



HAL
open science

Convergence in law implies convergence in total variation for polynomials in independent Gaussian, Gamma or Beta random variables

Ivan Nourdin, Guillaume Poly

► **To cite this version:**

Ivan Nourdin, Guillaume Poly. Convergence in law implies convergence in total variation for polynomials in independent Gaussian, Gamma or Beta random variables. 2013. hal-00821911

HAL Id: hal-00821911

<https://hal.science/hal-00821911>

Preprint submitted on 13 May 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.

Convergence in law implies convergence in total variation for polynomials in independent Gaussian, Gamma or Beta random variables

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

Abstract: Consider a sequence of polynomials of bounded degree evaluated in independent Gaussian, Gamma or Beta random variables. We show that, if this sequence converges in law to a nonconstant distribution, then (i) the limit distribution is necessarily absolutely continuous with respect to the Lebesgue measure and (ii) the convergence automatically takes place in the total variation topology. Our proof, which relies on the Carbery-Wright inequality and makes use of a diffusive Markov operator approach, extends the results of [9] to the Gamma and Beta cases.

Keywords: Convergence in law; convergence in total variation; absolute continuity; Carbery-Wright inequality; log-concave distribution; orthogonal polynomials.

2000 Mathematics Subject Classification:

1 Introduction and main results

The Fortet-Mourier distance between (laws of) random variables, defined as

$$d_{FM}(F, G) = \sup_{\substack{\|h\|_\infty \leq 1 \\ \|h'\|_\infty \leq 1}} |E[h(F)] - E[h(G)]| \quad (1.1)$$

is well-known to metrize the convergence in law, see, e.g., [5, Theorem 11.3.3]. In other words, as $n \rightarrow \infty$ one has that $F_n \xrightarrow{\text{law}} F_\infty$ if and only if $d_{FM}(F_n, F_\infty) \rightarrow 0$. But there is plenty of other distances that allows one to measure the proximity between laws of random variables. For instance, one may use the Kolmogorov distance:

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

Of course, if $d_{Kol}(F_n, F_\infty) \rightarrow 0$ then $F_n \xrightarrow{\text{law}} F_\infty$. But the converse implication is wrong in general, meaning that the Kolmogorov distance does *not* metrize the convergence in law. Nevertheless, it becomes true when the target law is *continuous* (that is, when the law of F_∞ has a density with respect to the Lebesgue measure), a fact which can be easily checked by using (for instance) the second Dini's theorem. Yet another popular distance for measuring the distance between laws of random variables, which is even stronger than the Kolmogorov distance, is the total variation distance:

$$d_{TV}(F, G) = \sup_{A \in \mathcal{B}(\mathbb{R})} |P(F \in A) - P(G \in A)|. \quad (1.2)$$

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

One may prove that

$$d_{TV}(F, G) = \frac{1}{2} \sup_{\|h\|_\infty \leq 1} |E[h(F)] - E[h(G)]|, \quad (1.3)$$

or, whenever F and G both have a density (noted f and g respectively)

$$d_{TV}(F, G) = \frac{1}{2} \int_{\mathbb{R}} |f(x) - g(x)| dx.$$

Unlike the Fortet-Mourier or Kolmogorov distances, it can happen that $F_n \xrightarrow{\text{law}} F_\infty$ for continuous F_n and F_∞ without having that $d_{TV}(F_n, F_\infty) \rightarrow 0$. For an explicit counterexample, one may consider $F_n \sim \frac{2}{\pi} \cos^2(nx) \mathbf{1}_{[0, \pi]}(x) dx$; indeed, it is straightforward to check that $F_n \xrightarrow{\text{law}} F_\infty \sim \mathcal{U}_{[0, \pi]}$ but $d_{TV}(F_n, F_\infty) \not\rightarrow 0$ (it is indeed a strictly positive quantity that does not depend on n).

