



HAL
open science

Berry-Esseen bounds for the chi-square distance in the Central Limit Theorem: a Markovian approach

Claire Delplancke, Laurent Miclo

► **To cite this version:**

Claire Delplancke, Laurent Miclo. Berry-Esseen bounds for the chi-square distance in the Central Limit Theorem: a Markovian approach. 2017. hal-01595748v2

HAL Id: hal-01595748

<https://hal.science/hal-01595748v2>

Preprint submitted on 8 Nov 2017

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.

Berry-Esseen bounds for the χ_2 -distance in the Central Limit Theorem: a Markovian approach

Claire Delplancke* Laurent Miclo†

November 8, 2017

Abstract

This article presents a new proof of the rate of convergence to the normal distribution of sums of independent, identically distributed random variables in chi-square distance, which was also recently studied in [7]. Our method consists of taking advantage of the underlying time non-homogeneous Markovian structure and studying the spectral properties of the non-reversible transition operator, which allows to find the optimal rate in the convergence above under matching moments assumptions. Our main assumption is that the random variables involved in the sum are independent and have polynomial density; interestingly, our approach allows to relax the identical distribution hypothesis.

Keywords: Berry-Esseen bounds, Central Limit Theorem, non-homogeneous Markov chain, Hermite-Fourier decomposition, χ_2 -distance.

Mathematics Subject Classification (MSC2010): 60F05, 60J05, 60J35, 47A75, 39B62.

1 Introduction and main result

The present article is devoted to the convergence in Central Limit Theorem with respect to χ_2 -distance, defined for two probability distributions θ, μ as:

$$\chi_2(\theta, \mu) := \left(\int \left(\frac{d\theta}{d\mu} - 1 \right)^2 d\mu \right)^{\frac{1}{2}}$$

if θ is absolutely continuous with respect to μ and $+\infty$ otherwise. In the following μ stands for the normal distribution. For a density function $f \in L^2(\mu)$ we use the shortened notation

$$\chi_2(f) := \chi_2(f \cdot \mu, \mu)$$

*Center of Mathematical Modeling, Santiago, Chile, cdelplancke@cmm.uchile.cl.

†Institut de Mathématiques de Toulouse, UMR 5219, Université de Toulouse and CNRS, France, laurent.miclo@math.univ-toulouse.fr.

to refer to the χ_2 -distance between the distribution with density f with respect to μ , denoted by $f \cdot \mu$, and μ itself. The χ_2 -distance bounds by above usual quantities like total variation distance and relative entropy:

$$d_{TV}(f, \mu) := \int |f - 1| d\mu \leq \chi_2(f), \quad \text{Ent}(f||\mu) := \int f \log f d\mu \leq \chi_2^2(f), \quad (1)$$

the first relation being a consequence of Cauchy-Schwarz inequality and the second of the inequality $\log x \leq x - 1$ for $x > 0$.

Let $(X_i)_{i \geq 1}$ be real i.i.d. random variables with density φ with respect to the normal distribution μ , and consider the renormalized sum

$$Y_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i, \quad n \geq 1. \quad (2)$$

We call f_n the density of Y_n with respect to μ . The main result of this article is the following.

Theorem 1.1 (Asymptotic bound on χ_2). *If the moments of X_1 and the moments of μ match up to order r for a given integer $r \geq 2$, and if the density φ of X_1 with respect to μ is polynomial and satisfies to Hypothesis **(H)** stated in Section 2, then*

$$\limsup_{n \rightarrow +\infty} n^{\frac{r-1}{2}} \chi_2(f_n) < +\infty.$$

The rate of convergence, which improves by a factor \sqrt{n} for each supplementary moment of X_1 that agrees with the corresponding moment of μ , is optimal: indeed by inequality (1), it implies a rate of at least $n^{(r-1)/2}$ in total variation distance, which is proved to be optimal in [3].

While this article was being written, the authors took notice of the recent article [7], which provides the same rate of convergence, as well as the constant in front of it, in the i.i.d. case, and under optimal assumptions.

Our method of proof starts from the simple observation that the sequence of renormalized sums $(Y_n)_{n \geq 1}$ is a non-homogeneous Markov chain, as it satisfies to the recursion equation

$$Y_{n+1} = \sqrt{1 - \frac{1}{n+1}} Y_n + \frac{1}{\sqrt{n+1}} X_{n+1}, \quad n \geq 1. \quad (3)$$

Although the result which we obtain in Theorem 1.1 is weaker than that of [7] in the i.i.d. case, this method presents two advantages:

- The result can be extended to independent, non necessarily identically distributed random variables $(X_i)_{i \geq 1}$, as this does not affect the Markovian character of $(Y_n)_{n \geq 1}$, provided that the $(X_i)_{i \geq 1}$ are distributed according to a finite set of distributions, each polynomial and satisfying to **(H)** (cf Remark 5.1). The idea to use a Markovian framework to deal with non identically distributed random variables has also recently been used in [4].

- It is original and rather straightforward. The action of the Markov transition semigroup amounts to a barycentric convolution, since

$$\left(\sqrt{1 - \frac{1}{n+1}}\right)^2 + \left(\frac{1}{\sqrt{n+1}}\right)^2 = 1.$$

From the point of view of functional analysis, our approach amounts to look at the bilinear barycentric convolution operator

$$(f, \varphi) \mapsto af * b\varphi,$$

where f, φ are two density functions in $L^2(\mu)$, $a, b \in \mathbb{R}$ such that $a^2 + b^2 = 1$, and $af * b\varphi$ denotes the density of the random variable $aU + bV$ which $U \sim f \cdot \mu$, $V \sim \varphi \cdot \mu$, as the linear operator (denoted $\mathcal{Q}_{a,\varphi}^*$ in the following):

$$\begin{aligned} L^2(\mu) &\rightarrow L^2(\mu) \\ f &\mapsto af * b\varphi. \end{aligned}$$

The function φ is fixed and represents the density of the innovations $(X_i)_{i \geq 1}$. Thanks to spectral analysis in the Hermite-Fourier domain, we derive an estimate of the operator norm, which is directly related to the rate of convergence in Theorem 1.1 and to Theorem 1.2 below.

More on the proof.– At the heart of the study is a formula describing the evolution of the χ_2 -distance under the action of the barycentric convolution. Set $(\bar{H}_n)_{n \in \mathbb{Z}^+}$ the set of renormalized Hermite polynomials, forming an orthonormal basis of $L^2(\mu)$:

$$\bar{H}_n(x) := \frac{(-1)^n}{\sqrt{n!}} e^{\frac{x^2}{2}} D^n \left(e^{-\frac{x^2}{2}} \right), \quad x \in \mathbb{R}, \quad n \in \mathbb{Z}^+,$$

where D denotes the derivation operator acting on smooth functions from \mathbb{R} to \mathbb{R} . We show the following result:

Theorem 1.2 (Barycentric convolution and χ_2). *Let r be a natural integer and f, φ be two densities in $L^2(\mu)$ whose moments match the moments of μ up to order r , and assume moreover that the density φ is polynomial. In particular, φ admits a decomposition on the Hermite basis of the form:*

$$\varphi = 1 + \sum_{k=r+1}^N \varphi_k \bar{H}_k. \quad (4)$$

Set:

$$a_\varphi := \left(1 + \frac{N}{r+1}\right)^{-\frac{1}{4}} \in (0, 1],$$

and for all $a \in (0, 1)$,

$$d_\varphi(a) := \sum_{k=r+1}^N \frac{|\varphi_k|}{\sqrt{k!}} \left(-2(r+1+N) \left(1 + \frac{N}{r+1}\right) \log a \right)^{k/2}.$$

If φ satisfies to Hypothesis **(H)** stated in Section 2, then for all $a \in (a_\varphi, 1)$, the following inequality stands:

$$\chi_2 \left(af * \sqrt{1 - a^2} \varphi \right) \leq a^{r+1} (1 + d_\varphi(a)) \chi_2(f) + (1 - a^2)^{\frac{r+1}{2}} \chi_2(\varphi). \quad (E_r)$$

Moreover,

$$\lim_{a \rightarrow 1} |\log a|^{-\frac{r+1}{2}} d_\varphi(a) < +\infty.$$

Equation (4), which involves Hermite coefficients $(\varphi_k)_{r+1 \leq k \leq N}$, is explained in Section 2. Hypothesis **(H)** gives conditions on these Hermite coefficients.

Theorem 1.2 does not preserve the symmetry of the equation $af * (1 - a^2)^{1/2} \varphi = (1 - a^2)^{1/2} \varphi * af$, in the first place because the assumption on f is weaker than the assumption on φ , and in the second place because the term $d_\varphi(a)$ in the upper-bound is non-vanishing in general. However, in the regimen of interest $a \rightarrow 1$ (which corresponds to $(n/n+1)^{1/2} \rightarrow 1$), one has

$$a^{r+1} = 1 - \frac{r+1}{2}(1-a) + o(a^2), \quad 1 + d_\varphi(a) = 1 + \mathcal{O}\left((1-a)^{\frac{r+1}{2}}\right),$$

hence in the Taylor expansion of the prefactor $a^{r+1}(1 + d_\varphi(a))$ the contribution from $d_\varphi(a)$ is negligible with respect to the contribution from a^{r+1} as soon as $r \geq 2$, i.e. if the centering and normalizing condition $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X_1^2] = 1$ is satisfied. Let us also mention that there exists an alternative inequality, holding without the polynomial assumption on φ , stated in Section 4.

Bound (E_r) bears a similarity to Shannon-Stam inequality for (absolute) entropy ([21, 23]):

$$\text{Ent} \left(aU + \sqrt{1 - a^2}V \right) \leq a \text{Ent}(U) + \sqrt{1 - a^2} \text{Ent}(V), \quad a \in [0, 1], \quad (5)$$

although, by the observation above, the coefficients in front of Ent are of different order with respect to the coefficients in front of χ for all natural integer r .

