



HAL
open science

An invariance principle under the total variation distance

Ivan Nourdin, Guillaume Poly

► **To cite this version:**

Ivan Nourdin, Guillaume Poly. An invariance principle under the total variation distance. *Stochastic Processes and their Applications*, 2015, 125 (6), pp.2190-2205. 10.1016/j.spa.2014.12.010. hal-00873311

HAL Id: hal-00873311

<https://hal.science/hal-00873311>

Submitted on 15 Oct 2013

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

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

An invariance principle under the total variation distance

Ivan Nourdin* (Université de Lorraine)
Guillaume Poly (Université du Luxembourg)

Abstract: Let X_1, X_2, \dots be a sequence of i.i.d. random variables, with mean zero and variance one. Let $W_n = (X_1 + \dots + X_n)/\sqrt{n}$. An old and celebrated result of Prohorov [16] asserts that W_n converges in total variation to the standard Gaussian distribution if and only if W_{n_0} has an absolutely continuous component for some n_0 . In the present paper, we give yet another proof and extend Prohorov's theorem to a situation where, instead of W_n , we consider more generally a sequence of homogeneous polynomials in the X_i . More precisely, we exhibit conditions for a recent invariance principle proved by Mossel, O'Donnell and Oleszkiewicz [14] to hold under the total variation distance. There are many works about CLT under various metrics in the literature, but the present one seems to be the first attempt to deal with homogeneous polynomials in the X_i with degree *strictly* greater than one.

Keywords: Convergence in law; convergence in total variation; absolute continuity; invariance principle.

1 Introduction and main results

Let X_1, X_2, \dots be independent copies of a random variable with mean zero and variance one. According to the central limit theorem, the normalized sums

$$W_n = \frac{X_1 + \dots + X_n}{\sqrt{n}} \tag{1.1}$$

converge in distribution to the standard normal law $N \sim N(0, 1)$. In fact, using, e.g., the second Dini's theorem it is straightforward to prove a much stronger result, namely that W_n converges to N in the Kolmogorov distance:

$$\lim_{n \rightarrow \infty} d_{Kol}(W_n, N) = 0, \tag{1.2}$$

where $d_{Kol}(U, V) = \sup_{x \in \mathbb{R}} |P(U \leq x) - P(V \leq x)|$.

In (1.2), can we replace the Kolmogorov distance d_{Kol} by the total variation distance d_{TV} , defined as $d_{TV}(U, V) = \sup_{A \in \mathcal{B}(\mathbb{R})} |P(U \in A) - P(V \in A)|$? In other words, do we also have

$$\lim_{n \rightarrow \infty} d_{TV}(W_n, N) = 0 \tag{1.3}$$

for W_n defined by (1.1)? The right answer is provided by an old and celebrated result of Prohorov [16]. To formulate it, first let us introduce the Lebesgue decomposition (of the distribution) of a

*supported in part by the (french) ANR grant 'Malliavin, Stein and Stochastic Equations with Irregular Coefficients' [ANR-10-BLAN-0121].

random variable. As is well known, each cumulative distribution function (cdf) F may indeed be represented in the form:

$$F(x) = u \int_{-\infty}^x g(y)dy + (1-u)G(x), \quad x \in \mathbb{R}, \quad (1.4)$$

where $u \in [0, 1]$, $g : \mathbb{R} \rightarrow [0, \infty)$ satisfies $\int_{\mathbb{R}} g(y)dy = 1$ and G is a singular cdf (corresponding to a distribution concentrated on a set of zero Lebesgue measure) with $G'(x) = 0$ for almost all x . The real number $u \in [0, 1]$ is uniquely determined by F ; the density function g is uniquely determined (up to a set of measure zero) if and only if $u > 0$.

Definition 1.1 *When X is a random variable with cdf F , we say that X is **singular** if $u = 0$ in (1.4). If $u > 0$, we say that X has an **absolutely continuous component** with density g .*

We can now state Prohorov's theorem [16]. A proof will be given in Section 2.4, to illustrate a possible use of our forthcoming results.

Theorem 1.2 (Prohorov) *The convergence (1.3) takes place if and only if there exists $n_0 \geq 1$ such that the random variable W_{n_0} has an absolutely continuous component.*

Prohorov's theorem has been the starting point of a fruitful line of research around the validity of the central limit theorem under various metrics and the estimation of their associated rates of convergence. Let us only give a small sample of references dealing with this rich and well studied topic. Convergence of densities in L^∞ are studied by Gnedenko and Kolmogorov [9]. On their side, Mamatov and Halikov [11] dealt with the multivariate CLT in total variation. Barron [2] studied the convergence in relative entropy, whereas Shimizu [17] and Johnson and Barron [10] studied the convergence in Fisher information. As far as rates of convergence are concerned, one can quote Mamatov and Sirazdinov [13] for the total variation distance and, very recently, Bobkov, Chistyakov and Götze for bounds in entropy [3], in Fisher information [4] and for Edgeworth-type expansions in the entropic central limit theorem [5]. Finally we mention [6, 7] for a variational approach of these issues with some variance bounds.

All these above-mentioned references have in common to 'only' deal with *sums* of independent random variables. In the present paper, in contrast, we will consider highly *non-linear* functionals of independent random variables. It is arguably a much harder framework to work with, precisely because all the nice properties enjoyed by sums of independent variables are no longer valid in this context (in particular, the use of characteristic functions is not appropriate).

Let us now turn to the details of the situation we are considering in the present article. Fix a degree of multilinearity $d \geq 1$ ($d = 1$ for linear, $d = 2$ for quadratic, etc.) and, for any $n \geq 1$, consider a homogeneous polynomial $Q_n : \mathbb{R}^{N_n} \rightarrow \mathbb{R}$ of the form

$$Q_n(\mathbf{x}) = \sum_{i_1, \dots, i_d=1}^{N_n} a_n(i_1, \dots, i_d) x_{i_1} \dots x_{i_d}, \quad \mathbf{x} = (x_1, \dots, x_{N_n}) \in \mathbb{R}^{N_n}. \quad (1.5)$$

In (1.5), it is implicitly supposed that $N_n \rightarrow \infty$ and also that $a_n(i_1, \dots, i_d)$ are real numbers vanishing on diagonals and symmetric in the indices. We further assume, in all what follows, that Q_n is properly normalized:

$$d! \sum_{i_1, \dots, i_d=1}^{N_n} a_n(i_1, \dots, i_d)^2 = 1, \quad n \geq 1, \quad (1.6)$$

and that all the terms of Q_n are asymptotically negligible, meaning in our context that

$$\lim_{n \rightarrow \infty} \max_{1 \leq i_1 \leq n} \sum_{\substack{N_n \\ i_2, \dots, i_d = 1}} a_n(i_1, \dots, i_d)^2 = 0. \quad (1.7)$$

As anticipated, property (1.7) will play a crucial role in the sequel. It is also the key to obtain the following invariance principle, due to Mossel, O'Donnell and Oleszkiewicz [14].