As we just saw, the convergence in total variation is very strong and therefore it cannot be expected from the mere convergence in law without additional structure. Let us give three representative results in this direction. Firstly, there is a celebrated theorem of Ibragimov (see, e.g., Reiss [10]) according to which, if F_n, F_∞ are continuous random variables with densities f_n, f_∞ that that are *unimodal*, then $F_n \xrightarrow{\text{law}} F_\infty$ if and only if $d_{TV}(F_n, F_\infty) \rightarrow 0$. Secondly, let us quote the paper [11], in which necessary and sufficient conditions are given (in term of the absolute continuity of the laws) so that the classical Central Limit Theorem holds in total variation. Finally, let us mention [1] or [6] for conditions ensuring the convergence in total variation for random variables in Sobolev or Dirichlet spaces. Although all the above examples are related to very different frameworks, they have in common the use of a particular structure of the involved variables; loosely speaking, this structure allows to derive a kind of “non-degeneracy” in an appropriate sense which, in turn, enables to reinforce the convergence, from the Fortet-Mourier distance to the total variation one.

Our goal in this short note is to exhibit yet another instance when convergence in law and in total variation are equivalent. More precisely, we shall prove the following result, which may be seen as an extension to the Gamma and Beta cases of our previous results in [9].

Theorem 1.1 *Assume that one of the following three conditions is satisfied:*

- (1) $X \sim N(0, 1)$;
- (2) $X \sim \Gamma(r, 1)$ with $r \geq 1$;
- (3) $X \sim \beta(a, b)$ with $a, b \geq 1$.

Let X_1, X_2, \dots be independent copies of X . Fix an integer $d \geq 1$ and, for each n , let m_n be a positive integer and let $Q_n \in \mathbb{R}[x_1, \dots, x_{m_n}]$ be a multilinear polynomial of degree at most d ; assume further that $m_n \rightarrow \infty$ as $n \rightarrow \infty$. Finally, suppose that F_n has the form

$$F_n = Q_n(X_1, \dots, X_{m_n}), \quad n \geq 1,$$

and that it converges in law as $n \rightarrow \infty$ to a non-constant random variable F_∞ . Then the law of F_∞ is absolutely continuous with respect to the Lebesgue measure and F_n actually converges to F_∞ in total variation.

In the statement of Theorem 1.1, by ‘multilinear polynomial of degree at most d ’ we mean a polynomial $Q \in \mathbb{R}[x_1, \dots, x_m]$ of the form

$$Q(x_1, \dots, x_m) = \sum_{S \subset \{1, \dots, m\}, |S| \leq d} a_S \prod_{i \in S} x_i,$$

for some real coefficients a_S and with the usual convention that $\prod_{i \in \emptyset} x_i = 1$.

Before providing the proof of Theorem 1.1, let us comment a little bit why we are ‘only’ considering the three cases (1), (2) and (3). This is actually due to our method of proof. Indeed, the two main ingredients we are using for showing Theorem 1.1 are the following.

- (a) We will make use of a Markov semigroup approach. More specifically, our strategy rely to the use of orthogonal polynomials which are also eigenvectors of diffusion operators. And in dimension 1, up to affine transformations only the Hermite (case (1)), Laguerre (case (2)) and Jacobi (case (3)) polynomials are of this form, see [7].
- (b) We will make use of the Carbery-Wright inequality (Theorem 2.1). The main assumption for this inequality to hold is the log-concavity property. This impose some further (weak) restrictions on the parameters in the cases (2) and (3).

The rest of the paper is organized as follows. In Section 2, we gather some useful preliminary results. Theorem 1.1 is shown in Section 3.

2 Preliminaries

From now on, we shall write m instead of m_n for the sake of simplicity.

2.1 Markov semigroup

In this section, we introduce the framework we shall need along the proof of Theorem 1.1. We refer the reader to [2] 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 X , for X satisfying either (1), (2) or (3). In these three cases, there exists a reversible Markov process on \mathbb{R}^m , with semigroup P_t , equilibrium measure μ and generator \mathcal{L} . The operator \mathcal{L} is selfadjoint and negative semidefinite. We define the Dirichlet form \mathcal{E} associated to \mathcal{L} :

$$\mathcal{E}(f, g) = - \int f \mathcal{L}g d\mu = - \int g \mathcal{L}f d\mu.$$

When $f = g$, we simply write $\mathcal{E}(f)$ instead of $\mathcal{E}(f, f)$. The carré du champ operator Γ will be also of interest; it is the operator defined as

$$\Gamma(f, g) = \frac{1}{2}(\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f).$$