For $r = 1$, bound (E_r) gives back Poincaré inequality for the Ornstein-Uhlenbeck semi-group $(P_t)_{t \geq 0}$, defined for $f \in L^2(\mu)$ as

$$P_t[f](x) = \mathbb{E} \left[f \left(e^{-t}x + \sqrt{1 - e^{-2t}}Z \right) \right], \quad Z \sim \mu; \quad x \in \mathbb{R}, \quad t \geq 0. \quad (6)$$

Indeed, in the case where φ is the density of the normal distribution μ , that is to say $\varphi = 1$, we adopt the convention $r = +\infty$ and $N = 0$ in equality (4), and set $a_\varphi = 0$ and $d_\varphi(a) = 0$ for all $a \in [0, 1]$. As we will see in Section 4, (E_1) then corresponds (up to a positivity assumption which can actually be discarded in the proof) to Poincaré inequality for $(P_t)_{t \geq 0}$: for $f \in L^2(\mu)$,

$$\text{Var}_\mu(P_t f) \leq e^{-2t} \text{Var}_\mu(f), \quad t \geq 0. \quad (7)$$

Results of the literature. – The first quantification result for the convergence of renormalized sums of i.i.d. variables was obtained independently by Berry and Esseen through

Kolmogorov distance ([6, 12]). Rate of convergence in total variation distance is first addressed in [22] and the optimal rate under matching moment assumptions and regularity condition is proved in [3] using Malliavin's calculus.

The rate of convergence in Theorem 1.1 relies crucially on the fact that the inequality (E_r) incorporates the matching moments assumption through exponent r on the barycentric coefficients. By comparison, Shannon-Stam inequality (5) implies (jointly with a result of monotonicity of relative entropy and Fisher information under convolution) the convergence $\text{Ent}(f_n) \rightarrow 0$ without rate ([10, 5]). The optimal rate is derived in [2] when the random variable X_1 satisfies a Poincaré inequality. [8] provides an asymptotic expansion of entropy involving the moments of X_1 . Similar developments occurred for Fisher information ([10], [9]), and for Rényi distances ([7]), which include the χ_2 -distance.

Other distances include Sobolev ([15]) and Wasserstein ([16, 24, 19, 20, 11]) distances. A typical assumption in Berry-Esseen theorems is the existence of moments up to a certain order for the random variable X_1 . In the present framework, the fact that φ is in $L^2(\mu)$ implies that moments of all order exist. This is consistent with the fact that the χ_2 -distance bounds by above the usual quantities (as shown in (1); a similar inequality holds for Wasserstein distance of order 1).

In another direction of research, Stein's method and Malliavin calculus revealed to be powerful tools to study, including in a quantitative way, the asymptotic normality of multidimensional random variables living in Gaussian chaoses. Let us cite, among an increasingly rich literature, the reference book [17]. It is interesting to remark the formal similarity between their objects and ours, although the results do not compare. Indeed, in the Stein-Malliavin framework, the typical random variable X_1 writes $X_1 = \varphi(Z)$, with Z a Gaussian random variable living in \mathbb{R}^n with law μ_n and $\varphi \in L^2(\mu_n)$. To compare to our framework, let us take $n = 1$ and φ a density in $L^2(\mu)$. The random variable $X_1 = \varphi(Z)$, where $Z \sim \mu$, bears no relation with the random variable X_1 which has *density* φ with respect to μ ; hence, the two approaches are not reducible one to another. The authors would like to thank Anthony Réveillac for interesting discussions on this subject.

Structure of the article.— The remaining of this paper is organized as follows. In Section 2, Hermite-Fourier decomposition (4) is detailed, and Hypothesis **(H)** is stated and commented. Section 3 is devoted to the explicit expression of the convolution operator and of its Hermite-Fourier decomposition. Theorem 1.2 is proved in Section 4 by spectral analysis in the Hermite-Fourier domain. Finally Section 5 is devoted to the proof of Theorem 1.1.

2 Hermite-Fourier decomposition of the density

First, let us introduce some notation. The symbol $\mathbf{1}$ stands for the function from \mathbb{R} to \mathbb{R} identically equal to 1, \mathbb{Z}^+ for the set of natural integers and $\binom{n}{k}$ for the binomial coefficient associated to natural integers $k \leq n$. For a function $f \in L^1(\mu)$, we denote indifferently

$$\mu(f) = \int f d\mu.$$

The space $L^2(\mu)$ is a Hilbert space, with scalar product and associated norm defined as

$$\langle f, g \rangle_{L^2(\mu)} := \int fg d\mu, \quad \|f\|_{L^2(\mu)} := \sqrt{\mu(f^2)}, \quad f, g \in L^2(\mu).$$

Set

$$\text{Var}_\mu(f) := \int (f - \mu(f))^2 d\mu.$$

Hermite polynomials $(H_n)_{n \in \mathbb{Z}^+}$ are defined as:

$$H_n(x) := (-1)^n e^{\frac{x^2}{2}} D^n \left(e^{-\frac{x^2}{2}} \right), \quad x \in \mathbb{R}, \quad n \in \mathbb{Z}^+,$$

where we recall that D stands for the derivation operator acting on smooth functions from \mathbb{R} to \mathbb{R} . Hermite polynomials are also characterized by the following equation: for all smooth functions $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$\int f H_n d\mu = \int D^n f d\mu. \quad (8)$$

They form an orthogonal basis of $L^2(\mu)$: for all $n, m \in \mathbb{Z}^+$,

$$\int H_m H_n d\mu = n! \delta_{n,m}.$$

In the paper it is more convenient to work with renormalized Hermite polynomials $(\bar{H}_n)_{n \in \mathbb{Z}^+}$:

$$\bar{H}_n = \frac{1}{\sqrt{n!}} H_n, \quad n \in \mathbb{Z}^+,$$

which form an orthonormal basis of $L^2(\mu)$. By convention set $\bar{H}_{-1} = 0$. One has:

$$D\bar{H}_n = \sqrt{n} \bar{H}_{n-1}, \quad n \in \mathbb{Z}^+. \quad (9)$$

As $H_0 = 1$, for all natural integer \bar{H}_n is of degree n . The basis $(\bar{H}_n)_{n \in \mathbb{Z}^+}$ is diagonal for the Ornstein-Uhlenbeck semigroup defined in (6):

$$P_t[\bar{H}_n] = e^{-nt} \bar{H}_n, \quad n \in \mathbb{Z}^+, \quad t \geq 0.$$

For all functions $g \in L^2(\gamma)$, call $(g_k)_{k \in \mathbb{Z}^+}$ its coefficients on the orthonormal Hermite basis,

$$g = \sum_{k \in \mathbb{Z}^+} g_k \bar{H}_k,$$

where the equality stands in $L^2(\mu)$, and denote indifferently $\mathcal{F}(g) := \vec{g} = (g_k)_{k \in \mathbb{Z}^+}$ the sequence of its coefficients, which belongs to the Hilbert space l^2 , defined as the set of real sequences $(u_k)_{k \in \mathbb{Z}^+}$ such that $\sum_{k \in \mathbb{Z}^+} u_k^2 < +\infty$. The application

$$L^2(\mu) \rightarrow l^2, \quad g \mapsto \mathcal{F}(g), \quad (10)$$

is an isometry of Hilbert spaces.

If $\varphi \in L^2(\mu)$ is a density, then $\varphi_0 = 1$. The matching moments assumption has a nice interpretation in terms of the coefficients: for all positive integer r ,

$$\left(\forall k \in \{1, \dots, r\}, \quad \int x^k \varphi(x) d\mu(x) = \int x^k d\mu(x) \right) \Leftrightarrow \left(\forall k \in \{1, \dots, r\}, \quad \varphi_k = 0 \right).$$

Indeed, Hermite polynomial \bar{H}_n being of degree n for all natural integers, one has the equivalence

$$\begin{aligned} & \left(\forall k \in \{1, \dots, r\}, \quad \int x^k \varphi(x) d\mu(x) = \int x^k d\mu(x) \right) \\ \Leftrightarrow & \left(\forall k \in \{1, \dots, r\}, \quad \int \bar{H}_k(x) \varphi(x) d\mu(x) = \int \bar{H}_k(x) d\mu(x) \right), \end{aligned}$$

and by orthogonality of $(\bar{H}_n)_{n \in \mathbb{Z}^+}$ it stands that for all $k \in \mathbb{Z}^+$,

$$\int \bar{H}_k d\mu = \int \bar{H}_k \bar{H}_0 d\mu = \delta_{0,k}.$$

The assumption that φ is a polynomial density whose moments agree with moments of μ up to r hence amounts to:

$$\exists N \in \mathbb{Z}^+, \quad N > r, \quad \varphi = 1 + \sum_{k=r+1}^N \varphi_k \bar{H}_k,$$

which corresponds to equality (4) above. From now on we set $K = r + 1$. When φ is the density of μ itself, that is to say $\varphi = \bar{H}_0 = 1$, we set $K = +\infty$ and $N = 0$ by convention.

Let us introduce the quantities

$$C_k := \left(1 + \frac{N}{K}\right)^{k/2}, \quad \gamma_k = \frac{1}{\sqrt{k!}} C_k |\varphi_k|, \quad k \in \mathbb{Z}^+.$$

We are now ready to state Hypothesis **(H)**, which is composed of two parts, **(H1)** and **(H2)** as follows.

