to\infty$*, *if and only if there exists $n_0 \ge 1$ such that the random variable S_{n_0} has an absolutely continuous component*.

Hereafter, we assume that the common law of the F_k's has an absolutely continuous component, and this is not a big loss in generality. In fact, due to the Prohorov's theorem, otherwise we can packet the sequence $\{F_k\}_k$ in groups of n_0 r.v.'s, so we can deal with

$$\bar{S}_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n \bar{F}_k \qquad \text{where } \bar{F}_k = \frac{1}{\sqrt{n_0}} \sum_{i=kn_0}^{(k+1)n_0} F_i.$$

Let us introduce an equivalent way to see probability laws having an absolutely continuous component. From now on, Leb_N denotes the lebesgue measure in $\mathbb{R}^N$.

Definition 2.3. *A probability law μ in $\mathbb{R}^N$ is said to be* locally lower bounded by the Lebesgue measure, *in symbols $\mu \succeq \mathrm{Leb}_N$*, *if there exist $\varepsilon_0 > 0$ and an open set $D_0 \subset \mathbb{R}^N$ such that*

$$\mu(A) \ge \varepsilon_0\, \mathrm{Leb}_N(A\cap D_0) \qquad \forall A \in \mathcal{B}\big(\mathbb{R}^N\big). \tag{2.2}$$

We have the following.

Proposition 2.4. *Let F be a r.v. in $\mathbb{R}^N$ and let μ_F denote its law. Then the following statements are equivalent*:

(i) $\mu_F \succeq \mathrm{Leb}_N$;

(ii) *F has an absolutely continuous component*;

(iii) *there exist three independent r.v.'s χ taking values in $\{0,1\}$, with $\mathbb{P}(\chi=1)>0$, and V, W in $\mathbb{R}^N$, with V absolutely continuous, such that*

$$\mathbb{P}\big(\chi V+(1-\chi)W\in dv\big)=\mu_F(dv). \tag{2.3}$$

Moreover, if one of the above conditions holds then the covariance matrix $C(F)$ of F is invertible.

The proof of Proposition 2.4 is postponed to Appendix A. As an immediate consequence of Proposition 2.4, if $\mu_F \succeq \mathrm{Leb}_N$ then $\underline{\lambda}(F)>0$, $\underline{\lambda}(F)$ denoting the smallest eigenvalue of $\widehat{C}(F)=C(F)^{-1}$. We denote through $\overline{\lambda}(F)$ the associated largest eigenvalue.

We are now ready to introduce the main contributions of this paper. We first give a new proof of the convergence in total variation in the CLT.

Theorem 2.5. *Suppose that $\mu_F \succeq \mathrm{Leb}_N$, $\mathbb{E}(F)=0$ and $\mathbb{E}(|F|^2)<\infty$. Then*

$$\lim_{n\to\infty} d_{\mathrm{TV}}(\mu_n,\Gamma)=0, \tag{2.4}$$

where μ_n denotes the law of S_n and Γ is the standard Gaussian law in $\mathbb{R}^N$.

This is done especially in order to set up the main arguments and results from abstract Malliavin calculus coming from representation (2.3) that are used throughout this paper. Let us stress that Nourdin and Poly in [12] have dealt with r.v.'s fulfilling properties that imply (2.3), to which they apply results from [2] about a finite dimensional Malliavin-type calculus.

Afterward, we deal with the estimate of the error. In fact, by means of additional requests of the existence of the moments of F up to order ≥ 3, we get the asymptotic expansion in powers of $n^{-1/2}$ of the law of S_n in total variation distance. We first obtain the following.

Theorem 2.6. *Suppose that $\mu_F \succeq \mathrm{Leb}_N$ and $\mathbb{E}(F)=0$. Let μ_n denote the law of S_n and Γ denote the standard Gaussian law in $\mathbb{R}^N$. Let $r\geq 2$. If $\mathbb{E}(|F|^{r+1})<\infty$ and all moments up to order r of $A(F)F$ agree with the moments of a standard Gaussian r.v. in $\mathbb{R}^N$ then*

$$d_{\mathrm{TV}}(\mu_n,\Gamma)\leq C\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\times\frac{1}{n^{(r-1)/2}}, \tag{2.5}$$

where $C>0$ depends on r, N, $\underline{\lambda}(F)$ and $\overline{\lambda}(F)$.

In the general case, that is the moments do not generally coincide, we get the following expansion. For $r\geq 2$ and $n\geq 1$, we define a measure on $\mathbb{R}^N$ through

$$\Gamma_{n,r}(dx)=\gamma(x)\left(1+\sum_{m=1}^{[r/3]}\frac{1}{n^{m/2}}\mathcal{K}_m(x)\right)dx, \tag{2.6}$$

where γ denotes the probability density function of a standard normal random variable in $\mathbb{R}^N$ and $\mathcal{K}_m(x)$ is a polynomial of order m – the symbol $[\cdot]$ stands for the integer part and for $r=2$ the sums in (2.6) nullify, so that $\Gamma_{n,2}(dx)=\gamma(x)\,dx=\Gamma(dx)$. Then we get the following.

Theorem 2.7. *Let $r\geq 2$ and $\mathbb{E}(|F|^{r+1})<\infty$. Then there exist polynomials $\mathcal{K}_m(x)$, $m=1,\dots,[r/3]$ (no polynomials are needed for $r=2$), such that, setting $\Gamma_{n,r}$ the measure in (2.6) and μ_n the law of S_n, one has*

$$d_{\mathrm{TV}}(\mu_n,\Gamma_{n,r})\leq C\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\times\frac{1}{n^{([r/3]+1)/2}},$$

where $C>0$ depends on r, N, $\underline{\lambda}(F)$ and $\overline{\lambda}(F)$.

The statement of Theorem 2.7 is not properly written, because no information is given about the polynomials $\mathcal{K}_m$'s. We observe that in next formula (4.38) we give a closed-form expression for the $\mathcal{K}_m$'s in terms of a linear combination of Hermite polynomials, whose coefficients can be explicitly written (so not involving inverse Fourier transforms).

Remark 2.8. Let $F\in\mathbb{D}^{2,p}$ with $p>N$, $\mathbb{D}^{k,p}$ denoting the set of the random variables which are derivable in Malliavin sense up to order k in L^p (see Nualart [13]). If $\mathbb{P}(\sigma_F>0)>0$, σ_F standing for the Malliavin covariance matrix of F (and note that this request is much weaker than the non-degeneracy of σ_F) then Theorem 2.16 in [3] gives that $\mu_F\succeq\mathrm{Leb}_N$ (and this property may be strict, that is F may not be absolutely continuous). So both Theorems 2.6 and 2.7 can be applied.

The rest of this paper is devoted to the proofs of the above results: Section 3 allows us to prove Theorem 2.5 and the remaining Theorems 2.6 and 2.7 are discussed in Section 4.

3. Convergence in the total variation distance

The aim of this section is to prove Theorem 2.5, whose proof requires some preparatives which will be useful also in the sequel.

3.1. Abstract Malliavin calculus based on a splitting method

We consider a random variable $F\in\mathbb{R}^N$ whose law μ_F is such that $\mu_F\succeq\mathrm{Leb}_N$. As proved in Proposition 2.4, the covariance matrix $C(F)$ of F is invertible. So, without loss of generality we can assume from now on that $C(F)$ is the identity matrix, otherwise we work with $A(F)F$, $A(F)$ being the inverse of $C(F)^{1/2}$.

We consider the following special splitting for the law of μ_F, giving, as a consequence, representation (2.3). We start from the class of localization functions $\psi_a:\mathbb{R}\to\mathbb{R}$, $a>0$, defined as

$$\psi_a(x)=1_{|x|\leq a}+\exp\bigg(1-\frac{a^2}{a^2-(|x|-a)^2}\bigg)1_{a<|x|<2a}. \tag{3.1}$$

Then $\psi_a \in C_c^\infty(\mathbb{R})$ (the subscript "c" standing for compact support), $0 \le \psi_a \le 1$ and we have the following property: for every $k, p \in \mathbb{N}$ there exists a universal constant $C_{k,p}$ such that for every $x \in \mathbb{R}_+$

$$\psi_a(x)\big|(\ln \psi_a)^{(k)}(x)\big|^p \le \frac{C_{k,p}}{a^{pk}}. \tag{3.2}$$

By the very definition, if $\mu_F \succeq \mathrm{Leb}_N$ then we may find $v_0 \in \mathbb{R}^N$, $r_0 > 0$ and $\varepsilon_0 > 0$ such that $\mathbb{P}(F \in A) \ge \varepsilon_0 \,\mathrm{Leb}_N(A \cap B_{r_0}(v_0))$. Then for every non-negative function $f : \mathbb{R}^N \to \mathbb{R}_+$ we have

$$\mathbb{E}\big(f(F)\big) \ge \varepsilon_0 \int_{\mathbb{R}^N} \psi_{r_0/2}\big(|v - v_0|\big) f(v)\, dv. \tag{3.3}$$

We denote

$$m_0 = \varepsilon_0 \int_{\mathbb{R}^N} \psi_{r_0/2}\big(|v - v_0|\big)\, dv. \tag{3.4}$$

Of course, $m_0 > 0$. But, up to choose ε_0 smaller, we also have $m_0 < 1$. So, we consider three independent random variables $\chi \in \{0, 1\}$ and $V, W \in \mathbb{R}^N$ with laws

$$\begin{aligned} &\mathbb{P}(\chi = 1) = m_0, \qquad \mathbb{P}(\chi = 0) = 1 - m_0, \\ &\mathbb{P}(V \in dv) = \frac{\varepsilon_0}{m_0} \psi_{r_0/2}\big(|v - v_0|\big)\, dv, \\ &\mathbb{P}(W \in dv) = \frac{1}{1 - m_0}\big(\mu_F(dv) - \varepsilon_0 \psi_{r_0/2}\big(|v - v_0|\big)\, dv\big). \end{aligned} \tag{3.5}$$

Then

$$\mathbb{P}\big(\chi V + (1 - \chi) W \in dv\big) = \mu_F(dv). \tag{3.6}$$

So, we have just proved the following.

Proposition 3.1. *If $\mu_F \succeq \mathrm{Leb}_N$ then representation (2.3) holds.*

From now on, we will work with the representation of μ_F in (3.6) so we always take

$$F = \chi V + (1 - \chi) W,$$

χ, V and W being independent and whose laws are given in (3.5).

We come now to the central limit theorem. We consider a sequence $\chi_k, V_k, W_k \in \mathbb{R}^N$, $k \in \mathbb{N}$ of independent copies of $\chi, V, W \in \mathbb{R}^N$ and we take $F_k = \chi_k V_k + (1 - \chi_k) W_k$. Then we look to

$$S_n = \frac{1}{n^{1/2}} \sum_{k=1}^n F_k = \frac{1}{n^{1/2}} \sum_{k=1}^n \big(\chi_k V_k + (1 - \chi_k) W_k\big).$$

In order to prove the CLT in the total variation distance, we will use the abstract Malliavin calculus settled in [4] and [3] associated to the basic noise

$$V = (V_1, \dots, V_n) = \big(\big(V_1^1, \dots, V_1^N\big), \dots, \big(V_n^1, \dots, V_n^N\big)\big) \in \mathbb{R}^{N \times n} \tag{3.7}$$

(this will be done for each fixed n). To begin, we recall the notation and some results from [3]. We work with functionals $X = f(V)$ with $f \in C_b^\infty(\mathbb{R}^{N\times n};\mathbb{R})$, the subscript "$b$" standing for bounded derivatives of any order. Then we set

$$\mathcal{S} = \{ f(V) : f \in C_b^\infty(\mathbb{R}^{N\times n};\mathbb{R}) \}$$

and for a functional $X \in \mathcal{S}$ we define the Malliavin derivatives

$$D_{(k,i)}X = \frac{\partial X}{\partial V_k^i} = \frac{\partial f}{\partial v_k^i}(V), \qquad k = 1,\dots,n, i = 1,\dots,N. \tag{3.8}$$

The Malliavin covariance matrix for a multidimensional functional $X = (X^1,\dots,X^d) \in \mathcal{S}^d$ is defined as

$$\sigma_X^{i,j} = \langle DX^i, DX^j \rangle = \sum_{k=1}^n \sum_{r=1}^N D_{(k,r)}X^i \times D_{(k,r)}X^j, \qquad i,j = 1,\dots,d. \tag{3.9}$$

We will denote by λ_X the lower eigenvalue of σ_X, that is,

$$\lambda_X = \inf_{|\xi|=1} \langle \sigma_X \xi, \xi \rangle = \inf_{|\xi|=1} \sum_{k=1}^n \sum_{i=1}^N \langle D_{(k,i)}X, \xi \rangle^2. \tag{3.10}$$

Moreover, we define the higher order derivatives just by iterating D. We consider a multiindex $\alpha = (\alpha_1,\dots,\alpha_m)$ with $\alpha_j = (k_j, i_j)$, $k_j \in \{1,\dots,n\}$, $i_j \in \{1,\dots,N\}$ and we set $|\alpha| = m$. Then we define

$$D_\alpha X = \frac{\partial^m X}{\partial V_{k_m}^{i_m} \cdots \partial V_{k_1}^{i_1}} = \partial_\alpha f(V) \tag{3.11}$$

with

$$\partial_\alpha f(v) = \frac{\partial^m f}{\partial v_{k_m}^{i_m} \cdots \partial v_{k_1}^{i_1}}(v).$$

We will work with the norms

$$|X|_{1,m}^2 = \sum_{1\le|\alpha|\le m} |D_\alpha X|^2, \qquad |X|_m^2 = |X|^2 + |X|_{1,m}^2, \tag{3.12}$$

$$\|X\|_{1,m,p} = \big\| |X|_{1,m} \big\|_p = \big(\mathbb{E}(|X|_{1,m}^p)\big)^{1/p}, \qquad \|X\|_{m,p} = \|X\|_p + \|X\|_{1,m,p}. \tag{3.13}$$

We define now the Ornstein–Uhlenbeck operator by

$$-LX = \sum_{k=1}^n \sum_{i=1}^N D_{(k,i)}D_{(k,i)}X + \sum_{k=1}^n \sum_{i=1}^N D_{(k,i)}X \partial_i \ln \psi_{r_0/2}(|V_k - v_0|). \tag{3.14}$$

These are the operators introduced in [4] and [3] in connection to the random variable V in (3.7) and taking the weights $\pi_k = 1$. We will use the results from [3] in this framework. In particular, as a straightforward consequence of Theorem 3.1 in [3] (take $\mathbf{\Theta} = 1$ therein) and Theorem 3.4 in [3] (see (3.28) therein), we can state integration by parts formulas and estimates for the weights. For later use, we resume in the following statement such facts.

Proposition 3.2. *$X \in \mathcal{S}^d$ be such that*

$$\left\| (\det \sigma_X)^{-1} \right\|_p < \infty \qquad \text{for every } p \geq 1.$$

Set γ_X the inverse of σ_X. Then the following integration by parts formula holds: for every $\phi \in C_b^\infty(\mathbb{R}^d; \mathbb{R})$, $Y \in \mathcal{S}$, $q \in \mathbb{N}$ and for every $\beta \in \{1, \dots, d\}^q$ one has

$$\mathbb{E}\big(\partial_\beta \phi(X) Y\big) = \mathbb{E}\big(\phi(X) H_\beta^q(X, Y)\big),$$

where $\partial_\beta \phi(x) = \partial_{x^{\beta_q}} \cdots \partial_{x^{\beta_1}} \phi(x)$ and the weights $H_\beta^q(X, Y)$ are recursively given by:

- *if $q = 1$, then*

$$H_\beta^1(X, Y) \equiv H_\beta(X, Y) = \sum_{r=1}^d \left(Y \gamma_X^{r,\beta} L X^r - \sum_{k=1}^n \sum_{i=1}^N D_{(k,i)}\big(Y \gamma_X^{r,\beta}\big) D_{(k,i)} X^r \right), \qquad \beta = 1, \dots, d;$$

- *if $q > 1$, then*

$$H_\beta^q(X, Y) = H_{\beta_q}\big(X, H^{q-1}_{(\beta_1, \dots, \beta_{q-1})}(X, Y)\big), \qquad \beta \in \{1, \dots, d\}^q.$$

Moreover, the following estimate holds: for every $\beta \in \{1, \dots, d\}^q$ and $m \in \mathbb{N}$ then

$$\begin{aligned} &\big| H_\beta^q(X, Y) \big|_m \leq C \mathbf{A}_{m+q}(X)^q |Y|_{m+q}, \\ &\text{where } \mathbf{A}_l(X) = \big(1 \vee (\det \sigma_X)^{-1}\big)^{l+1} \big(1 + |X|_{1,l+1}^{2d(l+2)} + |LX|_{l-1}^2\big), \end{aligned} \tag{3.15}$$

$|\cdot|_m$ being defined in (3.12).

We come now back to S_n, which we write as

$$S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n \big(\chi_k V_k + (1 - \chi_k) W_k\big).$$

For every $k = 1, \dots, n$ and $l, i = 1, \dots, N$, we have

$$D_{(k,i)} S_n^l = \frac{1}{\sqrt{n}} \chi_k \mathbf{1}_{l=i}.$$

As a consequence, we obtain

$$\sigma_{S_n} = \frac{1}{n}\sum_{k=1}^{n} \chi_k I, \tag{3.16}$$

where I denotes the identity matrix, and

$$\lambda_{S_n} = \frac{1}{n}\sum_{k=1}^{n} \chi_k. \tag{3.17}$$

The derivatives of order higher than two of S_n are null, so we obtain for every $q \in \mathbb{N}$

$$|S_n|_{1,q}^2 \le \frac{1}{n}\sum_{k=1}^{n} \chi_k \le 1, \qquad |S_n|_q^2 \le |S_n|^2 + \frac{1}{n}\sum_{k=1}^{n} \chi_k \le |S_n|^2 + 1, \tag{3.18}$$

and consequently

$$\|S_n\|_{1,q,p} \le 1, \qquad \|S_n\|_{q,p} \le \|S_n\|_p + 1. \tag{3.19}$$

In particular, $\|S_n\|_{1,q,p}$ is finite for every q, p whereas $\|S_n\|_{q,p}$ is finite according to $F \in L^p(\Omega)$.