Similarly to \mathcal{E} , when $f = g$ we simply write $\Gamma(f)$ instead of $\Gamma(f, f)$. Since $\int \mathcal{L}f d\mu = 0$, we observe the following link between the Dirichlet form \mathcal{E} and the carré du champ operator Γ :

$$\int \Gamma(f, g) d\mu = \mathcal{E}(f, g).$$

An important property which is satisfied in the three cases (1), (2) and (3) is that Γ is diffusive in the following sense:

$$\Gamma(\phi(f), g) = \phi'(f)\Gamma(f, g). \quad (2.4)$$

Besides, and it is another important property shared by (1), (2), (3), the eigenvalues of $-\mathcal{L}$ may be ordered as a countable sequence like $0 = \lambda_0 < \lambda_1 < \lambda_2 < \dots$, 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 $L^2(\mu)$. For completeness, let us give more details in each of our three cases (1), (2), (3) of interest.

(1) The case where $X \sim N(0, 1)$. We have

$$\mathcal{L}f(x) = \Delta f(x) - x \cdot \nabla f(x), \quad x \in \mathbb{R}^m, \quad (2.5)$$

where Δ is the Laplacian operator and ∇ is the gradient. As a result,

$$\Gamma(f, g) = \nabla f \cdot \nabla g. \quad (2.6)$$

We can compute that $\text{Sp}(-\mathcal{L}) = \mathbb{N}$ and that $\text{Ker}(\mathcal{L} + kI)$ (with I the identity operator) is composed of those polynomial functions $R(x_1, \dots, x_m)$ having the form

$$R(x_1, \dots, x_m) = \sum_{i_1+i_2+\dots+i_m=k} \alpha(i_1, \dots, i_m) \prod_{j=1}^m H_{i_j}(x_j).$$

Here, H_i stands for the Hermite polynomial of degree i .

(2) The case where $X \sim \Gamma(r, 1)$. The density of X is $f_X(t) = t^{r-1} \frac{e^{-t}}{\Gamma(r)}$, $t \geq 0$, with Γ the Euler Gamma function; it is log-concave for $r \geq 1$. Besides, we have

$$\mathcal{L}f(x) = \sum_{i=1}^m \left(x_i \partial_{ii} f + (r+1-x_i) \partial_i f \right), \quad x \in \mathbb{R}^m. \quad (2.7)$$

As a result,

$$\Gamma(f, g)(x) = \sum_{i=1}^m x_i \partial_i f(x) \partial_i g(x), \quad x \in \mathbb{R}^m. \quad (2.8)$$

We can compute that $\text{Sp}(-\mathcal{L}) = \mathbb{N}$ and that $\text{Ker}(\mathcal{L} + kI)$ is composed of those polynomial functions $R(x_1, \dots, x_m)$ having the form

$$R(x_1, \dots, x_m) = \sum_{i_1+i_2+\dots+i_m=k} \alpha(i_1, \dots, i_m) \prod_{j=1}^m L_{i_j}(x_j).$$

Here $L_i(X)$ stands for the i th Laguerre polynomial of parameter r defined as

$$L_i(x) = \frac{x^{-r} e^x}{i!} \frac{d^i}{dx^i} \{ e^{-x} x^{i+r} \}, \quad x \in \mathbb{R}.$$

(3) The case where $X \sim \beta(a, b)$. In this case, X is continuous with density

$$f_X(t) = \begin{cases} \frac{t^{a-1}(1-t)^{b-1}}{\int_0^1 u^{a-1}(1-u)^{b-1} du} & \text{if } t \in [0, 1] \\ 0 & \text{otherwise} \end{cases}.$$

The density f_X is log-concave when $a, b \geq 1$. Moreover, we have

$$\mathcal{L}f(x) = \sum_{i=1}^m \left((1-x_i^2)\partial_{ii}f + (b-a-(b+a)x_i)\partial_i f \right), \quad x \in \mathbb{R}^m. \quad (2.9)$$