(H1) If $K \leq N - 2$, for all $K \leq k \leq N - 2$,

$$(k + 2)\gamma_{k+2} \leq \gamma_k.$$

For non-vanishing φ_k this relation is equivalent to

$$\left| \frac{\varphi_{k+2}}{\varphi_k} \right| \leq \frac{1}{1 + \frac{N}{K}} \left(1 - \frac{1}{k+2}\right)^{1/2}$$

if $\varphi_k \neq 0$ and is implied by the simplest assumption

(H1') If $K \leq N - 2$, for all $K \leq k \leq N$,

$$|\varphi_{k+1}| \leq r |\varphi_k|; \quad r := \left(\left(1 - \frac{1}{K+2}\right)^{1/2} \frac{1}{1 + \frac{N}{K}} \right)^{1/2}.$$

Remark that assumption **(H1)** implies that $\gamma_k = 0 \Rightarrow \gamma_{k+2} = 0$.

The second condition **(H2)** has two parts:

(H2a) If $K \leq N - 1$,

$$\left(\frac{2(K+1)\gamma_{K+1}}{N} \right)^N \left(\frac{K-1}{2\gamma_N} \right)^{K-1} \vee \left(\frac{2K\gamma_K}{N-1} \right)^{N-1} \left(\frac{K-2}{2\gamma_{N-1}} \right)^{K-2} \leq \frac{1}{(N-K+1)^{N-K+1}}.$$

(H2b) If $K = N$,

$$\gamma_N \leq \frac{1}{2(N-2)^{(N-2)/2}}.$$

Loosely speaking, assumption **(H1)**, which is present only if $K \leq N - 2$, amounts to ask a geometric decrease for the coefficients (φ_k) . If $K = N$ or $K = N - 1$, assumption **(H2)** requires the leading coefficients φ_N and φ_{N-1} to be not too big, which is fair enough; it is more painful to write when $K < N - 1$, but it can be interpreted as the requirement that the coefficients $(\varphi_k)_{K \leq k \leq N}$ do not decay too fast.

We conclude this section by the following comment on the range of validity of Theorems 1.2 and 1.1.

Remark 2.1 (On the polynomial assumption). *We conjecture that inequality (E_r) holds without Hypothesis **(H)**: indeed, the fact that φ , as a density, is nonnegative already implies restrictions on the coefficients, which are in fact sufficient to prove (E_r) in the case of densities of the form $\varphi = 1 + c\bar{H}_2$ and $\varphi = 1 + c'\bar{H}_4$.*

Whether the polynomial assumption is necessary is less clear. For comparison, the hypothesis of Gaussian chaos of finite orders is needed in [18].

3 Explicit expression of the convolution operator

3.1 Convolution as a Markovian transition

Set $(X_i)_{i \geq 1}$ random variable of density $\varphi \in L^2(\mu)$. Throughout the paper, the notation f_n stands for the density of the renormalized sum

$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i, \quad n \geq 1,$$

which satisfy to the recursion equation:

$$Y_{n+1} = \sqrt{1 - \frac{1}{n+1}} Y_n + \frac{1}{\sqrt{n+1}} X_{n+1}, \quad n \geq 1. \quad (3)$$

For a parameter $a \in [0, 1]$, introduce the bilinear barycentric convolution operator K_a defined as

$$\forall f, \varphi \in L^2(\mu), \quad K_a(f, \varphi) := af * \sqrt{1 - a^2} \varphi.$$

Equation (3) in turn yields the corresponding recursion relation for the successive densities:

$$f_{n+1} = K_{a_{n+1}}(f_n, \varphi), \quad a_{n+1} := \sqrt{1 - \frac{1}{n+1}}, \quad n \geq 1. \quad (11)$$

On the other hand, relation (3) translates into the the fact that $(Y_n)_{n \geq 1}$ is an inhomogeneous Markov chain. The associated semigroup $(\mathcal{Q}_{p,q})_{q \geq p \geq 1}$ is defined for continuous bounded functions as:

$$\mathcal{Q}_{p,q}[f](x) := \mathbb{E}[f(Y_q) | Y_p = x], \quad x \in \mathbb{R}, \quad q \geq p \geq 1.$$

By relation (3), the explicit expression of $\mathcal{Q}_{n,n+1}$ is straightforward: for all $f \in L^2(\mu)$,

$$\mathcal{Q}_{n,n+1}[f] = Q_{a_{n+1},\varphi}[f],$$

where the operator $Q_{a,\varphi}$ is defined for all $a \in [0, 1]$ and density $\varphi \in L^2(\mu)$ as:

$$Q_{a,\varphi}[f](x) := \int f(ax + \sqrt{1-a^2}y)\varphi(y)dy, \quad x \in \mathbb{R}.$$

Denote $\mathcal{Q}_{n,n+1}^*$ (resp. $Q_{a,\varphi}^*$) the adjoint of $\mathcal{Q}_{n,n+1}$ (resp. $Q_{a,\varphi}$) in $L^2(\mu)$. The discrete version of Kolmogorov backward relation reads:

$$f_{n+1} := \mathcal{Q}_{n,n+1}^*[f_n], \quad n \geq 1.$$

Hence $\mathcal{Q}_{n,n+1}^*[f_n] = K_{a_{n+1}}(f_n, \varphi)$. More generally, for f, φ densities in $L^2(\mu)$, the following identity holds:

$$K_a(f, \varphi) = Q_{a,\varphi}^*[f], \quad a \in [0, 1].$$

The strategy used in the paper relies on this interpretation of barycentric convolution as the action of a Markovian transition, as explained in Section 4.1.

Proposition 3.1 (Explicit expression of the operators). *For all $a \in [0, 1]$, one has*

$$Q_{a,\varphi}[f](x) = \int f(ax + \sqrt{1-a^2}y)\varphi(y)d\mu(y), \quad (12)$$

$$Q_{a,\varphi}^*[f](x) = \int f(ax - \sqrt{1-a^2}y)\varphi(\sqrt{1-a^2}x + ay)d\mu(y) = K_a(f, \varphi)(x). \quad (13)$$

The formulas are to be understood in the following way: if f is bounded and continuous, they stand for all $x \in \mathbb{R}$; if $f \in L^2(\mu)$, they stand in the almost everywhere sense. One sees that the operators $Q_{a,\varphi}$ and $Q_{a,\varphi}^*$ are actually defined for all $f, \varphi \in L^2(\gamma)$ independently from them being densities; in what follows, $Q_{a,\varphi}$ and $Q_{a,\varphi}^*$ refers to this extended definition when required. Furthermore, $Q_{a,\varphi}$ and $Q_{a,\varphi}^*$ are bounded in $L^2(\mu)$ for all $\varphi \in L^2(\mu)$: indeed by Cauchy-Schwarz inequality, for all $f \in L^2(\mu)$,

$$\int (Q_{a,\varphi}[f])^2 d\mu \leq \|f\|_{L^2(\mu)}^2 \|\varphi\|_{L^2(\mu)}^2, \quad \int (Q_{a,\varphi}^*[f])^2 d\mu \leq \|f\|_{L^2(\mu)}^2 \|\varphi\|_{L^2(\mu)}^2.$$

Proof of Proposition 3.1. Formula (12) has already been given; let us prove formula (13). As $a \in [0, 1]$, there exists $\theta \in \mathbb{R}$ such that $\cos \theta = a$ and $\sin \theta = \sqrt{1-a^2}$. Denote R_θ the rotation of \mathbb{R}^2 with parameter θ and Γ the Gaussian distribution on \mathbb{R}^2 with identity as covariance matrix. Then, invariance of Γ under the action of R_θ implies that for all $f, g \in L^2(\gamma)$,

$$\begin{aligned} \int f Q_{a,\varphi}[g] d\mu &= \iint f(x)\varphi(y)g((\cos \theta)x + (\sin \theta)y) d\mu(x)d\mu(y) \\ &= \iint f(\Re(X))\varphi(\Im(X))g(\Re(R_\theta X)) d\Gamma(X) \\ &= \iint f(\Re(R_{-\theta} X))\varphi(\Im(R_{-\theta} X))g(\Re(X)) d\Gamma(X) = \int g Q_{a,\varphi}^*[f] d\mu. \end{aligned}$$

□

Remark 3.2 (Link with Ornstein-Uhlenbeck semigroup). *Recall the definition of the Ornstein-Uhlenbeck semigroup $(P_t)_{t \geq 0}$ given in (6). Hence, for all $a \in [0, 1]$ and $f, \varphi \in L^2(\mu)$, one has the two useful equalities:*

$$Q_{a,1}^*[f] = K_a(f, \mathbf{1}) = P_{-\log a}[f], \quad Q_{a,\varphi}^*[\mathbf{1}] = K_a(\mathbf{1}, \varphi) = P_{-\frac{1}{2}\log(1-a)^2}[\varphi].$$

Let us introduce the multiplicative reversibilization of the Markov transition operator $Q_{a,\varphi}$, defined as

$$M_{a,\varphi} := Q_{a,\varphi} Q_{a,\varphi}^*. \quad (14)$$

The concept of reversibilization of a non-reversible Markov operator traces back to [13] which deals with homogeneous, invariant Markov chains. The new operator $M_{a,\varphi}$ is now symmetric in $L^2(\mu)$, though in the general case, it is not Markovian: as $Q_{a,\varphi}$ is not invariant, then $Q_{a,\varphi}^*[\mathbf{1}] \neq \mathbf{1}$ and the mass conservation property $M_{a,\varphi}[\mathbf{1}] = \mathbf{1}$ does not hold. Nonetheless, we will see in Section 4.3 that spectral analysis of $M_{a,\varphi}$ gives quantitative information on the action of Markovian transition $Q_{a,\varphi}$ and convolution operator $Q_{a,\varphi}^*$.