Let us now compute LS_n. We have

$$\begin{aligned} -LS_n^l &= \sum_{k=1}^{n}\sum_{i=1}^{N} D_{(k,i)} D_{(k,i)} S_n^l + \sum_{k=1}^{n}\sum_{i=1}^{N} D_{(k,i)} S_n^l \partial_i \ln \psi_{r_0/2}\big(|V_k - v_0|\big) \\ &= \frac{1}{\sqrt{n}}\sum_{k=1}^{n} \chi_k \partial_l \ln \psi_{r_0/2}\big(|V_k - v_0|\big). \end{aligned}$$

We now estimate $\|LS_n\|_{q,p}$.

Lemma 3.3. *For every $q \in \mathbb{N}$, there exists a universal constant C_q such that*

$$\|LS_n\|_{q,p} \le \frac{C_q}{r_0^{q+1}}. \tag{3.20}$$

Proof. The basic fact in our calculus is that

$$\begin{aligned} \mathbb{E}\big(\partial_i \ln \psi_{r_0/2}(V_k - v_0)\big) &= \frac{\varepsilon_0}{m_0}\int_{\mathbb{R}^N} \partial_i \ln \psi_{r_0/2}\big(|v - v_0|\big) \times \psi_{r_0/2}\big(|v - v_0|\big)\, dv \\ &= \frac{\varepsilon_0}{m_0}\int_{\mathbb{R}^N} \partial_i \psi_{r_0/2}\big(|v - v_0|\big)\, dv = 0. \end{aligned}$$

We denote

$$Q_k = \nabla \ln \psi_{r_0/2}(V_k - v_0)$$

and we have

$$\mathbb{E}\big(Q_k^l\big) = \mathbb{E}\big(\partial_l \ln \psi_{r_0/2}\big(|V_k - v_0|\big)\big) = 0.$$

So $\sum_{k=1}^n \chi_k Q_k^l$, $n \in \mathbb{N}$, is a martingale and the Burkholder's inequality gives

$$\mathbb{E}\big(\big|LS_n^l\big|^p\big) = \mathbb{E}\left(\left|\frac{1}{\sqrt{n}}\sum_{k=1}^n \chi_k Q_k^l\right|^p\right) \le C\mathbb{E}\left(\left(\frac{1}{n}\sum_{k=1}^n \chi_k \big|Q_k^l\big|^2\right)^{p/2}\right) \le \frac{C}{n}\sum_{k=1}^n \mathbb{E}\big(\big|Q_k^l\big|^p\big).$$

By (3.2),

$$\mathbb{E}\big(\big|Q_k^l\big|^p\big) \le C\frac{1}{r_0^p}$$

so that

$$\|LS_n\|_p \le \frac{C}{r_0}.$$

We go further and we compute $D_{(k,i)}LS_n$. We have

$$-D_{(k,i)}LS_n^l = \frac{1}{\sqrt{n}}\sum_{k'=1}^n \chi_{k'} D_{(k,i)}\partial_l \ln \psi_{r_0/2}\big(|V_{k'} - v_0|\big) = \frac{1}{\sqrt{n}}\chi_k D_{(k,i)}\partial_l \ln \psi_{r_0/2}\big(|V_k - v_0|\big)$$

so that

$$\begin{aligned}|DLS_n|_1^2 &\le |LS_n|^2 + \frac{1}{n}\sum_{k=1}^n\sum_{i=1}^N \big|D_{(k,i)}\nabla \ln \psi_{r_0/2}\big(|V_k - v_0|\big)\big|^2 \\ &\le |LS_n|^2 + \frac{C}{n}\sum_{k=1}^n\sum_{i,j=1}^N \big|\partial_i\partial_j \ln \psi_{r_0/2}\big(|V_k - v_0|\big)\big|^2.\end{aligned}$$

Once again using (3.2), we obtain

$$\big\|\partial_i\partial_j \ln \psi_{r_0/2}\big(|V_k - v_0|\big)\big\|_p \le \frac{C}{r_0^2}$$

and consequently

$$\|LS_n\|_{1,p} \le \frac{C}{r_0^2}.$$

For higher order norms, the estimates are similar. □

We add a final property on the behavior of the Malliavin covariance matrix that will be used in next Section 4.2.

Lemma 3.4. *Suppose that* $\mu_F \succeq \mathrm{Leb}_N$. *There exists a universal constant* C *such that for every* $n \in \mathbb{N}$ *and every*

$$\varepsilon \le \varepsilon_* = 2^{-N} m_0^N \tag{3.21}$$

then

$$\mathbb{P}(\det \sigma_{S_n} \le \varepsilon) \le C \exp\left(-\frac{n}{4(1/m_0 - 1)}\right), \tag{3.22}$$

m_0 *being defined in* (3.4).

Proof. Using (3.17),

$$\mathbb{P}(\det \sigma_{S_n} \le \varepsilon) \le \mathbb{P}\big(\lambda_{S_n} \le \varepsilon^{1/N}\big) = \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n \chi_k \le \varepsilon^{1/N}\right) = \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n (\chi_k - m_0) \le \varepsilon^{1/N} - m_0\right).$$

Since $\varepsilon^{1/N} \le \frac{1}{2} m_0$, the above term is upper bounded by

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n (\chi_k - m_0) \le -\frac{1}{2} m_0\right) = \mathbb{P}\left(\frac{1}{n^{1/2}}\sum_{k=1}^n \frac{\chi_k - m_0}{v_{m_0}} \le -n^{1/2}\frac{m}{2v_{m_0}}\right)$$

with $v_{m_0} = (m_0(1-m_0))^{1/2} = \mathrm{Var}(\chi_k)$. We denote by $a = n^{1/2}\frac{m_0}{2v_{m_0}}$ and we use the Berry–Esseen theorem in order to upper bound this quantity by

$$C\int_{-\infty}^a \exp\big(-x^2/2\big)\,dx \le C' \exp\left(-\frac{a^2}{4}\right) = C' \exp\left(-\frac{n}{4(1/m_0 - 1)}\right).$$

□

3.2. Proof of Theorem 2.5

We need now a localized variant of Lemma 2.5 and Theorem 2.7 in [3]. So, we start with the basic definitions.

We consider a localizing r.v. Θ taking values in [0, 1] of the form

$$\Theta = \psi_a(Z), \qquad a > 0, Z \in \mathcal{S}, \tag{3.23}$$

ψ_a being defined in (3.1). We set $\mathbb{P}_\Theta$ and $\mathbb{E}_\Theta$ through

$$d\mathbb{P}_\Theta = \Theta\, d\mathbb{P} \quad \text{and} \quad \mathbb{E}_\Theta = \text{expectation w.r.t. } \mathbb{P}_\Theta.$$

For $X \in \mathcal{S}^d$, we define the localized Sobolev norms

$$\|X\|_{p,\Theta} = \mathbb{E}_\Theta\big(|X|^p\big)^{1/p}, \qquad \|X\|_{1,m,p,\Theta} = \mathbb{E}_\Theta\big(|X|_{1,m}^p\big)^{1/p} \quad \text{and} \quad \|X\|_{m,p,\Theta} = \mathbb{E}_\Theta\big(|X|_m^p\big)^{1/p},$$

$|X|_{1,m}$ and $|X|_m$ being given in (3.12), and we set

$$A_{p,\Theta}(X) = \|X\|_{3,p,\Theta} + \|LX\|_{1,p,\Theta}. \tag{3.24}$$

We also consider the law of a d-dimensional r.v. X under $\mathbb{P}_\Theta$: it is the measure in $\mathbb{R}^d$ defined as

$$\mu_{X,\Theta}(dx) = \mathbb{P}_\Theta(X \in dx).$$

We allow the case $a = +\infty$ in (3.23): this gives $\Theta \equiv 1$, so $\mathbb{P}_\Theta \equiv \mathbb{P}$ and no localization is taken into account.

Finally, for $k \in \mathbb{N}$, we define the distance d_k between two measures μ, ν in $\mathbb{R}^d$ as

$$d_k(\mu, \nu) = \sup\left\{\left|\int f\,d\mu - \int f\,d\nu\right| : \|f\|_{k,\infty} \le 1\right\}, \tag{3.25}$$

where $\|f\|_{k,\infty} = \sum_{0\le|\alpha|\le k} \|\partial_\alpha f\|_\infty$. Then we have $d_0 = d_{\mathrm{TV}}$ and $d_1 = d_{\mathrm{FM}}$ (Fortet–Mourier distance).

The following result is a localized version of Lemma 2.5 in [3]. Here, γ_δ denotes the density of the centred normal law of covariance $\delta \times I$ on $\mathbb{R}^d$, $\delta > 0$ (I denoting the identity matrix) and $f * \gamma_\delta$ denotes the convolution between f and γ_δ.

Lemma 3.5. *Let Θ be a localizing r.v. as in (3.23). Then, for every $\varepsilon > 0$, $\delta > 0$, $X \in \mathcal{S}^d$ and for every bounded and measurable function $f : \mathbb{R}^d \to \mathbb{R}$ one has*

$$\big|\mathbb{E}_\Theta\big(f(X)\big) - \mathbb{E}_\Theta\big(f * \gamma_\delta(X)\big)\big| \le C\|f\|_\infty\left(\mathbb{P}_\Theta(\sigma_X < \varepsilon) + \frac{\sqrt{\delta}}{\varepsilon^p}\big(1 + A_{p,\Theta}(X)\big)^a\right), \tag{3.26}$$

where $A_{p,\Theta}(X)$ is defined in (3.24) and $C, p, a > 0$ are suitable universal constants depending on the dimension d only.

Proof. The proof is identical to the one of Lemma 2.5 in [3] (the case $\Theta \equiv 1$ being the same result): just consider the localized measure $\mathbb{P}_\Theta$ instead of $\mathbb{P}$ in the proof of Lemma 2.5 in [3] (namely, replace the localizing variable Θ_ε therein with $\Theta_\varepsilon\Theta$). □

We state now a variant of Theorem 2.7 in [3] that takes into account localizations.

Theorem 3.6. *Let Θ, U be localizing r.v.'s as in (3.23) and let $X, Y \in \mathcal{S}^d$ be such that $A_{l,\Theta}(X)$, $A_{l,U}(Y) < \infty$, such quantities being defined in (3.24). Let $\mu_{X,\Theta}$ denote the law of X under $\mathbb{P}_\Theta$ and let $\mu_{Y,U}$ denote the law of Y under $\mathbb{P}_U$. Let $k \in \mathbb{N}$. Then there exist some universal constants $C, p, a, b > 0$ (independent of Θ, U, X, Y, k) such that*

$$\begin{aligned} d_0(\mu_{X,\Theta}, \mu_{Y,U}) \le{}& \frac{C}{\varepsilon^b}\big(1 + A_{l,\Theta}(X) + A_{l,U}(Y)\big)^a\big(d_k(\mu_{X,\Theta}, \mu_{Y,U})\big)^{1/(k+1)} \\ &+ C\mathbb{P}_\Theta(\det\sigma_X < \varepsilon) + C\mathbb{P}_U(\det\sigma_Y < \varepsilon). \end{aligned} \tag{3.27}$$

Proof. We take a bounded and measurable function f and we write

$$\begin{aligned} \big|\mathbb{E}_\Theta\big(f(X)\big) - \mathbb{E}_U\big(f(Y)\big)\big| \le{}& \big|\mathbb{E}_\Theta\big(f(X)\big) - \mathbb{E}_\Theta\big(f * \gamma_\delta(X)\big)\big| + \big|\mathbb{E}_U\big(f(X)\big) - \mathbb{E}_U\big(f * \gamma_\delta(Y)\big)\big| \\ &+ \big|\mathbb{E}_\Theta\big(f * \gamma_\delta(X)\big) - \mathbb{E}_U\big(f * \gamma_\delta(Y)\big)\big| \\ =:{}& I_\Theta(X) + I_U(Y) + I_{\Theta,U}(X, Y). \end{aligned}$$

By using (3.26), we get

$$I_\Theta(X) + I_U(Y) \le C\|f\|_\infty\Big(\mathbb{P}_\Theta(\sigma_X < \varepsilon) + \mathbb{P}_U(\sigma_Y < \varepsilon) + \frac{\sqrt{\delta}}{\varepsilon^p}\big(1 + A_{p,\Theta}(X) + A_{p,U}(Y)\big)^a\Big).$$

Moreover, by recalling that $\|f * \gamma_\delta\|_{k,\infty} \le C\delta^{-k/2}\|f\|_\infty$, we have

$$I_{\Theta,U}(X,Y) \le C\delta^{-k/2}\|f\|_\infty d_k(\mu_{X,\Theta}, \mu_{Y,U}).$$

Following the proof of Theorem 2.7 in [3], we now insert everything, optimize w.r.t. δ and we get the result. □

Remark 3.7. Lemma 3.5 and Theorem 3.6 are valid not only with the basic noise $V_1, \dots, V_n$ introduced in Section 3.1. Actually, both results remains true whenever the basic noise fulfils the abstract integration by parts framework developed in Section 2.1 of [3], the one considered in this paper being a particular case.

We are finally ready for the following.

Proof of Theorem 2.5. Let G denote a standard normal r.v. in $\mathbb{R}^N$. For each $K \ge 1$ set

$$\Theta_{n,K} = \psi_K(S_n), \qquad d\mathbb{P}_{\Theta_{n,K}} = \Theta_{n,K}\, d\mathbb{P} \quad \text{and} \quad \Theta_K = \psi_K(G), \qquad d\mathbb{P}_{\Theta_K} = \Theta_K\, d\mathbb{P},$$

ψ_K being defined in (3.1). Let $\mu_{n,K}$ be the law of S_n under $\mathbb{P}_{\Theta_{n,K}}$ and μ_K be the law of G under $\mathbb{P}_{\Theta_K}$, that is,

$$\mu_{n,K}(dx) = \mathbb{P}_{\Theta_{n,K}}(S_n \in dx) \quad \text{and} \quad \mu_K(dx) = \mathbb{P}_{\Theta_K}(G \in dx).$$

Consider a measurable function $f : \mathbb{R}^N \to \mathbb{R}$ such that $\|f\|_\infty \le 1$. We write

$$\begin{aligned}\big|\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f(G)\big)\big| &\le \big|\mathbb{E}\big(f(S_n)(1-\Theta_{n,K})\big)\big| + \big|\mathbb{E}\big(f(G)(1-\Theta_K)\big)\big| \\ &\quad + \big|\mathbb{E}\big(f(S_n)\Theta_{n,K}\big) - \mathbb{E}\big(f(G)\Theta_K\big)\big|.\end{aligned}$$

Using the Chebyshev's inequality,

$$\big|\mathbb{E}\big(f(S_n)(1-\Theta_{n,K})\big)\big| \le \|f\|_\infty \mathbb{P}\big(|S_n| \ge 2K\big) \le \frac{C}{K^2}\|f\|_\infty$$

and a similar estimates holds for $|\mathbb{E}(f(G)(1-\Theta_K))|$. We conclude that

$$\sup_{\|f\|_\infty \le 1}\big|\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f(G)\big)\big| \le \frac{C}{K^2} + \sup_{\|f\|_\infty \le 1}\big|\mathbb{E}\big(f(S_n)\Theta_{n,K}\big) - \mathbb{E}\big(f(G)\Theta_K\big)\big|.$$

We obtain

$$\limsup_{n\to\infty}\sup_{\|f\|_\infty \le 1}\big|\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f(G)\big)\big| \le \frac{C}{K^2} + \limsup_{n\to\infty} d_{\mathrm{TV}}(\mu_{n,K}, \mu_K)$$

for every $K \geq 1$. If we show that, for each fixed K, $d_{\mathrm{TV}}(\mu_{n,K}, \mu_K) \to 0$ as $n \to \infty$, the statement will follow by letting K go to $+\infty$. So, we study $d_{\mathrm{TV}}(\mu_{n,K}, \mu_K)$, for a fixed $K > 1$.

We use Theorem 3.6 with $\Theta = \Theta_{n,K}$, $X = S_n$, $U = \Theta_K$ and $Y = G$. Here, the noise includes the Gaussian r.v. G, so we add it to the underlying noise (recall Remark 3.7) in a standard way – we stress this trick because it will be used also in the sequel, for example, in Lemma 4.12.