As a result,

$$\Gamma(f, g)(x) = \sum_{i=1}^m (1-x_i^2)\partial_i f(x)\partial_i g(x), \quad x \in \mathbb{R}^m. \quad (2.10)$$

Here, the structure of the spectrum turns out to be a little bit more complicated than in the two previous cases (1) and (2). Indeed, we have that

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

Note in particular that the first nonzero element of $\text{Sp}(-\mathcal{L})$ is $\lambda_1 = a + b - 1 > 0$. 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

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

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

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

To end up with this quick summary, we stress that a Poincaré inequality holds true in the three cases (1), (2) and (3). This is well-known and easy to prove, by using the previous facts together with the decomposition

$$L^2(\mu) = \bigoplus_{\lambda \in \text{Sp}(-\mathcal{L})} \text{Ker}(\mathcal{L} + \lambda I).$$

Namely, with $\lambda_1 > 0$ the first nonzero eigenvalue of $-\mathcal{L}$, we have

$$\text{Var}_\mu(f) \leq \frac{1}{\lambda_1} \mathcal{E}(f). \quad (2.11)$$

2.2 Carbery-Wright inequality

The proof of Theorem 1.1 will rely, among others, on the following crucial inequality due to Carbery and Wright ([4, Theorem 8]). We state it here for convenience.

Theorem 2.1 (Carbery-Wright) *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 $\alpha > 0$,*

$$\left(\int Q^2 d\mu \right)^{\frac{1}{2k}} \times \mu\{x \in \mathbb{R}^m : |Q(x)| \leq \alpha\} \leq ck \alpha^{\frac{1}{k}}. \quad (2.12)$$

2.3 Absolute continuity

There is a celebrated result of Borell [3] according to which, if X_1, X_2, \dots are independent, identically distributed and X_1 has an absolute continuous law, then any nonconstant polynomial in the X_i 's has an absolute continuous law, too. In the particular case where the common law satisfies either (1), or (2) or (3) in Theorem 1.1, one can recover Borell's theorem as a consequence of the previous Carbery-Wright inequality. We provide the proof of this fact here, since it may be seen as a first step towards the proof of Theorem 1.1.

Proposition 2.2 *Assume that one of the three conditions (1), (2) or (3) of Theorem 1.1 is satisfied. Let X_1, X_2, \dots be independent copies of X . Consider two integers $m, d \geq 1$ and let $Q \in \mathbb{R}[x_1, \dots, x_m]$ be a polynomial of degree d . Then the law of $Q(X_1, \dots, X_m)$ is absolutely continuous with respect to the Lebesgue measure if and only if its variance is not zero.*

Proof. Write μ for the distribution of (X_1, \dots, X_m) and assume that $\text{Var}(Q(X_1, \dots, X_m)) > 0$. We shall prove that, if A is a Borel set of \mathbb{R} with Lebesgue measure zero, then $P(Q(X_1, \dots, X_m) \in A) = 0$. This will be done in three steps.

Step 1. Let $\varepsilon > 0$ and let B be a bounded Borel set. We shall prove that

$$\int \mathbf{1}_{\{Q \in B\}} \frac{\Gamma(Q)}{\varepsilon + \Gamma(Q)} d\mu = \int \left(\int_{-\infty}^Q \mathbf{1}_B(u) du \times \left\{ \frac{-\mathcal{L}Q}{\Gamma(Q) + \varepsilon} + 2 \frac{\Gamma(\Gamma(Q))}{(\Gamma(Q) + \varepsilon)^2} \right\} \right) d\mu. \quad (2.13)$$

Indeed, let $h : \mathbb{R} \rightarrow [0, 1]$ be C^∞ with compact support. We can write, using among other (2.4),

$$\begin{aligned} & \int \left(\int_{-\infty}^Q h(u) du \times \frac{-\mathcal{L}Q}{\Gamma(Q) + \varepsilon} \right) d\mu = \mathcal{E} \left(\int_{-\infty}^Q h(u) du \times \frac{1}{\Gamma(Q) + \varepsilon}, Q \right) \\ & = \int \left(h(Q) \frac{\Gamma(Q)}{\Gamma(Q) + \varepsilon} - 2 \int_{-\infty}^Q h(u) du \frac{\Gamma(\Gamma(Q))}{(\Gamma(Q) + \varepsilon)^2} \right) d\mu. \end{aligned}$$

Applying Lusin's theorem allows one, by dominated convergence, to pass from h to $\mathbf{1}_B$ in the previous identity; this leads to the desired conclusion (2.13).