Theorem 1.3 (Mossel, O'Donnell, Oleszkiewicz) *Fix an integer $d \geq 1$, and let Q_n be a sequence of homogeneous polynomials satisfying both (1.5), (1.6) and (1.7). Let $\mathbf{X} = (X_1, X_2, \dots)$ be a sequence of independent random variables with mean zero and variance one, belonging to $L^{2+\epsilon}(\Omega)$ for some $\epsilon > 0$ (the same ϵ for each X_i). Assume the same for $\mathbf{Y} = (Y_1, Y_2, \dots)$. Then*

$$\lim_{n \rightarrow \infty} d_{Kol}(Q_n(\mathbf{X}), Q_n(\mathbf{Y})) = 0. \quad (1.8)$$

Observe that one recovers (1.2) by considering, in (1.8), $d = 1$, $a_n(i) = \frac{1}{\sqrt{n}}$, $1 \leq i \leq n$ (which satisfies (1.7)) and $Y_1 \sim N(0, 1)$ (which leads to $Q_n(\mathbf{Y}) \sim N(0, 1)$ for any n).

In the light of the aforementioned results, it seemed natural to us to ask under which assumption the convergence (1.8) may be strengthened to the total variation distance as follows:

$$\lim_{n \rightarrow \infty} d_{TV}(Q_n(\mathbf{X}), Q_n(\mathbf{Y})) = 0. \quad (1.9)$$

Before detailing our answer, let us first do a quick digression. As anticipated, a main aspect of our approach will consist in introducing the following class of random variables.

Definition 1.4 *For any $p \in]0, 1]$ and $\alpha > 0$, the class $\mathcal{C}(p, \alpha)$ is the set of real random variables X satisfying*

$$X \stackrel{\text{law}}{=} \varepsilon(\alpha U + x_0) + (1 - \varepsilon)V, \quad (1.10)$$

where $x_0 \in \mathbb{R}$ is a real number, and $U \sim \mathcal{U}_{[-1, 1]}$, $\varepsilon \sim \mathcal{B}(p)$ and V (with no specified distribution) are three independent random variables.

Since it will play a crucial role in the sequel, let us first try to catch the meaning of (1.10). To this aim, we introduce yet another class of random variables.

Definition 1.5 *For any $c, \alpha > 0$, the class $\mathcal{G}(c, \alpha)$ is the set of real random variables X having an absolutely continuous component and whose density g , see (1.4), satisfies $g(x) \geq c$ for all $x \in [x_0 - \alpha, x_0 + \alpha]$ for some $x_0 \in \mathbb{R}$.*

The following result compares the two classes $\mathcal{C}(p, \alpha)$ and $\mathcal{G}(c, \alpha)$. Roughly speaking, it asserts that the class of random variables with an absolutely continuous component (that is, exactly the kind of random variables appearing in Prohorov's Theorem 1.2) coincides with $\cup_{p \in]0, 1], \alpha > 0} \mathcal{C}_{p, \alpha}$. Observe also that $\mathcal{G}(c, \alpha)$ is not empty if and only if $2c\alpha \leq 1$.

Proposition 1.6 *Fix $c, \alpha > 0$ and $p \in]0, 1]$. One has $\mathcal{G}(c, \alpha) \subset \mathcal{C}(2c\alpha, \alpha)$. Moreover, any random variable belonging to $\mathcal{C}(p, \alpha)$ has an absolutely continuous part.*

In Lemma 2.2 below, we state two further important properties of $\mathcal{C}(p, \alpha)$. Firstly, the sum of two independent random variables having an absolutely continuous component belong to $\cup_{c, \alpha > 0} \mathcal{G}(c, \alpha) \subset \cup_{p \in]0, 1], \alpha > 0} \mathcal{C}(p, \alpha)$. Secondly, if $0 < q \leq p \leq 1$ and $0 < \beta \leq \alpha$, then $\mathcal{C}(p, \alpha) \subset \mathcal{C}(q\beta/\alpha, \beta)$.

Now $\mathcal{C}(p, \alpha)$ has been introduced and is arguably well-understood, let us give a name to the set of sequences of independent and normalized random variables we will deal with throughout the sequel.

Definition 1.7 *Let $\alpha > 0$, $p \in]0, 1]$ and $\epsilon > 0$. A sequence $\mathbf{X} = (X_1, X_2, \dots)$ of random variables belongs to $\mathcal{D}(\alpha, p, 2 + \epsilon)$ if the X_i are independent, satisfy $\sup_i E|X_i|^{2+\epsilon} < \infty$ and if, for each i , $E[X_i] = 0$, $E[X_i^2] = 1$ and $X_i \in \mathcal{C}(p, \alpha)$.*

We are now in a position to state the main result of the present paper.

Theorem 1.8 *Fix an integer $d \geq 1$, and let Q_n be a sequence of homogeneous polynomials satisfying both (1.5), (1.6 and (1.7). Let \mathbf{X} and \mathbf{Y} belong to $\mathcal{D}(\alpha, p, 2 + \epsilon)$ for some $\epsilon, \alpha > 0$ and $p \in]0, 1]$. Then (1.9) holds true.*

A noticeable corollary of Theorem 1.8 is a new proof of Prohorov's Theorem 1.2, see Section 2.4. Another one is the following result.

Corollary 1.9 *Fix an integer $d \geq 2$, and let Q_n be a sequence of homogeneous polynomials satisfying (1.5)-(1.6)-(1.7). Let \mathbf{X} belong to $\mathcal{D}(\alpha, p, 2 + \epsilon)$ for some $\epsilon, \alpha > 0$ and $p \in]0, 1]$. If $Q_n(\mathbf{X})$ converges in law to W , then W has a density and $Q_n(\mathbf{X})$ converges to W in total variation.*

The statement of Corollary 1.9 would be clearly wrong without assuming (1.7). Consider, e.g., $Q_n(\mathbf{x}) = x_1$, $n \geq 1$ with X_1 singular. Another interesting consequence of Theorem 1.8 is provided by the next theorem.