Without loss of generality, we assume that G is defined on the same probability space and is independent of $V_1, \dots, V_n$. We consider as basic noise the one coming from $(G, V_1, \dots, V_n)$. For $X = \phi(G, V_1, \dots, V_n)$ with $\phi \in C_b^\infty(\mathbb{R}^{N(1+n)}; \mathbb{R})$, we set

$$D_{(0,i)}X = \frac{\partial}{\partial G^i}\phi(G, V_1, \dots, V_n)$$

and $D_{(k,i)}$ for $k = 1, \dots, n$ as in (3.8). The Ornstein–Uhlenbeck generator takes into account the contribution from the standard Gaussian G, so it becomes

$$\begin{aligned} -LX &= \sum_{i=1}^N D_{(0,i)} D_{(0,i)} X - \sum_{i=1}^N D_{(0,i)} X G^i \\ &\quad + \sum_{k=1}^n \sum_{i=1}^N D_{(k,i)} D_{(k,i)} X + \sum_{k=1}^n \sum_{i=1}^N D_{(k,i)} X \partial_i \ln \psi_{r_0/2}\big(|V_k - v_0|\big). \end{aligned}$$

And if X is a random vector in $\mathbb{R}^d$, the associated Malliavin covariance matrix is

$$\sigma_X^{i,j} = \sum_{k=0}^n \sum_{r=1}^N D_{(k,r)} X^i \times D_{(k,r)} X^j, \qquad i, j = 1, \dots, d.$$

It is standard to see that the above quantities bring to an abstract Malliavin calculus as developed in [3]. Of course, when the randomness does not depend on G then everything agrees with what developed in Section 3.1 and when the randomness does not depend on V then we get the standard Gaussian–Malliavin calculus. So, we use Remark 3.7 and we apply Theorem 3.6. In order to use (3.27), we need to study $A_{\Theta_{n,K}}(S_n)$ and $A_{\Theta_K}(G)$. By (3.18) and by recalling that $1_{\{\Theta_{n,K} \neq 0\}} |S_n| \leq 2K$, we obtain

$$\|S_n\|_{q,p,\Theta_{n,K}} + \|LS_n\|_{q-2,p,\Theta_{n,K}} \leq CK.$$

Standard computations give $\|G\|_{q,p,\Theta_K} + \|LG\|_{q-2,p,\Theta_K} \leq \|G\|_{q,p} + \|LG\|_{q-2,p} \leq C$, so we can write

$$A_{\Theta_{n,K}}(S_n) + A_{\Theta_K}(G) \leq CK,$$

$C > 0$ being independent of K and n. Moreover, σ_G is the identity matrix. And since $|\Theta_{n,K}| \leq 1$ and $\chi_k, k \in \mathbb{N}$ are i.i.d., the law of large numbers says that for $\varepsilon^{1/N} < \mathbb{E}(\chi_k) = m_0$ one has

$$\limsup_{n\to\infty} \mathbb{P}_{\Theta_{n,K}}(\det \sigma_{S_n} \leq \varepsilon) \leq \limsup_{n\to\infty} \mathbb{P}_{\Theta_{n,K}}\big(\lambda_{S_n} \leq \varepsilon^{1/N}\big) \leq \limsup_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n \chi_k \leq \varepsilon^{1/N}\right) = 0$$

in which we have used (3.17). We apply now Theorem 3.6 with $k = 1$ and $\varepsilon < 1 \wedge m_0^N$: by passing to the limit in (3.27) we obtain

$$\limsup_{n\to\infty} d_0(\mu_{n,K}, \mu_K) \le \frac{C}{\varepsilon^a}(1 + CK)^b \limsup_{n\to\infty} d_{\mathrm{FM}}(\mu_{n,K}, \mu_K)^{1/2}.$$

So, it remains to show that $d_{\mathrm{FM}}(\mu_{n,K}, \mu_K) \to 0$ as $n \to \infty$. Since $\psi_K \in C_c(\mathbb{R}^N)$, the CLT gives

$$\lim_n \mathbb{E}_{\Theta_{n,K}}\big(f(S_n)\big) = \lim_n \mathbb{E}\big(\psi_K(S_n) f(S_n)\big) = \mathbb{E}\big(\psi_K(G) f(G)\big) = \mathbb{E}_{\Theta_K}\big(f(G)\big)$$

for every $f \in C(\mathbb{R}^d)$. So, if we define the probability laws

$$\hat\mu_{n,K}(dx) = \frac{1}{\mathbb{E}(\Theta_{n,K})}\mu_{n,K}(dx) \quad \text{and} \quad \hat\mu_K(dx) = \frac{1}{\mathbb{E}(\Theta_K)}\mu_K(dx),$$

we get $\hat\mu_{n,K} \to \hat\mu_K$ weakly as $n \to \infty$. Since weak convergence of probability laws is equivalent to convergence in d_{FM} (see, e.g., Theorem 11.3.3 in [8]), we have $d_{\mathrm{FM}}(\hat\mu_{n,K}, \hat\mu_K) \to 0$ as $n \to \infty$. Finally, straightforward computations give

$$d_{\mathrm{FM}}(\mu_{n,K}, \mu_K) \le \big|\mathbb{E}(\Theta_{n,K}) - \mathbb{E}(\Theta_K)\big| + d_{\mathrm{FM}}(\hat\mu_{n,K}, \hat\mu_K) \to 0$$

as $n \to \infty$, and the statement follows. □

Remark 3.8. We note that if $C(F)$ was not the identity matrix then (3.16) and (3.17) would become

$$\sigma_{S_n} = \frac{1}{n}\sum_{k=1}^n \chi_k \widehat{C}(F) \quad \text{and} \quad \lambda_{S_n} = \underline{\lambda}(F)\frac{1}{n}\sum_{k=1}^n \chi_k$$

respectively, where $\widehat{C}(F) = C(F)^{-1}$ and $\underline{\lambda}(F)$ is the smallest eigenvalue of $\widehat{C}(F)$. This means that the estimates in (3.19) and (3.20) continue to hold up to a multiplying constant that now depends on $\underline{\lambda}(F)$ and $\overline{\lambda}(F)$ as well, the latter denoting the largest eigenvalue of $\widehat{C}(F)$.

4. Asymptotic expansion

The aim of this section is to prove Theorems 2.6 and 2.7. We first study the case of smooth functions and then, using a regularizing argument, we will be able to deal with general functions.

4.1. The development for smooth test functions

We recall that we are assuming that the r.v. F has null mean and non-degenerate covariance matrix, that we have set equal to the identity matrix. And we have set

$$F_i = \chi_i V_i + (1 - \chi_i) W_i$$

so that $S_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n F_i = \frac{1}{\sqrt{n}}\sum_{i=1}^n(\chi_i V_i + (1-\chi_i)W_i)$. Moreover, we consider $G_i = (G_i^1,\dots,G_i^N)$, $i\in\mathbb{N}$, some independent standard normal random variables in $\mathbb{R}^N$. For $k\in\{0,1,\dots,n\}$, we define

$$S_n^k = \frac{1}{\sqrt{n}}\Bigg(\sum_{i=1}^k F_i + \sum_{i=k+1}^n G_i\Bigg), \qquad \widehat{S}_n^k = \frac{1}{\sqrt{n}}\Bigg(\sum_{i=1}^{k-1} F_i + \sum_{i=k+1}^n G_i\Bigg) \tag{4.1}$$

in which we use the convention that the sums are null when done on the indexes $i\in\{i_0,\dots,i_1\}$ with $i_0>i_1$. Therefore, one has

$$S_n^n = S_n \quad \text{and} \quad S_n^0 = \frac{1}{\sqrt{n}}\sum_{i=1}^n G_i$$

and S_n^0 is a standard normal random variable in $\mathbb{R}^N$. Moreover,

$$S_n^k = \widehat{S}_n^k + \frac{F_k}{\sqrt{n}} \quad \text{and} \quad S_n^{k-1} = \widehat{S}_n^k + \frac{G_k}{\sqrt{n}}. \tag{4.2}$$

In the sequel, we will use the following notation. For a multiindex $\alpha=(\alpha_1,\dots,\alpha_r)\in\{1,\dots,N\}^r$ and $x=(x^1,\dots,x^N)$ we denote $x^\alpha=\prod_{i=1}^r x^{\alpha_i}$. We also denote by $\partial_\alpha = \partial_{x^{\alpha_1}}\cdots\partial_{x^{\alpha_r}}$ the derivative corresponding to α and by $|\alpha|=r$ the length of α. We allow α to be the null multiindex: in this case, we set $|\alpha|=0$, $\partial_\alpha f = f$ and $x^\alpha=1$.

Moreover, we will use the following form of the Taylor formula of order $r\in\mathbb{N}$: for $f\in C^{r+1}(\mathbb{R}^N)$,

$$f(x+y) = f(x) + \sum_{p=1}^r \frac{1}{p!}\sum_{|\alpha|=p}\partial_\alpha f(x)y^\alpha + U_r f(x,y) \tag{4.3}$$

with

$$U_r f(x,y) = \frac{1}{r!}\sum_{|\alpha|=r+1} y^\alpha \int_0^1 (1-\lambda)^r \partial_\alpha f(x+\lambda y)\,d\lambda. \tag{4.4}$$

We notice that for some $c_r>0$ it holds

$$\big|U_r f(x,y)\big| \le c_r |y|^{r+1}\|f\|_{r+1,\infty}, \tag{4.5}$$

where $\|\cdot\|_{r+1,\infty}$ is the usual norm on $C_b^{r+1}(\mathbb{R}^N)$: $\|f\|_{r+1,\infty} = \sum_{|\alpha|\le r+1}\|\partial_\alpha f\|_\infty$.

For a multiindex $\alpha=(\alpha_1,\dots,\alpha_r)\in\{1,\dots,N\}^r$, that is, $|\alpha|=r$, we now set

$$\Delta_\alpha = \mathbb{E}\big(F^\alpha\big) - \mathbb{E}\big(G^\alpha\big) = \mathbb{E}\Bigg(\prod_{i=1}^r F^{\alpha_i}\Bigg) - \mathbb{E}\Bigg(\prod_{i=1}^r G^{\alpha_i}\Bigg) \tag{4.6}$$

and $\theta_\alpha=1$ if r is even and $\alpha_{2j-1}=\alpha_{2j}$ for every $j=1,\dots,r/2$, otherwise $\theta_\alpha=0$. For $r=0$, we have $\alpha=\emptyset$ and we set $\Delta_\emptyset=0$ and $\theta_\emptyset=1$. It is clear that $\Delta_\alpha=0$ for $|\alpha|\le 2$ and, for $r\ge 3$, the assumption $\sup_{|\alpha|\le r}|\Delta_\alpha|=0$ means that all moments of F up to order r (and not only up to order 2) agree with the moments of a standard Gaussian random variable.

We now introduce the basic differential operators which appear in the asymptotic expansion: we set

$$\Psi_t=\sum_{p=0}^{t}\frac{(-1)^{(t-p)/2}}{2^{(t-p)/2}p!((t-p)/2)!}\sum_{|\alpha|=p}\sum_{|\beta|=t-p}\Delta_\alpha\theta_\beta\partial_\beta\partial_\alpha,\qquad t=0,1,2,\dots. \tag{4.7}$$

Recall that θ_β is null when $t-p$ is odd, so the sum actually runs on the indexes p such that $(t-p)/2\in\mathbb{N}$. The property $\Delta_\alpha=0$ if $|\alpha|\le 2$ gives that the above sum actually starts from $p=3$, so we have

$$\Psi_t=0\qquad\text{if } t=0,1,2\quad\text{and}$$
$$\Psi_t=\sum_{p=3}^{t}\frac{(-1)^{(t-p)/2}}{2^{(t-p)/2}p!((t-p)/2)!}\sum_{|\alpha|=p}\sum_{|\beta|=t-p}\Delta_\alpha\theta_\beta\partial_\beta\partial_\alpha,\qquad t\ge 3.$$

From now on, we use the convention $\sum_{p=3}^{t}(\cdot)=0$ if $t<3$. So, for example we can write

$$\Psi_t=\sum_{p=3}^{t}\frac{(-1)^{(t-p)/2}}{2^{(t-p)/2}p!((t-p)/2)!}\sum_{|\alpha|=p}\sum_{|\beta|=t-p}\Delta_\alpha\theta_\beta\partial_\beta\partial_\alpha,\qquad t=0,1,2,\dots.$$

We note that $\Psi_t=0$ for all t when $\Delta_\alpha=0$ for all α, that is when all the moments of F agree with the moments of the standard Gaussian law. And moreover, for every $t\ge 3$ and $q\ge 0$ there exists $C_{t,q}>0$ such that if $f\in C_b^{t+q}$ then

$$\|\Psi_t f\|_{q,\infty}\le C_{t,q}\sup_{|\alpha|=t}|\Delta_\alpha|\times\|f\|_{t+q,\infty}. \tag{4.8}$$

We also define the following objects ("remainders"): for $r\in\mathbb{N}$,

$$\begin{aligned}\mathcal{R}_{r,n}^k f={}&\sum_{p=3}^{r}n^{-([(r-p)/2]+1/2-(r-p)/2)}\frac{(-1)^{[(r-p)/2]+1}}{2^{[(r-p)/2]+1}p![(r-p)/2]!}\\&\times\sum_{|\alpha|=p}\sum_{|\beta|=2[(r-p)/2]+2}\Delta_\alpha\theta_\beta\int_0^1 s^{[(r-p)/2]}\mathbb{E}\Big(\partial_\beta\,\partial_\alpha f\Big(\widehat{S}_n^k+\sqrt{s}\,\frac{G_k}{\sqrt{n}}\Big)\Big)\,ds\\&+n^{(r+1)/2}\Big[\mathbb{E}\Big(U_r f\Big(\widehat{S}_n^k,\frac{F_k}{\sqrt{n}}\Big)\Big)-\mathbb{E}\Big(U_r f\Big(\widehat{S}_n^k,\frac{G_k}{\sqrt{n}}\Big)\Big)\Big],\end{aligned} \tag{4.9}$$

$U_r f$ being defined in (4.4). As usual, the first term of the above r.h.s. is set equal to zero if $r<3$. Moreover, $[(r-p)/2]+1/2-(r-p)/2\in\{0,1/2\}$, hence $n^{-([(r-p)2]+1/2-(r-p)/2)}\le 1$.

Remark 4.1. We note here if $F \in L^2$ then for every $f \in C_b^2$ one has

$$\mathcal{R}_{0,n}^k f = \frac{1}{\sqrt{n}} \mathcal{R}_{1,n}^k f.$$

And if $F \in L^3(\Omega)$ then for every $f \in C_b^3$ one has

$$\mathcal{R}_{0,n}^k f = \frac{1}{\sqrt{n}} \mathcal{R}_{1,n}^k f = \frac{1}{n} \mathcal{R}_{2,n}^k f. \tag{4.10}$$

In fact, for every $r \geq 0$, if $f \in C_b^{r+2}$ then

$$U_r f(x,y) = U_{r+1} f(x,y) - \frac{1}{(r+1)!} \sum_{|\alpha|=r+1} y^\alpha \partial_\alpha f(x).$$

Therefore, for $r = 0$, $F \in L^2$ and $f \in C_b^2$ we obtain

$$\begin{aligned}\mathcal{R}_{0,n}^k f = {} & \sqrt{n}\Big[\mathbb{E}\Big(U_1 f\Big(\widehat{S}_n^k, \frac{F_k}{\sqrt{n}}\Big)\Big) - \mathbb{E}\Big(U_1 f\Big(\widehat{S}_n^k, \frac{G_k}{\sqrt{n}}\Big)\Big)\Big] \\ & - \sqrt{n} \sum_{|\alpha|=1} \mathbb{E}\Big(\Big[\Big(\frac{F_k}{\sqrt{n}}\Big)^\alpha - \Big(\frac{G_k}{\sqrt{n}}\Big)^\alpha\Big] f\big(\widehat{S}_n^k\big)\Big).\end{aligned}$$

Since $\widehat{S}_n^k$ is independent of F_k and G_k and since $\Delta_\alpha = 0$ for $|\alpha| = 1$ we get $\mathbb{E}([(F_k)^\alpha - (G_k)^\alpha] f(\widehat{S}_n^k)) = \Delta_\alpha \mathbb{E}(f(\widehat{S}_n^k)) = 0$, so that

$$\mathcal{R}_{0,n}^k f = \frac{1}{\sqrt{n}} \mathcal{R}_{1,n}^k f.$$

As for (4.10), one uses $\Delta_\alpha = 0$ for $|\alpha| = 2$ and the statement is proved similarly.

Since $\mathbb{E}(f(S_n)) - \mathbb{E}(f(G)) = \mathbb{E}(f(S_n^n)) - \mathbb{E}(f(S_n^0))$, we study $\mathbb{E}(f(S_n^k)) - \mathbb{E}(f(S_n^{k-1}))$ for $k = 1, \dots, n$ and then apply a recurrence argument.

Lemma 4.2. *Let $n \in \mathbb{N}$, $1 \leq k \leq n$ and $r \in \mathbb{N}$. If $F \in L^{r+1}(\Omega)$ then for every $f \in C_b^{r+1}(\mathbb{R}^N)$ one has*

$$\mathbb{E}\big(f\big(S_n^k\big)\big) - \mathbb{E}\big(f\big(S_n^{k-1}\big)\big) = \sum_{p=3}^r \frac{1}{p! n^{p/2}} \sum_{|\alpha|=p} \mathbb{E}\big(\partial_\alpha f\big(\widehat{S}_n^k\big)\big)\Delta_\alpha + \frac{1}{n^{(r+1)/2}} \widetilde{\mathcal{R}}_{r,n}^k f, \tag{4.11}$$

where

$$\widetilde{\mathcal{R}}_{r,n}^k f = n^{(r+1)/2}\Big[\mathbb{E}\Big(U_r f\Big(\widehat{S}_n^k, \frac{F_k}{\sqrt{n}}\Big)\Big) - \mathbb{E}\Big(U_r f\Big(\widehat{S}_n^k, \frac{G_k}{\sqrt{n}}\Big)\Big)\Big].$$

Proof. We will use the Taylor formula (4.3). Since $S_n^k = \widehat{S}_n^k + \frac{F_k}{n^{1/2}}$ and F_k is independent of $\widehat{S}_n^k$, we obtain

$$\mathbb{E}\big(f\big(S_n^k\big)\big) = \mathbb{E}\big(f\big(\widehat{S}_n^k\big)\big) + \sum_{p=1}^{r} \frac{1}{p!n^{p/2}} \sum_{|\alpha|=p} \mathbb{E}\big(\partial_\alpha f\big(\widehat{S}_n^k\big)\big)\mathbb{E}\big(F_k^\alpha\big) + \mathbb{E}\left(U_r f\left(\widehat{S}_n^k, \frac{F_k}{n^{1/2}}\right)\right).$$

We now use that $S_n^{k-1} = \widehat{S}_n^k + \frac{G_k}{n^{1/2}}$: the same reasoning for G_k gives

$$\mathbb{E}\big(f\big(S_n^{k-1}\big)\big) = \mathbb{E}\big(f\big(\widehat{S}_n^k\big)\big) + \sum_{p=1}^{r} \frac{1}{p!n^{p/2}} \sum_{|\alpha|=p} \mathbb{E}\big(\partial_\alpha f\big(\widehat{S}_n^k\big)\big)\mathbb{E}\big(G_k^\alpha\big) + \mathbb{E}\left(U_r f\left(\widehat{S}_n^k, \frac{G_k}{n^{1/2}}\right)\right).$$

By recalling that $\Delta_\alpha = \mathbb{E}(F^\alpha) - \mathbb{E}(G^\alpha) = 0$ for $|\alpha| \leq 2$, the statement holds. □

Our aim is now to replace $\widehat{S}_n^k$ by S_n^{k-1} in the development (4.11). This opens the way to use a recurrence procedure.

Lemma 4.3. *Let $n \in \mathbb{N}$, $1 \leq k \leq n$ and $r \in \mathbb{N}$. If $F \in L^{r+1}(\Omega)$ then for every $f \in C_b^{r+2}(\mathbb{R}^N)$ one has*

$$\mathbb{E}\big(f\big(S_n^k\big)\big) - \mathbb{E}\big(f\big(S_n^{k-1}\big)\big) = \sum_{t=3}^{r} \frac{1}{n^{t/2}} \mathbb{E}\big(\Psi_t f\big(S_n^{k-1}\big)\big) + \frac{1}{n^{(r+1)/2}} \mathcal{R}_{r,n}^k f,$$

where Ψ_t and $\mathcal{R}_{r,n}^k$ are defined in (4.7) and (4.9), respectively.