3.2 Hermite-Fourier decomposition of the convolution operator

Let us now determine how the operators $Q_{a,\varphi}$, $Q_{a,\varphi}^*$ and $M_{a,\varphi}$ act with respect to the Hermite-Fourier decomposition defined in (10). For a bounded operator Q of $L^2(\mu)$, we call \vec{Q} the infinite matrix defined as:

$$\vec{Q} := (\vec{Q}(m, n))_{n, m \in \mathbb{Z}^+}; \quad (\vec{Q})(m, n) := \int Q(\bar{H}_n) \bar{H}_m d\mu, \quad n, m \in \mathbb{Z}^+.$$

The matrix \vec{Q} is the unique bounded operator of l^2 such that

$$\mathcal{F}(Q[f]) = \vec{Q} \vec{f}, \quad f \in L^2(\mu).$$

Set $\|R\|_{\text{op}}$ the operator norm of a bounded operator R on a Hilbert space \mathcal{H} , defined with evident notation as

$$\|R\|_{\text{op}} := \sup_{h \in \mathcal{H} \setminus \{0\}} \frac{\|Rh\|_{\mathcal{H}}}{\|h\|_{\mathcal{H}}}.$$

The following property reveals useful: for all bounded operator Q on $L^2(\mu)$, it stands that:

$$\|Q\|_{\text{op}} = \|\vec{Q}\|_{\text{op}}. \quad (15)$$

In the following proposition, we give the matrices $\vec{Q}_{a,\varphi}$, $\vec{Q}_{a,\varphi}^*$ and $\vec{M}_{a,\varphi}$ associated to the operators $Q_{a,\varphi}$, $Q_{a,\varphi}^*$ and $M_{a,\varphi}$.

Proposition 3.3 (Matrix form of operators). *For all $\varphi \in L^2(\mu)$ and $a \in [0, 1]$, one has:*

$$\forall m, n \in \mathbb{Z}^+, \quad \vec{Q}_{a,\varphi}(m, n) = \begin{cases} \binom{n}{m}^{\frac{1}{2}} a^m (1-a^2)^{\frac{n-m}{2}} \varphi_{n-m}, & m \leq n \\ 0, & m > n, \end{cases} \quad (16)$$

Denoting ${}^T N$ the transpose matrix of N , it stands that:

$$\vec{Q}_{a,\varphi}^* = {}^T \vec{Q}_{a,\varphi}.$$

Finally, the matrix $\vec{M}_{a,\varphi}$ is symmetric and

$$\forall l, i \in \mathbb{Z}^+, i \leq l, \quad \vec{M}_{a,\varphi}(l, l-i) = a^{2l-i} \sum_{k \geq 0} \binom{k+l}{k}^{1/2} \binom{k+l}{k+i}^{1/2} (1-a^2)^{\frac{2k+i}{2}} \varphi_{k+i} \varphi_k. \quad (17)$$

Proof. To begin with, one needs to compute the Hermite-Fourier decomposition of $Q_{a,\bar{H}_m}[\bar{H}_n]$, for $a \in [0, 1]$ and $n, m \in \mathbb{Z}^+$. Applying the properties of Hermite polynomials recalled in Section 2 yields:

$$Q_{a,\bar{H}_m}[\bar{H}_n] = \int \bar{H}_n(ax + \sqrt{1-a^2}y) \bar{H}_m(y) d\mu(y) = \frac{1}{\sqrt{n!}\sqrt{m!}} \int D^m(H_n(ax + \sqrt{1-a^2}y)) d\mu(y).$$

By the degree property, the integral vanishes for $n < m$. For $n \geq m$,

$$\begin{aligned} Q_{a,\bar{H}_m}[\bar{H}_n] &= (1-a^2)^{\frac{m}{2}} \frac{n \cdots (n-m+1)}{\sqrt{n!}\sqrt{m!}} \int H_{n-m}(ax + \sqrt{1-a^2}y) d\mu(y) \\ &= (1-a^2)^{\frac{m}{2}} \frac{n \cdots (n-m+1)}{\sqrt{n!}\sqrt{m!}} P_{-\log a}[H_{n-m}] \\ &= a^{n-m} (1-a^2)^{\frac{m}{2}} \frac{n \cdots (n-m+1)}{\sqrt{n!}\sqrt{m!}} H_{n-m} \\ &= a^{n-m} (1-a^2)^{\frac{m}{2}} \binom{n}{m}^{\frac{1}{2}} \bar{H}_{n-m}. \end{aligned}$$

By bilinearity, write

$$\begin{aligned} Q_{a,\varphi}(\bar{H}_n) &= \sum_{m \in \mathbb{Z}^+} \varphi_m Q_{a,\bar{H}_m}[\bar{H}_n] = \sum_{m=0}^n \varphi_m a^{n-m} (1-a^2)^{\frac{m}{2}} \binom{n}{m}^{\frac{1}{2}} \bar{H}_{n-m} \\ &= \sum_{m=0}^n \varphi_{n-m} a^m (1-a^2)^{\frac{n-m}{2}} \binom{n}{m}^{\frac{1}{2}} \bar{H}_m = \sum_{m=0}^n \vec{Q}_{a,\varphi}(m, n) \bar{H}_m, \end{aligned}$$

which proves (16). Furthermore, definition (14) implies that $\vec{M}_{a,\varphi} = \vec{Q}_{a,\varphi} {}^T \vec{Q}_{a,\varphi}$, and formula (17) follows by simple computation. \square

Remark 3.4 (On Hermite decomposition of the convolution operator).

- It is to be noted that convolution with barycentric coefficients only admits a nice decomposition; contrarily to what happens with Fourier transform associated to Lebesgue measure, the usual convolution has no explicit Hermite-Fourier representation.
- Thanks to formula (16) above, one finds that for all $\varphi \in L^2(\mu)$, $m \in \mathbb{Z}^+$ and $a \in [0, 1]$,

$$Q_{a,\varphi}^*(\bar{H}_m) = a^m \sum_{n \geq 0} \binom{m+n}{n}^{1/2} (1-a^2)^{\frac{n}{2}} \varphi_n \bar{H}_{m+n},$$

which allows to better understand the behaviour of the barycentric convolution: each nonvanishing coefficient on \bar{H}_m and \bar{H}_n in the respective decompositions of f and φ contribute to a coefficient on \bar{H}_{m+n} in the decomposition of $K_a(f, \varphi)$.

- If φ is polynomial, then $\vec{M}_{a,\varphi}$ is a band matrix.

We already noticed that $Q_{a,\varphi}$ and $Q_{a,\varphi}^*$, and by composition $M_{a,\varphi} = Q_{a,\varphi}Q_{a,\varphi}^*$, are bounded operators. In fact, they are Hilbert-Schmidt operators. By definition, a bounded operator R on the Hilbert space \mathcal{H} is Hilbert-Schmidt, if, $(e_n)_{n \in \mathbb{Z}^+}$ standing for an orthonormal basis of \mathcal{H} , one has:

$$\sum_{n \in \mathbb{Z}^+} \|R(e_n)\|_{\mathcal{H}}^2 < +\infty.$$

Proposition 3.5 (Hilbert-Schmidt operators). *For all $\varphi \in L^2(\mu)$ and $a \in [0, 1)$, the operators $Q_{a,\varphi}, Q_{a,\varphi}^*, M_{a,\varphi}$ are Hilbert-Schmidt, hence compact.*

Proof. Consider first $Q_{a,\varphi}^*$.

$$\begin{aligned} \sum_{n \in \mathbb{Z}^+} \|Q_{a,\varphi}^*(\bar{H}_n)\|_{L^2(\mu)}^2 &= \sum_{n \in \mathbb{Z}^+} \langle \bar{H}_n, M_{a,\varphi} \bar{H}_n \rangle_{L^2(\mu)} = \sum_{n \in \mathbb{Z}^+} \vec{M}_{a,\varphi}(n, n) \\ &= \sum_{n, k \in \mathbb{Z}^+} \binom{k+n}{k} (1-a^2)^k a^{2n} \varphi_k^2. \end{aligned}$$

Now, by the equality

$$\sum_{n \in \mathbb{Z}^+} \binom{k+n}{k} u^n = \frac{1}{(1-u)^{k+1}}, \quad k \in \mathbb{Z}^+, \quad u \in [0, 1),$$

we find that

$$\sum_{n \in \mathbb{Z}^+} \|Q_{a,\varphi}^*(\bar{H}_n)\|_{L^2(\mu)}^2 = \sum_{k \in \mathbb{Z}^+} \frac{(1-a^2)^k}{(1-a^2)^{k+1}} \varphi_k^2 = \frac{1}{1-a^2} \|\varphi\|_{L^2(\mu)}^2 < +\infty.$$

This implies that $Q_{a,\varphi}$ is Hilbert-Schmidt and in turn so is $M_{a,\varphi}$ by composition. \square

4 Proof of Theorem 1.2

4.1 Strategy for non-homogeneous Markov chains

Let us now explain the strategy to exploit the Markovian framework. In Remark 3.2, we noticed that if $\varphi = \mathbf{1}$, *i.e.* the X_i 's are normal, then $Q_{a,\mathbf{1}}^* = P_{-\log a}$. In this case, the renormalized sums $(Y_n)_{n \geq 1}$ are also normal, which corresponds to the fact that the Ornstein-Uhlenbeck semigroup $(P_t)_{t \geq 0}$ is invariant, and in fact reversible, with respect to μ . The semigroup $(P_t)_{t \geq 0}$ also enjoys a Poincaré inequality recalled in equation (7). If f is a density then $\text{Var}_\mu(f) = \chi^2(f)$, hence Poincaré inequality for the Ornstein-Uhlenbeck semigroup reads:

$$\chi_2(P_t[f]) \leq e^{-t} \chi_2(f), \quad t \geq 0.$$

Furthermore, $f_{n+1} = P_{-\log a_{n+1}}[f_n]$ by (11), hence:

$$\chi_2(f_{n+1}) = \chi_2(P_{-\log a_{n+1}}[f_n]) \leq a_n \chi_2(f_n), \quad n \geq 1,$$

and by straightforward calculation one gets the decrease of $\chi_2(f_n)$.