Theorem 1.10 *Let $\{a_n(i_1, \dots, i_d)\}_{1 \leq i_1, \dots, i_d \leq N_n}$ be an array of real numbers vanishing on diagonals, symmetric in the indices and satisfying (1.6). (We do not suppose (1.7).) Let $\mathbf{G} = (G_1, G_2, \dots)$ be a sequence of independent $N(0, 1)$ random variables, and let $N \sim N(0, 1)$. Then, the following four assertions are equivalent as $n \rightarrow \infty$.*

$$(a) \quad \sum_{i_1, i_2, \dots, i_d=1}^{N_n} a_n(i_1, i_2, \dots, i_d) G_{i_1} G_{i_2} \cdots G_{i_d} \xrightarrow{\text{law}} N.$$

$$(b) \quad d_{TV} \left(N, \sum_{i_1, i_2, \dots, i_d=1}^{N_n} a_n(i_1, i_2, \dots, i_d) G_{i_1} G_{i_2} \cdots G_{i_d} \right) \rightarrow 0.$$

(c) *For all $\mathbf{X} = (X_1, X_2, \dots)$ belonging to $\mathcal{D}(\alpha, p, 2 + \epsilon)$ for some $\epsilon, \alpha > 0$ and $p \in]0, 1]$, we have*

$$\sum_{i_1, i_2, \dots, i_d=1}^{N_n} a_n(i_1, i_2, \dots, i_d) X_{i_1} X_{i_2} \cdots X_{i_d} \xrightarrow{\text{law}} N.$$

(d) For all $\mathbf{X} = (X_1, X_2, \dots)$ belonging to $\mathcal{D}(\alpha, p, 2 + \epsilon)$ for some $\epsilon, \alpha > 0$ and $p \in]0, 1]$, we have

$$d_{TV} \left(N, \sum_{i_1, i_2, \dots, i_d=1}^{N_n} a_n(i_1, i_2, \dots, i_d) X_{i_1} X_{i_2} \cdots X_{i_d} \right) \rightarrow 0.$$

The rest of our paper is organised as follows. In Section 2 we prove all the results that are stated in this Introduction, except Theorem 1.8; in particular, we give a new proof of Prohorov's Theorem 1.2 in Section 2.4. Finally, the proof of our main result, namely Theorem 1.8, is provided in Section 3.

2 Proofs of all stated results except Theorem 1.8

2.1 Some useful lemmas

The following lemma will be used several times in the sequel.

Lemma 2.1 Fix $q \in [0, 1]$, and let Y, Z be two random variables satisfying $E[f(Y)] \geq qE[f(Z)]$ for all positive bounded function f . Then there exists two independent random variable W and $\zeta \sim \mathcal{B}(q)$, independent from Z , such that

$$Y \stackrel{\text{law}}{=} \zeta Z + (1 - \zeta)W. \quad (2.11)$$

Proof of Lemma 2.1. Our assumption ensures that the linear form $f \mapsto E[f(Y)] - qE[f(Z)]$ is positive. From the Riesz representation theorem, one deduces the existence of a positive finite Radon measure ν such that

$$E[f(Y)] = qE[f(Z)] + \int_{\mathbb{R}} f(x) d\nu(x). \quad (2.12)$$

Choosing $f \equiv 1$ in (2.12) gives $\nu(\mathbb{R}) = 1 - q$. If $\nu(\mathbb{R}) = 0$ then $q = 1$ and the proof of (2.11) is established. Otherwise, $\nu(\mathbb{R}) > 0$ and one can consider $W \sim \frac{1}{\nu(\mathbb{R})} d\nu(x)$, implying in turn (2.11) \square

In the following lemma, we gather useful properties of the classes $\mathcal{C}(p, \alpha)$ and $\mathcal{G}(c, \alpha)$.

Lemma 2.2 The following properties take place.

1. If $0 < q \leq p \leq 1$ and if $0 < \beta \leq \alpha$, then $\mathcal{C}(p, \alpha) \subset \mathcal{C}(q\beta/\alpha, \beta)$. In particular, $\mathcal{C}(p, \alpha) \subset \mathcal{C}(q, \alpha)$.
2. If X and Y both have an absolutely continuous component and if X is independent from Y , then there exists $c, \alpha > 0$ such that $X + Y \in \mathcal{G}(c, \alpha)$.
3. If X belongs to $\mathcal{C}(p, \alpha)$ with $\alpha > 0$ and $p \in]0, 1]$ and if Y is any random variable independent from X , then $X + Y$ belongs to $\mathcal{C}(q, \beta)$ for some $\beta > 0$ and $q \in]0, 1]$.
4. If $a \neq 0$ and b are two real numbers and if X belongs to $\mathcal{C}(p, \alpha)$ with $\alpha > 0$ and $p \in]0, 1]$, then $aX + b \in \mathcal{C}(p, |a|\alpha)$.

Proof. 1. Fix $0 < q \leq p \leq 1$ and $0 < \beta \leq \alpha$, and consider $X \in \mathcal{C}(p, \alpha)$. According to (1.10), we have, for any positive f ,

$$\begin{aligned} E[f(X)] &= p \int_{\mathbb{R}} f(x) \frac{1}{2\alpha} \mathbf{1}_{[x_0-\alpha, x_0+\alpha]}(x) dx + (1-p)E[f(V)] \\ &\geq \frac{q\beta}{\alpha} \int_{\mathbb{R}} f(x) \frac{1}{2\beta} \mathbf{1}_{[x_0-\beta, x_0+\beta]}(x) dx. \end{aligned}$$

The conclusion follows from Lemma 2.1.

2. Consider the decomposition (1.4) of the cdf F of X . This settles $u \in]0, 1]$ and $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ in a unique way. Settle similarly $v \in]0, 1]$ and $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ for Y . For any Borel set A , one has $P(X \in A) \geq u \int_A g(x) dx$ and the same for Y . We deduce

$$P(X + Y \in A) \geq uv \int_A (g \star h)(x) dx,$$

with \star denoting the usual convolution. Besides, $g \star h = \lim_{M \rightarrow \infty} g \star \inf(h, M)$ and the limit is increasing by positivity of g . Finally, we note that, since $g \in L^1$ and $\inf(h, M) \in L^\infty$, the convolution $g \star \inf(h, M)$ is continuous. Let $x_0 \in \mathbb{R}$ and $M > 0$ be such that $(g \star \inf(h, M))(x_0) > 0$. (Such a pair (x_0, M) necessarily exists, otherwise we would have $g \star h \equiv 0$ by taking the large M limit.) By continuity, there exists $c > 0$ and $\alpha > 0$ such that, for any $x \in]x_0 - \alpha, x_0 + \alpha[$, $(g \star h)(x) \geq (g \star \inf(h, M))(x) \geq c$. That is, $X + Y$ belongs to $\mathcal{G}(c, \alpha)$.

3. We have $X \stackrel{\text{law}}{=} \varepsilon(\alpha U + x_0) + (1 - \varepsilon)V$, with $x_0 \in \mathbb{R}$ a real number, and $U \sim \mathcal{U}_{[-1, 1]}$, $\varepsilon \sim \mathcal{B}(p)$ and V (with no specified distribution) three independent random variables. On the other hand, one can write $Y \stackrel{\text{law}}{=} \varepsilon Y + (1 - \varepsilon)Z$, with Z having the same law than Y and independent from Y, U, ε, V . Thus,

$$X + Y \stackrel{\text{law}}{=} \varepsilon(\alpha U + x_0 + Y) + (1 - \varepsilon)(V + Z).$$

The random variable $\alpha U + x_0 + Y$ has a density g given by

$$g(v) = \int_{\mathbb{R}} \frac{1}{2\alpha} \mathbf{1}_{[x_0-\alpha, x_0+\alpha]}(v - y) dP_Y(y) = \frac{1}{2\alpha} P(Y \in [v - x_0 - \alpha, v - x_0 + \alpha]).$$