Proof. Consider the generical term $\mathbb{E}(\partial_\alpha f(\widehat{S}_n^k))$ of (4.11). We recall that $\widehat{S}_n^k + G_k/\sqrt{n} = S_n^{k-1}$ and that $\widehat{S}_n^k$ and G_k are independent. So, we apply the backward Taylor's formula in (C.1) to $g(x) = \partial_\alpha f(\widehat{S}_n^k + x/\sqrt{n})$ with $|\alpha| = p \leq r$, and we expand up to order $[(r-p)/2]$. Hence we can write

$$\begin{aligned}
&\mathbb{E}\big(\partial_\alpha f\big(\widehat{S}_n^k\big)\big) \\
&= \sum_{q=0}^{[(r-p)/2]} \frac{(-1)^q}{2^q q! n^q} \sum_{|\beta|=2q} \theta_\beta \mathbb{E}\big(\partial_\beta \partial_\alpha f\big(S_n^{k-1}\big)\big) + \frac{1}{n^{[(r-p)/2]+1}} \widetilde{U}_{[(r-p)/2]} \partial_\alpha f\left(S_n^{k-1}, \frac{G_k}{\sqrt{n}}\right),
\end{aligned}$$

where

$$\widetilde{U}_q g\left(\widehat{S}_n^k, \frac{G_k}{\sqrt{n}}\right) = \frac{(-1)^{q+1}}{2^{q+1} q!} \sum_{|\beta|=2q+2} \theta_\beta \int_0^1 s^q \mathbb{E}\left(\partial_\beta g\left(\widehat{S}_n^k + \sqrt{s}\frac{G_k}{\sqrt{n}}\right)\right) ds.$$

By inserting in (4.11), we get

$$\begin{aligned}
&\mathbb{E}\big(f\big(S_n^k\big)\big) - \mathbb{E}\big(f\big(S_n^{k-1}\big)\big) \\
&= \sum_{p=3}^{r} \frac{1}{p!n^{p/2}} \sum_{|\alpha|=p} \Delta_\alpha \sum_{q=0}^{[(r-p)/2]} \frac{(-1)^q}{2^q q! n^q} \sum_{|\beta|=2q} \theta_\beta \mathbb{E}\big(\partial_\beta \partial_\alpha f\big(S_n^{k-1}\big)\big)
\end{aligned}$$

$$+\sum_{p=3}^{r}\frac{1}{p!n^{p/2}}\sum_{|\alpha|=p}\Delta_\alpha\frac{1}{n^{[(r-p)/2]+1}}\widetilde{U}_{[(r-p)/2]}\partial_\alpha f\left(\widehat{S}_n^k,\frac{G_k}{\sqrt{n}}\right)+\frac{1}{n^{(r+1)/2}}\widetilde{\mathcal{R}}_{r,n}^k f$$

$$=\sum_{p=0}^{r}\sum_{q=0}^{[(r-p)/2]}\frac{(-1)^q}{2^q p!q!n^{(p+2q)/2}}\sum_{|\alpha|=p}\sum_{|\beta|=2q}\mathbb{E}\big(\partial_\beta\partial_\alpha f(S_n^{k-1})\big)\theta_\beta\Delta_\alpha+\frac{1}{n^{(r+1)/2}}\mathcal{R}_{r,n}^k f$$

in which, for the last line, we have used (4.9) and in the first sum we can let the index p start from 0 because as $p=0,1,2$, $\Delta_\alpha=0$. Now, by considering the change of variable $t=p+q$ in the first term, we get

$$\begin{aligned}
&\mathbb{E}\big(f\big(S_n^k\big)\big)-\mathbb{E}\big(f\big(S_n^{k-1}\big)\big)\\
&=\sum_{t=0}^{r}\sum_{s=0}^{t}\frac{(-1)^{(t-s)/2}}{2^{(t-s)/2}s!((t-s)/2)!n^{t/2}}\sum_{|\alpha|=s}\sum_{|\beta|=t-s}\mathbb{E}\big(\partial_\beta\partial_\alpha f\big(S_n^{k-1}\big)\big)\Delta_\alpha\theta_\beta+\frac{1}{n^{(r+1)/2}}\mathcal{R}_{r,n}^k f\\
&=\sum_{t=0}^{r}\frac{1}{n^{t/2}}\mathbb{E}\left(\sum_{s=0}^{t}\frac{(-1)^{(t-s)/2}}{2^{(t-s)/2}s!((t-s)/2)!}\sum_{|\alpha|=s}\sum_{|\beta|=t-s}\partial_\beta\partial_\alpha f\big(S_n^{k-1}\big)\Delta_\alpha\theta_\beta\right)+\frac{1}{n^{(r+1)/2}}\mathcal{R}_{r,n}^k f\\
&=\sum_{t=0}^{r}\frac{1}{n^{t/2}}\mathbb{E}\big(\Psi_t f\big(S_n^{k-1}\big)\big)+\frac{1}{n^{(r+1)/2}}\mathcal{R}_{r,n}^k f.
\end{aligned}$$

Since $\Psi_t=0$ for $t\le 2$, the statement holds. □

For $k=1,\dots,n$, we define

$$\Psi_t^{(1)}=\Psi_t\quad\text{and}\quad\text{for } k\ge 2,\qquad \Psi_t^{(k)}=\Psi_t^{(k-1)}+\sum_{p=0}^{t}\Psi_p\Psi_{t-p}^{(k-1)},\qquad t=0,1,\dots.\tag{4.12}$$

Notice that $\Psi_t^{(k)}$ is a differential operator which is linked to the convolution w.r.t. t between $\Psi_\cdot$ and the preceding operator $\Psi_\cdot^{(k-1)}$. We also notice that $\Psi_t^{(k)}=0$ for $t=0,1,2$, as an immediate consequence of the fact that $\Psi_t=0$ for $t\le 2$. So, for $k\ge 2$ we can write

$$\Psi_t^{(k)}=\mathbf{1}_{\{t\ge 3\}}\Psi_t^{(k-1)}+\mathbf{1}_{\{t\ge 6\}}\sum_{p=3}^{t-3}\Psi_p\Psi_{t-p}^{(k-1)},\qquad t=0,1,\dots.\tag{4.13}$$

We also define the following reminder operators: for $r\in\mathbb{N}$,

$$\Phi_{r,n}^{(k)}f=\sum_{j=1}^{k-1}\sum_{t=0}^{r}\mathcal{R}_{r-t,n}^{k-j}\Psi_t^{(j)}f+\mathcal{R}_{r,n}^k f.\tag{4.14}$$

Note that, by definition, $\Phi_{r,n}^{(0)}=\mathcal{R}_{r,n}^0$ and $\Phi_{0,n}^{(k)}=\mathcal{R}_{0,n}^k$.

Lemma 4.4. *Let $n \in \mathbb{N}$, $1 \le k \le n$ and $r \in \mathbb{N}$. If $F \in L^{r+1}(\Omega)$ then for every $f \in C_b^{r+2}(\mathbb{R}^N)$ one has*

$$\mathbb{E}\big(f\big(S_n^k\big)\big) - \mathbb{E}\big(f\big(S_n^{k-1}\big)\big) = \sum_{t=3}^{r} \frac{1}{n^{t/2}} \mathbb{E}\big(\Psi_t^{(k)} f\big(S_n^0\big)\big) + \frac{1}{n^{(r+1)/2}} \Phi_{r,n}^{(k)} f,$$

$\Psi_t^{(k)}$ and $\Phi_{r,n}^{(k)}$ being given in (4.12) and (4.14), respectively.

Proof. We consider the development in Lemma 4.3:

$$\mathbb{E}\big(f\big(S_n^k\big)\big) - \mathbb{E}\big(f\big(S_n^{k-1}\big)\big) = \sum_{t=0}^{r} \frac{1}{n^{t/2}} \mathbb{E}\big(\Psi_t f\big(S_n^{k-1}\big)\big) + \frac{1}{n^{(r+1)/2}} \mathcal{R}_{r,n}^k f.$$

For $t \le r$, we apply such development up to order $r-t$ to $\mathbb{E}(\Psi_t f(S_n^{k-1}))$ and we get

$$\mathbb{E}\big(\Psi_t f\big(S_n^{k-1}\big)\big) = \mathbb{E}\big(\Psi_t f\big(S_n^{k-2}\big)\big) + \sum_{p=0}^{r-t} \frac{1}{n^{p/2}} \mathbb{E}\big(\Psi_p \Psi_t f\big(S_n^{k-2}\big)\big) + \frac{1}{n^{(r-t+1)/2}} \mathcal{R}_{r-t,n}^{k-1} \Psi_t f.$$

By inserting, we obtain

$$\begin{aligned}\mathbb{E}\big(f\big(S_n^k\big)\big) - \mathbb{E}\big(f\big(S_n^{k-1}\big)\big) &= \sum_{t=0}^{r} \frac{1}{n^{t/2}} \mathbb{E}\big(\Psi_t f\big(S_n^{k-2}\big)\big) + \sum_{t=0}^{r}\sum_{p=0}^{r-t} \frac{1}{n^{(t+p)/2}} \mathbb{E}\big(\Psi_p \Psi_t f\big(S_n^{k-2}\big)\big) \\ &\quad + \frac{1}{n^{(r+1)/2}} \sum_{t=0}^{r} \mathcal{R}_{r-t,n}^{k-1} \Psi_t f + \frac{1}{n^{(r+1)/2}} \mathcal{R}_{r,n}^k f\end{aligned}$$

and by a change of variable in the second sum above we get

$$\mathbb{E}\big(f\big(S_n^k\big)\big) - \mathbb{E}\big(f\big(S_n^{k-1}\big)\big) = \sum_{t=0}^{r} \frac{1}{n^{t/2}} \mathbb{E}\big(\Psi_t^{(2)} f\big(S_n^{k-2}\big)\big) + \frac{1}{n^{(r+1)/2}} \left[\sum_{t=0}^{r} \mathcal{R}_{r-t,n}^{k-1} \Psi_t f + \mathcal{R}_{r,n}^k f\right].$$

By iterating the same procedure up to step k, we obtain the statement. □

We now set

$$T_t^n = \sum_{k=1}^{n} \Psi_t^{(k)} \quad \text{and} \quad \mathcal{U}_r^n = \sum_{k=1}^{n} \Phi_{r,n}^{(k)} \tag{4.15}$$

$\Psi_t^{(k)}$ and $\Phi_{r,n}^{(k)}$ being given in (4.12) and (4.14), respectively.

Proposition 4.5. *Let $n \in \mathbb{N}$, $1 \le k \le n$ and $r \in \mathbb{N}$. If $F \in L^{r+1}(\Omega)$ then for every $f \in C_b^{r+2}(\mathbb{R}^N)$ one has*

$$\mathbb{E}\big(f\big(S_n^n\big)\big) - \mathbb{E}\big(f\big(S_n^0\big)\big) = \sum_{t=3}^{r} \frac{1}{n^{t/2}} \mathbb{E}\big(T_t^n f\big(S_n^0\big)\big) + \frac{1}{n^{(r+1)/2}} \mathcal{U}_r^n f$$

T_t^n and $\mathcal{U}_r^n$ are defined in (4.15).

Proof. Since $\mathbb{E}(f(S_n^n)) - \mathbb{E}(f(S_n^0)) = \sum_{k=1}^n (\mathbb{E}(f(S_n^k)) - \mathbb{E}(f(S_n^{k-1})))$, the statement immediately follows from Lemma 4.4. $\square$

We give now an explicit expression for the operators $\Psi_t^{(k)}$ in (4.12) and, as a consequence, for T_t^n in (4.15). For Ψ_t given in (4.7), $i = 1, 2, \dots,$ we set

$$\mathcal{A}_t^1 = \Psi_t \quad \text{and} \quad \text{for } i \geq 1, \qquad \mathcal{A}_t^{i+1} = \sum_{p=0}^{t} \Psi_p \mathcal{A}_{t-p}^i.$$

Since $\Psi_t = 0$ for $t = 0, 1, 2$, straightforward computations give that $\mathcal{A}_t^i = 0$ if $t < 3i$, so that we can also write

$$\mathcal{A}_t^1 = \Psi_t \quad \text{and} \quad \text{for } i \geq 1, \qquad \mathcal{A}_t^{i+1} = \sum_{p=3}^{t-3i} \Psi_p \mathcal{A}_{t-p}^i. \tag{4.16}$$

We can give an alternative representation for the $\mathcal{A}_t^i$'s. We set $\mathcal{M}$ the set of all multiindexes and for $\alpha, \beta \in \mathcal{M}$ (possibly with different length), we set $(\alpha, \beta) \in \mathcal{M}$ the associated concatenation. So, for $\gamma \in \mathcal{M}$ we define

$$A_\gamma = \{(\alpha, \beta) : (\alpha, \beta) = \gamma\}$$

and

$$\begin{aligned} c_\gamma^1 &= \sum_{(\alpha,\beta)\in A_\gamma} \frac{(-1)^{|\beta|/2}}{2^{|\beta|/2}|\alpha|!(|\beta|/2)!} \Delta_\alpha \theta_\beta \quad \text{and} \quad \text{for } i \geq 1, \\ c_\gamma^{i+1} &= \sum_{(\alpha,\beta)\in A_\gamma} c_\alpha^1 c_\beta^i, \qquad i \geq 1. \end{aligned} \tag{4.17}$$

Since $c_\gamma^1 = 0$ if $|\gamma| < 3$, by recurrence one gets $c_\gamma^i = 0$ if $|\gamma| < 3i$ for every i. Then straightforward computations give that, for $i \geq 1$,

$$\mathcal{A}_t^i = \sum_{\gamma:|\gamma|=t} c_\gamma^i \partial_\gamma \qquad \text{with } \{c_\gamma^i\}_{\gamma\in\mathcal{M}} \text{ given in (4.17)}. \tag{4.18}$$

It is immediate to see that for every $\gamma \in \mathcal{M}$ there exists C such that for every $i \geq 1$

$$|c_\gamma^i| \leq C \sup_{|\alpha|\leq|\gamma|} |\Delta_\alpha|^i. \tag{4.19}$$

As a consequence, for $t, q \geq 0$ there exists $C > 0$ (depending on t, q only) such that for every $i \geq 1$ and $f \in C_b^{t+q}(\mathbb{R}^N)$

$$\|\mathcal{A}_t^i f\|_{q,\infty} \leq C \sup_{|\alpha|\leq t} |\Delta_\alpha|^i \times \|f\|_{t+q,\infty} \leq C(1 + \mathbb{E}(|F|^t))^{i-1} \sup_{|\alpha|\leq t} |\Delta_\alpha| \times \|f\|_{t+q,\infty}. \tag{4.20}$$

Moreover, the $\mathcal{A}_t^i$'s give the following representation formula for the $\Psi_t^{(k)}$'s.

Proposition 4.6. *For every $k \geq 1$ the operator $\Psi^{(k)}$ given in (4.12) can be written as*

$$\Psi_t^{(k)} = \sum_{i=1}^{[t/3]} Q_{i-1}(k)\mathcal{A}_t^i, \qquad t = 0, 1, \dots,$$

where $Q_{i-1}(k)$ is defined as follows:

$$Q_0(k) = 1 \quad \textit{and} \quad \textit{for } l \geq 1, \qquad Q_l(k) = \sum_{j=l+1}^{k} Q_{l-1}(j-1).$$

In particular, $Q_l(k) = 0$ if $k \leq l$ and $Q_l(k) > 0$ otherwise.

Proof. We have already observed that if $[t/3] = 0$ then $\Psi_t^{(k)} = \Psi_t = 0$ for every k and if $[t/3] = 1$ then $\Psi_t^{(k)} = \Psi_t$ for every k (see (4.13)), so the formulas agree. We now assume that the formula is true for $[t/3] = j \geq 1$ and for every k, and we prove it for $[t/3] = j+1$ and for every k. We recall that $\Psi_t^{(k)} = \Psi_t^{(k-1)} + \sum_{p=3}^{t-3} \Psi_p \Psi_{t-p}^{(k-1)}$. But if $[t/3] = j+1$ then $[(t-p)/3] \leq j$ for any $p = 3, \dots, t-3$, so that by induction $\Psi_{t-p}^{(k-1)}$ fulfils the formula. Therefore, we can write

$$\Psi_t^{(k)} = \Psi_t^{(k-1)} + \sum_{p=3}^{t-3} \sum_{i=1}^{[(t-p)/3]} Q_{i-1}(k-1)\Psi_p A_{t-p}^i.$$

We do a change of variable in the last sum: the condition $i \leq [(t-p)]/3$ gives $3i \leq t-p$, that is, $p \leq t-3i$, and if $p \geq 3$ then $i \leq [t/3]-1$. So, by using also (4.16) we get

$$\begin{aligned}
\Psi_t^{(k)} - \Psi_t^{(k-1)} &= \sum_{i=1}^{[t/3]-1} Q_{i-1}(k-1) \sum_{p=3}^{t-3i} \Psi_p \mathcal{A}_{t-p}^i \\
&= \sum_{i=1}^{[t/3]-1} Q_{i-1}(k-1)\mathcal{A}_t^{i+1} = \sum_{i=2}^{[t/3]} Q_{i-2}(k-1)\mathcal{A}_t^i.
\end{aligned}$$

By summing

$$\Psi_t^{(k)} = \Psi_t + \sum_{i=2}^{[t/3]} \sum_{j=2}^{k} Q_{i-2}(j-1)\mathcal{A}_t^i = Q_0(k)\mathcal{A}_t^1 + \sum_{i=2}^{[t/3]} \sum_{j=2}^{k} Q_{i-2}(j-1)\mathcal{A}_t^i$$

and the statement holds for $Q_0(k) = 1$ and $Q_{i-1}(k) = \sum_{j=2}^{k} Q_{i-2}(j-1)$, $i \geq 2$. We now prove that $Q_l(k) = 0$ if $k \leq l$ and $Q_l(k) > 0$ for $k \geq l+1$. For $l = 1$, $Q_l(k) = k-1$, and the statement holds. If we assume that $Q_l(k)$ is not null for $k \geq l+1$ then

$$Q_{l+1}(k) = \sum_{j=2}^{k} Q_l(j-1)\mathbf{1}_{j-1\geq l+1} = \sum_{j=2}^{k} Q_l(j-1)\mathbf{1}_{j\geq l+2}$$

and this is null for $k \leq l+1$ and strictly positive if $k \geq l+2$. □

We now give an explicit formula for T_t^n, namely we write it in such a way that $n \mapsto T_t^n$ is a polynomial whose coefficients will be explicitly written. To this purpose, we need to handle polynomials of the type

$$n \mapsto \mathcal{S}_l(n-1) = \sum_{k=1}^{n-1} k^l, \qquad l \in \mathbb{N}, n \geq 1.$$

We recall the exact expansion for $\mathcal{S}_l(L) = \sum_{k=1}^{L} k^l$:

$$\mathcal{S}_l(L) = \frac{1}{l+1} \sum_{p=1}^{l+1} \binom{l+1}{p} B_{l+1-p} L^p, \tag{4.21}$$

where $\{B_m\}_m$ denotes the sequence of the (second) Bernoulli numbers (which are in fact defined as the numbers for which the above equality holds, see [1]), whose first numbers are given by

$$B_0 = 1, \qquad B_1 = \frac{1}{2}, \qquad B_2 = \frac{1}{6}, \qquad B_3 = 0, \qquad B_4 = -\frac{1}{30}, \qquad B_5 = 0, \qquad B_6 = \frac{1}{42},$$
$$B_7 = 0, \qquad B_8 = -\frac{1}{30}, \ldots.$$

Then straightforward computations give that for $l \in \mathbb{N}$ and $n \geq 1$,

$$\mathcal{S}_l(n-1) = \sum_{k=1}^{n-1} k^l = \sum_{q=0}^{l+1} b_{l,q} n^q,$$

where the sequence $(b_{l,q})_{q=0,\ldots,l+1}$ is given by

$$b_{l,q} = \frac{1}{l+1} \sum_{p=q\vee 1}^{l+1} \binom{l+1}{p} B_{l+1-p} \binom{p}{q} (-1)^{p-q}, \qquad q = 0, 1, \ldots, l+1 \text{ and } l \in \mathbb{N}, \tag{4.22}$$

in which B_l, $l \geq 0$, denote the (second) Bernoulli numbers. Just as an example:

- $l = 0$: $b_{0,0} = -1, b_{0,1} = 1$;
- $l = 1$: $b_{1,0} = 0$, $b_{1,1} = -\frac{1}{2}$, $b_{1,2} = \frac{1}{2}$;
- $l = 2$: $b_{2,0} = 0$, $b_{2,1} = \frac{1}{6}$, $b_{2,2} = -\frac{1}{2}$, $b_{2,3} = \frac{1}{3}$.