Step 2. Let us apply (2.13) to $B = A \cap [-n, n]$. Since $\int_{-\infty}^{\cdot} \mathbf{1}_B(u) du$ is zero almost everywhere, one deduces that, for all $\varepsilon > 0$ and all $n \in \mathbb{N}^*$,

$$\int \mathbf{1}_{\{Q \in A \cap [-n, n]\}} \frac{\Gamma(Q)}{\varepsilon + \Gamma(Q)} d\mu = 0.$$

By monotone convergence ($n \rightarrow \infty$) it comes that, for all $\varepsilon > 0$,

$$\int \mathbf{1}_{\{Q \in A\}} \frac{\Gamma(Q)}{\varepsilon + \Gamma(Q)} d\mu = 0. \quad (2.14)$$

Step 3. Observe that $\Gamma(Q)$ is a polynomial of degree at most $2d$, see indeed (2.6), (2.8) or (2.10). We deduce from the Carbery-Wright inequality (2.12), together with the Poincaré inequality (2.11), that $\Gamma(Q)$ is strictly positive almost everywhere. Thus, by dominated convergence ($\varepsilon \rightarrow 0$) in (2.14) we finally get that $\mu\{Q \in A\} = P(Q(X_1, \dots, X_m) \in A) = 0$. ■

3 Proof of Theorem 1.1

We are now in a position to show Theorem 1.1. We will split its proof in several steps.

Step 1. For any $p \in [1, \infty)$ we shall prove that

$$\sup_n \int |Q_n|^p d\mu_m < \infty. \quad (3.15)$$

(Let us recall our convention about m from the beginning of Section 2.) Indeed, using (for instance) Propositions 3.11, 3.12 and 3.16 of [8] (namely, a hypercontractivity property), one first observes that, for any $p \in [2, \infty)$, there exists a constant $c_p > 0$ such that, for all n ,

$$\int |Q_n|^p d\mu_m \leq c_p \left(\int Q_n^2 d\mu_m \right)^{p/2}. \quad (3.16)$$

(This is for obtaining (3.16) that we need Q_n to be *multilinear*.) On the other hand, one can write

$$\begin{aligned} \int Q_n^2 d\mu_m &= \int Q_n^2 \mathbf{1}_{\{Q_n^2 \geq \frac{1}{2} \int Q_n^2 d\mu_m\}} d\mu_m + \int Q_n^2 \mathbf{1}_{\{Q_n^2 < \frac{1}{2} \int Q_n^2 d\mu_m\}} d\mu_m \\ &\leq \sqrt{\int Q_n^4 d\mu_m} \sqrt{\mu_m \left\{ x : Q_n(x)^2 \geq \frac{1}{2} \int Q_n^2 d\mu_m \right\}} + \frac{1}{2} \int Q_n^2 d\mu_m, \end{aligned}$$

so that, using (3.16) with $p = 4$,

$$\mu_m \left\{ x : Q_n(x)^2 \geq \frac{1}{2} \int Q_n^2 d\mu_m \right\} \geq \frac{(\int Q_n^2 d\mu_m)^2}{4 \int Q_n^4 d\mu_m} \geq \frac{1}{4c_4}.$$

But $\{Q_n\}_{n \geq 1}$ is tight as $\{F_n\}_{n \geq 1}$ converges in law. As a result, there exists $M > 0$ such that, for all n ,

$$\mu_m \left\{ x : Q_n(x)^2 \geq M \right\} < \frac{1}{4c_4}.$$

We deduce that $\int Q_n^2 d\mu_m \leq 2M$ which, together with (3.16), leads to the claim (3.15).