As a matter of fact, g is a regulated function, since it is the difference of two increasing functions. In particular, the set \mathcal{E} of its discontinuous points is countable. As a consequence, $\text{Leb}(\mathcal{E}) = 0$, implying in turn $1 = \int_{\mathbb{R}} g(v) dv = \int_{\mathbb{R} \setminus \mathcal{E}} g(v) dv$, so that there exists $x_1 \notin \mathcal{E}$ satisfying $g(x_1) > 0$. Since g is continuous at x_1 , there exists $r > 0$ such that $g(v) \geq \frac{1}{2}g(x_1)$ for all $v \in [x_1 - r, x_1 + r]$. By Lemma 2.1, it comes that

$$\alpha U + x_0 + Y \stackrel{\text{law}}{=} \eta(rU + x_1) + (1 - \eta)T,$$

where $\eta \sim \mathcal{B}(p')$ for some $p' \in]0, 1]$, $U \sim \mathcal{U}_{[-1, 1]}$ and T are independent. Hence

$$X + Y \stackrel{\text{law}}{=} \varepsilon \eta(rU + x_1) + \varepsilon(1 - \eta)T + (1 - \varepsilon)(V + Z).$$

As a result, for any bounded positive function,

$$\begin{aligned} E[f(X + Y)] &= pp'E[f(rU + x_1)] + p(1 - p')E[f(T)] + (1 - p)E[f(V + Z)] \\ &\geq pp'E[f(rU + x_1)]. \end{aligned}$$

Finally, one deduces that $X + Y$ belongs to $\mathcal{C}(pp', r)$ by Lemma 2.1.

4. Obvious. □

2.2 Proof of Proposition 1.6

Let X be an element of $\mathcal{G}(c, \alpha)$. Let F denote its cdf, and consider g , G and u as in (1.4). Let $Y \sim g(x)dx$, $Z \sim dG(x)$ and $\eta \sim \mathcal{B}(u)$ be three independent random variables. Then $X \stackrel{\text{law}}{=} \eta Y + (1 - \eta)Z$. Using the assumption made on g , one obtains, for any positive function f ,

$$E[f(Y)] \geq 2c\alpha E[f(\alpha U + x_0)].$$

Observe that $0 < 2c\alpha \leq 1$ necessarily. We then deduce that $X \in \mathcal{C}(2c\alpha, \alpha)$ from Lemma 2.1.

Consider now a random variable X belonging to $\mathcal{C}(p, \alpha)$. We have, for any positive function f and according to the decomposition (1.10),

$$E[f(X)] = pE[f(\alpha U + x_0)] + (1 - p)E[f(V)]. \quad (2.13)$$

Let us consider the Lebesgue decomposition (u, g, G) of V , see (1.4):

$$E[f(V)] = u \int_{\mathbb{R}} f(x)g(x)dx + (1 - u) \int_{\mathbb{R}} f(x)dG(x).$$

Plugging into (2.13) yields

$$E[f(X)] = \int_{\mathbb{R}} f(x) \left\{ \frac{p}{2\alpha} \mathbf{1}_{[x_0 - \alpha, x_0 + \alpha]}(x) + (1 - p)u g(x) \right\} dx + (1 - p)(1 - u) \int_{\mathbb{R}} f(x)dG(x),$$

from which we deduce that X has an absolutely continuous component. □

2.3 Proof of Corollary 1.9

Let the assumption of Corollary 1.9 prevail and consider a sequence $\mathbf{G} = (G_1, G_2, \dots)$ composed of independent copies of a standard Gaussian random variable. By the Mossel, O'Donnell, Oleszkiewicz's invariance principle (1.8), one has that $Q_n(\mathbf{G})$ converges in law to W . It implies, by [15, Theorem 3.1] (see also [15, Lemma 2.4]), that W has a density and $d_{TV}(Q_n(\mathbf{G}), W) \rightarrow 0$ as $n \rightarrow \infty$. But $d_{TV}(Q_n(\mathbf{X}), Q_n(\mathbf{G})) \rightarrow 0$ by our Theorem 1.8, so

$$d_{TV}(Q_n(\mathbf{X}), W) \leq d_{TV}(Q_n(\mathbf{X}), Q_n(\mathbf{G})) + d_{TV}(Q_n(\mathbf{G}), W) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

This concludes the proof of Corollary 1.9. □

2.4 Proof of Theorem 1.2

We want to use Corollary 1.9 but the problem is that \mathbf{X} is not assumed to belong to \mathcal{D} in the statement of Theorem 1.2. To overcome this issue we shall need Lemma 2.2.

Let the assumptions and notation of Theorem 1.2 prevail.

First, when W_n is singular, then there exists a Borel set A such that $P(N \in A) = 0$ and $P(W_n \in A) = 1$; in particular, $d_{TV}(W_n, N) = 1$. Hence, if W_n is singular for all n , then (1.3) cannot hold.

Now, assume that W_{n_0} has an absolutely continuous component for some n_0 . To prove (1.3) is obviously equivalent to prove that, for all $k \in \{0, \dots, 2n_0 - 1\}$,

$$\lim_{n \rightarrow \infty} d_{TV}(W_{2n_0n+k}, N) = 0. \quad (2.14)$$

So, fix $k \in \{0, \dots, 2n_0 - 1\}$ and consider a sequence (Y_2, Y_3, \dots) of independent copies of W_{2n_0} . By Lemma 2.2 (points 2 and 4), observe that each Y_i belongs to $\mathcal{C}(p, \alpha)$, for some $p \in]0, 1[$ and $\alpha > 0$ (the same p and the same α for all $i \geq 2$); also, we have $E[Y_i] = 0$ and $E[Y_i^2] = 1$. On the other hand, let Y_1 be independent of Y_2, Y_3, \dots and have the same law than W_{2n_0+k} . By Lemma 2.2 (points 3 and 4), Y_1 belongs to $\mathcal{C}(q, \beta)$ for some $q \in]0, 1[$ and $\beta > 0$; also, we have $E[Y_1] = 0$ and $E[Y_1^2] = 1$. In fact, thanks to Lemma 2.2 (point 1), one may and will choose the same p and the same α for each Y_i , without making a difference between $i = 1$ and $i \geq 2$.