Then one has the following.

Proposition 4.7. *Let $n \geq 1$, $r \in \mathbb{N}$ and $F \in L^{q_r+1}(\Omega)$, where $q_r = \max(r, 2)$. For $t \leq r$, let T_t^n be defined as in (4.15). Then*

$$T_t^n = \sum_{i=1}^{[t/3]} P_i(n) \mathcal{A}_t^i, \qquad t = 0, 1, \ldots,$$

where $P_i(n) = 0$ *if* $n < i$ *and for* $n \geq i$,

$$P_i(n) = \sum_{p=0}^{i} a_{i,p} n^p, \qquad i = 1, \dots, n \tag{4.23}$$

with

$$\begin{aligned} &a_{1,0} = 0, \qquad a_{1,1} = 1 \quad \textit{and} \quad \textit{for } i \geq 1, \\ &a_{i+1,0} = \sum_{l=0}^{i} a_{i,l} b_{l,0} - \sum_{l=0}^{i} a_{i,l} \mathcal{S}_l(i-1), \qquad a_{i+1,p} = \sum_{l=p-1}^{i} a_{i,l} b_{l,p}, \qquad p = 1, \dots, i \end{aligned} \tag{4.24}$$

the sequence $(b_{l,p})_{p=0,\dots,l+1}$ *being defined in* (4.22) *and* $\mathcal{S}_l(i-1)$ *being given in* (4.21).

Proof. Since $T_t^n = \sum_{k=1}^n \Psi_t^{(k)}$, we get

$$T_t^n = \sum_{i=1}^{[t/3]} \sum_{k=1}^{n} Q_{i-1}(k) \mathcal{A}_t^i$$

so that $P_i(n) = \sum_{k=1}^n Q_{i-1}(k) = \sum_{j=2}^{n+1} Q_{i-1}(j-1) = Q_i(n+1)$. As a consequence, $P_i(n) = 0$ if $n+1 \leq i$, that is $n < i$. So, let $n \geq i$. We have $P_1(n) = \sum_{k=1}^n Q_0(k) = n$ and for $i \geq 2$,

$$P_i(n) = Q_i(n+1) = \sum_{j=2}^{n+1} Q_{i-1}(j-1) \mathbf{1}_{j-1 \geq i} = \sum_{k=i-1}^{n-1} Q_{i-1}(k+1) = \sum_{k=i-1}^{n-1} P_{i-1}(k). \tag{4.25}$$

Since $P_1(n) = n$, we get $a_{1,0} = 0$ and $a_{1,1} = 1$. In order to compute the sequence $(a_{i,l})_{l=0,\dots,i}$, we use a recurrence argument. For $i \geq 1$, one has

$$\begin{aligned} P_{i+1}(n) &= \sum_{k=i}^{n-1} P_i(k) = \sum_{k=i}^{n-1} \sum_{l=0}^{i} a_{i,l} k^l = \sum_{l=0}^{i} a_{i,l} \sum_{k=i}^{n-1} k^l = \sum_{l=0}^{i} a_{i,l} \big(\mathcal{S}_l(n-1) - \mathcal{S}_l(i-1) \big) \\ &= \sum_{l=0}^{i} a_{i,l} \mathcal{S}_l(n-1) - \sum_{l=0}^{i} a_{i,l} \mathcal{S}_l(i-1) = \sum_{l=0}^{i} a_{i,l} \sum_{p=0}^{l+1} b_{l,p} n^p - \sum_{l=0}^{i} a_{i,l} \mathcal{S}_l(i-1) \\ &= \sum_{p=0}^{i+1} n^p \sum_{l=0 \vee (p-1)}^{i} a_{i,l} b_{l,p} - \sum_{l=0}^{i} a_{i,l} \mathcal{S}_l(i-1) \end{aligned}$$

and (4.24) follows. □

We are now ready to prove our result on the asymptotic expansion for smooth functions. We set:

- for $m \geq 1$ and $f \in C_b^m(\mathbb{R}^N)$,

$$\mathcal{D}_m f = \sum_{\substack{t=3\vee m \\ t-m \text{ even}}}^{3m} \sum_{i=1\vee(t-m)/2}^{[t/3]} a_{i,(t-m)/2} \mathbb{E}\big(\mathcal{A}_t^i f(G)\big); \tag{4.26}$$

- for $r \geq 2$ and $f \in C_b^{r+1}(\mathbb{R}^N)$,

$$\begin{aligned} \mathcal{E}_r^n f &= n^{([r/3]+1)/2} \\ &\quad \times \Bigg[\sum_{m=[r/3]+1}^{r} \frac{1}{n^{m/2}} \sum_{\substack{t=3\vee m \\ t-m \text{ even}}}^{(3m)\wedge r} \sum_{i=1\vee(t-m)/2}^{[t/3]} a_{i,(t-m)/2} \mathbb{E}\big(\mathcal{A}_t^i f(G)\big) + \frac{1}{n^{(r+1)/2}} \mathcal{U}_r^n f \Bigg]. \end{aligned} \tag{4.27}$$

Then we have:

Theorem 4.8. *Let $r \geq 2$. If $F \in L^{r+1}(\Omega)$, then for every $f \in C_b^{r+3}(\mathbb{R}^N)$ one has*

$$\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f(G)\big) = \sum_{m=1}^{[r/3]} \frac{1}{n^{m/2}} \mathcal{D}_m f + \frac{1}{n^{([r/3]+1)/2}} \mathcal{E}_r^n f,$$

where $\mathcal{D}_m f$ and $\mathcal{E}_r^n f$ are defined in (4.26) *and* (4.27), *respectively.*

Remark 4.9. At this stage, we could prove that

$$\big|\mathcal{E}_r^n f\big| \leq C\big(1 + \mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1} \Bigg[\|f\|_{r+3,\infty} \sup_{|\alpha|\leq r} |\Delta_\alpha| + \|f\|_{r+1,\infty} \frac{1}{n^{(r-[r/3]-2)/2}} \Bigg], \tag{4.28}$$

C denoting a suitable constant depending on r and N only. But since we aim to deal with the distance in total variation, we need a representation and an estimate of the reminder in terms of f and not of its derivatives. So, we skip this point and we postpone the problem to next section.

Proof of Theorem 4.8. Take $r \geq 2$. We use Proposition 4.5: for every $n \in \mathbb{N}$ and $f \in C_b^{r+2}(\mathbb{R}^N)$ we have

$$\begin{aligned} \mathbb{E}\big(f\big(S_n^n\big)\big) - \mathbb{E}\big(f\big(S_n^0\big)\big) &= \sum_{t=3}^{r} \frac{1}{n^{t/2}} \sum_{i=1}^{[t/3]} P_i(n) \mathbb{E}\big(\mathcal{A}_t^i f(G)\big) + \frac{1}{n^{(r+1)/2}} \mathcal{U}_r^n f \\ &= \sum_{t=3}^{r} \frac{1}{n^{t/2}} \sum_{i=1}^{[t/3]} \sum_{p=0}^{i} a_{i,p} n^p \mathbb{E}\big(\mathcal{A}_t^i f(G)\big) + \frac{1}{n^{(r+1)/2}} \mathcal{U}_r^n f \\ &= \sum_{t=3}^{r} \sum_{p=0}^{[t/3]} \frac{1}{n^{t/2-p}} \sum_{i=1\vee p}^{[t/3]} a_{i,p} \mathbb{E}\big(\mathcal{A}_t^i f(G)\big) + \frac{1}{n^{(r+1)/2}} \mathcal{U}_r^n f. \end{aligned}$$

So, by recalling that $S_n = S_n^n$ and $G \stackrel{\mathcal{L}}{=} S_n^0$ we obtain

$$\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f(G)\big) = \sum_{t=3}^{r} \sum_{p=0}^{[t/3]} \frac{1}{n^{t/2-p}} \sum_{i=1\vee p}^{[t/3]} a_{i,p}\mathbb{E}\big(\mathcal{A}_t^i f(G)\big) + \frac{1}{n^{(r+1)/2}}\mathcal{U}_r^n f.$$

We set now $t-2p=m$, so $t-m$ is an even number. Now, $p \geq 0$ gives that $t \geq m$ and since $t \geq 3$ then $t \geq 3 \vee m$ and $m \leq r$; $p \leq [t/3]$ gives that $(t-m)/2 \leq [t/3]$. Therefore, the sum over $t \leq r$ must be done on the set $\{t : 3 \vee m \leq t \leq r, t-m \text{ even}, t-2[t/3] \leq m\}$. It is easy to see that this set equals to $\{t : 3 \vee m \leq t \leq (3m) \wedge r, t-m \text{ even}\}$. So, we obtain

$$\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f(G)\big) = \sum_{m=1}^{r} \frac{1}{n^{m/2}} \sum_{\substack{t=3\vee m\\ t-m \text{ even}}}^{(3m)\wedge r} \sum_{i=1\vee(t-m)/2}^{[t/3]} a_{i,(t-m)/2}\mathbb{E}\big(\mathcal{A}_t^i f(G)\big) + \frac{1}{n^{(r+1)/2}}\mathcal{U}_r^n f.$$

The statement now follows by using (4.26) (notice that $3m \leq r$ if $m \leq [r/3]$) and (4.27). □

4.2. Regularized functions and estimate of the reminder

Our problem is now to prove an estimate for the reminder in the development for a function f in terms of $\|f\|_\infty$ instead of $\|f\|_{r+1,\infty}$. To this purpose, we need some preliminary results.

For $\delta > 0$, we denote by γ_δ the density of the centred Gaussian law in $\mathbb{R}^N$ of variance δI and for $f : \mathbb{R}^N \to \mathbb{R}$ we denote $f_\delta = f * \gamma_\delta$. Using standard integration by parts on $\mathbb{R}^N$, one may prove that for each $r \in \mathbb{N}$ there exists an universal constant C (depending on N and r only) such that for every multiindex α with $|\alpha| = r$ one has

$$\|\partial_\alpha f_\delta\|_\infty \leq \frac{C}{\delta^{r/2}}\|f\|_\infty. \tag{4.29}$$

We give now some estimates following from Lemma 3.5 with $\Theta = 1$, which is actually Lemma 2.5 in [3].

Lemma 4.10. *Suppose that $\mu_F \succeq \mathrm{Leb}_N$. There exist universal constants $C > 0$ and $b > 4$, depending on N only, such that for every $\delta > 0$, $n \in \mathbb{N}$ and for every bounded and measurable function $f : \mathbb{R}^N \to \mathbb{R}$ one has*

$$\big|\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f_\delta(S_n)\big)\big| \leq C\|f\|_\infty\big(1 + \mathbb{E}\big(|F|\big)\big)\big(e^{-n/C} + \delta^{1/b}n^{(b-2)/(2b)}\big). \tag{4.30}$$

Proof. Let $K \geq 1$ and $\Psi_K \in C^\infty(\mathbb{R}^N)$ be such that $1_{B_K(0)} \leq \Psi_K \leq 1_{B_{K+1}(0)}$ and such that, for some $L > 0$, $\|\partial_\alpha \Psi_K\|_\infty \leq L$ for every multiindex α. Then we have

$$\big|\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f\big(\Psi_K(S_n)S_n\big)\big)\big| \leq \|f\|_\infty \mathbb{P}\big(|S_n| \geq K\big) \leq \|f\|_\infty \frac{\mathbb{E}(|S_n|)}{K} \leq \|f\|_\infty \frac{\sqrt{n}}{K}\mathbb{E}\big(|F|\big)$$

and in a similar way $|\mathbb{E}(f_\delta(S_n)) - \mathbb{E}(f_\delta(\Psi_K(S_n)S_n))| \le \|f\|\mathbb{E}(|F|)\sqrt{n}/K$. So we can write

$$\begin{aligned}
&\bigl|\mathbb{E}\bigl(f(S_n)\bigr) - \mathbb{E}\bigl(f_\delta(S_n)\bigr)\bigr| \\
&\quad \le \bigl|\mathbb{E}\bigl(f(S_n)\bigr) - \mathbb{E}\bigl(f\bigl(\Psi_K(S_n)S_n\bigr)\bigr)\bigr| + \bigl|\mathbb{E}\bigl(f_\delta(S_n)\bigr) - \mathbb{E}\bigl(f_\delta\bigl(\Psi_K(S_n)S_n\bigr)\bigr)\bigr| \\
&\qquad + \bigl|\mathbb{E}\bigl(f\bigl(\Psi_K(S_n)S_n\bigr)\bigr) - \mathbb{E}\bigl(f_\delta\bigl(\Psi_K(S_n)S_n\bigr)\bigr)\bigr| \\
&\quad \le 2\mathbb{E}\bigl(|F|\bigr)\|f\|_\infty \frac{\sqrt{n}}{K} + \bigl|\mathbb{E}\bigl(f\bigl(\Psi_K(S_n)S_n\bigr)\bigr) - \mathbb{E}\bigl(f_\delta\bigl(\Psi_K(S_n)S_n\bigr)\bigr)\bigr|.
\end{aligned}$$

As for the last term in the above right-hand side, we apply Lemma 3.5 with $\Theta = 1$ and $X = \Psi_K(S_n)S_n$: there exist some universal constants C, p, a depending only on N such that for every $\varepsilon > 0, \delta > 0$ and every $f \in L^\infty(\mathbb{R}^N)$ then

$$\begin{aligned}
&\bigl|\mathbb{E}\bigl(f\bigl(\Psi_K(S_n)S_n\bigr)\bigr) - \mathbb{E}\bigl(f_\delta\bigl(\Psi_K(S_n)S_n\bigr)\bigr)\bigr| \\
&\quad \le C\|f\|_\infty \times \biggl(\mathbb{P}(\det \sigma_{\Psi_K(S_n)S_n} < \varepsilon) + \frac{\sqrt{\delta}}{\varepsilon^p}\bigl(1 + \bigl\|\Psi_K(S_n)S_n\bigr\|_{3,p} + \bigl\|L\bigl(\Psi_K(S_n)S_n\bigr)\bigr\|_{1,p}\bigr)^a\biggr).
\end{aligned}$$

We note that we are forced to introduce the localization $\Psi_K(S_n)$ because in the above estimate it appears $\|\Psi_K(S_n)S_n\|_p$ with $p > 1$: since the r.v.'s are only square integrable, if we take $\Psi_K \equiv 1$ then in principle we do not know if such norm is finite.