The idea underlying our method consists in mimicking the reasoning above for the true operator $Q_{a,\varphi}^*$ acting on densities, which is neither reversible nor satisfies to the mass conservation property in the general case, as was explained above. For $a \in [0, 1]$ and f a density in $L^2(\mu)$, let us write by triangular inequality:

$$\begin{aligned} \chi_2(Q_{a,\varphi}^*[f]) &= \|Q_{a,\varphi}^*[f] - 1\|_{L^2(\mu)} \\ &\leq \|Q_{a,\varphi}^*[f - 1]\|_{L^2(\mu)} + \|Q_{a,\varphi}^*[\mathbf{1}] - 1\|_{L^2(\mu)}. \end{aligned}$$

The term

$$\|Q_{a,\varphi}^*[\mathbf{1}] - 1\|_{L^2(\mu)},$$

can be thought of as a measure of the divergence from invariance of the transition operator, an idea tracing back to [1].

Second, the centered term $\|Q_{a,\varphi}^*[f - 1]\|_{L^2(\mu)}$ rewrites:

$$\begin{aligned} \|Q_{a,\varphi}^*[f - 1]\|_{L^2(\mu)}^2 &= \int (Q_{a,\varphi}^*[f - 1])^2 d\mu = \int (f - 1)Q_{a,\varphi}Q_{a,\varphi}^*[f - 1]d\mu \\ &= \int (f - 1)M_{a,\varphi}[f - 1]d\mu, \end{aligned}$$

making appear the multiplicative reversibilization $M_{a,\varphi}$ of $Q_{a,\varphi}^*$ introduced above.

In terms of convolution, this amounts to consider separately $K_a(f - 1, \varphi)$ and $K_a(\mathbf{1}, \varphi)$.

Following this roadmap, Proposition 4.1 below deals with the default of invariance $\|Q_{a,\varphi}^*[\mathbf{1}] - 1\|_{L^2(\mu)}$ and Proposition 4.2 with the centered quantity $\|Q_{a,\varphi}^*[f - 1]\|_{L^2(\mu)}$. Theorem 1.2 is then proved in Section 4.4. For the sake of completeness, we conclude the part by stating an alternative bound to (E_r) in Section 4.5.

4.2 Improved Poincaré inequality for Ornstein-Uhlenbeck

The following result is an improvement of the usual Poincaré inequality for the Ornstein-Uhlenbeck semigroup (7) when more information is available on the function $f \in L^2(\mu)$ at play.

Proposition 4.1 (Improved Poincaré). *Let $r \in \mathbb{Z}^+$ and f be a function in $L^2(\mu)$ with Hermite decomposition of the form*

$$f = f_0 + \sum_{n \geq r+1} f_n \bar{H}_n,$$

Then for all $t \geq 0$,

$$\text{Var}_\mu(P_t f) \leq e^{-2(r+1)t} \text{Var}_\mu(f).$$

In particular, if φ is the density of a variable agreeing with the Gaussian moments up to r , then for all $a \in [0, 1]$,

$$\|Q_{a,\varphi}^*[\mathbf{1}] - 1\|_{L^2(\mu)} \leq (1 - a^2)^{\frac{r+1}{2}} \chi_2(\varphi).$$

Proof. Thanks to the properties of Hermite polynomials,

$$P_t f = \sum_{k=0}^{\infty} f_k P_t[\bar{H}_k] = h_0 + \sum_{k=r+1}^{\infty} f_k e^{-kt} \bar{H}_k;$$

$$\text{Var}_{\mu}(P_t f) = \sum_{k=r+1}^{\infty} f_k^2 e^{-2kt} \leq e^{-2(r+1)t} \sum_{k=r+1}^{\infty} f_k^2 = e^{-2(r+1)t} \text{Var}_{\mu}(f).$$

The second inequality of Proposition 4.1 follows from the first one by Remark 3.2. \square

4.3 Poincaré-like inequality for the convolution operator

For $a \in [0, 1]$ and a centered $g \in L^2(\mu)$ (that is $\mu(g) = 0$), let us consider the quantity $\|Q_{a,\varphi}^*[g]\|_{L^2(\mu)}$. By analogy with the Poincaré inequality for the Ornstein-Uhlenbeck semi-group, which is reversible with respect to μ , we call the following result a Poincaré-like inequality holding for the operator $Q_{a,\varphi}^*$, which in general is non-reversible.

Proposition 4.2 (Poincaré-like inequality). *Assume that φ is a polynomial density in $L^2(\mu)$ whose moments match the moments of μ up to order $r \in \mathbb{Z}^+$, and which satisfies to Hypothesis **(H)** stated in Section 2. Set $a_{\varphi} \in [0, 1)$ and $d_{\varphi} : (0, 1) \rightarrow \mathbb{R}$ as in Theorem 1.2. Then, for all function $g \in L^2(\mu)$ which writes as:*

$$g = \sum_{k=r+1}^{\infty} g_k \bar{H}_k,$$

and for all $a \in (a_{\varphi}, 1)$, it stands that:

$$\int (Q_{a,\varphi}^*[g])^2 d\mu \leq a^{r+1} (1 + d_{\varphi}(a))^2 \int g^2 d\mu.$$

The proof is cut out in a number of steps. We begin by the following lemma:

Lemma 4.3 (Gershgorin's theorem). *Let φ and g be as in Proposition 4.2, and set $K = r + 1$. Then, for all $a \in (0, 1)$,*

$$\int (Q_{a,\varphi}^*[g])^2 d\mu \leq \sup_{l \geq K} \Sigma_{a,\varphi}(l) \int g^2 d\mu, \quad (18)$$

where

$$\Sigma_{a,\varphi}(l) := \sum_{j=K}^{+\infty} \left| \vec{M}_{a,\varphi}(l, j) \right|, \quad l \geq K.$$

Proof. Let $r \in \mathbb{Z}^+$, $a \in (0, 1)$ and φ as in the statement of the proposition, and let $K = r + 1$. First, notice that V_K , the set of functions $g \in L^2(\mu)$ with Hermite decomposition

$$g = \sum_{k=K}^{\infty} g_k \bar{H}_k,$$

is stable under action of $Q_{a,\varphi}^*$ by Remark 3.4. The space V_K equipped with the $L^2(\mu)$ structure is again a Hilbert space. Let us call $Q_{a,\varphi|V_K}^*$ the restriction of $Q_{a,\varphi}^*$ to V_K . It is again bounded, with operator norm

$$\left\| Q_{a,\varphi|V_K}^* \right\|_{\text{op}} := \sup_{g \in V_K \setminus \{0\}} \frac{\|Q_{a,\varphi}^*[g]\|_{L^2(\mu)}}{\|g\|_{L^2(\mu)}}.$$

Hence, the desired majoration (18) is equivalent to the following bound on the operator norm:

$$\left\| Q_{a,\varphi|V_K}^* \right\|_{\text{op}}^2 \leq \sup_{l \geq K} \Sigma_{a,\varphi}(l). \quad (19)$$

The isometry between $L^2(\mu)$ and l^2 restricts to an isometry between V_K and l_K^2 , defined as the space of real sequences $(u_n)_{n \geq K}$ with $\sum_{n \geq K} u_n^2 < +\infty$. By this isometry, if $N_K = (N_K(i, j))_{i, j \geq K}$ stands for the infinite matrix associated to $Q_{a,\varphi|V_K}^*$, then:

$$\left\| Q_{a,\varphi|V_K}^* \right\|_{\text{op}} = \|N_K\|_{\text{op}}.$$

Furthermore, by the properties of block matrix multiplication, one sees that the matrix ${}^T N_K N_K$ is nothing else but the matrix $\vec{M}_{a,\varphi}$ defined in (17) (Section 3) restricted to l_K^2 , that is:

$${}^T N_K N_K = \left(\vec{M}_{a,\varphi}(i, j) \right)_{i, j \geq K}.$$

For a complex Banach space E , set $\mathcal{G}(E)$ the set of invertible operators on E . The spectral radius of a bounded operator M in E is then defined as

$$\rho(M) := \max \{ |\lambda|, \lambda \text{Id} - M \in \mathcal{G}(E) \}.$$

Moreover, if E is Hilbert and if T is a bounded operator of E with adjoint T^* , then

$$\|T\|_{\text{op}} = \sqrt{\rho(T^*T)}.$$

The operator N_K being a bounded operator of l_K^2 (by restriction of a bounded operator), the preceding equation applies:

$$\|N_K\|_{\text{op}} = \sqrt{\rho({}^T N_K N_K)}.$$

Let us recall a theorem of Gershgorin ([14]) related to finite complex auto-adjoint matrices A , where $A = (A_{i,j})_{1 \leq i, j \leq n}$ for a positive integer n . Denoting $\mathcal{G}_n(\mathbb{C})$ the set of invertible matrices of size n and $B(x, r)$ the complex ball of center $x \in \mathbb{C}$ and $r > 0$, one has:

$$\{ \lambda \in \mathbb{C}, \lambda \text{Id} - A \in \mathcal{G}_n(\mathbb{C}) \} \subset \bigcup_{l=1}^n B \left(A(l, l), \left| \sum_{1 \leq j \leq n, j \neq l} A(l, j) \right| \right).$$

As a consequence,

$$\rho(A) \leq \sup_{1 \leq l \leq n} |A(l, l)| + \left| \sum_{1 \leq j \leq n, j \neq l} A(l, j) \right| \leq \sup_{1 \leq l \leq n} \sum_{j=1}^n |A(l, j)|.$$

Gershgorin's theorem is stated for finite matrices, but the proof extends without difficulty to eigenvalues of operators on l_K^2 . The operator ${}^T N_K N_K$ being autoadjoint and compact by Proposition 3.5, its spectrum is included in the set of eigenvalues united with the singleton $\{0\}$, hence the formula above applies and yields majoration (19), which proves the lemma. \square

In order to derive an upper-bound of $\sup_{l \geq K} \Sigma_{a,\varphi}(l)$ from the explicit expression of $\vec{M}_{a,\varphi}$ stated in (17), we need two technical lemmas.