Step 2. We shall prove the existence of a constant $c > 0$ such that, for any $u > 0$ and any $n \in \mathbb{N}^*$,

$$\mu_m \{x : \Gamma(Q_n) \leq u\} \leq c \frac{u^{\frac{1}{2d}}}{\text{Var}_{\mu_m}(Q_n)^{\frac{1}{2d}}}. \quad (3.17)$$

Observe first that $\Gamma(Q_n)$ is a polynomial of degree at most $2d$, see indeed (2.6), (2.8) or (2.10). On the other hand, since X has a log-concave density, the probability μ_m is absolutely continuous with a log-concave density as well. As a consequence, Carbery-Wright inequality (2.12) applies and yields the existence of a constant $c > 0$ such that

$$\mu_m \{x : \Gamma(Q_n) \leq u\} \leq c u^{\frac{1}{2d}} \left(\int \Gamma(Q_n) d\mu_m \right)^{-\frac{1}{2d}}.$$

To get the claim (3.17), it remains one to apply the Poincaré inequality (2.11).

Step 3. We shall prove the existence of $n_0 \in \mathbb{N}^*$ and $\kappa > 0$ such that, for any $\varepsilon > 0$,

$$\sup_{n \geq n_0} \int \frac{\varepsilon}{\Gamma(Q_n) + \varepsilon} d\mu_m \leq \kappa \varepsilon^{\frac{1}{2d+1}}. \quad (3.18)$$

Indeed, thanks to the result shown in Step 2 one can write

$$\begin{aligned} \int \frac{\varepsilon}{\Gamma(Q_n) + \varepsilon} d\mu_m &\leq \frac{\varepsilon}{u} + \mu_m \{x : \Gamma(Q_n) \leq u\} \\ &\leq \frac{\varepsilon}{u} + c \frac{u^{\frac{1}{2d}}}{\text{Var}_{\mu_m}(Q_n)^{\frac{1}{2d}}}. \end{aligned}$$

But, by Step 1 and since $\mu_m \circ Q_n^{-1}$ converges to some probability measure η , one has that $\text{Var}_{\mu_m}(Q_n)$ converges to the variance of η as $n \rightarrow \infty$. Moreover this variance is strictly positive by assumption. We deduce the existence of $n_0 \in \mathbb{N}^*$ and $\delta > 0$ such that

$$\sup_{n \geq n_0} \int \frac{\varepsilon}{\Gamma(Q_n) + \varepsilon} d\mu_m \leq \frac{\varepsilon}{u} + \delta u^{\frac{1}{2d}}.$$

Choosing $u = \varepsilon^{\frac{2d}{2d+1}}$ leads to the desired conclusion (3.18).

Step 4. Let m' be shorthand for $m_{n'}$ and recall the Fortet-Mourier distance (1.1) as well as the total variation distance (1.3) from the Introduction. We shall prove that, for any $n, n' \geq n_0$ (with n_0 and κ given by Step 3), any $0 < \alpha \leq 1$ and any $\varepsilon > 0$,

$$d_{TV}(F_n, F_{n'}) \leq \frac{1}{\alpha} d_{FM}(F_n, F_{n'}) + 4\kappa \varepsilon^{\frac{1}{2d+1}} + 2\sqrt{\frac{2}{\pi}} \frac{\alpha}{\varepsilon} \sup_{n \geq n_0} \left(\int \Gamma(Q_n) d\mu_m + \int |\mathcal{L}Q_n| d\mu_m \right). \quad (3.19)$$

Indeed, 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 $g \in C_c^\infty$ be bounded by 1. It is immediately checked that

$$\|g * p_\alpha\|_\infty \leq 1 \leq \frac{1}{\alpha} \quad \text{and} \quad \|(g * p_\alpha)'\|_\infty \leq \frac{1}{\alpha}. \quad (3.20)$$