Bearing all the previous notation in mind, we can write

$$W_n \stackrel{\text{law}}{=} \sqrt{\frac{2n_0+k}{n}} Y_1 + \sqrt{\frac{2n_0}{n}} \sum_{k=2}^n Y_k.$$

The convergence (2.14) is now a direct consequence of Theorem 1.8 applied to $d = 1$, a sequence $\mathbf{X} = \mathbf{G} = (G_1, G_2, \dots)$ of independent $N(0, 1)$ variables, $a_n(1) = \sqrt{\frac{2n_0+k}{n}}$ and $a_n(i) = \sqrt{\frac{2n_0}{n}}$, $i = 2, \dots, n$. Note that

$$d_{FM}(W_n, N) = d_{FM}(Q_n(\mathbf{Y}), Q_n(\mathbf{X})) \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

by the usual CLT, so that it is not necessary to rely on Theorem 1.3 to conclude the proof of Theorem 1.8 (see Step 7 of Section 3) and so to assume the existence of an absolute q th moment for X_1 with q strictly greater than 2. \square

2.5 Proof of Theorem 1.10

Implication (a) \Rightarrow (b) is a reformulation of [12, Corollary 5.2.8]. Implications (b) \Rightarrow (a) and (d) \Rightarrow (c) are obvious. Implication (c) \Rightarrow (a) is because \mathbf{G} belongs to $\mathcal{D}(\alpha, p, 2 + \varepsilon)$ for some $\alpha, \varepsilon > 0$ and $p \in]0, 1[$ (use Proposition 1.6). Finally, implication (a) \Rightarrow (d) is a consequence of the following two facts. Firstly, if (a) takes place then, by a usual hypercontractivity argument, the sequence $F_n := \sum_{i_1, i_2, \dots, i_d=1}^{N_n} a_n(i_1, i_2, \dots, i_d) G_{i_1} G_{i_2} \cdots G_{i_d}$ (which is normalized so that $E[F_n^2] = 1$, see indeed (1.6)) satisfies $E[F_n^4] \rightarrow 3$. Secondly, one has, according to [12, (11.4.7) and (11.4.8) pp. 192-193]:

$$\max_{1 \leq i_1 \leq n} \sum_{i_2, \dots, i_d=1}^{N_n} a_n(i_1, \dots, i_d)^2 \leq \frac{1}{d!} \sqrt{E[F_n^4] - 3}.$$

These two facts together imply that, if (a) holds, then (1.7) is automatically satisfied. Thus, Theorem 1.8 implies the validity of assertion (d). \square

3 Proof of Theorem 1.8

Let the assumptions and notation of Theorem 1.8. Without loss of generality, for simplicity we assume that $N_n = n$.

The proof is divided into several steps.

Step 1. In the definition of $Q_n(\mathbf{X})$ one may and will replace each X_i by $\varepsilon_i(\alpha U_i + x_i) + (1 - \varepsilon_i)V_i$, where $\mathbf{e} = (\varepsilon_1, \varepsilon_2, \dots)$ is a sequence of independent Bernoulli random variables ($\varepsilon_i \sim \mathcal{B}(p_i)$), $\mathbf{U} = (U_1, U_2, \dots)$ is a sequence of independent $[-1, 1]$ -uniformly distributed random variables and $\mathbf{V} = (V_1, V_2, \dots)$ is a sequence of independent random variables; moreover, \mathbf{e} , \mathbf{U} and \mathbf{V} are independent. That is,

$$Q_n(\mathbf{X}) = \sum_{i_1, \dots, i_d=1}^n a_n(i_1, \dots, i_d) \{\varepsilon_{i_1}(\alpha_{i_1} U_{i_1} + x_{i_1}) + (1 - \varepsilon_{i_1})V_{i_1}\} \dots \{\varepsilon_{i_d}(\alpha_{i_d} U_{i_d} + x_{i_d}) + (1 - \varepsilon_{i_d})V_{i_d}\}.$$

Now, let us expand everything, and then rewrite $Q_n(\mathbf{X})$ as a polynomial in the U_i . We obtain

$$Q_n(\mathbf{X}) = A_n + B_n + C_n,$$

where

$$\begin{aligned} A_n &= \sum_{i_1, \dots, i_d=1}^n a_n(i_1, \dots, i_d) \alpha_{i_1} \dots \alpha_{i_d} \varepsilon_{i_1} \dots \varepsilon_{i_d} U_{i_1} \dots U_{i_d} \\ B_n &= Q_n(\mathbf{X}) - A_n - C_n \\ C_n &= \sum_{i_1, \dots, i_d=1}^n a_n(i_1, \dots, i_d) \{\varepsilon_{i_1} x_{i_1} + (1 - \varepsilon_{i_1})V_{i_1}\} \dots \{\varepsilon_{i_d} x_{i_d} + (1 - \varepsilon_{i_d})V_{i_d}\} \end{aligned}$$

satisfying

$$E[A_n B_n | \mathbf{e}, \mathbf{V}] = E[A_n C_n | \mathbf{e}, \mathbf{V}] = E[B_n C_n | \mathbf{e}, \mathbf{V}] = E[A_n | \mathbf{e}, \mathbf{V}] = E[B_n | \mathbf{e}, \mathbf{V}] = 0.$$

As a result,

$$\begin{aligned} \text{Var}[Q_n(\mathbf{X}) | \mathbf{e}, \mathbf{V}] &= E[A_n^2 | \mathbf{e}, \mathbf{V}] + E[B_n^2 | \mathbf{e}, \mathbf{V}] + \text{Var}[C_n | \mathbf{e}, \mathbf{V}] \\ &\geq E[A_n^2 | \mathbf{e}, \mathbf{V}] \\ &\geq \alpha^{2d} 3^{-d} d! \sum_{i_1, \dots, i_d=1}^n a_n(i_1, \dots, i_d)^2 \varepsilon_{i_1} \dots \varepsilon_{i_d}. \end{aligned}$$

To go one step further, let us decompose ε_i into $(\varepsilon_i - p) + p$ and use (1.6), so to obtain

$$\begin{aligned} &\text{Var}[Q_n(\mathbf{X}) | \mathbf{e}, \mathbf{V}] \\ &\geq d! \left(\frac{\alpha^2 p}{3} \right)^d + \alpha^{2d} 3^{-d} d! \sum_{k=1}^d \frac{\binom{d}{k}}{p^k} \sum_{i_1, \dots, i_d=1}^n a_n(i_1, \dots, i_d)^2 (\varepsilon_{i_1} - p) \dots (\varepsilon_{i_k} - p). \end{aligned}$$

Using the assumptions (1.6) and (1.7), we can write, for any fixed $k \in \{1, \dots, d\}$,

$$\begin{aligned}
& E \left[\left(\sum_{i_1, \dots, i_d=1}^n a_n(i_1, \dots, i_d)^2 (\varepsilon_{i_1} - p) \dots (\varepsilon_{i_k} - p) \right)^2 \right] \\
&= p^k (1-p)^k \sum_{i_1, \dots, i_k=1}^n \left(\sum_{i_{k+1}, \dots, i_d=1}^n a_n(i_1, \dots, i_d)^2 \right)^2 \\
&\leq \sum_{i_1=1}^n \left(\sum_{i_2, \dots, i_d=1}^n a_n(i_1, \dots, i_d)^2 \right)^2 \\
&\leq \max_{1 \leq j_1 \leq n} \sum_{j_2, \dots, j_d=1}^n a_n(j_1, \dots, j_d)^2 \times \sum_{i_1, \dots, i_d=1}^n a_n(i_1, \dots, i_d)^2 \\
&= \frac{1}{d!} \max_{1 \leq j_1 \leq n} \sum_{j_2, \dots, j_d=1}^n a_n(j_1, \dots, j_d)^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.
\end{aligned}$$

We deduce that, in probability,

$$\liminf_{n \rightarrow \infty} \text{Var}[Q_n(\mathbf{X}) | \mathbf{e}, \mathbf{V}] \geq d! \left(\frac{\alpha^2 p}{3} \right)^d. \quad (3.15)$$

Convention. From now on, and since all the quantities we are dealing with are measurable with respect to \mathbf{e} , \mathbf{U} and \mathbf{V} , we shall write $E_{\mathbf{U}}$ (resp. $E_{\mathbf{e}, \mathbf{V}}$) to indicate the mathematical expectation with respect to \mathbf{U} (resp. \mathbf{e} and \mathbf{V}). Note that $E_{\mathbf{U}}$ coincides with the conditional expectation $E[\cdot | \mathbf{e}, \mathbf{V}]$.