Lemma 4.4 (First technical lemma). *Let $N \geq K$ be positive integers, and call as in the preceding sections*

$$a_\varphi := \left(1 + \frac{N}{K}\right)^{-\frac{1}{4}}, \quad C_k := \left(1 + \frac{N}{K}\right)^{k/2}, \quad k \in \mathbb{Z}^+.$$

Let i, k be natural integers such that

$$0 \leq k \leq N - 1, \quad K \leq i + k \leq N, \quad 1 \leq i \leq N.$$

Then, for all $a \in (a_\varphi, 1)$ and for all $l \geq K$,

$$\begin{aligned} a^{-i} \binom{k+l}{k}^{1/2} \binom{k+l}{k+i}^{1/2} \mathbf{1}_{l \geq i+K} + a^i \binom{k+l+i}{k}^{1/2} \binom{k+l+i}{k+i}^{1/2} \\ \leq 2C_k C_{k+i} \binom{k+l}{k}^{1/2} \binom{k+l+i}{k+i}^{1/2}. \end{aligned}$$

Proof. Let $N \geq K$ be positive integers and i, k, l be natural integers such that

$$0 \leq k \leq N - 1, \quad K \leq i + k \leq N, \quad 1 \leq i \leq N, \quad l \geq K.$$

For two positive integers $m \geq n$, the notation $[m]_n$ stands for $[m]_n := m \cdots (m - n + 1)$. One has:

$$\binom{k+i+l}{k} \binom{k+l}{k}^{-1} = \frac{[k+i+l]_k}{[k+l]_k}.$$

If $i \geq k$, then

$$\binom{k+i+l}{k} \binom{k+l}{k}^{-1} \leq \left(\frac{l+N}{l+1}\right)^k \leq \left(\frac{l+N}{l+1}\right)^i.$$

If $i < k$, then

$$\binom{k+i+l}{k} \binom{k+l}{k}^{-1} = \frac{[k+i+l]_i [k+l]_{k-i}}{[k+l]_{k-i} [l+i]_i} = \frac{[k+i+l]_i}{[l+i]_i} \leq \left(\frac{l+N}{l+1}\right)^i.$$

In both cases,

$$\binom{k+i+l}{k} \binom{k+l}{k}^{-1} \leq \left(\frac{l+N}{l+1}\right)^i.$$

Furthermore, in the case where $l \geq i$,

$$\binom{k+l}{k+i} \binom{k+i+l}{k+i}^{-1} = \frac{[k+l]_k [l]_i}{[k+i+l]_i [k+l]_k} = \frac{[l]_i}{[k+i+l]_i} \leq \left(\frac{l}{l+1}\right)^i.$$

Hence, for all $a \in (0, 1)$,

$$\begin{aligned} & a^{-i} \binom{k+l}{k}^{1/2} \binom{k+l}{k+i}^{1/2} \mathbf{1}_{l \geq i+K} + a^i \binom{k+l+i}{k}^{1/2} \binom{k+l+i}{k+i}^{1/2} \\ & \leq \left(\frac{l}{l+1}\right)^{i/2} \binom{k+l}{k}^{1/2} \binom{k+l+i}{k+i}^{1/2} \left(a^{-i} \mathbf{1}_{l \geq i+K} + a^i \left(1 + \frac{N}{l}\right)^{i/2} \right) \\ & \leq \binom{k+l}{k}^{1/2} \binom{k+l+i}{k+i}^{1/2} f(a), \end{aligned} \tag{20}$$

where we defined

$$f(a) := a^{-i} + \left(1 + \frac{N}{K}\right)^{i/2} a^i, \quad a \in (0, 1).$$

As one checks easily, the inequality (20) still holds true if $l < i$, and $f'(a) \geq 0$ if and only if $a \geq a_\varphi = (1 + N/K)^{-1/4}$. This yields for all $a \in (a_\varphi, 1)$,

$$f(a) \leq f(1) = 1 + \left(1 + \frac{N}{K}\right)^{i/2} \leq 2 \left(1 + \frac{N}{K}\right)^{i/2} \leq 2C_k C_{k+i},$$

where C_k has been defined as $C_k = (1 + N/K)^{k/2}$ for all $k \in \mathbb{Z}^+$. \square

Lemma 4.5 (Second technical lemma). *Let $m > q$ be positive integers, and consider the polynomial $P = -\alpha X^m + \beta X^q - 1$ with $\alpha, \beta > 0$. Then $P \leq 0$ on \mathbb{R}^+ if and only if*

$$\left(\frac{\beta}{m}\right)^m \left(\frac{q}{\alpha}\right)^q \leq \frac{1}{(m-q)^{m-q}}.$$

Proof of Lemma 4.5. Let $P(x) = -\alpha x^m + \beta x^q - 1$ be as in the wording of the lemma. Then, for all $x \in \mathbb{R}$,

$$P'(x) = -m\alpha x^{m-1} + q\beta x^{q-1} = x^{q-1}(-m\alpha x^{m-q} + q\beta),$$

thus P attains its maximum on $[0, +\infty)$ at the point $x_0 = ((\beta q)/(\alpha m))^{1/(m-q)}$. Moreover,

$$\begin{aligned} P(x_0) &= x_0^q (-\alpha x_0^{m-q} + \beta) - 1 = \left(\frac{\beta q}{\alpha m}\right)^{\frac{q}{m-q}} \left(-\frac{\beta q}{m} + \beta\right) - 1 \\ &= \frac{\beta}{m} \left(\frac{\beta q}{\alpha m}\right)^{\frac{q}{m-q}} (m-q) - 1 = \left(\frac{\beta}{m}\right)^{\frac{m}{m-q}} \left(\frac{q}{\alpha}\right)^{\frac{q}{m-q}} (m-q) - 1, \end{aligned}$$

so that $P(x_0) \leq 0$ if and only if

$$\left(\frac{\beta}{m}\right)^m \left(\frac{q}{\alpha}\right)^q \leq \frac{1}{(m-q)^{m-q}}.$$

\square

We are now ready to show Proposition 4.2.

Proof of Proposition 4.2. Set $K = r + 1$, let $a \in (0, 1)$ and l a positive integer such that $l \geq K$. Then,

$$\begin{aligned}
\Sigma_{a,\varphi}(l) &= \sum_{j=K}^{+\infty} \left| \vec{M}_{a,\varphi}(l, j) \right| \\
&= \left| \vec{M}_{a,\varphi}(l, l) \right| + \sum_{i=1}^N \left(\left| \vec{M}_{a,\varphi}(l, l-i) \right| \mathbf{1}_{l-i \geq K} + \left| \vec{M}_{a,\varphi}(l, l+i) \right| \right) \\
&= a^{2l} \sum_{k \geq 0} \binom{k+l}{k} (1-a^2)^k \varphi_k^2 \\
&\quad + \sum_{i=1}^N a^{2l-i} \sum_{k \geq 0} \binom{k+l}{k}^{1/2} \binom{k+l}{k+i}^{1/2} (1-a^2)^{\frac{2k+i}{2}} \varphi_{k+i} \varphi_k \mathbf{1}_{l-i \geq K} \\
&\quad + \sum_{i=1}^N a^{2l+i} \sum_{k \geq 0} \binom{k+l+i}{k}^{1/2} \binom{k+l+i}{k+i}^{1/2} (1-a^2)^{\frac{2k+i}{2}} \varphi_{k+i} \varphi_k \\
&= a^{2l} \left(1 + \sum_{k=K}^N \binom{k+l}{k} (1-a^2)^k \varphi_k^2 + \sum_{0 \leq k < k+i \leq N} \left\{ (1-a^2)^{\frac{2k+i}{2}} |\varphi_k| |\varphi_{k+i}| \mathcal{C}_{a,\varphi}(l, i, k) \right\} \right),
\end{aligned}$$

where

$$\mathcal{C}_{a,\varphi}(l, i, k) := a^{-i} \binom{k+l}{k}^{1/2} \binom{k+l}{k+i}^{1/2} \mathbf{1}_{l \geq i+K} + a^i \binom{k+l+i}{k}^{1/2} \binom{k+l+i}{k+i}^{1/2}$$

is precisely the quantity addressed in Lemma 4.4. If $\varphi_k \varphi_{k+i} \neq 0$ and $i \geq 1$ then $i+k \geq K$, hence the assumptions of Lemma 4.4 hold and we get for all $a \in (a_\varphi, 1)$:

$$\begin{aligned}
\Sigma_{a,\varphi}(l) &\leq a^{2l} \left(1 + \sum_{k=K}^N \binom{k+l}{k} (1-a^2)^k \varphi_k^2 \right) \\
&\quad + 2 a^{2l} \left(\sum_{0 \leq k < k+i \leq N} \sum \left\{ (1-a^2)^{\frac{2k+i}{2}} |\varphi_k| |\varphi_{k+i}| C_k C_{k+i} \binom{k+l}{k}^{1/2} \binom{k+l+i}{k+i}^{1/2} \right\} \right).
\end{aligned}$$