Let $n, n' \geq n_0$ be given integers. Using Step 3 and (3.20) we can write

$$\begin{aligned}
& \left| \int g d(\mu_m \circ Q_n^{-1}) - \int g d(\mu_{m'} \circ Q_{n'}^{-1}) \right| = \left| \int g \circ Q_n d\mu_m - \int g \circ Q_{n'} d\mu_{m'} \right| \\
& \leq \left| \int (g * p_\alpha) \circ Q_n d\mu_m - \int (g * p_\alpha) \circ Q_{n'} d\mu_{m'} \right| \\
& \quad + \left| \int (g - g * p_\alpha) \circ Q_n \times \left(\frac{\Gamma(Q_n)}{\Gamma(Q_n) + \varepsilon} + \frac{\varepsilon}{\Gamma(Q_n) + \varepsilon} \right) d\mu_m \right| \\
& \quad + \left| \int (g - g * p_\alpha) \circ Q_{n'} \times \left(\frac{\Gamma(Q_{n'})}{\Gamma(Q_{n'}) + \varepsilon} + \frac{\varepsilon}{\Gamma(Q_{n'}) + \varepsilon} \right) d\mu_{m'} \right| \\
& \leq \frac{1}{\alpha} d_{FM}(F_n, F_{n'}) + 2 \int \frac{\varepsilon}{\Gamma(Q_n) + \varepsilon} d\mu_m + 2 \int \frac{\varepsilon}{\Gamma(Q_{n'}) + \varepsilon} d\mu_{m'} \\
& \quad + \left| \int (g - g * p_\alpha) \circ Q_n \times \frac{\Gamma(Q_n)}{\Gamma(Q_n) + \varepsilon} d\mu_m \right| + \left| \int (g - g * p_\alpha) \circ Q_{n'} \times \frac{\Gamma(Q_{n'})}{\Gamma(Q_{n'}) + \varepsilon} d\mu_{m'} \right| \\
& \leq \frac{1}{\alpha} d_{FM}(F_n, F_{n'}) + 4\kappa \varepsilon^{\frac{1}{2d+1}} + 2 \sup_{n \geq n_0} \left| \int (g - g * p_\alpha) \circ Q_n \times \frac{\Gamma(Q_n)}{\Gamma(Q_n) + \varepsilon} d\mu_m \right|.
\end{aligned}$$

Now, set $\Psi(x) = \int_{-\infty}^x g(s) ds$ and let us apply (2.4). We obtain

$$\begin{aligned}
& \left| \int (g - g * p_\alpha) \circ Q_n \times \frac{\Gamma(Q_n)}{\Gamma(Q_n) + \varepsilon} d\mu_m \right| \\
& = \left| \int \frac{1}{\Gamma(Q_n) + \varepsilon} \Gamma((\Psi - \Psi * p_\alpha) \circ Q_n, Q_n) d\mu_m \right| \\
& = \left| \int (\Psi - \Psi * p_\alpha) \circ Q_n \times \left(\Gamma(Q_n, \frac{1}{\Gamma(Q_n) + \varepsilon}) + \frac{\mathcal{L}Q_n}{\Gamma(Q_n) + \varepsilon} \right) d\mu_m \right| \\
& = \left| \int (\Psi - \Psi * p_\alpha) \circ Q_n \times \left(-\frac{\Gamma(Q_n, \Gamma(Q_n))}{(\Gamma(Q_n) + \varepsilon)^2} + \frac{\mathcal{L}Q_n}{\Gamma(Q_n) + \varepsilon} \right) d\mu_m \right| \\
& \leq \frac{1}{\varepsilon} \int |(\Psi - \Psi * p_\alpha) \circ Q_n| \times (\Gamma(\Gamma(Q_n)) + |\mathcal{L}Q_n|) d\mu_{m_n}. \tag{3.21}
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
|\Psi(x) - \Psi * p_\alpha(x)| & = \left| \int_{\mathbb{R}} p_\alpha(y) \left(\int_{-\infty}^x (g(u) - g(u - y)) du \right) dy \right| \\
& \leq \int_{\mathbb{R}} p_\alpha(y) \left| \int_{-\infty}^x g(u) du - \int_{-\infty}^x g(u - y) du \right| dy \\
& \leq \int_{\mathbb{R}} p_\alpha(y) \left| \int_{x-y}^x g(u) du \right| dy \leq \int_{\mathbb{R}} p_\alpha(y) |y| dy \leq \sqrt{\frac{2}{\pi}} \alpha. \tag{3.22}
\end{aligned}$$

The desired conclusion (3.19) now follows easily.