Now, on the set $\{|S_n| \le K\}$ we have $\det \sigma_{\Psi_K(S_n)S_n} = \det \sigma_{S_n}$, so that

$$\begin{aligned}
\mathbb{P}(\det \sigma_{\Psi_K(S_n)S_n} < \varepsilon) &\le \mathbb{P}(\det \sigma_{S_n} < \varepsilon) + \mathbb{P}\bigl(|S_n| > K\bigr) \le \mathbb{P}(\det \sigma_{S_n} < \varepsilon) + \frac{\mathbb{E}(|S_n|)}{K} \\
&\le \mathbb{P}(\det \sigma_{S_n} < \varepsilon) + \mathbb{E}\bigl(|F|\bigr)\frac{\sqrt{n}}{K}.
\end{aligned}$$

By taking $\varepsilon = \varepsilon_*/2$ as in Lemma 3.4, (3.22) gives

$$\mathbb{P}(\det \sigma_{\Psi_K(S_n)S_n} < \varepsilon) \le Ce^{-n/C} + \mathbb{E}\bigl(|F|\bigr)\frac{\sqrt{n}}{K}.$$

Therefore, we can write

$$\begin{aligned}
&\bigl|\mathbb{E}\bigl(f(S_n)\bigr) - \mathbb{E}\bigl(f_\delta(S_n)\bigr)\bigr| \\
&\quad \le C\|f\|_\infty\biggl(e^{-n/C} + \mathbb{E}\bigl(|F|\bigr)\frac{\sqrt{n}}{K} + \sqrt{\delta}\bigl(1 + \bigl\|\Psi_K(S_n)S_n\bigr\|_{3,p} + \bigl\|L\bigl(\Psi_K(S_n)S_n\bigr)\bigr\|_{1,p}\bigr)^a\biggr).
\end{aligned}$$

We use now Lemma B.1 in Appendix B: inequalities (B.1) and (B.2) give

$$\begin{aligned}
&\bigl\|\Psi_K(S_n)S_n\bigr\|_{3,p} + \bigl\|L\bigl(\Psi_K(S_n)S_n\bigr)\bigr\|_{1,p} \\
&\quad \le CK\bigl(1 + \|S_n\|_{1,3,4p}\bigr)^6 + CK\bigl(1 + \|S_n\|_{1,2,8p}\bigr)^5\bigl(1 + \|LS_n\|_{1,4p}\bigr) \\
&\quad \le CK\bigl(1 + \|S_n\|_{1,3,8p} + \|LS_n\|_{1,4p}\bigr)^6.
\end{aligned}$$

By using (3.19) and (3.20), we have

$$\left\|\Psi_K(S_n)S_n\right\|_{3,p} + \left\|L\big(\Psi_K(S_n)S_n\big)\right\|_{1,p} \le CK,$$

so that

$$\begin{aligned}\left|\mathbb{E}\big(f(S_n)\big) - \mathbb{E}\big(f_\delta(S_n)\big)\right| &\le C\|f\|_\infty\left(e^{-n/C} + \mathbb{E}\big(|F|\big)\frac{\sqrt{n}}{K} + \sqrt{\delta}K^a\right) \\ &\le C\|f\|_\infty\big(1+\mathbb{E}\big(|F|\big)\big)\left(e^{-n/C} + \frac{\sqrt{n}}{K} + \sqrt{\delta}K^a\right).\end{aligned}$$

We now optimize on K by taking it in order that $\sqrt{n}/K = \sqrt{\delta}K^a$. Straightforward computations give now (4.30), with $\frac{1}{b} = \frac{1}{2}(1 - \frac{a}{a+1}) < \frac{1}{4}$. □

Remark 4.11. We stress that when $C(F) \neq Id$ then the constant in (3.22) depends on $\underline{\lambda}(F)$. As a consequence, this dependence holds for the constant C appearing in (4.30) as well.

We now propose the following key result, allowing us to deal with the remaining terms.

Lemma 4.12. *Suppose that $\mu_F \succeq \mathrm{Leb}_N$. Let α and β denote multiindexes, with $|\alpha| = r$ and $|\beta| = m$. If $F \in L^m(\Omega)$, then there exists a constant C (which depends on N, r and m) such that for every $f \in L^\infty(\mathbb{R}^N)$, $\delta > 0$, $n \ge 1$ and $\lambda \in \mathbb{R}$ then*

$$\begin{aligned}\left|\mathbb{E}\left(\partial_\alpha f_\delta\left(\widehat{S}_n^k + \lambda\frac{F_k}{n^{1/2}}\right)F_k^\beta\right)\right| &\le C\|f\|_\infty\mathbb{E}\big(|F|^m\big)\big(1+\delta^{-r/2}e^{-n/C}\big), \\ \left|\mathbb{E}\left(\partial_\alpha f_\delta\left(\widehat{S}_n^k + \lambda\frac{G_k}{n^{1/2}}\right)G_k^\beta\right)\right| &\le C\|f\|_\infty\mathbb{E}\big(|G|^m\big)\big(1+\delta^{-r/2}e^{-n/C}\big),\end{aligned}$$

in which $f_\delta = f * \gamma_\delta$, γ_δ being the centred normal density in $\mathbb{R}^N$ with covariance matrix δI.

Proof. Without loss of generality, we suppose that n is even and we study separately the cases $k \le n/2$ and $k \ge n/2 + 1$ – if n was odd, it would be sufficient to study $k \le (n-1)/2$ and $k \ge (n-1)/2 + 1$.

Case 1: $k \le n/2$. We denote

$$A_k = \frac{1}{n^{1/2}}\left(\sum_{i=1}^{k-1} F_i + \sum_{i=k+1}^{n/2} G_i\right) + \lambda\frac{F_k}{n^{1/2}}, \qquad B = \frac{1}{n^{1/2}}\sum_{i=n/2+1}^{n} G_i$$

so that

$$\widehat{S}_n^k + \lambda\frac{F_k}{n^{1/2}} = A_k + B.$$

Notice that B is a Gaussian random variable with covariance $\frac{1}{2}I$ which is independent of A_k and of F_k. Using integration by parts with respect to B we may find a random variable H_α having all moments and

$$\mathbb{E}\left(\partial_\alpha f_\delta\left(\widehat{S}_n^k+\lambda\frac{F_k}{n^{1/2}}\right)F_k^\beta\right)=\mathbb{E}\big(\partial_\alpha f_\delta(A_k+B)F_k^\beta\big)=\mathbb{E}\big(f_\delta(A_k+B)F_k^\beta H_\alpha\big).$$

Since F_k and H_α are independent, H_α being a suitable function of $G_{n/2},\dots,G_n$, it follows that

$$\left|\mathbb{E}\left(\partial_\alpha f_\delta\left(\widehat{S}_n^k+\lambda\frac{F_k}{n^{1/2}}\right)F_k^\beta\right)\right|\le C\|f_\delta\|_\infty\mathbb{E}\big(|F_k|^m\big)\mathbb{E}\big(|H_\alpha|\big)\le C\|f\|_\infty\mathbb{E}\big(|F|^m\big).$$

Similarly, we obtain

$$\left|\mathbb{E}\left(\partial_\alpha f_\delta\left(\widehat{S}_n^k+\lambda\frac{G_k}{n^{1/2}}\right)G_k^\beta\right)\right|\le C\|f\|_\infty\mathbb{E}\big(|G|^m\big).$$

Case 2: $k>n/2$. We denote

$$A=\frac{1}{n^{1/2}}\sum_{i=1}^{n/2}F_i,\qquad B_k=\frac{1}{n^{1/2}}\left(\sum_{i=n/2+1}^{k-1}F_i+\sum_{i=k+1}^{n}G_i\right)+\lambda\frac{F_k}{n^{1/2}}$$

so that

$$\widehat{S}_n^k+\lambda\frac{F_k}{n^{1/2}}=A+B_k.$$

We notice that

$$A=\frac{1}{\sqrt{2}}S_{n/2},$$

so we can use the noise from the absolutely continuous r.v.'s $V_1,\dots,V_{n/2}$ "inside" $S_{n/2}$, as already seen in Section 3.1. We then proceed to use integration by parts w.r.t. the noise from A.

We notice that $\sigma_A=\frac{1}{2}\sigma_{S_{n/2}}$ and that the covariance matrix $\sigma_{S_{n/2}}$ of $S_{n/2}$ may degenerate. So, we use a localization: we consider a function $\phi\in C^1(\mathbb{R}_+)$ such that $\mathbf{1}_{(\varepsilon_*/2,\infty)}\le\phi\le\mathbf{1}_{(\varepsilon_*,\infty)}$ and $\|\nabla\phi\|_\infty\le 2/\varepsilon_*$ with ε_* given in (3.21). Then we write

$$\mathbb{E}\left(\partial_\alpha f_\delta\left(\widehat{S}_n^k+\lambda\frac{F_k}{n^{1/2}}\right)F_k^\beta\right)=\mathbb{E}\big(\partial_\alpha f_\delta(A+B_k)F_k^\beta\big)=I+J$$

with

$$I=\mathbb{E}\big(\partial_\alpha f_\delta(A+B_k)F_k^\beta\phi(\det\sigma_A)\big),$$
$$J=\mathbb{E}\big(\partial_\alpha f_\delta(A+B_k)F_k^\beta\big(1-\phi(\det\sigma_A)\big)\big).$$

We estimate I. Notice that $\phi(\det\sigma_A)\neq 0$ implies that $\det\sigma_A\geq\varepsilon_*/2$. We use the integration by parts with respect to A in Proposition 3.2, and we obtain

$$I=\mathbb{E}\big(f_\delta(A+B_k)F_k^\alpha H_\alpha^r\big(A,\phi(\det\sigma_A)\big)\big).$$

The estimate (3.15) for the weight gives that

$$\big|H_\alpha^r(A,\det\sigma_A)\big|\leq C\big(1\vee(\det\sigma_A)^{-1}\big)^{r(r+1)}\big(1+|A|_{1,r+1}^{2N(r+2)}+|LA|_{r-1}^2\big)^r\times\big|\phi(\det\sigma_A)\big|_r,$$

C denoting a universal constant. Since $\sigma_A=\frac{1}{2}\sigma_{S_n/2}=\frac{1}{2n}\sum_{k=1}^{n/2}\chi_k I$, all the Malliavin derivatives are null, so $|\phi(\det\sigma_A)|_r=|\phi(\det\sigma_A)|\leq 1$, so that

$$\big|H_\alpha^r(A,\det\sigma_A)\big|\leq C\big(1\vee(\det\sigma_A)^{-1}\big)^{r(r+1)}\big(1+|A|_{1,r+1}^{2N(r+2)}+|LA|_{r-1}^2\big)^r.$$

We pass now to expectation: by using the Hölder inequality, we may find some universal constants C,q,p such that

$$\mathbb{E}\big(\big|H_\alpha^r\big(A,\phi(\det\sigma_A)\big)\big|^2\big)\leq\frac{C}{\varepsilon_*^q}\big(1+\|A\|_{1,r+1,p}+\|LA\|_{r-1,p}\big)^q\leq C',$$

the latter inequality following from (3.19) and (3.20). Now, F_k and $H_\alpha^r(A,\phi(\det\sigma_A))$ are independent, so that

$$|I|\leq C\|f\|_\infty\mathbb{E}\big(|F|^m\big).$$

We estimate now J. By recalling again that F_k and σ_A are independent and by using (4.29) and (3.22), we obtain

$$\begin{aligned}|J|&\leq\|\partial_\alpha f_\delta\|_\infty\mathbb{E}\big(\big|F_k^\beta\big(1-\phi(\det\sigma_A)\big|\big)\leq\delta^{-r/2}\|f\|_\infty\mathbb{E}\big(|F_k|^m\big)\mathbb{P}(\sigma_{S_{n/2}}\leq\varepsilon_*)\big)\\&\leq C\delta^{-r/2}\|f\|_\infty\mathbb{E}\big(|F|^m\big)\times e^{-n/C}.\end{aligned}$$

By resuming, we get

$$\left|\mathbb{E}\left(\partial_\alpha f_\delta\left(\widehat{S}_n^k+\lambda\frac{F_k}{n^{1/2}}\right)F_k^\beta\right)\right|\leq C\|f\|_\infty\mathbb{E}\big(|F|^m\big)\big(1+\delta^{-r/2}e^{-n/C}\big).$$

And similarly, we prove that

$$\left|\mathbb{E}\left(\partial_\alpha f_\delta\left(\widehat{S}_n^k+\lambda\frac{G_k}{n^{1/2}}\right)G_k^\beta\right)\right|\leq C\|f\|_\infty\mathbb{E}\big(|G|^m\big)\big(1+\delta^{-r/2}e^{-n/C}\big).$$ □

We can now give a nice estimate for $\mathcal{U}_r^n f_\delta$ in terms of $\|f\|_\infty$. And this is enough for the moment.

Lemma 4.13. *Suppose that $\mu_F\succeq\mathrm{Leb}_N$. Let $r\geq 2$ and $F\in L^{r+1}(\Omega)$. For $f\in L^\infty(\mathbb{R}^N)$ and $\delta>0$, set $f_\delta=f*\gamma_\delta$, γ_δ being the centred normal density in $\mathbb{R}^N$ with covariance matrix δI.*

Then there exists $C>0$ depending on r and N only such that for every $f\in L^{\infty}(\mathbb{R}^N)$ one has

$$\begin{aligned}\left|\mathcal{U}_r^n f_\delta\right| \le\, & C\left(1+\mathbb{E}\left(|F|^{r+1}\right)\right)^{[r/3]\vee 1}\\ & \times \|f\|_\infty\left(1+\delta^{-(r+4)/2}e^{-n/C}\right)\Big(\sup_{|\alpha|\le r}|\Delta_\alpha|\times n^{(r-[r/3])/2}+n\Big).\end{aligned}\tag{4.31}$$

Proof. By using (4.15) and (4.14), we can write

$$\mathcal{U}_r^n f=\sum_{k=1}^n\left[\sum_{j=1}^{k-1}\sum_{t=3}^r \mathcal{R}_{r-t,n}^{k-j}\Psi_t^{(j)}f+\mathcal{R}_{r,n}^k f\right].$$

Since $g\mapsto \mathcal{R}_{t,n}^l g$ is linear, by using the expansion of $\Psi^{(k)}$ in Lemma 4.6 and by recalling that $Q_{i-1}(k)\ge 0$, we get

$$\left|\mathcal{U}_r^n f\right|\le \sum_{k=2}^n\sum_{j=1}^{k-1}\sum_{t=3}^r\sum_{i=1}^{[t/3]}Q_{i-1}(j)\left|\mathcal{R}_{r-t,n}^{k-j}\mathcal{A}_t^i f\right|+\sum_{k=1}^n\left|\mathcal{R}_{r,n}^k f\right|.$$

Since $r\ge 2$, (4.10) gives $\mathcal{R}_{0,n}^{\ell}=\frac{1}{n}\mathcal{R}_{2,n}^{\ell}$ and $\mathcal{R}_{1,n}^{\ell}=\frac{1}{\sqrt{n}}\mathcal{R}_{2,n}^{\ell}$. So, we isolate in the sum the terms with $t=r-1,r$ and we obtain

$$\begin{aligned}\left|\mathcal{U}_r^n f_\delta\right|\le \sum_{k=2}^n\sum_{j=1}^{k-1}\Bigg[&\mathbf{1}_{r\ge 5}\sum_{t=3}^{r-2}\sum_{i=1}^{[t/3]}Q_{i-1}(j)\left|\mathcal{R}_{r-t,n}^{k-j}\mathcal{A}_t^i f_\delta\right|\\ &+\mathbf{1}_{r\ge 4}\frac{1}{\sqrt{n}}\sum_{i=1}^{[(r-1)/3]}Q_{i-1}(j)\left|\mathcal{R}_{2,n}^{k-j}\mathcal{A}_{r-1}^i f_\delta\right|\\ &+\mathbf{1}_{r\ge 3}\frac{1}{n}\sum_{i=1}^{[r/3]}Q_{i-1}(j)\left|\mathcal{R}_{2,n}^{k-j}\mathcal{A}_r^i f_\delta\right|\Bigg]+\sum_{k=1}^n\left|\mathcal{R}_{r,n}^k f_\delta\right|.\end{aligned}\tag{4.32}$$

We have (recall formula (4.9))

$$\begin{aligned}\left|\mathcal{R}_{r,n}^k f_\delta\right|\le &\sum_{p=3}^r\sum_{|\alpha|=p}\sum_{|\beta|=2[(r-p)/2]+2}|\Delta_\alpha|\int_0^1\left|\mathbb{E}\left(\partial_\beta\,\partial_\alpha f_\delta\left(\widehat{S}_n^k+\sqrt{s}\frac{G_k}{\sqrt{n}}\right)\right)\right|ds\\ &+n^{(r+1)/2}\left[\left|\mathbb{E}\left(U_r f_\delta\left(\widehat{S}_n^k,\frac{F_k}{\sqrt{n}}\right)\right)\right|+\left|\mathbb{E}\left(U_r f_\delta\left(\widehat{S}_n^k,\frac{G_k}{\sqrt{n}}\right)\right)\right|\right]\end{aligned}$$

and by using Lemma 4.12 we get

$$\left|\mathcal{R}_{r,n}^k f_\delta\right|\le C\left(1+\mathbb{E}\left(|F|^{r+1}\right)\right)\|f\|_\infty\left(1+\delta^{-(r+1)/2}e^{-n/C}\right).\tag{4.33}$$

As for the other sums in the right-hand side of (4.32), for $s \geq 2$ we have

$$\big|\mathcal{R}_{s,n}^{k-j}\mathcal{A}_t^i f_\delta\big| \leq \sum_{|\gamma|=t}\big|c_\gamma^i\big| \times \big|\mathcal{R}_{s,n}^{k-j}\partial_\gamma f_\delta\big| \leq C\sup_{|\alpha|\leq t}|\Delta_\alpha|\big(1+\mathbb{E}\big(|F|^t\big)\big)^{i-1}\sum_{|\gamma|=t}\big|\mathcal{R}_{s,n}^{k-j}\partial_\gamma f_\delta\big|,$$

last inequality following from (4.19). We use again Lemma 4.12: for $|\gamma| = t$,

$$\big|\mathcal{R}_{s,n}^{k-j}\partial_\gamma f_\delta\big| \leq C\big(1+\mathbb{E}\big(|F|^{s+1}\big)\big)\|f\|_\infty\big(1+\delta^{-(s+t+1)/2}e^{-n/C}\big).$$

We apply such inequality with: $t \leq r-2$ and $s = r-t$, $t = r-1$ and $s = 2$, $t = r$ and $s = 2$. Then

$$\big|\mathcal{R}_{r-t,n}^{k-j}\mathcal{A}_t^i f_\delta\big| \leq C\sup_{|\alpha|\leq r}|\Delta_\alpha|\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\|f\|_\infty\big(1+\delta^{-(r+2)/2}e^{-n/C}\big),$$

$$\big|\mathcal{R}_{2,n}^{k-j}\mathcal{A}_{r-1}^i f_\delta\big| \leq C\sup_{|\alpha|\leq r}|\Delta_\alpha|\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\|f\|_\infty\big(1+\delta^{-(r+3)/2}e^{-n/C}\big),$$

$$\big|\mathcal{R}_{2,n}^{k-j}\mathcal{A}_r^i f_\delta\big| \leq C\sup_{|\alpha|\leq r}|\Delta_\alpha|\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\|f\|_\infty\big(1+\delta^{-(r+4)/2}e^{-n/C}\big).$$