Noticing that $C_k \geq 1$ for every positive integer K and that $C_0 = 1$ allows to recognize the development of a square:

$$\begin{aligned}
\Sigma_{a,\varphi}(l) &\leq a^{2l} \left(1 + \sum_{k=K}^N \binom{k+l}{k}^{1/2} (1-a^2)^{\frac{k}{2}} C_k |\varphi_k| \right)^2 \\
&\leq a^{2l} \left(1 + \sum_{k=K}^N \frac{C_k |\varphi_k|}{\sqrt{k!}} ((N+l)(1-a^2))^{\frac{k}{2}} \right)^2,
\end{aligned}$$

where we used that for all natural integer $k \leq N$,

$$\binom{k+l}{k} \leq \frac{(N+l)^k}{k!}.$$

We recognize the coefficient γ_k introduced in Section 2 to state Hypothesis **(H)**:

$$\gamma_k = \frac{C_k |\varphi_k|}{\sqrt{k!}}, \quad k \geq K,$$

so that

$$\Sigma_{a,\varphi}(l) \leq a^{-2N} a^{2(N+l)} \left(1 + \sum_{k=K}^N \gamma_k ((N+l)(1-a^2))^{\frac{k}{2}} \right)^2.$$

For all $a \in (0, 1)$ and $l \in \mathbb{Z}^+$, we perform the change of variables

$$u_{a,\varphi}(l) := -(l+N) \log a > 0,$$

so that

$$(N+l)(1-a^2) = (N+l) \left(1 - \exp\left(-2\frac{u_{a,\varphi}(l)}{l+N}\right) \right) \leq 2u_{a,\varphi}(l),$$

and introduce the function

$$h(u) = \exp(-u) \left(1 + \sum_{k=K}^N \gamma_k (2u)^{k/2} \right), \quad u \geq 0.$$

Then,

$$\Sigma_{a,\varphi}(l) \leq a^{-2N} h^2(u_{a,\varphi}(l)).$$

The last part of the proof is devoted to showing that the function h is non-increasing on $[0, +\infty)$; indeed in that case, we have for all $a \in (a_\varphi, 1)$ and $l \geq K$:

$$\Sigma_{a,\varphi}(l) \leq a^{-2N} h^2(u_{a,\varphi}(K)) = a^{2K} \left(1 + \sum_{k=K}^N \gamma_k (-2(K+N) \log a)^{k/2} \right)^2,$$

which, jointly with Lemma 4.3, proves Proposition 4.2. So let us study the variation of h . For all $u \geq 0$,

$$h'(u) = \exp(-u) \left(-1 - \sum_{k=K}^N \gamma_k (2u)^{k/2} + \sum_{k=K}^N \gamma_k k (2u)^{(k-2)/2} \right).$$

Let us consider separately the powers of $u^{\frac{1}{2}}$ ranging from K to $N-2$ (when existing) and the remaining powers:

$$\begin{aligned} -1 - \sum_{k=K}^N \gamma_k (2u)^{k/2} + \sum_{k=K}^N \gamma_k k (2u)^{(k-2)/2} &= \sum_{k=K}^{N-2} (-\gamma_k + (k+2)\gamma_{k+2}) (2u)^{k/2} \\ &\quad + K\gamma_K (2u)^{(K-2)/2} + \gamma_{K+1}(K+1)(2u)^{(K-1)/2} \\ &\quad - (1 + \gamma_N (2u)^{N/2} + \gamma_{N-1} (2u)^{(N-1)/2}). \end{aligned}$$

As Hypothesis **(H1)** holds, the sum $\sum_{k=K}^{N-2}$ is nonpositive. If $K \leq N - 1$, the remaining term writes

$$\begin{aligned} & \frac{1}{2} \left(-1 - 2\gamma_N(2u)^{N/2} + 2\gamma_{K+1}(K+1)(2u)^{(K-1)/2} \right) \\ & + \frac{1}{2} \left(-1 - 2\gamma_{N-1}(2u)^{(N-1)/2} + 2K\gamma_K(2u)^{(K-2)/2} \right), \end{aligned}$$

which is nonpositive thanks to Hypothesis **(H2a)** and Lemma 4.5. If $K = N$, the same arguments provide the nonpositivity of the remaining term, which reduces to

$$-1 - \gamma_N(2u)^{N/2} + N\gamma_N(2u)^{(N-2)/2}.$$

Finally, under Hypothesis **(H)**, we find that h has a nonpositive derivative on $[0, +\infty[$ hence is non-increasing, which completes the proof. \square

4.4 Proof of Theorem 1.2

Let us turn to the proof Theorem 1.2.

Proof of Theorem 1.2. Let $\varphi, f \in L^2(\mu)$ be as in the wording of the theorem. For all $a \in [0, 1]$,

$$\chi_2 \left(af * \sqrt{1-a^2}\varphi \right) = \|K_a(f, \varphi) - 1\|_{L^2(\mu)} \leq \|K_a(f-1, \varphi)\|_{L^2(\mu)} + \|K_a(\mathbf{1}, \varphi) - 1\|_{L^2(\mu)}.$$

According to Proposition 4.1, for all $a \in [0, 1]$,

$$\|K_a(\mathbf{1}, \varphi) - 1\|_{L^2(\mu)} = \|Q_{a,\varphi}^*[\mathbf{1}] - 1\|_{L^2(\mu)} \leq (1-a^2)^{\frac{r+1}{2}} \chi_2(\varphi),$$

while by Proposition 4.2, for all $a \in (a_\varphi, 1)$,

$$\|K_a(f-1, \varphi)\|_{L^2(\mu)} = \|Q_{a,\varphi}^*[f-1]\|_{L^2(\mu)} \leq a^{r+1} (1+d_\varphi(a)) \chi_2(f),$$

which proves the theorem. \square

4.5 Alternative bound

For the sake of completeness, let us conclude this section with a bound alternative to inequality (E_r) .

Proposition 4.6 (Alternative bound on χ_2 under convolution). *Let $f, \varphi \in L^2(\mu)$ be density with moments matching the Gaussian moments up to order $r \in \mathbb{Z}^+$, and moreover assume that f is $(r+1)$ -times derivable, with $D^{r+1}f \in L^2(\mu)$. Then, there exists a universal constant $c_r > 0$ such that $\forall a \in (0, 1)$,*

$$\begin{aligned} \chi_2 \left(af * \sqrt{1-a^2}\varphi \right) & \leq a^{r+1} \chi_2(f) + (1-a^2)^{\frac{r+1}{2}} \chi_2(\varphi) \\ & \quad + c_r (1-a^2)^{\frac{r+1}{2}} \left(\chi_2(f) \chi_2(\varphi) + \|D^{r+1}f\|_{L^2(\mu)} \|D^{-(r+1)}\varphi\|_{L^2(\mu)} \right), \end{aligned} \tag{21}$$

where D^{-1} stands for the operator which maps a function $\varphi \in L^2(\mu)$ onto its primitive with vanishing mean.

Contrarily to what happens for bound (E_r) , (21) stands for all $a \in (0, 1)$ and the polynomial assumption on φ is not required, making the relative roles of f and φ more symmetric. The drawback of bound (21) is that it involves the norm of the $(r + 1)$ -th derivative of f , which we fail to control in the framework of the Central Limit Theorem.

Proof of Proposition 4.6. Set $r \in \mathbb{Z}^+$, $K = r + 1$ and f, φ two densities as in the wording of the remark, so that

$$f = 1 + \sum_{k=K}^{+\infty} f_k \bar{H}_k, \quad \varphi = 1 + \sum_{k=K}^{+\infty} \varphi_k \bar{H}_k.$$

For all $a \in (0, 1)$, one has:

$$\begin{aligned} \chi_2 \left(af * \sqrt{1 - a^2} \varphi \right) &= \|K_a(f, \varphi) - 1\|_{L^2(\mu)} \\ &\leq \|K_a(f - 1, \mathbf{1})\|_{L^2(\mu)} + \|K_a(\mathbf{1}, \varphi - 1)\|_{L^2(\mu)} + \|K_a(f - 1, \varphi - 1)\|_{L^2(\mu)}. \end{aligned}$$

Now,

$$K_a(f - 1, \mathbf{1}) = P_{-\log a}[f - 1], \quad K_a(\mathbf{1}, \varphi - 1) = P_{-\frac{1}{2} \log(1 - a^2)}[\varphi - 1],$$

hence by the improved Poincaré inequality from Proposition 4.1 we get the two first terms of the bound. It remains to consider $\|K_a(f - 1, \varphi - 1)\|_{L^2(\mu)}$. For all $a \in (0, 1)$, one has by Remark 3.4:

$$K_a(f - 1, \varphi - 1) = \sum_{m=K}^{+\infty} \sum_{n=K}^{+\infty} \binom{m+n}{m}^{1/2} a^m (1 - a^2)^{n/2} f_m \varphi_n \bar{H}_{n+m},$$

hence

$$\begin{aligned} \|K_a(f - 1, \varphi - 1)\|_{L^2(\mu)}^2 &= \sum_{l=2K}^{+\infty} \left(\sum_{\substack{m+n=l \\ m, n \geq K}} \binom{l}{m}^{1/2} a^m (1 - a^2)^{n/2} f_m \varphi_n \right)^2 \\ &= (1 - a^2)^{K/2} \sum_{l=2K}^{+\infty} \left(\sum_{\substack{m+n=l-K \\ m \geq K, n \geq 0}} \binom{l}{m}^{1/2} a^m (1 - a^2)^{n/2} f_m \varphi_{n+K} \right)^2. \end{aligned}$$

By Cauchy-Schwarz inequality and Pascal formula, this rewrites again

$$\begin{aligned} (1 - a^2)^{K/2} \sum_{l=2K}^{+\infty} \left(\sum_{\substack{m+n=l-K \\ m \geq K, n \geq 0}} \binom{l-K}{m}^{1/2} \left(\frac{l \cdots (l-K+1)}{(n+K) \cdots (n+1)} \right)^{1/2} a^m (1 - a^2)^{n/2} f_m \varphi_{n+K} \right)^2 \\ \leq (1 - a^2)^{K/2} \sum_{l=2K}^{+\infty} \sum_{\substack{m+n=l-K \\ m \geq K, n \geq 0}} \frac{l \cdots (l-K+1)}{(n+K) \cdots (n+1)} f_m^2 \varphi_{n+K}^2. \end{aligned}$$