Step 5. We shall prove that

$$\sup_{n \geq n_0} \left(\int \Gamma(\Gamma(Q_n)) d\mu_m + \int |\mathcal{L}Q_n| d\mu_m \right) < \infty. \tag{3.23}$$

First, relying on the results of Section 2.1 we have that

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

Since \mathcal{L} is a bounded operator on the space $\bigoplus_{\alpha \leq \lambda_{2d}} \text{Ker}(\mathcal{L} + \alpha I)$ and Q_n is bounded in the space $L^2(\mu_m)$, we deduce immediately that $\sup_n \int \mathcal{L}(Q_n)^2 d\mu_m < \infty$, implying in turn that

$$\sup_n \int |\mathcal{L}(Q_n)| d\mu_m < \infty.$$

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

$$\sup_n \int \Gamma(\Gamma(Q_n)) d\mu_m < \infty.$$

The proof of (3.23) is complete.

Step 6: conclusion. The Fortet-Mourier distance d_{FM} metrizing the convergence in distribution, our assumption ensures that $d_{FM}(F_n, F_{n'}) \rightarrow 0$ as $n, n' \rightarrow \infty$. Therefore, combining (3.23) with (3.19), letting $n, n' \rightarrow \infty$, then $\alpha \rightarrow 0$ and then $\varepsilon \rightarrow 0$, we conclude that $\lim_{n, n' \rightarrow \infty} d_{TV}(F_n, F_{n'}) = 0$, meaning that F_n is a Cauchy sequence in the total variation topology. But the space of bounded measures is complete for the total variation distance, so the distribution of F_n must converge to some distribution, say η , in the total variation distance. Of course, η must coincide with the law of F_∞ . Moreover, let A be a Borel set of Lebesgue measure zero. By Proposition 2.2, we have $P(F_n \in A) = 0$ when n is large enough. Since $d_{TV}(F_n, F_\infty) \rightarrow 0$ as $n \rightarrow \infty$, we deduce that $P(F_\infty \in A) = 0$ as well, thus proving that the law of F_∞ is absolutely continuous with respect to the Lebesgue measure by the Radon-Nikodym theorem. The proof of Theorem 1.1 is now complete. \blacksquare

References

- [1] D.E. Aleksandrova, V.I. Bogachev and A. Yu Pilipenko (1999). On the convergence of induced measures in variation. *Sbornik: Mathematics* **190**, no. 9, 1229-1245.
- [2] D. Bakry, I. Gentil and M. Ledoux (2013). *Analysis and Geometry of Markov Diffusion Semigroups*. Forthcoming monograph.
- [3] Ch. Borell (1988). Real Polynomial Chaos and Absolute Continuity. *Probab. Th. Rel. Fields* **77**, 397-400.
- [4] 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**, 233-248.
- [5] R.M. Dudley (2003). *Real Analysis and Probability* (2nd Edition). Cambridge University Press, Cambridge.

- [6] D. Malicet and G. Poly (2013). Properties of convergence in Dirichlet structures. *J. Funct. Anal.* **264**, 2077-2096.
- [7] O. Mazet (1997). Classification des Semi-Groupes de diffusion sur \mathbb{R} associés à une famille de polynômes orthogonaux. In: *Séminaire de Probabilités XXXI*, pp. 40-53. Springer Berlin Heidelberg.
- [8] 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.
- [9] I. Nourdin and G. Poly (2013). Convergence in total variation on Wiener chaos. *Stoch. Proc. Appl.* **123**, pp. 651-674.
- [10] R. Reiss (1989). *Approximate Distributions of Order Statistics, with Applications to Non-parametric Statistics*. Springer-Verlag, New York.
- [11] S. Kh. Sirazhdinov and M. Mamatov (1962). On convergence in the mean for densities. *Theor. Probab. Appl.* **7**, no. 4, 424-428.