By inserting such estimates and (4.33) in (4.32), we get

$$\begin{aligned}
\big|\mathcal{U}_r^n f_\delta\big| &\leq \sum_{k=2}^{n}\sum_{j=1}^{k-1}\Bigg[\mathbf{1}_{r\geq 5}\sum_{t=3}^{r-2}\sum_{i=1}^{[t/3]}Q_{i-1}(j)\big|\mathcal{R}_{r-t,n}^{k-j}\mathcal{A}_t^i f_\delta\big| \\
&\quad + \mathbf{1}_{r\geq 4}\frac{1}{\sqrt{n}}\sum_{i=1}^{[(r-1)/3]}Q_{i-1}(j)\big|\mathcal{R}_{2,n}^{k-j}\mathcal{A}_{r-1}^i f_\delta\big| \\
&\quad + \mathbf{1}_{r\geq 3}\frac{1}{n}\sum_{i=1}^{[r/3]}Q_{i-1}(j)\big|\mathcal{R}_{2,n}^{k-j}\mathcal{A}_r^i f_\delta\big|\Bigg] + \sum_{k=1}^{n}\big|\mathcal{R}_{r,n}^k f_\delta\big| \\
&\leq C\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\|f\|_\infty\big(1+\delta^{-(r+4)/2}e^{-n/C}\big) \\
&\quad \times \Bigg(\sup_{|\alpha|\leq r}|\Delta_\alpha|\sum_{k=2}^{n}\sum_{j=1}^{k-1}\Bigg[\mathbf{1}_{r\geq 5}\sum_{i=1}^{[(r-2)/3]}Q_{i-1}(j) + \mathbf{1}_{r\geq 4}\frac{1}{\sqrt{n}}\sum_{i=1}^{[(r-1)/3]}Q_{i-1}(j) \\
&\quad + \mathbf{1}_{r\geq 3}\frac{1}{n}\sum_{i=1}^{[r/3]}Q_{i-1}(j)\Bigg] + n\Bigg).
\end{aligned}$$

Since $\sum_{k=2}^{n}\sum_{j=1}^{k-1}\sum_{i=1}^{L}Q_{i-1}(j) = \sum_{i=1}^{L}P_{i+1}(n-1)$ is a polynomial of order $L+1$ we obtain

$$\begin{aligned}
\big|\mathcal{U}_r^n f_\delta\big| &\leq C\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\|f\|_\infty\big(1+\delta^{-(r+4)/2}e^{-n/C}\big) \\
&\quad \times \Big(\sup_{|\alpha|\leq r}|\Delta_\alpha|\big[\mathbf{1}_{r\geq 5}n^{[(r-2)/3]+1} + \mathbf{1}_{r\geq 4}n^{[(r-1)/3]+1/2} + \mathbf{1}_{r\geq 3}n^{[r/3]}\big] + n\Big)
\end{aligned}$$

and the statement follows by noticing that

$$n^{[(r-2)/3]+1}\mathbf{1}_{r\geq 5}+n^{[(r-1)/3]+1/2}\mathbf{1}_{r\geq 4}+n^{[r/3]}\mathbf{1}_{r\geq 3}\leq Cn^{[r/3]+(r-3[r/3])/2}. \qquad \square$$

4.3. Estimate of the error in total variation distance

We want to get rid of the derivatives of f which appear in the coefficients $\mathcal{D}_m f$. In order to do it, we will use integration by parts w.r.t. the Gaussian law and then the Hermite polynomials come on. Again, we assume $\mu_F \succeq \mathrm{Leb}_N$ and F has null mean and identical covariance matrix.

We denote by H_m the Hermite polynomial of order m on $\mathbb{R}$, that is,

$$H_m(x)=(-1)^m e^{(1/2)x^2}\frac{d^m}{x^m}e^{-(1/2)x^2}. \tag{4.34}$$

For a multiindex $\alpha=(\alpha_1,\ldots,\alpha_r)\in\{1,\ldots,N\}^r$ we denote $\beta_i(\alpha)=\mathrm{card}\{j:\alpha_j=i\}$ so that $\partial_\alpha=\partial_{x_1}^{\beta_1(\alpha)}\ldots\partial_{x^N}^{\beta_d(\alpha)}$. And we define the Hermite polynomial on $\mathbb{R}^N$ corresponding to the multiindex α by

$$H_\alpha(x)=\prod_{i=1}^N H_{\beta_i(\alpha)}(x_i) \qquad \text{for } x=(x_1,\ldots,x_N). \tag{4.35}$$

With this definition, we have

$$\partial_\alpha e^{-(1/2)|x|^2}=(-1)^{|\alpha|}H_\alpha(x)e^{-(1/2)|x|^2}$$

and using integration by parts, for a centred Gaussian random variable $G\in\mathbb{R}^N$

$$\mathbb{E}\big(\partial_\alpha f(G)\big)=\mathbb{E}\big(f(G)H_\alpha(G)\big). \tag{4.36}$$

This means that we can compute $\mathbb{E}(\mathcal{A}_t^i f(G))$ by means of f and not of its derivatives. In fact, for $i\geq 1$ and $t\geq 0$, we define the polynomials $\mathcal{H}_t^i(x)$ as follows:

$$\mathcal{H}_t^i(x)=\sum_{\alpha:|\alpha|=t}c_\beta^i H_\alpha(x), \qquad c_\alpha^i \text{ defined in (4.17) and } H_\alpha \text{ given in (4.35)}. \tag{4.37}$$

Since $\mathcal{A}_t^i=\sum_{\alpha:|\alpha|=t}c_\alpha^i\partial_\alpha$, (4.36) gives

$$\mathbb{E}\big(\mathcal{A}_t^i f(G)\big)=\sum_{\alpha:|\alpha|=t}c_\gamma^i\mathbb{E}\big(\partial_\alpha f(G)\big)=\sum_{\alpha:|\alpha|=t}c_\alpha^i\mathbb{E}\big(f(G)H_\alpha(G)\big)=\mathbb{E}\big(f(G)\mathcal{H}_t^i(G)\big).$$

Therefore, for every $f\in C_b^m(\mathbb{R}^N)$ the coefficients $\mathcal{D}_m f$, $m\geq 1$, in (4.26) can be written as

$$\mathcal{D}_m f=\mathbb{E}\big(f(G)\mathcal{K}_m(G)\big), \qquad m\geq 1, \tag{4.38}$$

where $\mathcal{K}_m(x)=\sum_{\substack{t=3\vee m\\ t-m \text{ even}}}^{3m}\sum_{i=1\vee(t-m)/2}^{[t/3]}a_{i,(t-m)2}\mathcal{H}_t^i(x)$, $a_{i,l}$ given in (4.24) and $\mathcal{H}_t^i$ given in (4.37).

We are now ready to tackle our original problem: the exact expansion in total variation distance of the law μ_n of S_n. To this purpose, for $r \geq 2$ and $n \geq 1$ we define the following measure in $\mathbb{R}^N$:

$$\Gamma_{n,r}(dx) = \gamma(x)\left(1 + \sum_{m=1}^{[r/3]} \frac{1}{n^{m/2}}\mathcal{K}_m(x)\right)dx, \qquad \mathcal{K}_m(x) \text{ given in (4.38)}, \tag{4.39}$$

where $\gamma(x)$ denotes the probability density function of a standard normal random variable in $\mathbb{R}^N$. We stress that $\Gamma_{n,r}(dx) = \gamma(x)\,dx =: \Gamma(dx)$ not only for $r = 2$ but also when $\Delta_\alpha = 0$ for every $|\alpha| \leq r$. In fact, in the latter case, (4.19) gives $c_\alpha^i = 0$ for every $i \geq 1$ and $|\alpha| \leq r$, then from (4.37) we have $\mathcal{H}_t^i \equiv 0$ for every $i \geq 1$ and $t \leq r$ and from (4.38) we obtain $\mathcal{K}_m \equiv 0$ for every $m \leq r$.

Theorem 4.14. *Suppose $\mu_F \succeq \mathrm{Leb}_N$. Let $r \geq 2$ and $F \in L^{r+1}(\Omega)$. For $n \geq 1$, let μ_n denote the law of S_n and $\Gamma_{n,r}$ stand for the measure in (4.39). Then there exists a constant $C > 0$ depending on r and N only such that for every $n \in \mathbb{N}$,*

$$d_{\mathrm{TV}}(\mu_n, \Gamma_{n,r}) \leq C\big(1 + \mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\left[\sup_{|\alpha|\leq r} |\Delta_\alpha| \times \frac{1}{n^{([r/3]+1)/2}} + \frac{1}{n^{(r-1)/2}}\right].$$

Proof. We study $|\int f\,d\mu_n - \int f\,d\Gamma_{n,r}|$ for $f \in L^\infty(\mathbb{R}^N)$. From now on, C will denote a constant, possibly varying from line to line, that may depend only on N and r.

We take $\delta > 0$ and we consider the regularized function $f_\delta = f * \gamma_\delta$ where γ_δ is the centred Gaussian density of covariance matrix δI. We have

$$\left|\int f\,d\mu_n - \int f\,d\Gamma_{n,r}\right| \leq I_{n,\delta} + I'_{n,\delta} + J_{n,\delta}$$

with

$$I_{n,\delta} = \left|\int (f - f_\delta)\,d\mu_n\right|, \qquad I'_{n,\delta} = \left|\int (f - f_\delta)\,d\Gamma_{n,r}\right|, \qquad J_{n,\delta} = \left|\int f_\delta\,d\mu_n - \int f_\delta\,d\Gamma_{n,r}\right|.$$

By (4.30),

$$I_{n,\delta} \leq C\|f\|_\infty\big(1 + \mathbb{E}\big(|F|\big)\big)\big(e^{-n/C} + \delta^{1/b}n^{(b-2)/(2b)}\big),$$

where $b > 4$ is a suitable constant, independent of F and f. And using standard integration by parts on $\mathbb{R}^N$,

$$I'_{n,\delta} \leq C\|f\|_\infty\delta^{1/2}.$$

Moreover, since

$$\int f_\delta\,d\Gamma_{n,r} = \mathbb{E}\big(f_\delta(G)\big) + \sum_{m=1}^{[r/3]} \frac{1}{n^{m/2}}\mathcal{D}_m f_\delta,$$

Theorem 4.8 gives

$$J_{n,\delta}=\frac{1}{n^{([r/3]+1)/2}}\left|\mathcal{E}_r^n f_\delta\right|$$

with

$$\left|\mathcal{E}_r^n f_\delta\right|\leq n^{([r/3]+1)/2}\Bigg[\sum_{m=[r/3]+1}^{r}\frac{1}{n^{m/2}}\sum_{t=3\vee m}^{(3m)\wedge r}\sum_{i=1\vee(t-m)/2}^{[t/3]}|a_{i,(t-m)/2}| \\ \times\left|\mathbb{E}\big(\mathcal{A}_t^i f_\delta(G)\big)\right|+\frac{1}{n^{(r+1)/2}}\left|\mathcal{U}_r^n f_\delta\right|\Bigg].$$

But since $\mathbb{E}(\mathcal{A}_t^i f_\delta(G))=\mathbb{E}(f_\delta(G)\mathcal{H}_t^i(G))$, then

$$\left|\mathbb{E}\big(\mathcal{A}_t^i f_\delta(G)\big)\right|\leq\|f_\delta\|_\infty\mathbb{E}\big(\left|\mathcal{H}_t^i(G)\right|\big)\leq C\|f\|_\infty\big(1+\mathbb{E}\big(|F|^{t-1}\big)\big)\sup_{|\alpha|\leq t}|\Delta_\alpha|.$$

We use now Lemma 4.13: for $r\geq 2$, we apply (4.31) and we get

$$\left|\mathcal{E}_r^n f_\delta\right|\leq C\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\|f\|_\infty\big(1+\delta^{-(r+4)/2}e^{-n/C}\big)\Bigg[\sup_{|\alpha|\leq r}|\Delta_\alpha|+\frac{1}{n^{(r-[r/3]-2)/2}}\Bigg].$$

By replacing, we get

$$J_{n,\delta}\leq C\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\|f\|_\infty \\ \times\big(1+\delta^{-(r+4)/2}e^{-n/C}\big)\Bigg[\sup_{|\alpha|\leq r}|\Delta_\alpha|\times\frac{1}{n^{([r/3]+1)/2}}+\frac{1}{n^{(r-1)/2}}\Bigg].$$

By resuming, we can write

$$\left|\int f\,d\mu_n-\int f\,d\Gamma_{n,r}\right|\leq C\|f\|_\infty\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\Bigg[e^{-n/C}+\delta^{1/2}+\delta^{1/b}n^{(b-2)/(2b)} \\ +\big(1+\delta^{-(r+4)/2}e^{-n/C}\big)\Bigg(\sup_{|\alpha|\leq r}|\Delta_\alpha|\times\frac{1}{n^{([r/3]+1)/2}}+\frac{1}{n^{(r-1)/2}}\Bigg)\Bigg].$$

Now, we choose $\delta=\delta_n$ such that $\delta_n^{1/b}n^{(b-2)/(2b)}=\frac{1}{n^{(r-1)/2}}$. By observing that $n\mapsto\delta_n^{-(r+4)/2}\times e^{-n/C}$ is bounded and $\delta_n^{1/2}\leq\frac{1}{n^{(r-1)/2}}$, we get

$$\left|\int f\,d\mu_n-\int f\,d\Gamma_{n,r}\right|\leq C\|f\|_\infty\big(1+\mathbb{E}\big(|F|^{r+1}\big)\big)^{[r/3]\vee 1}\Bigg[\sup_{|\alpha|\leq r}|\Delta_\alpha|\times\frac{1}{n^{([r/3]+1)/2}}+\frac{1}{n^{(r-1)/2}}\Bigg]$$

and the result follows. □

We can now pass to the following.

Proof of Theorems 2.6 and 2.7. We apply Theorem 4.14 with F replaced by $A(F)F$, where $A(F)$ is the inverse of $C(F)^{1/2}$, $C(F)$ denoting the covariance matrix. And it is clear that now the constants appearing in the estimates will depend on $C(F)$ as well, through its most significant eigenvalues (the smallest and the largest one; see, e.g., in Remark 3.8 and 4.11). □

We conclude by explicitly writing $\mathcal{K}_m(x)$ for $m = 1, 2, 3$. From (4.38), we have:

$$\mathcal{K}_1(x) = a_{1,1}\mathcal{H}_3^1(x),$$
$$\mathcal{K}_2(x) = a_{1,1}\mathcal{H}_4^1(x) + a_{2,2}\mathcal{H}_6^2(x),$$
$$\mathcal{K}_3(x) = a_{1,0}\mathcal{H}_3^1(x) + a_{1,1}\mathcal{H}_5^1(x) + a_{2,2}\mathcal{H}_7^2(x) + a_{3,3}\mathcal{H}_9^3(x),$$

where $\mathcal{H}_t^i(x) = \sum_{|\gamma|=t} c_\gamma^i H_\gamma(x)$. Now, from (4.17) it is easy to see that

$$c_\gamma^1 = \begin{cases} \frac{1}{3!}\Delta_\gamma, & \text{if } |\gamma| = 3, \\ \frac{1}{4!}\Delta_\gamma, & \text{if } |\gamma| = 4, \\ -\frac{1}{3!2!}\Delta_{(\gamma_1,\gamma_2,\gamma_3)} 1_{\gamma_4=\gamma_5} + \frac{1}{5!}\Delta_\gamma, & \text{if } |\gamma| = 5, \end{cases}$$

$$c_\gamma^2 = \begin{cases} \frac{1}{(3!)^2}\Delta_{(\gamma_1,\gamma_2,\gamma_3)}\Delta_{(\gamma_4,\gamma_5,\gamma_6)}, & \text{if } |\gamma| = 6, \\ \frac{1}{3!4!}(\Delta_{(\gamma_1,\gamma_2,\gamma_3)}\Delta_{(\gamma_4,\gamma_5\gamma_6,\gamma_7)} + \Delta_{(\gamma_1,\gamma_2,\gamma_3,\gamma_4)}\Delta_{(\gamma_5,\gamma_6,\gamma_7)}), & \text{if } |\gamma| = 7, \end{cases}$$

$$c_\gamma^3 = \frac{1}{(3!)^3}\Delta_{(\gamma_1,\gamma_2,\gamma_3)}\Delta_{(\gamma_4,\gamma_5,\gamma_6)}\Delta_{(\gamma_7,\gamma_8,\gamma_9)} \qquad \text{if } |\gamma| = 9.$$

Moreover, $a_{1,0} = 0$, $a_{1,1} = 1$, $a_{2,2} = b_{1,2} = \frac{1}{2}B_0 = \frac{1}{2}$ and $a_{3,3} = a_{2,2}b_{2,3} = \frac{1}{2} \cdot \frac{1}{3}B_0 = \frac{1}{6}$. So, we can write

$$\mathcal{K}_1(x) = \frac{1}{3!}\sum_{|\gamma|=3}\Delta_\gamma H_\gamma(x),$$

$$\mathcal{K}_2(x) = \frac{1}{4!}\sum_{|\gamma|=4}\Delta_\gamma H_\gamma(x) + \frac{1}{2(3!)^2}\sum_{|\gamma|=6}\Delta_{(\gamma_1,\gamma_2,\gamma_3)}\Delta_{(\gamma_4,\gamma_5,\gamma_6)}H_\gamma(x),$$

$$\begin{aligned}\mathcal{K}_3(x) = &\sum_{|\gamma|=5}\left(-\frac{1}{3!2!}\Delta_{(\gamma_1,\gamma_2,\gamma_3)}1_{\gamma_4=\gamma_5} + \frac{1}{5!}\Delta_\gamma\right)H_\gamma(x) \\ &+ \frac{1}{2\times 3!4!}\sum_{|\gamma|=7}(\Delta_{(\gamma_1,\gamma_2,\gamma_3)}\Delta_{(\gamma_4,\gamma_5,\gamma_6,\gamma_7)} + \Delta_{(\gamma_1,\gamma_2,\gamma_3,\gamma_4)}\Delta_{(\gamma_5,\gamma_6,\gamma_7)})H_\gamma(x) \\ &+ \frac{1}{6\times(3!)^3}\sum_{|\gamma|=9}\Delta_{(\gamma_1,\gamma_2,\gamma_3)}\Delta_{(\gamma_4,\gamma_5,\gamma_6)}\Delta_{(\gamma_7,\gamma_8,\gamma_9)}H_\gamma(x).\end{aligned}$$

In the case $N = 1$, for $t \in \mathbb{N}$ set

$$\ell_t = \frac{\mathbb{E}(F^t)}{\mathrm{Var}(F)^{t/2}}.$$

Note that ℓ_t is strictly connected to the Lyapunov ratio $L_t = \frac{\mathbb{E}(|F|^t)}{\mathrm{Var}(F)^{t/2}}$. By recalling that for $G \sim \mathrm{N}(0,1)$ then $\mathbb{E}(G^t) = 0$ if t is odd and $\mathbb{E}(G^t) = (t-1)!!$ if t is even (with the convention $(-1)!! = 1$), we obtain $\Delta_t = \ell_t$ if t is odd and $\Delta_t = \ell_t - (t-1)!!$ if t is even. Remark that $\Delta_3 = \ell_3$ and $\Delta_4 = \ell_4 - 3$ are the skewness and the kurtosis, respectively. Hence, we obtain the polynomials in the classical Edgeworth expansion:

$$\mathcal{K}_1(x) = \frac{\ell_3}{6} H_3(x), \qquad \mathcal{K}_2(x) = \frac{(\ell_4 - 3)}{24} H_4(x) + \frac{\ell_3^2}{72} H_6(x),$$
$$\mathcal{K}_3(x) = \left(-\frac{\ell_3}{3!2!} + \frac{\ell_5}{5!}\right) H_5(x) + \frac{\ell_3(\ell_4 - 3)}{3!4!} H_7(x) + \frac{\ell_3^3}{6(3!)^3} H_9(x).$$

Appendix A: Probability measures which are locally lower bounded by the Lebesgue measure

We discuss here the proof of Proposition 2.4. For a random variable $F \in \mathbb{R}^N$ with law μ_F, we recall that $\mu_F \succeq \mathrm{Leb}_N$ if there exists an open set $D \subset \mathbb{R}^N$ and $\varepsilon > 0$ such that

$$\mu_F(A) := \mathbb{P}(F \in A) \geq \varepsilon \,\mathrm{Leb}_N(A \cap D) \qquad \forall A \in \mathcal{B}(\mathbb{R}^N). \tag{A.1}$$

Remark that we have already proved that if $\mu_F \succeq \mathrm{Leb}_N$ then (2.3) holds (see Proposition 3.1).