Now, there exists $c_K > 0$ such that $\forall x \geq K, \forall y \geq 0$,

$$(x + y + K) \cdots (x + y + 1) \leq c_K (x \cdots (x - K + 1) + (y + K) \cdots (y + 1)).$$

Applying this to $x = m$ and $y = n$, we find that

$$\begin{aligned} \|K_a(f-1, \varphi-1)\|_{L^2(\mu)}^2 &\leq c_K(1-a^2)^{K/2} \sum_{l=2K}^{+\infty} \sum_{\substack{m+n=l-K \\ m \geq K, n \geq 0}} \left(1 + \frac{m \cdots (m-K+1)}{(n+K) \cdots (n+1)}\right) f_m^2 \varphi_{n+K}^2 \\ &= c_K(1-a^2)^{K/2} \left(\sum_{m \geq K} f_m^2\right) \left(\sum_{n \geq K} \varphi_n^2\right) \\ &\quad + c_K(1-a^2)^{K/2} \left(\sum_{m \geq K} m \cdots (m-K+1) f_m^2\right) \left(\sum_{n \geq K} \frac{\varphi_n^2}{(n+K) \cdots (n+1)}\right), \end{aligned}$$

which is the Hermite representation of the expected quantity by formula (9). \square

5 Proof of Theorem 1.1

Finally, we conclude the article with the proof of our main theorem, Theorem 1.1, which follows on from the recursion formula (11) and barycentric convolution inequality for χ_2 -distance (E_r).

Proof of Theorem 1.1. In the framework of the theorem, denote

$$n_0 := \left\lceil \frac{1}{1-a_\varphi^2} \right\rceil \vee 2.$$

By the two aforementioned relations, we have for all integer $n \geq n_0$:

$$\chi_2(f_n) \leq \left(1 - \frac{1}{n}\right)^{\frac{r+1}{2}} \left(1 + d_\varphi \left(\sqrt{1-1/n}\right)\right) \chi_2(f_{n-1}) + \frac{1}{n^{\frac{r+1}{2}}} \chi_2(\varphi).$$

Remembering that $r \geq 2$, let us call for all $n \geq n_0$,

$$c_n := \left(1 - \frac{1}{n}\right)^{\frac{r+1}{2}} \left(1 + d_\varphi \left(\sqrt{1-1/n}\right)\right) = 1 - \frac{r+1}{2n} + \mathcal{O}\left(\frac{1}{n^{\frac{3}{2}}}\right), \quad d_n := \frac{1}{n^{\frac{r+1}{2}}} \chi_2(\varphi),$$

where we denote $v_n = \mathcal{O}(u_n)$ if $\limsup_{n \rightarrow +\infty} |v_n/u_n| < +\infty$. The preceding recursive inequality yields

$$\chi_2(f_n) \leq \left(\prod_{k=n_0}^n c_k\right) \chi_2(f_{n_0-1}) + \sum_{k=n_0}^n \left(\prod_{j=k+1}^n c_k\right) d_k.$$

Now,

$$\log \left(\prod_{k=n_0}^n c_k\right) = \sum_{k=n_0}^n \log c_k = - \sum_{k=n_0}^n \left(\frac{r+1}{2k} + \mathcal{O}\left(\frac{1}{k^{\frac{3}{2}}}\right)\right) = -\frac{r+1}{2} \log n + \mathcal{O}(1).$$

This leads to

$$\begin{aligned} \prod_{k=n_0}^n c_k &= \mathcal{O}\left(\frac{1}{n^{\frac{r+1}{2}}}\right); \quad \sum_{k=n_0}^n \left(\prod_{j=k+1}^n c_k\right) d_k = \left(\prod_{j=n_0}^n c_j\right) \sum_{k=n_0}^n \frac{d_k}{\prod_{j=n_0}^k c_j} \\ &= \mathcal{O}\left(\frac{1}{n^{\frac{r+1}{2}}}\right) \sum_{k=n_0}^n \mathcal{O}(1) \\ &= \mathcal{O}\left(\frac{1}{n^{\frac{r-1}{2}}}\right). \end{aligned}$$

Finally,

$$\chi_2(f_n) := \mathcal{O}\left(\frac{1}{n^{\frac{r+1}{2}}}\right) + \mathcal{O}\left(\frac{1}{n^{\frac{r-1}{2}}}\right) = \mathcal{O}\left(\frac{1}{n^{\frac{r-1}{2}}}\right),$$

which proves the theorem. \square

Remark 5.1 (Non i.i.d case). *If we suppose that the random variables $(X_i)_{i \geq 1}$ are independent and can be distributed according to densities $(\varphi_j)_{j \in J}$, where J is a finite set and each φ_j is polynomial and complies with **(H)**, then the integer n_0 above is replaced by*

$$n_0 := \max_{j \in J} \left\lceil \frac{1}{1 - a_{\varphi_j}^2} \right\rceil \vee 2,$$

and the quantities $(c_n, d_n)_{n \geq n_0}$ by

$$c_n := \left(1 - \frac{1}{n}\right)^{\frac{r+1}{2}} \max_{j \in J} \left(1 + d_{\varphi_j} \left(\sqrt{1 - 1/n}\right)\right) = 1 - \frac{r+1}{2n} + \mathcal{O}\left(\frac{1}{n^{\frac{3}{2}}}\right),$$

$$d_n := \frac{1}{n^{\frac{r+1}{2}}} \max_{j \in J} \chi_2(\varphi_j) = \mathcal{O}\left(\frac{1}{n^{\frac{r+1}{2}}}\right).$$

The rest of the proof is unchanged.

References

- [1] Marc Arnaudon and Laurent Miclo. A stochastic algorithm finding p -means on the circle. *Bernoulli*, 22(4):2237–2300, 2016.
- [2] Shiri Artstein, Keith M. Ball, Franck Barthe, and Assaf Naor. On the rate of convergence in the entropic central limit theorem. *Probab. Theory Related Fields*, 129(3):381–390, 2004.
- [3] Vlad Bally and Lucia Caramellino. Asymptotic development for the CLT in total variation distance. *Bernoulli*, 22(4):2442–2485, 2016.
- [4] Vlad Bally, Lucia Caramellino, and Guillaume Poly. Convergence in distribution norms in the CLT for non identical distributed random variables. *arXiv preprint arXiv:1606.01629*, 2017.
- [5] Andrew R. Barron. Entropy and the central limit theorem. *Ann. Probab.*, 14(1):336–342, 1986.
- [6] Andrew C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.*, 49:122–136, 1941.
- [7] S. G. Bobkov, G. P. Chistyakov, and F. Götze. Rényi divergence and the central limit theorem. *ArXiv e-prints*, 2016.
- [8] Sergey G. Bobkov, Gennadiy P. Chistyakov, and Friedrich Götze. Rate of convergence and Edgeworth-type expansion in the entropic central limit theorem. *Ann. Probab.*, 41(4), 2013.

- [9] Sergey G Bobkov, Gennadiy P Chistyakov, and Friedrich Götze. Fisher information and the central limit theorem. *Probability Theory and Related Fields*, 159(1-2):1–59, 2014.
- [10] Lawrence D. Brown. A proof of the central limit theorem motivated by the Cramér-Rao inequality. In *Statistics and probability: essays in honor of C. R. Rao*, pages 141–148. North-Holland, Amsterdam-New York, 1982.
- [11] Jérôme Dedecker, Florence Merlevède, and Emmanuel Rio. Rates of convergence for minimal distances in the central limit theorem under projective criteria. *Electron. J. Probab.*, 14:no. 35, 978–1011, 2009.
- [12] Carl-Gustav Esseen. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Math.*, 77:1–125, 1945.
- [13] James Allen Fill. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Probab.*, 1(1):62–87, 1991.
- [14] Semyon Gershgorin. Über die Abgrenzung der Eigenwerte einer Matrix. *Bulletin de l'Académie des Sciences de l'URSS*, (6), 1931.
- [15] Thierry Goudon, Stéphane Junca, and Giuseppe Toscani. Fourier-based distances and Berry-Esseen like inequalities for smooth densities. *Monatshefte für Mathematik*, 135(2):115–136, 2002.
- [16] I. A. Ibragimov. On the accuracy of approximation by the normal distribution of distribution functions of sums of independent random variables. *Teor. Veroyatnost. i Primenen.*, 11:632–655, 1966.
- [17] Ivan Nourdin and Giovanni Peccati. *Normal approximations with Malliavin calculus: from Stein's method to universality*, volume 192. Cambridge University Press, 2012.
- [18] Ivan Nourdin and Guillaume Poly. Convergence in total variation on Wiener chaos. *Stochastic Process. Appl.*, 123(2):651–674, 2013.
- [19] Emmanuel Rio. Upper bounds for minimal distances in the central limit theorem. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):802–817, 2009.
- [20] Emmanuel Rio. Asymptotic constants for minimal distance in the central limit theorem. *Electron. Commun. Probab.*, 16:96–103, 2011.
- [21] Claude E. Shannon and Warren Weaver. *The Mathematical Theory of Communication*. The University of Illinois Press, Urbana, Ill., 1949.
- [22] S. H. Siradžinov and M. Mamatov. On mean convergence for densities. *Teor. Veroyatnost. i Primenen.*, 7:433–437, 1962.
- [23] A. J. Stam. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control*, 2, 1959.
- [24] Hiroshi Tanaka. An inequality for a functional of probability distributions and its application to Kac's one-dimensional model of a Maxwellian gas. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 27:47–52, 1973.