Step 2. Set $p_\alpha(x) = \frac{1}{\alpha\sqrt{2\pi}} e^{-\frac{x^2}{2\alpha^2}}$, $x \in \mathbb{R}$, $0 < \alpha \leq 1$, and let $\phi \in C_c^\infty$ be bounded by 1. It is immediately checked that

$$\|\phi \star p_\alpha\|_\infty \leq 1 \leq \frac{1}{\alpha} \quad \text{and} \quad \|(\phi \star p_\alpha)'\|_\infty \leq \frac{1}{\alpha}. \quad (3.16)$$

We can write

$$\begin{aligned}
& |E[\phi(Q_n(\mathbf{X}))] - E[\phi(Q_n(\mathbf{Y}))]| \\
&\leq |E[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X}))]| + |E[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{Y}))]| + \frac{1}{\alpha} d_{FM}(Q_n(\mathbf{X}), Q_n(\mathbf{Y})),
\end{aligned}$$

where d_{FM} stands for the Fortet-Mourier distance, which is known to metrize the convergence in

law. Let us concentrate on the first two terms. We have, e.g., for the first term:

$$\begin{aligned}
& |E[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X}))]| \\
\leq & \left| E[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] < \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d\}}] \right| \\
& + |E[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \mathbf{1}_{\{E[Q_n(\mathbf{X})^2|\mathbf{e}, \mathbf{V}] > M\}}]| \\
& + \left| E[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d, E[Q_n(\mathbf{X})^2|\mathbf{e}, \mathbf{V}] \leq M\}}] \right| \\
\leq & 2P \left(\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] < \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d \right) + 2P(E[Q_n(\mathbf{X})^2|\mathbf{e}, \mathbf{V}] > M) \\
& + \left| E[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d, E[Q_n(\mathbf{X})^2|\mathbf{e}, \mathbf{V}] \leq M\}}] \right|.
\end{aligned}$$

We have, using the Markov inequality,

$$P(E[Q_n(\mathbf{X})^2|\mathbf{e}, \mathbf{V}] > M) \leq \frac{1}{M} E[E[Q_n(\mathbf{X})^2|\mathbf{e}, \mathbf{V}]] = \frac{1}{M}.$$

On the other hand,

$$\begin{aligned}
& \left| E \left[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d, E[Q_n(\mathbf{X})^2|\mathbf{e}, \mathbf{V}] \leq M\}} \right] \right| \\
\leq & E_{\mathbf{e}, \mathbf{V}} \left[\left| E_{\mathbf{U}}[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X}))] \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d, E[Q_n(\mathbf{X})^2|\mathbf{e}, \mathbf{V}] \leq M\}} \right| \right].
\end{aligned}$$

Step 3. In this step, we shall introduce the framework we are going to use for the rest of the proof. We refer the reader to [1] for the details and missing proofs. Fix an integer m and let μ denote the distribution of the random vector (X_1, \dots, X_m) , with X_1, \dots, X_m independent copies of $U \sim \mathcal{U}_{[-1,1]}$, There exists a reversible Markov process on \mathbb{R}^m , with semigroup P_t , equilibrium measure μ and generator \mathcal{L} given by

$$\mathcal{L}f(x) = \sum_{i=1}^m \left((1 - x_i^2) \partial_{ii} f - 2x_i \partial_i f \right), \quad x \in \mathbb{R}^m. \quad (3.17)$$

The operator \mathcal{L} is selfadjoint and negative semidefinite. We define the carré du champ operator Γ as

$$\Gamma(f, g)(x) = \frac{1}{2} (\mathcal{L}(fg)(x) - f(x)\mathcal{L}g(x) - g(x)\mathcal{L}f(x)) = \sum_{i=1}^m (1 - x_i^2) \partial_i f(x) \partial_i g(x). \quad (3.18)$$

When $f = g$ we simply write $\Gamma(f)$ instead of $\Gamma(f, f)$. An important property satisfied by Γ is that it is diffusive in the following sense:

$$\Gamma(\phi(f), g) = \phi'(f) \Gamma(f, g). \quad (3.19)$$

Besides, the eigenvalues of $-\mathcal{L}$ are given by

$$\text{Sp}(-\mathcal{L}) = \{i_1(i_1 - 1) + \cdots + i_m(i_m - 1) \mid i_1, \dots, i_m \in \mathbb{N}\}.$$

It may be ordered as a countable sequence like $0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots$, with a corresponding sequence of orthonormal eigenfunctions u_0, u_1, u_2, \dots where $u_0 = 1$; in addition, this sequence of eigenfunctions forms a complete orthogonal basis of $L^2(\mu)$. Also, note that the first nonzero element of $\text{Sp}(-\mathcal{L})$ is $\lambda_1 = 1 > 0$. Also, one can compute that, when $\lambda \in \text{Sp}(-\mathcal{L})$, then $\text{Ker}(\mathcal{L} + \lambda I)$ is composed of those polynomial functions $R(x_1, \dots, x_m)$ having the form

$$R(x_1, \dots, x_m) = \sum_{i_1(i_1+1)+\cdots+i_m(i_m+1)=\lambda} \alpha(i_1, \dots, i_m) J_{i_1}(x_1) \cdots J_{i_m}(x_m).$$

Here $J_i(X)$ is the i th Jacobi polynomial, defined as

$$J_i(x) = \frac{(-1)^i}{2^i i!} \frac{d^i}{dx^i} \{(1-x^2)^i\}, \quad x \in \mathbb{R}.$$

To end up with this quick summary, we recall the following Poincaré inequality, that is immediate to prove by using the previous facts together with the decomposition $L^2(\mu) = \bigoplus_{\lambda \in \text{Sp}(-\mathcal{L})} \text{Ker}(\mathcal{L} + \lambda I)$:

$$\text{Var}_\mu(f) \leq \int \Gamma(f) d\mu. \quad (3.20)$$

Step 4. We shall prove the existence of a constant $\kappa > 0$, depending on p, α and d but not on n , such that, for any $\delta > 0$,

$$\sup_{n \geq 1} E_{\mathbf{U}} \left[\frac{\delta}{\Gamma(Q_n)(\mathbf{X}) + \delta} \right] \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d\}} \leq \kappa \delta^{\frac{1}{2d+1}}. \quad (3.21)$$

The proof of (3.21) will rely on the Poincaré inequality (3.20) which, here, takes the following form:

$$\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] = \text{Var}_{\mathbf{U}}[Q_n(\mathbf{X})] \leq E_{\mathbf{U}}[\Gamma(Q_n)(\mathbf{X})]. \quad (3.22)$$

Another ingredient is the Carbery-Wright inequality, that we recall for sake of completeness.