We first prove the equivalence (i) $\Leftrightarrow$ (ii):

Lemma A.1. *$\mu_F \succeq \mathrm{Leb}_N$ if and only if there exists a non-negative measure μ with $\mu(\mathbb{R}^N) < 1$ and a non-negative lower semi-continuous function p with $\int_{\mathbb{R}^N} p(v)\,dv = 1 - \mu(\mathbb{R}^N)$ such that*

$$\mu_F(dv) = \mu(dv) + p(v)\,dv. \tag{A.2}$$

Proof. If (A.1) holds, we take $v_0 \in D$ and $r > 0$ such that $B_r(v_0) \subset D$. Then it suffices to take $p(x) = \varepsilon 1_{B_r(v_0)}(x)$ and $\mu(A) = \mathbb{P}(F \in A) - \int_A p(v)\,dv$, which turns out to be a non-negative measure.

Suppose now that (A.2) holds. Since p is non-negative and lower semicontinuous we may find an increasing sequence of non-negative and continuous functions $p_n, n \in \mathbb{N}$ such that $p_n \uparrow p$. It follows that $\int p_n \uparrow \int p = 1 - \mu(\mathbb{R}^N) > 0$, and we may find n such that $\int p_n > 0$. So there exists v_0 such that $p_n(v_0) > 0$. Since p_n is continuous, this implies that $p(v) \geq p_n(v) \geq \frac{1}{2} p_n(v_0)$ for $|v - v_0| < r$ for some small r. □

As a consequence, we get the final property in Proposition 2.4.

Lemma A.2. *If $\mu_F \succeq \mathrm{Leb}_N$, then the covariance matrix of F is invertible.*

Proof. We fix $v_0 \in \mathbb{R}^N$ and $\varepsilon > 0$ such that (A.1) holds with $D = B_r(v_0)$. We assume that $\mathbb{E}(F^i) = 0$ so that the covariance matrix is given by $C^{i,j}(F) = \mathbb{E}(F^i F^j)$. Then, for $\xi \in \mathbb{R}^N$ we write

$$\langle C(F)\xi, \xi\rangle = \mathbb{E}\big(\langle F, \xi\rangle^2\big) \geq \varepsilon \int_{B_r(v_0)} \langle v, \xi\rangle^2 \, dv.$$

We denote $A_\delta(\xi) = \{v : \langle v, \xi\rangle^2 \geq \delta|\xi|^2\}$ and we note that we may choose $\delta(v_0, r)$ such that

$$\inf_{|\xi|=1} \mathrm{Leb}_N\big(A_{\delta(v_0,r)}(\xi)\big) =: \eta(v_0, r) > 0.$$

Then

$$\inf_{|\xi|=1} \langle C(F)\xi, \xi\rangle \geq \varepsilon\eta(v_0, r)\,\mathrm{Leb}_N\big(B_r(v_0)\big).$$

□

We have already proved in Proposition 3.1 the implication (i) ⇒ (iii). Last implication (iii) ⇒ (ii) is trivial. In fact, let

$$\mathbb{P}\big(\chi V + (1-\chi)W \in dv\big) = \mathbb{P}(F \in dv),$$

where χ is a Bernoulli r.v. with parameter $p > 0$, V in $\mathbb{R}^N$ is absolutely continuous and W is a r.v. in $\mathbb{R}^N$. Setting μ_F, μ_V and μ_W, the law of F, V and W, respectively, then

$$\mu_F(dv) = p\mu_V(v)\,dv + (1-p)\mu_W(dv),$$

so F has an absolutely continuous component.

Appendix B: Estimates for the Sobolev norms in Lemma 4.10

This section is devoted to the proof the estimates used in Lemma 4.10, that is, the following.

Lemma B.1. *Let $d \geq 1$, $m \in \mathbb{N}$, $p \geq 1$. Then there exists $C > 0$ such that for every $K > 1$ and $X = (X^1, \dots, X^d)$ the following estimates holds*:

$$\big\|\Psi_K(X)X\big\|_{m,p} \leq CK\big(1 + \|X\|_{1,m,(m+1)p}\big)^{m+1}, \tag{B.1}$$

$$\big\|L\big(\Psi_K(X)X\big)\big\|_{m,p} \leq CK\big(1 + \|X\|_{1,m+1,4(m\vee 2)p}\big)^{2m+3}\big(1 + \|LX\|_{m,4p}\big), \tag{B.2}$$

where $\Psi_K(X)$ denote any function in $C^\infty(\mathbb{R}^d)$ such that $1_{B_K(0)} \leq \Psi_K \leq 1_{B_{K+1}(0)}$ and whose derivatives are uniformly bounded, that is there exists $L > 0$ such that $|\partial_\alpha \Psi_K| \leq L$ for every multiindex α.

Proof. For a multiindex α, one has

$$D_\alpha\big(\Psi_K(X)X^i\big) = D_\alpha\Psi_K(X)X^i + \sum_{\beta,\gamma\in A_\alpha, |\beta|\geq 1} D_\gamma\Psi_K(X)D_\beta X^i,$$

where the condition "$\beta,\gamma\in A_\alpha$" means that β,γ is a partition of α. Moreover, one has

$$D_\gamma \Psi_K(X)=\sum_{\ell=1}^{|\gamma|}\sum_{|\rho|=\ell}\partial_\rho\Psi_K(X)\sum_{\beta_1,\dots,\beta_\ell\in\mathcal{B}_\gamma}D_{\beta_1}X^{\rho_1}\cdots D_{\beta_\ell}X^{\rho_\ell},$$

where "$\beta_1,\dots,\beta_\ell\in\mathcal{B}_\gamma$" means that $\beta_1,\dots,\beta_\ell$ are non-empty multiindexes of γ running through the list of all of the (non-empty) "blocks" of γ. Then, for $|\gamma|\le m$ we obtain

$$\bigl|D_\gamma\Psi_K(X)\bigr|\le C1_{|X|\le K+1}\Bigl(1+\sum_{1\le|\rho|\le m}|D_\rho X|\Bigr)^m. \tag{B.3}$$

So, for $|\alpha|=m$ we have

$$\bigl|D_\alpha\bigl(\Psi_K(X)X\bigr)\bigr|\le CK\bigl(1+|X|_{1,m}\bigr)^{m+1}$$

and (B.1) follows. Consider now $L(\Psi_K(X)X^l)$. We have

$$-L\bigl(\Psi_K(X)X^l\bigr)=-L\Psi_K(X)X^l-\Psi_K(X)LX^l+\sum_{k=1}^n\sum_{i=1}^d D_{(k,i)}\Psi_K(X)D_{(k,i)}X^l.$$

We use now the inequality $\|XY\|_{m,p}\le C\|X\|_{m,2p}\|Y\|_{m,2p}$. But concerning the first term of the right-hand side of the equality above, we take care of the derivatives of Ψ_K as done to obtain formula (B.3) and we get

$$\begin{aligned}\bigl\|L\bigl(\Psi_K(X)X\bigr)\bigr\|_{m,p}&\le C\bigl\|L\Psi_K(X)\bigr\|_{m,2p}\bigl(\|X1_{|X|<K+1}\|_{2p}+\|X\|_{1,m,2p}\bigr)\\&\le CK\bigl\|L\Psi_K(X)\bigr\|_{m,2p}\bigl(1+\|X\|_{1,m,2p}\bigr).\end{aligned}$$

So, we obtain

$$\begin{aligned}\bigl\|L\bigl(\Psi_K(X)X\bigr)\bigr\|_{m,p}\le{}& C\bigl(K\bigl\|L\Psi_K(X)\bigr\|_{m,2p}\bigl(1+\|X\|_{1,m,2p}\bigr)\\&+\bigl\|\Psi_K(X)\bigr\|_{m,2p}\|LX\|_{m,2p}+\bigl\|\Psi_K(X)\bigr\|_{1,m,2p}\|X\|_{1,m,2p}\bigr).\end{aligned}$$

(B.3) gives that

$$\bigl\|\Psi_K(X)\bigr\|_{m,2p}\le C\bigl(1+\|X\|_{1,m,2mp}\bigr)^m, \tag{B.4}$$

so we can write

$$\bigl\|L\bigl(\Psi_K(X)X\bigr)\bigr\|_{m,p}\le CK\bigl(1+\|X\|_{1,m,2mp}\bigr)^{m+1}\bigl(1+\bigl\|L\Psi_K(X)\bigr\|_{m,2p}+\|LX\|_{m,2p}\bigr).$$

It remains to estimate $\|L\Psi_K(X)\|_{m,2p}$. Since

$$L\Psi_K(X)=\sum_{j=1}^d\partial_j\Psi_K(X)LX^j-\frac12\sum_{i,j=1}^d\partial_i\partial_j\Psi_K(X)\langle DX^i,DX^j\rangle$$

we have

$$\big\|L\Psi_K(X)\big\|_{m,2p} \le C\big(\big\|\nabla\Psi_K(X)\big\|_{m,4p}\|LX\|_{m,4p} + \big\|\nabla^2\Psi_K(X)\big\|_{m,4p}\|DX\|_{m,8p}^2\big).$$

An inequality analogous to (B.4) can be proved for $\nabla\Psi_K$ and $\nabla^2\Psi_K$, so we obtain

$$\begin{aligned}\big\|L\Psi_K(X)\big\|_{m,2p} &\le C\big(\big(1+\|X\|_{1,m,4mp}\big)^m\|LX\|_{m,4p} + \big(1+\|X\|_{1,m,4mp}\big)^m\|X\|_{1,m+1,8p}^2\big)\\ &\le C\big(1+\|X\|_{1,m+1,4(m\vee 2)p}\big)^{m+2}\big(1+\|LX\|_{m,4p}\big).\end{aligned}$$

Therefore, we can write

$$\begin{aligned}\big\|L\big(\Psi_K(X)X\big)\big\|_{m,p} &\le CK\big(1+\|X\|_{1,m,2mp}\big)^{m+1}\big(1+\|X\|_{1,m+1,4(m\vee 2)p}\big)^{m+2}\\ &\quad\times\big(1+\|LX\|_{m,4p}+\|LX\|_{m,2p}\big)\\ &\le CK\big(1+\|X\|_{1,m+1,4(m\vee 2)p}\big)^{2m+3}\big(1+\|LX\|_{m,4p}\big)\end{aligned}$$

and the statement holds. □

Appendix C: A backward Taylor formula for the Gaussian law

Proposition C.1. *Let G denote a centred normal distributed r.v. in $\mathbb{R}^N$. Then for every $L\in\mathbb{N}$ and $g\in C_b^{2(L+1)}(\mathbb{R}^N)$ one has*

$$g(0) = \sum_{\ell=0}^{L}\frac{(-1)^\ell}{2^\ell \ell!}\sum_{|\alpha|=2\ell}\theta_\alpha\mathbb{E}\big(\partial_\alpha g(G)\big) + \frac{(-1)^{L+1}}{2^{L+1}L!}\sum_{|\alpha|=2(L+1)}\theta_\alpha\int_0^1 s^L\mathbb{E}\big(\partial_\alpha g(\sqrt{s}\,G)\big)\,ds, \quad \text{(C.1)}$$

where for $|\alpha|=0$ then $\theta_\alpha=1$ and for $|\alpha|=r>0$, $\theta_\alpha=1$ if r is even and $\alpha_{2j-1}=\alpha_{2j}$ for every $j=1,\dots,r/2$, otherwise $\theta_\alpha=0$.

Proof. Let W denote a Brownian motion in $\mathbb{R}^N$. By Itô's formula, one has $\mathbb{E}(g(W_1)) = g(W_t) + \frac{1}{2}\int_t^1\mathbb{E}(\Delta g(W_s))\,ds$, so we can write

$$\mathbb{E}\big(g(W_t)\big) = g(W_1) - \frac{1}{2}\sum_{|\alpha|=2}\theta_\alpha\int_t^1\mathbb{E}\big(\partial_\alpha g(W_s)\big)\,ds. \quad \text{(C.2)}$$

Taking $t=0$, this gives

$$g(0) = \mathbb{E}\big(g(W_1)\big) - \frac{1}{2}\sum_{|\alpha|=2}\theta_\alpha\int_0^1\mathbb{E}\big(\partial_\alpha g(W_s)\big)\,ds.$$

By iteration, we write

$$g(0)=\mathbb{E}\big(g(W_1)\big)-\frac{1}{2}\sum_{|\alpha|=2}\theta_\alpha\mathbb{E}\big(\partial_\alpha g(W_1)\big)-\frac{1}{2}\sum_{|\alpha|=2}\theta_\alpha\int_0^1\big[\mathbb{E}\big(\partial_\alpha g(W_s)\big)-\mathbb{E}\big(\partial_\alpha g(W_1)\big)\big]\,ds$$

and by using (C.2) we get

$$g(0)=\mathbb{E}\big(g(W_1)\big)-\frac{1}{2}\sum_{|\alpha|=2}\theta_\alpha\mathbb{E}\big(\partial_\alpha g(W_1)\big)+\frac{1}{4}\sum_{|\alpha|=4}\theta_\alpha\int_0^1 u\mathbb{E}\big(\partial_\alpha g(W_u)\big)\,du.$$

Further iterations then give

$$g(0)=\sum_{\ell=0}^{L}\frac{(-1)^\ell}{2^\ell \ell!}\sum_{|\alpha|=2\ell}\theta_\alpha\mathbb{E}\big(\partial_\alpha g(W_1)\big)+\frac{(-1)^{L+1}}{2^{L+1}L!}\sum_{|\alpha|=2L+2}\theta_\alpha\int_0^1 s^L\mathbb{E}\big(\partial_\alpha g(W_s)\big)\,ds.$$

(C.1) now follows because, for every $s\in[0,1]$, $\sqrt{s}\,G$ and W_s have the same law. □

References

[1] Abramowitz, M. and Stegun, C.A. (1972). Bernoulli and Euler polynomials and the Euler–Maclaurin formula. In *Handbook of Mathematical Functions with Formulas*, *Graphs*, *and Mathematical Tables* 9th ed. New York: Dover.

[2] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Cham: Springer. MR3155209

[3] Bally, V. and Caramellino, L. (2014). On the distances between probability density functions. *Electron*. *J*. *Probab*. **19** 1–33. MR3296526

[4] Bally, V. and Clément, E. (2011). Integration by parts formula and applications to equations with jumps. *Probab*. *Theory Related Fields* **151** 613–657. MR2851695

[5] Bhattacharaya, R.N. and Ranga Rao, R. (2010). *Normal Approximation and Asymptotic Expansions*. *SIAM Classics in Applied Mathematics* **64**. Philadelphia: SIAM.

[6] Bhattacharya, R.N. (1968). Berry–Esseen bounds for the multi-dimensional central limit theorem. *Bull*. *Amer*. *Math*. *Soc*. **74** 285–287. MR0226706

[7] Bobkov, S.G., Chistyakov, G.P. and Götze, F. (2014). Berry–Esseen bounds in the entropic central limit theorem. *Probab*. *Theory Related Fields* **159** 435–478. MR3230000

[8] Dudley, R.M. (2002). *Real Analysis and Probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. MR1932358

[9] Nourdin, I., Nualart, D. and Poly, G. (2013). Absolute continuity and convergence of densities for random vectors on Wiener chaos. *Electron*. *J*. *Probab*. **18** 19 pp. MR3035750

[10] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus*: *From Stein's Method to Universality*. *Cambridge Tracts in Mathematics* **192**. Cambridge: Cambridge Univ. Press. MR2962301

[11] Nourdin, I. and Poly, G. (2013). Convergence in total variation on Wiener chaos. *Stochastic Process*. *Appl*. **123** 651–674. MR3003367

[12] Nourdin, I. and Poly, G. (2015). An invariance principle under the total variation distance. *Stochastic Process*. *Appl*. **125** 2190–2205. MR3322861

[13] Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. *Probability and Its Applications* (*New York*). Berlin: Springer. MR2200233

[14] Prohorov, Yu.V. (1952). A local theorem for densities. *Doklady Akad*. *Nauk SSSR* (*N.S.*) **83** 797–800. MR0049501

[15] Ranga Rao, R. (1961). On the central limit theorem in R_k. *Bull*. *Amer. Math*. *Soc*. **67** 359–361. MR0133150

[16] Sirazhdinov, S.Kh. and Mamatov, M. (1962). On convergence in the mean for densities. *Theory Probab. Appl*. **7** 424–428.

Received July 2014 and revised February 2015