Theorem 3.1 (see [8, Theorem 8]) *There exists an absolute constant $c > 0$ such that, if $Q : \mathbb{R}^m \rightarrow \mathbb{R}$ is a polynomial of degree at most k and μ is a log-concave probability measure on \mathbb{R}^m , then, for all $\alpha > 0$,*

$$\left(\int Q^2 d\mu \right)^{\frac{1}{2k}} \times \mu\{x \in \mathbb{R}^m : |Q(x)| \leq \alpha\} \leq c k \alpha^{\frac{1}{k}}. \quad (3.23)$$

Observe that the density of \mathbf{U} is log-concave, as an indicator function of a convex set. Let us now proceed with the proof of (3.21). For any strictly positive u , and provided $\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d$, one has

$$E \left[\frac{\delta}{\Gamma(Q_n)(\mathbf{X}) + \delta} \right] \leq \frac{\delta}{u} + P(\Gamma(Q_n)(\mathbf{X}) \leq u) \leq \frac{\delta}{u} + c u^{\frac{1}{2d}}, \quad (3.24)$$

where $c > 0$ denotes a constant only depending on d , α and p and where the last inequality follows from the Carbery-Wright inequality (3.23), the inequality (3.22) and the fact that $\Gamma(Q_n)$ is a polynomial of order $2d$, see (3.18). Finally, choosing $u = \delta^{\frac{2d}{2d+1}}$ in (3.24) leads to the desired conclusion (3.21).

Step 5. We shall prove that

$$\sup_{n \geq 1} (E_{\mathbf{U}}[\Gamma(\Gamma(Q_n))(\mathbf{X})] + E_{\mathbf{U}}|(\mathcal{L}Q_n)(\mathbf{X})|) \mathbf{1}_{\{E[Q_n(\mathbf{X})^2 | \mathbf{e}, \mathbf{V}] \leq M\}} \leq c(M) \quad (3.25)$$

where $c(M)$ is a constant only depending on M (whose value may change from one line to another within this step). First, relying on the results of Step 3 we have that, for any n ,

$$Q_n \in \bigoplus_{\alpha \leq \lambda_{2d}} \text{Ker}(\mathcal{L} + \alpha I).$$

Since \mathcal{L} is a bounded operator on the space $\bigoplus_{\alpha \leq \lambda_{2d}} \text{Ker}(\mathcal{L} + \alpha I)$, we deduce immediately that $\sup_{n \geq 1} E_{\mathbf{U}}[(\mathcal{L}Q_n)(\mathbf{X})^2] \mathbf{1}_{\{E_{\mathbf{U}}[Q_n^2(\mathbf{X})] \leq M\}} \leq c(M)$. Besides, one has $\Gamma = \frac{1}{2}(\mathcal{L} + 2\lambda I)$ on $\text{Ker}(\mathcal{L} + \lambda I)$ and one deduces for the same reason as above that

$$\sup_{n \geq 1} E_{\mathbf{U}}[\Gamma(\Gamma(Q_n))(\mathbf{X})] \mathbf{1}_{\{E_{\mathbf{U}}[Q_n^2(\mathbf{X})] \leq M\}} \leq c(M).$$

The proof of (3.25) is complete.

Step 6. We shall prove that, for any $n \geq 1$, any $0 < \alpha \leq 1$, any $\delta > 0$ and any $M > 0$,

$$|E_{\mathbf{U}}[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X}))]| \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X}) | \mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d, E[Q_n(\mathbf{X})^2 | \mathbf{e}, \mathbf{V}] \leq M\}} \leq 2\kappa \delta^{\frac{1}{2d+1}} + \sqrt{\frac{2}{\pi}} \frac{\alpha}{\delta} c(M). \quad (3.26)$$

Using Step 4, one has

$$\begin{aligned} & \left| E_{\mathbf{U}} \left[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \frac{\Gamma(Q_n)(\mathbf{X}) + \delta}{\Gamma(Q_n)(\mathbf{X}) + \delta} \right] \right| \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X}) | \mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d, E[Q_n(\mathbf{X})^2 | \mathbf{e}, \mathbf{V}] \leq M\}} \\ & \leq 2 E_{\mathbf{U}} \left[\frac{\delta}{|\Gamma(Q_n)(\mathbf{X}) + \delta|} \right] \mathbf{1}_{\{\text{Var}[Q_n(\mathbf{X}) | \mathbf{e}, \mathbf{V}] \geq \frac{d!}{2} \left(\frac{\alpha^2 p}{3}\right)^d\}} \\ & \quad + \left| E_{\mathbf{U}} \left[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \frac{\Gamma(Q_n)(\mathbf{X})}{\Gamma(Q_n)(\mathbf{X}) + \delta} \right] \right| \mathbf{1}_{\{E[Q_n(\mathbf{X})^2 | \mathbf{e}, \mathbf{V}] \leq M\}} \\ & \leq 2\kappa \delta^{\frac{1}{2d+1}} + \left| E_{\mathbf{U}} \left[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \frac{\Gamma(Q_n)(\mathbf{X})}{\Gamma(Q_n)(\mathbf{X}) + \delta} \right] \right| \mathbf{1}_{\{E[Q_n(\mathbf{X})^2 | \mathbf{e}, \mathbf{V}] \leq M\}}. \quad (3.27) \end{aligned}$$

Now, set $\Psi(x) = \int_{-\infty}^x \phi(s)ds$ and let us apply (3.19). We obtain

$$\begin{aligned}
& \left| E_{\mathbf{U}} \left[(\phi - \phi \star p_\alpha)(Q_n(\mathbf{X})) \frac{\Gamma(Q_n)(\mathbf{X})}{\Gamma(Q_n)(\mathbf{X}) + \delta} \right] \right| \\
&= \left| E_{\mathbf{U}} \left[\frac{1}{\Gamma(Q_n)(\mathbf{X}) + \delta} \Gamma((\Psi - \Psi \star p_\alpha) \circ Q_n, Q_n)(\mathbf{X}) \right] \right| \\
&= \left| E_{\mathbf{U}} \left[((\Psi - \Psi \star p_\alpha) \circ Q_n)(\mathbf{X}) \left(\Gamma(Q_n, \frac{1}{\Gamma(Q_n) + \delta})(\mathbf{X}) + \frac{(\mathcal{L}Q_n)(\mathbf{X})}{\Gamma(Q_n)(\mathbf{X}) + \delta} \right) \right] \right| \\
&= \left| E_{\mathbf{U}} \left[((\Psi - \Psi \star p_\alpha) \circ Q_n)(\mathbf{X}) \left(-\frac{\Gamma(Q_n, \Gamma(Q_n))(\mathbf{X})}{(\Gamma(Q_n)(\mathbf{X}) + \delta)^2} + \frac{(\mathcal{L}Q_n)(\mathbf{X})}{\Gamma(Q_n)(\mathbf{X}) + \delta} \right) \right] \right| \\
&\leq \frac{1}{\delta} E_{\mathbf{U}} \{ |((\Psi - \Psi \star p_\alpha) \circ Q_n)(\mathbf{X})| \times [\Gamma(\Gamma(Q_n))(\mathbf{X}) + |(\mathcal{L}Q_n)(\mathbf{X})|] \}. \tag{3.28}
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
|\Psi(x) - \Psi \star p_\alpha(x)| &= \left| \int_{\mathbb{R}} p_\alpha(y) \left(\int_{-\infty}^x (\phi(u) - \phi(u-y)) du \right) dy \right| \\
&\leq \int_{\mathbb{R}} p_\alpha(y) \left| \int_{-\infty}^x \phi(u) du - \int_{-\infty}^x \phi(u-y) du \right| dy \\
&\leq \int_{\mathbb{R}} p_\alpha(y) \left| \int_{x-y}^x \phi(u) du \right| dy \leq \int_{\mathbb{R}} p_\alpha(y) |y| dy \leq \sqrt{\frac{2}{\pi}} \alpha. \tag{3.29}
\end{aligned}$$

The desired conclusion (3.26) now follows easily from (3.25), (3.27), (3.28) and (3.29).

Step 7: Concluding the proof. Combining the results of all the previous steps, we obtain, for any $n \geq 1$, any $0 < \alpha \leq 1$, any $\delta > 0$ and any $M > 0$,

$$\begin{aligned}
& \sup_{\phi \in C_c^\infty: \|\phi\|_\infty \leq 1} |E[\phi(Q_n(\mathbf{X}))] - E[\phi(Q_n(\mathbf{Y}))]| \\
&\leq \frac{1}{\alpha} d_{FM}(Q_n(\mathbf{X}), Q_n(\mathbf{Y})) + \frac{4}{M} + 2P \left(\text{Var}[Q_n(\mathbf{X})|\mathbf{e}, \mathbf{V}] < \frac{d!}{2} \left(\frac{\alpha^2 p}{3} \right)^d \right) \tag{3.30}
\end{aligned}$$

$$+ 2P \left(\text{Var}[Q_n(\mathbf{Y})|\mathbf{e}, \mathbf{V}] < \frac{d!}{2} \left(\frac{\alpha^2 p}{3} \right)^d \right) + 4\kappa \delta^{\frac{1}{2d+1}} + 2\sqrt{\frac{2}{\pi}} \frac{\alpha}{\delta} c(M). \tag{3.31}$$

In (3.30)-(3.31), take the limit $n \rightarrow \infty$. Due to (3.15) on one hand and Theorem 1.3 on the other hand (plus the fact that the Fortet-Mourier distance d_{FM} metrizes the convergence in distribution), one obtains

$$\limsup_{n \rightarrow \infty} \sup_{\phi \in C_c^\infty: \|\phi\|_\infty \leq 1} |E[\phi(Q_n(\mathbf{X}))] - E[\phi(Q_n(\mathbf{Y}))]| \leq \frac{4}{M} + 4\kappa \delta^{\frac{1}{2d+1}} + 2\sqrt{\frac{2}{\pi}} \frac{\alpha}{\delta} c(M).$$

The desired conclusion (1.9) then follows by letting (in this order) $\alpha \rightarrow 0$, $\delta \rightarrow 0$ and $M \rightarrow \infty$. ■

References

- [1] D. Bakry, I. Gentil and M. Ledoux (2013). *Analysis and Geometry of Markov Diffusion Semigroups*. Forthcoming monograph.
- [2] A.R. Barron (1986). Entropy and the central limit theorem. *Ann. Probab.* **14**, no. 1, pp. 336-342.
- [3] S. G. Bobkov, G. P. Chistyakov and F. Götze (2013): Berry-Esseen bounds in the entropic central limit theorem. *Probab. Theory Related Fields*, to appear
- [4] S. G. Bobkov, G. P. Chistyakov and F. Götze (2013): Fisher information and the central limit theorem. *Probab. Theory Related Fields*, to appear
- [5] S. G. Bobkov, G. P. Chistyakov and F. Götze (2013): Rate of convergence and Edgeworth-type expansion in the entropic central limit theorem. *Ann. Probab.* **41**, no. 4, pp. 2479-2512.
- [6] T. Cacoullos, V. Papathanasiou, and S. A. Utev. (1994): Variational inequalities with examples and an application to the central limit theorem. *Ann. Probab.* **22**, no. 3, pp. 1607-1618.
- [7] T. Cacoullos, V. Papathanasiou, and S. A. Utev. (1993): Another characterization of the normal law and a proof of the central limit theorem connected with it. *Theory of Probability and Its Applications* **37**, no. 4, pp. 581-588.
- [8] A. Carbery and J. Wright (2001): Distributional and L^q norm inequalities for polynomials over convex bodies in \mathbb{R}^n . *Math. Research Lett.* **8**, pp. 233-248.
- [9] B.V. Gnedenko and A.N. Kolmogorov (1968): *Limit distributions for sums of independent random variables*. Revised edition Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills., Ont. 1968 ix+293 pp. 60.30
- [10] O.T. Johnson and A.R. Barron (2004): Fisher information inequalities and the central limit theorem. *Probab. Theory Related Fields* **129**, no.3, pp. 391-409.
- [11] M. Mamatov and M.K. Halikov (1964): Global limit theorems for distribution functions in the higher-dimensional case. *Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk* no. 1, 13-21 (in Russian).
- [12] I. Nourdin and G. Peccati (2012). *Normal Approximations with Malliavin Calculus. From Stein's Method to Universality*. Cambridge University Press.
- [13] S.H. Sirazdinov and M. Mamatov (1962): On mean convergence for densities. *Teor. Veroyatnost. i Primenen.* **7**, pp. 433-437 (in Russian).
- [14] E. Mossel, R. O'Donnell and K. Oleszkiewicz (2010). Noise stability of functions with low influences: Variance and optimality. *Ann. Math.* **171**, pp. 295-341.
- [15] I. Nourdin and G. Poly (2013). Convergence in total variation on Wiener chaos. *Stoch. Proc. Appl.* **123**, pp. 651-674.

- [16] Y. Prohorov (1952): On a local limit theorem for densities. *Doklady Akad. Nauk SSSR (N.S.)* **83**, pp. 797-800. In Russian.
- [17] R. Shimizu (1975): On Fisher's amount of information for location family. In *G.P.Patil et al (ed.), Statistical Distributions in Scientific Work* **3**, pp. 305-312. Reidel.